



Terahertz and infrared nonlocality and field saturation in extreme-scale nanoslits

FERRAN VIDAL-CODINA,^{1,5} LUIS MARTÍN-MORENO,² CRISTIAN CIRACÌ,³  DAEHAN YOO,⁴ NGOC-CUONG NGUYEN,¹ SANG-HYUN OH,^{4,6} AND JAIME PERAIRE¹

¹*Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

²*Instituto de Ciencia de Materiales de Aragón and Departamento de Física de la Materia Condensada, CSIC-Universidad de Zaragoza, E-50009 Zaragoza, Spain*

³*Center for Biomolecular Nanotechnologies, Istituto Italiano di Tecnologia, Via Barsanti 14, 73010 Arnesano (LE), Italy*

⁴*Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, Minnesota 55455, USA*

⁵*fvidal@mit.edu*

⁶*sang@umn.edu*

Abstract: With advances in nanofabrication techniques, extreme-scale nanophotonic devices with critical gap dimensions of just 1-2 nm have been realized. The plasmonic response in these extreme-scale gaps is significantly affected by nonlocal electrostatics, quenching field enhancement and blue-shifting the resonance with respect to a purely local behavior. The extreme mismatch in lengthscales, ranging from millimeter-long wavelengths to atomic-scale charge distributions, poses a daunting computational challenge. In this paper, we perform computations of a single nanoslit using the hybridizable discontinuous Galerkin method to solve Maxwell's equations augmented with the hydrodynamic model for the conduction-band electrons in noble metals. This method enables the efficient simulation of the slit while accounting for the nonlocal interactions between electrons and the incident light. We study the impact of gap width, film thickness and electron motion model on the plasmon resonances of the slit for two different frequency regimes: (1) terahertz frequencies, which lead to 1000-fold field amplitude enhancements that saturate as the gap shrinks; and (2) the near- and mid-infrared regime, where we show that narrow gaps and thick films cluster Fabry-Pérot (FP) resonances towards lower frequencies, derive a dispersion relation for the first FP resonance, in addition to observing that nonlocality boosts transmittance and reduces enhancement.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

A nanometer-scale slit in a metal film is one of basic building blocks to construct plasmonic waveguides [1–4], sensors [5–7], metasurfaces [8], nanolasers [9], and nanoelectrodes [10]. The interactions of electromagnetic waves with a single metallic slit or periodic slits have been extensively studied over broad spectral ranges covering visible, infrared, terahertz, and microwave regimes [11]. With continual advances in nanofabrication techniques, researchers are able to produce sub-10 nm slits with wafer-scale throughput and uniformity [12–15].

As electromagnetic waves pass through narrow slits, their local field intensity can be significantly enhanced. Such high field energy density in the slit, in turn, can lead to interesting effects such as enhanced optical transmission [16], nonlinear phenomena [17,18], and surface-enhanced infrared absorption [19,20], among others. As the slit width is reduced to few nanometers and below, quantum mechanical effects [21–26], in particular the nonlocal smearing of conduction

electrons characterized by the Thomas-Fermi screening length (~ 0.1 nm in gold) cannot be ignored [22,27–32].

Under the push of an external electric field, electrons in a metal system tend to accumulate at its surface. In the classical local-response description, electrons do not interact with each other so that they can occupy identical states, generating an infinitesimally thin surface distribution layer, which in turn leads to a discontinuity in the macroscopic normal field component. However, Coulomb interactions and, more importantly, the Pauli exclusion principle prevent electrons from occupying the same state, leading to a strong electron-electron repulsion. Thus an electron gas resists compression by the applied electromagnetic field, giving rise to a smeared charge distribution. An external field applied at a point \mathbf{r}_0 then causes a response not only at that point, but also in the neighboring volume, producing microscopically a nonlocal response effect. Because abrupt spatial variations in the materials properties at the interface do not lead to equally abrupt variations in charge/field distributions, the nonlocal effect poses one of fundamental barriers to the realization of ultimate field confinement and enhancement in the ultra-narrow gaps [27].

Previous studies on nonlocal effects mainly focused on visible and near-infrared regimes. Extending such investigations to longer wavelength regimes, in particular, millimeter-wave or terahertz (THz) and infrared (MIR) regime (wavelengths 0.75–20 μm), is important for two reasons. First, researchers have demonstrated substantial field enhancements in metal-insulator-metal (MIM) structures that are resonant in the MIR [20,33] and THz regimes [34–39]. Naturally, one could ask how the nonlocal effect limits field confinement and enhancement in such cases. Second, besides its fundamental importance, nonlocality is also critical for the accurate design and modeling of antennas, metamaterials, and sensors for practical applications. Despite the rapid growth of MIR and THz nanophotonics [40], the enormous mismatch between the long free-space wavelengths and the atomic-scale nature of nonlocality poses severe challenges to both experimental and numerical studies. As the wavelength increases, computational modeling of nonlocality in the MIR or THz regimes, where the mismatch between the wavelength (several microns to millimeters) and the Thomas-Fermi length (~ 0.1 nm) can be larger than eight orders of magnitude, exceeding the capability of many commercial finite-difference time-domain (FDTD) or finite-element method (FEM) solvers.

The hybridizable discontinuous Galerkin (HDG) method is a high-order accurate finite-element numerical scheme that can address aforementioned challenges. Our previous work showed the capability of the HDG method for efficiently modeling long-wavelength MIR and THz radiation in free space and plasmonic resonators [15,41–43]. We have also equipped the HDG method with the hydrodynamic model for metals [44,45]. In this paper, we leverage this novel framework to address the daunting challenge of modeling THz nonlocality, which requires resolving the nonlocal charge distribution in the metal and the interaction of a millimeter-long wave within single-digit nanometer gaps. Here we apply the HDG method to study light transmission through a single slit in the MIR and THz regimes. We quantify the impact of nonlocality in field enhancement, localization, as well as field saturation phenomena [46,47].

2. Results and discussion

The structure analyzed in this article consists of an infinitely extended gold film of thickness T suspended in free space, with an infinitely long aperture or gap of width G , illuminated from below by a Gaussian beam orthogonally polarized to the slit, see Fig. 7 (in the Appendix). For the simulations we consider film thicknesses between 150 and 500 nm and perfectly smooth gaps ranging from 1 to 50 nm, which will allow us to identify the influence of nonlocal electrostatics without accounting for purely quantum phenomena, such as electron tunneling, that gain importance below ~ 1 nm [23,31]. In addition, two different frequency regimes are considered: the low terahertz regime (0.35–1.5 mm wavelengths) and the near to mid infrared

regime (0.75–20 micron wavelengths), where Fabry-Pérot (FP) resonances are excited. Simulation parameters and results exhibit differences between these two regimes, hence they are treated separately in the discussion.

2.1. Simulation methods

The results presented in this work couple two sets of equations: (1) the time-harmonic Maxwell's equations are solved in free space; and (2) the time-harmonic Maxwell's equations augmented with the hydrodynamic model for metals [27,48] are solved in the gold film. For illustration purposes, the influence of the nonlocal parameter β is accounted through an auxiliary b factor described as $\beta^2 = b\beta_0^2$, where $\beta_0^2 = 3/5v_F^2$ is the widely used value of the nonlocal parameter in terms of the Fermi velocity v_F [49]. Note that $b = 0$ corresponds to a purely local description of electron motion in metals based on the Drude model.

