Plasmonics and Plasmonic Metamaterials

Analysis and Applications

World Scientific Series in Nanoscience and Nanotechnology

Series Editor: Mark Reed (*Yale University*)

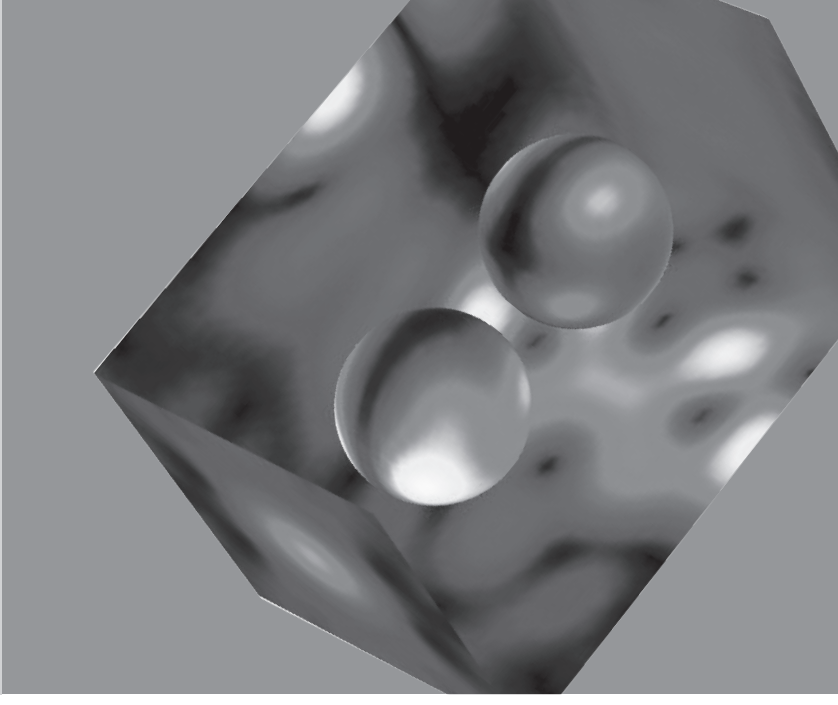
Vol. 1 Molecular Electronics: An Introduciton to Theory and Experiment
Juan Carlos Cuevas (*Universidad Autónoma de Madrid, Spain*) *and Elke Scheer* (*Universität Konstanz, Germany*)

Vol. 2 Nanostructures and Nanomaterials: Synthesis, Properties, and Applications, 2nd Edition
Guozhong Cao (*University of Washington, USA*) *and Ying Wang* (*Louisiana State University, USA*)

Vol. 3 Molecular Cluster Magnets
edited by Richard Winpenny (*The University of Manchester, UK*)

Vol. 4 Plasmonics and Plasmonic Metamaterials: Analysis and Applications
edited by Gennady Shvets (The University of Texas, Austin, USA) and Igor Tsukerman (The University of Akron, USA)

Volume
4
World Scientific Series in
Nanoscience and Nanotechnology



Plasmonics and Plasmonic Metamaterials

Analysis and Applications

Editors

Gennady Shvets
The University of Texas, Austin, USA

Igor Tsukerman
The University of Akron, USA

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Cover Credits:

The scanning electron microscopy image of a molded, 3D plasmonic crystal that exhibits angle-dependent surface sensitivity, an important characteristic for biosensing (Chapter 8). Plasmonic substrates with 3D topography support both surface plasmon polaritons and localized plasmons simultaneously. *Image credit:* Jiun-Chan Yang, Julia Y. Lin, and Teri W. Odom (Northwestern University)

Intensity of the Poynting vector in a metamaterial cell containing a pair of gold spheres. OpenMaX (Chapter 14) simulation of a periodic structure. *Image credit:* Christian Hafner (ETH Zürich).

Electric field around the Li-Stockman-Bergman nanoparticle cascade at a plasmonic resonance. *Image credit:* Jianhua Dai, Igor Tsukerman, František Èajko (the University of Akron), and Mark I. Stockman (Georgia State University).

World Scientific Series in Nanoscience and Nanotechnology — Vol. 4
PLASMONICS AND PLASMONIC METAMATERIALS
Analysis and Applications

ISBN-13 978-981-4355-27-8
ISBN-10 981-4355-27-5

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore.

CONTENTS

This page intentionally left blank

PREFACE

This volume combines two "hot" directions of current research in optics and photonics. *Plasmonic* effects — collective oscillations of electrons in metals coupled with electromagnetic waves (in optics, the waves of light) — play a dominant role in confinement, guiding and strong local enhancement of waves, which may be difficult or even impossible to achieve by any other means. *Metamaterials* are artificial structures judiciously designed, on a subwavelength scale, to control the propagation of electromagnetic waves and to devise physical properties not available otherwise. One striking example is artificial magnetism at high frequencies — a physical effect that does not exist in nature.

This book is a collective effort of many leading researchers in plasmonics, metamaterials and related areas. Fourteen chapters cover all aspects of the problems involved: analysis, fabrication, simulation and applications of various plasmonic structures and metamaterials. Reviews of recent developments, as well as new advances, are presented.

In the Analysis section, Khurgin & Sun deal with the fundamental questions of plasmon-enhanced absorption and emission of optical radiation: how the optical properties of atoms, molecules or other quantum-size objects are affected by plasmonic particles nearby. Khurgin & Sun develop an analytical model for surface plasmon modes of the particle and estimate the impact of these modes on the enhancement of optical radiation absorbed or emitted by the optically active objects, electroluminescence, and photoluminescence. Due to its analytical nature and relative simplicity, Khurgin & Sun's model opens up new avenues for optimization of optical nanostructures.

In Chapter 2, Fung *et al.* investigate very rich and intriguing properties of chiral plasmonic media with twisted anisotropic elements. The band structures and transmission properties of such media are analyzed.

Interestingly, chirality offers a way to achieve negative refraction — light bending the "wrong" way at material interfaces. (This happens when phase and group velocities of the wave are at an obtuse angle to one another in one of the materials.) Fung *et al.* show that chiral plasmonic media can support both positive and negative refraction; criteria for achieving negative refraction are established.

In Chapter 3, Petschulat *et al.* transfer the multipole analysis of classical electrodynamics to the field of optical metamaterials. The authors argue that artificial metamaterials "can be treated as ordinary matter but featuring beyond the electric dipole interaction also second order multipole interactions". As an example, the near-field response of split-ring resonators and other structures is used to reveal the multipolar contributions of the "meta-molecules". The review of classical multipole analysis and averaging procedures is useful in its own right.

Experimental studies and applications — a major part of this book — starts with Chapter 4 by De Leon and Berini, "Amplification and Lasing with Surface Plasmon Polaritons". Surface plasmon polaritons (SPPs) are light waves coupled with electron oscillation in metals at metal-dielectric interfaces. Losses in the metal, however, hamper many practical applications of SPP; thus amplification of these waves is of great practical interest. De Leon and Berini review major theoretical and experimental accomplishments in this field, in connection with amplified spontaneous emission, stimulated emission, and lasing. This contribution belongs to the general area of *active plasmonics* that is developing very rapidly.

Yu & Capasso, in Chapter 5, review recent work on plasmonic beam shaping of mid-infrared and far-infrared quantum cascade lasers (QCL). The reader unfamiliar with QCL will undoubtedly find an introductory review in the chapter useful. The authors explain how "metal and semiconductor microstructures can effectively tailor the dispersion properties of mid- and far-infrared surface plasmon polaritons". Plasmonic Bragg gratings, spoof surface plasmon polariton structures, and channel polariton structures are discussed along with their various applications. Yu & Capasso also identify future directions in wavefront engineering with plasmonics: producing diffraction-free Bessel beams and beams with orbital angular momentum that can be used to rotate small particles. New materials and systems in the mid-IR and THz ranges are also identified — for example, semiconductors for THz light and conductive oxides such as indium tin oxide (ITO) and aluminum zinc oxide (AZO) for the near-infrared wavelengths and beyond.

Chapter 6 by Altug *et al.* is devoted to integrated plasmonic systems for ultrasensitive infrared nanospectroscopy and biodetection. Infrared spectroscopy has exceptional identification capabilities via vibrational fingerprints of the bio- and chemical molecules in the mid-IR. Altug *et al.* show that diffractively coupled plasmonic nanoantennas provide strong plasmonic field enhancement leading to surface-enhanced spectroscopy with zepto-mole level sensitivities. A low-cost fabrication technique for high-throughput fabrication of infrared plasmonic antenna arrays is also introduced. Finally, Altug *et al.* describe a novel biosensing system merging nanoplasmonics and nanofluidics to overcome fundamental mass transport limitations imposed by conventional microfluidic approaches.

We already mentioned surface plasmon polaritons (SPPs) in connection with Chapter 4. Chapter 7 by Leosson and others deals with fabrication and optical characterization of long-range SPP devices. The design involves thin metal stripes or nanowires in a polymer matrix, and the devices include thermo-optically controlled interferometers, compact extinction modulators with low polarization-dependent loss and optically pumped plasmonic amplifiers.

Chapter 8 by Lin *et al.* is, like Chapter 6, devoted to biosensing, more specifically to highly sensitive label-free sensing. The high sensitivity is afforded by 3D plasmonic crystals, and the team has been able to distinguish surface sensitivity from bulk sensitivity. Plasmonic crystal structures that support both surface plasmon polaritons and localized surface plasmons can be used to provide complementary information about bulk and surface properties at the same time, which could make label-free, multi-modal readout systems possible for high throughput biodiagnostics.

Over the last decade, the curious and peculiar phenomenon of negative refraction (already mentioned in connection with Chapter 2) has generated unprecedented interest and excitement in applied physics, optics and engineering. In Chapter 9, Xiao & Shalaev review and report significant recent advancements that push the operational wavelengths of negative-index metamaterials into the optical range, allow tunability, and alleviate losses by introducing active media in the design. This research should facilitate practical applications of negative-index materials.

Future computing systems will increasingly include optical interconnections that can be expected to rely on the plasmonic confinement and guidance of light. In Chapter 10, Feng & Fainman deal with the design and fabrication of plasmonic chips capable of manipulating light from nano- to microscale. The authors use standard micro- and nanofabrication

tools to create complex geometries of composite metal-dielectric plasmonic structures and to realize the integration of plasmonics with the CMOS-compatible Silicon-On-Insulator (SOI) platform.

Integration of nanophotonic and electronics components on the same chip is also a major goal pursued by Holmgaard & Bozhevolnyi (Chapter 11). They consider modeling, design, and characterization of dielectric-loaded plasmonic waveguide components at telecommunication wavelengths. Holmgaard & Bozhevolnyi show that such waveguides may allow propagation of a single surface plasmon polariton mode with sub-wavelength confinement and low propagation loss. Coupling between two parallel waveguides, as well as modulation and wavelength-separation of signals are investigated.

Chapter 12 deals with an entirely different aspect of plasmonics: strong field enhancement and confinement may lead to strong electrodynamic forces acting on nanoparticles. Crozier *et al.* show in this chapter how such forces can be used to propel and manipulate nanoparticles on surfaces. The focus is on the applications to surface enhanced Raman spectroscopy.

The final part of this volume is devoted to methods of computer simulation in plasmonics, metamaterials, and in related problems of photonics. The numerical method described by Eremina *et al.* in Chapter 13 involves a superposition of fields due to discrete sources. Two illustrative application examples of light scattering by nano-objects are considered.

Finally, in Chapter 14 Hafner *et al.* review key computational techniques for electromagnetic problems, with a focus on plasmonics. Particular attention is paid to various symmetries that may be exploited in practical simulations. Several complicated problems are solved using the finite element method (FEM) as well as the multiple multipole program (MMP) developed by the authors.

***

We do hope that this book will prove useful to researchers and developers in industry and academia, as well as to graduate and advanced undergraduate students specializing in photonics, optics and related areas of applied physics. Thoughtful comments and critique from these readers will be greatly appreciated.

We thank World Scientific Publishing for all the effort and cooperation on this volume. Zvi Ruder was instrumental in getting the book project off the ground. Ryan Bong and Kostas Ikonomopoulos handled all aspects of their editorial work quite patiently and tirelessly.

Most of all, we are grateful to all the authors for their timely and inspirational research that is advancing the fields of plasmonics and metamaterials.

Gennady Shvets, the University of Texas at Austin
Igor Tsukerman, the University of Akron

This page intentionally left blank

ACKNOWLEDGMENT FROM I.T.

The editorial work on this book was performed during my one-year sabbatical leave. I thank the University of Akron for granting that leave and for providing partial financial support. I am immensely grateful to the many hosts who made the leave productive and utterly unforgettable: Weng Cho Chew, Che Ting Chan, Oszkár Bíró, Christian Magele, Kurt Preis, Sergey Bozhevolnyi, Greg Rodin, Stéphane Clénet, Guglielmo Rubinacci, Paolo Di Barba, Vadim Markel, Charles Epstein, Graeme Milton, Sascha Schnepp, Thomas Weiland, and others. Financial support from various research fellowship programs in the host institutions is greatly appreciated.

My participation in this book project would have been impossible without the incredible patience, love and support of my wife Anna. She is proof positive that miracles do happen in life sometimes.

PLASMONIC ENHANCEMENT OF OPTICAL PROPERTIES BY ISOLATED AND COUPLED METAL NANOPARTICLES

Greg Sun† and Jacob B. Khurgin*

†*Department of Physics, University of Massachusetts, Boston, USA*
greg.sun@umb.edu

**Department of Electrical and Computer Engineering,*
Johns Hopkins University, USA
jakek@jhu.edu

In this chapter we present a simple and comprehensive explanation of the mechanisms that can so dramatically modify the optical properties of atoms, molecules, or other quantum-size objects placed in the vicinity of metal nanoparticles. We develop a simple model that describes surface plasmon modes supported by the metal nanoparticles and describes them using just three key parameters — effective volume, Q-factor, and radiative decay rate. We subsequently apply this model to the tasks of estimating the enhancement of optical radiation, electroluminescence, and photoluminescence absorbed or emitted by the optically active objects in the presence of an isolated single nanoparticle. Using the example of gold nanospheres embedded in GaN dielectric, we show that enhancement for each case depends strongly on the nanoparticle size enabling optimization for each combination of absorption cross section, original radiative efficiency, and separation between the object and metal sphere. We then expand the model for single metal nanoparticles to coupled metal nanostructures. We show that complex structures can be treated as coupled multipole modes with highest enhancements obtained due to the superposition of these modes mainly in small particles. This model allows for optimization of the structures for the largest possible field enhancements, which depends on the quality factor Q of the metal and can be as high as Q^2 for two spherical particles. The "hot spot" can occur either in the nano-gaps between the particles or near the smaller

particles. We trace the optimal field enhancement mechanism to the fact that the extended dipole modes of larger particles act as the efficient antennas while the modes in the gaps or near the smaller particles act as the compact sub-wavelength cavities. The physically-transparent, comprehensive analytical approach developed in this chapter not only offers a quick route for optimization but also can be conveniently extended to incorporate large numbers of particles in various complex arrangements.

1. Introduction

Nanometer-scale metallic structures have been used as a means of dramatic modification of the optical properties of various optically active objects of similar or smaller dimensions such as atoms, molecules, or quantum dots (QDs)[1–8] (In this chapter we shall use the generic term "molecule" to denote them all). This rather significant modification is conceptually well understood as a product of the collective oscillations of free-electrons in the metal called surface plasmons (SPs) that induce strong localized electric fields near the surface of nanostructured metals[9] whose intensities exceed that of the average fields impinging on the structure by orders of magnitude. This phenomenon has been successfully used to demonstrate spectacular enhancement of sensitivity in Raman sensing[10–14] as well as in fluorescence measurements,[15–18] and has been proposed as a method to increase the efficiency of solar cells,[19] detectors,[20] and various nonlinear optical devices. Experimental studies have yielded many results and have been supported and complemented by vigorous theoretical-computational efforts of many research groups. But with all this massive effort no definite guidelines for the experimentalists have yet emerged and quite often the experimental work in the field still follows the trial and error pattern. This state of affairs, in our view, is the direct consequence of preoccupation with numerical models. These rather time-consuming numerical methods provide accurate results for each particular metallic nanostructure, but more often than not fail to present a broad picture and provide physical insight into the nature of enhancement and its limits. That makes optimization of the enhancement a daunting task. Over the last few years we have undertaken the task of changing this situation and developing a set of tools that would enable experimentalists to quickly ascertain the expected enhancement of optical properties using metal nanostructures.

Thus, in this chapter, we present a simple, analytical theory that not only provides a rigorous understanding of the origin of the optical

enhancement by metal nanoparticles, but also paves a clear path for optimization of the nanoparticle structure in order to maximize a given effect. We begin with the electro-static model to describe the SP modes that are supported by the metal nanospheres and their associated characteristics. Based on the effective volume method pioneered by Maier,[21] this model adequately takes into account the radiative decay of the lower order SP modes that is often neglected in the electrostatic approach. In addition the model properly incorporates the nonradiative decay of the higher order SP modes. At first we apply this model to calculate the enhancement of optical properties of molecules placed in close proximity to a single isolated metal nanoparticle, and study the enhancement of optical absorption and emission including electroluminescence (EL) and photoluminescence (PL). Besides proper consideration of higher order SP modes, the salient and novel feature of our approach is the fact that it is self-consistent in the sense that we properly include the changes imposed on the enhancement by the active molecules themselves. Once the model for an isolated particle has been developed, we use it to demonstrate the procedure for optimization and establish limits on how much enhancement one can realistically obtain for a given combination of molecule and metal in a surrounding dielectric medium.

Based on the established model of the SP modes of isolated metal nanospheres, we then proceed to expand this analytical description to coupled metal nanoparticles. It has become clear that enhancement is the largest in the so-called "hot spots"[22,23] occurring when the metal is structured in a rather sophisticated way with sharp peaks or small gaps. Maximum enhancement is limited by the metal loss. A single metal nanoparticle[24–34] having a simple smooth shape (sphere, ellipsoid, or nanorod) usually provides the electric field enhancement no larger than a Q-factor of the metal,[35,36] where $Q = \varepsilon''/\varepsilon'$ is the ratio of real and imaginary parts of dielectric function of the metal, and is no larger than a factor of 10–20 in the visible and near IR. But far more significant (up to three orders of magnitude) enhancement can occur in the intricately structured and arranged nanoparticles when the field gradually couples from the larger particles or regions serving as antennae into the smaller regions that serve as field-concentrating hot spots.[37] However, for the most part, the theoretical description of the field enhancement in the complex plasmonic nanostructures relies heavily on time-consuming numerical simulations, thus the basic physics behind the enhancement tends to become obscured making optimization rather difficult. With our fully

analytical "coupled modes model" for plasmonic optical field enhancement in complex metal nanostructures, we show that whether the enhancement is achieved near the small feature (nanotip) or inside the nanogap, the enhancement is proportional to Q^m where m is the effective number of sequential coupling transitions occurring between the light being coupled into the structure and it being concentrated around or inside the smallest surface feature.

This chapter is organized as follows. In Section 2, we provide a general description of the SP modes associated with an isolated metal nanosphere along with its effective mode volume, radiative and nonradiative decay rates. We then apply this model to the enhancement of optical absorption, EL and PL with examples of isolated Au nanosphere embedded in GaN dielectric. In Section 3, we developed an analytical coupled-mode theory for complex metal nanostructures by expanding the SP modes of single isolated metal nanoparticle to include the coupling effect of closely spaced metal nanospheres. We employ this theory to study the field enhancement in the gaps of coupled metal nanospheres of different geometries and compare it with that of optimized single isolated metal nanospheres. Finally, in Section 4, we summarize our conclusions.

The theory can be adapted rather easily to the elliptical particles of various eccentricities,[38,39] but, other than the shift in the resonance frequency the conclusions, at least qualitatively, will not change relative to the spherical particle, while the simplicity will be lost. Therefore we shall restrict ourselves to spherical particles and their combinations to present a simple, analytical model.

2. Optical Enhancement due to Isolated Metal Nanospheres

This section is intended to establish an analytical description of the SP modes around the metal nanospheres and bring together several optical effects that can be enhanced by placing molecules in close proximity to isolated metal nanoparticles including electric field (thus optical absorption), EL, and PL.

2.1. *Surface plasmon modes of an isolated metal sphere*

The geometry of a single metal nanosphere with a radius a being placed in a dielectric media with the dielectric constant ε_D is shown in Fig. 1 where the spherical polar coordinate system with the z-axis parallel to the

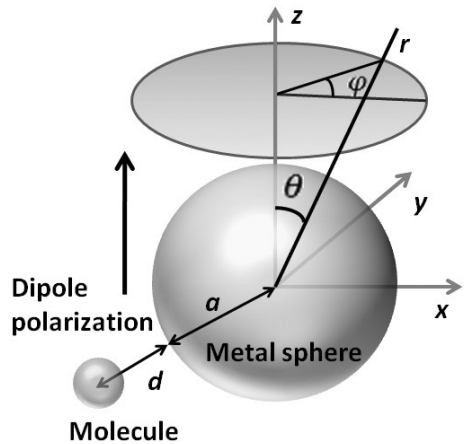


Fig. 1. Illustration of the spherical coordinate system used to describe the metal sphere dipole polarized along z-axis with a radius a separated from a molecule by a distance d.

dipole polarization is used. In the absence of external fields and charges, the electric potential of the eigenmodes under the electrostatic approximation should satisfy the Laplace equation

$$\nabla^2 \Phi = 0. \tag{1}$$

The solution for the l-th mode can be given as[40]

$$\Phi_l = \begin{cases} C \left(\frac{r}{a}\right)^l P_l(\cos\theta), & r < a \\ C \left(\frac{a}{r}\right)^{l+1} P_l(\cos\theta), & r \geq a \end{cases} \tag{2}$$

where $P_l(\cos\theta)$ is the Legendre polynomial and C is a normalization constant. Obviously, spherical symmetry allows additional solutions of the Laplace equation with angular dependence given by associated Legendre polynomials, $P_l^m(\cos\theta), -l \leq m \leq l$, but for as long as we consider the situation of a single active molecule in the vicinity of the metal nanosphere we can associate the polar axis with that molecule and then consider only $m = 0$ modes, disregarding their $2l + 1$ degeneracy.

While the continuity of the tangential electric field $E_\theta = -\frac{\partial \Phi_l}{r \partial \theta}$ at the boundary $r = a$ is automatically satisfied by the potential in Eq. (2), the continuity of the normal component of the electrical displacement $D_r = -\varepsilon \frac{\partial \Phi_l}{\partial r}$ requires that the dielectric functions of the metal $\varepsilon_M(\omega)$ and the

dielectric $\varepsilon_D(\omega)$ must be related as

$$l\varepsilon_M(\omega_l) + (l+1)\varepsilon_D(\omega_l) = 0 \tag{3}$$

where ω_l is the resonance frequency of the l-th mode. Using the Drude model

$$\varepsilon_M(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + j\omega\gamma} \tag{4}$$

in which ω_p is the plasma frequency and γ is the metal loss, and assuming frequency-independent ε_D, we arrive at the mode frequency

$$\omega_l = \omega_p \sqrt{\frac{l}{l + (l+1)\varepsilon_D}}. \tag{5}$$

With the metal dispersion taken into account,[41] Fig. 2 shows the mode frequency that ranges from $\hbar\omega_1 = 1.967\,\text{eV}$ to $\hbar\omega_\infty = 2.262\,\text{eV}$ along with its mode decay rate $\hbar\gamma$ due to metal loss that increases from 0.2 to 0.7 eV for Au nanosphere embedded in GaN. We shall refer to these solution as multipoles with the lowest ($l = 1$) solution being a dipole solution $\varepsilon_M(\omega_1) + 2\varepsilon_D(\omega_1) = 0$. As the mode order increases, the multipole solutions gradually

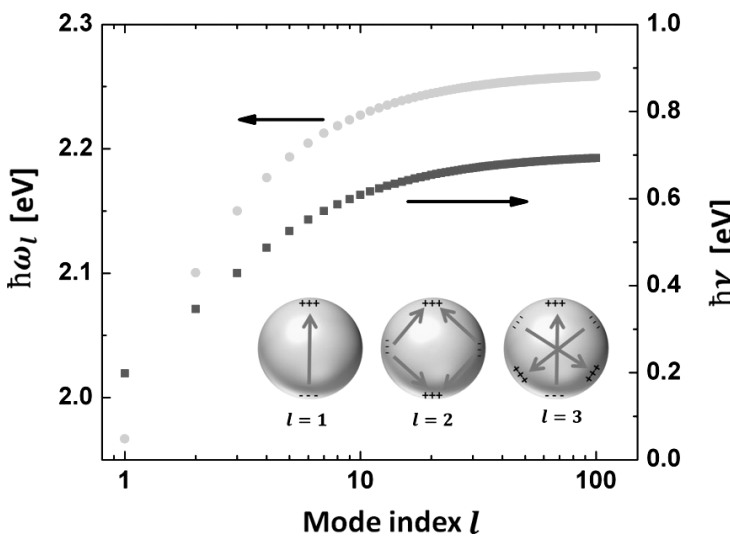


Fig. 2. Mode frequency $\hbar\omega_l$ and mode decay rate due to metal loss $\hbar\gamma$ vs. mode index for Au nanosphere embedded in GaN. Insert: illustration of the charge distribution on the metal surface along with the dipole moment for three lower order SP modes ($l = 1, 2, 3$).

approach the resonance of the surface plasmon polariton (SPP) on the flat metal-dielectric interface, $\varepsilon_M(\omega_\infty) + \varepsilon_D(\omega_\infty) = 0$ where $\omega_\infty = \omega_{\text{SPP}}$ is the resonance SPP frequency.

The electric field of the l-th mode can now be obtained as

$$\begin{aligned} & E_l(r,\theta) \\ & \quad = -\nabla\Phi \\ & \quad = \begin{cases} E_{\text{in},l} \\ E_{\text{out},l} \end{cases} \\ & \quad = \begin{cases} E_{\max,l}\left(\dfrac{r}{a}\right)^{l-1}\left[-\dfrac{l}{l+1}P_l(\cos\theta)\hat{r} + \dfrac{1}{\sin\theta}[P_{l+1}(\cos\theta) - \cos\theta P_l(\cos\theta)]\hat{\theta}\right], \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad r < a \\ E_{\max,l}\left(\dfrac{a}{r}\right)^{l+2}\left[P_l(\cos\theta)\hat{r} + \dfrac{1}{\sin\theta}[P_{l+1}(\cos\theta) - \cos\theta P_l(\cos\theta)]\hat{\theta}\right], \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad r \geq a \end{cases} \end{aligned} \tag{6}$$

It reaches maximum $E_{\max,l} = \frac{C}{a}(l+1)$ just outside of the metal sphere at $r = a$ and $\theta = 0$. The radial dependence of the electric field shows that as the mode order increases the mode gets more and more "compressed" closer to the surface of a nanoparticle. In fact for very large l the dependence on distance d from the surface can be approximated by an exponential one, $e^{-ld/a}$, which is a characteristic of SPP. This effect is better described by the parameter of effective mode volume $V_{\text{eff},l}$ that can be calculated as follows.

The surface charge density for the l-th mode is related to the normal component ($\hat{r}$) of the electric field in Eq. (6) at $r = a$ as

$$\begin{aligned} \sigma_l(\theta) &= \varepsilon_0(\varepsilon_M - 1)E^{\hat{r}}_{\text{in},l}(a,\theta) - \varepsilon_0(\varepsilon_D - 1)E^{\hat{r}}_{\text{out},l}(a,\theta) \\ &= \frac{2l+1}{l+1}\varepsilon_0 E_{\max,l}P_l(\cos\theta), \end{aligned} \tag{7}$$

where ε_0 is the permittivity of free space and we have used the relation Eq. (3). The effective volume of the l-th mode can be now defined through the mode energy[21] which is an integral of $\Phi_l\sigma_l/2$ over the sphere surface as

$$\begin{aligned} U_l &= \frac{1}{2}\iiint \Phi_l\sigma_l d^2r = 2\pi a^2\int_0^\pi \Phi_l\sigma_l a\cos\theta\sin\theta d\theta \\ &= \frac{1}{2}\varepsilon_0\varepsilon_D E^2_{\max,l}V_{\text{eff},l} \end{aligned} \tag{8}$$

to arrive at

$$V_{\text{eff},l} = \frac{4\pi a^3}{(l+1)^2 \varepsilon_D}, \tag{9}$$

which is always less than the volume of the nanosphere. As the mode order index l increases, the effective volume decreases with $(l+1)^{-2}$ as the SP energy gets concentrated in a very small volume close to the surface of the nanosphere and is contained within a narrow angle around axis z.

The dipole moment evaluated as an integral of the surface charge density [Eq. (7)] multiplied by its displacement $a\cos\theta$ along z-axis over the sphere surface vanishes for all higher order modes ($l \geq 2$)(Insert in Fig. 2), except the $l = 1$ mode whose dipole

$$p_1 = 2\pi a^2 \int_0^\pi \sigma_1(\theta) a \cos\theta \sin\theta \, d\theta = 2\pi a^3 \varepsilon_0 E_{\max,1}. \tag{10}$$

This dipole mode usually referred to as a localized SP mode of the nanosphere is the only solution coupled to the external fields for as long as the nanosphere diameter is much smaller than the wavelength, while all higher order modes remain uncoupled to external radiation modes. As a result, these higher order modes capable of compressing high intensity of field near the metal surface cannot be exploited to achieve high energy concentration for isolated symmetric spherical particles, at least in dipole approximation, since external fields cannot be coupled into any of the higher order modes. For the same reason, the dipole mode is also the only one subjected to the radiative dumping. Using the standard expression for the dipole radiating power in the dielectric with an index of refraction $n = \sqrt{\varepsilon_D}$,

$$P_{\text{rad}} = -\left(\frac{dU}{dt}\right)_{\text{rad}} = \frac{n^3 \omega_1^4}{12\pi\varepsilon_0 \varepsilon_D c^3} p_1^2, \tag{11}$$

where c is the speed of light in free space, it is easy to show that the radiative decay rate of the dipole mode

$$\gamma_{\text{rad}} = -\frac{1}{U_1}\left(\frac{dU_1}{dt}\right)_{\text{rad}} = \frac{2\omega_1}{3\varepsilon_D}\left(\frac{2\pi a}{\lambda_1}\right)^3 = \frac{2\omega_1}{3\varepsilon_D}\chi^3 \tag{12}$$

where ω_1 is the dipole oscillating frequency, λ_1 is the corresponding wavelength in the dielectric, and $\chi = 2\pi a/\lambda_1$ is the normalized metal sphere radius. Simultaneously, all the modes also experience nonradiative decay due to the imaginary part of the metal dielectric function at roughly the

same rate

$$\gamma_{\mathrm{nrad},l} = -\frac{1}{U_1}\left(\frac{dU_1}{dt}\right)_{\mathrm{nrad}} \approx \gamma \tag{13}$$

equal to the metal loss in the Drude model.

The latter fact can be rather easily understood from a rather straightforward argument. If all the relevant dimensions are significantly smaller than wavelength, then according to Maxwell's equations, it implies that the magnetic field becomes vanishingly small as we approach the so-called "electrostatic limit". In this case, as we follow the evolution of the energy contained in the medium, when the phase is 0 or π the energy is entirely in the form of potential energy of charges and potentials, but when the phase changes to $\pi/2$ or $3\pi/2$ the energy has nowhere to go but into the kinetic energy of moving electrons, and the electrons in the metal do get scattered by defects and phonons. Thus in a truly sub-wavelength structure (or in a propagating plasmon with a large wavevector), about half of the time all the energy is contained in the kinetic motion of electrons. Hence, on average, the energy loss rate is about one half of the energy loss rate in the metal 2γ, with little regard to any particular geometry. The decay rate can be summarized for all modes as

$$\gamma_l = \begin{cases} \gamma_{\mathrm{rad}} + \gamma, & l = 1 \\ \gamma, & l \geq 2. \end{cases} \tag{14}$$

The only SP mode that has a nonzero radiative out-coupling efficiency is the dipole mode

$$\eta_{dp} = \frac{\gamma_{\mathrm{rad}}}{\gamma_{\mathrm{rad}} + \gamma} = \frac{2Q\chi^3}{3\varepsilon_D + 2Q\chi^3} \tag{15}$$

where the Q-factor for the dipole mode

$$Q = \frac{\omega_1}{\gamma}. \tag{16}$$

Figure 3 shows the frequency dependence of the Q-factor for both Au and Ag obtained with the actual dispersions for the two metals.[41] Being a less lossy metal, the Q-factor of Ag is consistently greater than that of Au over a wide frequency range. For an Ag sphere embedded in GaN with the SP mode frequency ranging from $\hbar\omega_1 = 2.341\,\mathrm{eV}$ to $\hbar\omega_\infty = 2.912\,\mathrm{eV}$, the Q-factor is between $30 \sim 35$. For an Au sphere in GaN with $\hbar\omega_1 = 1.967\,\mathrm{eV}$ to $\hbar\omega_\infty = 2.262\,\mathrm{eV}$, the Q-factor is only between $3 \sim 10$.

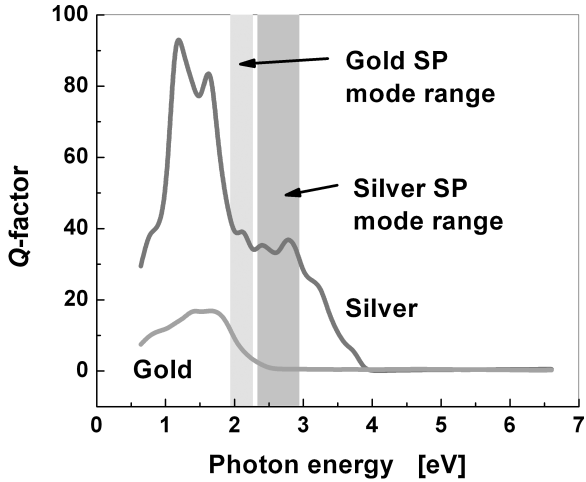


Fig. 3. The frequency dependence of the Q-factor for Au and Ag. The ranges of SP mode energy for Ag and Au spheres embedded in GaN are highlighted.

The higher order modes with their energy concentrated in smaller mode volumes and not subjected to radiative damping are capable of confining high density of energy for the enhancement of optical processes, but unfortunately, they simply produce none. In the case of optical absorption, the energy of optical excitation does not get coupled into these higher order modes because they are effectively decoupled from the radiation modes with their vanishing dipole moments. For optical emission, the energy from a nearby excited molecule gets coupled efficiently into these higher order modes, but then becomes trapped in these modes because they do not decay radiatively, and eventually dissipates due to metal loss. In the latter case, the higher order modes actually present another energy loss channel for the emission process that competes with the enhancement by the dipole mode.

2.2. *Absorption enhancement*

We now apply the above established model for the SP modes to study increased optical absorption that results when molecules are placed near isolated metal nanoparticles. This issue is important for two reasons — first of all it has been proposed to use metal nanoparticles to enhance the efficiency of photo-detectors[42,43] and solar panels,[44,45] and second, absorption is the first step in the photo-luminescence emission of PL sensors

and thus enhanced absorption means improved sensitivity. We use the example of Au nanosphere embedded in GaN dielectric to illustrate the calculation procedure for obtaining enhancement of optical absorption that takes into account all the radiative, nonradiative, and absorption losses. Most importantly, we show how to optimize enhancement for a given combination of metal and absorber characterized by its original absorption strength.

The way to evaluate unambiguously the field enhancement is to compare maximum field to that of a tightly focused light beam in the absence of metal spheres. Therefore we consider now a Gaussian beam with a numerical aperture characterized by a far field half angle θ_a gets focused onto a diffraction limited spot at the apex of the cone with a far-field solid angle Ω whose radius at the waist

$$w = \frac{\lambda_{\text{ex}}}{\pi\theta_a} \tag{17}$$

where λ_{ex} is the wavelength of the optical excitation in the dielectric at the frequency ω_{ex} as shown in Fig. 4 (The propagation direction has been rotated by $\pi/2$ relative to Fig. 1).[46] In the absence of the metal nanospheres, the field at the focal spot E_{foc} can be related to the power $|s_+|^2$ carried by

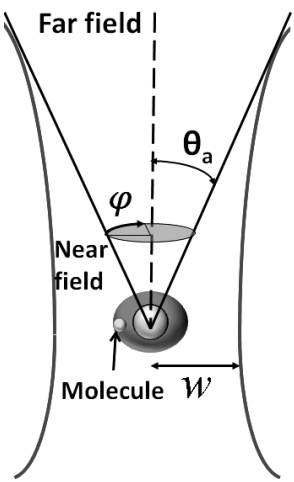


Fig. 4. Illustration of a metal nanosphere placed at the apex of a focused Gaussian beam with a numerical aperture characterized by the far-field half angle θ_a. A molecule is situated in the surrounding SP field of the nanosphere. In the absence of the metal sphere, the beam will be focused onto a diffraction-limited spot with radius w at the waist.

the incident wave as

$$|s_+|^2 = \frac{n}{Z_0}\pi\left(\frac{w}{2}\right)^2 E_{\text{foc}}^2 \tag{18}$$

where Z_0 is the impedance of free space.

In the presence of a metal sphere, the incident light can be coupled into the dipole mode but not the higher order modes because all $l \geq 2$ modes with vanishing dipole moments are not coupled with external fields. We therefore need only to include the dipole mode in the calculation of enhancement of optical absorption by an isolated metal sphere.

The optical absorption can be described as a two-step process in the presence of the SP modes supported by the metal sphere as illustrated in Fig. 5. First, the energy from free space modes contained in the focused beam gets coupled into the SP dipole mode with an in-coupling coefficient k_Ω. This is a process reciprocal to the radiative decay of the dipole mode into the free space radiation modes that are contained within the far field solid angle Ω which is characterized by the angle dependent decay rate

$$\gamma_{\text{rad}}^{\Omega} = \gamma_{\text{rad}} \int_0^{\Omega} f(\theta, \varphi) d\Omega \tag{19}$$

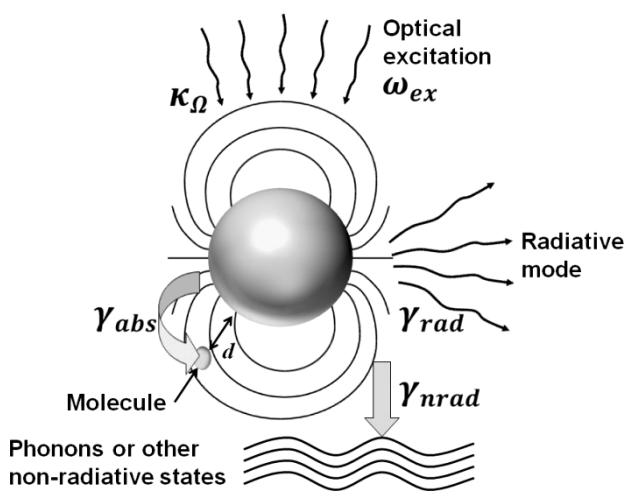


Fig. 5. Illustration of the enhancement process of optical absorption by a molecule placed at distance d from the metal sphere.

where the normalized angular dependence in the spherical system (Fig. 4) is

$$f(\theta, \varphi) = \frac{3}{8\pi}(1 - \sin^2\theta\cos^2\varphi) \tag{20}$$

Integrating (19) over small angles θ_a we obtain

$$\gamma_{\text{rad}}^{\Omega} \approx \frac{3}{8}\gamma_{\text{rad}}\theta_a^2, \tag{21}$$

and then using the reciprocity discussed by Haus[47] that relates the in-coupling coefficient k_Ω from the light cone with the radiative decay rate $\gamma_{\text{rad}}^{\Omega}$ of the SP mode into the same cone, we obtain

$$k_\Omega = \sqrt{\gamma_{\text{rad}}^{\Omega}} \approx \frac{\theta_a}{2}\sqrt{\frac{3\gamma_{\text{rad}}}{2}}. \tag{22}$$

The second step of the optical absorption is that the energy in the dipole mode actually gets absorbed by molecules with an absorption cross section σ_a located at a distance d away from the metal sphere. The rate of decay of the dipole mode due to absorption by the active molecules is

$$\gamma_{\text{abs}} = \frac{c}{n}\frac{N_a\sigma_a}{V_{\text{eff},1}}\left(\frac{a}{a+d}\right)^6 \tag{23}$$

where N_a is the number of molecules per sphere. Now we arrive at the rate equation for the amplitude $A_1 = \sqrt{U_1}$ of the dipole mode

$$\frac{dA_1}{dt} = j(\omega_{\text{ex}} - \omega_1)A_1 - \frac{\gamma_{\text{rad}} + \gamma_{\text{nrad}} + \gamma_{\text{abs}}}{2}A_1 + k_{\text{in}}^{\Omega}S_+ \tag{24}$$

The steady state solution of (24) yields

$$A_1 = \sqrt{\frac{3}{2}}\frac{\sqrt{\gamma_{\text{rad}}}\theta_a S_+}{(\gamma_{\text{rad}} + \gamma_{\text{nrad}} + \gamma_{\text{abs}}) + j2(\omega_1 - \omega_{\text{ex}})}. \tag{25}$$

Since the optical absorption strength is proportional to the electric field intensity (thus field squared), using the relations Eqs. (8) and (18), we obtain the absorption enhancement factor for a molecule located at a distance d (normalized $\chi_d = 2\pi d/\delta_1$) from the metal sphere

$$\begin{aligned} F_A &= \left|\frac{E_{\max,1}}{E_{\text{foc}}}\right|^2 \left(\frac{a}{a+d}\right)^6 \\ &= \left(\frac{\omega_1}{\omega_{\text{ex}}}\right)^2 \frac{2}{[Q^{-1} + Q_a^{-1} + 2\chi^3/3\varepsilon_D]^2 + \delta_{\text{ex}}^2}\left(\frac{\chi}{\chi + \chi_d}\right)^6 \end{aligned} \tag{26}$$

where we have introduced the absorption Q-factor as

$$Q_a = \frac{\omega_1}{\gamma_{\rm abs}} = \frac{\lambda_1^2}{N_a\sigma_a}\frac{\chi^3}{2\pi\varepsilon_D}\left(\frac{\chi+\chi_d}{\chi}\right)^6 \tag{27}$$

and the normalized excitation detuning $\delta_{\rm ex} = 2(1-\omega_{\rm ex}/\omega_1)$. It should be pointed out that the absorption enhancement depends on the absorption cross section of the molecule — better absorbers receive less enhancement as the absorption provides another channel for the SP energy to decay.

The result of absorption enhancement F_A at resonance ($\omega_{\rm ex} = \omega_1$) is shown in Fig. 6 as a function of sphere radius a by fixing $d = 5\,$nm for several values of the total absorption cross section $N_a\sigma_a$ with $Q \approx 10$ for Au nanosphere embedded in GaN. The origin of a strong dependence of the absorption enhancement factor F_A on sphere radius lies in the importance of a radiative decay term that is not taken account of in the simple electrostatic calculations. It is clear that the absorption of the active molecules becomes important when it approaches the loss in the metal — this can be confirmed by optimizing Eq. (26) at resonance for small $d \sim 0$. Then we find the optimal enhancement factor

$$F_{A,\rm opt} = \frac{2Q^2}{[1+4Q(\pi N_a\sigma_a/3\lambda_1^2)^{1/2}]^2} \tag{28}$$

at the normalized metal sphere radius $\chi_{\rm opt} = (3\pi\varepsilon_D^2 N_a\sigma_a/\lambda_1^2)^{1/6}$. Thus absorption in the active molecules themselves will become main limiting

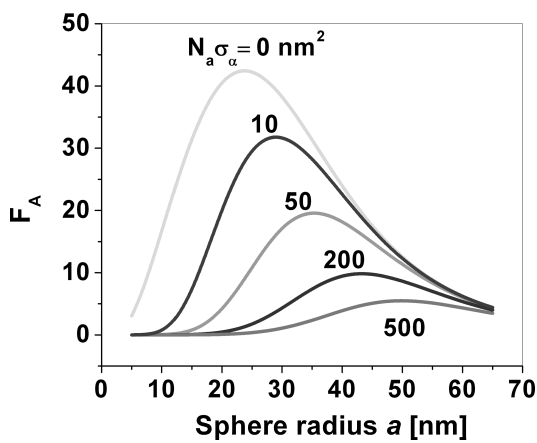


Fig. 6. Absorption enhancement as a function of sphere radius a for the Au/GaN system wherein the absorbing molecules with a range of total absorption cross section are placed at $d = 5\,$nm from the Au sphere.

factor when $N_a\sigma_a \geq \frac{3\lambda_1^2}{16\pi Q^2} \approx 40\,\mathrm{nm}^2$ for the Au/GaN system and would not change substantially over the most of visible and near IR spectrum because an increase in wavelength would be balanced by a corresponding increase in Q-factor as metals become less absorptive at longer wavelengths. For a relatively small number of organic molecules, even with absorption cross sections as high as $0.1\,\mathrm{nm}^2$ in the case of Rhodamine 6G[48] the absorption on active molecules will play no role in determining the limit of field enhancement. At the same time, if one considers detectors based on semiconductors, specifically semiconductor quantum dots, their cross sections are on the scale of a few nm^2 [49] — thus the factor $N_a\sigma_a$ can easily approach $100\,\mathrm{nm}^2$ and absorption becomes more relevant. As one can see from Fig. 6 for $N_a\sigma_a \sim 200\,\mathrm{nm}^2$ only a factor-of-few enhancement is attainable and even then only for relatively small distance from the metal surface.

To better emphasize the points made in this part we have performed optimization of the field enhancement for a wide range of total absorption cross section and spacing between the molecules and metal spheres. The optimized results, shown in Fig. 7, clearly indicate that the strongest absorption enhancement (by almost 2 orders of magnitude) is attainable for the small number of weakly absorbing molecules placed close to the metal spheres, while for the large number of strong absorbers spread out only a dozen nm further away from the spheres the field enhancement is far more modest if it exists at all.

2.3. *Electroluminescence enhancement*

The study of EL enhancement in the presence of metal nanoparticles has practical implications to improving the efficiency of light emitting devices. Our investigation is aimed to provide a simple answer to this question: for an emitter with radiative efficiency, what kind of enhancement can be achieved with a given metal and what should be the optimal parameters of the nanoparticles?

In the absence of metal, the original radiative decay rate $1/\tau_{\mathrm{rad}}$ of an electrically excited molecule emitting energy at the frequency ω_{EL} competes with its nonradiative decay rate $1/\tau_{\mathrm{nrad}}$, yielding an original radiative efficiency

$$\eta_{\mathrm{rad}} = \frac{\tau_{\mathrm{rad}}^{-1}}{\tau_{\mathrm{rad}}^{-1} + \tau_{\mathrm{nrad}}^{-1}}. \tag{29}$$

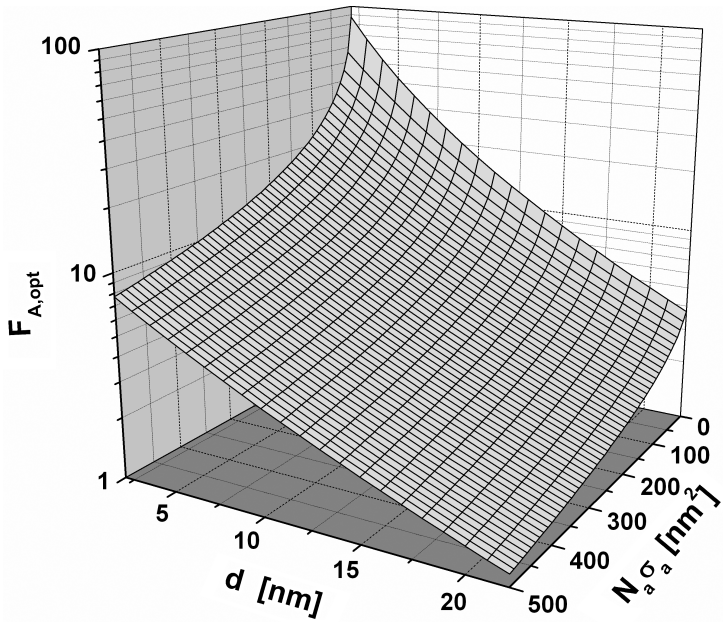


Fig. 7. Absorption enhancement in the Au/GaN system optimized for absorbing molecules with a total absorption cross section $N_a\sigma_a$ placed at a distance d from the Au sphere.

The energy transfer rate into radiation, according to the Fermi Golden rule, is proportional to the density of states (DOS) in the mode with a wavelength λ_{EL} in the dielectric

$$\rho_{\mathrm{rad}}(\omega_{\mathrm{EL}}) = \frac{1}{3\pi^2}\left(\frac{2\pi}{\lambda_{\mathrm{EL}}}\right)^3 \frac{1}{\omega_{\mathrm{EL}}} \tag{30}$$

whose polarization is lined up with the dipole of the molecule.

In the presence of the metal sphere, the molecule has an additional channel to relax by transferring energy into the SP modes supported by the metal sphere. Because SP modes are tightly confined with small effective mode volumes, their DOS are high compared to those in radiation modes. As a result, the energy transfer rate into the SP modes is enhanced by a factor which is essentially the DOS ratio, known as the Purcell Factor.[50] For the l-th order SP mode with the decay rate γ_l, we can obtain its effective DOS through its normalized Lorentzian linewidth factor

$$L_l(\omega) = \frac{\gamma_l/2\pi}{(\omega-\omega_l)^2 + \gamma_l^2/4} \tag{31}$$

and effective mode volume $V_{\text{eff},l}$ as

$$\rho_l(\omega, d) = \frac{L_l(\omega)}{V_{\text{eff},l}} \left(1 + \frac{d}{a}\right)^{-2l-4} \tag{32}$$

at a distance d from the surface of the metal sphere. The Purcell factor can then be evaluated as

$$F_{p,l}(\omega) = \frac{\rho_l(\omega, d)}{\rho_{\text{rad}}(\omega)} = \frac{3\pi\varepsilon_D(l+1)^2\omega L_l(\omega)}{4\chi^3} \left(\frac{\chi}{\chi + \chi_d}\right)^{2l+4}. \tag{33}$$

The EL enhancement can be treated as a two-step process as shown in Fig. 8. First, the electrically excited molecule with the original radiative decay rate $1/\tau_{\text{rad}}$ relaxes by emitting energy at the frequency ω_{EL} into each of the SP modes at the rate of $F_{p,l}/\tau_{\text{rad}}$, which is enhanced by the Purcell factor $F_{p,l}$ in Eq. (33). Simultaneously, the molecule also relaxes into nonradiative modes at its original nonradiative rate of $1/\tau_{\text{nrad}}$. Second, the energy transferred to SP modes needs to be coupled out into the radiation continuum. Among them only the dipole mode ($l = 1$) will couple out to radiation modes at the rate γ_{rad} while all higher order modes ($l \geq 2$) dissipate nonradiatively.

The Purcell factors for various SP modes are shown in Fig. 9 where the molecule is taken to be at resonance with the dipole mode $\omega_{\text{EL}} = \omega_1$ — the only mode that couples out to radiation. For small molecule-metal separation d, significant amount of energy is coupled into high order modes ($l \geq 2$) in which it decay only nonradiatively. This phenomenon

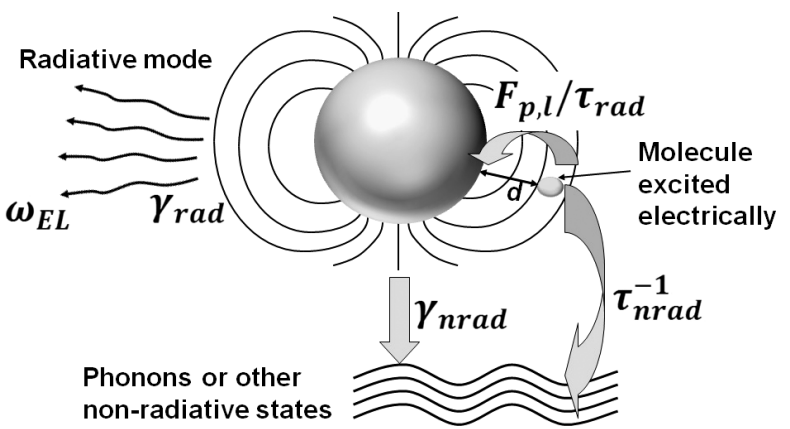


Fig. 8. Illustration of enhancement of EL from a molecule placed at a distance d from the metal sphere.

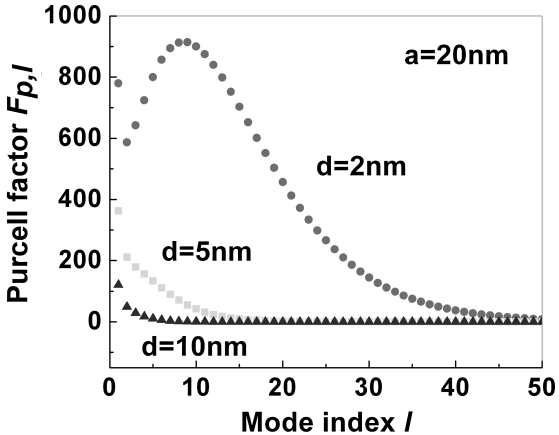


Fig. 9. Purcell factors associated with various SP modes for the molecule placed at $d = 2, 5, 10$ nm from the Au nanosphere with a radius of 20 nm in GaN.

is better illustrated with the cumulative effect of the higher order modes known as the "luminescence quenching ratio"

$$f_q = \frac{\sum_{l=2}^{\infty} F_{p,l}}{F_{p,1}} = \sum_{l=2}^{\infty} \frac{\pi\gamma_1 L_l(\omega_1)(l+1)^2}{8(1+\chi_d/\chi)^{2l-2}} \approx \frac{1}{4}\left(1+\frac{\gamma_{\rm rad}}{\gamma}\right)\frac{1}{1+\bar{\delta}^2_{l\geq 2}}\sum_{l=2}^{\infty}\frac{(l+1)^2}{(1+\chi_d/\chi)^{2l-2}} \tag{34}$$

where some average detuning, $\bar{\bar{\delta}}_{l\geq 2} = \frac{2(\bar{\omega}_{l\geq 2}-\omega_1)}{\gamma}$ has been taken for these non-radiative modes in the Lorentzian linewidth $L_l(\omega_1)$. The above summation Eq. (34) can be approximated analytically to show a d^{-3} dependence of the quenching ratio for small separation $d \ll a$, that can also be obtained in the vicinity of flat metal surface by the method of image charges, $f_q \approx \frac{1}{2}(\frac{\chi}{\chi_d})^3(1+\frac{2}{3\varepsilon_D}Q\chi^3)\frac{1}{1+\bar{\delta}^2_{l\geq 2}}$. The result is shown in Fig. 10 for EL quenching ratios vs. the molecule-sphere separation for a range of metal sphere size. Clearly, in order to limit the quenching ratio, the molecule cannot be placed too close to the metal nanoparticle.

Realizing that while energy of the emitting molecule can transfer into all SP modes with the Purcell enhancement factor $F_{p,l}$, but only the dipole mode ($l = 1$) couples out into radiation with the efficiency η_{dp}, we can thus write the EL efficiency in the presence of the metal sphere as

$$\eta_{\rm sp} = \frac{\tau_{\rm rad}^{-1} + F_{p,1}\tau_{\rm rad}^{-1}\eta_{dp}}{\tau_{\rm nrad}^{-1} + (1+\sum_{l=1}^{\infty} F_{p,l})\tau_{\rm rad}^{-1}}. \tag{35}$$

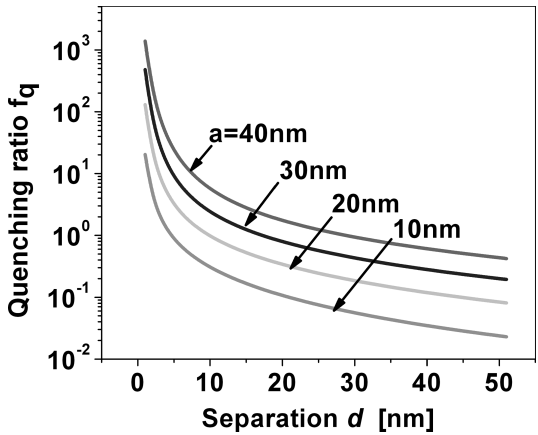


Fig. 10. EL quenching ratios vs. molecule-metal sphere separation for an Au sphere of different size.

Taking the ratio of $\eta_{\rm SP}$ to the original radiative efficiency $\eta_{\rm rad}$ in Eq. (29), we arrive at the EL enhancement factor as

$$F_L = \frac{\eta_{\rm SP}}{\eta_{\rm rad}} = \frac{1 + F_{p,1}\eta_{dp}}{1 + F_{p,1}}. \tag{36}$$

It is clear from Eq. (36) that the quenching effect of the higher order modes reduces attainable enhancement, but even when it is negligibly small, the out-coupling efficiency of the dipole mode must be greater than the original radiative efficiency, $\eta_{dp} > \eta_{\rm rad}$, in order to enhance luminescence.

Using the example of an isolated Au sphere in GaN, we have calculated the enhancement factor for a molecule emitting at the dipole mode resonance $\omega = \omega_1$ with $\eta_{\rm rad} = 0.01$ (Fig. 11). We can see that the enhancement factor exhibits strong dependence not only upon the nano-sphere size but also upon the separation between the molecule and sphere. The optimal nano-sphere size occurs at the radius where it is small enough to yield small effective mode volume for an enhanced Purcell factor, yet is still sufficiently large to assure strong radiative out-coupling of the dipole mode. If high order modes can be neglected, it is always better to have molecules positioned as close as possible to the metal sphere in order to take advantage of the large Purcell factor of the dipole mode. But with high order modes, energy of those molecules placed too close to the metal sphere also gets coupled into these nonradiative modes, and simply dissipates as metal loss. As a result, an optimized separation can be located to allow for significant

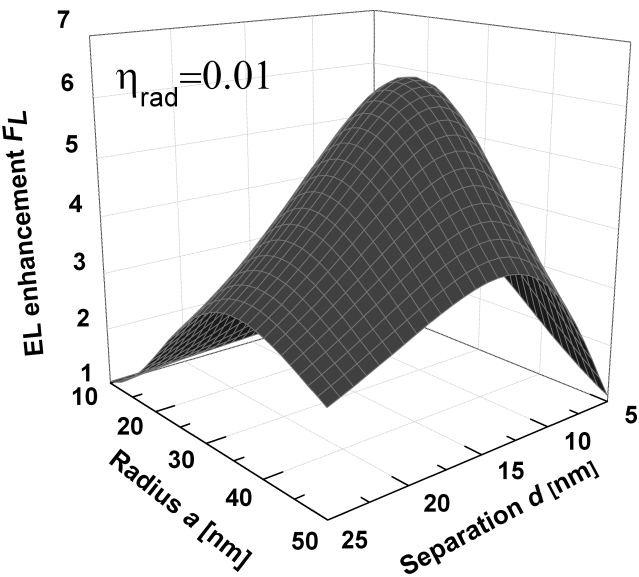


Fig. 11. Enhancement dependence on the Au nano-sphere radius and on the molecule-sphere separation for emitting at the dipole mode resonance $\omega_{\rm EL} = \omega_1$ with $\eta_{\rm rad} = 0.01$.

coupling into the dipole mode while adequately suppresses the luminescence quenching by high order modes.

The optimized enhancement $F_{L,\rm opt}$ for a wide range of original radiative efficiencies obtained at optimal nano-sphere size $a_{\rm opt}$ and separation $d_{\rm opt}$ is shown in Fig. 12. It can be seen that for molecules with low original radiative efficiencies the matter of most concern is to compete with the high nonradiative decay rate of the molecule by transferring energy into the dipole mode even at the expense of simultaneous coupling into high order modes, which clearly favors small nano-spheres and close separations with high Purcell factors. But for higher radiative efficiencies, the less critical becomes the concern of transferring energy from the emitter into the dipole mode, and accordingly the more important grows the concern for the efficient energy transfer from the dipole mode into free-space radiation modes and for the energy loss caused by coupling into the high order modes. Obviously, this situation favours larger nanoparticles that have greater radiative decay rate $\gamma_{\rm rad}$, as well as larger separations that reduce the energy coupling into the high order modes.

Before embarking on the treatment of photoluminescence one cannot help but notice a certain asymmetry in the way the higher order modes

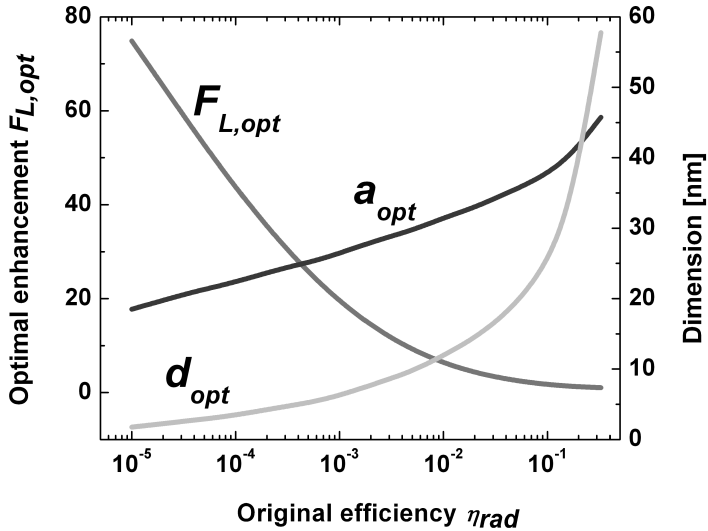


Fig. 12. Optimized enhancement, Au nano-sphere radius, and molecule-sphere separation over a range of original radiative efficiency.

influence emission and absorption. While the higher order modes act as additional loss channels in the case of emission, they appear to play no role at all in the case of absorption. In other words, while the luminescence gets quenched by the higher order modes, especially when the emitter is close to the nanoparticle, there is no equivalent effect of "absorption quenching", and one can, in principle place the absorber right next to the metal without any detrimental effect on the absorption efficiency. This statement is, in general, erroneous because in our treatment of absorption we have made an assumption that once the energy has been transferred from the dipole SP mode to the molecular excitation, it can no longer be transferred back into the energy of higher order SP modes. Often this is indeed the case as the SP-excited molecular excitation quickly decays into the lower energy excitations and thermal modes at a time scale that is much faster than the rate of emission into the SP mode. But in some cases, for instance in solar cell, the excited electron-hole pairs have a long lifetime and can in principle transfer their energy to higher order SP modes, thus effectively quenching the absorption process. While consideration of the absorption quenching in the specific materials is beyond the scope of this chapter, the reader should pay attention to it and in general avoid getting too close to the metal.

2.4. *Photoluminescence enhancement*

The treatment of luminescence enhancement in the previous section is adequate in estimating efficiency improvement for EL where the excitation energy in the form of an electric dipole gets coupled directly into the SP modes. But for PL, the situation is more complicated. There are two enhancement mechanisms at work. One takes place during optical excitation, and the other during light emission. In the absence of metal nanoparticles, the optical excitation typically in the form of a focused laser beam is being absorbed by the molecules. With an isolated metal nanoparticle placed in the vicinity of the molecules, the beam can now get coupled first into the confined SP dipole mode, concentrating the optical energy near the molecule and thus enhancing its absorption rate. The same SP modes (albeit at different frequency) can also enhance the energy-emission efficiency of the excited molecules through the Purcell effect. A proper estimate of PL enhancement by metal nanoparticles therefore must include both energy absorption and emission processes that are enhanced simultaneously by the SP modes. To make the matter even more complicated, all of these processes depend on the frequency relationships between the optical excitation, the molecule emission, and the SP modes. A clear understanding of the enhancement mechanism is obviously important for improving the performance of sensors relying upon PL. With the models developed for optical absorption (Section 2.2) and EL emission (Section 2.3) by metal nanoparticles, we can now calculate how much total enhancement one can realistically obtain in a PL process for a molecule placed near a given metal embedded in a given dielectric medium. We once again can provide an analytical approach for optimizing the metal nanostructure in order to achieve maximum enhancement. The salient feature of our approach is that it shows the attainable PL enhancement can be optimized for each particular molecule characterized by the absorption cross section $\sigma_a(\omega_{\rm ex})$ at the excitation frequency $\omega_{\rm ex}$, the radiative efficiency $\eta_{\rm rad}(\omega_{\rm PL})$ at the PL-emission frequency $\omega_{\rm PL}$, and their product — the PL cross section $\sigma_{\rm PL}(\omega_{\rm ex},\omega_{\rm PL}) = \sigma_a(\omega_{\rm ex})\eta_{\rm rad}(\omega_{\rm PL})$.

The enhancement of a PL process is illustrated in Fig. 13 which is essentially the combination of optical absorption and EL emission processes illustrated in Figs. 5 and 8, respectively. Optical excitation at the frequency of $\omega_{\rm ex}$ in the form of a laser beam is focused into the region where a metal nanosphere and a molecule separated by a distance d are located. The excitation beam couples only into the dipole mode around the metal sphere with an in-coupling coefficient k_Ω. Energy inside the dipole mode is then

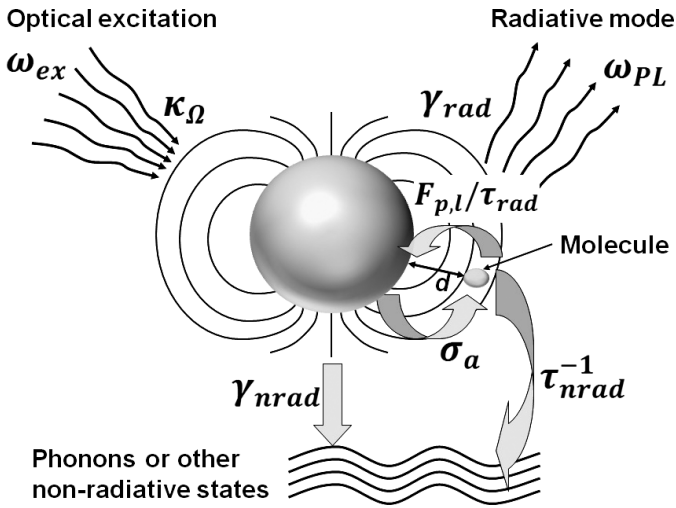


Fig. 13. Illustration of the enhancement of a PL process by the in-coupling of the optical excitation into the SP mode surrounding a metal sphere and by the out-coupling of the SP mode into the radiative mode.

absorbed at the rate of $\gamma_{\rm abs}$ by the active molecule with an absorption cross section σ_a. This process competes with radiative and nonradiative decays of the SP mode with decay rates of $\gamma_{\rm rad}$ and $\gamma_{\rm nrad}$, respectively. The optically excited molecule with the original radiative decay rate $1/\tau_{\rm rad}$ subsequently relaxes by emitting energy at the frequency $\omega_{\rm PL}$ into all SP modes at the rate of $F_{p,l}/\tau_{\rm rad}$ enhanced by the Purcell factor $F_{p,l}$. Among them only the dipole mode ($l = 1$) will couple out to radiation modes with the efficiency η_{dp} given in Eq. (15) while all higher order modes ($l \geq 2$) dissipate nonradiatively. It is clear that strong PL enhancement occurs when the frequencies of both optical excitation and emission are close to the resonance of the dipole mode ω_1. It is thus optimal to have the frequency relationship $\omega_{\rm ex} \geq \omega_1 \geq \omega_{\rm PL}$ for PL measurement.

Combining the two sequential enhancement processes given by Eqs. (26) and (36) we arrive at the resulting PL enhancement factor

$$F_{\rm PL}(\omega_{\rm ex}, \omega_{\rm PL}) = F_A(\omega_{\rm ex}) F_L(\omega_{\rm PL}). \tag{37}$$

Once again, this factor reaches maximum at some optimized nanoparticle size a. This fact can be traced to the two mutually exclusive roles that the nanoparticle plays in enhancing the PL process — that of antenna for efficient in- and out-coupling of energy and that of a nanocavity for energy

concentration. An efficient antenna requires a large dipole, while a high concentration of energy calls for a small nanocavity. Simultaneously, the impact of higher order modes on the luminescence enhancement is such that an optimal molecule-sphere separation maximizes the overall enhancement. Therefore, for each combination of $N_a\sigma_a, \eta_{\rm rad}$, there exist optimum $a_{\rm opt}$ and $d_{\rm opt}$ that maximize PL enhancement.

It is easy to see from Eq. (37) that the best PL enhancement requires both absorption and luminescence enhancement to be large, which calls for a small number of molecules placed reasonably close to the metal sphere, $d < a$, with a small absorption cross section, $Q_a \gg Q$, and a small original radiative efficiency such that $F_{p,1}(1+f_q)\eta_{\rm rad} \ll 1$, and both excitation and PL frequencies are close to SP resonance, $\omega_{\rm ex}, \omega_{\rm PL} \sim \omega_1$. Under these favorable conditions, we obtain this simple expression for PL enhancement

$$F_{\rm PL} \approx 2Q_1^2(1+2Q_1^2) \approx 4Q_1^4 \tag{38}$$

where we have introduced the Q-factor for the dipole mode which decays at the rate γ_1 as $Q_1 = \frac{\omega_1}{\gamma_1} = \frac{\omega_1}{\gamma_{\rm rad}+\gamma} < Q$. For small metal spheres, $\gamma_{\rm rad} \ll \gamma$, the maximum enhancement factor $F_{\rm PL,max} \approx 4Q^4$ in line with what a simple electrostatic analysis predicts. For an Au nanosphere embedded in GaN ($Q \approx 10$), $F_{\rm PL,max} \approx 4 \times 10^4$. This is a huge enhancement, but in reality, the PL enhancement is not nearly as significant when the finite absorption cross section and original radiative efficiency of the molecules that are spaced a finite distance away from the metal sphere are taken into account.

Consider the example of InGaN QDs situated at $d = 5\,{\rm nm}$ away from an Au sphere in GaN with $N_a\sigma_a = 1\,{\rm nm}^2$ and $\eta_{\rm rad} = 0.01$. The resulting PL enhancement factor $F_{\rm PL}$ along with its enhancement contributions from absorption F_A and luminescence F_L is shown in Fig. 14 where the optical excitation and PL frequencies are very close to the resonance of SP mode, $\omega_{\rm ex}, \omega_{\rm PL} \sim \omega_1$. There exists an optimized size of the metal sphere for which maximum enhancement is achieved. Also, as one can see, the two contributions are roughly of the same order. Similar to the EL enhancement F_L which peaks at an optimal molecule-sphere separation $d_{\rm opt}$, the PL enhancement can also be optimized with regard to this separation. Figure 15(a) shows this dependence on the size of metal nanosphere and the molecule sphere separation from which we extract optimized values of $a_{\rm opt} = 27\,{\rm nm}$ and $d_{\rm opt} = 6.5\,{\rm nm}$ for both excitation and PL frequencies in resonance with the dipole mode, $\omega_{\rm ex}, \omega_{\rm PL} \sim \omega_1$. With a

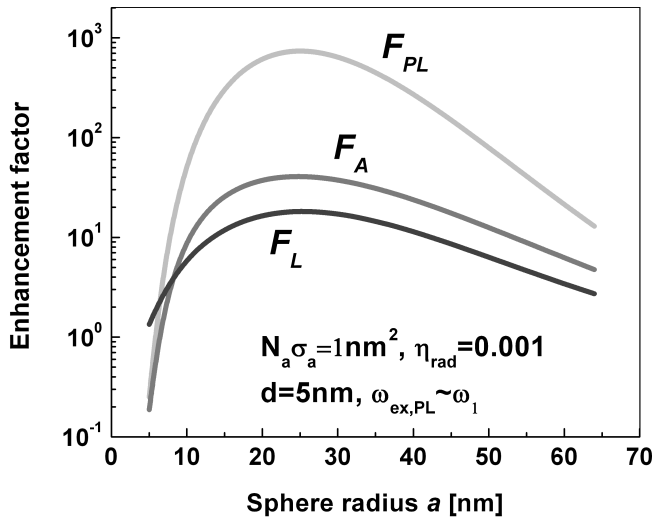
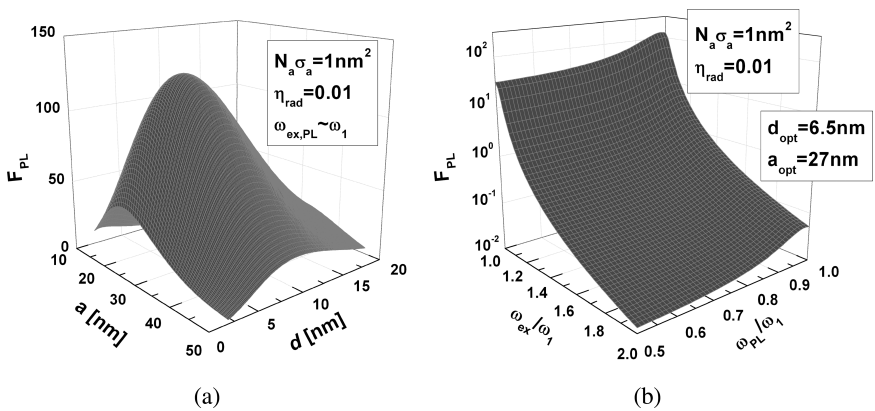


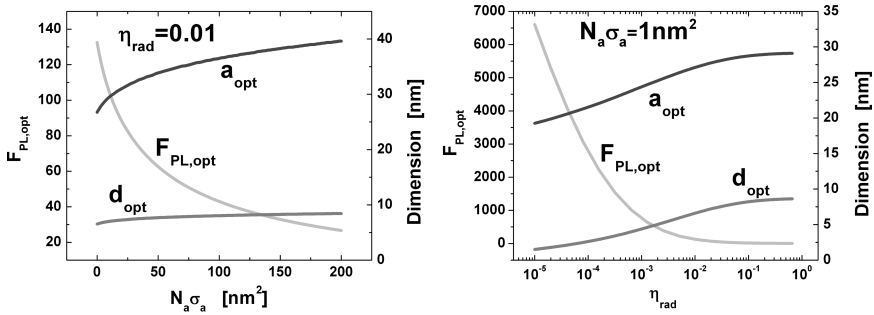
Fig. 14. Absorption F_A, emission F_L, and total enhancement factors $F_{\rm PL}$ vs. metal sphere radius for Au/GaN.



(a) (b)

Fig. 15. (a) PL enhancement vs. Au sphere radius a and molecule-sphere separation d for $N_a\sigma_a = 1\,\mathrm{nm}^2$ and $\eta_{\rm rad} = 0.01$ near resonance $\omega_{\rm ex}, \omega_{\rm PL} \sim \omega_1$, (b) PL enhancement vs. frequency detuning ratio of the optical excitation $\omega_{\rm ex}/\omega_o$ and PL emission $\omega_{\rm PL}/\omega_o$ at $a_{\rm opt} = 27\,\mathrm{nm}$ and $d_{\rm opt} = 6.5\,\mathrm{nm}$ — optimized values at $\omega_{\rm ex}, \omega_{\rm PL} \sim \omega_1$.

and d fixed at these values, Fig. 15(b) shows the PL enhancement as a function of the detuning of optical excitation $\omega_{\rm ex}/\omega_1$ and that of PL emission $\omega_{\rm PL}/\omega_1$ for the same example — which clearly suggests that it is more critical to have the excitation frequency near resonance with the



(a) (b)

Fig. 16. PL enhancement near resonance $\omega_{\rm ex}, \omega_{\rm PL} \sim \omega_1$, optimized at Au sphere radius and molecule-sphere separation as a function of (a) absorption cross section $N_a\sigma_a$ at fixed $\eta_{\rm rad} = 0.01$, and (b) original radiative efficiency $\eta_{\rm rad}$ at fixed $N_a\sigma_a = 1\,{\rm nm}^2$.

dipole mode, indicating that the matter of most concern is the transfer of optical excitation into the system in order for the PL process to take place.

The dependence of PL enhancement optimized at $a_{\rm opt}$ and $d_{\rm opt}$ on the total absorption cross section $N_a\sigma_a$ Fig. 16(a) and on the original radiative efficiency $\eta_{\rm rad}$ Fig. 16(b) shows that PL enhancement is strong only if the active molecules are both weak absorbers and inefficient emitters being positioned in close proximity to the metal nanoparticles, $d_{\rm opt} < a_{\rm opt}$. Note that condition $\sigma_a\eta_{\rm rad} \to 0$ under which the maximum enhancement is achieved is always satisfied for Raman scattering, which can be seen as nothing but PL with a negligibly small cross section, hence $F_{\rm Raman} \approx F_{\rm PL,max}$ as discussed above. Indeed the experimentally verified enhancement of Raman scattering is always significantly larger than that for PL.[10–14]

3. Enhancement due to Coupled Metal Nanoparticles

The enhancement results that we have obtained with the example of a single Au sphere embedded in GaN are far from spectacular, especially for molecules with not so small absorption cross sections and/or original radiative efficiencies. The obstacle in achieving higher enhancement lies in the fact that it requires the single metal nanoparticle to be both an efficient antenna and resonator. We encountered this challenge while analyzing single spherical nanoparticles and attempting to maximize the enhancement by optimizing the nanoparticle size, because the large particles acted as good antennae but poor resonators, and the small particles provided excellent confinement but acted as very poor antennae.

Can we use the different SP modes of the single metal nanosphere to solve such a dilemma? The higher order modes with their energy concentrated in smaller mode volumes and not subjected to radiative damping should in principle provide an excellent confinement mechanism for the enhancement of optical processes; unfortunately, these modes do not couple well into radiating modes outside the nanoparticle because of their vanishing dipole moments. At the same time, the $l = 1$ dipole mode does couple to the outside rather well, but its effective volume is relatively large and thus the enhancement of optical processes cannot be all that high. In other words, the $l = 1$ mode acts as an efficient antenna, while $l \geq 2$ modes act as efficient resonators. But the different SP modes of an isolated symmetric structure are orthogonal and thus decoupled from one another. Thus, the higher order modes of isolated metal nanosphere cannot be utilized to improve enhancement.

It is therefore only natural then to follow the techniques used in microwave engineering, where no one ever dreams of combining antenna and cavity into one element, but use two distinct elements, antenna and resonator coupled to each other. Combining two or more nanoparticles (Fig. 17) allows us to engineer the schemes in which efficient antennas are coupled into the resonators with high confinement. One can think of two ways of attaining this. In case of two spheres of equal dimensions the dipole modes in both spheres act as antennae and the superposition of higher order

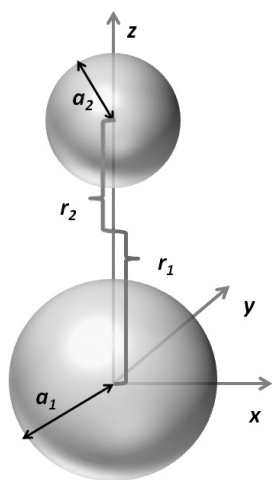


Fig. 17. Illustration of the geometry of two coupled metal spheres that are separated by $r_0 = r_1 + r_2$.

modes act as resonators allowing efficient coupling of the radiation into the gap region. This is the dimer case considered in Ref. (51). In the case of two extremely dissimilar spheres or nanolens[52] where for the most part only the dipole modes participate in energy concentration, the dipole mode of the larger sphere acts as an antenna and the dipole mode of the smaller particle acts as a resonator. Obviously when the spheres are of different size but are still on the same scale, the field enhancement mechanism is a combination of both aforementioned effects and all the higher order modes should be taken into account.

In this section we explore this enhancement mechanism for a variety of nanoparticle sizes and their relative placements by developing a fully analytical "coupled modes model" for plasmonic optical field enhancement in complex metal nanostructures. Using the model we show that whether the enhancement is achieved near the small feature (nanotip) or inside the nanogap, the enhancement is proportional to Q^m where m is the effective number of sequential coupling transitions occurring between the light being coupled into the structure and it being concentrated around or inside the smallest surface feature that serve as field-concentrating hot spots.[53] We further develop the optimization routine for maximum field enhancement which does not rely on extensive numerical computations and allows one to quickly analyze a number of different structures.

3.1. *Coupled mode theory*

We now develop a theory for treating the coupling of two closely spaced metal nanospheres whose SP modes are overlapping with each other. For a fair comparison, we obtain the field enhancement relative to that of a tightly focused Gaussian beam in the absence of metal spheres [Fig. 18(a)] as has been done for the isolated metal spheres in Section 2.2. The energy transfer process from external excitation to various SP modes of the coupled spheres is illustrated in Fig. 18(b). First, the dipole mode of each sphere acts as antenna allowing energy to be transferred in, it then gets coupled into higher order modes of both spheres through their coupling with the dipole mode of the different sphere.

The coupling energy between the two modes can be obtained as an integral of the electric potential $\Phi_{l_1}^{(1)}$ of the l_1-th mode of sphere 1 multiplied by the surface charge density $\sigma_{l_2}^{(2)}$ of the l_2-th mode of 2 evaluated over the surface of the sphere 2,

$$U_{l_1 l_2}^{(12)} = \oiint \Phi_{l_1}^{(1)} \sigma_{l_2}^{(2)} ds^{(2)} = -4k_{l_1 l_2}^{(12)} A_{l_1}^{(1)} A_{l_2}^{(2)} \tag{39}$$

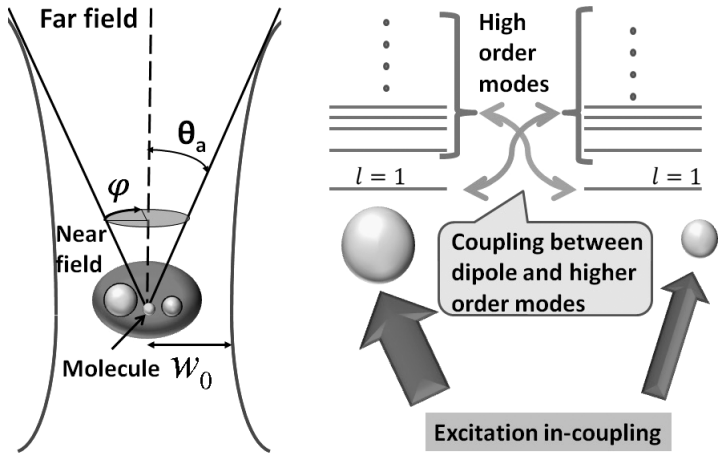


Fig. 18. Illustration of (a) the metal spheres placed at the apex of a focused Gaussian beam with a numerical aperture characterized by the far-field half angle θ_a and (b) the coupling of optical excitation into the dipole modes of both spheres and their subsequent coupling into the higher order modes.

where $A_l^{(i)}$ is the amplitude of the l-th mode in the i-th sphere evaluated as the square root of its energy $U_l^{(i)}$ as given in Eq. (8). Of all the coupling coefficients $k_{l_1 l_2}^{(12)}$ we are mostly interested in the coupling between the dipole mode ($l_1 = 1$) in one sphere and all the modes with index l in the other sphere $k_{1l}^{(12)}$ because they are the only ones associated with energy transfer. The coupling between higher order modes in two spheres only shifts the resonant frequencies of those modes by a small amount, typically smaller than broadening γ and can be neglected in this analysis.

Using Eq. (2) for the electric potential of one sphere and Eq. (7) for the surface charge density of the other, we can carry out the integral Eq. (39) analytically to obtain

$$\begin{aligned} k_{1l}^{(12)} &= \frac{l+1}{2}\left(\frac{a_1}{r_0}\right)^{3/2}\left(\frac{a_2}{r_0}\right)^{l+1/2} \\ k_{1l}^{(21)} &= \frac{l+1}{2}\left(\frac{a_2}{r_0}\right)^{3/2}\left(\frac{a_1}{r_0}\right)^{l+1/2} \end{aligned} \tag{40}$$

and between the dipole modes of two spheres, $k_{11}^{(12)} = k_{11}^{(21)} = k_{11} = (\frac{a_1 a_2}{r_0^2})^{3/2}$.

To determine the electric field in the gap between the particles we follow the energy as it couples first from the incoming light and into the

dipole mode of each spherical particle with the in-coupling coefficient $k_{\Omega}^{(1,2)}$ obeying Eq. (22) for each sphere. Taking into account of the energy transfer from the dipole mode of one sphere to all modes of the other, we can simply modify Eq. (24) to obtain the energy balance equations for the dipole modes of spheres 1 and 2 as

$$\begin{aligned}\frac{dA_1^{(1)}}{dt} &= j(\omega_1-\omega)A_1^{(1)} - j\sum_{l=1}^{\infty}\omega_{1l}k_{1l}^{(12)}A_l^{(2)} - \frac{1}{2}\gamma_1^{(1)}A_1^{(1)} + k_{\Omega}^{(1)}S_+ \\ \frac{dA_1^{(2)}}{dt} &= j(\omega_1-\omega)A_1^{(2)} - j\sum_{l=1}^{\infty}\omega_{1l}k_{1l}^{(21)}A_l^{(1)} - \frac{1}{2}\gamma_1^{(2)}A_1^{(2)} + k_{\Omega}^{(2)}S_+\end{aligned} \tag{41}$$

where $\omega_{1l} = \sqrt{\omega_1\omega_l}$. We have neglected the absorption loss of the dipole modes due to nearby absorbing molecules in Eq. (41), thus the dipole mode decays radiatively and nonradiatively $\gamma_1^{(i)} = \gamma_{\text{rad}}^{(i)} + \gamma, i = 1, 2$.

The subsequent transfer of energy from the dipole modes to higher order modes $l \geq 2$ can be treated similarly as

$$\begin{aligned}\frac{dA_l^{(1)}}{dt} &= j(\omega_l-\omega)A_l^{(1)} - j\omega_{1l}k_{1l}^{(21)}A_1^{(2)} - \frac{\gamma}{2}A_l^{(1)} \\ \frac{dA_l^{(2)}}{dt} &= j(\omega_l-\omega)A_l^{(2)} - j\omega_{1l}k_{1l}^{(12)}A_1^{(1)} - \frac{\gamma}{2}A_l^{(2)}\end{aligned}, \tag{42}$$

where the decay of these modes is only nonradiative with the rate γ. It should be pointed out that the coupling between the higher order modes of the two different spheres is neglected because the overlaps of these modes are small.

At steady state, Eq. (42) relates the electric field of the l-th higher order mode of sphere 2 to that of dipole mode of sphere 1 as

$$E_{\max,l}^{(2)} = \frac{\omega_{1l}}{(\omega_l-\omega)+j\frac{\gamma}{2}}\left(\frac{l+1}{2}\right)^2\left(\frac{a_1}{r_0}\right)^{3/2}\left(\frac{a_2}{r_0}\right)^{l+1/2}\left(\frac{a_1}{a_2}\right)^{3/2}E_{\max,1}^{(1)}. \tag{43}$$

A similar expression exists between $E_{\max,l}^{(1)}$ and $E_{\max,1}^{(2)}$. Let us take a quick look at Eq. (43) for two extreme cases. In the first case "symmetric dimer"[51] we consider two spheres of equal radii $a_2 = a_1 = a$, with negligibly small gap $r_0 \approx 2a$ and neglect the detuning relative to broadening which immediately

brings us to

$$|E^{(2)}_{\max,l}| < Q\frac{(l+1)^2}{2^{l+3}}|E^{(1)}_{\max,1}|, \quad Q = \frac{\omega}{\gamma}. \tag{44}$$

For the noble metals in the optical and near IR regions the value of Q ranges from 10 to 15 in case of Au and can be as high as 40 for bulk Ag, although in the nanoparticles actual Q is always lower due to the surface scattering.

If we define the cut-off mode as the one whose maximum field is equal to 1/2 of the field of the dipole mode, we obtain that for realistic Q's of less than 20 no more than 10 modes will get excited and once one takes detuning and gap into account that number will become even less. In the case of "extreme nanolens[52]" $a_2 \ll a_1$ only the larger sphere would act as an antenna and for all the modes in the smaller sphere

$$|E^{(2)}_{\max,l}| < Q\frac{(l+1)^2}{2}\left(\frac{a_2}{a_1}\right)^{l-1}|E^{(1)}_{\max,l}|. \tag{45}$$

From Eq. (45) it immediately follows that in the "extreme nanolens" regime only the $l = 1$ dipole mode gets excited since the term $(a_2/a_1)^{l-1}$ is essentially zero for all $l > 1$ modes, and the field enhancement relative to single sphere, in the limit of zero gap between the particles is about

$$\left|1 + \frac{E^{(2)}_{\max,1}}{E^{(1)}_{\max,1}}\right| \sim (1 + 4Q^2)^{1/2} \tag{46}$$

in general agreement with Ref. (52). For the symmetric dipole one can perform summation of Eq. (44) to obtain the maximum enhancement relative to single sphere with zero gap

$$2\left|1 + \frac{\sum_{l=2}^{\infty} E^{(2)}_{\max,l}}{E^{(1)}_{\max,1}}\right| \sim 2\left[1 + \left(\frac{9Q}{8}\right)^2\right]^{1/2} \tag{47}$$

where the factor of two in front comes from having two antennae and the factor of $9Q/8$ comes from all higher order modes of the other sphere, all added in phase. One cannot help referring to this phenomenon as "spatial mode-locking". So, from the most simple considerations the "symmetric dimer" and "extreme nanolens" can provide roughly the same field enhancement. For other cases, in order to optimize the field enhancement one should obtain solutions for arbitrary radii ratio and also

for finite gap size. Obviously for arbitrary radii ratio one must include all the modes in the calculation.

3.2. *Solution for the field enhancement*

We proceed as follows. The total electric field at the location $r_1 = r_0 - r_2$ in the gap (Fig. 17) is the summation of all modes from both spheres as

$$E(r_1) = E^{(1)}_{\max,1}\left[\left(\frac{a_1}{r_1}\right)^3 + \sum_{l=2}^{\infty}\frac{\omega_{1l}k^{(1)}_{1l}}{(\omega_l-\omega)+j\frac{\gamma}{2}}\frac{l+1}{2}\left(\frac{a_1}{r_2}\right)^{3/2}\left(\frac{a_2}{r_2}\right)^{l+1/2}\right] + E^{(2)}_{\max,1}\left[\left(\frac{a_2}{r_2}\right)^3 + \sum_{l=2}^{\infty}\frac{\omega_{1l}k^{(2)}_{1l}}{(\omega_l-\omega)+j\frac{\gamma}{2}}\frac{l+1}{2}\left(\frac{a_2}{r_1}\right)^{3/2}\left(\frac{a_1}{r_1}\right)^{l+1/2}\right] \tag{48}$$

where the first term is the combination of the dipole mode of sphere 1 and the higher order modes of sphere 2, and the energy of all these modes is coupled in through the $l = 1$ mode of sphere 1, and vice versa for the second term.

Combining Eqs. (41) and (42) at steady state and using Eqs. (8) and (18), we can relate

$$\begin{pmatrix} E^{(1)}_{\max,1} \\ E^{(2)}_{\max,1} \end{pmatrix} = \frac{\omega}{\sqrt{2}} M^{-1}_{2\times 2}\begin{pmatrix} E_{\text{foc}} \\ E_{\text{foc}} \end{pmatrix} \tag{49}$$

where the elements in the 2×2 matrix $M_{2\times 2}$ are

$$\begin{aligned} m_{11} &= j(\omega-\omega_1) + \sum_{l=2}^{\infty}\frac{\omega^2_{1l}[k^{(12)}_{1l}]^2}{j(\omega-\omega_l)+\frac{\gamma}{2}} + \frac{1}{2}\gamma^{(1)}_1 \\ m_{12} &= j\omega_1 k_{11}\left(\frac{a_2}{a_1}\right)^{3/2} = j\omega_1\left(\frac{a_2}{r_0}\right)^3 \\ m_{21} &= j\omega_1 k_{11}\left(\frac{a_1}{a_2}\right)^{3/2} = j\omega_1\left(\frac{a_1}{r_0}\right)^3 \\ m_{22} &= j(\omega-\omega_1) + \sum_{l=2}^{\infty}\frac{\omega^2_{1l}[k^{(21)}_{1l}]^2}{j(\omega-\omega_l)+\frac{\gamma}{2}} + \frac{1}{2}\gamma^{(2)}_1. \end{aligned} \tag{50}$$

Finally, substituting Eq. (49) into Eq. (48) in conjunction with Eq. (40), we arrive at the field enhancement factor which is defined as the ratio of

the electric field in the presence of the metal spheres to that of the focusing spot in the absence of the metal spheres

$$\begin{aligned} F &= \left|\frac{E(r_1)}{E_{\text{foc}}}\right| \\ &= \frac{\omega}{\sqrt{2}}\frac{1}{|M_{2\times 2}|} \\ &\quad \times \left|(m_{22}-m_{12})\left(\frac{a_1}{r_1}\right)^3\left[1+\frac{1}{4}\frac{a_2 r_1^3}{r_0^2 r_2^2}\sum_{l=2}^{\infty}\frac{\omega_{1l}(l+1)^2}{j(\omega_l-\omega)+j\frac{\gamma}{2}}\left(\frac{a_2^2}{r_0 r_2}\right)^l\right]\right. \\ &\quad \left. +\ (m_{11}-m_{21})\left(\frac{a_2}{r_2}\right)^3\left[1+\frac{1}{4}\frac{a_1 r_2^3}{r_0^2 r_1^2}\sum_{l=2}^{\infty}\frac{\omega_{1l}(l+1)^2}{j(\omega_l-\omega)+j\frac{\gamma}{2}}\left(\frac{a_1^2}{r_0 r_1}\right)^l\right]\right|. \end{aligned} \tag{51}$$

Note the presence of different phases in the denominators — this is a direct consequence of delay associated with the energy transfer from one nanoparticle to another, i.e., the retardation effect. Therefore, this quasi-electric-static model is valid for as long as the dimensions of each individual particle are small compared to the wavelength, while the total size of the system of nanoparticles can be of the order of wavelength and even larger. We shall now simplify Eq. (51) by examining the field enhancement at the mid gap position $r = r_0/2$ of two equal spheres $a = a_1 = a_2$ that are excited at the dipole mode frequency $\omega = \omega_1$. We use the fact that coupling coefficients are small, $[k_{1l}^{(ij)}] \approx 0$, $Q^{-1} \gg 2k_{11}$, and in the limit of very small metal spheres, $Q^{-1} \gg \chi^3$ where χ is the normalized radius defined in Eq. (12). Next, we realize that the terms from higher order modes ($l \geq 2$) in Eq. (51) are significant only for those lower indexes l whose frequency detuning from ω_1 is small, we thus approximate $Q^{-1} \gg 2(1-\omega_l/\omega_1)$. In the limit of zero gap, $r_0 \approx 2a$, we have

$$F \approx 2\sqrt{2}Q\left|1-j\frac{9Q}{8}\right| \approx \frac{9\sqrt{2}}{4}Q^2. \tag{52}$$

In comparison with the field enhancement by a single metal sphere which is proportional to Q, we now have additional contributions from higher order modes that have a relationship of Q^2. In the other case of "extreme nanolens" $a_2 \ll a_1$, when the field is focused in the vicinity of the smaller

sphere, Eq. (27) reduces to

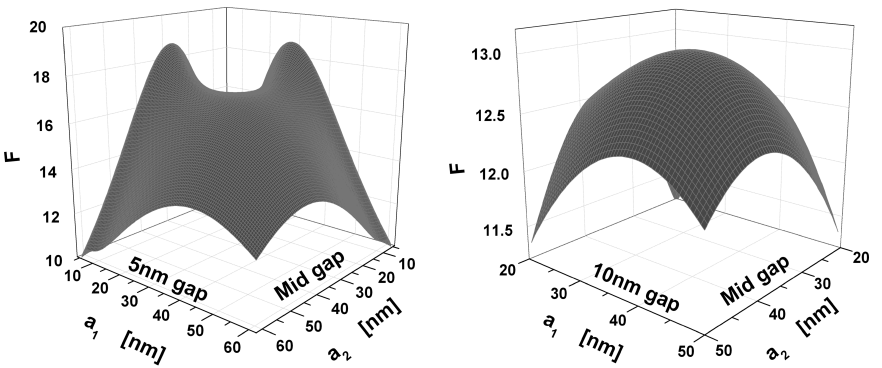
$$F \approx \sqrt{2}Q\left|1 - j2Q\left(\frac{a_1}{r_0}\right)^3\right|\left(\frac{a_2}{r_2}\right)^3. \tag{53}$$

Once again in the limit of zero gap, $r_2 \approx a_2$, and $r_0 \approx a_1$, we have $F \approx \sqrt{2}Q(1+4Q^2)^{1/2} \approx 2\sqrt{2}Q^2$. As we have already noted above, both extremes significantly overestimate the enhancement, and once detuning between the resonance frequencies of different modes and the presence of the gap between the particles are taken into account, the actual enhancement will be significantly less which is indeed the case as shown in the next section.

3.3. *Enhancement results and discussion*

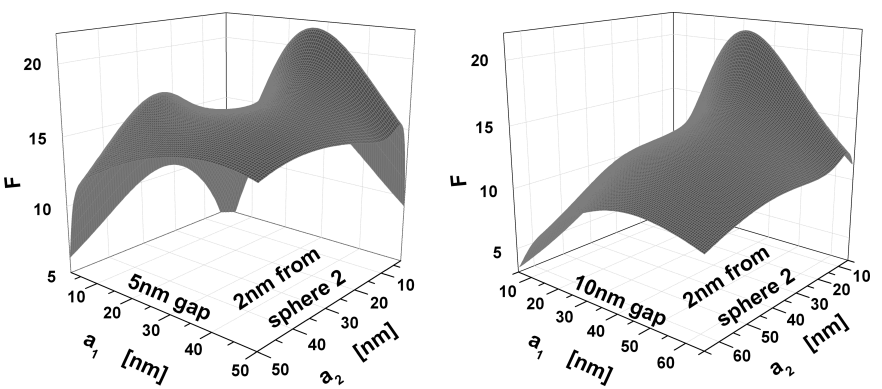
We have applied this model to Au metal spheres embedded in GaN dielectric with $Q \approx 10$ at the dipole frequency $\hbar\omega_1 = 1.967\,\text{eV}$ ($\lambda_1 = 630\,\text{nm}$).[41] While it is not difficult to evaluate the field enhancement anywhere in the near field of the two spheres, we shall present our results in the gap of the two spheres since that is where the strongest enhancement occurs. We first calculate the enhancement at the frequency of optical excitation in resonance with the dipole frequency $\omega = \omega_1$. For gaps less than 2 nm, the quantum effects such as electron tunneling and screening significantly reduce the enhancement,[59] we shall therefore limit our model to the coupled metal nanospheres with their separation gap greater than 2 nm. As has been demonstrated earlier for isolated single spheres, the enhancement has a strong dependence on the nanoparticle size, the enhancement in the coupled structure here shown in Fig. 19 also depends quite sensitively on the sizes of both spheres. The results in Fig. 19 are for enhancement at the mid gap between the two spheres with two different gaps. For smaller gap (5 nm), there are two symmetrical peaks in Fig. 19(a) indicating that the maximum enhancement is obtained with two unequal metal spheres, i.e., the "nanolens" case. In this situation, the larger sphere primarily acts as an antenna for energy to be coupled into the system while the smaller one behaves like a cavity for energy to be concentrated. As the gap increases, the two peaks merge into one as shown in Fig. 19(b) (10 nm gap) calling for spheres of equal size. This is easy to understand because at large distances the larger antenna sphere cannot effectively excite the smaller sphere and it is preferable to have both spheres of equal size.

Next we calculate the enhancement at the location that is fixed at 2 nm from sphere 2 with the radius a_2 as shown in Fig. 20. For smaller



(a) (b)

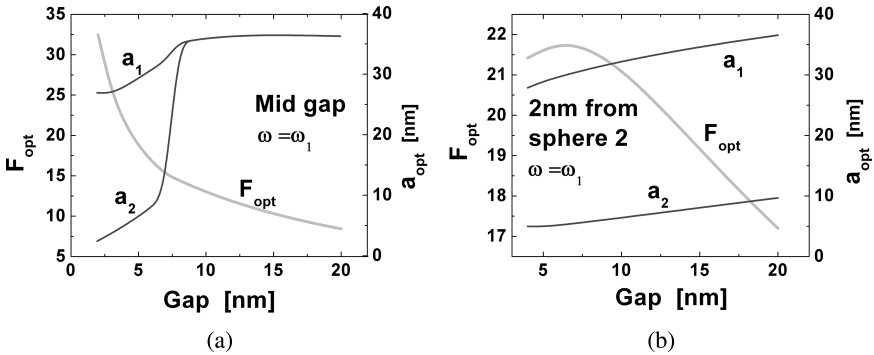
Fig. 19. Enhancement F at mid gap of the two Au spheres embedded in GaN as a function of their radii a_1 and a_2 with (a) 5 nm and (b) 10 nm gap.



(a) (b)

Fig. 20. Enhancement F at 2 nm from sphere 2 in the gap of the two Au spheres embedded in GaN as a function of their radii a_1 and a_2 with (a) 5 nm and (b) 10 nm gap.

gap of 5 nm, the location of 2 nm from sphere 2 is close to the mid gap, similarly for the result in Fig. 18(a), two peaks emerge but this time they are asymmetrical, but for larger gap of 10 nm, only one peak appears. In both cases, since the position of enhancement is closer to sphere 2, it consistently favors sphere 2 to be smaller for the field to be focused in its vicinity, i.e., the "nanolens" case.



(a) (b)

Fig. 21. Maximum field enhancement at (a) the mid gap and (b) 2 nm away from sphere 2 obtained by optimizing the radii of both spheres as a function of their separation gap.

We can now perform optimization of nanoparticle sizes to obtain peak enhancement F_{opt} at mid gap [Fig. 21(a)] and at 2 nm away from sphere 2 [Fig. 21(b)] for a range of gap. In comparison between the results obtained for the two cases, it can be said that in general the optimal enhancement is somewhat greater for the locations that are closer to one of the spheres than at mid gap for the same gap. Except at very small gaps, the optimum enhancement at mid gap is attained for two spheres whose dimensions are not far apart. The maximum enhancement is in the order of a factor of 20 for the gaps of about 5 nm and approaches 30 for the 2 nm gaps, while the Eqs. (52) and (53) predict enhancements as high as almost 300 for $Q \sim 10$ and zero gap. The discrepancy is easily explained by the presence of the term in Eq. (53)

$$\left(\frac{a_1}{r_0}\right)^3 \left(\frac{a_2}{r_2}\right)^3 \sim \left(\frac{a_1}{a_1 + a_2 + \delta}\right)^3 \left(\frac{a_2}{a_2 + \delta/2}\right)^3 \tag{54}$$

where δ is the size of the gap, and a similar term is in Eq. (52). When the relative size of the gap is rather small in comparison to the radii of the spheres, one can see the fact that both parentheses in Eq. (54) appear in cube causes significant reduction in the enhancement factor. For instance, when both parentheses in Eq. (54) equal to 0.7 the field enhancement factor gets reduced by tenfold.

Clearly, reduction in the gap size will increase the enhancement, but then it is not clear whether one can get enough "working space" with small gaps. Furthermore, we want to avoid tunneling between the nanospheres

that can occur when fields are too high, hence, as mentioned above we are not going to explore gaps that are less than 2 nm.

The above enhancement optimization at the dipole mode frequency $\omega = \omega_1$ can be further improved by optimizing the frequency. Indeed, the mode coupling can shift the resonance, and in case of strong coupling, it splits into two resonances which can be analyzed by examining the determinant of the 2×2 matrix $M_{2\times 2}$ in Eq. (50) that is in the denominator of Eq. (51). The two split frequencies at the two minima from the dipole resonance ω_1 can be obtained as

$$\frac{\omega}{\omega_1} \approx 1 \mp k_{11} \left(1 - \frac{1}{4k_{11}^2 Q^2}\right) \tag{55}$$

when coupling coefficient $k_{11} > 1/2Q$, i.e., the splitting must be greater than the broadening of the dipole mode. The splitting will be further shifted by the coupling with higher order modes. Figure 22 shows the frequency dependence of the enhancement at several gaps with optimized metal sphere radii given in Fig. 21, where the peak enhancement has clearly shifted towards lower frequency $\omega < \omega_1$. The curve for 5 nm gap in Fig. 22(a) and all those in Fig. 22(b) exhibit only one peak with no splitting because their optimized radii are all unequal ($a_1 \neq a_2$) (Fig. 21) which yield small coupling coefficient k_{11} in the range of $0.023 \sim 0.038$ less than $1/2Q \approx 0.05$. For the frequency dependence of 10 and 20 nm gaps in Fig. 22(a), the sizes of two spheres are optimized at the same radius as shown in Fig. 21 ($a_1 = a_2$), their coupling coefficients $k_{11} = 0.084$ for 10 nm gap and $k_{11} = 0.06$ for

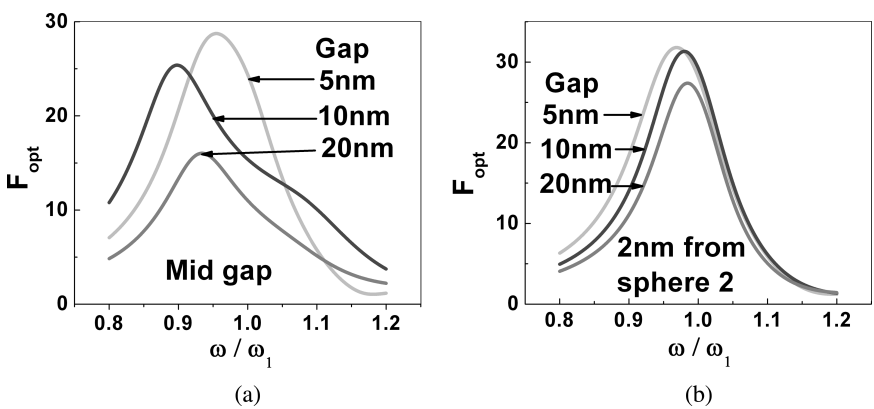


Fig. 22. Frequency dependence of the enhancement at (a) the mid gap and (b) 2 nm away from sphere 2 for a range of gap at optimized sphere radii given in Fig. 21.

20 nm gap are both greater than $1/2Q$, as a result, a shoulder on the higher frequency side $\omega > \omega_1$ can be resolved revealing the higher split of the dipole resonance. The amount of splitting depends on the coupling strength which is determined by the separation gap — the smaller the gap, the stronger the coupling, and thus the greater the splitting. As the gap increases, the enhancement decreases as a result of the reduced coupling between the spheres.

Now let us compare these results with those of single spheres. To have a fair comparison, we obtain optimal enhancement at the locations of equal separation from metal surface for both cases. This means that for a single sphere we are evaluating the field enhancement at a separation distance d from its surface that is equal to either half the gap or 2 nm, and we are neglecting the absorption decay rate γ_{abs}. The field enhancement factor for a single isolated sphere is related to the absorption enhancement in Eq. (26) at $\omega = \omega_1$ as,

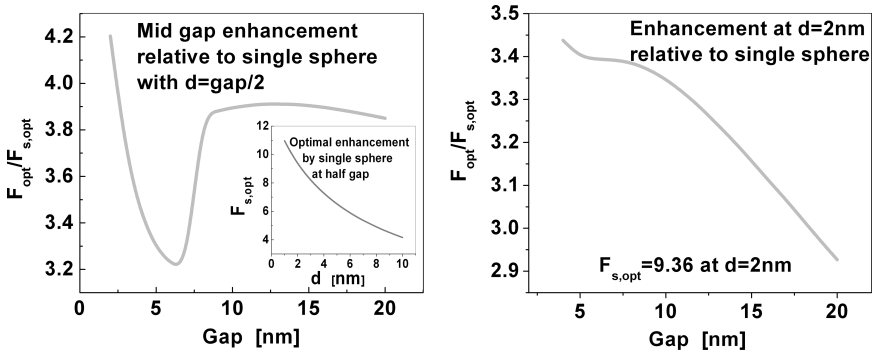
$$F_S(\omega_1) = \sqrt{F_A(\omega_1)} = \frac{\sqrt{2}Q}{1 + 2Q\chi^3/3\varepsilon_D}\left(\frac{\chi}{\chi + \chi_d}\right)^3 \tag{56}$$

which, at the optimized radius $\chi_{\text{opt}} = (3\varepsilon_D\chi_d/2Q)^{1/4}$, reaches its maximum value

$$F_{S,\text{opt}} = \frac{\sqrt{2}Q}{[1 + (2Q\chi_d^3/3\varepsilon_D)^{1/4}]^4}. \tag{57}$$

The optimal enhancement $F_{S,\text{opt}}$ for a single Au sphere embedded in GaN is shown in the insert of Fig. 23(a) for separations up to 10 nm corresponding to 20 nm gap which ensures the entire dimension of two spheres with gap remains smaller than a quarter of the wavelength. The ratio of $F_{\text{opt}}/F_{S,\text{opt}}$ versus the gap is shown in Fig. 23(a) for mid gap enhancement, and in Fig. 23(b) for the case of 2 nm separation from the sphere. It can be stated that the enhancement in the gaps of coupled spheres always outperforms that of single spheres. The improvement over single sphere is about a factor of $3 \sim 4$. This factor is substantially smaller than the factor of $2Q \sim 20$ obtained in Eqs. (46) and (47) in the limit of zero gap, but once again it can be explained by the strong cubic dependence of the field enhancement on the gap width.

For optical absorption and emission with properties directly proportional to the energy density, i.e., electric field squared (E^2), the improvement is roughly a factor of 10. For the surface enhanced Raman scattering (SERS) process whose intensity is proportional to E^4, an additional factor



(a) (b)

Fig. 23. The ratio of maximum field enhancement by the coupled spheres to that by a single sphere vs. the gap for the case of (a) mid gap and (b) 2 nm separation from one of the spheres. Insert in (a): maximum enhancement by a single sphere vs. the separation which is equal to half the gap between the two spheres.

of 10 can be recovered. Now, two orders of magnitude is substantial gain and thus using coupled nanoparticles is definitely worthwhile.

4. Implications

In this chapter, we have established an analytical model based on the effective-mode-volume method to predict the enhancement of optical absorption, EL and PL by molecules placed in the vicinity of isolated and coupled metal nanoparticles. The enhancement effect is the result of highly localized SP mode that is supported by the metal nanoparticles. The theory takes into account the radiative and nonradiative decays of the SP modes of various orders as well as the perturbation of field in the presence of active absorbing molecules. Using the example of Au nanospheres embedded in GaN dielectric, we show that the degree of enhancement for each case strongly depends on the size of the metal sphere enabling optimization for combinations of molecular properties such as absorption cross section, original radiative efficiency, and separation between the molecule and metal sphere. For EL and PL, because of the luminescence quenching effect induced by the high order modes, there exists an optimized molecule-metal sphere separation that maximizes the enhancement. The general conclusion is that the enhancement effect is the most significant for relatively weak and diluted absorbers and rather inefficient emitters that are placed in close proximity to the metal nanoparticles. Specifically for optical absorption,

metal nanoparticles can dramatically improve the performance of optical sensors in which the analyte molecules are few and their original absorption is low. At the same time, when the original absorption is already significant, as is the case of most optical detectors and photovoltaic devices, the enhancement is weak or nonexistent due to metal loss. For EL, we must conclude then that nanoparticles are excellent means of enhancing weak optical emission while their ability to improve performance of relatively efficient emitters, such as LEDs, is limited at best. The key conclusion for PL is that metal nanoparticles provide large enhancement only for small quantities of molecules with originally low PL cross section. Hence metal nanoparticles can be indispensable in improving sensors but are of limited use in other applications, where PL is already reasonably (a few percent) efficient.

Building upon the model established for the single isolated metal sphere,[40,46,55] we have developed a rigorous analytical approach to the field enhancement in complex systems of coupled metallic nanoparticles. The main conclusion for this part of research is the definite evidence that shows using systems of coupled nanoparticles allows one to achieve larger field enhancements than the ones attainable with single particle. The simple explanation of this effect is the fact that in order to achieve large enhancement one needs to have both an efficient antenna to interact with outside fields and a small effective volume. Single nanoparticles cannot possibly satisfy these two requirements, although a certain degree optimization is possible. But having more than a single nanoparticle immediately opens a possibility of using a large dipole mode as an efficient antenna and then transfer the energy into one or more tightly confined modes in which the high energy concentration gets achieved. Using an example of two coupled spherical nanoparticles we have shown that there are two ways the concentration can be achieved. In the case of symmetric dimer the superposition of quadrupole and higher order modes of both spheres has high energy concentration in the gap between the spheres and this combined "supermode" acts as a small cavity coupled to the dipole antenna. In the case of highly asymmetric "nanolens" the smaller particle acts a small cavity while the larger particle acts as a dipole antenna. Our theoretical analysis has shown that for both "dimer" and "nanolens" the electric-field enhancement of the order of Q^2 near the metal surface can be achieved versus Q in a simple particle.

With $Q \approx 10$ for Au in GaN, this enhancement would translate into 4 orders of magnitude enhancement of the absorption and up to 8 orders

of magnitude for luminescence and Raman scattering. This maximum enhancement is reduced, however, once the detuning between different modes, the finite size of the gap, and the distance from the metal surface are taken into account and the optimized field enhancement on the order of 30 appears to be a realistic maximum, which is larger than the enhancement attainable with the single sphere by a factor of about $3 \sim 4$ and can be translated to about ten-fold improvement for the processes of optical absorption and emission and about 100-fold for SERS.

Our analysis confirms the fact that more complex metallic nanostructures do offer significant advantage over the single nanoparticles and provides a simple "engineering" explanation in which the large enhancement is achieved in a smaller "cavity" mode that is coupled to a larger "antenna" mode. While the systems of two and even three nanoparticles has been previously studied numerically,[51,52] thus in this respect our analysis only offers better physical insight and a more straightforward way to optimization. The main contribution of our work, however, consists of developing an analytical method that can be easily applied to far more complex systems than dimers and trimers. Consider for example assembly of 5 spherical nanoparticles in a cluster shown in Fig. 24. Numerical analysis, let alone optimization of this cluster, requires a lot of effort. However, using our method all one needs to do is to set up a system of coupled linear equations, involving no more than a few modes per sphere which is far less daunting task than a full numerical solution especially in 3-dimensional case.

Furthermore, our method allows one to make a rough estimate of the field enhancement achievable in the nanocluster and its location without

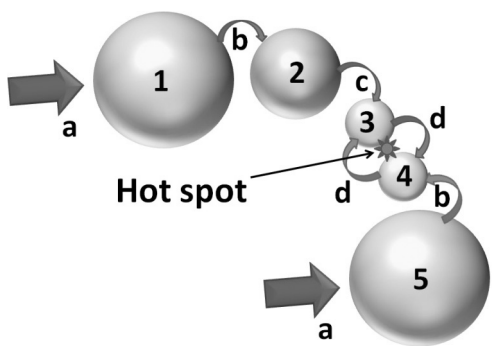


Fig. 24. Complex cluster of 5 spherical nanoparticles and the steps leading to the field enhancement in the hot spot inside the gap.

performing any numerical calculations. Let us follow the progress of field enhancement step by step in the system shown in Fig. 24. Clearly the large spheres 1 and 5 act as dipole antenna and in the first step (a) the incident light couples in the dipole modes of spheres 1 and 5. In the next step (b) the energy gets coupled into the smaller spheres 2 and 4, then, in the third step (c) the dipole mode of the sphere 3 gets excited from sphere 2. Finally, in the last step (d) the higher order modes of spheres 3 and 4 get excited by the dipole modes, just like in any dimer and a high field concentration is achieved in the hot-spot inside the gap between these spheres. Ideally, then given the four "enhancement steps" one can achieve the field enhancement of the order of Q^4, i.e., 10^4 but given our experience with dimers it is reasonable to expect maximum field enhancement in the range from $(Q/3)^4$ to maybe $(Q/2)^4$ i.e., 100–600. The exact value of enhancement depends on the exact radii and relative positions of nanoparticles, and thus should vary widely within the above boundaries. Given the fact that the number of steps is limited to no more than $3 \sim 4$ given the fact that the largest nanoparticles should be sub-wavelength and the smallest one cannot be much less than a few nanometers, the actual maximum enhancement of absorption of light is expected to be anywhere from 10^4 to 5×10^5 while for Raman scattering the electromagnetic enhancement can be anywhere between 10^8 and 10^{11} — in line with experimental data.[10–14] Note that it was first demonstrated in[52] that the enhancement does scale up as the power of Q in case of nanolens — but here we have extended this result to an arbitrary combination of nanospheres and gaps between them.

Thus, the coupled mode approach to the field enhancement by complexes of metal nanoparticles developed in this chapter, provides the scientific community with a powerful tool for understanding, estimating, analyzing, and optimizing the metal nanostructures for a wide variety of applications.

References

1. P. J. Schuck, D. P. Fromm, A. Sundaramurthy, G. S. Kino, and W. E. Moerner, *Phys. Rev. Lett.* **94**, 017402 (2005).
2. S. Kühn, U. Håkanson, L. Rogobete, and V. Sandoghdar, *Phys. Rev. Lett.* **97**, 017402 (2006).
3. R. Carminati, J.-J. Greffet, C. Henkel, and J. M. Vigoureux, *Opt. Comm.* **261**, 368 (2006).
4. L. Rogobete, H. Schniepp, V. Sandoghdar, and C. Henkel, *Opt. Lett.* **28**, 1736 (2003).

5. C. Girard, O. J. F. Martin, and A. Dereux, *Phys. Rev. Lett.* **75**, 3098 (1995).
6. M. Thomas, J.-J. Greffet, R. Carminati, and J. R. Arias-Gonzalez, *Appl. Phys. Lett.* **85**, 3863 (2004).
7. G. Baffou, C. Girard, E. Dujardin, G. C. des Francs, and O. J. F. Martin, *Phys. Rev. B* **77**, 121101(R) (2008).
8. G. C. des Francs, C. Girard, T. Laroche, G. Leveque, and O. J. F. Martin, *J. Chem. Phys.* **127**, 034701 (2007).
9. D. Maystre, in *Electromagnetic Surface Modes*, A. D. Boardman, ed. (Wiley, 1982), Chap. 17.
10. M. Moskovitz, *Rev. Mod. Phys.* **57**, 783 (1985).
11. M. Moskovits, L.-L. Tay, J. Yang, and T. Haslett, *Top. Appl. Phys.* **82**, 215 (2002).
12. K. Kneipp, Y. Wang, H. Kneipp, L. T. Perelman, I. Itzkan, R. R. Dasari, and M. S. Feld, *Phys. Rev. Lett.* **78**, 1667 (1997).
13. S. Nie and S. R. Emory, *Science* **275**, 1102 (1997).
14. A. M. Michaels, M. Nirmal, and L. E. Brus, *J. Am. Chem. Soc.* **121**, 9932 (1999).
15. M. Thomas, J.-J. Greffet, R. Carminati, and J. R. Arias-Gonzalez, *Appl. Phys. Lett.* **85**, 3863 (2004).
16. L. Novotny, *Appl. Phys. Lett.* **69**, 3806 (1996).
17. R. M. Bakker, V. P. Drachev, Z. Liu, H.-K. Yuan, R. H. Pedersen, A. Boltasseva, J. Chen, J. Irudayaraj, A. V. Kildishev, V. M. Shalaev, *New Journal of Physics*, **10**, 125022-1-16 (2008).
18. K. Okamoto, I. Niki, and A. Scherer, *Appl. Phys. Lett.* **87**, 071102 (2005).
19. S. Pillai, K. R. Catchpole, T. Trupke, and M. A. Green, *J. Appl. Phys.* **101**, 093105 (2007).
20. S. C. Lee, S. Krishna, and S. R. J. Brueck, *Opt. Express* **17**, 23160 (2009).
21. S. A. Maier, *Optics Express* **14**, 1957 (2006).
22. V. M. Shalaev, *Phys. Rep.* **272**, 61 (1996).
23. F. Brouers, S. Blacher, A. N. Lagarkov, A. K. Sarychev, P. Gadenne, and V. M. Shalaev, *Phys. Rev.* B **55**, 13234 (1997).
24. T. D. Corrigan, S.-H. Guo, R. J. Phaneuf, and H. Szmacinski, *J. Fluorescence* **15**, 777 (2005).
25. J.-H. Song, T. Atay, S. Shi, H. Urabe, and A. V. Nurmikko, *Nano Lett.* **5**, 1557 (2005).
26. P. Anger, P. Bharadwaj, and L. Novotny, *Phys. Rev. Lett.* **96**, 113002 (2006).
27. S. Kühn, U. Håkanson, L. Rogobete, and V. Sandoghdar, *Phys. Rev. Lett.* **97**, 017402 (2006).
28. J. S. Biteen, N. S. Lewis, H. A. Atwater, H. Mertens, and A. Polman, *Appl. Phys. Lett.* **88**, 131109 (2006).
29. Y. Chen, K. Munechika, and D. S. Ginger, *Nano Lett.* **7**, 690 (2007).
30. S. Kühn, G. Mori, M. Agio, and V. Sandoghdar, *Mol. Phys.* **106**, 893 (2008).
31. H. Mertens and A. Polman, *Appl. Phys. Lett.* **89**, 211107 (2006).
32. H. Mertens, J. S. Biteen, H. A. Atwater, and A. Polman, *Nano Lett.* **6**, 2622 (2006).

33. T. H. Taminiau, R. J. Moerland, F. B. Segerink, L. Kuipers, and N. F. van Hulst, *Nano Lett.* **7**, 28 (2007).
34. T. H. Taminiau, F. D. Stefani, F. B. Segerink, and N. F. van Hulst, *Nat. Photonics* **2**, 234 (2008).
35. J. B. Khurgin and G. Sun, *Appl. Phys. Lett.* **94**, 191106 (2009).
36. J. B. Khurgin, G. Sun, R. A. Soref, *Appl. Phys. Lett.* **94**, 101103 (2009).
37. Y. Zou, P. Steinvurzel, T. Yang, and K. B. Crozier, *Appl. Phys. Lett.* **94**, 171107 (2009).
38. T. Kalkbrenner, U. Håkanson, and V. Sandoghdar, *Nano Lett.* **4**, 2309 (2004).
39. O. Sqalli, I. Utke, P. Hoffmann, and F. Marquis-Weible, *J. Appl. Phys.* **92**, 1078 (2002).
40. G. Sun, J. B. Khurgin, and C. C. Yang, *Appl. Phys. Lett.* **95**, 171103 (2009).
41. P. B. Johnson and R. W. Christy, *Phys. Rev. B* **6**, 4370 (1972).
42. M. Westphalen, U. Kreibig, J. Rostalski, H. Luth, and D. Meissner, *Sol. Energy Mat. Sol. Cells* **61**, 97 (2000).
43. B. P. Rand, P. Peumans, and S. R. Forrest, *J. Appl. Phys.* **96**, 7519 (2004).
44. L. Tang, S. E. Kocabas, S. Latif, A. K. Okyay, D.-S. Ly-Gagnon, K. C. Saraswat, and D. A. B. Miller, *Nature Photonics* **2**, 226 (2008).
45. D. M. Schaadt, B. Feng, and E. T. Yu, *Appl. Phys. Lett.* **86**, 063106 (2005).
46. J. B. Khurgin and G. Sun, *J. Opt. Soc. Am.* B **26**, B83 (2009).
47. H. A. Haus, *Waves and Fields in Optoelectronics*, 1st ed. (Prentice-Hall, Englewood Cliffs, New Jersey, 1984).
48. M. A. Ali, J. Moghaddassi, and S. A. Ahmed, *J. Opt. Soc. Am.* B **8**, 1807 (1991).
49. S. W. Osborne, P. Bloos, P. M. Smowton, and Y. C. Xin, *J. Phys: Condens. Matter* **16**, S3749 (2004).
50. M. Purcell, *Phys. Rev.* **69**, 681 (1946).
51. P. Nordlander, C. Oubre, E. Prodan, K. Li, and M. I. Stockman, *Nano Lett.* **4**, 899 (2004).
52. K. Li, M. I. Stockman, and D. J. Bergman, *Phys. Rev. Lett.* **91**, 227402 (2003).
53. Y. Zou, P. Steinvurzel, T. Yang, and K. B. Crozier, *Appl. Phys. Lett.* **94**, 171107 (2009).
54. J. Zuloaga, E. Prodan, and P. Nordlander, *Nano Lett.* **9**, 887 (2009).
55. J. B. Khurgin, G. Sun, and R. A. Soref, *Appl. Phys. Lett.* **94**, 071103 (2009).

CHIRAL PHOTONIC AND PLASMONIC STRUCTURES

Kin Hung Fung*, Jeffrey Chi Wai Lee and Che Ting Chan†

Department of Physics,
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
**khfung@mit.edu*
†*phchan@ust.hk*

We discuss the wave propagation inside a chiral metallic/plasmonic medium defined by stacking and twisting an anisotropic medium in one direction. An anisotropic transfer matrix method for calculating the band structures and transmission properties of such a medium is described. The optical properties of some dielectric media that have the same kind of chiral structures are also discussed. When a bulk plasma dielectric function is introduced to one of the axis of the anisotropic medium, a W-shape band appears above the plasmonic bandgap. Such a chiral metallic/plasmonic medium supports both positive and negative refraction, and the criteria for achieving a negative refraction is addressed.

1. Introduction

In the past decade, negative refraction and negative index materials[1] and, in particular, their applications on imaging have attracted a lot of attention.[2–7] There are multiple ways to achieve negative refraction. One possible route is to achieve a negative refractive index using sub-wavelength resonant structures to realize simultaneously negative permittivity and

*Present Address: Massachusetts Institute of Technology, 77 Massachusetts Avenue 3-365, Cambridge, MA 02139, USA.

permeability.[8–10] Another route is to achieve negative refraction for one polarization by employing a chiral medium[11–25] with the magnitude of the chiral parameter (or optical activity) κ larger than $\sqrt{\varepsilon\mu}$,[12] where κ is defined in the constitutive relations:[26]

$$\begin{aligned} \overrightarrow{D} &= \varepsilon\overrightarrow{E} + i\kappa\sqrt{\varepsilon_0\mu_0}\overrightarrow{H} \\ B &= -i\kappa\sqrt{\varepsilon_0\mu_0}\overrightarrow{E} + \mu\overrightarrow{H} \end{aligned}, \tag{1}$$

ε (ε_0) and μ (μ_0) are the relative (vacuum) electric permittivity and magnetic permeability, respectively. In recent years, there has been a significant progress in the research of chiral negative refraction in terms of theoretical consideration and fabrication. A recent review with a lot of references can be found for example in Ref. (11). This section gives a brief introduction to some previously proposed ideas.

Natural materials do not have strong chiral electromagnetic properties. However, strong chirality can be induced by chiral resonators embedded in a host, or by dispersing dipole resonators into a chiral medium. In principle, a chiral system that has a resonance gap can potentially give negative refraction.[13] For example, negative refraction can be realized in a system with a non-zero κ together with an effective resonant response in ε, which can be written in the characteristic form:

$$\varepsilon(\omega) = 1 + \frac{A}{\omega_0^2 - \omega^2}. \tag{2}$$

If the structure is non-chiral, i.e., $\kappa = 0$, the band structure of the medium with constitutive relations given by Eqs. (1) and (2) shows a band gap in the frequency regime where $\varepsilon(\omega) < 0$ (i.e., $\omega_0 < \omega < \sqrt{A + \omega_0^2}$) [see Fig. 1(a)]. The dispersion in such non-chiral system is degenerate for the two circular polarizations. If $\kappa \neq 0$, the two circular polarizations are no longer degenerate. The chirality splits the degenerate U-shaped band into two bands, showing a W-shaped dispersion [see Fig. 1(b)]. The iso-frequency contour near the bottom of the W-shaped band of this chiral medium comprises two concentric circles (as the chiral medium is isotropic) and the radius of the inner circle decreases as frequency increases. This shrinking circle gives negative refraction.

In most of the previous theoretical studies, the chiral medium is assumed to be homogeneous and isotropic. In some man-made chiral media that are fabricated with techniques such as layer-by-layer top-down methods,[21,27–30] laser holographic lithography[31] or two photon writing,[32,33] the chiral medium is usually (bi-)anisotropic and the anisotropy may

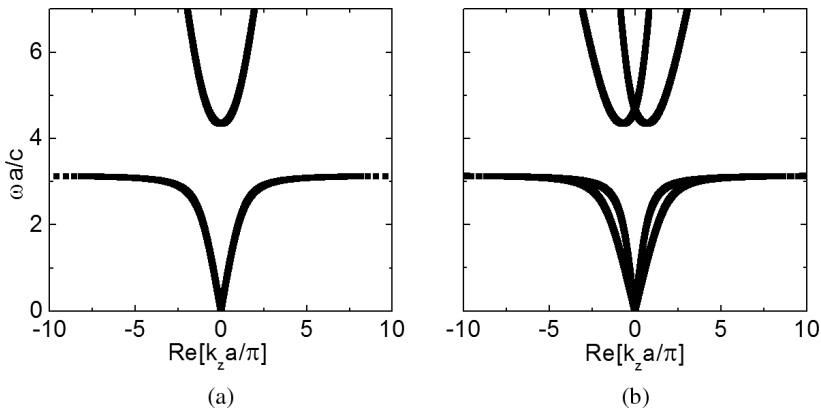


Fig. 1. (a) Band structure of a non-chiral medium with an electric resonance ($\kappa = 0$). Light does not propagate in the frequency regime $\omega_0 < \omega < \sqrt{A + \omega_0^2}$ as $\varepsilon(\omega)$ is negative in this regime. (b) Same as (a) except the medium is chiral ($\kappa \neq 0$). The chirality of the medium breaks the degeneracy of the two circular polarizations and gives rise to the W-shaped group of bands.

introduce complication in realizing the negative refraction. Also, many studies on the chirality based on using the empirical parameter κ in Eq. (1) did not explicitly specify the origins of the chirality.

In this chapter, we discuss the optical properties of chiral systems in which the chirality is derived directly from the structure of the system. Instead of assuming a phenomenological chiral parameter κ, we consider the wave propagation properties of a system in which the constituent material is non-chiral, but the structure has some "twisting" in one direction which gives rise to chirality. This brings the theoretical structure closer to the systems that can be made using the aforementioned fabrication technologies. We will also compare our systems with the isotropic chiral medium discussed above. As our system is neither homogeneous nor isotropic, its behavior is expected to differ somewhat from those described by Eq. (1) and (2). In Section 2, we give a brief review of a transfer matrix method that is known in the literature to be very useful for studying this class of systems, and in Section 3, we give numerical results for some specific chiral photonic and plasmonic crystals to illustrate their properties. In Section 4, we give analytical results that help us to understand the conditions required to achieve unusual dispersions in systems discussed in Section 3.

2. Transfer Matrix Method for Anisotropic Medium

We first give a brief description of the 4×4 transfer matrix method[34,36] that is useful for treating the chiral medium. The transmission spectrum as well as the band structure of a periodic chiral medium can be obtained by this method. For the sake of simplicity, we first consider a chiral medium that is homogeneous in two of the dimensions, i.e., in the x- and y-directions, and the constitutive relations depend on frequency (ω) and z only. We further assume that the incident wave propagates in the $x-z$ plane, i.e., there is no y component in the wavevector of the incident wave. The Maxwell's equations give

$$\frac{d}{dz}\overrightarrow{F}(z) = i\frac{\omega}{c}\overleftrightarrow{A}(z)\overrightarrow{F}(z), \tag{3}$$

where

$$\overrightarrow{F}(z) = \begin{pmatrix} E_x(z) \\ E_y(z) \\ Z_0 H_y(z) \\ -Z_0 H_x(z) \end{pmatrix}, \tag{4}$$

where $Z_0 = \sqrt{\mu_0/\varepsilon_0}$, and $E_x(z)$, $E_y(z)$, $H_x(z)$, and $H_y(z)$ are the x- and y-components of the electric and associated magnetic fields. The 4×4 matrix $\overleftrightarrow{A}(z)$ is given by

$$\overleftrightarrow{A}(z) = \begin{pmatrix} -\frac{\varepsilon_{31}}{\varepsilon_{33}}\frac{ck_x}{\omega} & \left(\frac{\mu_{23}}{\mu_{33}} - \frac{\varepsilon_{32}}{\varepsilon_{33}}\right)\frac{ck_x}{\omega} & \mu_{22} - \frac{\mu_{23}\mu_{32}}{\mu_{33}} - \frac{1}{\varepsilon_{33}}\left(\frac{ck_x}{\omega}\right)^2 & \frac{\mu_{23}\mu_{31}}{\mu_{33}} - \mu_{21} \\ 0 & -\frac{\mu_{13}}{\mu_{33}}\frac{ck_x}{\omega} & \frac{\mu_{13}\mu_{32}}{\mu_{33}} - \mu_{12} & \mu_{11} - \frac{\mu_{13}\mu_{31}}{\mu_{33}} \\ \varepsilon_{11} - \frac{\varepsilon_{13}\varepsilon_{31}}{\varepsilon_{33}} & \varepsilon_{12} - \frac{\varepsilon_{13}\varepsilon_{32}}{\varepsilon_{33}} & -\frac{\varepsilon_{13}}{\varepsilon_{33}}\frac{ck_x}{\omega} & 0 \\ \varepsilon_{21} - \frac{\varepsilon_{23}\varepsilon_{31}}{\varepsilon_{33}} & \varepsilon_{22} - \frac{\varepsilon_{23}\varepsilon_{32}}{\varepsilon_{33}} - \frac{1}{\mu_{33}}\left(\frac{ck_x}{\omega}\right)^2 & \left(\frac{\mu_{32}}{\mu_{33}} - \frac{\varepsilon_{23}}{\varepsilon_{33}}\right)\frac{ck_x}{\omega} & -\frac{\mu_{31}}{\mu_{33}}\frac{ck_x}{\omega} \end{pmatrix}, \tag{5}$$

where ε_{ij} and μ_{ij} are the components of the relative electric permittivity tensor,

$$\varepsilon(z,\omega) = \begin{pmatrix} \varepsilon_{11}(z,\omega) & \varepsilon_{12}(z,\omega) & \varepsilon_{13}(z,\omega) \\ \varepsilon_{21}(z,\omega) & \varepsilon_{22}(z,\omega) & \varepsilon_{23}(z,\omega) \\ \varepsilon_{31}(z,\omega) & \varepsilon_{32}(z,\omega) & \varepsilon_{33}(z,\omega) \end{pmatrix}, \tag{6}$$

and the relative magnetic permeability tensor,

$$\mu(z,\omega) = \begin{pmatrix} \mu_{11}(z,\omega) & \mu_{12}(z,\omega) & \mu_{13}(z,\omega) \\ \mu_{21}(z,\omega) & \mu_{22}(z,\omega) & \mu_{23}(z,\omega) \\ \mu_{31}(z,\omega) & \mu_{32}(z,\omega) & \mu_{33}(z,\omega) \end{pmatrix}, \tag{7}$$

and k_x is the component of wave vector parallel to the layer interface. Equation (3) can be solved with a transfer matrix method when the ε_{ij} and μ_{ij} are piecewise constant in z. If ε_{ij} and μ_{ij} are functions that change continuously, we can approximate each continuous function by a piecewise-constant function and we can take the intervals to be as small as we please. In that case, the transfer matrix method gives approximate solutions to Eq. (3) by assuming ε_{ij} and μ_{ij} to be piecewise constant in z.

For given piecewise constant functions $\varepsilon_{ij}(z)$ and $\mu_{ij}(z)$, the matrix $\overleftrightarrow{A}(z)$ is also piecewise constant. Suppose that the chiral medium occupies the interval (z_0, z_n) and the functions are constant at each of the subintervals (z_i, z_{i+1}), the solution to Eq. (3) in the subinterval (z_i, z_{i+1}) is

$$\overrightarrow{F}(z_{i+1}) = \exp\left[\frac{i\omega(z_{i+1}-z_i)}{c}\overleftrightarrow{A}_i\right]\overrightarrow{F}(z_i), \tag{8}$$

where $\overleftrightarrow{A}_i$ is the value of $\overleftrightarrow{A}(z)$ in such a subinterval and the solutions to Eq. (3) in the whole domain is

$$\overrightarrow{F}(z_n) = \left\{\prod_{i=0}^{n-1}\exp\left[\frac{i\omega(z_{i+1}-z_i)}{c}\overleftrightarrow{A}_i\right]\right\}\overrightarrow{F}(z_0). \tag{9}$$

Transmission spectrum is obtained by imposing the boundary conditions

$$\overrightarrow{F}(z_0) = a_1\begin{pmatrix}\cos\theta\\0\\1\\0\end{pmatrix} + a_2\begin{pmatrix}0\\1\\0\\\cos\theta\end{pmatrix} + r_1\begin{pmatrix}\cos\theta\\0\\-1\\0\end{pmatrix} + r_2\begin{pmatrix}0\\1\\0\\-\cos\theta\end{pmatrix}, \tag{10}$$

and

$$\overrightarrow{F}(z_n) = t_1 \begin{pmatrix} \cos\theta \\ 0 \\ 1 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \cos\theta \end{pmatrix}, \tag{11}$$

where a_1 and a_2 are, respectively, the coefficients of TM and TE components of the incident wave with incident angle θ. Here, r_i's and t_i's are four unknown complex numbers to be obtained by solving Eq. (9) together with Eqs. (10) and (11) and they are the TM and TE components of the reflected and transmitted waves, respectively.

Equation (9) can also be used for obtaining the band structure of a periodic system. The Bloch theorem imposes the condition

$$\overrightarrow{F}(a) = e^{ik_z a}\overrightarrow{F}(0) \tag{12}$$

for a system with period a. Combining with Eq. (9), it gives

$$[\overleftrightarrow{T}(k_x, \omega) - e^{ik_z a}]\overrightarrow{F}(0) = \overrightarrow{0}, \tag{13}$$

where $\overleftrightarrow{T}(k_x, \omega) = \left\{\prod_{i=0}^{n-1} \exp\left[\frac{i\omega(z_{i+1}-z_i)}{c}\overleftrightarrow{A}_i\right]\right\}$ and the dispersion is given by

$$\det(\overleftrightarrow{T}(k_x, \omega) - e^{ik_z a}\overleftrightarrow{I}) = 0. \tag{14}$$

Here, k_z are obtained for fixed values of k_x, ω and a.

3. Chiral Media from Discrete Screw Operations

Employing the transfer matrix formulation presented above, we are able to solve Maxwell's equations in a chiral medium defined by a twisting operation. Consider a chiral medium formed by a discrete screw operation which results in a discrete twisting medium. The simplest configuration of this kind of chiral medium is the one that is obtained by stacking a number of layers of anisotropic medium, each making an angle θ with the adjacent layer. Aligning coordinate axes with the principle axes of the anisotropic medium in the xy-plane, the relative permittivity and permeability tensors

of the medium is

$$\overleftrightarrow{\varepsilon}(z) = \overleftrightarrow{R}\left(\frac{\pi}{n}\left\lfloor\frac{nz}{a}\right\rfloor\right)\cdot\overleftrightarrow{\varepsilon_0}\cdot\overleftrightarrow{R}^{-1}\left(\frac{\pi}{n}\left\lfloor\frac{nz}{a}\right\rfloor\right) \tag{15}$$

$$\overleftrightarrow{\mu}(z) = \overleftrightarrow{R}\left(\frac{\pi}{n}\left\lfloor\frac{nz}{a}\right\rfloor\right)\cdot\overleftrightarrow{\varepsilon_0}\cdot\overleftrightarrow{R}^{-1}\left(\frac{\pi}{n}\left\lfloor\frac{nz}{a}\right\rfloor\right) \tag{16}$$

where $\overleftrightarrow{\varepsilon_0} = \begin{pmatrix}\varepsilon_1 & 0 & 0\\ 0 & \varepsilon_2 & 0\\ 0 & 0 & \varepsilon_3\end{pmatrix}$, $\overleftrightarrow{R}(\phi) = \begin{pmatrix}\cos\phi & -\sin\phi & 0\\ \sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{pmatrix}$, a is the period of the medium, and π/n is the angle of rotation as one goes from one layer to another. The function $\lfloor x\rfloor$ is defined as the largest integer smaller than x. The rotation matrix $\overleftrightarrow{R}(\phi)$ represents a right-handed rotation of the medium with an angle of ϕ, i.e., as one goes along the positive z-direction, the anisotropic medium is rotated in the counter-clockwise direction when we view it from the positive side of the z-axis. Therefore, the medium is chiral when $\varepsilon_1 \neq \varepsilon_2$. From now on, unless otherwise specified, we shall assume that the medium is non-magnetic, i.e., $\mu = 1$, and also $\varepsilon_3 = 1$.

