Extremely confined gap plasmon modes: when nonlocality matters

Sergejs Boroviks [1,2,3], Zhan-Hong Lin [2], Vladimir A. Zenin [1], Mario Ziegler [2], Andrea Dellith[2], P. A. D. Gonçalves [1], Christian Wolff [1], Sergey I. Bozhevolnyi [1,4], Jer-Shing Huang [2,5,6,7] & N. Asger Mortensen [1,4 ✉]

Historically, the field of plasmonics has been relying on the framework of classical electrodynamics, with the local-response approximation of material response being applied even when dealing with nanoscale metallic structures. However, when the confinement of electromagnetic radiation approaches atomic scales, mesoscopic effects are anticipated to become observable, e.g., those associated with the nonlocal electrodynamic surface response of the electron gas. Here, we investigate nonlocal effects in propagating gap surface plasmon modes in ultrathin metal–dielectric–metal planar waveguides, exploiting monocrystalline gold flakes separated by atomic-layer-deposited aluminum oxide. We use scanning near-field optical microscopy to directly access the near-field of such confined gap plasmon modes and measure their dispersion relation via their complex-valued propagation constants. We compare our experimental findings with the predictions of the generalized nonlocal optical response theory to unveil signatures of nonlocal damping, which becomes appreciable for few-nanometer-sized dielectric gaps.

[1] Center for Nano Optics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. [2] Leibniz Institute of Photonic Technology, Albert-Einstein Straße 9, 07745 Jena, Germany. [3] Nanophotonics and Metrology Laboratory, Swiss Federal Institute of Technology Lausanne (EPFL), Station 11, CH 1015 Lausanne, Switzerland. [4] Danish Institute for Advanced Study, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. [5] Institute of Physical Chemistry and Abbe Center of Photonics, Friedrich-Schiller-Universität Jena, 07743 Jena, Germany. [6] Research Center for Applied Sciences, Academia Sinica, 128 Sec. 2, Academia Road, Nankang District 11529 Taipei, Taiwan. [7] Department of Electrophysics, National Yang Ming Chiao Tung University, 1001 University Road, 30010 Hsinchu, Taiwan. ✉email: asger@mailaps.org

One of the appealing features of plasmonics—the possibility to squeeze light beyond the diffraction limit and guide electromagnetic energy in waveguides with sub-wavelength dimensions[1-3]—has remained at the heart of the nanophotonics community for several decades[4]. Such tight focusing is only possible due to unique properties of the surface-plasmon polariton (SPP), being a collective oscillation of the free charge-carriers in metals coupled with light fields[5]. Various configurations, designs, and materials for plasmonic systems have been studied from both fundamental and applied perspectives, attempting to realize a diversity of functional devices, ranging from plasmonic integrated circuitry[6] to metasurface-based flat optical components[7].

In particular, many devices are based on metal–dielectric–metal (MDM) heterostructures, which support gap surface plasmon (GSP) modes[8]. GSPs can lead to substantial electric-field enhancements inside the dielectric gap, which can strongly enhance linear and nonlinear optical processes[9], and may be exploited for various applications of plasmon-enhanced light–matter interactions, e.g., surface-enhanced Raman spectroscopy (SERS)[10] or Purcell enhancement of the emission rate of single-photon sources[11,12]. Furthermore, MDM heterostructures may be engineered to achieve larger mode propagation lengths with better field confinements when compared with other SPP modes[13,14]. The GSP modes supported by such waveguides were extensively studied by various experimental methods, including far-field techniques[15,16] and scanning near-field optical microscopy (SNOM)[17,18].

The above-mentioned ambitions, along with the promise of realizing plasmonic waveguides for integration, or even replacement of conventional photonic waveguides, largely rely on the aspiration of significantly mitigating Ohmic losses inherent to metals[19,20]. Indeed, some progress in the development of traditional and alternative plasmonic materials has been made[21,22]. Moreover, with advances in colloidal synthesis methods of traditional plasmonic materials, it became possible to grow mono-crystalline gold (Au) flakes with high aspect ratio—few tens of nanometers in thickness and up to several hundreds of microns in lateral size[23,24]. Along with the improvement in nanofabrication techniques, such as focused ion beam (FIB) milling, these progresses allowed the fabrication of plasmonic nanocircuitry[25,26] and plasmonic nanoantennas[27,28] with superior quality. Furthermore, monocrystalline Au flakes present as a "playground" material platform for experimental studies of fundamental aspects of plasmonics[29-36].

Thus far, most of the novel aspects in plasmonics have emerged and developed from considerations rooted in classical electrodynamics and the local-response approximation (LRA) for the interaction of light with the free-electrons in metals[5]. However, in recent years there has been increasing attention to quantum plasmonics[37-39] and the importance of quantum corrections to classical electrodynamics in plasmonic nanostructures[40-47], including nonlocal effects impacting GSP modes in MDM structures with ultranarrow gaps[48,49] and hyperbolic metamaterials[50]. This has contributed to a general recognition that in addition to bulk losses in plasmonic metals, there is an additional contribution in metallic surfaces associated with nonlocal effects[20,41,45].