The numerical experiments in this article have been carried out with the hybridizable discontinuous Galerkin (HDG) method, a fast and high-order accurate numerical method that has been developed to solve the 3-D time-harmonic Maxwell's equations for both the local [41] and the nonlocal [45] model. The HDG method is specifically tailored to resolve tight field localizations among multiple length scales, and has therefore been used to successfully simulate ultra-narrow coaxial apertures periodically patterned on optically thin gold films [15,42,43,45]. The HDG method belongs to the class of discontinuous Galerkin (DG) methods, which are unstructured, high accurate, locally conservative, exhibit low dissipation and dispersion and are high-order, meaning the numerical error in the approximation can be made insensitive to the mesh discretization, as opposed to other finite-difference time-domain (FDTD) or finite element method (FEM) commercial solvers. For the aforementioned reasons, DG methods have been extensively used to simulate plasmonics phenomena [50–54]. The HDG method differs from other DG methods in the introduction of auxiliary variables defined on the faces of the discretization—also known as hybrid variables, hence the name hybridizable—, which enable the solution of smaller linear systems of equations and attain optimal convergence rates. Thus, the HDG method is a strong candidate for solving extreme nanophotonics phenomena [44,45,52]. For a more thorough review on the HDG method and its implementation to the infinite slit problem, we refer the reader to Appendix A.

Leveraging the symmetry of the structure, we simulate only the 2-D cross-section of half of the slit, thus prescribing perfect electric conductor conditions on the boundary along the symmetry plane. Perfectly matched layers (PMLs) [55,56] are used to truncate the computational domain on all non-symmetry boundaries, precluding the outgoing waves from re-entering the domain. The structure is illuminated at normal incidence from below with a monochromatic Gaussian beam of waist $w_0 = 3\lambda$. Since we are interested in different frequency regimes, meshes of distinct resolution and degree of anisotropy have been used. In all cases, the solutions are approximated with cubic polynomials on quadrilateral elements. The computational domain is chosen such that the paraxial approximation of Gaussian beam remains valid and that the outgoing waves are totally attenuated in the PMLs region. The quantities of interest are the amplitude enhancement of the electric field along the upper tip of the aperture, hereinafter referred to as Q , and the transmittance. The expressions used to compute these outputs, along with the discussion on the meshes, governing equations, illumination settings, boundary conditions and parameter values are presented in the Appendix A.

2.2. Nonlocal electrodynamics in the terahertz regime

The single slit exhibits extraordinary optical transmission and huge field enhancements as the frequency approaches zero [11,46,57], and the enhancement has been shown to saturate for shrinking gap size [47,58], with a limiting saturation enhancement Q_{lim} that is inversely proportional to the film thickness and the wave frequency. These studies have been carried

out for a slit on a micrometer-thick film described with both the Drude model and as a perfect conductor, and here we extend the study to metal films thinner than a micron and for gaps ranging between 1 and 50 nm and modeled with both the Drude and the hydrodynamic model. The simulation of electromagnetic wave propagation for low terahertz frequencies is computationally challenging due to the huge mismatch in length scales –mm long wavelengths are squeezed into nanometer-wide apertures. In addition, simulations with the nonlocal model require a very fine resolution near the metal-vacuum interfaces so as to capture the extreme localization of the electron density and decay of the electric field. For each wavelength, we simulate a computational domain of $12\lambda \times 6\lambda$ surrounded by a 2λ -thick PMLs in each direction, ensuring the relevant plasmonic phenomena are properly represented. The longest wavelength considered for this regime is 1.5 mm (0.2 THz).

The simulated field enhancement $|E_x|/|E_{\text{inc}}|$ results for the 2 nm aperture, 300 nm film at 0.2 THz are shown in Fig. 1(a). The enhancement is almost constant along the slit with an extreme localization on the upper corners, thus we focus there to better appreciate the differences between the local and nonlocal model, namely a lower enhancement, a more diffused plasmonic hotspot and a smoother metal-dielectric transition. The induced electron density is depicted in Fig. 1(b) in logarithmic scale, where we observe the decay of several orders of magnitude within few Ångströms, forming a boundary-layer structure at the metal-dielectric interface. A detail of the mesh used to compute these solutions is also shown in Fig. 1(b), and we refer the reader to Fig. 8 (in the Appendix) for additional images on the mesh dimensions, characteristics and resolution. To further investigate the nonlocal effects at the metal-dielectric interface, we show the field enhancement profile along the middle of the slit in Fig. 1(c), for 300 nm film at 0.2 THz as before, according to the axes at the inset. The $b = 0$ case corresponds to describing the metal with the Drude model, where the metal acts like a mirror and thus allows minimal penetration of the electric field, resulting in an abrupt decay of the enhancement at the interface. For positive b the hydrodynamic model allows the propagation of a new longitudinal mode that relaxes the field enhancement decay inside the metal, forming a boundary-layer structure characterized by

the longitudinal skin depth $\delta_L = \beta/\Re\sqrt{\frac{\omega_p^2}{\epsilon_\infty} - \omega(\omega + i\gamma)}$ [59,60]. Note that this is much smaller than the usual (transverse mode) skin depth. For $\gamma = 0$, $\epsilon_\infty = 1$ and $\omega \ll \omega_p$, the depth at which the field has decayed by a factor of $1/e$ equals $\delta_L \simeq \beta/\omega_p$ [61]. Our simulation results coincide with the analytic decay, for instance with $b = 1$ and $G = 1$ nm, we have $\delta_{L,\text{exact}} = 0.786$ Å and $\delta_{L,\text{simulation}} = 0.783$ Å. The depth where the field enhancement $|E_x|/|E_{\text{inc}}|$ plateaus is roughly equivalent to 9, 12.5 and 15.5 Å for $b = 0.5, 1, 1.5$ respectively. Consequently, the effective aperture width is enlarged [62] compared to the classical local approximation, where the electric field decay occurs sharply at the interface, and larger values of β translate into increased skin depth inside the metal.

To study the electric field structure we monitor the field enhancement on the upper tip of the slit Q –details available in Appendix section A.2. The enhancement values for all gaps are depicted in Fig. 2(a) only for the nonlocal model ($b = 1$) and the 150 and 500 nm film –results are qualitatively identical for the local model– where we see that narrower gaps increase field enhancement. Nonetheless, we also observe that for a fixed thickness and frequency, shrinking the gap does not lead to arbitrarily large enhancements, but rather approaches a limiting value. This phenomenon is known as field enhancement saturation [47], and the limiting enhancement has been described as $Q_{\text{lim}} = \frac{\lambda\Delta H}{2\pi T}$, where T is the film thickness and ΔH is the difference in magnetic enhancement between the upper and lower boundaries of the film

$$\Delta H = \left| \frac{\int_{S_0} |\mathbf{H}|}{\int_{S_0} |\mathbf{H}_{\text{inc}}|} - \frac{\int_{S_1} |\mathbf{H}|}{\int_{S_1} |\mathbf{H}_{\text{inc}}|} \right|, \quad (1)$$