Depending on the signs of ε_1 and ε_2, we consider two classes of chiral medium, namely the dielectric chiral medium and plasmonic chiral medium. In the dielectric chiral medium, the real parts of both ε_1 and ε_2 are positive while in a plasmonic chiral medium, one of ε_1 or ε_2 has a positive real part and the other has a negative real part. For the plasmonic chiral media, while it is possible for the real parts of both ε_1 and ε_2 to be negative, there is no propagating wave solution in the chiral medium and the configuration is less interesting and thus we do not discuss it here.

We first consider the dielectric chiral medium. The complex band structures along the z-direction of a typical dielectric chiral medium are shown in Figs. 2(a) and 2(b). At both the boundary and center of the Brillouin zone, there are ranges of frequency in which only one of the polarizations has a real solution of wavevector k. This means that only one of the polarizations has propagating solutions in these frequency ranges. Such band gaps are called the polarization gaps. From the calculated transmission spectra shown in Figs. 2(c) and 2(d), the polarizations of the modes shown in Fig. 2(a) are found to possess a strong circularly polarized component. The polarization gaps forbid the propagation of one of the circular polarizations, rendering the transmitted light to have a strong circularly polarized character. This chiral medium gives a circular polarization gap because $\varepsilon_1 \neq \varepsilon_2$. For a right-handed (RH) circularly polarized light with wavelength equals to twice of the period, there are two ways to align the electric field. The first way is to align its electric

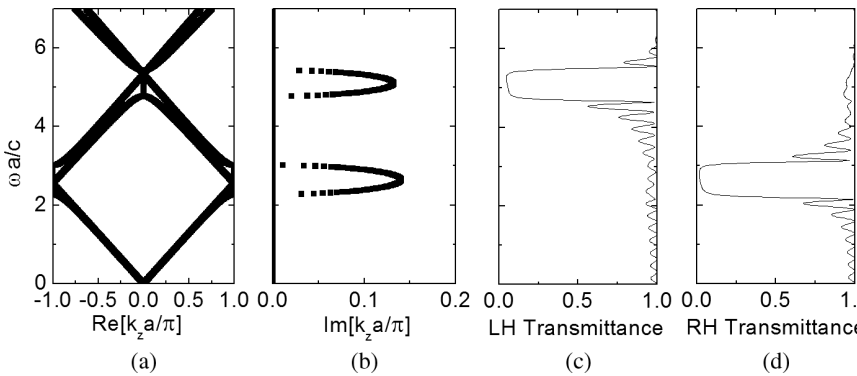
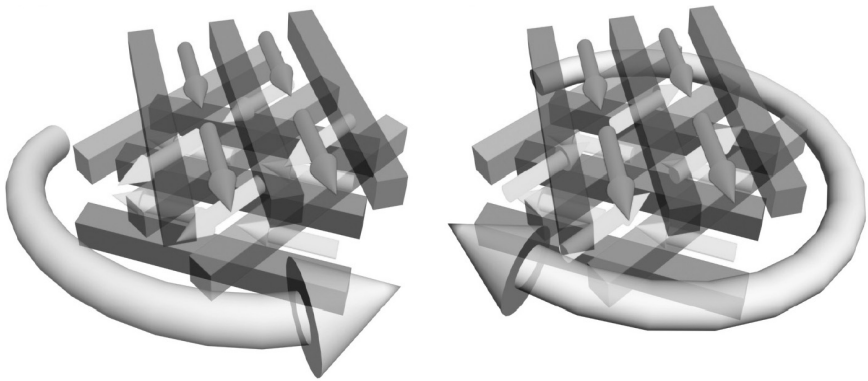


Fig. 2. (a) and (b) Complex band structure of a chiral medium with epsilon given by Eq. (15) in which $\varepsilon_1 = 1$, $\varepsilon_2 = 2$ and $n = 3$. (c) and (d) Corresponding transmission spectra for left-handed and right-handed circularly polarized light.

field parallel to the direction of the axes with ε_2. The other way is to align its electric field perpendicular to the first direction. In this case, the permittivity sampled by the propagating light is ε_1, which is smaller than ε_2 and results in an eigenmode of a higher frequency. Such a difference in frequency results in a polarization gap.

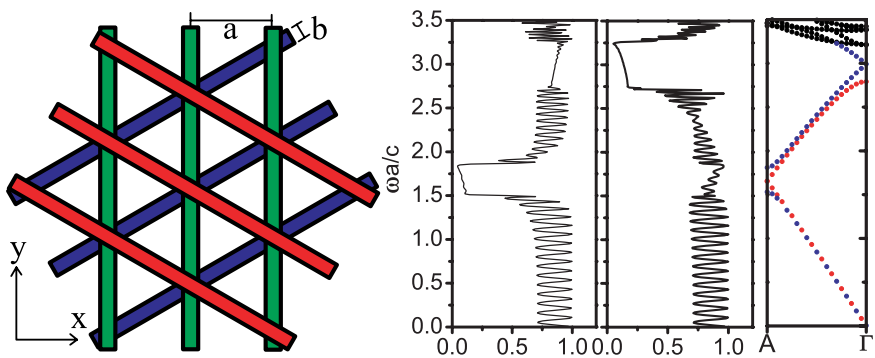
There are two polarization gaps shown in Fig. 2(a), one is for left-handed (LH) circularly polarized light and the other is for RH light. The frequency of the LH polarization gap is pinned at approximately twice that of the RH gap. This can be explained as follows. There are two ways for the electric field of the light to follow the discrete twisting of the anisotropic layers. The first one is for the electric field to follow the sense of twisting and twists at the same rate (60° per layer) as one moves up in the z-direction. As discussed previously, this gives rise to the RH polarization gap. The other way is to twist in the opposite sense but the electric field then has to twist twice as fast (120° per layer), which results in, approximately, one half of the former wavelength (i.e., double in frequency). In the case shown in Fig. 3, such ways of matching the field and the principle axis correspond, respectively, to a 60° twist in the anti-clockwise direction and a 120° twist in the clockwise direction. As a result, there are two polarization gaps for structures generated by discrete twisting operations, in which the high frequency gap pinned at double the frequency of the lower frequency gap.

This kind of chiral medium can be realized by using a kind of photonic crystal (PC) structure as in Fig. 4(a).[29,37] The structure of a chiral woodpile PC is the same as the structure of the rectangular rods shown in Fig. 3. The dielectric responses of the rods in the axial direction are different from



(a) (b)

Fig. 3. (a) Schematic diagram showing the electric field of a right-handed circularly polarized light that follows the rotation of the layers (indicated by the square rods). The circular arrows indicate the turning direction of the electric field as one goes upwards. (b) Same as (a) but for the case of left-handed circularly polarized light in the same structure.



(a) (b) (c) (d)

Fig. 4. (a) Top view of a chiral woodpile PC. The relative positions of the rods are marked by colors: white rods are on top of gray rods, while gray rods are on top of black rods. (b) Transmission spectra of RH circularly polarized light propagating in the z-direction. (c) Same as (b) but for LH circularly polarized light. (d) shows the band structure of the chiral woodpile PC along the z-direction. The dielectric constant of the rods is 9. b is the width of the woodpile. The period in the z-direction and the distance between rods in each layer are also a ($a = 2.6b$).

those in the transverse directions. Hence, the individual layers of the chiral woodpile PC constitute an anisotropic layer and the PC is chiral as there is a twisting in the z-direction. The calculated transmission spectra of RH and LH circularly polarized light of a chiral woodpile PC with 16 unit cells are shown in Figs. 4(b) and 4(c). The salient features of these transmission spectra are basically the same as those shown in Figs. 2(c) and 2(d), i.e., there are two polarization gaps for the two circular polarizations and thus the physics can be understood using the simple model as specified by Eq. (15). However, the frequency of the LH polarization gap is slightly less than twice of that of the RH polarization gap. Such a difference comes from some additional scattering effects due to the inhomogeneity in the xy-plane which does not exist in the chiral medium with constitutive parameters given by Eqs. (15) and (16).

If the angle between adjacent layers is 45° instead of 60°, the chiral medium repeats itself every four layers instead of three. The band structure and the transmission spectra of the $n = 4$ case are shown in Fig. 5. If we keep the period (lattice constant a along z-direction) of the medium unchanged, there is almost no change in the frequency of the first polarization gap. This is because from a geometrical point of view, the wavelength of the light in the first polarization gap roughly equals to twice of the lattice constant of the medium. As the period of the chiral medium is unchanged, the frequency of the first polarization gap depends mainly on the relative permittivity of the medium, which is more or less the same as the previous case. Therefore, the frequency of the first polarization gap is mainly pinned by the primitive period along the z-axis.

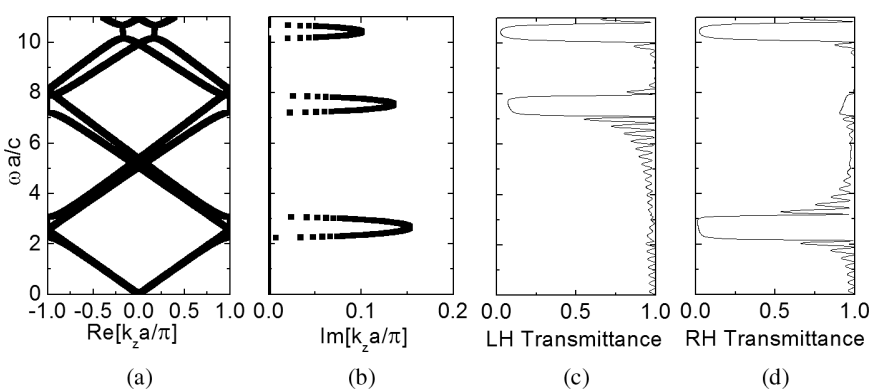


Fig. 5. Same as Fig. 2 except that $n = 4$.

The frequency of the second polarization gap in the $n = 4$ case is higher than that in the $n = 3$ case. In fact, in the $n = 4$ case, the frequency of the second polarization gap is nearly triple that of the first polarization gap. As discussed before, there are two ways for the electric field of the light to follow the twist of the layers. In this case, the "slower" way is to twist by 45° in the anti-clockwise direction as one moves from one layer to the adjacent layer, while the "faster" way is to twist by 135° in the clockwise direction. The ratio between the twisting rate of the two cases is 3 and, therefore, the frequency of the second polarization gap is approximately three times of that of the first polarization gap.

An interesting feature in all the transmission spectra shown in Figs. 2, 4, and 5 is the apparent suppression of Fabry-Perot effect inside the polarization gap, while Fabry-Perot effect induced oscillations are conspicuous in the pass bands. This is a special property of circular polarization gaps. The Fabry-Perot effect comes from the interference of light reflected from different boundaries of a medium. A slab of chiral medium has two surfaces: the entry surface where the incident light first hits the slab and the exit surface where the light leaves the slab. The suppression of Fabry-Perot in the chiral medium suggests that the exit surface is almost non-reflecting. This is due to the fact that circularly polarized light changes its circular polarization when it is reflected. Upon reflection from an isotropic surface, a LH circularly polarized light is reflected as a RH circularly polarized light and vice versa. Inside a polarization gap that allows only LH circularly polarized light to pass through, the light reflected by the exiting surface becomes RH circularly polarized and is forbidden from propagating back to the entry surface. Therefore, there is nothing to interfere with the light reflected by the entry surface. This explains the almost non-reflecting phenomena as long as the suppression of the Fabry-Perot effect in reflection and transmission spectra. In our chiral medium, the exit surface is not isotropic and the propagating mode in the polarization gap is not purely circularly polarized. Therefore, the suppression of the Fabry-Perot effect is not complete and there are still some small oscillations shown in the transmission spectra inside the gap, but the magnitude is much smaller than those in the pass bands in which both polarizations can propagate.

We now move on to consider plasmonic chiral media, in which ε is negative along one of the principle axes. A spiral PC made of thin metallic wire belongs to this kind of chiral medium, which has recently been realized in infrared frequencies.[33,35] As the metal wire is thin, current can only go along the tangential direction of the wire and the behavior is metallic along

this direction. For the normal and binormal directions (i.e., perpendicular to both tangential and normal directions), the system behaves effectively like a dielectric. However, the mathematical treatment of a metallic or plasmonic spiral is difficult. Although we can obtain numerical solutions using methods like the finite-difference-time-domain algorithm, the interpretation of the results would not be straightforward. Here, we make use of the numerical formulation discussed in Section 2 to study systems that have plasmonic responses. The advantage of using our model is that the physics is easier to interpret and it allows us to obtain analytical solutions if the twisting is continuous. The constitutive parameters of our model systems are also given by Eqs. (15) and (16) with one of the eigenvalues of $\ddot{\varepsilon}(z)$ taken to be negative. Without loss of generality, we choose ε_1 to be negative. We use a plasma model to model the dispersive features of the plasmonic response. The dielectric function ε_1 is taken to be

$$\varepsilon_1(\omega) = 1 - \frac{\omega_p^2}{\omega^2}. \tag{17}$$

We will compare the results obtained for different values of ω_p.

With a particular choice of ω_p, say $\frac{\omega_p a}{c} = 5$, $\varepsilon_2 = 1$ and $n = 3$, there is a low frequency gap and the bottom of the lowest pass band is W-shaped [see Fig. 6(a)], which can potentially give rise to a negative refraction. The band structure for a wider frequency regime is also shown in Fig. 6(b). The transmission spectra [Figs. 6(c) and 6(d)] show that the lowest pass band

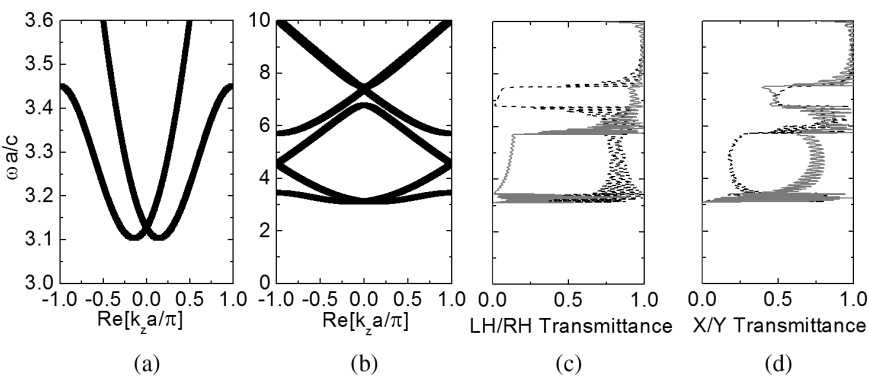


Fig. 6. (a) Band structure near the lower band edge of the lowest pass band of a metallic chiral medium with $\frac{\omega_p a}{c} = 5$. (b) Same as (a) but for a wider frequency range. (c) and (d) Transmission spectra for RH (solid gray curve) circular polarization, LH (black dashed curve) circular polarization, x-polarization (solid gray curve), and y-polarization (black dashed curve).

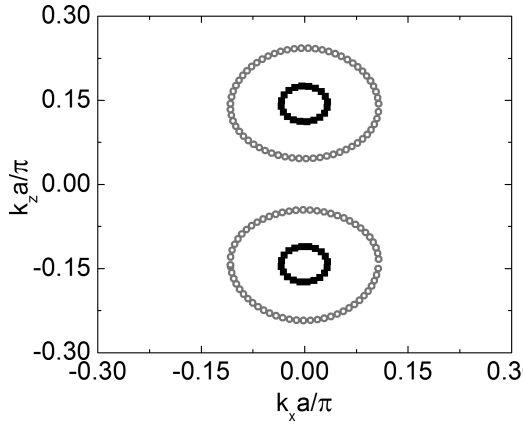


Fig. 7. Iso-frequency contours at $\frac{\omega a}{c} = 3.104$ (black solid squares) and $\frac{\omega a}{c} = 3.114$ (gray open circles). The group velocity of light in the chiral medium is given by $\nabla_k \omega$ and is easily seen from the iso-frequency contours.

is associated with LH circularly polarized incident wave. The closed-form solutions of such kind of twisting chiral medium are very complicated and difficult to analyze. Fortunately, it becomes relatively simple when the twisting is continuous ("spiral"), and these solutions can give us some insight to the plasmonic chiral system. More details are discussed in the next section.

Since our structure is not isotropic, the band structure in Fig. 6(a) cannot tell us the sign of refraction at oblique incidence. In order to understand the sign of refraction, we plot the iso-frequency contours near the bottom of the W-shaped band in Fig. 7. Now, we consider a light beam propagating in a direction parallel to the xz-plane. The iso-frequency contours, manifesting as two disjoint ellipses in Fig. 7, clearly show that negative refraction is not omni-directional in this system. If we use the xy-plane as the interface between the vacuum and the chiral medium, the x component of the wavevector, k_x, is conserved during refraction. The direction of group velocity obtained from the iso-frequency contours suggests that the sign of refraction of the medium is positive (since the refracted light propagates in the same direction as the group velocity, $\nabla_k \omega$). On the other hand, negative refraction can be achieved when the yz-plane is used as the interface.

When we lower the plasma frequency, say $\frac{\omega_p a}{c} = 3$, the W-shaped band has a redshift and its negative slope region becomes narrower, although

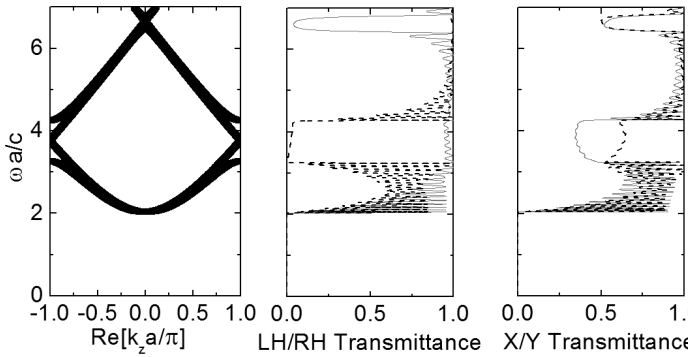


Fig. 8. (a) Same as Fig. 6(b) except that $\frac{\omega_p a}{c} = 3$. The "negative band" is much narrower. (b) Transmission spectra of LH light (solid gray curve) and RH light (black dashed curve) corresponding to the band structure shown in (a). (c) Same as (b) except that it is for x-polarized (solid gray curve) and y-polarized (black dashed curve) light. There is polarization gap for RH light in the frequency range between 3.3 and 4.25. This is due to the fact that in this frequency regime, the dielectric tensor is positive definite and the material itself behaves like a dielectric.

the bottom of the lowest pass band is still W-shaped (see Fig. 8) near the zone center. This is consistent with the results from the isotropic chiral system, which suggests that the W-shaped band comes from the plasmonic behavior of the medium. There is also a polarization gap for RH light in the frequency range between 3.3 and 4.25. The polarization gap comes from the fact that the material behaves like a dielectric when $\omega > \omega_p$ and the polarization gap is derived from such dielectric-like behavior.

On the other hand, if ω_p is large, say $\frac{\omega_p a}{c} = 10$, we enter a new regime. The band structure is no longer of W-shape [see Fig. 9(a)]. From the transmission spectra shown in Figs. 8(b) and 8(c), the lowest pass band of the chiral medium is y-polarized, i.e., light with electric field along the ε_2 direction at the entering surface is able to penetrate into the medium. These suggest that when the plasma frequency is high (which gives a large negative value of ε_1 at low frequency) the chiral medium does not give an effective chiral response. This is due to the fact that light polarized along the ε_1 direction cannot penetrate into the structure deeply when ε_1 is too negative. In this case, circularly polarized light can only see the surface of the chiral medium. Therefore, the chiral medium does not give an observable chiral response.

The iso-frequency contour of this chiral medium is shown in Fig. 10. The iso-frequency contours with this ω_p no longer take the shapes of two

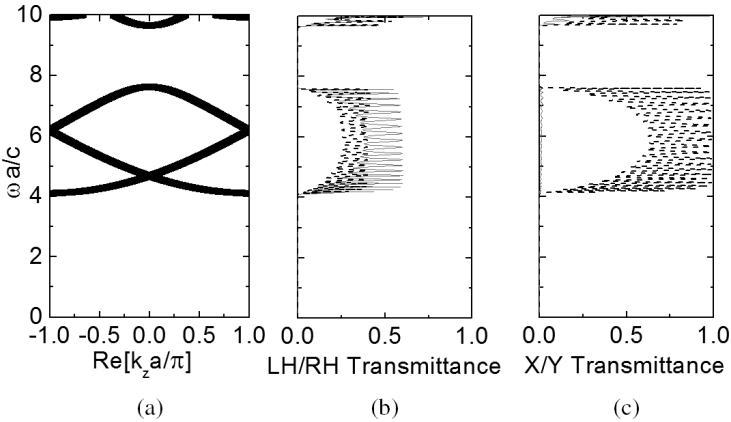


Fig. 9. Same as Fig. 7 except that $\frac{\omega_p a}{c} = 10$.

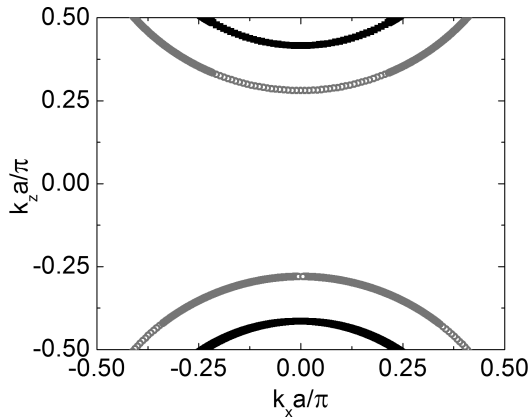


Fig. 10. Same as Fig. 7 except that $\frac{\omega_p a}{c} = 10$. $\frac{\omega a}{c} = 4.3$ (black solid squares) and $\frac{\omega a}{c} = 4.4$ (gray open circles). Compared to Fig. 7, the iso-frequency contours no longer take the shapes of ellipses but the medium can also support negative refraction.

separated ellipses. If we follow the same analysis that we did for the case with $\frac{\omega_p a}{c} = 5$, negative refraction can also be seen in this medium. To achieve negative refraction in this medium, we have to use the yz-plane as the interface.

We end this section with discussions on the effect of increasing n. The continuous twisting chiral medium is the limiting case for $n \to \infty$. Hence, we expect that as n increases, the behavior of the system converges to that

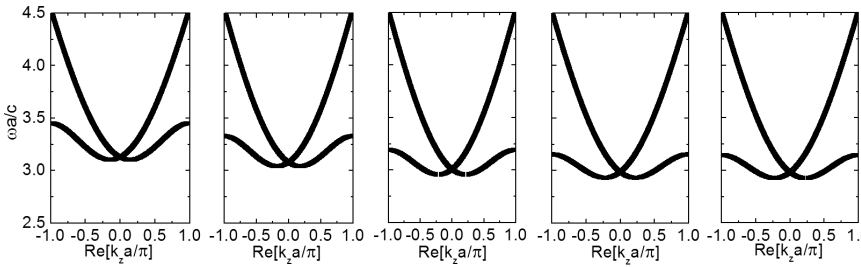


Fig. 11. Same as Fig. 6(a) except for different values of n. (a) $n = 3$. (b) $n = 4$. (c) $n = 8$. (d) $n = 16$. (e) $n = 32$. This figure shows the evolution of the bands as n increases. The frequencies of the bands decrease with increasing n. The bottom of the band converges to a value of 2.923. To illustrate the rate of convergence, it should be noted that the bottom of the bands are at 2.932, 2.926 and 2.924 for $n = 16$, 32 and 64, respectively.

of the continuous twisting case. We again consider the case of $\frac{\omega_p a}{c} = 5$ in which the W-shaped band can be seen clearly. The band structures for various values of n are shown in Fig. 11. As n increases, the bands shift to lower frequencies. The bottom of the band converges to a value of $\frac{\omega a}{c} = \frac{\omega_p a}{2c}\sqrt{2 - \left(\frac{\omega_p a}{2\pi c}\right)^2} \approx 2.923$. This expression will be derived in Section 4.

4. Chiral Media from Continuous Screw Operation

In this section, we will focus on the chiral media generated by continuous screw operations. Solutions for these continuously twisting chiral media can be obtained analytically for $k_x = 0$ and these solutions give us insight to the phenomenon inside the chiral medium. It also helps us understand more about those chiral media obtained from discrete screw operations.

The dielectric tensor of a continuous twisting chiral medium is given by

$$\overleftrightarrow{\varepsilon}(z) = \overleftrightarrow{R}\left(\frac{\pi z}{a}\right) \cdot \overleftrightarrow{\varepsilon_0} \cdot \overleftarrow{R^{-1}}\left(\frac{\pi z}{a}\right) \tag{18}$$

This medium is a right-handed medium and the corresponding $\overleftrightarrow{A}(z)$ defined in Eq. (5) is

$$\overleftrightarrow{A}(z) = \overleftrightarrow{R_4}\left(\frac{\pi z}{a}\right) \overleftrightarrow{A_0} \overleftarrow{R_4^{-1}}\left(\frac{\pi z}{a}\right), \tag{19}$$

where

$$\overleftrightarrow{A_0} = \begin{pmatrix} 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & \mu \\ \varepsilon_1 & 0 & 0 & 0 \\ 0 & \varepsilon_2 & 0 & 0 \end{pmatrix}, \tag{20}$$

$$\overleftrightarrow{R_4}(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 & 0 \\ \sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & \cos\phi & -\sin\phi \\ 0 & 0 & \sin\phi & \cos\phi \end{pmatrix}, \tag{21}$$

with $\overleftrightarrow{A}(z)$ given by Eq. (19) can be solved by a rotation of coordinates. By substituting $\overleftrightarrow{F}(z) = \overleftrightarrow{R_4}(\frac{\pi z}{a})\overleftrightarrow{G}(z)$ into Eq. (3), we get

$$\frac{d\overleftrightarrow{G}(z)}{dz} = i\frac{\omega}{c}\overleftrightarrow{A_0}\overleftrightarrow{G}(z) - \overleftrightarrow{R}_4^{-1}\left(\frac{\pi z}{a}\right)\frac{d}{dz}\left(\overleftrightarrow{R_4}\left(\frac{\pi z}{a}\right)\right)\overleftrightarrow{G}(z). \tag{22}$$

Using the fact that

$$-\overleftrightarrow{R_4^{-1}}\left(\frac{\pi z}{a}\right)\frac{d}{dz}\left(\overleftrightarrow{R_4}\left(\frac{\pi z}{a}\right)\right) = \frac{\pi}{a}\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}. \tag{23}$$

for right-handed media, Eq. (22) can be simplified as[38,39]

$$\frac{d}{dz}(\overleftrightarrow{G}(z)) = \overleftrightarrow{B_0}\overleftrightarrow{G}(z), \tag{24}$$

where

$$\overleftrightarrow{B_0} = \begin{pmatrix} 0 & \frac{\pi}{a} & \frac{i\omega\mu}{c} & 0 \\ -\frac{\pi}{a} & 0 & 0 & \frac{i\omega\mu}{c} \\ \frac{i\omega\varepsilon_1}{c} & 0 & 0 & \frac{\pi}{a} \\ 0 & \frac{i\omega\varepsilon_2}{c} & -\frac{\pi}{a} & 0 \end{pmatrix}. \tag{25}$$

It should be noted that if the medium is left-handed, the corresponding matrix in Eq. (23) should have been transposed. Solutions to Eq. (24) take

the form

$$\overrightarrow{G}(z) = \exp(\overleftrightarrow{B_0} z)\overrightarrow{G}(0) \tag{26}$$

and the analytical solutions can be obtained by diagonalizing the matrix $\overleftrightarrow{B_0}$. The properties of the eigenmodes are determined by the eigenvalues of the matrix $\overleftrightarrow{B_0}$. If a particular eigenvalue of the matrix $\overleftrightarrow{B_0}$ is purely imaginary, the corresponding eigenmode is a propagating wave. Otherwise, it is an evanescent wave.

Assuming $\mu = 1$, the dispersion relation given by Eq. (14) can be written as[40]

$$k'^4 - (2 + \omega'^2\varepsilon_1 + \omega'^2\varepsilon_2)k'^2 + (\omega'^2\varepsilon_1 - 1)(\omega'^2\varepsilon_2 - 1) = 0, \tag{27}$$

where $k' = \frac{ka}{\pi}$, $\omega' = \frac{\omega a}{\pi c}$, $k = \frac{\lambda}{i}$ and λ is the eigenvalue of the matrix $\overleftrightarrow{B_0}$. In order to have a W-shaped band, there must be four propagating solutions. The necessary condition is

$$(2 + \omega'^2\varepsilon_1 + \omega'^2\varepsilon_2)^2 - 4(\omega'^2\varepsilon_1 - 1)(\omega'^2\varepsilon_2 - 1) \geq 0 \tag{28}$$

which gives

$$\omega'^2[\omega'^2(\varepsilon_1 - \varepsilon_2)^2 + 8(\varepsilon_1 + \varepsilon_2)] \geq 0 \tag{29}$$

Since $\omega' \geq 0$, Eq. (29) has a restriction on the parameters only when

$$\varepsilon_2 + \varepsilon_1 < 0 \quad \text{and} \quad \omega'^2 \geq \frac{-8(\varepsilon_2 + \varepsilon_1)}{(\varepsilon_2 - \varepsilon_1)^2}. \tag{30}$$

With the condition given by Eq. (30) satisfied, Eq. (27) can be factorized into the product of two quadratic polynomials. As the polynomial in Eq. (27) does not contain the odd power terms of k, the factorized Eq. (27) is in the form of

$$(k^2 - A^2)(k^2 - B^2), \tag{31}$$

with A^2 and B^2 to be real numbers. The solutions are real when both A^2 and B^2 are positive and the requirement is

$$2 + \omega'^2\varepsilon_1 + \omega'^2\varepsilon_2 \geq 0 \tag{32}$$

and

$$(\omega'^2\varepsilon_1 - 1)(\omega'^2\varepsilon_2 - 1) \geq 0. \tag{33}$$

Equation (32) can be expressed as

$$\frac{1}{\omega'^2} \geq -\frac{\varepsilon_1 + \varepsilon_2}{2}, \tag{34}$$

and Eq. (33) is equivalent to

$$\frac{1}{\omega'^2} \leq \min(\varepsilon_1, \varepsilon_2) \quad \text{or} \quad \frac{1}{\omega'^2} \geq \max(\varepsilon_1, \varepsilon_2). \tag{35}$$

For simplicity, let us assume $\varepsilon_1 \leq \varepsilon_2$ (which includes the case of $\varepsilon_1 = 1 - \frac{\omega_p^2}{\omega^2}$ and $\varepsilon_2 = 1$). Then, Eq. (34) becomes

$$\frac{1}{\omega'^2} \leq \varepsilon_1 \quad \text{or} \quad \frac{1}{\omega'^2} \geq \varepsilon_2. \tag{36}$$

To keep the following analysis simple, we only focus on ε_1 and ε_2 that are either constant or strictly increasing functions of ω. Let us first consider positive values of ε_1 and ε_2, which corresponds to a chiral photonic crystal. In this case, only Eq. (36) is useful and it becomes

$$\omega' \leq \frac{1}{\sqrt{\varepsilon_2}} \quad \text{or} \quad \omega' \geq \frac{1}{\sqrt{\varepsilon_1}}, \tag{37}$$

which is consistent with the fact that, when ε_1 and ε_2 are constant, there is a polarization bandgap (in the range $\frac{1}{\sqrt{\varepsilon_2}} \leq \omega' \leq \frac{1}{\sqrt{\varepsilon_1}}$) which cannot support propagating wave of one kind of circular polarizations.[40] This also explains the polarization bandgap above the W-band.

In the frequency range where ε_1 is negative but $\varepsilon_1 > -\varepsilon_2$, Eqs. (29) and (34) still have no restriction to ω' and the only requirement to have four solutions is Eq. (36). Such requirement is $\omega' \leq \frac{1}{\sqrt{\varepsilon_2}}$ for $\varepsilon_2 > 0$ [as Eq. (36) has no restriction on ω' for $\varepsilon_2 < 0$]. This explains the zone-boundary frequency ($\omega' = \frac{1}{\sqrt{\varepsilon_2}}$) of W-band. It suggests that the modes in the W-shaped near to the zone boundary follow the twisting of the chiral medium with their electric field aligning along the ε_2 direction. This can be easily verified with the eigenvector of matrix shown in Eq. (25).

Now, we consider the frequency range where $\varepsilon_1 < -\varepsilon_2$. Again, Eq. (36) gives the upper bound for ω' and the restriction is at the former frequency range. In addition, we have restrictions given by Eqs. (30) and (34). The exact requirement is generally complicated but, for the case of $\varepsilon_1 = 1 - \frac{\omega_p'^2}{\omega'^2}$

and $\varepsilon_2 = 1$ (where $\omega'_p = \frac{\omega_p a}{\pi c}$), Eqs. (30) and (34) can be simplified as

$$\omega'^2 \geq \frac{\omega'^2_p}{2} - 1 \text{ and } \omega'^2 \geq \left(\frac{\omega'_p}{2}\right)^2 \left[2 - \left(\frac{\omega'_p}{2}\right)^2\right], \tag{38}$$

respectively. To have the W-shaped band, the intersection between Eqs. (38) and (30) must not be an empty set. The condition is

$$\omega'_p < 2, \tag{39}$$

which corresponds to $\varepsilon_1 < -3$. In this case, the lower bound of the W-shaped band is given by

$$\omega'^2 \geq \omega'^2_L \equiv \left(\frac{\omega'_p}{2}\right)^2 \left[2 - \left(\frac{\omega'_p}{2}\right)^2\right], \tag{40}$$

i.e.,

$$\omega \geq \omega_L \equiv \frac{\omega_p}{2}\sqrt{2 - \left(\frac{\omega_p a}{2\pi c}\right)^2}, \tag{41}$$

which agrees with the numerical results in Section 3.

5. Conclusions

In this chapter, we briefly reviewed the idea of using chiral media to achieve negative refraction and studied some interesting band structure and propagation properties of chiral photonic and plasmonic structures. We introduced a chiral photonic crystal (defined by a twisting operation on anisotropic medium) that can suppress the Fabry-Perot effect. We have also shown that negative refraction can be achieved in a metallic/plasmonic chiral medium defined by the same twisting operation. Such chiral media supporting negative refraction can also be formed by continuous twisting operations. However, the sign of refraction depends on the entry interface of the chiral medium. Also, if the dielectric tensor is too negative in one particular direction of the plasmonic chiral media, the chiral response is weak in the medium even though negative refraction is still possible. Analytical solutions are obtained for the continuously twisting chiral medium and are used to understand the band structures and the conclusion can be extended to the discrete twisting chiral media.

Acknowledgments

This work was supported by Hong Kong RGC grant 600308.

References

1. K. M. Krowne and Y. Zhang (Eds.), *Physics of Negative Refraction and Negative Index Materials* (Springer-Verlag, New York, 2007).
2. V. G. Veselago, *Sov. Phys. Usp.* **10**, 509 (1968).
3. J. B. Pendry, *Phys. Rev. Lett.* **85**, 3966 (2000).
4. G. W. 't Hooft, *Phys. Rev. Lett.* **87**, 249701 (2001).
5. J. B. Pendry, *Phys. Rev. Lett.* **87**, 249702 (2001).
6. N. Fang, H. Lee, C. Sun, and X. Zhang, *Science* **308**, 534–537 (2005).
7. R. A. Shelby, D. R. Smith, and S. Schultz, *Science* **292**, 77 (2001).
8. D. R. Smith and D. Schurig, *Phys. Rev. Lett.* **90**, 077405 (2003).
9. R. A. Shelby, D. R. Smith, S. C. Nemat-Nasser, and S. Schultz, *Appl. Phys. Lett.* **78**, 489 (2001).
10. D. R. Smith, W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, *Phys. Rev. Lett.* **84**, 4184 (2000).
11. For a recent review, see T. G. Mackay, A. Lakhtakia. *SPIE Reviews* **1**, 018003 (2010).
12. S. Tretyakov, I. Nefedov, A. Sihvola, S. Maslovski, and C. Simovsko, *J. Electromagn. Waves Appl.* **17**, 695–706 (2003).
13. J. B. Pendry, *Science* **306**, 1353 (2004).
14. T. G. Mackay and A. Lakhtakia, *Phys. Rev. E* **69**, 026602 (2004).
15. Y. Jin and S. He, *Opt. Express* **13**, 4974 (2005).
16. C. Monzon and D. W. Forester, *Phys. Rev. Lett.* **95**, 123904 (2005).
17. V. Yannopapas, *J. Phys.: Condens. Matter* **18**, 6883 (2006).
18. Q. Cheng and T. J. Cui, *Phys. Rev. B* **73**, 113104 (2006).
19. H. Chen, K. H. Fung, H. Ma, and C. T. Chan, *Phys. Rev. B* **77**, 224304 (2008).
20. K. H. Fung, J. C. W. Lee, and C. T. Chan, *arXiv*: 0811.1438 (2008).
21. S. Zhang, Y. S. Park, J. S. Li, X. C. Lu, W. L. Zhang, and X. Zhang, *Phys. Rev. Lett.* **102**, 023901 (2009).
22. E. Plum, J. Zhou, J. Dong, V. A. Fedotov, T. Koschny, C. M. Soukoulis, and N. I. Zheludev, *Phys. Rev. B* **79**, 035407 (2009).
23. J. F. Zhou, J. F. Dong, B. N. Wang, T. Koschny, M. Kafesaki, and C. M. Soukoulis, *Phys. Rev. B* **79**, 121104 (2009).
24. G. Mackay and A. Lakhtakia, *Phys. Rev. B* **79**, 235121 (2009).
25. B. Wang, J. Zhou, T. Koschny, and C. M. Soukoulis, *Appl. Phys. Lett.* **94**, 151112 (2009).
26. I. V. Lindell, A. H. Sihvola, S. A. Tretyakov, and A. J. Viitanen, *Electromagnetic Waves in Chiral and Bi-isotropic Media* (Artech House, Boston, 1994).
27. J. H. Lee, J. C. W. Lee, W. Leung, M. Li, K. Constant, C. T. Chan, and K. M. Ho, *Adv. Mater.* **20**, 3244 (2008).

28. A. V. Rogacheva, V. A. Fedotov, A. S. Schwanecke, and N. I. Zheludev, *Phys. Rev. Lett.* **97**, 177401 (2006).
29. M. Thiel, G. Freymann, and M. Wegener, *Opt. Lett.* **32**, 2547 (2007).
30. N. Kanda, K. Konishi, and M. Kuwata-Gonokami, *Opt. Express* **15**, 11117 (2007).
31. Y. K. Pang, J. C. W. Lee, H. Lee, W. Y. Tam, C. T. Chan, and P. Sheng, *Opt. Express* **13**, 7615 (2005).
32. M. Thiel, M. Decker, M. Deubel, M. Wegener, S. Linden, and G. Freymann, *Adv. Mater.* **19**, 207 (2007).
33. J. K. Gansel, M. Thiel, M. S. Rill, M. Decker, K. Bade, V. Saile, G. Freymann, S. Linden, and M. Wegener, *Science* **325**, 1513 (2009).
34. D. W. Berreman and T. J. Scheffer, *Phys. Rev. Lett.* **25**, 577 (1970).
35. H. Q. Li, (private communication).
36. For details, see Wulin Jia and Shuyi Zhang, *Phys. Lett. A* **358**, 487 (2006); J. Hao and L. Zhou, *Phys. Rev. B* **77**, 094201 (2008).
37. J. C. W. Lee and C. T. Chan, *Appl. Phys. Lett.* **90**, 051912 (2007).
38. W. S. Weiglhofer and A. Lahktakia, *Optik* **96**, 179 (1994).
39. A. Lakhtakia and W. S. Weiglhofer, *Proc. R. Soc. Lond. A* **448**, 419 (1995). *Corrections:* **454**, 3275 (1998).
40. P. G. de Gennes and J. A. Prost, *The Physics of Liquid Crystals*, 2nd ed., (Clarendon Press, Oxford, 1993).

MULTIPOLE METAMATERIALS

Joerg Petschulat*,§, Carsten Rockstuhl†, Christoph Menzel†, Arkadi Chipouline*, Andreas Tünnermann*,‡ , Falk Lederer†, and Thomas Pertsch*

*Friedrich-Schiller-Universität Jena, Institute of Applied Physics, Max Wien Platz 1, Jena 07743, Germany
§joerg.petschulat@uni-jena.de

†Friedrich-Schiller-Universität Jena, Institute of Condensed Matter Theory and Solid State Optics, Max Wien Platz 1, Jena 07743, Germany

Optical properties of plasmonic nanostructures and metamaterials are often accessed by evaluating their interaction with light by means of rigorous numerical methods. Such analysis allows the reliable prediction of any measurable quantity, whereas insights into the physical mechanisms that govern the observable effects require an intense interpretation of these quantities. Therefore, analytical methods are required that simplify the description of plasmonic entities to a certain extent but yet allow the disclosure of their physical peculiarities. We outline in this chapter the basics of such an analytical model which we coined the *multipole approach to metamaterials*. In this parametric model the elementary constituents that form plasmonic nanostructures are conceptually replaced by coupled dipoles. By describing the evolution of these dipoles in terms of differential equations, we disclose the dynamics of complex nanostructures. Furthermore, by introducing averaged quantities derived from the dipole dynamics, such as an electric and magnetic dipole and an electric quadrupole density, the light propagation in a medium comprising a dense array of these nanostructures is fully accessible. This contribution is written with the intention to familiarize readers with this framework and to allow its application to many related problem that may emerge in the field of plasmonics and metamaterials.

‡also with Fraunhofer Institute of Applied Optics and Precision Engineering Jena, Albert Einstein Strasse 7, Jena 07745, Germany.

1. Introduction

We present a physically thorough interpretation of the optical properties of plasmonic metamaterials on the basis of a classical multipole expansion. The main intention of our work is to provide an analytical framework to discuss the optical parameters of metamaterials and to enable any scientist to apply it to a particular problem. The approach requires the adjustment of only a few physically motivated parameters to a measured or simulated optical response from a metamaterial to fully disclose its properties and to describe the light propagation in it. An essential feature in all such considerations is the transition from analyzing the properties of individual building blocks towards that of a bulk media if they are densely packed. Therefore, we re-examine at first the spatial averaging procedure for mesoscopic meta-molecules, the plasmonic nanostructures constituting the building blocks of each metamaterial, as opposed to that of real molecules. Our findings will facilitate the understanding of linear effective material properties and various effects supported by the structure. We exemplarily consider the effect of asymmetric transmission, being a pivotal and counter-intuitive example for a far field response supported by suitably tailored metamaterials. Finally, this work will be completed by presenting a tool for explaining and understanding the radiation patterns for particular metamaterials as well. We will show how basic textbook electrodynamics can be applied in order to understand and describe the linear optical properties of metamaterials in a qualitative, quantitative, and physical manner.

2. Spatial Averaging for Meta-Molecules — Recalling the Role of Multipole Moments

To get started we revisit an important electrodynamic principle essential for optics; the spatial averaging procedure. The means to describe the transition between microscopic and macroscopic Maxwell's equations has been established in the first half of the last century.[1,2] Many averaging techniques have been applied and introduced to achieve this goal.[3,4] The result of this procedure is the physical description of light-matter interactions for macroscopic materials on the basis of simplified microscopic models. In that respect, the Drude-Lorentz model for metals and various others have become common in the framework of optics and material sciences today. In order to categorize and treat metamaterials as artificial macroscopic matter consisting of mesoscopic building blocks, this procedure will be recapitulated and the key differences to the averaging for natural

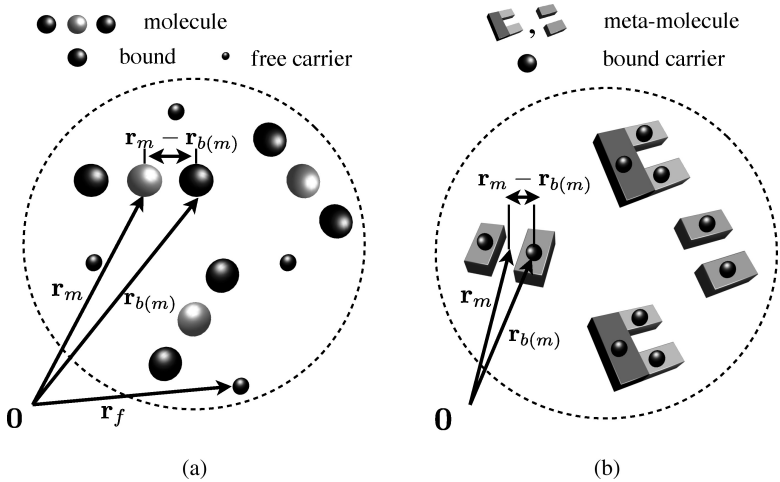


Fig. 1. (a) Typical averaging volume including bound molecular carriers and free carriers. (b) Bound carriers in a metamaterial represented by the microscopic free electrons but bound to the mesoscopic meta-molecules.

matter will be highlighted in this section. Thereby we briefly follow the averaging procedure according to G. Russakoff,[4] since it is probably one of the most common routines for averaging applied in various textbooks.[5]

The main idea of the spatial averaging centers around the consideration of *bound* and *free* carriers [Fig. 1(a)]. Analogously, free electrons of metal can be considered as bound carriers in artificial meta-molecules, see Fig. 1(b). Free carriers in metamaterials correspond to meta-molecules which are delocalized at least in one dimension. Since in this chapter localized plasmonic excitations, i.e., resonant carrier motions in spatially delimited meta-molecules are considered, free carriers analogues in metamaterials have been neglected, but can be in principle taken into account as well. In Fig. 1(b) two pioneering meta-molecules, i.e., the split-ring resonator (SRR)[6] and the cut-wire pair (CW),[7] have been sketched to visualize important geometries of interest. Now, the task is to spatially average over a distinct volume in which these meta-molecules are densely embedded. This procedure consists of two steps. At first, the molecular carrier distributions of each molecule will be spatially smeared-out. Second, the entire material will be built up as the arithmetic mean of all smoothed molecular contributions in the averaging volume. Thereby the spatial smoothing is often referred to the averaging itself since it is the most extensive part of the entire procedure. It will be shown that the smoothed molecular contributions emerge as multipole moments. Spatial smoothing,

i.e., averaging of any temporally and spatially dependent function $\psi(\mathbf{r},t)$ can be expressed as the convolution with a test function $G(\mathbf{r})$[8]

$$\langle\psi(\mathbf{r},t)\rangle = \int_{\mathbb{R}^3} d^3\mathbf{r}'\psi(\mathbf{r}-\mathbf{r}',t)G(\mathbf{r}'). \tag{1}$$

The test function $G(\mathbf{r})$ can be chosen arbitrarily with the restriction that it can be normalized to unity. Applying the averaging [Eq. (1)] to time and space derivatives yields for continuously differentiable functions $\psi(\mathbf{r},t)$

$$\frac{\partial}{\partial\mathbf{r}_i}\langle\psi(\mathbf{r},t)\rangle = \left\langle\frac{\partial}{\partial\mathbf{r}_i}\psi(\mathbf{r},t)\right\rangle, \tag{2}$$

$$\frac{\partial}{\partial t}\langle\psi(\mathbf{r},t)\rangle = \left\langle\frac{\partial}{\partial t}\psi(\mathbf{r},t)\right\rangle. \tag{3}$$

Thus, the microscopic electric field $\mathfrak{E}(\mathbf{r},t)$ and the magnetic induction $\mathfrak{B}(\mathbf{r},t)$ can be straightforwardly rewritten applying the respective averaged quantities. Then the homogeneous microscopic Maxwell's equations[5] read as

$$\nabla\times\langle\mathfrak{E}(\mathbf{r},t)\rangle + \frac{\partial}{\partial t}\langle\mathfrak{B}(\mathbf{r},t)\rangle = 0, \tag{4}$$

$$\nabla\cdot\langle\mathfrak{B}(\mathbf{r},t)\rangle = 0, \tag{5}$$

which can be transformed into macroscopic Maxwell's equations by substituting the microscopic fields $\mathfrak{E}$, $\mathfrak{B}$

$$\langle\mathfrak{E}(\mathbf{r},t)\rangle = \mathbf{E}(\mathbf{r},t), \tag{6}$$

$$\langle\mathfrak{B}(\mathbf{r},t)\rangle = \mathbf{B}(\mathbf{r},t). \tag{7}$$

The remaining inhomogeneous equations

$$\nabla\times\langle\mathfrak{B}(\mathbf{r},t)\rangle - \frac{1}{c^2}\frac{\partial}{\partial t}\langle\mathfrak{E}(\mathbf{r},t)\rangle = \mu_0\langle\mathfrak{j}(\mathbf{r},t)\rangle, \tag{8}$$

$$\nabla\cdot\langle\mathfrak{E}(\mathbf{r},t)\rangle = \frac{1}{\epsilon_0}\langle\mathfrak{r}(\mathbf{r},t)\rangle, \tag{9}$$

contain next to the averaged macroscopic fields $\mathbf{E}(\mathbf{r},t)$, $\mathbf{B}(\mathbf{r},t)$ also the averaged current $\langle\mathfrak{j}(\mathbf{r},t)\rangle$ and charge $\langle\mathfrak{r}(\mathbf{r},t)\rangle$ densities that have to be considered in detail. By applying the averaging procedure [Eq. (1)] to the charge density $\mathfrak{r}(\mathbf{r},t)$ one obtains

$$\langle\mathfrak{r}(\mathbf{r},t)\rangle = \int_{\mathbb{R}^3} d^3\mathbf{r}'\mathfrak{r}(\mathbf{r}-\mathbf{r}',t)G(\mathbf{r}'). \tag{10}$$

Any microscopic carrier distribution corresponding to a natural or artificial material may then be separated into two contributions: *bound* and *free* carriers, according to Fig. 1(a, b), where both can be considered as point charges in the form of

$$\begin{aligned}\mathfrak{r}(\mathbf{r},t) &= \mathfrak{r}_{\text{bound}}(\mathbf{r},t) + \mathfrak{r}_{\text{free}}(\mathbf{r},t) \qquad (11)\\ &= \sum_{m=1}^{N_{\text{molecules}}} \sum_{b(m)=1}^{N_{\text{bound}}} q_{b(m)}\delta(\mathbf{r}-\mathbf{r}_m-\mathbf{r}_{b(m)}) \\ &\quad + \sum_{f=1}^{N_{\text{free}}} q_f \delta(\mathbf{r}-\mathbf{r}_f). \qquad (12)\end{aligned}$$

Bound and free carriers are understood as charges of atoms and molecules or valence electrons, respectively. For the case of meta-molecules bound carriers are associated with ensembles of electrons that can be driven into resonance upon external illumination. Therefore noble metals are typically used as building blocks because of their high density of quasi-free conduction band carriers. Averaging Eq. (12) reads as

$$\begin{aligned}\langle \mathfrak{r}(\mathbf{r},t)\rangle &= \sum_{m=1}^{N_{\text{molecules}}} \sum_{b(m)=1}^{N_{\text{bound}}} q_{b(m)} G(\mathbf{r}-\mathbf{r}_m-\mathbf{r}_{b(m)}) \\ &\quad + \sum_{f=1}^{N_{\text{free}}} q_f G(\mathbf{r}-\mathbf{r}_f). \qquad (13)\end{aligned}$$

In a final step it is assumed that the typical dimensions over which the averaging is performed are much larger than the molecular dimensions. If this applies a Taylor expansion of the bound contribution of Eq. (12) up to the second order around $|\mathbf{r}-\mathbf{r}_m|$ provides for the m-th molecule

$$\begin{aligned}\langle \mathfrak{r}_{\text{bound}}(\mathbf{r},t)\rangle_m &= \sum_{b(m)=1}^{N_{\text{bound}}} q_{b(m)} G(\mathbf{r}-\mathbf{r}_m-\mathbf{r}_{b(m)}) \qquad (14)\\ &\approx \sum_{b(m)=1}^{N_{\text{bound}}} q_{b(m)} \Bigg\{ G(\mathbf{r}-\mathbf{r}_m) - \mathbf{r}_{b(m)}\cdot\nabla G(\mathbf{r}-\mathbf{r}_m) \\ &\quad + \frac{1}{2}\sum_{\alpha,\beta}[\mathbf{r}_{b(m)}]_\alpha [\mathbf{r}_{b(m)}]_\beta \frac{\partial^2}{\partial r_\alpha \partial r_\beta} G(\mathbf{r}-\mathbf{r}_m)\Bigg\}. \qquad (15)\end{aligned}$$

Now, in the series expansion (15) one can introduce the electric multipole moments yielding an averaged carrier density that can be expressed as

$$\langle \mathfrak{r}_{\text{bound}}(\mathbf{r},t)\rangle_m \approx \tilde{q}_m G(\mathbf{r}-\mathbf{r}_m) - \mathbf{p}_m \cdot \nabla G(\mathbf{r}-\mathbf{r}_m) + \frac{1}{6}\sum_{\alpha,\beta}(Q_m)_{\alpha\beta}\frac{\partial^2}{\partial r_\alpha \partial r_\beta} G(\mathbf{r}-\mathbf{r}_m), \tag{16}$$

where $\tilde{q}_m$, $\mathbf{p}_m$, and $\hat{\mathbf{Q}}_m$ represent the *bound*, i.e., *molecular* quantities: charge, electric dipole, and electric quadrupole moment associated with the m-th molecule, respectively

$$\tilde{q}_m = \sum_{b(m)=1}^{N_{\text{bound}}} q_{b(m)}, \tag{17}$$

$$\mathbf{p}_m = \sum_{b(m)=1}^{N_{\text{bound}}} q_{b(m)}\mathbf{r}_{b(m)}, \tag{18}$$

$$(Q_m)_{\alpha\beta} = 3\sum_{b(m)=1}^{N_{\text{bound}}} q_{b(m)}[r_{b(m)}]_\alpha [r_{b(m)}]_\beta. \tag{19}$$

In passing we mention that a primitive quadrupole tensor has been applied in order to obtain Maxwell's equations that are origin independent.[9,10] Considering Eq. (15) one would have arrived at exactly the same expressions, if the charge distribution would have been

$$\mathfrak{r}_{\text{bound}}(\mathbf{r},t) = \sum_{m=1}^{N_{\text{molecules}}} \left[\tilde{q}_m \delta(\mathbf{r}-\mathbf{r}_m) - \nabla\cdot\mathbf{p}_m\delta(\mathbf{r}-\mathbf{r}_m) + \frac{1}{6}\sum_{\alpha,\beta}\frac{\partial^2}{\partial r_\alpha \partial r_\beta}(Q_m)_{\alpha\beta}\delta(\mathbf{r}-\mathbf{r}_m)\right]. \tag{20}$$

With the averaging procedure according to Eq. (10) and Eqs. (11, 12) one obtains

$$\langle \mathfrak{r}_{\text{bound}}(\mathbf{r},t)\rangle = \sum_{m=1}^{N_{\text{molecules}}} \left[\langle\tilde{q}_m \delta(\mathbf{r}-\mathbf{r}_m)\rangle - \nabla\cdot\langle\mathbf{p}_m\delta(\mathbf{r}-\mathbf{r}_m)\rangle + \frac{1}{6}\sum_{\alpha,\beta}\frac{\partial^2}{\partial r_\alpha \partial r_\beta}\langle(Q_m)_{\alpha\beta}\delta(\mathbf{r}-\mathbf{r}_m)\rangle\right]. \tag{21}$$

The explicit application of the convolution [Eq. (10)] results in Eq. (15), but solely by considering Eq. (21) one has an intuitive physical interpretation of the averaging procedure at hand. Any microscopic molecular or atomic carrier ensemble will result in a collection of electric point multipoles in the averaged macroscopic domain. For the first inhomogeneous Eq. (9) the averaged charge density (21) gives

$$\epsilon_0 \nabla \cdot \langle \mathfrak{E}(\mathbf{r},t)\rangle = \langle \mathfrak{r}(\mathbf{r},t)\rangle,$$
$$\sum_\alpha \frac{\partial}{\partial r_\alpha}\left[\epsilon_0 \langle \mathfrak{E}_\alpha(\mathbf{r},t)\rangle + P_\alpha(\mathbf{r},t) - \frac{\partial}{\partial r_\beta} Q_{\alpha\beta}(\mathbf{r},t)\right] = \rho(\mathbf{r},t), \tag{22}$$

where the *macroscopic charge density*

$$\rho(\mathbf{r},t) = \sum_{m=1}^{N_{\text{molecules}}} \langle \tilde{q}_m \delta(\mathbf{r}-\mathbf{r}_m)\rangle + \sum_{f=1}^{N_{\text{free}}} \langle \tilde{q}_f \delta(\mathbf{r}-\mathbf{r}_f)\rangle, \tag{23}$$

the *macroscopic polarization*

$$\mathbf{P}(\mathbf{r},t) = \sum_{m=1}^{N_{\text{molecules}}} \langle \mathbf{p}_m \delta(\mathbf{r}-\mathbf{r}_m)\rangle, \tag{24}$$

and the *macroscopic quadrupole density*

$$Q_{\alpha\beta}(\mathbf{r},t) = \frac{1}{6}\sum_{m=1}^{N_{\text{molecules}}} \langle (Q_m)_{\alpha\beta} \delta(\mathbf{r}-\mathbf{r}_m)\rangle, \tag{25}$$

have been introduced. Hence, the electric displacement $\mathbf{D}(\mathbf{r},t)$ can be formally introduced on the basis of Eq. (22) as

$$\mathbf{D}(\mathbf{r},t) = \epsilon_0 \mathbf{E}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t) - \nabla \cdot \hat{\mathbf{Q}}(\mathbf{r},t), \tag{26}$$
$$\text{with: } [\nabla \cdot \hat{\mathbf{Q}}(\mathbf{r},t)]_\alpha \equiv \sum_\beta \frac{\partial}{\partial r_\beta} Q_{\alpha\beta}(\mathbf{r},t).$$

To finalize the transition from microscopic to macroscopic Maxwell's equations, the microscopic current density needs to be averaged. Averaging

microscopic free and molecular currents in the form of

$$\mathrm{j}(\mathbf{r},t) = \langle \mathrm{j}_{\mathrm{bound}}(\mathbf{r},t)\rangle + \langle \mathrm{j}_{\mathrm{free}}(\mathbf{r},t)\rangle \tag{27}$$

$$= \sum_{m=1}^{N_{\mathrm{molecules}}} \sum_{b(m)=1}^{N_{\mathrm{bound}}} q_{b(m)}[\mathbf{v}_m + \mathbf{v}_{b(m)}]\delta(\mathbf{r} - \mathbf{r}_m - \mathbf{r}_{b(m)}) + \sum_{f=1}^{N_{\mathrm{free}}} q_f \mathbf{v}_f \delta(\mathbf{r} - \mathbf{r}_f), \tag{28}$$

will lead, similar to the charge density Eq. (23), to

$$\langle \mathrm{j}(\mathbf{r},t)\rangle = \sum_{m=1}^{N_{\mathrm{molecules}}} \sum_{b(m)=1}^{N_{\mathrm{bound}}} q_{b(m)}[\mathbf{v}_m + \mathbf{v}_{b(m)}]G(\mathbf{r} - \mathbf{r}_m - \mathbf{r}_{b(m)}) + \sum_{f=1}^{N_{\mathrm{free}}} q_f \mathbf{v}_f G(\mathbf{r} - \mathbf{r}_f). \tag{29}$$

Now, the bound carrier contribution of Eq. (29) consists of two parts, the velocity of the m-th molecule $\mathbf{v}_m$ together with the relative velocity of the bound charges $\mathbf{v}_{b(m)}$ and the free carrier motion $\mathbf{v}_f$ which are simply added in the non-relativistic regime. Again the expression can be analyzed in terms of a Taylor expansion up to second order

$$\langle \mathrm{j}_{\mathrm{bound}}(\mathbf{r},t)\rangle_m = \sum_{b(m)=1}^{N_{\mathrm{bound}}} q_{b(m)}[\mathbf{v}_m + \mathbf{v}_{b(m)}]G(\mathbf{r} - \mathbf{r}_m - \mathbf{r}_{b(m)}) \tag{30}$$

$$\approx \sum_{b(m)=1}^{N_{\mathrm{bound}}} q_{b(m)}[\mathbf{v}_m + \mathbf{v}_{b(m)}] \left[G(\mathbf{r} - \mathbf{r}_m) - \mathbf{r}_{b(m)} \cdot \nabla G(\mathbf{r} - \mathbf{r}_m) + \sum_{\alpha\beta} [r_{b(m)}]_\alpha [r_{b(m)}]_\beta \frac{\partial^2}{\partial r_\alpha \partial r_\beta} G(\mathbf{r} - \mathbf{r}_m) \right]. \tag{31}$$

Similarly to the procedure applied for the charge density, the molecular current density (31) can be transformed into[4]

$$\langle \mathrm{j}_\alpha(\mathbf{r},t)\rangle = j_\alpha(\mathbf{r},t) + \frac{\partial}{\partial t}\left[P_\alpha(\mathbf{r},t) - \frac{\partial}{\partial r_\beta} Q_{\alpha\beta}(\mathbf{r},t) \right] + [\nabla \times \mathbf{M}(\mathbf{r},t]_\alpha$$

$$+\sum_{\beta}\frac{\partial}{\partial r_\beta}\sum_{m=1}^{N_{\text{molecules}}}\langle[(p_m)_\beta(v_m)_\alpha-(p_m)_\alpha(v_m)_\beta]\delta(\mathbf{r}-\mathbf{r}_m)\rangle$$

$$+\frac{1}{6}\sum_{\beta\gamma}\frac{\partial^2}{\partial r_\beta\partial r_\gamma}\sum_{m=1}^{N_{\text{molecules}}}\langle[(Q_m)_{\alpha\beta}(v_m)_\gamma-(Q_m)_{\gamma\beta}(v_m)_\alpha]\delta(\mathbf{r}-\mathbf{r}_m)\rangle, \tag{32}$$

where the *macroscopic current density*

$$\mathbf{j}(\mathbf{r},t)=\sum_{f=1}^{N_{\text{free}}}\langle q_f\mathbf{v}_f\delta(\mathbf{r}-\mathbf{r}')\rangle+\sum_{m=1}^{N_{\text{molecules}}}\langle q_m\mathbf{v}_m\delta(\mathbf{r}-\mathbf{r}')\rangle, \tag{33}$$

the *macroscopic magnetization*

$$\mathbf{M}(\mathbf{r},t)=\sum_{m=1}^{N_{\text{molecules}}}\langle\mathbf{m}_m\delta(\mathbf{r}-\mathbf{r}')\rangle, \tag{34}$$

as well as the *molecular magnetic dipole moment*

$$\mathbf{m}_m=\frac{1}{2}\sum_{b(m)=1}^{N_{\text{bound}}}q_{b(m)}[\mathbf{r}_{b(m)}\times\mathbf{v}_{b(m)}], \tag{35}$$

have been introduced. The last two terms of Eq. (32) are usually negligible since for many optical systems the molecular velocity is assumed to be zero or at least much smaller than the velocities of the bound carriers.[5] These contributions have been taken into account just for being consistent, but are dropped in the following by formally setting $\mathbf{v}_m=0$. It is convenient to write the current density in this form because it is then easily possible to evaluate the remaining inhomogeneous Maxwell's equation (8)

$$\frac{1}{\mu_0}\nabla\times\langle\mathfrak{B}(\mathbf{r},t)\rangle-\epsilon_0\frac{\partial}{\partial t}\langle\mathfrak{E}(\mathbf{r},t)\rangle=\langle\mathfrak{j}(\mathbf{r},t)\rangle,$$

$$\frac{1}{\mu_0}\nabla\times\mathbf{B}(\mathbf{r},t)-\frac{\partial}{\partial t}[\epsilon_0\mathbf{E}+\mathbf{P}(\mathbf{r},t)-\nabla\times\hat{\mathbf{Q}}(\mathbf{r},t)]=\mathbf{j}(\mathbf{r},t)+\nabla\times\mathbf{M}(\mathbf{r},t),$$

$$\nabla\times\left[\frac{1}{\mu_0}\mathbf{B}(\mathbf{r},t)-\mathbf{M}(\mathbf{r},t)\right]-\frac{\partial}{\partial t}\mathbf{D}(\mathbf{r},t)=\mathbf{j}(\mathbf{r},t), \tag{36}$$

which is equal to Eq. (6) if the macroscopic magnetic field is defined by

$$\mathbf{H}(\mathbf{r},t)=\frac{1}{\mu_0}\mathbf{B}(\mathbf{r},t)-\mathbf{M}(\mathbf{r},t). \tag{37}$$

This concise discussion of the averaging procedure provides an easy understanding of the macroscopic Maxwell's equations for natural media. Moreover the multipole moments up to second order are defined and their contribution to the macroscopic *constitutive relations*, i.e., Eqs. (26, 37) has been retraced.

Performing the averaging procedure and determining the constitutive relations in the above manner is the method of choice in various textbooks but not the only possibility. For several reasons the magnetic dipole response can be included within the macroscopic current density. Using another definition of the macroscopic current density that is based on Eq. (32) one obtains

$$\mathbf{j}(\mathbf{r},t) = \left\langle \sum_{f=1}^{N_{\text{free}}} q_f \mathbf{v} \delta(\mathbf{r} - \mathbf{r}_f) \right\rangle + \nabla \times \mathbf{M}(\mathbf{r},t). \tag{38}$$

The current density [Eq. (38)] with the related macroscopic charge density ρ must satisfy the continuity equation

$$\frac{\partial}{\partial t}\rho(\mathbf{r},t) + \nabla \cdot \mathbf{j}(\mathbf{r},t) = 0. \tag{39}$$

With the current density [Eq. (38)] the charge density remains equal to that in Eq. (23), because $\nabla \cdot [\nabla \times \mathbf{M}] = 0$. The main difference to the definition of the current density given before [Eq. (33)], concerns the constitutive relation for the magnetic field compared with Eq. (37)

$$\mathbf{H}(\mathbf{r},t) = \frac{1}{\mu_0}\mathbf{B}(\mathbf{r},t). \tag{40}$$

Hence, there is a certain degree of freedom in defining the macroscopic current density. The influence of such modifications will result in different constitutive relations, e.g., Eq. (40) and thus in different boundary conditions at the interface of the macroscopic media compared with the standard boundary conditions.[11] In the same manner as the magnetic dipole moment has been included in $\mathbf{j}$ also the electric multipole contributions could be included yielding another constitutive relation for the electric field[12] when compared with Eq. (26). So it is important to keep this ambiguity in mind when reading articles or books where differing constitutive relations or averaging procedures have been used.

In order to conveniently describe a macroscopic medium, where the averaging is performed, material parameters can be introduced. This is

typically done in the Fourier domain to which the transformations may be defined as

$$\mathbf{X}(t) = \frac{1}{2}\left[\int_{\omega\in\mathbb{R}} d\omega \mathbf{X}(\omega)e^{-i\omega t} + \text{c.c.}\right], \tag{41}$$

$$\mathbf{X}(\mathbf{r},t) = \frac{1}{2}\left[\int_{\omega\in\mathbb{R}} d\omega \int_{\mathbf{k}\in\mathbb{R}^3} d\mathbf{k}\mathbf{X}(\mathbf{k},\omega)e^{i\mathbf{kr}-i\omega t} + \text{c.c.}\right]. \tag{42}$$

Applying the temporal Fourier transformation [Eq. (41)] to Eq. (26) and Eq. (37) the constitutive relations become

$$\mathbf{D}(\mathbf{r},\omega) = \epsilon_0\mathbf{E}(\mathbf{r},\omega) + \mathbf{P}(\mathbf{r},\omega) - \nabla\cdot\hat{\mathbf{Q}}(\mathbf{r},\omega), \tag{43}$$

$$\mathbf{H}(\mathbf{r},\omega) = \frac{1}{\mu_0}\mathbf{B}(\mathbf{r},\omega) - \mathbf{M}(\mathbf{r},\omega). \tag{44}$$

Now that all multipole moments are related to carrier dynamics induced by the impinging fields, a linear dependence of all moments on the electric field can be assumed.[a] Due to the presence of spatial derivatives in the dielectric displacement, e.g., in front of the electric quadrupole tensor in Eq. (43), the general response will show *non-local effects*. Hence, for a linear relation between the field and the moments, the application of the Fourier transform (42) will immediately result in a k-dependence in the constitutive relations. Thus, in general second order moments will induce *spatial dispersion* and for Eqs. (43, 44) the following expressions can be introduced

$$\mathbf{D}(\mathbf{k},\omega) = \epsilon_0\hat{\epsilon}(\mathbf{k},\omega)\mathbf{E}(\mathbf{k},\omega), \tag{45}$$

$$\mathbf{B}(\mathbf{k},\omega) = \mu_0\hat{\mu}(\mathbf{k},\omega)\mathbf{H}(\mathbf{k},\omega). \tag{46}$$

There $\hat{\epsilon}(\mathbf{k},\omega)$ corresponds to the *electric permittivity* while $\hat{\mu}(\mathbf{k},\omega)$ represents the *magnetic permeability* for the averaged medium. These equations complete the averaging process since they allow us to treat any macroscopic, i.e., averaged medium, by the knowledge of its microscopic constituents. Such constituents are averaged multipole moments that account for light matter interactions with bound and free molecular carriers. Having arrived at the macroscopic level, e.g., Eqs. (45, 46), the basics of the averaging procedure are no longer of interest for many applications, since for various macroscopic media these values are obtained experimentally.[13]

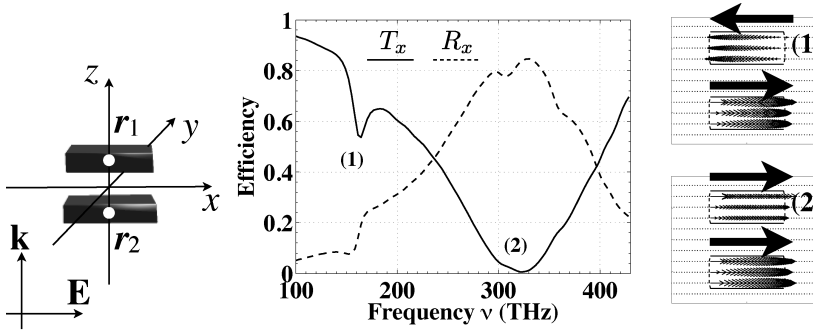
[a]This assumption of course only holds in the linear optical regime. For the considerations of multipole induced nonlinearities the dependence is different.[22]

The purpose of the intense recapitulation in this section was to recall the fundamental principles which will facilitate the understanding of ensembles of mesoscopic nanostructures in the following. These nanostructures are acting as artificial atoms or molecules, termed meta-molecules, forming a new material class that will be considered in complete analogy to the previous procedure throughout this work [Fig. 1(a, b)].

The most appealing difference in the averaging of nanoparticles representing artificial meta-molecules in comparison to ordinary molecules is that all multipole moments in nanoparticles are induced by an electromagnetic field. Hence, all moments in meta-molecules vanish without an external field. In contrast, real molecules can possess permanent multipole contributions even in the absence of any external fields. A prominent example for a material with a permanent molecular quadrupole moment is carbon dioxide being the first experimentally measured molecular quadrupole.[14] In meta-molecules higher order multipoles can be achieved at will. To obtain second order effects in meta-molecules it has been shown above that effects such as *non-locality*, i.e., *spatial dispersion* in the Fourier domain, must exist. For this purpose each meta-molecule must comprise a certain dimension in propagation direction. On the other side, meta-molecules are supposed to be as small as possible to make the averaging and the multipole expansion meaningful. Thus, meta-molecules should fulfill the sub-wavelength restriction.[15] These two competing issues led to the concept of *mesoscopic structures*[16] which applies to meta-molecules exhibiting the above-mentioned optical effects, i.e., artificial magnetism occurring at optical wavelengths.[17,18] The application of the averaging procedure, i.e., Eq. (10), for mesoscopic structures despite of natural atoms or molecules is of course an approximation that describes the structure better if its dimensions decrease. From a physical point of view the important difference between natural molecules and meta-molecules is the possibility to control the carrier dynamics in meta-materials and the increased ratio between the meta-molecules dimensions and the wavelength when compared with natural molecules. However, it will be shown in the following that this concept holds to describe and predict the optical properties of certain metamaterials sufficiently well.

3. Light Propagation in Metamaterials Including Multipole Moments Up to the Second Order

As repeated in the previous section the electromagnetic response of meta-molecules can be transformed into multipolar contributions. The remaining missing part for this task is the access to the bound carrier dynamics in



(a) (b) (c)

Fig. 2. (a) Sketch of the CW meta-molecule with the charge carriers representing the dynamics denoted by white dots. Each of the dots accounts for two carriers. A positively and a negatively charged one being at the same position without an external electric field. (b) The numerically obtained far-field reflectance (dashed) and transmittance (solid) for a periodic arrangement of CW pairs. The spectra have been calculated for plane wave propagation in z-direction presuming x-polarized electric fields. The numbers denote the two fundamental resonances: (1) anti-symmetric and (2) symmetric resonance. (c) Top: Electric field (E_x) inside the CW pairs for resonance (1). Bottom: Electric field for resonance (2).

meta-molecules $\mathbf{r}_{b(m)}$. With an underlying understanding of these carrier motions all multipole contributions of interest [Eqs. (18, 19, 35)] as well as the dispersion relation and the material responses according to Eqs. (45, 46) are obtained. In Fig. 2(a) the cut-wire (CW) pair meta-molecule is sketched. The white dots in Fig. 2(a) account for the carrier distribution which is used to model the microscopic currents. We apply our formalism to such a structure since the CW was decisive for the field of optical metamaterials; it is well understood, sufficiently simple and overall it is very educative for our purpose to outline the forte of the above developed formalism. The reflection and transmission spectra for a periodical array of CW meta-molecules have been computed by means of the rigorous Fourier Modal Method (FMM)[19] and are shown in Fig. 2(b). For the CW pairs two distinct resonances emerge that can be associated with a *symmetric* and an *anti-symmetric* field distribution. The near-fields were also obtained by FMM and are shown in Fig. 2(c). In order to account for these dynamics, a suitable carrier configuration of two negatively charged (superscript "−") and two positively charged (superscript "+") carriers as sketched in Fig. 2(a) is assumed

$$\begin{aligned} \mathbf{r}_1^+ &= [0,0,z_0]^T, \qquad \mathbf{r}_1^- = [\xi_1(t),0,z_0]^T, \\ \mathbf{r}_2^+ &= [0,0,-z_0]^T, \quad \mathbf{r}_2^- = [\xi_2(t),0,-z_0]^T. \end{aligned} \tag{47}$$

Thereby only negatively charged carriers are considered to perform oscillations $\xi_{1,2}(t)$ which are induced by the microscopic electric field as[b]

$$\frac{\partial^2}{\partial t^2}\xi_1(t) + \gamma_1 \frac{\partial}{\partial t}\xi_1(t) + \omega_{01}^2 \xi_1(t) + \sigma\xi_2(t) = -\frac{q_1}{m_1}\mathfrak{E}_x(z+z_0,t), \quad (48)$$

$$\frac{\partial^2}{\partial t^2}\xi_2(t) + \gamma_2 \frac{\partial}{\partial t}\xi_2(t) + \omega_{02}^2 \xi_2(t) + \sigma\xi_1(t) = -\frac{q_2}{m_2}\mathfrak{E}_x(z-z_0,t). \quad (49)$$

In addition to the well-known terms in the oscillator equations (48, 49) such as the damping constants γ_j, the eigenfrequencies ω_{0j}, the two equations are coupled denoted by the coupling constant σ. This coupling accounts for the near-field interaction that takes place between the two metal wires forming the CW meta-molecule. Without this coupling the symmetric and the antisymmetric modes would be degenerated in their resonance frequencies. A decoupling could be achieved by separating both wires far beyond the typical near-field coupling distances, which are in the order of $\approx$20 nm for optical frequencies. This case would clearly violate the subwavelength restriction in propagation direction.[15] With the oscillator equations it is now possible to calculate the bound carrier dynamics for an isolated CW meta-molecule and to find the desired molecular quantities: electric dipole (18), electric quadrupole (19), and magnetic dipole (35) moments

$$\mathbf{p}_{\mathrm{CW}} = \sum_{l=1}^{N} q_l \mathbf{r}_l = \begin{pmatrix} -q[\xi_1(t)+\xi_2(t)] \\ 0 \\ 0 \end{pmatrix}, \quad (50)$$

$$\hat{\mathbf{Q}}_{\mathrm{CW}} = 3\sum_{l=1}^{N} q_l [\mathbf{r}_l]_\alpha [\mathbf{r}_l]_\beta = -3q \begin{pmatrix} \xi_1^2(t) - \xi_2^2(t) & 0 & z_0[\xi_1(t)-\xi_2(t)] \\ 0 & 0 & 0 \\ z_0[\xi_1(t)-\xi_2(t)] & 0 & 0 \end{pmatrix}, \quad (51)$$

$$\mathbf{m}_{\mathrm{CW}} = \frac{1}{2}\sum_{l=1}^{N} q_l \mathbf{r}_l \times \frac{\partial}{\partial t}\mathbf{r}_l = -\frac{qz_0}{2} \begin{pmatrix} 0 \\ \frac{\partial}{\partial t}[\xi_1(t)-\xi_2(t)] \\ 0 \end{pmatrix}. \quad (52)$$

For the sake of simplicity a symmetric CW meta-molecule is assumed with $\gamma_1 = \gamma_2 \equiv \gamma$, $\omega_{01} = \omega_{02} \equiv \omega_0$, and $q_1 = q_2 \equiv q$. It is obvious from Eqs. (50–52) that all moments depend either on the sum

[b]Note that the interaction of the magnetic field with charged carriers is neglected since a non-relativistic dynamic is assumed ($\partial_t |\mathbf{r}|/c \ll 1$).

or the difference of the oscillator elongations ξ_j. Explicitly, the electric dipole resonance depends on the sum, while both second order multipole moments are proportional to the difference. Hence, for the symmetric mode ($\xi_1 = \xi_2$) the electric dipole moment is maximized and both second order moments vanish. In turn, for a completely anti-symmetric elongation ($\xi_1 = -\xi_2$) second order moments are maximal and the electric dipole moment is zero. Hence, both oscillator eigenmodes, i.e., the symmetric as well as the anti-symmetric, can be successfully translated into multipole moments of first and second order. The respective wave equation which describes the propagation of electromagnetic fields inside a macroscopic material comprising up to second order multipoles upon temporal Fourier transformation [Eq. (41)] reads as

$$\Delta\mathbf{E}(\mathbf{r},\omega) + \frac{\omega^2}{c^2}\mathbf{E}(\mathbf{r},\omega) + \frac{\omega^2}{c^2\epsilon_0}\left[\mathbf{P}(\mathbf{r},\omega) - \nabla\cdot\hat{\mathbf{Q}}(\mathbf{r},\omega) + \frac{i}{\omega}\nabla\times\mathbf{M}(\mathbf{r},\omega)\right] = 0. \quad (53)$$

For x-polarized transverse electrical fields propagating in z-direction and by Fourier transformation and substitution of the multipole moments (50–52), Eq. (53) simplifies to the scalar equation

$$\frac{\partial^2}{\partial z^2}E_x(z,\omega) + \frac{\omega^2}{c^2}E_x(z,\omega) + \frac{\omega^2}{c^2\epsilon_0}\left[P_x(z,\omega) - \frac{\partial}{\partial z}Q_{xz}(z,\omega) - \frac{i}{\omega}\frac{\partial}{\partial z}M_y(z,\omega)\right] = 0. \quad (54)$$

Prior to solving this equation for particular field distributions, there is one important point to be mentioned. The macroscopic quantities $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{M}$ have been obtained by basically summing up over all molecular quantities (24, 25, 34). Hence, it is assumed that the corrections due to interactions of meta-molecules can be neglected since the meta-molecules are well separated. This is barely fulfilled for realistic metamaterials since even a weak coupling between adjacent meta-molecules causes a shift of resonance frequencies when compared to the isolated meta-molecule. To include this effect, either a more sophisticated averaging procedure[c] or adapted

[c]In principle the Clausius Mosotti formalism can be applied where the Lorentz field interaction between meta-molecules can be considered dependent on the metamaterials volume density.

oscillator parameters can be used. Because properties of isolated meta-molecules are hardly accessible since typically ensembles of metamaterials are considered this interaction will be incorporated in the form of adapted oscillator parameters, being the numerically and experimentally accessible ensemble values.

Determining the averaged quantities [Eqs. (24, 25, 34)] from the molecular moments [Eqs. (50–52)] in the Fourier domain [Eq. (41)]

$$\mathbf{P}(\mathbf{r},\omega) = \eta \mathbf{p}_{\mathrm{CW}}(\mathbf{r},\omega), \tag{55}$$

$$\hat{\mathbf{Q}}(\mathbf{r},\omega) = \frac{\eta}{6}\hat{\mathbf{Q}}_{\mathrm{CW}}(\mathbf{r},\omega), \tag{56}$$

$$\mathbf{M}(\mathbf{r},\omega) = \eta \mathbf{m}_{\mathrm{CW}}(\mathbf{r},\omega), \tag{57}$$

with the number density η and substituting them into Eq. (54) yields

$$\frac{\partial^2}{\partial z^2}E_x(z,\omega) + \frac{\omega^2}{c^2}E_x(z,\omega) - \frac{\omega^2 \eta q}{c^2 \epsilon_0}\left\{[\xi_1(\omega)+\xi_2(\omega)] - z_0\frac{\partial}{\partial z}[\xi_1(\omega)-\xi_2(\omega)]\right\} = 0. \tag{58}$$

The first term $\xi_1(\omega)+\xi_2(\omega)$ in the brackets represents the electric dipole interactions whereas the second term $\propto \partial_z[\xi_1(\omega)-\xi_2(\omega)]$ arises from both the electric quadrupole and the magnetic dipole interactions. Now, the remaining task is to evaluate these two contributions as functions of the averaged electric field $E_x(z,\omega)$. Thus, from Eqs. (48, 49) and by replacing the microscopic with the macroscopic field according to Eq. (6) one ends up with

$$\xi_1(\omega)+\xi_2(\omega) = -\frac{q}{m}\frac{E_x(z+z_0,\omega)+E_x(z-z_0,\omega)}{A+\sigma}, \tag{59}$$

$$\xi_1(\omega)-\xi_2(\omega) = -\frac{q}{m}\frac{E_x(z+z_0,\omega)-E_x(z-z_0,\omega)}{A-\sigma}, \tag{60}$$

$$\text{with: } A(\omega) \equiv \omega_0^2 - \omega^2 - i\omega\gamma. \tag{61}$$

Now, if an ansatz for the propagating modes inside the metamaterial like

$$E_x(z,\omega) = E_0 e^{ik_z(\omega)z}, \tag{62}$$

is substituted into the wave equation (58), the dispersion relation can be obtained as

$$k_z^2(\omega) = \frac{\omega^2}{c^2}\left\{1 + \frac{\eta q^2 2}{\epsilon_0 m}\frac{\cos[k_z(\omega)z_0]}{A(\omega)+\sigma} + \frac{\eta q^2 2}{\epsilon_0 m}\frac{z_0 k_z(\omega)\sin[k_z(\omega)z_0]}{A(\omega)-\sigma}\right\}. \tag{63}$$

Analyzing the functional form of the implicit equation (63) it can be easily verified that this equation has various solutions k_z for a fixed frequency, representing different branches of the dispersion relation. For CW meta-molecules having larger dimensions higher order modes have to be considered. For CW meta-molecules with only a small extent in the propagation direction only the fundamental mode occurs. Hence, for such metamaterials operating in the fundamental mode regime the trigonometric functions in Eq. (63) may be expanded in a Taylor series up to the first non-constant contribution

$$\cos[k_z(\omega)z_0] \approx 1 - \frac{k_z^2(\omega)z_0^2}{2}, \quad \sin[k_z(\omega)z_0] \approx k_z(\omega)z_0, \tag{64}$$

which provides an explicit form of the dispersion relation $k_z(\omega)$ for the fundamental mode

$$k_z^2(\omega) = \frac{\omega^2}{c^2}\frac{1+\frac{C}{A(\omega)+\sigma}}{1-\frac{\omega^2}{c^2}\frac{Cz_0^2}{2}\frac{A(\omega)+3\sigma}{A(\omega)^2-\sigma^2}}, \quad \text{with: } C \equiv \frac{2\eta q^2}{m\epsilon_0}. \tag{65}$$

In addition to the dispersion relation Eq. (65), also the effective material properties according to Eqs. (45, 46) may be assigned

$$\begin{aligned}\epsilon_x(z,\omega) &= 1 + \epsilon_0^{-1}[P_x(z,\omega) - \frac{\partial}{\partial z}Q_{xz}(z,\omega)]E_x^{-1}(z,\omega),\\ \epsilon_x(k_z,\omega) &= 1 + \frac{2\eta q^2}{m\epsilon_0}\frac{\cos[k_z(\omega)z_0]}{A(\omega)+\sigma} + k_z(\omega)z_0\frac{\eta q^2}{m\epsilon_0}\frac{\sin[k_z(\omega)z_0]}{A(\omega)-\sigma}.\end{aligned} \tag{66}$$

Considering Eq. (66) it can be inferred that spatial dispersion occurs, as indicated by the intrinsic $k_z(\omega)$-dependence of the permittivity. By applying the approximations (64), the permittivity associated with the fundamental

mode becomes

$$\epsilon_x(k_z,\omega) = 1 + \frac{C}{A(\omega)+\sigma} + k_z^2(\omega) z_0^2 \frac{C\sigma}{A^2(\omega)-\sigma^2}. \tag{67}$$

For the magnetic permeability

$$\mu_y(z,\omega) = B_y(z,\omega)[B_y(z,\omega) - \mu_0 M_y(z,\omega)]^{-1} \tag{68}$$

the magnetic induction has to be replaced by the electric field with $B_y(z,\omega) = [k_z(\omega)/\omega]E_x(z,\omega)$ which finally results again in a spatially dispersive expression

$$\mu_y(k_z,\omega) = \left\{1 - \frac{\omega^2}{c^2}\frac{\eta q^2}{m\epsilon_0}\frac{z_0 \sin(k_z(\omega)z_0)}{k_z(\omega)[A(\omega)-\sigma]}\right\}^{-1}. \tag{69}$$

Upon substitution of the approximations (64) one ends up with

$$\mu_y(\omega) = \left\{1 - \frac{\omega^2}{c^2}\frac{C z_0^2}{2[A(\omega)-\sigma]}\right\}^{-1}, \tag{70}$$

which, in contrast to the permittivity, is not spatially dispersive, i.e., $\mu(k_z,\omega) = \mu(\omega)$ for the applied approximation of $k_z z_0 \ll 1$.

In order to apply the developed formalism to a realistic CW metamaterial,[20] numerical FMM simulations of a slab of periodically arranged CW meta-molecules have been performed. The lateral periods of $\Lambda_x = 600\,\text{nm}$ and $\Lambda_y = 500\,\text{nm}$ have been used. The CW meta-molecule has a single wire extension in x- and y-direction of 400 nm and 150 nm, respectively, and a thickness of 40 nm in z-direction. Both wires are 25 nm separated, yielding a total slab thickness of 105 nm. The far field reflection and transmission spectra for such a slab are shown in Fig. 2(b). On the basis of these spectral results the effective parameters for normal incidence can be estimated.[21] These are approximative since they are associated to homogeneous slab for which spatial dispersion is neglected. Considering the spatial dispersion in Eqs. (63, 66, 69) one can anticipate that this is solely connected to the parameter z_0. This is due to several reasons. At first, $2z_0$ is the dimension of meta-molecules in propagation direction, enabling the observation of second order multipoles. It can be easily proven that for vanishing z_0, second order multipoles become zero. Second, z_0 is related to the phase accumulation of the electric field evolution upon propagating through the metamaterial according to Eqs. (59, 60). This field difference is essential in order to excite any anti-symmetric currents that are required for second order excitations, i.e., magnetic dipoles. Due to the dependence of multipole moments on the

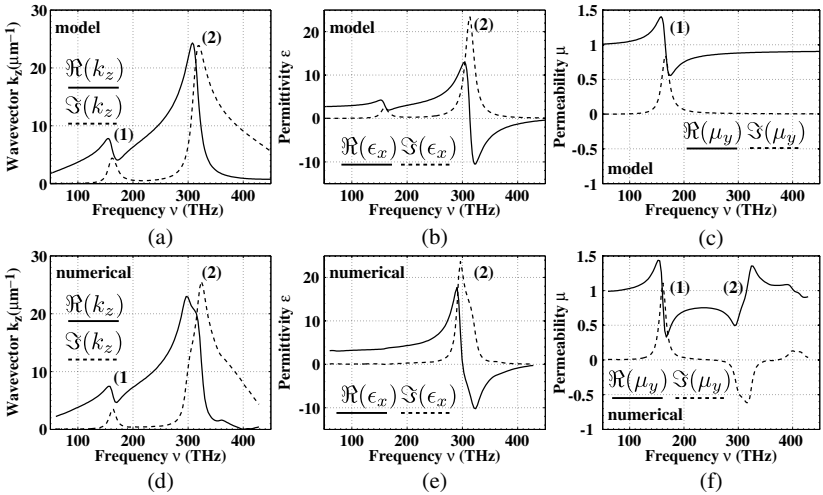


Fig. 3. In the first row the (a) dispersion relation, (b) permittivity, and (c) permeability values for the multipole model are shown, whereas on the second row (d, e, f) the respective numerically obtained observables are shown for a slab of periodically arranged CW meta-molecules. Solid lines represent the real part, while dashed lines label the imaginary part.

fields, the non-local response translates into the dispersion relation as well as into the effective material parameters.

In order to evaluate the dispersion relation (65), the permittivity (67) and the permeability (70) the unknown values: ω_0, γ, σ and C have to be determined. This can be achieved by fitting one of the quantities: k_z, ϵ, or μ to the respective numerical result. The remaining two quantities are then fixed and can be computed without further adaption. For the results shown in Fig. 3(a–f) the numerical dispersion relation has been used to fix the parameters of the model, whereas the permittivity and the permeability follow directly. When comparing analytical and numerical results major discrepancies can be only observed in the magnetic permeability [Fig. 3(c, f)]. There an anti-resonance feature of the numerical data is not present in the multipole model. To conclude this comparison for the CW metamaterial, we wish to state that all relevant numerically observed effects can be simultaneously observed in the multipole model.

After the comprehensive presentation of the approach at the example of the CW meta-molecule, which allowed us to calculate experimentally accessible and relevant quantities, the model can be applied to describe various other meta-molecules.[20,22] A major advantage of the presented

modeling is that numerically obtained results can be unambiguously linked to a physical meaning beyond the phenomenological interpretation of observed effects in terms of multipolar excitations. It has even been shown that quantitative results can be achieved. But since the model is based on parameters that have to be determined, the prediction of new meta-molecule properties is only possible to a certain extent, as will be shown in the following section.

Before this is done, a brief explanation why the presented model for a certain metamaterial has to be fitted to either experimental or numerical data is given. For ordinary materials operating in the electric dipole limit a similar oscillator ansatz is typically chosen, where usually one uncoupled oscillator is sufficient. In order to fix the material parameters, i.e., the permittivity, this model has to be compared to experimental measurements, preferably by means of reflection and transmission measurements at thin films.[13] As an example, the plasma frequency can be considered as one of such parameters being specific for each noble metal. However, depending on the microscopic carrier dynamics these parameters are different for each material due to the specific intrinsic carrier mobility. Performing the averaging for metamaterials rather than natural materials, the intrinsic carrier dynamics can be adjusted by geometrically controlling the meta-molecule's shape. As a result, the oscillator parameters will be different for each metamaterial. Strictly speaking, the presented model requires that for each metamaterial these parameters have to be determined as well. By changing the mesoscopic geometry, the entire metamaterial is modified, and thus requires a new fitting of the parameters. To illustrate that the presented model is just a consequent continuation of the description of natural materials, we set $z_0 = 0$, which is equivalent to considering only one oscillator in the origin which makes σ meaningless. With this step all second order multipole moments [Eqs. (51, 52)] vanish. The magnetic permeability [Eq. (70)] becomes unity and the electric permittivity takes the form of

$$\epsilon(\omega) = 1 + \frac{C}{A(\omega)} = 1 + \frac{C}{\omega_0^2 - \omega^2 - i\omega\gamma}, \tag{71}$$

which, depending on the presence of ω_0 in $A(\omega)$, provides the well-known Drude metal or a polar material featuring a Lorentzian resonance. Although the CW meta-molecule described in this section is rather simple, several important physical features have already been observed, such as optical

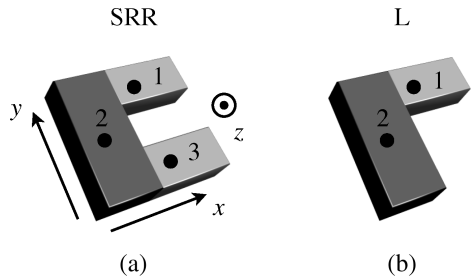


Fig. 4. Illustration of (a) the SRR, and (b) the L meta-molecule.[28] Both meta-molecules are arranged in the xy-plane, hence the propagation vector is assumed to be normal to the surface, i.e., parallel to the z-direction. The dots denote the carrier configuration assumed for the analytical model.

magnetism, the occurrence of spatial dispersion and higher order dispersion branches, which help to facilitate the understanding of metamaterials by means of fundamental electrodynamic principles.

4. Multipolar Properties of Planar Meta-Molecules

In the following, metamaterials composed of planar meta-molecules are investigated.[23] The major part of this section is devoted to the question, to what extent our model allows us to predict the properties of new metamaterials. Hence, we start with an original meta-molecule which later on will be modified. As will be shown, the parameters of the modified meta-molecule will be predicted correctly by the model. We will extend these modifications to cover effects associated with asymmetric transmission for circularly polarized light.[24–27] The planar SRR,[6] Fig. 4(a), serves as the original structure. The polarization eigenstates of this meta-molecule are shown to be linearly polarized. Next, it will be revealed that for particular geometrical modifications [Fig. 4(b)] of this meta-molecule the polarization eigenstates change toward elliptical polarization.

In order to describe the internal carrier dynamics for the SRR, conceptually consisting of three connected straight wire pieces, the following microscopic carrier configuration is assumed

$$\begin{aligned} \mathbf{r}_1^+ &= [0, y_0, 0]^T, & \mathbf{r}_1^- &= [\xi_1(t), y_0, 0]^T, \\ \mathbf{r}_2^+ &= [-x_0, 0, 0]^T, & \mathbf{r}_2^- &= [-x_0, \xi_2(t), 0]^T, \\ \mathbf{r}_3^+ &= [0, -y_0, 0]^T, & \mathbf{r}_3^- &= [\xi_3(t), -y_0, 0]^T. \end{aligned} \tag{72}$$

A set of three coupled microscopic oscillators is proposed to account for the coupling between the conductively connected wires

$$\begin{aligned}
\frac{\partial^2}{\partial t^2}\xi_1(t) + \gamma_1\frac{\partial}{\partial t}\xi_1(t) + \omega_{01}^2\xi_1(t) + \sigma_{21}\xi_2(t) &= -\frac{q_1}{m}\mathfrak{E}_x(z,t),\\
\frac{\partial^2}{\partial t^2}\xi_2(t) + \gamma_2\frac{\partial}{\partial t}\xi_2(t) + \omega_{02}^2\xi_2(t) + \sigma_{21}\xi_1(t) - \sigma_{23}\xi_3(t) &= -\frac{q_2}{m}\mathfrak{E}_y(z,t),\\
\frac{\partial^2}{\partial t^2}\xi_3(t) + \gamma_3\frac{\partial}{\partial t}\xi_3(t) + \omega_{03}^2\xi_3(t) - \sigma_{23}\xi_2(t) &= -\frac{q_3}{m}\mathfrak{E}_x(z,t).
\end{aligned} \tag{73}$$

It is assumed that wire 1 and wire 3 are identical, but different than the SRR base wire 2. Hence, Eq. (73) can be simplified with: $\omega_{01} = \omega_{03} \equiv \omega_x$, $\omega_{02} \equiv \omega_y$, $\gamma_1 = \gamma_3 \equiv \gamma_x$, $\gamma_2 \equiv \gamma_y$. $\sigma_{12} = \sigma_{23} \equiv \sigma$, and $q_1 = q_3 \equiv q_x$, $q_2 \equiv q_y$. Next, the microscopic multipole contributions Eqs. (50–52) have been evaluated for the carrier distribution [Eq. (72)]. The resulting molecular SRR multipole moments are

$$\mathbf{p}_{\mathrm{SRR}} = \begin{pmatrix} q_x[\xi_1(t) + \xi_3(t)] \\ q_y\xi_2(t) \\ 0 \end{pmatrix}, \tag{74}$$

$$\hat{\mathbf{Q}}_{\mathrm{SRR}} = 3\begin{pmatrix} -q_x[\xi_1^2(t) + \xi_3^2(t)] & -q_x y_0[\xi_1(t) - \xi_3(t)] + q_y x_0\xi_2(t) & 0 \\ -q_x[\xi_1(t) - \xi_3(t)] + q_y x_0\xi_2(t) & -q_y\xi_2^2(t) & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{75}$$

$$\mathbf{m}_{\mathrm{SRR}} = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ q_x y_0\frac{\partial}{\partial t}[\xi_1(t) - \xi_3(t)] + q_y x_0\frac{\partial}{\partial t}\xi_2(t) \end{pmatrix}. \tag{76}$$

For arbitrary carrier dynamics $\xi_j(t)$, $j \in [1,2,3]$, molecular electric dipole moments in x- and y-direction, a magnetic moment in z-direction and electric quadrupole contributions Q_{ij}, $[i,j] \in [x,y]$ are induced. Hence, even after the transition to averaged quantities [Eqs. (56, 57)] the wave equation (54), describing x-polarized electrical fields propagating in z-direction, has no contributions stemming from second order multipole moments. As can be easily verified, these conclusions hold for y-polarized electric fields

propagating in z-direction as well. Thus, the entire optical response is described by the macroscopic polarization

$$\mathbf{P}(z,\omega) = \eta \mathbf{p}_{\mathrm{SRR}}(z,\omega) = -\eta \begin{pmatrix} q_x[\xi_1(\omega) + \xi_3(\omega)] \\ q_y \xi_2(\omega) \\ 0 \end{pmatrix}. \tag{77}$$

Similar to the CW meta-molecule the oscillator elongations according to the set of equations (73) are simply obtained in Fourier domain [Eq. (41)]. For x-polarization one obtains

$$\xi_1(\omega) = \xi_3(\omega) = -\tfrac{q_x}{m} \tfrac{1}{A_x(\omega)} E_x(z,\omega), \tag{78}$$

$$\xi_2(\omega) = 0, \tag{79}$$

whereas for y-polarized electric fields we have

$$\xi_1(\omega) = -\xi_3(\omega) = -\frac{q_y}{m} \frac{\sigma}{A_x(\omega)A_y(\omega) - 2\sigma^2} E_y(z,\omega), \tag{80}$$

$$\xi_2(\omega) = -\frac{q_y}{m} \frac{A_x(\omega)}{A_x(\omega)A_y(\omega) - 2\sigma^2} E_y(z,\omega), \tag{81}$$

$$\text{with: } A_{x,y}(\omega) \equiv \omega_{x,y}^2 - \omega^2 - i\omega\gamma_{x,y}. \tag{82}$$

Considering Eqs. (77–81) it becomes obvious that x-polarized electrical fields induce only a polarization in x-direction, because $\xi_2(\omega) = 0$. In turn, y-polarized light induces a polarization solely in y-direction due to the annihilation of the carrier dynamics in wires 1 and 3 [$\xi_1(\omega) = -\xi_3(\omega)$]. In other words, the electric permittivity [Eq. (40)] is *diagonal* and consequently the polarization eigenstates are *linearly polarized*

$$\hat{\epsilon}(\omega) = \begin{pmatrix} \epsilon_{xx}(\omega) & 0 & 0 \\ 0 & \epsilon_{yy}(\omega) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{83}$$

$$\epsilon_{xx}(\omega) = 1 + \frac{2\eta q_x^2}{m\epsilon_0} \frac{1}{A_x(\omega)}, \tag{84}$$

$$\epsilon_{yy}(\omega) = 1 + \frac{\eta q_y^2}{m\epsilon_0} \frac{A_x(\omega)}{A_x(\omega)A_y(\omega) - 2\sigma^2}. \tag{85}$$

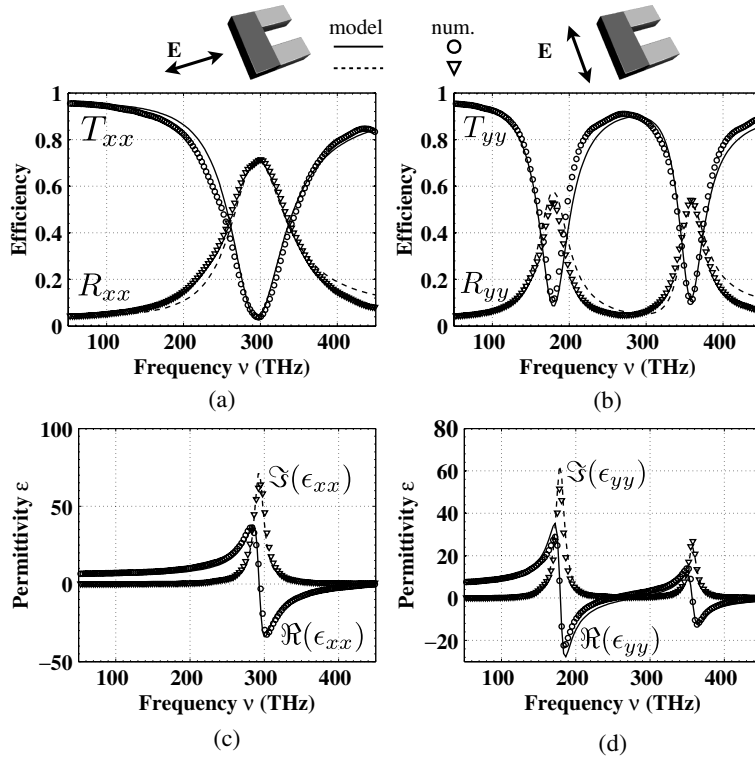


Fig. 5. The SRR far field spectra obtained by numerical simulation (T circles, R triangles) and for the developed model (T solid, R dashed) for (a) x-polarization and (b) for y-polarization. (c, d) The corresponding complex permittivities obtained by the parameter retrieval of the numerical spectra in comparison to predictions from Eq. (83) for the respective polarization direction.

In turn, the magnetic permeability is unity for both polarization directions as expected for media comprising electric dipole interaction only. With Eqs. (84, 85) it is now possible to calculate the reflection and transmission coefficients by applying a standard transfer matrix formalism for films.[29] For reference, the SRR spectrum has been computed at first numerically with the FMM and secondly by calculating and fitting the spectra of a metamaterial slab incorporating the permittivity derived above. Spectra for polarization in x- and y-directions are shown in Fig. 5(a, b), respectively. The SRR metamaterial has been selected similar to the one investigated in the literature.[6,23] The SRR wires are 200 nm long, 50 nm wide, the SRR

base has a length of 100 nm and a width of 80 nm. The entire SRR thickness is 25 nm and it is made of gold.[13] A periodical arrangement with a period of 400 nm in both x- and y-direction has been assumed. The thickness of the metamaterial corresponds to the thickness of the SRR. In addition, the permittivity has been obtained from the numerical spectra by the parameter retrieval that has been used for the CW pairs before. Comparing this retrieved permittivity to the one that has been found in the spectral fit for the reflection and transmission spectra almost perfect coincidence is observed [Fig. 5(c, d)]. Hence, the developed formalism can be applied as a kind of parameter retrieval procedure. In addition to the permittivity, the formalism may be useful in order to quantify the coupling between the SRR wire pieces.

Next, the SRR meta-molecule is modified in the following yielding the L meta-molecule, see Fig. 4(b) where one of the SRR arms (3) has been removed. This modification prevents the mutual cancellation of the carrier dynamics in wires 1 and 3 when compared to the original SRR meta-molecule. The solutions for the oscillator elongations are

$$\xi_1^x(\omega) = -\frac{q_x}{m}\frac{A_y(\omega)}{A_x(\omega)A_y(\omega) - \sigma^2}E_x(z,\omega), \tag{86}$$

$$\xi_2^x(\omega) = -\frac{q_x}{m}\frac{\sigma}{A_x(\omega)A_y(\omega) - \sigma^2}E_x(z,\omega), \tag{87}$$

for polarization in x-direction and

$$\xi_1^y(\omega) = -\frac{q_y}{m}\frac{\sigma}{A_x(\omega)A_y(\omega) - \sigma^2}E_y(z,\omega), \tag{88}$$

$$\xi_2^y(\omega) = -\frac{q_y}{m}\frac{A_x(\omega)}{A_x(\omega)A_y(\omega) - \sigma^2}E_y(z,\omega), \tag{89}$$

in y-direction, respectively. Since one of the SRR arms has been removed one oscillator elongation can be set to zero. Hence, an electric field polarized in x-direction induces a cross-polarization in y-direction and *vice versa*. For the permittivity that is calculated from the polarization [Eq. (77)] upon setting $\xi_3(\omega) = 0$ these cross-polarizabilities result in non-zero off-diagonal

elements

$$\hat{\epsilon}(\omega) = \begin{pmatrix} \epsilon_{xx}(\omega) & \epsilon_{xy}(\omega) & 0 \\ \epsilon_{yx}(\omega) & \epsilon_{yy}(\omega) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{90}$$

$$\epsilon_{xx}(\omega) = 1 + \frac{q_x^2 \eta}{m\epsilon_0} \frac{A_y(\omega)}{A_x(\omega)A_y(\omega) - \sigma^2}, \tag{91}$$

$$\epsilon_{yy}(\omega) = 1 + \frac{q_y^2 \eta}{m\epsilon_0} \frac{A_x(\omega)}{A_x(\omega)A_y(\omega) - \sigma^2}, \tag{92}$$

$$\epsilon_{yx}(\omega) = \epsilon_{xy}(\omega) = \frac{q_x q_y \eta}{m\epsilon_0} \frac{\sigma}{A_x(\omega)A_y(\omega) - \sigma^2}. \tag{93}$$

If $A_x(\omega) \neq A_y(\omega)$, which is fulfilled for different geometries of wires 1 and 2, the permittivity tensor cannot be diagonalized. Thus, the polarization eigenstates are in general *elliptically polarized*.[24] In order to verify these analytical considerations the spectral response of the L structure has been numerically modeled by using FMM. For media which are described by a permittivity according to Eq. (90) the reflectivity and transmissivity have to be calculated requiring more sophisticated methods. Here a homogeneous slab with this permittivity and the thickness corresponding to the thickness of the meta-molecules has been simultaneously modeled.[19] In a first step, the parameters of the SRR that have been found before are applied while considering the vanishing of one wire. The results for both polarization directions are shown in Fig. 6(a, b, d, e). As can be seen, the resonance positions as well as the resonance widths are predicted with only minor differences in the resonance strength, see Fig. 6(a, c). Second, the model parameters have been fitted the numerical spectra, directly. This improves the coincidence between the spectra predicted by the model and the numerical spectra toward almost perfect agreement. Considering the cross-polarized quantities $T_{ij},\ R_{ij},\ i \neq j$, agreement with sufficient accuracy is observed even for the SRR parameters [Fig. 6(b,e)]. As a further outcome of the spectral adaption, the entire permittivity tensor can be directly obtained [Fig. 6(c, f, i)] which is, for media that do not possess linearly polarized eigenstates, a non-trivial task. In passing, it is mentioned that the anti-resonance observed in Fig. 6(i) can be fully explained by the mutual interplay of the coupled carriers.[23] Finally, it will be proven that asymmetric transmission[24–27] for circularly polarized light occurs, as expected from the shape of the permittivity tensor [Eq. (90)]. The connection between left-("−") and right-("+") handed circularly and linearly polarized light may

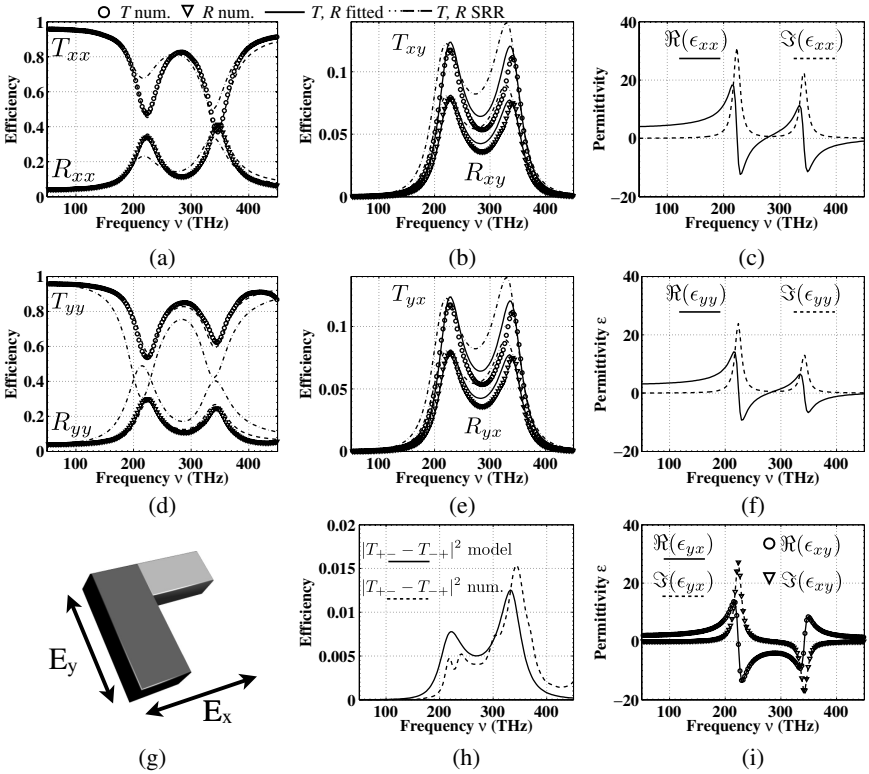


Fig. 6. The (a) far field spectra, (b) the cross-polarized spectral quantities T_{xy}, R_{xy}, and (c) the permittivity for x-polarization. (d, e, f) The respective results for y-polarization. The spectra predicted by the SRR values in (a, b, d, e) are labeled by dashed dotted lines, whereas the directly fitted values correspond to the solid lines. The permittivities (c, f, i) have been calculated with Eq. (90) using the directly fitted values. (g) The L structure made of two wires with different widths together with the two polarization directions of interest are shown. (h) The asymmetric transmission in terms of a difference between T_{+-} and T_{-+} is shown for the numerical and the adapted transmissivities.

be written as[27]

$$\begin{pmatrix} T_{++} & T_{+-} \\ T_{-+} & T_{--} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} (T_{xx} + T_{yy}) + i(T_{xy} - T_{yx}) & (T_{xx} - T_{yy}) - i(T_{xy} + T_{yx}) \\ (T_{xx} - T_{yy}) + i(T_{xy} + T_{yx}) & (T_{xx} + T_{yy}) - i(T_{xy} - T_{yx}) \end{pmatrix}. \tag{94}$$

Considering Eq. (94) it becomes obvious that for non-zero, cross-polarized transmission coefficients T_{xy}, T_{yx} the difference $T_{+-} - T_{-+}$ is non-zero, too, requiring that $T_{xy} + T_{yx} \neq 0$. For the investigated L meta-molecule the results for the modulus of the difference between T_{+-} and T_{-+} are shown in Fig. 6(h). The two lines correspond to the results for the numerical transmission and the ones predicted by the model. Again all main features are correctly described by the developed formalism, see Fig. 6(h). In passing, it is mentioned that typically the effects of asymmetric transmission in planar meta-molecules are rather small compared to meta-molecules that additionally exhibit chiral properties due to an extension and particular shape in the propagation direction.[27]

To conclude this section, it has been shown that planar meta-molecules upon normal incidence possess electric dipole response only, which simplifies the model significantly. These findings are consistent with experimental observations, where magnetic effects, i.e., a magnetic resonance, could be only observed for angular incidence.[6] Furthermore, the transition toward elliptically polarized eigenstates has been performed by geometrical variations of the meta-molecule. In doing so the optical response could be estimated solely from knowing the parameters of an original structure,[23] here the SRR. The formalism has been utilized to determine the permittivity tensor which can be accessed by comparing with experimental or theoretical far field intensities without requiring the spectral phase. The developed description supports and simplifies the physical understanding of asymmetric transmission effects by the simple mutual interaction of electric dipoles. Beyond this qualitative understanding it has also been shown that quantitative agreement can be achieved.

It is furthermore mentioned that the same procedure has also been examined for an S-shaped meta-molecule where similar results have been obtained.[23] As an important physical property, all derived permittivities are consistent with the Casimir-Onsager relations[30–32] which require the symmetry in the permittivity tensor $\epsilon_{xy} = \epsilon_{yx}$ which can be proven by considering Eq. (93) or Fig. 6(i).

5. Multipole Near-Field Decomposition for Meta-Molecules

In the last section the linear properties of meta-molecules are investigated beyond their far field wave propagation properties, i.e., k, ϵ, μ. Therefore the near-field distributions of isolated meta-molecules are considered and

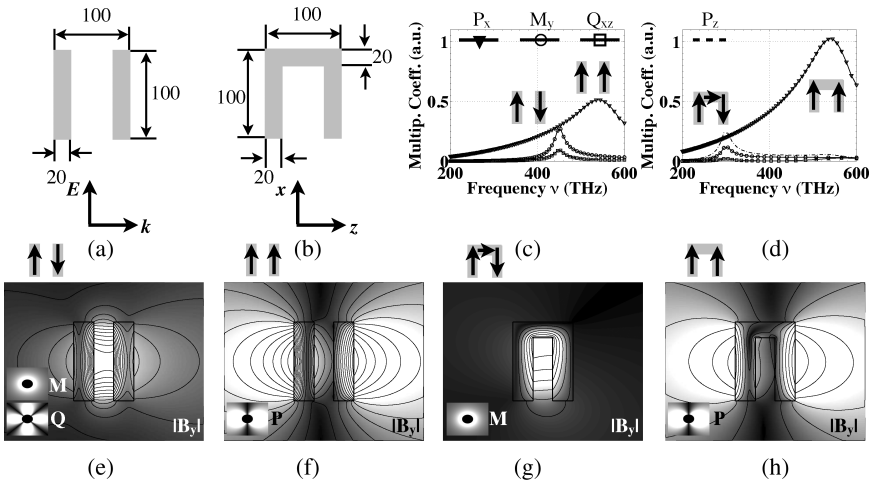


Fig. 7. Dimensions and excitation conditions for (a) the CW and (b) the SRR. Spectrally resolved evaluation of the overlap integral between the analytical multipole field and the numerical near-field for (c) the CW and (d) the SRR. (f–h) Magnetic field distribution for the resonance positions labeled in (c, d) for the two meta-molecules. The insets show the y-component of the magnetic field for (e) top: the magnetic dipole, (e) bottom: the electric quadrupole, (f): the electric dipole, (g): the magnetic dipole, and (h): the electric dipole moment.

their radiation patterns are rigorously analyzed by a multipolar field decomposition technique. Starting point is the vector potential[5,23]

$$\mathbf{A}(\mathbf{r},t) = \frac{\mu_0}{4\pi}\int_{V_0} d\mathbf{r}'^3 \frac{\mathrm{j}(\mathbf{r}-\mathbf{r}', t-\frac{|\mathbf{r}-\mathbf{r}'|}{c})}{|\mathbf{r}-\mathbf{r}'|}, \tag{95}$$

which is driven by the microscopic current j and is a solution of

$$\Delta\mathbf{A}(\mathbf{r},t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A}(\mathbf{r},t) = -\mu_0 \mathrm{j}(\mathbf{r},t). \tag{96}$$

The microscopic electric and magnetic fields can be calculated in the Fourier domain (41) assuming the Lorentz gauge[5] according to

$$\mathfrak{E}(\mathbf{r},\omega) = i\frac{c^2}{\omega} e\nabla\times\nabla\times\mathbf{A}(\mathbf{r},\omega), \tag{97}$$

$$\mathfrak{B}(\mathbf{r},\omega) = \nabla\times\mathbf{A}(\mathbf{r},\omega). \tag{98}$$

Now, the microscopic current density $\mathrm{j}(\mathbf{r},t)$ is expanded in a Taylor series up to second order. The first order term yields the vector potential for the electric dipole radiation. The second order expansion term consists

of two contributions: the electric quadrupole and the magnetic dipole radiation.[5,23,33] Hence, both moments represent second order multipole contributions.

In order to investigate the strength of the respective multipole moment contributing to the entire radiation pattern the scattered field for meta-molecules will be decomposed into multipole contributions in the following. There, the overlap integral between a certain field component for the respective multipole moment and the numerically calculated or experimentally observed field pattern will be performed. Caused by the orthogonality of both the radial and the azimuthal part of the multipolar field components, it suffices to elaborate the overlap integral on a ring outside the meta-molecule. Recently, this method has been proposed[34] for two-dimensional meta-molecules but can be straightforwardly extended toward three dimensions. Results for the evaluation of the overlap integral for the near-fields of an isolated CW and SRR are shown in Fig. 7(c, d). It can be recognized that for the symmetric resonance, corresponding to the resonance in the electric permittivity, the electric dipole moments are parallel to the x-polarized incident electric field. For the anti-symmetric resonance, in turn, second order multipoles are maximized. To illustrate the underlying numerically calculated near-field distributions the y-component of the magnetic field patterns for the frequencies corresponding to the maxima in Fig. 7(c, d) are shown in Fig. 7(e–h) for both meta-molecules. The insets correspond to the magnetic field components calculated from Eq. (98) for the respective multipole moment[34] by expanding Eq. (95). It can be observed that the numerical near-field patterns can be excellently described by the excitation of elementary multipoles.

With these considerations it could be proven that besides calculating optical far field observables multipole contributions can be successfully applied to describe the near-field optical response of meta-molecules as well. Furthermore, this rigorous method unambiguously justifies the role of multipole excitations that has been assumed in the previous sections.

6. Summary and Outlook

The primary goal of this chapter was to show how elementary textbook electrodynamics can be applied to enter the field of metamaterials where emphasis has been put on the understanding of induced carrier dynamics, the appearance of higher order dispersion relation branches, the appearance of electric and magnetic effects as well as a clear picture to read near-field

patterns of meta-molecules. Unlike other electrodynamical approaches that have been successfully established,[11] here the language of multipoles has been shown to be an appropriate and versatile tool for this task.

Starting from a recapitulation of the transition between microscopic and macroscopic Maxwell's equations, the role of multipolar contributions in the description of light-matter interaction has been presented. From this point of view, artificial metamaterials can be treated as ordinary matter but featuring beyond the electric dipole interaction also second order multipole interactions. For the determination of the carrier dynamics being essential for computing the multipole moments a simple oscillator model has been proposed. Thus, characteristic quantities, such as the dispersion relation, can be obtained and compared to numerical or experimental data. It has also been shown that the approach can be utilized to predict the optical response of meta-molecules that refer to an original, known meta-molecule but are subjected to particular geometrical modifications. Therefore meta-molecule modifications changing the polarization eigenstates from linear to elliptical have been investigated for the purpose of observing asymmetric transmission behavior. While these two examples present applications for the far field description, the near-field response of the SRR and the CW meta-molecule has been considered in the last section. There, an approach has been introduced to unambiguously reveal multipolar contributions from the near-field patterns. All presented investigations suggest that the linear optical response of metamaterials can be understood in terms of elementary multipoles. Besides, it is mentioned that the seminal work of Pendry[17] starts similarly with the considerations of averaged fields before investigating wave propagation characteristics, such as optical magnetism of artificial matter. Here, metamaterials are discussed in the context of multipole moments, which is a different approach yielding consistent results with respect to the functional shape of the wave propagation parameters as well as for conclusions on the occurrence of magnetic effects.

Beyond these linear optical properties, multipole induced nonlinearities, known from the pioneering works in nonlinear optics,[35] can be investigated[22] where the comparison with linear optical properties suffices to predict the second order nonlinear optical response as well.

Acknowledgments

Financial support by the Federal Ministry of Education and Research (PhoNa, MetaMat) and the State of Thuringia within the ProExcellence program (MeMa) is acknowledged.

References

1. H. A. Lorentz, *Proc. Roy. Acad. Amsterdam.* **5**, 254 (1902).
2. L. Rosenfeld, *Theory of Electrons.* (Amsterdam: North-Holland Pub., 1951).
3. P. Mazur and B. R. A. Nijboer, *Physica.* **19**, 971 (1953).
4. G. Russakoff, *Am. Journ. Phys.* **38**, 1188 (1970).
5. J. D. Jacksson, *Classical Electrodynamics.* (Wiley, New York, 1975).
6. S. Linden, C. Enkrich, M. Wegener, J. Zhou, T. Koschny, and C. M. Soukoulis, *Science* **306**, 1351 (2004).
7. V. M. Shalaev, W. Cai, U. K. Chettiar, H.-K. Yuan, A. K. Sarychev, V. P. Drachev, and A. V. Kildishev, *Opt. Lett.* **30**, 3356 (2005).
8. C. C. Yan, *Found. Phys.* **25**, 491 (1995).
9. E. B. Graham, J. Pierrus, and R. E. Raab, *J. Phys. B: At. Mol. Opt. Phys.* **25**, 4673 (1992).
10. O. L. de Lange and R. E. Raab, *Phys. Rev. E.* **71**, 036620 (2005).
11. S. Serdyukov, I. Semchenko, S. Tretyakov, and A. Sihvola, *Electromagnetics of Bianisotropic Media.* (Gordon and Breach, Amsterdam, 2001).
12. P. S. Pershan, *Phys. Rev.* **130**, 919 (1963).
13. P. B. Johnson and R. W. Christy, *Phys. Rev. B.* **6**, 4370 (1972).
14. A. Buckingham and R. Disch, *Proc. R. Soc. London* **273**, 275 (1963).
15. X. Zhang, M. Davanco, Y. Urzhumov, G. Shvets, and S. R. Forrest, *Phys. Rev. Lett.* **101**, 267401 (2008).
16. C. Menzel, T. Paul, C. Rockstuhl, T. Pertsch, S. Tretyakov, and F. Lederer, *Phys. Rev. B.* **81**, 035320 (2010).
17. J. B. Pendry, A. J. Holden, D. J. Robbins, and W. J. Stewart, *IEEE Trans. Microwave Theory Tech.* **47**, 2075 (1999).
18. M. Burresi, D. V. Oosten, T. Kampfrath, H. Schoenmaker, R. Heifeman, A. Leinse, and L. Kuipers, *Science* **326**, 550 (2009).
19. L. Li, *J. Opt. A: Pure Appl. Opt.* **5**, 345 (2003).
20. J. Petschulat, C. Menzel, A. Chipouline, C. Rockstuh, A. Tünnermann, F. Lederer, and T. Pertsch, *Phys. Rev. A.* **78**, 043811 (2008).
21. D. R. Smith, D. C. Vier, T. Koschny, and C. M. Soukoulis, *Phys. Rev. E.* **71**, 036617 (2005).
22. J. Petschulat, C. Menzel, A. Chipouline, C. Rockstuh, A. Tünnermann, F. Lederer, and T. Pertsch, *Phys. Rev. A.* **80**, 063828 (2009).
23. J. Petschulat, C. Menzel, A. Chipouline, C. Rockstuhl, T. Paul, A. Tünnermann, F. Lederer, and T. Pertsch, *Phys. Rev. B.* **82**, 075102 (2010).
24. S. V. Zhukovsky, A. V. Novitsky, and V. M. Galynsky, *Opt. Lett.* **34**, 1988 (2009).
25. V. A. Fedotov, A. S. Schwanecke, N. I. Zheludev, V. V. Khardikov, and S. L. Prosvirnin, *Nano Lett.* **7**, 1996 (2007).
26. R. Singh, E. Plum, C. Menzel, C. Rockstuhl, A. K. Azad, R. A. Cheville, F. Lederer, W. Zhang, and N. I. Zheludev, *Phys. Rev. B.* **80**, 153104 (2009).
27. J. Zhou, J. Dong, B. Wang, T. Koschny, M. Kafesaki, and C. M. Soukoulis, *Phys. Rev. B.* **79**, 121104(R) (2009).

28. B. K. Canfield, S. Kujala, M. Kauranen, K. Jefimovs, T. Vallius, and J. Turunen, *Appl. Phys. Lett.* **86**, 183109 (2005).
29. M. Born and E. Wolf, *Principles of Optics*. (Cambridge, New York, 1999).
30. L. Onsager, *Phys. Rev.* **37**, 405 (1931).
31. H. B. G. Casimir, *Rev. Mod. Phys.* **17**, 343 (1945).
32. S. Tretyakov, A. Shivola, and B. Jancewicz, *J. Electromagn. Waves Appl.* **16**, 573 (2002).
33. R. E. Raab and O. L. D. Lange, *Multipole Theory in Electromagnetism*. (Clarendon, Oxford, 2005).
34. J. Petschulat, J. Yang, C. Menzel, C. Rockstuhl, A. Chipouline, P. Lalanne, A. Tünnermann, F. Lederer, and T. Pertsch, *Opt. Express* **18**, 14454–14466 (2010).
35. N. Bloembergen, *Nonlinear Optics*. (Benjamin Press, New York, 1965).

AMPLIFICATION AND LASING WITH SURFACE-PLASMON POLARITONS

Israel De Leon* and Pierre Berini†

*School of Information Technology and Engineering,
University of Ottawa, 161 Louis Pasteur,
Ottawa, Ontario, K1N 6N5, Canada
ideleon@site.uottawa.ca

†Department of Physics, University of Ottawa,
150 Louis Pasteur, Ottawa,
Ontario, K1N 6N5, Canada
berini@site.uottawa.ca

§Spectalis Corp., PO Box 72029, Kanata North RPO,
Ottawa, Ontario, K2K 2P4, Canada

Surface plasmon-polaritons offer useful properties that find applications in a broad range of scientific and engineering fields. However, many applications face practical limitations imposed by the intrinsic energy losses experienced by this wave at optical and near-infrared wavelengths. During the last decade, the topic of surface plasmon-polariton amplification has experienced a tremendous growth as it offers a viable venue to eliminate the wave's losses without compromising other key attributes. This review summarises the major theoretical and experimental progress achieved in the amplification of this wave. It discusses the topics of amplified spontaneous emission, stimulated emission, and lasing within the context of surface plasmon-polaritons.

1. Introduction

Surface plasmon-polaritons (SPPs) are electromagnetic surface waves formed through the interaction of photons with free electrons at the surface of some metals bounded by a vacuum or a dielectric.[1] These surface waves

offer interesting properties that have found applications in a broad range of scientific and engineering fields such as spectroscopy,[2,3] nano-photonics,[4,5] imaging,[6] non-linear optics,[7] and biosensing.[8] SPPs dissipate their energy through a natural interaction with the atomic structure of the metal; thus, they are intrinsically lossy. In many cases, excessive losses encountered at visible and near-infrared wavelengths render applications impractical. To alleviate this problem researchers have proposed lowering the SPP losses through optical amplification, which in principle can be achieved by adding optical gain to the dielectric or dielectrics bounding the metal.

The amplification of SPPs has been a subject of study for three decades and has experienced an enormous growth during the last one. This chapter reviews the existing body of literature on the topic and summarises the most important theoretical and experimental progress. It also discusses potential applications. The material presented here is representative of the progress achieved to date.

A variety of structures can support SPPs and a main distinction is made between *localised* SPPs supported by metallic nanoparticles (with sizes smaller than the free-space wavelength) and *propagating* SPPs supported by planar metallic surfaces. Also, metallic nano-cavities can support SPP resonant modes that might be categorised separately. Due to their particular properties, SPPs supported by the different structures find different applications and gain requirements. This review is organised according to these three types of plasmonic structures.

Naturally, there exist more complex metallic structures that support SPPs. They are generally formed by composite metallic geometries designed to tailor the structure's optical response, in which SPPs play a crucial role. Such structures fall under the category of metamaterials.[14,15] Metamaterials incorporating gain media have received considerable attention[16–21] because they offer a possibility to overcome practical restrictions imposed by losses in phenomena such as transparency, super-resolution, and enhanced reflectance and transmittance over small bandwidths. Nevertheless, a proper review of active metamaterials falls out of the scope of this chapter and is not included.

2. Planar Metallic Surfaces

Most of the studies conducted on SPP amplification consider propagating SPPs supported by planar structures. Specific plasmonic modes are supported by different planar geometries; yet, for the cases discussed in

this section, we shall categorise these geometries in three groups. The first group comprises structures that support *single-interface* SPPs. For example, an interface between an optically thick (or semi-infinite) metal and a dielectric, which supports a non-radiative SPP mode, or an optically thin metal film in Kretschmann-Raether configuration,[1] which supports a leaky SPP mode on one of the metal surfaces. The second group comprises the thin metal film[9–11] and stripe[12,13] bounded by index-symmetric dielectrics. These structures support low-loss modes with symmetric field distribution termed *long-range* SPPs (LRSPPs) and high-loss modes with asymmetric field distribution termed *short-range* SPP (SRSPP). Finally, the third group comprises structures consisting of a thin dielectric film cladded by metals. This type of structure, also called metal-insulator-metal structure, supports coupled SPP modes with symmetric and asymmetric field distributions that exhibit strong field confinement but also strong attenuation.[22] In the remainder of this section, the material is organised as much as possible according to these three groups of structures.

2.1. *Single metal-dielectric interface*

The first theoretical study on SPP amplification was reported by Plotz and coworkers[23] in 1979 within the context of enhanced total internal reflection. The authors considered a ∼50 nm-thick silver film attached to a glass prism on one side (Kretschmann-Raether configuration) and in contact with a gain medium on the other, and explored the reflectivity of the glass-metal interface for different gain values. The gain medium, a fluorescent dye, was assumed to be homogeneous and inverted. Their analysis shows that the usual reflectance drop occurring at the SPP resonance angle decreases monotonically as the gain in the dye region increases, reaching eventually a reflectance that is larger than one; i.e., enhanced total reflection. When this condition is met, a singularity can be generated in the reflectivity expression by choosing an appropriate thickness for the metal film. The authors indicated that this singularity is physically associated with the coherent out-coupling of SPPs of infinite propagation length.

Almost a decade later, Sudarkin and Demkovic[24] complemented the work by Plotz and coworkers, and proposed for the first time the possibility of an SPP laser. The authors studied a similar structure and showed that for a specific set of complex dielectric constants describing the metal and gain media, there is an optimum metal-film thickness for which the ratio of the SPP field intensity at the metal surface to the intensity of the excitation

source is maximum. They noted that for silver films thicker than those considered in Ref. (23) the reflectance does not increase monotonically with the gain. Instead, it decreases with increasing gain (a phenomenon later observed experimentally by Seidel *et al.*[25]) until the gain reaches a specific value. As the gain increases beyond this point, the reflectance begins to increase. Despite the non-monotonic reflectance behaviour, the analysis shows that the SPP field increases and the SPP resonance linewidth narrows monotonically with gain increments, as expected for SPP loss reduction. The authors also showed that the Fresnel formulae applied to this can lead to erroneous results of reflectivity when the gain is high and the diameter of the excitation beam is comparable to the SPP range plasmonic structure.

Nezhad and colleagues[26] studied SPP amplification in planar metallic structures bounded by a homogeneous gain medium. The authors derived analytical expressions for the material gain needed for lossless propagation, zero wavefront tilt, and bound propagation limit of single-interface SPPs. Considering silver and an InGaAsP-based gain medium, they showed that lossless propagation of single-interface SPPs is possible at $\lambda = 1550\,\text{nm}$ when the material gain is about $1260\,\text{cm}^{-1}$. The authors also studied numerically amplification of LRSPPs supported by a 40 nm-thick silver film and a 400 nm-wide stripe of the same thickness embedded in an InGaAsP-based gain medium. For these geometries, it was found that the material gain required for lossless propagation is an order of magnitude smaller than for single-interface SPPs. They also investigated the effect of non-uniform gain close to the metal stripe by modelling the gain medium as a layer of thickness h located at a distance d from one of the metal surfaces. Their results showed that for $50 < h < 175\,\text{nm}$ and $0 < d < 200\,\text{nm}$, the gain needed for lossless propagation increases monotonically as the layer moves away from the stripe. The required gain values that varied from 800 to $4500\,\text{cm}^{-1}$ depending on the size and position of the gain layer, which reflects the importance of a good gain-mode overlap.

Avrutsky[27] reported a theoretical study on the behaviour of gain-assisted SPPs and their interaction with nanoscale gratings near the resonance angular frequency, $\omega_r = \omega_p/\sqrt{2}$, ($\omega_p$ being the plasma frequency of the metal). The study considers SPPs supported at the interface between silver and a dielectric medium with gain in the spectral range $330 < \lambda < 370\,\text{nm}$, which lies around the resonance wavelength of silver, $\lambda_r = 337\,\text{nm}$. The author conducted numerical calculations showing that a material gain of $8 \times 10^4\,\text{cm}^{-1}$ can compensate the losses of single-interface SPPs at $\lambda = 349\,\text{nm}$, revealing the unusual properties that result from the

singularity in the dispersion relations that are masked in the passive case by large optical losses. These properties, which include extremely small group velocity (reaching one kilometre per second), large SPP effective index ($n_{\text{eff}} \approx 29$), and high field localisation allow SPPs to interact with features that are an order of magnitude smaller in size than the wavelength. Indeed, these unique properties show potential for characterisation of sub-nanometer surface features and ultra-sensitive optical sensors.

Amplification of SPPs was first demonstrated at far-infrared wavelengths a decade ago by Cho and coworkers.[28,29] In Ref. (29) the authors reported a distributed-feedback (DFB) SPP laser emitting at $\lambda \approx 17\,\mu$m. In part, this plasmonic laser owes its success to the low absorption of metals at far-infrared wavelengths. The device consisted of an alternate sequence of gold (300 nm-thick) and titanium/gold (10/300 nm-thick) stripes deposited on a quantum-cascade (QC) active material. The DFB mechanism was provided by the refractive index difference of the alternating metallic pattern. With this approach, the thickness of epitaxial growth was reduced by a factor of 2.25 and the confinement factor was increased by a factor of 1.8 compared to the respective values of a conventional DFB QC laser with the same active medium. At low temperatures, the laser showed a reduction in the lasing threshold by a factor of two compared to conventional QC lasers operating at similar wavelengths; this allowed record high power of 38 mW at a temperature of 5 K. Furthermore, the laser offered high modal purity with side suppression ratio of approximately 30 dB and operated at a maximum temperature of 240 K with thermal wavelength tenability of $\sim$1 nm/K.

The first demonstration of SPP stimulated emission at visible wavelengths was reported by Seidel and colleagues[25] in 2005. They employed structures consisting of $\sim$40 and $\sim$65 nm-thick silver films attached to a glass-prism on one side and in contact with an optically excited dye solution on the other [see Fig. 1(a)]. The authors studied the change in reflectance at the glass-silver interface induced by the gain medium (excited dye solution) as a function of the incidence angle. Experiments were conducted using two flowing dye solutions, rhodamine 101 and cresyl violet in ethanol, both with a molecular density of $N = 7 \times 10^{17}\,\text{cm}^{-3}$. Transverse magnetic polarised light from a helium-neon laser at $\lambda = 633$ nm was used to probe the SPPs while the dye molecules were excited via SPP pumping at $\lambda = 580$ nm using light from a dye laser incident on the glass-silver interface. The pump beam carried a power of 10 mW and was focused to a spot of 60 μm in diameter; it overlapped completely with the probe beam which was focused to an area

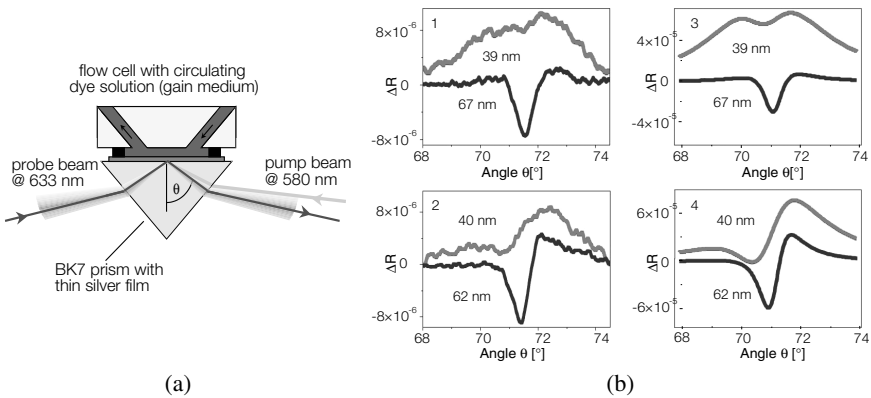


Fig. 1. Adapted figure with permission from Ref. (25). Copyright (2005) by the American Physical Society. (a) Structure and pump/probe arrangement employed by Siedel *et al.* in Ref. (25) for the analysis of SPP stimulated emission. (b) Reflectance change due to SPP stimulated emission as a function of the incidence angle: measurements using rhodamine 101 (1) and cresyl violet (2) dyes in ethanol. The theoretical estimation corresponding to the cases in (1) and (2) are shown in (3) and (4), respectively. The silver-film thickness is indicated for each curve.

approximately twice as large. The probe and pump beams were modulated at 23 Hz and 27 MHz, respectively, and a phase-sensitive detection scheme using two lock-in amplifiers was employed to measure the reflectance change due to SPP stimulated emission. For the 40 nm-thick films, the effect of stimulated SPPs was observed as an increased emission of light into the reflected probe beam; on the other hand, for the 65 nm-thick films it was manifested by the narrowing and deepening of the reflectance dip occurring at the SPP resonance angle [see Fig. 1(b)]. The experimental results agreed qualitatively with a theoretical model that considers the gain medium as a 4-level system and a number of factors characteristic of the experiment, such as the exponential decay of population inversion away from the metal surface and the transverse profile of the probe and pump beams.

Noginov and colleagues[30] studied experimentally SPP amplification in a structure similar to that used by Seidel in Ref. (25). In this case, the gain medium consisted of a 10 μm-thick layer of Rhodamine-6G (R6G) doped polymethyl-methacrylate (PMMA) with a molecular density of $N = 2.2 \times 10^{22}\,\mathrm{cm}^{-3}$, which was dried on the back side of a 40 nm-thick silver film attached on its front side to a glass ($n = 1.784$) prism. The gain medium was optically pumped from the back side using pulsed light at $\lambda = 532$ nm with pulse energy of 18 mJ, pulse duration t_p=10 ns, and

repetition rate $f_p = 10\,\text{Hz}$. TM polarised light from a helium-neon laser at $\lambda = 594\,\text{nm}$ was coupled through the prism to excite (probe) SPPs at the silver-PMMA/R6G interface. The reflected probe light was spectrally filtered using a monochromator tuned to the probe's wavelength and then measured as a function of the incidence angle using a photomultiplier tube. Stimulated emission of SPPs was observed as an increase in the reflected probe light compared to that measured when the gain medium is not pumped, consistent with previous observations reported by Seidel and colleagues.[25] By fitting the reflectance measurements to the Fresnel expression describing the reflectance of the structure, the authors estimated the material gain of the doped polymer as $420\,\text{cm}^{-1}$ and a SPP loss reduction of 35%. The estimated material gain was nearly eight times smaller than that expected from the spectroscopic parameters of the dye. The authors pointed out that the discrepancy was probably due to luminescence quenching at the high dye concentrations, and amplified spontaneous emission into photonic and plasmonic modes.