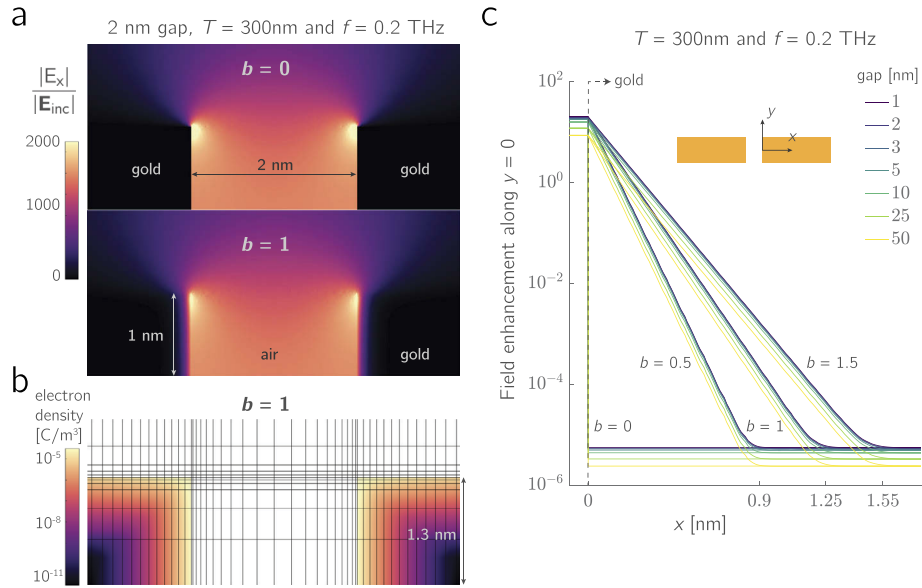


Fig. 1. Numerical results for a 300 nm thick film at 0.2 THz: (a) Field enhancement $|E_x|/|E_{inc}|$ on the upper tip of the 2 nm gap for the local (top) and nonlocal (bottom) model for $b = 1$. (b) Induced electron density on the upper corners in logscale for $b = 1$, with mesh discretization overlaid. (c) $|E_x|/|E_{inc}|$ profile along $x > 0, y = 0$ for different gaps and values of the parameter b , with axes system described in inset. Decreasing b reduces the electric field penetration inside the gold film, and the limiting case is the local model ($b = 0$) for which a sharp decay occurs at the interface due to the infinitesimal charge concentration.

under the same illumination conditions when there is no gap, that is the structure is an infinite gold film without an aperture [47].

In our simulations, we have found the limiting enhancement to be almost constant across models $Q_{lim}^{b=0}/Q_{lim}^{b=1} \approx 1$, mainly because the nonlocal effects are not noticeable in the absence of a gap where light is channeled.

For each film thickness and frequency, we evaluate its limiting enhancement and compare it with the actual enhancement Q observed when simulating the slit for various gap sizes. The results for $f = 0.2$ THz are shown in Fig. 2(b) for both the local and the nonlocal model with $b = 1$. Note that as gap decreases and film thickness increases, the ratio Q/Q_{lim} approaches 1, but at a different rate amongst electron models. For the local model, Q/Q_{lim} is between 95-97% for the 1 nm gap for all thicknesses, whereas for the 10 nm gap saturation ranges from 66% ($T = 150$ nm) to 86% ($T = 500$ nm), thus differences among thicknesses vanish as the gap reduces. This effect is not as pronounced for the nonlocal model, see for instance the 1 nm with saturation values between 83-94%, whereas the 10 nm gap exhibits saturation values in the range of 64-85%, very similar to those of the local model. Due to the electron density profile that smears inside the metal, the gap is effectively enlarged in the nonlocal calculations, therefore resulting in lower enhancements and lower saturation values.

The numerical results are therefore in good agreement with the theoretical behavior of the perfectly conducting slit, and show that for this quasi-static regime, extraordinary optical transmission and field enhancement saturation may also be observed if a more realistic physical model for electron motion is used, albeit achieving lower saturation values than for perfect electric conductor and Drude model representations of the metal.

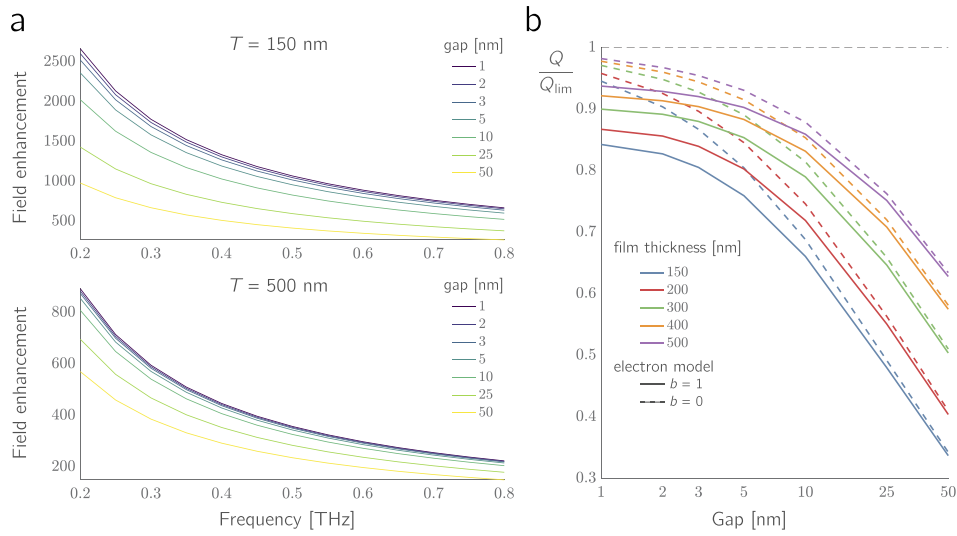


Fig. 2. (a) Field enhancement spectra of the nonlocal model for a 150 (top left) and 500 (bottom left) nm gold film, exhibiting saturation as gap shrinks. (b) Field enhancement as a function of gap size for different film thicknesses at 0.2 THz. Note the rate at which enhancement approaches Q_{lim} for shrinking gap width depends on film thickness. (c) Saturation field enhancement spectra for different film thicknesses, where we observe that the nonlocal model predicts lower saturation enhancements than the local model since the gap is effectively at enlarged. Differences between models are more acute for thinner films.

2.3. Nonlocal electrodynamics in the infrared regime

For a given slit gap and film thickness, increasing the frequencies towards the infrared regime excites several resonant Fabry-Pérot (FP) modes along the vertical direction of the slit, that lead to extraordinary optical transmission and large field enhancements [63]. In this section, we investigate the impact of nonlocality, gap width and film thickness on the FP resonances, devoting special attention to the first Fabry-Pérot resonance (FP1). Since the simulations span frequencies between 30 and 600 THz, for each wavelength we prescribe a computational domain of $12\lambda \times 6\lambda$, that together with a PMLs of width 2λ in each direction ensures a faithful representation of the underlying physics.

Firstly, we identify the separate influence of the aperture width and the film thickness on the location and distribution of the subsequent FP resonances in the mid IR, using the nonlocal $b = 1$ model. Setting the thickness to 500 nm, we compute the field enhancement profile for the 1 and 10 nm gap, see Fig. 3 (solid and dash). Narrower gaps squeeze FP resonances to lower frequencies and exhibit larger enhancements, see that below 300 THz the first seven FP resonances of the 1 nm gap are excited compared to only the first two FP resonances of the 10 nm gap. Then, we set the gap to 1 nm and evaluate the field enhancement profile for a 150 and 500 nm gold film, see Fig. 3 (solid and dotted). Thicker films squeeze FP resonances to lower frequencies and exhibit lower enhancements, see that below 300 THz the first seven FP resonances of the 500 nm film are excited compared to only the first two FP resonances of the 150 nm film. Hence, we can conclude that narrower slits and thicker films cluster the FP resonances towards the lower end of the spectrum.

We now focus solely on FP1, whose peak field enhancement Q and transmittance as a function of the film thickness are depicted in Fig. 4 for several gaps using only the nonlocal model ($b = 1$), since the results for the Drude model are qualitatively similar. For a given thickness, smaller

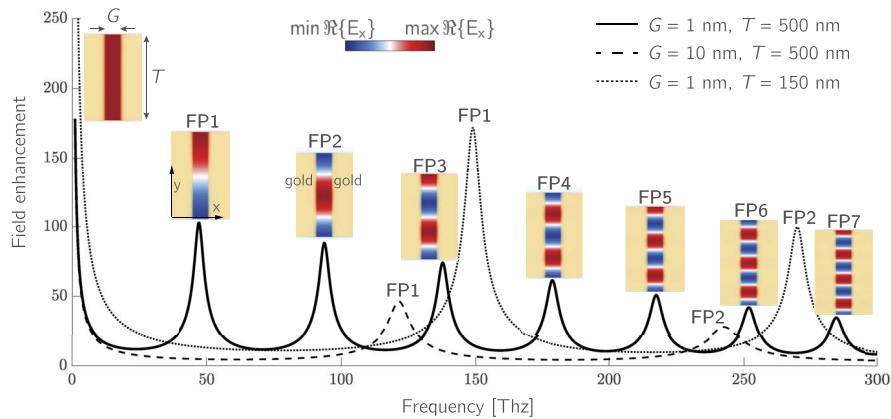


Fig. 3. Field enhancement Q for frequencies between 1 and 300 THz with three different gap/thickness combinations and $b = 1$. For 1 nm gap and 500 nm film thickness, we show $\Re\{E_x\}$ at several FP resonances with gold film highlighted to illustrate the nature of the resonant modes.

gaps exhibit more enhancement (since there is more confinement) and less transmittance (since there is less open area). If we account for thickness variation, the simulation results show that thinner films increase the field enhancement for gaps below 3 nm, and that the effect is reversed for gaps larger than 3 nm. On the other hand, transmittance is determined by both absorption—that augments as we consider thicker films—and field enhancement—whose behavior has just been described—and interact as follows: for gaps below 6 nm, film thinning from 500 to 150 nm leads to an increase of field enhancement (1.6-fold for the 1 nm gap) that is positively combined with absorption reduction, resulting in a transmittance boost (8-fold for the 1 nm gap); however, for gaps larger than 6 nm we observe that as the film thickens from 150 to 500 nm the increased absorption is outbalanced by the boost in field enhancement (3.3-fold for the 50 nm gap), which leads to more transmittance (2.5-fold for the 50 nm gap) for thicker films. To conclude, the combination of field enhancement and absorption, which interact constructively for narrow gaps and destructively for large gaps, drive transmittance.

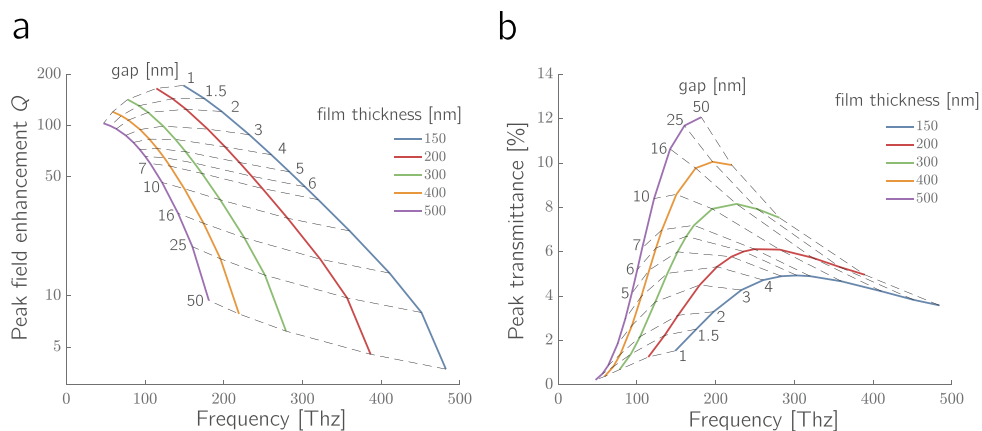


Fig. 4. Tracking of FP1 resonance of the nonlocal model ($b = 1$) for several gaps and thicknesses: (a) Field enhancement. (b) Transmittance.

Incorporating nonlocal electrodynamics in the simulation of the slit induces a blue-shift in the FP resonances with respect to local simulations, and the shift does not depend on the thickness of the gold film, see Figs. 5(a) and 5(b) for FP1. The resonance blue-shift in noble metals has already been identified in previously published work [27,31,64] as a relevant consequence of nonlocal effects. Furthermore, nonlocality is responsible of reducing the field enhancement and boosting the transmission, see Figs. 5(c) and 5(d). This behavior is directly linked to the penetration depth δ shown in Fig. 1(b). Indeed, the sharp decay of the electric field inside the metal is smoothed by nonlocality ($b > 0$), effectively enlarging the gap seen by the incident wave. This enlargement leads to lower enhancements (less confinement of the wave) and increased transmissions (larger open-area to propagate through), and since δ does not depend on the original aperture width, its effects become less noticeable as the gap widens – 5 nm seems to be the cutoff for pronounced nonlocal effects.

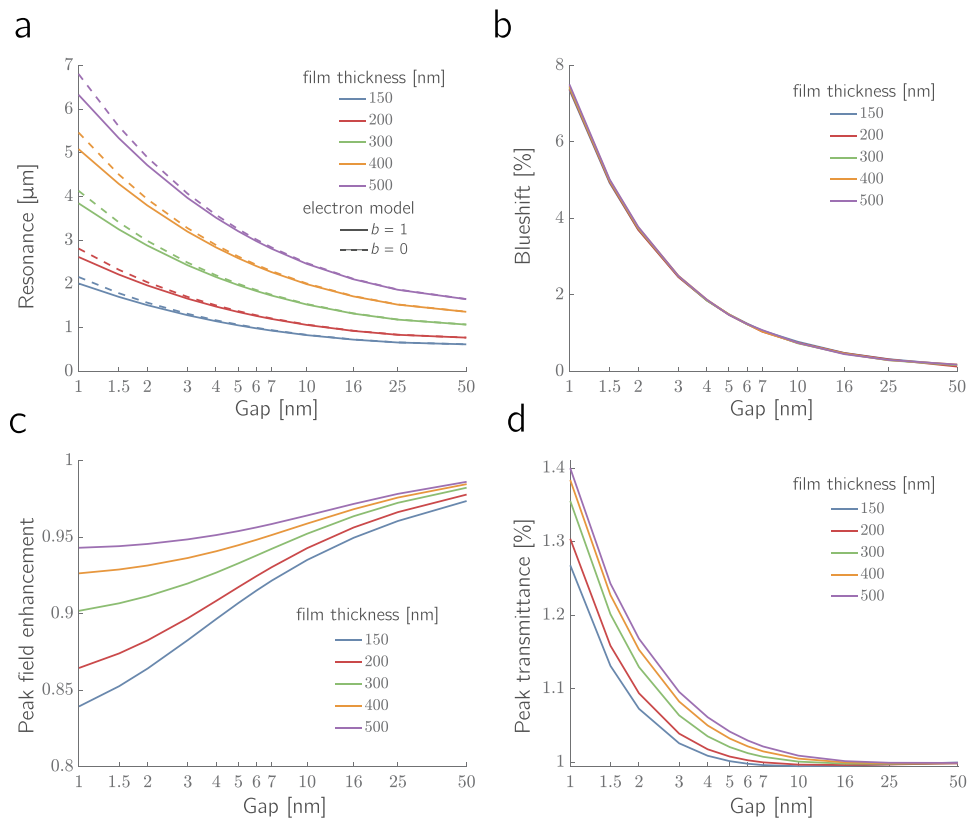


Fig. 5. Effects of nonlocality for all gaps and thicknesses: (a) FP1 resonant wavelengths. (b) Relative blue-shift with respect to FP1 local resonance. (c) Reduction in peak field enhancement at resonance. (d) Boosting in peak transmittance at resonance.