In a subsequent study,[31] Noginov and colleagues measured the fluorescence decoupled from the prism side as the gain medium was pumped using the same experimental arrangement than in Ref. (30). For this experiment, the structures employed were similar to those reported in Ref. (30) but with a thinner ($3\,\mu$m-thick) gain medium. In addition to the characteristic angular profile of decoupled SPPs, the authors observed attributes of stimulated emission in the measured fluorescence. They reported spectral narrowing at high pump intensities and a non-linear intensity behaviour of the decoupled light with respect to the pump intensity that exhibited a distinct threshold point. On the basis of a theoretical model, the authors claimed that the observations correspond to those expected for amplified SPP propagation when the losses are overcome entirely.

Grandidier and colleagues[32] demonstrated an increase in propagation length of SPPs supported by a polymer loaded metallic waveguide at a telecommunication wavelength. The waveguides consisted of 600 nm-thick by 400 nm-wide PMMA stripes doped with lead sulfide quantum dots (PbS QDs) fabricated on a 40 nm-thick gold film. The waveguides were $64\,\mu$m in length and the QD density was $N \approx 9 \times 10^{16}\,\text{cm}^{-3}$. The QD emission spectrum exhibited a broad peak near $\lambda = 1550\,\text{nm}$, the probe light wavelength used in the experiments. The gain medium was pumped using continuous wave (CW) light at $\lambda = 532\,\text{nm}$ homogeneously distributed over the waveguide. The SPP propagation length as a function of the pump power was measured via leakage radiation microscopy.[33] They

showed a maximum increase in SPP propagation of 27%, achieved with a pump power of $1\,\mathrm{kW/cm^2}$ and a corresponding gain in the polymer of $\sim 600\,\mathrm{cm^{-1}}$. They also measured the angular momentum spread of the amplified SPP showing a narrowing in the angular momentum spectrum with increasing pump power. The authors pointed out that the narrowing is indicative of loss compensation since the full-width at half-maximum (FWHM) measure of the angular spectrum, $\Delta k_{\|}$, is inversely proportional to the SPP propagation length $\Delta k_{\|} = 1/l_{\mathrm{spp}}$.

Oulton and colleagues[34] demonstrated laser-like behaviour in hybrid-plasmon modes supported by a thick silver film coupled to a cadmium sulphide (CdS) nanowire that served as the gain medium. Experiments were conducted on structures as shown in Fig. 2(a); the CdS nanowire had a diameter d and was separated from the silver film by a magnesium fluoride insulating layer of thickness h. The hybrid plasmon can have a good portion of its energy concentrated in this insulating layer.[35] The gain medium was pumped using pulsed light at $\lambda_p = 405\,\mathrm{nm}$ ($t_p = 100\,\mathrm{fs}$; $f_p = 80\,\mathrm{MHz}$). The emitted light coupled out of the structure was studied for a number of structures with different insulating layer thicknesses and CdS nanowire diameters. The amplified spontaneous emission (ASE)

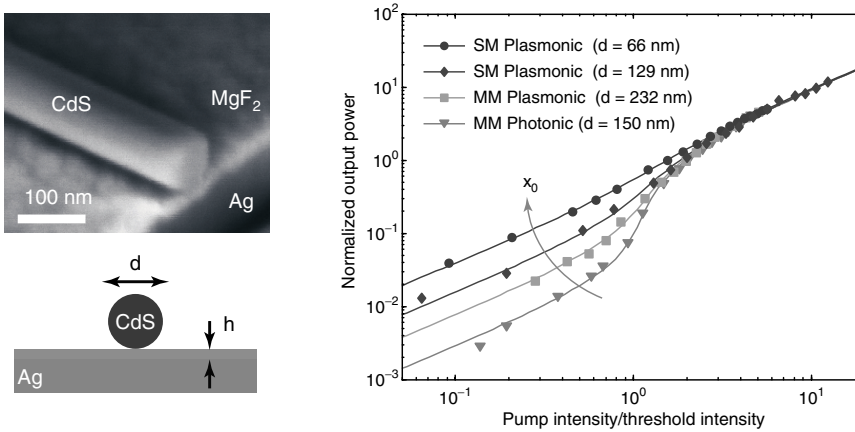


Fig. 2. Adapted by permission from Macmillan Publishers Ltd: Nature [Ref. (34)], copyright (2009). (a) Hybrid-plasmon active structure employed by Oulton *et al.* in Ref. (34). (b) Output power versus pump intensity characteristics of four structures with different CdS nanowire diameter, d. The mode size is reduced and the spontaneous emission into the lasing mode increase in the direction given for the parameter x_0. SM: single mode lasing, MM: multi-mode lasing.

threshold was estimated as the pump intensity for which the Fabry-Perot plasmonic modes of the cavity are discernible in the measured light spectrum. For structures with $h = 5\,\text{nm}$ and $100 < d < 400\,\text{nm}$, the authors reported ASE thresholds that occur between 10 and $60\,\text{MW/cm}^2$. For pump intensities above these thresholds a non-linear output-power versus pump-intensity relation developed suggesting laser action. A six-fold enhancement in the spontaneous emission rate when compared to a photonic laser consisting of a CdS nanowire on quartz was observed for $d \approx 120\,\text{nm}$. This enhancement is usually associated with high-Q cavities but in this laser it was attributed to the high mode confinement. The authors pointed out that the transition from spontaneous emission to lasing in the output-power versus pump-intensity relation of the hybrid-plasmon laser did not show the typical 'kink' characteristic of photonic lasers but this transition was obscured for highly confined plasmonic modes [see Fig. 2(b)]. This behaviour is a signature that most of the spontaneous light is emitted into the lasing mode.[36]

Bolger and colleagues[37] studied experimentally SPP ASE and the limitations that it imposes on SPP stimulated emission. Experiments were conducted on gold films of various thicknesses (25, 50, and 100 nm) deposited on a silica substrate and cladded with a $1\,\mu$m-thick gain medium in the form of PMMA doped with PbS QDs. The QD concentration was 5 wt.%. The SPP mode supported at the polymer-gold interface was used for the analysis. Pairs of gratings separated by several distances were fabricated on the gold film in order to couple light in and out of the plasmonic structure. The gain medium is pumped with a helium-neon laser at $\lambda = 633\,\text{nm}$ focused to a $100\,\mu$m-diameter spot located between the two gratings and the SPP mode was probed via QD-photoluminescence generated using a second helium-neon laser focused onto one of the gratings. The SPP emission spectra coupled out of the structures peaked at $\lambda \approx 1160\,\text{nm}$ and exhibited a significant linewidth narrowing as a function of the pump intensity. The emission linewidth reached its minimum value with a pump intensity of $\sim 10\,\text{W/cm}^2$ and remained so for higher pump intensities. The authors estimated the pump intensity for the SPP ASE on-set as $\sim 5\,\text{W/cm}^2$. They reported a maximum increase in propagation length of 30% achieved with a pump intensity of $\sim 1\,\text{W/cm}^2$. For higher pump intensities the propagation length decreased, eventually reaching values comparable to those measured in the absence of the pump. This was attributed to a reduction of available gain caused by the competition

of SPP ASE and SPP stimulated emission, a well known phenomenon in high gain optical amplifiers.[38]

2.2. *Thin metal film/stripe*

Alam and colleagues[39] proposed a structure to overcome the SPP losses at telecommunication wavelengths via optical amplification using quantum wells as the gain medium. It consisted of a 1 μm-wide and 10 nm-thick silver stripe on an AlGaInAs multiple quantum well (QW) structure on a semi-infinite InP substrate. The metal stripe was covered by a dielectric superstrate of finite thickness. The authors conducted a numerical analysis of the structure's optical properties varying systematically parameters such as superstrate and QW barrier refractive indices, and metal and superstrate dimensions. They showed that a 400 nm-thick superstrate with refractive index matched to that of the barrier regions allows a well-guided SPP mode despite the structure's large refractive index asymmetry. This design also provided a good balance between SPP mode attenuation ($\sim$12 dB/mm) and mode-gain overlap ($\sim$6.5%). They found that a material gain of 402 cm^{-1} was required for lossless SPP propagation in this configuration assuming that modal loss was due solely to metal absorption. The authors pointed out the difficulty of growing a crystalline superstrate on top of a metallic film which can lead to additional losses of up to 35 dB/mm for a poly-crystalline superstrate.

Okamoto and colleagues[40] observed fluorescence enhancement of laser dyes deposited on two-dimensional metallic gratings exhibiting a plasmonic bandgap at the dye emission wavelength. The enhancement was measured relative to emission on a planar silver film. The structures consisted of 150 nm-thick corrugated silver films forming a two-dimensional hexagonal lattice with a 550 nm and 507 nm pitch; the former was covered with a 25 nm layer of evaporated 4-dicyanomethylene-2-methyl-6-p-dimethyl-aminostyryl-4H-pyran (DCM) dye, while the latter was covered with a 50 nm layer of methyl red-doped PMMA. The fluorescence enhancements observed were 3- and 150-fold for DCM and methyl red, respectively. The authors suggested a similar structure that supports low-loss LRSPPs in order to devise a plasmonic bandgap laser at visible wavelengths. Such a structure consisted of a 20 nm-thick corrugated silver film covered with a DCM-doped tris(8-hydroxyquinoline)aluminium (Alq_3) as a gain medium. They argued that this medium can provide sufficient gain to overcome the LRSPP propagation loss, calculated as 71 cm^{-1}, at the peak emission

of DCM-doped-Alq_3 ($\lambda = 620\,\text{nm}$). In subsequent studies, Okamoto and colleagues investigated numerically the behaviour of plasmonic bandgaps appearing in the SPP dispersion relations of corrugated thin silver films and determined design conditions for improved performance of a LRSPP plasmonic bandgap laser.[41,42]

Winter and colleagues[43] studied in detail the plasmonic bandgap laser structure suggested by Okamoto in Ref. (40) and questioned the possibility of LRSPP lasing. They discussed the need for considering the power dissipation of excited molecules into SRSPPs also supported by the structure. Photoluminescence measurements on (effectively) symmetric corrugated structures with silver-film thicknesses varying from 20 to 90 nm showed that a significant amount of power is indeed coupled into both SRSPPs and LRSPPs. The authors conducted numerical calculations to estimate the fraction of power dissipated by excited molecules into LRSPPs and SRSPPs supported by a planar symmetrically cladded silver-film as a function of the film thickness and for specific separations, d, between the molecule and the metal surface. The results showed that for a film thickness of 20 nm and $d = 20\,\text{nm}$ excited molecules dissipate about 80% of their power into SRSPPs and only about 7% into LRSPPs. The authors estimated the gain available for LRSPPs and SRSPPs as fractions of the total gain in correspondence to the fractions of power dissipated by an excited molecule at a particular location into the respective modes. Their results suggested that lasing in the LRSPP would be possible only with silver films thinner than about 15 nm.

De Leon and Berini[44] presented a theoretical model for SPP amplification in one-dimensional planar metallic structures incorporating an optically-pumped dipolar gain medium. The model takes into account the gain non-uniformity close to the metal surface caused by a position-dependent dipole lifetime and inhomogeneous pump irradiance distribution. The pump light was assumed monochromatic and its irradiance distribution throughout the structure was computed rigorously using a transfer matrix method. The dipole lifetime as a function of the metal-dipole separation was calculated by taking into account five excited state decay channels; four of them were associated with dipole coupling to LRSPPs, SRSPPs, radiative modes of the structure (RAD) and electron-holes (EH) in the metal; and the fifth, internal conversion (IC), accounted for the dipole's quantum efficiency [see Fig. 3(a)]. The nonuniform gain distribution was then obtained by applying locally rate equations for the standard four-level pumping model, with the lifetime and irradiance taking on their

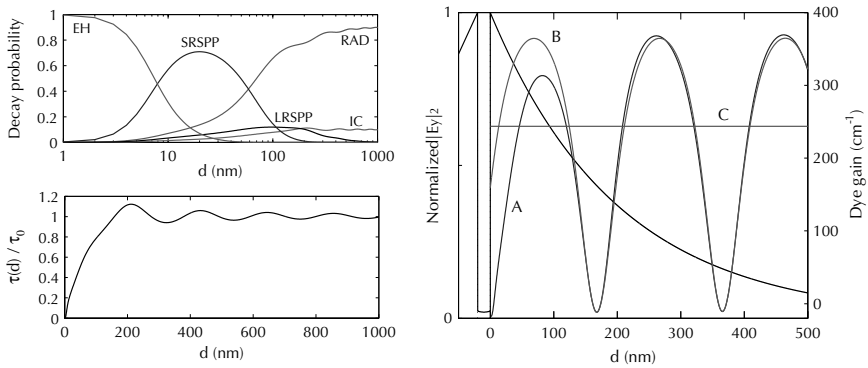


Fig. 3. Adapted from Ref. (44). (a) Top panel: Spontaneous emission probability of excited R6G molecules into the different energy decay channels supported by the plasmonic structure considered by De Leon and Berini in Ref. (44). Bottom panel: Spontaneous emission lifetime as a function of the molecule-metal separation, $\tau(d)$, normalised to the molecule's lifetime away from the structure, τ_0. (b) The LRSPP mode ($|E_y|^2$) and dye gain distributions for the three gain models in Ref. (44); i.e., model proposed by the authors (A), considering pump position dependance but uniform lifetime (B), and uniform gain model (C).

position-dependent values. It was then discretised and incorporated into a multilayer waveguide mode solver from which mode power gains were computed for the modes of the system. The authors studied amplification of LRSPPs on a 20 nm-thick silver film bounded by Cytop ($n = 1.343$) on one side and by an index-matched gain medium on the other. The gain medium consisted of a 5 μm-thick layer of R6G molecules dissolved in a mixture of 57% methanol and 43% ethanol with molecular density of $N = 1.8 \times 10^{18}\,\mathrm{cm}^{-3}$. They predicted that net amplification of the LRSPP is possible at visible wavelengths, $\lambda = 560\,\mathrm{nm}$, using a reasonable pump irradiance of $210\,\mathrm{kW/cm}^2$ at $\lambda = 532\,\mathrm{nm}$. The mode power gain computed with the proposed model was $2.9\,\mathrm{cm}^{-1}$, while the values obtained using two other modes, one that accounts for the pump position dependance but assumes uniform dipole lifetime and one that assumes a uniform gain, were $26.5\,\mathrm{cm}^{-1}$ and $22.3\,\mathrm{cm}^{-1}$, respectively. The large discrepancy was ascribed to the smaller gain-mode overlap present when the lifetime quenching near the metal surface was taken into account [see Fig. 3(b)].

In a subsequent publication, De Leon and Berini[45] studied the amplification of LRSPPs and single-interface SPPs using moderately and heavily concentrated R6G dye molecules in solution as the gain medium,

assuming molecular densities of $N = 3 \times 10^{18}\,\mathrm{cm}^{-3}$ and $2.4 \times 10^{19}\,\mathrm{cm}^{-3}$, respectively. They considered a symmetrically-cladded active structure similar to that studied in Ref. (44) but with a thinner ($1\,\mu$m) gain layer. Furthermore, for the analysis of single-interface SPPs the silver film was taken as infinitely thick. They employed the model proposed previously in Ref. (44) and accounted for the concentration-dependent photophysical properties of the dye to estimate the SPP mode power gain as a function of the pump irradiance. Numerical calculations showed that lossless propagation of single-interface SPPs supported by the structure could be achieved only with the high dye concentration using an intense pump irradiance of $3.45\,\mathrm{MW/cm}^2$, whereas for LRSPPs it was achieved with both dye concentrations using pump levels of few hundreds of $\mathrm{kW/cm}^2$. The authors discussed the relevance of using the proposed theoretical model. They pointed out the importance of the model in analysing plasmonic amplifiers in the gain saturation regime (near full population inversion) and showed that, for the cases under analysis, neglecting the effects of pump and/or dipole lifetime position-dependence led to errors in the required pump irradiance of up to 800%. On the other hand, they showed for dipoles with a low quantum efficiency that the lifetime distribution is well approximated as a constant; thus, the effects of the metal on the dipole lifetime can be neglected in such a case.

Ambati and colleagues[46] observed stimulated emission of LRSPPs at telecommunication wavelengths using a propagating pump-probe arrangement. They conducted experiments on structures consisting of a $8\,\mu$m-wide by 20 nm-thick gold stripe embedded in erbium-doped glass. The insertion loss of a 8 mm-long waveguide was measured as 39 dB, which includes the ground-state absorption of erbium ions, the LRSPP propagation loss, and other loss mechanisms such as coupling and scattering losses. Presumably, structures of the same length were used for their stimulated emission experiments. The pump ($\lambda = 1480\,\mathrm{nm}$) and probe ($\lambda = 1532\,\mathrm{nm}$) signals were combined by a fibre multiplexer and coupled simultaneously into the LRSPP modes of the metal stripe using a butt-coupled single mode fibre (SMF) and a polarisation controller. The output at the probe wavelength was extracted using the same type of fibre and a fibre demultiplexer. The LRSPP stimulated emission under pulsed and CW pumping was investigated. For the pulsed analysis, a $500\,\mu$s pump pulse was followed by a $150\,\mu$s probe pulse with a time delay between them of $30\,\mu$s and the probe signal enhancement was measured as a function of the pump power. The authors reported enhancements of 0.74 dB with a pump power of 92 mW

and of 1.73 dB with a pump power of 266 mW in pulsed and CW operation, respectively. They noticed that the ground level of the detected signal did not change appreciably even under the highest pump power, suggesting negligible amplified spontaneous emission.

De Leon and Berini[47] reported a direct measurement of gain in the LRSPP mode supported by a gold stripe at a near-infrared wavelength. The gold stripe was 20 nm-thick and 1 μm-wide, and was deposited on a 15 μm-thick SiO_2 layer thermally grown on a Si wafer. It was covered with a $\sim$100 μm-thick index-matched gain layer consisting of IR140 dye molecules dissolved in a static mixture of 69.6% dimethyl sulphoxide and 30.4% ethylene glycol. The gain medium, with molecular density of $N = 6 \times 10^{17}\,\mathrm{cm}^{-3}$, was pumped at normal incidence using pulsed light ($t_p = 8$ nm; $f_p = 10$ Hz) at $\lambda = 808$ nm polarised along the stripe length. The LRSPP mode of the stripe was probed using CW light at $\lambda = 882$ nm, close to the peak emission of the dye, via end-fire coupling using a polarisation maintaining fibre [see Fig. 4(a)]. It was out-coupled in a similar fashion using the same type of fibre and then sent to the detection system. Gain measurements were conducted on a structure of length $l = 2.70$ mm by varying the amplifier length (by varying the length of the pumped region) over the range $0.457 < l_a < 1.676$ mm while keeping a constant pump energy density of $E_p = 20 \pm 5\,\mathrm{mJ/cm}^2$ and an input power of $P_i = 2.10$ mW. The fibre-to-fibre gain as a function of l_a was obtained by measuring the output power, P_o, with and without the probe signal [see Fig. 4(b)] and then subtracting the first measurement (the noise) from the second one (amplified signal plus noise). The gain of the amplifier section as a function of l_a was deduced from these measurements using known values of coupling efficiency (92%; measured via cutback) and propagation loss in passive sections ($\sim$4.7 dB/mm; measured *in situ*). Then a linear model was fitted to the linear region of the amplifier-gain versus amplifier-length curve [see Fig. 4(c)], and a small-signal LRSPP gain of 8.55 dB/mm was deduced from its slope. The authors also investigated the noise characteristics of the amplifier by capturing the time-averaged ASE distribution at the output facet of a structure with $l = 1.8$ mm and $l_a = 1.05$ mm using a slow-response infrared camera. They observed a spontaneous emission reduction of 6.3 dB in the vicinity of the waveguide relative to the emission in the bulk gain, as depicted in Fig. 4(d) by the two ASE intensity distribution cuts perpendicular to the metal plane. They pointed out that LRSPP amplifiers benefit from the low spontaneous emission into LRSPPs, which

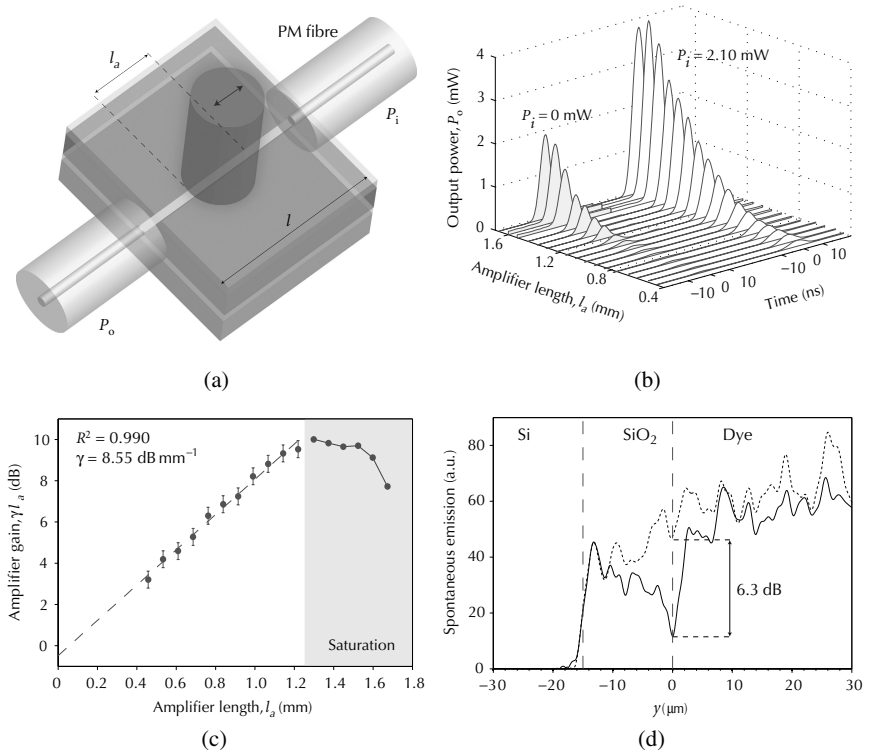


Fig. 4. Adapted from Ref. (47). (a) Active plasmonic structure employed by De Leon and Berini in Ref. (47). Pump and probe signal coupling arrangements; the pump polarization is indicated by the red arrow; sample length, l; amplifier length, l_a; input power, P_i; output power, P_o. (b) Measurements of P_o as a function of l_a with (right) and without (left) probe signal. (c) Measurements of amplifier gain versus l_a; the dashed line indicates the linear fit. (d) Spontaneous emission intensity distribution cuts perpendicular to the metal plane through the centre of the waveguide (solid) and 10 μm to the left of the waveguide (dashed).

is suppressed by other energy decay processes with larger probabilities such as dipole coupling to SRSPPs, EH pairs, and unguided radiative modes.

Gather and colleagues[48] also reported gain in the LRSPP mode supported by a gold film at a visible wavelength. The 4 nm-thick gold film was deposited on an 20 μm-thick transparent polymer ($n = 1.5$), then a 1 μm-thick index-matched fluorescent polymer (MDMO-PPV:PSF:HBO) covered the gold surface and provided the optical gain; finally an additional 20 μm-thick transparent polymer layer completed the structure. The gain

layer was pumped from the top using pulsed light at $\lambda = 532\,\text{nm}$ ($t_p = 5\,\text{nm}$; $f_p = 1\,\text{Hz}$) polarised along the propagation direction of the LRSPP. The LRSPP propagation loss without pumping was measured through scattered light imaging and was found to be approximately $60\,\text{cm}^{-1}$ at $\lambda = 600\,\text{nm}$. The gain of the fluorescent polymer was measured separately in a dielectric slab geometry and was found to be $59\,\text{cm}^{-1}$. A gain guided LRSPP waveguide was created by focusing the pump light into a $200\,\mu\text{m}$ wide by 2 mm long stripe and the light at the output was collected using a $\times 10$ microscope objective; then the emission spectrum of the collected light was monitored as a function of the pump energy, E_p. For TM light, ASE was observed as the reduction of the emission linewidth, from 67 nm (FWHM) at $E_p = 6\,\mu\text{J}$ to 13 nm at $E_p = 50\,\mu\text{J}$, accompanied by a sudden increase of signal intensity. The ASE pump threshold was estimated as $E_p = 13\,\mu\text{J}$. ASE was not observed for TE light suggesting that the signal was associated mainly to the LRSPP supported by the structure. Employing the variable stripe length method[49] the authors estimated the LRSPP gain as $8\,\text{cm}^{-1}$ (3.47 dB/mm) when the gain layer is pumped at $12.5\,\text{mJ/cm}^2$. They also performed numerical calculations similar to those in Ref. (44) to estimate the spontaneous emission fraction coupled into LRSPPs in the structure as a function of the metal-molecule separation, d. It was found a maximum LRSPP excitation efficiency of 0.17 at $d \approx 100\,\mu\text{m}$ that decays exponentially with increasing distance reaching 0.05 at $d \approx 1\,\mu\text{m}$.

2.3. *Metal-insulator-metal structure*

Maier[50] reported a numerical study of gain assisted SPP propagation in metal-insulator-metal structures at telecommunication wavelengths. The analysed structures are formed by two gold semi-infinite regions separated by semiconductor gain core ($n = 3.4$) with thicknesses varying from 5 to 500 nm. It was shown that a sufficiently large gain can completely compensate the losses of the symmetric (with respect to E_y) coupled SPP mode supported by the metal-semiconductor interfaces. The gain coefficient for complete loss compensation was estimated as $1625\,\text{cm}^{-1}$ and $4830\,\text{cm}^{-1}$ for structures with core thicknesses of 500 nm and 50 nm, respectively. The author pointed out the possibility of reducing the gain requirements by about an order of magnitude using gain media with low refractive index, such as quantum dot- or dye-doped polymers or glasses.

3. Metallic Nanocavities

Lasing in metallic nanocavities was first demonstrated by Hill and colleagues[51] at telecommunication wavelengths. Such a nanolaser consisted of an InP-InGaAs-InP double heterostructure gain medium shaped in the form of a circular pillar of diameter $d \approx 300\,\text{nm}$. The pillar is grown on a p-InGaAsP layer on an Si-InP substrate and the entire structure was covered by a thick gold layer [Fig. 5(a)]. The gain medium was electrically pumped with electrons and holes injected through the pillar top and lateral contacts, respectively [see Fig. 5(a)]. The CW lasing properties were

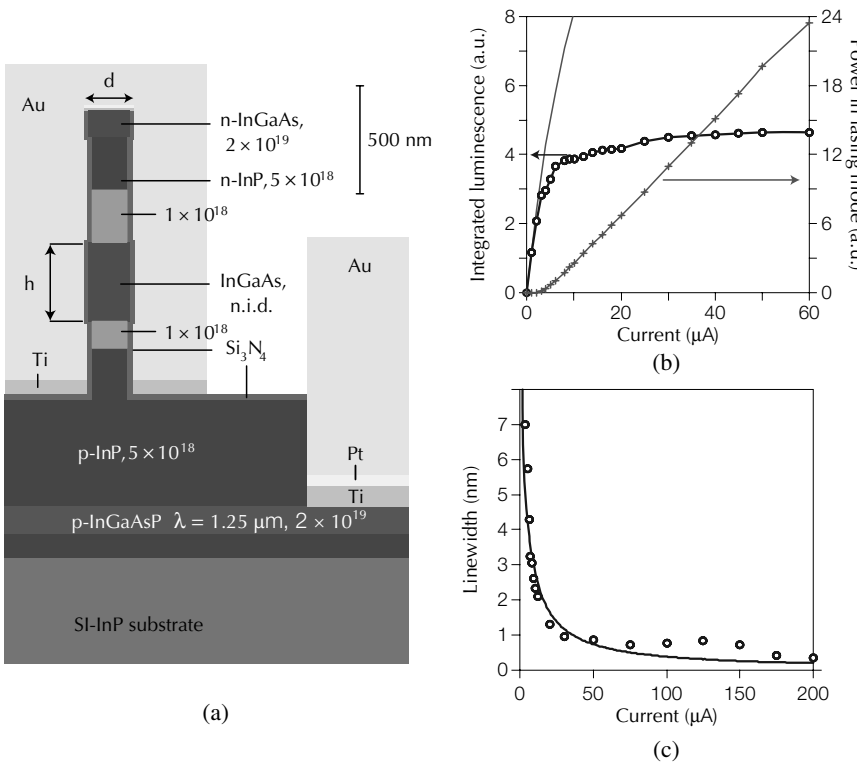


Fig. 5. Adapted by permission from Macmillan Publishers Ltd: Nature Photonics [Ref. (51)], copyright (2009). (a) Schematic of the nanolaser structure constructed by Hill *et al.* in Ref. (51). (b) Characteristics of the lasing-mode output (crosses), integrated luminescence (circles), and integrated luminescence from a non-lasing device (no markers) versus pump current at a temperature of 10 K. (c) Lasing mode linewidth (circles) versus pump current at a temperature of 10 K. The solid line is the theoretical estimate.

analysed at cryogenic temperatures by measuring light leaking through the base of the structure. The laser provided single mode emission over a broad wavelength range, 1418 nm $< \lambda <$ 1452 nm, by varying the pillar's diameter over 30 nm. At temperatures of 10 K and 77 K, respectively, the laser exhibited lasing thresholds of 3.5 μA [see Fig. 5(b)] and 6 μA and quality factors of 200 and 140. Spectral linewidth measurements at 10 K as a function of the pump current revealed a sharp reduction after the lasing threshold, reaching FWHM values of approximately 1 nm for pump currents ranging from 30 to 150 μA and 0.3 nm for higher currents [see Fig. 5(c)].

In a subsequent study,[52] Hill and colleagues reported lasing at telecommunication wavelengths in similar metallic nanocavities to those used in Ref. (51); however, in this case the semiconductor structure was shaped in the form of a rectangular pillar and covered with a thick layer of silver. The structures were 300 nm in height with widths varying from 90 nm to 350 nm and lengths in the order of microns; hence, supporting Fabry-Perot lasing modes. The structures supported a TM0 transverse mode similar to that considered by Maier in Ref. (50) but distorted by the inhomogeneous core of the cavity. The authors characterised the laser's CW operation at a temperature of 78 K, reporting a lasing threshold current of 40 μA and quality factor of $\sim$370 for a 130 nm wide and 3 μm long device. In addition, they demonstrated pulsed laser operation at room temperature reporting a lasing threshold peak current of $\sim$5 mA and a quality factor of $\sim$340 for a 310 nm wide and 6 μm long device.

The main disadvantage of metal-coated nanocavities is their high lasing threshold which requires operation at cryogenic temperatures. To alleviate this problem, Mizrahi *et al.*[53] proposed a subwavelength resonator design for low lasing threshold based on a circular pillar geometry similar to that reported in Ref. (51). They provided numerical calculations showing that shielding the gain region from the metallic walls using a low index dielectric reduces the threshold gain considerably and enables room-temperature lasing at $\lambda = 1550$ nm using an InGaAsP gain medium. Employing this design, Nezhad and colleagues[54] demonstrated pulsed room-temperature operation of metal-coated laser at $\lambda = 1430$ nm. The structure consisted in an InGaAsP gain region shielded by a 200 nm thick SiO_2 layer and encapsulated by a 70 nm layer of sputtered aluminium. The gain medium was optically pumped; measurements of pump intensity versus laser output were conducted using pulsed pump light with pulse duration of 12 ns and repetition rate of 300 kHz at $\lambda = 1064$ nm. The ASE and lasing thresholds occurred at pump intensities of 400 and 700 W/cm^2, respectively. Although

this laser cavity does not support a purely plasmonic mode, its physical (and modal) size is smaller than the wavelength, having a diameter of 1.1 μm and a hight of 1.35 μm.

4. Metallic Nanoparticles

In 2003, Bergman and Stockman[55] proposed the concept of surface plasmon amplification by stimulated emission of radiation (*spaser*) to create a nanoscopic coherent SPP source. Indeed, a *spaser* could serve as a convenient light source for plasmonic applications, eliminating the need of a coupling mechanism between photons and SPPs. The authors conducted a theoretical and numerical analysis considering a metallic nanoparticle surrounded by a gain medium in the form of quantum dots. The nanoparticle supports a localised SPP resonance. Hence, the nanoparticle itself provides the feedback, which together with the inverted gain medium form a structure capable of amplifying coherently the surface plasmon modes supported by the system, much like the photonic modes of a cavity are amplified in a conventional *laser*.

Noginov and colleagues[56] reported SPP laser-like behaviour in metallic nanoparticles according to the concept of *spaser* introduced earlier by Bergman and Stockman.[55] The spaser design, shown in Fig. 6(a), consisted of 44 nm-diameter nanoparticles formed by a 14 nm-diameter gold core

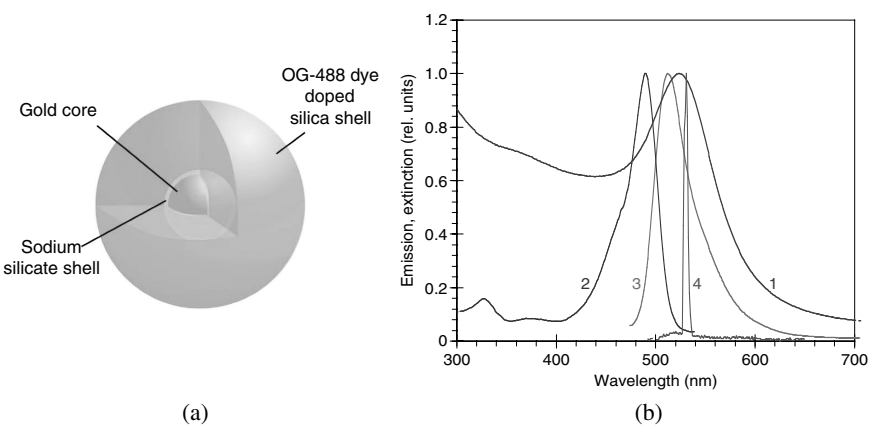


Fig. 6. Adapted by permission from Macmillan Publishers Ltd: Nature [Ref. (56)], copyright (2009). (a) Spaser-based nanolaser design employed by Noginov and colleagues in Ref. (56). (b) Normalised extinction (1), excitation (2), spontaneous emission (3), and stimulated emission (4) spectra of nanoparticle shown in (a).

embedded in a 15 nm thick dye-doped silica shell ($n = 1.46$). The dye employed (Oregon Green 488) emits in a band centred about $\lambda \approx 510$ nm and overlaps well with the broad SPP resonance spectrum supported by the gold nanoparticle, which peaks at $\lambda \approx 520$ nm [see spectra 1 and 3 in Fig. 6(b)]. The number of dye molecules per nanoparticle was estimated as 2.7×10^3. An aqueous nanoparticle suspension with nanoparticle density $N = 3 \times 10^{11}\,\mathrm{cm}^{-3}$ was placed in a 2 mm path-length cuvette and pumped at the dye's peak absorption wavelength, $\lambda = 488$ nm [spectrum 2 in Fig. 6(b)] using pulsed light ($t_p = 5$ ns) from an optical parametric oscillator. The authors conducted experiments revealing laser-like attributes of the nanoparticles' emission such as narrow linewidth and non-linear characteristics of pump versus emission power. The nanoparticle emission spectrum [spectrum 4 in Fig. 6(b)] peaks at $\lambda = 531$ nm, close to the SPP resonance maximum. The authors pointed out that reducing the nanoparticle density did not change considerably the emission characteristics. Hence, they concluded that the laser-like behaviour was produced by individual nanoparticles [as predicted in Ref. (55)] and not by a collective feedback mechanism in the nanoparticle suspension or by a feedback provided by the cuvette walls.

5. Concluding Remarks

During the past years there has been good progress in understanding the physical principles and approaching a practical implementation of SPP amplifiers and lasers. Despite the fact that active plasmonics is at an early stage, a number of experiments have already demonstrated propagation of SPPs with reduced losses and even gain in several active structures. These achievements open avenues for numerous potential applications including ultra sensitive biosensors, nanolithography, highly integrated optical devices, nanoscopic light sources, and low noise optical amplifiers.

At the same time, there is still much to be uncovered from both the fundamental and the practical point of views. For instance, more investigation is needed to understand the implications of enhanced spontaneous emission and amplified spontaneous emission in SPP amplifiers, and to describe realistic SPP amplification scenarios including phenomena such as SPP gain saturation and amplification with inhomogeneous and anisotropic gain media. Also, developing efficient electrically-pumped gain materials with sufficient gain at room temperature for lasers and amplifiers with sub-wavelength optical confinement still remains a challenge. It is clear that

active plasmonics has a long road ahead and its future is bright offering an ample panorama of research activities.

References

1. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings.* (Springer, 1988).
2. P. L. Stiles, D. J. Dieringer, N. C. Shah, and R. P. Van Duyne, *Annual Rev. Anal. Chem.* **1**, 601 (2008).
3. T. Neumann, M. L. Johansson, D. Kambhampati, and W. Knoll, *Adv. Funct. Mater.* **12**, 575 (2002).
4. E. Ozbay, *Science* **311**, 189 (2006).
5. W. L. Barnes, A. Dereux, and T. W. Ebbesen, *Nature* **424**, 824 (2003).
6. S. Kawata, Y. Inouye, and P. Verma, *Nature Photon.* **3**, 388 (2009).
7. H. J. Simon, D. E. Mitchell, and J. G. Watson, *Phys. Rev. Lett.* **33**, 1531 (1974).
8. J. N. Anker, *et al. Nature Mater.* **6**, 442 (2008).
9. J. J. Burke, G. I. Stegeman, and T. Tamir, *Phys. Rev. B* **33**, 5186 (1986).
10. D. Sarid, *Phys. Rev. Lett.* **47**, 1927 (1981).
11. F. Yang, J. R. Sambles, and G. W. Bradberry, *Phys. Rev. B* **44**, 5855 (1991).
12. P. Berini, *Phys. Rev. B.* **61**, 10484 (2000).
13. P. Berini, *Adv. Opt. Photon.* **1**, 484 (2009).
14. D. R. Smith, J. B. Pendry, and M. C. K. Wiltshire, *Science* **305**, 788 (2004).
15. V. M. Shalaev, *Nature Photon.* **1**, 41 (2007).
16. S. Zhang, D. A. Genov, Y. Wang, M. Liu, and X. Zhang, *Phys. Rev. Lett.* **101**, 047401 (2008).
17. A. Fang, Th. Koschny, M. Wegener, and C. M. Soukoulis, *Phys. Rev. B* **79**, 241104 (2009).
18. S. A. Ramakrishna and J. B. Pendry, *Phys. Rev. B.* **67**, 201101 (2003).
19. N. I. Zheludev, S. L. Prosvirnin, N. Papasimakis, and V. A. Fedotov, *Nature Photon.* **2**, 351 (2008).
20. Z. G. Dong, H. Liu, T. Li, Z. H. Zhu, S. M. Wang, J. X. Cao, S. N. Zhu, and X. Zhang, *Phys. Rev. B.* **80**, 235116 (2009).
21. A. Kovyakov, A. R. Zakharian, K. M. Gundu, and S. A. Darmanyan, *App. Phys. Lett.* **94**, 151111 (2009).
22. J. A. Dionne, L. A. Sweatlock, H. A. Atwater, and A. Polman, *Phys. Rev. B.* **73**, 035407 (2006).
23. G. Plotz, H. Simnon, and J. Tucciarone, *J. Opt. Soc. Am.* **69**, 419 (1979).
24. A. N. Sudarkin and P. A. Demkovich, *Sov. Phys. Tech. Phys.* **34**, 764 (1988).
25. J. Seidel, S. Grafstrom, and L. Eng, *Phys. Rev. Lett.* **94**, 177401 (2005).
26. M. P. Nezhad, K. Tetz, and Y. Fainman, *Opt. Express.* **12**, 4072 (2004).
27. I. Avrutsky, *Phys. Rev. B* **70**, 155416 (2004).
28. C. Sirtori, C. Gmachl, F. Capasso, J. Faist, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, *Opt. Phys. Lett.* **23**, 1366 (1998).
29. A. Tredicucci, C. Gmachl, F. Capasso, A. L. Hutchinson, D. L. Sivco, and A. Y. Cho, *App. Phys. Lett.* **76**, 2164 (2000).

30. M. A. Noginov *et al.*, *Opt. Express.* **16**, 1385 (2008).
31. M. A. Noginov *et al.*, *Phys. Rev. Lett.* **101**, 226806 (2008).
32. J. Grandidier *et al.*, *Nano Lett.* **9**, 2935 (2009).
33. J. Grandidier, S. Massenot, G. Colas des Francs, A. Bouhelier, J. C. Weeber, L. Markey, A. Dereux, J. Renger, M. U. Gonzalez, and R. Quidant, *Phys. Rev. B* **78**, 245419 (2008).
34. R. F. Oulton *et al.*, *Nature* **461**, 629 (2009).
35. R. F. Oulton, V. J. Sorger, D. A. Genov, D. F. Pile, and X. Zhang, *Nature Photon.* **2**, 496 (2008).
36. G. Bjork and Y. Yamamoto, *IEEE J. Quantum Elect.* **27**, 2386 (1991).
37. P. M. Bolger, W. Dickson, A. V. Krasavin, L. Liebscher, S. G. Hickey, D. V. Skryabin, and A. V. Zayats, *Opt. Lett.* **35**, 1197 (2010).
38. L. W. Casperson, *J. App. Phys.* **48**, 256 (1977).
39. M. Z. Alam, J. Meier, J. S. Aitchison, and M. Mojahedi, *Opt. Express.* **15**, 176 (2007).
40. T. Okamoto, F. H'Dhili, and S. Kawata, *App. Phys. Lett.* **85**, 3968 (2004).
41. T. Okamoto, J. Simonen, and S. Kawata, *Phys. Rev. B* **77**, 115425 (2008).
42. T. Okamoto, J. Simonen, and S. Kawata, *Opt. Express.* **17**, 8294 (2009).
43. G. Winter, S. Wedge, and W. L. Barnes, *New J. Phys.* **8**, 125 (2006).
44. I. De Leon and P. Berini, *Phys. Rev. B.* **78**, 161401(R) (2008).
45. I. De Leon and P. Berini, *Opt. Express.* **17**, 20191 (2009).
46. M. Ambati *et al.*, *Nano Lett.* **8**, 3998 (2008).
47. I. De Leon and P. Berini, *Nature Photon.* **4**, 382 (2010).
48. M. C. Gather, K. Meerholz, N. Danz, and K. Leosson, *Nature Photon.* **4**, 457 (2010).
49. K. L. Shaklee, R. E. Nahory, and R. F. Leheny, *J. Luminescence* **7**, 284 (1973).
50. S. A. Maier, *Opt. Commun.* **258**, 295 (2006).
51. M. T. Hill *et al.*, *Nature Photon.* **1**, 589 (2007).
52. M. T. Hill *et al.*, *Opt. Express.* **17**, 11107 (2009).
53. A. Mizrahi, V. Lomakin, B. A. Slutsky, M. P. Nezhad, L. Feng, and Y. Fainman, *Opt. Letters* **33**, 1261 (2008).
54. M. P. Nezhad, A. Simic, O. Bondarenko, B. Slutsky, A. Mizrahi, L. Feng, V. Lomakin, and Y. Fainman, *Nature Photon.* **4**, 395 (2010).
55. D. J. Bergman and M. I. Stockman, *Phys. Rev. Lett.* **90**, 027402 (2003).
56. M. A. Noginov *et al.*, *Nature* **460**, 1110 (2009).

WAVEFRONT ENGINEERING OF QUANTUM CASCADE LASERS USING PLASMONICS

Nanfang Yu* and Federico Capasso†

School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, U.S.A.
**nyu@fas.harvard.edu, †capasso@seas.harvard.edu*

We review recent work on beam shaping of mid-infrared and far-infrared (terahertz) quantum cascade lasers using plasmonics. Essentials of quantum cascade lasers (QCLs) are discussed; these include the operating principle based on bandstructure engineering, and beam quality problems associated with laser waveguide design. We explain how metal and semiconductor microstructures can effectively tailor the dispersion properties of mid- and far-infrared surface plasmon polaritons, and therefore can be used as important building blocks for optical devices in these frequencies. The physical principles of three structures are discussed: plasmonic Bragg gratings, designer (spoof) surface plasmon polariton structures, and channel polariton structures. We demonstrate the effectiveness of these structures by realizing various functionalities in QCLs, ranging from beam collimation, polarization control, to multi-beam emission, and spatial wavelength demultiplexing. Plasmonics offers a monolithic, compact, and low-loss solution to the problem of poor beam quality of QCLs and may have a large impact on applications such as sensing, light detection and ranging (LIDAR), free-space optical communication, and heterodyne detection of chemicals. The plasmonic designs are scalable and applicable to near-infrared active or passive optical devices.

1. Introduction

1.1. *Surface plasmons and Zenneck waves*

Plasmonics involves manipulation of surface plasmon polaritons (SPPs), which are collective oscillations of electrons at a metal-dielectric interface interacting with electromagnetic fields.[1–6] In contrast to conventional optical devices, plasmonic structures control light at a subwavelength scale owing to the large spatial frequency components associated with their sharp geometries. SPPs have been intensively studied in the last two decades, and the focus has been primarily on the visible and near-infrared regimes.[4–6] In these spectral ranges, SPPs are characterized by subwavelength confinement to the interfaces and enhanced optical near-fields; strong coupling between metallic structures and SPPs leads to efficient light manipulation. However, when coming to mid-infrared (mid-IR) and terahertz (THz) frequencies, the physical picture changes: surface waves interact weakly with surface charge oscillations and typically extend many wavelengths away from the metallic surface.

The distinction between the two types of SPPs can be explained in terms of the SPP dispersion property, which is the in-plane wavevector of SPPs as a function of frequency. The dispersion curve for SPPs on a planar interface is written as[1–6]

$$\beta(\omega) = k_o(\varepsilon_d\varepsilon_m/(\varepsilon_d + \varepsilon_m))^{1/2}. \quad (1)$$

Here ε_d and ε_m are, respectively, the complex permittivity of the dielectric and metal defining the interface, and k_o is the vacuum wavevector. The separation between the curve defined by $\beta(\omega)$ and the light line in the dielectric, $k_o\varepsilon_d^{1/2}$, determines the SPP confinement, since $|\kappa| = (\beta^2 - k_o^2\varepsilon_d)^{1/2}$ represents the decay rate of near-field amplitude ($|E|$) normal to the interface in the dielectric. While ε_d is, in general, weakly dependant on ω away from the restrahlen bands, ε_m is strongly dispersive. As such, $\beta(\omega)$ deviates from $k_o\varepsilon_d^{1/2}$ as ω increases from zero and the deviation reaches maximum when $\text{Real}(\varepsilon_m) = -\varepsilon_d$ [i.e., the absolute value of the denominator in Eq. (1) is minimized]. The frequency satisfying this condition is called the asymptotic SPP frequency; it is usually in the visible or ultraviolet for metals and it corresponds to the best SPP confinement. At mid-IR wavelengths and above, ε_m is at least one order of magnitude greater than ε_d so that it is hard to distinguish the light line in the dielectric and the SPP dispersion curve: $\beta(\omega) \approx k_o\varepsilon_d^{1/2}$. As a result, $|\kappa|$ is very small, corresponding to poorly confined surface waves, known as Sommerfeld or

Zenneck waves,[7–9] which were first introduced in the context of radio waves propagating on the surface of the Earth.

In short, it is physically meaningful to identify two regimes of SPPs: "surface plasmons" refer to SPPs with subwavelength confinement near the asymptote of the dispersion curve, while "Zenneck waves" correspond to SPPs loosely bonded to the surface at the low-frequency end of the dispersion curve. Despite the fact that there are no surface plasmons in the real sense for a planar metal surface in the mid- and far-infrared, the difficulty is not insurmountable: we will show in later sections that by properly choosing the plasmonic media (e.g., doped semiconductors in the THz), and utilizing suitably engineered plasmonic structures, one can tailor the dispersion properties of SPPs and substantially improve the SPP confinement in the long wavelength range, thereby recovering the surface plasmon regime. Plasmonic devices based on these concepts have been successfully integrated with mid-IR and THz quantum cascade lasers (QCLs) to solve their poor beam quality problem and to realize new functionalities. Before discussing the plasmonic design, we would like to briefly review the basics of QCLs, in particular, the quantum design, the waveguide design, the beam quality issue, and existing methods for QCL collimation.

1.2. *Quantum design and waveguide design of quantum cascade lasers*

In 1971 Kazarinov and Suris predicted in a seminal paper[10] that light amplification is possible in intersubband transitions, i.e., transitions between quantized energy states within one energy band of a semiconductor. The first quantum cascade laser, based on such principle, was invented[11] at Bell Labs in 1994 by Faist, Capasso, Sivco, Hutchinson, and Cho. Since then continuous improvements in the design, material quality, fabrication and thermal management have led to many important achievements, such as extension of the operating wavelength to as short as $3.5\,\mu$m and as long as $19\,\mu$m,[12,13] record optical power of $\sim$120 W pulsed[14] and $\sim$3.0 W continuous-wave (CW) operation[15,16] at room temperature, development of single-mode distributed feedback QCLs,[17,18] and broad band QCLs based on external cavities with tuning ranges of $\sim$200 cm^{-1} CW and $\sim$400 cm^{-1} pulsed.[19–24]

The first THz QCL emitting at 4.4 THz ($\lambda_o \sim 67\,\mu$m) was demonstrated by Köhler *et al.* at the Scuola Normale Superiore in Pisa, Italy, in

collaboration with Cambridge University.[25] Intensive research over the past few years has produced rapid improvements.[26–28] At present, spectral coverage has been demonstrated from 0.84 to 5.0 THz.[29,30] Record operating temperature is 186 K, pulsed,[31] and 117 K, CW.[32] The output power can routinely reach a few tens of mW, up to a couple of hundred mW for both pulsed and CW operations.[33]

The QCL relies on a radically different process for light emission compared to diode lasers. Instead of using opposite charge carriers in semiconductors (i.e., electrons and holes) at the bottom of their respective conduction and valence bands, which recombine to produce light of frequency $\nu \approx E_g/h$, where E_g is the energy bandgap and h is Planck's constant, QCLs use only one type of charge carriers (i.e., electrons) that undergo quantum jumps between energy levels E_n and E_{n-1} to create a laser photon of frequency $(E_n - E_{n-1})/h$. These energy levels do not exist naturally in the constituent materials of the active core but are artificially created by structuring the active core in ultra-thin layers known as quantum wells of nanometric thickness. The motion of electrons perpendicular to the layer interfaces is quantized and characterized by energy levels whose difference is determined by the thickness of the wells and by the height of the energy barriers separating them.

In QCLs, an electron remains in the conduction band after emitting a laser photon. The electron can therefore easily be recycled by being injected into an adjacent identical active region, where it emits another photon, and so forth. To achieve this cascading emission of photons, active regions are alternated with doped electron injectors and an appropriate bias voltage is applied. The active-region/injector stages give rise to an energy staircase in which photons are emitted at each of the steps. The number of stages typically ranges from 20 to 50 for lasers designed to emit in the mid-IR range, and it is typically a couple of hundreds for THz QCLs. This cascade effect is responsible for the high power that QCLs can attain.

Mid-IR QCLs are typically based on a dielectric waveguide structure in which the active core (i.e., active-region/injector stages) is embedded between semiconductors with smaller refractive indices; see Fig. 1(a). The thickness of the active core, t, is usually of the order of the wavelength in the laser material, λ_o/n, where n is the mode refractive index and λ_o is the wavelength in vacuum. As such, the output of edge-emitting mid-IR QCLs has the intrinsic problem of large beam divergence caused by diffraction at the small emission aperture. The divergence angle in the fast axis direction, i.e., perpendicular to the material layers, can be estimated by λ_o/t and is

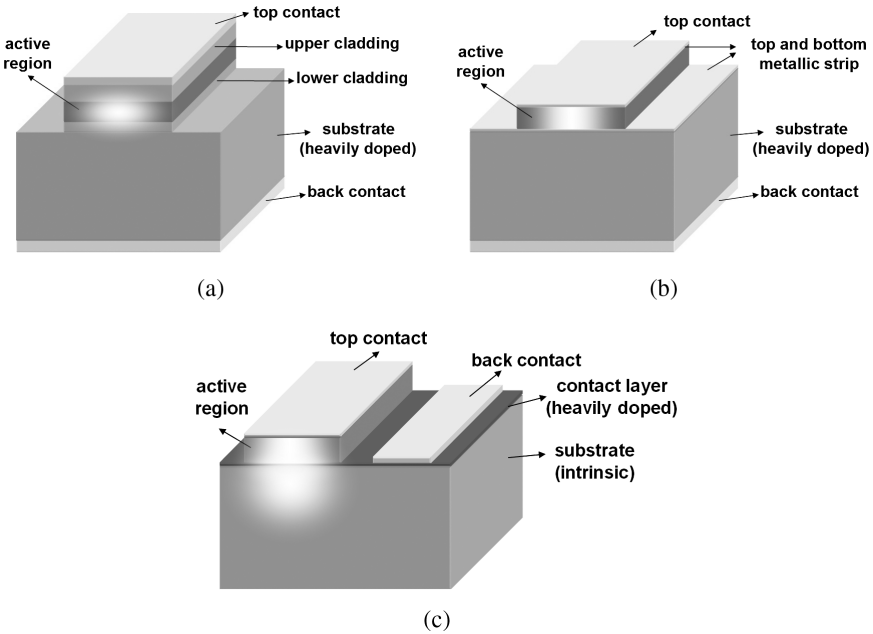


Fig. 1. Waveguide designs. (a) Dielectric waveguide for mid-IR QCLs. (b) Double-metal waveguide for THz QCLs. (c) Single plasmon waveguide for THz QCLs. The confinement of the laser mode is based on surface waves bound to the top metallization and on the quasi-metallic confinement provided by a thin, heavily doped semiconductor contact layer placed below the active core. Schematics of laser waveguide modes are shown in white.

typically a few tens of degrees. The divergence in the horizontal direction is dependent on the waveguide width and is typically larger than 10 degrees.

THz QCLs with the highest operating temperature and lowest threshold to date take advantage of the high optical confinement (near 100%) and heat removal properties of a double-metal waveguide design, in which the laser active core, about 10 μm thick, is sandwiched between two metal strips;[26–28] see Fig. 2(b). However, this leads to non-Fresnel reflection at the subwavelength laser apertures, which gives rise to inefficient power outcoupling and poor beam quality. For example, the power reflectivity of laser modes at the aperture can be up to 90%.[34] The laser emission is extremely divergent with divergence angle $\sim 180^\circ$ perpendicular to the semiconductor layers, and this represents a particularly serious problem for applications such as the far-infrared heterodyne detection of chemicals since the output of THz QCLs (local oscillator) must be focused into a small area Schottky diode mixer.[35]

Collimation of QCLs is conventionally conducted using refractive lenses or metallic mirrors. These optical components are usually bulky and need careful alignment; they are certainly expensive because of a limited choice of infrared transparent materials (germanium, zinc selenide, polyethylene, etc), thicker optical coatings compared to visible/near-IR optics, and relatively fewer suppliers. In a similar approach, micro optical components such as silicon microlenses[36] and metallic horn antennas[37] have been mounted onto laser apertures to reduce beam divergence and to increase power output. However, this method requires meticulous manipulation and alignment of small optical components, which affects device yield and robustness.

QCLs that intrinsically produce small divergence beams are highly desirable. A number of approaches have been demonstrated in this respect. First, tapered laser waveguides or wide-ridge waveguides have been demonstrated to emit diffraction limited laser beams with significantly reduced lateral divergence,[14,38–40] but this does not solve the problem of vertical divergence. Second, QCLs have been integrated with grating outcouplers. This approach relies on constructive interference between multiple surface emissions to reduce beam divergence. For example, one-dimensional (1D) integrated grating out-couplers have helped to create mid-IR[41–44] and THz QCLs[45–47] with greatly reduced beam divergence in the direction along laser ridges. Surface-emitting ring or disc QCLs with 2D second-order gratings have demonstrated promising results of 2D collimation.[48–50] Third, QCLs have been processed into edge-emitting or surface-emitting photonic crystal structures.[51–54] Such devices operate on photonic band-gap modes (microcavity) or photonic band-edge modes (distributed feedback) and have been demonstrated to achieve controllable far-field emission patterns. The drawback of the surface emission scheme is that it results in device structures with reduced mode confinement and therefore increased laser threshold current density, which will likely lead to reduced maximum device operating temperatures in CW operation.

2. Methods to Tailor the Dispersion Properties of Mid-IR and THz Surface Plasmon Polaritons

The dispersion properties of SPPs can be modified by employing a variety of mechanisms. We are going to explore three of them; the plasmonic Bragg gratings rely on collective interference between a large number of periodically spaced elements, the spoof SPP structures exploit the effective

medium regime, and the channel polariton grooves are based on localized metal-insulator-metal waveguide modes.

Consider a metallic surface patterned with a Bragg grating with periodicity Λ. Like electron waves interacting with a periodic potential or light waves interacting with a photonic crystal, SPPs with certain frequencies experience strong Bragg diffraction from the grating (i.e., reflection or coupling out of the plane) so that they are forbidden to propagate on the plane.[55–58] In the dispersion diagram, small bandgaps open up in the vicinity of $\beta = m\pi/\Lambda$ (Fig. 2), m being an integer. Assuming we operate in a region close to one of these bandgaps where the dispersion curve $\omega(\beta)$ deviates from the light line (see Fig. 2), the decay rate of SPPs normal to the surface, $|\kappa| = (\beta^2 - k_o^2\varepsilon_d)^{1/2}$, will be increased and consequently the SPP confinement will be improved. $\beta = \pi/\Lambda$ corresponds to a first-order grating, which reflects SPPs; the second-order Bragg diffraction, $\beta = 2\pi/\Lambda$, primarily couples SPPs into free-space waves in a direction normal to the interface. We shall see later that plasmonic structures based on $\beta \approx 2\pi/\Lambda$ (or $\beta \approx 0$ in the reduced zone scheme, corresponding to frequency ω_{IR} in Fig. 2) can successfully shape the output of mid-IR QCLs.

Bragg gratings modify SPP dispersion in narrow frequency bands and the degree of modification is limited. To what extent can one engineer the dispersion properties of SPPs? The asymptotic SPP frequency seems to be intrinsically determined by the properties of the interface materials: by inserting a simplified Drude model $\varepsilon_m = 1 - \omega_p^2/\omega^2$ in the formula

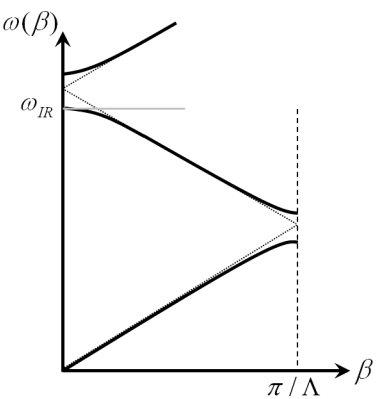


Fig. 2. Schematic dispersion curve of SPPs on a metallic surface sculpted with a second-order grating (solid curve). Gray horizontal curve: operating frequency. Black dotted curve: light line. Black dashed curve: boundary of the first Brillouin zone.

Real$(\varepsilon_m) = -\varepsilon_d$, one obtains the asymptotic SPP frequency $\omega_{\text{spp}} = \omega_p/(1+\varepsilon_d)^{1/2}$, where $\omega_p = (Ne^2/(\varepsilon_o m^*))^{1/2}$ is the bulk plasma frequency; e, m^*, and N are the elementary charge, the effective mass, and the density of the electrons, respectively. The expression for ω_p implies that the asymptotic SPP frequency can be adjusted via changing N. One such approach is to use doped semiconductors whose carrier concentrations are much lower than those in metals and are able to be tuned by external excitations.[58–61] This holds great promise for THz plasmonics: for example, it is estimated that GaAs with a doping level higher than about $n = 10^{17}\,\text{cm}^{-3}$ turns metallic for 3-THz frequency light.

A completely different method to engineer the SPP dispersion is to perforate metallic surfaces with subwavelength features. By introducing localized electromagnetic modes strongly interacting with the structured surfaces, one is able to reduce the asymptotic SPP frequency by orders of magnitude into infrared and GHz frequency ranges. Such perforated surfaces are called designer plasmonic structures or "spoof" SPP structures, as introduced by Pendry, Martín-Moreno, and García-Vidal[62,63] and observed on structured metals at THz[64] frequencies. Interestingly, similar structures were studied by microwave engineers as early as 1950s[65,66] but have been largely unknown to the optics community.

Consider an array of grooves with subwavelength periodicity, p, sculpted on the surface of a perfect electric conductor; see Fig. 3(a). The asymptotic frequency is given by[63]

$$\omega_{\text{spp}} = \pi c/2d \tag{2}$$

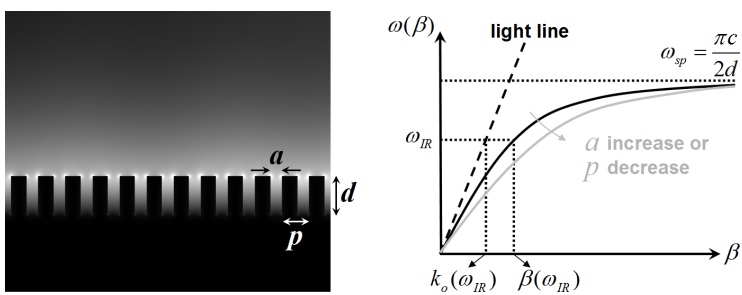


Fig. 3. Physics of the spoof SPPs. (a) Geometry of the spoof SPP structure and calculated electric-field distribution ($|E|$). (b) Schematic dispersion diagram for infrared spoof SPPs on a corrugated surface of a perfect electric conductor.

where d is the groove depth, and c is the speed of light in vacuum. Physically, Eq. (2) (or equivalently $d = \lambda_o/4$) corresponds to the first-order standing wave inside the groove cavities: the phase accumulated during a roundtrip of the groove cavity mode is 2π (wavevector times $2d$ contributes π, while the other π originates from reflection of the mode at the bottom of the grooves). As SPPs strongly couple to the resonant cavity modes, the group velocity of the SPPs markedly reduces to near zero.

An alternative way to understand Eq. (2) is to treat each groove as a waveguide shorted at one end. According to transmission line theory, the input impedance of such terminated waveguide is[67]

$$\eta = i\eta_o \tan(k_o d) \tag{3}$$

where d, the groove depth, is now the length of the waveguide, and $\eta_o \sim 377\,\Omega$ is the free-space impedance. Equation (3) predicts that the impedance will be resistive if $d = \lambda_o/4$, while it will be inductive (capacitive) if $d < \lambda_o/4$ ($d > \lambda_o/4$). Resistive impedance means that current and voltage are in phase and therefore net power will flow into the grooves and finally be dissipated through ohmic absorption, which is characteristic of spoof SPPs near the asymptotic frequency.

Another type of spoof SPP structures comprise square indentations of side a arranged on a $d \times d$ lattice perforated through a metallic film. Unlike the previous structure, where TEM-like modes inside the groove cavities have no cut-off frequency, here each square indentation can be treated as a square metallic waveguide and electromagnetic modes with frequencies higher than $\pi c/a$ will couple into waveguide modes. Therefore, the asymptote of the SPP dispersion is given by the cut-off frequency of a square waveguide of *width* a as[62]

$$\omega_{\text{spp}} = \pi c/a. \tag{4}$$

This is in contrast to the previous case, where the *depth* of the grooves d determines the asymptote. Note that Eq. (4), or equivalently $a = \lambda_o/2$, implies that we are no longer in the effective medium regime. Therefore, in practice one would prevent operating in a region too close to the asymptote.

The plasmonic Bragg grating and the spoof SPP structures both involve interaction between SPPs and an assembly of surface textures. The difference is that the spoof SPP structures explore the effective medium regime and therefore their effectiveness does not rely on collective behavior of a large number of elements as in the case of plasmonic Bragg gratings,

where interference of waves coherently scattered by many periods is vital for their functionality. Because of this, spoof SPP structures are more flexible to use: by stacking short sections of them with different geometries or by adiabatically tailoring the geometry of the constituent elements, complex optical devices can be built to manipulate SPPs. We will see a couple of examples later.

The third plasmonic structure involves interaction between SPPs and a single element: a channel sculpted into the metallic surface. Such channel waveguides have been studied extensively in recent years.[68–73] They support the so-called channel polaritons that are localized to the immediate vicinity of the channel and travel along it, provided that the channel is sufficiently deep and that inclined walls of the channel are steep enough. The dispersion properties of these channel modes are different from those of SPPs on planar interfaces: channel polaritons have larger in-plane wavevectors or greater effective mode indices.

In the simplest sense, channel waveguides are similar to metal-insulator-metal (MIM) waveguides and the properties of channel polaritons can be understood from the standpoint of transmission line theory. Let us treat the two walls of a channel groove as a pair of parallel perfect conducting strips. Assuming that the channel depth d is much larger than its width a, and that the propagating modes are transverse electromagnetic (TEM) waves, waveguide theory[67] predicts that the capacitance per unit length is $C = \varepsilon_o d/a$, while the inductance per unit length is $L = \mu_o a/d$. As a result, the characteristic impedance of the channel grooves, modeled as coplanar strip transmission lines, will be

$$\eta_{wg} = (L/C)^{1/2} = \eta_o a/d, \tag{5}$$

which is much smaller than the free-space impedance η_o. Such a difference in impedance ensures that channel polariton modes are well localized in the channels and do not leak into SPPs on planar surfaces outside the channels. For constant width a, a deeper channel leads to a more distinctive difference between η_{wg} and η_o and therefore modes are better confined. We will show later that channel polaritons are employed to effectively spread THz energy over a large area on device facets, thereby reducing beam divergence in the far-field. To calculate the dispersion properties of channel polaritons, readers can refer to Ref. (68), which provides thorough theoretical and numerical studies of channel polaritons, or refer to the series of papers by Bozhevolnyi and others,[71–74] which describe an effective-index method for semi-quantitative calculations.

3. One-Dimensional Collimators for Mid-IR QCLs

The 1D collimator[75] consists of a slit aperture and a plasmonic grating patterned on the metal-coated laser facet [Fig. 4(a)]. The structure is based on experimental and theoretical works on the beaming of optical radiation by aperture-groove structures.[76,77] The slit aperture and the grating grooves are oriented perpendicular to the laser polarization. The aperture has subwavelength vertical dimension; it couples part of the laser emission into SPPs propagating along the z-direction on the device facet, where the plasmonic second-order grating coherently scatters their energy into the far-field, thereby breaking the diffraction limit set by the small emission aperture of the original laser. Plasmonic collimation is essentially an antenna array effect. We use as a figure of merit the concept of antenna directivity to characterize collimation in the vertical direction, and it is defined as $D = 10\ \log_{10}(2\pi I_{\text{peak}}/I_{\text{total}})$, where I_{peak} and I_{total} are, respectively, the peak intensity and the total intensity of the far-field.

We performed systematic 2D simulations to help design a structure with optimal beam collimation characteristics. It is critical to fine tune the

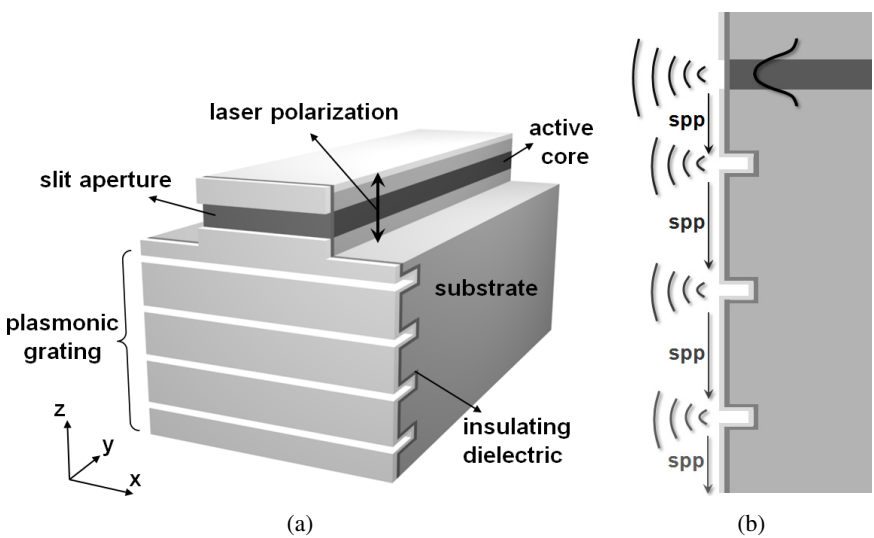


Fig. 4. (a) Schematic of a 1D collimated laser. (b) Cross sections of the device. The grooves are sculpted directly into the laser facet, followed by conformal coatings of an insulating layer and an optically thick metal layer. Focused ion beam (FIB) milling is used to finally open the slit aperture on the active core.

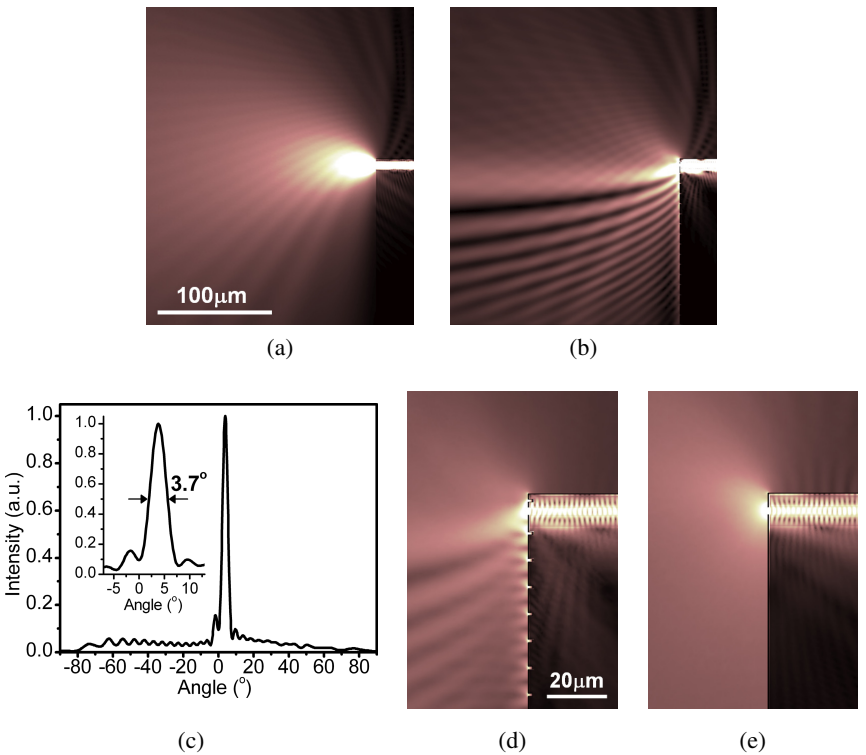


Fig. 5. (a) Simulated intensity distribution ($|E|^2$) of an original unpatterned QCL emitting at $\lambda_o = 9.9\,\mu$m. The simulation plane is perpendicular to the laser materials layers and along the symmetry plane of the waveguide ridge. (b) Simulated intensity distribution of the QCL patterned with a 1D collimator containing 15 grating grooves. (c) Calculated far-field intensity distribution in the vertical direction for the device shown in (b); inset: zoom-in view of the central lobe. (d) Simulated electric-field distribution ($|E|$) around the slit and the first 7 grating grooves. (e) Simulated electric-field distribution of a QCL with only the slit aperture.

groove width and depth so that the propagation distance of SPPs on the facet is comparable to the length of the grating in the z-direction. This ensures that the plasmonic grating scatters almost all the energy of the SPPs into the far field and that maximum number of scattering elements are involved to minimize far-field divergence.

Figure 5(a) shows the simulated intensity distribution ($|E|^2$) of an unpatterned QCL; the laser beam is observed to diverge rapidly in the vertical direction after exiting the laser waveguide. As a comparison,

Fig. 5(b) is a simulation for the device after integration with a plasmonic collimator consisting of 15 grating grooves with optimized geometry. It shows that light emits from the entire patterned laser facet. Figure 5(c) is the calculated vertical far-field profile ($|E|^2$) for the collimated device: the full-width at half-maximum (FWHM) divergence angle of the central lobe is reduced from $\sim 60^\circ$ of the original device to 3.7°. The optical background is relatively uniform as a function of angle and has an average value less than 10% of the peak intensity. Figure 5(d) shows the electric-field distribution ($|E|$) around the slit and the first few grating grooves; the SPPs are observed to localize within about one free-space wavelength from the device facet, indicating an efficient interaction with the grating. This is in sharp contrast to the weak coupling between mid-IR SPPs and planar gold surfaces, where the decay distance of SPPs in the air is estimated to be $\sim 10\lambda_o$. The latter case is illustrated in Fig. 5(e), which shows that SPs are loosely bound to the metal/air interface for a device with a slit aperture and without the grating.

As discussed above, the reduction in far-field divergence is essentially an antenna array effect. Simulations show that indeed the divergence angle is inversely proportional to the number of the grating grooves N at least up to $N = 60$ [Fig. 6(a)]. The peak intensity in the far-field is proportional to N^2 up to $N = 25$ [Fig. 6(b)]; for $N > 25$, the effect of loss and lateral spreading of SPPs gives rise to a weaker variation of the peak intensity with N.

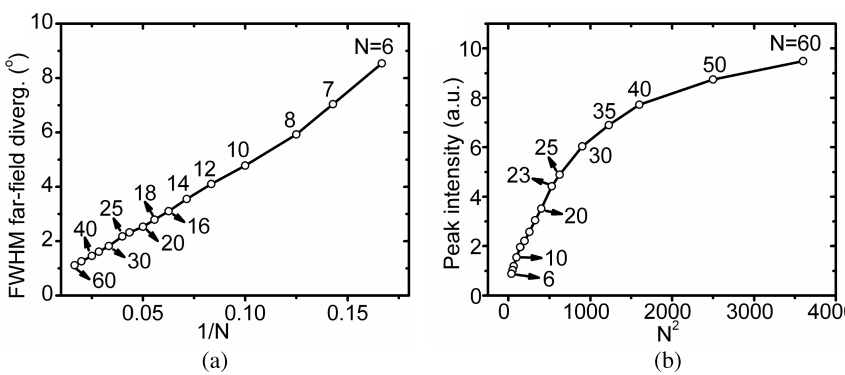


Fig. 6. (a) Simulation results showing the relation between far-field divergence and $1/N$, the inverse of the number of grating grooves. (b) Simulation results showing the relation between far-field peak intensity and N^2.

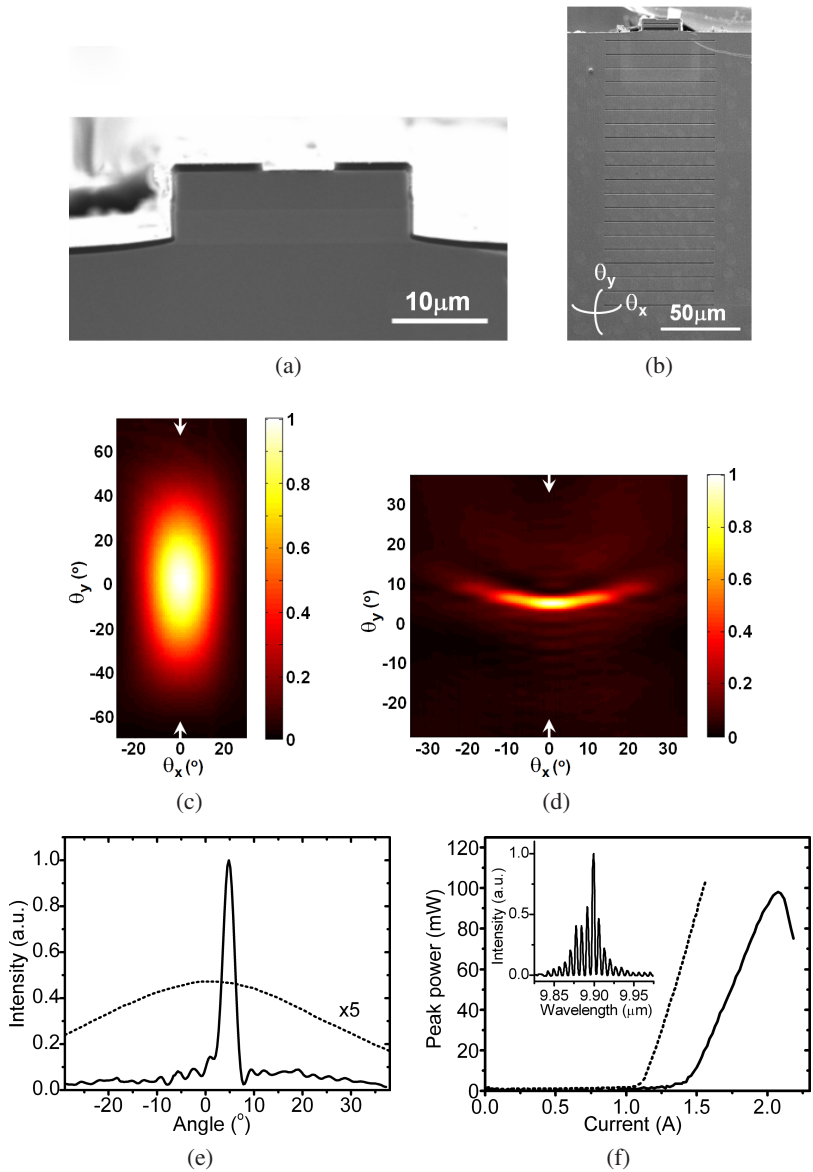


Fig. 7. (a) and (b) SEM images of the facet of a $\lambda_o = 9.9\,\mu$m QCL before and after patterning a 1D plasmonic collimator. (c) and (d) Measured 2D far-field emission patterns for (a) and (b), respectively. (e) Vertical line scans of (c) (dotted curve) and (d) (solid curve) along the arrows. (g) The dotted and solid curves are light output versus drive current (LI) characteristics of the unpatterned and patterned devices, respectively. Inset: spectrum of the collimated device taken at $I = 1.8$ A.

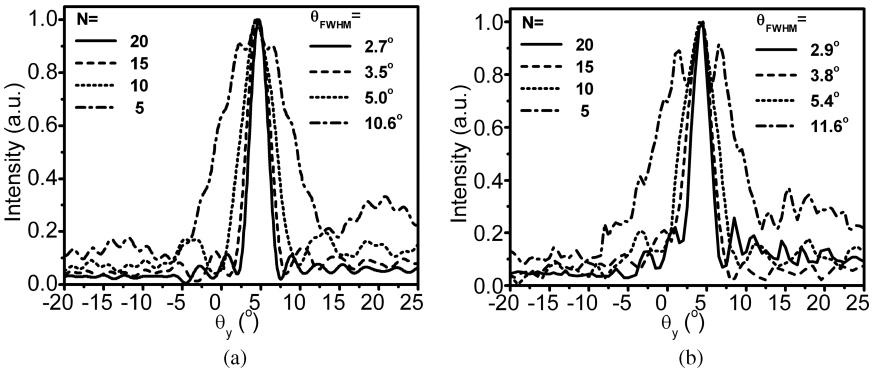


Fig. 8. (a) and (b) Simulated and measured vertical far-field intensity profiles of devices with 1D collimators containing different number of grating grooves.

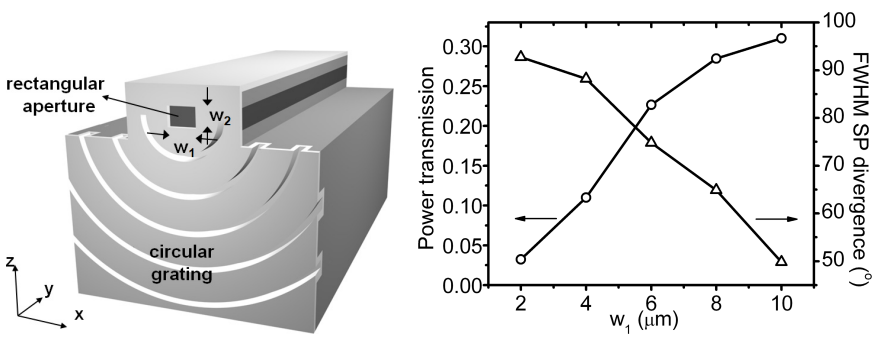
Figures 7(a) and 7(b) show scanning electron microscope (SEM) images of a $\lambda_o = 9.9\,\mu$m QCL before and after patterning a 1D plasmonic collimator. The corresponding 2D far-field emission patterns are presented in Figs. 7(c) and 7(d), respectively, demonstrating a strong reduction of the beam divergence in the vertical direction. The slight curvature of the far-field pattern in Fig. 7(d) is an edge effect due to the finite lateral size of the slit aperture. The vertical line scans of the 2D far-field are provided in Fig. 7(e), which shows that the divergence angle is reduced from about 62° to about 2.9° for the device with 20 grating grooves. The average intensity of the optical background is below 10% of the peak value. We note that there is no noticeable damage to the collimator after high peak-power operation (∼100 mW peak power, pulsed mode). The beam quality factor M^2 of the device is determined to be about 2.5 based on measurements of the variation of the beam waist along the propagation direction. M^2 stands for the factor by which the divergence angle of a laser beam is larger than that of a Gaussian beam, assuming the two have the same beam waist. We calculated based on Figs. 7(c) and 7(d) that the directivity D is about 17.7 dB for the collimated device, while D is only about 7.2 dB for the original unpatterned laser. The lateral beam divergence after fabrication of the collimator is similar to that of the original laser. Figure 7(f) shows the light output versus current (LI) characteristics before and after defining the collimator, demonstrating a maximum output power of about 100 mW. We showed experimentally that the divergence angle is roughly inversely proportional to N; see Fig. 8. Devices with larger N have smaller average background and larger slope efficiencies.

4. Two-Dimensional Collimators for Mid-IR QCLs

For 2D collimation,[78,79] it is crucial to efficiently propagate SPPs in two dimensions on the laser facet. We choose a design consisting of a rectangular metallic aperture and a circular plasmonic second-order grating; see Fig. 9. SPPs propagate preferentially in the vertical direction due to the TM polarization of QCLs. To efficiently launch SPPs into the lateral direction, the width of the aperture w_1 has to be subwavelength, but this will limit power output. A trade-off is therefore to be sought between collimation and power throughput.

We performed systematic 3D simulations to study the evolution of the power outflow and the lateral spreading of SPPs as a function of w_1 [Fig. 9(b)]. We increase w_1 in steps while maintaining the vertical aperture size w_2 to be approximately equal to the thickness of the laser active core ($\sim 2\,\mu$m). It is found that, for a small aperture ($2\times 2\,\mu\text{m}^2$), the SP spreading angle is very close to 90°, consistent with the $(\cos\theta)^2$ rule of a dipole emitter. When w_1 is nearly equal to $\lambda_o = 8.06\,\mu$m, the power throughput is improved by a factor of $\sim$10 compared to $w_1 = 2\,\mu$m, while the lateral spreading of SPs is still significant.

We fabricated 2D plasmonic collimators on $\lambda_o = 8.06\,\mu$m buried heterostructure (BHT) QCLs grown by MOCVD. A BHT device has regrowth regions on the lateral sides of the laser active core [Fig. 10(a)]; it is more efficient in heat removal compared with ridge devices, which is critical for CW operations. Their active core has a cross section of $2.1\,\mu\text{m} \times 9.7\,\mu\text{m}$



(a) (b)

Fig. 9. Schematic of a QCL integrate with a 2D plasmonic collimator. (b) Simulated FWHM SPP lateral spreading angle and power throughput as a function of the lateral aperture size w_1 for a device with 20 circular grating grooves.

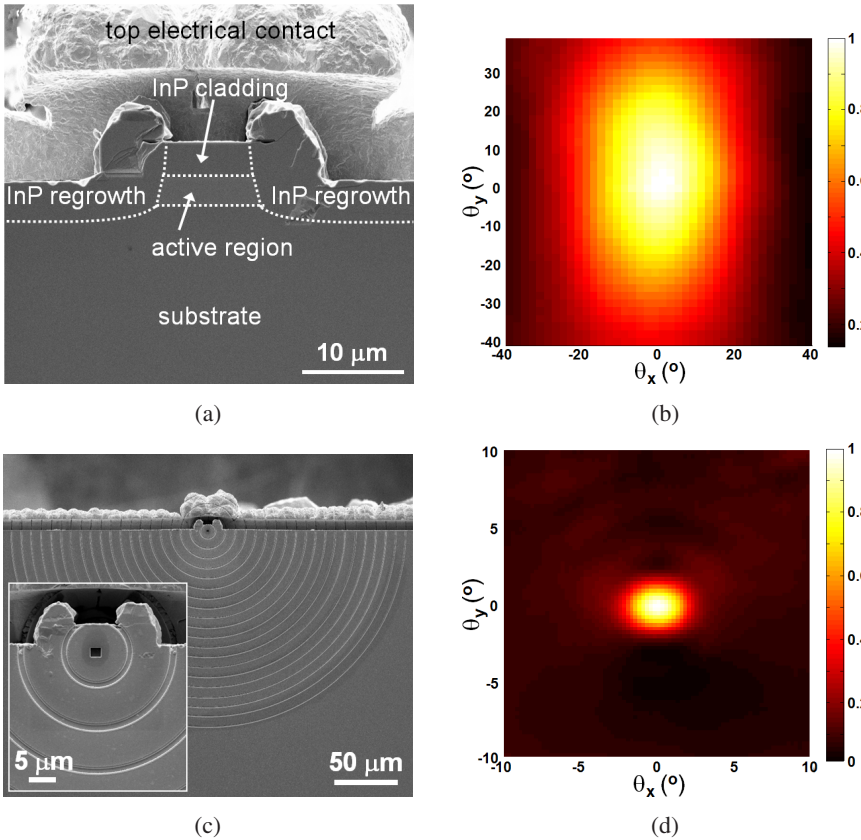


Fig. 10. (a) SEM image showing the facet of an unpatterned $\lambda_o = 8.06\,\mu$m buried heterostructure QCL. (b) Measured emission pattern of the device in (a). (c) SEM image of the device with a 2D plasmonic collimator. Inset: zoom-in view. (d) Measured emission pattern of the device in (c).

in the vertical and lateral directions and the divergence angles before fabrication of the collimator are $\theta_{||} = 42°$ and $\theta_{\perp} = 74°$ [Fig. 10(b)].

Figure 10(c) shows one representative device patterned with 20 circular grooves and a small aperture with $w_1 \times w_2 = 2.8 \times 1.9\,\mu\text{m}^2$. The device exhibits greatly reduced divergence angles of $\theta_{||} = 3.7°$ and $\theta_{\perp} = 2.7°$. The main lobe of the far field contains about half of the total emitted power. The directivity D is about 8 dB for the unpatterned lasers, while it is about 27 dB after collimation. The beam quality factor M^2 of the device is about 2.0 in both the vertical and lateral directions.

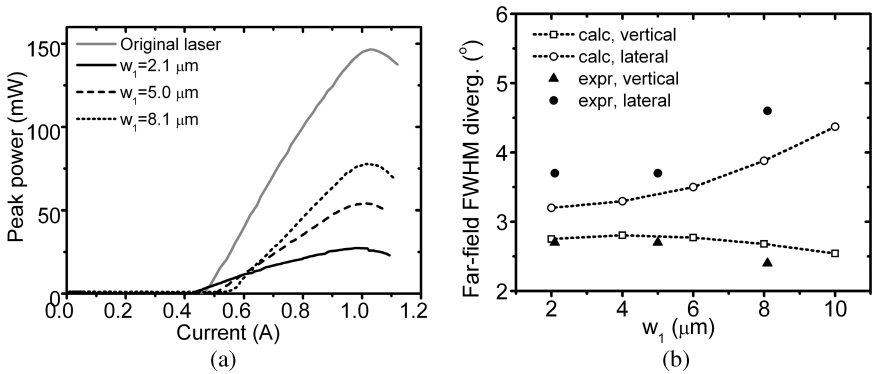


Fig. 11. (a) LI characteristics for the device with a 2D plasmonic collimator. (b) Divergence angles in the vertical and lateral directions as a function of the lateral aperture size w_1.

We increased the lateral aperture size w_1 of the device stepwise using FIB milling to investigate its impact on $\theta_\perp$, $\theta_{||}$, and the power output. The vertical aperture size w_2 was kept constant at 1.9 μm. The LI characteristics for different aperture sizes are shown in Fig. 11(a). Larger w_1 corresponds to larger slope efficiency and therefore a higher maximum output power. As w_1 increases, the device experiences a gradual increase of the lasing threshold, due to decreased effective facet reflectivity.

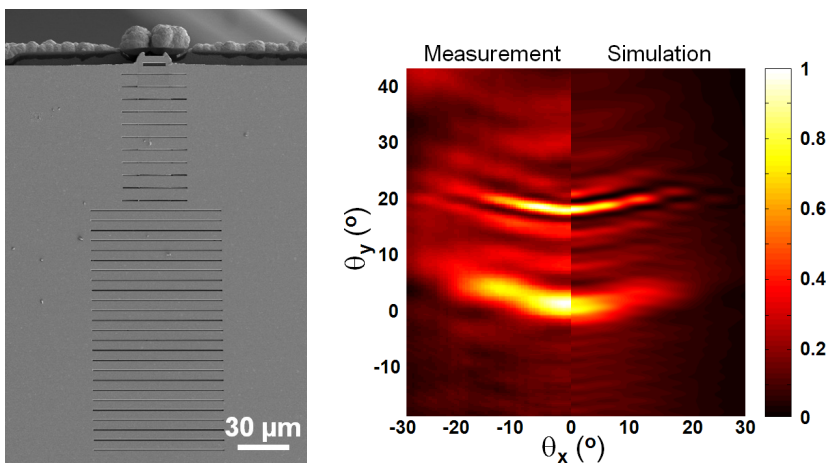
Figure 11(b) summarizes the measured far-field beam divergence angles for different aperture sizes. As w_1 increases, the lateral divergence angle $\theta_{||}$ increases due to less efficient lateral spreading of SPPs; on the other hand, $\theta_\perp$ remains essentially constant because the vertical propagation of SPPs is insensitive to the change in aperture geometry. For the widest aperture with $w_1 = 8.1\,\mu$m, the maximum output power is about 50% of that of the original unpatterned laser, while the divergence angles ($\theta_\perp = 2.4°$ and $\theta_{||} = 4.6°$) are still significantly smaller than those of the original device.

5. Multi-Beam QCLs

In plasmonic second-order gratings, the grating period is approximately equal to the SPP wavelength, yielding a collimated beam normal to the laser facet. If instead the grating period differs from the SPP wavelength, constructive interference will occur in a direction away from the normal of the facet.[80] We utilized this concept to split the emission of a single wavelength QCL into two 1D collimated beams in different directions. Two

successive plasmonic gratings with different periods and lengths can be defined on the device facet. By tuning the grating period, the emission direction of the beam originating from each individual grating can be controlled; by tuning the grating length (i.e., the number of grooves per grating), we can independently control the intensity of the emitted beams. Note that since mid- and far-infrared SPPs are able to propagate for significant distances comparable to at least hundreds of free-space wavelengths on a flat metal surface, several gratings can be patterned on the laser facet to produce multiple laser beams with controlled intensity and direction.

Figure 12 shows experimental and simulation results for a single-wavelength QCL emitting two beams. The grating closer to the aperture is designed to produce a beam propagating normal to the facet. The grating farther away from the aperture has a smaller grating period [Fig. 12(a)]. It provides a reciprocal lattice vector larger than the wavevector of the SPPs, and therefore light diffracted from the grating is deflected away from the surface normal toward the top electrical contact of the device. The deflection angle is approximately 20° from the surface normal; see 2D far-field emission patterns in Fig. 12(b). Because the intensity of SPPs



(a) (b)

Fig. 12. (a) SEM image of a $\lambda_o = 8.06\,\mu$m QCL patterned with two plasmonic gratings. The grating closer to the aperture contains 11 grooves and has a 7.8-μm period. The other grating contains 25 grooves and has a 6-μm period; it has larger lateral dimensions to account for the lateral spreading of SPPs. (b) Measured (left half) and simulated (right half) far-field emission patterns of the device.

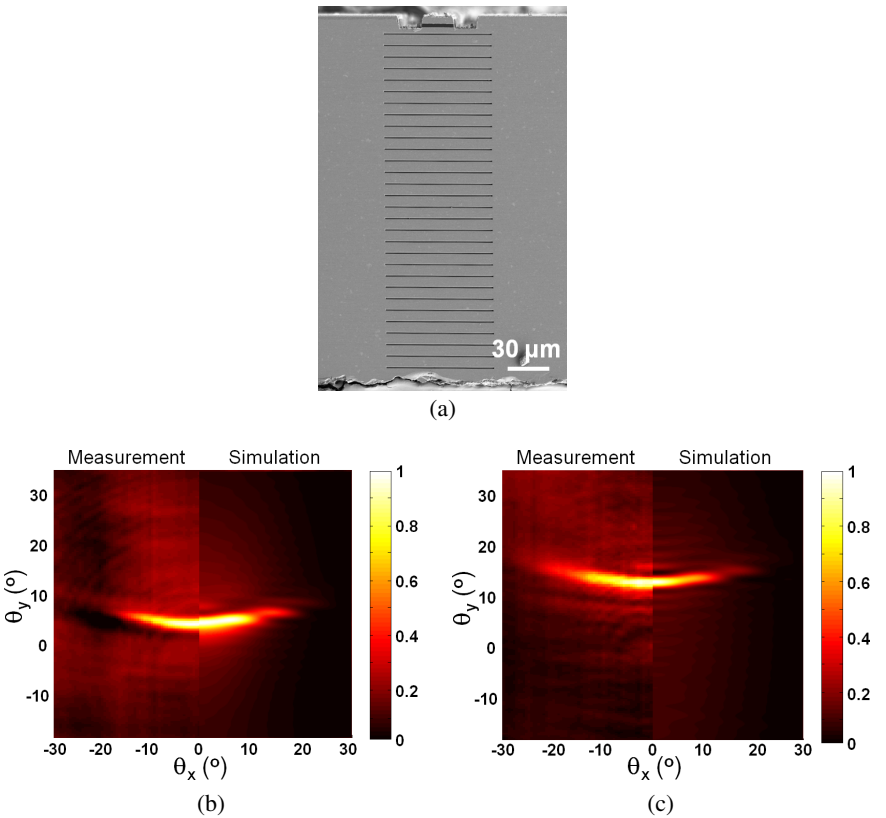


Fig. 13. (a) SEM image of a dual-wavelength QCL patterned with a demultiplexer. (b) Measured (left half) and simulated (right half) emission pattern of the $\lambda_o = 9.3\,\mu$m component of the laser emission. (c) Measured (left half) and simulated (right half) emission pattern of the $\lambda_o = 10.5\,\mu$m component of the laser emission.

decreases as they propagate through the grating grooves due to scattering from the grooves and losses within the metal, the second grating was chosen to contain more grooves than the first one to allow the two emitted beams to have similar peak intensity; see Fig. 12(b).

To demonstrate lasers emitting spatially demultiplexed wavelengths, we used mid-IR QCLs with two stacks of active regions: one based on a two-phonon design emitting at $\sim$10.5 μm and the other based on a bound-to-continuum (BTC) design emitting around $\sim$9.3 μm. It is essential to devise a plasmonic grating efficient in scattering light at both wavelengths. We chose a design that is the trade-off between the optimal structures for the two wavelengths.

Simulations and experimental results for a dual-wavelength device are presented in Fig. 13. A single grating with a period of $8.5\,\mu$m was used; it deflects the $\lambda_o = 9.3\,\mu$m component away from the facet normal toward the top contact by $\sim 3^\circ$ and the $\lambda_o = 10.5\,\mu$m component by $\sim 12^\circ$. The measured direction and divergence angle of the two beams are in good agreement with simulations; see Figs. 13(b) and 13(c).

The plasmonic beam splitter for single-wavelength lasers can be potentially useful for applications requiring a reference beam and a probe beam such as interferometry and holography. Our designs should be scalable and applicable to near-IR lasers and passive optical components. For example, the spatial wavelength demultiplexing component for dual-wavelength lasers can function as a demultiplexer for optical fibers.

6. Mid-IR QCLs with Integrated Plasmonic Polarizers

Light sources with a desirable polarization state are of high interests for many applications. Polarization multiplexing is used in optical fiber communications and satellite communications to enhance data transmission rate; circularly-polarized light is of great importance in chemistry and biology for detecting chiral molecules. However, semiconductor lasers are mostly linearly-polarized either TE or TM, which is determined by the optical selection rules of the gain medium.[81,82]

We are going to discuss plasmonic structures that can realize circularly-polarized laser emissions.[83] But before that, we shall first see how to control the linear polarization direction of a QCL, which is a key component for achieving circular polarization. To project the polarization of a QCL onto a direction defined by θ, we use an aperture-grating structure similar to the 1D plasmonic collimator but rotated by θ from the vertical direction; see Fig. 14(a). The slit aperture is illuminated by laser waveguide modes polarized in the vertical direction. Only the component of the laser polarization perpendicular to the slit efficiently generates charge oscillations across the two edges of the slit, and therefore radiates directly into the far-field and couples to SPs on the device facet. The output of the device as a result of the interference between the direct emission from the aperture and scattered light from the grating grooves will be polarized along the θ direction.

Figure 14(a) shows an SEM image of a QCL facet patterned with a linear polarizer with $\theta = 45^\circ$. Its emission divergence is reduced along the $\theta = 45^\circ$ direction; see Fig. 14(b). To determine the polarization of the

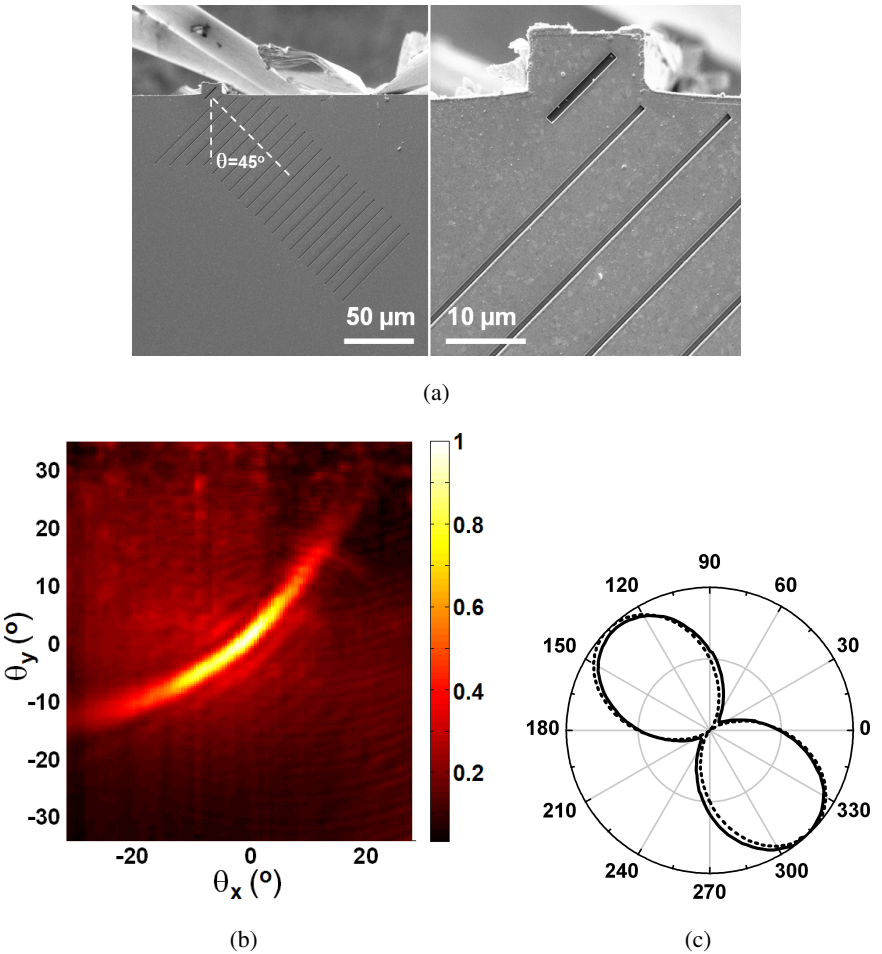


Fig. 14. (a) SEM images of a QCL integrated with a linear polarizer. The orientation of the slit aperture and the grating is 45° with respect to the vertical direction. The right panel is the zoom-in view of the left panel. (b) Measured emission pattern of the device in (a). (c) Measured device output as a function of the rotation angle of the wire-grid polarizer. The solid curve is experimental data and the dotted curve is calculation assuming a 45° linearly-polarized light.

emission, an MCT (mercury-cadmium-telluride) detector with a wire-grid polarizer placed in front was positioned in the far-field where the laser intensity is maximum. As the axis of the wire-grid polarizer was rotated, the detected power showed two maxima corresponding to 135° and −45° and a high cross-polarization suppression ratio, demonstrating the ability

to efficiently project the initial vertical polarization of the laser onto the 45° direction; see Fig. 14(c).

A circularly-polarized beam can be constructed by coherently combining two linearly-polarized beams satisfying three conditions: the two beams should have perpendicular polarizations, 90° phase difference, and the same amplitude. The first condition can be met by defining two aperture-grating structures on the device facet with their respective orientation angles equal to $\theta = \pm 45°$; see Fig. 15(a). The distance between the aperture and the nearest groove [d_1 and d_2 in Fig. 15(a)] determines the phase of the scattered light. The second condition can be satisfied by choosing $|d_1 - d_2| \sim \lambda_{sp}/4$, so that the two beams produced by the left and the right gratings have a 90° phase difference.

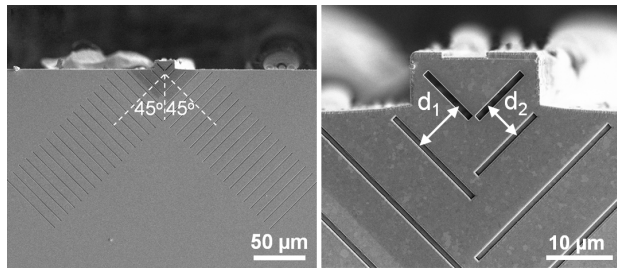
The efficiency of coupling laser emission into SPPs on the facet is a periodic function of d, the separation between the aperture and the first grating groove. The efficiency reaches maxima when the reflected SPPs by the grating and the SPPs launched at the aperture and propagating towards the device top contact are out of phase. This condition translates to

$$2\pi + 2dk_{sp} = (2m + 1)\pi \tag{6}$$

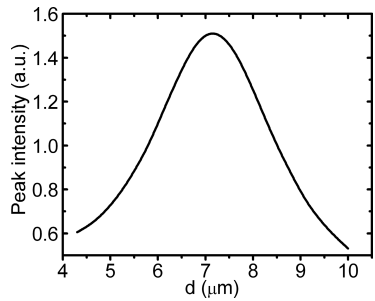
where the first term, 2π, is the phase picked up by SPPs on reflection at a second-order grating;[84] the second term originates from path differences of the two SPP components; m is an integer. Equation (6) predicts that the coupling efficiency of SPs on the facet is the highest for $d = 3\lambda_{sp}/4$, which is confirmed by our simulations; see Fig. 15(b). We chose d_1 and d_2 to be respectively larger and smaller than $d = 3\lambda_{sp}/4$, and $|d_1 - d_2| \sim \lambda_{sp}/4$, to satisfy the third condition for creating circular polarization.

The device shown in Fig. 15 emits light that contains a circularly-polarized component. Each grating on the device facet contributes to a 1D collimated beam and the far-field pattern is cross-shaped; see Fig. 15(c). Figure 15(d) is the measured device power when a wire-grid polarizer was rotated in front of the detector, which was placed in the far-field corresponding to the center of the cross. Analysis shows that this part of the laser beam consists of a right circularly-polarized component and a linearly-polarized component. The former is about 1.5 times larger than the latter in amplitude. The direct emissions from the two apertures have no phase difference and they contribute to the linearly-polarized component.

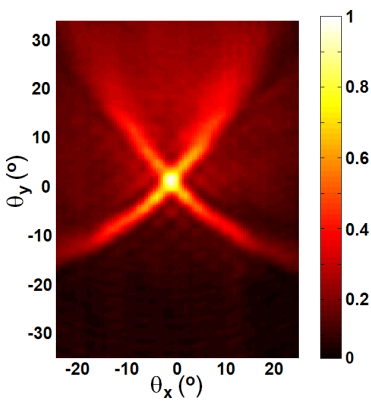
Several strategies can be employed to boost the circularly-polarized emission. By adding spoof SPP grooves adjacent to the laser aperture, one can couple more light into SPPs and suppress the direct emission into



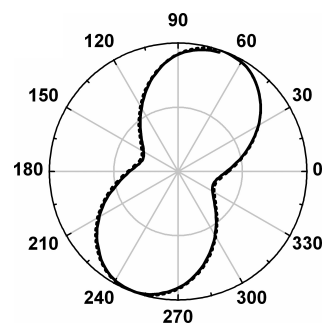
(a)



(b)



(c)



(d)

Fig. 15. (a) SEM images of a QCL capable of producing circularly-polarized emissions. The right panel is the zoom-in view of the left panel. (b) Calculated peak intensity of the beam created by one grating versus the distance between the aperture and the first grating groove d. Laser wavelength is assumed to be $\lambda_o = 9.9\,\mu$m. (c) Measured emission pattern of the device in (a). (d) Measured optical power of the central spot in (c) while a wire-grid polarizer was rotated in front of the detector. The solid curve is experimental data and the dotted curve is a fitting assuming coherent superposition of a circularly-polarized beam and a linearly-polarized beam.

the far-field; more detail on this will be discussed later in the section on collimation of THz QCLs. Furthermore, instead of using straight grating grooves, circular ones will give rise to 2D collimated beams from the left and right gratings, which will have better spatial overlap in the far-field.

7. Bean Shaping of THz QCLs Using Plasmonics

We demonstrate the power of THz plasmonics by addressing one of the key problems in THz science and technology. It will be shown in this section that collimators based on plasmonic second-order gratings and designer spoof SPP structures greatly improve the beam quality of THz QCLs.[85–87]

7.1. *Semiconductor plasmonic second-order grating collimator for THz QCLs*

We first adapted a design that is based on a plasmonic second-order grating; see Fig. 16(a). Like the ones used in mid-IR QCLs, the collimator utilizes the antenna array effect and works by transferring laser emission into SPPs on the device facet and coherently scattering their energy into the far-field. All the grating grooves are sculpted directly into the metallic GaAs facet ($\varepsilon_{GaAs} \approx -200 + 42i$ for $\lambda_o = 100\,\mu$m, GaAs n-doped to $1.6 \times 10^{18}\,\text{cm}^{-3}$, and $T = 80\,\text{K}$) using FIB milling. The limited facet area allows only five grating grooves to be defined [Fig. 16(a)]. To enhance the interaction between the grating and the SPPs, we used fairly large grating grooves (depth and width ∼15%–20% of λ_o) so that the energy of the SPPs can be scattered into the far-field as efficiently as possible. The effect of the second-order grating on the dispersion of SPPs is illustrated by the improved SPP confinement; compare Figs. 16(b) and 16(c).

The measured vertical far-field of the device has a FWHM divergence angle of ∼15°, which agrees reasonably well with simulation [Fig. 16(d)]. Thanks to the improved emission profile, the collected power of the device increases by a factor of ∼5 compared to the original unpatterned device at 80 K [Fig. 16(e)]. However, there is a fairly large background outside the central lobe, and this is due to the dominance of the direct emission from the laser aperture over the secondary emissions from the grating grooves. The latter are comparatively weak because the inefficient coupling into SPPs: simulation indicates that the power coupled into SPPs confined to the laser facet is merely 15%, while the remaining 85% is radiated directly into the far-field.

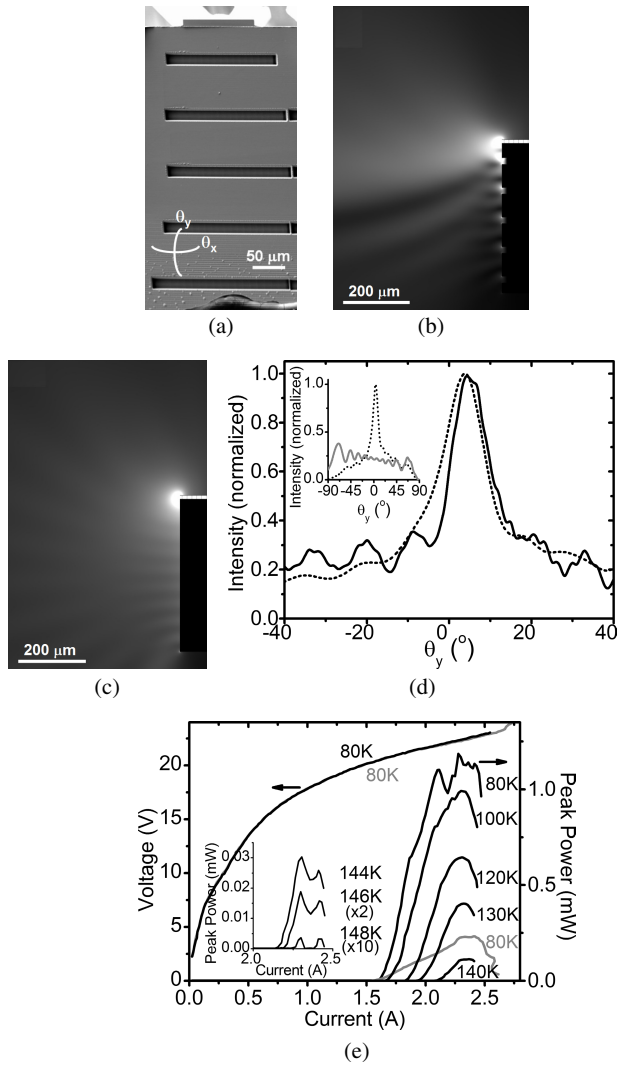


Fig. 16. Results for a $\lambda_o = 100\,\mu$m THz QCL with a second-order-grating collimator directly sculpted into the GaAs facet. (a) SEM image of the device facet. The laser waveguide is 125-μm wide, 1.2-mm long and sits on a 450-μm thick GaAs substrate (n-doped to $\sim 1.6 \times 10^{18}\,\mathrm{cm}^{-3}$). The grating has groove width and depth of $18\,\mu$m and $14\,\mu$m, respectively, and has periodicity of $88\,\mu$m. The center-to-center distance between the laser aperture and the nearest groove is $58\,\mu$m. (b) and (c) Simulated electric-field distribution ($|E|$) of the device with the second-order-grating collimator and of the original device, respectively. (d) The black solid curve and black dotted curve are measured and calculated vertical far-field of the device, respectively. Insets show calculated vertical far-field of the original (gray curve) and the collimated (black dotted curve) devices in the half space. (e) LIV characteristics and temperature performances of the device before (gray) and after (black) defining the collimator.

7.2. *Metasurface collimator for THz QCLs*

An improved design involves adding spoof SPP grooves adjacent to the laser aperture and between the second-order grating grooves, giving rise to a "metasurface" collimator. A schematic of such design for a ∼3.19 THz ($\lambda_o = 94\,\mu$m) laser and its cross-section are shown in Figs. 17(a) and 17(b), respectively. Dark and light colors are used in Fig. 17(b) to identify the second-order grating and the spoof SPP grooves, respectively.

Metasurfaces are an extension of the concept of metamaterials[88–91] to planar optics where the dispersion properties of SPPs are spatially tailored by structuring metallic surfaces so that SPPs can be controlled at will. Metasurfaces or metafilms have been under rapid progress in recent years and have found interesting applications such as subwavelength imaging,[92] waveguiding,[93,94] phase masks,[95] and the localization,[93,94] confinement,[96] and slowing of light.[97]

In the original unpatterned device, the laser emits both directly into the far-field and into Zenneck waves on the facet; both components have a wavevector close to k_o. The wavevector of the laser waveguide mode is several times larger, $n_{\text{eff}}k_o$ ($n_{\text{eff}} \approx 3.5$ is the effective mode index). A wavevector mismatch then exists at the aperture between the laser waveguide mode, the SPP mode on the facet, and free-space modes. In our collimator, the spoof SPP grooves adjacent to the aperture increase the in-plane wavevector of the SPPs, reducing the wavevector mismatch. More light is therefore coupled out from the laser cavity, and a larger percentage of it is channeled into SPPs on the facet instead of being directly emitted into the far-field. Most importantly, the spoof SPP grooves greatly increase the confinement of SPPs, thus improving the scattering efficiency of the second-order grating. Constructive interference between the scattered waves and the direct emission from the laser aperture gives rise to a low-divergence beam normal to the facet in the far-field.

Figure 17(c) shows that the in-plane wavevector increases by ∼25% for the spoof SPP grooves adjacent to the laser aperture (dash-dotted curve). The depth of the spoof SPP grooves was chosen to be in the range $6-11\,\mu$m to provide sufficient confinement without introducing large optical losses, which are mainly due to absorption in the plasmonic medium and rise sharply as the groove depth increases.[98] Figure 17(d) shows that the confinement of SPPs is improved to a few tens of micrometers, representing a reduction in decay length by approximately one order of magnitude compared to Zenneck waves on a planar interface (∼300 μm).

The simulated electric-field distribution of the device is presented in Fig. 18(a). Waves scattered from the laser facet by the collimator are

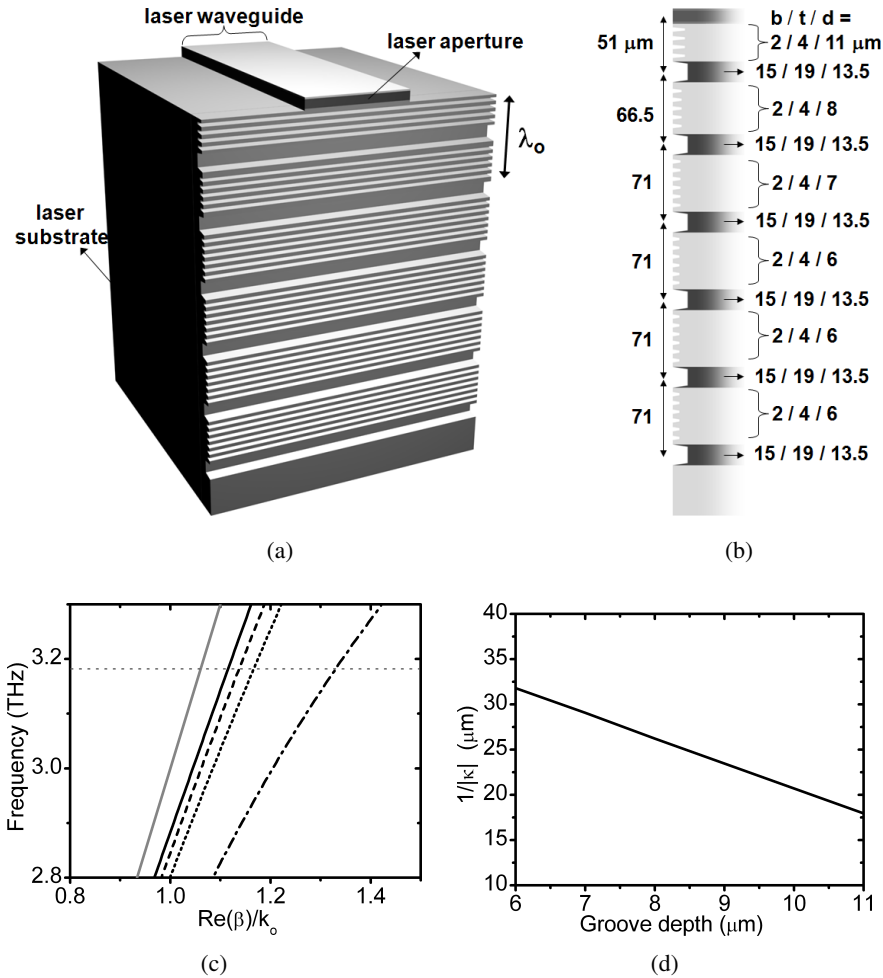


Fig. 17. (a) Schematic of a THz QCL integrated with a metasurface collimator. (b) Cross-section of the design for a $\lambda_o = 94\,\mu$m device. The width of the bottom and top of the grooves, and the groove depth are labeled as b, t, and d, respectively. The periodicity of the spoof SPP grooves is $8\,\mu$m. (c) The gray solid curve is the dispersion diagram of Zenneck waves on a planar interface between semiconductor and air, which is very close to the light line. The black curves are the dispersion diagrams corresponding to the different sections of the spoof SPP structures. Solid curve: $b/t/d = 2/4/6\,\mu$m; dashed curve: $b/t/d = 2/4/7\,\mu$m; dotted curve: $b/t/d = 2/6/8\,\mu$m; dash-dotted curve: $b/t/d = 2/4/11\,\mu$m. The horizontal dotted line indicates the lasing frequency. (d) 1/e decay length of the spoof SPP electric field ($|E|$) normal to the interface into the air as a function of d.

clearly observed in the near- and meso-field, and confined spoof SPPs can be seen on the facet. Simulations indicate that, compared to the original unpatterned device, the power throughput of the collimated device is increased by approximately 25%. In addition, it is estimated that about half of the laser output is coupled into SPP modes while the remaining half

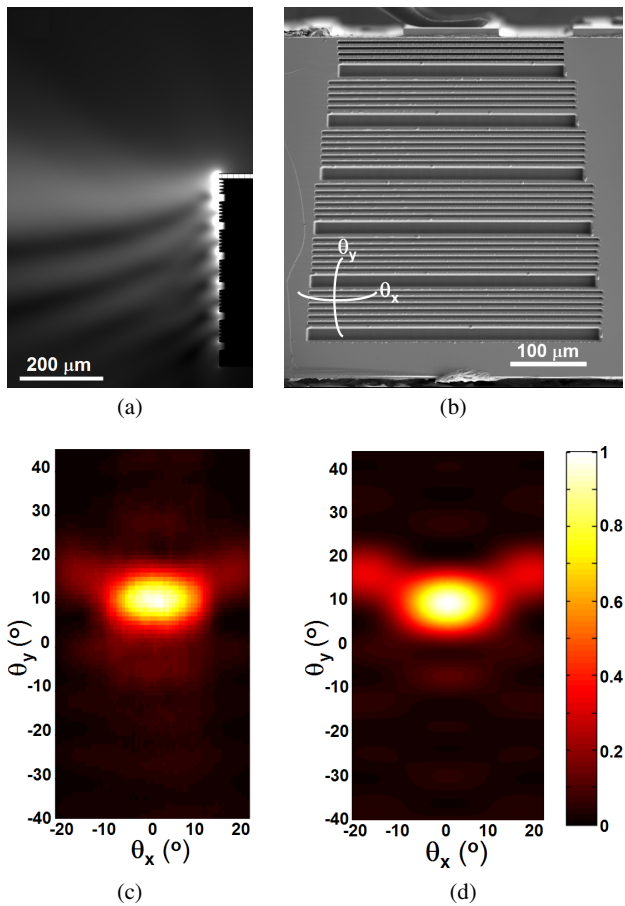


Fig. 18. Simulation and experimental results for a $\lambda_o = 94\,\mu$m THz QCL patterned with a metasurface collimator. (a) Simulated electric-field distribution (|E|) of the device. (b) SEM image of the device facet. (c) and (d) Measured and simulated emission patterns of the device. (e) Vertical line-scans of (c) (solid curve) and (d) (dotted curve) along $\theta_x = 0°$. (f) LIV characteristics of the device. The gray solid, black dotted, and black solid curves are for the unpatterned device, the device with only the second-order grating and the device with the metasurface collimator, respectively. Inset: spectrum of the collimated device measured at $I = 2.5$ A.

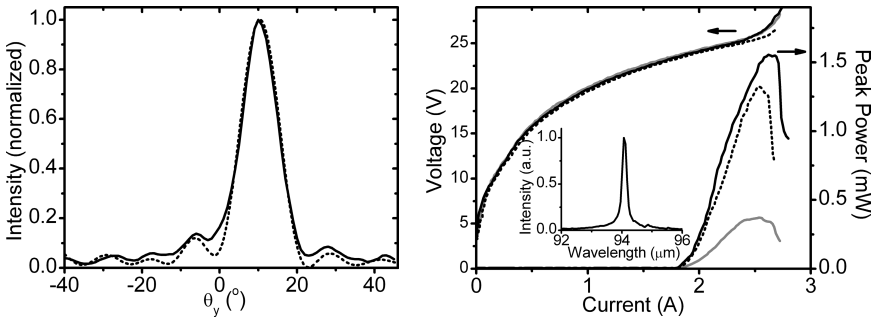


Fig. 18. (*Continued*)

is emitted directly into the far-field; this is a significant improvement over the design with only the second order grating (Fig. 16).

Figure 18(b) shows an SEM image of the facet of a fabricated device. The collimator occupies a small footprint with dimensions $\sim 4\lambda_o \times 4.5\lambda_o$. The central beam has vertical and lateral FWHM divergence angles of $\sim 11°$ and $\sim 19°$, respectively, in good agreement with the 3D full-wave simulation results; compare Figs. 18(c) and 18(d). The optical background has average intensity below 10% of the central lobe peak intensity [Fig. 18(e)], owing to a more balanced direct emission from the aperture and secondary emissions from the facet. The beam quality of our collimated device is better than or comparable to those obtained in THz QCLs with higher-order gratings, photonic crystals, or mounted micro optical elements.[36,37,45–49,52,54] The small divergence beam emitted from our device is compatible with the receiver front end of modern submillimeter heterodyne detection systems.[99,100]

The collected power of the collimated device increases by a factor of ~ 4.5 compared to the original unpatterned device; see Fig. 18(f). The increase in power collection efficiency is ~ 3.7 for the device with only the second-order grating. This difference in measured power is primarily a result of the increased total power throughput originating from the reduced wavevector mismatch. Based on the far-field measurements, the power in the main lobe of the collimated device is about two times larger than that of the device with only the second-order grating. The maximum operating temperature of the patterned device is 140 K, the same as that of the original device. Figure 18(f) shows that the lasing threshold did not change after defining the collimator.

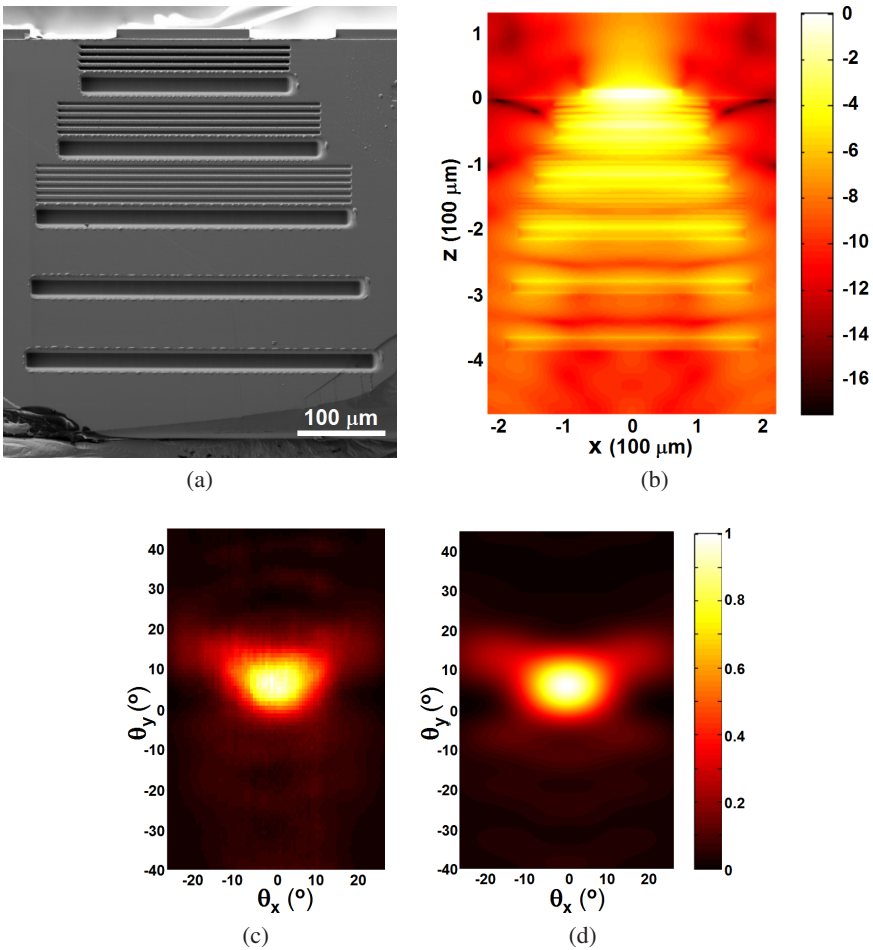
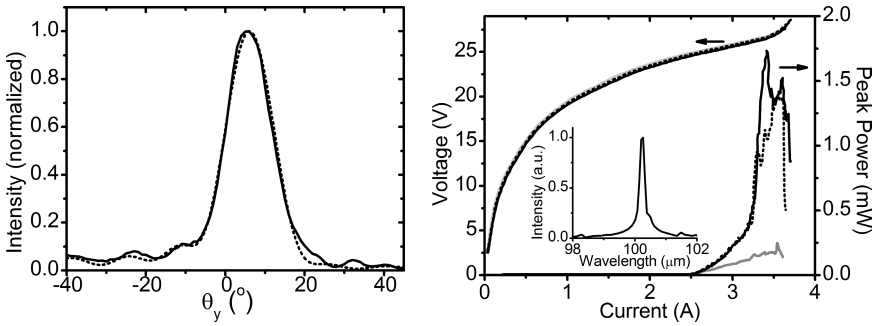
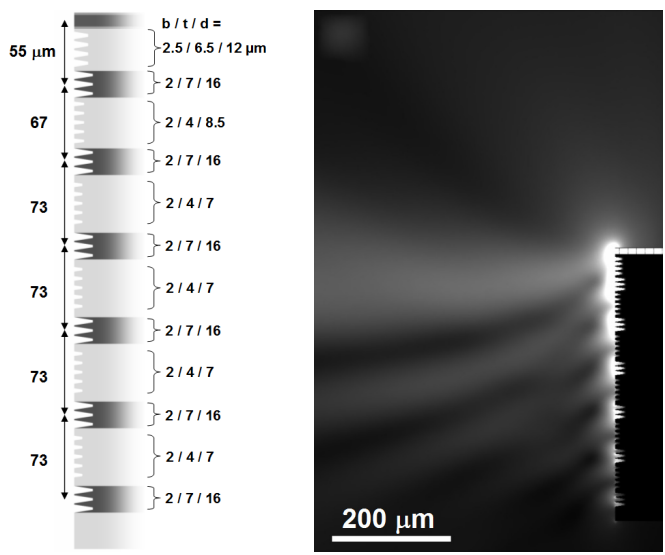


Fig. 19. Simulation and experimental results for a $\lambda_o \sim 100\,\mu$m THz QCL patterned with a hybrid collimator. (a) SEM image of the device facet. (b) Simulated near-field distribution ($\log_e(|E|/\max(|E|))$) of the device. The near-field is monitored on a plane $1\,\mu$m above the device facet. (c) and (d) Measured and simulated emission patterns of the device. (e) Vertical line-scans of (c) (solid curve) and (d) (dotted curve) along $\theta_x = 0°$. (f) LIV characteristics of the device. The gray solid, black dotted, and black solid curves are for the unpatterned device, the device with only the second-order grating and the device with the complete collimator, respectively. Inset: spectrum of the collimated device measured at $I = 3.4\,$A.



(e) (f)

Fig. 19. (*Continued*)

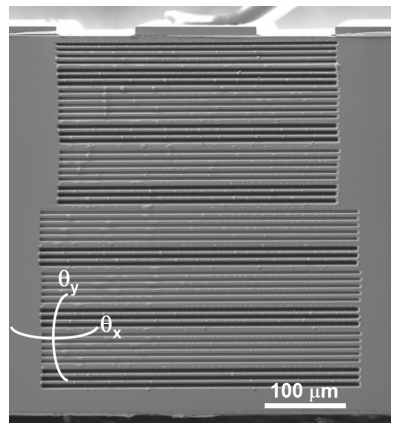


(a) (b)

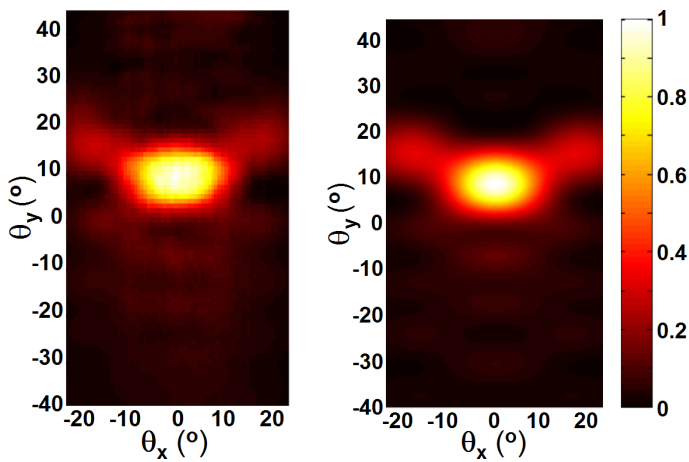
Fig. 20. (a) Cross-section of a metasurface collimator design consisting entirely of spoof SPP grooves. All spoof SPP grooves have a periodicity of $8\,\mu$m. $\lambda_o = 100\,\mu$m. (b) Simulated electric-field distribution (|E|) of the device.

The device presented in Fig. 19 has spoof SPP grooves patterned in the vicinity of the laser aperture: adjacent to the laser aperture and in the first two periods of the second-order grating, giving rise to a "hybrid" collimator consisting of both metasurface and conventional plasmonic structures. The device exhibits performance comparable to the previous one in terms of

vertical beam divergence and power throughput. The device starts to lase under a higher-order transverse mode but switches to the fundamental mode when the pump current is above ∼3.25 A, leading to a kink in the LI curve [Fig. 19(f)].

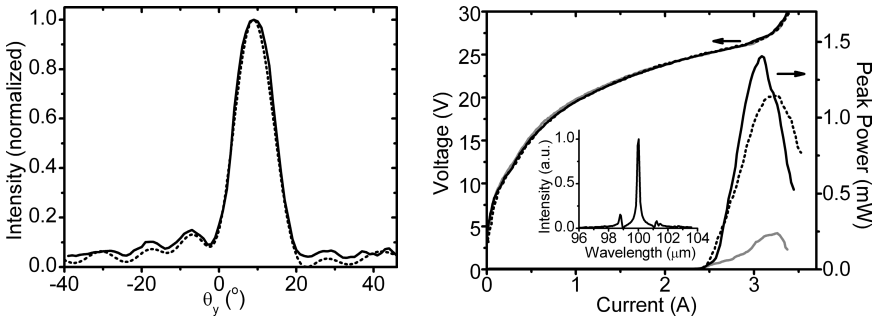


(a)



(b) (c)

Fig. 21. Experimental results for a $\lambda_o = 100\,\mu$m QCL fabricated according to the design in Fig. 20. (a) SEM image of the device facet. The device has a 1.2-mm-long, 150-μm-wide, and 10-μm-thick waveguide. (b) and (c) Measured and simulated emission profiles of the device. (d) Line-scans of (b) (solid curve) and (c) (dotted curve) along $\theta_x = 0°$. (f) LIV characteristics of the device. The gray solid, black dotted, and black solid curves are for the unpatterned device, the device with only the deep spoof SPP grooves and the device with the complete structure, respectively.



(d) (e)

Fig. 21. (*Continued*)

We can go one step further to create a metasurface collimator consisting entirely of spoof SPP structures; see Fig. 20(a). Spoof SPP grooves with different depths periodically modulate the dispersion of the SPPs on the device facet, creating an effective second-order grating. The shallow grooves [light in Fig. 20(a)] improve SPP confinement. The grooves adjacent to the laser aperture help increase device power throughput by coupling more laser output into spoof SPPs on the facet. The fabricated device shows very similar performance as the one in Fig. 18. The central beam has vertical and lateral FWHM divergence angles of $\sim 12^\circ$ and $\sim 16^\circ$, respectively; the optical background is below 10% of the peak intensity. The measured beam directivity is increased from ~ 5 dB for the original unpatterned laser, to ~ 16 dB for the device with the collimator. The power collection efficiency is enhanced by a factor of ~ 6; see Fig. 21.

7.3. *Reduction of the lateral beam divergence*

Reduction of beam divergence occurs in both the vertical and lateral directions for the three devices presented above. In the vertical direction, the collimation is explained by the antenna array effect. In the lateral direction, the collimation is due to an increased effective emission area: the spoof SPP patterns were intentionally fabricated to be a few times wider than the laser waveguide, which helps spread SPPs laterally. Physically, laser emission first excites the central section of the long spoof SPP grooves; the channel polaritons generated then propagate along the grooves in a way analogous to TEM waves traveling along coplanar strip transmission

lines. Simulations confirm that the lateral divergence is roughly inversely proportional to the width of the plasmonic pattern; see Fig. 22.

The advantage of the channel polariton scheme compared to the ring-shaped 2D collimators discussed in Section 4 is that there is no need to use small apertures for 2D spreading of SPPs. 2D collimation with high power

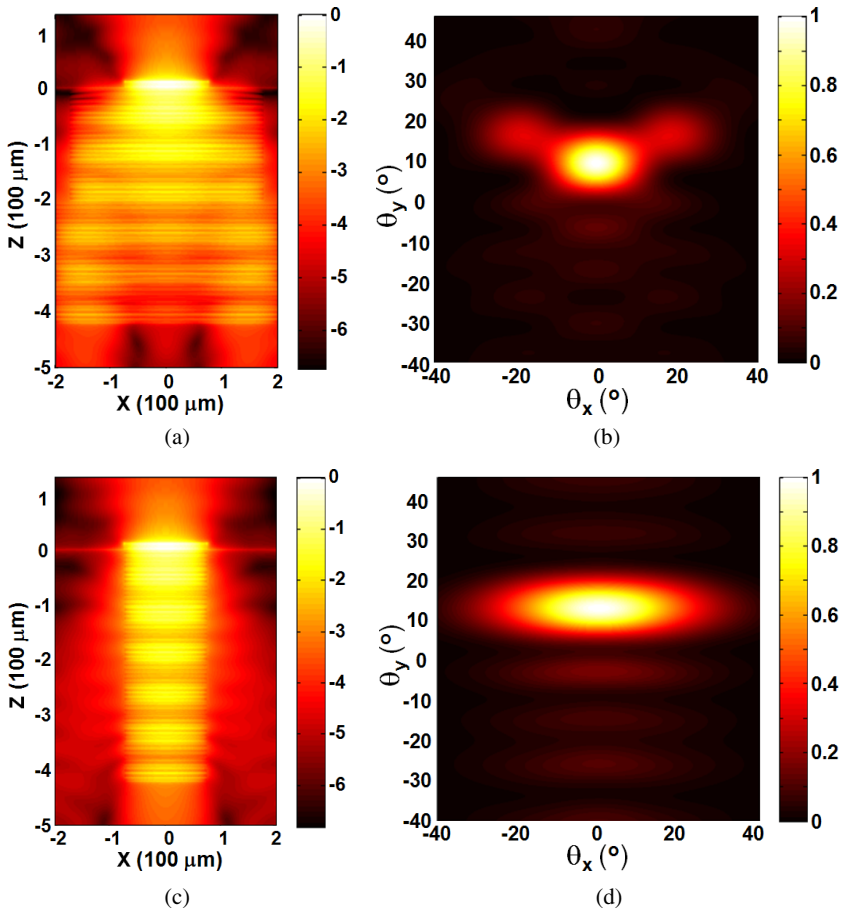
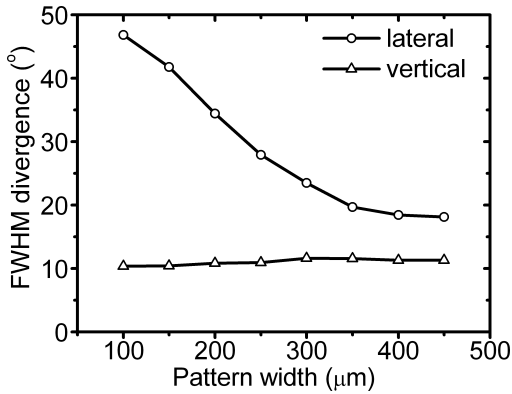


Fig. 22. Reducing lateral far-field divergence. (a) Simulated near-field distribution ($\log_e(|E|/\max(|E|))$) for the device in Fig. 21. The near-field is monitored on a plane 1 μm above the device facet. (b) Simulated far-field intensity corresponding to the near-field distribution in (a). (c) Simulated near-field distribution of another device whose collimator has narrower lateral dimensions. The collimator width is ~400 μm in (a) and 150 μm in (c); other design parameters are the same. (d) Simulated far-field intensity corresponding to the near-field distribution in (c). (e) Lateral and vertical far-field divergence versus width of the collimator.



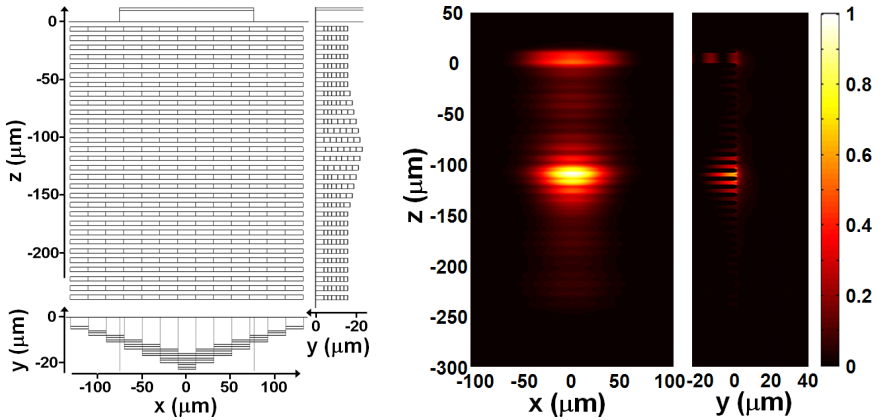
(e)

Fig. 22. (*Continued*)

throughput can be achieved by using a 1D design comprising of straight grooves and widely open apertures.

7.4. *Metasurface lens*

More sophisticated designer plasmonic structures can be created by engineering the effective mode index or impedance of SPPs. Figure 23 shows a graded-index near-field plasmonic lens that is able to focus the output of a THz QCL into a subwavelength spot on the facet. The lens comprises deep rectangular indentations surrounded by successively shallower ones [Fig. 23(a)], all sculpted on the laser facet that is assumed to be a perfect electric conductor. The indentation at the center of the pattern has the largest depth $d \sim \lambda_o/4$, satisfying the cavity resonance condition discussed in Section 2. In the vertical direction (z-direction in Fig. 23), the group velocity of SPPs gradually decreases towards the central indentation; consequently the electromagnetic energy emitted at the laser aperture is funneled into the close vicinity of the central indentation where the group velocity is close to zero. The lateral profile of the indentations is shown in the lower panel of Fig. 23(a); as channel polaritons propagate sideways, they experience distributed reflection due to mode impedance change, as such the lateral propagation of channel polaritons is discouraged. The plasmonic lens is based on periodically spaced indentations; in principle, using indentations of the same depth and tuning their periodicities and/or widths can achieve the same effect.



(a) (b)

Fig. 23. (a) Design of a plasmonic lens based on spoof SPP structures. xz-plane is parallel to the laser facet and y-axis is perpendicular to the laser facet. (b) Simulated near-field distributions ($|E|$) of the plasmonic lens defined on the facet of a $\lambda_o = 100\,\mu$m QCL. The left and right panels are the top view and the cross-section view, respectively. The top view is monitored on a plane $1\,\mu$m above the device facet and the cross-section view is through the symmetric plane of the structure.