If instead of representing the resonant wavelength against the gap size as in Fig. 5(a), we use the film thickness T , a linear dependence $\lambda = \alpha_0 + \alpha_1 T$ arises. To that end, we use the resonant wavelength values from the simulations to regress the coefficients (α_0 , α_1), and show both the data (cross) and the linear fit (solid line) in Fig. 6(a), with coefficient of determination $R^2 > 0.9999$ for all gap widths and b values. The coefficient of determination is the proportion of variance in the resonant wavelength that may be predicted from the thickness value. Since only a tiny percentage of variance is not explained by the linear models, we may safely conclude the relationship between resonant wavelength and thickness is linear. The values of the linear fit

parameters for several gaps and b values are collected in Table 1 (in the Appendix), along with information on how R^2 is computed. Moreover, the dispersion relation for FPI can be derived following the results of [4] for a metal-insulator-metal waveguide. Namely, for each unique gap the wavelength of the propagating plasmon $\lambda_p = 2\pi/k$ along the slit at FPI is twice the thickness, hence the wave vector reads $k = \pi/T$. The dispersion relation is therefore expressed as $\lambda = \alpha_0 + \alpha_1 \pi/k$, and is shown in Fig. 6(b) for several gaps and electron models.

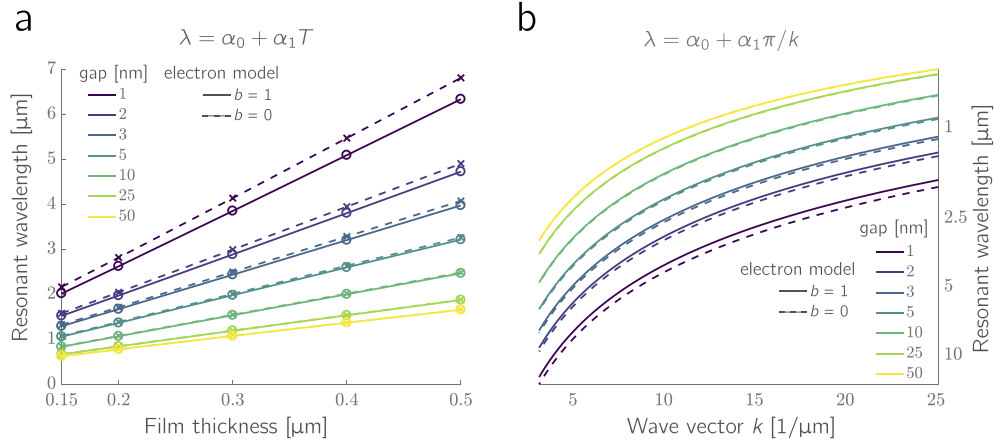


Fig. 6. FPI resonance analysis for both the Drude model and the hydrodynamic model with $b = 1$: (a) Computational data and linear fit ($R^2 > 0.9999$) of FPI wavelength vs film thickness. (b) Dispersion relation using the coefficients obtained in the resonance-thickness regression.

Table 1. Linear fit coefficients for $\lambda = \alpha_0 + \alpha_1 T$, where T is film thickness in μm and λ is the FPI resonant wavelength in μm .

G [nm]	Nonlocal auxiliary factor b							
	0		0.5		1		1.5	
	α_0	α_1	α_0	α_1	α_0	α_1	α_0	α_1
1	0.162	13.273	0.157	12.591	0.155	12.340	0.154	12.158
1.5	0.150	10.905	0.148	10.522	0.145	10.380	0.145	10.270
2	0.144	9.505	0.142	9.254	0.141	9.155	0.140	9.083
3	0.138	7.862	0.136	7.724	0.135	7.669	0.135	7.626
4	0.135	6.898	0.134	6.808	0.133	6.770	0.133	6.742
5	0.134	6.249	0.133	6.182	0.132	6.158	0.132	6.137
6	0.134	5.776	0.133	5.726	0.132	5.705	0.132	5.690
7	0.133	5.415	0.133	5.374	0.132	5.357	0.132	5.346
10	0.135	4.694	0.134	4.67	0.134	4.659	0.134	4.652
16	0.141	3.963	0.14	3.95	0.140	3.946	0.139	3.941
25	0.151	3.458	0.150	3.45	0.150	3.448	0.150	3.445
50	0.184	2.956	0.184	2.952	0.184	2.950	0.183	2.951

3. Conclusions

This computational study shows that nonlocal description of electron motion in metals has a significant effect on the optical properties of an ultra-narrow slit in the THz and MIR regimes.

First, for frequencies in the low-THz regime we extend the field saturation studies in [47] to both the local and nonlocal model, as well as to thinner films and narrower gaps. We have numerically verified that, for a given frequency and film thickness, field enhancement saturates as the gap is reduced and approaches the limiting enhancement value proposed in [47]. We have demonstrated that nonlocal saturation is less pronounced, since the theoretical limiting enhancement values are similar in both cases whereas nonlocal enhancement is significantly weaker due to the effective gap enlargement. In addition, we have numerically shown the impact of the parameter β on the field penetration at the metal-dielectric interface that characterizes nonlocality, and verified that the numerical values for the skin depth coincide with the analytical ones.

Furthermore, for MIR frequencies we observe a blue-shift of the gap plasmon for nonlocal calculations with respect to Drude model predictions. In addition, the nonlocal model predicts a decrease in enhancement and a boost in transmittance, due to the effectively increased gap width that stems from the nonzero penetration depth of the electric field into the metal. These results are consistent with previously published work [27,31], whereby a nonlocal representation of electron motion is indispensable to achieve accurate predictions for gaps below 5 nm.

On the numerical perspective, we have shown that the HDG method equipped with the hydrodynamic model for methods is a powerful scheme to simulate plasmonic structures. Indeed, it enables us to resolve the squeezing of millimeter-long wavelengths into nanometer-wide cavities, along with the atomic-scale charge distribution layers that characterize nonlocal electrodynamics. This is even more dramatic for the THz regime, where both high order solvers and adaptive anisotropic meshes are required to accurately represent the solution in the presence of such lengthscale mismatch. The results in this paper showcase the strength of the HDG methods towards more efficient simulations of extreme-scale nanophotonic phenomena - terahertz nonlocality. Future applications of the HDG method include investigations of THz optoelectronics [39], nonlocality in the nonlinear regime [65] and extreme field confinement via polaritons in two-dimensional materials [66–68].

Appendix A. Hybridizable discontinuous Galerkin method

A.1. Governing equations

The hybridizable discontinuous Galerkin (HDG) method [41,43,45] is particularly suited to simulate wave propagation problems for a number of reasons: (1) it can be used on unstructured and anisotropic meshes, thus enabling the simulation of complex geometries; (2) it has low dissipation and dispersion, and the numerical error may be reduced to the level where it is insensitive to the geometry discretization; (3) it gives rise to a linear system that contains a reduced number of degrees of freedom defined on the faces of the discretization; and (4) enables the natural treatment of both boundary conditions and material contrasts at interfaces.

In this section, we describe the equations used in the simulation of the infinite 2-D slit shown in Fig. 7(a). The HDG method was used to simulate the full 2-D electromagnetic response on a computational domain $\Omega \subset \mathbb{R}^2$ that contains the slit. The domain Ω can be split according to whether the material is free space Ω_f or a metal Ω_m , thus $\Omega = \Omega_f \cup \Omega_m$, with boundary $\partial\Omega = \Gamma$. In Ω_f the scattering is governed by the time-harmonic Maxwell's equations for a single frequency ω , that is

$$\nabla \times \mathbf{E} - i\omega\mathbf{H} = 0, \quad \nabla \times \mathbf{H} + i\omega\mathbf{E} = 0. \quad (2)$$

namely the atomically-thin charge distribution layer. The histogram of the mesh elements' areas, as well as several mesh details for 0.2 THz (the one with largest discrepancy between lengthscales) are depicted in Fig. 8. It can be observed that the mesh elements' areas span twelve orders of magnitude, thus highlighting the difficulties of the simulations and the necessity of high-order methods and anisotropic meshes to properly capture the solution.