8. Conclusions and Future Perspectives

We have demonstrated that, by monolithically integrating on the emission facet of mid-IR and THz QCLs, properly designed plasmonic structures that tailor the dispersion properties of SPPs, the device performance can be significantly enhanced and new functionalities can be realized.

Simulation and experimental results in Figs. 19 and 23 have not been published and other part of the chapter is a review of previously published materials.

Future interesting direction in wavefront engineering with plasmonics includes creating light sources capable of producing special beams such as Bessel beams,[101] which are diffraction free, beams that are radially or azimuthally polarized,[102] and beams with orbital angular momentum,[103,104] which can be used to rotate small particles.

In this paper we discussed only far-field engineering of infrared light sources, but the possibilities that plasmonics can offer in the near-field are also numerous starting from optical antennas.[105,106] Plasmonic antennas and periodic/aperiodic antenna arrays can be integrated on optical components such as semiconductor lasers and optical fibers. By properly engineering the extinction spectra and the scattering cross sections of the

antenna structures, one can create near-field optical devices with applications ranging from high-resolution probes for imaging/spectroscopy,[107,108] to broad band lenses for solar cells,[109,110] and novel modulators/detectors/sensors.[111]

We further envision using soft-lithography based techniques for inexpensive, high-throughput fabrication of plasmonic structures so that we are not restricted to FIB milling, which is mostly suitable for proof-of-concept demonstrations. There have been encouraging progresses in this direction in the past few years. For example, nanoskiving[112] in combination with replica molding[113] has been developed to produce planar arrays of single- or multi-component (e.g., metals, semiconductors, insulators) structures with subwavelength features and high aspect ratio.[114] A soft-lithography based pattern transfer technique, "decal transfer",[115] has been demonstrated for convenient and reliable fabrication of plasmonic structures on non-conventional surfaces such as the facet of optical devices or curved surfaces.

New materials systems are worth exploring in the mid-IR and THz frequency ranges. For example, semiconductors are suitable plasmonic media for THz light. Due to smaller carrier densities, the plasma frequency is much lower in semiconductors compared to that of metals. As a result, the confinement of SPPs to planar interfaces is largely improved, from typically several millimeters for metals to a few hundred micrometers for semiconductors. In addition, doped semiconductors offer great versatility because their optical properties (e.g., permittivity, anisotropy) can be tuned by applying optical excitation, electrical potentials,[116,117] or magnetic fields,[1,60,61,118] or by controlling the temperature.[58,61] Of special interest are narrow-bandgap semiconductors such as InSb, which possesses a few interesting properties that can be employed to effectively tailor the SPP dispersion in the THz regime. First, its optical phonon frequency falls into the THz frequency range;[119] second, thanks to its small effective electron mass, the cyclotron frequency can be brought into the THz range by using moderate magnetic fields ($\sim$1 Tesla);[61] third, its carrier concentration can be tuned over a wide range by adjusting temperature due to its low thermal excitation energy.[58]

Conductive oxides such as indium tin oxide (ITO), and aluminum zinc oxide (AZO), have recently been recognized as promising plasmonic materials for the near-infrared wavelengths and beyond[120] owing to their low optical losses: the imaginary part of the dielectric permittivity of these conductive oxides is smaller compared to silver at wavelengths longer than about one micrometer.[121] Furthermore, these oxide semiconductors are

non-stoichiometric: their optical properties can be largely tuned by using different growth methods and annealing conditions, or by choosing different doping levels in the oxide (e.g., Sn in In_2O_3)[120,121]

Acknowledgment

We gratefully acknowledge the following people who contributed to the work presented in this paper: R. Blanchard, M. A. Kats, J. A. Fan, Q. J. Wang, C. Pflügl, L. Diehl, M. A. Belkin, S. P. Khanna, L. Li, A. G. Davies, E. H. Linfield, T. Edamura, S. Furuta, M. Yamanishi, H. Kan, M. Geiser, M. Fischer, A. Wittmann, J. Faist, and N. Antoniou. We acknowledge funding from AFOSR under contract No. FA9550-09-0505-DOD, the EPSRC (UK), and Hamamatsu Photonics (Japan). The authors acknowledge support from the National Science Foundation, Harvard Nanoscale Science and Engineering Center (NSEC), and the Center for Nanoscale Systems (CNS) at Harvard University. Harvard CNS is a member of the National Nanotechnology Infrastructure Network (NNIN). The computations in this Letter were run on the Odyssey cluster supported by the Harvard Faculty of Arts and Sciences (FAS) Sciences Division Research Computing Group.

References

1. A. D. Boardman, *Electromagnetic Surface Modes* (Wiley, New York, 1982).
2. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer, Berlin, 1988).
3. S. A. Maier, *Plasmonics: Fundamentals and Applications* (Springer-Verlag, New York, 2007).
4. W. L. Barnes, A. Dereux, and T. W. Ebbesen, *Nature* **424**, 824 (2003).
5. E. Ozbay, *Science* **311**, 189 (2006).
6. H. A. Atwater, *Sci. Am.* **296**, 56 (2007).
7. J. Zenneck, *Ann. Physik* **23**, 846 (1907).
8. A. Sommerfeld, *Ann. Physik* **28**, 665 (1909).
9. J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Companies, 1941).
10. R. F. Kazarinov and R. A. Suris, *Sov. Phys. Semicond.* **5**, 707 (1971).
11. J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, *Science* **264**, 553 (1994).
12. C. Gmachl, F. Capasso, D. L. Sivco, and A. Y. Cho, *Rep. Prog. Phys.* **64**, 1533 (2001).
13. F. Capasso, C. Gmachl, R. Paiella, A. Tredicucci, A. L. Hutchinson, D. L. Sivco, J. N. Baillargeon, A. Y. Cho, and H. C. Liu, *IEEE J. Sel. Top. Quantum Electron.* **6**, 931 (2000).
14. Y. Bai, S. Slivken, S. R. Darvish, A. Haddadi, B. Gokden, and M. Razeghi, *Appl. Phys. Lett.* **95**, 221104 (2009).

15. M. Razeghi, *Proc. SPIE* **7230**, 723011 (2009).
16. A. Lyakh, R. Maulini, A. Tsekoun, R. Go, C. Pflügl, L. Diehl, Q. J. Wang, F. Capasso, and C. K. N. Patel, *Appl. Phys. Lett.* **95**, 141113 (2009).
17. J. Faist, C. Gmachl, F. Capasso, C. Sirtori, D. L. Sivco, J. N. Baillargeon, A. L. Hutchinson, and A. Y. Cho, *Appl. Phys. Lett.* **70**, 2670 (1997).
18. C. Gmachl, J. Faist, J. N. Baillargeon, F. Capasso, C. Sirtori, D. L. Sivco, S. N. G. Chu, and A. Y. Cho, *IEEE Photon. Technol. Lett.* **9**, 1090 (1997).
19. G. Wysocki, R. F. Curl, F. K. Tittel, R. Maulini, J. M. Bulliard, and J. Faist, *Appl. Phys. B* **81**, 769 (2005).
20. R. Maulini, A. Mohan, M. Giovannini, J. Faist, and E. Gini, *Appl. Phys. Lett.* **88**, 201113 (2006).
21. M. B. Pushkarsky, I. G. Dunayevskiy, M. Prasanna, A. G. Tsekoun, R. Go, and C. K. N. Patel, *Proc. Natl. Acad. Sci. USA* **103**, 19630 (2006).
22. G. Wysocki, R. Lewicki, R. F. Curl, F. K. Tittel, L. Diehl, F. Capasso, M. Troccoli, G. Hofler, D. Bour, S. Corzine, R. Maulini, M. Giovannini, and J. Faist, *Appl. Phys. B* **92**, 305 (2008).
23. A. Wittmann, A. Hugi, E. Gini, N. Hoyler, and J. Faist, *IEEE J. Quantum Electron.* **44**, 1083 (2008).
24. A. Hugi, R. Terazzi, Y. Bonetti, A. Wittmann, M. Fischer, M. Beck, J. Faist, and E. Gini, *Appl. Phys. Lett.* **95**, 061103 (2009).
25. R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi, *Nature* **417**, 156 (2002).
26. B. S. Williams, *Nature Photon.* **1**, 517 (2007).
27. G. Scalari, C. Walther, M. Fischer, R. Terazzi, H. Beere, D. Ritchie, and J. Faist, *Laser Photon. Rev.* **1**, 1 (2008).
28. M. A. Belkin, Q. J. Wang, C. Pflügl, A. Belyanin, S. P. Khanna, A. G. Davies, E. H. Linfield, and F. Capasso, *IEEE J. Sel. Topics Quantum Electron.* **15**, 952 (2009).
29. A. W. M. Lee, Q. Qin, S. Kumar, B. S. Williams, Qing Hu, and J. L. Reno, *Appl. Phys. Lett.* **89**, 141125 (2006).
30. J. Faist, G. Scalari, C. Walther, and M. Fischer, Materials Research Society (MRS) Spring Meeting, San Fransisco, California, CC7.2 (2007).
31. S. Kumar, Q. Hu, and J. L. Reno, *Appl. Phys. Lett.* **94**, 131105 (2009).
32. B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno, *Opt. Express* **13**, 3331 (2005).
33. B. S. Williams, S. Kumar, Q. Hu, and J. L. Reno, *Electron. Lett.* **42**, 89 (2006).
34. S. Kohen, B. S. Williams, and Q. Hu, *J. Appl. Phys.* **97**, 053106 (2005).
35. M. Hajenius, P. Khosropanah, J. N. Novenier, J. R. Gao, T. M. Klapwijk, S. Barbieri, S. Dhillon, P. Filloux, C. Sirtori, D. A. Ritchie, and H. E. Beere, *Opt. Lett.* **33**, 312 (2008).
36. A. W. M. Lee, Q. Qin, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, *Opt. Lett.* **32**, 2840 (2007).
37. M. I. Amanti, M. Fischer, C. Walther, G. Scalari, and J. Faist, *Electron. Lett.* **43**, 573 (2007).

38. M. Troccoli, C. Gmachl, F. Capasso, D. L. Sivco, and A. Y. Cho, *Appl. Phys. Lett.* **80**, 4103 (2002).
39. L. Nähle, J. Semmel, W. Kaiser, S. Höfling, and A. Forchel, *Appl. Phys. Lett.* **91**, 181122 (2007).
40. Q.-Y. Lu, W.-H. Guo, W. Zhang, L.-J. Wang, J.-Q. Liu, L.-Li, F.-Q. Liu, and Z.-G. Wang, *Appl. Phys. Lett.* **96**, 051112 (2010).
41. D. Hofstetter, J. Faist, M. Beck, and U. Oesterle, *Appl. Phys. Lett.* **75**, 3769 (1999).
42. W. Schrenk, N. Finger, S. Gianordoli, L. Hvozdara, G. Strasser, and E. Gornik, *Appl. Phys. Lett.* **77**, 2086 (2000).
43. D. Hofstetter, J. Faist, M. Beck, A. Müller, and U. Oesterle, *Physica E* **7**, 25 (2000).
44. C. Pflügl, M. Austerer, W. Schrenk, S. Golka, G. Strasser, R. P. Green, L. R. Wilson, J. W. Cockburn, A. B. Krysa, and J. S. Roberts, *Appl. Phys. Lett.* **86**, 211101 (2005).
45. J. A. Fan, M. A. Belkin, and F. Capasso, *Opt. Express* **14**, 11672 (2006).
46. S. Kumar, B. S. Williams, Q. Qin, A. W. M. Lee, and Q. Hu, *Opt. Express* **15**, 114 (2007).
47. M. I. Amanti, M. Fischer, G. Scalari, M. Beck, and J. Faist, *Nature Photon.* **3**, 586 (2009).
48. E. Mujagić, L. K. Hoffmann, S. Schartner, M. Nobile, W. Schrenk, M. P. Semtsiv, M. Wienold, W. T. Masselink, and G. Strasser, *Appl. Phys. Lett.* **93**, 161101 (2008).
49. L. Mahler, A. Tredicucci, F. Beltram, C. Walther, J. Faist, B. Witzigmann, H. E. Beere, and D. A. Ritchie, *Nature Photon.* **3**, 46 (2009).
50. E. Mujagić, M. Nobile, H. Detz, W. Schrenk, J. Chen, C. Gmachl, and G. Strasser, *Appl. Phys. Lett.* **96**, 031111 (2010)
51. R. Colombelli, K. Srinivasan, M. Troccoli, O. Painter, C. F. Gmachl, D. M. Tennant, A. M. Sergent, D. L. Sivco, A. Y. Cho, and F. Capasso, *Science* **302**, 1374 (2003).
52. Y. Chassagneux, R. Colombelli, W. Maineult, S. Barbieri, H. E. Beere, D. A. Ritchie, S. P. Khanna, E. H. Linfield, and A. G. Davies, *Nature* **457**, 174 (2009).
53. G. Xu, Y. Chassagneux, R. Colombelli, G. Beaudoin, and I. Sagnes, *Opt. Lett.* **35**, 859 (2010).
54. Y. Chassagneux, R. Colombelli, W. Maineult, S. Barbieri, S. P. Khanna, E. H. Linfield, and A. G. Davies, *Appl. Phys. Lett.* **96**, 031104 (2010).
55. S. C. Kitson, W. L. Barnes, and J. R. Sambles, *Phys. Rev. Lett.* **77**, 2670 (1996).
56. J.-C. Weeber, Y. Lacroute, A. Dereux, E. Devaux, T. Ebbesen, C. Girard, M. U. González, and A.-L. Baudrion, *Phys. Rev. B* **70**, 235406 (2004).
57. J. Gómez Rivas, M. Kuttge, P. Haring Bolivar, H. Kurz, and J. A. Sánchez-Gil, *Phys. Rev. Lett.* **93**, 256804 (2004).
58. J. Gómez Rivas, M. Kuttge, H. Kurz, P. Haring Bolivar, and J. A. Sánchez-Gil, *Appl. Phys. Lett.* **88**, 082106 (2006).
59. J. Allen, D. C. Tsui, and R. A. Logan, *Phys. Rev. Lett.* **38**, 980 (1977).

60. M. S. Kushwaha, *Surf. Sci. Rep.* **41**, 1 (2001).
61. X. Wang, A. A. Belyanin, S. A. Crooker, D. M. Mittleman, and J. Kono, *Nature Phys.* **6**, 126 (2010).
62. J. B. Pendry, L. Martín-Moreno, and F. J. García-Vidal, *Science* **305**, 847 (2004).
63. F. J. García-Vidal, L. Martín-Moreno, and J. B. Pendry, *J. Opt. A: Pure Appl. Opt.* **7**, S97 (2005).
64. C. R. Williams, S. R. Andrews, S. A. Maier, A. I. Fernández-Domínguez, L. Martín-Moreno, and F. J. García-Vidal, *Nature Photon.* **2**, 175 (2008).
65. G. Goubau, *J. Appl. Phys.* **21**, 1119 (1950).
66. W. Rotman, *Proc. IRE* **39**, 952 (1951).
67. J. A. Kong, *Electromagnetic Wave Theory* (EMW Publishing, Cambridge, 2000).
68. I. V. Novikov, and A. A. Maradudin, *Phys. Rev. B* **66**, 035403 (2002).
69. D. K. Gramotnev and D. F. P. Pile, *Appl. Phys. Lett.* **85**, 6323 (2004).
70. S. I. Bozhevolnyi, V. S. Volkov, E. Devaux, J.-Y. Laluet, and T. W. Ebbesen, *Nature* **440**, 508 (2006).
71. S. I. Bozhevolnyi, V. S. Volkov, E. Devaux, and T. W. Ebbesen, *Phys. Rev. Lett.* **95**, 046802 (2005).
72. S. I. Bozhevolnyi, *Opt. Express* **14**, 9467 (2006).
73. S. I. Bozhevolnyi and J. Jung, *Opt. Express* **16**, 2676 (2008).
74. I. P. Kaminow, W. L. Mammel, and H. P. Weber, *Appl. Opt.* **13**, 396 (1974).
75. N. Yu, J. Fan, Q. J. Wang, C. Pflügl, L. Diehl, T. Edamura, M. Yamanishi, H. Kan, and F. Capasso, *Nature Photon.* **2**, 564 (2008).
76. H. J. Lezec, A. Degiron, E. Devaux, R. A. Linke, L. Martín-Moreno, F. J. García-Vidal, and T. W. Ebbesen, *Science* **297**, 820 (2002).
77. L. Martín-Moreno, F. J. García-Vidal, H. J. Lezec, A. Degiron, and T. W. Ebbesen, *Phys. Rev. Lett.* **90**, 167401 (2003).
78. N. Yu, R. Blanchard, J. Fan, T. Edamura, M. Yamanishi, H. Kan, and F. Capasso, *Appl. Phys. Lett.* **93**, 181101 (2008).
79. N. Yu, R. Blanchard, J. Fan, Q. J. Wang, C. Pflügl, L. Diehl, T. Edamura, M. Yamanishi, H. Kan, and F. Capasso, *Opt. Express* **16**, 19447 (2008).
80. N. Yu, M. A. Kats, C. Pflügl, M. Geiser, Q. J. Wang, M. A. Belkin, F. Capasso, M. Fischer, A. Wittmann, J. Faist, T. Edamura, S. Furuta, M. Yamanishi, and H. Kan, *Appl. Phys. Lett.* **95**, 161108 (2009).
81. A. Yariv and P. Yeh, *Photonics: Optical Electronics in Modern Communications (6th ed.)* (Oxford University Press, Oxford, 2007).
82. H. C. Liu and F. Capasso, *Intersubband Transitions in Quantum Wells: Physics and Device Applications I* (Academic Press, San Diego, 2000).
83. N. Yu, Q. J. Wang, C. Pflügl, L. Diehl, F. Capasso, T. Edamura, S. Furuta, M. Yamanishi, and H. Kan, *Appl. Phys. Lett.* **94**, 151101 (2009).
84. F. López-Tejeira, Sergio G. Rodrigo, L. Martín-Moreno, F. J. García-Vidal, E. Devaux, T. W. Ebbesen, J. R. Krenn, P. Radko, S. I. Bozhevolnyi, M. U. Gonzalez, J. C. Weeber, and A. Dereux, *Nature Phys.* **3**, 324 (2007).
85. N. Yu, Q. J. Wang, M. A. Kats, J. A. Fan, S. P. Khanna, L. Li, A. G. Davies, E. H. Linfield, and F. Capasso, *Nature Mater.* **9**, 730 (2010).

86. N. Yu, Q. J. Wang, M. A. Kats, J. A. Fan, S. P. Khanna, L. Li, A. G. Davies, E. H. Linfield, and F. Capasso, *Electron. Lett.* (in press).
87. N. Yu and F. Capasso, *J. Opt. Soc. Am. B* (in press).
88. N. Engheta and R. W. Ziolkowski, *Metamaterials: Physics and Engineering Explorations* (Wiley-IEEE Press, 2006).
89. W. Cai and V. M. Shalaev, *Optical Metamaterials, Fundamentals and Applications* (Springer, New York, 2009).
90. J. B. Pendry, D. Schurig, and D. R. Smith, *Science*, **312**, 1780 (2006).
91. U. Leonhardt, *Science*, **312**, 1777 (2006).
92. I. I. Smolyaninov, Y.-J. Hung, and C. C. Davis, *Phys. Rev. B*, **76**, 205424 (2007).
93. J. Beermann, I. P. Radko, A. Boltasseva, and S. I. Bozhevolnyi, *Opt. Express*, **15**, 15234 (2007).
94. I. P. Radko, V. S. Volkov, J. Beermann, A. B. Evlyukhin, T. Søndergaard, A. Boltasseva, and S. I. Bozhevolnyi, *Laser Photon. Rev.* **3**, 575 (2009).
95. B. Kanté, J.-M. Lourtioz, and A. de Lustrac, *Phys. Rev. B*, **80**, 205120 (2009).
96. M. Navarro-Cía, M. Beruete, S. Agrafiotis, F. Falcone, M. Sorolla, and S. A. Maier, *Opt. Express*, **17**, 18184 (2009).
97. Q. Gan, Z. Fu, Y. J. Ding, and F. J. Bartoli, *Phys. Rev. Lett.* **100**, 256803 (2008).
98. B. Wang, L. Liu, and S. He, *J. Appl. Phys.* **104**, 103531 (2008).
99. M. C. Gaidis, H. M. Pickett, C. D. Smith, S. C. Martin, R. P. Smith, and P. H. Siegel, *IEEE Trans. Microwave Theory. Tech.* **48**, 733 (2000).
100. P. H. Siegel and R. J. Dengler, *IEEE Trans. Antennas Propag.* **39**, 40 (1991).
101. J. Durnin, J. J. Miceli, Jr., and J. H. Eberly, *Phys. Rev. Lett.* **58**, 1499 (1987).
102. Z. Bomzon, V. Kleiner, and E. Hasman, *Appl. Phys. Lett.* **79**, 1587 (2001).
103. R. Beth, *Phys. Rev.* **50**, 115 (1936).
104. M. Padgett, J. Courtial, and L. Allen, *Phys. Today* May iss., 35 (2004).
105. E. Cubukcu, N. Yu, E. J. Smythe, L. Diehl, K. B. Crozier, and F. Capasso, *IEEE J. Sel. Top. Quantum Electron.* **14**, 1448 (2008).
106. N. Yu, E. Cubukcu, L. Diehl, D. Bour, S. Corzine, J. Zhu, G. Höfler, K. B. Crozier, and F. Capasso, *Opt. Express* **15**, 13272 (2007).
107. J. N. Farahani, H.-J. Eisler, D. W. Pohl, M. Pavius, P. Flückiger, P. Gasser, and B. Hecht, *Nanotechnology* **18**, 125506 (2007).
108. S. Palomba and L. Novotny, *Nano Lett.* **9**, 3801 (2009).
109. Y. A. Akimov, W. S. Koh, and K. Ostrikov, *Opt. Express* **17**, 10195 (2009).
110. R. A. Pala, J. White, E. Barnard, J. Liu, and M. L. Brongersma, *Adv. Mater.* **21**, 3504 (2009).
111. E. J. Smythe, M. D. Dickey, J. Bao, G. M. Whitesides, and F. Capasso, *Nano Lett.* **9**, 1132 (2009).
112. Q. Xu, R. M. Rioux, M. D. Dickey, and G. M Whitesies, *Nano Lett.* **41**, 1566 (2008).
113. Y. Xia and G. M. Whitesides, *Angew. Chem. Int. Ed.* **37**, 550 (1998).

114. D. J. Lipomi, M. A. Kats, P. Kim, S. H. Kang, J. Aizenberg, F. Capasso, and G. M. Whitesides, *ACS Nano* ASAP (2010).
115. E. J. Smythe, M. D. Dickey, G. M. Whitesides, and F. Capasso, *ACS Nano* **3**, 59 (2009).
116. H.-T. Chen, W. J. Padilla, J. M. O. Zide, A. C. Gossard, A. J. Taylor, and R. D. Averitt, *Nature*, **444**, 597 (2006).
117. J. A. Dionne, K. Diest, L. A. Sweatlock, and H. A. Atwater, *Nano Lett.* **9**, 897 (2009).
118. E. D. Palik and J. K. Furdyna, Infrared and microwave magnetoplasma effects in semiconductors, *Rep. Prog. Phys.* **33**, 1193 (1970).
119. D. J. Lockwood, G. Yu, and N. L. Rowell, *Solid State Commun.* **136**, 404 (2005).
120. P. R. West, S. Ishii, G. V. Naik, N. K. Emani, V. M. Shalaev, and A. Boltasseva, *Laser Photon. Rev.* Published Online: 1 Mar 2010.
121. I. Hamberg and C. G. Granqvist, *J. Appl. Phys.* **60**, R123 (1986).

PLASMONICS FOR ULTRASENSITIVE NANOSPECTROSCOPY AND OPTOFLUIDIC-PLASMONICS BIOSENSORS

Hatice Altug*,†,‡,§, Ahmet A. Yanik*,†, Ronen Adato*,
Serap Aksu‡, Alp Artar* and Min Huang*

*Department of Electrical and Computer Engineering,
†Photonics Center and
‡Materials Science and Engineering Division,
Boston University, 8 St. Mary St.,
Boston, MA, 02215, USA
§altug@bu.edu

Plasmonics, localizing light to the sub-wavelength dimensions and dramatically enhancing local fields, is enabling new possibilities for realization of advanced biosensors that can detect and analyze small quantities of biomolecules and dangerous pathogens. In this chapter, we will focus on integrated plasmonic systems for ultrasensitive infrared nanospectroscopy and biodetection. Infrared absorption spectroscopy, which directly accesses vibrational fingerprints of the biomolecules/ chemicals at mid-IR frequencies, is an important identification and analysis tool. However, small absorption cross sections of the molecules strictly hinder the usage of this technique for identifying small quantities of biological specimen. We will demonstrate diffractively coupled plasmonic nanoantennas enabling ultra-sensitive surface-enhanced spectroscopy with zeptomole level sensitivities. In addition, we will introduce a low-cost fabrication technique for high-throughput fabrication of such engineered infrared plasmonic antenna arrays. Finally, we will describe a novel biosensing system merging nanoplasmonics and nanofluidics to overcome fundamental mass transport limitations imposed by conventional microfluidic approaches. Our detection platform, manipulating

light as well as directing flow on the through the nanoholes, enables targeted analyte delivery and dramatically improves sensor response time.

1. Introduction

Biosensors that can detect and analyze small quantities of molecules (i.e., proteins, enzymes) as well as dangerous pathogens (i.e., viruses, spores) are very important for clinical applications, biomedical research and national defense. For example, screening of large variety of proteins and determination of their functions can enable early diagnostics of complex diseases such as cancer and Alzheimer's as well as their treatment by discovering effective drugs. Large-scale study of proteins is significant due to such far-reaching implications, however its realization is challenging due to the limitation of current detection techniques. In human proteome, there are more than 100,000 kinds of proteins[1] and some expressed in few copies. Protein functions have to be deciphered without using labeling techniques since labels interfere with the protein interactions.[2] Unlike deoxyribonucleic acid (DNA) replication with polymerase chain reaction (PCR), there is no equivalent method for protein amplification, a major limitation for the identification of low abundant proteins. Similarly, rapid detection and surveillance of infectious pathogens remains to be a challenge for point-of-care applications in public health and national defense. Deliberate release of viruses as a biological warfare agent can cause millions of deaths, if the outbreak is not detected at its onset. Current detection techniques rely on extensive sample processing and require advanced equipment and infrastructure. Furthermore, most common tests (such as flu) are for the detection of single type of pathogens.

Rapid, sensitive and multiplexed detection platforms are needed to screen large variety of biomolecules and bio-agents simultaneously. The sensing technology should also be portable and require minimal sample preparation for point-of-care applications. To meet this need, during the last decade electrical,[3] mechanical[4] and optical[5] sensing techniques have been proposed. Among these, photonic based approaches are very promising. Optical biosensors allow remote transduction of the biomolecular binding signal from the sensing volume. Unlike mechanical and electrical sensors, they are also compatible with physiological solutions and are not sensitive to the changes in the ionic strengths of the analyte solutions.[6]

1.1. *Plasmonic nano-biosensors*

Optical biosensors exploiting dielectric and metallic nanostructures with strong photonic and plasmonic resonances are particularly appealing. Dielectric approaches include photonic crystals[7] supporting either defect,[8] waveguide[9] or slab modes;[10] total internal reflection based micro-resonators supporting whispering gallery modes;[11] and the dielectric waveguides supporting optical modes.[12] Although some of these dielectric micro-resonators have been shown to be highly sensitive,[13] their practical use outside the laboratory has also been hindered due to the precise alignment requirements.[14] Plasmonics, on the other hand, offer much versatile and powerful detection methodologies by localizing light below the diffraction limit and dramatically enhancing the local fields.[15] Nanostructured metals can interact with light through the excitation of surface plasmons, which are collective electron oscillations leading to surface-bound electromagnetic fields.[16] Depending on the geometry, composition and the environment, metallic nanostructures can support two distinct forms of surface plasmon resonances (SPR): (1) localized surface plasmons (LSPs) and (2) propagating surface plasmon polaritons (SPPs).[17] LSPs, for example, are created in "isolated" plasmonic nanoparticles with dimensions smaller than the incident wavelength of light. SPPs, on the other hand, are realized in semi-infinite metal sheets. The contributions of these plasmonic modes strongly determine the optical characteristics of the metallic nanostructures and their suitability for specific applications such as spectroscopy, sensing and waveguiding.[18]

Plasmonic nanoparticles,[19] through the excitation of LSPs, can act as very efficient optical antennas and serve as optical transmitters and receivers. By producing dramatically enhanced near-field intensities confirmed to sub-diffraction limited volumes, they can enable light-matter interaction at an extreme.[20] These abilities are strongly controlled by the shape as well as the arrangement of particles in arrays.[17] Therefore, in recent years, we have seen a surge of novel particle designs[21] ranging from nanorods and nanoshells to bow-ties as well as innovative particle arrangements such as self-similar chains[22] (Fig. 1). Large field intensities realized in plasmonic structures are crucial in obtaining orders of magnitude enhanced signals from Raman[23] and infra-red vibrational spectroscopies, which are two powerful techniques for identification of unknown biochemical samples.[24] In addition, the resonance frequencies of LSPs are strongly dependent on the local refractive index, which can be utilized for label-free sensing of biomolecules and biomolecular binding kinetics.[25]

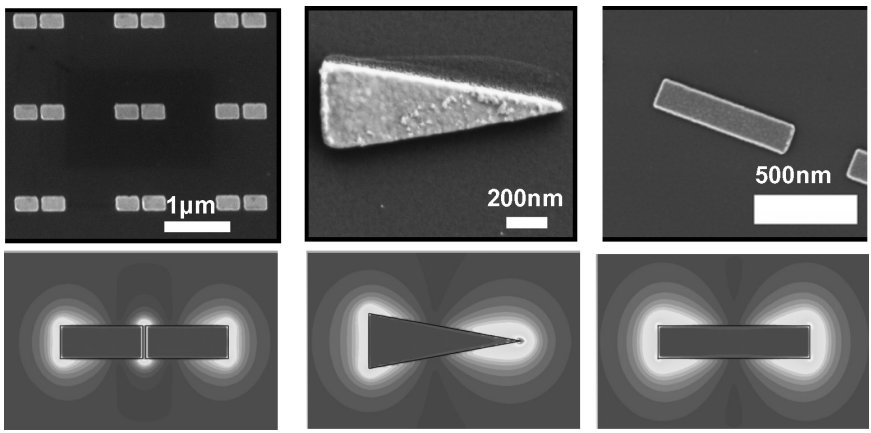


Fig. 1. Scanning electron microscope images (top row) of various shapes of plasmonic nanoparticles and their corresponding near-field intensity distribution in space (bottom row) calculated by three-dimensional finite difference time domain simulation method.

One of the most studied plasmonic nanostructures involving SPPs are nanohole arrays in metal films.[26] At specific frequencies and for certain geometries and periodicities, the light transmission in subwavelength nanoaperture arrays can be orders of magnitude greater than that predicted by Bethe, in his calculations for a single hole in a perfectly conducting metal screen.[27,28] The mechanism involved in this phenomenon, so called extraordinary light transmission (EOT) effect, has been under debate.[29] Recent numerical and experimental studies[30,31] indicate a strong role of SPPs as well as LSPs supported around the rims of the nanoholes.

In EOT effect, incident light goes through different electromagnetic states as it travels from the incidence surface to the out-coupling surface.[31] Three-dimensional finite difference time domain (3-D FDTD) simulations in Fig. 2 summarize excitation, transfer and out-coupling of surface plasmons for the rectangular nanohole arrays defined on an optically thick metal film. Here, light is incident from the bottom surface and interacts first with the substrate/metal interface. Initially, SPPs are created on this interface at resonance wavelengths (λ_{sp}), where the momentum matching condition is satisfied

$$\lambda_{sp} = (i^2 + j^2)^{-1/2}\, d \left(\frac{\varepsilon_{Au}\varepsilon_d}{\varepsilon_{Au} + \varepsilon_d} \right)^{1/2} \quad (1)$$

Here, d is the lattice constant, and ε_d (ε_{Au}) is the dielectric constant of the surrounding dielectric (gold film), (i, j) are the diffraction grating orders. At

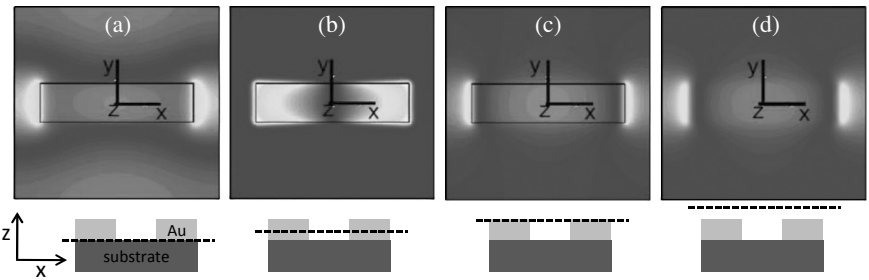


Fig. 2. (a) Excitation (b–c) transfer and (d) out-coupling of surface plasmons from rectangular nanohole arrays on optically thick metal film for a y-polarized incident plane wave. Intensity of the dominant H-field components [which are H_x for Fig. 1(a), (c–d) and H_z for Fig. 1(b)] are plotted at interfaces indicated below (with dashed horizontal line).

resonance, a symmetric standing wave pattern is evident in the field profile due to the interference of the counter propagating SPP waves in y-direction [Fig. 2(a)]. Hot spots around the rims of the cavities are due to excitation of LSPs. These localized hot spots scatter the SPPs into attenuated waveguide modes of the sub-wavelength dimension cavities [Fig. 2(b)]. The waveguide modes primarily mediate the coupling between the two metal surfaces and transfer the electromagnetic field to the top (metal/air) interface. At metal/air interface, SPP excitations are prohibited due to the effective index difference causing momentum mismatch [Fig. 2(c)]. Finally, at metal/air interface, LSP and waveguide modes are converted directly to the out-coupling photons [Fig. 2(d)].

Resonance wavelength (λ_{sp}) of SPP mediated resonant optical transmission strongly depends on the effective refractive index of the surrounding medium [ε_d in Eq. (1)]. This dependence has motivated the use nanohole arrays for bio-chemical detection.[32] Nanohole arrays, acting as a grating, can excite SPPs with a normally incident light as described above. This coupling scheme eliminates the need for precise alignment requirements and allows large-scale multiplexing of plasmonic nanohole array sensors.

Within the last decade, there has been significant progress in our understanding of plasmons and its application for bio-detection; however many fundamental questions and practical challenges still remain. For example, in biosensing applications based on refractive index change, the detection limit depends on both the LSPR sensitivity to the local dielectric environment and the resonance line-width.[33] Narrower line-widths allow smaller shifts to be detected.[34] So far, most of the effort has been concentrated on optimizing

the nanoparticle geometries to improve the near-field enhancements and to control the resonance frequencies/line-widths.[35,36] Further improvements in plasmon line-widths (lifetimes) and near-field enhancements require innovative approaches such as "*radiative engineering*" of plasmonic losses. Another problem in biosensing applications is the "*size mismatch*" between the nanoscale optical sensing volumes and the micronscale fluidic circuits utilized for analyte delivery. As a result, performances of the biosensors are often limited in a fluidic environment by the inefficient analyte (mass) transport, instead of their intrinsic detection capabilities.[37] In this book chapter, we will present some of the unique detection systems addressing these challenges. We will start with an ultra-sensitive, collectively enhanced infrared absorption spectroscopy (CEIRAS) enabling direct detection of molecular specific signatures of zepto-mole levels of proteins.[38] Next, we will present a high-throughput fabrication method for low-cost production of these engineered plasmonic substrates.[39] Finally, to overcome size mismatch problem, we will introduce a novel hybrid biosensing platform merging nanoplasmonics and nanofluidics.[40]

2. Mid-Infrared Plasmonics for Ultrasensitive Nanospectroscopy

Infrared spectroscopy is a unique tool for identifying and characterizing molecular bonds.[41] For most organic and inorganic molecules (such as proteins, chemical toxins and gases), vibrational and rotational modes are spectroscopically accessible within the mid-infrared (mid-IR; 3–20 μm) regime of the electromagnetic spectrum. Characteristic vibrational modes are associated with unique IR absorption spectral bands that are bond-specific. Because of that, the IR wavelength range is also known as "finger print" region. While the absorption cross-sections of mid-IR active modes of proteins are nearly 10 orders of magnitude larger than the corresponding Raman cross-sections, signal levels are still small compared to that of fluorescence-label based methods.[42] By exploiting extremely strong localized fields of plasmons at infrared frequencies,[43] dramatic signal enhancement in mid-IR spectroscopy can be obtained. In analogy to SERS,[44] this method is called surface enhanced infrared absorption spectroscopy (SEIRAS).[45] Until recently, the bulk of SEIRAS studies have revolved around the enhancements achieved on chemically prepared metal surfaces.[46] In these substrates, however, due to the random nature of the surface, the signal enhancement factors[47] have been limited to only 10–100.

Plasmonic excitations on engineered substrates on the other hand, could result in much higher and reproducible enhancements.

Interest in plasmonic characteristics of metal nanoparticles in engineered arrays has grown significantly within the last decade.[48] With the recent proposals on diffractively coupled arrays supporting "*remarkably narrow resonances*", nanoparticle research has taken a new twist.[49] Recently, researchers show that the collective plasmonic excitations in these coupled arrays could also result in much larger near-field enhancements[50,51] opening up new opportunities in bio-detection applications. However, research has been mainly focused on visible and near-IR frequency range of the spectrum for applications such as SERS.[52] In a series of recent publications, we have recently extended these diffractively coupled collective plasmonic excitations to the mid-IR frequency range and demonstrated an ultra-sensitive absorption spectroscopy tool.[38,51] By spectrally overlapping these plasmonic excitations with the vibrational modes of the protein, we achieved detection of vibrational signals from proteins at zeptomole levels. In the measurements, we are particularly interested in amide-I and amide-II bands of the proteins (at $1660\,\mathrm{cm}^{-1}$ and $1537\,\mathrm{cm}^{-1}$), the spectroscopic signatures of the peptide backbone.

2.1. *Radiative engineering with collective plasmons on antenna arrays*

Plasmonic behavior of the nanoparticles in an ensemble can strongly differ from that of the individual constituent nanoparticles.[51,53] This phenomenon can be understood using coupled dipole approximation (CDA).[54] For an individual nanoparticle, the acting field is only the incident field exciting the LSPRs ($\mathbf{E}_{\mathrm{act}} = \mathbf{E}_{\mathrm{inc}}$). A nanoparticle responds to this electric field with an induced dipole moment, $\mathbf{p} = \alpha_p \mathbf{E}_{\mathrm{act}}$. In an ensemble, on the other hand, the acting field on the individual particle includes both (i) the incident field and (ii) the sum of the retarded dipolar fields due to the other nanoparticles:[55]

$$\mathbf{E}_{\mathrm{act},i} = \mathbf{E}_{\mathrm{inc},i} + \sum_{\substack{j \neq i \\ j=1}}^{N} e^{ikr_{ij}} \mathbf{C}_{ij} \mathbf{p}_j \tag{2}$$

Here, $\mathbf{C}_{ij}$ is the phase term written out explicitly to emphasize its importance, and $\mathbf{C}_{ij}$ is the dipolar interaction matrix without that term.[38,51] Indices i and j label the ith and jth particles, r_{ij} is the distance between them, and N is the total number of particles. The

sum in Eq. (2) strongly depends on the phase delay experienced by the light propagating among particles. For a periodically arranged nanoparticle array, the scattered fields add in phase at specific wavelengths when $kr_{ij} = 2\pi m$, where m is an integer. These correspond to the appearance of grating orders. For wavelengths shorter/longer than these transition wavelengths, the grating order is evanescent/radiative. Interesting physical phenomena leading to the narrowing of the plasmonic resonances and the enhanced near-fields are observed around these transition wavelengths. A quantitative understanding of the phenomena can be developed for an infinite chain of identical nanoparticles excited by normally incident light. In this case, dipolar moments of the constituent particles are the same $\mathbf{p}_j = \mathbf{p}_i = \alpha_p \mathbf{E}_{\mathrm{act},i}$ and Eq. (2) can be simplified to:

$$\mathbf{E}_{\mathrm{act},i} = \mathbf{E}_{\mathrm{inc},i} + \underbrace{\left(\sum_{\substack{j \neq i \\ j=1}}^{N} e^{ikr_{ij}} \mathbf{C}_{ij} \right)}_{S} \alpha_p \mathbf{E}_{\mathrm{act},i} \tag{3}$$

Following this relation, local electric field can be expressed as $\mathbf{E}_{\mathrm{act},i} = (1 - \alpha_p S)^{-1} \mathbf{E}_{\mathrm{inc},i}$, where S is the retarded dipole sum defined in the parentheses in Eq. (3). Accordingly, an effective polarizability for nanoparticles can be defined as:

$$\alpha_{\mathrm{eff}} = \frac{1}{1/\alpha_p - S} \tag{4}$$

Therefore, we can relate dipole moment to incident field by $\mathbf{p}_i = \alpha_p \mathbf{E}_{\mathrm{act},i} = \alpha_{\mathrm{eff}} \mathbf{E}_{\mathrm{inc},i}$. This equation shows that the polarizabilities of the nanoparticles in an ensemble are controlled by the retarded dipole sum S, which is only a function of geometrical parameters. A maximum both in the imaginary part and modulus part of the particle's complex polarizability, thus a peak in extinction spectrum corresponding to the array resonance, is expected when the real part of the denominator $(1/\alpha_p - S)$ vanishes. Creation of collective resonances can be explained using Eq. (4) for an infinite chain of nanorod particles.

In Fig. 3(a) the nanoparticle polarizability (α_p) and the retarded dipole sum (S) are shown with respect to the wavelength of the incident light, which is normally incident and polarized perpendicular to the chain axis. The particles are modeled as gold ellipsoids, with the dielectric function computed from a Lorentz-Drude model.[56] In order to account for the finite size of the particle, α_p is computed using the modified long

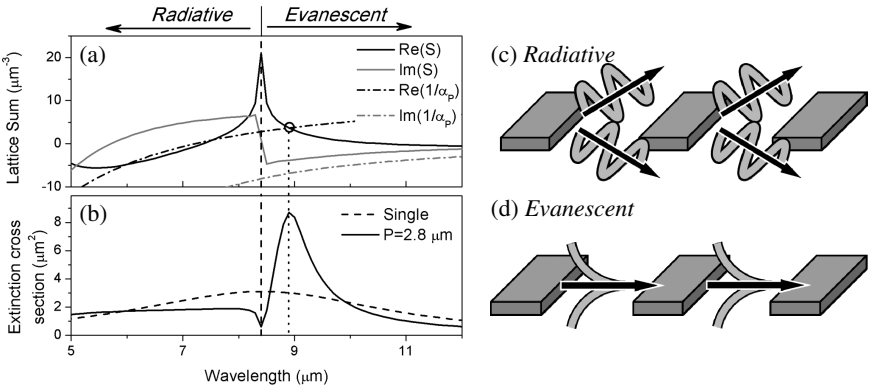


Fig. 3. Adapted from Ref. (51). (a) Lattice sum and (b) extinction spectra are shown for 1D nanoparticle chains. The particles are gold ellipsoids with dimensions $1.6 \times 0.3 \times 0.1\,\mu$m and the background refractive index is 3. A modified long wavelength approximation (MWLA) is used for the single particle polarizability. The particles are arranged on a $2.8\,\mu$m periodic chain. Light is normally incident and polarized perpendicular to the chain axis. The real part of S diverges at the grating transition wavelength [black vertical dashed line in (a)] while the imaginary part exhibits a drastic sign change. At wavelengths below (above) the transition wavelength, a grating order has radiative (c) [evanescent (d)] character.

wavelength approximation.[57] Computation of S involves evaluating an infinite summation. This was done numerically with the sum terminated at $N = 400$ particles. As shown in Fig. 3, the real part of the inverse nanoparticle polarizability monotonously increases starting from a negative value and crosses to the positive domain at the LSPR wavelength of a single nanoparticle. The lattice sum S, shows a more complex behavior. For large particle separations, $\mathbf{C}_{ij}$ is dominated by the far-field term, which is a real positive number. Hence, for a periodic chain where r_{ij} is an integer multiple of the inter-particle spacing, the real part of the lattice sum S diverges at the diffraction condition ($kr_{ij} = 2\pi m$). This is visible as a sharp maximum in the figure at the grating transition wavelength (dashed vertical line). Im(S) exhibits a rapid sign change around this grating transition wavelength. Imaginary part of the lattice sum, Im(S), is positive (negative) when the grating order is radiative (evanescent) resulting in increased (decreased) radiative damping. The sudden appearance of the new grating order causes a dramatic increase in the radiated power from the array, which is closely associated with the Wood anomalies and Rayleigh's explanation.

As shown in Fig. 3(a), cancelation of the real terms in the denominator in Eq. (4) occurs at a wavelength (dotted vertical line) slightly longer than

the grating transition wavelength, where the real part of $1/\alpha_p$ (green curve) crosses the real part of lattice sum (red curve) as indicated by circle in Fig. 3(a). A maximum both in the imaginary and the modulus of the particle's complex polarizability (corresponding to resonance in extinction spectrum) is observed at this crossing wavelength, which is slightly red-shifted from the LSP resonance of the individual nanoparticle. More interestingly, the imaginary part of S is negative at the array resonance wavelength [Fig. 3(a)] and partially cancels the imaginary parts of $1/\alpha_p$. This partial cancelation (which, physically corresponds to the suppression of the radiative damping) results in linewidth narrowing of the far-field extinction resonance.[51,58] In addition, enhanced intensities in the near-field are expected as a result of stronger field confinement in the array plane.[38]

This behavior is well captured by the 3D-FDTD simulations presented in Fig. 4. An order of magnitude stronger near-field intensities than those of individual antennas are observed at slightly shorter wavelengths than $\lambda_{Si(1,0)}$. Here, $\lambda_{Si(1,0)}$ is the critical wavelength where the $(1,0)$ grating order of the silicon interface becomes radiative at a given lattice periodicity ($\lambda_{Si(1,0)} = n_{Si}d$).

Far-field response of collective plasmons is shown in Fig. 5 for a periodic array of rod shaped gold nanoparticle antennas fabricated by electron beam lithography (EBL) on silicon substrate [Fig. 5(a)]. The arrays consist of 1100 nm long rods with varying periodicities ranging from 1.5 μm to 2.0 μm, as well as a random set that serves as a control to probe the individual

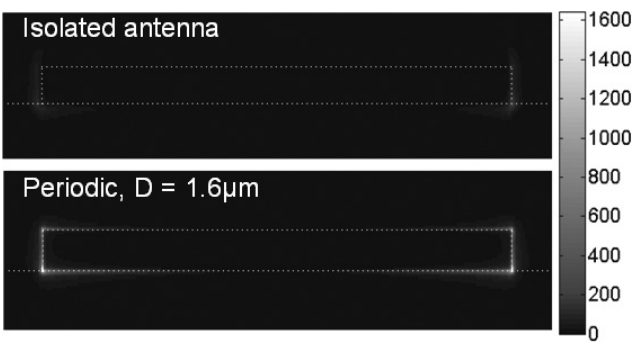
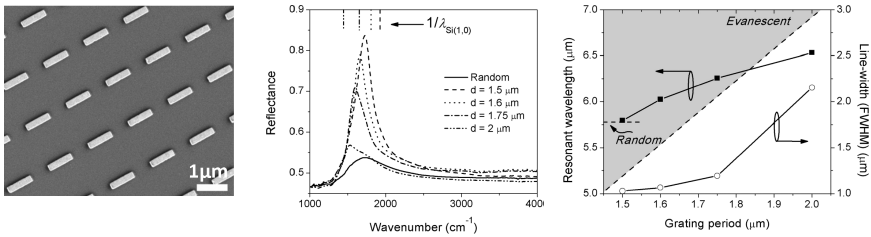


Fig. 4. Adapted from Ref. (38). Cross-sections of the intensity distribution taken through the edge of the rod are shown for periodic ($d = 1.6\,\mu$m) and isolated antenna. Collective resonances in periodic arrangements of nanoantennas can give rise to nearly an order of magnitude larger near-field intensity enhancements in comparison to the isolated antenna.



(a) (b) (c)

Fig. 5. Adapted from Ref. (38). (a) Scanning electron microscope (SEM) images of a periodic array. (b) Reflectance spectra of periodic nanoantenna arrays. All the antennas are 1100 nm long. The wavenumbers corresponding to $1/\lambda_{Si(1,0)}$ for a given periodicity are indicated by the dashed lines at the top of the figure. (c) The variation in peak position (black squares) and linewidth (white circles) with the grating period are shown. The dashed line indicates the wavelength at which the grating order transitions from evanescent to radiative in character. The grating order is evanescent in the shaded region.

antenna response. The resulting reflectance spectra are shown in Fig. 5(b). For periodicities smaller than 2 μm, collective resonance peaks are above the $(1, 0)$ grating order cutoff of the silicon interface [dashed line in Fig. 5(c)] where the field is evanescent. As shown in Fig. 5(c) (red triangles), this results in narrower plasmonic far field resonances with respect to individual antenna. A progressive blue shift is expected for greater separations as the coupling among antennas is reduced, and the optical response converges back to that of an isolated particle. At 2 μm periodicity, where the grating order has changed from being evanescent to radiative, plasmonic resonances have a broader linewidth. As shown in Fig. 5(c), in comparison, 1.6 μm periodic array offers the best combination of narrow linewidth ($\sim$1 μm versus 2.75 μm for the individual particle behavior) and spectral overlap with the protein amide-I band at 1660 cm^{-1}.

2.2. *Collectively enhanced infrared absorption spectroscopy*

Direct identification of the vibrational signatures of the protein monolayers is achieved by utilizing well optimized collective excitations. Here, a 2 nm thin film of silk fibroin protein layer is applied uniformly to the nanoantenna substrates by spin coating. Atomic force microscopy (AFM) is used to confirm the uniformity and the thickness of the film [Fig. 6(a)].

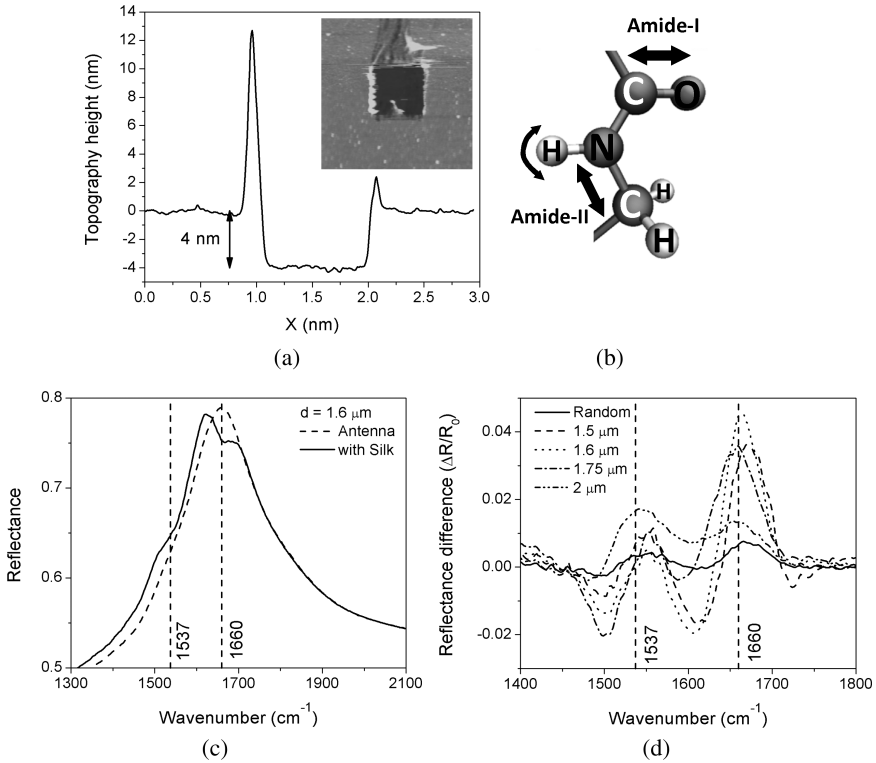


Fig. 6. Adapted from Ref. (38). (a) Silk film thickness is measured by atomic force microscopy for a 4 nm thick film. (b) Amide-I and II vibrational modes of the protein back bone. (c) Reflectance spectra from the 1.6 μm periodic array before (dashed line) and after coating of 2 nm thick protein film (solid line). Dashed vertical lines indicate the positions of the protein amide-I and II absorption peaks. (d) Difference absorption spectra of the arrays whose spectral characteristics before protein coating are given in Fig. 4(b).

As shown in Fig. 6(c), protein absorption bands are clearly noticeable within the optical spectra collected from the protein-coated antenna arrays with 1.6 μm periodicity. Dips in the plasmonic response as a result of the amide I and II absorption bands are indicated in the figure at 1660 and 1537 cm^{-1}, respectively. Capacitive loading of the antenna with the protein layer results in slight red shifting of the plasmonic resonances[59,60] (dashed curve in the figure). This shift is corrected using a polynomial fitting procedure in difference spectrum measurements ($\Delta R/R_0 = R_{before}/R_0 - R_{after}/R_0$, where R_0 is the reflection signal from reference mirror). Figure 6(c) shows the difference spectra of the periodic and the randomized

arrays. The enhanced absorption signals well above the noise level are observed for the antenna arrays with narrow linewidth (1.6 μm periodicity). Control measurements were performed on the bare silicon substrate with protein films of the same thickness in a region far from any fabricated nanoantennas. No absorption features were observed from the control samples. The observed signal ($\Delta R/R_0$) of 6.8% in the periodic structures is nearly an order of magnitude higher than that of the randomized array, which is only 0.9%. This improvement is in agreement with the enhanced near-field intensities predicted by the FDTD simulations (Fig. 4).

In order to calculate the near-field enhancement factor for the CEIRAS signal, we compared the "enhanced" signal collected from the 1.6 μm periodic array to the expected reflectance signal from a 2 nm thick silk film on bare silicon substrate. Since the signal from 2 nm thick protein films on bare silicon is below the noise level, instead we performed IR reflection absorption spectroscopy (IRRAS) measurements at a grazing angle (80°) on a 100 nm thick gold layer deposited on a silicon substrate. The expected value of the absorption signal for a normally incident light is obtained to be 4.7×10^{-2}%. For an accurate estimate of the enhancement factor, we also need to include the following factors. The enhanced signal mainly comes from a small quantity of molecules at the close vicinity of the N^2 nanorod tips ($N = 63$ is the number of rows and columns of the antenna array). Secondly, in contrast to the studies conducted with self-assembled monolayers (SAMs),[43] here the protein molecules are physisorbed, which results in the following differences. (i) Unlike the SAMs, the transition dipole moments of the physisorbed proteins have no fixed orientation with respect to the metal surface normal. Accordingly, we expect that approximately one third of the transition dipoles of the molecules at the nanorod tips contribute to the absorption signal. (ii) Additionally, unlike in chemisorption, lack of molecular binding to the metal surface rules out any contribution of the chemical effects. (iii) Finally, given that the silk film is only 2 nm thick, it is unlikely that the physical deposition method results in the same degree of uniform coverage over the 70 nm high vertical sidewalls of our nanorods as would be possible with a SAM method. With these in consideration, we estimated the signal enhancement to be within the range of 10^4–10^5. Such large enhancements allow us to detect absorption signals even in the raw spectral measurements from monolayer protein films with a commercial FTIR microscope. The detection volume of the antenna is calculated by considering the lateral area of the nanotips and the thicknesses of the monolayer protein film. Given that the density of

the silk protein[61] is 1.4 g/cm^3 and the molecular weight[62] is approximately 375 kDa, we estimate that the measured absorption signals are obtained from about 300 zeptomoles for the entire array, corresponding to only 145 silk molecules per antenna. The large signal to noise ratios achieved in our measurements indicate that with plasmonics we should be able to observe vibrational signatures from even smaller quantities of protein molecules down to a few tens of zeptomoles.

3. High Throughput Fabrication of Plasmonics with Nanostencil Lithography

As shown in Section 2, large field enhancements in plasmonics could lead to ultra-sensitive infrared spectroscopy. The strong light matter interactions and extreme light manipulation with plasmonics is not only important for spectroscopy and biosensing but also for other novel device applications such as superlensing, ultra-fast detection, cloaking. These advances however, are critically dependent on our ability to structure metals in a controllable way at sub-100 nm resolution. The most common top-down nanopatterning techniques with high resolution are electron and focused ion beam lithography (FIB). Both EBL and FIB offer tremendous flexibility in creating large variety of nanostructure geometries and patterns at high resolution. However, their major drawback is the low throughput. Due to their serial nature, each nanostructure has to be created one at a time, which is both slow and expensive. As a result, in recent years there has been a surge of new fabrication techniques for high throughput nanopatterning.[63–67] Among them, one innovative approach is nanostencil lithography (NSL). NSL is a shadow-mask patterning technique[68–70] that can allow fabrication of structures below 100 nanometer resolution. The method relies on direct deposition of materials through a pre-patterned mask. The deposited material could be metallic, dielectric and organic. The mask, which acts as a stencil, is fabricated on suspended silicon nitride membrane using EBL (or FIB) and dry etching methods. The stencils containing large numbers of nanoapertures/nanoslits with a variety of shapes, sizes and arrangements can be fabricated on wafer scale for high throughput nanofabrication.[69] When placed in contact on a desired substrate, direct deposition of materials (such as noble metal) enables lift-off free production of nanoparticles and nanowires with high reliability and uniformity. Since NSL does not require any resists, it has the advantage of reducing the number of fabrication steps and allowing the patterning

on different types of substrates. Another advantage of NSL is that the masks can be reused to pattern the same nanostructures multiple times with minimal effort.

In a recent work, we demonstrated high throughput fabrication of infrared plasmonic nanorod antenna arrays using nanostencil lithography.[39] We showed that the technique offers the flexibility and the resolution to manufacture plasmonic substrates supporting spectrally sharp collective excitations at mid-infrared wavelengths. The extinction spectra of our antenna arrays are comparable to that of the arrays fabricated by EBL. More importantly, we showed nanostencil masks can be reused multiple times to create series of nanoantenna arrays leading to same optical responses. Our observation is confirmed by optical measurements as well as scanning electron microscopy images. Finally, we demonstrated fabrication of nanostructures in various shapes with a single metal deposition step on different substrates, including non-conducting surfaces (i.e., CaF_2 and glass). This nanofabrication scheme, by enabling the reusability of stencil and offering flexibility on the substrate choice and nano-pattern design, could significantly speed up the transition of plasmonic devices into the real-world applications.

3.1. *Nanostencil lithography technique*

Nanostencil technique, summarized in Fig. 7, consists of three consecutive steps: (i) fabrication of the free standing membrane, (ii) patterning on the membrane and (iii) direct deposition of metallic plasmonic devices on a substrate. Processing steps for free standing membrane fabrication is illustrated in Fig. 7(a). An important consideration here is the mechanical strength of the membranes. Highly robust Low Pressure Chemical Vapor Deposition (LPCVD) SiN_x films are an excellent choice.[40,71] Fabrication process starts with 550 μm thick silicon wafers coated with 400 nm thick LPCVD SiN_x on double sides. By optical lithography, dry and wet etching methods, we form $\sim$200 μm $\times$ 200 μm and 400 nm thick free standing SiN_x membranes. The second stage is nanoaperture patterning on the membrane with EBL and dry-etching methods. Here, the EBL process is needed only once for the creation of the mask, since the mask can be used multiple times. The resulting structure acts as a stencil. The top and the cross sectional views of a fabricated stencil masks are shown in Fig. 7(b).

Final stage of the fabrication method involves direct deposition of the plasmonic structures to the desired surface. To get high quality structures,

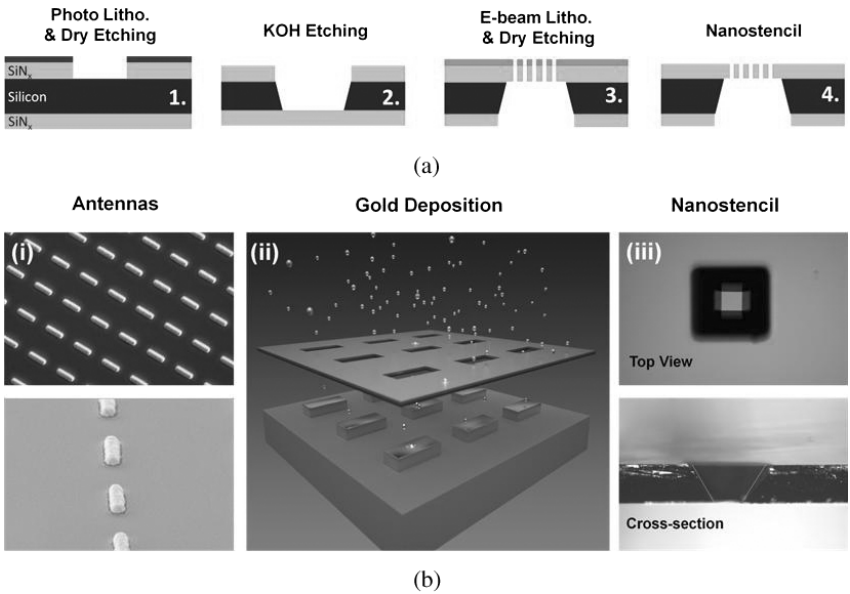
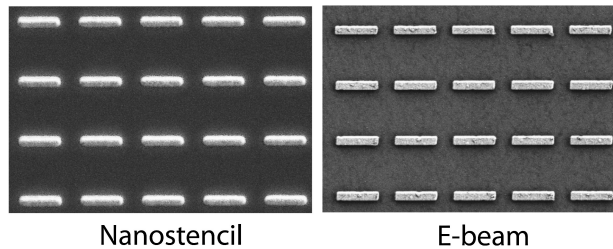


Fig. 7. Adapted from Ref. (39). (a) Fabrication of free standing membrane and nanostencil is illustrated from 1 to 4. Dry and wet etching processes and EBL are used for achieving precisely defined pattern of nanoapertures/slits on the membrane. (b) (i) Top and angled SEM images of the gold nanorod with 1100 nm length, 230 nm width and 100 nm height are shown. (ii) Gold deposition scheme with reusable mask is illustrated. (iii) Top view and cross sectional images of the stencil are shown.

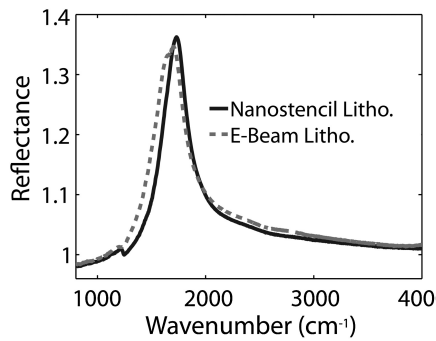
the gap between substrate and the stencil must be minimized. Here, we directly place the stencil on the substrate and secure it tightly by mechanical force using clips so that the patterned SiN_x layer is nominally kept in contact with the substrate. Then, directional gold deposition is performed for 100 nm gold film without depositing any prior adhesion layer (such as Ti or Cr) [Fig. 7(b)]. Unlike EBL, an adhesion layer is not necessary since NSL does not require metal lift-off processes. When the mask is removed from the substrate, it leaves plasmonic nanostructures on the substrate with the shapes complimentary to the nanoapertures.

3.2. *High quality plasmonic resonances with NSL*

Using NSL, we show successful fabrication of various plasmonic nanostructures, including nanowires and nanoparticles, in different arrangements and on different substrates. Figure 8(a) displays nanorod arrays fabricated using EBL and NSL on silicon with periods 1.5 μm, width 230 nm and height



(a)



(b)

Fig. 8. Adapted from Ref. (39). (a) SEM images of nanorods with 1100 nm length, 230 nm width and 100 nm height fabricated using NSL (left) and EBL (right) are shown for comparison. Scale bars represent 1 μm. NSL technique for large area patterning of nanorod arrays provides comparable quality with arrays fabricated using EBL. (b) Nanorods fabricated using two different techniques give identical reflection spectrum with resonances at $\sim$1700 cm^{-1}.

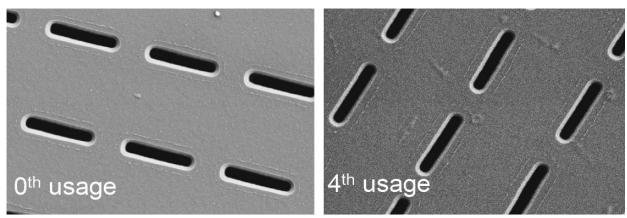
100 nm. No irregularities on the periodicity or the physical dimensions are detected for the nanorod arrays fabricated using NSL. Round edged nanoapertures on the mask caused particles to have rounded tips. Due to unavoidable gap between the mask and the substrate, scattering of gold particles within 20 nm vicinity of the nanorods has also been observed. But, as we demonstrate below, this scattering has negligible effect on the optical quality of the structures.

The reflection spectra of the NSL fabricated arrays [Fig. 8(b)] show strong resonance at the designed wavelength of 5.77 μm. Both the intensities as well as the linewidth of the resonance are comparable for the same size arrays fabricated by EBL. This observation clearly demonstrates the high optical quality of the plasmonic nanoantenna arrays obtained with NSL. As we have introduced in Section 2, collective excitation of the antenna

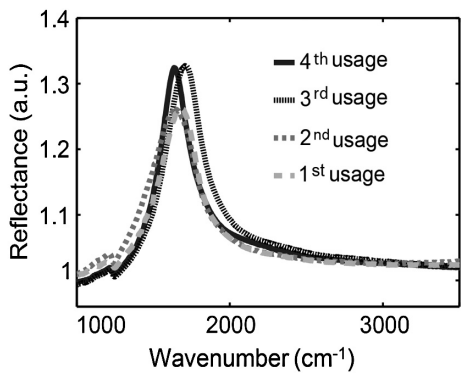
arrays leading to enhanced near-field intensities are highly suitable for ultrasensitive vibrational nanospectroscopy.

3.3. *High throughput nanofabrication with NSL*

The unique advantage of the nanostencil lithography is that stencils can be reused multiple times. This capability is particularly useful when high throughput replication of the optimized nanoparticle arrays is desired. We achieve this by first dipping the used stencil in wet metal etchant and then rinsing in DI-water. Here, gold etchant is used to clean the stencil including the remnants inside the nanoapertures after deposition. After the deposited metal is stripped away, the stencil can be reused.



(a)



(b)

Fig. 9. Adapted from Ref. (39). (a) SEM images of the same mask are shown before its first usage and after the fourth usage. Apertures have dimensions of 1050 nm length, 200 nm width and 100 nm height. No sign of degradation and deformation is observed on the mask after fourth usage. (b) Reflectance spectra are shown for different nanorod arrays obtained from four consecutive depositions using the same mask. The resulting spectra show negligible deviations on resonances (3.5%) around 1700 cm^{-1}.

SEM images in Fig. 9(a) show the stencil mask right before the initial deposition, and also after the fourth usage (with subsequent cleaning). The sizes of the apertures, 1050 nm and 200 nm in width and height, are almost the same for both cases. There is no sign of degradation and deformation after the fourth usage, indicating that the stencil can be further reused. Figure 9(b) shows the reflection spectrum of a series of nanorod arrays fabricated on different silicon chips by using the same stencil. The resulting spectra for all the structures have similar spectral profiles. They show strong resonances around 1700 cm^{-1} with deviations in the spectral peak position less than 3.5%, which could be due to the uncontrollable variations of the thicknesses of the deposited metals in the evaporation chamber. Our observations clearly indicate that with a single stencil, optimized designs can be replicated many times with high degree of plasmonic antenna uniformity and high optical quality. This high throughput fabrication capability is in stark contrast to the serial nature of the e-beam lithography. As a result, we believe that NSL can facilitate a significant progress towards the wide usage of plasmonics in real-world applications.

4. Integrated Nanoplasmonic-Nanofluidic Biosensors Molding the Flow of Light and Fluidics

In this section, we introduce a unique plasmonic biosensing system which merges nanoplasmonics and nanofluidics on the same platform. We show how our approach can effectively address the mass transport problem, a fundamental limitation for any surface biosensor. As we have highlighted in the introduction, in recent years label free biosensors combined with innovative signal transduction methods are proposed to push the detection limits down to femtomolar concentrations of analytes.[72–74] Concurrently, researchers are integrating such sensitive and compact nano-sensors with micro-fluidics for automated sample handling.[75,76] While microfluidics can enable portable and lab-on-a-chip systems, recent calculations indicate that performances of surface biosensors can be seriously limited in a fluidic environment by the inefficient analyte (mass) transport instead of their intrinsic detection capabilities.[37,77] As the analytes are collected by the functionalized surface, a depletion zone forms around the sensing area where the analyte transport is diffusive. Depletion zone expands with time until the convective flow ceases its further growth. Random nature of the mass transport in the depletion zone severely limits the delivery of the analytes from the convective flow to the sensor surface. At low

concentrations, this limitation causes impractically long detection times.[77] One can compress the extent of these depletion zones by increasing the convective flow rate.[37] However, such a passive control scheme often results in moderate improvements of the device performances due to the shorter times required for analytes to stream pass the sensing surface. Stirring of the depletion zone using various mixing strategies also seem to result in moderate enhancements in device performances.[78,79] Innovative approaches are needed to overcome the mass transport limitations. One of the main conceptual constraints in previous approaches is that microfluidics and biosensing are always considered as different parts of a sensor platform completing each other but not a fully merged single entity.

Recently, we demonstrated a hybrid biosensing system merging nanoplasmonics and nanofluidics in a single platform.[40,71] Our system employs suspended plasmonic nanohole arrays sealed in a multi-layered microfluidic chambers (Fig. 10). For sensing, we use the extraordinary light transmission effect. At the same time, we utilize the nanoscale openings to actively steer the convective flow to the surface [Fig. 10(b)]. This is contrary to the conventional approach in which the convective flow stream passes over the sensor [Fig. 10(a)]. Using our platform, we showed 14-fold improvement in the mass transport rate constants. Considering that this rate constant appears in the exponential term, such an improvement implies superior analyte delivery to the biosensor surface at low concentrations. In

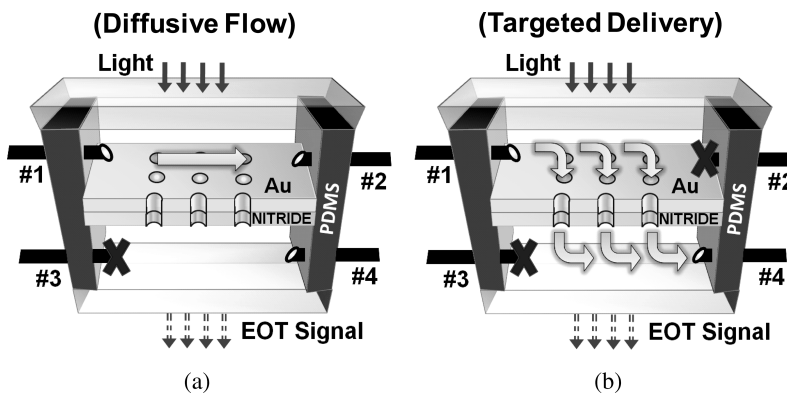


Fig. 10. Adapted from Ref. (40). Multilayered microfluidic scheme allows 3-D control of the convective flow enabling (a) passive (diffusive) and (b) active (targeted delivery) transport of the analytes to the sensing surface. Perpendicular steering of the convective flow is achieved by allowing flow only through one of the inlet/outlets of the top/bottom channels (b).

addition, a unique property of our sensing platform is that it also offers an extra degree of freedom in microfluidic circuit engineering by connecting separate layers of microfluidic circuits through biosensors. Through this platform, it is possible to create "multilayered lab-on-chip systems" allowing three dimensional control of the fluidic flow.

4.1. *Targeted versus conventional fluidics*

To compare the flow profile of our proposed scheme with the conventional approach, we employed the Navier-Stokes equations and finite element method using COMSOL^{TM}. Boundaries of the simulation domain are defined as 200 μm in height and 500 μm in width for each of the channels with 100 μm inlets openings. We employ a refined mesh near the sensor area where the flow rate is high. Figure 11 shows the steady state solutions of the velocity field profile of the fluidic flow. Transfer rate (TR), reflecting the ratio of the perpendicular flow to the inlet flow, is used to quantify the performance of the delivery scheme. When the analyte is injected and collected from the bottom channel, a very weak perpendicular flow analyte towards the top channel is observed [Fig. 11(a)]. Although liquid

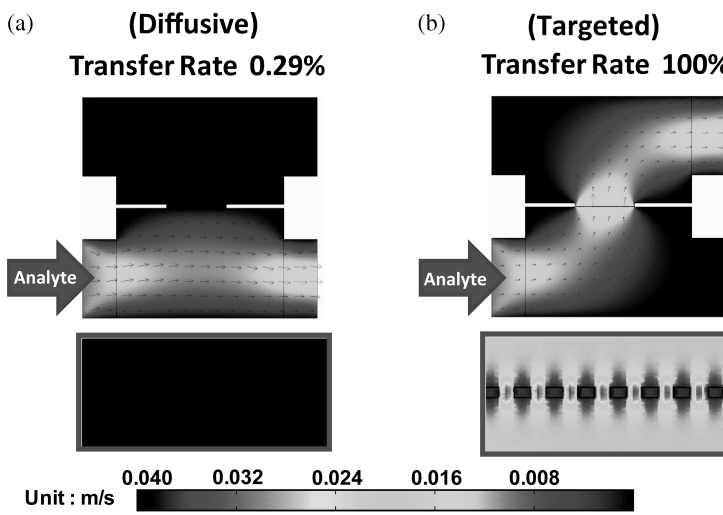


Fig. 11. Adapted from Ref. (40). Microfluidic simulations obtained by solving Navier-Stokes equations demonstrate (a) low transfer rates for the passive transport scheme due to the weaker perpendicular flow of the analytes, while (b) much more efficient mass transport towards the surface is observed for the targeted delivery scheme.

flow through the top outlet is allowed in the simulations, the flow rate at this outlet is calculated to be extremely small (only 0.29% of the input flow). On the other hand, active directing of the convective stream towards the surface and through the nanoholes (inset) results in a transfer flow rate of 100% [Fig. 11(b)]. Accordingly, in an arrangement where the fluidics is flowed in and out from different sides of the biosensor, all the fluidic current is actively transferred through the nanohole openings.

4.2. *Lift-off fabrication of plasmonic nanohole arrays*

In order to implement experimentally the proposed scheme, we need to fabricate plasmonic nanohole arrays on a free standing membrane. Obviously, one important consideration here is the mechanical strength of the membranes as they need to stand the relatively high pressures generated by the perpendicular flow. Mechanically highly robust Low Pressure Chemical Vapor Deposition (LPCVD) silicon nitride (SiNx) films are an excellent choice. The fabrication procedure is similar to the fabrication of nanostencils described in Section 3. First, we form 50 nm thick free standing SiN_x membranes (Fig. 12). Then, the membranes are coated with positive e-beam resist poly(methyl methacrylate) (PMMA), and e-beam lithography is performed. Nanohole pattern (with hole diameters of 220 nm and a periodicity of 600 nm) is transferred to the suspended SiN_x film through a dry etching process [Fig. 12(a)]. E-beam resist is later removed with an oxygen plasma cleaning process leaving only a patterned SiN_x film with air on both sides [Fig. 12(c)]. Finally, we use a directional e-beam evaporator to deposit Ti (5 nm) and Au (125 nm) metal layers defining the suspended plasmonic sensors with nanohole openings [Fig. 12(b)]. This deposition process is observed to be extremely reliable; large areas of metallic nanohole arrays are repeatedly obtained without clogging the openings. Only a small shrinking in nanohole diameter (<4%) is observed after gold deposition due to a slight coverage of the nanohole sidewalls [Fig. 12(c–d)].

Our fabrication technique conveniently results in lift-off free production of plasmonic nanohole arrays with single layer e-beam lithography and etching step. In addition, we eliminate the need for operationally slow focused ion beam lithography, which introduces optically active ions. Fabricated arrays are mounted in a custom designed multilayered microfluidic channel system using poly(dimethylsiloxane). The system also includes multi-inlets/outlets for analyte delivery and fluidic control.

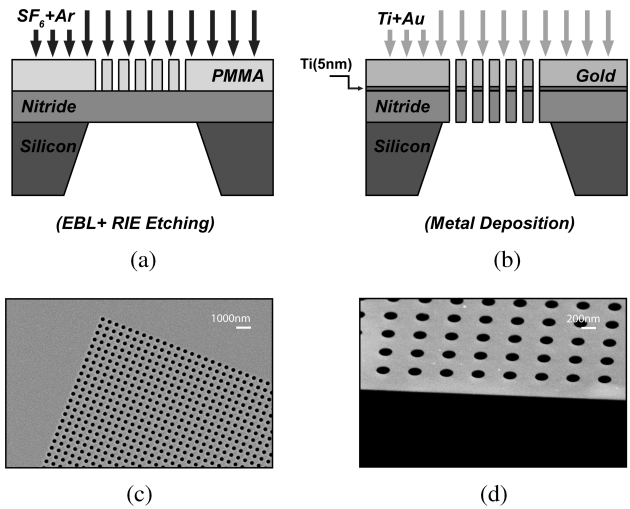
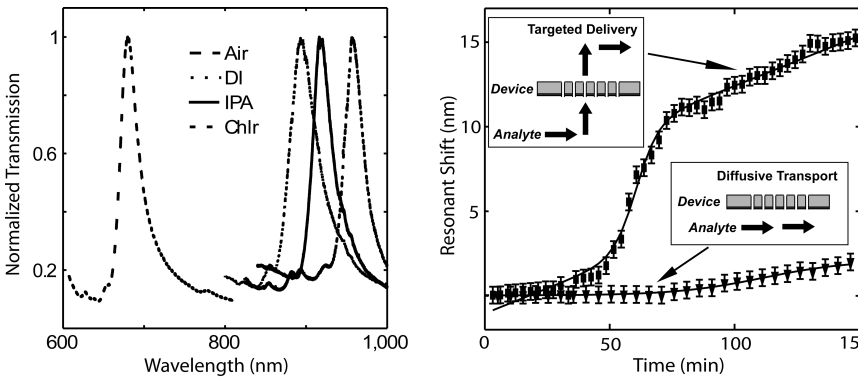


Fig. 12. Adapted from Ref. (40). Suspended plasmonic nanohole arrays are fabricated using a single layer lithography process consisting of following steps: (a) e-beam lithography and pattern transfer with RIE etching, (b) surface cleaning with oxygen plasma and metal deposition. (c–d) Scanning electron images of the nanoholes are shown from different angles for nanohole arrays with 200 nm in diameter and 600 nm in periodicity.

4.3. *Active analyte delivery with sub-wavelength fluidics*

Active delivery of the analytes to the sensing surface is demonstrated in spectral measurements as shown in Fig. 13(a). Initially, both the top and the bottom channels are filled with a low refractive index liquid, deionized (DI) water ($n_{\mathrm{DI}} = 1.333$), at a high flow rate ($550\,\mu$L/min). Once the channels are filled with DI water completely, the plasmonic resonance shifts from $\lambda_{\mathrm{air}} = 679$ nm (air on both sides) to $\lambda_{\mathrm{DI}} = 889$ nm (DI on both sides). This corresponds to a bulk refractive index sensitivity of $\Delta\lambda/\Delta n = 630\,\mathrm{nm}/RIU$, which is also confirmed in independent measurements [Fig. 13(a)] performed in acetone ($n_{\mathrm{acetone}} = 1.356$), IPA ($n_{\mathrm{IPA}} = 1.377$) and chloroform ($n_{\mathrm{chloroform}} = 1.49$). Given that plasmons at the Ti/SiN_x interface are suppressed by the losses, this shift only reflects the spectral response of the EOT resonance to the changing refractive index in the top channel.

The spectrum, obtained once the channels are filled with DI water, is used as a background for the following measurements. To quantify the analyte transport efficiency of the both delivery schemes, a lower viscosity analyte solution (IPA) with higher refractive index is introduced from the bottom inlet and time-dependent spectral measurements are performed.



(a) (b)

Fig. 13. Adapted from Ref. (40). (a) Bulk refractive index sensitivity of the plasmonic nanohole arrays are obtained in different solutions indicates ∼630 nm/RIU. (b) Efficiencies of the passive (triangles) and targeted (squares) delivery of the analytes are compared in real time measurements. Solid lines are fit to sigmoid function. 14-fold improvement in mass transport rate constant is observed for the targeted delivery scheme.

In the diffusive transport scheme, IPA solution is pumped into the bottom channel (4 μL/min) and collected from the bottom side while the top outlet is kept open. For targeted delivery of the convective current to the surface, IPA is directed from down-to-top direction by enabling flow between $3 \rightarrow 2$ (inlets/outlets are numbered as in Fig. 10). As shown in Fig. 13(b), the directed delivery results in a larger resonance shift compared to the passive one, indicating a much more efficient analyte delivery to the biosensor area. Experimentally observed resonance shifts are (least squares) fitted to a sigmoid function of form $A_b + (A_t - A_b)/(1 + e^{-k(t-t_0)})$. This is superposed to a linearly increasing background with $C_t(t - t_0) + C_0$ due to increasing refractive index of the bulk medium in the top channels as IPA concentration increases. The mass transport rate constants are obtained as $k_{\text{pass}} = 0.0158\ \text{min}^{-1}$ and $k_{\text{act}} = 0.2193\ \text{min}^{-1}$ for the passive and targeted transport schemes, respectively. This corresponds to more than 14-fold improvement in rate constants, which is crucial for enhancing the performance in immunoassay based applications.

5. Conclusion and Outlook

In this chapter we presented an ultrasensitive nanospectroscopy tool and a novel biosensing system overcoming mass transport limitation. For surface enhanced spectroscopy techniques, obtaining strong near-field

enhancements are crucial. We showed that when the tailored nanorod antennas are ensembled uniquely, one can create spectrally sharp plasmonic resonances. Collective behavior of plasmonic nanoantennas in such arrays enables much stronger near-field enhancements than what is achievable with an individual nanoantenna. Using nanorod antenna arrays, we demonstrated up to 100,000 fold stronger absorption signals from the peptide backbone of the proteins and demonstrated direct detection of vibrational signals at zeptomole level sensitivities. Our method, adaptable to enhance the IR fingerprints of other biomolecules, can be a general purpose toolkit for ultra-sensitive bio-analysis and identification. To produce our infrared plasmonic antenna array substrates in a high throughput fashion, we also introduced a novel and low-cost fabrication method based on nanostencil lithography. Finally, we demonstrated a hybrid nanoplasmonic-nanofluidic platform, in which nanoholes are employed as optical waveguides transmitting light below the subdiffraction limit, as well as nanofluidic channels transporting analytes through them, for effective mass transport. Unlike previous approaches where the analytes simply stream pass over the surface, our platform enables efficient delivery of the analytes to the biosensor surface. This 3-D dimensional fluidic scheme allows us to overcome the mass transport problem, a fundamental limitation causing impractically long detection times at low analyte concentration. To realize this hybrid platform, we also introduced a lift-off free nanofabrication scheme allowing us to manufacture plasmonic nanohole biosensors with ease and high reliability/uniformity. Currently, we are integrating these novel detection systems on a single platform for complete functional studies of proteins.

Acknowledgments

This work is supported in part by National Science Foundation CAREER Award (ECCS-0954790), Office of Naval Research Young Investigator Award, Massachusetts Life Science Center New Investigator Award, NSF Engineering Research Center on Smart Lighting (EEC-0812056), Boston University Photonics Center and Army Research Laboratory.

References

1. V. Dhingraa, M. Gupta, T. Andacht, and Z. F. Fu, *International Journal of Pharmaceutics* **299**, 1 (2005).
2. Y. S. Sun, J. P. Landry, Y. Y. Fei, C. D. Zhu, J. T. Luo, X. B. Wang, and K. S. Lam, *Langmuir* **24**, 13399 (2008).

3. Y. Cui, Q. Q. Wei, H. K. Park, and C. M. Lieber, *Science* **293**, 1289 (2001); Y. L. Bunimovich, Y. S. Shin, W.-S. Yeo, M. Amori, G. Kwong, and J. R. Heath, *J. Am. Chem. Soc.* **128**, 16323 (2006).
4. J. Lee, J. Jang, D. Akin, C. A. Savran, and R. Bashir, *Applied Physics Letters* **93**, 013901 (2008); J. Fritz, M. K. Baller, H. P. Lang, H. Rothuizen, P. Vettiger, E. Meyer, H. Güntherodt, C. Gerber, and J. K. Gimzewski, *Science* **288**, 316 (2000).
5. S. M. Borisov and O. S. Wolfbeis, *Chem. Rev.* **108**, 423 (2008); J. Homola, *Anal. Bioanl.Chem.* **377**, 528 (2003).
6. E. Stern, R. Wagner, F. J. Sigworth, R. Breaker, T. M. Fahmy, and M. R. Reed, *Nano Lett.* **7**, 3405 (2007).
7. H. Altug, D. Englund, and J. Vuckovic, *Nature Physics* **2**, 484 (2006).
8. H. Altug and J. Vuckovic, *Optics Letters* **30**, 982 (2005); M. R. Lee and P. Fauchet, *Optics Express* **15**, 4530 (2007); E. Chow, A. Grot, L. W. Mirkarimi, M. Sigalas, and G. Girolami, *Optics Letters* **29**, 1093 (2004).
9. N. Skivesen, A. Têtu, M. Kristensen, L. H. Frandsen, and P. I. Borel, *Optics Express* **15**, 3169 (2007); S. Xiao and N. A. Mortensen, *Journal of Optics A: Pure and Applied Optics* **9**, S463 (2007).
10. I. D. Block, N. Ganesh, M. Lu, and B. T. Cunningham, *IEEE Sensors* **8**, 274 (2008); O. Levi, M. M. Lee, J. Zhang, S. R. J. Brueck, S. Fan, and J. S. Harris, *Proc. of SPIE.* **6447**, 64470P (2007).
11. F. Vollmer, D. Braun, A. Libchaber, M. Khoshsima, I. Teraoka, and S. Arnold, *Applied Physics Letters* **80**, 4057 (2002); I. M. Zhu, J. D. White, P. Suter, S. Dale, and X. Fan, *Optics Express* **15**, 9139 (2007).
12. R. Yan, S. P. Mestas, G. Yuan, R. Safaisini, D. S. Dandy, and K. L. Lear, *Lab on a Chip* **9**, 2163 (2009).
13. A. M. Armani *et al.*, *Science* **317**, 783 (2007).
14. M. Cai, O. Painter, and K. J. Vahala, *Phys. Rev. Lett.* **85**, 74 (2000).
15. S. Lal, S. Link, and N. J. Halas, *Nature Photonics* **1**, 641 (2007).
16. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings. Springer Tracts in Modern Physics 111*, (Springer-Verlag, New York, 1988).
17. S. Maier, *Plasmonics: Fundamentals and Applications* (Springer, New York, 2007).
18. H. Atwater, *Scientific American*, **April Issue**, 55 (2007).
19. T. W. Odom and C. L. Nehl. *ACS Nano* **2**, 612 (2008).
20. H. Atwater, *Journal of Applied Physics* **98**, 011101 (2005).
21. N. Xia and N. J. Halas, *MRS Bulletin* **30**, 338 (2005).
22. K. Li, M. I. Stockman, and D. J. Bergman, *Phys. Rev. Lett.* **91**, 227402 (2003).
23. J. P. Camden, J. Dieringer, J. Zhao, and R. P. Van Duyne, *Accounts of Chemistry Research* **41**, 1653 (2008).
24. S. K. Gray, *Plasmonics* **2**, 143 (2007).
25. S. Lee, K. M. Mayer, and J. F. Hafner, *Analytical Chemistry* **81**, 4450 (2009); A. J. Haes, W. P. Hall, L. Chang, W. L. Klein, and R. P. Van Duyne, *Nano Letters* **4**, 1029 (2004); P. K. Jain, X. Huang, I. H. El-Sayed, and M. A. El-Sayed, *Plasmonics* **2**, 107 (2007).

26. T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolf, *Nature* **391**, 667 (1998).
27. C. Genet and T. W. Ebbesen, *Nature* **445**, 39 (2007).
28. H. A. Bethe, *Physics Review* **66**, 163 (1944).
29. W. L. Barnes, W. A. Murray, J. Dintinger, E. Devaux, and T. W. Ebbesen, *Phys. Rev. Lett.* **92**, 107401 (2004); H. J. Lezec and T. Thio, *Optics Express* **12**, 3629 (2004); A. Degiron and T. W. Ebbesen, *Optics Express* **12**, 3694 (2004); G. Gay, O. Alloschery, B. Lesegno, C. O'Dwyer, J. Weiner, and H. J. Lezec, *Nature Physics* **2**, 262 (2006); P. Lalanne and J. P. Hugonin, *Nature Physics* **2**, 551 (2006).
30. H. Liu and P. Lalanne, *Nature* **452**, 728 (2008).
31. A. A. Yanik, X. Wang, S. Erramilli, M. K. Hong, and H. Altug, *Applied Physics Letters* **93**, 081104 (2008); A. Ali Yanik, R. Adato, S. Erramilli, and H. Altug, *Optics Express* **17**, 20900 (2009).
32. Y. Liu, J. Bishop, L. Williams, S. Blair, and J. Herron, *Nanotechnology* **15**, 1368 (2004); P. R. H. Stark, A. E. Halleck, and D. N. Larson, *Methods* **37**, 37 (2005); K. A. Tetz, L. Pang, and Y. Fainman, *Optics Letters* **31**, 1528 (2006); J. Dintinger, S. Klein, and T. W. Ebbesen, *Advanced Materials* **18**, 1267 (2006); T. Rindzevicius, Y. Alaveryan, A. Dahlin, F. Hook, D. Sutherland, and M. Kall, *Nano Letters* **5**, 2335 (2005); J. C. Sharpe, J. Mitchell, L. Lin, N. Sedoglavich, and R. J. Blaikie, *Analytical Chemistry* **80**, 2244 (2008); A. Lesuffleur, H. Im, N. Lindquist, K. S. Lim, and S. Oh, *Optics Express* **16**, 219 (2008); A. Artar, A. Ali Yanik, and H. Altug, *Appl. Phys. Lett.* **95**, 051105 (2009).
33. J. Homola, S. S. Yee, and G. Gaauglitz, *Sens. Actuators B Chem.* **54**, 3 (1999).
34. I. M. White and X. Fan, *Opt. Express* **16**, 1020 (2008).
35. F. Wang and Y. R. Shen, *Phys. Rev. Lett.* **97**, 206806 (2006).
36. K. S. Lee and M. A. El-Sayed, *J. Phys. Chem. B* **110**, 19220 (2006).
37. T. M. Squires, R. J. Messinger, and S. R. Manalis, *Nature Biotechnology* **26**, 417 (2008).
38. R. Adato, A. A. Yanik, J. J. Amsden, D. L. Kaplan, F. G. Omenetto, M. K. Hong, S. Erramilli, and H. Altug, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 19227 (2009).
39. S. Aksu, A. Yanik, R. Adato, A. Artar, M. Huang, and H. Altug, *Nano Letters* **10**, 2511 (2010).
40. A. A. Yanik, M. Huang, A. Artar, T. Y. Chang, and H. Altug, *Appl. Phys. Lett.* **96**, 021101 (2010).
41. F. Siebert and P. Hildebrandt, *Vibrational Spectroscopy in Life Science* (Wiley-VCH, New York, 2007).
42. J. L. McHale, *Molecular Spectroscopy* (Prentice Hall, New Jersey, 1999).
43. J. Kundu, F. Le, P. Nordlander, and N. J. Halas, *Chemical Physics Letters* **452**, 115 (2008); F. Neubrech, A. Pucci, T. W. Cornelius, S. Karim, A. Garcia-Extarri, and J. Aizpurua, *Physical Review Letters* **101**, 157403 (2008); R. Bukasov and J. S. Shumaker-Parry, *Analytical Chemistry* **81**, 4531 (2009); S. M. Williams, A. D. Stafford, K. R. Rodriguez, T. M. Rogers, and J. V. Coe, *J. Phys. Chem. B* **107**, 11871 (2003).

44. K. Willets and R. P. Van Dyne, *Annu. Rev. Phys. Chem.* **58**, 267 (2006).
45. M. Osawa, In *Handbook of Vibrational Spectroscopy.* (JM Chalmers and PR Griffiths, eds.) 785, (Wiley, Chichester).
46. T. R. Jensen, R. P. Van Duyne, S. A. Johnson, and V. A. Maroni, *Applied Spectroscopy* **54**, 371 (2000); A. Pucci, *Physica Status Solidi B* **242**, 2704 (2005).
47. K. Ataka and J. Heberle, *Analytical Bioanal. Chemistry* **388**, 47 (2007).
48. M. Meier, A. Wokaun, and P. F. Liao, *Journal of the Optical Society of America B* **2**, 931 (1985).
49. S. Zou, N. Janel, and G. C. Schatz, *Journal of Chemical Physics* **120**, 10871 (2004).
50. S. Zou and G. C. Schatz, *Chemical Physics Letters*, 62 (2005); Y. Chu, E. Schonbrun, T. Yang, and K. B. Crozier, *Applied Physics Letters* **93**, 181108 (2008).
51. R. Adato, A. A. Yanik, C.-H. Wu, G. Shvets, and H. Altug, *Optics Express* **18**, 4526 (2010).
52. J. M. Montgomery, A. Imre A, U. Welp, V. V. Vlasov, and S. K. Gray, *Optics Express* **17**, 8669 (2009); A. S. Grimault, A. Vial, and A. L. de la Chapelle, *Applied Physics B: Lasers and Optics* **84**, 111 (2005); Y. A. Urzhumov and G. Shvets, *Proc. of SPIE* **5927**, 59271D (2005).
53. B. Lamprecht *et al.*, *Phys. Rev. Lett.* **84** (2000).
54. V. A. Markel, *J. Phys. B: At. Mol. Opt. Phys.* **38**, L115 (2005).
55. B. Auguié and W. L. Barnes, *Phys. Rev. Lett.* **101**, 143902 (2008).
56. E. D. Palik, ed. *Handbook of Optical Constants of Solids II* (Academic, Orlando, Fla., 1985).
57. M. Meier and A. Wokaun, *Opt. Lett.* **8**, 581 (1983).
58. V. G. Kravets, F. Schedin, and A. N. Grigorenko, *Phys. Rev. Lett.* **101**, 087403 (2008).
59. A. Alu and N. Engheta, *Nature Photonics* **2**, 307 (2008).
60. M. Schnell, A. Garcia-Etxarri, A. J. Huber, K. Crozier, J. Aizpurua, and R. Hillenbrand, *Nature Photonics* **3**, 287 (2009).
61. J. O. Warwicker, *Acta Crystal.* **7**, 565 (1954).
62. E. S. Sashina, A. M. Bocheck, N. P. Novoselov, and D. A. Kirichenko, *Russian Journal of Applied Chemistry* **79**, 869 (2006).
63. A. J. Boltasseva, *Opt. A: Pure Appl. Opt.* **11**, 114001 (2009).
64. J. Prikulis, P. Hanarp, L. Olofsson, D. Sutherland, and M. Kall, *Nano Lett.* **4**, 1003 (2004).
65. C. L. Haynes and R. P. Van Duyne, *J. Phys. Chem. B* **105**, 5599 (2001).
66. S. Aksu and H. Altug, *Mater. Res. Soc. Symp. Proc.* 1208E (2010).
67. J. Yao, A.-P. Le, S. K. Gray, J. S. Moore, J. A. Rogers, and R. G. Nuzzo, *Advanced Materials* **22**, 1102 (2010).
68. L. Gross, R. R. Schlittler, G. Meyer, A. Vanhaverbeke, and R. Allenspach, *Appl. Phys. Lett.* **90**, 093121 (2007).
69. O. Vazquez-Mena, G. Villanueva, V. Savu, K. Sidler, M. A. F. van den Boogaart, and J. Brugger, *Nano Lett.* **8**, 3675 (2008).

70. X. M. Yan, A. M. Contreras, M. M. Koebel, J. A. Liddle, and G. A. Somorjai, *Nano Lett.* **5**, 1129 (2005).
71. M. Huang, A. A. Yanik, T. Y. Chang, and H. Altug, *Optic Express* **17**, 24224 (2009).
72. A. N. Shipway, E. Katz, and I. Willner, *Phys. Chem.* **1**, 18 (2000).
73. R. Raiteria, M. Grattarola, and R. Berge, *Materials Today* **5**, 22 (2002).
74. D. Erickson, S. Manda, Allen H. J. Yang, and B. Cordovez, *Microfluidics and Nanofluidics* **4**, 33 (2007).
75. P. S. Waggoner and H. G. Craighead, *Lab on a Chip* **7**, 1238 (2007).
76. C. Monat, P. Domachuk, and B. J. Eggleton, *Nature Photonics* **1**, 106 (2007).
77. P. E. Sheehan and L. J. Whitman, *Nano Lett.* **5**, 803 (2005).
78. S. K. Yoon, G. W. Fichtl, and P. J. A. Kenis, *Lab Chip* **6**, 1516 (2006).
79. R. A. Vijayendran, K. M. Motsegood, D. J. Beebe, and D. E. Leckband, *Langmuir* **19**, 1824 (2003).

LONG-RANGE SURFACE PLASMON POLARITON WAVEGUIDES AND DEVICES

Kristjan Leosson[*,‖], Malte C. Gather[†], Petur G. Hermannsson[*], and Alexandra Boltasseava[‡,§,¶]

[*]*Department of Physics, Science Institute, University of Iceland, Dunhagi 3, IS107 Reykjavik, Iceland*

[†]*Wellman Center, Harvard Medical School, 40 Blossom Street, Boston, Massachusetts 02114, USA*

[‡]*DTU Fotonik — Department of Photonics Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*

[§]*School of Electrical and Computer Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette IN 47907, USA*

[¶]*Erlangen Graduate School in Advanced Optical Technologies (SAOT), Friedrich-Universität Erlangen-Nürnberg, 91052 Erlangen, Germany*

[‖]*kleos@hi.is*

Long-range surface plasmon polariton waveguides form a class of plasmonic waveguide geometries characterized by low propagation loss and weak confinement, relative to other metal-dielectric waveguides. These properties allow for fabrication of a range of plasmonic devices that are compatible with conventional fiber optics while possessing several unique features, including the possibility of strong index modulation, large evanescent-field volumes for optical sensing purposes, and control of waveguide properties by passing electrical current through the waveguide core. Moreover, the low-loss waveguide geometry offers a platform for realizing plasmonic amplification, using conventional organic gain materials like conjugated fluorescent polymers. Here, we describe fabrication and optical characterization of long-range surface plasmon waveguides and devices based on thin metallic stripes or

metallic nanowires embedded in a polymer matrix, including thermo-optically controlled interferometric devices, compact extinction modulators with low polarization-dependent loss and optically pumped plasmonic amplifiers.

1. Introduction

The controlled excitation and manipulation of surface plasmon polaritons (SPPs) on metal-dielectric interfaces[1] is a topic which has attracted great attention within the fields of nano-optics, bio-optics, photonics and photovoltaics during the past decade.[2–5] SPPs represent quasi-two-dimensional transverse-magnetic (TM) excitations with maximum field amplitudes at the metal-dielectric interface, decaying exponentially into both neighboring media.[1] Tight field intensity confinement to the metal surface (typically on the order of or smaller than the wavelength in the corresponding media) makes SPPs very sensitive to surface irregularities and/or local refractive index changes. Hence, SPPs have been used for surface analysis, mainly using the so-called surface plasmon resonance (SPR) biosensing technique.[6] The same property makes manipulation of SPPs in the surface plane possible on nanostructured surfaces and efficient control of SPP propagation has been achieved using different surface configurations.[1,2,7–12] Surface plasmons on nanostructured surfaces can exhibit a variety of novel effects, including extraordinary light transmission and giant field enhancement as well as subwavelength waveguiding. However, when compared to conventional dielectric waveguides used in integrated optics, SPPs suffer from strong damping due to absorption in metals (ohmic loss) and their propagation length is usually limited to distances of tens or hundreds of micrometers in the visible and near-infrared wavelength range.

One particular type of metal-dielectric structure for guiding and manipulation of surface plasmon polaritons that has been the focus of recent research is the dielectric-metal-dielectric structure supporting so-called *long-range surface plasmon polariton* (LRSPP) modes.[13,14] LRSPP propagation on thin metal films of finite width embedded in a uniform dielectric [Fig. 1(a)] was originally demonstrated and theoretically analyzed by Berini and coworkers,[15–17] forming the basis for a range of plasmonic devices characterized by low field confinement (similar to that achievable by conventional low-index-contrast dielectric waveguides) and relatively large propagation lengths compared to other plasmonic waveguide geometries.[18] The symmetric waveguide structure supports propagation of coupled SPPs

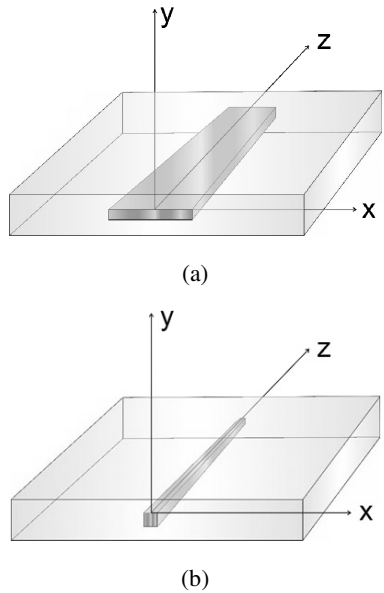


Fig. 1. The two main waveguide architectures discussed in this paper: (a) thin metal stripe and (b) metal nanowire, embedded in a dielectric medium.

at the two metal-dielectric interfaces resulting in a small field concentration within the metal film and, hence, low propagation loss. Long-range modes can also be formed from coupled plasmonic corner modes (also known as wedge plasmon polaritons[19–22]) in metallic nanowire waveguides [Fig. 1(b)]. Nanowires with a square-symmetric cross-sectional profile are of particular interest due to their polarization-independent waveguiding properties.[23,24] The geometry of metal stripe waveguides and nanowire waveguides can be tailored to make them compatible with common fiber-optic standards used, e.g., in telecom applications.[25,26] A comprehensive review on LRSPP waveguides and devices was recently published by Berini.[18]

In most cases, losses due to absorption in the metal present a serious problem, limiting the application of plasmonic devices. It has long been recognized that these losses can only be mitigated by introducing optical gain at the interface supporting SPP propagation.[27] Numerous recent theoretical and experimental studies have focused on incorporating optical gain into a dielectric material adjacent to a metal surface, in order to reduce or overcome absorption losses, realize plasmonic lasers, lossless metamaterials and to further our understanding of spontaneous and

stimulated emission from excited dipoles in the vicinity of a metal surface or metallic nanoparticles.[28–44] This represents an exciting new area of research in plasmonics and light-matter interaction in general.

2. Dispersion Relations, Field Confinement and Propagation Loss

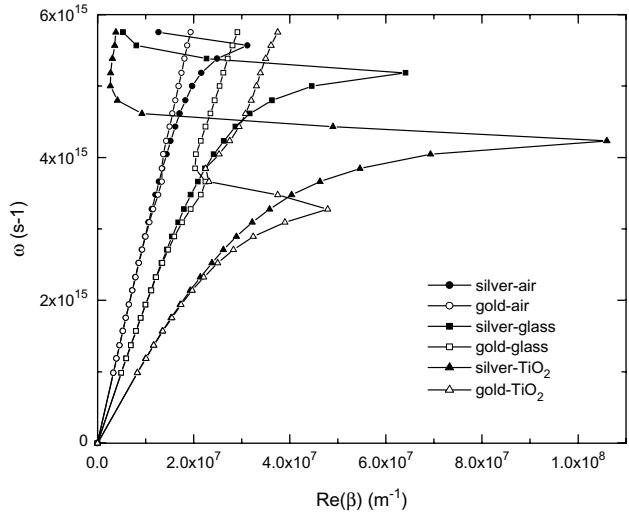
In this section, dispersion relations and related properties of SPPs and LRSPPs will be discussed. Instead of adopting the common approach of describing SPPs at the interface between an ideal (Drude) metal and a dielectric with a purely real refractive index, we will focus our discussion on real and commonly used materials, including also the use of dielectrics with optical gain (negative imaginary component of the dielectric constant).

As a starting point for our discussion, we consider the well-known dispersion relation for surface plasmon polaritons at the interface between two materials with complex dielectric constants $\hat{\varepsilon}_1$ and $\hat{\varepsilon}_2$

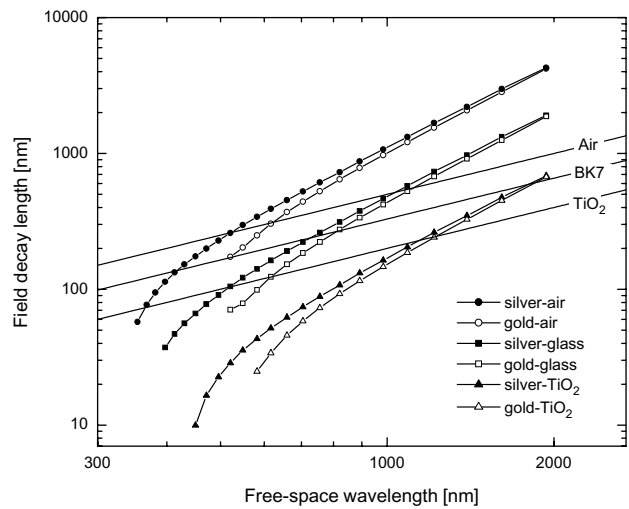
$$\beta = \frac{\omega}{c}\sqrt{\frac{\hat{\varepsilon}_1\hat{\varepsilon}_2}{\hat{\varepsilon}_1 + \hat{\varepsilon}_2}} \tag{1}$$

The dependence of the complex propagation constant β on the frequency of the electromagnetic radiation ω, contains all relevant information on the confinement, wavelength, phase velocity and propagation loss or gain of the guided wave.[45] Aside from requiring that an oscillation of electric charge density can occur at the interface, we make no assumptions at this point about the expressions for the complex dielectric functions of the two materials. We note only that in the case of surface-confined waves (field amplitudes exponentially decaying away from the interface) where the denominator becomes small ($|\hat{\varepsilon}_1 + \hat{\varepsilon}_2| \ll |\hat{\varepsilon}_1\hat{\varepsilon}_2|$), the norm of the complex propagation constant $\beta = \beta' + i\beta''$ increases, with an associated decrease of the group velocity $\nu_g \equiv \delta\omega/\delta\beta'$. This property is characteristic of "plasmonic" behavior and defines the surface plasma frequency of an ideal (Drude) metal,[46] where $\nu_g \to 0$.

Chemical resistance makes gold a particularly attractive metal for applications in plasmonics, since surface plasmons are, by nature, highly sensitive to structural or chemical modifications of the metal surface. Gold, however, deviates significantly from the behavior predicted by the simple Drude model of an ideal metal due to strong interband transitions occurring for photon energies $> 1.8\,\text{eV}$ (free-space wavelength $< 700\,\text{nm}$). Figure 2(a) shows examples of dispersion curves on several metal-dielectric



(a)



(b)

Fig. 2. (a) Dispersion curves for surface plasmon polaritons on different metal-dielectric interfaces ($n \approx 1$, $n \approx 1.5$, $n \approx 2.5$), illustrating the different degrees of "plasmonic" behavior for different material systems. (b) Confinement of the SPP field for the same metal-dielectric interfaces (lines with symbols), compared to the classical diffraction limit in the respective dielectric materials (solid lines).

interfaces. Experimentally determined values for the dielectric functions of gold and silver are taken from the work of Johnson and Christy.[47] When considering only the real part of the propagation constant, the dispersion of a surface plasmon polariton supported by a gold-air interface does not differ significantly from photonic dispersion. Increasing the refractive index of the dielectric (e.g., by replacing air with glass or a higher-index material like TiO_2) makes the SPP on the gold surface more "plasmonic" in the sense that the real part of the propagation constant can become considerably larger than that of the dielectric material alone. Silver, however, exhibits more plasmon-like dispersion. As an example, for the silver-TiO_2 interface,[a] the dispersion curve indicates that the SPP wavelength becomes approximately 60 nm at resonance (1/3 of the wavelength in TiO_2) and that the group velocity of the surface wave falls below $0.02c$. At the resonance wavelength, however, the traveling wave is overdamped and has a propagation length smaller than the SPP wavelength.

The interest in plasmonics has, to a large extent, been motivated by the possibility of confining optical energy to a volume smaller than that defined by Abbe's diffraction limit in dielectric materials, given by $\lambda/2n$ where λ is the free-space wavelength.[4] The strong confinement of the electromagnetic field at a metal-dielectric interface is associated with the surface plasmon resonance, as illustrated in Fig. 2(b) where the $1/e$ SPP field decay length into the dielectric material is compared to the diffraction limit in the corresponding material. For large wavelengths, the extent of the SPP field actually *exceeds* the diffraction limit, but the field decay length decreases substantially with increasing optical frequency. As the "plasmonic" character of the metal-dielectric interface becomes more pronounced, the field becomes more confined, e.g., close to one order of magnitude below the diffraction limit for a silver-TiO_2 interface at a free-space wavelength of 450 nm. Conversely, the smallest extent of the SPP field on a gold-air interface is only about 30% below the diffraction limit.

A serious challenge in the study and application of propagating plasmons with sub-diffraction-limited confinement is the propagation loss of the SPP, which in all cases rises dramatically as the optical frequency approaches the surface plasmon resonance. For all the cases plotted in Fig. 2(b), the SPP propagation length is below 1 μm at maximum confinement, increasing at most up to a few tens of μm when SPP confinement

[a]The refractive index of the rutile phase of TiO_2 ($n \approx 2.5$ in the visible wavelength range) was used for the purposes of this calculation.

is similar to the diffraction limit. This does not compare favorably to, e.g., silicon photonics where highly integrated devices with near-diffraction-limited mode sizes of around 200 nm (at 1.5-μm free-space wavelength) and propagation lengths of the order of centimeters have been realized.[48–50] Numerous different sub-diffraction-limited plasmonic waveguide geometries have been studied in order to improve the tradeoff between SPP field confinement and propagation loss.[4] Although none of these designs are, at present, likely to rival silicon waveguides as a platform for very large scale integrated photonics, highly concentrated optical fields are nevertheless of substantial technological interest, e.g., for sub-diffraction-limited lithography, plasmon-assisted magnetic recording, catalysis, broadband photovoltaics, optical nanoprobes, ultrafast switching and enhancement of non-linear effects.[5]

Multiple metal-dielectric interfaces (or metal wedges) can support SPP supermodes which may be regarded as coupled single-interface (or single-wedge) SPPs. In the simplest case of a planar three-layer structure (e.g., metal-dielectric-metal or dielectric-metal-dielectric), the dispersion relation for the coupled modes can be derived from the requirement of continuity of the tangential components of the **E** and **B** fields at the two interfaces (assuming non-magnetic media with $\mu_i = \mu_0$), yielding:[46]

$$e^{-2k_2 d} = \frac{k_2/\hat{\varepsilon}_2 + k_1/\hat{\varepsilon}_1}{k_2/\hat{\varepsilon}_2 - k_1/\hat{\varepsilon}_1} \cdot \frac{k_2/\hat{\varepsilon}_2 + k_3/\hat{\varepsilon}_3}{k_2/\hat{\varepsilon}_2 - k_3/\hat{\varepsilon}_3} \tag{2}$$

where d is the thickness of the middle layer and layers 1 and 3 are considered semi-infinite. The components of the wave vectors parallel to the interface are given by $k_i^2 = \beta^2 - \omega^2 \mu_0 \hat{\varepsilon}_i$ and Eq. (2) therefore provides an expression for the dispersion relation for the coupled SPPs which can be evaluated numerically. For the case of a perfectly symmetric structure with $\hat{\varepsilon}_1 = \hat{\varepsilon}_3$, Eq. (2) becomes

$$e^{-k_2 d} = \pm \frac{k_2/\hat{\varepsilon}_2 + k_1/\hat{\varepsilon}_1}{k_2/\hat{\varepsilon}_2 - k_1/\hat{\varepsilon}_1} \tag{3}$$

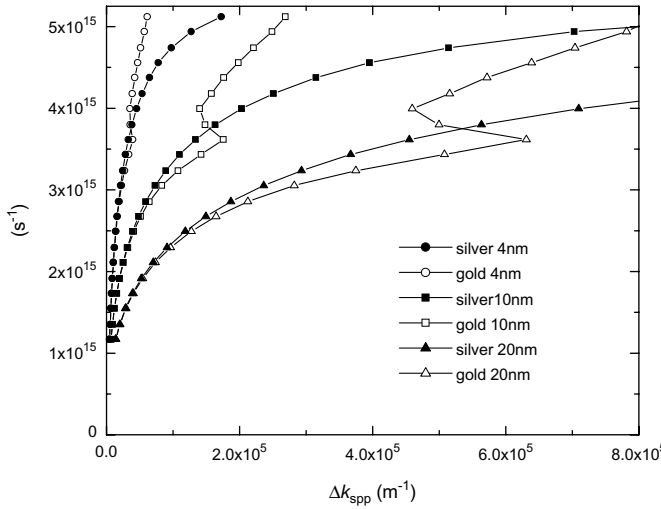
and has bound (non-radiative) solutions corresponding to symmetric and antisymmetric $B_x(z)$ field distributions with respect to the middle layer. For the case of a metal-dielectric-metal structure, one of the solutions does not exhibit cutoff as the dielectric thickness is decreased,[46] giving stronger confinement and larger wave vectors than the single metal-dielectric interfaces considered in Figs. 2(a) and 2(b). For the case of a symmetric dielectric-metal-dielectric structure, the solutions of Eq. (3)

correspond to long-range and short-range SPP supermodes. As the metal layer is made thinner, the long-range mode becomes less confined and its propagation loss decreases, whereas the opposite is true for the short-range mode.

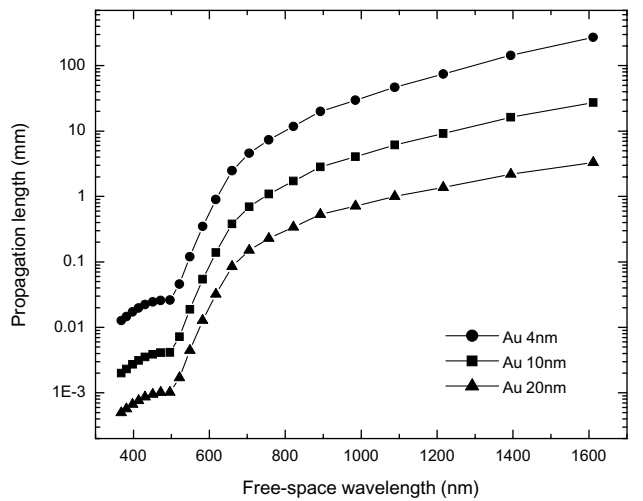
Figure 3(a) shows the dispersion relations corresponding to LRSPP modes on thin gold and silver films of different thicknesses, embedded in a homogeneous dielectric with a purely real refractive index of around 1.55, typical of several optical polymers. Film thicknesses down to 4 nm have been plotted, although it should be noted that fabricating continuous metal films with a thickness below 10 nm is not necessarily straightforward, as discussed in Section 3 below. Also, describing metal films with a thickness approaching 10 atomic layers with a bulk dielectric constant can only be considered an approximation, but a detailed investigation of its validity is outside the scope of the present paper. For clarity, Fig. 3(a) shows only the additional component of the wave vector Δk_{spp} introduced by the metal film (compared to the dielectric material alone), as the dispersion lies very close to the light line in the dielectric. Bound solutions may exist for asymmetric dielectric-metal-dielectric structures as well, as described in detail in Ref. (51). Weakly bound LRSPP modes can only tolerate a small asymmetry but as the additional wave vector component provided by the metal film increases (due to frequency, film thickness or the value of the complex permittivity of the metal), the LRSPP mode becomes more robust. Tunable asymmetry in cladding index has been used, e.g., as a means of introducing extra loss (leaky modes) for modulating LRSPP transmission in temperature-controlled glass-metal-polymer structures.[52]

As in the case of SPPs on a single metal-dielectric interface, the LRSPP propagation distance drops dramatically as the plasmon resonance condition is approached, as shown in Fig. 3(b). In the case of gold films, the LRSPP propagation distance drops nearly three orders of magnitude as the free-space wavelength decreases from 800 nm to 500 nm, e.g., from approximately 10 mm to 25 μm in the case of a 4-nm thick gold film, while the $1/e$ field decay length (half-$1/e^2$ mode-size) decreases from 3 μm to 1 μm, respectively.

Due to their relatively low propagation loss, LRSPP waveguides present a suitable geometry for incorporating optical gain into a plasmonic waveguide structure. It has been shown experimentally that conventional gain materials like conjugated fluorescent polymers[44] or dye solutions[41] can provide sufficient gain to fully overcome absorption losses in LRSPP waveguides, while SPP geometries require orders of magnitude larger modal



(a)



(b)

Fig. 3. (a) Dispersion curves for long-range surface plasmon polariton (LRSPP) waveguides consisting of thin gold or silver films embedded in a homogeneous dielectric with a refractive index corresponding to common optical polymers. For clarity, only the component of the wave vector introduced by the metal film is plotted. (b) Calculated propagation lengths for LRSPPs on Au films of different thickness.

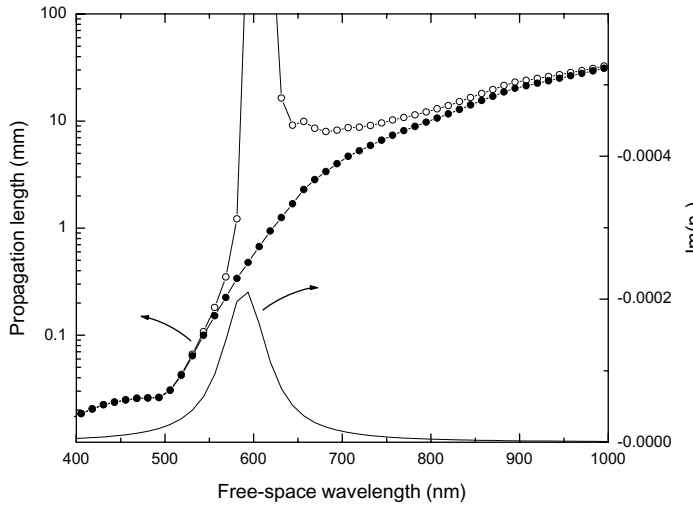


Fig. 4. Calculated propagation length for a 4-nm thick Au film in a passive dielectric (solid circles) and an asymmetric structure (open circles) where the dielectric on one side of the film has a negative imaginary component of the complex refractive index peaking around 600 nm (solid line).

gain to realize significant loss compensation. Figure 4 shows the value of the LRSPP propagation distance in a polymer-gold-polymer structure, calculated from the complex propagation constant given by Eq. (2). One of the cladding polymers (both semi-infinite for the purposes of this calculation) is assumed to exhibit optical gain around 600 nm, represented by the hypothetical gain profile shown in the figure, resembling a typical fluorescence spectrum from a conjugated polymer blend.[44] As a first approximation, the real part of the effective index of the two polymer layers is assumed to be equal, although a modification of the imaginary part of the refractive index must necessarily be reflected in a change in the real part, via the Kramers-Kronig relations. In the calculation shown in Fig. 4, Im(n_3) was scaled to give peak LRSPP gain of about 1 cm^{-1}. A thin (250 nm) layer of polymer with the same complex refractive index, coated on a low-index substrate would exhibit optical gain of about 25 cm^{-1} at the maximum gain wavelength of 600 nm, which is well within the range measured for conjugated fluorescent polymers.[53] Optical gain might, however, be substantially enhanced near a metal surface, as compared to the purely dielectric case, due to the high local field intensity[42] and increased photonic density of states associated with the SPP.

3. Fabrication

In this section, we describe the fabrication process for LRSPP waveguides and discuss different methods for nanoscale patterning of such waveguides. We consider both metal stripes and nanowires embedded in a homogeneous dielectric environment consisting of optical polymers, as well as ultra-thin metal films with a dielectric cladding containing a layer of fluorescent polymer for providing optical gain.

3.1. *Metal stripe LRSPP waveguides*

Planar plasmonic waveguides consisting of a nanometer-thin metal film embedded in a transparent dielectric can be fabricated using standard wafer-scale photolithographic processes.[54,55] The dielectric-metal-dielectric sandwich structure can be realized, e.g., by wafer bonding of similar materials, by polymer coating of glass wafers (where symmetry in refractive index is obtained at a specific temperature) or by CVD glass deposition or spin-coating of polymer or hybrid organic/inorganic cladding layers from solution. To ensure good lithographic quality of the metal waveguides, the bottom cladding needs to be resistant to photoresist solvent, developer and remover, a condition which is fulfilled by many thermoset or UV-cured optical polymers.

A commonly used LRSPP fabrication approach (Fig. 5) first reported by Nikolajsen *et al.*[56] involves spin-coating a silicon wafer with a layer of benzocyclobutane (BCB) resin (Cyclotene 3022, DOW Chemical Company). The lower cladding is soft-cured at 210°C which ensures improved adhesion to the top cladding layer. The bottom cladding layer is then

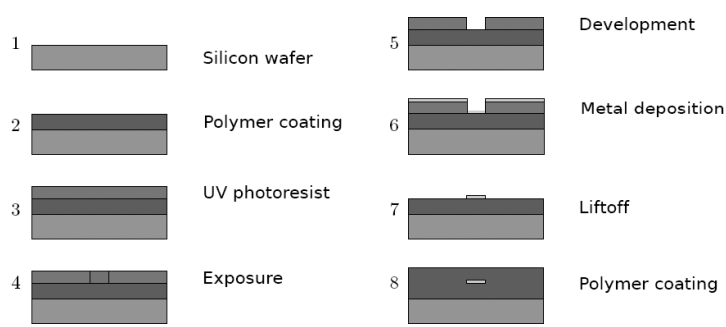


Fig. 5. Schematic illustrations showing the fabrication steps for LRSPP stripe waveguides.

coated with a layer of negative photoresist. Waveguides and passive devices are patterned using standard UV lithography and deposition of thermally evaporated or sputter-coated gold layer (typically 10–20 nm) and a subsequent lift-off process. For such thin layers, gold usually adheres well to the BCB surface, even withstanding ultrasonic treatment during lift-off. This is important since nm-thin layers of commonly used adhesion-promoting metals like Cr or Ti are very detrimental to waveguide performance. The waveguides are covered by a second layer of BCB, followed by a hard-cure of the entire structure at 250°C. Since the symmetry of the structure is very important for the LRSPP properties,[55] it was carefully controlled that the cladding layers had the same refractive index in spite of different thermal treatment. Also, the cladding layers were made thick enough to accommodate the electromagnetic field of the LRSPP and avoid leakage to the silicon substrate (typical total cladding thickness 25–30 μm).

Waveguides for signal transmission, including curved waveguides, splitters and directional couplers, can be fabricated directly by UV lithography. Optical circuits of very high structural quality can be fabricated this way, due to the fact that the core waveguide layer is essentially two-dimensional, as opposed to single-mode cores in conventional dielectric waveguides that typically have similar lateral and vertical dimensions. Structural quality is of importance, e.g., for equal power splitting for interferometric devices and precise control of mode coupling in closely spaced waveguides. LRSPP signals can be switched and modulated by using the waveguide stripes as heating elements in order to introduce a phase delay in interferometric components or achieving extinction by introducing mode leakage.[25,57,58] This involves an additional lithography and metal deposition step for defining electrical contacts, and a suitable process for exposing the contacts, such as reactive-ion etching or laser ablation.

Added functionality (such as wavelength filtering or grating-assisted coupling) can be added by nanostructuring LRSPP waveguides. Electron-beam lithography (EBL) has been used to fabricate reflective Bragg gratings[59] and photonic crystal structures[60,61] by symmetric or asymmetric thickness modulation of the LRSPP waveguides. Conventional UV lithography can be used to fabricate higher-order width-modulated LRSPP waveguides gratings in a single fabrication step.[62] Nanoimprint lithography (NIL) in combination with photolithography provides a means of high-throughput fabrication of nanostructured LRSPP waveguides, as first reported in Ref. (63) and illustrated in Fig. 6. Nanoimprint lithography is a particularly promising technique for high-throughput fabrication of nanostructured

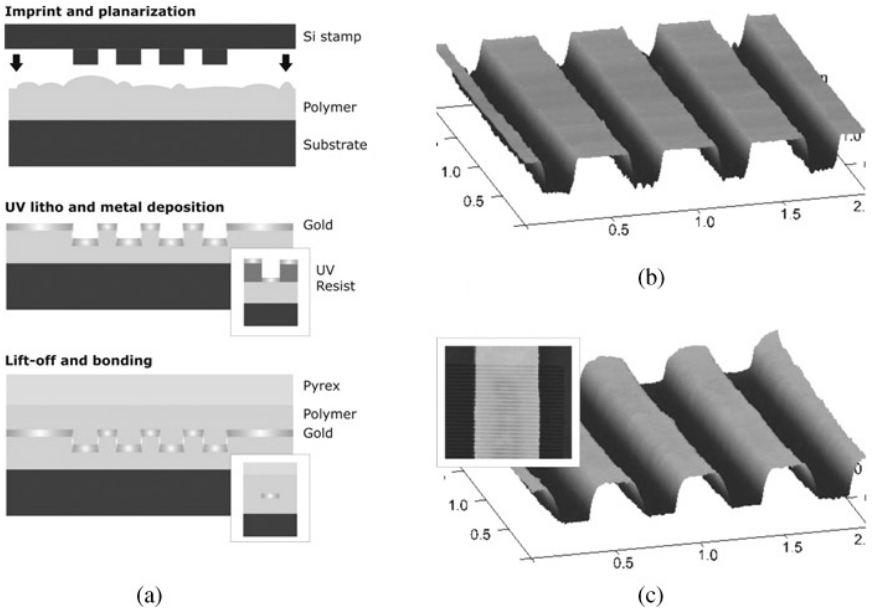
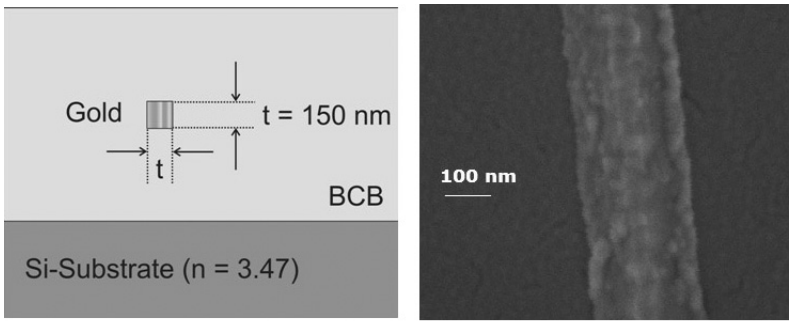


Fig. 6. (a) Schematic of a fabrication process for LRSPP waveguides with nanoimprinted gratings: Silicon stamp fabricated by electron-beam lithography and etching is pressed into NIL resist spun on a silicon substrate to define the grating structure. After thermal nanoimprint and separation, stripe waveguides are defined by UV lithography, gold deposition and lift-off (insets show the transverse cross-section of the UV-patterned waveguide). A symmetric top cladding is applied by thermal bonding to a glass substrate coated with the same NIL resist type. (b) Atomic force microscope (AFM) image of the grating on the silicon master stamp (ridge height is ∼27 nm). (c) AFM image of the imprinted grating, protrusion depth 28 nm [horizontal axes in (b) and (c) are in μm]. Inset shows an AFM image of an 8-μm-wide metal stripe deposited on the nanoimprinted polymer surface.

optical components.[64] In addition, the nanoimprint process simultaneously planarizes the polymer surface, resulting in a reduction of propagation losses associated with surface scattering. Nanoimprinted LRSPP gratings have been employed, e.g., for accurate refractive-index sensing in microfluidic channels.[65]

3.2. *Nanowire plasmonic waveguides*

Planar SPP or LRSPP waveguide geometries are polarization-selective, due to the fact that only TM-polarized waves can fulfill the required electromagnetic boundary conditions on a metal-dielectric interface.[1]



(a) (b)

Fig. 7. (a) Schematic of a plasmonic waveguide based on a symmetric nanowire suitable for transmitting wavelengths around 1550 nm. (b) SEM image of a fabricated 150-nm-wide gold nanowire on a polymer surface (coated with a thin layer of gold to avoid charge buildup during imaging).

The polarization properties of non-planar metal-dielectric geometries are, however, more complex. A cylindrical wire, for example, supports a radially polarized TM wave which becomes increasingly more confined as the radius of the cylinder is decreased.[66] A metal nanowire with a square cross-sectional profile supports four supermodes which can be constructed from proper combinations of individual corner modes.[24] This includes a radially polarized mode which can be focused without cutoff, similar to the cylindrical wire, as well as two degenerate long-range supermodes which become less confined as the nanowire cross section decreases. These long-range supermodes are transverse-magnetic modes (with respect to the propagation direction) with the main electrical field components pointing in the x and the y, respectively, directions and can therefore be combined to transmit an input signal at any polarization.

Plasmonic nanowire waveguides can be fabricated using a process similar to the stripe waveguides discussed in Section 3.1 above, except that the wire dimensions are below the limit of conventional i-line (365 nm) photolithography. Instead, we used EBL to define wires having widths of 100−200 nm (Fig. 7). We used standard e-beam resists such as PMMA or ZEP520 with a thickness about 30% larger than the intended wire thickness. Precise monitoring of the width of the exposed pattern and the actual wire thickness is required to obtain waveguide profiles with proper symmetry. It was found that a thin adhesion layer of Ti or Cr was necessary to maintain structural integrity of the gold nanowires during

liftoff, unlike the thin stripe waveguides discussed above. The thickness of this layer must be kept at a minimum, however, since numerical modeling indicates that even a 1-nm thick adhesion layer has a noticeable effect on the propagation loss and polarization dependent loss (PDL) in nanowire waveguides.[67] The PDL can be balanced by using a slightly asymmetric structure, at the cost of increased propagation loss. The cross-sectional profile of the fabricated nanowires was found to be slightly trapezoidal due to a shadowing effect (which has been used intentionally, e.g., for the fabrication of gold nanocones[68]). Numerical modeling indicates, however, that the resulting wire shape does not significantly affect the propagation characteristics.[69]

3.3. *Ultra-thin LRSPP waveguides with fluorescent polymer cladding*

Propagation losses in LRSPP waveguides can be reduced or fully overcome by embedding the waveguide in a dielectric medium that is doped with an organic dye or consists of an electroluminescent polymer.[44] If the organic medium is optically pumped at sufficiently high energy, amplification of the LRSPP mode propagating along a metal film can be observed, as a result of stimulated emission from the excited dipoles in the organic medium into the LRSPP mode. Naturally, net plasmonic gain can only be realized if the propagation loss of the structure does not exceed the gain provided by the organic medium. Most organic dyes and polymers emit in the visible range of the spectrum and become increasingly less efficient as the wavelength is increased beyond the red end of the spectrum. For gold LRSPP waveguides, it is therefore most suitable to work at wavelengths around 600 nm, being reasonably close to the plasmon resonance wavelength and providing high photoluminescence (PL) quantum efficiency of the optical material.

For fabrication of LRSPP waveguides with gain material, we selected a UV-curable hybrid organic/inorganic material (Ormocer, Micro Resist Technology GmbH) as the dielectric cladding. On one side of the LRSPP waveguide, a thin layer of this material was doped with an organic dye (pyrromethene 597 or 650, Exciton Inc.) or replaced by an index-matched layer of diluted fluorescent conjugated polymer [MDMO-PPV:PSF:HBO blend, for processing details see Ref. (44)]. The index matching of the gain layer to the underlying cladding is critical for ensuring low propagation loss and to avoid dielectric waveguiding in the gain layer. The thickness of the

gain layer must be chosen to provide sufficient modal gain for the LRSPP mode without introducing too much absorption at the pump wavelength which would prevent efficient excitation of the gain material close to the gold surface.

In order to realize LRSPP waveguides with propagation lengths up to 1 mm at 600 nm free-space wavelength, it is necessary to fabricate gold films down to about 4 nm thickness. With standard deposition techniques and substrates (e.g., glass or silicon), gold forms a discontinuous and thermally unstable film even at considerably larger thickness.[70] Ultra-thin gold deposited on the Ormocer surface, however, seems to give continuous films, as indicated by AFM measurements (Fig. 8), and visible-light LRSPP propagation over distances of the order of 100 μm on planar films and in stripe waveguides has been clearly observed.[44,71] Further work is being carried out in this area to determine the structural characteristics and thermal stability of such ultra-thin gold films.

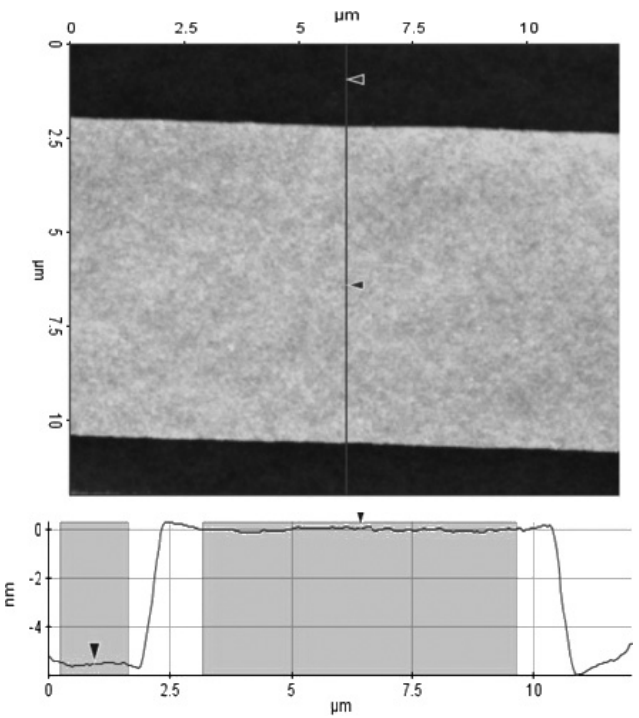


Fig. 8. AFM image (non-contact mode, top view and cross sectional profile) of a 5.5-nm thick and 8 μm wide gold stripe on an Ormocer-coated substrate.

4. Optical Properties

We consider in this section the experimentally determined properties of LRSPP waveguides and devices, ranging from basic issues of propagation loss and mode size in simple stripe waveguides, to nanostructured waveguides, thermo-optically controlled devices, mode coupling and finally the successful demonstration of plasmonic gain in an LRSPP waveguide.

4.1. *Waveguide properties*

Straight LRSPP waveguides consisting of 10–15 nm thick gold stripes embedded in an optical polymer (fabricated using the processing steps described in the previous section) show propagation losses close to theoretically predicted values over a wide range of wavelengths, as shown in Fig. 9. As the film thickness is decreased, scattering due to metal roughness and edge roughness becomes increasingly important. In addition, a finite cladding layer thickness will eventually pin the mode size and prevent a further decrease in propagation loss. By using sufficiently thick cladding layers and special lithographic processes for minimizing waveguide edge

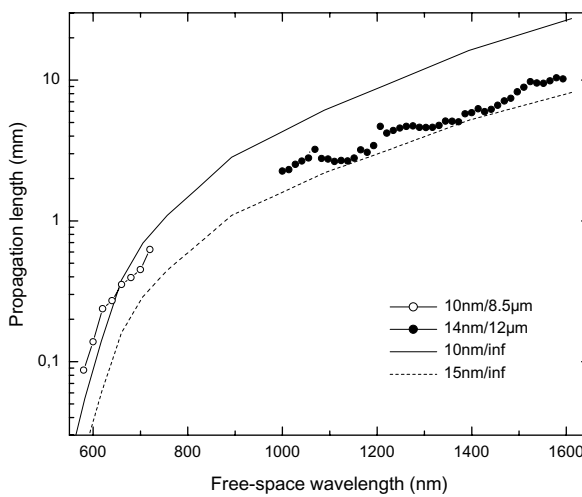


Fig. 9. Experimentally determined propagation lengths for wide gold stripe waveguides determined by cutback measurements (solid circles, 1000–1600 nm) or scattered light imaging (open circles, 580–720 nm). The thickness (width) of the measured waveguides was 14 nm (8.5 μm) and 10 nm (12 μm), respectively. Also shown are the calculated curves based on experimental values of the dielectric function of gold, calculated for planar (i.e., infinitely wide) films of 10 nm (solid line) and 15 nm (dashed line) thickness.

roughness, Park *et al.* achieved sub-dB/cm propagation loss (propagation length > 4.3 cm) in 11-nm thick silver stripe waveguides at a wavelength of 1310 nm.[72]

In practice, it is often advantageous to choose the dimensions of the metal stripe so that the LRSPP field distribution matches closely that of conventional single-mode optical fibers for a specific wavelength range. This facilitates efficient end-fire excitation and experimentally determined coupling losses below 1 dB per facet are reported in the literature.[54–56] Detailed characterization of propagation loss and coupling loss for a wide range of wavelengths (1000–1650 nm) and stripe widths can be found in Ref. (73). In the case of gold LRSPP waveguides for use in the telecom wavelength range 1460–1675 nm (S, C, L and U bands), a film thickness of about 15 nm provides a good tradeoff between confinement and propagation loss. While propagation loss is, in most cases, detrimental to device performance, LRSPP waveguides have the unique property that absorption loss provides a way to directly monitor the optical power carried by the waveguide.[74]

Strongly polarization-dependent optical devices, such as those based on LRSPP stripe waveguides, are limited to applications where the polarization state of the input light is well known, e.g., in laser modules or polarization-maintaining fiber systems. In a conventional fiber-optic network, however, the polarization state is generally unknown and varies with time. The development of polarization-insensitive plasmonic waveguides is therefore of fundamental as well as technological interest. As discussed in Section 3.2. above, square-symmetric nanowire waveguides supporting long-range supermodes can be fabricated using EBL. Figure 10(a) shows an example of the direction and magnitude of the electric field component perpendicular to the propagation direction for the y-polarized mode in a symmetric wire (by symmetry, the x-polarized mode can be obtained by 90° rotation). Figures 10(b) and 10(c) show the output facet of a nanowire waveguide transmitting two polarization of the input light, demonstrating that the measured mode size and transmitted intensity of the two modes are nearly identical. Depending on the exact application, gold wire edge lengths of 150–200 nm are the most suitable for guiding light around 1550 nm wavelength, where propagation lengths close to 10 mm have been demonstrated. The trade-off between coupling loss (to standard single-mode fibers) and propagation loss in straight waveguides is, however, optimal if the plasmonic waveguide is embedded inside a dielectric waveguide core, as described in Ref. (69).