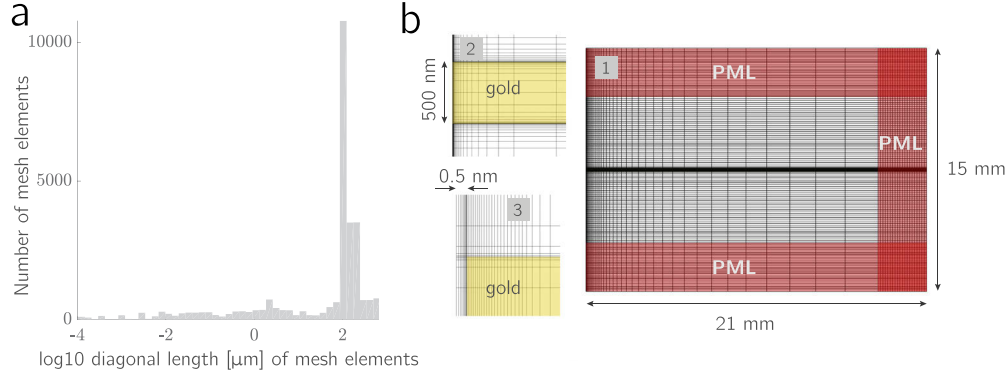


Fig. 8. Mesh for $G = 1$ nm and 0.2 THz to illustrate anisotropy: (a) Histogram of the log10 diagonal length of the mesh elements. (b) General mesh view, highlighting PML and gold regions, with three different levels of zoom.

Transmittance is computed as the ratio between the transmitted power across a y -constant half-plane A located 1 micron above the gold film with $x \in [0, 500]$ nm and the incident power across a y -constant plane A_0 located 1 micron below the gold film with $x \in [0, 500]$ nm, namely

$$\frac{\int_{A_1} |\Re [\mathbf{E} \times \mathbf{H}^*] \cdot \mathbf{n}|}{\int_{A_0} |\Re [\mathbf{E}_{\text{inc}} \times \mathbf{H}_{\text{inc}}^*] \cdot \mathbf{n}|}, \quad (4)$$

The amplitude enhancement Q for on the upper tip of the gap S , similarly computed as the following ratio

$$Q = \frac{\int_{S_1} |E_x|}{\int_{S_1} |E_{\text{inc}}|}. \quad (5)$$

Appendix B. Coefficients for linear fit of the dispersion relation

In this section, we report the coefficients α_0, α_1 that govern the dispersion relation (λ, k) for the slit FP1 resonance, arising from the linear fit between resonant wavelength and film thickness, see Fig. 6 in the manuscript. The coefficients are collected in Table 1 for several gaps and electron models, $b = 0, 0.5, 1, 1.5$.

As mentioned in the manuscript, the R^2 coefficient of the fit is above 0.9999 for all cases. To calculate this coefficient, for each gap and electron model b , we numerically obtain the resonant wavelength for each of the five thickness values considered (150, 200, 300, 400 and 500 nm), which can be expressed as five tuples $(T_i, \lambda_i)_{i=1}^5$. The coefficient of determination is therefore computed as

$$R^2 = 1 - \frac{\sum_{i=1}^5 (\mathcal{L}(T_i) - \lambda_i)^2}{\sum_{i=1}^5 (\bar{\lambda} - \lambda_i)^2} \quad (6)$$

where $\bar{\lambda}$ is the mean value of resonant wavelength for the particular gap and b , that is $\bar{\lambda} = \sum_{i=1}^5 \lambda_i / 5$, and $\mathcal{L}(T_i)$ is the model that we use, in this case a linear fit $\mathcal{L}(T_i) = \alpha_0 + \alpha_1 T_i$. Note the R^2 value

ranges between $-\infty$ (the fit can be arbitrarily bad) to 1 (perfect fit), and that $R^2 = 0$ corresponds to using a model that always predicts the mean value $\bar{\lambda}$. We can thus conclude that our linear fit it is almost perfect.

Funding

Ministerio de Ciencia, Innovación y Universidades (MAT2017-88358-C3-1-R); Gobierno de Aragón (Q-MADS); National Science Foundation (EECS 1809240, EECS 1809723); Air Force Office of Scientific Research (FA9550-16-1-0214, FA9550-19-1-0240).

Acknowledgments

S-H.O. acknowledges support the Sanford P. Bordeau Endowed Chair at the University of Minnesota.

F.V.-C., N.-C.N., and J.P. were responsible for the development of the HDG method and all computer simulations. L.M.-M., C.C., D.Y. and S.-H.O. were responsible for the theoretical analysis. All authors contributed to writing the paper.

Disclosures

The authors declare no conflicts of interest.

References

1. S. A. Maier, *Plasmonics: Fundamentals and Applications* (Springer, 2007).
2. S. Lal, S. Link, and N. J. Halas, "Nano-optics from sensing to waveguiding," *Nat. Photonics* **1**(11), 641–648 (2007).
3. J. Dionne, L. Sweatlock, H. Atwater, and A. Polman, "Plasmon slot waveguides: Towards chip-scale propagation with subwavelength-scale localization," *Phys. Rev. B* **73**(3), 035407 (2006).
4. H. T. Miyazaki and Y. Kurokawa, "Squeezing visible light waves into a 3-nm-thick and 55-nm-long plasmon cavity," *Phys. Rev. Lett.* **96**(9), 097401 (2006).
5. J. Bravo-Abad, L. Martín-Moreno, and F. García-Vidal, "Transmission properties of a single metallic slit: From the subwavelength regime to the geometrical-optics limit," *Phys. Rev. E* **69**(2), 026601 (2004).
6. D. R. Ward, N. K. Grady, C. S. Levin, N. J. Halas, Y. Wu, P. Nordlander, and D. Natelson, "Electromigrated nanoscale gaps for surface-enhanced Raman spectroscopy," *Nano Lett.* **7**(5), 1396–1400 (2007).
7. J. J. Baumberg, J. Aizpurua, M. H. Mikkelsen, and D. R. Smith, "Extreme nanophotonics from ultrathin metallic gaps," *Nat. Mater.* **18**(7), 668–678 (2019).
8. A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Planar photonics with metasurfaces," *Science* **339**(6125), 1232009 (2013).
9. J. Y. Suh, C. H. Kim, W. Zhou, M. D. Huntington, D. T. Co, M. R. Wasielewski, and T. W. Odom, "Plasmonic bowtie nanolaser arrays," *Nano Lett.* **12**(11), 5769–5774 (2012).
10. C. T. Ertsgaard, N. J. Wittenberg, D. J. Klemme, A. Barik, W.-C. Shih, and S.-H. Oh, "Integrated nanogap platform for sub-volt dielectrophoretic trapping and real-time raman imaging of biological nanoparticles," *Nano Lett.* **18**(9), 5946–5953 (2018).
11. F. J. García-Vidal, L. Martín-Moreno, T. W. Ebbesen, and L. Kuipers, "Light passing through subwavelength apertures," *Rev. Mod. Phys.* **82**(1), 729–787 (2010).
12. H. Im, K. C. Bantz, N. C. Lindquist, C. L. Haynes, and S.-H. Oh, "Vertically oriented sub-10-nm plasmonic nanogap arrays," *Nano Lett.* **10**(6), 2231–2236 (2010).
13. N. C. Lindquist, P. Nagpal, K. M. McPeak, D. J. Norris, and S.-H. Oh, "Engineering metallic nanostructures for plasmonics and nanophotonics," *Rep. Prog. Phys.* **75**(3), 036501 (2012).
14. X. Chen, H.-R. Park, M. Pelton, X. Piao, N. C. Lindquist, H. Im, Y. J. Kim, J. S. Ahn, K. J. Ahn, N. Park, D.-S. Kim, and S.-H. Oh, "Atomic layer lithography of wafer-scale nanogap arrays for extreme confinement of electromagnetic waves," *Nat. Commun.* **4**(1), 2361 (2013).
15. D. Yoo, N. C. Nguyen, L. Martín-Moreno, D. A. Mohr, S. Carretero-Palacios, J. Shaver, J. Peraire, T. W. Ebbesen, and S.-H. Oh, "High-throughput fabrication of resonant metamaterials with ultrasmall coaxial apertures via atomic layer lithography," *Nano Lett.* **16**(3), 2040–2046 (2016).
16. L. Martín-Moreno, F. J. García-Vidal, H. J. Lezec, K. M. Pellerin, T. Thio, J. B. Pendry, and T. W. Ebbesen, "Theory of extraordinary optical transmission through subwavelength hole arrays," *Phys. Rev. Lett.* **86**(6), 1114–1117 (2001).
17. S. G. Rodrigo, S. Carretero-Palacios, F. García-Vidal, and L. Martín-Moreno, "Metallic slit arrays filled with third-order nonlinear media: Optical kerr effect and third-harmonic generation," *Phys. Rev. B* **83**(23), 235425 (2011).