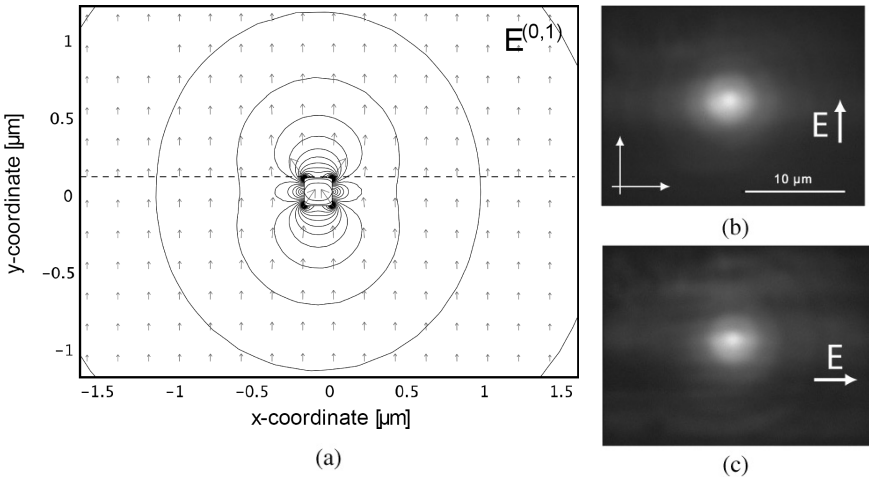


Fig. 10. (a) Magnitude and direction for the y-polarized long-range supermode of a metallic nanowire embedded in a homogeneous dielectric, determined by finite-element calculations. The x-polarized mode is obtained by 90° rotation. Measurements of (b) y-polarized and (c) x-polarized light at 1550 nm free-space wavelength, imaged at the output facet of an LRSPPs gold nanowire waveguide sample.

4.2. *LRSPP mode coupling*

Mode coupling is observed between closely spaced metal stripe waveguides (analogous to evanescent-wave coupling in dielectric waveguides) and parallel-stripe LRSPP directional couplers have been studied in some detail.[55,73,75,76] Due to the weak guiding, lateral coupling in LRSPP stripe waveguides can be modeled with suitable accuracy using a simple effective-index approach. Investigated directional couplers (DCs) consisted of two waveguides that approach each other by means of low-loss cosine s-bends, followed by a straight interaction region of variable length and coupling gap, separating again with identical cosine bends. Coupling lengths (corresponding to full power transfer from one waveguide to the other) from 0.8 mm to several mm can be achieved in DCs based on waveguides separated by a few micrometers.[73]

The coupling length of a DC is determined by the difference in the propagation constants of even and odd combinations of the LRSPP supermodes of each waveguide and is, thereby, wavelength dependent. LRSPP DCs with suitable interaction lengths can therefore be used, e.g., as simple add/drop multiplexers for two wavelength bands. As an example, Fig. 11 shows typical transmission spectra for the direct and coupled arms of

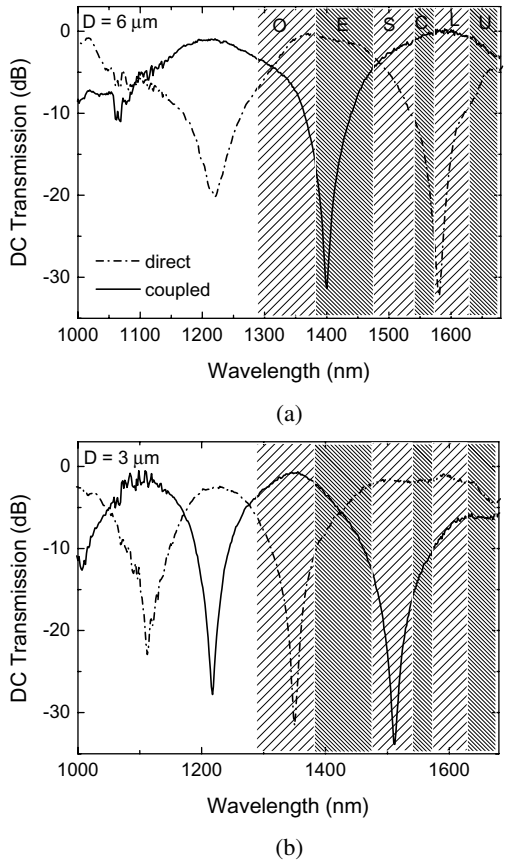
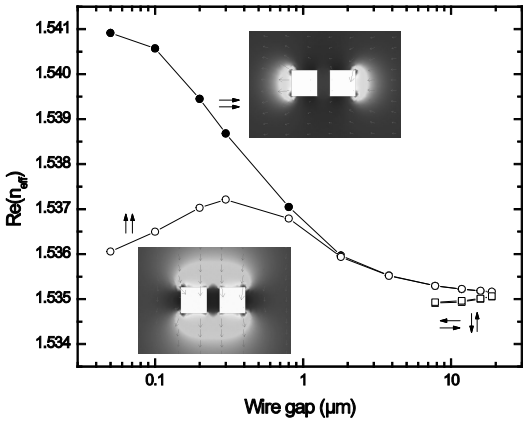


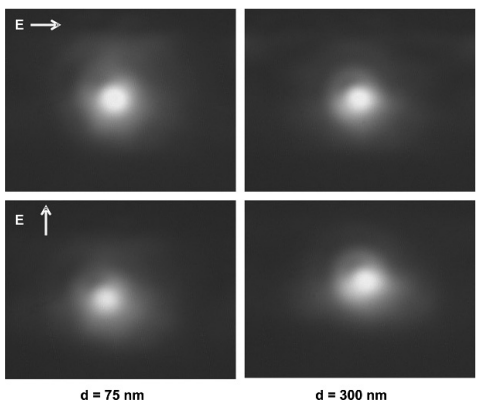
Fig. 11. Measured transmission spectra through the direct and the coupled arms of LRSPP directional couplers based on 8-μm-wide gold stripe waveguides with (a) 3-μm separation and (b) 6-μm separation. The parallel interaction length is 8 mm [From Ref. (73)].

the DCs based on 8-μm-wide waveguides with an interaction length of 8 mm and two different separation distances. Transmission spectra are normalized to the straight stripe transmission of the same length, demonstrating that no additional loss is introduced by the DC. Spectra were recorded over a very broad wavelength range from 1000 to 1680 nm and showed efficient (approximately 30 dB) power transfer from one arm to another when dividing the signals between different telecom wavelength bands.

An interesting feature of LRSPP waveguides is that the spacing of coupled waveguides can be made much smaller than the mode size. In the



(a)



(b)

Fig. 12. (a) Calculated mode coupling in closely spaced gold nanowire waveguides with 200 nm×200 nm core size. The cladding index was assumed to be 1.535 and the simulation wavelength was 1550 nm (b) Experimental mode profiles for parallel gold nanowires (150 nm × 150 nm) with a spacing of 75 nm (left panels) and 300 nm (right panels), for both polarizations.

case of stripe waveguides, this can be achieved by vertical coupling[75,77] which also opens up the possibility for coupling of multiple waveguide layers, fabricated by repeated dielectric and metal coating/patterning steps. The same is possible by lateral or vertical coupling of two nanowire waveguides as illustrated in Fig. 12(a), showing finite-element simulations of two 200 nm × 200 nm gold nanowire waveguides with variable spacing. When separated by a large distance, each wire supports two orthogonally

polarized long-range supermodes. When the wires are closer together, these will couple to form symmetric or antisymmetric modes of the combined system. The splitting of the odd and even combinations of x and y-polarized LRSPP supermodes is observed down to a wire spacing of about $7\,\mu$m, below this spacing the odd solutions are no longer bound. Below a spacing of about $1\,\mu$m however, a splitting of the symmetric combinations of the two orthogonally polarized modes is observed. Figure 12(b) confirms experimentally that double wire structures transmit both horizontal and vertical input polarizations for small wire spacing (150 nm × 150 nm wires, 1550 nm wavelength, 75 nm and 300 nm wire-to-wire gap size). Details of the mode structure in this case are well below the resolution of the imaging system, in spite of a high numerical aperture of 0.95.

4.3. *Reflection gratings*

Typical LRSPP reflection and transmission spectra for stripe waveguides with periodically thickness-modulated gratings, having different modulation amplitude and length, are presented in Fig. 13. As described in earlier studies[59] the investigated gratings showed a behavior similar to that expected from conventional Bragg reflectors, with the Bragg wavelength easily estimated from simple considerations. A more detailed theoretical analysis on reflection, transmission and out-of-plane scattering in LRSPP gratings can be found in Ref. (78). For thickness-modulated gratings, the grating strength can be increased either by increasing the grating length or making higher ridges thus increasing the effective refractive index modulation. Both the gap in transmission and the peak in reflection increase with increasing ridge height, as clearly seen in Fig. 13. Short but efficient gratings with maximum power reflection up to 60% can be realized by making shallow (only tens of nanometers) metal ridges.[59] It should be noted that LRSPP Bragg gratings fabricated by varying the width of a metal stripe give a weaker modulation, thus exhibiting very narrow bandwidths (below 1 nm) and requiring grating lengths up to several millimeters.[79]

Figure 14 shows a comparison between LRSPP thickness-modulated[59] and constant-thickness[63] gratings with the same grating period, fabricated by EBL and NIL, respectively. Both structures show a clear dip in transmission that becomes more pronounced as the grating length increases. As expected, the thickness-modulated gratings provide considerably stronger reflection, exhibiting a transmission dip of more than 10 dB for a grating length of $160\,\mu$m. Obtaining a similar dip in transmission for gratings

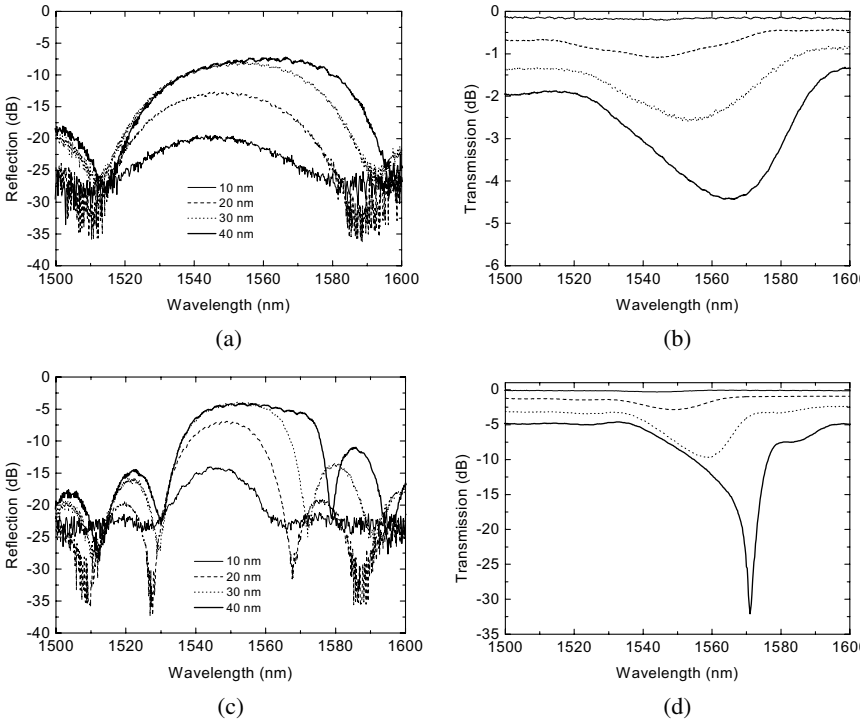
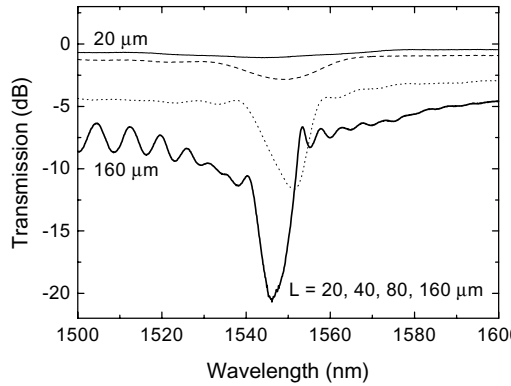


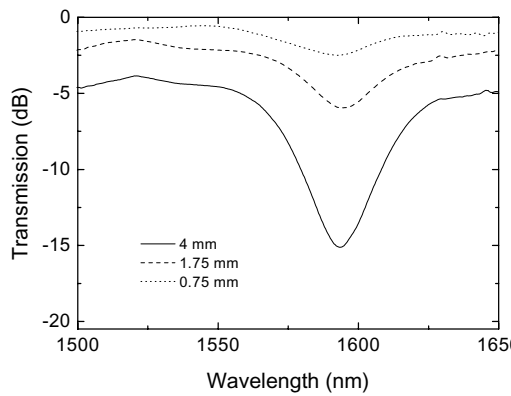
Fig. 13. Optical performance of symmetrically thickness-modulated LRSPP gratings with two different lengths [20 μm (a,b) and 40 μm (c,d)]: measured reflection (a,c) and transmission (b,d) spectra for different heights of the ridges (total height on both sides of the waveguide plane ranges from 10 nm to 40 nm). Grating period is 500 nm [from Ref. (59)].

defined by nanoimprint lithography requires a grating length of about 4 mm. The difference in the transmission dip position, or Bragg wavelength, for the same grating period is caused by the difference in the refractive index of the surrounding polymer. The observed asymmetry in the spectra of the thickness modulated gratings is related to different loss mechanisms in such gratings, namely absorption in the metal and out-of-plane scattering.[78]

LRSPP Bragg gratings offer the possibility of tuning grating parameters (for example, grating size and bandwidth) over a very broad range, which, for example, cannot be done in the case of fiber Bragg gratings due to the photorefractive effect limitation. Moreover, the fabrication procedures for making LRSPP gratings are quite robust, offer good reproducibility, scalability and rapid wafer-level processing using the nanoimprint



(a)



(b)

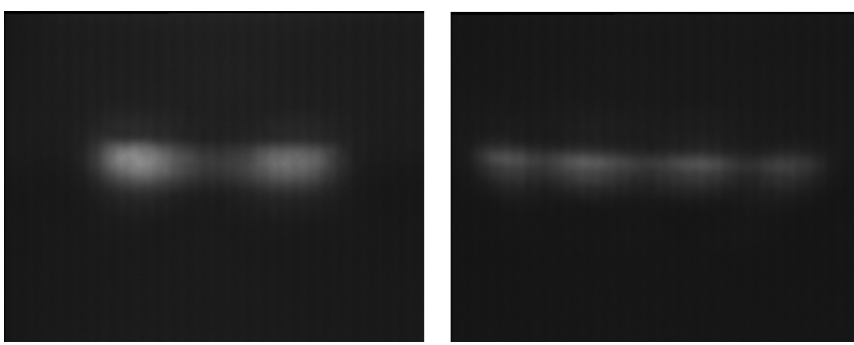
Fig. 14. Transmission spectra for different grating geometries. (a) Thickness-modulated gold stripes with 10-nm ridges on each side of the waveguide plane and (b) corrugated gold stripe of constant thickness fabricated by nanoimprint lithography. The grating period is 500 nm in both cases but the transmission dips occur at different positions due to a difference in the refractive index of the cladding polymers.

technique, which additionally allows straightforward integration into microfluidic and lab-on-a-chip systems.[65] Plasmonic gratings based on periodically thickness-modulated films or corrugated metal surfaces provides a good example of how plasmonics can be applied to practical optical devices by providing a robust wavelength-selective building block for fabricating more complex device geometries ranging from DFB lasers to light collection and extraction structures to optical sensors with a large evanescent-wave sensing volume.

4.4. *Interferometric devices*

Integrated optical devices relying on interference for power splitting, signal modulation or switching are easily realized using LRSPP waveguides. Examples of reported devices include multi-mode interferometers (MMIs) and Mach-Zehnder interferometers (MZIs). Figure 15 illustrates how two-fold or four-fold power splitting is realized with 40- and 60-μm-wide LRSPP waveguides supporting several lateral modes (at 1550 nm wavelength). The output intensity distributions for MMIs of different widths and lengths correspond well to intensity distribution patterns obtained from effective index modeling.[55] Furthermore, the positions of experimentally observed two- and four-fold images in the MMI devices shown in Fig. 15 were found to be in good agreement with simple estimates using the self-imaging principle.[80]

Metallic waveguides have the interesting and unique property of being able to carry electrical current, which can be utilized, e.g., for thermo-optic[25] or electro-optic[81] control of the optical signal, as well as for detaching target analytes from waveguide surfaces in optical biosensors.[65] A design for a thermally controllable MMI structure was proposed in Ref. (57) and is shown in Fig. 16. By passing a current through the edge sections of a multimode waveguide, the mode interference can be tailored to obtain 1×2 switching. Considerably more detailed control of mode interference can be envisioned in MMI structures with a larger number



(a) (b)

Fig. 15. Two-fold and four-fold power splitting in LRSPP multi-mode interferometers (MMIs) consisting of a 15-nm thick gold stripe with (a) width 40 μm, length 3.3 mm and (b) width 60 μm, length 5.5 mm. The MMIs were excited in the center of the wide stripes by single-mode LRSPP waveguides.

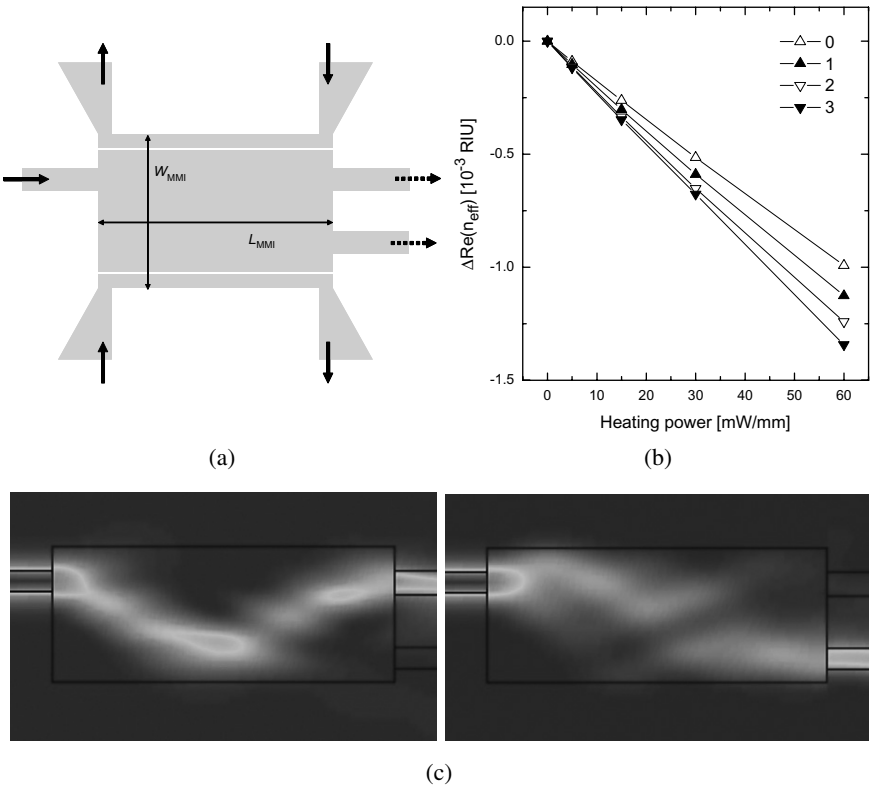


Fig. 16. (a) Schematic illustration of a thermally controllable MMI structure. The modes are excited asymmetrically from the off-center waveguide on the left. The MMI interference pattern is modified by changing the refractive index at the MMI edges as electrical current is passed through the edges of the multimode waveguide. (b) Temperature-induced change in effective mode index calculated for the four lateral modes supported by a multimode LRSPP waveguide with a polymer cladding. The change in index is more pronounced for higher order modes. (c) Simulation of switching of the LRSPP signal from one output waveguide to another in a 40-μm $\times$ 2450-μm MMI structure.

of individually heated sections, since only the immediate lateral vicinity of the heated area is affected, as the substrate wafer (e.g., silicon) acts as an efficient heat sink. Heating power is also used efficiently since the maximum refractive change coincides with the maximum field intensity of the guided modes, occurring at the surface of the metal. The highly localized thermo-optic control is conveniently illustrated in the case of the LRSPP directional-coupler switch reported in Ref. (25).

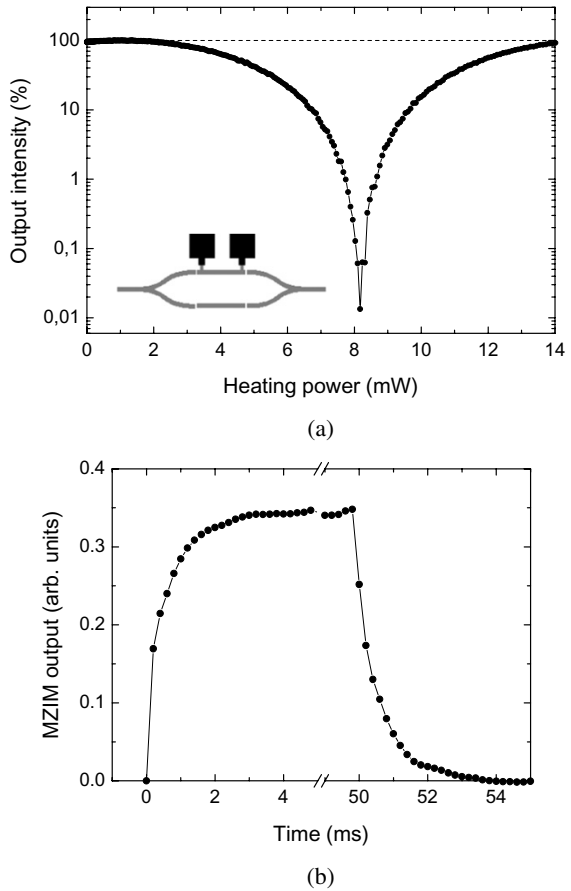


Fig. 17. (a) LRSPP extinction by destructive interference in a Mach-Zehnder interferometer (MZI) structure, illustrated in the inset. A phase delay in one of the MZI arms is introduced by passing electrical current through the LRSPP waveguide, heating the surrounding cladding material. (b) Time response of the MZI, with heating (switch off) and cooling (switch on) times of approximately 2 ms and 1 ms, respectively.

A different method of mode interference involves the 50/50 splitting of a single-mode input signal, introduction of a phase difference between the two signals, followed by recombination giving constructive and destructive interference with respect to the guided mode of a single-mode output waveguide.[25] As shown in Fig. 17(a), an extinction ratio of at least 4 orders of magnitude can be realized this way by applying only 8 mW of heating power over a 5.7-mm long waveguide section. The high extinction ratio

confirms the precise 50/50 power splitting ratio of the LRSPP y-splitter structure and the low driving power and the short timescale of the thermal response [approximately 1 ms, Fig. 17(b)] indicates efficient thermo-optic control.

4.5. *Extinction modulators*

LRSPP signals can also be modulated by introducing mode leakage, e.g., using in-line extinction modulators consisting of heated nanowire waveguides[26] or stripe waveguides in symmetric[58] or asymmetric[52] cladding environment. The asymmetric (polymer/glass) structure potentially uses the smallest driving power but requires temperature regulation to maintain maximum transmission. In the symmetric case, transmission does not depend on the ambient temperature but a polymer cladding (negative thermo-optic coefficient) is required to introduce mode leakage by heating of the waveguide core.[67]

The extinction properties of a symmetric gold nanowire attenuator device (length 2 mm, 180 nm × 180 nm cross section) at 1550 nm wavelength are shown for x- and y-polarized input signals in Fig. 18(a). The device geometry is shown in Fig. 18(b). For properly fabricated nanowires, low

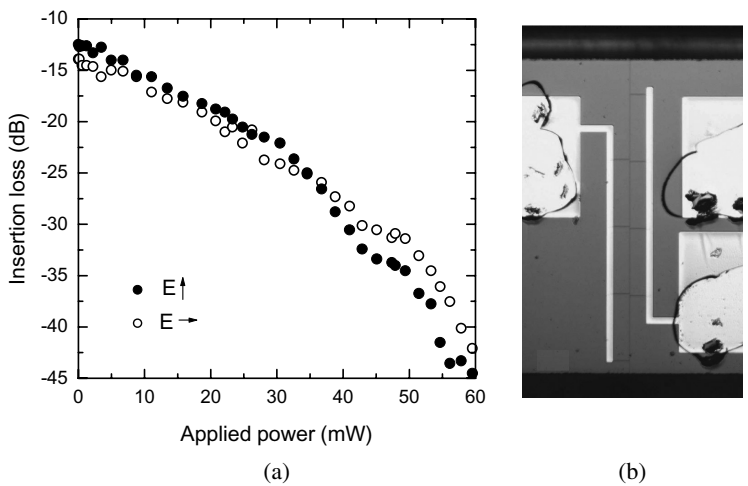


Fig. 18. (a) Extinction characteristics for x-polarized and y-polarized light (1550 nm wavelength) propagating through a 2-mm long gold nanowire waveguide device depicted in (b). The wire (cross-sectional dimensions 180 × 180 nm) is visible as a thin dark line between the parallel electrodes in the photograph.

PDL can be realized over the whole attenuation range. Extinction ratios over 30 dB are observed for both polarizations using 60 mW of heating power. Slight deviations from square symmetry, however, can result in strong polarization dependence, especially at higher extinction ratios.[26] Also, the off-state insertion loss should be decreased to about 1 dB in order to realize a practical device. This should be possible by using short (0.5 mm) plasmonic nanowire devices with additional dielectric confinement to reduce coupling loss to single-mode fibers.[69] Such devices have not yet been tested experimentally.

4.6. *LRSPP amplification*

Optical amplification of plasma waves on metallic surfaces has the potential to create new opportunities in plasmonics and metamaterials research, as previously noted. We have experimented with ultrathin LRSPP waveguides where dye-doped or conjugated fluorescent polymers form part of the dielectric cladding, as described in Sec. 3.3. Figure 19(a) shows emission

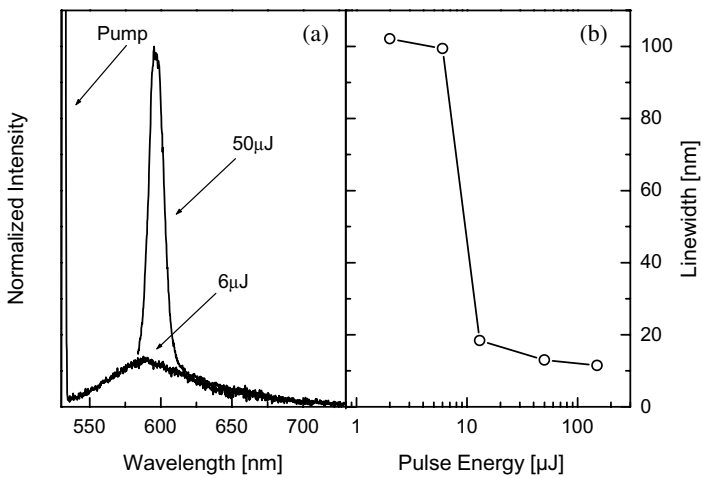


Fig. 19. (a) TM-polarized emission spectra collected from the end facet of a gold-film LRSPP waveguide sample with a conjugated fluorescent polymer layer forming a part of the top cladding. Spectra are shown for pump pulse energies above and below the threshold for amplified spontaneous emission. To facilitate comparison, the spectra are normalized to give similar fluorescence background intensities. TE-polarized spectra only show a broad fluorescence peak at all pulse energies. (b) The collapse of the emission linewidth (FWHM of the TM-polarized emission spectrum), occurring for pulse energies above 10 μJ in the sample under investigation (Au film thickness approximately 4 nm).

spectra collected from the end facet of diced LRSPP samples with a gain layer. The fluorescent layer was pumped in a stripe geometry with a pulsed 532-nm laser. Gain measurements were carried out using the variable stripe length method,[53] with a pump stripe length up to 2 mm. At low pulse energies, only a broad fluorescence peak from the cladding material is observed for both TE and TM polarizations. At higher pulse energies, however, the TM-polarized emission spectrum becomes dominated by a narrow line, indicative of amplified spontaneous emission. No such amplification is observed in the TE polarized spectrum, or when the sample is pumped in a region with no gold film.

The above results can only be accounted for by assuming that chromophores in the gain layer excited by the pump pulse transfer energy to the (TM-polarized) LRSPP mode through stimulated emission. In the absence of feedback, the emitted signal above the transparency threshold represents amplified spontaneous emission. Plotting the FWHM linewidth of the emission spectrum [Fig. 19(b)] as a function of pulse energy shows a clear threshold behavior. Furthermore, measurements of the emitted light intensity as a function of the length of the pump stripe indeed confirm that the emission follows the expected exponential increase characteristic of amplified spontaneous emission. From this data, a net gain coefficient of approximately $8\,\mathrm{cm}^{-1}$ at a pulse energy of $12.5\,\mu\mathrm{J}$ was extracted.[71] This corresponds to a 5-fold intensity increase over a 2-mm propagation distance, or a 3000-fold increase over 1 cm. The observation of a net gain coefficient in the LRSPP waveguide geometry confirms that the high material gain provided by the fluorescent polymer is sufficient to overcome the propagation loss of the LRSPP waveguide. The propagation loss depends strongly on the metal film thickness but for a 4-nm-thick gold film it was determined to be approximately $60\,\mathrm{cm}^{-1}$ at 600 nm wavelength when the gain layer is properly index-matched to the other cladding materials.

The efficiency of energy transfer from the gain layer to surface plasmon modes in the metal layer was also investigated using the Green's function method, confirming that dipoles oriented perpendicular to the gold surface in the range 30–500 nm from the surface of a 4-nm thick gold film transfer 10–15% of spontaneously emitted energy directly to the LRSPP mode, while the remaining 85–90% is emitted as fluorescence or dissipated in the metal film.[44] The distance range correlates well with the extent of the LRSPP mode which (for the film thickness and wavelength used here) has a mode size of about $1\,\mu$m (FWHM).

5. Conclusions and Outlook

In this paper, we have discussed the properties of LRSPP waveguides and their application in various types of integrated optical devices. Plasmonic waveguiding along metal stripes or metal nanowires offers interesting possibilities in designing novel photonic devices which might be especially useful in integrated biophotonics, where large sensing volumes and the possibility of using functionalized gold surfaces for waveguiding are of importance. Because of the possibility of realizing strong index modulation and efficient mode extinction, certain types of plasmonic components might prove useful in combination with a dielectric waveguide platform in cases where component footprint is of critical concern. Furthermore, LRSPP waveguides provide a useful starting point for studying plasmonic amplification and plasmonic lasing, since their propagation loss can be overcome using common organic gain materials in convenient device geometries.

Acknowledgments

KL acknowledges support by the Icelandic Research Fund (grant no. 070021021), the Eimskip Research Fund, University of Iceland Research Fund and the Selma & Kaj Langvad fund. AB acknowledges support from the Danish Research Council for Technology and Production Sciences (grant no. 274-07-0057), the 2009 SAOT Young Researcher Award, ARO grant W911NF-09-1-0516 and ARO-MURI award W911NF-09-1-0539. The authors also wish to thank Sergey I. Bozhevolnyi, Thomas Nikolajsen, Thomas Søndergaard, Peter M. W. Skovgaard, John Erland Østergaard, Jørn M. Hvam, Anders Kristensen, Rasmus H. Pedersen, Tiberiu Rosenzveig, Klaus Meerholz and Norbert Danz for fruitful collaboration.

References

1. H. Raether, *Surface Plasmons*. (Springer, Berlin, 1988).
2. W. L. Barnes, A. Dereux, and T. W. Ebbesen, *Nature* **424**, 824 (2003).
3. S. Lal, S. Link, and N. J. Halas, *Nature Photonics* **1**, 641 (2008).
4. D. K. Gramotnev and S. I. Bozhevolnyi, *Nature Photonics* **4**, 83 (2010).
5. J. A. Schuller, E. S. Barnard, W. Cai, Y. C. Jun, J. S. White, and M. L. Brongersma, *Nature Materials* **9**, 193 (2010).
6. J. Homola, *Surface Plasmon Resonance Based Sensors*. (Springer, Berlin, 2006).
7. S. Maier and H. A. Atwater, *Journal of Applied Physics* **98**, 1 (2005).

8. I. I. Smolyaninov, D. L. Mazzoni, and C. C. Davis, *Phys. Rev. Lett.* **77**, 3877 (1996).
9. S. I. Bozhevolnyi and F. A. Pudonin, *Phys. Rev. Lett.* **78**, 2823 (1997).
10. H. Ditlbacher, J. R. Krenn, G. Schider, A. Leitner, and F. R. Aussenegg, *Appl. Phys. Lett.* **81**, 1762 (2002).
11. S. I. Bozhevolnyi, J. Erland, K. Leosson, P. M. W. Skovgaard, and J. M. Hvam, *Phys. Rev. Lett.* **86**, 3008 (2001).
12. S. I. Bozhevolnyi, V. S. Volkov, and K. Leosson, *Phys. Rev. Lett.* **89**, 186801 (2002).
13. M. Fukui, V. C. Y. So, and R. Normandin, *Phys. Status Solidi B* **91**, K61, (1979).
14. D. Sarid, *Phys. Rev. Lett.* **47**, 1927 (1981).
15. P. Berini, *Phys. Rev. B* **61**, 10484 (2000).
16. P. Berini, *Optics Letters* **24**, 1011 (1999).
17. R. Charbonneau, Berini, Pierre, E. Berolo, and Lisicka-Skrzek, Ewa. *Optics Letters* **25**, 844 (2000).
18. P. Berini, *Advances in Optics and Photonics* **1**, 484 (2009).
19. L. Dobrzynski and A. A. Maradudin, *Physical Review B* **6**, 3810 (1972).
20. D. F. P. Pile, T. Ogawa, D. K. Gramotnev, T. Okamoto, M. Haraguchi, M. Fukui, and S. Matsuo, *Appl. Phys. Lett.* **87**, 061106 (2005).
21. E. Moreno, S. G. Rodrigo, S. I. Bozhevolnyi, L. Martin-Moreno, and F. J. Garcia-Vidal, *Phys. Rev. Lett.* **100**, 023901 (2008).
22. A. Boltasseva, V. S. Volkov, R. B. Nielsen, E. Moreno, S. G. Rodrigo, and S. I. Bozhevolnyi, *Optics Express* **16**, 5252 (2008).
23. K. Leosson, T. Nikolajsen, A. Boltasseva, and S. I. Bozhevolnyi, *Optics Express* **14**, 314 (2006).
24. J. Jung, T. Søndergaard, and S. I. Bozhevolnyi, *Phys. Rev. B* **76**, 035434 (2007).
25. T. Nikolajsen, K. Leosson, and S. I. Bozhevolnyi, *Appl. Phys. Lett.* **85**, 5833 (2004).
26. K. Leosson, T. Rosenzveig, P. G. Hermannsson, and A. Boltasseva, *Optics Express* **16**, 15546 (2008).
27. A. N. Sudarkin and P. A. Demkovich, *Soviet Physics Uspekhi* **34**, 764 (1989).
28. C. Sirtori, C. Gmachl, F. Capasso, J. Faist, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, *Optics Letters* **23**, 1366 (1998).
29. M. P. Nezhad, K. Tetz, and Y. Fainman, *Optics Express* **12**, 4072 (2004).
30. D. J. Bergman and M. I. Stockman, *Phys. Rev. Lett.* **90**, 027402 (2003).
31. N. M. Lawandy, *Appl. Phys. Lett.* **85**, 5040 (2004).
32. J. Seidel, S. Grafstroem, and L. Eng, *Phys. Rev. Lett.* **94**, 177401 (2005).
33. M. A. Noginov, G. Zhu, M. Bahoura, *et al. Optics Letters* **31**, 3022 (2006).
34. O. Popov, A. Zilbershtein, and D. Davidov, *Appl. Phys. Lett.* **89**, 191116 (2006).
35. M. Ambati, S. H. Nam, E. Ulin-Avila, D. A. Genov, G. Bartal, and X. Zhang, *Nano Letters* **8**, 3998 (2008).
36. J. Grandidier, G. Colas-des-Francs, S. Massenot, *et al. Nano Letters* **9**, 2935 (2009).

37. M. A. Noginov, G. Zhu, M. Mayy, B. A. Ritzo, N. Noginova, and V. A. Podolskiy, *Phys. Rev. Lett.* **101**, 226806 (2008).
38. M. T. Hill *et al.*, *Optics Express* **17**, 11107 (2009).
39. M. A. Noginov *et al.*, *Nature* **460**, 1110 (2009).
40. R. F. Oulton *et al. Nature* **461**, 629 (2009).
41. I. De Leon and P. Berini, *Nature Photonics* **4**, 382 (2010).
42. S. Xiao, V. P. Drachev, A. V. Kildishev, X. Ni, U. K. Chettiar, H.-K. Yuan, and V. M. Shalaev, *Nature* **466**, 735 (2010).
43. G. Colas des Francs, P. Bramant, J. Grandidier, A. Bouhelier, J.-C. Weeber and A. Dereux, *Optics Express* **18**, 16327 (2010).
44. M. C. Gather, K. Meerholz, N. Danz, and K. Leosson, *Nature Photonics* **4**, 457 (2010).
45. K. Okamoto, *Fundamentals of Optical Waveguides*. (Elsevier Science & Technology, 2006).
46. S. Maier, *Plasmonics: Fundamentals and Applications*. (Springer, New York, 2007).
47. P. B. Johnson and R. W. Christy, *Phys. Rev. B* **6**, 4370 (1972).
48. P. Dumon, W. Bogaerts, V. Wiaux, *et al. Photonics Technology Letters* **16**, 1328 (2004).
49. H. Yamada, T. Chu, and S. Ishida, *IEEE Journal of Selected Topics in Quantum Electronics* **12**, 1371 (2006).
50. B. Jalali, *Physica Status Solidi (a)* **205**, 213 (2008).
51. B. Prade, J. Y. Vinet, and A. Mysyrowicz, *Phys. Rev. B* **44**, 13556 (1991).
52. G. Gagnon, N. Lahoud, G. A. Mattiussi, and P. Berini, *Journal of Lightwave Technology* **24**, 4391 (2006).
53. M. D. McGehee, R. Gupta, S. Veenstra, E. K. Miller, M. A. Díaz-García, and A. J. Heeger, *Physical Review B* **58**, 7035 (1998).
54. R. Charbonneau, N. Lahoud, G. Mattiussi, and P. Berini, *Optics Express* **13**, 977 (2005).
55. A. Boltasseva, T. Nikolajsen, K. Leosson, K. Kjaer, M. S. Larsen, and S. I. Bozhevolnyi, *Journal of Lightwave Technology* **23**, 413 (2005).
56. T. Nikolajsen, K. Leosson, I. Salakhutdinov, and S. I. Bozhevolnyi, *Applied Physics Letters* **82**, 668 (2003).
57. P. G. Hermannsson and K. Leosson, *Proc. SPIE,* **6988**, 69880A (2008).
58. T. Nikolajsen, K. Leosson, and S. I. Bozhevolnyi, *Optics Communications* **244**, 455 (2005).
59. A. Boltasseva, S. I. Bozhevolnyi, T. Nikolajsen, and K. Leosson, *Journal of Lightwave Technology* **24**, 912 (2006).
60. S. I. Bozhevolnyi, A. Boltasseva, T. Søndergaard, T. Nikolajsen, and K. Leosson, *Optics Communications* **250**, 328 (2005).
61. A. Boltasseva, T. Søndergaard, T. Nikolajsen, K. Leosson, S. I. Bozhevolnyi, and J. M. Hvam, *Journal of the Optical Society of America B* **22**, 2027 (2005).
62. S. Jetté-Charbonneau, R. Charbonneau, N. Lahoud, G. A. Mattiussi, and P. Berini, *IEEE Journal of Quantum Electronics* **41**, 1480 (2005).
63. R. H. Pedersen *et al.*, *Microelectronic Engineering* **84**, 895 (2007).
64. L. J. Guo, *Journal of Physics D: Applied Physics* **37**, R123 (2004).

65. Y. Y. Joo, S. H. Song, and R. Magnusson, *Optics Express* **17**, 10606 (2009).
66. J. Takahara, S. Yamagishi, H. Taki, A. Morimoto, and T. Kobayashi, *Optics Letters* **22**, 475 (1997).
67. T. Rosenzveig, P. G. Hermannsson, A. Boltasseva, and K. Leosson, *Applied Physics A* **100**, 341 (2010).
68. J. M. Kontio, H. Husu, J. Simonen, M. J. Huttunen, and J. Tommila, *Optics Letters* **34**, 1979 (2009).
69. T. Rosenzveig, P. G. Hermannsson, and K. Leosson, *Plasmonics* **5**, 75 (2010).
70. Y.-S. Shon, H. Y. Choi, M. S. Guerrero, and C. Kwon, *Plasmonics* **4**, 95 (2009).
71. M. C. Gather and K. Leosson, *CLEO/IQEC Proceedings* (2009), JTuD122.
72. S. Park, J. J. Ju, J. T. Kim *et al.*, *Optics Express* **17**, 697 (2009).
73. A. Boltasseva and S. I. Bozhevolnyi, *IEEE Journal of Selected Topics in Quantum Electronics* **12**, 1233 (2006).
74. S. I. Bozhevolnyi, T. Nikolajsen, and K. Leosson, *Optics Communications* **255**, 51 (2005).
75. T. Srivastava and A. Kumar, *Applied Optics* **49**, 2397 (2010).
76. R. Charbonneau, C. Scales, I. Breukelaar, S. Farard, N. Lahoud, G. Mattiussi, and P. Berini, *J. Lightwave Technol.* **24**, 477 (2006).
77. H. S. Won, K. C. Kim, S. H. Song, C.-H. Oh, P. S. Kim, S. Pak, and S. I. Kim, *Appl. Phys. Lett.* **88**, 011110 (2006).
78. T. Søndergaard, S. I. Bozhevolnyi, and A. Boltasseva, *Phys. Rev. B* **73**, 045320 (2006).
79. S. Jetté-Charbonneau, R. Charbonneau, N. Lahoud, G. A. Mattiussi, and P. Berini, *Optics Express* **13**, 4672 (2005).
80. L. B. Soldano and E. C. M. Pennings, *J. Lightwave. Technol.* **13**, 615 (1995).
81. P. Berini, R. Charbonneau, S. Jetté-Charbonneau, N. Lahoud, and G. Mattiussi, *J. Appl. Phys.* **101**, 113114 (2007).

SURFACE PLASMON BIOSENSING WITH 3D PLASMONIC CRYSTALS

Julia Y. Lin, Jiun-Chan Yang, Hanwei Gao, and Teri W. Odom*

Department of Chemistry, Northwestern University,
2145 Sheridan Road, Evanston, Illinois, 60208, USA
**todom@northwestern.edu*

This chapter introduces a new type of engineered surface — plasmonic crystals with three-dimensional (3D) unit cells — for biosensing. Nanostructured metal films exhibit surface plasmon (SP) resonances with exceptional sensitivities but are limited to bulk refractive index (RI) sensing. Here we describe the biosensing capabilities of two different plasmonic crystal structures that support both surface plasmon polaritons (SPPs) and localized surface plasmons (LSPs): (i) a nano-pyramidal grating and (ii) a 3D nanohole array. These plasmonic crystals not only exhibit bulk RI sensitivities comparable to commercial SP sensors but also have surface sensitivities that can be tuned by manipulating the free-space light excitation angle θ. We show how tailoring the surface sensitivity of the plasmonic crystals optimizes real-time measurements of biomolecular interactions.

1. Introduction

Plasmonic nanostructures are attracting increased attention for chemical and biological sensing because of their potential in portable, miniaturized devices for point-of-use analysis. Rapid progress in the development of optical sensors based on metal nanostructures has made great strides in medical diagnostics,[1,2] environmental monitoring,[3,4] and food safety.[5] This chapter describes a class of nanostructured metal surfaces with well-defined, 3D topographical features that are promising for use in sensitive biodetection. We will discuss how free-space light excitation of surface

plasmons at different angles θ can tune the local refractive index (RI) sensitivity of these plasmonic crystals and how this information can be used to optimize real-time biosensing.

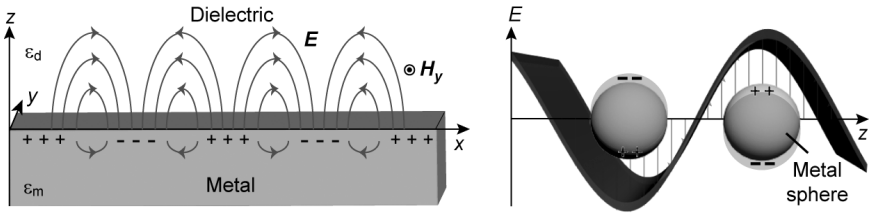
1.1. *Label-free sensing*

Biosensors are used to monitor interactions between biomolecules and to provide quantitative information regarding their binding affinity and reaction kinetics. Common detection schemes involve labeling the receptor and ligand of interest with molecular tags, such as fluorophores, radiolabels or enzymatic labels, which then report binding events and/or amplify signals needed for quantification.[6,7] The use of labels has drawbacks, however, because the tag must be chemically attached to the biomolecule. This attachment step adds extra time and cost to the sensing process and can induce conformational changes in a receptor that interferes with ligand binding, leading to false negatives.[7] In addition, the hydrophobic nature of most fluorescent compounds allow non-specific binding that leads to false positives.[7] An ideal biosensor should be label-free, be sensitive enough to detect intrinsic properties of molecules (i.e., mass and RI), and have the throughput and cost to be appropriate in screening processes.[7]

Optical transduction techniques, such as ellipsometry and surface plasmon resonance (SPR), have shown promise in the direct detection of biomolecular interactions on metal surfaces.[8–10] The first label-free measurement was demonstrated in 1983 using SPR to detect immunochemical reactions on Ag films.[10] Since then, Au has been the standard material for optical sensing because of its inertness to oxidation and well-established surface chemistry compared to Ag. The first commercial SPR instrument was introduced in 1990 by Biacore.[11] Subsequent technological advancements have produced sensors with significantly improved bulk sensitivities, and recent multiplexed sensing capabilities have been introduced by patterning surfaces for high throughput applications.[11,12]

1.2. *Surface plasmon sensing*

The best-known commercial optical biosensors use SPR to detect biomolecules in a label-free and real-time format.[11] SPR measurements are based on the optical properties of metals, where the resonant interaction of light with the conduction electrons at a metal-dielectric interface generates surface plasmons (SPs) [Fig. 1(a)]. When light (e.g., $\lambda = 633\,\text{nm}$) illuminates a thin, planar Au film through a prism, typically in the



(a) (b)

Fig. 1. Types of surface plasmons. (a) Surface plasmon polaritons (SPPs) at a metal-dielectric interface and (b) localized surface plasmons (LSPs) excited on metal nanoparticles by free-space light.

Kretschmann geometry,[13] a sharp minimum appears in the reflectance spectrum (intensity vs. θ) at the resonance condition; this dip is referred to as the SP resonance. The resonance angle is sensitive to variations in RI and shifts when the metal surface is exposed to an analyte.[10] To date, SPR has been the standard in bulk RI sensing.

In contrast with SPR, SP sensors based on metal nanostructures use free-space light excitation to measure resonant *wavelength*, which offers significant improvements in signal resolution compared to resonant *angle* θ. RI sensing using SPs encompasses both localized surface plasmons (LSPs) and surface plasmon polaritons (SPPs). Although LSP- and SPP-sensing are similar because they both respond to changes in RI close to the metal surface, there are important differences between the two sensing modalities. First, the plasmon excitation conditions for LSPs and SPPs are distinct. LSPs can be excited by simply illuminating a metal nanoparticle with light near its resonance wavelength ($\lambda_{\mathrm{LSP}} \approx 550\,\mathrm{nm}$ for Au nanoparticles with diameter $d \leq 100\,\mathrm{nm}$).[14,15] To excite SPPs, the energy and momentum of free-space light must match that of the SP at the metal surface.[6,16] Second, the distances that the SP fields extend into the local environment [away from the metal] are very different. SPP fields decay exponentially and extend hundreds of nanometers into the dielectric region, which makes them useful for detecting changes in RI of bulk solutions. The electromagnetic fields from LSPs, however, are localized within tens of nanometers from the nanostructure surface[17] and produce sensing volumes that are extremely sensitive to molecular binding.[18]

Current research efforts in SP sensing aim to design nanostructures that can improve key analytical figures of merit [e.g., limits of detection (LOD), sensitivity, dynamic range] to not only match but to surpass those of commercial SPR systems. For example, two major challenges

include: (i) improving the surface sensitivities of SPP resonance sensors (to detect biomolecules) and (ii) increasing signal-to-noise from LSP measurements (to detect single molecules).

1.2.1. *Discrete nanostructures: particles and holes*

1.2.1.1. Localized surface plasmons (LSPs)

LSPs are collective oscillations of electrons confined to a finite volume, such as a metal particle[14] or a hole[19] [Fig. 1(b)]. The resonant condition of these nanoparticles can be tuned from ultraviolet (UV) to near-infrared (NIR) wavelengths by tailoring their size, shape, material and dielectric environment.[14] Thus, in the context of sensing, molecules that adsorb on a nanoparticle surface will produce a shift in the LSP wavelength. Other metal nanostructures that support LSPs and that are commonly used in chemical and biological sensing include arrays of particles supported on substrates and planar nanohole films.[20,21]

1.2.1.2. Why use LSPs and discrete nanostructures in sensing?

Discrete nanostructures support highly localized electromagnetic fields that produce sensing volumes with dimensions in the order of the size of a molecule, which makes them favorable for biosensing.[20] Compared to conventional SPR, however, LSP biosensors universally exhibit poor data quality as a consequence of low signal-to-noise (S/N).[22] S/N is defined as the maximum resonance wavelength shift divided by the noise amplitude from the measurement. Sensing based on tracking resonance peak positions has led to S/N = 50 in protein binding experiments;[23] similarly, monitoring incremental shifts in the extinction spectra at a fixed wavelength has produced S/N $\leq$ 100, which is 1–2 orders of magnitude lower than that measured by SPR.[24] Applying the appropriate peak fitting algorithms to resolve resonance wavelength shifts, however, can increase S/N ratios so they are more comparable to SPR.[22,24]

1.2.1.3. Discrete nanostructures on surfaces: bulk RI sensitivity

Surface-supported nanostructures are a versatile RI sensing platform because their optical properties can be tuned by manipulating structural parameters through nanofabrication techniques. Nanosphere lithography (NSL)[25] and colloidal lithography[26] have generated nanoparticle and hole

substrates that have been used in LSP sensing, such as detecting changes in protein conformation.[27]

Sensing with arrays of particles or nanoholes is advantageous over individual structures because the signal from similar nanostructures can be averaged to increase intensity,[28] and the optimization of spectroscopic measurements on relatively large substrate areas (mm^2) can improve S/N.[29] In addition, the RI sensitivity does not change significantly between a single nanostructure or an ensemble of the same structures on a surface.[30] The major drawback of surface-patterned nanostructures, however, is that the particles are often supported on surfaces (e.g., glass, $n = 1.5$) with RIs different from the environment surrounding them. The large difference between the RI of the substrate and the bulk surrounding reduces the overall sensitivity of the particles.[31] Strategies to enhance the bulk sensitivity of these plasmonic structures include creating ordered, periodic arrays of nanoparticles to narrow the spectral bandwidth of the extinction peak,[28,32] decreasing particle and substrate interaction by reducing contact area,[33] or fabricating the particles on low RI surfaces.[34] In the case of a film perforated with nanoscale holes, the analyte could be selectively placed within the nanohole.[35–37]

1.2.1.4. Sensing biomolecules in real-time

Real-time measurements of molecular binding events have been demonstrated on surface-supported nanostructures under sub-monolayer coverage (ca. 100 molecules).[23,24,30,38,39] Discrete nanostructures introduce topographical perturbations to flat substrates, which can affect the binding kinetics of biomolecules.[39] Specifically, the local structure (high curvature or sharp edges) of nanoparticles or holes has been observed to alter the rates at which biomolecules bind or unbind, producing different on and off rate constants than those determined by SPR.[24,37] In contrast, nanostructures with surfaces that are roughly planar, such as nanorods,[39] have less of an effect on the kinetic analysis of binding interactions. When designing LSP sensors for quantitative bioanalysis, it is important to consider structural parameters, such as nanostructure height,[40] shape,[37] and size,[41] so that the binding constants can be determined accurately.

1.2.1.5. LSP sensing with single nanoparticles: promise and potential

LSP sensing at the single nanoparticle level has received attention because discrete nanostructures are more compatible with the array-based and

multiplexed analysis required for miniaturized, point-of-care devices than SPR.[42] The ultimate goal is to use the highly confined LSP fields for single molecule detection, recently demonstrated in a resonant microcavity[43] and on an Au bipyramid nanoparticle.[44]

The development of characterization tools, such as dark field scattering spectroscopy, for the analysis of individual nanoparticles has enabled single particle LSP resonance methods.[30] The high noise level is a significant problem in these measurements,[29] which leads to reduced temporal and peak shift resolution[45] as well as much worse LOD compared to SPR measurements ($LOD_{SPR} = 10\,pg\ protein/cm^2$).[6] Despite these drawbacks, however, single nanoparticles have significant potential for sensing trace analytes because of reduced sample volumes to produce a response.[30]

1.2.2. *Continuous nanostructures: gratings and hole arrays*

1.2.2.1. Surface plasmon polaritons (SPPs) and extraordinary optical transmission

At a continuous metal-dielectric interface, SPs propagate as two-dimensional (2D) charge density waves [Fig. 1(a)]. A method to excite surface plasmon polaritons (SPPs) is to couple free-space light with the conduction electrons of a noble metal surface patterned with 2D periodic modulations.[6,16] The in-plane momentum of the metal SPs is greater than that of the incident light because of the discontinuity between the electric permittivity across the interface. Hence, a periodic array can provide the additional momentum to compensate for the mismatch and generate SPPs. SPPs play an important role in extraordinary optical transmission on metallic nanohole arrays,[46] which results in strong plasmon fields localized within the voids of the metal that could impact biosensing.[47]

1.2.2.2. Why use SPPs and nanostructured films in sensing?

Nanostructured films, such as one-dimensional (1D) diffraction gratings and two-dimensional (2D) subwavelength hole arrays, have generated interest in sensing applications because (i) they act as both a light coupler and optical transducer, which facilitates device miniaturization compared to the SPR prism configuration, and (ii) they are simple and inexpensive to fabricate, which makes them attractive for the development of low-cost SPP-based devices. Despite their compatibility with high throughput fabrication, diffraction gratings are not as widely used in SPR as prisms

because the sensitivity of most grating-based sensors is lower than that of conventional SPR sensors.[48,49]

1.2.2.3. Continuous nanostructures: bulk refractive index sensitivity

Several parameters of nanostructured films can be tailored to improve the bulk RI sensitivity of grating-based sensors, such as grating profile,[50] periodicity ($a_0 = 400$–600 nm), and feature height.[51] For example, a metal film with a sinusoidal profile has a 6-fold higher sensitivity than a planar film.[50] Moreover, diffractive gratings can be engineered to sustain more than one SPP mode response. Monitoring multiple SPPs provides information about the RI at different distances from the sensor surface, which can increase sensitivity; also, the response to bulk changes from one SPP resonance can be correlated with other resonances. Diffractive gratings with three different grating periods[52] and metal films perforated with an array of thin slits[53] exhibit multiple SPP modes that have been used to monitor changes in bulk RI. For example, variations in the background RI and changes in the thickness and RI of a bovine serum albumin multilayer (4 layers, thickness = 5.1 nm) could be determined by tracking three resonances concurrently.[52]

1.2.2.4. Detecting single biomolecule events using SPPs: promise and potential

Biorecognition of immunochemical (antigen-antibody) interactions were performed on early prototypes of Au diffraction line gratings.[54,55] These engineered films, however, were found to be much more sensitive to changes in bulk RI compared to local RI. Thus, nanostructured surfaces need to be developed to make possible SPP sensing from sub-monolayer coverage down to the single molecule level.

The remaining sections in the chapter will focus on 2D gratings with 3D unit cells that have started to address this important problem. These nanostructured surfaces are promising because they support different types of plasmons that can distinguish between changes in bulk and local surface RI. For example, hybrid hole-particle arrays are architectures that support LSP resonances, SPPs and Wood's anomalies, which can couple to produce spectral features that are sensitive to the surrounding environment. Quantitative determination of protein layer thickness using these hybrid substrates exceeded previously reported structures.[56]

2. Engineered Surfaces for Biosensing: Plasmonic Crystals with 3D Unit Cells

This section highlights recent breakthroughs in biosensing based on 3D nanofabricated surfaces, mostly taken from the authors' lab. Advances in large-area nanofabrication techniques have enabled engineered surfaces to be designed with uniquely tunable optical properties. Plasmonic crystals are an example of new nanopatterned substrates that consist of 2D arrays of 3D unit cells and are ideal platforms to test for biosensing. We have discovered that noble metal films patterned with well-defined 3D topography offer two advantages over commercial 1D gratings: (i) an excitation angle (θ)-sensitive response,[57] which introduces a *new* design parameter for biosensor development, and (ii) spectral features from both LSPs and SPPs,[58] which enable multi-modal readout.

We have designed two types of 3D plasmonic crystals that can be used for biosensing: nanopyramidal gratings and 3D nanohole arrays. To fabricate the patterned substrates, we first define a nanoscale photoresist pattern (square array on pitch $a_0 = 400\,\text{nm}$ with diameter $d = 100\,\text{nm}$) on Si (100) surfaces using soft interference lithography,[34] a near-field photolithography technique that uses a poly(dimethylsiloxane) (PDMS) mask. Next, the photoresist pattern is transferred into a Si template by PEEL (Phase-shifting photolithography, Etching, E-beam deposition and Lift-off), a procedure that converts photoresist nanopatterns into functional materials such as metals or semiconductors [Fig. 2(a)].[59,60] Deposition of a thin layer of Cr followed by lift-off of the photoresist produces holes in a Cr film. This perforated film acts as an etch mask to form the pyramidal pits in the underlying Si (100). Removal of the Cr film yields a Si template to create plasmonic crystals with 3D unit cells.

For these nanostructured surfaces to be integrated into portable devices and point-of-use diagnostics, their fabrication must be scalable and low-cost. A simple molding technique[61,62] [Fig. 2(a)] has the potential to replicate and mass-produce the nanopyramidal gratings and 3D nanohole arrays in stiff and transparent polyurethane (PU) substrates. First, a composite PDMS stamp[63] was fabricated from the Si plasmonic crystal template (master). Molding this PDMS stamp, which contains the inverted topographical features of the master, against PU generates a replica of the original pattern (cm^2) in PU on glass. Either large-area 2D nanopyramidal gratings or 3D nanohole arrays can result from the line-of-sight ($\alpha = 0°$) or angled deposition (α) of Au onto the PU [Figs. 2(b–d)].

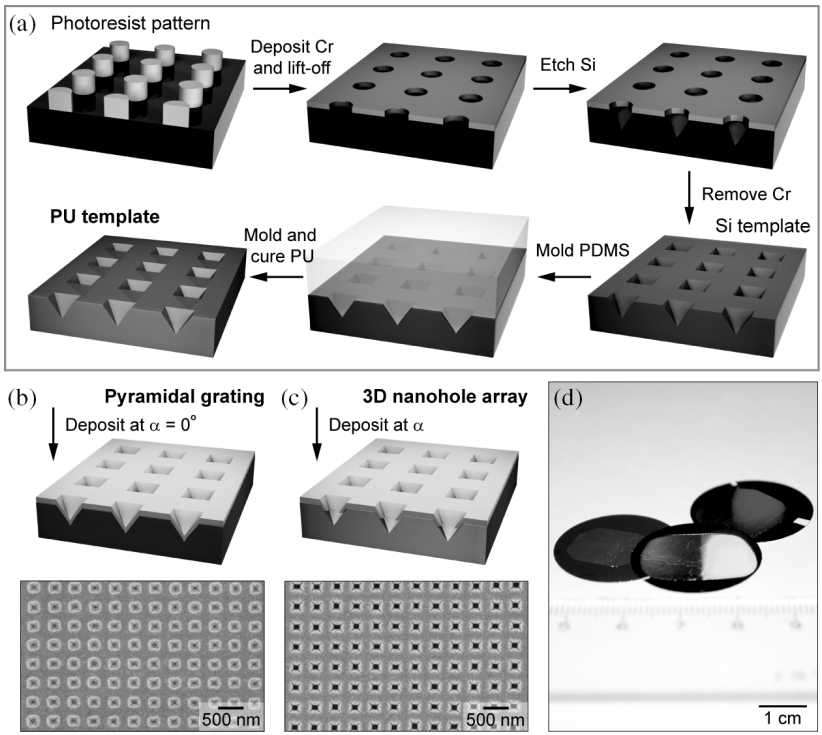


Fig. 2. Fabrication of plasmonic crystals by (a) PEEL (Phase-shifting photolithography, Etching, E-beam and Lift-off) and molding in polyurethane (PU). (b) 2D nanopyramidal grating generated by line-of-sight deposition and (c) 3D nanohole array created by angled deposition of Au onto patterned PU substrates. (d) Optical micrograph of Au plasmonic crystals patterned over macroscale areas. [Reproduced with permission and adapted from Refs. (57), (58) and (65)]

3. Bulk Refractive Index Sensitivity and Dependence on Angle of Incidence

In order to compare the performance of plasmonic sensors made from different nanostructures and materials, a metric referred to as the figure of merit (FOM) was introduced to assess sensitivity to changes in the RI of the surrounding bulk solution. The FOM is defined as the shift in energy of the resonance per refractive index unit (RIU) divided by the full width at half maximum (FWHM) of the resonance.[64] Both prism-coupled SPR (planar film) and SPP resonance (1D grating) sensors exhibit extremely sensitive bulk responses at fixed angles of incidence θ, with calculated

FOMs of 83 and 80, respectively.[18] Independent of θ, plasmonic crystals of 3D nanopyramids had bulk RI sensitivities that could generate a FOM greater than 85.[65]

Because plasmonic crystals can be patterned over large areas (cm^2), we took advantage of the available sample area to manipulate excitation angles from $\theta = 20$–$70°$ to optimize the angle-dependent SPP resonances for biosensing. For 2D square gratings, the SPP Bloch waves are given by the Bragg coupling condition:[66]

$$\frac{\omega}{c}\sqrt{\frac{\varepsilon_m \varepsilon_d}{\varepsilon_m + \varepsilon_d}} = |k_0 \sin\theta + iG_x + jG_y| \tag{1}$$

where ω, c, and k_0 are the angular frequency, speed and momentum of free-space light. ε_d represents the relative permittivity of the adjacent dielectric and can also be expressed in terms of RI ($n = \sqrt{\varepsilon_d}$) if the material is non-absorptive. Furthermore, the magnitude of the Bragg vectors for a square array are defined as $|G_x| = |G_y| = 2\pi/a_0$ and related to the pitch a_0 of the first Brillouin zone. Each of these variables can tune the optical properties of the grating, which are characterized by the integer pair (i, j), indices that define the order of the Bloch wave modes.

The nanopyramidal gratings exhibit one narrow plasmon resonance from the metal-dielectric $(-1, 0)$ SPP Bloch mode at all angles of excitation ($\theta = 13°$ to $64°$) [Fig. 3(a,b)]. This mode was monitored as a function of RI to determine the bulk sensitivity of the pyramidal grating. For Ag nanopyramidal gratings, the bulk sensitivity was determined to be 410 nm/RIU [Fig. 3(c)], which is comparable to other grating platforms on a 400-nm pitch.

The reflection spectra from nanopyramidal plasmonic crystals also revealed that the SPP resonances became narrower at higher excitation angles [Fig. 3(b)]; at high θ, SPPs have longer wavelengths which lead to increased propagation distances and coherent coupling between more unit cells. Thus, although the FOM for an Ag pyramidal grating was 34 at a low excitation angle of $\theta = 13°$ (FWHM = 13.6 or 0.0378 eV), and increased to 85 at a high angle of $\theta = 64°$ (FWHM = 5.2 nm or 0.0085 eV) [Fig. 3(d)]. This result is the highest value reported for transducers at optical frequencies and suggests that designing plasmonic crystals that produce extremely narrow SPP resonances with FWHM < 6 nm is important for high performance sensors.

Spectral bandwidths of the plasmonic resonances can also be tuned by making modifications to both unit cell size and shape.[16] By simple

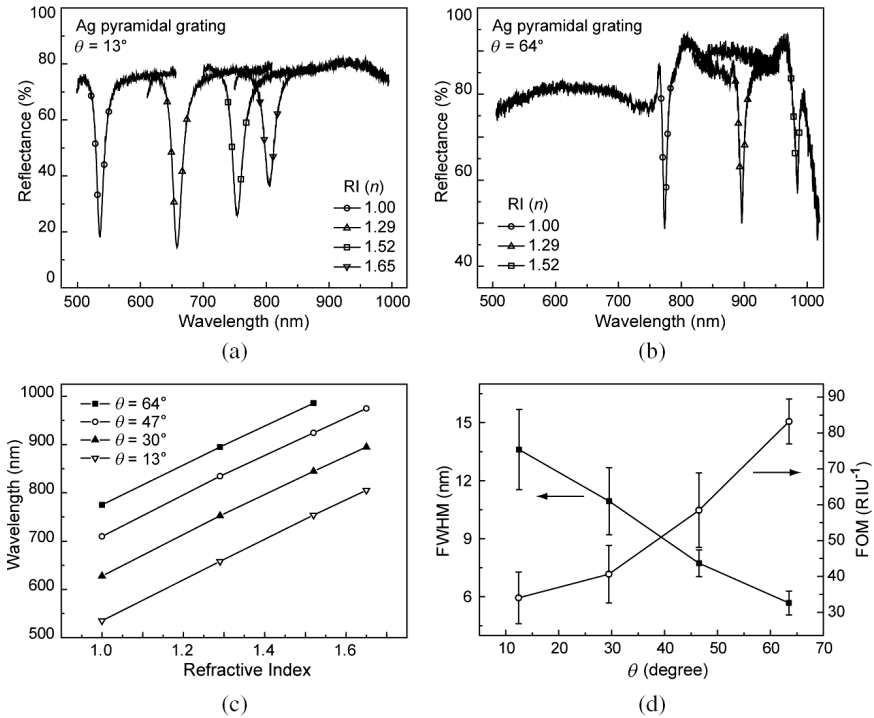


Fig. 3. Angle-dependent bulk RI sensitivity of an Ag nanopyramidal gratings. Reflectance measurements at (a) low excitation angle $\theta = 13°$ and (b) high excitation angle $\theta = 64°$ under different refractive indices which show (c) large resonance peak shifts (410 nm/RIU) of the $(-1, 0)$ SPP mode. (d) FOM and FWHM of the Ag nanopyramidal gratings exhibited strong dependence on angle. [Reproduced with permission from Ref. (65). Copyright 2008 PNAS]

modifications to the fabrication process, we generated unit cells with different dimensions and geometry. Nanopyramidal gratings with unit cell edge lengths of 220 nm and 320 nm exhibited SPP resonances with FWHMs of 7.1 nm and 12.8 nm, respectively, under the same excitation conditions. Changing the shape of the unit cells from 3D square-base pyramids to 2D circular holes in a metal film ($d = 100$ nm) increased the FWHM of the resonance and hence adversely affected the FOM. These observations suggest that greater bulk RI sensitivities could be achieved by nanostructured films with smaller unit cells and more optimized geometries.

4. Molecular Sensing Using 3D Plasmonic Crystals

4.1. *Nanopyramidal gratings: sensing under reflection conditions*

Because the response to changes in bulk environments for plasmonic crystals is independent of excitation angle, bulk RI sensitivity may not be a good indicator for local detection events on the molecular scale. We investigated whether the sensitivity of nanopyramidal gratings could be further tuned to respond to variations in RI localized at the metal-dielectric interface. To determine whether an Au nanopyramidal grating could detect a surface response, we assembled N polyelectrolyte bilayers with known thickness and RI on the surface in a layer-by-layer manner.[57] Each bilayer was composed of poly(dimethyldiallylammonium chloride) (PDDA), a positively charged polymer, and double stranded DNA, a negatively charged species with a characteristic thickness of ca. 6 nm (PDDA layer = 4 nm and DNA layer = 2 nm).[67] Figure 4 shows that the surface sensitivity of the plasmonic crystal strongly depends on incident θ where the slopes of the resonance wavelength shift $\Delta\lambda$ per N bilayer relation decreased as θ increased unlike the case for *bulk* sensitivities as in Fig. 3.[65]

The angle-dependent or surface sensitivity can be understood by considering the evanescent nature of the SPP field. The decay length of SPPs is directly related to θ, which determines the resonance wavelength of the

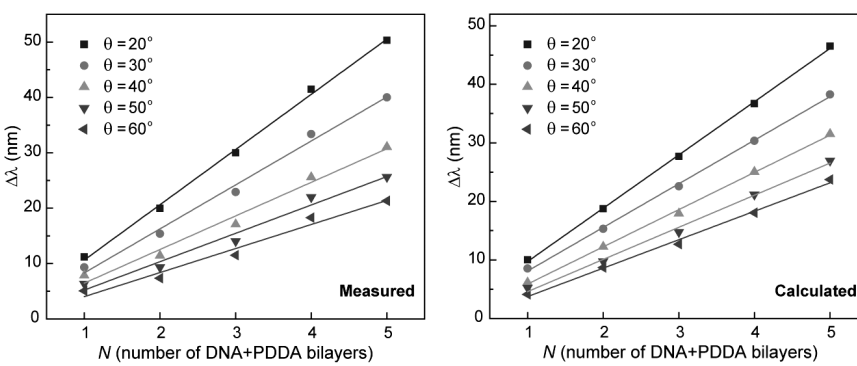


Fig. 4. Angle-dependent surface sensitivity of molded Au nanopyramidal gratings. (a) Experimental demonstration of how surface sensitivity was angle-dependent. The slopes of the $\Delta\lambda$ vs. N relation decrease as θ increases, a trend that was not observed for bulk sensitivity. (b) Calculated results were in agreement with (a). [Adapted with permission from Ref. (57). Copyright 2010 American Chemical Society]

plasmonic crystal.[68] As θ increases, the SPP resonance red-shifts to longer wavelengths, which in turn increases the decay length and reduces $\Delta\lambda$. Thus, excitation at low θ produced SPP fields with shorter decay lengths that could probe molecular monolayers with greater sensor resolution. Biotin-avidin interactions on Au nanopyramidal gratings confirmed this trend for a protein monolayer. The resonance wavelength shift of the sensor was larger at low excitation angles ($\theta = 13°$) compared to high angles ($\theta = 64°$).[65] Thus, the angle-dependence of the surface sensitivity intrinsic to grating-based sensors is an important parameter to consider in the development of improved sensing platforms.

The surface sensitivity of plasmonic crystals can also be tuned by plasmonic material selection. Although conventional materials, such as Au and Ag, have earned their place in plasmonic sensing,[17,69,70] the search for other materials that could be used for sensing is important as the limits of current sensing technologies are approached. Also, different materials enable the opportunity to expand the types of chemical reactions not easily accessible through Au-thiol chemistry, such as catalysis.[71–73] The fabrication of wafer-scale plasmonic crystals enables the screening of multiple plasmonic metals to test for potential biosensor materials. We evaluated the sensitivity of three different metals (Ag, Au and Pd) to changes in local RI using biotin-avidin as a model system.

Optical characterization at different excitation angles showed that the Ag pyramidal grating produced a narrower $(-1, 0)$ SPP resonance than either Au or Pd at all θ [Figs. 5(a–c)]. Surprisingly, the wavelength shift of the resonance from the Ag grating (2.7 nm) after avidin binding was half of that observed from a Au grating (5.0 nm) [Fig. 5(d)]. Notably, this 5.0-nm shift observed from biotin-avidin binding on a Au pyramidal grating was larger than the resonance wavelength shifts reported for the protein-receptor interaction previously,[21,74] which again shows that the angular excitation of the nanopyramidal array can lead to improved surface sensitivity.

The combination of Pd and the molded plasmonic crystal substrates pave the way for nontraditional and less efficient plasmonic materials to be interfaced with emerging optical sensing technologies. Low SPP coupling efficiency has made Pd an unlikely material for use in sensing; however, the wavelength shift on biotinylated Pd gratings was 4.9 nm, which is comparable to the shift observed on an Au grating and greater than the value from an Ag grating [Fig. 5(d)]. This result is significant because the broad and asymmetric resonance of Pd did not obscure the large

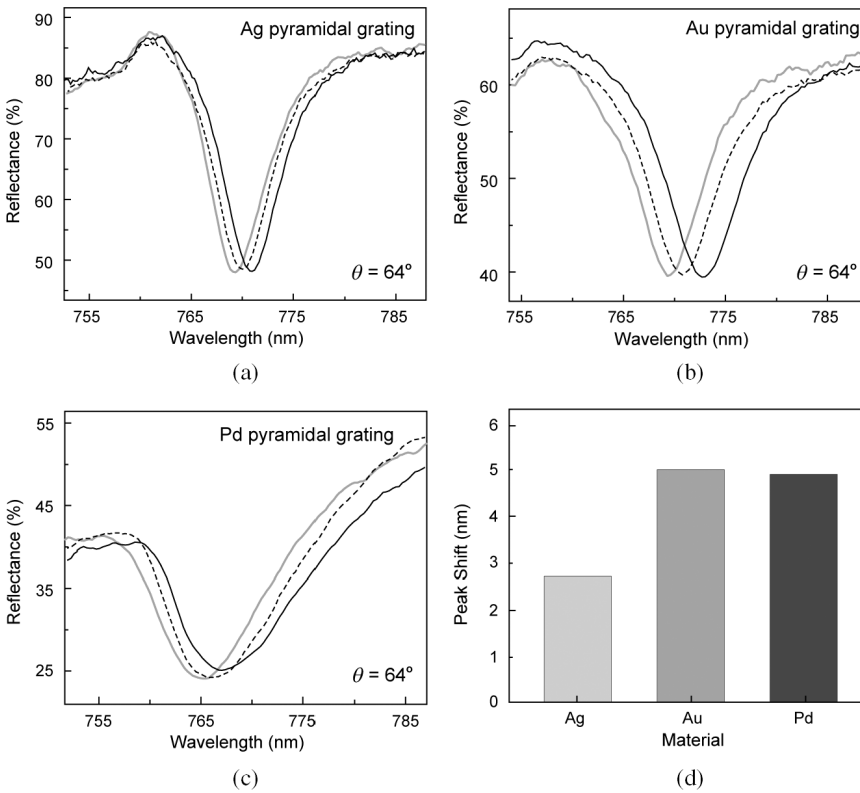


Fig. 5. Surface sensitive detection of biomolecules on different plasmonic materials (Ag, Au, and Pd). (a–c) Reflectance spectra of (a) Ag, (b) Au, and (c) Pd pyramidal gratings at different stages of biofunctionalization: Clean surface (light gray line), biotinylated surface (dotted black line), and avidin binding (solid black line). (d) The resonance wavelength shift of the $(-1, 0)$ SPP mode was dependent on material. Ag produced the smallest peak shift upon avidin binding to the biotinylated surface. [Adapted with permission from Ref. (65). Copyright 2008 PNAS]

resonance shifts, and Pd chemistry is similar to the surface chemistry of Au for the self-assembly of thiolated molecules. Hence, Pd may be a promising material for sensing monolayer coverage of surface-bound proteins. These results also reveal that Ag may not be preferable for detecting molecules despite exhibiting the highest bulk sensitivity among Au and Pd.

4.2. *3D Nanohole arrays: sensing under transmission conditions*

3D nanohole arrays are unique plasmonic crystals with enhanced optical properties that provide new opportunities for biosensing because they

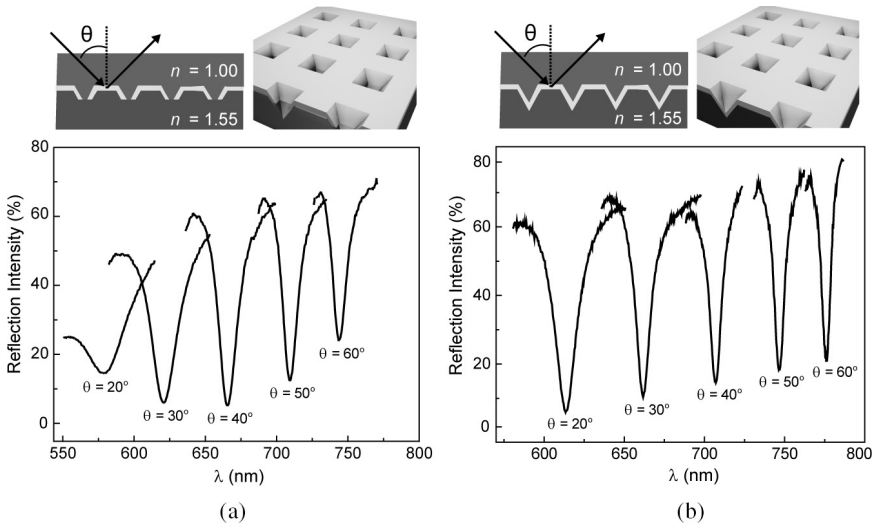


Fig. 6. Zero-order reflectance measurements from (a) an Au 3D nanohole array and (b) an Au nanopyramidal grating under angular excitation in air ($n = 1.00$). Both types of 2D plasmonic crystals display a narrow $(-1, 0)$ SPP resonance under reflection excitation conditions. These resonances behave similarly as θ increases. [Adapted with permission from Ref. (57). Copyright 2010 American Chemical Society]

support, simultaneously, two types of surface plasmons.[58] The reflection spectrum of a 3D nanohole array [Fig. 6(a)] is similar to that of a nanopyramidal grating [Fig. 6(b)], where a very narrow $(-1, 0)$ SPP resonance is observed at all θ and the resonance red-shifts as θ increases.

Under transmission conditions, however, 3D nanohole arrays can show strong spectral signatures from both SPPs and LSPs, which can be excited simultaneously on the same substrate at excitation angles $\theta = 45°$.[58] We have taken advantage of the inherent properties of SPP and LSP signals to obtain more information regarding the bulk and surface sensitivities of plasmonic crystals. The response of the SPs to changes in the local environment was again characterized by the layer-by-layer assembly of polyelectrolytes (PDDA and DNA). The LSPs and SPPs displayed characteristic responses to changes in local RI, where the LSP resonance wavelength shift saturated as the polyelectrolyte multilayer approached a thickness of 60 nm (11 layers), but the SPP response continued to increase linearly [Fig. 7(b)]. This result confirmed that both types of SPs could be independently considered to provide separate and different bulk and local RI information for future multiplexing experiments. Significantly, the decay lengths of the 3D nanohole arrays were longer than those supported by

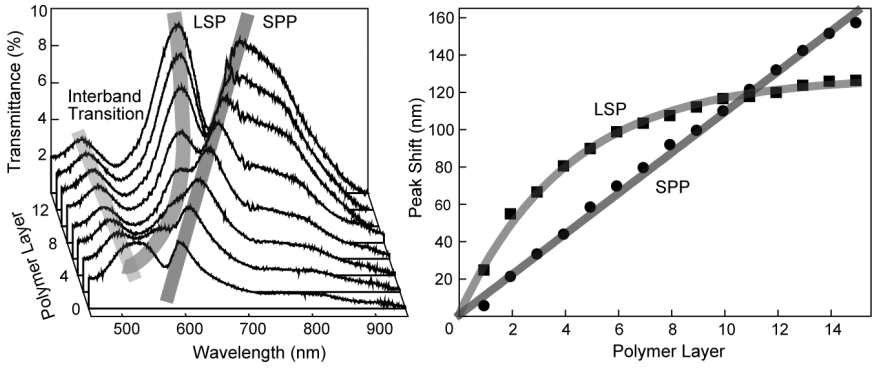


Fig. 7. Optical response of LSP and SPP resonances to changes in local RI determined from layer-by-layer assembly of polyelectrolytes on 3D nanohole arrays. (a) Zero-order transmission spectra at different thicknesses of deposited polymer measured at $\theta = 45°$. (b) Resonance wavelength peak shift plotted against the number of polyelectrolyte layers on the metal surface. The LSP response saturates after 11 layers (ca. 60 nm thickness), but SPP response increases linearly. [Reprinted with permission from Ref. (58). Copyright 2010 American Chemical Society]

metal nanoparticles[75] or arrays of particles (<30 nm).[76] These LSP fields could improve the resolution of smaller biomolecules on nanostructured gratings.

5. Biosensing with 3D Plasmonic Crystals in Real-Time

We have discussed above that 3D plasmonic crystals display surface sensitivities that can resolve molecular monolayers from endpoint assays. For quantitative bioanalysis, however, sensing platforms should provide information about dynamic interactions between biomolecules. Besides discovering that the excitation angle can be tailored to improve surface sensitivity for biosensing, we found that θ could be tuned to increase S/N for real-time binding measurements.

We performed kinetic studies using a model system. First, a receptor-ligand pair (lectin from *Arachis hypogaea* and Galβ(1,3)Gal-NAc disaccharide ligand) with low binding affinity (5 μM)[77] was chosen in order to assess the real-time performance of the sensor quantitatively. To prepare the sensor, we modified the Au surface of a molded plasmonic crystal with a mixed self-assembled monolayer (SAM) presenting both maleimide and tri(ethylene glycol) functionality at 5% surface coverage. The surface was

then treated with a thiol-terminated Galβ(1,3)Gal-NAc which reacts with the maleimide group and becomes covalently immobilized. Because the maleimide was expressed at 5% density on the grating, the population of Galβ(1,3)Gal-NAc on the surface should be the same. This surface coverage was optimized to avoid crowding upon binding of the lectin and to maintain the non-fouling properties of the tri(ethylene glycol).[78]

By monitoring the interaction of lectin and Galβ(1,3)Gal-NAc [Fig. 8(a)], the real-time response of the nanopyramidal grating was

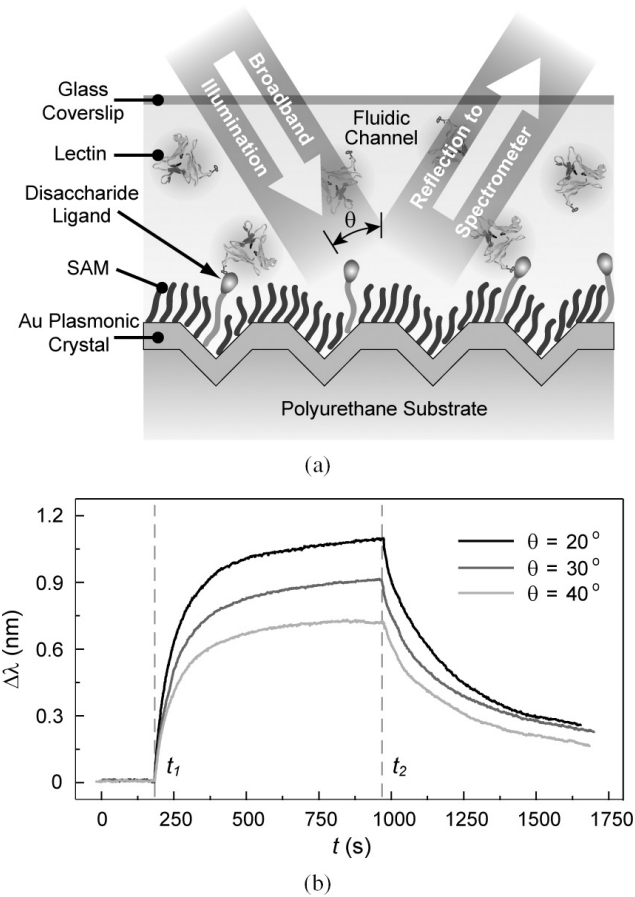


Fig. 8. Optimizing signal-to-noise for real-time biodetection. (a) Scheme of the sensing experiment. (b) Binding curves acquired under different θ exhibit angle-dependent surface sensitivity in a real-time format. [Reproduced with permission from Ref. (57). Copyright 2010 American Chemical Society]

evaluated from the kinetic binding curve analysis. The binding curves were acquired by monitoring the $(-1, 0)$ SPP resonance of the pyramidal grating at 2 s intervals. This time resolution was determined from the calibration of the sensing surface, which was achieved by injecting NaCl solutions of different concentrations (and RI) and monitoring the $\Delta\lambda$. From these measurements, the bulk RI resolution of the sensor was determined to be 9.7×10^{-7} RIU, comparable to a commercial SPR.

To obtain sensorgrams for lectin, we injected a 5 μM concentration of lectin into a microfluidic cell at a constant flow rate of 60 μL/min to ensure that protein binding was not diffusion-limited. Protein association was observed after the injection at t_1 and prolonged washing of the sensor with phosphate buffer after surface saturation began to show protein dissociation after t_2. These binding curves were collected under excitation angles $\theta = 20°$, 30° and 40° [Fig. 8(b)]. The sensor response was greatest at $\theta = 20°$ and decreased with increasing θ, consistent with the angle-dependent surface sensitivity determined for these plasmonic crystals discussed in Section 4.1.

The angle-dependent sensitivity of the pyramidal gratings was used to optimize the S/N of the sensor for real-time detection.[57] Identifying an angle of incidence that would maximize S/N is important because the FWHM and $\Delta\lambda$ are related. That is, although narrow resonances are obtained at larger θ, the shifts in peak position are smaller. Thus, both parameters are critical to determine S/N since the observed noise levels from measurements taken at all three incident angles were comparable. We discovered that the optimal signal-to-noise ratio ($\mathrm{S/N} = 1.1027\,\mathrm{nm}/1.5 \times 10^{-3}\,\mathrm{nm} = 698$) was determined by the lowest excitation angle ($\theta = 20°$), which produced the greatest resonance wavelength shift but not necessarily the narrowest resonance.

The kinetic constants for the association (k_a) and dissociation (k_d) of the interaction were obtained by fitting the lectin binding curves using the Langmuir model.[79] Notably, the dynamic binding constants ($k_a \approx 1810\,\mathrm{M^{-1}s^{-1}}$, $k_d \approx 2.56 \times 10^{-3}\,\mathrm{s^{-1}}$) were similar at different angles of incidence $\theta = 20°$, 30° and 40°, which suggests that the surface sensitivities at higher angles for the pyramidal gratings are also sufficient for biosensing. This result further suggests that the quality of the nanostructured substrates was relatively consistent from batch-to-batch preparation, and the average distribution of Galβ(1,3)Gal-NAc on the surface was reproducible using the self-assembly functionalization scheme at 5% surface density. Moreover, these kinetic constants agreed with those measured by commercial SPR instruments ($k_a = 1720\,\mathrm{M^{-1}s^{-1}}$, $k_d = 1.73 \times 10^{-3}\,\mathrm{s^{-1}}$),

which suggests that the pyramidal nanostructures in the plasmonic crystal do not significantly perturb protein binding.

6. Conclusions and Outlook

In this chapter, we reviewed our recent work on plasmonic crystals with 3D unit cells and their contributions to SP-based biosensing. We have demonstrated that plasmonic crystals are invaluable platforms for sensing because of their surface sensitivity. The state-of-the-art (Biacore) systems and commercially available line diffraction gratings are currently characterized by their bulk sensitivities, which does not necessarily indicate their sensitivity to binding events that occur within tens of nanometers, the average size of molecules and proteins, from the sensor surface. We were able to distinguish surface sensitivity from bulk sensitivity with the plasmonic crystals by (1) tuning the excitation angle of incident light and (2) modifying their unit cell geometries and creating engineered surfaces characterized by narrow SPP resonances. Although the surface sensitivities we have found for the plasmonic crystals are comparable to the bulk sensitivities for conventional sensors, we anticipate that further development of these types of substrates will improve sensitivity for surface interactions. In addition, the prospect for 3D unit cells in sensing is particularly promising. A 3D nanostructured film, such as the 3D nanohole array, is especially advantageous for sensing because it supports both SPPs and local field enhancements from LSPs. Simultaneous excitation of both SPPs and LSPs on the same substrate can be used to provide complementary information about bulk and surface RI at the same time, which could make label-free, multi-modal readout systems possible for high throughput biodiagnostics. Furthermore, taking advantage of the enhanced fields by placing the biorecognition element in the high sensitivity region generated by LSPs could bring sensing on continuous film transducers from monolayers of large molecules to single and small molecules.

Acknowledgments

This work was supported by the National Science Foundation under NSF Award Number CMMI-0826219, the Nanoscale Science and Engineering initiative under NSF Award Number EEC-0647560, and the Center of Cancer Nanotechnology Excellence at Northwestern University (NCI U54CA119341).

References

1. J. W. Chung, S. D. Kim, R. Bernhardt, and J. C. Pyun, *Sensors and Actuators B: Chemical* **111**, 416 (2005).
2. R. F. Dutra and L. T. Kubota, *Clinica. Chimica. Acta.* **376**, 114 (2007).
3. E. Mauriz, A. Calle, J. Manclús, A. Montoya, and L. Lechuga, *Analytical and Bioanalytical Chemistry* **387**, 1449 (2007).
4. J. S. Lee, M. Han, and C. Mirkin, *Angewandte Chemie International Edition* **46**, 4093 (2007).
5. A. D. Taylor, J. Ladd, Q. Yu, S. Chen, J. Homola, and S. Jiang, *Biosensors and Bioelectronics* **22**, 752 (2006).
6. J. Homola, *Chemical Reviews* **108**, 462 (2008).
7. M. A. Cooper, *Nat. Rev. Drug Discov.* **1**, 515 (2002).
8. Z. H. Wang and G. Jin, *Analytical Chemistry* **75**, 6119 (2003).
9. G. Gauglitz, *Analytical and Bioanalytical Chemistry* **381**, 141 (2005).
10. B. Liedberg, C. Nylander, and I. Lunström, *Sensors and Actuators* **4**, 299 (1983).
11. R. B. M. Schasfoort and A. J. Tudos, *Handbook of Surface Plasmon Resonance*. The Royal Society of Chemistry: Cambridge, UK, 2008.
12. J. Dostálek and J. Homola, *Sensors and Actuators B: Chemical* **129**, 303 (2008).
13. E. Kretschmann and H. Raether, *Z. Naturforsch.* **23A**, 2135 (1968).
14. K. L. Kelly, E. Coronado, L. L. Zhao, and G. C. Schatz, *The Journal of Physical Chemistry B* **107**, 668 (2002).
15. P. K. Jain, K. S. Lee, I. H. El-Sayed, and M. A. El-Sayed, *The Journal of Physical Chemistry B* **110**, 7238 (2006).
16. H. Gao, W. Zhou, and T. W. Odom, *Advanced Functional Materials* **20**, 529 (2010).
17. A. J. Haes, S. Zou, G. C. Schatz, and R. P. Van Duyne, *The Journal of Physical Chemistry B* **108**, 6961 (2004).
18. P. Kvasnicka and J. Homola, *Biointerphases* **3** FD4 (2008).
19. J. Prikulis, P. Hanarp, L. Olofsson, D. Sutherland, and M. Käll, *Nano Letters* **4**, 1003 (2004).
20. J. N. Anker, W. P. Hall, O. Lyandres, N. C. Shah, J. Zhao, and R. P. Van Duyne, *Nat. Mater.* **7**, 442 (2008).
21. T. Rindzevicius, Y. Alaverdyan, A. Dahlin, F. Höök, D. S. Sutherland, and M. Käll, *Nano Letters* **5**, 2335 (2005).
22. A. B. Dahlin, J. O. Tegenfeldt, and F. Höök, *Analytical Chemistry* **78**, 4416 (2006).
23. G. J. Nusz, S. M. Marinakos, A. C. Curry, A. Dahlin, F. Hook, A. Wax, and A. Chilkoti, *Analytical Chemistry* **80**, 984 (2008).
24. C. R. Yonzon, E. Jeoung, S. Zou, G. C. Schatz, M. Mrksich, and R. P. Van Duyne, *Journal of the American Chemical Society* **126**, 12669 (2004).
25. J. C. Hulteen and R. P. Van Duyne, *Journal of Vacuum Science & Technology A* **13**, 1553 (1995).
26. H. Fredriksson, Y. Alaverdyan, A. Dmitriev, C. Langhammer, D. Sutherland, M. Zäch, and B. Kasemo, *Advanced Materials* **19**, 4297 (2007).

27. W. P. Hall, J. N. Anker, Y. Lin, J. Modica, M. Mrksich, and R. P. Van Duyne, *Journal of the American Chemical Society* **130**, 5836 (2008).
28. E. M. Hicks, S. Zou, G. C. Schatz, K. G. Spears, R. P. Van Duyne, L. Gunnarsson, T. Rindzevicius, B. Kasemo, and M. Käll, *Nano Letters* **5**, 1065 (2005).
29. A. B. Dahlin, S. Chen, M. P. Jonsson, L. Gunnarsson, M. Käll, and F. Höök, *Analytical Chemistry* **81**, 6572 (2009).
30. A. D. McFarland and R. P. Van Duyne, *Nano Letters* **3**, 1057 (2003).
31. M. M. Miller and A. A. Lazarides, *The Journal of Physical Chemistry B* **109**, 21556 (2005).
32. S. Zou and G. C. Schatz, *Journal of Chemical Physics* **121**, 12606 (2004).
33. A. Dmitriev, C. Hägglund, S. Chen, H. Fredriksson, T. Pakizeh, M. Käll, and D. S. Sutherland, *Nano Letters* **8**, 3893 (2008).
34. B. Brian, B. Sepúlveda, Y. Alaverdyan, L. M. Lechuga, and M. Käll, *Opt. Express* **17**, 2015 (2009).
35. A. Dahlin, M. Zäch, T. Rindzevicius, M. Käll, D. S. Sutherland, and F. Höök, *Journal of the American Chemical Society* **127**, 5043 (2005).
36. B. D. Andreas, P. J. Magnus, and H. Fredrik, *Advanced Materials* **20**, 1436 (2008).
37. L. Feuz, P. Jönsson, M. P. Jonsson, and F. Höök, *ACS Nano* **4**, 2167 (2010).
38. G. Raschke, S. Kowarik, T. Franzl, C. Sönnichsen, T. A. Klar, J. Feldmann, A. Nichtl, and K. Kürzinger, *Nano Letters* **3**, 935 (2003).
39. K. M. Mayer, S. Lee, H. Liao, B. C. Rostro, A. Fuentes, P. T. Scully, C. L. Nehl, and J. H. Hafner, *ACS Nano* **2**, 687 (2008).
40. T. A. Bendikov, A. Rabinkov, T. Karakouz, A. Vaskevich, and I. Rubinstein, *Analytical Chemistry* **80**, 7487 (2008).
41. N. Nath and A. Chilkoti, *Analytical Chemistry* **76**, 5370 (2004).
42. T. Endo, K. Kerman, N. Nagatani, H. M. Hiepa, D.-K. Kim, Y. Yonezawa, K. Nakano, and E. Tamiya, *Analytical Chemistry* **78**, 6465 (2006).
43. A. M. Armani, R. P. Kulkarni, S. E. Fraser, R. C. Flagan, and K. J. Vahala, *Science* **317**, 783 (2007).
44. K. M. Mayer, F. Hao, S. Lee, P. Nordlander, and J. H. Hafner, *Nanotechnology* **21**, 255503 (2010).
45. A. Curry, G. Nusz, A. Chilkoti, and A. Wax, *Appl. Opt.* **46**, 1931 (2007).
46. T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, *Nature* **391**, 667 (1998).
47. A. G. Brolo, R. Gordon, B. Leathem, and K. L. Kavanagh, *Langmuir* **20**, 4813 (2004).
48. C. Thirstrup, W. Zong, M. Borre, H. Neff, H. C. Pedersen, and G. Holzhueter, *Sensors and Actuators B: Chemical* **100**, 298 (2004).
49. J.-C. Yang, J. Ji, J. M. Hogle, and D. N. Larson, *Nano Letters* **8**, 2718 (2008).
50. C. J. Alleyne, A. G. Kirk, R. C. McPhedran, N.-A. P. Nicorovici, and D. Maystre, *Opt. Express* **15**, 8163 (2007).
51. K. H. Yoon, M. L. Shuler, and S. J. Kim, *Opt. Express* **14**, 4842 (2006).
52. P. Adam, J. Dostálek, and J. Homola, *Sensors and Actuators B: Chemical* **113**, 774 (2006).

53. M. H. Lee, H. Gao, and T. W. Odom, *Nano Letters* **9**, 2584 (2009).
54. D. C. Cullen, R. G. W. Brown, and C. R. Lowe, *Biosensors* **3**, 211 (1987).
55. C. R. Lawrence, N. J. Geddes, D. N. Furlong, and J. R. Sambles, *Biosensors and Bioelectronics* **11**, 389 (1996).
56. M. E. Stewart, N. H. Mack, V. Malyarchuk, J. A. N. T. Soares, T.-W. Lee, S. K. Gray, R. G. Nuzzo, and J. A. Rogers, *Proceedings of the National Academy of Sciences* **103**, 17143 (2006).
57. H. Gao, J.-C. Yang, J. Y. Lin, A. D. Stuparu, M. H. Lee, M. Mrksich, and T. W. Odom, *Nano Letters* **10**, 2549 (2010).
58. J.-C. Yang, H. Gao, J. Y. Suh, W. Zhou, M. H. Lee, and T. W. Odom, *Nano Letters* **10**, 3173 (2010).
59. J. Henzie, M. H. Lee, and T. W. Odom, *Nat Nano* **2**, 549 (2007).
60. T. W. Odom, *MRS Bulletin* **35**, 66 (2010).
61. Y. Xia, E. Kim, X.-M. Zhao, J. A. Rogers, M. Prentiss, and G. M. Whitesides, *Science* **273**, 347 (1996).
62. Y. Xia and G. M. Whitesides, *Angewandte Chemie International Edition* **37**, 550 (1998).
63. T. W. Odom, J. C. Love, D. B. Wolfe, K. E. Paul, and G. M. Whitesides, *Langmuir* **18**, 5314 (2002).
64. L. J. Sherry, S.-H. Chang, G. C. Schatz, R. P. Van Duyne, B. J. Wiley, and Y. Xia, *Nano Letters* **5**, 2034 (2005).
65. H. Gao, J. Henzie, M. H. Lee, and T. W. Odom, *Proceedings of the National Academy of Sciences* **105**, 20146 (2008).
66. W. L. Barnes, W. A. Murray, J. Dintinger, E. Devaux, and T. W. Ebbesen, *Physical Review Letters* **92**, 107401 (2004).
67. R. Pei, X. Cui, X. Yang, and E. Wang, *Biomacromolecules* **2**, 463 (2001).
68. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. Springer-Verlag: Berlin and New York, 1988.
69. K.-S. Lee and M. A. El-Sayed, *The Journal of Physical Chemistry B* **110**, 19220 (2006).
70. R. Elghanian, J. J. Storhoff, R. C. Mucic, R. L. Letsinger, and C. A. Mirkin, *Science* **277**, 1078 (1997).
71. K. Lin, Y. Lu, J. Chen, R. Zheng, P. Wang, and H. Ming, *Opt. Express* **16**, 18599 (2008).
72. S. M. Williams, K. R. Rodriguez, S. Teeters-Kennedy, A. D. Stafford, S. R. Bishop, U. K. Lincoln, and J. V. Coe, *The Journal of Physical Chemistry B* **108**, 11833 (2004).
73. P. Tobiska, O. Hugon, A. Trouillet, and H. Gagnaire, *Sensors and Actuators B: Chemical* **74**, 168 (2001).
74. A. De Leebeeck, L. K. S. Kumar, V. de Lange, D. Sinton, R. Gordon, and A. G. Brolo, *Analytical Chemistry* **79**, 4094 (2007).
75. C. J. Murphy, A. M. Gole, S. E. Hunyadi, J. W. Stone, P. N. Sisco, A. Alkilany, B. E. Kinard, and P. Hankins, *Chemical Communications* 544 (2008).
76. T. Sannomiya, P. K. Sahoo, D. I. Mahcicek, H. H. Solak, C. Hafner, D. Grieshaber and J. Vörös, *Small* **5**, 1889 (2009).

77. J. D. Milton, D. G. Fernig, and J. M. Rhodes, *Glycoconjugate Journal* **18**, 565 (2001).
78. B. T. Houseman, E. S. Gawalt, and M. Mrksich, *Langmuir* **19**, 1522 (2002).
79. R. I. Masel, *Principles of Adsorption and Reaction on Solid Surfaces*. 1st ed. (Wiley-Interscience: New York, 1996).

TUNABLE AND ACTIVE OPTICAL NEGATIVE INDEX METAMATERIALS

Shumin Xiao and Vladimir M. Shalaev*

School of Electrical and Computer
Engineering and Birck Nanotechnology Center,
Purdue University, West Lafayette, Indiana 47907, USA
**shalaev@ecn.purdue.edu*

Optical negative-index metamaterials (NIMs) have attracted a significant amount of research attention because of their unique properties and important applications including superlenses and cloaking. Recently, after numerous fundamental studies, research interest has turned to the realization of real NIM applications. These applications are strongly limited by the long and fixed working wavelength and the significant losses in NIM designs. In this chapter, we demonstrate our recent progress in the design, fabrication, and characterization of optical NIMs. First, we demonstrate NIMs operating in optical wavelength range. Then we show a method to adjust the refractive index of dielectric layers in a NIM and thereby tune the resonant wavelength of NIM by tens of nanometers. Finally, and most importantly for real-world NIM applications, we introduce active materials into NIMs and show that loss compensation can significantly improve the essential parameters of NIMs. We believe our studies will be very interesting and promising for real NIM applications in the near future.

1. Introduction

Negative index metamaterials (NIMs) are artificially tailor-made composites of noble metals (silver or gold) and dielectrics to provide a negative refractive index that cannot be readily obtained in nature.[1–4] This kind of composite material has attracted a significant amount of research

attention due to interesting potential applications. For example, in 2000, superlensing using NIMs was predicted by John Pendry.[5] A near-field metamaterial superlens was successfully demonstrated in experiment soon after.[6] Recently, NIM research has been further inspired by the possibility of hiding materials from detection (cloaking).[7,8] The fast progress in the area of optical NIMs (ONIMs) and maturing nanofabrication techniques now enable us to mount a systematic approach toward the goal of full, system-level integration of light and NIMs, since these 'left-handed' materials provide a different approach for controlling light. However, several significant obstacles have to be overcome before NIMs can be applied outside of research labs in real devices. For example, the structure needs to be easily assembled with tunable elements so the response can be tuned during operation. Also, the loss in the nanostructure, which primarily arises from losses in the metal components, needs to be suppressed in order to improve the overall NIM performance.

A large number of exciting potential applications such as the superlens, have to work in the visible range, necessitating the design of NIMs that operate at wavelengths shorter than 700 nm. Progress in visible NIMs, which began at wavelengths as long as 1.5 μm, has most recently been pushed to a wavelength of only 715 nm[9] for applications in the visible and "near-visible" ranges. For real-world devices, progress in realizing negative indices in bulk metamaterials is especially important.[10] In addition, recent theoretical work proposing a negative index on a wrapped, coaxial waveguide has already pushed the NIM operating wavelength into the blue light range.[11] However this proposed design hasn't been realized due to fabrication limitations. Unfortunately, in real applications, wavelengths for three-dimensional metamaterials are still too long for applications such as superlens-based biological sensing and imaging devices; this is especially unfortunate since shorter excitation wavelengths would produce images with higher optical resolution. Therefore, pushing the operational wavelength of negative refractive index materials to shorter wavelength is a critical step to realize more exciting metamaterial applications.

Devices based on NIMs could come to include filters, modulators, amplifiers, detectors and resonators, among others.[12,13] The usefulness of such devices could be extended tremendously if the NIMs' response characteristics could be dynamically tuned. To provide full operational functionality for optical NIM devices, they need to be engineered by integrated tunable elements, which will be demonstrated in Section 4 in this chapter.

The most significant challenge in ONIM design is dealing with the losses incurred by the metallic elements in the structure. Eliminating losses in optical NIMs is critical for enabling their numerous potential applications. Loss is an extremely severe, detrimental factor in NIM designs. The performance of all NIM applications is significantly limited by the inherent and strong energy dissipation in metals, especially in the near infrared and visible wavelength ranges.[14,15] For example, to realize sub-wavelength resolution in a superlens, losses should decrease exponentially to obtain a linear improvement in the spatial resolution.[16] Generally, the losses are orders of magnitude too large for the envisioned applications, and the reduction of losses with optimized designs appears to be out of reach. One way of addressing this issue is to incorporate gain media into NIM designs.[17–20] However, whether NIMs with low loss can be achieved has been the subject of theoretical debate.[21,22]

This chapter will focus on how to solve the problems mentioned above. Section 2 discusses how to fabricate real ONIMs using current fabrication technologies. Section 3 discusses the realization of dynamically tunable metamagnetism by incorporating a liquid-crystal material into the nanostructure of the device. Section 4 discusses the first loss-free and active ONIMs in the world, which is a significant step forward in realizing the many unique applications of optical metamaterials. Finally, Section 5 summarizes the whole chapter and provides an outlook for the future of ONIM research.

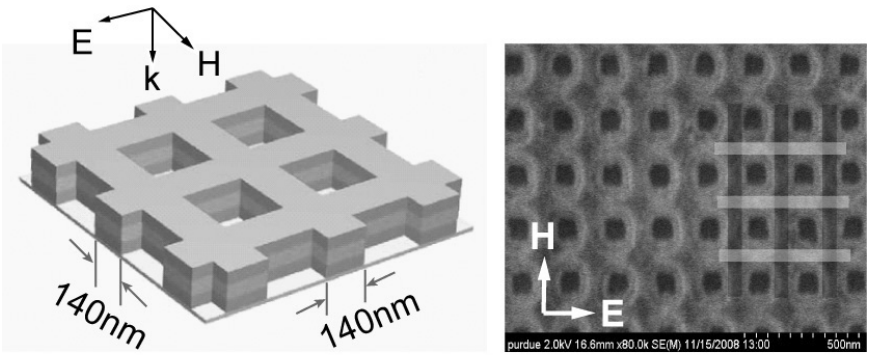
2. Negative Index Metamaterials in the Optical Range

In the optical range, it is easy to obtain a negative permittivity since all the noble metals have negative permittivity at frequencies below the plasma frequency. However, the permeability of all natural material at optical frequencies is strictly unity. So the main challenge to obtain relatively a negative refractive index at optical frequencies is to design structure that would yield a negative permeability.[23,24] Usually, in the visible optical frequencies, such a magnetic response is realized in practice via the excitation of the anti-symmetric mode between two coupled metallic nanostrips. The resonant frequency is strongly dependent on the size, the shape and thickness of the nano-scale structures.[25] Therefore, the key limitations in obtaining a magnetic resonance at shorter optical wavelengths are actually the limitations in modern fabrication techniques. For example, fishnet-based NIMs are created using e-beam lithography followed by a lift-off process.

According to our simulations, the geometry of the fishnet structure needs to be scaled down in order to shorten the working wavelength to visible optical range. However scaled-down geometry is limited by the minimum achievable dimensions of e-beam lithography because the e-beam resist will collapse for extremely small features due to forward and backward electron scattering (or reflection) during electron-beam writing. Therefore, the optimization of a design and its fabrication process become essential to successfully creating NIMs in visible optical range. The optimization of the design is used to increase the smallest dimension inside the nanostructure or pattern. The fabrication optimization is used to improve the quality of the whole structure. In this section, we solve these problems on a reduced unit cell in a fishnet designed for NIM structures. After carefully studying the transmittance, reflectance, and corresponding numerical simulations, we succeeded in optimizing our design and fabrication processes to produce a sample that behaves as a NIM with the real part of its refractive index $n' = -0.25$ at 580 nm with a figure of merit ($FOM = |n'|/n''$) of 0.3.

2.1. *Sample fabrication*

In our work, the fishnet nanostructure is still used because it has generally been proven to be a vigorous and scalable design at visible wavelengths,[2,9,26,27] including the successful demonstration of a negative refractive index at 715 nm.[9] The structure and geometry of the fishnet are shown in Fig. 1(a). The fishnet structure consists of two sets of nanostrips and is periodic in both of the mutually perpendicular directions: along the electric and magnetic field vector directions. One pair of the strips act, as off resonant wires and, at the wavelength where the magnetic resonance occurs in another pairs of strips, they simply provide a broad negative permittivity background. As a result, the fishnet structure works with resonant magnetic strips combined with non-resonant electric strips. Before the real fabrication, the geometry of the fishnet design was optimized for yellow light using numerical simulations through a commercial finite-element package (COMSOL Multiphysics). These simulations give the optimum parameters for the highest FOM for a design with unit cell sizes of around 220 nm for both periodicities and a nanostrip with a bottom width of about 140 nm. In the design, the structure is isotropic and expected to obtain the same optical properties in both directions along the two sets of strips. The metal coverage ratio, or pattern ratio, is as high as about 70%.



(a) (b)

Fig. 1. (a) Schematic of a metamaterial consisting of a fishnet structure. The geometric parameters are 220-nm periodicity along two directions, and nanostrip bottom widths are 140 nm in both directions. The layers and their thicknesses within the sample are (top to bottom): Al_2O_3 (10 nm), Ag (43 nm), Al_2O_3 (45 nm), Ag (43 nm), Al_2O_3 (10 nm) — substrate. (b) A top view from a representative FESEM image of the structure. The sets of "magnetic strips" and "electric strips" are shown in darker and lighter highlights respectively. The primary polarization is shown in the panel and is defined as the electric field lying in the plane of the sample and perpendicular to the magnetic strips. Reprinted with permission from Ref. (30). Copyright (2009) OSA.

The sample was fabricated using electron-beam lithography on positive photo-resist poly(methyl-methacrylate) photo-resist (PMMA) followed by electron-beam evaporation of the metal/dielectric layers and lift-off. A glass coated with a 15-nm-thick indium–tin–oxide (ITO) layer is used as substrate in order to conduct the electrons during the e-beam writing. Usually, fabrication of such a high-density fishnet is very challenging because the forward and backward scattering of electrons within the resist will destroy the designed structures and cause the resist to collapse. To solve this problem, a very low electron-beam current and a longer baking time for the PMMA is used to reduce the dose for electrons required to ionize the PMMA molecules. As a result, the scattering of electrons during e-beam writing will also be reduced. Also, a Leica VB6 system with a 100 kV accelerating voltage was used to pattern the nanostructure on the PMMA. We note that using a high accelerating voltage of 100 kV will decrease the backward scattering because electrons after acceleration by a larger voltage will have higher energy and therefore travel through the PMMA more easily. Additionally, the surface roughness of the metal films can broaden the resonance, which will greatly decrease the final FOM.[15] Due to this, a

low deposition rate of 0.2 Å/s was used during the deposition of the metal layers in order to obtain a smoother final film to reduce the loss.[16]

The stacked layers of the fabricated structure include two 43-nm perforated silver layers separated by a 45-nm layer of alumina. In addition, a 10-nm-thick layer of alumina was deposited above and below the structure to protect the silver from deterioration and improve the adhesion. Because the value of the real part of the permittivity of silver at shorter wavelengths is less negative than the value at longer wavelengths, a larger ratio of metal to dielectric layer thicknesses is needed. As a result, the thickness of each silver layer is slightly larger compared to that of previous samples, which operated at longer wavelength.[3] The structure represents a square lattice with a periodic unit cell of 220 nm. The fishnet's stacked layers have a trapezoidal shape for the cross section due to the limitation of e-beam evaporation during fabrication.[25] As a result, the width of the strip near the substrate (the bottom width) is larger than the top width of the strip near the air interface (the top width). Figure 1(b) shows the field-emission scanning electron microscope (FESEM) image of the fabricated structure viewed from the top. The sets of "magnetic strips" and "electric strips" are shown in darker and lighter highlights, respectively. The primary polarization is defined as the polarization where the electric field vector of the incident light is perpendicular to the set of magnetic strips. The voids in the structure incline to be much more rounded than the simulated design; this is due to much more back scattering of electrons at the corners of the voids in e-beam writing. Several relative smaller electric strips can also be observed in Fig. 1(b). These smaller widths are generated due to the irregular shape of the electron-beam spot in e-beam writing. The 5 nm electron-beam spot is not always ideally circular in shape when the beam current is as low as 400 nA used in our experiment.

The sample was optically characterized using far-field transmittance and reflectance spectra obtained with the incident light polarized as the primary polarization defined in Fig. 1(b). The details of this experimental setup are described elsewhere.[25] The results are shown as solid lines in Fig. 2. For the primary polarization, there are two regular dips in reflection and two regular peaks in transmission at 570 nm and 526 nm. The first dip in reflection is caused by the magnetic resonance around 570 nm. The magnetic resonance arises from the coupling of current inside the upper and lower layers in the magnetic strips. The second dip at 526 nm comes from the electric resonance from the electric strips. Note that, since the electric resonance occurs at 526 nm, the electric strips act simply like a dilute metal

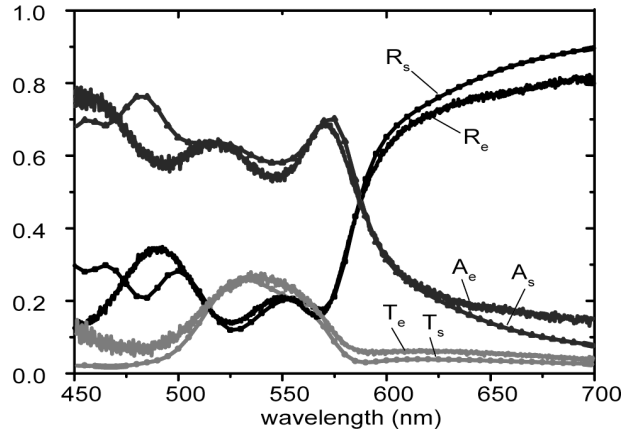


Fig. 2. A comparison of the experimental far-field transmission (Te), reflection (Re) and absorption (Ae) spectra of the fishnet metamaterial sample along with simulated results (Ts, Rs, and As) at the primary linear polarization as shown in Fig. 1(b). Reprinted with permission from Ref. (30). Copyright (2009) OSA.

and provide a broad negative permittivity background at the wavelength larger than 526 nm.

2.2. *Simulations*

Numerical simulations were performed for this sample in order to confirm our experimental results and retrieve the index of refraction. In the simulations the real parameters and dimensions of the fabricated sample are used. As it is discussed before, the electron damping will increase due to imperfections in the silver, including grain boundaries and size effects, and due to the geometrical effects of surface roughness on the magnetic strips.[15] So in the simulation, we used experimental data for the permittivity of silver and the electron collision rate was modeled to be three times larger than that of bulk silver at the plasmon resonances.[15,28] The simulation results are shown as dashed lines in Fig. 2. The simulation and experimental results match very well to each other, including a sharp magnetic resonance near 570 nm and a broad electrical resonance near 526 nm, except a deviation in the wavelength range shorter than 500 nm. The mismatch in this shorter-wavelength range is caused by differences between the actual losses in the silver at the shorter wavelengths in comparison to our model for the silver loss. Note that these differences exist only well outside of the range of negative-index behavior.

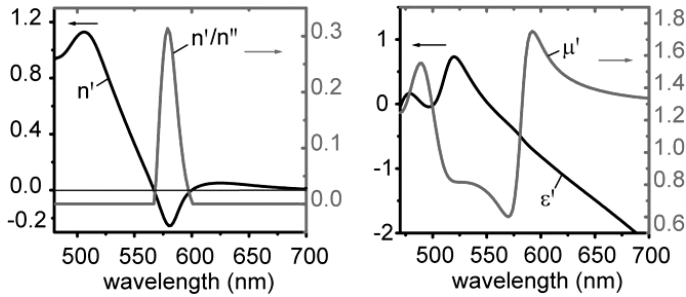


Fig. 3. (a) Real part of the effective refractive index and FOM; the FOM is set to zero if $n > 0$. (b) Real part of the effective permeability μ' and permittivity ε'. Reprinted with permission from Ref. (30). Copyright (2009) OSA.

Since the spectroscopic measurements and the numerical simulation over the wide wavelength range of interest in Fig. 2 match each other, our numerical model is valid and trustable. Therefore, the effective refractive index can be calculated using the numerical results and a standard retrieval procedure.[29] These results are shown in Fig. 3. In Fig. 3(a), a negative refractive index in the range between 567 nm and 602 nm is observed. The lowest n' occurs at 580 nm with the value of -0.25. Figure 3(b) shows the effective real parts of permeability and permittivity for the structure. The real part of permeability is positive throughout the entire wavelength range, while the permittivity decays towards the longer wavelengths from -0.23 to -0.81 in the SNNIM band from 567 nm to 602 nm. Even though the magnetic resonance is weak and the real part of the permeability is positive through the entire wavelength range, a clear magnetic resonance caused by the anti-symmetric mode is still present. But even in this optimized design the magnetic resonance is still not strong enough to obtain a negative permeability at 570 nm. Fortunately, the permittivity will be very negative at 570 nm and the loss is low enough (ε'' and μ'' are relatively small). As a result, the requirement of the single-negative NIM (SNNIM) $\varepsilon'\mu'' + \mu'\varepsilon'' \leq 0$ is fulfilled with $\varepsilon' < 0$ and $\mu' > 0$ near the magnetic resonance.[3] The minimum permeability value ($\mu' \approx 0.67$) is obtained at a wavelength of 570 nm along with $\varepsilon' \approx -0.39$. These results in SNNIM behavior with $n' \approx -0.18$. The figure of merit (FOM, shown in Fig. 3 and defined as n'/n'') is set to zero if $n' > 0$, and is $\sim$0.27 at 570 nm. The maximum FOM is achieved at 580 nm (FOM $\sim$ 0.3), with $\mu' \approx 0.98$.

As we mentioned before, the sample is specially designed with isotropic structure and is expected to show isotropic optical behavior. In order

to examine this, optical characterization and numerical simulations for the secondary polarization of incident light is also performed. In this polarization, the electric field is in plane and parallel to the magnetic strips (see Fig. 1). From the far-field spectra (not shown here), we found that the magnetic peak becomes much broader and relatively weaker than in the primary polarization case. These effects can be explained by a difference in the width of the upper silver layers for the electric and magnetic strips, as shown in the FESEM image of Fig. 1(b). From the FESEM image, we found that the widths of the electric strips vary more than the widths of the magnetic strips. In fact, the top width of the electric strips at some positions is much smaller than that of magnetic strips. Such anisotropy in the structure is likely the result of the anisotropic shape of the electron beam during the e-beam writing process. Therefore, the variation in width for the electric strips leads to a detuning of the resonant frequency position since the magnetic resonance is very sensitive to the geometric parameters and the shape of the strips. As a result, the magnetic resonance becomes extremely broad and relatively weak with an increase in the effective permeability, which is more positive than that for the primary polarization. Hence, the sufficient requirement for $\varepsilon'\mu'' + \mu'\varepsilon'' \leq 0$ cannot be met and the effective refractive index is positive in the whole wavelength range for the secondary polarization. Therefore, only a single negative index is obtained in the primary polarization, and the structure is highly anisotropic, which is different from the results expected from simulations. Still, this is the first experimental demonstration of a yellow-light negative refractive index.

3. Active Optical Negative Index Metamaterials

The success in creating optical NIMs in the visible optical range using nano-structured metal-dielectric composites has triggered intense basic and applied research in physics over past several years,[1–4,30] and now drives even more applications such as enhanced quantum interference[31] and "trapped rainbow" light storage.[32] However, the performance of all such applications is significantly limited by the inherent and substantial losses in NIMs, especially in the near-infrared and visible wavelength ranges. These losses originate from several sources, such as the intrinsic absorption of nanostructured metallic inclusions[14] and the surface roughness at metal-dielectric interfaces.[15] Generally, the losses are orders of magnitude too large for the envisioned applications. The overwhelming losses in

the nanoscale building blocks of the material plague the entire field of metamaterials and are one of the major restrictions preventing metamaterials from leaving the domain of academic research and entering industrial applications. In addition, due to limitations in material properties, the reduction of losses with optimized designs appears out of reach.[19] Therefore, efficient loss-compensation in optical metamaterials is the key issue for real applications.

Several methods have been proposed to compensate losses in metamaterials.[33,34] Compared with the complicated nonlinear response[33] and electromagnetically induced transparency,[34] incorporating an active material is a more effective and straightforward way to minimize or eliminate the absorption loss in NIMs.[18,19] Even though a negative refractive index n results in losses,[35] those losses do not have to be at the same wavelength where $n < 0$.[36] This means that a NIM with no loss is possible within a certain spectral interval. Many types of gain models and device configurations have been theoretically discussed in metamaterials.[37–41] However, compared with the hot theoretical discussions, an experimental demonstration of loss compensation in NIMs has until now been absent due to limitations in materials and fabrication capabilities. For example, the high levels of gain required for this method were previously considered impossible to obtain in experiments using either organic dye molecules or semiconductor quantum dots.[41–44] The obstacle of low available gain levels can be overcome by using a thick active host layer as the surrounding medium in a design, as was shown in the recent demonstration of the spaser.[45–47] Unfortunately, in a NIM-based device the active material cannot be too thick in order to preserve the negative refractive index. Therefore, the incorporation of gain in an optical NIM design is mired in difficulty, even though such an achievement would produce low- or no-loss optical NIMs for use in a large number of breakthrough applications.

Fortunately, it has been proven that the effective gain of active media in plasmonic devices is normally much higher than its bulk counterpart because of the local field enhancement inherent to the plasmonic response.[38–42] This in turn makes it realistically possible to compensate losses in NIM designs by the incorporation of active media in the areas of the high local fields that are inherent to plasmonic devices.

In fact, we have overcome the difficulties of incorporating an active medium into a NIM structure and have experimentally demonstrated a gain-assisted negative-index metamaterial with no loss. We have also confirmed our experimental results through numerical simulations.

We have shown that the large loss around the magnetic resonance of a given fishnet nanostructure can be compensated by the moderate gain offered by organic dyes, which is further enhanced by the localized electric field enhancement.[38,42] As a result, the active medium within the NIM gives rise to an effective gain much higher than its bulk counterpart. In our experiments, the transmission spectrum from the optical NIM sample is amplified by pumping the active medium incorporated within the sample. The structure is carefully designed to ensure that the active medium feels the highest local field while simultaneously preserving the negative-index property of the structure. Our experimental results, along with the numerical simulations, directly demonstrate the first lossless and active NIM.

3.1. *Sample fabrication*

The structure used in our experiments was the well-known fishnet structure, which already has been used to demonstrate a negative index in a wide wavelength range from the infrared to yellow light.[23] Rhodamine 800 (Rh800) dye doped in SU-8 epoxy is used as the gain medium. The Rh800 dye is selected here because of its high quantum efficiency and high solubility in the organic host. The measured spectral absorption and emission peaks of Rh800 in epoxy are 690 nm and 724 nm, respectively.[48] To utilize the enhancement of the plasmonic near fields, we replaced the typical alumina spacer of the initial fishnet with the selected Rh800/SU-8 material. Successful fabrication for this gain-assisted fishnet nanostructure is the main challenge since the nanostructure will be easily damaged in the replacement process. If the nanostrips collapse during this process, the magnetic response of the strips will disappear and no negative index will be obtained. Due to this issue, a special fabrication process was designed and is schematically shown in Fig. 4.

First, a $80 \times 80\,\mu m^2$ regular fishnet sample is fabricated using standard electron-beam lithography and lift-off processes on a glass substrate coated with a 15-nm-thick indium-tin-oxide (ITO) layer. The fabrication details of the initial fishnet are discussed in Section 2. The structure consists of a 50-nm alumina layer sandwiched by two 50-nm perforated silver layers, which are protected by 10-nm alumina layers at both the top and bottom surfaces.

The in-plane fishnet structure at different stages of fabrication is shown in Fig. 5. The initial fishnet structure is shown in Fig. 5(a), where

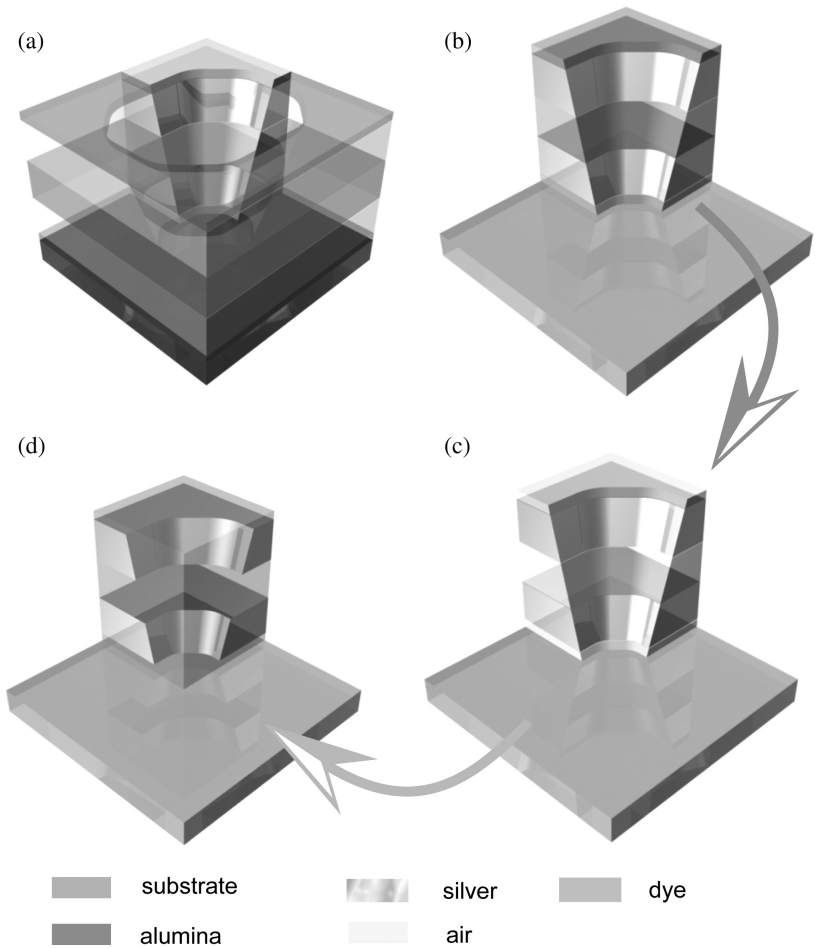


Fig. 4. Schematic of the fabrication process. (a) Unit cell of the fishnet structure with alumina as the spacer material between two silver layers; (b) A quarter of the fishnet structure with an alumina spacer; (c) After etching the alumina, the fishnet structure has air or solvent as the spacer with alumina pillars as support; (d) After coating with Rh800/epoxy, the fishnet structure has the dye/epoxy material in the spacer region and above the fishnet structure. Reprinted with permission from Ref. (49).

both lateral directions share the same periodicity of 280 nm. The bottom widths of the nanostrips are 170 nm in the x-direction and 209 nm in the y-direction. After initial fabrication, the next step is to etch the alumina in a tetramethylammonium hydroxide (TMAH) solution. The etching time is precisely controlled to create thin alumina pillars between the perforated

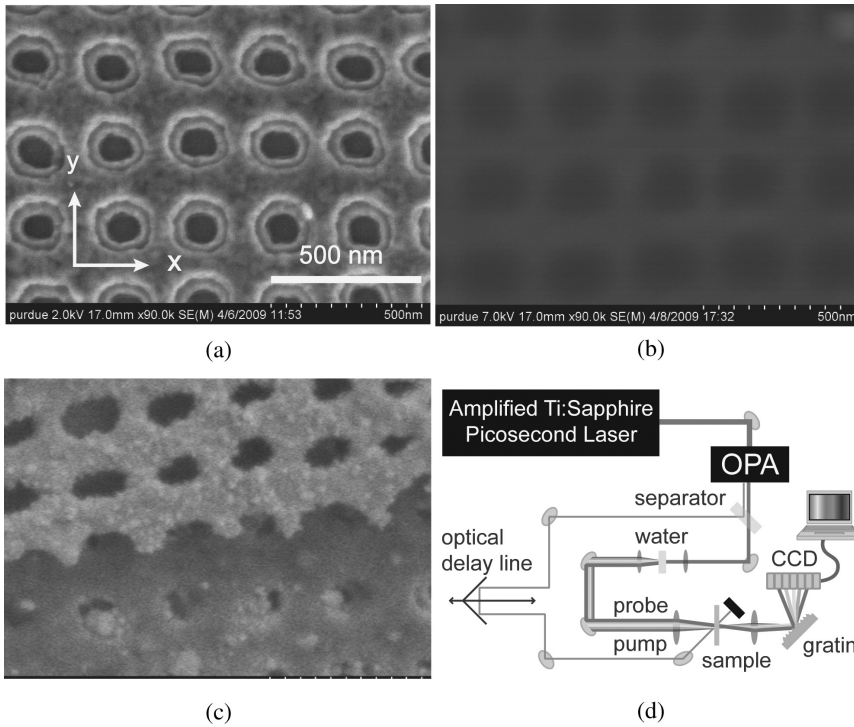


Fig. 5. FESEM images of the fishnet structure at different fabrication stages. All of the FESEM images share the same scale bar. (a) Fishnet structure with alumina spacer; (b) After etching the alumina and coating with Rh800/epoxy, the fishnet structure has the polymer/dye medium as the spacer and on the top; (c) Tilt-view SEM image of the structure after coating with Rh800/epoxy and after a part of the top layer of silver has been removed by focused ion beam milling; (d) The pump-probe experimental set-up.

metal layers in order to avoid any collapse of the top silver layer. Then an 800-nm-thick Rh800-doped epoxy film is spin-coated onto the sample. The dye concentration is 2×10^{-2} M, which corresponds to a concentration of Rh800 molecules of $1.2 \times 10^{19}\,\mathrm{cm}^{-3}$. During the spin-coating process, the Rh800/epoxy solution will penetrate into the fishnet holes and fills the etched spacer layer. Since the original thickness of the Rh800/epoxy is a rather large 800 nm, the final step in fabrication is to use reactive ion etching (RIE) to etch the thickness of Rh800/epoxy back down to 220 nm, leaving only about 60 nm on top of the fishnet structure. Figure 5(b) shows the FESEM image of the polymer-coated fishnet structure. Although the structure is fuzzy because the surface is covered by the Rh800/epoxy layer, the lateral profile of the fishnet structure is still observable. To confirm that

our fishnet structures are damage-free after the etching and spin-coating processes, focused ion beam (FIB) milling was used to remove the epoxy layer and part of the top layer of silver. Figure 5(c) shows the tilt-view FESEM image of the structure after FIB milling. The intact silver fishnet can be easily seen, showing no cracks or other defects in the structure. This confirms that the silver fishnet structure is indeed intact after etching the alumina spacer layer and spin-coating the Rh800/SU-8 layer.

3.2. *Sample characterization*

The fabricated sample was first optically characterized without pumping the gain material by measuring the far-field transmission and reflection spectra using normally incident light at the primary polarization with the electric vector of the incident light along the narrower strip (x-axis) in Fig. 5(a). The blue, red and green solid lines in Fig. 6(a) represent the transmission, reflection and absorption spectra of the sample, respectively, and a clear resonance around 725 nm can be observed. The resonance peak matches very well with the fluorescence peak of Rh800 dye, ensuring that the dye offers the largest gain. Because the impedance of the sample is nearly matched, the reflection is low around the resonance. As a result, the low transmission is solely due to high absorption. The experimental

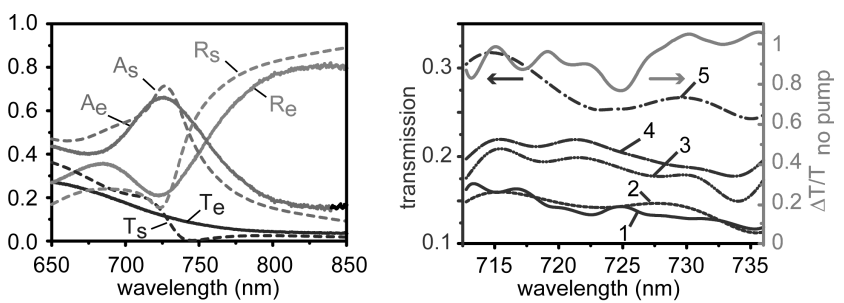


Fig. 6. Experimental results and simulation. (a) Experimental far-field transmission (T_e), reflection (R_e) and absorptance (A_e) spectra of the sample along with simulated results (T_s, R_s, and A_s) at the primary linear polarization as shown in Fig. 2(a); (b) The transmission spectra without pumping (line 1), with the optimized delay between pump and probe (probe pulse is 54 ps later than the pump) and 1 mW pumping power (line 5), with the optimized delay and 0.12 mW pumping power (line 3), with the optimized delay and 0.16 mW pumping power (line 4), and with 1 mW pump power preceding the probe by 6 ps (line 2). The wavelength-dependent relative transmission change from the pump-probe experiment is shown by the solid line at the top. Reprinted with permission from Ref. (49).

spectra match very well with numerical simulations [dotted lines in Fig. 6(a)]. Here, a Lorentz-oscillator absorption model for the material properties of the Rh800/epoxy combination is used in simulation. The model parameters were obtained from our experimental results. The overall, effective parameters of the sample were retrieved using our simulation data and the bi-anisotropic parameter retrieval method described in Ref. (50). The real part of the refractive index was negative from 720 nm to 760 nm, with the strongest value of -0.86 occurring at 740 nm, and the maximum FOM (FOM $= -n'/|n''|$, for $n' < 0$) was 1 at 737 nm.

We then characterized the gain-assisted response of the sample. The compensation of the structure's loss was investigated using a pump-probe experiment whose set-up is shown in Fig. 5(d). A 690-nm incident beam from an optical parametric amplifier (OPA) pumped by a 800-nm femtosecond Ti:sapphire laser was focused onto the sample and acted as the pump beam. A super-continuum white light source, generated by pumping water with a femtosecond pulse from the 800-nm laser, was normally focused to a 70-μm spot on the sample to act as the probe beam. The probe beam overlapped with the pump beam on the sample surface with a very small angle. The spot size for the pump light was 200 μm in order to make sure that the pumped area covers the whole probe beam on the sample surface. The pulse durations and repetitions of both beams were 2 ps and 1 kHz, respectively. The time delay between the two beams was adjustable via an optical delay line. The transmission spectrum of the probe light was collected from a 30- to 40-μm spot at the center of the pumped fishnet structure and analyzed with a spectrometer. The polarizations of both the pump and probe pulses were along the primary polarization.

The experimental results of the pump-probe measurements are shown in Fig. 6(b). Because of the limitations of the spectrometer grating, we focused our study on the peak wavelength region of the Rh800 dye. To exclude the possibility of damage-induced transparency and confirm our observations, we repeated the measurements with and without pumping 10 times with an acquisition time of 15 s. This results in a collection error of less than 10%. We first measured the transmission spectrum without pumping, which is shown in Fig. 6(b). The transmission value is almost the same as in Fig. 6(a) in the same wavelength region and confirms that the measurement technique is valid. We then measured the transmission of the sample with the pump laser turned on with 1 mW average power, which is below the threshold for damage to the sample but five times larger than the gain-saturation level for the Rh800/epoxy. The delay time

was optimized to make sure the dye offered the maximum gain. We observed that the transmission spectrum significantly increases when the gain medium is pumped, as shown in Fig. 6(b). The relative change between the transmission values with and without pumping [$\Delta T/T_{\text{nopump}} = (T_{\text{pump}} - T_{\text{nopump}})/T_{\text{nopump}}$] is plotted as the red solid line in Fig. 6(b). The transmission was 100% stronger in a wide wavelength range from 712 nm to 736 nm. These measurements are far outside the experimental error, which was less than 10%. The dashed line 2 in Fig. 6(b) shows the transmission spectrum that was measured by adjusting the probe beam pulse to arrive before the pump pulse. This spectrum matches very well to the spectrum without pumping, so the influence of the setup itself and local heating due to optical pumping can be excluded by this control experiment. The same measurements were performed on a control sample with pure epoxy as the spacer; this control sample showed no changes in its transmission spectrum in all of the measurements. Therefore, we can confirm that the changes in transmission spectrum of the gain-assisted with an optimized delay time are only due to the decrease of loss in the fishnet sample.

The experiment was also repeated for pump pulses at lower intensities. The results are shown in lines 3 and 4 in Fig. 6(b). Under the much lower pumping, intensities, the required population inversion in the dye molecules is not achieved and thus the dye does not offer the needed gain; instead, in this situation we see only a reduction the absorption of the dye so that the transmission change is much smaller than at the higher pump intensity.

The loss-compensation mechanism in the sample is relatively straightforward to understand. When Rh800 is excited by a pump pulse with sufficiently high power, a population inversion is formed in the energy levels of the dye molecules. This provides amplification for a properly delayed probe light whose wavelength is coincident with the stimulated emission wavelength of the dye molecules. As the loss is compensated with high gain, the transmission through the sample is greatly increased. Once the pump is turned off or the probe pulse arrives on the sample earlier than the pump pulse, the probe light no longer experiences any amplification during its propagation through the sample, and the transmission spectrum drops back down to its initial values.

3.3. *Simulations*

We also performed numerical simulations to understand the effects of gain on the corresponding transmission spectra and to retrieve the refractive

index and FOM of the sample excited by a pump pulse. In our study, the refractive index of the Rh800/epoxy gain medium was described as $n_a(\lambda) = 1.65 - i\frac{100}{4\pi}\lambda g_{\rm sim}(\lambda)$, where $g_{\rm sim}(\lambda)$ is the gain lineshape, which is wavelength-dependent and was obtained from the experiment on a Rh800/epoxy film with pumping power larger than the saturation power. The dependence of the relative transmission on the effective gain efficiency $g_{\rm sim}$ around the resonance was first calculated and is shown in Fig. 7(a). Similar to our experimental observations, with increasing $g_{\rm sim}$ we see an increase in the relative transmission through the sample. When $g_{\rm sim} = 2800\,{\rm cm}^{-1}$ around 725 nm (the center wavelength of the fluorescence of Rh800), the simulated transmission matches the experimental results well

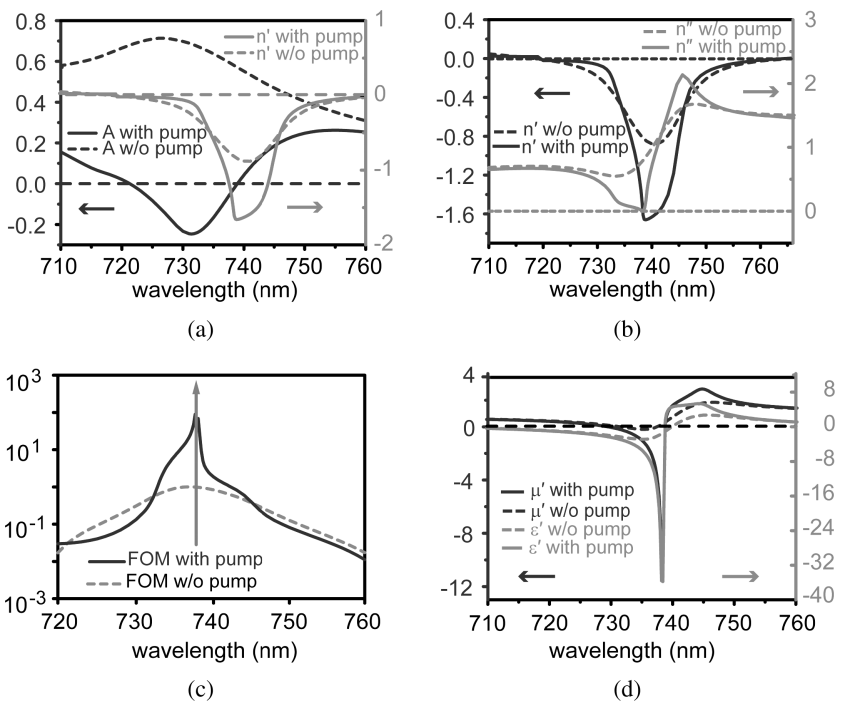


Fig. 7. Simulation and retrieved parameters. (a) The simulated refractive index n' and absorptance A (in the forward direction) as functions of wavelength with and without gain; (b) The retrieved effective refractive index $n = n' + i\,n''$ with and without gain; (c) The retrieved effective FOM with and without gain (the FOM is set to zero when the real part of the refractive index is positive); (c) The effective permittivity and permeability (real parts) with and without gain. Reprinted with permission from Ref. (49).

[Fig. 6(b)]. In the experiment described above, the concentration of Rh800 molecules in the epoxy film is $1.2 \times 10^{19}\,\mathrm{cm}^{-3}$. Therefore, using the known parameters for Rh800 dye, this concentration of excited molecules with an emission cross section of $3.4 \times 10^{-16}\,\mathrm{cm}^{-2}$ at 725 nm corresponds to a gain value of $2000\,\mathrm{cm}^{-1}$. The value of g_{sim} in our simulations is larger than this calculated value, which could be due to a chemical enhancement mechanism known to play an important role, for example, in surface-enhanced Raman scattering.[51] Another possible reason could be related to a feedback mechanism similar to that existing in a spaser.[46]

The spectra of the sample experience significant changes accompanying the loss compensation. First, the transmission is enhanced by 100%, as was mentioned before. As the same time, the impedance at the interface between the sample and the air becomes mismatched due to the change of the refractive index of the sample after the loss compensation. Thus the spectral reflection also increases, reaching 107.6% at 730 nm. The sum of the field intensities in transmission and reflection is nearly 1.23 times larger than that of the incident beam at this wavelength. This proves that the incident light is indeed amplified in the sample, resulting in a negative absorptance A(A = 1 – transmission – reflection), and the sample is therefore active. The absorbance of the sample with and without pumping is plotted in Fig. 7(a). It can be seen that the NIM remains active within a spectral range between 722 nm and 738 nm, while the refractive index is actually negative in a broader range between 720 nm and 760 nm.

We also retrieved the effective refractive index and FOM of the sample with pumping from the simulated spectra using the bi-anisotropic retrieval method mentioned above. The results are shown in Fig. 7. Both the real part n' and the imaginary part n'' of the refractive index drop significantly near the resonance. At 737 nm, n' changes with the addition of gain from -0.66 to -1.017, while n'' drops from 0.66 to 0.039 [see Fig. 7(b)]. This results in an increase in the FOM from 1 to 26, as is shown in Fig. 7(c). This value is so far the largest reported FOM obtained in any NIM in the optical region; it is even much larger than the values reported for two-dimensional metamaterial waveguides.[52] A much larger FOM is expected at 738 nm, with a value on the order of 10^6 ($n' = -1.26$, $n'' = 1 \times 10^{-6}$), and the structure is practically loss-free with $A \leq 0$ even at macroscopically large sizes. We need to note that it is hard to imperceptibly tune the gain and other parameters to the optimized performance point experimentally; but it is still easy to achieve FOM values that are up to several tens within about a 3-nm wavelength range near the resonance. We found that with further

increases in the gain, the retrieved n'' for our sample does not necessary go to the negative domain but instead starts to increase. This happens because the impedance mismatch grows at this level of gain, leading to increased front-side reflection and asymmetry in the front- and back-side reflections. But the NIM still remains active with $A < 0$ in a relatively broad range between 722 nm and 738 nm. With further improvements to the structure's design to decrease the impedance mismatch of the sample after loss compensation, a NIM with a macroscopically large FOM in a much broader spectral range can be achieved. However, we also need to realize that the use of the standard definition of the FOM $= |n'|/|n''|$ (for $n' < 0$) is normally reserved for non-bi-anisotropic metamaterials, since in bi-anisotropic materials the losses depend on the bi-anisotropy parameters. Hence the FOM ratio becomes somewhat less useful in quantifying the loss in bi-anisotropic NIMs. In that sense, the really important finding here is related to the presence of a negative index material with negative absorptance $A < 0$.

The effective dielectric permittivity and magnetic permeability of the sample can be dramatically improved by reducing the loss. Figure 7(d) shows the effective real parts of permittivity ε' and permeability μ' for the sample with and without gain. Both ε' and μ' become sharper and both the electric and magnetic resonances become stronger with the addition of gain. This is in good agreement with theoretical considerations.[38,40]

In retrieving the effective parameters of a sample, the trapezoidal cross-section and the substrate effect in fishnet-type NIMs can induce non-zero effective bi-anisotropy parameters,[53] which will further affect the retrieved characteristics of the sample with gain. One method for validating our retrieval procedure, which is built on the classical Kramers-Kronig (K-K) relation,[54,55] has been used for the available (truncated) spectral range. In the K-K relation, the frequency-dependent n' and n'' are expressed using the integral function $n'(\upsilon) = 1 - \frac{2}{\pi}\int_{\omega_1}^{\omega_2} \omega n''(\omega)(\omega^2 - \upsilon^2)^{-1} d\omega$. Here, ω_1 and ω_2 define the integral's low- and high-frequency edges. A numerical integration scheme using the subgrid mid-points $\upsilon_l = \frac{1}{2}(\omega_l + \omega_{l+1})$ obtained from the initially non-uniform spectral grid ω is used in calculating the integral. As a result, the values of $n'_l = n'(\upsilon_l)$ are calculated using $n'_l = 1 - \frac{1}{\pi}\sum_1^{p_{\max}-1} \Delta\omega_p$ $[\omega_{p+1} n''_{p+1}/(\omega^2_{p+1} - \upsilon_l^2) + \omega_p n''_p/(\omega_p^2 - \upsilon_l^2)]$, with $\Delta\omega_p = \omega_{p+1} - \omega_p$ and $n''_p = n''(\omega_p)$.

Figure 8 compares the values obtained with the truncated K-K numerical convolution (upper line with squares) and the n'' values retrieved with our general scheme based on Refs. (50) and (56) (black line). The figure

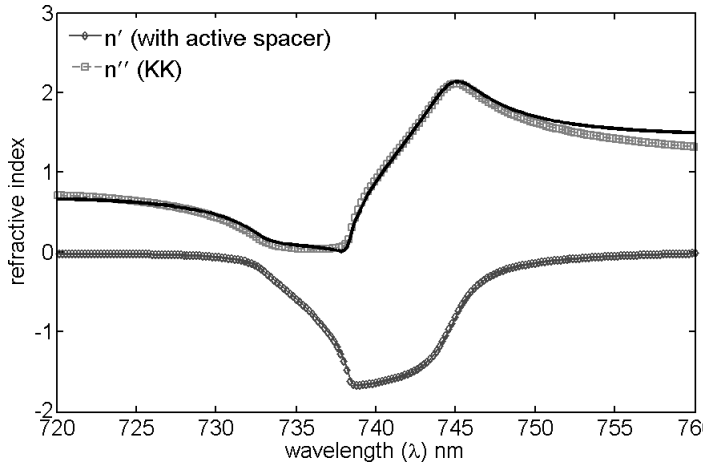


Fig. 8. Comparison of numerically retrieved results for the imaginary part of the effective refractive index (n''). The n'' values obtained with the truncated K-K numerical convolution (upper line with squares) and the n'' values retrieved with our general scheme (black line) show qualitatively and quantitatively similar results. Reprinted with permission from Ref. (49).

indicates sufficient qualitative and quantitative consistency in the positions, magnitudes, and widths of the major resonant features of n'', including the important spectral band of negative refraction. We note that there is some expected mismatch at the edges of the wavelength range due to the truncation of the spectral range. However, this does not affect the conclusion at the wavelength region around the resonance. Therefore, this result confirms the correct choice of the branch in our retrieval scheme.

The effectiveness of the high local fields in the gain-assisted fishnet structure results in much higher total energy produced by the gain media. Since the effective extinction coefficients at 737 nm are $\alpha \approx 6.75 \times 10^3\,\mathrm{cm}^{-1}$ and $\alpha = 1.13 \times 10^5\,\mathrm{cm}^{-1}$ for the sample with and without gain, respectively, the effective amplification is $\alpha = -1.07 \times 10^5\,\mathrm{cm}^{-1}$, which is 46 times larger than the "seed" value obtained without accounting for the increased local fields in the plasmonic structure. We used this "seed" value in our simulations at this wavelength. From the results of our numerical modeling, however, we see that the high local fields in the fishnet structure result in a total (spatially integrated) energy produced by the gain medium that is about 45 times larger than that produced by a homogeneous gain material slab of the same volume. This value is in good agreement with the factor of 46 found above and confirms that the loss compensation in our sample arises

from the local-field enhancement of the structure when a gain medium is used as the spacer layer.[38,40]

This study has for the first time demonstrated conclusively an optical metamaterial that is not limited by the inherent loss of its metal component. This breakthrough will significantly impact the fields of metamaterials and transformation optics, and it will likely drive the development of a family of new applications for optical metamaterials.

4. Tunable Optical Metamaterials

As mentioned in the previous sections, the working wavelength of NIMs has been pushed into the visible light range, and the inherent, overwhelming loss in passive metamaterials can be compensated. Thus ONIMs are positioned for full integration and real-world applications. One drawback still remains, however. The resonant nature of NIMs results in a very narrow operational bandwidth with a center frequency that is decided by the geometry and dimensions of elements comprising the NIM structure. Material is required to possess $\varepsilon' < 0$ and $\mu' < 0$ at the same frequency simultaneously. However, the resonance band for negative permeability is rather narrow, while the band for negative permittivity is rather broad. As the result, the key issue to implement negative index materials is to control the magnetic response. Because of the limitations of current nano-fabrication techniques, it is almost impossible to fabricate a nanostructure exactly the same as the designed plan. Thus, the resonant wavelengths obtained in optical metamaterials almost certainly will be different from the desired resonances. Therefore, externally controlled, tunable behavior is very essential for ONIM applications.[57,58]

Several methods have been proposed or demonstrated to control the negative permeability of a metamaterial. Changing the geometry and position of the building-block components[25,59] is the method that is the most popular for metamaterial designers and yet the most difficult in the optical range. The wavelength of the magnetic response can also be adjusted by changing the properties of the substrate.[60] However, changes in the substrate also imply a high possibility to incur changes in the properties of other, adjacent devices. Thus, this method requires very accurate design and implementation. Importantly, these methods are static and passive, and they perform worse than dynamic response modulation schemes.

One way to dynamically adjust the environment and resonant wavelength of a metamaterial is to use a liquid crystal material. Liquid crystals

(LCs) have large, broadband, optically anisotropic refractive indices that can be precisely controlled by temperature or external electromagnetic fields.[61] Therefore, LCs should be a good candidate for the dynamic control of permeability in metamaterials. Recently, LC-based, tunable metamaterials have been proposed by Khoo *et al.* and Werner *et al.* through randomly dispersed core-shell nanoparticles,[62] and incorporating a superstrate and a substrate,[63] respectively. Zhao *et al.* and Zhang *et al.* have also experimentally shown that a negative-permeability, uni-planar split-ring resonator (SRR) array infiltrated with LCs could be reversibly controlled by external electric and magnetic fields[64,65] in the centimeter (microwave) frequency range. Up to now, however, no experimental demonstrations have shown tunable behavior for the permeability of metamaterials in the optical wavelength range, and in particular within the visible range.

In this section, we discuss a tunable, negative-permeability metamaterial for the visible light range. By changing the ambient temperature, the wavelength of the magnetic response has been effectively changed from 650 nm to 632 nm. This is accomplished by causing nematic LCs incorporated into the material to undergo a phase transition from the ordered phase to the isotropic phase by adjusting the external temperature. When the temperature is raised above the phase transition temperature T_c (about 35°C for the liquid crystal 5CB), the LC material changes from an initially ordered, anisotropic phase to a disordered, isotropic phase. The resonant wavelength change is up to 3%. So far, this is the best result in wavelength tuning of metamaterials observed to date.

4.1. *Sample fabrication*

In these experiments, we used a 1D grating-type structure to obtain an initial sample with a significant magnetic response; we refer to this as a metamagnetism sample.[25] As in the previous sections, electron-beam lithography and lift-off techniques were used to fabricate the metamagnetism samples for our studies. Figure 9(a) shows the cross-sectional schematic of the metamagnetism samples. The figure shows a pair of thin, silver strips (thickness 35 nm) separated by an alumina spacer (thickness 60 nm). The whole sandwich stack is trapezoidally shaped due to fabrication limitations. The structure is periodic with a periodicity of 300 nm. Figures 9(b) and 9(c) show the FESEM and the atomic force microscope (AFM) images of the sample. Two 10-nm alumina layers were deposited on the top and bottom surfaces of the Ag-Al_2O_3-Ag sandwich stacks to increase the fabrication

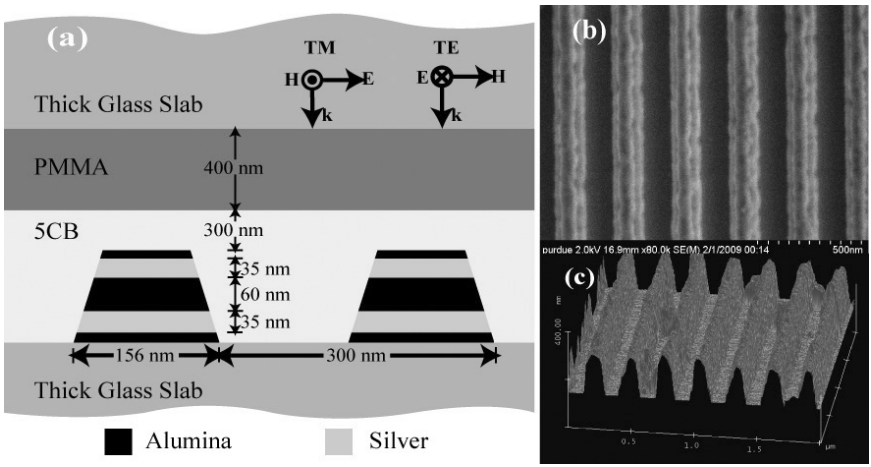
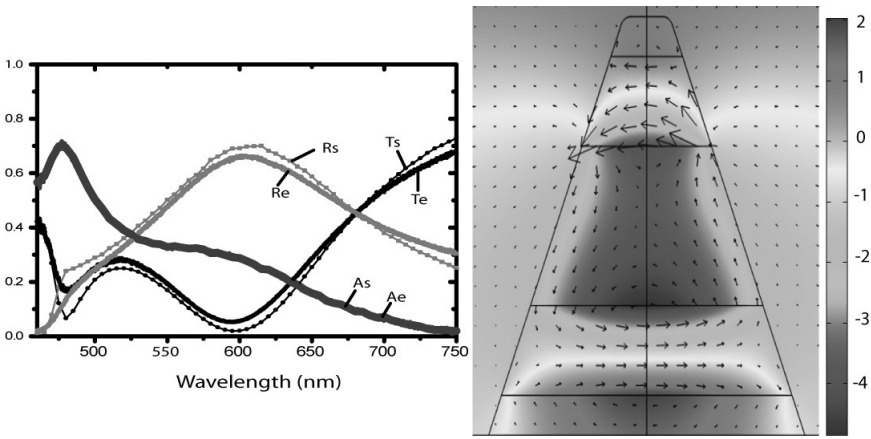


Fig. 9. Structure of the coupled nanostrip sample. (a) The cross-sectional schematic of arrays of coupled nanostrips; (b) The FESEM image of a typical sample; (c) The AFM image of the sample. Reprinted with permission from Ref. (67). Copyright (2009) American Institute of Physics.

stability. In order to investigate the tunability of the magnetic resonance, a layer of liquid crystal (5CB) was coated on top of the nanostrips at a temperature of 50°C. The thickness of the 5CB layer was about 400 nm. Then, a glass substrate, which was initially coated with a 300-nm layer of robust PMMA, was sandwiched on top of the sample. The PMMA layer was used to align the extraordinary axis of the liquid crystal molecules parallel to the surface of the substrate and perpendicular to the direction of magnetic strips.[66] After forming the sandwiched structure, the whole sample was cooled down to room temperature (about 20°C).

Before coating it with LCs, the sample was first optically characterized via far-field transmittance and reflectance spectral measurements using normally incident light at TM polarization (the primary polarization, i.e., the incident magnetic field is polarized along the set of strips, see Fig. 9. The resulting spectra are shown as solid lines in Fig. 10(a), which clearly shows a transmission dip and a reflection peak at about 595 nm. Here the dip and peak are related to the magnetic resonance resulting from an anti-symmetric current flow in the upper and lower strips. This resonance can induce a local minimum in the transmission spectrum and a local maximum in reflection. The corresponding absorption spectrum has also been measured and is shown in Fig. 10(a), where a clear absorption enhancement can be observed



(a) (b)

Fig. 10. (a) Transmission (T), reflection (R) and absorption (A), (including diffractive scattering) spectra under TM polarization for a typical coupled nanostrip sample with three characteristic wavelengths marked. Solid lines (Te, Re, and Ae) show the experimental data, and dashed lines (Ts, Rs, and As) represent simulated results. (b) Simulated electric displacement and magnetic field distributions in the cross-section of the nanostrips at the magnetic resonance wavelength. Reprinted with permission from Ref. (67). Copyright (2009) American Institute of Physics.

at about 595 nm. This enhanced absorption also supports the claim that this is a magnetic resonance.

In addition to experimental characterization, we also investigated the properties of the metamagnetism sample by numerical simulations, again using a commercial finite-element package (COMSOL Multiphysics). The material properties of silver were taken from our previous experimental data,[28] with the inclusion of an adjustable, wavelength-dependent loss factor.[68] The simulated results are shown as dashed lines in Fig. 10(a). We find that the theoretical and experimental results are consistent with each other. From our simulation results, we retrieved a negative effective permeability of about -1.5 at a wavelength of 595 nm. This negative permeability also indicates that indeed the initial sample has a significant magnetic resonance. To illustrate the nature of the magnetic and electric resonances, we also calculated the field distribution at the resonance wavelength in the cross section of a representative coupled nanostrip pair. The results are shown in Fig. 10(b), where the arrows represent electric displacement and the color map represents magnetic field. At the wavelength of magnetic resonance, we find that the electric displacement

forms a loop and produces an artificial magnetic moment, which results in negative permeability.[68]

4.2. *Tunable behavior*

The metamagnetism sample was then coated with LCs and optically characterized by obtaining its transmittance spectrum using TM light with normally incident light and the primary linear polarization. The results are shown in Fig. 11. When the sample was measured at a temperature of about 20°C, a transmission dip was observed at around 650 nm. Compared with the spectral measurements from the same sample before being coated with LCs [Fig. 10(a)], a distinct red-shift is observed due to the increased effective permittivity of the dielectric environment. When the temperature was increased to 50°C, the transmission dip shifted to 632 nm.

Meanwhile, we also fabricated another sample for a control experiment. This sample was fabricated under the same conditions and using the same geometry as the above one. Instead of sandwiching a layer of LC in the structure, however, it was coated by a 10-nm layer of Al_2O_3 only. The transmission spectrum has also been measured and magnetic resonance was observed. However, when the ambient temperature is increased from room temperature to 50°C, no detectable shift of the resonance was found.

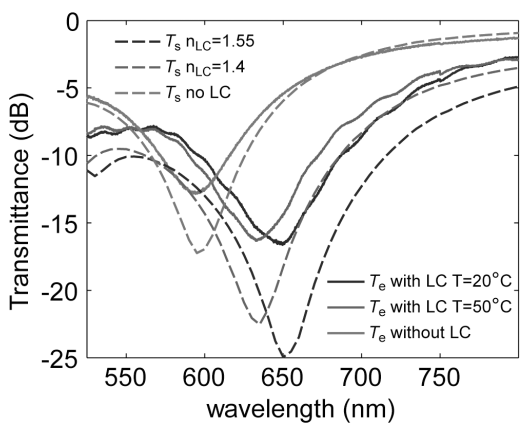


Fig. 11. Demonstration of a thermally tunable magnetic response in a metamaterial. Solid lines show the experimental data, and dashed lines represent simulated results without LCs, with LCs in the structure at 20°C and at 50°C. Reprinted with permission from Ref. (67). Copyright (2009) American Institute of Physics.

Thus, we conclude that the change in magnetic resonance wavelength is induced by change in the refractive index of LCs. The LC is in nematic phase at room temperature. In the case of planar alignment by PMMA, the direction of the 5CB molecules is parallel to the surface of the glass substrate and perpendicular to the direction of the magnetic strips. Thus the effective permittivity of 5CB for the primary polarization is close to $n_e = 1.73$, with some differences due to the thermal fluctuations. When the temperature is increased beyond 35°C, the 5CB liquid crystal molecules will transition from the nematic phase to the isotropic state. In the isotropic state, the LC molecules are randomly oriented and the material exhibits an isotropic refractive index that can be expressed as $n_{\text{iso}}^2 = \frac{2}{3}n_o^2 + \frac{1}{2}n_e^2 = 2.5$. The difference in refractive index between the two states is around $\Delta n = (1.73 - 1.58) = 0.15$. Normally incident light at the primary polarization experiences this change in refractive index during the phase transition. The decrease of the effective permittivity for the field at the primary polarization results in a blue shift of the magnetic resonance.[63] In the control sample, however, the change in the refractive index for Al_2O_3 is negligible for such small a temperature increase, and hence no spectral shift can be resolved.

4.3. *Simulations*

We also simulated the dependence of the magnetic resonance wavelength on the LC phase transition using the structural parameters of the original sample. The results of these simulations are shown as dashed lines in Fig. 11. By matching the simulation results to our experimental spectra, we found the effective refractive indices for LCs were only 1.55 and 1.4 for nematic and isotropic phase, respectively. Using these effective refractive indices, the resonance wavelength shift occurring as a result of the refractive index change $\Delta n = 0.15$ in the LC layer was also observed in simulation.

The simulation results match the experimental observations very well, although the obtained effective refractive indices of the 5CB in our simulation for both the nematic and isotropic phases are lower than those of a real LC medium. This is due to the small air gap between the nanostrip surface and the LC medium. Because the nanostrips have narrow and deep gaps between two adjacent strips, it is impossible to fully fill the spaces between adjacent strips. Therefore, some air will be left inside and reduce the effective refractive index of the dielectric environment.

It is worth noting here that there are differences in amplitude between the experimental results and the numerical simulations. The LC layer is assumed to be perfectly uniform in the numerical simulation. However, the LC molecules can form into different droplets[69,70] in real experiments, and some air bubbles can also be formed inside the LC material. The resultant scattering from the LC droplets and the air bubbles will therefore reduce the amplitude of resonant peaks and dips in the experimental results.

Overall, by using a metamagnetism sample coated with a layer of liquid crystal molecules, we have demonstrated a temperature-controlled, tunable negative permeability metamaterial working in the visible light regime. Our results show that the magnetic resonance can be externally controlled by adjusting the environmental temperature, which is used to change the refractive index of the LCs through a phase transition. For a further study, loss-free tunable metamaterials and electrical driven tunable metamaterials will be interesting topic that can of course enhance the performance of the real metamaterial devices.

5. Conclusions and Outlook

In the preceding sections, we have presented our recent progress in trying to solve some of the critical problems in the field of optical NIMs. We designed and demonstrated loss-free and tunable NIMs in the visible optical range. Although these research results have successfully solved some of the seemingly fatal disadvantages of optical NIMs and will certainly speed up the realization of a series of real applications, new challenges are already visible ahead of us. These challenges are related to the further commercialization of NIM-based devices. For example, commercial devices will demand a negative index at a range of incidence angles, and loss-free metamaterials must be operated with electrically driven control. These demanding challenges will open up a new route of optical NIM research both for deeper theoretical designs and more advanced experimental realizations.

Acknowledgments

This work was supported in part by ARO-MURI awards 50342-PH-MUR and W911NF-09-1-0539 and by NSF PREM grant # DMR 0611430. The authors are grateful to Dr. V. P. Drachev, Dr. A. V. Kildishev, Dr. U. K. Chettiar, and Dr. H.-K. Yuan for useful discussions related to this work.

References

1. J. B. Pendry, D. Schurig, and D. R. Smith, *Science* **312**, 1780 (2006).
2. S. Zhang, W. J. Fan, B. K. Minhas, A. Frauenglass, K. J. Malloy, and S. R. Brueck, *Phys. Rev. Lett.* **94**, 037402 (2005).
3. G. Dolling, M. Wegener, C. M. Soukoulis, and S. Linden, *Opt. Lett.* **32**, 53 (2007).
4. V. M. Shalaev, W. Cai, U. K. Chettiar, H. K. Yuan, A. K. Sarychev, V. P. Drachev, and A. V. Kildishev, *Opt. Lett.* **30**, 3356 (2005).
5. J. B. Pendry, *Phys. Rev. Lett.* **85**, 3966 (2000).
6. N. Fang, H. Lee, C. Sun, and X. Zhang, *Science* **308**, 534 (2005).
7. J. B. Pendry, D. Schurig, and D. R. Smith, *Science* **312**, 1780 (2006).
8. I. I. Smolyaninov, V. N. Smolyaninov, A. V. Kildishev, and V. M. Shalaev, *Phys. Rev. Lett.* **102**, 213901 (2009).
9. U. K. Chettiar, S. Xiao, A. V. Kildishev, W. Cai, H. K. Yuan, V. P. Drachev, and V. M. Shalaev, *MRS Bulletin* **33**, 921 (2008).
10. J. Valentine, S. Zhang, T. Zentgraf, E. U. Avila, D. A. Genov, G. Bartal, and X. Zhang, *Nature* **455**, 376 (2008).
11. S. P. Burgos, R. D. Waele, A. Polman, and H. A. Atwater, *Nat. Materials* **9**, 407 (2010).
12. X. Hu, Q. Zhang, and S. He, *Microwave and Optical Technology Letters* **51**, 2519 (2009).
13. I. Al-Naib, C. Jansen, and M. Koch, *J. Phys. D: Appl. Phys.* **41**, 205002 (2008).
14. A. Pinchuk, U. Kreibig, and A. Hilger, *Surf. Sci.* **557**, 269 (2004).
15. V. P. Drachev, U. K. Chettiar, A. V. Kildishev, H.-K. Yuan, W. Cai, and V. M. Shalaev, *Opt. Exp.* **16**, 1186 (2008).
16. V. A. Podolskiy and E. E. Narimanov, *Opt. Lett.* **30**, 75 (2005).
17. M. A. Noginov, G. Zhu, M. Bahoura, J. Adegoke, C. E. Small, B. A. Ritzo, V. P. Drachev, and V. M. Shalaev, *Opt. Lett.* **31**, 3022 (2006).
18. S. A. Ramakrishna and J. B. Pendry, *Phys. Rev. B* **67**, 201101 (2003).
19. T. A. Klar, A. V. Kildisher, V. P. Drachev, and V. M. Shalaev, *IEEE J. Sel. Top. Quantum Electrons* **12**, 1106 (2006).
20. A. K. Sarychev and G. Tartakovsky, *Phys. Rev. B* **75**, 085436 (2007).
21. M. I. Stockman, *Phys. Rev. Lett.* **98**, 177404 (2007).
22. P. Kinsler and M. W. McCall, *Phys. Rev. Lett.* **101**, 167401 (2008).
23. A. V. Kildishev, W. Cai, U. K. Chettiat, H.-K. Yuan, A. K. Sarychev, V. P. Drachev, and V. M. Shalaev, *J. Opt. Soc. Am. B* **23**, 423 (2006).
24. C. Enkrich, M. Wegener, S. Linden, S. Burger, L. Zschiedrich, F. Schmidt, J. F. Zhou, Th. Koschny, and C. M. Soukoulis, *Phy. Rev. Lett.* **95**, 203901 (2005).
25. W. Cai, U. K. Chettiar, H.-K. Yuan, V. C. Silva, A. V. Kildishev, V. P. Drachev, and V. M. Shalaev, *Opt. Express* **15**, 3333 (2007).
26. S. Zhang, W. J. Fan, N. C. Panoiu, K. J. Malloy, R. M. Osgood, and S. R. Brueck, *Phys. Rev. Lett.* **95**, 137404 (2005).
27. U. K. Chettiar, A. V. Kildishev, H.-K. Yuan, W. Cai, S. Xiao, V. P. Drachev, and V. M. Shalaev, *Opt. Lett.* **32**, 1671 (2007).

28. P. B. Johnson and R. W. Christy, *Phys. Rev. B* **6**, 4370 (1972).
29. D. R. Smith, S. Schultz, P. Markos, and C. M. Soukoulis, *Phys. Rev. B* **65**, 195104 (2002).
30. S. Xiao, U. C. Chettiar, A. V. Kildishev, V. P. Drachev, and V. M. Shalaev, *Opt. Lett.* **34**, 3478 (2009).
31. Y. Yang, J. Xu, H. Chen, and S. Zhu, *Phys. Rev. Lett.* **100**, 043601 (2008).
32. K. L. Tsakmarkidis, A. D. Boardman, and O. Hess, *Nature* **450**, 397 (2007).
33. N. M. Litchinitser and V. M. Shalaev, *Nature Photon.* **3**, 75 (2009).
34. S. Zhang, D. A. Genov, Y. Wang, M. Liu, and X. Zhang, *Phys. Rev. Lett.* **101**, 047401 (2008).
35. M. I. Stockman, *Phys. Rev. Lett.* **98**, 177404 (2007).
36. P. Kinsler and M. W. McCall, *Phys. Rev. Lett.* **101**, 167401 (2008).
37. M. Wegener, J. L. Garcia-Pomar, C. M. Souloulis, N. Meinzer, M. Ruther, and S. Linden, *Opt. Express* **16**, 19785 (2008).
38. A. Fang, Th. Koschny, M. Wegener, and C. M. Soukoulis, *Phys. Rev. B* **79**, 241104 (2009).
39. N. I. Zheludev, S. L. Prosvirnin, N. Papasimakis, and V. A. Fedotov, *Nat. Photon.* **2**, 351 (2008).
40. Y. Sivan, S. Xiao, U. K. Chettiar, A. V. Kildishev, and V. M. Shalaev, *Opt. Express* **26**, 24060 (2009).
41. M. A. Noginov, V. A. Podolskiy, G. Zhu, M. Mayy, M. Bahoura, J. A. Adegoke, B. A. Ritzo, and K. Reynolds, *Opt. Express* **16**, 1385 (2008).
42. V. I. Klimov, A. A. Mikhailovsky, S. Xu, A. Malko, J. A. Hollingsworth, C. A. Leatherdale, H.-J. Eisler, and M. G. Bawendi, *Science* **290**, 314 (2000).
43. F. Hide, B. J. Schwartz, M. A. Dias-Garcia, and A. J. Heeger, *Synth. Met.* **91**, 35 (1997).
44. E. Desurvire, (John Wiley and Sons, New York, 1994).
45. N. I. Zheludev, S. L. Prosvirnin, N. Papasimakis, and V. A. Fedotov, *Nat. Photo.* **2**, 351 (2008).
46. M. I. Stockman, *J. Opt.* **12**, 024004 (2010).
47. E. Plum, V. A. Fedotov, P. Kuo, D. P. Tsai, and N. I. Zheludev, *Opt. Express* **17**, 8548 (2009).
48. www.exciton.com/pdfs/ld800.pdf.
49. S. Xiao, V. P. Drachev, A. V. Kildishev, X. Ni, U. K. Chettiar, H.-K. Yuan, and V. M. Shalaev, *Nature* **466**, 735 (2010).
50. C. E. Kriegler, M. S. Rill, S. Linden, and M. Wegener, *IEEE J. Sel. Top. Quant. Electron.* **16**, 367 (2010).
51. D. P. Fromm, A. Sundaramurthy, A. Kinkhabwala, P. J. Schuck, G. S. Kino, and W. E. Moerner, *J. Chem. Phys.* **124**, 061101 (2006).
52. H. J. Lezec, J. A. Dionne, and H. A. Atwater, *Science* **316**, 430 (2007).
53. Z. Ku, K. M. Dani, P. C. Upadhya, and S. R. Brueck, *J. Opt. Soc. Am. B* **26**, B34 (2009).
54. J. D. Jackson, *Classical Electrodynamics*, 2nd edition, (Wiley, New York, 1975) Sec. 7.10.
55. J. J. H. Cook, K. L. Tsakmakidis, and O. Hess, *J. Opt. A: Pure Appl. Opt.* **11**, 114026 (2009).

56. A. V. Kildishev, W. Cai, U. K. Chettiar, H.-K. Yuan, A. K. Sarychev, V. P. Drachev, and V. M. Shalaev, *J. Opt. Soc. Am. B* **23**, 423 (2006).
57. V. M. Shalaev, *Nat. Photo.* **1**, 41 (2007).
58. R. A. Shelby, D. R. Smith, and S. Schultz, *Science* **292**, 77 (2001).
59. S. Linden, C. Enkrich, M. Wegener, J. Zhou, T. Koschny, and C. M. Soukoulis, *Science* **306**, 1351 (2004).
60. Z. Y. Sheng and V. V. Varadan, *J. App. Phys.* **101**, 014909 (2007).
61. I. C. Khoo, *Phys. Report* **471**, 221 (2009).
62. I. C. Khoo, D. H. Werner, X. Liang, A. Diaz, and B. Weiner, *Opt. Lett.* **31**, 2592 (2006).
63. D. H. Werner, D.-H. Kwon, I.-C. Khoo, A. K. Kildishev, and V. M. Shalaev, *Opt. Express* **15**, 3342 (2007).
64. Q. Zhao, L. Kang, B. Du, B. Li, J. Zhou, H. Tang, X. Liang, and B. Zhang, *App. Phys. Lett.* **90**, 011112 (2007).
65. F. Zhang, Q. Zhao, L. Kang, D. P. Gaillot, X. Zhao, and J. Zhou, *App. Phys. Lett.* **92**, 193104 (2008).
66. Y. Zhu, P. Zhu, Q. Wei, Z. Lu, Y. Wei, and Z. Luo, *Phys. Let. A* **176**, 141 (1993).
67. S. Xiao, U. K. Chettiar, A. V. Kildishev, V. M. Drachev, and V. M. Shalaev, *App. Phys. Lett.* **95**, 033114 (2009).
68. H.-K. Yuan, U. K. Chettiar, W. Cai, A. V. Kildishev, A. Boltasseva, V. P. Drachev, and V. M. Shalaev, *Opt. Express* 15, 1078 (2007).
69. O. Yaroshchuk, R. Kravchuk, A. Dobrovolskyy, L. Qiu, and O. Laventovich, *Liquid Crystals* **31**, 859 (2004).
70. B. D. Hamlington, B. Steinhaus, J. J. Feng, D. Link, M. J. Shelley, and A. Q. Shen, *Liquid Crystals* **34**, 861 (2007).

MANIPULATION OF PLASMONICS FROM NANO TO MICRO SCALE

Liang Feng* and Yeshaiahu Fainman†

Department of Electrical and Computer Engineering,
University of California, San Diego
9500 Gilman Drive, La Jolla, CA 92093-0407, USA
**lifeng@ucsd.edu*
†*fainman@ece.ucsd.edu*

Nanophotonics is finding myriad applications in information technology, health care, lighting and sensing. Plasmonics has been recently experiencing rapid development due to its myriad applications in nanophotonics. In this chapter, we explore the electrodynamics of plasmonic fields on different structured metallic chips and demonstrate how to manipulate light from nano to micro scale on the structured nanophotonic chips. We investigate on-chip plasmonic metamaterials with novel material responses and functionalities, develop the design methodology for plasmonic chips compatible with the conventional Fourier optical devices operating on diffraction limited plasmonic waves, as well as construct sophisticated chip-scale integration of optical elements with variable scales to achieve deeply subwavelength localization of optical fields.

1. Introduction

Plasmonics,[1] one of the most important fields in nanophotonics, has great potential to revolutionize many applications in nanophotonics, including bio-sensing,[2,3] imaging,[4–7] lighting,[8–11] photolithography[12–14] and magnetic recording.[15,16] Surface plasmon polaritons (SPPs) are strong interactions between electromagnetic fields and free electron oscillations

on the metal-dielectric interfaces. Different plasmonic fields can be strongly localized and confined on different structured interfaces. At these interfaces, there is an intrinsic SPP field localization that can facilitate nanoscale optical field manipulation, and there exists a potential for full optical confinement on the nanoscale in all three dimensions[17] beyond the diffraction limit of optical fields. Nanoscale optical field confinement enabled by plasmonic phenomena can reduce average power requirements for nonlinear optical activities and enable sub-diffraction limited imaging and waveguiding. These unique features of plasmonic fields can therefore result in chip-scale nanophotonic devices that are much more compact than those achievable with the current conventional bulky optical technologies. Plasmonics are also expected to bridge the gap between photonics and CMOS-based electronics technologies. However, there are still a number of fundamental issues being addressed in the attempt to make practical plasmonic optical elements for future technologies.

To accomplish the practical functionalities for plasmonic optical elements, it is highly desirable to excite and control propagation of SPP fields in a systematic fashion as it is possible with optical fields both in free space and dielectric waveguides. It requires the design methodology available from the conventional Fourier optical devices, on-chip plasmonic metamaterial with novel material responses and functionalities, as well as sophisticated chip-scale integration of different optical elements. This chapter explores the electrodynamics of SPP fields on different structured metallic chips and demonstrates how to manipulate SPPs from nano- to micro-scale on these plasmonic chips. The objective of this chapter is to engineer metallic structures and plasmonic optical elements on the scale from nanometers to tens of microns. The chapter will first focus on (i) the development of metallic metamaterials that exhibit novel plasmonic properties and optical responses which reduce to (ii) wavelength scale plasmonic photonic crystals. Further increasing the scale leads to (iii) the demonstration of Fourier Plasmonics that realizes conventional Fourier optics to miniaturize bulky optical devices on a plasmonic chip, and finally (iv) describe a nanophotonic device that implements an on-chip deeply sub-wavelength nanoscale optical field localization.

2. Form-Birefringent Metal and Its Plasmonic Anisotropy

Optical anisotropy and birefringence originate from atomic scale dipole moments that vary in space depending on the crystal lattice structure,

and have been used to create various devices which found applications in displays, communication, storage, and computing.[18,19] However, since optical properties of natural metals are always isotropic, in practice, it is difficult to realize optical plasmonic anisotropy. In this section, we propose and experimentally demonstrate a metal-dielectric composite metamaterial that is engineered on nanoscale for manipulation of inherent optical properties of metals, acting as an anisotropic uniaxial form-birefringent metal. This form-birefringent metal transfers its material birefringence to the SPP's.[20]

The form birefringent, schematically depicted in Fig. 1(a), is a two-dimensional (2D) array of air nanoholes in a thin metal film implementing a deeply subwavelength periodic structure with periods differing in x and y directions. These deep sub-wavelength features on nanoscale can not be resolved by light but in electrostatic field approximation act similarly to an anisotropic crystal, where the sub-wavelength scale geometry can have average polarizability varying in three dimensions of the medium.[21–25] The effective dielectric constant tensor of this metamaterial can be therefore described as:

$$\varepsilon_{\text{eff}} = \begin{pmatrix} \varepsilon_x & 0 & 0 \\ 0 & \varepsilon_y & 0 \\ 0 & 0 & \varepsilon_z \end{pmatrix}, \tag{1}$$

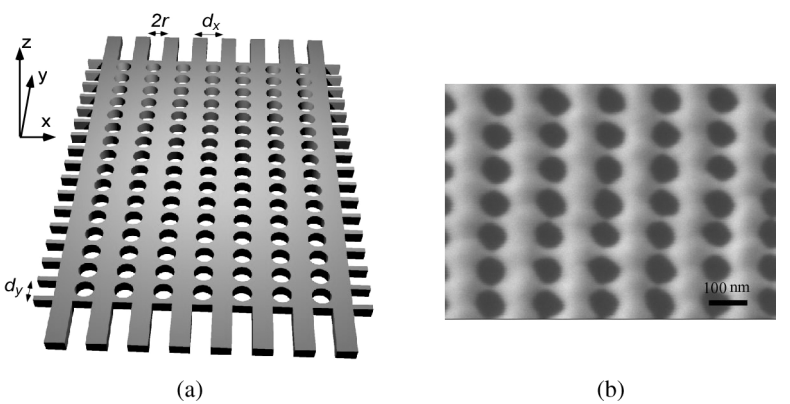


Fig. 1. "Form birefringence" metal acts as an anisotropic plasmonic metamaterial. (a) Schematic of a 55 nm-thick Au anisotropic plasmonic metamaterial with the design parameters: $d_x = 140\,\text{nm}$, $d_y = 93\,\text{nm}$, and $r = 35\,\text{nm}$. (b) SEM micrograph of the fabricated structure using FIB milling.

and using the effective medium theory the three diagonal elements are approximately expressed by[26]

$$\begin{aligned}
\varepsilon_x &= \varepsilon_m \left(1 - \frac{\pi r^2}{2d_y r}\right) + \varepsilon_0 \frac{\pi r^2}{2d_y r} \\
\varepsilon_y &= \varepsilon_m \left(1 - \frac{\pi r^2}{2d_x r}\right) + \varepsilon_0 \frac{\pi r^2}{2d_x r} \\
\varepsilon_z &= \varepsilon_m \left(1 - \frac{\pi r^2}{d_x d_y}\right) + \varepsilon_0 \frac{\pi r^2}{d_x d_y},
\end{aligned} \tag{2}$$

where ε_m and ε_0 are the dielectric constants of metal (e.g., Au) and dielectric medium in the nanoholes (e.g., air), respectively, r is the radius of nanoholes, and d_x and d_y are the periods of the deeply subwavelength structure in the x and y directions, respectively. It is worth noting that Eq. (2) is approximately valid once the designed structure is in the quasi-static regime with all the diffraction orders falling in the evanescent regime (i.e., the period is much shorter than the wavelength). Therefore, it is evident that the different effective dielectric constants will inevitably result in different SPP eigenmodes with their corresponding dispersion relations along the x and the y axes, thus creating optical anisotropy for SPP fields.

To validate our proposed anisotropic plasmonic metal-dielectric metamaterial, we use focused-ion-beam milling to fabricate fine deeply subwavelength features shown in Fig. 1(a) using an Au film on a glass substrate. The SEM micrograph of the fabricated structure is shown in Fig. 1(b). The excited SPP modes in the fabricated metamaterial are characterized in reciprocal space by analyzing reflection images [see Fig. 2(a)] obtained from the glass substrate side of the samples inserted in an oil immersion microscope objective (MO) with high numerical aperture (NA = 1.4).

Figure 2 shows the numerically simulated and experimentally measured index ellipsoids of SPPs on the form birefringent metal in reciprocal space at the wavelength around 540 nm. A converging unpolarized spherical wave is used as an illumination function to code a wide-band angular spectrum of the illumination function to a constant radial position within the illumination circle. The measurement setup directly captures the corresponding Fourier transform, which represents the SPP's index ellipsoid in the k space. Some of these spatial frequencies satisfy the phase matching condition and thus excite SPPs. The bright circle in Figs. 2(b) and 2(c) represents the boundary of the illumination at the angle corresponding to total internal reflection (TIR) when the transverse component of the

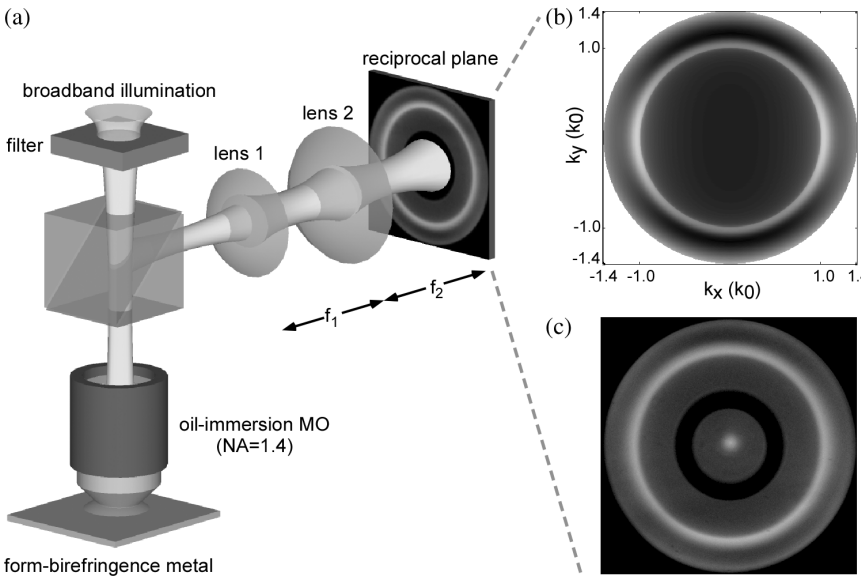


Fig. 2. (a) Measurement setup, composed of an oil immersion MO and two lenses, to directly image SPP's index ellipsoid in reciprocal space in the far field. (b) Simulated SPP's index ellipsoid in a representation of reflected power distribution in reciprocal space. (c) Experimental imaging of reflection from the metamaterial in reciprocal space. The dips (dark ellipses) in these index ellipsoids outside the bright circle in (b) and (c) represent total absorption due to the excitation of SPPs and can thus be extracted as SPPs' effective wavevectors.

illumination wavevector is k_0. Outside this TIR circle, an elliptical dark ring caused by excitation of SPPs manifests the optical plasmonic anisotropy of the excited SPP waves. From this index ellipsoid, we can also retrieve experimentally the effective wavevectors of SPP modes propagating along x and y optical axes as well as other azimuthal angles.

With considering the boundary continuity,[1] the eigenvectors of SPPs along x and y axes yield

$$k_{\mathrm{spp}}^{x,y} = k_0 \sqrt{\frac{\varepsilon_0 \varepsilon_z (\varepsilon_{x,y} - \varepsilon_0)}{\varepsilon_{x,y} \varepsilon_z - \varepsilon_0 \varepsilon_0}}, \tag{3}$$

where k_0 represents the wavevector of light in free space. The experimentally measured wavevectors at different wavelengths from 540 nm to 680 nm from Fig. 2 are integrated into Fig. 3 for comparison with the analytical calculations using Eq. (3) and numerical simulations. The plasmonic

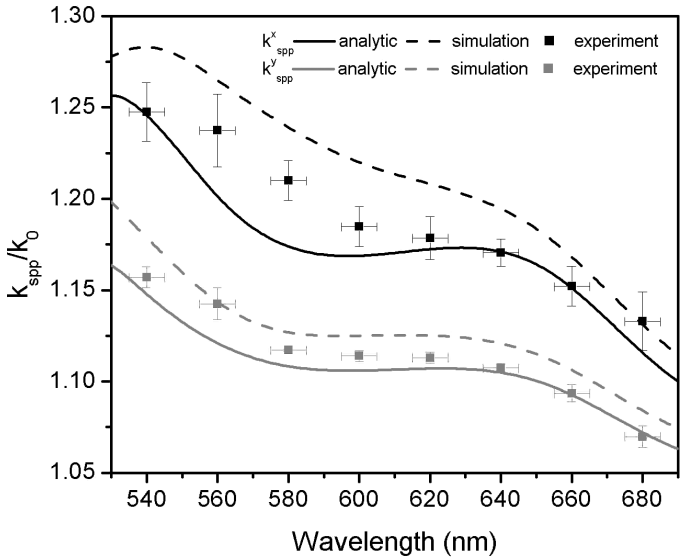


Fig. 3. SPP dispersion relations. Upper and lower sets of symbols represent SPP's wavevectors along the x and y directions, respectively. Solid curve, dashed curve and square dot correspond to the data retrieved from analytical calculations according to Eqs. (1)–(3), numerical simulations and experimental measurements, respectively.

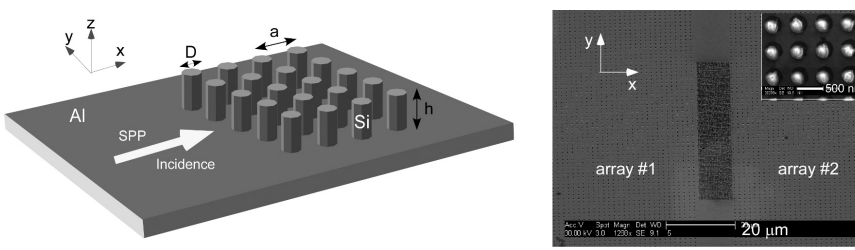
anisotropy in the range of interest attributes to the effective dipole moments engineered in an "artificial-atomic" scale with deep subwavelength features that vary in space depending on the crystal lattice of our form birefringent metal. The effective SPP indices in different directions vary since the plasmonic field interactions between unit cells strongly change with different duty cycles and dipole orientations, and thus manifest different averaged plasmonic E-fields in both individual unit cells and holes.

3. Plasmonic Photonic Crystal

When the feature size of the designed optical structure is on the length scale of the wavelength, the introduced periodicity in photonic crystals can result in a strong resonant manipulation of light, named Bragg resonance. Enhanced photon scattering can open a broad bandgap where the propagation of light is not allowed in photonic crystals.[27–29] Similarly, due to Bragg resonance at the boundaries of Brillouin zones, SPP fields in a specified frequency range are also forbidden to propagate through a periodic array of in-plane metallic bumps, which has been used to make

waveguides with sub-wavelength confinement,[30,31] high-efficiency Bragg reflectors[32] and resonators[33] for SPP fields. In this section, in contrast to these pure metallic plasmonic photonic crystals (PPCs), we introduce a metallo-dielectric PPC, where the photonic lattice is implemented in the dielectric material on the metal-dielectric interface. The band structure of the PPC has been theoretically calculated using the plane wave expansion (PWE) method[34] and validated in experiments, exhibiting a complete wide two-dimensional (2D) bandgap for SPP fields.[35]

Figure 4(a) shows the structural geometry of our PPC design consisting of a 2D array of high dielectric constant Si cylinders surrounded by low dielectric constant air on the surface of Al. The Si cylinders are 200 nm in diameter and 400 nm in height, and they are arranged in a square mesh with a lattice constant $a = 500$ nm. Since Si-Al and air-Al interfaces have different effective indices for in-plane SPP waves, the equivalent potential diagram in quantum mechanics would correspond to an electron propagating in a real crystal or light propagating in photonic crystals. This periodic potential contrast can cause strong scatterings near the boundaries of Brillouin zones, thus modifying SPP's dispersion and resulting in a complete bandgap caused by Bragg resonances. For experimental validation, we apply the overlay e-beam lithography and integrate the Si-based PPC with two 2D nanohole arrays on a single "chip" for excitation and detection as shown in Fig. 1(b). Nanohole array #1 will allow us to excite SPP fields that then will be incident onto the PPC and the transmitted SPP field will be scattered by nanohole array #2 into free space for detection and assessment of the PPC properties.[36,37]



(a) (b)

Fig. 4. (a) Schematic diagram of PPC's geometry with design parameters $a = 500$ nm, $D = 200$ nm, and $h = 400$ nm; (b) SEM micrographs of the integrated structure consisting of nanohole array #1, a PPC, and nanohole array #2 in an Al film.

The electric and magnetic field components of SPP fields are not independent as the eigenstate for one component is also the eigenstate for the other two components. Due to the boundary continuity, the E_z component has been chosen to form eigenfunctions and used to calculate the band structure of the PPC using the PWE method. By expanding $E_z(r)$ and $\varepsilon_{\text{spp}}(r)$ in reciprocal space using Bloch theory, the eigenfunction of our PPC can be derived as:[34]

$$\sum_{\vec{G}} \varepsilon^{r}_{\text{spp}}(|\vec{G} - \vec{G'}|)|\vec{k} + \vec{G}||\vec{k} + \vec{G'}|E_z(\vec{G}) = \frac{\omega^2}{c^2} E_z(\vec{G}). \tag{4}$$

The calculated band diagram [see Fig. 5(a)] shows a complete 0.14 $\omega a/2\pi c$-wide 2D bandgap from 0.28 $\omega a/2\pi c$ to 0.42 $\omega a/2\pi c$. The complete 2D bandgap results from high contrast between air and Si cylinders in $\varepsilon_{\text{spp}}(r)$.

This band structure is next validated by full wave 3-dimensional (3D) finite element simulations (with 400 nm-high Si cylinders on the Al surface) as shown in Figs. 5(b–d). An SPP wave, propagating from left to right, is launched at the left side of the simulation region. We observe [see Figs. 5(b, d)] that the SPP wave can propagate inside the PPC in the second and the first bands at the frequencies of 0.48 $\omega a/2\pi c$ and 0.22

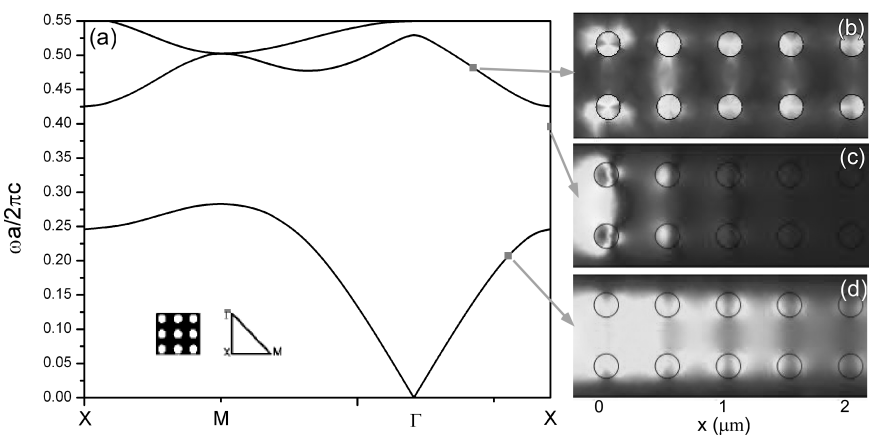


Fig. 5. (a) Calculated band structure of the proposed PPC using the PWE method with 289 plane waves. Full wave 3D finite element simulations show $|H_y|$ of SPP waves propagating inside the proposed PPC at different frequencies: (b) 0.48 $\omega a/2\pi c$ (corresponding to the 2nd band); (c) 0.39 $\omega a/2\pi c$ (corresponding to the bandgap); (d) 0.22 $\omega a/2\pi c$ (corresponding to the 1st band).

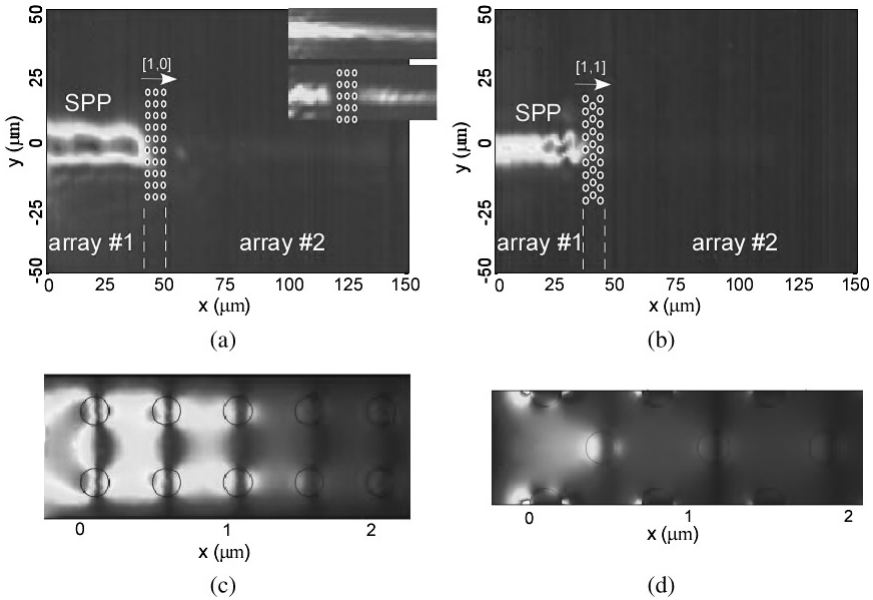


Fig. 6. (a), (b) Measured SPP time-averaged intensity maps at 1520 nm. (c), (d) Numerical mappings of $|H_y|$ of SPPs propagating inside the proposed PPC at 1520 nm.

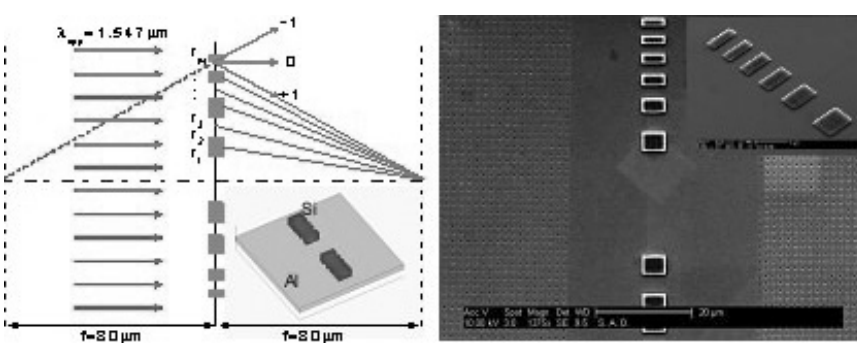
$\omega a/2\pi c$. But in the bandgap occurring at the frequency of 0.39 $\omega a/2\pi c$ [see Fig. 5(c)], strong attenuation can be observed and the SPP field propagation is forbidden.

The PPCs in Figs. 6(a) and 6(b) are fabricated for operation with normally incident SPP fields to propagate along $[1, 0]$ (Γ-X) and $[1, 1]$ (Γ-M) orientations of the PPC lattices, respectively. The measured transmittances are about 5.4% and 5.8%, respectively, about 6 times less than the transmittance observed with a 400 nm high Si slab on an Al surface. To have visual comparisons, the insets in Fig. 3(a) show the image of unperturbed SPP propagation on a continuous nanohole array (top) and SPP's transmission through the first passing band of another PPC with $a = 1000$ nm. For both PPC orientations shown in Figs. 3(a) and 3(b), the incident SPP waves can not propagate through the PPC, since the wavelength of 1520 nm (0.33 $\omega a/2\pi c$) lies within the complete band gap of the PPC as expected from Fig. 5(a). Numerical simulations in Figs. 3(c) and 3(d) also confirm our measured results, showing strong attenuation inside the PPC for both lattice orientations.

4. Fourier Plasmonics

Most of conventional optical systems that use components and devices in free space are designed with Fourier optics to manipulate propagation of light in free-space, interface between free-space modes with planar optical nanophotonics, as well as propagation on a chip. These approach is used when the systems have feature sizes on the order of tens of microns to millimeters, much larger than the wavelength of light.[38] It is therefore highly desirable to miniaturize the bulky optical devices on plasmonic chips to improve the devices' performance, e.g., resolution that is inherently associated with SPPs. However, it is still a challenge to excite and control propagation of SPP fields in a systematic fashion as it is possible with optical fields both in free space and dielectric waveguides. In this section, we report, as an example, a novel focusing approach using an in-plane SPP Fresnel zone plate (FZP).[39] We provide design rules to construct such an FZP and demonstrate experimentally the diffractive focusing of SPP fields in the plasmonic structures fabricated following the developed design rules.

A conventional optical binary amplitude FZP consists of a series of concentric rings, known as Fresnel zones that alternate in transmittance between transparent (i.e., 1) and opaque (i.e., 0). An adaptation of FZP for in-plane SPP focusing is illustrated schematically in Fig. 7(a). Consider the diffraction of an SPP wave impinging on this structure from the left. Constructive interference of SPP fields is obtained at a focal distance f



(a) (b)

Fig. 7. (a) Schematic diagram of SPP FZP geometry and design parameters; (b) SEM micrographs of SPP FZP fabricated by deposition of amorphous Si on Al/air interface.

from the FZP when Fresnel zone radii r_m satisfy

$$r_m = \sqrt{m\lambda_{\text{spp}} f + \frac{m^2\lambda_{\text{spp}}^2}{4}}, \tag{5}$$

where m is an integer, and λ_{spp} is the SPP wavelength. For a planar FZP, r_m are boundary positions where transmittance changes between transparent and opaque. In our experiments, a 12-zone SPP FZP with zone radii given by Eq. (5) was constructed at an Al/air interface for operation at a free space wavelength of 1.55 μm (corresponding to $\lambda_{\text{spp}} = 1.547\,\mu$m) and a focal length $f = 80\,\mu$m. As in previous work,[13,20] nanohole arrays were integrated on both sides of the device for launching SPP waves and visualizing their propagation as shown in Fig. 7(b).

The result of this average power measurement, shown in Fig. 8(a), represents the intensity map of SPP over the output nanohole array. The +1st-order FZP focal point is clearly seen at $x = 27\,\mu$m, 83 μm away from the FZP (located outside the image at $x = -110\,\mu$m, 20 μm to the left of the white dashed line indicating the edge of the nanohole array), in good agreement with the 80 μm expected from Eq. (5). Fringes due to the diverging-1st-order of the FZP are also visible. It is seen in Fig. 8(b) that, after a post-processing to compensate the radiative loss due to nanohole arrays, SPP intensity at the focal point may be expected to be about 3 times that of the input SPP wave.

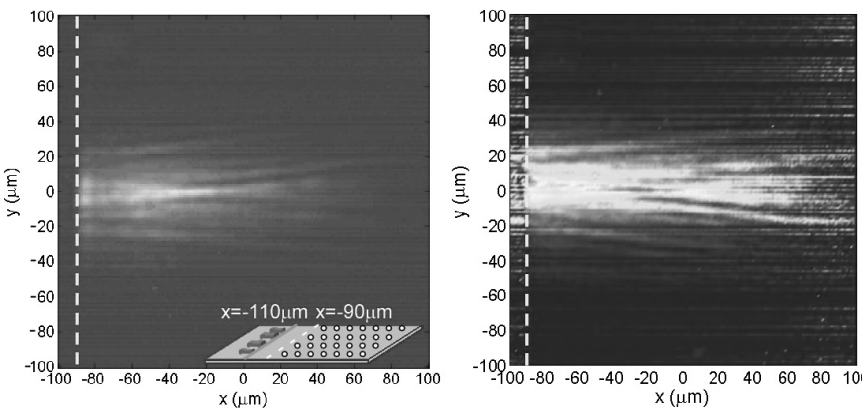


Fig. 8. (a) Measured SPP time-averaged intensity map over the nanohole array to the right of the FZP. (b) Post-processed image obtained from (a) by compensating for radiative loss.

The effectiveness of this in-plane SPP FZP is limited by the fact that even its opaque zones are partially transparent. Experimentally, about 30% transmission was observed through the 'opaque' zones, consistent with our finite-element numerical models. This effect can be taken into account by generalizing the standard expression for FZP diffracted field[38] to include contributions from both the transparent and the opaque zones,

$$u(x,y) = \frac{\exp(ik_{\mathrm{spp}}x)}{i\lambda_{\mathrm{spp}}x}\iint_T u_0(y_0)\exp\left(\frac{ik_{\mathrm{spp}}}{2x}(y-y_0)^2\right)dy_0 + \frac{\exp(ik_{\mathrm{spp}}x)}{i\lambda_{\mathrm{spp}}x}\iint_O \sqrt{0.3}u_0(y_0)\exp\left(\frac{ik_{\mathrm{spp}}}{2x}(y-y_0)^2\right)dy_0. \quad (6)$$

SPP intensity $|u(x,y)|^2$ at the focal plane ($f = 80\,\mu$m) calculated from Eq. (6) is plotted in Fig. 9(a), and agrees well with the measured data in Fig. 9(b). The slight deviation may come from the slight convergence of the incident SPP fields. Methods of conventional Fourier optics can thus be successfully applied to describe in-plane diffraction of SPP waves, leading to a new direction of research on Fourier Plasmonics to miniaturize conventional optical components on small plasmonic chips.

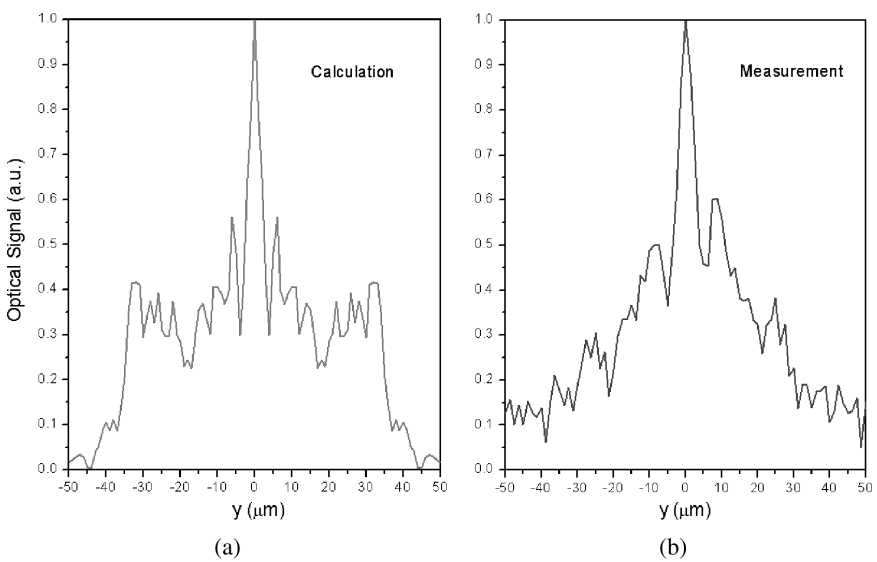


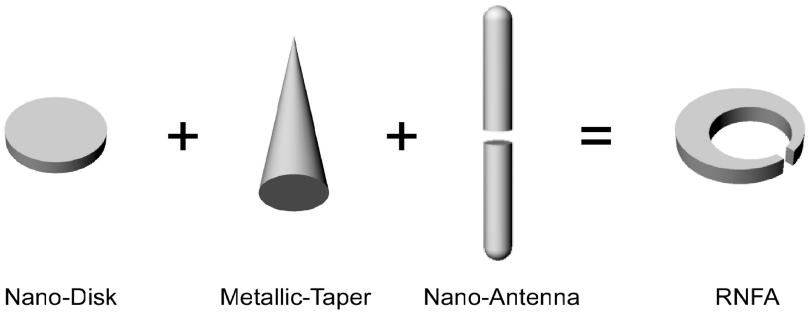
Fig. 9. SPP intensity profile in the focal plane computed by the Fresnel diffraction method (a) and found experimentally (b).

5. Nanoscale Optical Field Localization

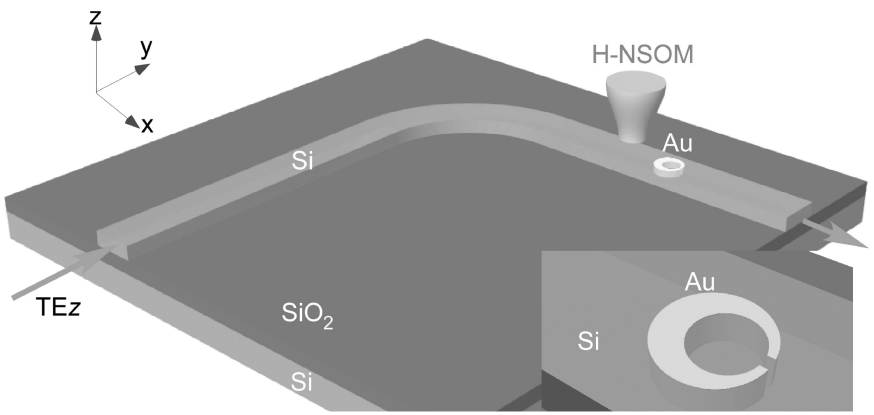
Nanoscale field confinement enabled by plasmonic phenomena has great potential to revolutionize many applications in nanophotonics that rely on deeply subwavelength field localization including bio-sensing, nanoscale imaging, and heat assisted magnetic recording. Various schemes using SPPs have been suggested.[40–46] In this section, we introduce a nanophotonic chip that integrates Si-nanophotonics with a novel plasmonic resonant nano-focusing-antenna (RNFA) geometry[47] for efficient field nano-focusing and localization. The RNFA that simultaneously uses three physical mechanisms: localized surface plasmons (LSPs),[48] thin metallic wedge localization of SPP with high effective index,[49–53] and TEM field localization[54–56] [see Fig. 10(a)]. This novel RNFA geometry manipulates light with feature sizes on both nano and micro scales: the nano focusing defined by the nano structure engineering in a quasi-static regime with its resonance dynamically determined with the size of the whole plasmonic component.

While LSPs on the nanodisk structure already lead to strong field confinement, a significantly smaller field localization spot size can be achieved by introducing the field edge localization by sharpening the nanodisk. The effective index near the sharp edge can be very high (corresponding to a very low potential) thereby resulting in strong plasmonic field localization. The plasmonic fields on two sharp thin metallic wedges of the RNFA in Fig. 10(a) are strongly coupled and they support nearly uniformly distributed TEM type fields regardless of the gap size. These TEM fields are similar to the strong fields obtained at the feed points or small air gaps in dipole, bow-tie, and other small antennas. The gap operates as a capacitor that is known to confine strong fields even in the static regime. Such a nano-capacitor strongly confines the plasmonic fields, already enhanced due to the LSP and the sharp thin metallic wedge localizations. Our experimental apparatus consists of a lossless Si waveguide integrated with the gold RNFA nanostructure to deliver TE_z-like electromagnetic wave to the RNFA and excite the LSPs [see Fig. 10(b)]. The experiments are designed for operation in the near infrared telecommunication optical spectrum range ($\sim$1550 nm). Top faces of the RNFA and waveguide are leveled to efficiently excite resonant LSP around the RNFA as well as easily detect the generated LSP outside the Si waveguide using heterodyne near-field scanning optical microscope (H-NSOM)[56] [see insets of Fig. 10(b)].

Finite element simulations summarized in Fig. 11 show a comparison between the calculated resonant responses of the densities of the electric field intensity achieved inside the RNFA gap and the LSP localization



(a)



(b)

Fig. 10. (a) The proposed RNFA nanostructure. (b) The experimental setup consisting of monolithically integrated single gold RNFA, placed at the center of a 1 μm-wide Γ-shaped Si waveguide fabricated using SOI technology.

spots in a simple gold nanodisk, clearly demonstrating about six times resonant enhancement [see Fig. 11(a)]. At the resonant frequency, the LSP field is strongly confined at the edge of the nano-disk and decays exponentially in Fig. 11(b). While in Fig. 11(c), the field is strongly and uniformly localized inside the nano-capacitor formed by the gap with the spot size of about 25 nm along the x direction. Note that the RNFA and nanodisk structures share similar eigenfrequencies and have resemblance in the spatial modal structure. These similarities between the eigenstates have the same nature as the current distribution in a microwave dipole wire antenna, which is nearly identical to the case of free standing wire

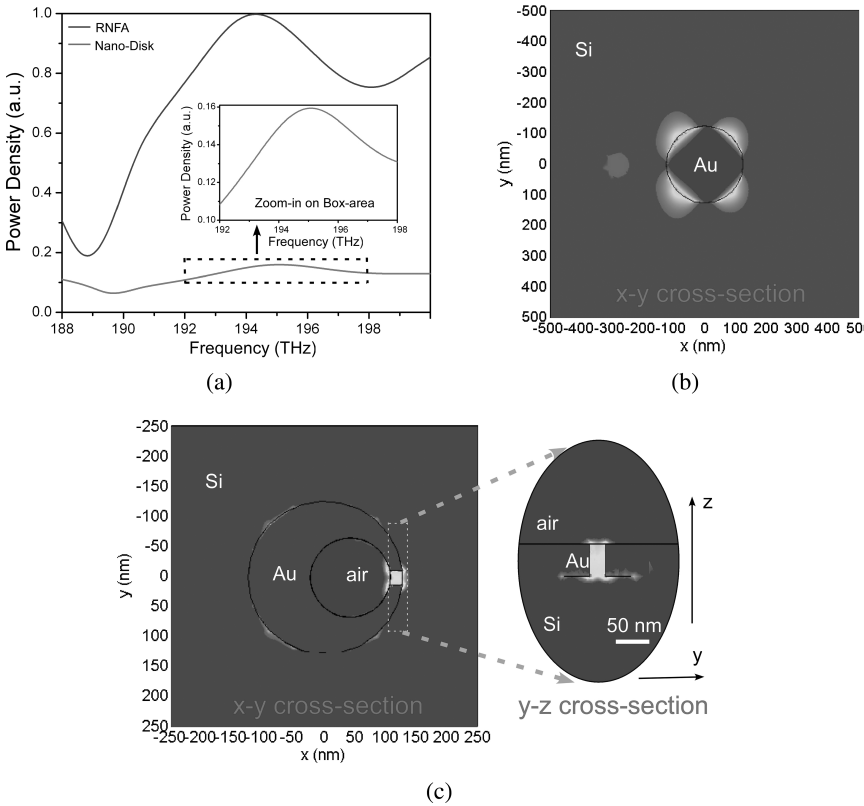
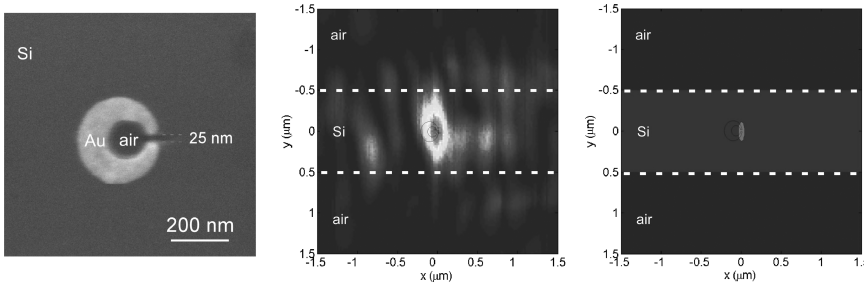


Fig. 11. (a) The upper and lower curves show the spectra of the power densities inside the gap of the RNFA nanostructure and the field localization regions of the nanodisk, respectively. (b) Map of the LSP electric field intensity distribution on the top surface of the waveguide and nanodisk interface at the resonant frequency of 195 THz (wavelength of 1538.5 nm). (c) Map of the electric field intensity distribution around the RNFA-Si interface (left) and the transverse plane cross-section at the center of the gap of the RNFA structure (right) at the resonant frequency of 194.5 THz (wavelength of 1542 nm).

and an antenna of the same size fed through a small gap. The small gap in RNFA does not change significantly the source-excited or source-free field/current distributions as well as eigenfrequencies of such structures. Moreover, for the complicated RNFA structure, its eigenfrequencies can be solely characterized by developing closed form approximate expressions for a simple nanodisk structure. While the LSP field impedance and resonant frequencies do not change, the presence of the sharp thin metallic wedges and the small gap results in more than 6-fold enhancement of the field



(a) (b) (c)

Fig. 12. (a) SEM micrograph of the RNFA geometry. (b) Raw H-NSOM image of high resolution intensity mapping measured at the resonant frequency of 194.2 THz. (c) Processed image of the field intensity in the same area after deconvolution signal processing to extract the effect of the NSOM probe on the measured data.

density [see Fig. 11(a)] localized in all 3 dimensions inside the nano-capacitor with a 25 nm gap.

The near field H-NSOM measurements are performed on the fabricated RNFA structure [see Fig. 12(a)], first with a tip scanning step size of 100 nm to locate the resonance by simultaneously scanning the optical frequencies of the input field in the range of 191 THz to 196 THz. The data show a strong LSP localization at the frequency of 194.2 THz ($\lambda_0 = 1544.9$ nm). Next we use the input optical field at the resonant frequency to detect a high resolution near field intensity map with H-NSOM scanning step size of 10 nm [see Fig. 12(b)]. The strong field localization obtained experimentally corresponds to the resonance of LSPs on the RNFA. The measured spot size along the x direction taking into account the finite probe aperture size was about 220 nm. Since the expected nanofocusing spot size is only tens of nanometers (corresponding to the size of the nano-capacitor), the measurement of the actual size even with current H-NSOM techniques[56] is challenging owing to the size and the electromagnetic interaction between the probe and the measured structure. We consider these near field interaction effects in the RNFA characterization experiments by including a deconvolution procedure into our H-NSOM process. In the restored near field intensity image shown in Fig. 12(c), the estimated effect of the probe on the convolution with the measured localized optical field is an effective aperture of about 155 ± 5 nm, consistent with the estimates achieved for other plasmonic nanostructures we tested in H-NSOM measurement (e.g., 150 nm for nanodisks). An elliptical focusing spot was obtained, which is associated with a slightly weaker localization in the y direction. The field

localization along the x direction estimated from the restored measured data in Fig. 12(c) is about 75 ± 5 nm.

6. Conclusions and Outlook

The complexity of future computing systems will increasingly rely on optical interconnections that may reach out scales of intra-chip communications, where the efficient on-chip manipulation of light on different scales is necessary. This chapter presents the investigation of chip-scale resonant plasmonic nanostructures, in order to manipulate light on the plasmonic chips from nano to micro scales.

The main points include our ability to use standard micro and nano fabrication tools to create complex geometries of composite metal-dielectric plasmonic structures and realize the integration of plasmonics with the CMOS-compatible SOI platform. The feature sizes of these composite materials can be engineered on nanoscale to be deeply subwavelength, which enables creation of materials with unique functionalities that is not possible to achieve with existing technologies. These structures can also be created on a microscale equal to or much larger than the wavelength, where the possibility of miniaturization of bulky optical devices on plasmonic chips is demonstrated. We also introduced a plasmonic geometry utilizing both localized quasi-static effects on nanoscale and dynamical resonance control on microscale. The realization implemented using Si-photonics SOI platform significantly improves the coupling efficiency to the integrated plasmonic element.

Acknowledgments

This work was supported by NSF, DARPA and NSF CIAN ERC.

References

1. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer-Verlag, 1988).
2. K. A. Willets and R. P. Van Duyne, *Annu. Rev. Phys. Chem.* **58**, 267 (2007).
3. L. Pang, G. Hwang, B. Slutsky, and Y. Fainman, *Appl. Phys. Lett.* **91**, 123112 (2007).
4. J. R. Pendry, *Phys. Rev. Lett.* **85**, 3966 (2000).
5. N. Fang, H. Lee, C. Sun, and X. Zhang, *Science* **308**, 534 (2005).
6. Z. Liu, H. Lee, Y. Xiong, C. Sun, and X. Zhang, *Science* **315**, 1686 (2007).
7. S. Kawata, A. Non, and P. Verma, *Nature Photon.* **2**, 438 (2008).

8. D. J. Bergman and M. I. Stockman, *Phys. Rev. Lett.* **90**, 027402 (2003).
9. M. A. Noginov, G. Zhu, A. M. Belgrave, R. Bakker, V. M. Shalaev, E. E. Narimanov, S. Stout, E. Herz, T. Suteewong, and U. Wiesner, *Nature* **460**, 1110 (2009).
10. R. F. Oulton, V. J. Sorger, T. Zentgraf, R. M. Ma, C. Gladden, L. Dai, G. Bartal, and X. Zhang, *Nature* **461**, 629 (2009).
11. L. Cao, J. S. White, J. S. Park, J. A. Schuller, B. M. Clemens, and M. L. Brongersma, *Nature Mater.* **8**, 643 (2009).
12. W. Srituravanich, L. Pan, Y. Wang, C. Sun, D. B. Bogy, and X. Zhang, *Nature Nanotech.* **3**, 733 (2008).
13. P. Zijlstra, J. W. M. Chon, and M. Gu, *Nature* **459**, 410 (2009).
14. W. Srituravanich, N. Fang, C. Sun, Q. Luo, and X. Zhang, *Nano Lett.* **4**, 1085 (2004).
15. W. A. Challener, T. W. McDaniel, C. D. Mihalcea, K. R. Mountfield, K. Pelhos, and I. K. Sendur, *Jpn. J. Appl. Phys.* **42**, 981 (2003).
16. W. A. Challener, C. Peng, A. V. Itagi, D. Karns, W. Peng, Y. Peng, X. Yang, X. Zhu, N. J. Gokemeijer, Y. T. Hsia, G. Ju, R. E. Rottmayer, M. A. Seigler, and E. C. Gage, *Nature Photon.* **3**, 220 (2009).
17. W. L. Barnes, A. Dereux, and T. W. Ebbesen, *Nature* **424**, 824 (2003).
18. A. Yariv and P. Yeh, *Optical Waves in Crystals: Propagation and Control of Laser Radiation* (Wiley and Sons, New York, 2003).
19. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation Interference and Diffraction of Light* (Cambridge, University Press, 1999).
20. L. Feng, Z. Liu, V. Lomakin, and Y. Fainman, *Appl. Phys. Lett.* **96**, 041112 (2010).
21. I. Richer, P. C. Sun, F. Xu, and Y. Fainman, *Appl. Opt.* **34**, 2421 (1995).
22. R. C. Tyan, A. A. Salvekar, H. P. Chou, C. C. Cheng, A. Scherer, F. Xu, P. C. Sun, and Y. Fainman, *J. Opt. Soc. Am. A* **14**, 1627 (1997).
23. L. Pang, M. Nezhad, U. Levy, C. H. Tsai, and Y. Fainman, *Appl. Opt.* **44**, 2377 (2005).
24. F. Xu, R. C. Tyan, P. C. Sun, Y. Fainman, C. C. Cheng, and A. Scherer, *Opt. Lett.* **21**, 1513 (1996).
25. U. Levy, M. Abashin, K. Ikeda, A. Krishnamoorthy, J. Cunningham, and Y. Fainman, *Phys. Rev. Lett.* **98**, 243901 (2007).
26. S. M. Rytov, *Sov. Phys. JETP* **2**, 466 (1956).
27. E. Yablonovitch, *Phys. Rev. Lett.* **58**, 2059 (1987).
28. S. John, *Phys. Rev. Lett.* **58**, 2486 (1987)
29. J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light* (Princeton NJ, Princeton University Press, 2008).
30. S. I. Bozhevolnyi, J. E. Erland, K. Leosson, P. M. W. Skovgaard, and J. M. Hvam, *Phys. Rev. Lett.* **86**, 3008 (2001).
31. C. Marquart, S. I. Bozhevolnyi, and K. Leosson, *Opt. Express* **13**, 3303 (2005).
32. M. U. Gonzalez, J. C. Weeber, A. L. Baudrion, A. Dereux, A. L. Stepanov, J. R. Krenn, E. Devaux, and T. W. Ebbesen, *Phys. Rev. B* **73**, 155416 (2006).

33. J. C. Weeber, A. Bouhelier, G. Colas des Francs, L. Markey, and A. Dereux, *Nano Lett.* **7**, 1352 (2007).
34. K. M. Ho, C. T. Chan, and C. M. Soukoulis, *Phys. Rev. Lett.* **65**, 3152 (1990).
35. L. Feng, M. H. Lu, V. Lomakin, and Y. Fainman, *Appl. Phys. Lett.* **93**, 231105 (2008).
36. K. A. Tetz, R. Rokitski, M. Nezhad, and Y. Fainman, *Appl. Phys.Lett.* **86**, 111110 (2005).
37. R. Rokitski, K. A. Tetz, and Y. Fainman, *Phys. Rev. Lett.* **95**, 177401 (2005).
38. J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill, New York, 1996).
39. L. Feng, K. A. Tetz, B. Slutsky, V. Lomakin, and Y. Fainman, *Appl. Phys. Lett.* **91**, 081101 (2007).
40. Z. W. Liu, J. M. Steele, W. Srituravanich, Y. Pikus, C. Sun, and X. Zhang, *Nano Lett.* **5**, 1726 (2005).
41. Z. W. Liu, J. M. Steele, H. Lee, and X. Zhang, *Appl. Phys. Lett.* **88**, 171108 (2006).
42. L. Yin, V. K. Vlasko-Vlasov, J. Pearson, J. M. Hiller, J. Hua, U. Welp, D. E. Brown, and C. W. Kimball, *Nano Lett.* **5**, 1399 (2005).
43. R. Kiyan, C. Reinhardt, S. Passinger, A. L. Stepanov, A. Hohenau, J. R. Krenn, and B. N. Chichkov, *Opt. Express* **15**, 4205 (2007).
44. E. Verhagen, L. Kuipers, and A. Polman, *Nano Lett.* **7**, 334 (2007).
45. K. Li, X. Li, M. I. Stockman, and D. J. Bergman, *Phys. Rev. B* **71**, 115409 (2005).
46. H. Shin and S. Fan, *Phys. Rev. Lett.* **96**, 073907 (2006).
47. E. Hutter and J. H. Fendler, "Exploitation of localized surface plasmon resonance," *Adv. Mater.* **16**, 1685 (2004).
48. M. I. Stockman, *Phys. Rev. Lett.* **93**, 137404 (2004).
49. E. Moreno, S. G. Rodrigo, S. I. Bozhevolnyi, L. Martin-Moreno, and F. J. Garcia-Vidal, *Phys. Rev. Lett.* **100**, 023901 (2008).
50. D. K. Gramotnev, D. F. P. Pile, M. W. Vogel, and X. Zhang, *Phys. Rev. B* **75**, 035431 (2007).
51. E. Verhagen, L. Kuipers, and A. Polman, *Nano Lett.* **7**, 334 (2007).
52. C. Ropers, C. C. Neacsu, T. Elsaesser, M. Albrecht, M. B. Raschke, and C. Lienau, *Nano Lett.* **7**, 2784 (2007).
53. M. Cinchetti, A. Gloskovskii, S. A. Nepjiko, G. Schonhense, H. Rochholz, and M. Kreiter, *Phys. Rev. Lett.* **95**, 047601 (2005).
54. A. Sundaramurthy, P. J. Schuck, N. R. Conley, D. P. Fromm, G. S. Kino, and W. E. Moerner, *Nano Lett.* **6**, 355 (2006).
55. P. Muhlschlegel, H.-J. Eisler, O. J. F. Martin, B. Hecht, and D. W. Pohl, *Science* **308**, 1607 (2005).
56. A. Nesci and Y. Fainman, *Proc. SPIE* **5181**, 62 (2003).

DIELECTRIC-LOADED PLASMONIC WAVEGUIDE COMPONENTS

Tobias Holmgaard* and Sergey I. Bozhevolnyi†

*Department of Physics and Nanotechnology, Aalborg University, Skjernvej 4A, DK-9220, Aalborg Øst, Denmark
holmgaard@nano.aau.dk

†Institute of Sensors, Signals and Electrotechnics (SENSE), University of Southern Denmark, Niels Bohrs Allé 1, DK-5230 Odense M, Denmark

This chapter concerns modeling, design, and characterization of dielectric-loaded plasmonic waveguide components at telecommunication wavelengths, aimed at establishing a new technology platform for integratable nanophotonic components by combining photonics and electronics on the same chip. The dielectric-loaded surface plasmon-polariton waveguides (DLSPPWs) consist of dielectric ridges with rectangular nano-meter-sized cross-sections, deposited on smooth metal films. The SPPs are strongly confined to the ridges due to the large index contrast between the dielectric ridge region and the surrounding air. Theoretical studies of the DLSPPW structure show that single mode propagation with sub-wavelength confinement and low propagation loss can be achieved by proper design of the waveguide dimensions, which is confirmed by near-field optical imaging of fabricated DLSPPWs, performed with a scanning near-field optical microscope. The performance basic waveguide components along with several different passive wavelength selective components is demonstrated. The coupling between two parallel DLSPPWs is investigated by characterizing directional couplers, where the two waveguides are brought into close proximity of one another by means of S-bends. Periodic modulation of the transmission is achieved by realizing waveguide-ring resonators and wavelength filtering is obtained by realizing Bragg gratings. In addition an approach for designing components capable of physically separating two signals of different wavelengths is introduced.

1. Introduction

The remarkable increase in operational speed and component integration of chip-scale electronics observed in the last decades are pushing the data transport capabilities of conventional circuits to the limit. As the electronic components become increasingly small, so does the metal electrodes that connect them, thereby increasing the RC-delay, which limits the operational speed of the circuit. This limited bandwidth will become a severe limitation in the ever increasing demand for faster, smaller, and more efficient computer chips. Photonic interconnects (waveguides) on the contrary offer an almost unimaginable large bandwidth as compared to electronic interconnects, however, due to the diffraction limit of light a sizing mismatch exists between conventional photonic waveguides and electronic circuits which complicates the implementation. One of the potential solutions to circumvent this problem is to utilize photonic circuits based on plasmonics, i.e., on surface plasmon polaritons (SPPs). Plasmonic waveguides are not, in general, restricted in size by the diffraction limit, and thus offer the potential of becoming the next chip-scale technology.[1,2]

Waveguiding of SPP modes can be achieved in many ways, and during the last decades several approaches have been successful in demonstrating guiding of SPPs. Channel SPP waveguides, where lateral confinement is achieved by fabricating a square or V shaped groove in a flat metal surface, have been subject to extensive theoretical and experimental investigation in recent years.[3,4] Rectangular metal stripes, where lateral confinement is achieved by shrinking the lateral extension of a thin metal film, have been investigated theoretically[5] and characterized experimentally.[6] Lateral SPP confinement has also been achieved by utilizing nanostructured periodic variations in the metal surface surrounding the waveguide,[7] or by using chains of closely spaced metal nanoparticles as waveguides.[8]

Lateral confinement of SPPs can alternatively be achieved by utilizing an approach similar to conventional integrated optics, where one drastically reduces a waveguide mode size by employing high-index-contrast waveguides, i.e., waveguides having a core refractive index significantly different from that of the surrounding media.[9] A configuration which can be used for SPP waveguiding consists of a dielectric ridge placed on a metal surface (Fig. 1), where the lateral SPP confinement is achieved due to the fact that the SPP effective index is much larger for the SPP mode supported by the metal-dielectric interface (even for thin dielectric layers) than for that supported by the metal-air interface.[10,11] Borrowing an analogous term from integrated optics[12] this

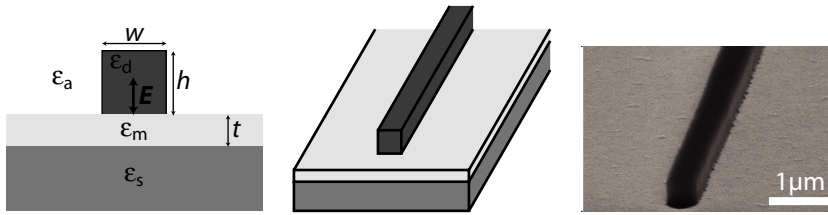


Fig. 1. (a) Illustration of a DLSPPW in the cross-section. The rectangular dielectric ridge of height h, width w, and with dielectric constant ε_{d} is placed on top of a smooth gold film with thickness t and dielectric function ε_{m}. The gold film is supported by a glass substrate with dielectric constant ε_{s}. The electric field is p-polarized. (b) Schematic illustration of a straight DLSPPW and (c) Scanning electron microscope image of a DLSPPW fabricated using deep-UV lithography.

waveguide configuration has been denoted the dielectric-loaded SPP waveguide (DLSPPW) configuration. Early experimental investigations of this waveguide structure have already indicated that the DLSPPWs can be used for efficient guiding of SPP modes.[10,11] The SPP mode is mostly confined to the dielectric ridge, and by properly designing the dimensions of this ridge, one can achieve single-mode guiding of sub-wavelength confined modes with propagation lengths comparable to those of, e.g., CPPs. One of the great assets of this technology in regard to the aforementioned SPP waveguides is that it is technologically simple, and that it can be adapted to large scale industrial fabrication using photolithography. It is furthermore a very versatile technology, as one can choose different types of dielectrics with a desired property, e.g., non-linearity or electro-optically active, and as one still have the possibility of structuring the underlying electrodes.

2. Design of Waveguide Dimensions

When considering the design of waveguides there are several parameters to account for, e.g., propagation loss. The attenuation of the SPP signals inherits from damping of electron oscillations, and this loss channel is one of the critical parameters when designing plasmonic circuits. Another critical parameter is the mode confinement. SPPs feature a strong inherent confinement in the direction perpendicular to the metal-dielectric interface due to the exponential decay away from the interface.[1] At unstructured surfaces, however, there exists no lateral confinement of the SPP modes,

which is essential in order to obtain guiding of SPPs, as small mode widths enable large bend angles with acceptable losses and increases the integration of plasmonic components on a chip. In order to realize plasmonic circuits one must thus design waveguides with a strong lateral (and vertical) confinement, with sufficiently low propagation loss, and which, in general, are needed to support single mode propagation only.

The effective index method (EIM) is one of the standard approaches for mode analysis of photonic waveguides, known to yield accurate mode effective indexes and field profiles for linearly polarized waveguide modes being far from cutoff.[12] The EIM is attractive due to the simplicity of the approach and the low complexity of the calculations, which keeps the computational time low, even for detailed analysis of, e.g., wavelength dependency of the mode effective index and propagation loss. In the consideration of dielectric rib waveguides, known from integrated optical circuits, the EIM provides accurate solutions under the conditions that; the aspect ratio h/w is small, i.e., the ridge is much wider than its height, and that the effective indexes for the ridge region and the regions outside the ridge are close to each other.[13] The plasmonic dielectric ridge waveguide structure, considered here, does not obey these conditions as it is more or less quadratic, and as the difference in refractive index between the dielectric ridge and the surrounding air is large. The EIM has, however, proven to be accurate in the mode analysis of other types of plasmonic waveguides such as metal stripes embedded in a dielectric[14] and rectangular and V-shaped grooves in a smooth metal film,[15] and is thus expected to be applicable in the analysis of DLSPPWs. In order to establish whether the EIM can provide accurate results in the analysis of DLSPPWs, a study of the accuracy has been performed by comparing with results obtained with the finite element method (FEM) for different dimensions of the dielectric ridge.[16] Here it was established that for tightly bound DLSPPW modes, far from cutoff, the deviation in calculated mode index and propagation loss is insignificant, and thus the EIM is applicable for modeling of the DLSPPW structures considered here, which is also supported by the comparison to near-field measurements presented below.

2.1. *Mode confinement and propagation loss*

A thorough investigation of the waveguiding properties, i.e., mode effective index, propagation loss, mode confinement, and single-mode requirement, has been performed for wavelengths in the telecommunication range.[16]

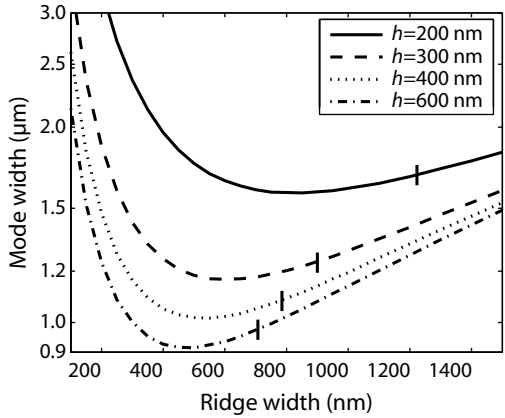


Fig. 2. The mode width (logarithmic scale) depicted versus the ridge width for four different values of ridge thickness h. The vertical lines crossing the curves mark the end of the single-mode regime. The calculations are performed at $\lambda = 1550\,\text{nm}$ by using the EIM. [Adapted from Ref. (16)]

In regard to the mode confinement and propagation loss[a], it is advantageous to keep the height of the waveguide as large as possible (Fig. 2), as allowed by the requirement of single-mode propagation, as most of the mode power in this case will be concentrated in the high-index ridge region. The choice of width, however, is slightly more complicated and is subject to a tradeoff between confinement and propagation loss. Although the mode effective index increases monotonously with an increase in ridge width, the mode width has a distinct minimum (Fig. 2). This can be envisaged as the field is squeezed into the ridge subregion when the width is decreased until a certain point, where the field no longer can be confined to the ridge. It then starts to spread out in the surrounding air, where the decay length is larger, thereby increasing the mode width. Considering the propagation loss, it is advantageous to keep the width of the waveguide as small as possible as an increasing part of the mode power will be carried outside the dielectric ridge, where the propagation loss is smaller. These considerations have led to a determination of the optimum waveguide parameters $h = 600\,\text{nm}$ and $w = 500\,\text{nm}$, which ensures single-mode propagation, propagation lengths of $\sim 50\,\mu\text{m}$, and sub-wavelength lateral confinement (Fig. 2). It should be noted

[a]Strictly speaking the propagation losses can be smaller for very low ridges, but this domain is not considered as the mode confinement drastically deteriorates for low ridges.[16]

that the DLSPPW with these dimensions supports a TE mode, however, by controlling the polarization when exciting DLSPPW modes, this can be ignored as it is not coupled to.

3. Sample Preparation and Near-Field Characterization

The experimental investigations presented in the following are primarily related to near-field optical imaging of samples containing various DLSPPW components. As described above, SPPs are strongly confined to a metal-dielectric interface with exponentially decreasing (evanescent) fields extending only in the order of hundreds of nanometers into the dielectric, which makes investigation with conventional far-field imaging systems impossible. With the further confinement due to the nanostructured high-index dielectric ridges, a high-resolution near-field imaging technique is necessary in the characterization of DLSPPWs. The near-field images presented in this chapter are obtained by a collection mode scanning near-field optical microscope (SNOM), operating in the telecommunication range.

3.1. *Sample configuration and fabrication method*

The DLSPPW samples investigated in this work consist of poly-methyl-methacrylate (PMMA) ridges ($n = 1.493$) deposited on $\sim$60 nm thin gold films, thermally evaporated on 170 μm thick glass substrates. The thickness of the gold film is chosen in order to enable excitation of SPPs from the glass substrate side using the Kretschmann-Raether configuration,[17] while keeping the radiation losses back into the glass substrate as low as possible. After the initial substrate cleaning and gold deposition, a PMMA film was spin coated from 950 K molecular weight PMMA resist (Allresist Gmbh, Ref. AR-P 671.04) and soft-baked. Photolithography was performed with home made and commercially fabricated chromium masks on a Süss Microtech MJB4 mask aligner, equipped with UV250 optics, using the vacuum mode, and the resist was finally developed in methyl-isobutyl-ketone.

Due to in coupling considerations the DLSPPWs are extended by funnel shaped tapering structures [Fig. 3(a)]. In the initial sample investigations four different dimensions of the in coupling funnels were characterized, with a fixed width of 10 μm and a length varying from 10 μm to 25 μm in steps of 5 μm. This was done to determine the optimum funnel dimensions for efficiently exciting DLSPPW modes.[18,19] The width of the fabricated

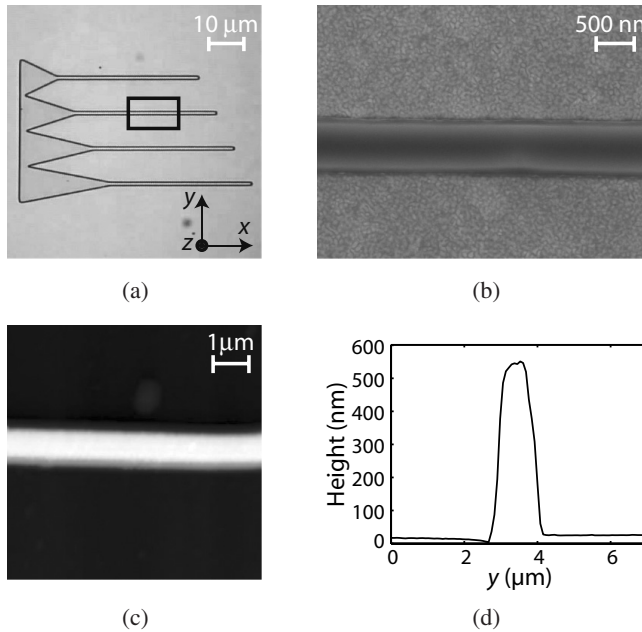


Fig. 3. (a) Microscope image of a DLSPPW block with four different tapering structures. (b) SEM image of a straight section of a waveguide, revealing a waveguide width of $w \simeq 500\,\text{nm}$. (c) AFM image of a straight section of a waveguide, and (d) Cross sectional profile of the AFM image revealing a waveguide height of $h \simeq 550\,\text{nm}$. [Adapted from Ref. (18)]

PMMA ridge waveguides has been investigated using a scanning electron microscope (SEM) [Fig. 3(b)], and is found to be $w \simeq 500\,\text{nm}$ in accordance with the optimum waveguide parameters found with the EIM. The height of the waveguides has been investigated with an atomic force microscope (AFM) [Figs. 3(c, d)], and is determined to be $h \simeq 550\,\text{nm}$, which is only slightly lower than the desired optimum height of $h = 600\,\text{nm}$. In addition to the straight waveguides several different DLSPPW components have been fabricated, such as waveguide S-bends, Y-splitters, waveguide-ring resonators, directional couplers, Mach-Zehnder interferometers, and Bragg gratings.

3.2. *Near-field imaging of DLSPPW components*

The fabricated DLSPPW samples are investigated using a SNOM operating in collection mode at telecommunication wavelengths. In order to excite

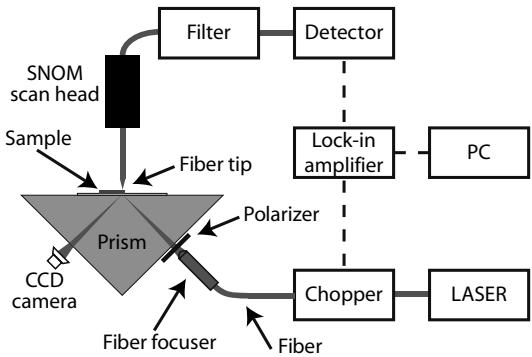


Fig. 4. The experimental setup of the collection mode SNOM imaging system used to obtain near-field optical and topographical images of the DLSPPW structures. [Adapted from Ref. (18)]

DLSPPW modes, the samples are placed on an equilateral prism using index matching immersion oil between the sample substrate and the prism. A pigtail style fiber focuser is used to focus a *p*-polarized Gaussian laser beam at the gold surface, opposite to the DLSPPW structures, to a spot size of $\sim 15\,\mu$m (Fig. 4). Using two tunable lasers the telecommunication wavelength range from 1425 nm to 1625 nm can be covered. SPPs are excited directly in the dielectric funnel, by employing a high-index prism and focusing the spot directly below the funnel and matching the lateral component of the incident wavevector to that of SPPs propagating along the gold-PMMA interface, by adjustment of the incident angle.[17,18]

Near-field optical and topographical images of the DLSPPW components are recorded simultaneously by raster scanning an uncoated tapered fiber tip across the sample surface. The fiber tip is etched to a tip diameter of a few hundred nanometers using hydrofluoric acid. When scanning the etched fiber across the sample surface, the tip scatters the optical near-field at an effective detection point ~ 100 nm from its extremity point inside the fiber.[20] A part of this scattered field is coupled into fiber modes, and is thus directed to the detector, where the intensity is recorded by a femtowatt InGaAs photoreceiver. As the optical signal incident on the detector is very weak a lock-in amplifier is used in conjunction with a chopper (inserted just after the laser before coupling into SPPs) in order to amplify the signal and reduce the disturbances from ambient radiation. In addition to this, a filter is inserted before the detector in order to minimize the disturbance from ambient illumination. The movement of the fiber tip in the

xy-plane and the tip-sample distance is controlled by a SNOM scan head. The SNOM scan head is fabricated by Danish Micro Engineering A/S, and it operates according to the shear force principle, where the tip oscillates in the plane parallel with the sample surface. Excitation and detection of the tip oscillations are made by means of piezo crystals. Coarse adjustment of the scan area can be performed as the SNOM scan head is fixed on a 3D scan stage which can be controlled by micrometer screws.

4. Excitation and Propagation of Guided Modes

In order to characterize various DLSPPW components it is imperative to realize strong excitation of DLSPPW modes with a large signal to noise ratio. The excitation scheme, where DLSPP modes are excited directly in the tapering structure, is attractive as no copropagating SPP fields at the gold-air interface are excited due to the much higher effective index of the dielectric funnel region. This is observed by SNOM characterization of a $10 \times 25\,\mu\text{m}^2$ funnel (Fig. 5). An averaged cross-sectional profile of the near-field optical image taken at the straight waveguide section just after the end of the funnel shows a highly confined DLSPPW mode with

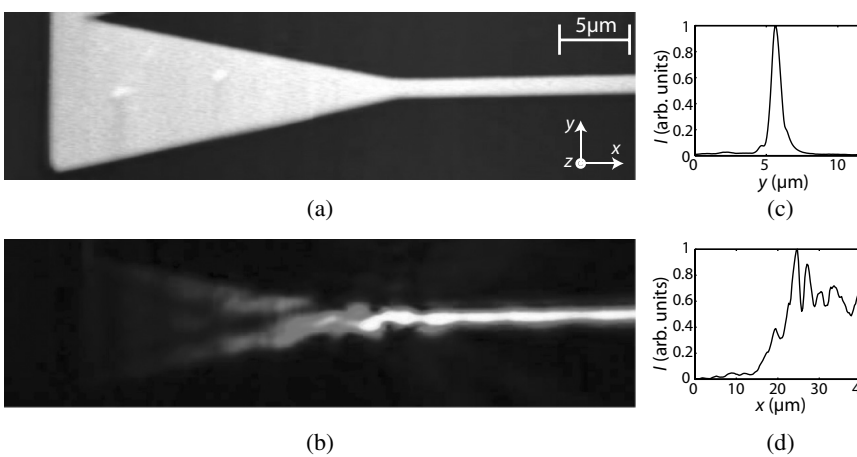


Fig. 5. SNOM images of a $10 \times 25\,\mu\text{m}^2$ funnel in coupling, where SPPs are excited at the gold-PMMA interface inside the tapering region, for the free space excitation wavelength $\lambda = 1550\,\text{nm}$. (a, b) Topographical and near-field optical images, respectively, obtained with the SNOM. (c) Averaged cross sectional profile of the near-field optical image taken at the straight waveguide section just after the end of the funnel. (d) Averaged profile of the near-field optical image taken parallel with the waveguide through the middle of the funnel and waveguide regions. [Adapted from Ref. (18)]

no copropagating SPPs at the gold-air interface and almost no apparent scattering originating from the end of the funnel [Fig. 5(c)]. A profile of the near-field optical signal taken parallel to the waveguide shows a strong buildup of the DLSPP mode in the tapering region [Fig. 5(d)]. This effect is caused by several factors all contributing to a stronger optical signal toward the end of the tapering region. First the excitation of SPPs by the attenuated TIR method causes a strong initial increase in SPP intensity, as the propagation loss initially has a small effect. Second the tapering structure has a focusing effect toward the funnel end due to TIR of DLSPPs, which will also result in a stronger optical signal toward the funnel end. Finally the imaging process further enhances this trend as near-field optical modes with lower effective indexes are picked up more strongly with the tapered fiber, and as the mode effective index of a dielectric ridge decreases with width[16] this also contributes to a stronger optical signal toward the funnel end. Due to the strong excitation of DLSPPWs without causing copropagating SPPs and scattering of SPPs, this excitation scheme is found highly attractive and has thus been used in the achievement of the results presented in the following.[18]

4.1. *Mode confinement and propagation loss*

The propagation and confinement of DLSPPW modes in straight waveguides have been characterized by considering a waveguide section $\sim 50\,\mu$m after the tapering region using SNOM imaging at the free space excitation wavelength $\lambda = 1550\,$nm [Figs. 6(a, b)]. The near-field optical image shows a strongly confined DLSPPW mode, and the absence of mode beating confirms that the designed DLSPPW structure indeed only supports a single TM mode. The mode width is investigated by making an averaged cross-sectional profile of the near-field optical signal, revealing a full width at half maximum (FWHM) of $\sim 743\,$nm [Fig. 6(e)], indicating subwavelength confinement. The waveguide termination region has also been investigated by using SNOM imaging, and it is found that even after propagation over more than $\sim 100\,\mu$m a strong DLSPPW signal exists as can be observed from the strong scattering at the waveguide termination [Figs. 6(c, d)]. An interference pattern can be observed in the near-field optical signal, and by considering a profile taken along the waveguide the interference period is found to be $\Lambda \sim 700\,$nm [Fig. 6(f)]. This interference pattern is caused by interference between the forward propagating DLSPPW (TM_{00}) mode with mode effective index $N_{\text{eff}} \sim 1.21$ (EIM calculations at $\lambda = 1550\,$nm

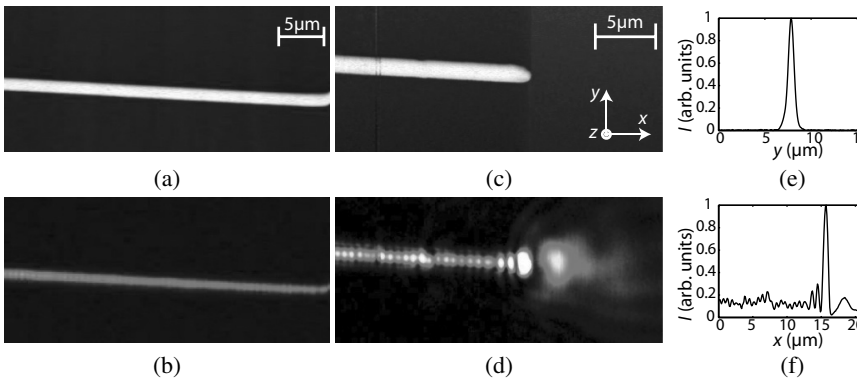


Fig. 6. SNOM images of a straight DLSPPW section and a waveguide termination, recorded at the free space excitation wavelength $\lambda = 1550\,\text{nm}$. (a, b) Topographical and near-field optical images, respectively, of a straight DLSPPW section. (c, d) Topographical and near-field optical images, respectively, of a waveguide termination. (e) Cross sectional profile of the near-field optical image in (b) yielding a FWHM $\sim 743\,\text{nm}$. (f) Profile of the near-field optical image in (d) taken parallel with the waveguide, revealing an interference pattern with a period of $\Lambda \sim 700\,\text{nm}$. [Adapted from Ref. (18)]

yields $N_{\text{eff}} = 1.22$) and a backward propagating light wave with effective index close to 1. The origin of the backward propagating light wave is the waveguide termination, where the DLSPPW mode is scattered from. No significant back reflection of the DLSPPW mode inside the ridge for the termination is expected due to the, in this respect, relatively small difference in mode effective index between the DLSPPW mode and an SPP mode at the gold-air interface, and the fact that the termination of the dielectric ridge is not completely abrupt. This is supported from FEM simulations of a DLSPPW termination, where the reflectivity has been estimated to be $\sim 5\%$.[18]

The propagation loss in the fabricated DLSPPWs has been investigated by using near-field imaging with a SNOM to characterize a straight waveguide section at different wavelengths in the telecommunication range (Fig. 7). From EIM calculations the DLSPPWs are expected to show similar properties at the investigated telecommunication wavelengths, however, with an increase in propagation length with increasing wavelength, at the expense of poorer mode confinement (lower mode effective index) [Figs. 7(e, f)]. This trend is confirmed from contemplation of the obtained near-field optical images, where an exponential fit to an averaged profile taken along the waveguide yields propagation lengths of $L = 46\,\mu\text{m}$, $L = 52\,\mu\text{m}$, and $L = 65\,\mu\text{m}$ at the free space excitation wavelengths

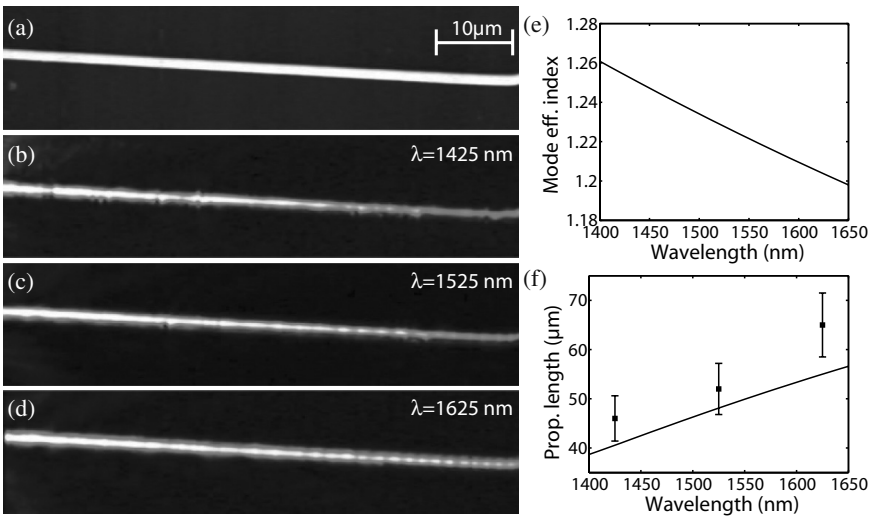


Fig. 7. SNOM images of a straight waveguide at different free space excitation wavelengths. (a) Topographic image, (b) Near-field optical image for $\lambda = 1425\,\text{nm}$, where the propagation length is determined to $L = 46\,\mu\text{m}$, (c) Near-field optical image for $\lambda = 1525\,\text{nm}$, where $L = 52\,\mu\text{m}$, and (d) Near-field optical image for $\lambda = 1625\,\text{nm}$, where $L = 65\,\mu\text{m}$. (e, f) Mode effective index and propagation length, respectively, as a function of wavelength, calculated by utilizing the EIM (fully drawn line) and measured with the SNOM (squares). [Adapted from Ref. (18)]

$\lambda = 1425\,\text{nm}$, $\lambda = 1525\,\text{nm}$, and $\lambda = 1625\,\text{nm}$, respectively [Figs. 7(b, d)]. It is observed that the measured propagation lengths in all cases are longer than those expected from the EIM calculations conducted for the DLSPPWs having the designed ridge width and height. This can be caused by several effects, where the most likely are found to be discrepancy in waveguide dimensions, e.g., smaller waveguide width than designed or a slightly trapezoidal shape of the waveguide, which can result in quite large deviations in propagation length.[16] Another possible cause could be discrepancy between the actual gold refractive index and that used in the EIM calculations.

5. Waveguide Bends and Splitters

To obtain high integration of circuits based on DLSPPWs it is necessary to realize several compact basic components such as waveguide bends and splitters. The design of S-bends, investigated in the following, is based on sine curves,[21] allowing for continuous bend curvature and thereby adiabatic

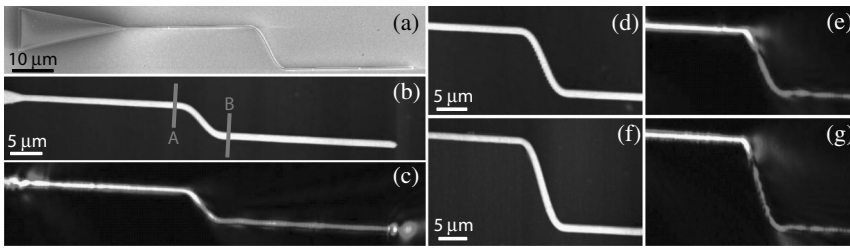


Fig. 8. Plasmonic tapers and S-bends: (a) SEM image of a taper waveguide structure comprising a 10-μm-long S-bend with an offset of 10 μm; (b) Topographical and (c) Near-field optical ($\lambda = 1.55\,\mu$m) images of the SPP mode propagation in a 10-μm-long S-bend with an offset of 5 μm (note the out-coupling of the SPP waveguide mode into a free-propagating and diverging SPP wave at the waveguide termination). (d, e) As (b, c) but for a waveguide offset of 10 μm. (f, g) As (d, e) but for a waveguide offset of 15 μm. [Adapted from Ref. (22)]

modification of the DLSPPW mode throughout the bend. Using the SNOM several 10-μm-long S-bends with different offsets d, ranging from 2 to 15 μm has been characterized. Excellent performance is observed for small offsets ($d \leq 5\,\mu$m) [Figs. 8(b, c)] which gradually deteriorates with the offset increase due to the radiation (into SPP waves) out of the bend [Figs. 8(d, e)]. However, even the S-bend with the largest offset ($d = 15\,\mu$m), and the smallest curvature radius of 1.95 μm, transmits a noticeable amount of DLSPPW mode radiation [Figs. 8(f, g)] confirming the strong confinement of the DLSPPW mode.

Bend transmission data (Fig. 9) has been calculated using full vectorial three-dimensional finite element method (3D-FEM) numerical simulations and obtained directly from the SNOM images by averaging the optical signal over a few micrometers just before and after the S-bend [see lines A and B in Fig. 8(b)]. The calculated and measured transmission values were found in good agreement with each other (Fig. 9), and this agreement is also observed in the optical images when comparing the field intensity distributions obtained with the 3D-FEM simulations (see insets in Fig. 9) with the near-field images obtained with the SNOM (Fig. 8). Note that the transmission shown in Fig. 9 reflects not only the pure bend loss (due to the mode radiation out of the bend) but also the DLSPPW mode propagation loss (due to absorption in gold) accumulated over the length of 10-μm-long S-bends. Given the DLSPPW propagation length (whether it being $\sim$47 μm found with the 3D-FEM simulations or $\sim$50 μm obtained from the SNOM images), the pure bend loss is found to be very small (<10%) for $d \leq 3\,\mu$m, whereas it constitutes the main loss contribution for $d > 6\,\mu$m.[22]

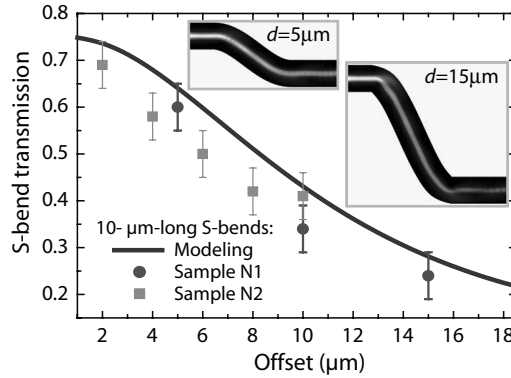


Fig. 9. S-bend transmission dependence on the offset distance between the arms for the 10-μm-long S-bends ($\lambda = 1.55\,\mu$m) determined experimentally from SNOM images (similar to those shown in Fig. 8) and by modeling with 3D-FEM simulations. Insets show calculated intensity distributions in two S-bends with different offsets that should be compared to the optical images shown in Figs. 8(c, g). [Adapted from Ref. (22)]

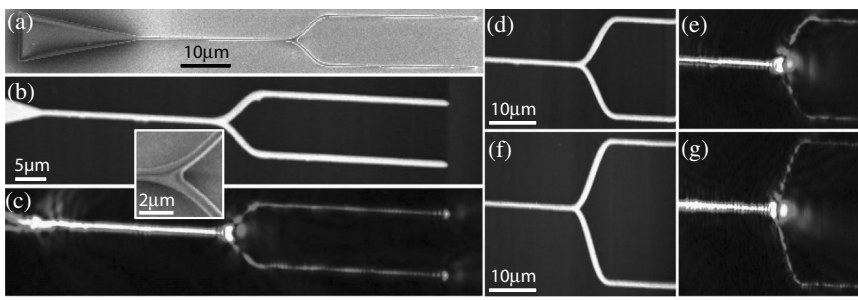


Fig. 10. Plasmonic Y-splitters composed of two mirrored S-bends: (a) SEM image of a taper waveguide structure comprising a 10-μm-long Y-splitter with the arm separation of 10 μm; (b) Topographical and (c) Near-field optical ($\lambda = 1.55\,\mu$m) images of the Y-splitter shown in (a) along with an inset showing an SEM image of the junction area. (d, e) As (b, c) but for the arm separation of 20 μm. (f, g) As (d, e) but for the arm separation of 30 μm. [Adapted from Ref. (22)]

The performance of 10-μm-long Y-splitters has been studied in a similar manner by carrying out the SNOM characterization of Y-splitters fabricated with different arm separations ranging from 3 to 30 μm (Fig. 10). It is found that the Y-splitters feature pronounced rounding (∼300-nm-radius) of the junction area [inset in Figs. 10(b, c)] due to limited resolution of the UV-lithography technique used for their fabrication. Such a rounding prevents adiabatic separation of the waveguides and introduces

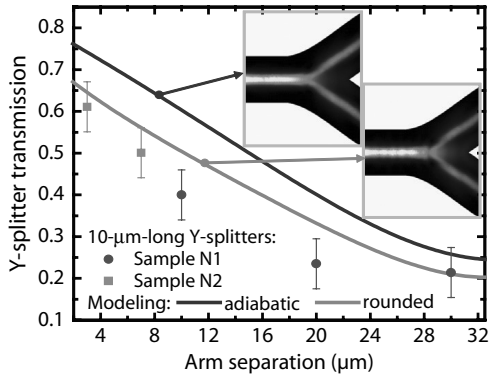


Fig. 11. Transmission of 10-μm-long Y-splitters ($\lambda = 1.55\,\mu$m) as a function of arm separation determined experimentally from the SNOM images similar to the ones shown in Fig. 10 and by modeling with 3D-FEM simulations. Insets show the calculated intensity distributions in 10-μm-wide Y-splitters having adiabatic arm separation and 300-nm-radius rounding of the junction area respectively [the latter should be compared to the optical image in Fig. 10(c)]. [Adapted from Ref. (22)]

a mismatch between the incoming mode field (into the junction area) and the field of outgoing modes in the waveguide branches, resulting thereby in a transmission loss due to radiation out of the junction as SPP waves propagating in the forward direction on the smooth gold surface [Figs. 10(c, e and g)]. The occurrence of this radiation along with its influence on the Y-splitter transmission was further verified with the 3D-FEM simulations (Fig. 11). Simulation of Y-splitters with a 300-nm-radius rounding of the junction area [derived from the SEM image shown as inset in Figs. 10(b, c)] shows a $\sim$20% decrease in transmission related to the adiabatic splitters, which is in good agreement with the experimentally obtained values. Note that the transmission of adiabatic Y-splitters is expected to be very close to that of the corresponding S-bends (cf. Figs. 9 and 11), impelling one to improve on the resolution in DLSPPW fabrication, e.g., by exploiting optical proximity correction techniques.[22]

6. Coupling between Waveguides

In order to realize compact plasmonic devices, wavelength-selective components for filtering and separation of wavelengths are needed. Such wavelength-selective components are essential in the realization of demultiplexers and could be based on, e.g., directional couplers (DCs).

In guided-wave photonics there are generally two approaches to analyze coupling between waveguides. The coupled-mode approach, where the presence of a second identical waveguide is included by considering it a perturbation to the unperturbed field profiles associated with a mode incident in the first waveguide, and the super-mode approach, where the phase difference between the two modes supported by the entire structure is considered. The coupled-mode approach has the advantage of being computationally simple, however, since the field profiles used in the calculation of the coupling are those of two waveguides in isolation, they are approximations valid for weak coupling only, i.e., large separation between the waveguides.[13] This is not the case for the plasmonic DCs considered here, and thus the super-mode approach is applied in the following.

The investigated DCs basically consists of four sections, being the funnel shaped excitation region, an in-coupling S-bend bringing the two waveguides into close proximity, a parallel interaction region, and an out-coupling S-bend separating the two waveguides again [Fig. 12(a)]. In the notation used throughout, the initial and final center-to-center separation

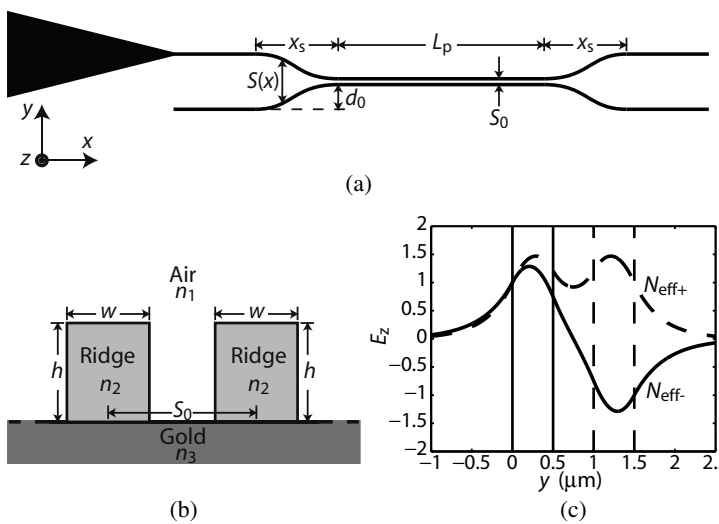


Fig. 12. DC layout and design parameters. (a) DC layout consisting of a funnel used for exciting DLSPPWs in the experiments, an S-bend "in-coupling" region, a parallel interaction section where the waveguide modes overlap, and an S-bend "out-coupling" region. (b) Cross-sectional view of the DC at the interaction region, where the center-to-center separation of the two waveguides is S_0. (c) Cross-sectional profile of the EIM calculated symmetric ($N_{\text{eff}+}$) and anti-symmetric ($N_{\text{eff}-}$) DLSPPW modes supported by the DC structure. [Adapted from Ref. (23)]

of the two waveguides is $2d_0 + S_0$, the separation at the parallel section is S_0, the length of the parallel section is L_p, and the length of the S-bends is x_s. The S-bends are based on cosine curves, where the center-to-center separation is given as

$$S(x) = S_0 + d_0 \left[1 + \cos\left(\frac{\pi x}{x_\mathrm{s}}\right)\right], \tag{1}$$

in the case of the in-coupling S-bend ($x = 0$ at the beginning of the bend). The DLSPPWs consists of a rectangular dielectric ridge of height h, width w, and refractive index n_2 deposited on a metal film of refractive index n_3, and bound by air, n_1, above [Fig. 12(b)].

When considering DCs by means of the EIM, the first step in the calculation is to find the mode effective index of a three layer air-dielectric-metal geometry with infinite height of the air and metal layer, and height h of the dielectric layer, corresponding to setting $w \sim \infty$. In this calculation p-polarized modes are assumed, corresponding to SPP modes. The mode effective index found in this calculation is then used to represent the two waveguide regions of width w in the second step. Outside these regions the mode effective index of an SPP propagating along a smooth metal film with air above is used. This five layer geometry is then solved for s-polarized modes, and a symmetric ($N_{\mathrm{eff}+}$) and anti-symmetric ($N_{\mathrm{eff}-}$) mode is found [Fig. 12(c)]. These two modes are orthogonal and uncoupled, thus if the propagation losses are ignored, the only effect on the modes due to propagation over a distance L, is the phase shift

$$\phi_\pm = k_0 N_{\mathrm{eff}\pm} L = \frac{2\pi}{\lambda} N_{\mathrm{eff}\pm} L, \tag{2}$$

where λ is the vacuum wavelength. When the accumulated phase difference between the two modes $\Delta\phi = \phi_+ - \phi_-$ reaches a value of π the mode power has shifted completely from one waveguide to the other. The length corresponding to this shift is denoted the coupling length L_c and can be found as

$$\Delta\phi = \frac{2\pi}{\lambda} \Delta N_\mathrm{eff} L_\mathrm{c} = \pi \Rightarrow L_\mathrm{c} = \frac{\lambda}{2\Delta N_\mathrm{eff}}. \tag{3}$$

where $\Delta N_\mathrm{eff} = N_{\mathrm{eff}+} - N_{\mathrm{eff}-}$. Thus knowing the dependencies of the effective indexes of the symmetric and anti-symmetric modes on, e.g., wavelength and waveguide separation, one can analyze the coupling properties of the DC.

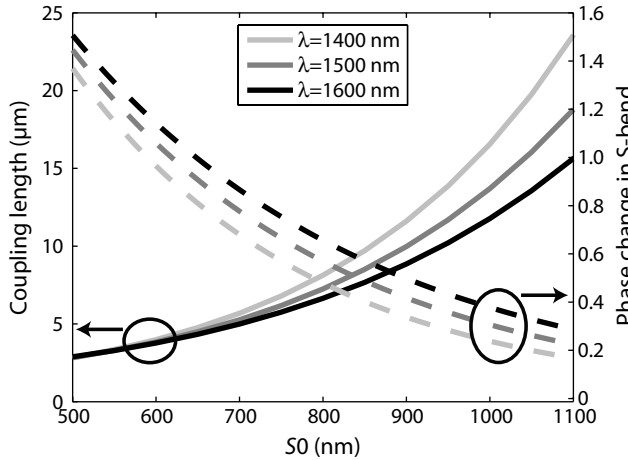


Fig. 13. Coupling length and phase change in S-bend dependence on final waveguide separation (S_0) calculated utilizing the EIM. In the phase change calculation the S-bend parameters $d_0 = 3\,\mu$m and $x_s = 10\,\mu$m are used. [Adapted from Ref. (23)]

In the calculations gold is used for the metal film, which is considered optically thick, with the complex refractive index taken from Ref. (24). The dielectric used is PMMA with refractive index $n_2 = 1.493$, $h = 600$ nm, and $w = 500$ nm.[16] By varying the gap between the two waveguides in the calculation, a dependence of the coupling length on S_0 can be established (Fig. 13). For a DC where the two waveguides are brought into contact with one another, i.e., $S_0 = w$, a very short coupling length with weak dependence on wavelength is expected as most of the mode power is concentrated inside the double width waveguide. However, in general the coupling lengths are expected to be larger for shorter wavelengths due to the better confinement to the ridge, and thus less coupling to an adjacent ridge. An increase in S_0 increases the coupling length exponentially, as the DLSPPW mode fields decrease exponentially outside the ridge. These trends are confirmed in the EIM calculations of the coupling length (Fig. 13).

When extending the calculations to considering coupling between the two waveguides in the S-bends, with non-uniform separation between the waveguides, a slightly modified approach is needed. In order to find the phase change due to coupling in the S-bend, a dependence of the coupling constant κ, given as

$$\kappa = \frac{\pi}{2L_c} = \frac{\pi \Delta N_{\text{eff}}}{\lambda}, \tag{4}$$

on the separation $S(x)$ must be established by means of the EIM. As the coupling length shows an exponential dependence on S, so does κ, and by fitting the EIM results of κ, the analytical expression

$$\kappa(S) = A\exp(BS) \tag{5}$$

can be established, with A and B being wavelength dependent factors. Using this, and the expression for the waveguide separation throughout the S-bend given by Eq. (1), the phase change due to an S-bend can be found as

$$\Delta\phi_{\mathrm{s}} = 2\int_0^{x_{\mathrm{s}}} \kappa(S(x))\mathrm{d}x. \tag{6}$$

Solving this equation, using the S-bend parameters $d_0 = 3\,\mu$m and $x_{\mathrm{s}} = 10\,\mu$m (used in the fabricated DC samples), confirms that the coupling increases with wavelength, but decreases exponentially with S_0 (Fig. 13) as was the case in the parallel section. The results clearly illustrate the importance of including the coupling between the waveguides in the in- and out-coupling S-bends, particularly for small S_0 values.

By being able to describe the coupling in the parallel interaction region and in the in- and out-coupling S-bends, one is able to analyze the performance of the whole DC device and thus being able to describe the influence of the wavelength, the initial separation, the separation in the parallel section, and the length of the parallel section. The total accumulated phase difference can be calculated as

$$\Delta\phi_{\mathrm{tot}} = 2\Delta\phi_{\mathrm{s}} + \pi\frac{L_{\mathrm{p}}}{L_{\mathrm{c}}}, \tag{7}$$

where the first term is related to the two S-bends and the second term to the parallel section.[23]

6.1. *Near-field characterization*

In order to validate the applicability of the EIM in the analysis of DLSPPW based DCs, the above results are compared to results obtained from fabricated DC structures, by application of a SNOM. The DC structures are fabricated using the S-bend parameters $d_0 = 3\,\mu$m and $x_{\mathrm{s}} = 10\,\mu$m while varying S_0. The length of the parallel section is kept constant $L_{\mathrm{p}} = 25\,\mu$m. The recorded topographical images confirm the fabrication of DCs of high quality with a very low degree of defects [Fig. 14(a)], and the near-field optical images show the coupling between the straight (the one excited)

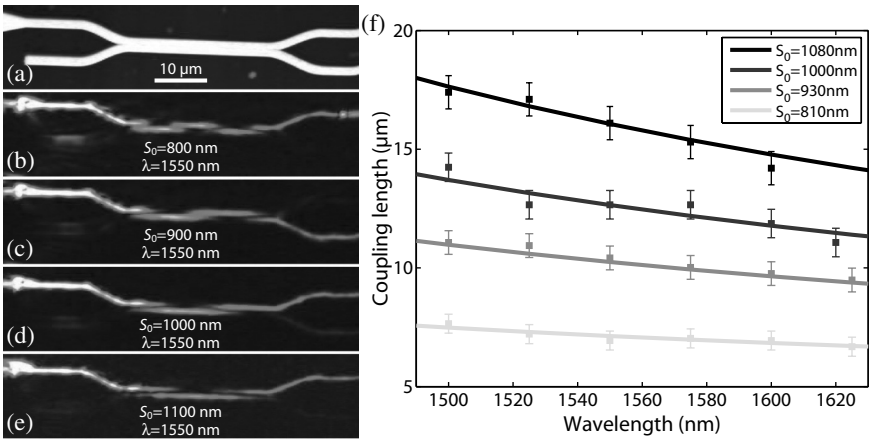


Fig. 14. Coupling length dependence on wavelength and waveguide separation. (a) Topographical ($S_0 = 1000$ nm) and (b–e) Near-field optical (at $\lambda = 1550$ nm) images of DCs measured with a SNOM. (b) $S_0 = 800$ nm, (c) $S_0 = 900$ nm, (d) $S_0 = 1000$ nm, and (e) $S_0 = 1100$ nm. (f) Coupling length as a function of wavelength measured with the SNOM (squares) along with EIM calculated dependence (lines) fitted to the measurements by slight adjustment of S_0. [Adapted from Ref. (23)]

and cross arm [Figs. 14(b–e)]. The rapid increase in coupling length with separation is clearly observed in the recorded near-field images, and by performing a wavelength scan of each of the DCs, the dispersion with respect to wavelength can be found. By determining the coupling lengths from profiles of the near-field optical images, and finding the best fit of EIM calculated curves (fitting S_0) it is observed that the EIM describes the behavior of the DCs with regard to separation and wavelength well [Fig. 14(f)].

Observing the performance of one of the fabricated DCs with $S_0 = 800$ nm, by contemplation of the near-field images at two different wavelengths $\lambda = 1500$ nm [Fig. 15(b)] and $\lambda = 1600$ nm [Fig. 15(c)], shows that although the coupling is different at the two wavelengths, the output of the DC is almost identical [Fig. 15(f)]. Note that the bright spot at the output of the straight arm is caused by scattering due to a defect, which occurred during near-field investigation. This performance is in strong contrast to the performance of another of the fabricated DCs with $S_0 = 1000$ nm. In this case close to all of the mode power is transmitted through the straight arm at $\lambda = 1500$ nm [Fig. 15(d)], whereas a 50/50 splitting occurs at $\lambda = 1600$ nm [Fig. 15(e)], which is clearly illustrated by the cross-sectional profiles at the DC output [Fig. 15(f)]. Inclusion of the S-bends in the

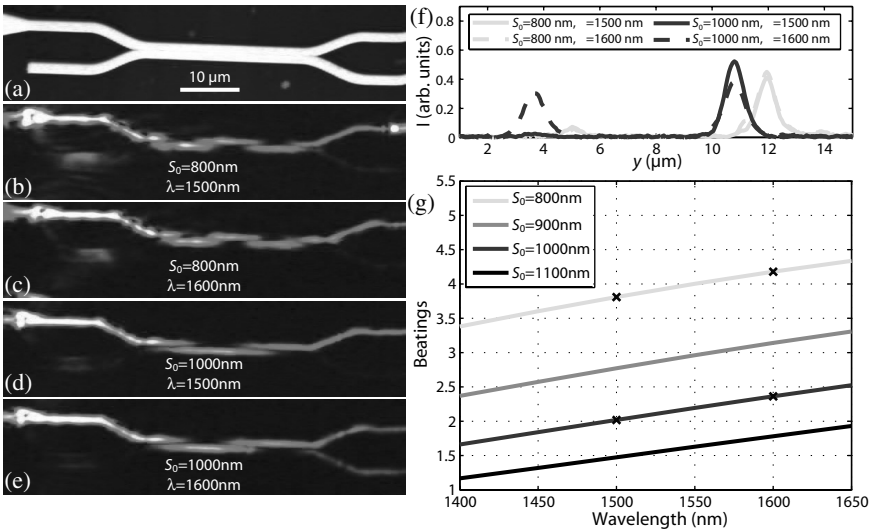


Fig. 15. Near-field investigation and EIM calculation of number of beatings in DCs with different waveguide separations. (a) Topographical ($S_0 = 1000\,\text{nm}$) and near-field optical (b–e) images of DCs measured with a SNOM. (b) $S_0 = 800\,\text{nm}$ and $\lambda = 1500\,\text{nm}$, (c) $S_0 = 800\,\text{nm}$ and $\lambda = 1600\,\text{nm}$, (d) $S_0 = 1000\,\text{nm}$ and $\lambda = 1500\,\text{nm}$, and (e) $S_0 = 1000\,\text{nm}$ and $\lambda = 1600\,\text{nm}$. (f) Cross-sectional profiles of the near-field optical images (b–e) taken at the DC output, i.e., after the separation of the waveguides. (g) EIM calculation (lines) of the total number of beatings throughout the DCs (including the two S-bends) as a function of wavelength, with parameters identical to those used in the design of the fabricated sample. The four cases investigated with the SNOM shown in the near-field images (b–e) are marked with crosses. [Adapted from Ref. (23)]

EIM calculation enables the calculation of the total number of beatings throughout the whole DC, which can easily be compared to the near-field optical images, and explain the above observation. The total number of beatings is found by using Eq. (7) to calculate the total phase change in the DC, and dividing by π [Fig. 15(g)]. An even number of beatings means transmission through the straight arm, whereas an odd number of beatings results in transmission through the cross arm. From the EIM results it is clear that the $S_0 = 800\,\text{nm}$ curve intersects 4 beatings at $\lambda = 1550\,\text{nm}$, i.e., in the middle of the band bound by $\lambda = 1500\,\text{nm}$ and $\lambda = 1600\,\text{nm}$, and thus here transmits all mode power in the straight arm [see Fig. 14(b)]. In the case of the $S_0 = 1000\,\text{nm}$ DC, however, the curve intersects 2 beatings right at $\lambda = 1500$. This clearly demonstrates a need for careful design, when applying DCs for physical separation of different wavelengths, as the output from the $S_0 = 800\,\text{nm}$ DC is the same at $\lambda = 1500\,\text{nm}$ and $\lambda = 1600\,\text{nm}$.

In addition the results presented in Fig. 15 show a very good correspondence between EIM calculations and the results obtained by performing SNOM measurements on fabricated DC structures, thus confirming that the EIM is applicable in the analysis of DCs based on DLSPPWs.[23]

6.2. *Design of wavelength selective DC*

Physical separation of two distinct wavelengths propagating in the same multi-channel waveguide can be achieved with a DC by carefully designing the length of the parallel section L_p and the separation S_0. In order to achieve separation of two signals with the wavelengths λ_1 and λ_2, both incident in the straight channel, the total number of beatings through the DC at the two wavelengths B_1 and B_2 must be positive integers and the difference in beatings must be one,

$$1 = B_2 - B_1, \tag{8}$$

where it is assumed that $\lambda_1 < \lambda_2$. Expressing B_1 and B_2 in terms of L_p, L_c, and $\Delta\phi_\mathrm{s}$ using Eq. (7) yields

$$B_1 = \frac{L_\mathrm{p}}{L_{\mathrm{c}1}} + \frac{2\Delta\phi_{\mathrm{s}1}}{\pi}, \tag{9}$$

and

$$B_2 = \frac{L_\mathrm{p}}{L_{\mathrm{c}2}} + \frac{2\Delta\phi_{\mathrm{s}2}}{\pi}, \tag{10}$$

where $L_{\mathrm{c}1}$ and $\Delta\phi_{\mathrm{s}1}$ depend on λ_1 and S_0, and $L_{\mathrm{c}2}$ and $\Delta\phi_{\mathrm{s}2}$ depend on λ_2 and S_0 (assuming d_0 and x_s fixed). By insertion in Eq. (8) one can express the length of the parallel section in terms of the coupling lengths and phase changes in the S-bends as

$$L_\mathrm{p} = \frac{\left(1 - 2\frac{\Delta\phi_{\mathrm{s}2} - \Delta\phi_{\mathrm{s}1}}{\pi}\right)(L_{\mathrm{c}1}L_{\mathrm{c}2})}{L_{\mathrm{c}1} - L_{\mathrm{c}2}}. \tag{11}$$

Choosing the two wavelengths that must be separated in the DC, one can find the S_0 dependency of L_p. For small separations $L_{\mathrm{c}1}$ approaches $L_{\mathrm{c}2}$ (Fig. 13) and L_p thus increases rapidly as the denominator approaches zero. For large S_0, $L_{\mathrm{c}1}$ and $L_{\mathrm{c}2}$ increase exponentially and thus L_p increases rapidly also. Thus a minimum in L_p is expected and indeed found when plotting L_p versus S_0 for the choice of wavelengths $\lambda_1 = 1400\,\mathrm{nm}$ and $\lambda_2 = 1600\,\mathrm{nm}$ and the S-bend parameters $d_0 = 3\,\mu\mathrm{m}$ and $x_\mathrm{s} = 10\,\mu\mathrm{m}$ (Fig. 16).