18. J. B. Lassiter, X. Chen, X. Liu, C. Ciraci, T. B. Hoang, S. Larouche, S.-H. Oh, M. H. Mikkelsen, and D. R. Smith, "Third-harmonic generation enhancement by film-coupled plasmonic stripe resonators," *ACS Photonics* **1**(11), 1212–1217 (2014).
19. X. Chen, C. Ciraci, D. R. Smith, and S.-H. Oh, "Nanogap-enhanced infrared spectroscopy with template-stripped wafer-scale arrays of buried plasmonic cavities," *Nano Lett.* **15**(1), 107–113 (2015).
20. L. Dong, X. Yang, C. Zhang, B. Cerjan, L. Zhou, M. L. Tseng, Y. Zhang, A. Alabastri, P. Nordlander, and N. J. Halas, "Nanogapped au antennas for ultrasensitive surface-enhanced infrared absorption spectroscopy," *Nano Lett.* **17**(9), 5768–5774 (2017).
21. J. Zuloaga, E. Prodan, and P. Nordlander, "Quantum description of the plasmon resonances of a nanoparticle dimer," *Nano Lett.* **9**(2), 887–891 (2009).
22. N. J. Halas, S. Lal, W.-S. Chang, S. Link, and P. Nordlander, "Plasmons in strongly coupled metallic nanostructures," *Chem. Rev.* **111**(6), 3913–3961 (2011).
23. R. Esteban, A. G. Borisov, P. Nordlander, and J. Aizpurua, "Bridging quantum and classical plasmonics with a quantum-corrected model," *Nat. Commun.* **3**(1), 825 (2012).
24. H. Duan, A. I. Fernández-Domínguez, M. Bosman, S. A. Maier, and J. K. Yang, "Nanoplasmonics: classical down to the nanometer scale," *Nano Lett.* **12**(3), 1683–1689 (2012).
25. D. C. Marinica, M. Zapata, P. Nordlander, A. K. Kazansky, P. M. Echenique, J. Aizpurua, and A. G. Borisov, "Active quantum plasmonics," *Sci. Adv.* **1**(11), e1501095 (2015).
26. R. Gordon and A. Ahmed, "Reaching the limits of enhancement in (sub) nanometer metal structures," *ACS Photonics* **5**(11), 4222–4228 (2018).
27. C. Ciraci, R. T. Hill, J. J. Mock, Y. Urzhumov, A. I. Fernández-Domínguez, S. A. Maier, J. B. Pendry, A. Chilkoti, and D. R. Smith, "Probing the ultimate limits of plasmonic enhancement," *Science* **337**(6098), 1072–1074 (2012).
28. C. Ciraci, X. Chen, J. J. Mock, F. McGuire, X. Liu, S.-H. Oh, and D. R. Smith, "Film-coupled nanoparticles by atomic layer deposition: Comparison with organic spacing layers," *Appl. Phys. Lett.* **104**(2), 023109 (2014).
29. J. B. Pendry, A. Aubry, D. R. Smith, and S. A. Maier, "Transformation optics and subwavelength control of light," *Science* **337**(6094), 549–552 (2012).
30. A. Wiener, A. I. Fernández-Domínguez, A. P. Horsfield, J. B. Pendry, and S. A. Maier, "Nonlocal effects in the nanofocusing performance of plasmonic tips," *Nano Lett.* **12**(6), 3308–3314 (2012).
31. W. Zhu, R. Esteban, A. G. Borisov, J. J. Baumberg, P. Nordlander, H. J. Lezec, J. Aizpurua, and K. B. Crozier, "Quantum mechanical effects in plasmonic structures with subnanometre gaps," *Nat. Commun.* **7**(1), 11495 (2016).
32. F. J. García de Abajo, "Nonlocal effects in the plasmons of strongly interacting nanoparticles, dimers, and waveguides," *J. Phys. Chem. C* **112**(46), 17983–17987 (2008).
33. C. Huck, F. Neubrech, J. Vogt, A. Toma, D. Gerbert, J. Katzmann, T. Härtling, and A. Pucci, "Surface-enhanced infrared spectroscopy using nanometer-sized gaps," *ACS Nano* **8**(5), 4908–4914 (2014).
34. M. Seo, H. Park, S. Koo, D. Park, J. Kang, O. Suwal, S. Choi, P. Planken, G. Park, N. Park, Q. Park, and D. Kim, "Terahertz field enhancement by a metallic nano slit operating beyond the skin-depth limit," *Nat. Photonics* **3**(3), 152–156 (2009).
35. X. Chen, H.-R. Park, M. Pelton, X. Piao, N. C. Lindquist, H. Im, Y. J. Kim, J. S. Ahn, K. J. Ahn, N. Park, D.-S. Kim, and S.-H. Oh, "Atomic layer lithography of wafer-scale nanogap arrays for extreme confinement of electromagnetic waves," *Nat. Commun.* **4**(1), 2361 (2013).
36. W. Gao, J. Shu, K. Reichel, D. V. Nickel, X. He, G. Shi, R. Vajtai, P. M. Ajayan, J. Kono, D. M. Mittleman, and Q. Xu, "High-contrast terahertz wave modulation by gated graphene enhanced by extraordinary transmission through ring apertures," *Nano Lett.* **14**(3), 1242–1248 (2014).
37. H.-R. Park, S. Namgung, X. Chen, N. C. Lindquist, V. Giannini, Y. Francescato, S. A. Maier, and S.-H. Oh, "Perfect extinction of terahertz waves in monolayer graphene over 2-nm-wide metallic apertures," *Adv. Opt. Mater.* **3**(5), 667–673 (2015).
38. B. C. Pein, W. Chang, H. Y. Hwang, J. Scherer, I. Coropceanu, X. Zhao, X. Zhang, V. Bulovic, M. Bawendi, and K. A. Nelson, "Terahertz-driven luminescence and colossal Stark effect in CdSe–CdS colloidal quantum dots," *Nano Lett.* **17**(9), 5375–5380 (2017).
39. B. C. Pein, C. K. Lee, L. Shi, J. Shi, W. Chang, H. Y. Hwang, J. Scherer, I. Coropceanu, X. Zhao, X. Zhang, V. Bulovic, M. G. Bawendi, A. P. Willard, and K. A. Nelson, "Terahertz-driven Stark spectroscopy of CdSe and CdSe: CdS core: shell quantum dots," *Nano Lett.* **19**(11), 8125–8131 (2019).
40. J. D. Caldwell, L. Lindsay, V. Giannini, I. Vurgaftman, T. L. Reinecke, S. A. Maier, and O. J. Glembocki, "Low-loss, infrared and terahertz nanophotonics using surface phonon polaritons," *Nanophotonics* **4**(1), 44–68 (2015).
41. N. C. Nguyen, J. Peraire, and B. Cockburn, "Hybridizable discontinuous Galerkin methods for the time-harmonic Maxwell's equations," *J. Comput. Phys.* **230**(19), 7151–7175 (2011).
42. H.-R. Park, X. Chen, N. C. Nguyen, J. Peraire, and S.-H. Oh, "Nanogap-enhanced terahertz sensing of 1 nm thick ($\lambda/10^6$) dielectric films," *ACS Photonics* **2**(3), 417–424 (2015).
43. D. Yoo, F. Vidal-Codina, C. Ciraci, N.-C. Nguyen, D. R. Smith, J. Peraire, and S.-H. Oh, "Modeling and observation of mid-infrared nonlocality in effective epsilon-near-zero ultranarrow coaxial apertures," *Nat. Commun.* **10**(1), 4476 (2019).
44. L. Li, S. Lanteri, N. A. Mortensen, and M. Wubs, "A hybridizable discontinuous Galerkin method for solving nonlocal optical response models," *Comput. Phys. Commun.* **219**, 99–107 (2017).