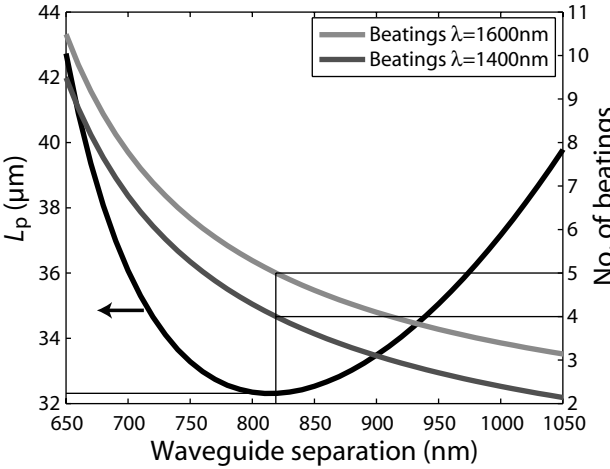


Fig. 16. Illustration of the EIM design approach for achieving physical separation of the two wavelengths $\lambda_1 = 1400\,\text{nm}$ and $\lambda_2 = 1600\,\text{nm}$. The length of the parallel section is shown as a function of the waveguide separation on the left axis, calculated using Eq. (11). The number of total beatings through the DC for the two wavelengths is plotted on the right axis as a function of waveguide separation using Eqs. (9) and (10). The optimum (minimized device length) separation and length of parallel section are marked by the thin lines. [Adapted from Ref. (23)]

Using Eq. (9) and Eq. (10) one can now determine the number of beatings at the two wavelengths for each S_0 value. The design task is then simplified to choosing the B_1 value which makes the device as short as possible, and the optimum S_0 and L_p can be found. In the design example shown in (Fig. 16) $B_1 = 4$ is found to best match the minimum in L_p, yielding the optimum device parameters $S_0 = 819\,\text{nm}$ and $L_p = 32.3\,\mu\text{m}$ giving a total device length $L_d = 52.3\,\mu\text{m}$ (due to the two S-bends), which is comparable to the propagation length currently achieved with the DLSPPW technology.[18] It is thus illustrated that the design of very efficient wavelength separation components, with spectra similar to that of Fig. 3 in Ref. (25), is possible by utilizing the EIM in the analysis of DCs.[23]

7. Waveguide-Ring Resonators

The waveguide-ring resonator (WRR) is a photonic component well known to yield periodic modulation of the output signal, which makes it attractive in plasmonic circuits due to the possibility of filtering out undesired wavelengths. The DLSPPW based WRRs studied here have been designed

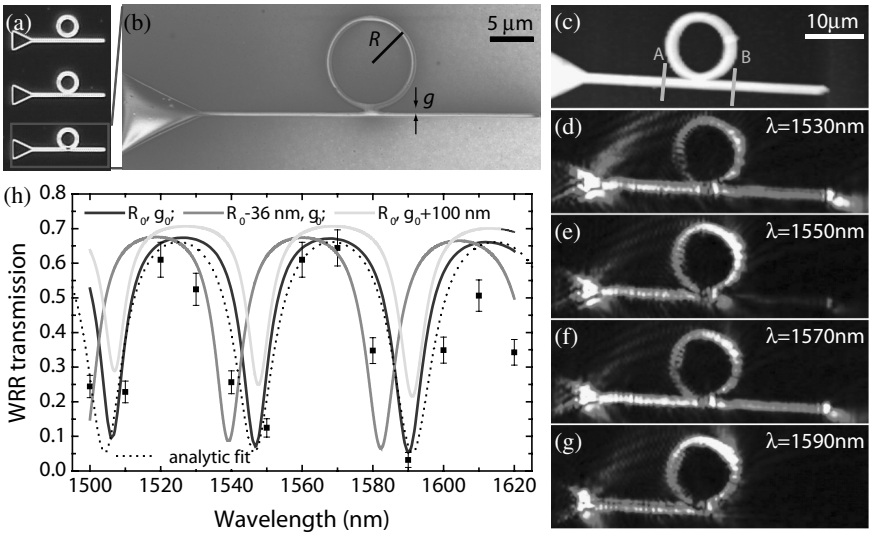


Fig. 17. Plasmonic (WRR). (a) Dark-field microscope image of fabricated WRRs with different gaps. (b) SEM image along with (c) Topographical and (d–g) Near-field optical [λ = (d) 1530, (e) 1550, (f) 1570 and (g) 1590 nm] SNOM images of the investigated WRR. (h) Transmission WRR spectra determined experimentally from SNOM images [similar to (d–g)] using an analytic fit by Eq. (12) and with 3D-FEM simulations for 10-μm-long input-to-output propagation [marked by lines A and B in (c)] and for three sets of WRR parameters: ring radius $R_0 = 5486$ nm and gap width $g_0 = 250$ nm, $R = 5450$ nm and $g = 250$ nm, $R = 5486$ nm and $g = 350$ nm. [Adapted from Ref. (26)]

to have 5-μm-radius rings separated by different gaps from straight waveguides [Fig. 17(a)]. Using the SNOM imaging it was found that while the WRRs with the gaps (the minimum distance between the edge of the waveguide and the edge of the ring [Fig. 17(b)]) $g \cong 2.53$ and $1.53\,\mu$m, did not exhibit noticeable ring-DLSPPW mode excitation, the WRR with the gap $g \cong 0.53\,\mu$m [Fig. 17(c)] featured very efficient ring excitation and well-pronounced wavelength-dependent behavior [Figs. 17(d–g)]. The SEM image of the $g \cong 0.53\,\mu$m WRR reveals that the gap is not completely resolved as some residual PMMA resistance still exists in the gap [Fig. 17(b)], which is likely to cause slightly more coupling to the ring than expected for that gap size. Note that the WRR transmission [Fig. 17(h)] was evaluated using the input and output waveguide cross sections separated by $10\,\mu$m [Fig. 17(c)], a propagation length that alone introduces $\sim$20% of loss as estimated from the DLSPPW propagation length of $\sim 50\,\mu$m.

In general, the WRR transmission is expected to be periodic with respect to the phase accumulated by the ring-DLSPPW mode per circulation and can be described with the following expression:[27]

$$T = \exp(-l/L_{\mathrm{SP}}) \frac{\alpha^2 + t^2 - 2\alpha t \cos\theta}{1 + \alpha^2 t^2 - 2\alpha t \cos\theta}, \tag{12}$$

where $\theta = (2\pi/\lambda) N_{\mathrm{eff}}(\lambda) 2\pi R$. The first factor in Eq. (12) reflects the power loss incurred by the propagation over the distance $l = 10\,\mu$m related to the DLSPPW propagation length $L_{\mathrm{SP}} \sim 50\,\mu$m. α is the field attenuation factor per roundtrip around the ring, t is the mode transmission through the coupling region in the straight waveguide,[27] λ is the light wavelength in air, R is the ring radius, and $N_{\mathrm{eff}}(\lambda)$ is the DLSPPW effective index, whose dispersion has to be taken into account. When fitting the experimental data [Fig. 17(h)], we used the calculated[16] mode dispersion resulting in $N_{\mathrm{eff}}(\lambda) \cong 1.61 - 0.25\lambda(\mu\mathrm{m})$, finding other parameters to be at $\alpha \cong 0.55$, $t \cong 0.7$, and $R \cong 5.43\,\mu$m. Note that the fitted mode loss in the ring α implies that the bend loss amounted to $\sim$0.36 dB/rad was indeed relatively small as expected. Finally, it is seen that the investigated WRR features nearly *complete extinction* (reaching $\sim$13 dB at 1590 nm) with a $\sim$20-nm bandwidth.

Full three-dimensional finite element (3D-FEM) simulations of the WRR transmission using the ring radius R and ring-waveguide gap g as fitting parameters are presented in Fig. 17(h). It is seen that the WRR transmission calculated for $R_0 = 5486$ nm and $g_0 = 250$ nm, being close to the analytic fit, reproduces well the experimental features. More importantly, the simulations predict the observed trend of increasing the contrast with wavelength much better than the fit by virtue of taking consistently into account the DLSPPW dispersion influencing the phase delay, insertion loss, and coupling strength. Note that the positions of the minima are very sensitive to the ring radius, whereas their level is strongly influenced by the gap width (controlling the coupling strength), indicating the tolerance level in the WRR design and fabrication.[26,28]

8. Bragg Gratings

Wavelength selection can also be realized with diffraction gratings, e.g., by using in-line BG-based filters well known in integrated optics and introduced recently for long range SPP waveguides.[29,30] We made use of the dependence of DLSPPW mode index on the dielectric ridge width,[16] and

fabricated 600-nm-period step-in-width (from 500 nm to 2 μm) BGs with 300-nm-long intervals between wide BG sections and containing different number of periods: $N = 10$, 20, 30, and 50. The DLSPPW mode index variation was estimated[16] between ∼1.2 and 1.4 at $\lambda \sim 1.55\,\mu$m, so that the BG transmission minimum was expected at $\lambda \cong 2\Lambda n_{\mathrm{av}} = 1.56\,\mu$m ($\Lambda$ and n_{av} being the grating period and average DLSPPW mode index). The advantage of this design is that the BG fabrication does not require additional processing (as, for example, step-in-depth gratings[30]). However, UV lithography resolution is more critical here than in the above case of the WRRs because of small-sized (∼300 nm) features throughout the whole BG length, resulting in more pronounced proximity effects, seen as a rounding of the gratings, affecting the geometry of fabricated BGs [see inset in Fig. 18(a)]. Still, the 30-period BG exhibited well-defined Bragg reflection and extinction in transmission with the wavelength interval of 1540–1600 nm (Fig. 18). Note that relatively low transmission levels (∼0.3) observed outside of the band gap are partially accounted for by the DLSPPW mode propagation loss reducing alone the transmission to ∼0.7. In comparison, the 20-period BG performed similarly but with considerably weaker Bragg reflection (∼0.15) and extinction (∼0.13), while the 50-period BG featured about the same Bragg reflection but very low transmission (<0.15) at all wavelengths.

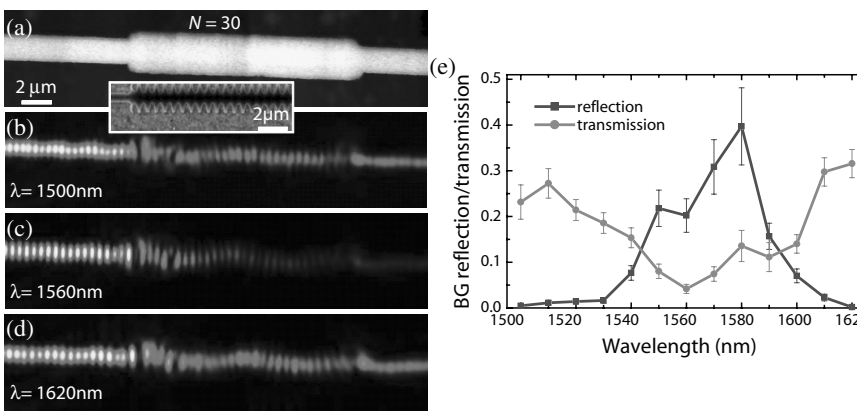


Fig. 18. Plasmonic Bragg grating (BG). (a) Topographical and (b–d) Near-field optical [λ = (b) 1500, (c) 1560 and (d) 1620 nm] SNOM images of an 18-μm-long BG with the period of 600 nm along with an inset showing an SEM image of the BG section. (e) BG reflection and transmission spectra determined experimentally from SNOM images [similar to (b–d)]. [Adapted from Ref. 26].

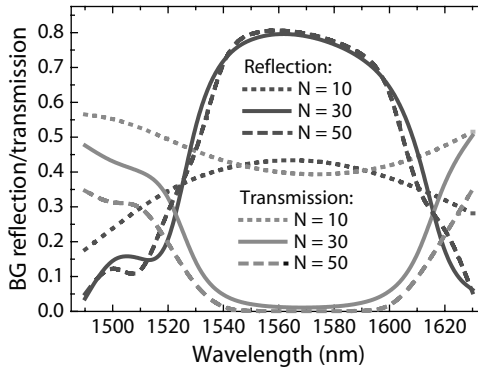


Fig. 19. Reflection and transmission BG spectra calculated with 3D-FEM simulations for the 600-nm-period BGs with different numbers of grating periods: $N = 10$, 30 and 50. [Adapted from Ref. (26)]

The 3D-FEM simulations of BG reflection and transmission (Fig. 19) showed all the main features observed: the well defined band gap (seen in both reflection and transmission) within the range of 1540–1600 nm, the similar reflection for $N = 30$ and $N = 50$ but lower transmission for the latter, and the rather weak band-gap effects for $N = 10$. However, the simulated BG characteristics are much better with respect to the insertion loss (most probably due to weaker out-of-plane scattering), a circumstance that we attribute to the structural imperfections.[26]

9. Discussion

The near-field optical characterization of basic waveguiding and various wavelength selective components demonstrate the achievement of integratable components capable of, e.g., filtering out undesired wavelength intervals and physically separating signals with different wavelengths. These achievements are important steps in the establishment of a DLSPPW technology platform and towards realization of ultra-compact integrated plasmonic circuits operating simultaneously as an electric circuit due to the underlying electrodes with enhanced functionalities for optical signal processing and on-electronic-chip interconnects. Although the transmission losses observed in the DLSPPWs are somewhat larger than in other conventional integrated photonic devices, the DLSPPW technology platform presents some intriguing novel applications due to the diversity in the choice of the dielectric and the presence of the metal electrodes below

the dielectric ridges. The metal involved can be turned into an asset of the technology as the DLSPPW mode features a field maximum right at the metal surface, which makes it very sensitive to, e.g., thermo-optical or electro-optical effects induced by the electrodes. This could lead to electronic control of nanophotonic devices such as wavelength selective components, variable optical attenuators, and integrated power monitors, which is already demonstrated by thermo-optic modulation of a DLSPPW signal.[31] As the DLSPPW technology allows for a choice of the dielectric, one can also envision the usage of the strong electromagnetic field at the metal-dielectric interface by realizing non-linear photonic elements at low light intensities.

Acknowledgments

The authors gratefully acknowledge Thomas Søndergaard, Zhuo Chen, Laurent Markey, Alain Dereux, Alexey V. Krasavin, Pádraig Bolger, and Anatoly V. Zayats for their contributions to this work. The financial support of the EC FP6 STREP PLASMOCOM project (IST 034754) is acknowledged.

References

1. W. L. Barnes, A. Dereux, and T. W. Ebbesen, *Nature* **424**, 824 (2003).
2. D. K. Gramotnev and S. I. Bozhevolnyi, *Nat. Photonics* **4**, 83 (2010).
3. I. V. Novikov and A. A. Maradudin, *Phys. Rev. B* **66**, 035403 (2002).
4. S. I. Bozhevolnyi, V. S. Volkov, E. Devaux, J.-Y. Laluet, and T. W. Ebbesen, *Nature* **440**, 508 (2006).
5. P. Berini, *Phys. Rev. B* **61**, 10484 (2000).
6. T. Nikolajsen, K. Leosson, I. Salakhutdinov, and S. I. Bozhevolnyi, *Appl. Phys. Lett.* **82**, 668 (2003).
7. S. I. Bozhevolnyi, J. Erland, K. Leosson, P. M. W. Skovgaard, and J. M. Hvam, *Phys. Rev. Lett.* **86**, 3008 (2001).
8. S. A. Maier, P. G. Kik, H. A. Atwater, S. Meltzer, E. Harel, B. E. Koel, and A. A. G. Requicha, *Nat. Mat.* **2**, 229 (2003).
9. C. Manolatou, S. G. Johnson, S. Fan, P. R. Villeneuve, H. A. Haus, and J. D. Joannopoulos, *J. Lightwave Technol.* **17**, 1682 (1999).
10. B. Steinberger, A. Hohenau, H. Ditlbacher, A. L. Stepanov, A. Drezet, F. R. Aussenegg, A. Leitner, and J. R. Krenn, *Appl. Phys. Lett.* **88**, 094104 (2006).
11. C. Reinhardt, S. Passinger, B. N. Chichkov, C. Marquart, I. P. Radko, and S. I. Bozhevolnyi, *Opt. Lett.* **31**, 1307 (2006).
12. H. Kogelnik, *Theory of Dielectric Waveguides — In Integrated Optics.* (Springer-Verlag, Berlin, 1979). Edited by T. Tamir.

13. A. B. Buckman, *Guided-Wave Photonics*. (Saunders College Publishing, 1992), 1st edition.
14. A. Boltasseva, T. Nikolajsen, K. Leosson, K. Kjaer, M. S. Larsen, and S. I. Bozhevolnyi, *J. Lightwave Technol.* **23**, 413 (2005).
15. S. I. Bozhevolnyi, *Opt. Express* **14**, 9467 (2006).
16. T. Holmgaard and S. I. Bozhevolnyi, *Phys. Rev. B* **75**, 245405 (2007).
17. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. (Springer-Verlag, Berlin, 1988), 1st edition.
18. T. Holmgaard, S. I. Bozhevolnyi, L. Markey, A. Dereux, A. V. Krasavin, P. Bolger, and A. V. Zayats, *Phys. Rev. B* **78**, 165431 (2008).
19. T. Holmgaard, S. I. Bozhevolnyi, L. Markey, and A. Dereux, *Appl. Phys. Lett.* **92**, 011124 (2008).
20. I. P. Radko, S. I. Bozhevolnyi, and N. Gregersen, *Appl. Opt.* **45**, 4054 (2006).
21. A. Kumar and S. Aditya, *Microwave Opt. Technol. Lett.* **19**, 289 (1998).
22. T. Holmgaard, Z. Chen, S. I. Bozhevolnyi, L. Markey, A. Dereux, A. V. Krasavin, and A. V. Zayats, *Opt. Express* **16**, 13585 (2008).
23. T. Holmgaard, Z. Chen, S. I. Bozhevolnyi, L. Markey, and A. Dereux, *J. Lightwave Technol.* **27**, 5521 (2009).
24. E. D. Palik, *Handbook of Optical Constants of Solids*. (Academic, New York, 1985), 1st edition.
25. Z. Chen, T. Holmgaard, S. I. Bozhevolnyi, A. V. Krasavin, A. V. Zayats, L. Markey, and A. Dereux, *Opt. Lett.* **34**, 310 (2009).
26. T. Holmgaard, Z. Chen, S. I. Bozhevolnyi, L. Markey, A. Dereux, A. V. Krasavin, and A. V. Zayats, *Appl. Phys. Lett.* **94**, 051111 (2009).
27. A. Yariv, *Electron. Lett.* **36**, 321 (2000).
28. T. Holmgaard, Z. Chen, S. I. Bozhevolnyi, L. Markey, and A. Dereux, *Opt. Express* **17**, 2968 (2009).
29. S. Jetté-Charbonneau, R. Charbonneau, N. Lahoud, G. Mattiussi, and P. Berini, *Opt. Express* **13**, 4674 (2005).
30. S. I. Bozhevolnyi, A. Boltasseva, T. Søndergaard, T. Nikolajsen, and K. Leosson, *Opt. Commun.* **250**, 328333 (2005).
31. J. Gosciniak, S. I. Bozhevolnyi, T. B. Andersen, V. S. Volkov, J. Kjelstrup-Hansen, L. Markey, and A. Dereux, *Opt. Express* **18**, 1207 (2010).

MANIPULATING NANOPARTICLES AND ENHANCING SPECTROSCOPY WITH SURFACE PLASMONS

Kenneth B. Crozier*, Kai Wang, Yizhuo Chu, Mohamad Banaee, and Ethan Schonbrun

School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
**kcrozier@seas.harvard.edu*

Field enhancement from surface plasmon structures presents new opportunities for optical manipulation and surface enhanced Raman spectroscopy (SERS). We demonstrate the propulsion of gold nanoparticles using surface plasmon polaritons. SPPs are excited on a thin gold film. The resultant evanescent field draws nanoparticles toward the film, where they are propelled along by the optical scattering force. We describe our work on a novel SERS substrate consisting of a metal nanoparticle array separated from a gold film by a thin silicon dioxide spacer. We show that the double plasmon resonances of these structures enable strong field enhancement at both pump and Stokes frequencies.

1. Introduction

Surface plasmon nanostructures present a means to concentrate light into deep-subwavelength scales. The field enhancement that results can be highly advantageous for a number of applications in nanoscience. In this Chapter, we describe experimental work on two applications of surface plasmon field enhancement. The first application is in optical manipulation. Here, field enhancement offers a way to increase the optical force on micro- and nanoparticles for a given input laser power, compared to conventional

(non-surface plasmon) optical tweezers. The second application is surface enhanced Raman spectroscopy (SERS). In this case, field enhancement enables signals from molecules attached to surface plasmon nanostructures to be greatly enhanced. We introduce these applications in more detail below.

The focal spots of conventional optical tweezers are limited by diffraction to being no smaller than roughly half a wavelength. The gradient force that can be achieved for a given laser power is therefore fundamentally limited. In addition, if one is trapping a nanoparticle, then the size of the region in which the particle is trapped is much larger than the particle. In response to this, recently there has been considerable interest in the use of surface plasmons for optical manipulation.[1–5] This is motivated by their ability to produce strongly enhanced fields confined to subwavelength dimensions. This work is being complemented by related work being undertaken on optical manipulation using waveguides,[6–8] diffractive optics,[9] optical fibers,[10] photonic crystals,[11,12] silica microspheres[13] and micro-mirrors.[14] In this Chapter, we describe the propulsion of gold nanoparticles using surface plasmon polaritons excited on a gold film. Prior to this study,[15] a few related papers had been present in the literature. Garces-Chavez *et al.* excited SPPs on a gold film and observed the self-organization of dielectric particles over a large area through optical and thermal convection forces.[16] Using a photonic force microscope, Volpe *et al.* measured the enhancement of optical forces on dielectric particles by SPP, where the particles were 500 nm above the gold film and therefore did not disturb the SPPs very much.[17] In this work, we present experimental results in which particle velocities are measured as a function of the laser power and incident angle of illumination. The results are compared to the predictions of numerical electromagnetic modeling and found to be in reasonable agreement.

The second application we turn our attention to in this Chapter is that of SERS. SERS refers to the phenomenon whereby molecules absorbed to a nanostructure exhibit Raman scattering that is much stronger than it would otherwise be.[18] Raman spectroscopy is a powerful technique, as it enables molecules to be identified by their characteristic vibrational spectra. It represents a highly sensitive molecular detection method with great potential for chemical sensor applications.[19] Other applications include the monitoring of water pollution[20] and the detection of biohazards.[21] One of its drawbacks, however, is the fact that Raman scattering cross sections are usually very small, by comparison with fluorescence, for

example. SERS is therefore an important technique, enabling molecules to be identified at lower concentrations. The enhancement factor in SERS arises from two mechanisms: electromagnetic enhancement (EM) and chemical enhancement. The electromagnetic mechanism is related to the enhancement in the near field intensity as a result of the excitation of surface plasmons.[22] Chemical enhancement is interpreted as being due to electronic coupling between the adsorbed molecules and the metal surface.[23] In this Chapter, we describe a study[24] undertaken on improving the EM enhancement factor. The SERS EM enhancement factor is proportional to the product of the field intensity (square of electric field, E^2) enhancements at the excitation and Raman scattering frequencies. Giant enhancement of the local electromagnetic fields can be produced through the excitation of localized surface plasmons (LSPs), collective electron oscillations in the nanoparticles.[25] It has been shown that, for a substrate with a single LSP resonance, locating the LSP resonance between laser and Raman scattering frequencies maximizes the SERS enhancement factor.[26] Much research has been carried out using substrates for SERS with single LSP resonances. In this study, we present results on SERS substrates with double resonances. The substrates have plasmon resonances at the laser and Stokes frequencies, enabling simultaneous enhancement of both. In this Chapter, we describe the mechanism by which the substrate exhibits double resonances. Experimental results are presented of the extinction spectra of fabricated devices. SERS spectra are measured from benzenethiol monolayers formed on the substrates, and the enhancement factor (EF) is found.

2. Propulsion of Gold Nanoparticles with Surface Plasmon Polaritons

In this work, we employ the evanescent fields in the water medium above a gold film to exert optical forces on gold nanoparticles. Evanescent fields can be achieved in other ways, however. Illumination of a glass-water interface at an angle greater than the critical angle, for example, results in total internal reflection (TIR) and evanescent fields in the water medium. It is therefore natural to compare this case, illustrated in [Fig. 1(a)], to the SPP case [Fig. 1(b)]. In both cases, a p-polarized plane wave is incident from within the glass. The evanescent field in the water exerts a force on the particle F_{EM} that tends to draw the particle toward the glass, and propels it along the interface. In the SPP case, the angle of incidence α_{SPP} is

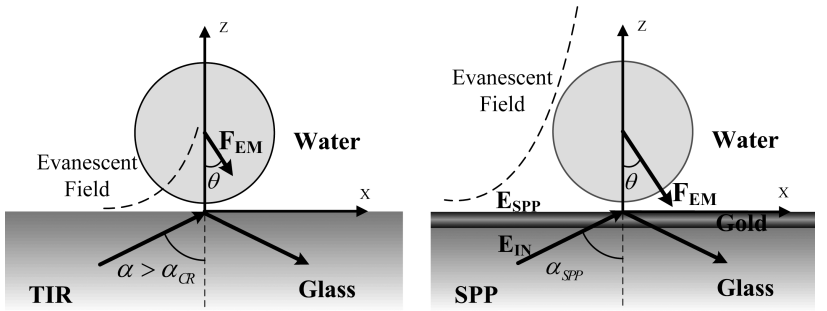


Fig. 1. Schematic of optical manipulation configurations. (a) Total internal reflection (TIR) configuration. Light is incident from glass substrate at an angle α greater than the critical angle α_{CR} for TIR, resulting in an evanescent field in the water medium above the glass. Nanoparticle in water medium experiences optical force F_{EM} in the xz plane, at an angle θ with the normal to the surface. (b) Surface plasmon polariton (SPP) configuration. Light is incident from glass substrate at an angle α_{SPP} appropriate for the excitation of surface plasmon polaritons at the gold-water interface, resulting in evanescent fields in the water and an optical force F_{EM}.

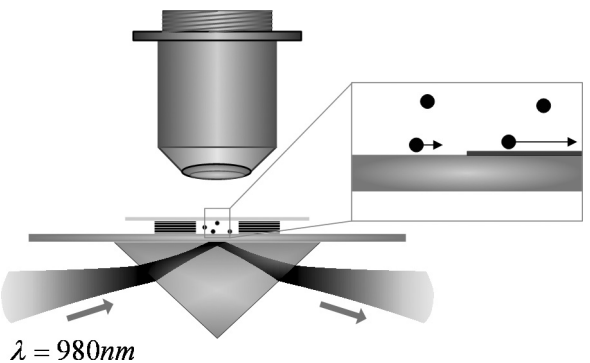


Fig. 2. Experimental setup for nanoparticle manipulation using surface plasmon polaritons (SPPs). Near-infrared laser (980 nm wavelength) is incident on glass prism, on which a glass substrate is positioned, with index matching oil in between them. Microfluidic cell is formed on top of glass substrate. On some parts of the glass substrate, a gold film has been deposited. Nanoparticles within water medium in cell experience optical forces that pull them onto gold film, and propel them along it.

chosen to enable phase matching to the SPP propagating at the gold-water interface.

We realize the TIR and SPP configurations described in Fig. 1 using the experimental system of Fig. 2. Light from a pigtailed laser diode

(wavelength: 980 nm) is expanded and spatially filtered, enabling the beam to be controlled to have a diameter of ∼6 mm. A lens with a focal length of 20 cm is used to loosely focus the beam, which is then reflected by a mirror into a glass prism. The resulting $1/e^2$ intensity focused spot has an elliptical shape of $\sim 40\,\mu\text{m} \times 80\,\mu\text{m}$ on the surface of the sample, as estimated by Gaussian beam theory. A glass substrate is placed on the prism, with oil in between them for index matching. The glass substrate contains regions in which a gold film (55 nm thick) is present. A microfluidic cell, consisting of an SU8 spacer and a glass coverslip, is formed on top of the glass substrate. The microfluidic cell contains gold colloids in water. The SU8 spacer is made only $\sim 5\,\mu$m thick to minimize the effects of thermal convection, a method described by Garces-Chavez *et al.*[16] A layer of silicon dioxide is deposited on the gold film by atomic layer deposition to reduce the adhesion of gold particles to the film. The gold film is deposited by thermal evaporation. The gold particles, obtained from Ted Pella Corporation, are 250 nm diameter spheres with a reported standard deviation of 20 nm. To observe the trapping process, the light scattered by the gold particles is imaged by a microscope objective onto a charge coupled device (CCD) camera.

Before describing the experimental results obtained using the system illustrated in Fig. 2, we turn first to predict the forces exerted on particles in the TIR and SPP configurations. This is done using the Maxwell stress tensor method which is described, for example, by Novotny and Hecht.[27] In this method, the force on the particle is calculated from the electric and magnetic fields on a surface enclosing it. These are found by the finite difference time domain (FDTD) method, using the FullWave software from RSoft. The simulation parameters are chosen to match the experimental conditions. The gold particle has a diameter of 250 nm, and the gold film is 55 nm thick. A gap of 10 nm, equal to the Debye length, is assumed to exist between particle and film. While we do not know the precise separation between the gold nanoparticle and film, for the calculations we assume it to be the Debye length, as this is the characteristic length of the electrostatic interaction. To model the particle, gap, and film accurately, we use a non-uniform simulation grid, ranging from 2 nm near material interfaces to 10 nm further from these boundaries.

Simulation results are shown in Fig. 3. In this figure, the optical forces and their directions are plotted for the TIR and SPP cases. The optical force is normalized to the incident illumination intensity. The angle θ represents the angle the force vector makes with normal to the surface, as shown in

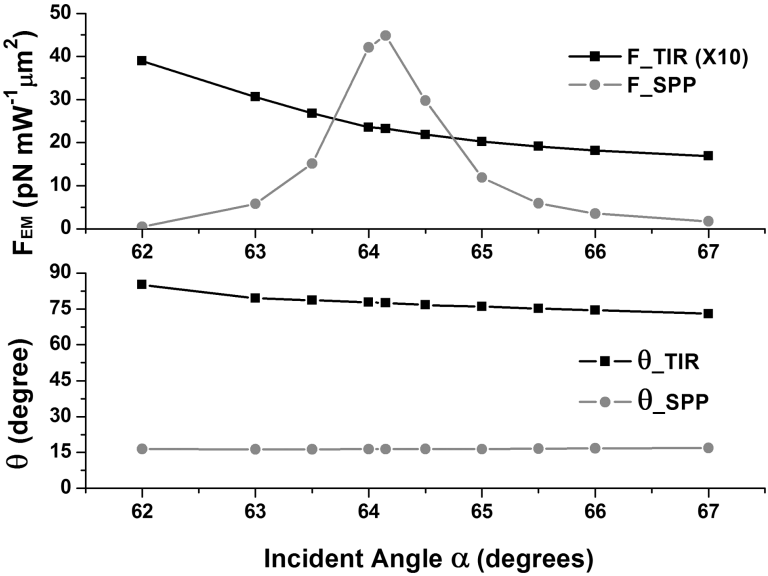


Fig. 3. Optical forces and their directions as a function of incident angles, as found by 3D FDTD simulations. Optical force is normalized to incident intensity.

Fig. 1. The plots of Fig. 3 are shown as a function of the angle of incidence of the illumination. It should be noted that, to facilitate comparison with the SPP case, the figure plots the force in the TIR case multiplied by 10. Excitingly, the simulations of Fig. 3 predict that the optical force is ∼22 times larger for the SPP case than the TIR case, for illumination at the SPP angle. This is larger than what would be predicted just from the enhancement in intensity of the SPP case by comparison to the TIR case, $|E_{\mathrm{SPP}}/E_{\mathrm{TIR}}|^2$, which is ∼10 times. Interestingly, the simulation results also indicate that for the SPP case the optical force is directed much more toward the sample than for the TIR case.

We now discuss the experimental results. SPPs are excited on a gold film. The accompanying evanescent fields draw particles onto the film, where they are then propelled by optical forces. The scattered light images obtained from the CCD camera enable this process to be observed. A typical sequence of scattered light images of particles recorded by the CCD camera at a rate of 30 frames per second (fps) is shown in Fig. 4. The particle indicated by the arrow moves across the image field from right to left in successive frames. The other two particles, circled with dashed lines, are stuck on the surface, and do not move. As we will describe below, particle

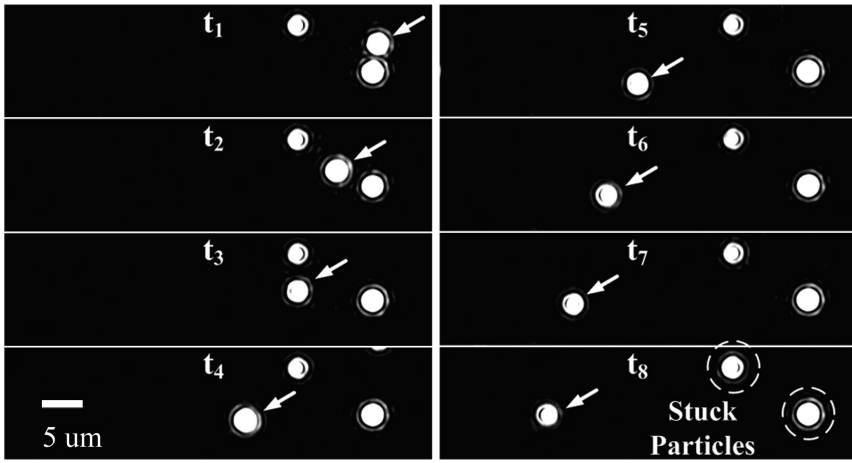


Fig. 4. CCD camera images, showing time sequence of scattered light images of gold nanoparticles. Time interval between images is 1/6 s. The images have dimensions of $70\,\mu\text{m} \times 20\,\mu\text{m}$. The gold nanoparticle denoted by the arrow moves right to left. Two nanoparticles, indicated by dashed circles, are stuck to the sample surface.

velocity can be determined from the data contained in the CCD images. Experiments are also carried out under TIR, using the evanescent fields present on the areas of sample with no gold film present. In that case, Brownian motion is seen to dominate, and propulsion is not observed. We note that the particles do not remain in the evanescent field for sufficient time for a velocity measurement to be taken. This is found to be the case over the range of laser powers and incident angles available to us in the experiment. Presumably, however, it would be possible to propel particles using TIR fields using laser powers higher than those available to us.

In the experimental setup (Fig. 2), the mirror used to direct the laser beam into the prism is mounted on a rotation stage. This enables the angle of incidence to be varied. In order to understand the propulsion mechanism, we vary the angle of incidence and measure the corresponding particle velocities. The outcome of this experiment is shown in Fig. 5. A fixed illumination field intensity of $6 \times 10^6\,\text{W/m}^2$ incident on the gold film is used. It can be seen that the velocity peaks for illumination angles near 64.5°. In order to understand this phenomenon, we measure the reflectance of the film as a function of incident angle. The results are shown in Fig. 5, along with the theoretical calculation of the reflectance. It can be seen that there is a dip in reflectance at about the same angle as the peak in velocity. The enhanced fields that result from the coupling of power to the

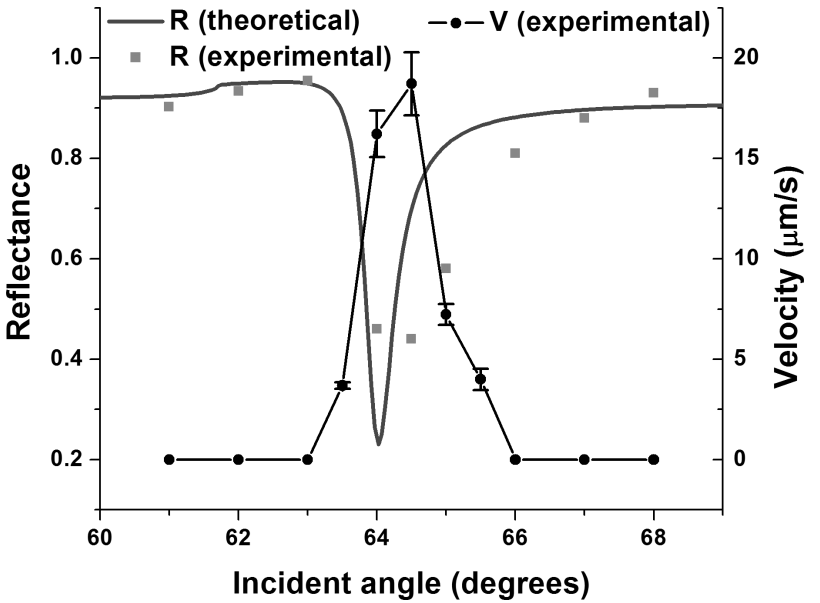


Fig. 5. Experimentally-measured particle velocity (black filled circles) as a function of incident angle, for fixed illumination intensity. Experimentally-measured reflectance is also shown (black filled squares) as a function of angle. Theoretically-calculated reflectance is shown as gray curve.

SPP results in an increased force on the particle, and therefore a larger velocity.

From the measured particle velocity, the force on the particle can be found. This permits a direct comparison between the measured and theoretically-predicted forces in the x-direction, i.e., along the plane of the substrate. To find the force, we multiply the particle velocity v by the drag coefficient μ, i.e., $F_x = v\mu$. When the particle is far from the surface, the drag coefficient is given by $\mu = 6\pi\eta(T)a$, where $\eta(T)$ is the water viscosity and a is the particle radius. This would normally be straightforward to calculate, but for the fact that SPP excitation results in temperature variation. To account for this, we determine the water temperature by analyzing the particle's Brownian motion in the y-direction. No optical force is present along the y-direction because the plane of incidence of the illumination is the xz plane. Analysis of the particle motion in the y-direction allows the diffusion coefficient D to be found (Fig. 6), from which the temperature can be determined. The drag coefficient can then be found, using tabulated values for the viscosity of water as a function

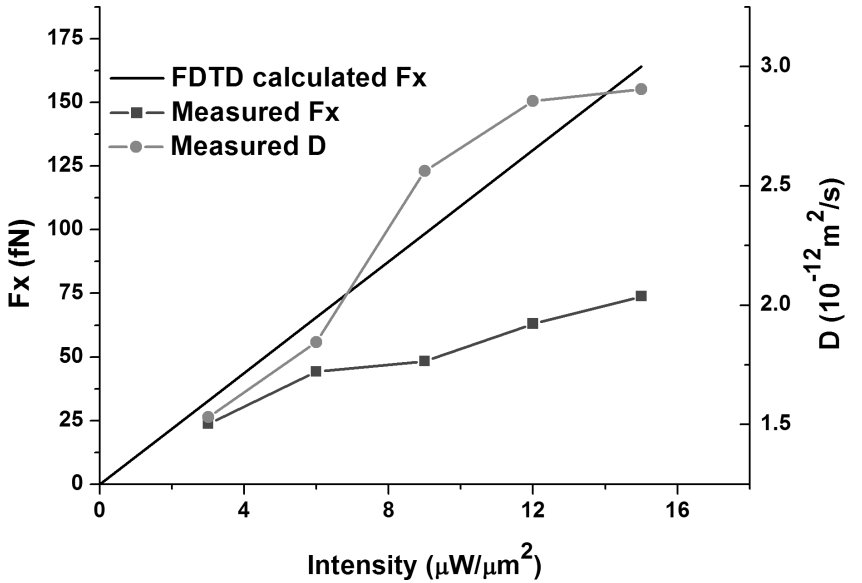


Fig. 6. Experimentally-measured optical force F_x in the plane of the substrate and FDTD calculated optical force F_x as a function of incident intensity. Experimentally-measured diffusion coefficient D is also shown. Incident angle is fixed at 64.5°. Gold particles have diameters of 250 nm.

of temperature.[28] This allows a direct comparison between measured and theoretically-predicted forces, as shown in Fig. 6. It should be noted that there are no fitting parameters in the theoretical calculation of the force. It can be seen that the experimental results for F_x are in reasonable agreement with the theoretical predictions.

Possible explanations for the differences between theory and experiment are discussed below.

First, the non-zero divergence of the input beam decreases the coupling efficiency of the input beam to the SPPs. In the FDTD simulations, the incoming beam is a plane wave. The coupling efficiency γ in this case can be determined from the reflectance R: $\gamma = 1 - R = 0.77$. For the experiments, the actual coupling efficiency γ_{act} can be estimated from the measured reflectance shown in Fig. 5 to be $\gamma_{\text{act}} = 0.56$. This smaller coupling efficiency could account for the smaller measured optical forces in experiments.

Second, an important factor that has to be taken into consideration is the effect of the surface on particles' dynamics when particle is very close to it. According to Faxen's law, the effective drag coefficient experienced

by a particle will have an enhancement factor β ($\mu_{\text{eff}} = \mu\beta, \beta \geq 1$) when it gets closer to a surface. The enhancement factor β depends on the particle's size and the distance between the particle and the surface. It will affect the measured optical force in two ways. Firstly, the temperature T_{est} determined above is likely to be an underestimate of the actual temperature T_{act}. This is due to the fact that, when Faxen's law is taken into account, the Einstein relationship is modified to be $D = k_{\mathbf{B}}T_{\text{act}}/\mu_{\text{eff}} = k_{\mathbf{B}}T_{\text{act}}/6\pi a\eta(T_{\text{act}})\beta$. When Faxen's law is not taken into account, therefore, the optical force F_x^{est} is therefore underestimated as $F_x^{\text{est}} = k_B T_{\text{est}} v/D < F_x^{\text{act}} = k_B T_{\text{act}} v/D$. Secondly, it is entirely possible that the enhancement factor β increases with intensity because the stronger optical force leads to a decrease in the average particle-surface distance. This would result in a larger discrepancy between the measured and predicted optical force at higher input power levels.

The results demonstrate the benefits of field enhancement from surface plasmons for optical manipulation. In the next section, we consider another application of field enhancement: spectroscopy.

3. Double Resonance Substrates for Surface-Enhanced Raman Spectroscopy

Localized surface plasmons (LSPs) on metal nanoparticles are currently being investigated with interest for SERS. The enhanced fields on the surfaces of nanoshells have been shown to be effective for SERS.[29] Arrays of noble metal nanoparticles on glass substrates[30] have been produced by nanosphere lithography for SERS. The EM SERS enhancement factor is the product of the intensity enhancement (E^2) at the laser frequency with the intensity enhancement (E^2) at the Stokes frequency. McFarland *et al.* carried out an extensive study of metal nanoparticles using wavelength-scanned SERS.[26] It was found that locating the plasmon resonance between laser and Stokes frequencies leads to the maximum SERS enhancement. In this section, we present results from our study[24] that considers the possibility of SERS substrates with double resonances, to enable strong intensity enhancement to be simultaneously achieved at laser and Stokes frequencies. We begin by discussing the method by which our substrate exhibits double resonances. We then describe the fabrication method, and present results on the measurement of the plasmon resonances via white light reflection spectroscopy. SERS results are then described, including determination of the enhancement factor.

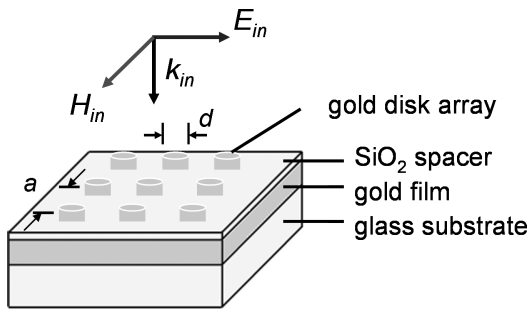


Fig. 7. Double resonance SERS substrate: schematic illustration. Substrate consists of a two dimensional array of gold disks above a gold film, separated from it by a silicon dioxide spacer layer. Substrate is illuminated at normal incidence by a plane wave, whose electric field (E_{in}), magnetic field (H_{in}) and propagation vectors (k_{in}) are shown.

The double resonance substrate we investigate here is shown as Fig. 7. It consists of an array of metal disks fabricated above a metal film, separated from it by a thin (e.g., few tens of nanometers) silicon dioxide spacer layer. It should be noted that only a small portion of the device is illustrated in Fig. 7; the overall extent of the actual device is such that there are more than 100 periods along each of the axes. The metal of the device illustrated in Fig. 7 is gold, although later in this section we present results using silver. Illumination of the substrate by plane waves at normal incidence results in the excitation of two major plasmon resonances. These consist of LSPs excited on the gold disks, and SPPs excited at the interface between the gold film and the silicon dioxide spacer layer. It would not be possible to excite SPPs with plane waves at normal incidence on a flat gold film (i.e., without the disk array) as, by definition, the SPP is bound to the surface and has an in-plane momentum vector larger than the wavevector of the illumination. On the other hand, the disk array provides a reciprocal lattice vector that permits coupling to the SPP. As described further below, when the geometric parameters of the disks, spacer layer and film are chosen appropriately, the LSPs and SPPs hybridize. The double resonances are hybrid states exhibiting characteristics of both LSPs and SPPs.

We have previously carried out a study,[31] employing experiments and numerical electromagnetic simulations, on the interaction between LSPs and SPPs in the structure illustrated in Fig. 7. To understand the effects of coupling between LSPs and SPPs, a series of devices was fabricated, in which the geometric parameters were varied in a systematic fashion to shift the resonance frequency of either the LSP, or the SPP. For example,

by varying the disk diameter, while keeping the disk period constant, the LSP resonance frequency was shifted from being smaller, to being larger, than that of the SPP. An anti-crossing behavior was observed, indicating strong coupling between LSPs and SPPs. Excitingly, this study also predicted, using numerical electromagnetic simulations, considerable field enhancement, by comparison to gold disks on a glass substrate (with no underlying gold film). It was predicted that the intensity enhancement (square of electric field, E^2) on the surface of a single gold disk separated from a gold film by a spacer layer is one order of magnitude higher than that of a gold disk with the same dimensions on a glass substrate. The simulations also predicted that the intensity enhancement of each disk of an array above a gold film is twice as large as that of a single disk above a gold film. This is due to the fact that, when the array periodicity is appropriately chosen, the strong coupling between the LSP and the SPP enhances the near field intensity. These predictions motivated us to carry out SERS using these types of substrates.

The starting substrate is a glass slide coated with a layer of indium tin oxide (ITO). A gold film is thermally evaporated to a thickness of 100 nm, with a 5 nm thick chromium layer used for adhesion. A silicon dioxide layer is then deposited by plasma-enhanced chemical vapor deposition (PECVD). For device shown in Fig. 8, this layer is 23 nm thick. Electron beam lithography, thermal evaporation of gold and lift off are then performed,

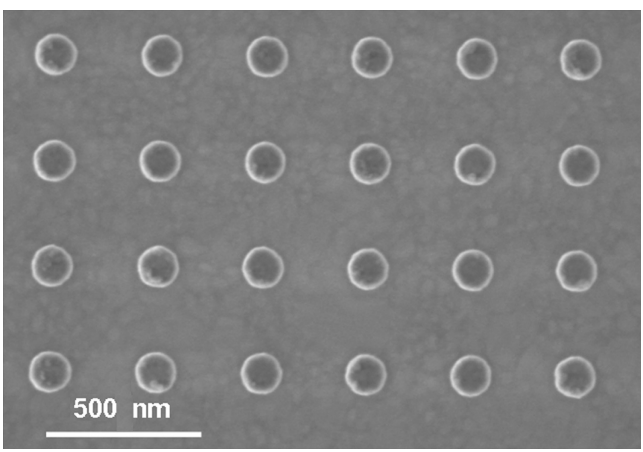


Fig. 8. Scanning electron micrograph (SEM) of fabricated double resonance SERS substrate, consisting of a gold disk array on a silicon dioxide spacer layer, and a gold film. Overall extent of array is 100 μm × 100 μm; SEM only shows a small part of it.

yielding the device shown in Fig. 8. The overall size of the gold disk array is $100 \times 100\,\mu$m. The array periods vary from 350 to 780 nm. The gold disks are 40 nm thick, with diameters ranging from 120 to 140 nm. The silver device, described later in this Chapter, consists of a 100 nm thick silver film, an SiO_2 spacer, and a square array of silver disks. The SiO_2 spacer is 27 nm thick, and the silver disks are 40 nm thick, with diameters ranging from 140 to 150 nm. The array period is 780 nm.

To determine the plasmon resonances of the fabricated devices, reflection spectroscopy is performed. The reflectance of the array is measured using a microscope. The plasmon resonances in this device are k-vector dependent, and the angle of illumination is therefore an important parameter. Collimated illumination is not possible with the microscope employed in this study. However, the use of a low numerical aperture objective (NA 0.1) minimizes the spread of the illumination angles. From the measured reflectance R, the extinction cross section associated with each disk is found using: $C_{\mathrm{REF}} = (1 - R)a^2$, where a is the array period.

Experimental results are shown in Fig. 9 for three fabricated devices. It should be noted that benzenethiol monolayers are formed on these devices prior to the reflection measurements being made. The first device has a

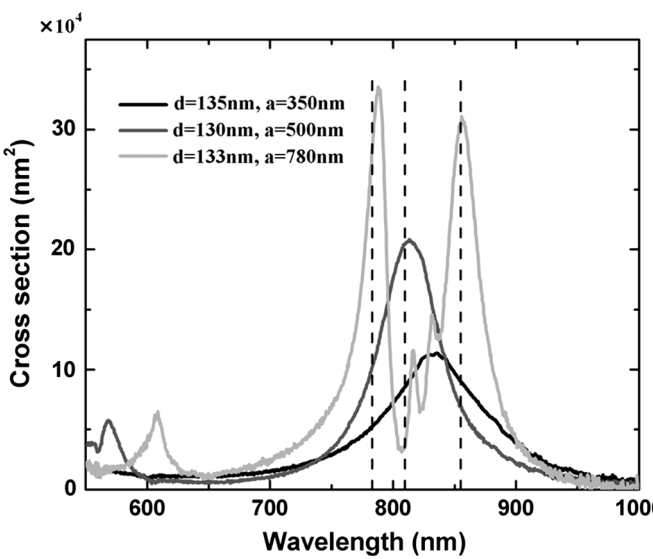


Fig. 9. Experimentally-measured extinction cross sections of three fabricated SERS substrates, with different disk diameters (d) and periods (a). Extinction cross sections are found from reflection spectra.

period a of 350 nm, and a disk diameter d of 135 nm. A resonance can be seen at a wavelength of 833 nm, and corresponds to LSPs excited on the disks. The second device has a period a of 500 nm, and a disk diameter d of 130 nm. A resonance, corresponding to the $(p, q) = (1, 0)$ SPP mode, can be seen at $\lambda = 568$ nm. Here, the (p, q) refers to the momentum vector $G = (2\pi/a)\sqrt{(p^2 + q^2)}$ provided by the lattice that allows coupling to the SPP. An additional resonance, corresponding to the LSP, can be seen at $\lambda =$ 813 nm. The third device has a period a of 780 nm, and a disk diameter d of 133 nm. In this device, the LSP and $(1, 0)$ SPP would normally occur about the same wavelength, were they uncoupled. The strong coupling between the LSP and SPP, however, leads to a repulsion of the resonances, and the spectrum shown in Fig. 9. In Fig. 9, the positions of the laser used in the SERS measurement ($\lambda = 783$ nm) and 420 cm^{-1} and 1074 cm^{-1} Raman lines are shown as vertical dashed lines. It can be seen that, for the third device, the plasmon resonances match the laser wavelength and 1074 cm^{-1} Raman line very well.

In order to investigate the near field properties of the structures, finite-difference time-domain (FDTD) simulations are carried out to calculate the near field spectra and the near field intensity distribution. The method of Ref. (31) is employed. Disk arrays with periods of 500 nm and 780 nm are simulated. For both periods, the gold disks are 133 nm in diameter and 40 nm thick. The SiO_2 spacer is 25 nm thick, and has a refractive index of 1.44. In Fig. 10(a), near-field intensity spectra are shown, for a monitor point at the edge of the gold disk, at the interface between the gold disk and SiO_2 spacer. The monitor point is on the x-axis, the polarization direction of the incident illumination. Consistent with the experimental results, the structure with period of 500 nm has an LSP resonance at $\lambda = 827$ nm and a (1,0) SPP resonance at $\lambda = 586$ nm. The structure with period of 780 nm has double resonances at $\lambda = 794$ nm and $\lambda = 860$ nm which are the hybridized resonances. We calculate the scattered field intensity ($|E_s^2|$) distribution on the cross sections perpendicular to the substrate for excitation at 586 nm and 827 nm for the 500 nm period array, and at 794 nm and 860 nm for the 780 nm period array. These are shown as Fig. 10(c–f). The position of the cross section is given in Fig. 10(b). It can be seen that, for the 500 nm period array at $\lambda = 827$ nm, the field is mainly confined around the gold nanoparticle, which is the character of the LSP. In contrast, for the 500 nm period array at $\lambda = 586$ nm, the evanescent field above the gold film indicates the excitation of the SPP mode on the interface of the gold film and the SiO_2 spacer. For the 780 nm period array, at both

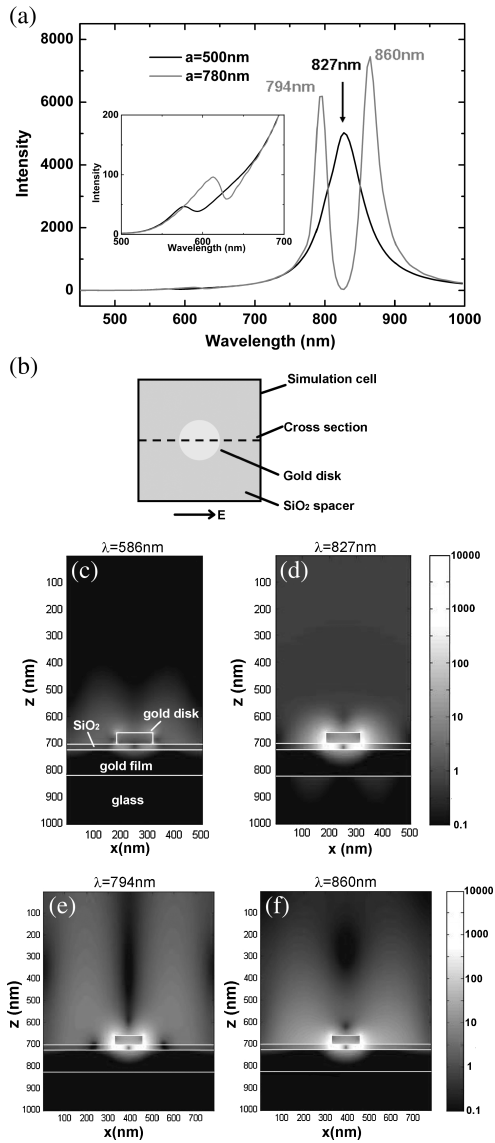


Fig. 10. (a) Simulated near field intensity spectra. Black: weak coupling case with array period of 500 nm. Gray: strong coupling case with array period of 780 nm. Inset: expanded view of near field spectra over the range $\lambda = 500$ nm to 700 nm. (b) Top view of the structure in the simulation. (c–f) Intensity distributions of scattered field on the cross section perpendicular to substrate for illumination at (c) $\lambda = 586$ nm and at (d) $\lambda = 827$ nm for the 500 nm period array, at (e) $\lambda = 794$ nm and at (f) $\lambda = 860$ nm for the 780 nm period array.

$\lambda = 794\,\text{nm}$ and 860 nm, enhanced near fields in the vicinity of the gold nanoparticle can be seen, but with the addition of an extended evanescent field above the gold film, which is the character of the SPP. These results indicate that the two resonances for strong coupling case carry both the characters of LSPs and SPPs.

The SERS measurements are carried out on a Renishaw confocal Raman microscope with excitation at $\lambda = 783\,\text{nm}$. The microscope is equipped with an objective lens with a numerical aperture of 0.12, and the signal is recorded by a thermoelectrically cooled CCD array. Measurements taken on the three devices of Fig. 9 are shown as Fig. 11. From this figure, it is clear that the third substrate ($d = 133\,\text{nm}$, $a = 780\,\text{nm}$) shows particularly strong SERS enhancement for the 1074 cm^{-1} Raman line. In fact, despite the fact that this device has a longer period, meaning that the unit cell area is larger and there are fewer molecules within the laser spot, the total SERS signal is much larger than that of the other two devices. This is because the plasmon resonances of this structure match well to the laser wavelength and the $1074\,\text{cm}^{-1}$ Raman line, resulting in a SERS enhancement factor (EF) that is 7–11 times larger for this structure than the other two. The SERS EF is discussed in further detail below.

As noted by LeRu *et al.*,[32] a variety of definitions have been adopted for the enhancement factor in SERS. Here, the definition we use is the factor

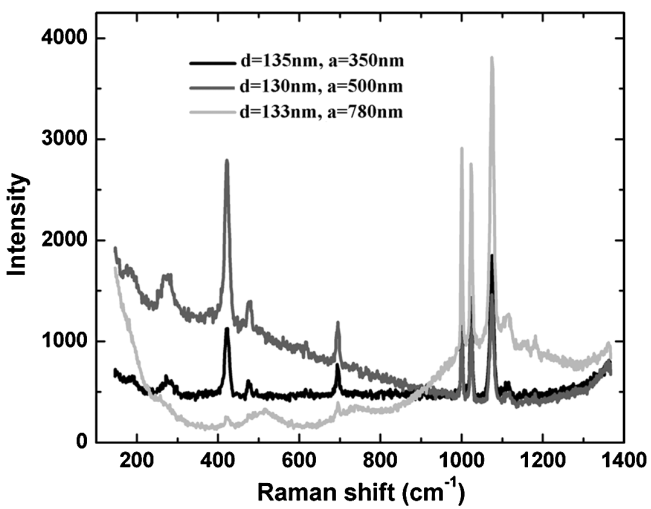


Fig. 11. SERS spectra obtained from three substrates on which benzenthiol monolayers are formed. Substrates have different disk diameters (d) and periods (a).

by which the Raman scattering from each molecule on the nanostructure is enhanced, compared to the Raman scattering from each molecule in a bulk solution. Specifically, the SERS enhancement factor is determined from $\mathrm{EF} = (I_{\mathrm{SERS}}/N_{\mathrm{SERS}})/(I_{\mathrm{RF}}/N_{\mathrm{RF}})$. I_{SERS} and I_{RF} are the intensities of a specific Raman line for the SERS substrate and reference sample, respectively. N_{SERS} and N_{RF} are the numbers of probed molecules in the laser spot for the SERS substrate and reference sample, respectively. The intensity is calculated from the integral of the Raman line of the measured spectrum, from which the background has been subtracted, and normalized by the laser power and CCD integration time. For N_{SERS}, we assume the benzenethiol surface density is $6.8 \times 10^{14}\,\mathrm{cm}^{-2}$ [Ref. (26)]. The surface area taken for each gold disk accounts for the top surface as well as the side walls. N_{RF} is calculated using the molecular weight and volume density of benzenthiol, together with the detection volume.

From the data of Fig. 11, the SERS EFs of the $1074\,\mathrm{cm}^{-1}$ Raman line are found to be 6.5×10^6 for the 350 nm period array, 1.0×10^7 for the 500 nm period array and 7.2×10^7 for the 780 nm period array. The 780 nm period array exhibits a distinctly strong signal around $1100\,\mathrm{cm}^{-1}$, corresponding to the position of its longer wavelength resonance. It can be seen that the $421\,\mathrm{cm}^{-1}$ Raman line is weaker for the 780 nm period array than the other devices. Nonetheless, the EF of 1.2×10^6 for the 780 nm period array may be sufficient for many applications, although it is smaller than the EFs of the 350 nm and 500 nm period arrays (2.3×10^6 and 1.1×10^7, respectively).

For comparison, we fabricate a double resonance substrate consisting of silver disks above a silver film. In Fig. 12, the measured extinction cross section and SERS spectrum of benzenethiol are shown. The excitation laser is at $\lambda = 782$ nm, which matches with the high energy resonance. The strong enhanced SERS signal around 850 nm corresponds to the low energy resonance. The EFs of the 421 and $1074\,\mathrm{cm}^{-1}$ Raman lines are found to be 7.7×10^7 and 8.4×10^8, respectively.

For comparison, we fabricate single resonance SERS substrates, consisting of arrays of gold disks on glass. The arrays are square lattices, with periods of 300 nm or 450 nm. Each disk is 40 nm thick. The disk diameters range from 170 nm to 200 nm. SERS measurements are made on the fabricated devices using benzenethiol, and the largest SERS EF for the $1074\,\mathrm{cm}^{-1}$ line is found to be 3.9×10^5, achieved for a device with a period of 450 nm and disk diameter of 180 nm. The double resonance substrate we demonstrate, therefore, has a SERS EF more than two orders of magnitude higher than this comparison device.

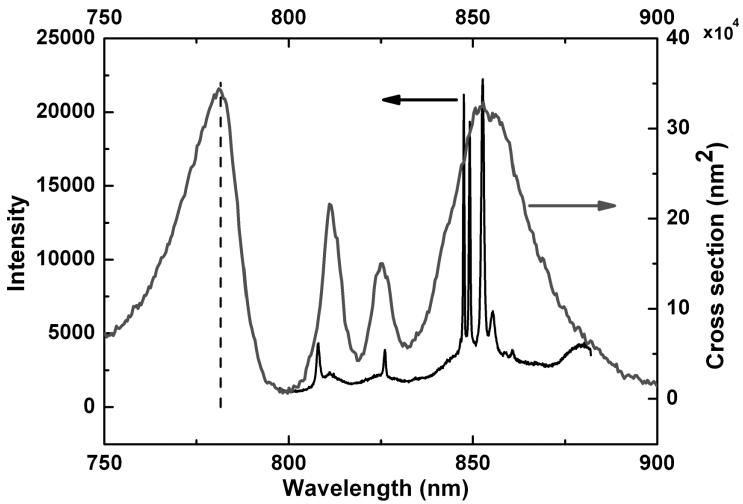


Fig. 12. Measured SERS spectrum (left axis) and extinction cross section (right axis) for device consisting of array of silver disks above silver film.

To conclude, we have shown that a SERS substrate consisting of an array of metal disks above a metal film can be designed to achieve strong enhancement at laser and Stokes frequencies simultaneously. The SERS EF of greater than 10^8 approaches the requirement for single molecule SERS,[32] and is achieved using simple disks. The use of dimer structures — consisting of pairs of nanoparticles separated by small gaps — could allow the SERS EF to be increased still further.

4. Conclusions and Outlook

Optical tweezers and optical spectroscopy are two common tools used for the manipulation and analysis of materials. They are fundamentally difficult to apply on the nanoscale, however, due to the mismatch between the wavelength of light, and the dimensions of nanostructures. Surface plasmons offer a means for overcoming this limitation as they enable light to be concentrated into dimensions smaller than the wavelength. In this Chapter, we have demonstrated that surface plasmons enable nanoparticles to be propelled using optical forces. We have shown that SERS substrates with resonances at laser and Stokes frequencies enable larger signal enhancements to be achieved than comparable single resonance substrates. We anticipate that the concepts studied in this work will assist advance in the fields of nanomanipulation and ultra-sensitive spectroscopy.

Acknowledgments

This work was supported by the Defense Advanced Research Projects Agency (N/MEMS S&T Fundamentals, grant number HR0011-06-1-0044, and SERS S&T Fundamentals, grant number FA9550-08-1-0285), and by the National Science Foundation (grant number ECCS-0747560). Fabrication work was carried out at the Harvard Center for Nanoscale Systems, which is funded by the National Science Foundation.

References

1. L. Novotny, R. X. Bian, and X. S. Xie, *Phys. Rev. Lett.* **79**, 645 (1997).
2. R. Quidant, D. Petrov, and G. Badenes, *Opt. Lett.* **30**, 1009 (2005).
3. M. Righini, G. Volpe, C. Girard, D. Petrov, and R. Quidant, *Phys. Rev. Lett.* **100**, 186804 (2008).
4. A. N. Grigorenko, N. W. Roberts, M. R. Dickinson, and Y. Zhang, *Nat. Photonics* **2**, 365 (2008).
5. M. Righini, P. Ghenuche, S. Cherukulappurath, V. Myroshnychenko, F. J. Garcia de Abajo, and R. Quidant, *Nano Lett.* **9**, 3387 (2009).
6. S. Kawata, and T. Tani, *Opt. Lett.* **21**, 1768 (1996).
7. S. Gaugiran, S. Geti, J. M. Fedeli, G. Colas, A. Fuchs, F. Chatelain, and J. Derouard, *Opt. Express* **13**, 6956 (2005).
8. A. H. J. Yang, S. D. Moore, B. S. Schmidt, M. Klug, M. Lipson, and D. Erickson, *Nature* **457**, 71 (2009).
9. E. Schonbrun, C. Rinzler, and K. B. Crozier, *Appl. Phys. Lett.* **92**, 071112 (2008).
10. S. D. Collins, R. J. Baskin, and D. G. Howitt, *Appl. Opt.* **38**, 6068 (1999).
11. S. Lin, J. Hu, L. Kimerling, and K. B. Crozier, *Opt. Lett.* **34**, 3451 (2009).
12. M. Mandal, X. Serey, and D. Erickson, *Nano Lett.* **10**, 99 (2010).
13. S. Arnold, D. Keng, S. I. Shopova, S. Holler, W. Zurawsky, and F. Vollmer, *Opt. Express* **17**, 6230 (2009).
14. F. Merenda, J. Rohner, J.-M. Fournier, and R.-P. Salathe, *Opt. Express* **15**, 6075 (2007).
15. K. Wang, E. Schonbrun, and K. B. Crozier, *Nano Lett.* **9**, 2623 (2009).
16. V. Garces-Chavez, R. Quidant, P. J. Reece, G. Badenes, L. Torner, and K. Dholakia, *Phys. Rev. B* **73**, 085417 (2006).
17. G. Volpe, R. Quidant, G. Badenes, and D. Petrov, *Phys. Rev. Lett.* **96**, 238101 (2006).
18. K. Kneipp, M. Moskovits, and H. Kneipp, *Surface-Enhanced Raman Scattering. Physics and Applications* (Springer, New York, 2007).
19. G. M. Murray and G. E. Southard, *IEEE Instrum. Meas. Mag.* **5**, 12 (2002).
20. M. Mulvihill, A. Tao, K. Benjauthrit, J. Arnold, and P. Yang, *Angew. Chem. Int. Ed.* **47**, 6456 (2008).
21. X. Zhang, M. A. Young, O. Lyandres, and R. P. Van Duyne, *J. Am.Chem. Soc.* **127**, 4484 (2005).

22. M. Kerker, D. Wang, and H. Chew, *Appl. Opt.* **19**, 4159 (1980).
23. A. Otto, *J. Raman Spectrosc.* **36**, 497 (2005).
24. Yizhuo Chu, Mohamad G. Banaee, and Kenneth B. Crozier, *ACS Nano* **4**, 2804 (2010).
25. S. A. Maier. *Plasmonics: Fundamentals and Applications* (Springer, New York, 2007).
26. A. D. McFarland, M. A. Young, J. A. Dieringer, and R. P. VanDuyne, *J. Phys. Chem. B* **109**, 11279 (2005).
27. L. Novotny and B. Hecht. *Principles of Nano-Optics* (Cambridge University Press, Cambridge UK, 2007).
28. J. R. Watson, R. S. Basu, and J. V. J. Sengers, *Phys. Chem. Ref. Data* **9**, 1255 (1980).
29. C. E. Talley, J. B. Jackson, C. Oubre, N. K. Grady, C. Hollars, S. M. Lane, T. R. Huser, P. Nordlander, and N. J. Halas, *Nano Letters* **5**, 1569 (2005).
30. C. L. Haynes and R. P. Van Duyne, *J. Phys. Chem. B* **105**, 5599 (2001).
31. Y. Chu and K. B. Crozier, *Optics Letters* **34**, 244 (2009).
32. E. C. LeRu, E. Blackie, M. Meyer, and P. G. Etchegoin, *J. Phys. Chem. C.* **111**, 13794 (2007).

ANALYSIS OF LIGHT SCATTERING BY NANOOBJECTS ON A PLANE SURFACE VIA DISCRETE SOURCES METHOD

Elena Eremina*, Yuri Eremin† and Thomas Wriedt‡

*University of Bremen,
Badgasteiner Str. 3, 28359 Bremen, Germany
eremina@iwt.uni-bremen.de

†Faculty of Applied Math. & Computer Sci.,
Moscow State University, Lenin's Hills,
119991 Moscow, Russia
eremin@cs.msu.su

‡Institut für Werkstofftechnik,
Badgasteiner Str. 3, 28359 Bremen, Germany
thw@iwt.uni-bremen.de

In the last years light scattering by nanostructures is of interest in different areas of science and technology. Analysis of light scattered by nanostructures is an effective tool for a better understanding of their properties. In this work the Discrete Sources Method (DSM) is applied to model light scattering by nanoparticles on a surface. One of attractive features of the DSM is an ability to account for all the features of the modeled system, such as complex refractive index with frequency dispersion of particles and a substrate, scattering interaction of particle and an interface. To demonstrate the variety of possible applications for the DSM, we concentrated on two practical applications. First is light scattering by a nanorod on a surface, which requires the use of a general 3D version of the DSM. The second case discussed in this chapter is light scattering by a nanoshell, which allows the accounting for the axial symmetry of the problem and essential reduction of calculation time. In both cases light scattering characteristics and their dependence on

nanostructure characteristics like size, symmetry, incident angle, particle orientation, refractive index and wavelength are analyzed and discussed.

1. Introduction

Nanoparticles and nanostructures are widely seen as having great potential in benefiting many areas of science and technology. Nanotechnology can be found in areas as diverse as pharmaceutical, water decontamination, information and communication technologies, production of computers and materials with new advanced properties. With the development and further miniaturization of nanostructures, there is a clear need to take a closer look into their special features.

There are diverse practical applications e.g., optical antennas,[1] photonic crystals,[2] biosensors[3–5] and many others,[6–15] which are based on light scattering analysis of nanoparticles on or near a substrate.[16,17] The examination of light scattered by nanoparticles is a suitable way to investigate their properties, such as size, shape, composition and imperfections.[18–25] Compared to experimental techniques, mathematical modeling and computer simulation are inexpensive and flexible tools for the investigation, analysis and prediction of the scattering behavior of nanoparticles. Once realized, computer models can be easily adjusted to different features of a designed structure or special schemes of an experimental setup. Additionally, modeling of the nanoparticles properties and analyzing their light scattering behavior is widely used for a correct interpretation of measured data.[26,27]

Depending on application different excitation can be used: polarized or unpolarized light, evanescent wave, plane wave or Gaussian laser beam. Evanescent light scattering is a base for very sensitive sensor applications, as it allows detection of scattering response from tiny objects, where observation by other means is impossible or highly complicated. In experimental setups using evanescent wave scattering, a laser beam propagates in the glass-prism and hits the prism surface with an angle larger than the angle of total internal reflection. As a result, an evanescent field appears close to the surface and decays in the lower refractive index medium. Particles, deposited above the prism (often dispersed in water) start to convert evanescent wave into the scattered field in the vicinity of the surface, and the intensity can be registered by special detectors. The scattering characteristics contain important information regarding particle and media features as well as particle distance from the surface.[28,29]

For successful mathematical modeling, which would take into account all the features of the scattering problem, the choice of a suitable simulation method is very important. The advantage of pure numerical methods, such as the Finite Difference Time Domain (FDTD) or the Finite Elements Method (FEM) is the simplicity of implementation. They are applied directly to Maxwell equations. There was a strong trend towards FDTD solvers in the last decades. FDTD is a simple technique, because it does not require profound knowledge of Maxwell theory. It is based on simple mathematical operations, which can be handled even by very simple computers. Time domain formulations have big advantages when nonlinear materials are present, but they are not really well suited for dispersive materials with strong nonlinearities in their frequency response. These models are not accurate enough in many cases.[30] Additionally, a conventional FDTD scheme does not account for infinite plane interfaces or uses special trick to approximate it.[31] The simplest way to get rid of problems connected with materials dispersion or strong skin effects near the metal films is to work in the frequency domain. This approach leads to direct methods, such as the FEM. The FEM implementation leads to matrix equations with large sparse matrices. The approach allows to reach a very high accuracy, which is valuable when one is exploring nanostructures which have not been fabricated yet. But direct application of the FEM to infinite structures with plasmonic features can cause problems related to a truncation of the simulation domain where surface wave exists.[32] Reviews on this topic can be found in.[33,34]

In this work we present mathematical modeling of light scattered by nanoparticles on the surface based on the Discrete Sources Method (DSM).[35] The DSM has been developed in the early sixties and has recently been applied to multiple problems in nanooptics. The results obtained by the DSM have been repeatedly validated by comparison with other numerical methods and experimental results.[27,36,37] The numerical scheme of the method allows us to account for all the features of the modeled system, such as complex refractive index with frequency dispersion of particles and a substrate, scattering interaction of particle and an interface.

In this chapter we present two practical applications which motivate the development of an exact mathematical model to create an effective computer model for light scattering simulation based on the DSM. The first application is a light scattering by a nanorod.[7,8,12,13,19–21,38] Due to the special shape of the nanorod and its position on the surface, from the mathematical point of view we have to consider a 3D problem without

any symmetry. The theory of the DSM for the general case is discussed in the first part of this work. Another application presented in the paper is light scattering by a nanoshell.[4,5,10,11,23,24,26,29,39–42] In this case, due to the spherical shape of the particle, there is a good chance to use the axial symmetry of the system (particle and the interface). This reduces the dimension of the problem in the use of the DSM and essentially reduces the calculation time. Detailed theory for the axially symmetrical case is described in the second part of this work. The detailed descriptions of the problems itself as well as the mathematical statement and details of the numerical schemes of the DSM based models are given in corresponding chapters. For better readability and understanding we kept some repetitions in the description of the theory.

2. Light Scattering by a Nanorod

2.1. *Introduction*

In this section we will discuss light scattering by a nanorod. In general, optical scattering effects, significant for particles larger than a few tens of nanometers, give more efficient detection schemes for monitoring of the local environments of single particle plasmon resonance.[38] For sensor and other applications, it is very important to understand how optical absorption and scattering depend on particle geometry, material, and excitation in controlling their relative contribution for selected applications.[7,12,21] From the viewpoint of optical scattering, noble metal nanorods have attracted additional attention due to their higher tunability of resonance frequency and multiple resonance peaks in comparison to spherical particles.[19] Even a small elongation of a spherical nanoparticle demonstrates a drastic increase in scattering efficiency.[20] Another feature of a nanorod is the strong dependence of the plasmon resonance intensity on the particle orientation with respect to an exciting field.[8,21]

In this part of the contribution we present the DSM based model of a nanorod deposited on the surface. To model a nanorod the shape of elongated spheroid was taken.

2.2. *Asymmetrical DSM model for nanorod*

Let us introduce a Cartesian coordinate system $Oxyz$ by choosing its origin O at the prism-water interface Σ and Oz is directed into domain filled with water D_0 and the plane $z = 0$ corresponds to the Σ plane (Fig. 1).

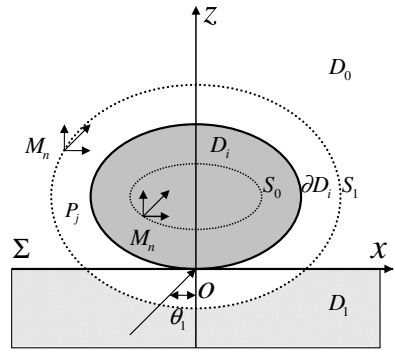


Fig. 1. Model geometry.

Consider an axially symmetrical penetrable particle with an interior domain D_i and smooth boundary ∂D_i deposited on the plane prism surface so that the axis of symmetry is parallel to the interface and belongs to the Ozx plane. We denote the prism domain by D_1. We assume that the exciting field $\{\mathbf{E}^0, \mathbf{H}^0\}$ is a plane wave propagating from D_1 at the angle θ_1 with respect to the z-axis. Then the mathematical statement of the scattering problem can be formulated as follows:

$$\text{rot } \mathbf{H}_\zeta = jk\varepsilon_\zeta \mathbf{E}_\zeta; \quad \text{rot } \mathbf{E}_\zeta = -jk\mu_\zeta \mathbf{H}_\zeta \quad BD_\zeta, \quad \zeta = 0, 1, i, \tag{1a}$$

$$\begin{aligned} \mathbf{n}_p \times (\mathbf{E}_i(p) - \mathbf{E}_0(p)) &= 0, \\ \mathbf{n}_p \times (\mathbf{H}_i(p) - \mathbf{H}_0(p)) &= 0, \end{aligned} \quad p \in \partial D_i; \tag{1b}$$

$$\begin{aligned} \mathbf{e}_z \times (\mathbf{E}_0(p) - \mathbf{E}_1(p)) &= 0, \\ \mathbf{e}_z \times (\mathbf{H}_0(p) - \mathbf{H}_1(p)) &= 0, \end{aligned} \quad p \in \Sigma; \tag{1c}$$

and radiation conditions for the scattered fields at infinity.

Here, $\mathbf{n}_p$ is the outward unit normal vector to the surface ∂D_i, $k = \omega/c$ and $\{\mathbf{E}_\zeta, \mathbf{H}_\zeta\}$ stands for the total field in the corresponding domain $D_\zeta, \zeta = 0, 1, i$. Note that the total field in D_0 is a superposition of the exciting field and the scattered one. If $\text{Im } \varepsilon_\zeta, \mu_\zeta \leq 0$ (the time dependence for the fields is chosen as $\exp\{j\omega t\}$) and the particle surface is smooth enough: $\partial D_i \subset C^{(2,\alpha)}$, then the above boundary-value scattering problem is uniquely solvable.[43]

We construct an approximate solution to the scattering problem [Eq. (1)] based on the Discrete Sources Method.[35] First the plane wave $\{\mathbf{E}^0, \mathbf{H}^0\}$ diffraction problem on the plane interface Σ is solved. This gives the external excitation fields $\{\mathbf{E}^0_\zeta, \mathbf{H}^0_\zeta\}, \zeta = 0, 1$ in domains $D_{0,1}$, which

satisfy the transmission conditions [Eq. (1c)] at the plane interface. Now we can start to construct the approximate solution of the boundary value problem [Eq. (1)] for the scattered field $\{\mathbf{E}_\zeta^s, \mathbf{H}_\zeta^s\}$ in $D_{0,1}$ and the total field inside D_i.

In the frame of asymmetrical DSM the approximate solution is constructed by representing the electromagnetic fields as a finite linear combination of the fields of dipoles deposited in the corresponding domains. Their fields satisfy analytically the transmission conditions enforced at the plane interface Σ.[44] Then the approximate solution fulfils Maxwell equations in the domains $D_\zeta, \zeta = 0, 1, i$, the radiation conditions and the transmission conditions at Σ. Thus, the scattering problem is reduced to the problem of approximating the exciting field on the scattering surface ∂D_i. The amplitudes of discrete sources are to be determined from the boundary conditions at the particle surface, which can be written in following form

$$\mathbf{n}_p \times (\mathbf{E}_i - \mathbf{E}_0^s) = \mathbf{n}_p \times \mathbf{E}_0^0; \quad \mathbf{n}_p \times (\mathbf{H}_i - \mathbf{H}_0^s) = \mathbf{n}_p \times \mathbf{H}_0^0 \tag{2}$$

where $\{\mathbf{E}_0^0, \mathbf{H}_0^0\}$ is the transmitted plane wave inside D_0.

To construct the fields scattered by an obstacle we will use system of localized electric dipoles.[45] The field of the dipoles analytically satisfy the transmission conditions for the tangential components of the electric and magnetic fields enforced at the plane interface Σ. In this case the vector potential of the dipole localized in M_0 accepts the form

$$\mathbf{A}(M) = \int_V \overleftrightarrow{\mathbf{G}}(M, M_0)\mathbf{j}(P)d\tau_P \tag{3}$$

where $\overleftrightarrow{\mathbf{G}}(M, M_0)$ is the Green tensor of a half-space:[32]

$$\overleftrightarrow{\mathbf{G}}(M, M_0) = \begin{bmatrix} G_{11} & 0 & 0 \\ 0 & G_{11} & 0 \\ \frac{\partial g}{\partial x_M} & \frac{\partial g}{\partial y_M} & G_{33} \end{bmatrix}. \tag{4}$$

The corresponding tensor components can be written as

$$\begin{aligned} G_{\alpha\alpha}(M, M_0) &= \frac{k_0}{j} h_0^{(2)}(k_0 R_{MM_0}) \\ &\quad + \int_0^\infty J_0(\lambda r) v_{\alpha\alpha}(\lambda, z) \exp\{-\eta_0 z\} \lambda \, d\lambda, \quad \alpha = 1, 3; \\ g(M, M_0) &= \int_0^\infty J_0(\lambda r) v_{31}(\lambda, z_0) \exp\{-\eta_0 z\} \lambda \, d\lambda. \end{aligned} \tag{5}$$

Here $R_{MM_0} = r^2 + (z - z_0)^2$; $r^2 = \rho^2 + \rho_0^2 - 2\rho\rho_0 \cos(\varphi - \varphi_0)$ $J_0(\cdot)$ is the cylindrical Bessel function, $h_0^{(2)}(\cdot)$ is the spherical Hankel function, (ρ, φ, z) are the cylindrical coordinates of the point M. The associating spectral functions v_{11}, v_{33}, v_{31} are given by

$$\begin{aligned} v_{11}(\lambda, z_0) &= \frac{\mu_1\eta_0 - \mu_0\eta_1}{\mu_1\eta_0 + \mu_0\eta_1}\frac{1}{\eta_0}\exp\{-\eta_0 z_0\}; \\ v_{33}(\lambda, z_0) &= \frac{\varepsilon_1\eta_0 - \varepsilon_0\eta_1}{\varepsilon_1\eta_0 + \varepsilon_0\eta_1}\frac{1}{\eta_0}\exp\{-\eta_0 z_0\}; \quad z \geq 0, \;\; z_0 > 0; \\ v_{31}(\lambda, z_0) &= \frac{2(\mu_1\varepsilon_1 - \mu_0\varepsilon_0)}{(\mu_1\eta_0 + \mu_0\eta_1)(\varepsilon_1\eta_0 + \varepsilon_0\eta_1)}\exp\{-\eta_0 z_0\}. \end{aligned} \tag{6}$$

where $\eta_\zeta^2 = \lambda^2 - k_\zeta^2, k_\zeta^2 = k^2\varepsilon_\zeta\mu_\zeta, \zeta = 0, 1$.

We will construct the representation for the scattered fields in $D_{0,1}$ on a system of electric dipoles, deposited in some points M_n. Let $\{M_n\}_{n=1}^{\infty}$ be a dense set of points, which are distributed over an auxiliary surface $S_0 \in C^{(2,\alpha)}$ deposited inside D_i and co-axial with ∂D_i (Fig. 1). Let us consider three linearly independent electric dipoles $\{\mathbf{e}^l\}_{l=1}^3$ situated in every point M_n and orientated in accordance with the cylindrical coordinate system. Following [Eq. (3)] the corresponding potentials for these dipoles accept the form:

$$\begin{aligned} \mathbf{A}_n^1(M) &= \Bigg\{ G_{11}(M, M_n)\cos(\varphi - \varphi_n); -G_{11}\sin(\varphi - \varphi_n); \\ &\qquad \frac{\partial g(M, M_n)}{\partial\rho}\cos(\varphi - \varphi_n) - \frac{1}{\rho}\frac{\partial g}{\partial\varphi}\sin(\varphi - \varphi_n)\Bigg\}; \\ \mathbf{A}_n^2(M) &= \Bigg\{ G_{11}(M, M_n)\sin(\varphi - \varphi_n); G_{11}\cos(\varphi - \varphi_n); \\ &\qquad \frac{\partial g(M, M_n)}{\partial\rho}\sin(\varphi - \varphi_n) + \frac{1}{\rho}\frac{\partial g}{\partial\varphi}\cos(\varphi - \varphi_n)\Bigg\}; \\ \mathbf{A}_n^3(M) &= \{0; 0; G_{33}(M, M_n)\}. \end{aligned} \tag{7}$$

Then for the scattered fields in $D_{0,1}$ which satisfy the transmission conditions at Σ, the following representation is valid:

$$\begin{aligned} \mathbf{E}_\zeta^N(M) &= \sum_{n=1}^{N_0}\sum_{l=1}^{3} p_{nl}^0 \text{ rot rot } \mathbf{A}_n^l(M), \\ \mathbf{H}_\zeta^N(M) &= \frac{j}{k\mu_\varsigma}\text{rot } \mathbf{E}_\zeta^N(M), \quad \zeta = 0, 1; \;\; M \in D_{0,1} \end{aligned} \tag{8}$$

Let us emphasize that the scattered fields in D_0 and D_1 are represented by a unitary set of dipoles amplitudes $\{p_{nl}^0\}$.

Consider now the construction of the approximated solution for the total field inside an obstacle D_i. We will choose the surface S_i, which confines ∂D_i (Fig. 1), and a dense set of points $\{M_n^i\}_{n=1}^{\infty}$, which are distributed over S_i. Let us introduce the following system of vector potentials corresponding to the external electrical dipoles:

$$\mathbf{A}_{ni}^l(M) = h_0^{(2)}(k_i R_{MM_n^i})\mathbf{e}_l; \quad M_n^i \in \mathrm{S}_i.$$

Then the representation for the total field inside the particle is:

$$\begin{aligned} \mathbf{E}_i^N(M) &= \sum_{n=1}^{N_i}\sum_{l=1}^{3} p_{nl}^i \text{ rot rot } \mathbf{A}_{ni}^l(M), \\ \mathbf{H}_i^N(M) &= \frac{j}{k\mu_i} \text{ rot } \mathbf{E}_i^N(M), \quad M \in D_i \end{aligned} \tag{9}$$

The representation for the internal field [Eq. (9)] essentially differs from those presented in,[35] because we have in this case 3D problem. Deposition of S_i outside D_i region seems to be more efficient especially for metal particles of small diameter. The representations [Eqs. (8) and (9)] satisfy all the conditions of the boundary value problem [Eq. (1)] except the boundary conditions enforced at the particle surface ∂D_i. The unknown amplitudes of the discrete sources $\{\{p_{nl}^0\}_{n=1}^{N_0}\}_{l=1}^3; \{\{p_{nl}^i\}_{n=1}^{N_i}\}_{l=1}^3$ are to be defined from [Eq. (2)]. Because the representation for the scattered field [Eq. (8)] satisfies the transmission conditions for the fields at the interface Σ, it allows us to take into account all interactions between particle and prism surface in an analytical way.

The completeness of the system of dipoles guarantees the convergence of the approximate solution to the exact one.[35]

2.3. *DSM numerical scheme*

In this section, we describe the details of the DSM numerical scheme, which was realized from the new DSM based model of the last section. We also focus on the differences between the new model and the conventional one.[35]

As has been mentioned, the approximate DSM solution [Eqs. (8) and (9)] satisfies most of the conditions of the scattering problem [Eq. (1)]. So, the determination of the unknown amplitudes of the DS is performed from the fitting of the boundary conditions [Eq. (2)] enforced at the particle

surface ∂D_i in $L_2(\partial D_i)$ norm.[35] To do this, we need an expression for the exciting field $\{\mathbf{E}_0^0, \mathbf{H}_0^0\}$ in the D_0 domain. In our case the P,S-polarized transmitted field in D_0 takes the form

$$\mathbf{E}_0^0 = T^{P,S}\mathbf{e}^{P,S}\exp(-j\mathbf{kr}), \quad \mathbf{H}_0^0 = n_0\mathbf{k}\times\mathbf{E}_0^0; \tag{10}$$

here $\mathbf{k} = (-\sin\theta_1\cos\varphi_1; -\sin\theta_1\sin\varphi_1; -\cos\theta_1)^T$ is a wave vector;

$$\mathbf{e}^P = (-\cos\theta_1\cos\varphi_1; -\cos\theta_1\sin\varphi_1; \sin\theta_1)^T;$$
$$\mathbf{e}^S = (\sin\varphi_1; -\cos\varphi_1; 0)^T;$$
$$T^P = \frac{2n_1\cos\theta_1}{n_0\cos\theta_1 + n_1\cos\theta_0}; \quad T^S = \frac{2n_1\cos\theta_1}{n_1\cos\theta_1 + n_0\cos\theta_0};$$

$n_\zeta = \sqrt{\varepsilon_\zeta\mu_\zeta}$ is the refractive index in D_ζ, θ_0 is the refraction angle of the wave transmitted into D_0. Snell's law in this case requires: $\sin\theta_0 = n_1/n_0\sin\theta_1$. Since $|n_1| > |n_0|$ from a certain incident angle $\theta_1 \geq \theta_C$, $\theta_C = \arcsin(n_0/n_1)$, $|\sin\theta_0| > 1$. Then in the upper half-space an evanescent wave appears, which is concentrated along the interface and exponentially decays along the normal to the interface. To take into account this circumstance following branch for $\cos\theta_0$ should be chosen as: $\cos\theta_0 = -j\sqrt{\sin^2\theta_0 - 1}$.

Various schemes for the DS amplitude determination are at our disposal. It has been established that more stable results can be obtained by using pseudo-inversion of an over-determined system of linear equations obtained by following the generalized point-matching technique.[44] Select a set of matching points on the particle surface $\{P_j\}_{j=1}^J \in \partial D_i$, homogeneously covering the particle surface ∂D_i. Then the linear system to be used for determining the DS amplitudes can be found from matching the boundary conditions [Eq. (2)] at the set $\{P_j\}_{j=1}^J$. This procedure leads to an over determined linear system with a dimension of $4J\times 3(N_0+N_i)$. The DS amplitudes are evaluated by a pseudo inversion technique.[44] The DSM scheme enables us to estimate the error in the solution by calculating the surface residual [Eq. (2)] in $L_2(\partial D_i)$ norm.

Let us emphasize the main differences between the new DSM scheme and the axisymmetrical DSM one applied to axial symmetric structures on a substrate.[35]

1. The approximate solution [Eqs. (8) and (9)] is independent of the polarization of the external excitation in contrast to the axisymmetrical

DSM. This allows examination of the scattering problem for the whole set of polarizations and incident angles simultaneously.

2. The representation for the scattered field [Eq. (9)] involves dipole sources only. This simplifies the numerical evaluation of the Sommerfeld integrals [Eq. (5)], and provides a more stable scheme for the matrix elements computation compared to the axisymmetrical DSM where multipole sources are used.[35]

During computation the number of matching points is increased until the required accuracy of the results is achieved. The DS number usually is 4–6 times less than the number of matching points. The resulting errors are estimated by computation of a surface residual in least-square norm. As a rule in a range of parameters under consideration relative surface residual less than 0.2% ensures that the error of the results is less than 0.1%.

After the amplitudes of the DS are determined, one can calculate the far field pattern $\mathbf{F}(\theta,\varphi)$ of the scattered field, which is determined at the upper part of the unit semisphere $\Omega = \{0^\circ \le \theta < 90^\circ, 0^\circ \le \phi \le 360^\circ\}$ and is given by

$$\mathbf{E}(M)/|\mathbf{E}_0^0(z=0)| = \frac{\exp\{-jk_0 r\}}{r}\mathbf{F}(\theta,\varphi) + O(1//r^2), \quad r = |M| \to \infty, \;\; z > 0.$$

We use an asymptotic approximation for the Sommerfeld integrals,[46] which leads to the following representation for the components of the scattering diagram

$$\begin{aligned} F_\theta(\theta_1,\theta,\varphi) &= jk\sum_{n=1}^{N_0}\{p_{n1}^0 f_{\theta 1} + p_{n2}^0 f_{\theta 2} - p_{n3}^0 f_{\theta 3}\}, \\ F_\varphi(\theta_1,\theta,\varphi) &= jk\sum_{n=1}^{N_0}\{-p_{n1}^0 f_{\varphi 1} + p_{n2}^0 f_{\varphi 2}\}. \end{aligned} \tag{11}$$

where

$$\begin{aligned} f_{\theta 1} &= \cos\theta\cos(\varphi-\varphi_n)\gamma^+ + \{\cos\theta\bar{v}_{11}(\theta) + \sin^2\theta\bar{v}_{33}(\theta)\}\cos\theta\cos(\varphi-\varphi_n)\gamma^-, \\ f_{\theta 2} &= \cos\theta\sin(\varphi-\varphi_n)\gamma^+ + \{\cos\theta\bar{v}_{11}(\theta) + \sin^2\theta\bar{v}_{33}(\theta)\}\cos\theta\sin(\varphi-\varphi_n)\gamma^-, \\ f_{\theta 3} &= \sin\theta\gamma^+ + \cos\theta\bar{v}_{33}(\theta)\sin\theta\gamma^-, \\ f_{\varphi 1} &= \sin(\varphi-\varphi_n)\gamma^+ + \cos\theta\bar{v}_{11}(\theta)\sin(\varphi-\varphi_n)\gamma^-, \end{aligned}$$

$$f_{\varphi 2} = \cos(\varphi - \varphi_n)\gamma^+ + \cos\theta \bar{v}_{11}(\theta)\cos(\varphi - \varphi_n)\gamma^-;$$

$$\gamma^+ = \exp\{ik_0(\rho_n \sin\theta\cos(\varphi - \varphi_n) + z_n\cos\theta)\};$$

$$\gamma^- = \exp\{ik_0(\rho_n \sin\theta\cos(\varphi - \varphi_n) - z_n\cos\theta)\}.$$

Here (ρ_n, φ_n, z_n) are the cylindrical coordinates of the points $M_n \in S_0$. Spectral functions involved into representation [Eq. (11)] accept the form

$$\bar{v}_{11}(\theta) = \frac{\mu_1\cos\theta - \mu_0\psi}{\mu_1\cos\theta + \mu_0\psi}, \quad \psi = \sqrt{\varepsilon_1 - \sin^2\theta}, \quad \bar{v}_{33}(\theta) = \frac{\varepsilon_1\cos\theta - \varepsilon_0\psi}{\varepsilon_1\cos\theta + \varepsilon_0\psi},$$

$$\bar{v}_{31}(\theta) = \frac{2(\mu_1\varepsilon_1 - \mu_0\varepsilon_0)}{(\varepsilon_1\cos\theta + \varepsilon_0\psi)(\mu_1\cos\theta + \mu_0\psi)}.$$

Hence, the components of the far field pattern [Eq. (11)] do not contain Sommerfeld integrals and can be represented as finite linear combinations of elementary functions. This circumstance ensures a low cost computation of the scattering characteristics in the far zone.

2.4. *Numerical results*

In this section some results obtained on the base of DSM model are presented. We will consider the intensity of scattered unpolarized light

$$I(\theta_1, \theta, \varphi) = \frac{1}{2}\left(|I^P(\theta_1, \theta, \varphi)|^2 + |I^S(\theta_1, \theta, \varphi)|^2\right) \tag{12}$$

where $I^{P,S}(\theta_1, \theta, \varphi) = |F_\theta^{P,S}(\theta_1, \theta, \varphi)|^2 + |F_\phi^{P,S}(\theta_1, \theta, \varphi)|^2$ and $F_{\theta,\varphi}^{P,S}(\theta_1, \theta, \varphi)$ are the components of the far field pattern for P and S polarized incident plane wave in a spherical coordinate system θ, φ: [Eq. (11)]. We are mainly interested in the objective scattering cross-section (objective response), which represents the integrated intensity scattered into the prescribed solid angle Θ

$$\sigma(\lambda) = \int_\Theta I(\theta_1, \theta, \varphi)d\omega, \tag{13}$$

where $\Theta = \{0 \le \theta \le \theta_{NA}; 0 \le \varphi \le 360^\circ\}$, θ_{NA} is an angle, which corresponds to the numerical aperture (NA) of the objective lens in accordance with $\theta_{NA} = \arcsin(NA/n_0)$. To be close to the reality, NA = 0.75 has been chosen.

As a shape model of nanorod we consider an elongated spheroid immersed in the water. By the diameter D of spheroid we mean diameter of an equivolume sphere. For nanorod material we consider gold (Au) and

silica glass for the substrate material. To represent incidence, an unpolarized light is mostly used. The incident angle is $\theta_1 = 67°$ which is close to the critical angle $\theta_c = \arcsin(n_0/n_1)$ for our system. Let us mention, that n_0 and n_1 depend on the exciting wavelength λ, and $\theta = 0°$ corresponds to the normal direction to the prism surface. Different orientations of the spheroid with respect to the incident plane are examined. In the results presented below the $\varphi = 180°$ plane corresponds to the incidence case where P-polarized $\mathbf{E}^0$ belongs to the zx-plane and $\varphi = 90°$ plane corresponds to the incidence case when P-polarized $\mathbf{E}^0$ belongs to the zy-plane. In Fig. 2 the objective response [Eq. (13)] versus wavelength for spheroids of diameter $D = 45\,\text{nm}$ with different aspect ratios r are presented for an observing plane orientation $\varphi = 180°$. From the results one can see, that the objective response for $r = 2$ demonstrates the maximal value.

The computer objective responses for spheroids with $D = 50\,\text{nm}$ and $D = 60\,\text{nm}$ for different aspect ratios are presented in Figs. 3 and 4 respectively. Even when we take a finer step in aspect ratio values, we can conclude that maximum enhancement of the objective response for spheroid is reached at $r = 2$.

In Fig. 5 the comparison for several spheroids of different D for the aspect ratio $r = 2$ is given. From the results we can see that if the objective response is higher, the diameter is bigger. Together with the increasing of the nanorod size, the peak of the objective response slightly moves to the

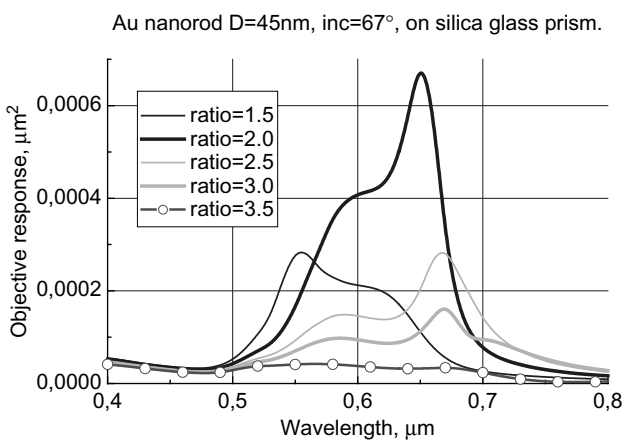


Fig. 2. The objective response [Eq. (13)] versus wavelength for gold (Au) spheroids of diameter $D = 45\,\text{nm}$ for different aspect ratios r.

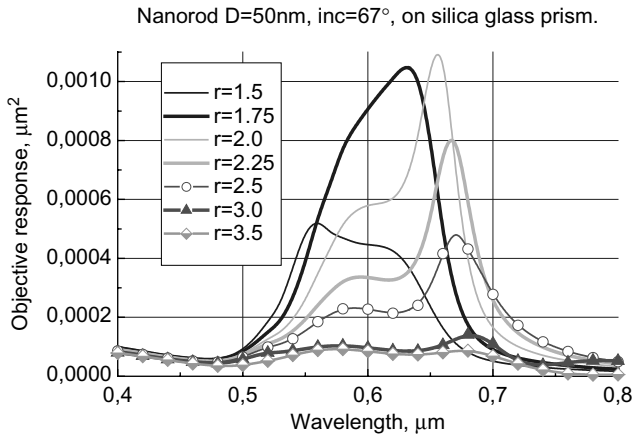


Fig. 3. The objective response versus wavelength for gold spheroids of $D = 50\,\text{nm}$, different r.

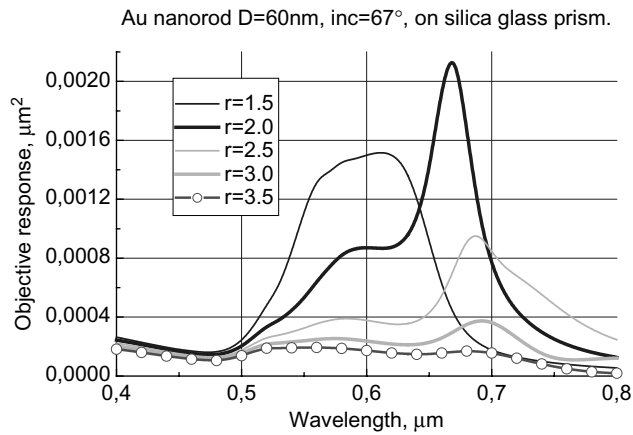


Fig. 4. The objective response versus wavelength for gold spheroids of $D = 60\,\text{nm}$, different r.

area of longer wavelengths. But for all the nanorods presented in Fig. 5 the curves keep the same shape with two pronounced maxima.

Let us now vary the position of the incident plane. In Fig. 6 the objective response versus wavelength for a gold spheroid of $D = 50\,\text{nm}$ and $r = 2$ is given for several different orientations of the incident plane φ compared to the orientation average, where maximum response is achieved for $\varphi = 90°$ and minimum for $\varphi = 180°$. In general, with rotation of

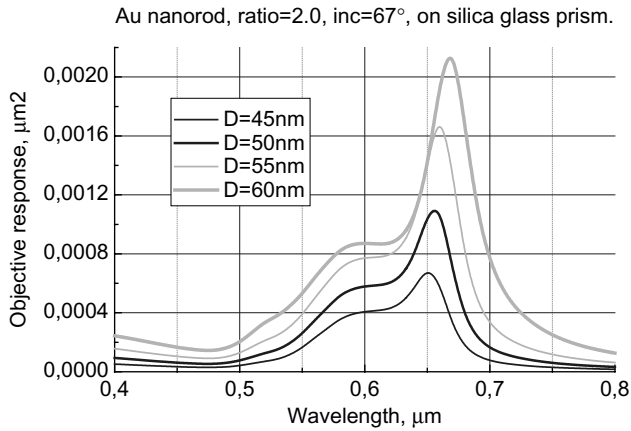


Fig. 5. The objective response versus wavelength for gold spheroid with aspect ratio $r = 2$ for different D.

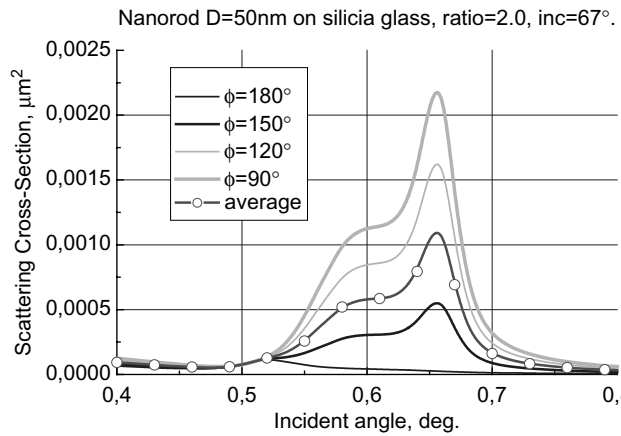


Fig. 6. The objective response versus wavelength for a gold spheroid of $D = 50\,\text{nm}$ and $r = 2$ for different orientations of the incident plane compared to an average orientation.

the incidence plane from 180° to 90°, the objective response of a spheroid increases. In Fig. 7 similar results are also seen for aspect ratio $r = 3$. The behavior of the curves is similar to those presented in Fig. 6. In addition, the results for a spheroid of $D = 60\,\text{nm}$ and $r = 2$ as presented in Fig. 8 are in good agreement with the previous observations.

In Fig. 9 the objective response versus wavelength for a gold spheroid of $D = 50\,\text{nm}$ and $r = 2$ is given for incidence planes 180° to 90° at two

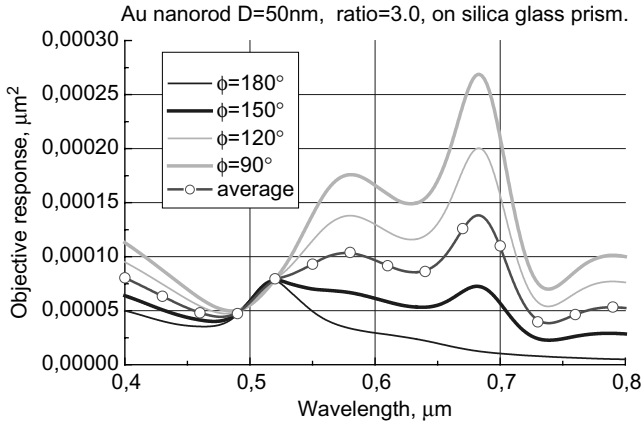


Fig. 7. The objective response versus wavelength for a gold spheroid of $D = 50\,\text{nm}$ and $r = 3$ for different orientations of the incident plane compared to an average orientation.

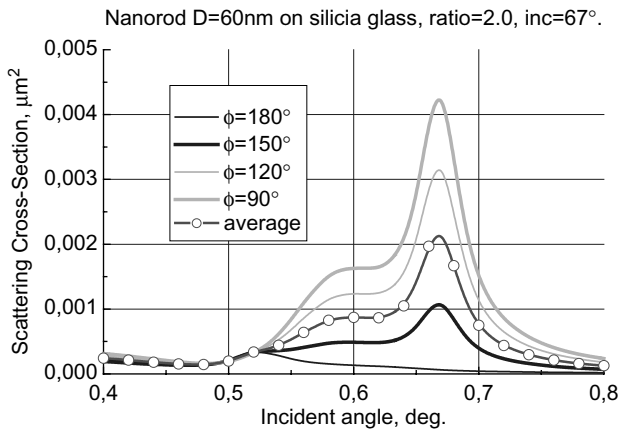


Fig. 8. The objective response versus wavelength for a gold spheroid of $D = 60\,\text{nm}$ and $r = 2$ for different orientations of the incident plane compared to an average orientation.

different incident angles: $\theta_1 = 67°$ and $\theta_1 = 70°$. From the presented results, we can conclude that the behavior of the response curve depends neither on the spheroid's aspect ratio nor on the incident angle while a maximum value is achieved for the incident plane orientation $\varphi = 90°$.

Let us now investigate the behavior of the scattered intensity of polarized light when the objective response achieves its maximum value. For this purpose, two spheroids with $D = 50\,\text{nm}$ and $D = 60\,\text{nm}$ were

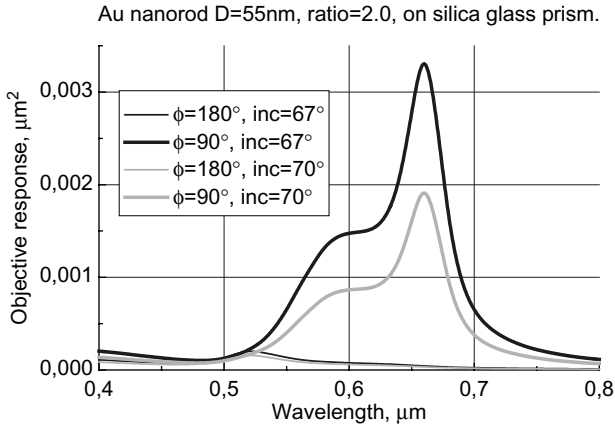


Fig. 9. The objective response versus wavelength for a gold spheroid of $D = 55\,\text{nm}$ and $r = 2$ for two different orientations of the incident plane and two incident angles.

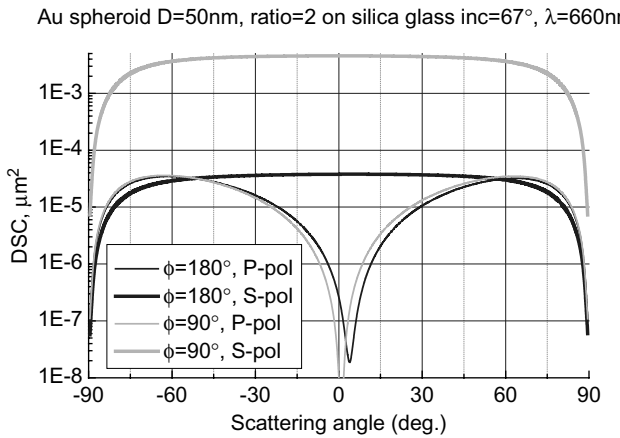


Fig. 10. The scattering intensity of P and S-polarized components versus scattering angle for a gold spheroid of $D = 50\,\text{nm}$ and $r = 2$ for two different orientations of the incident plane, $\lambda = 660\,\text{nm}$.

taken. The intensity of the polarized light versus scattering angle for both spheroids is presented in Figs. 10 and 11 respectively. The wavelengths chosen fit to the position of the response maximum value. One can see that while curves for P-polarized light do not greatly depend on the orientation of the incident plane, the intensity of S-polarization drastically increases with a rotation of the incidence plane from 180° to 90°. Thus, the main

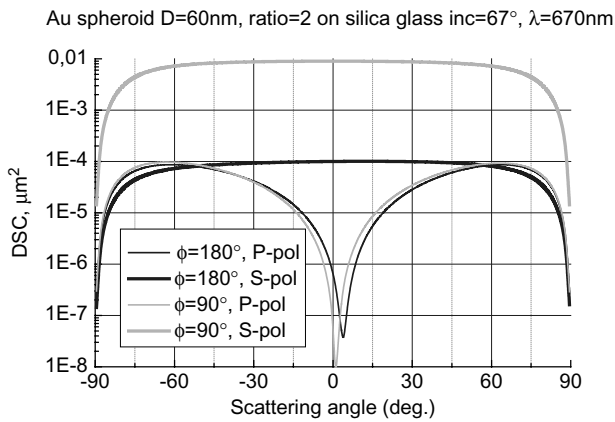


Fig. 11. The scattering intensity of P and S-polarized components versus scattering angle for a gold spheroid of $D = 60$ nm and $r = 2$ for different orientations of the incident plane, $\lambda = 670$ nm.

contribution to scattered intensity is provided by the S-polarized component of the exciting plane wave.

To summarize our research using computer simulation analysis:

- The spectral peak shifts to the area of longer wavelengths with the increasing aspect ratio of a spheroid.
- The spectral peak value increases when vector $\mathbf{E}^0$ is parallel to the longer axis of a spheroid. An excitation of the longitude mode[8] is responsible for this effect.
- The spectral peak reaches its maximum value for a spheroid when the aspect ratio is approximately $r = 2$.

2.5. *Conclusion*

In this part of the contribution the DSM has been applied to the scattering analysis of a nanorod deposited on a prism surface. The main goal of the investigation was to clarify the scattering behavior of nanorods with respect to their aspect ratio and orientation under excitation. The results obtained by simulations support the conclusion that an aspect ratio $r = 2$ provides the maximum value of the scattered field enhancement. It has been found that the response increases when the incident plane rotates from 180° to 90° where the main contribution to scattering enhancement was provided by the S-component of the incident light. So, the longitude mode contribution is responsible for this enhancement.

3. Light Scattering by a Nanoshell

3.1. *Introduction*

Noble metal nanoparticles attracted great interest in recent years due to their special properties in the optical range.[3,5] Nowadays nanoparticles have practical value in many applications such as chemical and biochemical sensors, medical diagnostics and therapeutics, biological imaging, nanophotonics.[10,11,23,24] When the dimension of a metal nanoparticle is small enough compared to the wavelength of incident light, surface plasmon can be excited due to a collective motion of free electrons at the surface of the metal nanoparticle that resonantly couples with the oscillating electric field of the light. As a result of surface plasmon excitation, there are strong enhancements of absorption, scattering, and local electric field in the vicinity of particle and in the far zone. This effect strongly depends on particle size, shape, material composition as well as type of the local environment.[29,38,39,44,47] Optical absorption and scattering have dominant effects at different size regions, which can be used specifically for certain applications. The dependence of the plasmon resonance on surrounding media is useful in the imaging of attached biosystems.[15]

Nanoshells are nanoparticles composed of a dielectric core coated with an ultra-thin noble metallic shell. They are considered to be an improvement of solid noble particles.[5] In this part we present the description of the DSM based model for a spherical nanoshell deposited on a surface.

3.2. *DSM model for nanoshell*

Consider an axial symmetric penetrable particle covered with a layer. We will mostly keep the notations used before, with some exceptions: now the particle interior domain be D_i and layer domain be D_l. Let us define the smooth particle and layer boundaries as $\partial D_{i,l}$ (Fig. 12). Then mathematical statement [Eq. (1)] should be supplemented with $\zeta = l$ for [Eq. (1a)] and boundary conditions [Eq. (1d)] instead of [Eq. (1b)]:

$$\text{rot } \mathbf{H}_\zeta = jk\varepsilon_\zeta \mathbf{E}_\zeta; \quad \text{rot } \mathbf{E}_\zeta = -jk\mu_\zeta \mathbf{H}_\zeta \quad BD_\zeta, \quad \zeta = 0, 1, i, l, \tag{1a'}$$

$$\mathbf{n}_i \times (\mathbf{E}_i - \mathbf{E}_l) = 0, \quad \mathbf{n}_i \times (\mathbf{H}_i - \mathbf{H}_l) = 0, \quad \text{at } \partial D_i$$

$$\mathbf{n}_l \times (\mathbf{E}_l - \mathbf{E}_0) = 0, \quad \mathbf{n}_l \times (\mathbf{H}_l - \mathbf{H}_0) = 0, \quad \text{at } \partial D_l \tag{1d}$$

$$\left.\begin{aligned} &\mathbf{e}_z \times (\mathbf{E}_0(p) - \mathbf{E}_1(p)) = 0, \\ &\mathbf{e}_z \times (\mathbf{H}_0(p) - \mathbf{H}_1(p)) = 0, \end{aligned}\right\} \quad p \in \Sigma;$$

We will now describe the main steps of the algorithm and outline the features which differ from the nanorod model described before. The main difference is that in case of a layered particle we have an axisymmetric system (particle including the surface) and can essentially use this symmetry. The approximate solution should be now constructed for the scattered field $\{\mathbf{E}^s_\zeta, \mathbf{H}^s_\zeta\}$ in $D_{0,1}$ and the total field $\{\mathbf{E}_{i,l}, \mathbf{H}_{i,l}\}$ inside both core and shell of the particle.

The amplitudes of discrete sources are determined instead of Eq. (2) from the boundary conditions at the core and shell surfaces:

$$\begin{aligned} \mathbf{n}_i \times \begin{bmatrix} (\mathbf{E}_i(p) - \mathbf{E}_1(p)) = 0 \\ (\mathbf{H}_i(p) - \mathbf{H}_1(p)) = 0 \end{bmatrix}, & \qquad p \in \partial D_i \\ \mathbf{n}_l \times \begin{bmatrix} (\mathbf{E}_l(p) - \mathbf{E}^s_0(p)) = \mathbf{n}_l \times \mathbf{E}^0_0 \\ (\mathbf{H}_l(p) - \mathbf{H}^s_0(p)) = \mathbf{n}_l \times \mathbf{H}^0_0 \end{bmatrix}, & \qquad p \in \partial D_l. \end{aligned} \tag{14}$$

For the approximate solution the system of multipole is employed. Following the DSM outlines, an approximate solution to the scattering problem is constructed taking into account the rotational symmetry of the scattering problem geometry (particle plus interface) and the polarization of the exciting field.[35]

Consider P-polarized plane wave excitation. In this case the refracted plane wave into D_0 accepts the form

$$\begin{gathered} \mathbf{E}^0_0 = T^P_{1,0}(-\mathbf{e}_x \cos\theta_0 + \mathbf{e}_z \sin\theta_0)\chi_0, \quad \mathbf{H}^0_0 = -T^P_{1,0} n_0 \mathbf{e}_y \chi_0, \\ \chi_0 = \exp[-jk_0(x \sin\theta_0 + z\cos\theta_0)]. \end{gathered} \tag{15}$$

Here $\eta_\xi = \sqrt{\varepsilon_\xi \mu_\xi}$ is associating refractive index of D_ξ, θ_0 is a refraction angle for the plane wave. The refraction coefficient $T^P_{1,0}$ can be written as $T^P_{1,0} = \frac{2n_1 \cos\theta_1}{n_1 \cos\theta_0 + n_0 \cos\theta_1}$. Beyond the critical angle $\theta_1 > \theta_c = \arcsin(n_0/n_1)$ in the upper half-space an evanescent wave appears, which is propagated

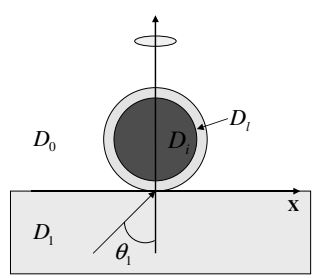


Fig. 12. Axial symmetric penetrable particle covered with a layer.

along the interface and damping in the normal ($z > 0$) direction $\chi_0 = \exp\{-jk_0 x \sin\theta_0\} \exp\{-k_0 z\sqrt{\sin^2\theta_0 - 1}\}$.[48]

Let us take into account the axial symmetry of the problem geometry and the polarization of the external excitation. For this purpose we will deposit the multipoles at the axis of symmetry (z-axis). To account for the polarization of the exciting field [Eq. (3)] we implement plane wave resolution in Fourier series with respect to the azimuthal angle φ. Then for Fourier harmonics of the external excitation [Eq. (3)] in cylindrical coordinates system the following representations are valid:

$$\begin{aligned}
\mathbf{E}^0_{0(m)} &= \{\mathbf{e}^{0,P}_{m\rho}(\xi)\cos(m+1)\varphi;\ \mathbf{e}^{0,P}_{m\varphi}(\xi)\sin(m+1)\varphi;\ \mathbf{e}^{0,P}_{mz}(\xi)\cos(m+1)\varphi\},\\
\mathbf{H}^0_{0(m)} &= \{\mathbf{h}^{0,P}_{m\rho}(\xi)\sin(m+1)\varphi;\ \mathbf{h}^{0,P}_{m\varphi}(\xi)\cos(m+1)\varphi;\ \mathbf{h}^{0,P}_{mz}(\xi)\sin(m+1)\varphi\}.
\end{aligned} \tag{16}$$

where $\xi = (\rho, z)$ stands for a point located in the half plane $\varphi = \text{const}$. To take the polarization of the external excitation into account we should use some linear combination of electrical and magnetic multipoles deposited along the axis of symmetry of the particle. For the scattered field representation outside a particle the following vector potentials are used:[35]

$$\begin{aligned}
\mathbf{A}^{e,0}_{mn} &= \{G^e_m(\xi, z_n)\cos(m+1)\varphi;\ -G^e_m(\xi, z_n)\sin(m+1)\varphi;\\
&\quad - g_{m+1}(\xi, z_n)\cos(m+1)\varphi\}\\
\mathbf{A}^{h,0}_{mn} &= \{G^h_m(\xi, z_n)\sin(m+1)\varphi;\ G^h_m(\xi, z_n)\cos(m+1)\varphi;\\
&\quad - g_{m+1}(\xi, z_n)\sin(m+1)\varphi\},\\
\mathbf{A}^{e,0}_{n} &= \{0; 0; G^h_0(\xi, z_n)\}
\end{aligned} \tag{17}$$

here $G^{e,h}_m$, g_m are the Fourier harmonics corresponding to the Green tensor[35] components, which accept the form

$$\begin{aligned}
G^{e,h}_m(\xi, z_n) &= Y^0_m(\xi, z_n) + \int_0^\infty J_m(\lambda\rho) v^{e,h}_{11}(\lambda, z_n)\exp\{-\eta_0 z\}\lambda^{1+m} d\lambda,\\
g_m(\xi, z_n) &= \int_0^\infty J_m(\lambda\rho) v_{31}(\lambda, z_n)\exp\{-\eta_0 z\}\lambda^{1+m} d\lambda,\\
Y^0_m(\xi, z_n, k_0) &= \frac{k_0}{i} h^{(2)}_m(k_0 R_{\xi z_n})(k_0\rho/R_{\xi z_n})^m, \quad R^2_{\xi z_n} = \rho^2 + (z - z_n)^2,
\end{aligned} \tag{18}$$

Here J_m is the cylindrical Bessel function, $h_m^{(2)}$ is the spherical Hankel function and the coordinates of multipoles $\{z_n\} \in Oz$ are distributed inside $D_i \cup D_l$. Associating spectral functions $v_{11}^{e,h}$, v_{31} are given by Eq. (6).

Following the previous directions, for the representation of the total field inside the core and shell of the particle we use the following vector potentials:

$$\begin{aligned}
\mathbf{A}_{mn}^{e,\kappa} &= \{Y_m^\kappa(\xi, z_n^\kappa)\cos(m+1)\varphi;\ -Y_m^\kappa(\xi, z_n^\kappa)\sin(m+1)\varphi; 0\} \\
\mathbf{A}_{mn}^{h,\kappa} &= \{Y_m^\kappa(\xi, z_n^\kappa)\sin(m+1)\varphi;\ Y_m^\kappa(\xi, z_n^\kappa)\cos(m+1)\varphi; 0\}; \qquad (19) \\
\mathbf{A}_{n}^{e,\kappa} &= \{0; 0; Y_0^\kappa(\xi, z_n^\kappa)\}.
\end{aligned}$$

where $\kappa = i, l\pm$ and $Y_m^i(\xi, z_n^i) = Y_m^i(\xi, z_n^i, k_i)$,

$$\begin{aligned}
Y_m^{l+}(\xi, z_n^{l+}) &= h_m^{(2)}(k_l R_{\xi z_n^{l+}})(k_l\rho/R_{\xi z_n^{l+}})^m, \\
Y_m^{l-}(\xi, z_n^{l-}) &= j_m(k_l R_{\xi z_n^{l-}})(k_l\rho/R_{\xi z_n^{l-}})^m.
\end{aligned}$$

Here j_m are the spherical Bessel functions, $\{z_n^{i,l\pm}\} \in Oz$, and $\{z_n^{i,l\pm}\} \subset D_i$. So, the fields inside the particle's shell are presented as superposition of outgoing $(l+)$ and standing $(l-)$ waves.

Now we can formulate our approximate solution to the scattering problem for the P-polarized excitation as

$$\begin{aligned}
\mathbf{E}_\zeta^N &= \sum_{m=0}^{M}\sum_{n=1}^{N_m^\zeta}\left\{p_{mn}^\zeta \frac{j}{k\varepsilon_\zeta\mu_\zeta}\nabla\times\nabla\times\mathbf{A}_{mn}^{e,\zeta} + q_{mn}^\zeta\frac{1}{\varepsilon_\zeta}\nabla\times\mathbf{A}_{mn}^{h,\zeta}\right\} \\
&\quad + \sum_{n=1}^{N_0^\zeta} r_n^\zeta \frac{j}{k\varepsilon_\zeta\mu_\zeta}\nabla\times\nabla\times\mathbf{A}_n^{e,\zeta}; \qquad (20) \\
\mathbf{H}_\zeta^N &= \frac{j}{k\mu_\zeta}\nabla\times\mathbf{E}_\zeta^N; \quad \zeta = 0, i, \pm l. \\
\mathbf{E}_l^N &= \mathbf{E}_{l+}^N + \mathbf{E}_{l-}^N; \quad \mathbf{H}_l^N = \mathbf{H}_{l+}^N + \mathbf{H}_{l-}^N
\end{aligned}$$

Last term in [Eq. (20)] corresponds to vertical electric dipoles because $\mathbf{E}_0^0$ vector belongs to the incident plane. Let us emphasize that the scattered field $\{\mathbf{E}_0^N, \mathbf{H}_0^N\}$ in the domains $D_{0,1}$ is represented in terms of the unitary set of amplitudes $\{p_{mn}^0, q_{mn}^0, r_n^0\}$, because the transmission conditions at the interface Σ are satisfied analytically by the Green tensor components [Eq. (18)]. This fact is the key feature because it allows us to take into account all interactions between obstacle and interface analytically.

Consider now S-polarization of the exciting plane wave. In this case the exciting field accepts the following form

$$\mathbf{H}_0^0 = T_{1,0}^S n_0(-\mathbf{e}_x\cos\theta_0 + \mathbf{e}_z\sin\theta_0)\chi_0, \quad \mathbf{E}_0^0 = T_{1,0}^S\mathbf{e}_y\chi_0,$$
$$T_{1,0}^S = \frac{2n_1\cos\theta_1}{n_1\cos\theta_1 + n_0\cos\theta_0}. \tag{21}$$

The corresponding Fourier harmonics for the external excitation can be written as

$$\mathbf{E}_{0(m)}^0 = \{\mathbf{e}_{m\rho}^{0,S}(\xi)\sin(m+1)\varphi;\ \mathbf{e}_{m\varphi}^{0,S}(\xi)\cos(m+1)\varphi;\ \mathbf{e}_{mz}^{0,S}(\xi)\sin(m+1)\varphi\},$$
$$\mathbf{H}_{0(m)}^0 = \{\mathbf{e}_{m\rho}^{0,S}(\xi)\cos(m+1)\varphi;\ \mathbf{e}_{m\varphi}^{0,S}(\xi)\sin(m+1)\varphi;\ \mathbf{e}_{mz}^{0,S}(\xi)\cos(m+1)\varphi\}. \tag{22}$$

To construct the fields we should use the following electric and magnetic potentials for the scattered field in $D_{0,1}$ and total fields inside the core and shell of the particle:

$$\begin{aligned}
\mathbf{A}_{mn}^{e,0}(\xi,z_n) &= \{G_m^e(\xi,z_n)\sin(m+1)\varphi;\ G_m^e(\xi,z_n)\cos(m+1)\varphi;\\
&\quad - g_{m+1}(\xi,z_n)\sin(m+1)\varphi\},\\
\mathbf{A}_{mn}^{h,0}(\xi,z_n) &= \{G_m^h(\xi,z_n)\cos(m+1)\varphi; -G_m^h(\xi,z_n)\sin(m+1)\varphi;\\
&\quad - g_{m+1}(\xi,z_n)\cos(m+1)\varphi\},\\
\mathbf{A}_n^{e,h,0}(\xi,z_n) &= \{0;0;G_0^{h,e}(\xi,z_n)\}.\\
\mathbf{A}_{mn}^{e,\kappa} &= \{Y_m^\kappa(\xi,z_n^\kappa)\sin(m+1)\varphi;\ Y_m^\kappa(\xi,z_n^\kappa)\cos(m+1)\varphi;0\};\\
&\qquad\qquad \kappa = i, l\pm;\\
\mathbf{A}_{mn}^{h,\kappa} &= \{Y_m^\kappa(\xi,z_n^\kappa)\cos(m+1)\varphi; -Y_m^\kappa(\xi,z_n^\kappa)\sin(m+1)\varphi;0\};\\
\mathbf{A}_n^{h,\kappa} &= \{0;0;Y_0^\kappa(\xi,z_n^\kappa)\}.
\end{aligned} \tag{23}$$

Finally, for S polarization the approximate solution accepts the following representation

$$\begin{aligned}
\mathbf{E}_\zeta^N &= \sum_{m=0}^{M}\sum_{n=1}^{N_m^\zeta}\left\{p_{mn}^\zeta\frac{j}{k\varepsilon_\zeta\mu_\zeta}\nabla\times\nabla\times\mathbf{A}_{mn}^{e,\zeta} + q_{mn}^\zeta\frac{1}{\varepsilon_\zeta}\nabla\times\mathbf{A}_{mn}^{h,\zeta}\right\}\\
&\quad + \sum_{n=1}^{N_0^\zeta} r_n^\zeta\frac{1}{\varepsilon_\zeta}\nabla\times\mathbf{A}_n^{e,\zeta};\\
\mathbf{H}_\zeta^N &= \frac{j}{k\mu_\zeta}\nabla\times\mathbf{E}_\zeta^N; \quad \zeta = 0, i, \pm l\\
\mathbf{E}_l^N &= \mathbf{E}_{l+}^N + \mathbf{E}_{l-}^N; \quad \mathbf{H}_l^N = \mathbf{H}_{l+}^N + \mathbf{H}_{l-}^N
\end{aligned} \tag{24}$$

In contrast to [Eq. (20)] the last term in [Eq. (24)] is associated with vertical magnetic dipoles because now $\mathbf{H}_0^0$ vector belongs to the incident plane. Completeness of the system of the lowest-order distributed multipoles used in [Eqs. (20) and (24)][45] assures the convergence of the approximate solution to the exact one in least square norm at the core-shell surfaces:

$$\sqrt{\left\|\mathbf{n}_i \times \begin{matrix}\mathbf{E}_i^N - \mathbf{E}_l^N \\ \mathbf{H}_i^N - \mathbf{H}_l^N\end{matrix}\right\|_{L_2(\partial D)}^2 + \left\|\mathbf{n}_l \times \begin{matrix}\mathbf{E}_l^N - \mathbf{E}_0^N - \mathbf{E}_0^0 \\ \mathbf{H}_l^N - \mathbf{H}_0^N - \mathbf{H}_0^0\end{matrix}\right\|_{L_2(\partial D)}^2} \to 0, \quad N \to \infty \tag{25}$$

The last circumstance guarantees closeness of the approximate solution to the exact one in any closed subset of D_0.[35]

3.3. *Numerical scheme of the DSM*

Let us now describe the computational algorithm in details. As we outlined above the representations [Eqs. (20) and (24)] satisfy all the conditions of the scattering problem [Eq. (1)] except the transmission conditions at the core and shell surfaces [Eq. (14)]. These conditions are used to determine the unknown amplitudes of discrete sources $\{p_{mn}^{0,i,\pm l}, q_{mn}^{0,i,\pm l}, r_n^{0,i,\pm l}\}$. Since the scattering problem geometry is axially symmetrical with respect to the z-axis representations for the scattered and total fields [Eqs. (20) and (24)] accept the form of finite sum of the Fourier series with respect to φ. Besides, we have made resolution of the exciting field in the Fourier series. This means that fulfilling the transmission conditions [Eq. (14)] at the surfaces $\partial D_{i,l}$ can be reduced to a sequential solution of the 1D transmission problems for the Fourier harmonics of the fields. So, instead of matching the fields on the scattering surfaces [see Eq. (14)], we can match their Fourier harmonics separately thus reducing the approximation problem on the surfaces $\partial D_{i,l}$ to a set of 1D problems enforced at the particle and layer generatrices $\Im_{i,l}$. By solving these problems one can determine the required DS amplitudes $\{p_{mn}^{0,i,\pm l}, q_{mn}^{0,i,\pm l}, r_n^{0,i,\pm l}\}$.

Various numerical schemes for the amplitudes evaluation have been suggested. It has been found that more stable results can be obtained by using the Generalized Point-Matching Technique and a pseudo-solution of the associated over-determined system of linear equations.[44] As DSM is a direct method, it allows solving the scattering problem for the entire set of incident angles θ_1 and for both polarizations (P and S) simultaneously. Besides, this numerical scheme provides an opportunity to control the

actual convergence of the approximate solution to the exact one by *a posteriori* evaluating of surface residual [Eq. (25)].[35]

We use an extended version of DSM algorithm described in Ref. (28). In the frame of this extension we use different numbers of the DS for the representation of the scattered field outside and total field inside core and shell of the particle. The numbers of DS are chosen proportionally to the value of refractive index of the corresponding media. For the internal domains (higher refractive index $\sqrt{\varepsilon_{i,l}\mu_{i,l}} > \sqrt{\varepsilon_0\mu_0}$) we use a higher number of DS than for the scattered field $N_{i,l}^m > N_s^m$. Besides the number of discrete sources depends on the rank of Fourier harmonics $N_t^m > N_s(m)$. For higher harmonics we use a lower number of multipoles $N_\xi^{m+1} \leq N_\xi^m$. This circumstance enables us to acquire a more accurate simulation result, provides a monotone decrease of the surface residual and reduces the demand on computer resources up to 30% for larger particles compared to the conventional DSM model.[44] The order of multipoles (M) is *a priori* defined from the condition that the plane wave approximation by corresponding Fourier series should be less than 0.1%.

The θ, φ-components of the far field pattern corresponding to representation [Eq. (20)] have a form:

$$\begin{aligned} F_\theta^{\mathrm{P}}(\theta,\varphi) &= jk_0 \sum_{m=0}^{M} (jk_0 \sin\theta)^m \cos(m+1)\varphi \\ &\quad \times \sum_{n=1}^{N_m^0} \{p_{mn}^0 [D_n^e \cos\theta + jk_0 \sin^2\theta F_n] + q_{nm}^0 D_n^h\} \\ &\quad - jk_0 \sin\theta \sum_{n=1}^{N_0^0} r_n^0 D_n^h, \\ F_\varphi^{\mathrm{P}}(\theta,\varphi) &= -jk_0 \sum_{m=0}^{M} (jk_0 \sin\theta)^m \sin(m+1)\varphi \\ &\quad \times \sum_{n=1}^{N_m^0} \{p_{mn}^0 D_n^e + q_{nm}^0 [D_n^h \cos\theta + jk_0 \sin^2\theta F_n]\}. \end{aligned} \tag{26}$$

For S-polarized excitation following representation [Eq. (24)] one gets

$$\begin{aligned} F_\theta^{\mathrm{S}}(\theta,\varphi) &= jk_0 \sum_{m=0}^{M} (jk_0 \sin\theta)^m \sin(m+1)\varphi \\ &\quad \times \sum_{n=1}^{N_m^0} \{p_{mn}^0 [D_n^e \cos\theta + jk_0 \sin^2\theta F_n] - q_{nm}^0 D_n^h\}, \end{aligned}$$

$$F_{\varphi}^{\mathrm{S}}(\theta,\varphi) = jk_0 \sum_{m=0}^{M} (jk_0 \sin\theta)^m \cos(m+1)\varphi \times \sum_{n=1}^{N_m^0} \{p_{mn}^0 D_n^e - q_{nm}^0 [D_n^h \cos\theta + jk_0 \sin^2\theta F_n]\} + jk_0 \sin\theta \sum_{n=1}^{N_0^0} r_n^0 D_n^e, \quad (27)$$

where the corresponding spectral functions $D_n^{e,h}, F_n$ are

$$D_n^{e,h}(\theta) = \exp\{jk_0 z_n \cos\theta\} + jk_0 \cos\theta v_{11}^{e,h}(k_0 \sin\theta, z_n),$$
$$F_n(\theta) = jk_0 \cos\theta v_{31}(k_0 \sin\theta, z_n).$$

After the unknown amplitudes of DS are determined, the far field patterns for P/S polarizations are represented as finite linear combinations of elementary functions. This fact ensures a low cost computer analysis of the scattering characteristics in the far zone.

3.4. *Results and discussion*

In this section, we will present numerical results obtained using the DSM model and discuss them briefly. As external excitation unpolarized light in the wavelength range of $400 \leq \lambda \leq 1050$ nm is used. We investigate particles with different core materials, which were chosen from those which are often used for nanoshell synthesis. As the layer material we used gold (Au) and silver (Ag), the ambient medium is water, and the substrate material is BK7 glass. The numerical aperture of objective lens is NA = 0.5. The refractive index data for noble particles and for substrate materials were taken from public source.[49]

For most of the presented results we took the fixed incident angle of 63°, which was close to the critical angle $\theta_c \cong 62°$ in the range of λ. In all the presented results the core diameter is decreasing while the shell thickness is increasing, keeping the external diameter of the core-shell particle (D) constant.

We would like to start with the investigation of the shell thickness influence on the scattering response. For this we chose Au_2S as core material with a gold shell. We calculated the objective response for the biosensors diameter $D = 40$ nm, but with different thicknesses of shell. The obtained results are shown on Fig. 13.

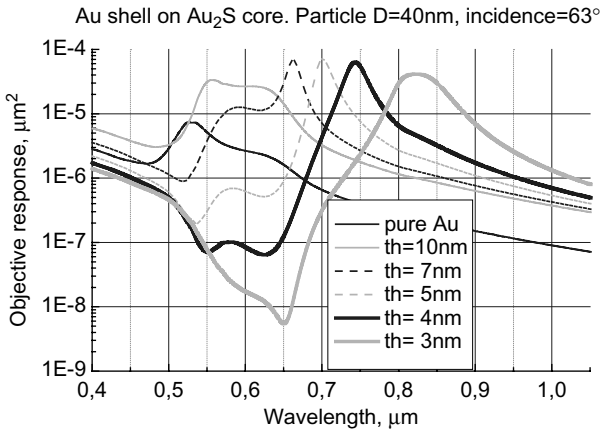


Fig. 13. Objective scattering response versus wavelength for biosensors $D = 40$ nm with different shell thicknesses.

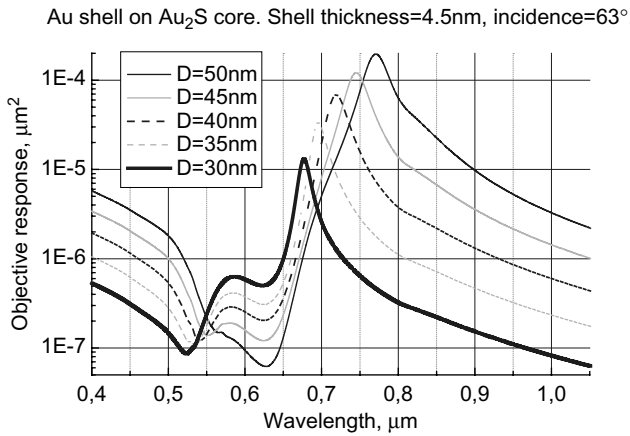


Fig. 14. Objective scattering response versus wavelength for biosensors of different diameters with fixed shell thickness.

From the results it is clear that increasing shell thickness leads to the shift of the maximum to the area of higher wavelengths. The width of the peak also differs depending on the thickness of a shell. In Fig. 14 the results for the fixed thickness of shell (4.5 nm) are presented for different particle diameters. One can see that with increasing diameters not only does the intensity of the corresponding peaks increase, but the intensity maximum also shifts to the area of longer wavelength as the peak gets wider.

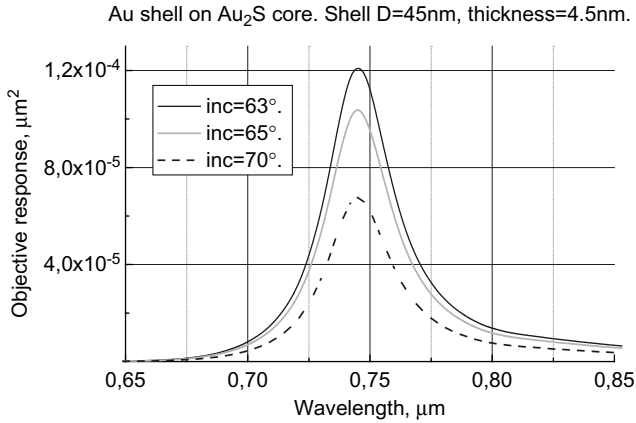


Fig. 15. Objective scattering response versus wavelength for biosensors of $D = 40\,\text{nm}$ with shell thickness of 4.5 nm for different incident angles.

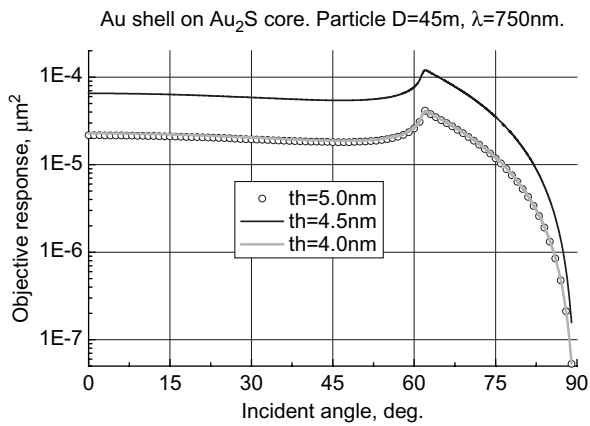


Fig. 16. Objective scattering response versus scattering angle for biosensors of $D = 45\,\text{nm}$ under wavelength of 750 nm for different shell thicknesses.

The optimal diameter which provides relatively high intensity together with a still not so wide peak corresponds to $D = 45\,\text{nm}$. Now let us fix both particle diameter and shell thickness and get back to the choice of our incident angle that stayed fixed up to now.