45. F. Vidal-Codina, N.-C. Nguyen, S.-H. Oh, and J. Peraire, "A hybridizable discontinuous Galerkin method for computing nonlocal electromagnetic effects in three-dimensional metallic nanostructures," *J. Comput. Phys.* **355**, 548–565 (2018).
46. J. Lin and H. Zhang, "Scattering and field enhancement of a perfect conducting narrow slit," *SIAM J. Appl. Math.* **77**(3), 951–976 (2017).
47. J. Lin, S.-H. Oh, H.-M. Nguyen, and F. Reitich, "Field enhancement and saturation of millimeter waves inside a metallic nanogap," *Opt. Express* **22**(12), 14402–14410 (2014).
48. G. Toscano, S. Raza, A.-P. Jauho, N. A. Mortensen, and M. Wubs, "Modified field enhancement and extinction by plasmonic nanowire dimers due to nonlocal response," *Opt. Express* **20**(4), 4176–4188 (2012).
49. S. Raza, S. I. Bozhevolnyi, M. Wubs, and N. A. Mortensen, "Nonlocal optical response in metallic nanostructures," *J. Phys.: Condens. Matter* **27**(18), 183204 (2015).
50. K. R. Hiremath, L. Zschiedrich, and F. Schmidt, "Numerical solution of nonlocal hydrodynamic Drude model for arbitrary shaped nano-plasmonic structures using Nédélec finite elements," *J. Comput. Phys.* **231**(17), 5890–5896 (2012).
51. G. Toscano, J. Straubel, A. Kwiatkowski, C. Rockstuhl, F. Evers, H. Xu, N. A. Mortensen, and M. Wubs, "Resonance shifts and spill-out effects in self-consistent hydrodynamic nanoplasmonics," *Nat. Commun.* **6**(1), 7132 (2015).
52. N. Schmitt, C. Scheid, J. Viquerat, and S. Lanteri, "Simulation of three-dimensional nanoscale light interaction with spatially dispersive metals using a high order curvilinear DGTD method," *J. Comput. Phys.* **373**, 210–229 (2018).
53. M. Moeferd, T. Kiel, T. Sproll, F. Intravaia, and K. Busch, "Plasmonic modes in nanowire dimers: A study based on the hydrodynamic drude model including nonlocal and nonlinear effects," *Phys. Rev. B* **97**(7), 075431 (2018).
54. K. Busch, M. König, and J. Niegemann, "Discontinuous Galerkin methods in nanophotonics," *Laser Photonics Rev.* **5**(6), 773–809 (2011).
55. J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.* **114**(2), 185–200 (1994).
56. S. G. Johnson, "Notes on perfectly matched layers (pmls)," Lect. notes, Mass. Inst. Technol. Mass. **29** (2008).
57. J. Lin and F. Reitich, "Electromagnetic field enhancement in small gaps: a rigorous mathematical theory," *SIAM J. Appl. Math.* **75**(5), 2290–2310 (2015).
58. Y.-M. Bahk, S. Han, J. Rhie, J. Park, H. Jeon, N. Park, and D.-S. Kim, "Ultimate terahertz field enhancement of single nanoslits," *Phys. Rev. B* **95**(7), 075424 (2017).
59. A. Moreau, C. Ciraci, and D. R. Smith, "Impact of nonlocal response on metalodielectric multilayers and optical patch antennas," *Phys. Rev. B* **87**(4), 045401 (2013).
60. E. J. Dias, D. A. Iranzo, P. Gonçalves, Y. Hajati, Y. V. Bludov, A.-P. Jauho, N. A. Mortensen, F. H. Koppens, and N. Peres, "Probing nonlocal effects in metals with graphene plasmons," *Phys. Rev. B* **97**(24), 245405 (2018).
61. C. Ciraci, J. B. Pendry, and D. R. Smith, "Hydrodynamic model for plasmonics: a macroscopic approach to a microscopic problem," *ChemPhysChem* **14**(6), 1109–1116 (2013).
62. T. V. Teperik, P. Nordlander, J. Aizpurua, and A. G. Borisov, "Robust subnanometric plasmon ruler by rescaling of the nonlocal optical response," *Phys. Rev. Lett.* **110**(26), 263901 (2013).
63. Y. Takakura, "Optical resonance in a narrow slit in a thick metallic screen," *Phys. Rev. Lett.* **86**(24), 5601–5603 (2001).
64. S. Raza, N. Stenger, S. Kadkhodazadeh, S. V. Fischer, N. Kotesha, A.-P. Jauho, A. Burrows, M. Wubs, and N. A. Mortensen, "Blueshift of the surface plasmon resonance in silver nanoparticles studied with eels," *Nanophotonics* **2**(2), 131–138 (2013).
65. C. Ciraci, M. Scalora, and D. R. Smith, "Third-harmonic generation in the presence of classical nonlocal effects in gap-plasmon nanostructures," *Phys. Rev. B* **91**(20), 205403 (2015).
66. T. Low, A. Chaves, J. D. Caldwell, A. Kumar, N. X. Fang, P. Avouris, T. F. Heinz, F. Guinea, L. Martín-Moreno, and F. Koppens, "Polaritons in layered two-dimensional materials," *Nat. Mater.* **16**(2), 182–194 (2017).
67. D. A. Iranzo, S. Nanot, E. J. Dias, I. Epstein, C. Peng, D. K. Efetov, M. B. Lundeberg, R. Parret, J. Osmond, J.-Y. Hong, J. Kong, D. R. Englund, N. M. R. Peres, and F. H. L. Koppens, "Probing the ultimate plasmon confinement limits with a van der waals heterostructure," *Science* **360**(6386), 291–295 (2018).
68. I.-H. Lee, D. Yoo, P. Avouris, T. Low, and S.-H. Oh, "Graphene acoustic plasmon resonator for ultrasensitive infrared spectroscopy," *Nat. Nanotechnol.* **14**(4), 313–319 (2019).
69. M. A. Ordal, R. J. Bell, R. W. Alexander, L. L. Long, and M. R. Querry, "Optical properties of fourteen metals in the infrared and far infrared: Al, Co, Cu, Au, Fe, Pb, Mo, Ni, Pd, Pt, Ag, Ti, V, and W.," *Appl. Opt.* **24**(24), 4493–4499 (1985).