Figure 15 presents the objective response for the biosensor of $D = 45\,\text{nm}$ and a shell thickness of 4.5 nm for different incident angles. From the results one can see that the pattern stays the same, but the intensity decreases with increasing incident angle. Figure 16 shows the results for the

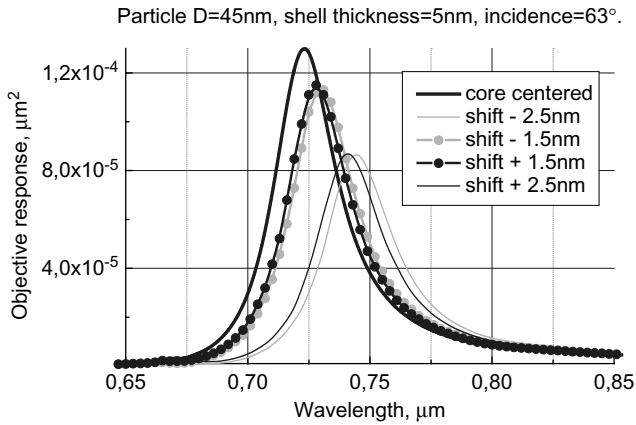


Fig. 17. Objective scattering response versus wavelength for a biosensor of $D = 45$ nm with shell thickness of 5 nm for different asymmetry models.

particles of $D = 45$ nm with different shell thicknesses for a fixed wavelength of 750 nm (which corresponds to the maximum of scattering response), but for different incident angles. The response curves first is very close to a straight line with a local maximum in the area of a critical angle (about 62°) and then there is a rather fast decrease of the intensity with an increase of the incident angle. From these results it is obvious that the response for the biosensor with shell thickness of 4.5 nm is much higher than for particles with shell thicknesses of 4 nm and 5 nm, in spite of the variation of thickness is just about 1% of the total particle diameter.

Now, after we have shown how sensitive the objective response is to biosensor properties, we would like to analyze another important aspect: the core-shell asymmetry of the nanoshell. Biosensor asymmetry, which is seldom taken into account in theoretical investigations, occurs rather often.[24,25] Figure 17 presents the scattering response from biosensor of $D = 45$ nm with a shell thickness of 5 nm under the incidence of 63° for different core position inside the particles. A shift in this case means, that the center of core is shifted up (+) or down (−) inside the biosensor. From the simulations one can see that even such a small core-shift of 2.5 nm shifts the maximum of the response about 30 nm and essentially decreases its intensity. So, a possible asymmetry of the biosensors should be taken into account for the design of sensitive systems, where the shift of the resonance peak at 30 nm could lead to the wrong interpretation of results.

In Fig. 18 the objective response for a silver nanoshell with two different core materials SiO_2 and PSL and two shell thicknesses of 3 and 4 nm versus

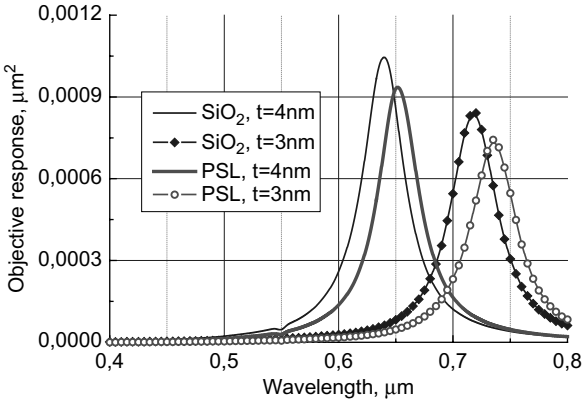


Fig. 18. Objective scattering response versus wavelength for Ag nanoshell $D = 45\,\text{nm}$ and different core materials for different shell thicknesses.

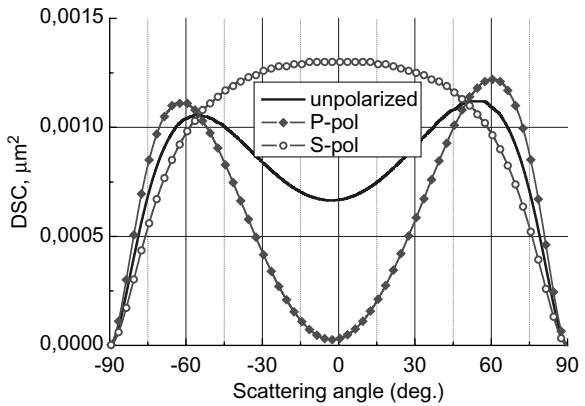


Fig. 19. Scattering diagram versus scattering angle of Ag nanoshell with SiO_2 core $D = 45\,\text{nm}$ and $t = 3\,\text{nm}$ for different polarizations, $\lambda = 720\,\text{nm}$.

wavelengths is presented. From the diagrams it is easy to see that an increase of the refractive index of the core material reduces the resonance intensity. We can also observe that even a slight deviation in shell thickness changes the diagram more than a variation of the core material. Reducing the shell thickness by just 1 nm shifts the peak by nearly 80 nm and shifts the intensity peaks into the "biological window".

We will concentrate now on a silver nanoshell with the most promising SiO_2 core and a shell thickness of 3 nm and have a closer look at its scattering behavior. Next, the incidence wavelength is fixed at 720 nm, corresponding to the resonance wavelength of this particle. In Fig. 19 the

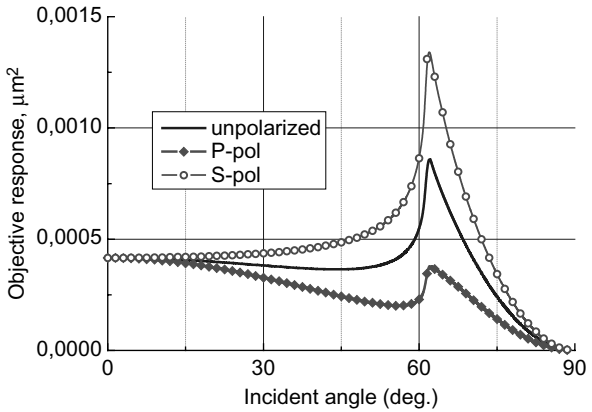


Fig. 20. Objective scattering response versus incident angle for Ag nanoshell with SiO_2 core $D = 45\,\text{nm}$ and $t = 3\,\text{nm}$ for different polarizations, $\lambda = 720\,\text{nm}$.

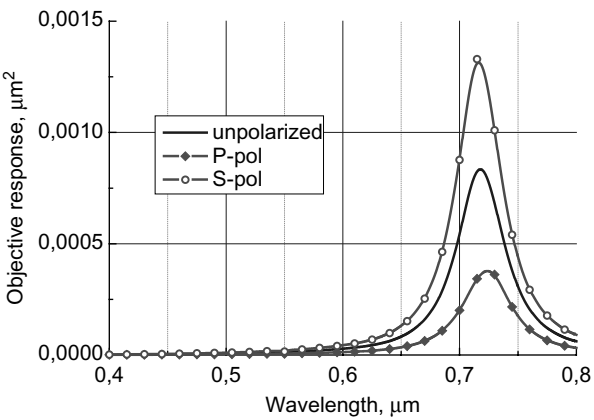


Fig. 21. Objective scattering response versus wavelength for Ag nanoshell with SiO_2 core $D = 45\,\text{nm}$ and $t = 3\,\text{nm}$ for different polarizations.

scattering diagram versus the scattering angle is presented for different polarizations of external excitation. From the diagram it is obvious that S-polarization provides the main contribution to the scattering. In Fig. 20 the objective response versus incident angle is presented for different polarizations of the excitation. Next, the incidence wavelength is fixed at 720 nm, corresponding to the resonance wavelength of this particle. From the diagram one can see that the intensity of S-polarized light at the critical angle is approximately 3.5 times greater than for P-polarized light.

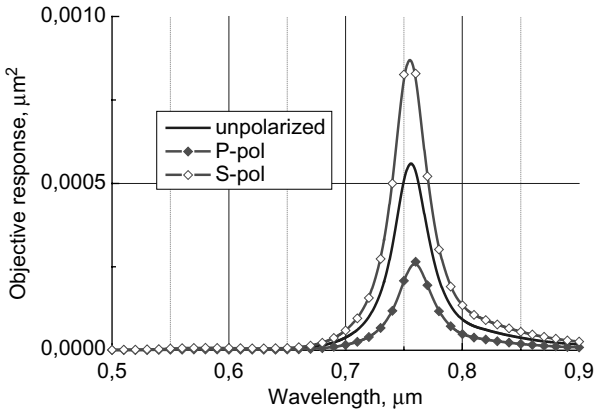


Fig. 22. Objective scattering response versus incident angle for Au nanoshell with SiO_2 core $D = 45\,\text{nm}$ and $t = 3\,\text{nm}$ for different polarizations.

Coming back to the scattering spectra, we now analyze the influence of incident polarization on a resonance peak of a nanoshell. In Figs. 21 and 22 objective responses versus incident wavelength for silver (Fig. 21) and gold (Fig. 22) nanoshells are presented. Here again we see that the intensity in case of S-polarized excitation is much higher than for unpolarized light. We could conclude that using S-polarized light instead of unpolarized light could provide higher enhancement making the detection of the resonance easier.

3.5. *Conclusion*

In this part of the contribution we presented the scattering analysis of nanoshell on a glass prism based on the DSM. We analyzed the dependence of the scattering behavior of such a biosensor on its parameters: diameter, shell thickness, angle of incidence and core-shell asymmetry. The results based on the DSM model allow the prediction of the optimal parameters for biosensors synthesis, which provide maximum enhancement and additionally a shift of the resonance into the window of biological transparency. For a gold nanoshell, the optimal diameter is $D = 45\,\text{nm}$ with a shell thickness of 4.5 nm at an incident angle of 63°. It has also been shown that the scattering behavior strongly depends on the core-shell asymmetry, which may lead to incorrect interpretation of experimental results if it is not included in the simulation model. It has also been detected, that for different core materials we investigated, the higher the refractive index of the core

was, the lower was the intensity of the resonance peak. The S-polarization of the incident light provides the main contribution to the scattering by a nanoshell and it is possible to efficiently increase the resonance intensity by using just S-polarized excitation. At the same time, we must conclude that the core material does not strongly influence the position of resonance peak and does not allow the manipulation of the resonance peak position, as was possible by varying other parameters, such as shell thickness, particle size and core-shell asymmetry.

4. Summary

In this chapter we presented the application of the DSM for two problems of nanooptics. The model of a nanorod is a 3D problem, while the model of nanoshell employs an axially symmetrical model reducing the computational demand. The different models are responsible for essential differences in the DSM numerical algorithm, which were described in the corresponding sections. Based on the DSM models the numerical results for light scattering by nanoobjects were presented and discussed. From the results analysis, a conclusion about the influence of the internal structure of the particles, their material and shape on their scattering behavior presents itself. We believe that mathematical modeling and computer simulation can serve as independent tools for the investigation and analysis of scattering behavior of nanostructures in line with experimental studies which play an important role in development of nanooptical technologies.

References

1. P. Bharadwaj, B. Deutsch, and L. Novotny, *Advances in Optics and Photonics* **1**, 438 (2009).
2. J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light* (Princeton NJ: Princeton University Press, 2008).
3. I. Abdulhalim, M. Zourov, and A. Lakhtakia, *Electromagnetics* **28**, 214 (2008).
4. T. Endo, D. Ikeda, Y. Kawakami, Y. Yanagida, and T. Hatsuzawa, *Analytica Chimica Acta* **661**, 200 (2010).
5. G. Raschke, S. Brogl, A. S. Susha, A. L. Rogach, T. A. Klar, J. Feldmann, B. Fieres, N. Petkov, T. Bein, A. Nichtl, and K. Kürzinger, *Nanoletters* **4**, 1853 (2004).

6. B. Evlyukhin, G. Brucoli, L. Martín-Moreno, S. Bozhevolnyi, and F. J. García-Vidal, *Phys. Rev. B* **76**, 075426 (2007).
7. Y.-F. Chau, D. P. Tsai, G.-W. Hu, L.-F. Shen, and T.-J. Yang, *Optical Engineering* **46**, 039701 (2007).
8. K. Imura, T. Nagahara, and H. Okamoto, *J. Phys. Chem. B* **109**, 13214 (2005).
9. H. M. Hiep, H. I Yoshikawa, and E. Tamiya, *Anal. Chem.* **82**, 1221 (2010).
10. B. N. Khlebtsov, V. A. Khanadeev, V. A. Bogatyrev, L. A. Dykman, and N. G. Khlebtsov, *Nanotechnologies in Russia* **3**, 442 (2008).
11. L. R. Bickford, G. Agollah, R. Drezek, and T.-K. Yu, *Breast Cancer Research and Treatment* **120**, 547 (2010).
12. V. V. Kislyuk and O. P. Dimitriev, *J. Nanoscience and Nanotechnology* **8**, 131 (2008).
13. X. Huang, I. H. El-Sayed, and M. A. El-Sayed, **624**, 343 (2010).
14. Z. Yuan and H. Jiang, **624**, 309 (2010).
15. L. R. Hirsch, J. B. Jackson, A. Lee, N. J. Halas, and J. L. West, *Anal. Chem.* **75**, 2377 (2003).
16. E. Simsek, *Optics Express* **18**, 1722 (2010).
17. W. L. Barnes, A. Dereux, and T. W. Ebbesen, *Nature* **424**, 824 (2003).
18. J. J. Mock, M. Barbic, D. R. Smith, D. A. Schultz, and S. Schultz, *J. Chem. Phys.* **116**, 6755 (2002).
19. O. P. Varnavski, M. B. Mohamed, M. A. El-Sayed, and T. Goodson III, *J. Phys. Chem. B* **107**, 3101 (2003).
20. K.-S. Lee and M. A. El-Sayed, *J. Phys. Chem. B* **109**, 20331 (2005).
21. J. Pérez-Juste, I. Pastoriza-Santos, L. M. Liz-Marzán, and P. Mulvaney, *Coordination Chemistry Reviews* **249**, 1870 (2005).
22. J. J. Mock, D. R. Smith, and S. Schultz, *Nanoletters* **3**, 485 (2003).
23. R. D. Averitt, D. Sarkar, and N. J. Halas, *Phys. Rev. Lett.* **78**, 4217 (1997).
24. J.-E. Jönsson, O. J. Karlsson, H. Hassander, and B. Törnell, *Macromolecules* **34**, 1512 (2001).
25. J. B. Lassiter, M. W. Knight, N. A. Mirin, and N. J. Halas, *Nano Letters* **9**, 4326 (2009).
26. A. W. H. Lin, N. A. Lewinski, M.-H. Lee, and R. A. Drezek, *Journal of Nanoparticle Research* **8**, 681 (2006).
27. L. Helden, E. Eremina, Y. Eremin, N. Riefler, C. Hertlein, C. Bechinger, and T. Wriedt, *Appl. Optics* **45**, 7299 (2006).
28. E. Eremina, Y. Eremin, and T. Wriedt, *Opt. Comm.* **273**, 278 (2007).
29. E. Eremina, Y. Eremin, and T. Wriedt, *J. Comput. Theor. Nanosci.* **5**, 2186 (2008).
30. Ch. Hafner. *Phys. Stat. Sol. B* **244**, 3435 (2007).
31. I. R. Capoglu and G. S. Smith, *IEEE Trans. Antennas Propag.* **52**, 3805 (2006).
32. J. Alegret, P. Johansson, and M. Käll, *New Journal of Physics* **10**, 105004 (2008).
33. T. Wriedt, *JQSRT* **11**, 833 (2009).

34. J. Parsons, C. P. Burrows, J. R. Sambles, and W. L. Barnes, *Journal of Modern Optics* **57**, 356 (2010).
35. Yu. A. Eremin, *J. Comm. Technology and Electronics* **45**, 269 (2000).
36. N. Riefler, E. Eremina, C. Hartlein, L. Helden, Y. Eremin, and T. Wriedt, *JQSRT* **106**, 464 (2007).
37. Yu. A. Eremin, J. C. Stover, and N. V. Orlov, *Opt. Eng.* **38**, 1296 (1999).
38. E. K. Payne, K. L. Shuford, S. Park, G. C. Schatz, and C. A. Mirkin, *J. Phys. Chem. B* **110**, 2150 (2006).
39. R. Ashayer, M. Green, and S. H. Mannan, *Journal of Nanoparticle Research* **12**, 1489 (2010).
40. L. R. Hirsch, A. M. Gobin, A. R. Lowery, F. Tam, R. A. Drezek, N. J. Halas, and J. L. West, *Annals of Biomedical Engineering* **34**, 15 (2006).
41. S. Schelm and G. B. Smith, *JOSA A* **22**, 1288 (2005).
42. D. J. Wu, X. J. Liu, L. L. Liu, and W. P. Qian, *Applied Physics A* **92**, 279 (2008).
43. D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory* (Springer, Berlin, 1992).
44. Y. Eremin, N. Orlov, and A. Sveshnikov, in: T. Wriedt (Ed.), *Generalizes Multipole Techniques for Electromagnetic and Light Scattering* Vol. 39, Elsevier Science, Amsterdam, 1999.
45. A. Doicu, Yu. A. Eremin, and T. Wriedt, *Acoustic and Electromagnetic Scattering Analysis Using Discrete Sources* (Academic Press, London, 2000).
46. Yu. A. Eremin and A. G. Sveshnikov, *Comput. Maths. Math. Phys.* **39**, 1967 (1999).
47. E. Eremina, Y. Eremin, and T. Wriedt, *Opt. Comm.* **267**, 524 (2006).
48. J. A. Kong, *Electromagnetic Wave Theory* (EMW Publ., Cambridge, MA, 2000).
49. http://refractiveindex.info/

COMPUTATIONAL TECHNIQUES FOR PLASMONIC ANTENNAS AND WAVEGUIDES

Christian Hafner*, Jasmin Smajic†, and Ralf Vogelgesang‡

*Electromagnetic Fields and Microwave Electronics Laboratory,
ETH Zurich, 8092 Zurich, Switzerland
christian.hafner@ifh.ee.ethz.ch

†ABB Corporate Research Ltd., Department of Electrotechnology,
Baden-Dättwil, Switzerland

‡Department Kern, Max-Planck-Institute for Solid State Research,
70569 Stuttgart, Germany

First, the main computational techniques for solving electromagnetic problems with focus on plasmonics, antennas, and waveguides are reviewed. Special attention is paid to special symmetries, such as symmetry with respect to one or several planes, cylindrical symmetry, rotational symmetry, periodic symmetry in one or more directions, and combinations of these symmetries. Then, tricky problems, such as a metallic tip — excited by a fundamental wire mode — and a finite chain of metallic spheres, embedded in a dielectric cylinder are investigated using the finite element method (FEM) and the multiple multipole program (MMP) in order to illustrate the procedures.

1. Introduction

Plasmonics has attracted much interest in nano science for a variety of reasons.[1] Quite simply, many of the structures now being studied could not have been manufactured a few decades ago, and microscopy techniques with adequate resolution for the study of their plasmonic excitations in real-space are coming of age only in the last few years.[2,3] The possibilities of modern bottom-up as well as top-down approaches to artificial material designs have elevated many topics from academic constructs

to conceptions and innovations of real products. One key phenomenon is the possibility of extreme field enhancements at interfaces between materials of drastic dielectric contrast, leading to minute mode volumes. Controlled manufacturing of nanometer-sharp edges, corner, and gaps facilitates novel approaches to hyper-responsive detectors and super-resolution microscopes.

Being collective phenomena, surface plasmon polaritons also exhibit fascinating concerted interactions over extended distances, such as super-transmission through subwavelength hole arrays or the new class of "metamaterials". Arguably this ushered in an "optical engineering of space", with phenomena like negative refraction, transformation optics, and seemingly science-fictional concepts like cloaking. These have now become subject to rational study through tuning of the properties of the constituent "metaatoms" — their composition, shape, dimensions, relative locations, and environment.

Such artificial metamedia may still be considered continuous at the scale of optical wavelengths. But the exciting outlook of optical circuitry as the consequent merging of electronics with optics requires other concepts. It will certainly feature specific functional elements of fundamentally subwavelength size. First proposals for active optical elements like amplifiers, transistors, or gates have been demonstrated — at this point, of course, without concerns for commercial applicability.

A number of crucial aspects of passive elements like interconnects (e.g., waveguides) and couplers (e.g., antennas) also remain as challenges. Reducing heat dissipation in long distance plasmonic waveguides calls for new designs that concentrate most of the optical field in non-dissipative dielectric materials. The generalization of impedance concepts from high-frequency electronics to plasmonic optical circuitry requires adequate concepts for impedance matching.

A prerequisite to tackling any of these challenges by simulations, is the reliable numerical computation. To a large extent it is safe to exclude quantum mechanical effects like spin, single particle properties, or quantum confinement. Within Maxwell's theory of electrodynamics, metals providing plasmonic effects near optical frequencies (OF) are simply described by strongly frequency dependent complex permittivities $\varepsilon(\omega)$ — usually with a negative real part. Beside this, the equations used for describing the electromagnetic fields are the same as at arbitrarily low frequencies. Therefore, one may benefit in plasmonics from the well established numerical methods and software packages for computational electromagnetics.

It is important to note that metals at frequencies below plasmon resonance are usually described by a complex epsilon with negative real part, but in the regime of radio frequencies (RF), microwaves (MW) and millimetre waves one usually assumes that the real part of the relative permittivity is 1 rather than negative, which is obviously wrong when one considers the Drude approximation

$$\varepsilon(\omega) = \varepsilon_0 \left(1 - \frac{\omega_p^2}{\omega(\omega + i/\tau)}\right). \tag{1}$$

Good numerical results may be obtained under the *wrong* assumption $\mathrm{Re}(\varepsilon(\omega)) = \varepsilon_0$ only if $\mathrm{Im}(\varepsilon(\omega))$ is strongly dominant, i.e., $|\mathrm{Re}(\varepsilon(\omega))| \ll |\mathrm{Im}(\varepsilon(\omega))|$, which holds at sufficiently low frequencies $\omega \ll \omega_p$. Then, metals are usually described by their electric conductivity σ and the complex permittivity is set $\varepsilon(\omega) = \varepsilon_0 + i\sigma/\omega$.

At optical frequencies, the condition $\omega \ll \omega_p$ does not hold and in interesting plasmonic situations, the imaginary part of the permittivity should be as small as possible, because it then is responsible for undesired material loss. It should also be mentioned that the Drude approximation is rather inaccurate for most of the attractive metals near plasmon frequency. For accurate simulations one therefore should work with measured data of $\varepsilon(\omega)$ or use more advanced and more accurate material models such as higher order Drude-Lorentz models and critical point analysis.[4] The latter are important when working with time domain codes such as finite difference time domain (FDTD),[5,6] time domain versions of the finite integral technique (FIT),[7] finite volume time domain (FVTD),[8] or time domain versions of finite element methods (FEM).[9] When selecting a time domain solver for plasmonics, one therefore should verify if higher order Drude-Lorentz models are supported. If this is not the case, one must subdivide the frequency domain of interest into sufficiently small intervals and approximate the metal in each interval with available low order Drude or Lorentz models and run the solver once for each interval, which obviously is cumbersome and time consuming. In any case, time domain solvers that allow the user to define only material properties without frequency dependence cannot be applied when the real part of the permittivity is negative because the update schemes in such codes will not be stable.

In principle it is very easy to write frequency domain codes that can handle materials with both positive and negative real parts of the

permittivity, but some of the available software packages might not have been tested and might not allow the user to specify negative real parts of the permittivity. Furthermore, it is important to note that the losses may be rather substantial in plasmonics. In order to compute them, one must make sure that the corresponding software package can handle lossy materials, i.e., complex permittivity. Since loss free computations are much easier than lossy ones, there are rather many software packages that cannot handle losses or only sufficiently small losses by means of some approximations. This especially holds when eigenvalue problems are involved, namely, when modes of waveguides are computed. As a result, one must be careful when selecting a software package for the analysis of plasmonic configurations. In the following, the main methods are outlined and hints for selecting appropriate methods are given. After this, the main concepts and difficulties of modeling plasmonic antennas and waveguides are presented. Finally, it is demonstrated how advanced structures consisting of antenna-waveguide combinations may be handled.

2. Time Domain Solvers

Time domain solvers are directly based on the standard formulation of Maxwell's equations in 3D space and time. It is important to note that this is done in the traditional way of considering the electromagnetic field as a time-dependent field in 3D space rather than as a field in 4D space-time.

The main advantage of time domain solutions is that one may benefit from Fourier transforms to obtain the frequency response over a wide frequency range as soon as the time response is known for a time signal such as a pulse that includes frequency components within the desired frequency range. Furthermore, most of the time domain techniques take advantage of some update scheme that allows one to efficiently compute the field at time $t + \mathrm{d}t$ from the field at time t and before. Since the time step $\mathrm{d}t$ is finite, a finite difference approach is applied along the time axis. This holds for all time domain solvers that are currently used and not only for finite difference solvers. The non-finite-difference solvers only handle the special field dependence in a different way.

Furthermore, it is assumed in most of the available software packages that the entire field is zero everywhere for time $t < 0$, which is essential for the simple initialization of time domain codes. In order to obtain a non-zero field for $t > 0$, one must necessarily impress a time-dependent non-zero field at least in one source point as the excitation of the field.

This implies that one cannot directly tackle eigenvalue problems, which have no explicit excitation. In fact, eigenvalue problems such as electromagnetic resonators are usually defined most naturally in frequency domain. Other features that are most naturally defined in frequency domain are dispersive material properties and periodic symmetries that will be outlined below.

2.1. *Eigenvalue problems*

Instead of handling eigenvalue problems without excitation directly, one can convert them into scattering problems by introducing a fictitious excitation and measure the response of the system in at least one sensor point. Such a configuration may then be treated as an ordinary scattering problem. This procedure mimics the approach of measuring eigenvalues, for example, resonance frequencies of a resonator by introducing a fixed-amplitude input port that feeds the system and an output port where the response of the system is measured as a function of the frequency. Eigenfrequencies are then detected as peaks in the response. Obviously, the fictitious excitation plays the role of the input port whereas the sensor point corresponds to the output port. This technique has several drawbacks. Namely, (1) an eigenvalue may not be excited when the fictitious excitation is poorly placed, (2) it may not be detected when the sensor point is poorly placed, (3) it may be impossible to separate eigenvalues that are close to each other, and (4) sometimes incorrect eigenvalues are detected.[10]

The problems (1)–(3) may be reduced by changing symmetries, e.g., by using more than a single excitation or by a random initialization of the field at $t = 0$, combined with a sufficiently large set of sensor points. (4) is especially dependent on how the frequency dependence of the response is obtained from the time dependence that is obtained from the time-domain technique. Fast methods such as the discrete Fourier transform (DFT) that are not memory consuming usually provide a frequency response with many ripples on both sides of a resonance peak that need smart filtering routines for avoiding the wrong detection of the ripples as eigenfrequencies. Harmonic inversion[11] avoids this but it requires the storage of the entire time response in the sensor points, which is memory consuming.

2.2. *Dispersive materials*

Another problem for all time domain solvers is caused by the strong material dispersion, i.e., the frequency dependence of the permittivity of metals at

optical frequencies. For handling frequency dependent equations — such as the material description $\overrightarrow{D} = \varepsilon \overrightarrow{E}$ — convolution integrals need to be solved. Various techniques for handling convolution integrals are available[5,6] but they only work efficiently — without drastically increasing the memory requirement and computation time — when the material properties may be approximated by certain simple formulae, such as the Drude formula given above. Such approximations lead to errors that are difficult to estimate and make the validation of the results cumbersome.

It should also be mentioned that time domain solvers are only stable when certain stability criteria are met.[5,6] For non-dispersive materials, the stability criteria are relatively simple — they essentially require the time steps to be short enough — but for dispersive materials it is hard to know under what conditions the algorithm is stable. In advanced material models one may have several parameters and the stability may depend on all of them. When these parameters are determined in a way that a good fit with measured material data is achieved, stability is not always guaranteed.

2.3. *Periodic symmetries*

Periodic symmetries are often encountered in computational optics, namely when gratings, photonic crystals, metamaterials, and periodic waveguides are studied. Periodic symmetries allow one to drastically reduce the memory requirement and computation time. Therefore, one should take them into account whenever possible.

It is important to note that periodic symmetries lead to a symmetry decomposition that is essentially a spatial Fourier decomposition along each periodic symmetry axis. This fits very well to frequency domain treatment. For example, when a grating with period Dx in x direction is considered, one obtains $\mathrm{Field}(x + Dx, y, z) = \exp(i \cdot Cx \cdot Dx) \cdot \mathrm{Field}(x, y, z)$. Since the symmetry factor $\exp(i \cdot Cx \cdot Dx)$ is complex, complex notation of the field is required. In principle the field in time domain is real valued, but it may be extended to a complex field, which doubles the memory requirement and increases the computation time by an even higher factor.

One may easily take advantage of the periodicity $\mathrm{Field}(x + Dx, y, z) = \exp(i \cdot Cx \cdot Dx) \cdot \mathrm{Field}(x, y, z)$ by introducing periodic boundary conditions, i.e., by reducing the computational domain to the interval $0 < x < Dx$ and setting $\mathrm{Field}(Dx, y, z) = \exp(i \cdot Cx \cdot Dx) \cdot \mathrm{Field}(0, y, z)$ on the boundary

at $x = Dx$ and $\text{Field}(0, y, z) = \exp(-i \cdot Cx \cdot Dx) \cdot \text{Field}(Dx, y, z)$ on the boundary at $x = 0$.

The main problem that remains is that the constant Cx is complex and frequency dependent. For example, when a grating is illuminated by a plane wave, Cx is the projection of the wave vector of the incident wave on the x axis, i.e., $Cx = k(\omega) \cdot \cos(\alpha)$, there the wavenumber k depends on the frequency. Thus, one has two options: (1) One assumes that the incident plane wave propagates in a certain direction and is time-harmonic. Then, the problem must be solved for each frequency separately, which spoils the advantage of time domain computations that can simultaneously compute an entire spectrum in the frequency domain. (2) One assumes for one computation that Cx is constant but the incident wave is not time-harmonic. Then one implicitly has a superposition of time-harmonic waves incoming from different directions for different frequencies. This allows one to reduce the computation time but requires post-processing of the data[12] for obtaining standard representations of the frequency dependence.

2.4. *Grid refinement*

Probably the most important feature of plasmonic structures is that even particles or features of size much smaller than the wavelength may cause strong field localization and sharp resonances. In order to capture strong resonances, the number of time steps must be high enough. Although this may cause long computation time, it is not a fundamental problem. Resolving small features that play an essential role for the electromagnetic properties of the structure requires a very fine local discretization. According to Refs. (5) and (6), the grid size for plasmonic objects should be 0.5 nm or less, which is much shorter than typical wavelengths of around 500 nm. Fine spatial discretization does not only lead to huge memory requirement, it also causes very short time steps because of the stability criterion.

For structures with small features that need to be discretized, procedures that allow local mesh refinements are favorable and highly desirable. In principle this leads to unstructured grids that are typical for finite element methods. In standard finite differences (FD), one prefers structured grids, usually equally spaced, orthogonal grid lines. For local refinements one may apply subgridding,[5] which requires appropriate interpolation

techniques and usually destroys the second order accuracy of standard FD operators on regular grids.

2.5. *Finite differences and finite integrals*

Finite difference time domain (FDTD) algorithms are essentially based on Yee's[13] leapfrog scheme that provides second order accuracy at a very low computational cost on an orthogonal regular grid. Because of the simplicity of this scheme, it is very easy to implement standard FDTD algorithms and therefore FDTD has become very popular.

As mentioned above, local mesh refinement is highly desirable for plasmonic structures. Therefore, the extension to irregular grids is desirable. In principle, FDTD starts with Maxwell's equations in time domain. After introducing a coordinate system (u, v, w), the resulting coupled system of differential equations is approximated by replacing all derivatives in u, v, w by finite differences. This process is the easiest in Cartesian coordinates (x, y, z). Deriving FD schemes in non-Cartesian coordinates is more difficult and the resulting FD update schemes are more difficult to implement and evaluate. For this reason, all well-known FDTD packages are based on Cartesian grids.

Typical commercial FDTD solvers are Lumerical[14] and OptiFDTD.[15]

In order to simplify the derivation of update schemes for non-Cartesian grids, Weiland[7] proposed to start from integral notation of Maxwell's equations, i.e., to integrate the curl equations using Stokes' law, which leads to integrals over closed loops. This technique is called finite integral technique (FIT). When applied to Cartesian grids, both FDTD and FIT lead to identical update schemes. FIT was first implemented in the MAFIA code[16] and later on in CST Microwave Studio.[17]

2.6. *Finite volume*

Finite volume time domain (FVTD) is also based on integral formulations of Maxwell's equations, but instead of the line integrals used in FIT one works with volume integral formulations,[8] which are popular from fluid dynamics.

FVTD works very naturally on an unstructured mesh like FEM. For the generation of unstructured meshes — which is the most essential burden of all methods working on such meshes — many commercial and free software packages are available from FEM and other codes, for example,

HyperMesh[18] and GMSH.[19] Thus, FVTD works on identical meshes as FEM, but it replaces the way how the electromagnetic field is handled. The main advantage of FVTD compared with standard time domain FEM is that standard FEM works with a large sparse matrix equation that must be solved for each time step, whereas FVTD works with a much faster update scheme (similar to the FDTD scheme), where only the local field and the field of the neighbor elements is involved.

Since FVTD works on unstructured meshes, the discretization may be well adapted to small features, which is essential for plasmonics. In order to avoid very short time steps (because of stability issues similar to FDTD), local time stepping should be implemented,[8] i.e., shorter time steps are applied to areas with finer discretization.

The most crucial problem of FVTD is that its update scheme seems to be only stable when it is dissipative, i.e., when some energy is lost. This is also a fundamental difference to FDTD, which is known to provide phase errors.

In principle, FVTD may be considered as a first order DG-FEM. Since DG-FEM also works with higher order schemes, which are very promising for plasmonics, the research focus has shifted from FVTD to DG-FEM. To our knowledge, no commercial FVTD solvers for electromagnetics are available.

2.7. *FEM and DG-FEM*

During the last two decades the Finite Element Method (FEM) in the time domain (TD) received considerable attention because of its potential to solve transient nonlinear problems on an irregular grid.

The FEM implementations in the TD have been developed in two essentially different ways. The first approach solves Maxwell's equations directly in leapfrog fashion. This method yields an explicit scheme for directly obtaining the field of the current time step based on the field of the previous steps without solving a large sparse system of equations in each time step.[9,20] The second approach has arisen from adapting well-developed frequency domain FEM implementations to the time domain. This method yields an implicit scheme for solving the second order vector wave equation obtained from Maxwell's equations in a classical FEM manner by solving a large sparse system of equations in each time step.[9,21]

The basic advantage of the explicit TD FEM schemes is the absence of a need for solving a large linear sparse system in each time step.

Its fundamental disadvantage is a problematic stability condition which requires small time steps and consequently long CPU time.[9] The implicit TD FEM schemes are usually absolutely stable but they require large RAM and long CPU time to store and solve a linear sparse system in each time step.[9]

To truncate the computational domain in the TD FEM modeling classical techniques, i.e., absorbing boundary conditions (ABC)[9] and perfectly matched layers (PML)[9,22] are available.

In order to make the implementation of interfaces between different materials and boundary conditions (for the truncation of the computational domain) simpler the Discontinuous Galerkin (DG) method was adopted for computational electromagnetics.[23] This method can be understood as a hybrid between FEM TD and FVTD, which incorporates many good features of both.

As opposed to standard FEM, the DG FEM equations are local to the generating element, and the corresponding solution of this element is not reconstructed by looking at its neighboring elements.[23]

3. Frequency Domain Solvers

Since frequency domain solvers may handle dispersive materials, eigenvalue problems, and periodic symmetries in a very natural way, applying frequency domain solvers to plasmonics is rather straightforward, but one should keep in mind that electromagnetic field solvers are often restricted in order to simplify their implementation. Therefore, there is no guarantee that a specific software package can efficiently solve certain problems, for example, periodic plasmonic waveguides, where dispersive materials, non-linear eigenvalue problems, and periodic symmetries are involved all together. Such a problem will be considered in detail below.

The time domain solvers that were outlined in the previous section discretize the entire domain where the field is to be computed by meshing or gridding procedures. In plasmonics, the natural space to be discretized is infinite. Since one cannot discretize infinite space entirely, special treatment is required. In most cases one truncates the space by means of special "fictitious" boundaries and imposes special absorbing boundary conditions (ABCs)[9] on them or one defines a layer of absorbing material along these boundaries. The absorbing layer may have hypothetical material properties that provide zero reflection in theory. Such layers are called perfectly matched layers (PMLs).[5,6,9,22] The aim of ABCs and PMLs is to avoid any

reflection of all possible outgoing waves that hit the fictitious boundaries. The implementation of ABCs and PMLs is particularly problematic when natural boundaries (interfaces between different materials) are assumed to extend to infinity.

This is obviously the case when waveguide structures are considered, which either have cylindrical or periodic symmetry. Such infinite structures are usually handled by a separate (analytical) treatment of the field dependence along the waveguide. In the cylindrical case one then only needs to discretize a 2D cross section. Within this cross section one may truncate the infinite 2D space using ABCs or PMLs. This is not really trivial because ABCs and PMLs may cause a non-linear, lossy eigenvalue problem which is much more difficult than linear, loss-free eigenvalue problems that may be efficiently handled, for example, by FEM. Therefore, there are many software packages that cannot efficiently handle open waveguide problems or cannot handle such problems at all.

The modeling of periodic waveguides is more complicated than the modeling of cylindrical ones, because a 3D "unit cell" that is one period thick must be discretized instead of the 2D cross section. Since this is a less common task, most of the software packages are not well prepared for computing periodic waveguide structures.

In reality, infinite waveguides do not exist, of course. Thus one is tempted to consider the modeling of infinite structures as academic, but waveguides are so long in most cases that they cannot be modeled by a brute force 3D discretization. Even for short waveguide sections, brute force computations become very time consuming as we will see in the following.

The situation becomes even more difficult, when a transition between two waveguide structures or a termination of a waveguide structure shall be modeled. This does not only require an appropriate modeling technique but also a good understanding of electromagnetic wave propagation. Such a structure will be considered at the end of this chapter. Before practical examples are studied, the most important frequency domain solvers shall be outlined.

3.1. *Finite differences*

The implementation of FD in frequency domain is very similar to FDTD and one usually derives a very similar update scheme for the field on a regular grid.[24] This requires a certain number of iterations for each

frequency until steady state is observed. The required number of iterations M_I depends on the maximum number of grid steps M_G in one direction and on the desired accuracy. Since M_I must be a multiple of M_G, it is typically of the same order as for solving the same problem in time domain. This is the reason why it is often stated in FDTD papers that this method is much more efficient than frequency domain techniques, because it delivers an entire frequency spectrum instead of a single frequency value within the same computation time. In fact, this argument is rather strong when comparing FDTD with FD in frequency domain but it is not valid when comparing FDTD with other frequency domain techniques. However, one may easily understand why FD is currently applied rarely in frequency domain.

3.2. *Finite elements*

FEM in the frequency domain (FD) has a much longer tradition than its TD version, which was outlined above. Therefore, it is not surprising that the FD version has a much more extensive base of algorithms and a bigger research community. Moreover, numerous commercial FD FEM solvers are available.

Modern FD FEM simulation tools for computational electromagnetics solve a second order vector wave equation derived from Maxwell's equations. To remove spurious solutions arising from inaccurate numerical treatment of material interfaces, edges, and corners, the solver is usually based on vector (Whitney) elements with vector shape functions.[9,25,26] The accuracy is improved (controlled) by employing higher order elements, i.e., elements with higher order shape functions (quadratic, cubic, etc.).

For a better representation of complicated real-life geometries, curved finite elements, i.e., elements with curved faces and edges were developed. Curved finite elements combined with higher order interpolation functions are a powerful tool of the modern design based on virtual prototypes, as they are capable of accurately solving geometrically very complex 3D problems with a relatively small number of elements within acceptable CPU time.

Since FD FEM is a domain method, the computational domain has to be truncated with an artificial boundary. To eliminate nonphysical reflection from the artificial boundary, the so-called absorbing boundary condition (ABC) is assigned to the domain truncation boundary. The general ABC developed for FD FEM was reported in Ref. (27). The basic drawback

of ABCs is that they can absorb perfectly only those waves that have perpendicular incidence to the boundary. To eliminate this problem the perfectly matched layer (PML) technique[28] has been adopted also by the FD FEM community and sometimes even combined with ABCs.[29]

The FD FEM can be applied equally well to propagation problems, scattering problems, and eigenvalue problems (waveguides, resonators, etc.). Since a large square sparse matrix is obtained after the FEM discretization of the vector wave equation, the corresponding eigenvalues and eigenvectors can be extracted directly without any need for defining a fictitious source or excitation.[9] This can be seen as an important advantage over the FDTD and FVTD.

The research related to the FD FEM is presently oriented towards efficient preconditioning methods for iterative solution of the large sparse linear systems with a special focus on parallel computing.[30]

3.3. *Method of moments*

The problem of truncating infinite space is especially important when the radiation of antennas is considered, because one then usually wants to know the far field pattern, i.e., the field distribution far away from the antenna. Discretizing a sufficiently large area would be inefficient.

A way out is based on the observation that the sources of the electromagnetic waves, i.e., charges and currents, are located on the antennas, i.e., on finite structures. Once the sources are known, the electromagnetic field may be computed everywhere in space from Coulomb and Ampère integrals or by means of Green's functions. Thus, one considers the sources as the primary field in the method of moments (MoM)[31] and discretizes only the areas where charges and currents are present. In practice antenna structures are often much bigger than the skin depth and one may often assume that an antenna is made of perfect electric conductors (PECs) and that the sources are surface charges and surface currents on the PEC surfaces. Thus, the discretization domain may be reduced to the 2D PEC surfaces, which is an important advantage. Because of this, one is tempted to claim that MoM is a boundary discretization technique like BEM (see below) or that BEM is a variant of MoM rather than an independent method. In fact, when only loss-free materials are present, it may be difficult to distinguish BEM from MoM. As soon as lossy materials are present, i.e., non-perfect conductors or non-perfect dielectrics, MoM codes must discretize the entire domains where currents flow, whereas BEM

codes discretize only the boundaries of these domains. Because of this, the entire domains containing lossy materials must be discretized when MoM is applied. This obviously increases the computational burden very strongly.

In order to fully understand the inefficiency of MoM for modeling lossy materials, one must know that MoM leads to dense matrix equations whereas FEM and other domain discretization methods lead to sparse matrices that may be solved much more quickly even when the number of unknowns or degrees of freedom is very high.

Since lossy materials are essential in plasmonics, it is reasonable to look for true boundary discretization methods that also work when lossy materials are present.

3.4. *BEM*

The boundary element method (BEM) is based on the numerical discretization of an integral equation (boundary integral form) obtained from Maxwell's equations. The integral equation can have different forms, depending on the field formulation used. It contains the unknown vector field simultaneously as a function under integral operators as well as a free term.[32]

BEM discretizes boundary integral equations and therefore requires only a surface mesh of a 3D problem. Thus, the number of unknowns of the BEM matrix is orders of magnitude smaller than for FEM, but the BEM matrix is dense and requires completely different numerical treatment than the sparse FEM matrix. To reduce the RAM required for solving large 3D problems, the fully populated BEM matrix must be compressed (for example, by using the fast multipole algorithm[33]), preconditioned, and iteratively solved.[34]

The basic advantage of BEM is that it does not require an artificial truncation boundary with adjoined ABC or PML. The main drawbacks are its difficult mathematical background, numerically demanding matrix assembly (singular integrals), large dense linear system of equations, and limitation to linear materials.[33,34]

3.5. *Semi-analytic boundary discretization*

In MoM, the physical sources of the field are discretized and their strengths are computed from boundary conditions such as zero tangential field

on a PEC surface. When the physical sources are discretized by simple point sources, singularities are obtained on the surfaces. There are various methods to handle such singularities, but finally one ends up with codes that are either inefficient or provide only accurate fields at some distance from the surfaces.

A very useful idea to overcome these difficulties is to move the sources of the field away from the boundaries, i.e., when one has two domains on both sides of the boundary, one moves the sources for modeling the field in domain 1 into domain 2 and vice versa. Since the electromagnetic field of a source places anywhere outside a certain domain does still fulfill the Maxwell's equations inside this domain, one approximates the field in each domain by a superposition of fields that are analytical solutions of the Maxwell's equations and originate from fictitious or auxiliary sources. This is called method of auxiliary sources (MAS),[35] method of fictitious sources,[36] or method of fundamental solutions.

It is clear that each auxiliary source simulates a certain charge and current distribution on the boundary. This distribution is strongly localized when the source is close to the surface and it becomes broader and smoother with increasing distance. The resulting MAS matrix becomes more and more ill-conditioned with increasing distance. At the same time, the errors decrease with distance as long as the high condition number does not significantly reduce the error of the matrix solver. Thus, setting auxiliary sources in a way that accurate results are obtained efficiently requires some experience. Furthermore, MAS codes usually work with a point matching technique or a collocation method. Then, an appropriate correlation between the locations of the sources and the matching points is important, which also requires some experience and influences the condition number of the MAS matrix.

When one considers, for example, a boundary with a circular section, the auxiliary sources move with increasing distance toward the center of curvature. When they reach the center of curvature, the condition number becomes infinite — no matter how the matching points are distributed. As a result, no matrix solver can solve the MAS matrix equation. This can be avoided by a simple trick: the cluster of auxiliary sources is replaced by a multipole expansion. For the field outside a circular or spherical domain, a single multipole expansion is sufficient. This corresponds to the well known analytic Mie solution.[37] For more complicated geometries, it is often reasonable to work with more than one multipole expansion per domain. This is the main concept of the multiple multipole program (MMP),[38]

which may be considered as a generalization of MAS or as a combination of MAS and the Mie approach. In fact, MMP also contains other types of expansions such as plane waves and linear combinations of various expansions, called connections. Furthermore, MMP replaces simple point matching techniques by generalized point matching and error minimization that leads to overdetermined systems of equations that can be solved accurately even when very high condition numbers are encountered. As a result, MMP is very flexible and versatile. The MMP results shown below were computed with the most recent MMP implementation in the OpenMaX software package,[39] which is an open source project based on MaX-1.[40]

4. Plasmonic Antennas

In principle plasmonic antennas act as well-known RF antennas. Therefore, all software designed for the analysis of RF antennas seems to be appropriate also for plasmonic antennas. Unfortunately, this is not the case at all, because RF antennas exhibit so small material loss that most of the RF antenna software assumes no loss or considers loss as minor disturbance. Furthermore, plasmonic antennas may resonate even when the structure is small compared with the wavelength — which may lead to quasi-static effects — and finally, one is often interested in an accurate nearfield calculation in plasmonics, which is not the case in the standard RF antenna design. Therefore, prominent MoM codes for the RF antenna design are not well prepared for the plasmonic antennas analysis. Since mesh refinement is highly desirable, we focus on FEM and MMP in the following.

Plasmonic antennas usually interact with molecules or quantum dots, i.e., very small structures that cannot be described by macroscopic Maxwell models. Within a Maxwell model, such small structures may be simplified as follows. Essentially, one has two cases: emission and absorption.

An emitting molecule or quantum dot can be modeled as a dipole source within the plasmonic antenna structure. This dipole then plays the role of the excitation of the antenna. Such an active, i.e., radiating antenna can be modeled by any software that allows one to embed dipole sources. In FEM, one may define a given electric field along the edge of a sufficiently small element in order to simulate an electric dipole source. Since dipoles are special cases of multipoles, dipole excitations in MMP are trivial.

In the case of a "passive" receiver antenna, one usually assumes that the structure is illuminated by a source that is placed far away and may be well approximated by a plane wave. This is also the case for RF antennas. Therefore, plane wave excitations are available for all reasonable Maxwell solvers. In optics one often works with focused beams that cannot be approximated by a single plane wave. Gaussian beams are frequently used in software packages for optics. It should be mentioned that Gaussian beams do not exactly fulfill Maxwell's equations. Therefore, Gaussian beams may limit the accuracy that may be obtained. MMP includes Gauss beams but it also provides complex origin multipole and Bessel expansions that can be very close to Gaussian beams and fulfill Maxwell's equations analytically. In fact, complex origin multipoles are simple extensions of standard multipoles that are highly valuable for the efficient modeling of axisymmetric structure as we will demonstrate in the last section.

An important issue in the passive analysis is that one often is interested in the behavior of the structure for various angles of incidence. This is a rather difficult problem for FDTD and similar codes. Methods leading to matrix equations with right hand sides that contain the excitation can be implemented in such a way that multiple right hand sides, i.e., several excitations can be handled simultaneously at not much higher cost than a single excitation. This is a highly valuable feature that is available in prominent MoM codes and also in MMP.

In fact, multiple right hand sides are also valuable for active antennas when one wants to study the impact of the location of the dipole source on the radiation pattern.

When passive antennas are studied, the absorbing molecule usually has a weak interaction with the antenna, i.e., the field at the position of the molecule is almost the same as when the molecule is not there at all. One then can simply evaluate the field at the location of the molecule without explicitly modeling the molecule. In principle, the molecule can absorb energy and emit it at another frequency. When working in frequency domain, this does not provide any problem because absorption and emission take place at different frequencies. Therefore, the two processes are well separated in the frame of Maxwell simulations.

In principle, antennas are always 3D structures. Since 3D modeling is expensive, one often encounters simplified 2D models in the literature "for getting a first idea". Unfortunately, the 2D models often provide results that are rather far away from full 3D simulations. This is demonstrated for

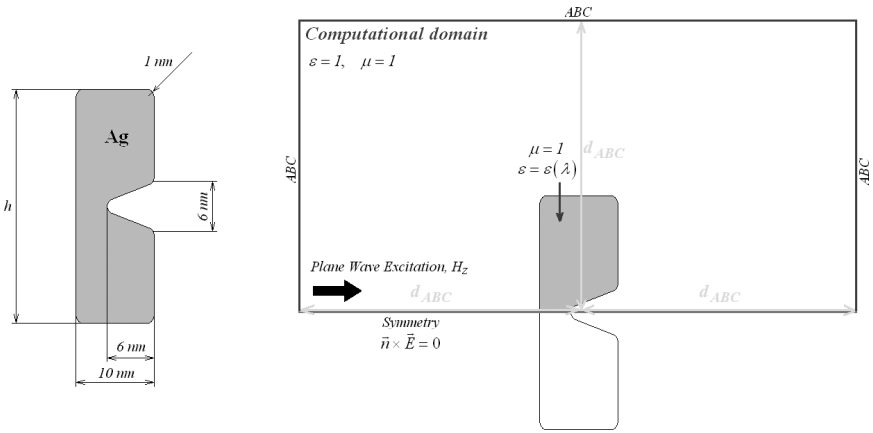


Fig. 1. Metallic V-groove antenna: Geometric arrangement (left) and FEM modeling (right). Due to the antennas symmetry only one half of the model is computed.

a small plasmonic antenna consisting of a metallic patch with a groove as shown in Fig. 1.

4.1. *Metallic patch antenna with small groove*

The geometry is presented in Fig. 1. For reasons of simplicity, a 2D model is considered, i.e., the antenna is assumed to be infinitely long in the direction perpendicular to the figure plane. The antenna is made of silver and its permittivity is modeled by using the Drude model

$$\varepsilon_{\mathrm{D}}(\omega) = \varepsilon_{\infty} - \frac{\omega_{\mathrm{p}}^2}{\omega(\omega + i\gamma)},\ \varepsilon_{\infty} = 5,\ \omega_{\mathrm{p}} = 1.4433 \cdot 10^{16},\ \gamma = 10^{14}$$

that fits very well with the measured permittivity of silver within the wavelength range of 350–600 nm.

The FD FEM model presented in Fig. 1 is rather simple and consists of the silver domain of the antenna surrounded by air. The air box is terminated by first order ABCs.[9]

For a given wavelength the structure is illuminated by a plane wave and the scattered field is computed. To find the antenna's resonances, i.e., the wavelengths with high level of the electric field enhancement in the groove, the field at a certain position in the groove is recorded for each wavelength within the range of 350 nm–600 nm. The peaks of this function correspond to the resonances.

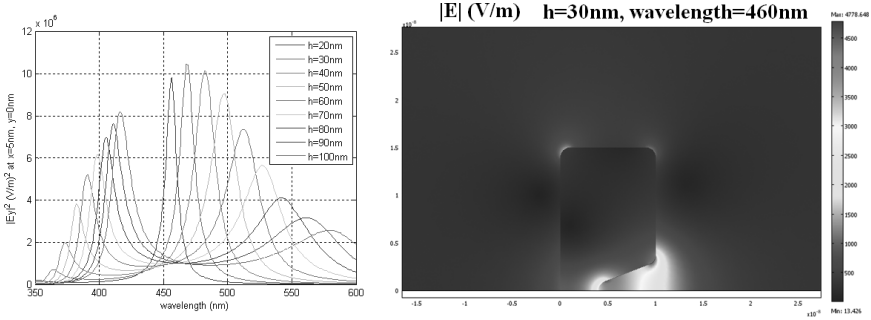


Fig. 2. Metallic V-groove antenna: Resonance curves (left) and electric field plot at resonance (right). The resonance curves of the antennas with different height (h) were computed by using the FEM scattering approach with a plane wave as an excitation.

The results of the computation of the antennas with different heights are given in Fig. 1. Evidently, each curve has two resonant peaks in the wavelength range of interest and these peaks are logically shifted towards longer wavelengths for larger antennas.

The resonant field of the antenna with height $h = 30$ nm at a wavelength of 460 nm is presented in Fig. 2 (right). The field is confined to the narrow volume of the groove due to the surface plasmon effect with a very strong enhancement. This behavior is not observable at lower frequencies where a field enhancement is observed only near convex edges and corners of metallic bodies.

To achieve a high level of accuracy the FEM mesh has to be very fine over the silver-to-air interface.[23] The element edge size was set there to 0.5 nm.

The 2D model of the optical V-groove antenna is assumed to be infinitely long. However, every structure has its finite length. In order to show the importance of 3D modeling, the antenna with $h = 30$ nm and length $L = 50$ nm is also simulated by a procedure similar to the 2D analysis. The results are presented in Fig. 3. Evidently, the 3D effects reduce the field enhancement factor at the most pronounced resonance and the resonance wavelength is shifted. Additionally, the weaker resonance at 360 nm is radically changed into two weak resonance peaks.

4.2. *Axisymmetric antenna structures*

In many cases, it is much more reasonable to assume rotational symmetry, i.e., to consider axisymmetric simplifications of antenna structures, which

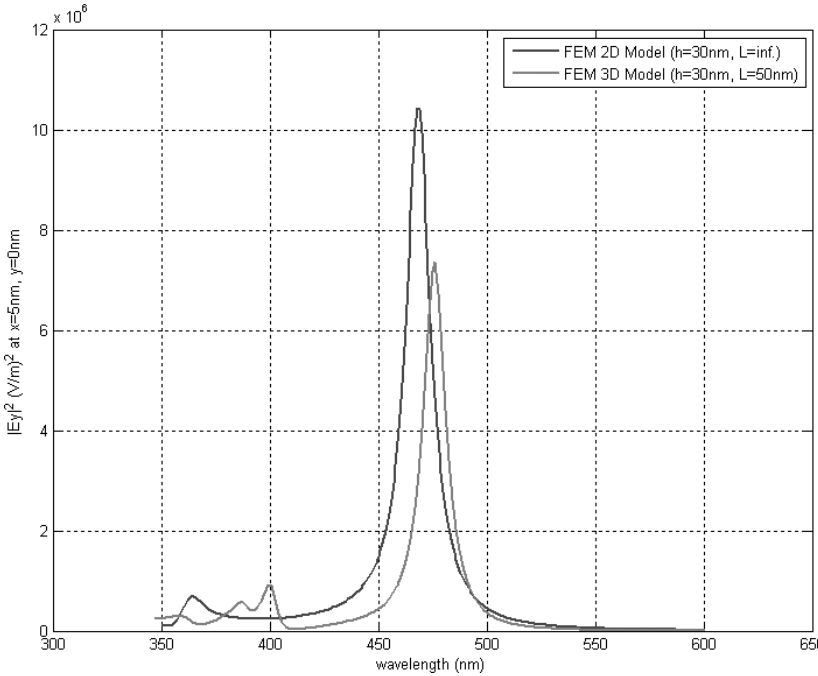


Fig. 3. Metallic V-groove antenna: The results of the 2D antenna (infinite length) and 3D antenna (50 nm long) are compared. Evidently, the 3D effects significantly change the behavior of the antenna.

are also called bodies of revolution (BOR). In this case, one only needs to model a section of the structure through its axis and one may separate the angular dependence of the field around the axis. Thus, one obtains a 2D model as well. The main problem is that the angular dependence depends on the excitation of the antenna. For a dipole excitation on the axis, polarized along the axis, the radiated field is axisymmetric, i.e., its angular dependence $f(\varphi)$ is constant and may be set equal to 1. When the dipole excitation is still on the axis but perpendicular to it, one obtains $f(\varphi) = A\cos(\varphi) + B\sin(\varphi)$ with coefficients A and B depending on the direction of the dipole and on the field component that is considered. The same holds for a polarized plane wave propagating along the axis. For a general excitation, a symmetry decomposition is required, which is essentially a Fourier decomposition in φ with $\cos(n\varphi)$ and $\sin(n\varphi)$ dependences for the order n terms. One must perform separate 2D computations for each Fourier order, which may be cumbersome. Automatic symmetry decompositions

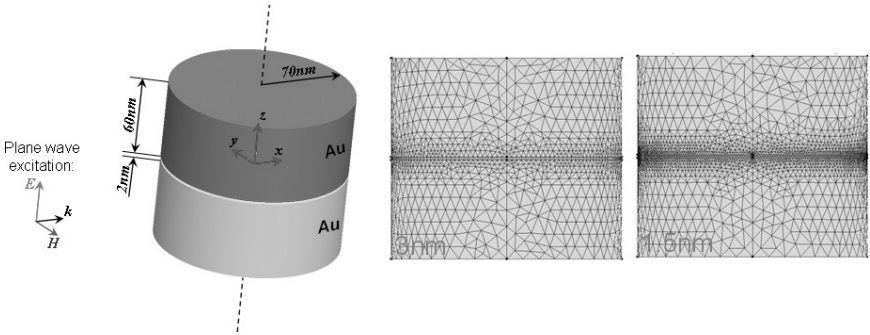


Fig. 4. The geometry of the antenna (left) and the impact of the element size in the gap to the overall mesh (right). If the element edge size in the gap is set to 3 nm the number of generated elements was 507,000 and the corresponding RAM and CPU time were 9 GB and 6 minutes, respectively. However, for a good accuracy the element edge size should be at least 1.5 nm which means 970,000 elements, 17 GB of RAM, and 16 minutes of CPU time (4 CPU cores: Intel Xeon 2.5 GHz).

for axisymmetric structures are only available rarely, for example, in JCMwave.[41]

To illustrate this, an axisymmetric plasmonic antenna consisting of a pair of metallic disks is considered. This structure with an academically thin gap of 2 nm was used for a FEM simulation. The reason for the ultra thin gap was essentially checking the software because the numerical effort increases very much with decreasing gap.

The geometry of the antenna with two gold disks and the corresponding FEM mesh are shown in Fig. 4. Evidently, the impact of the element edge size in the gap to the mesh size, required RAM memory, and CPU time is enormous.

Since our simulation was performed in the frequency domain, we had to sweep the entire frequency range. The total CPU time was around 53 hours (4 CPU cores: Intel Xeon 2.5 GHz).

The current MMP implementation in OpenMaX does not allow one to take the rotational symmetry into account because it is broken by the plane wave that is incident from the side. Therefore, a full 3D modeling is required and only two symmetry planes may be taken into account as in the FEM model. The main difference is that MMP does not define PEC and PMC planes as FEM and FD and other codes. It symmetrizes the expansions, which is numerically more efficient for boundary discretization methods. In order to avoid field singularities near the edges, the corners are rounded with 5 nm radius of curvature, which is also realistic from the practical

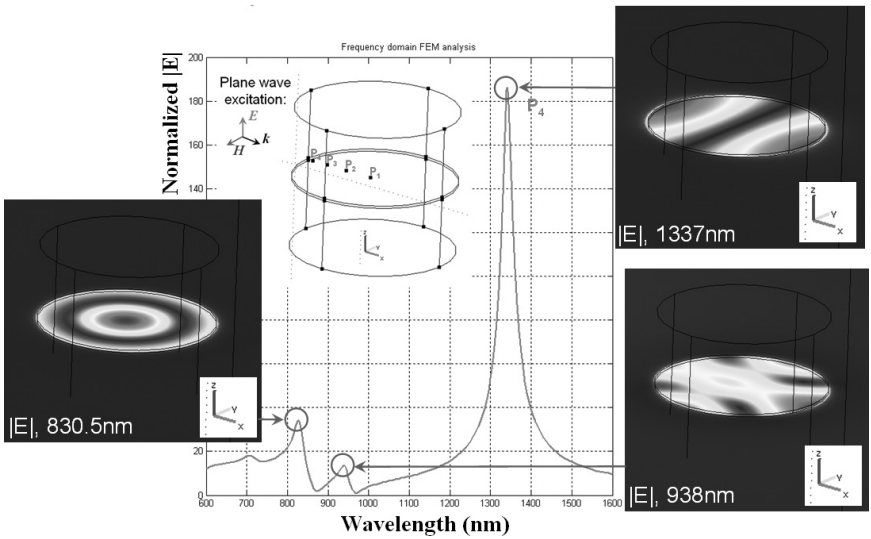


Fig. 5. Resonant search of the optical antenna with two golden disks. Evidently, three different resonances were found. The corresponding electric field patterns in the gap are depicted.

point of view. Figure 6 shows the electric field intensity on the surface of half a disk and in a cut plane. From this, one may clearly see that the first two modes shown in Fig. 5 are excited by the incident plane wave at the considered wavelength of 613 nm.

The structure consisting of two plasmonic disks with a bit bigger gap becomes interesting when quantum dots are included in the gap. In this case, an incident wave provides energy to the quantum dots at some wavelength and the quantum dots emit energy at some longer wavelength. For modeling the emission, the radiating quantum dot may be considered as a dipole. When it is located in the center of the gap, i.e., on the axis of the rotational symmetry, one has two main orientations: perpendicular or parallel to the axis. In both cases, the exciting dipole field is symmetry adapted, i.e., the angular dependence is either $\cos(\varphi)$ (perpendicular) or 1 (parallel). Then, the scattered field also must have the same angular dependence, which allows only one section of the structure, i.e., one essentially has a 2D model rather than a 3D one.

The dipole parallel to the axis couples very well to the axisymmetric mode shown in Fig. 5 and provides an axisymmetric radiation pattern as shown in Fig. 7 for a pair of disks with 10 nm gap and 30 nm radius of

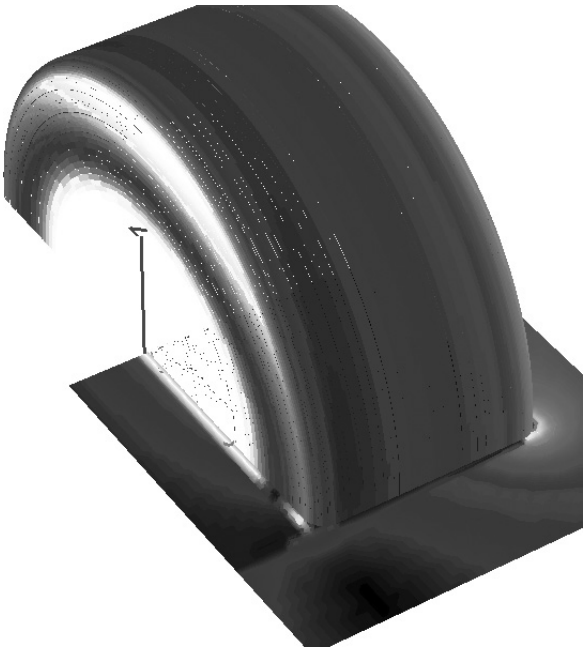


Fig. 6. Passive axisymmetric antenna: Pair of gold disks with 2 nm gap in between, electric field intensity on the surface of one disk and in a cross section plane. Plane wave illumination with electric field parallel to the symmetry axis. Note that the field is not axisymmetric.

curvature. The simulation of this structure is much faster than for the structure shown in Fig. 6 for two reasons: (1) rotational symmetry may be accounted for without any Fourier decomposition in the φ direction, (2) larger radius of curvature produces a much smoother field pattern. Therefore, Fig. 7 could be computed with only 303 unknowns, i.e., a very small MMP matrix, whereas 3347 unknowns were required for computing the model shown in Fig. 6.

4.3. *Chains of plasmonic particles*

We now study a more complicated plasmonic structure that may be considered as a bridge from isolated antennas — made of single or a few coupled discrete elements — to infinitely long, periodic or cylindrical waveguides. Quite a while ago,[42] resonantly coupled plasmonic particles have been recognized for their potential as long distance devices for optical energy transporting. Their subwavelength cross-sectional

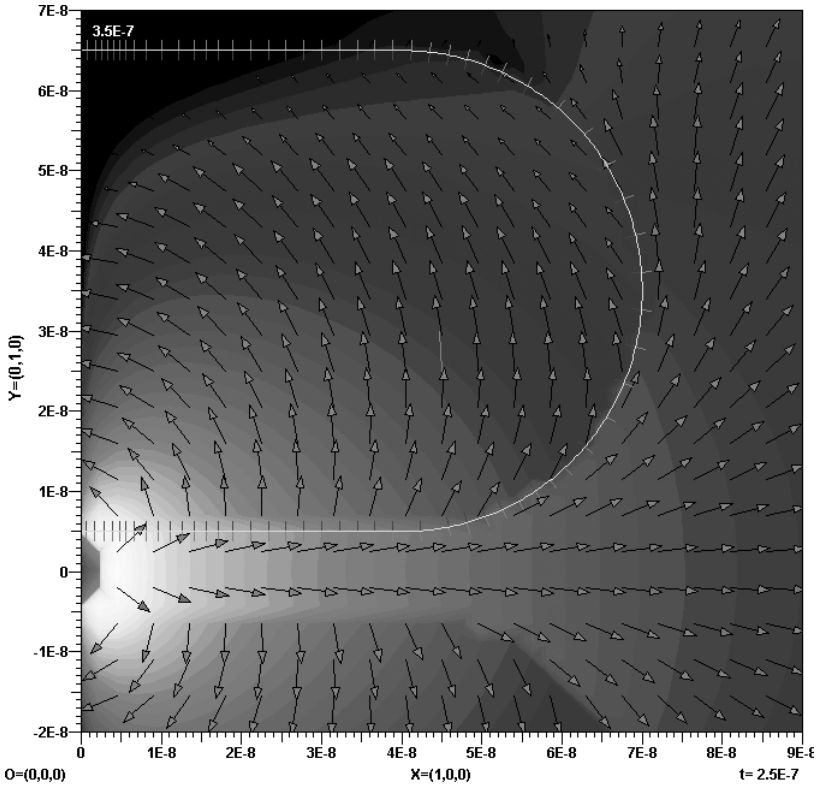


Fig. 7. Active axisymmetric antenna: Pair of gold discs with 10 nm gap, illuminated by a vertical electric dipole in the center of the gap at 350 nm. The field is axisymmetric. The time average of the Poynting vector field is shown in a section. The field in the area close to the dipole source becomes very high. In order to avoid undesired scaling, it is set to a much smaller value around the dipole.

dimensions promise delivery (or pick-up) of radiation below the diffraction limit.[43–46]

Arguably, the simplest conceivable case is that of a chain of plasmonic spheres in a homogeneous, isotropic medium. The successful ansatz for the full solution consists of a combination of Mie solutions for each sphere, with maximum order and degree suitably chosen according to wavelengths, diameter and interparticle gap size, as well as the desired degree of accuracy. It can be solved exactly, for example, by the T-matrix approach.[47] Even simplified versions like the coupled dipole method or semi-analytical solutions prove extremely helpful — once they have been

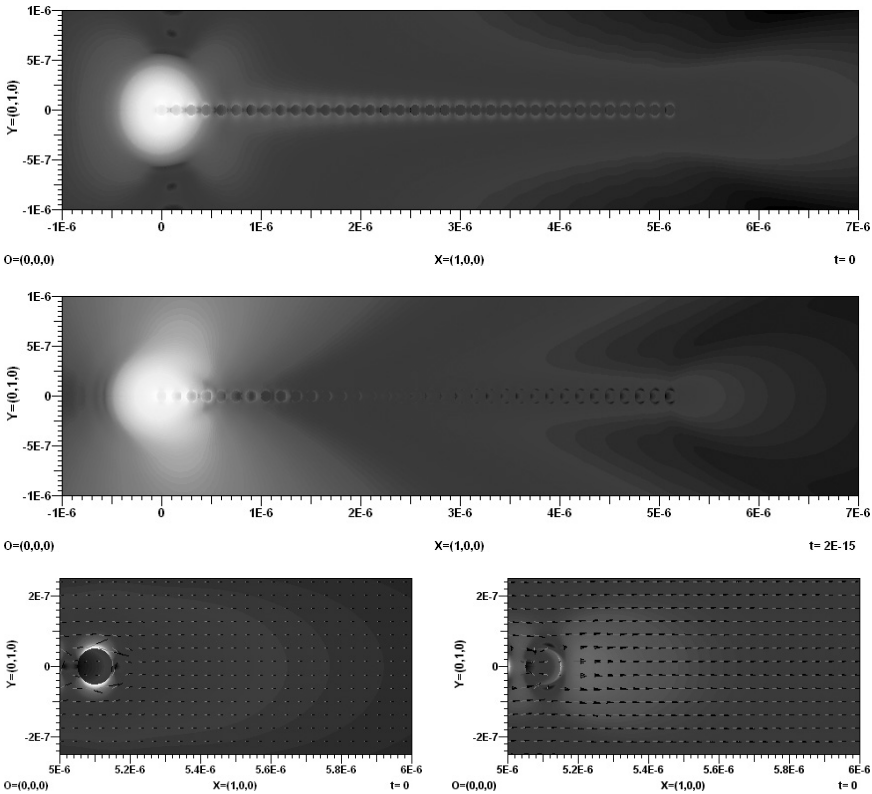


Fig. 8. MMP simulations of a chain of 35 Au spheres of 100 nm diameter with a period of 150 nm, in air. A 555 nm excitation field is modeled with a Bessel expansion at a complex origin chosen such that it closely resembles a Gaussian beam tightly focused on the left-most sphere. Propagation direction is along z, i.e., into the displayed plane. Shown are the magnitudes of the time averaged Poynting field. The top panel displays (logarithmic color scale) an overview of the transverse waveguide mode, excited by a beam polarized mainly along the y-axis. The middle panel shows the same for the longitudinal mode, excited with polarization along the x-axis. The bottom panels display (in linear color scale) the area to the right end of the chain for the transverse (left) and longitudinal (right) mode. As the arrows indicate, energy is radiated into free space.

validated against some benchmark method — especially for evaluating extremely long ($>10^5$ particles) but finite chains.[48,49]

In view of possible structural extensions, we use MMP (OpenMaX[39]) with the same ansatz, i.e., a Bessel expansion for the interior and a multipole expansion for the exterior of each sphere. Figure 8 shows an approximately 5 μm long chain of Au spheres, 100 nm in diameter with a period of 150 nm. Upon linearly polarized focused excitation of the left end of the chain,

optical energy is transported along the chain by transverse or longitudinal waveguide modes. A zoom-in onto the right reveals how the chain end acts like an antenna emitting radiation into free space. Indeed, the excitation wavelength of 555 nm used here was determined by optimizing the emerging energy flow out of the chain at the location of the first unoccupied chain period.

This illustrates the basic properties of periodic plasmonic waveguides constructed from individual antennas. In this sense, one may consider them as the consequent extension of coupled antenna arrays into long periodic structures. The use of spheres is owed mainly to the relative computational ease. In practice, any kind of building block may be employed. For example, RF dipole antennas have been considered for the microwave regime.[50] Besides particle shape, the periodic spacing and relative orientation of successive constituents are obvious options to improve the performance of plasmonics chains.

With an eye on eventual device integration, the crucial characteristics of waveguides are the losses that are incurred during in- and out-coupling as well as propagation. The latter is described by the damping constant, which is straightforward to evaluate numerically from simulations of sufficiently long structures. Typical values for 3 dB loss distances range between tens of nm (off-resonance) and a few microns, when in resonance. From Fig. 8, the transverse and longitudinal modes are estimated as $3\,\mathrm{dB}/2.8\,\mu\mathrm{m}$ and $3\,\mathrm{dB}/1.8\,\mu\mathrm{m}$, respectively.

One way to reduce losses is to alter the particle shape.[51] Here, we discuss a different approach illustrating some advanced MMP simulation techniques along the way: We consider enclosing the entire plasmonic chain in a cylindrical dielectric, vaguely reminiscent of peas in a pod. Surrounding each plasmonic particle individually with its own dielectric shell has been studied before, specifically in the context of chemical self-assembly of plasmonic clusters from suspension, where the dielectric shell plays a critical role in forming stable connections.[52,53] Besides obvious mechanical advantages, a dielectric pod that holds all particles in one container also provides a new means to manipulate the fields outside the plasmonic particles in a way more akin to dielectric waveguides.

Figure 9(a) shows the principal structure. A finite number of plasmonic particles form a periodic chain that is enclosed by a rod-like dielectric pod. To model this kind of structure within the MMP framework, we need additional expansions to appropriately describe the fields that meet at the new interface. Here, for a large range of parameters, it turns out

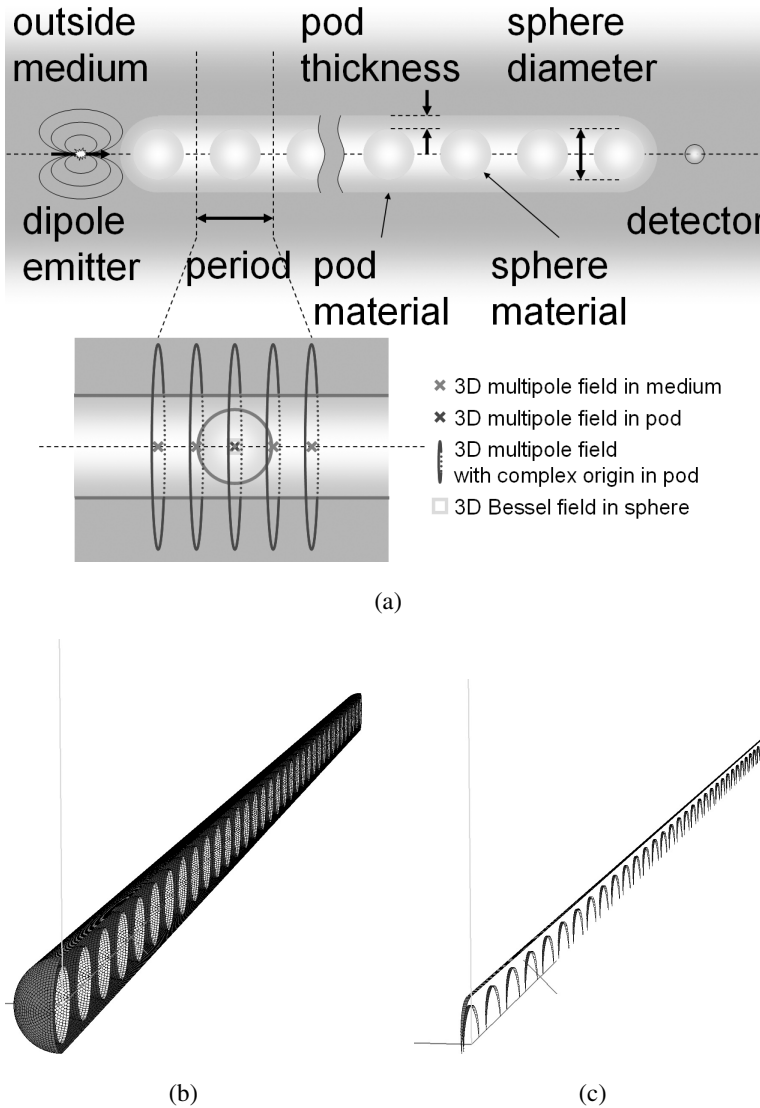


Fig. 9. Modeling plasmonic chains in OpenMaX.[39] (a) The structure of plasmonic chains embedded in a pod-like dielectric medium. Adjustable geometry parameters and the placement of the multiple multipole expansions required to represent the partial fields of one periodic cell are indicated. Considering a dipole emitter alleviates the numerical burden considerably through the use of symmetry decomposition. The location of the "detector" represents an output port of this model. (b,c) Comparison of surface meshings for full models (b), required when non-axisymmetric excitations are used, as in Fig. 8, and (c) the quasi-one-dimensional irreducible slice of axisymmetric models.

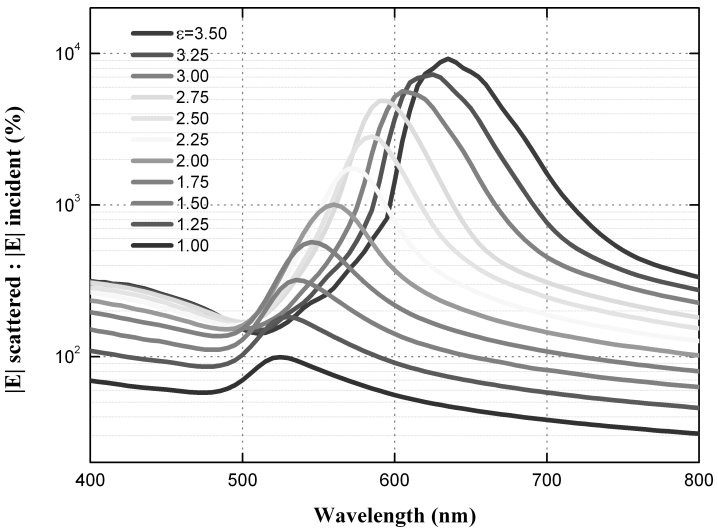
to be sufficient to include four multipole expansions per periodic cell on the central axis to represent the fields scattered into the outside medium. And another four expansions of ring-like form [either a discrete set of point multipoles distributed on a ring, or a single multipole expansion with suitable complex-origin such that its ring-like singularity encircles the waveguide; see Ref. (39)] represent the fields inside the pod.

In total, therefore we have 10 expansions per periodic cell — which can be set programmatically — and a similar number of manually placed expansions at the ends. The $5\,\mu$m long structure discussed below, for example, requires solving for a total of 3086 unknowns. This approach to modeling each periodic cell in an identical manner leads to a rather stable solution matrix, whose condition number hardly depends on the length of the plasmonic chain. As long as it fits in the computer's memory, an accurate brute-force solution is obtained within minutes on current desktop hardware.

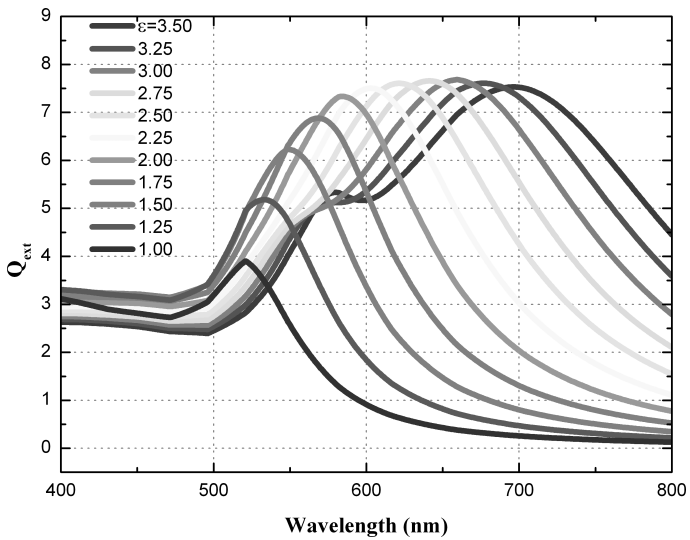
An alternative ansatz, suitable in principle for arbitrary lengths of chains, is to artificially consider only the input and output ends of the chain explicitly and model the inner part with the infinite waveguide modes. Such strategies are discussed in Section 6.

For the present case of a finite chain waveguide, we consider now a dipole emitter as excitation for two reasons. First, the model geometry becomes axisymmetric. With a possible phase factor of $\exp(\mathrm{i}m\phi)$ in the fields, symmetry decomposition lifts much of the computational burden,[54] in contrast to the perpendicular excitation with focused radiation. As indicated in Fig. 9(c), the irreducible element is just a thin slice along the optical axis. The second motivation for using a dipole emitter comes again from possible device applications. Such sub-diffraction-limit waveguide structures are most desired in delivery (or pick-up) of optical energy to (or from) nanoscopic structures such as molecules or quantum dots. Coupling to macroscopic fields, as demonstrated in Fig. 8, is of less interest, as this can be achieved with mature conventional technology. Therefore, we consider the performance of the hybrid waveguide composed of a plasmonic particle chain inside a dielectric pod as coupler between two point ports: a dipole emitter on the "input" side (located at the first empty position to the left of the periodic chain) and a point-like detector for the fields on the output side (located at the first empty position to the right).

Depending on wavelength and waveguide structure, a more or less significant portion of the emitted optical energy is coupled into the waveguide and carried over the distance of $5\,\mu$m to the detector port. Figure 10(a) shows the scattered E-field at the output port relative to the



(a)



(b)

Fig. 10. (a) Transport efficiency of a plasmonic chain waveguide of 35 Au spheres, embedded in a dielectric pod as presented in Fig. 9. The scattered electric field magnitude transported by a longitudinal mode to the "output" port of the waveguide is shown. It is normalized by the field magnitude at the same location in absence of any waveguide. For each curve, a non-dispersive, positive dielectric constant is assumed. (b) Extinction efficiency of a single 100 nm diameter gold sphere in homogeneous media of different dielectric constants.

E-field at the same port in absence of any waveguide in the visible spectrum for dielectric pods of variable dielectric constant. Two aspects are of main interest: first, how does such a waveguide perform optimally, or, in other words, how much energy is transported at what peak wavelength? The second question is whether it is possible to relate, at least qualitatively, these waveguide resonances to resonances of its constituent particles?

The case $\varepsilon = 1$ corresponds to the same waveguide as discussed above, namely isolated gold spheres in air. Not surprisingly, a strong similarity can be drawn to the extinction spectrum of a single Mie-scattering gold sphere in air, as shown in Fig. 10(b). This similarity, however, diminishes, as the index of refraction is increased towards realistic values of candidate pod materials, such as glasses ($\varepsilon = 2.0 - 2.3$), sapphire ($\varepsilon = 3.0 - 3.4$), zinc oxide ($\varepsilon = 3.8 - 4.4$), or rutile ($\varepsilon = 6.1 - 6.3$). The apparent red-shift of the spectra is slightly less pronounced for the chain than for isolated gold spheres. Qualitatively, this may be related to an "effective medium" interpretation, according to which the average dielectric constant surrounding the structures is much lower for the thinly encased chain. Thus it appears the extinction resonances provide a qualitative understanding of the waveguide peak wavelength — at least, when the inter-particle distances are sufficiently large to exclude strong dimer interactions between neighboring particles.

This is not strictly the case, however, as Fig. 11 shows. Here the period of an otherwise unaltered waveguide is varied. Clearly, the peak power transmission wavelength red-shifts with decreasing period. As the distance between particles becomes smaller, their interaction strength becomes larger, resulting in the observed red-shift for the waveguide mode.

Besides the spectral behavior, the transmission efficiency is of utmost importance. As we can conclude from the results displayed in Figs. 10–11, this varies by orders of magnitude, which in general is a welcome feature. It promises high selectivity and considerable opportunity for fine tuning in the design. To understand the relative contributions of insertion and propagation losses, we look at Fig. 12, which displays the transmission of dipole radiation by a 150 nm period waveguide enclosed in a sapphire pod. Evaluating the fields just outside the waveguide along its length, we extract 3 dB loss distances of 1.1 μm and 0.7 μm for the longitudinal and transverse mode, respectively. As these are quite comparable to the values obtained for bare plasmonic particle chains, we conclude that the propagation losses are hardly affected by the presence of a dielectric pod enclosure.

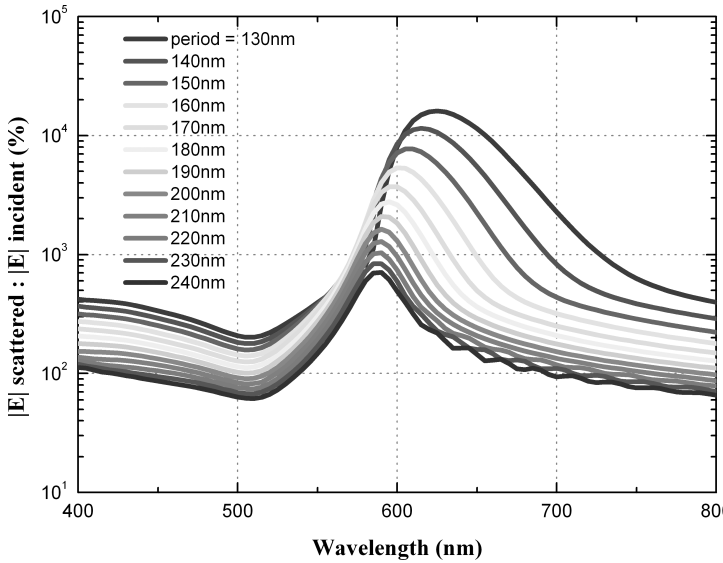


Fig. 11. Transport efficiency of a plasmonic chain waveguide of 35 Au spheres, embedded in a sapphire pod, as in Fig. 10(a). Here, the period is varied (and hence also the total length). The dispersion of sapphire is taken into account.

Nevertheless, as a function of the pod material dielectric constant, we find field enhancements of two orders of magnitude between $\varepsilon = 1.0$ and 3.5 [Fig. 10(a)]. This is only partly explained by a change in the behavior of isolated spheres subject to a varying background medium — the peak extinction of isolated spheres hardly changes by a factor of two [Fig. 10(b)]. One arrives at the conclusion that the insertion and extraction losses are dramatically altered by the dielectric pod.

5. Plasmonic Waveguides

It is clear that the brute force modeling of long structures, such as long metallic cylinders or long arrays of plasmonic particles becomes more inefficient the longer the structure is. Therefore, it is reasonable to assume infinite length and consider cylindrical or periodical symmetry along the axis. The symmetry decomposition then is similar to the symmetry decomposition of axisymmetric structures. For cylindrical symmetry along the z axis, one may separate the z dependence $\exp(i\gamma z)$, where γ is the propagation constant that is complex in the case of lossy waveguides. It is

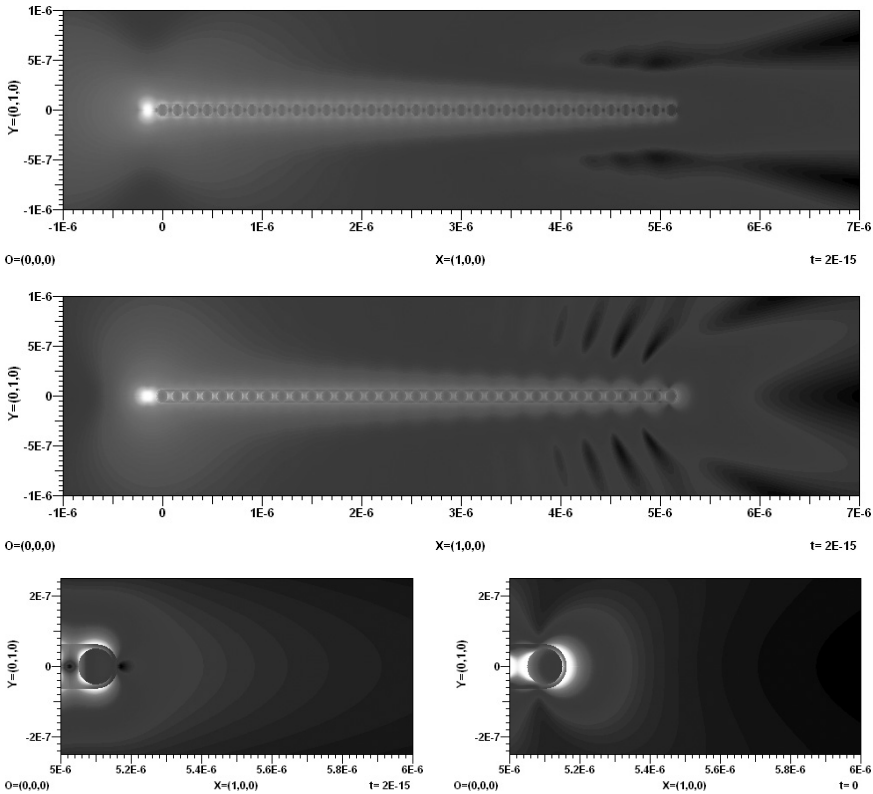


Fig. 12. MMP simulations of a 5 μm long plasmonic/dielectric hybrid chain waveguide similar to that displayed in Fig. 8, but with a sapphire pod enclosure. The outside medium is air. Near-resonant, dipolar excitation of 617 nm is modeled with a point-like emitter to the left of the waveguide. The magnitudes of the E-field are shown. The top/middle panel displays (in logarithmic color scale) an overview of the transverse/longitudinal waveguide mode, excited by a dipole along the y/x-axis, respectively. The bottom panels display (in linear color scale) the area to the right end of the chain for the transverse (left) and longitudinal (right) mode.

important to note that this separation is natural and relatively simple in frequency domain only.

It was already mentioned in the introduction that the x dependence of the field can be separated very similarly in the case of a periodic waveguide that is periodic in x direction, i.e., one can write $\exp(i \cdot Cx \cdot Dx)$ and obtains Cx instead of γ.

In principle, both γ and Cx are frequency dependent. Guided modes only exist within certain frequency ranges for certain values of $\gamma(\omega)$ or

$Cx(\omega)$, where ω is the angular frequency. In the case of well-known loss-free waveguides made of materials without dispersion, the propagation constant becomes real above a certain cutoff frequency ω_c. If so, the mode is considered to be strictly guided and it propagates without any damping. Below cutoff frequency, the mode is not propagating. It then is usually not considered at all because the field decays along the axis of the waveguide. In the case of open waveguide structures, one may not only have guided and evanescent modes but also radiating modes. The latter do not fulfill the Sommerfeld radiation condition that is assumed to be valid in many software packages. Furthermore, radiating modes are always characterized by a complex propagation constant that makes the evaluation of such modes much more cumbersome.[55]

Since plasmonic structures are rather lossy, it is important to compute complex propagation constants anyway. This considerably reduces the number of useful Maxwell solvers. Furthermore, the strong dispersion of metals also does not allow one to benefit from efficient linear eigenvalue solvers. In Ref. (55) various procedures for handling such complex eigenvalue problems were presented. In the following, the complex eigenvalue is presented for four examples of increasing complexity.

5.1. *Axisymmetric, cylindrical waveguide*

Waveguides such as optical fibers and circular wires that are not only cylindrical but also axisymmetric are relatively easy. The symmetry decomposition in angular direction allows one to classify modes depending on their symmetry properties, i.e., one can use the symmetry number m of the Fourier decomposition — $\cos(m\varphi)$ and $\sin(m\varphi)$ — to group all possible modes into classes with a specific symmetry number. For optical fibers it is well known that fundamental, polarized modes with symmetry number $m = 1$ that may propagate at arbitrarily low frequencies, whereas metallic wires have fundamental Sommerfeld waves,[56] which are axisymmetric and have the symmetry number $m = 0$. Sommerfeld waves are well known from RF applications, where the conductivity of metals is very high. At optical frequencies, the conductivity of metals is low and their permittivity is often negative. Then plasmon-polariton waves may propagate. The analytic modeling of them is exactly the same as for Sommerfeld waves, i.e., the field inside the wire is modeled by a 2D Bessel expansion and the field outside the wire is modeled by a 2D multipole expansion.[57,58] The order of these expansions is equal to the symmetry number m. For $m = 0$ one obtains a

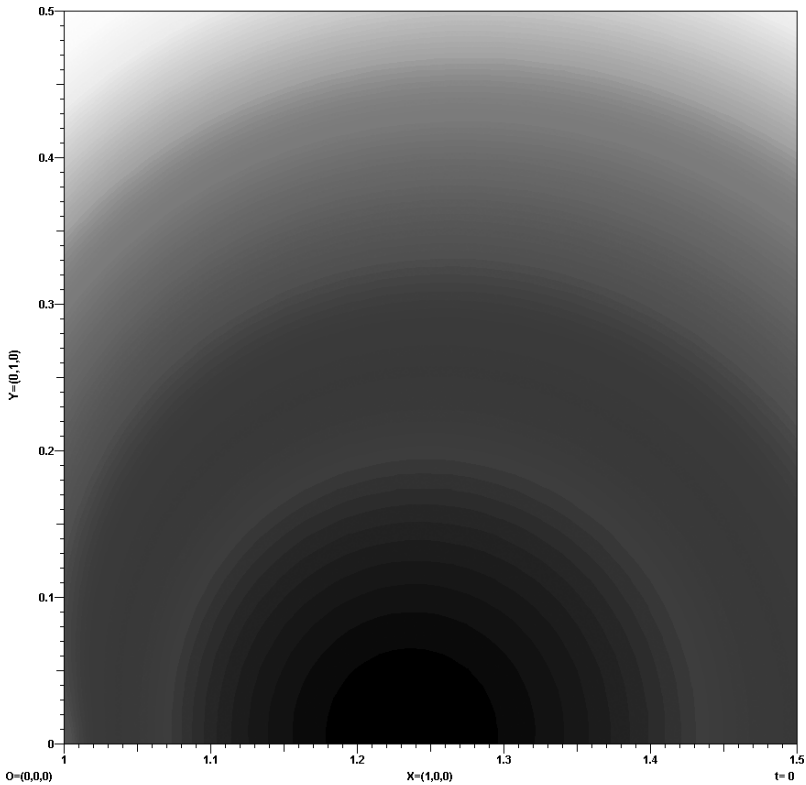


Fig. 13. MMP eigenvalue search function at $\lambda = 600\,\text{nm}$ for the circular silver wire with 100 nm diameter in the complex γ/k_0 plane. The minimum of the search function determines the location of the mode.

very simple 2 by 2 matrix equations but this finally leads to a transcendental equation that must be solved numerically for finding the eigenvalue, i.e., the propagation constant for a given frequency. Since the propagation constant is complex, this is really not easy but it can be done with appropriate codes.

Figure 13 shows the MMP eigenvalue search function in the complex γ/k_0 plane at 600 nm. Measured data for silver[59] were use. Within the observed area $1 < \text{Re}(\gamma/k_0) < 1.5$, $0 < \text{Im}(\gamma/k_0) < 0.5$ only one mode is present. The search function is very smooth and the minimum is located near the real axis. For computing the search function, only a 2×2 matrix equation must be solved. Therefore, the minimum may be found within a few iterations very rapidly. Since MMP offers Bessel expansions and

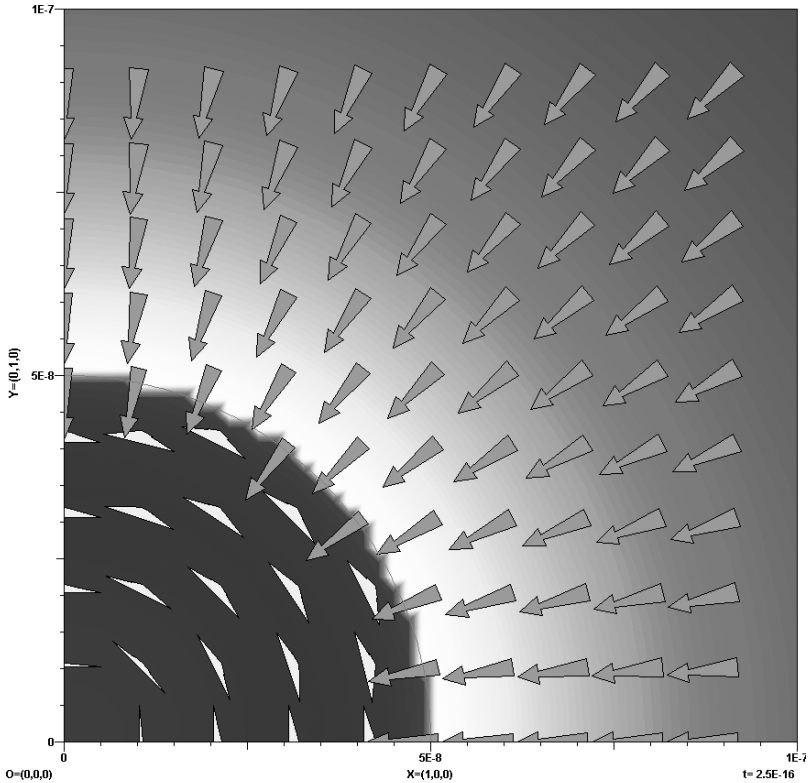


Fig. 14. Electric field of the fundamental silver wire mode. The baselines of the arrows indicate the strength of the longitudinal component, which is perpendicular to the plane. Inside the wire, the longitudinal component is dominant.

multipole expansions, the MMP model is very close to the analytical one — only the eigenvalue search procedure is different.

Figure 14 shows that the field corresponding to this mode is strongly concentrated near the surface of the wire. In order to couple it with other wires — for instance, to construct a simple double-wire transmission line — the distance between the wires must be very small. This is different from Sommerfeld wire modes at RF that exhibit a field that is much less localized.

Finding higher order modes or modes with higher symmetry numbers is not difficult but time consuming. Finally, only the fundamental wire mode propagates with reasonably low damping. Thus, this is the only mode of practical interest.

5.2. *Non-axisymmetric, cylindrical waveguide*

When considering a combination of a dielectric fiber with metallic wires, one can recognize that a fundamental mode of two symmetric wires has a polarized field similar to the fundamental HE11 mode of optical fibers. Therefore similar modes should exist for fibers with two symmetric metal patches as shown in Fig. 15. The analysis of this structure is very demanding because it supports many different modes with a complicated behavior. Furthermore, triple points are obtained where the metal patches, the fiber core, and the surrounding medium come together. Triple points may cause severe numerical problems and require either a very fine local discretization or a cumbersome special treatment.

It should also be mentioned that the rotational symmetry is now broken. Therefore, it is no longer possible to categorize the existing modes. Since there are still two symmetry planes (xz and yz planes) one can build four symmetry adapted classes of modes. The most interesting ones are those similar to the HE11 fiber mode with electric field perpendicular to the yz plane and parallel to the xz plane. Because of the difficulties caused by the triple points, these modes are analyzed with two completely different methods, FEM and MMP.

The FEM model of the cladded optical fiber is shown in Fig. 15. A high level of accuracy can be achieved only if the ABCs are set far enough from the fiber. Thus, a relative large air domain around the fiber has to be meshed.

Since MMP is a boundary discretization technique, the surrounding space must not be truncated by an ABC or PML, which provides a significant advantage for open waveguide structures. The main numerical problems are then caused by the triple points where the dielectric core, the metal cladding and the surrounding air come together. In these triple points the field becomes singular and this drastically reduces the local convergence. Therefore, a very fine discretization is required for all numerical methods that do not provide special procedures for triple point singularities — that also depend on the mode that is considered as one may see from Fig. 16.

MMP provides several eigenvalue search functions that may be plotted in the complex eigenvalue plane, which is the plane of the complex propagation constant in the case of a cylindrical waveguide. Once a mode is detected as a local minimum of the eigenvalue search function, it may be efficiently traced with smart routines for increasing or decreasing wavelengths in the complex eigenvalue plane. A complicated shape of the

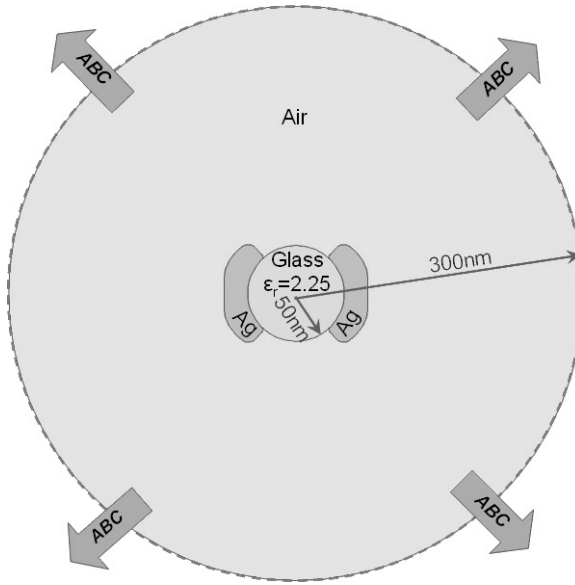


Fig. 15. The FEM eigenvalue problem of the optical fiber cladded with silver. A relatively large air region around the fiber is terminated by the ABCs.

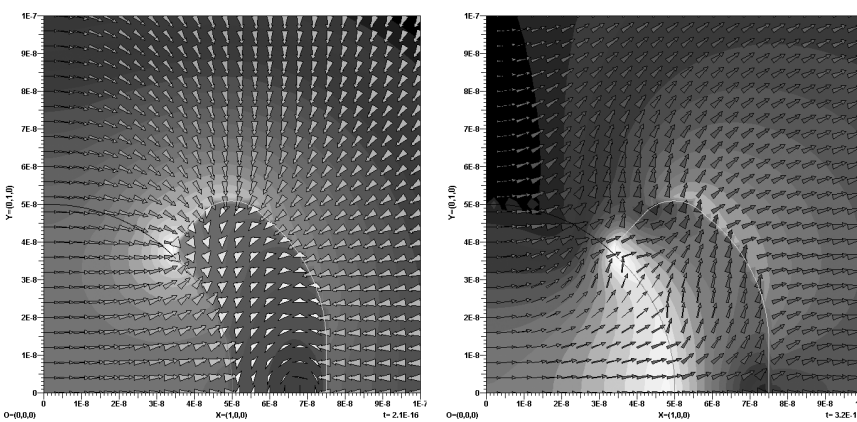


Fig. 16. Electric field of the fundamental mode at 500 nm (left) and of a higher order mode at 380 nm (right) with the same polarization.

search function is observed at 350 nm, as illustrated in Fig. 17. Within the plotted area, one detects three modes. Two of them are rather lossy (large imaginary part). As one may see, the corresponding trace are complicated and cross each other. Furthermore, there are sections where the propagation

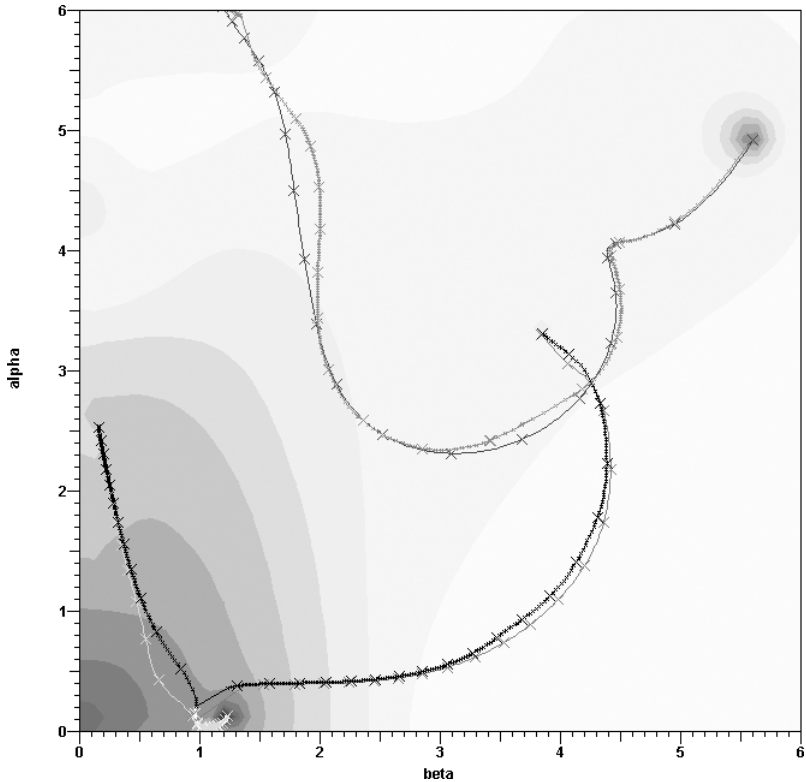


Fig. 17. MMP eigenvalue search function at 350 nm for a circular dielectric fiber coated with two symmetric silver patches. The fundamental HE11 like mode is not present in the observation windows because the real part of its propagation constant is too high. Three higher order modes are detected and traced for increasing wavelengths. The traces for two models with different accuracy are shown.

constant changes considerably even if the frequency varies only slightly, i.e., where the derivative of the propagation constant with respect to the frequency is very high. In Fig. 17, the derivative is higher the larger the distance between neighbor x markers is. Large derivatives are typical for slow waves, but here, the loss is significant that the term slow wave is questionable.

In order to clarify the accuracy of the MMP computations, several models with different multipole orders were computed. Two of them are shown in Fig. 17. As one may see, the differences are the biggest in the "slow wave" areas, which indicates that an accurate computation requires highly accurate models in these areas.

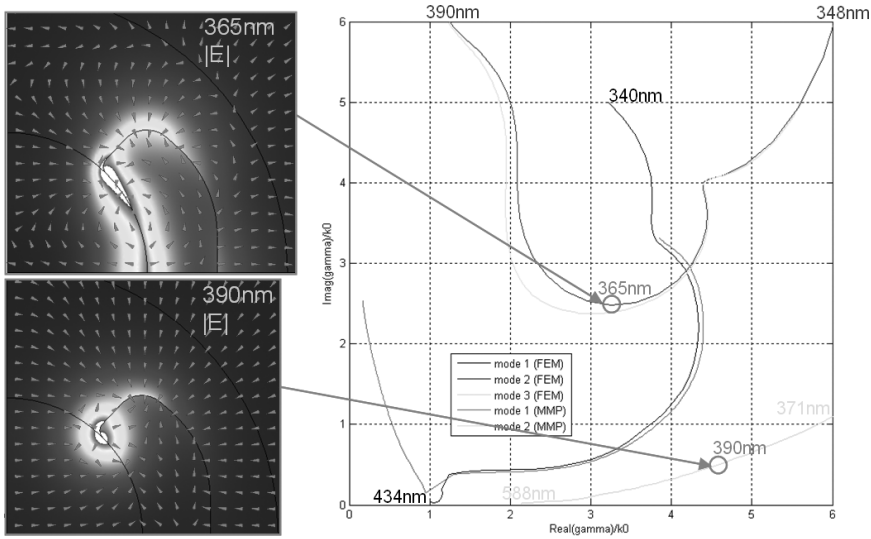


Fig. 18. The comparison between the MMP and FEM eigenvalue traces in the complex γ plane. A very good agreement of the results is obtained. The results disagree only in those regions where the resonance peaks are very flat (high losses), i.e., where an accurate eigenvalue extraction is not possible.

One may also observe that the mode indicated by the black and gray lines switches to another mode at longer wavelength. This happens when the real part of the normalized propagation constant becomes smaller than 1. Such modes are not strictly guided, i.e., couple with plane waves in free space. One may also say that this mode has a cutoff and does not propagate at longer wavelengths. In order to verify that this is not an artifact of the MMP simulation, a comparison with FEM is performed.

The obtained FEM results agree well with the corresponding MMP results as one may see in Fig. 18. There is an observable disagreement only for the mode with high losses (between the dark gray and light gray curves) in the "slow wave" area.

A mode with very low losses (the curve near the real axis) has been also detected. This mode is not visible in Fig. 17 because it is out of the observation window at 350 nm. This fundamental mode has a field pattern similar to the HE11 mode of optical fiber without cladding and similar to the fundamental two wire mode known from transmission lines. Numerically, it is much easier to evaluate than the other modes. Fortunately, it is usual that the fundamental modes are the easiest ones from a numerical point of view.

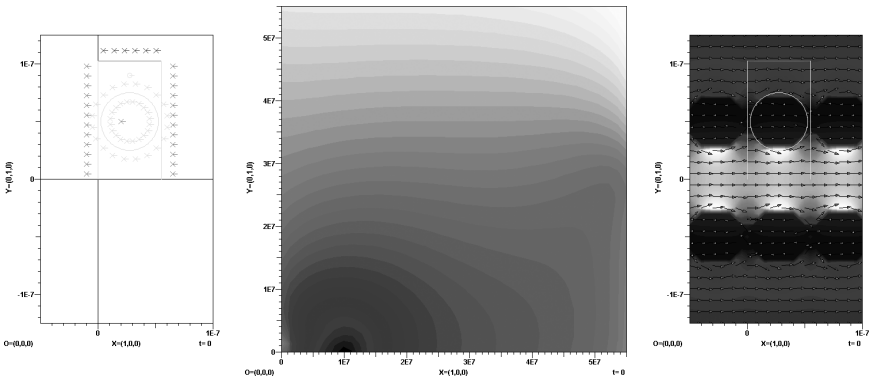


Fig. 19. The rudimentary metallic (silver) photonic crystal waveguide. The MMP model (left), the eigenvalue search function (Error/Amplitude) in the complex γ plane (middle), and the Poynting vector of the "guiding" mode (right) at the wavelength of 600 nm are presented.

5.3. *Periodic, cylindrical waveguide*

Periodic waveguides with radiation and material losses are very challenging for eigenvalue solvers. As an example, a rudimentary metallic photonic crystal waveguide is defined in Fig. 19. The radiation losses of this "waveguide" are significant as it contains only one row of silver rods along the "guiding" channel. The permittivity of silver is represented by previously defined Drude model.

The geometry of the waveguide is presented in Fig. 19 along with the MMP expansions used for modeling of the field in different regions. Due to the large radius of the rods several lower older multiple expansions are distributed along the circular boundary rather than very high order expansions placed in the centre.

The structure is periodic in the x direction. The periodic constant Cx is chosen as an eigenvalue for a given wavelength. The constant Cx is the propagation constant of the waveguide.

In our MMP modeling a special attention to the radiation condition should be paid. Therefore, a fictitious horizontal boundary and $(x + y = lE - 7/\alpha)$ Rayleigh expansion (gray square) were defined.

A fictitious excitation (monopole) was set in the circle to excite the field in the air of the periodic cell. As an eigenvalue search function the average error over the transmitted power was chosen. The behavior of this function at the wavelength of 600 nm in the complex Cx plane is shown in Fig. 14 (middle).

The dark blue dip of the search function represent our eigenvalue. The corresponding Poynting vector field is presented in Fig. 14 (right).

Since the resonance dip is close to the real Cx axis, the radiation and material losses are not significant. This is rather surprising as we have taken only two photonic crystal rows around the guiding channel.

5.4. *Periodic, axisymmetric waveguide*

In the previous section, a finite chain of metallic spheres within a dielectric cylinder were considered. In order to understand the wave propagation along this structure, an eigenvalue analysis for the corresponding infinite, periodic structure is considered.

Figure 20 shows the MMP eigenvalue search function for a simple chain of gold spheres along the x axis in the complex Cx plane. The search function is periodic along its $\mathrm{Re}(Cx)$ axis. Therefore, only one finite interval along the $\mathrm{Re}(Cx)$ axis must be considered. In $\mathrm{Im}(Cx)$ one may find infinitely many modes but only the modes with low loss, i.e., small $\mathrm{Im}(Cx)$ are of practical interest and produce observable physical objects. One may see that all modes are rather lossy and exhibit a complicated frequency dependence because of the strong material dispersion of gold.

When the gold spheres are embedded in a dielectric rod, the MMP search function becomes rather simple at 500 nm and the trace of the

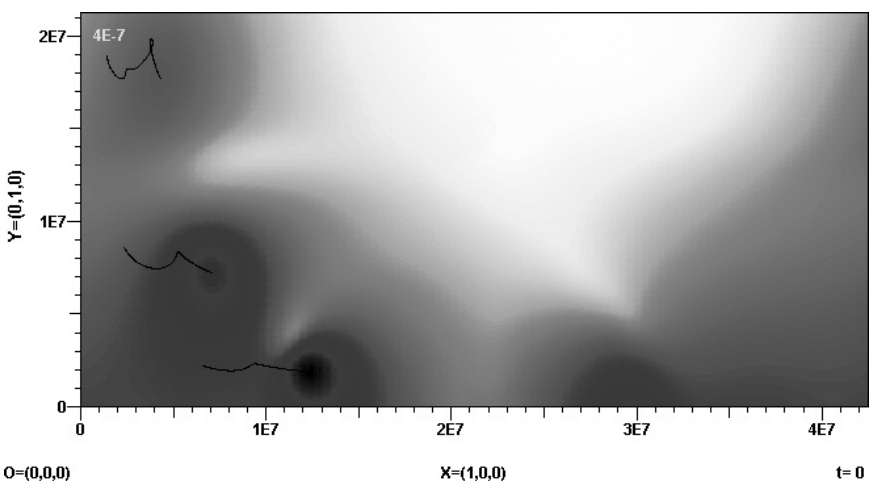


Fig. 20. MMP eigenvalue search function at 400 nm in the complex Cx plane for a chain of gold spheres with traces of the eigenvalues for increasing wavelength.

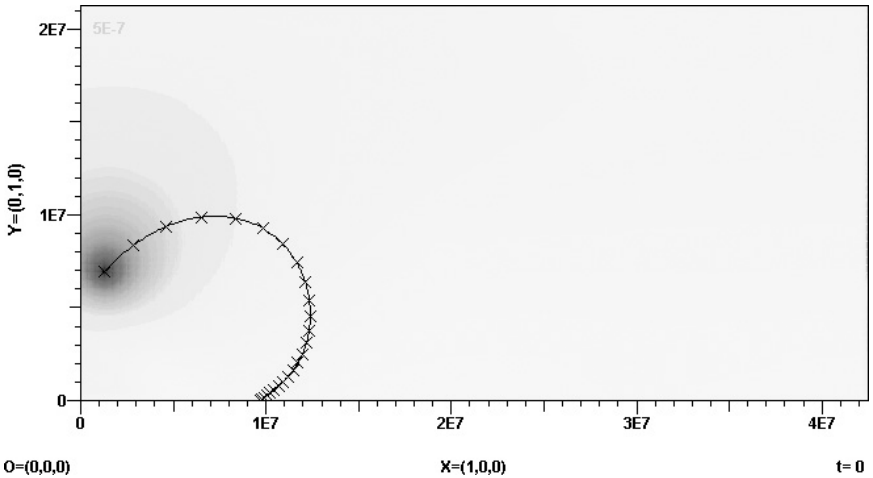


Fig. 21. MMP eigenvalue search function at 500 nm in the complex Cx plane for a chain of gold spheres within a sapphire rod with trace of the eigenvalue with low loss (at 617 nm, where the trace hits the real axis) for increasing wavelength.

corresponding mode follows a smooth line that hits the real axis near 617 nm, which indicates that the mode has very low loss at wavelength of 617 nm and a bit less (see Fig. 21). At 550 nm it is already rather lossy. Since the mode disappears for wavelengths longer than 617 nm, it seems to have a cutoff. This situation is rather delicate. Figure 22 shows that the energy propagates almost parallel to the x axis. Only a part of the spheres is illuminated and the energy mostly surrounds them but there is still some energy penetrating the spheres, which must lead to power loss. Since $\mathrm{Re}(Cx)$ is almost 0, the energy conservation law is violated. Most likely, the solver continues tracing the mode after cutoff is reached. It then continues searching although the numerical model becomes less and less inaccurate. Correctly detecting a cutoff is very cumbersome because of the limited numerical accuracy — especially in demanding situations as periodic waveguides that require a full 3D analysis.

It is interesting to note that the low loss mode has a wavelength roughly equal to four periods, i.e., $4Dx$. This may be seen from Figs. 23 and 24. An intriguing idea may come to mind, which views each successive sphere as an individual oscillator, driven by its predecessor in the chain. As the optimal phase difference for driving an oscillator is 90 degrees, it might appear only natural that 4 such units should make up one full

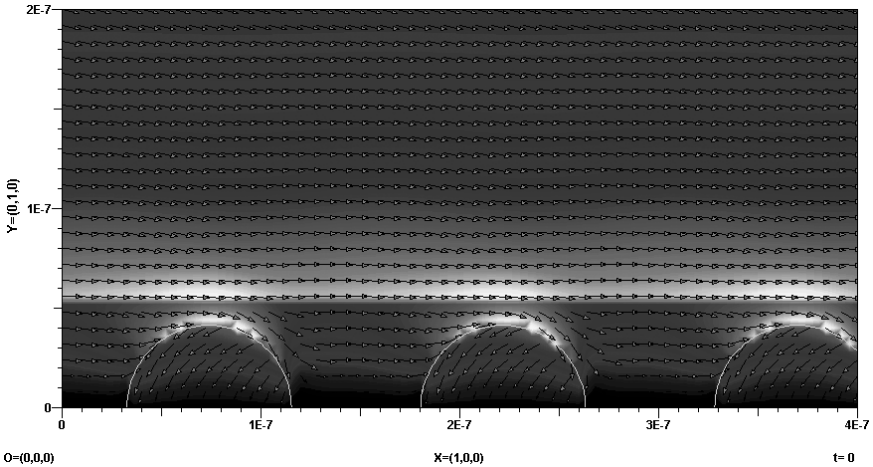


Fig. 22. Time average of the Poynting vector field for the low loss mode at 617 nm in a longitudinal section of the gold chain within the sapphire rod.

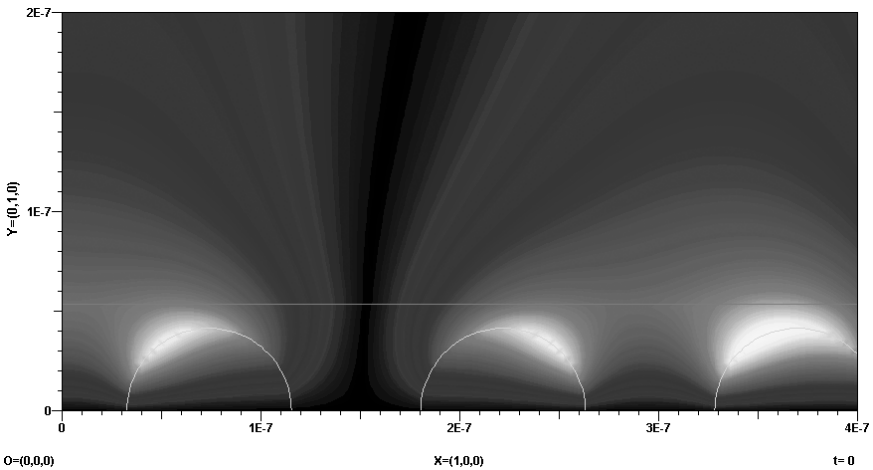


Fig. 23. Magnetic field intensity for the low loss mode at 617 nm.

wavelength, i.e., a 360 degree total phase shift. However, in view of the results presented in Fig. 11, where the chain period is varied by nearly an octave but the corresponding change in spectral peak position is not nearly as much, this present finding appears hardly more than a coincidental curiosity.

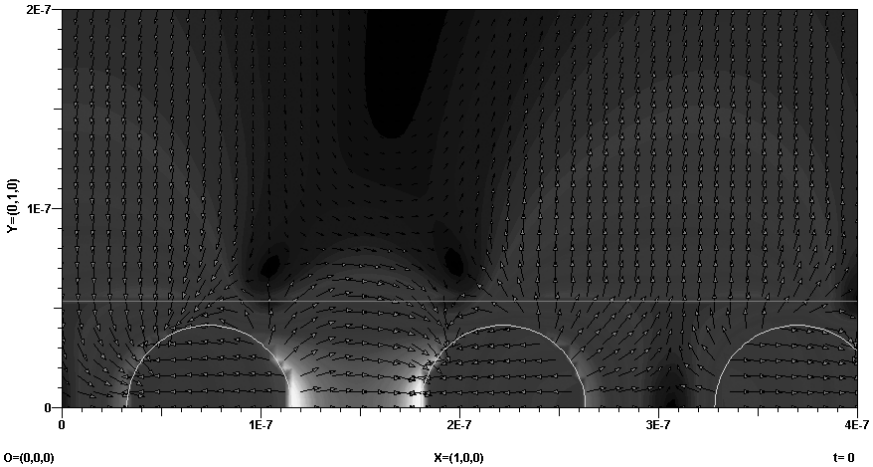


Fig. 24. Electric field for the low loss mode at 617 nm.

6. Advanced Structures

In order to demonstrate the treatment of truncated plasmonic waveguides, we consider a circular metallic wire with an ultra sharp tip at its end. Such wires are attractive for scanning nearfield optical microscopy (SNOM).[60–62] Assume that one of the metal claddings of the fiber in the previous section is moved away from the fiber. Then it may guide the fundamental wire mode that was also considered in the previous section. If the wire is then cut, the mode will be partly reflected and a part of the energy will radiate from the end of the wire. By forming a sharp tip at the end of the wire, one may additionally obtain a very strongly localized electric field in the tip area, which is valuable for SNOM applications. Since the wire may be long compared with the wavelength, it is reasonable to assume that it extends to infinity along the y axis as shown in Fig. 25.

From the eigenvalue analysis of the wire, it is known that an axisymmetric mode exists that can propagate with relatively low loss. If the wire is long enough, higher order modes with stronger attenuation will be negligible near the end of the wire. Furthermore, as long as the rotational symmetry is not broken by the tip, no mode conversion will take place and non-axisymmetric modes must not be considered at all. Thus, one may assume that a single mode — propagating in y direction — excites the structure. In 3D MMP models, the expansions describing this mode — obtained from the previous eigenvalue analysis — may be embedded in a so-called connection.

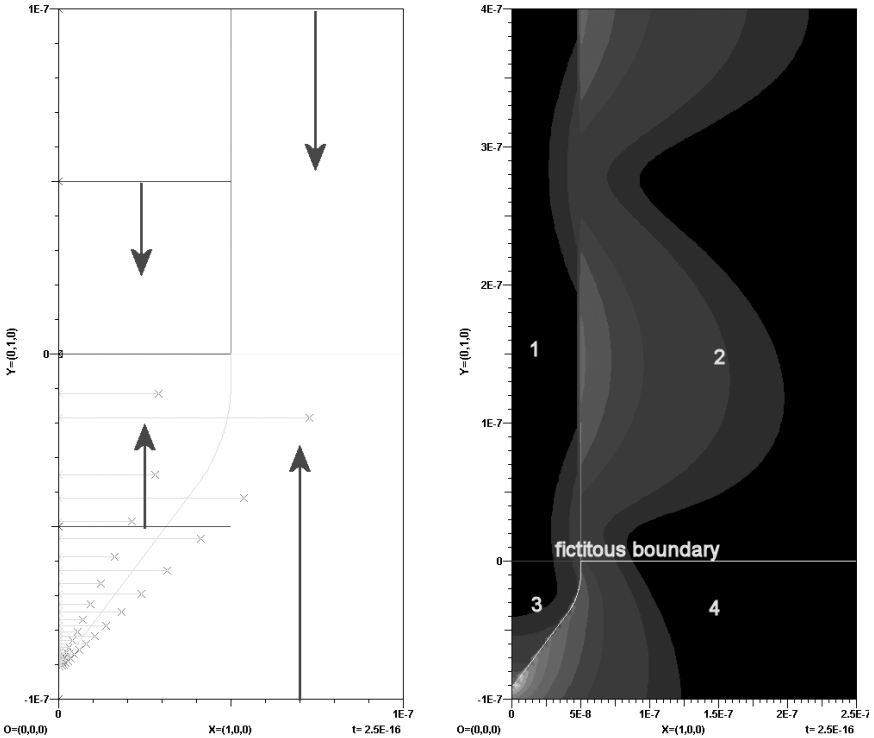


Fig. 25. Metallic SNOM tip. Left: cross section of the model with four domains and MMP expansions. The 4 complex origin multipoles are indicated by arrows showing the main direction of radiation. Green lines ending in red crosses indicate the ring multipoles that are used for modeling the tip area (domains 3 and 4). The incident and reflected modes in domains 1 and 2 are not shown. Right: Intensity plot of the electric field density, domain numbers and fictitious boundary that separates the upper and lower half spaces.

The connection is then a new expansion that can be used as any other expansion. Since the connection contains the field of the mode both inside and outside the wire, it may be considered as a multi-domain expansion that fulfills not only the Maxwell's equations but also the boundary conditions between the two domains. As a result, the boundary between the core and surrounding of the wire needs not be modeled explicitly. This is very important because this boundary extends to infinity.

In the tip area, the cylindrical symmetry of the wire is broken. Thus, the field of the wire mode is only useful for modeling the field in the upper half space at some distance from the tip. In order to separate this area, a fictitious boundary must be introduced that truncates both the core of

the wire and the surrounding medium. Since the latter extends to infinity, this fictitious boundary must also extend to infinity. For numerical reasons, it must be truncated. As a consequence the field on the interface between upper and lower half space cannot be matched at some distance from the wire. As long as the field decays reasonably with the distance from the wire, this is acceptable.

Now, the field in the four domains (1) upper half space, inside the wire, (2) upper half space outside the wire, (3) lower half space inside the tip, (4) lower half space inside the tip, must be modeled with appropriate expansions. In the domains 1 and 2, one expects not only the incident, guided mode, but also a reflected mode, strongly damped higher order modes and radiating modes. The reflected mode can be modeled with the same connection as the incident mode, rotated by 180 degrees. This expansion also needs no boundary definition between domains 1 and 2. In a general mode matching approach one now would have to evaluate all possible higher order and radiating modes, which would be extremely cumbersome. In MMP one may model all of these modes together with multipole expansions sitting in the lower half space and radiating into the domains 1 and 2 respectively. These multipole expansions should be distributed along the fictitious boundary. Since they do not fulfill the boundary condition between domains 1 and 2, this boundary must now be discretized, which requires another truncation because it also extends to infinity. Fortunately, the higher order and radiating modes decay rather rapidly with the distance from the tip. Therefore, it is sufficient to model a short piece of the boundary between domain 1 and 2 as shown in Fig. 25.

Instead of setting several multipoles along the fictitious boundary for modeling the higher order and radiating modes, it is sufficient to place a single complex origin multipole expansion on the axis for domain 1 and another one for domain 2. In fact, complex origin multipoles are excellently suited for modeling the field of truncated waveguides.

Now, the field in the domains 3 and 4 must be modeled. First of all, the incident waves and parts of the radiating modes may be modeled with two complex origin multipoles sitting on the axis in the upper half space and radiating into domains 3 and 4.

In addition to the complex origin multipoles one must now place two sets of multipoles along the surface of the tip, one inside the wire radiating into domain 4 and one in domain 4 radiating into domain 3. These sets may be generated automatically with pole setting routines of OpenMaX.

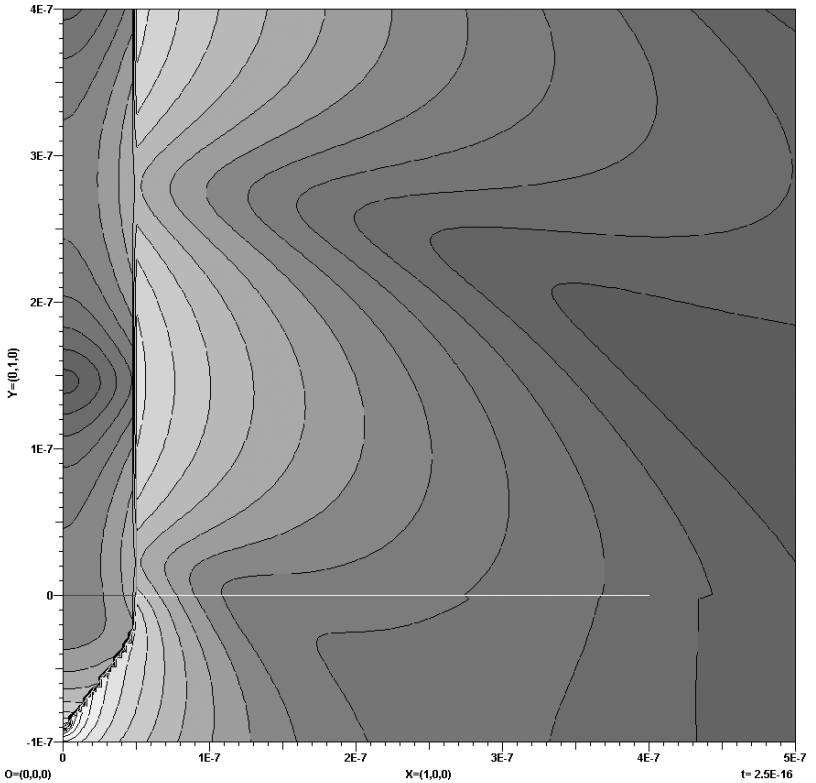


Fig. 26. Electric field density in a cross section of the metallic tip, excited by a guided mode from top. Logarithmic field representation and iso lines make the discontinuities along the fictitious boundary visible. The discontinuities are due to inaccurate modeling at larger distances from the tip, where the field is very small. The electric energy density between neighbor iso lines varies by the factor 2.

Since the structure is axisymmetric, it is reasonable to benefit from ring multipoles,[38] which integrate multipole expansions along circular rings around the axis. Using a combination of 2 connections, 4 complex origin multipoles, and 2 sets of automatically generated ring multipole expansions, one obtains the solution shown in Fig. 26 with accurate field computations in the tip area, where the field enhancement is very high as one may see from Fig. 27.

The model leads to a small 258×1398 matrix and to an average mismatching error of 9.35% along the boundaries. The maximum relative error is 3.4%. Since it is observed at positions where the field is relatively small, this is acceptable. As one can see from Fig. 26 the error clearly

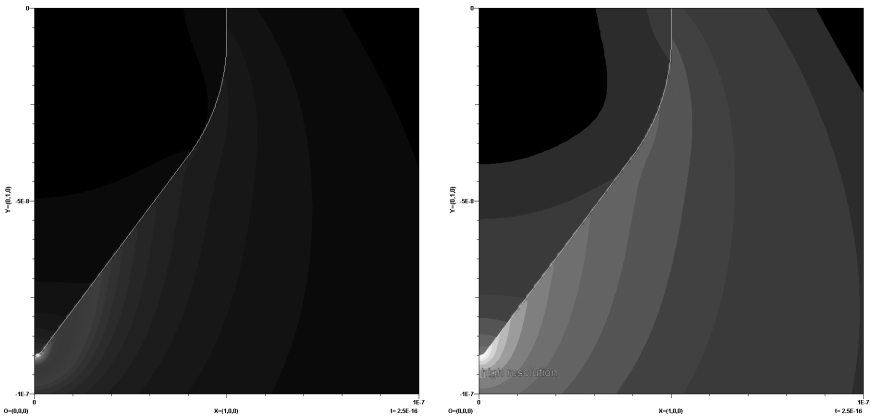


Fig. 27. Electric field intensity near the metallic tip. Left: linear scale for showing how pronounced the field enhancement is at the tip. Right: logarithmic scale. In the close vicinity of the tip a finer mesh was used for the representation of the field.

increases along the fictitious boundary with the distance from the wire, which is expected. On an older single core AMD Athlon 64 (2.6 GHz), the matrix setup time is 2.3 sec and the solve time 0.3 sec. The relatively long matrix setup time is caused by the ring multipole expansions that require numerical integrations of multipole expansions. In principle, these expansions might be replaced by appropriate sets of complex origin multipoles to reduce the matrix setup time.

7. Conclusions

Basic concepts of RF antennas and waveguides are not directly applicable to plasmonic structures due to a number of crucial differences. Most importantly, the skin depth is not negligible. Penetration of fields into the metallic structures leads to reduced phase velocities along the metal-dielectric interfaces, which in turn gives rise to geometric resonances that differ drastically from the usual rule of integral multiples of half the vacuum wavelength.

For understanding and designing plasmonic waveguides and antennas, appropriate numerical methods are highly valuable. In order to obtain efficient calculations, taking symmetries into account is very important, especially in the case of periodic plasmonic waveguides that lead to demanding complex eigenvalue problems with periodic symmetry. As a consequence, frequency domain methods are favorable. Furthermore, strong

material dispersion and considerable material loss are also more naturally and efficiently handled in frequency domain methods. Finally, strong field localizations cause substantial difficulties for methods based on regular grids therefore; methods based on unstructured meshes, such as FEM, or mesh-free methods and boundary discretization techniques, such as MMP, are best suited.

Various applications illustrate how plasmonic antenna and waveguide problems may be solved using FEM and MMP and demonstrate that a good agreement of the results may be obtained.

References

1. W. L. Barnes, A. Dereux, and T. W. Ebbesen, *Nature* **424**, 824 (2003).
2. R. Esteban *et al.*, *Nano Lett.* **8**, 3155 (2008).
3. R. Vogelgesang and A. Dmitriev, *Analyst* **135**, 1175 (2010).
4. P. G. Etchegoin, E. C. Le Ru, and M. Meyer, *J. Chem. Phys.* **125**, 164705 (2006).
5. A. Taflove and S. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. (Artech House, Norwood, MA, 2005).
6. A. Taflove, *The Finite-Difference Time-Domain Method*. (Artech House, Norwood, MA, 1995).
7. T. Weiland, *AEU-Int. J. Electron. Commun.* **31**, 116 (1977).
8. D. Baumann, C. Fumeaux, C. Hafner, and E. P. Li, *Optics Express* **17**, 15186 (2009).
9. J. Jin, *The Finite Element Method in Electromagnetics*. (John Wiley & Sons, New York, 2002).
10. G. A. Stark *et al.*, *Int. J. Numer. Model.-Electron. Netw. Device Fields* **22**, 201 (2009).
11. V. A. Mandelshtam and H. S. Taylor, *J. Chem. Phys.* **107**, 6756 (1997).
12. Meep: Free FDTD simulation software package, by MIT (http://ab-initio.mit.edu/wiki/index.php/Meep).
13. K. S. Yee, *IEEE Trans. Antennas Propag.* **AP14**, 302 (1966).
14. Lumerical, by Lumerical (http://www.lumerical.com/).
15. OptiFDTD, by Optiwave (http://www.optiwave.com/products/fdtd_overview.html).
16. MAFIA, by CST (http://www.temf.de/FIT-MAFIA.27.0.html).
17. CST Microwave Studio, by CST (http://www.cst.com/Content/Products/MWS/Overview.aspx).
18. HyperMesh, by A. HyperWorks (http://www.altairhyperworks.com/HW-Temp1Product.aspx?product_id=7).
19. GMSH, by C. Geuzaine, J.-F. Remacle (http://geuz.org/gmsh/).
20. T. V. Yioultsis, N. V. Kantartzis, C. S. Antonopoulos, and T. D. Tsiboukis, *IEEE Trans. Magn.* **34**, 3288 (1998).
21. G. Mur, *IEEE Trans. Magn.* **28**, 1130 (1992).

22. D. Jiao and J. M. Jin, *IEEE Trans. Antennas Propag.* **50**, 1615 (2002).
23. D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *SIAM J. Numer. Anal.* **39**, 1749 (2002).
24. C.-P. Yu and H.-C. Chang. In *Electromagnetic Theory and Applications for Photonic Crystals*, ed. K. Yasumoto. Boca Raton, FL (CRC Press, Taylor & Francis Group, 2006).
25. H. Whitney, *Geometric Integration Theory.* (Princeton University Press, Princeton, NJ, 1957).
26. J. C. Nedelec, *Numer. Math.* **35**, 315 (1980).
27. J. P. Webb and V. N. Kanellopoulos, *Microw. Opt. Technol. Lett.* **2**, 370 (1989).
28. J. P. Berenger, *J. Comput. Phys.* **127**, 363 (1996).
29. J. Jian-Ming and C. Weng Cho, *Microw. Opt. Technol. Lett.* **12**, 192 (1996).
30. Y. Zhu and A. Cangellaris, *Multigrid Finite Element Methods for Electromagnetic Field Modeling.* (IEEE Press, John Wiley & Sons, Hoboken, NJ, 2006).
31. D. B. Davidson, *Computational Electromagnetics for RF and Microwave Engineering* (Cambridge University Press, 2005).
32. A. Buffa and R. Hiptmair, In *Topics In Computational Wave Propagation — Direct And Inverse Problems* (Springer-Verlag Berlin, 2003) p. 83.
33. L. Greengard and V. Rokhlin, *J. Comput. Phys.* **73**, 325 (1987).
34. J. Smajic, Z. Andjelic, and M. Bebendorf, *IEEE Trans. Magn.* **43**, 1269 (2007).
35. K. Tavzarashvili *et al.*, *J. Comput. Theor. Nanosci.* **4**, 667 (2007).
36. F. Zolla and R. Petit, *J. Opt. Soc. Am. A-Opt. Image Sci. Vis.* **13**, 796 (1996).
37. G. Mie, *Ann. Phys.-Berlin* **25**, 377 (1908).
38. C. Hafner, *Post-modern Electromagnetics: Using Intelligent Maxwell Solvers.* (John Wiley & Sons, Chichester, 1999).
39. OpenMaX, by C. Hafner (http://OpenMaX.ethz.ch/).
40. C. Hafner, *MaX-1: A Visual Electromagnetics Platform.* (John Wiley & Sons, Chichester, 1998).
41. JCMwave, by JCMwave (http://www.jcmwave.com/).
42. M. Quinten, A. Leitner, J. R. Krenn, and F. R. Aussenegg, *Opt. Lett.* **23**, 1331 (1998).
43. M. L. Brongersma, J. W. Hartman, and H. A. Atwater, *Physical Review B* **62**, 16356 (2000).
44. S. A. Maier *et al.*, *Adv. Mater.* **13**, 1501 (2001).
45. S. A. Maier *et al.*, *Nat. Mater.* **2**, 229 (2003).
46. S. A. Maier and H. A. Atwater, *J. Appl. Phys.* **98**, 011101 (2005).
47. L. L. Zhao, K. L. Kelly, and G. C. Schatz, *J. Phys. Chem. B* **107**, 7343 (2003).
48. D. S. Citrin, *Nano Lett.* **5**, 985 (2005).
49. S. L. Zou and G. C. Schatz, *Nanotechnology* **17**, 2813 (2006).
50. S. A. Maier, M. L. Brongersma, and H. A. Atwater, *Appl. Phys. Lett.* **78**, 16 (2001).
51. S. A. Maier, P. G. Kik, and H. A. Atwater, *Appl. Phys. Lett.* **81**, 1714 (2002).

52. Y. J. Kang, K. J. Erickson, and T. A. Taton, *J. Am. Chem. Soc.* **127**, 13800 (2005).
53. S. Lin *et al.*, *Adv. Mater.* **17**, 2553 (2005).
54. T. Sannomiya, J. Voros, and C. Hafner, *J. Comput. Theor. Nanosci.* **6**, 749 (2009).
55. C. Hafner and J. Smajic, *J. of Modern Optics* 58, 467–479, March 2011.
56. A. Sommerfeld, *Partial Differential Equations in Physics (English version).* (Academic Press Inc., New York, 1949).
57. C. A. Pfeiffer, E. N. Economou, and K. L. Ngai, *Physical Review B* **10**, 3038 (1974).
58. J. C. Ashley and L. C. Emerson, *Surface Science* **41**, 615 (1974).
59. P. B. Johnson and R. W. Christy, *Physical Review B* **6**, 4370 (1972).
60. R. Esteban, R. Vogelgesang, and K. Kern, *Nanotechnology* **17**, 475 (2006).
61. R. Esteban, R. Vogelgesang, and K. Kern, *Physical Review B* **75** (2007).
62. W. Nakagawa, L. Vaccaro, H. P. Herzig, and C. Hafner, *J. Comput. Theor. Nanosci.* **4**, 692 (2007).

Index