The dominant microscopic mechanism contributing to non-local losses is Landau damping[51], which becomes especially pronounced at large plasmonic mode propagation vectors, and thus manifesting also in MDM structures with very small dielectric gaps. While ab initio approaches in principle account for such quantum nonlocal effects[52], quantum corrections to the LRA may also be explored semi-analytically using either hydro-dynamic models or a surface-response formalism (SRF)[45]. In particular, the use of Feibelman d-parameters[53] has recently been revived in the context of nanogap structures and related plasmonic phenomena[44,54-59]. At the same time, there has been modest progress in experimental investigations of extremely confined propagating GSP modes as far as nonlocal corrections to the GSP dispersion in nanometer-sized gaps are concerned. The most relevant characterization of GSP modes is associated with observations of large splitting of symmetric and antisymmetric eigenmodes in side-by-side aligned single-crystalline Au nanorod dimers with atomically defined gaps reaching ~5 Å[60]. While the experimentally measured resonance wavelengths could be accounted for without involving quantum nonlocal effects, experimentally measured resonance quality factors for GSP-based resonances were noted to be significantly lower than those predicted by the classical, local-response theory[60]. This qualitative observation tallies well with the analysis of the Landau damping influence on the GSP dispersion for nanometer-sized gaps that revealed little difference between local and nonlocal considerations of the real part of the GSP propagation constant, while showing progressively strong increase of the GSP propagation loss for gaps below 10 nm[51]. Given the importance of GSP-based configurations for a wide range of plasmon-mediated light–matter interactions[9], it is crucial to experimentally establish a benchmark for the nonlocal (i.e., Landau or surface collision) damping associated with extremely confined GSP modes.

In this work, we present an experimental study of extremely confined GSP modes in MDM structures fabricated out of monocrystalline Au flakes and atomic-layer-deposited (ALD) ultrathin aluminum oxide (Al_2O_3) films. The use of crystalline metal and high-quality dielectric material is crucial for reducing bulk and surface-roughness related losses to the minimum, thus opening a doorway to explore nonlocality[48]. The concept of the experiment is schematically illustrated in Fig. 1. Using a scattering-type SNOM (s-SNOM)[61], we obtain near-field (NF) maps of propagating GSP modes, which exhibit, to the best of our knowledge, a record-high experimentally demonstrated effective-mode index, reaching values of approximately 6.2 in the case of a ~2 nm-thick gap at $\lambda_0 = 1550$ nm excitation wavelength. Further analysis of the experimental data from a range of samples suggests signatures of gap-dependent nonclassical damping, being especially pronounced in samples with dielectric gap thicknesses
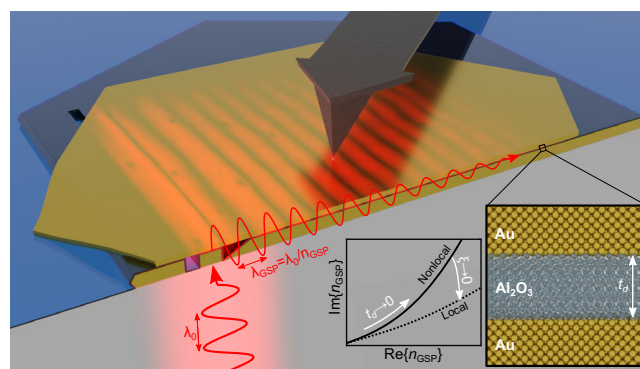
Fig. 1 Schematic illustration of the experiment. The red curves schematically illustrate the electric-field profile of the excitation light (bottom left) and of the propagating GSP mode (in the sample) with the group-index n_{GSP} (corresponding to a GSP-wavelength λ_{GSP} at the free-space wavelength λ_0). Insets show the parametric plot (varying the gap-size, t_d) of the GSP dispersion trends as the nonlocal length scale $\xi \longrightarrow 0$, while at the bottom right it is shown a close-up of the MDM waveguide comprised of the monocrystalline gold (Au) flakes separated by a thin dielectric gap formed by atomic-layer deposition of aluminum oxide (Al_2O_3).

smaller than 5 nm. We show that such observations are compatible with a nonclassical interpretation in terms of the so-called generalized nonlocal optical response (GNOR) theory[41], and provide an estimate of the diffusion constant that accounts for nonlocal damping. Incidentally, although the microscopic origin of the diffusion constant in GNOR represents carrier-scattering, it can also impersonate other sources of nonlocal damping (e.g., Landau damping) in a phenomenological fashion.

Results

Quantum nonlocal corrections to the GSP dispersion relation. The dispersion relation of the fundamental GSP mode in a MDM structure is generally given by[48]

$$\tanh\left(\frac{\kappa_{\mathrm{d}} t_{\mathrm{d}}}{2}\right) = -\frac{\kappa_{\mathrm{m}} \varepsilon_{\mathrm{d}}}{\kappa_{\mathrm{d}} \varepsilon_{\mathrm{m}}}\left(1 + \delta_{\mathrm{nl}}\right), \quad (1\mathrm{a})$$

which reduces to the LRA expression of Economou[5] in the absence of nonlocal corrections (i.e., $\delta_{\mathrm{nl}} \to 0$). Here, t_{d} is the thickness of the dielectric gap, whereas the out-of-plane components of the wavevectors are defined as

$$\kappa_j = \sqrt{q^2 - \varepsilon_j k_0^2} \quad \text{for} \quad j \in \{\mathrm{d, m}\}, \quad (1\mathrm{b})$$

where q is the GSP propagation constant, $k_0 = \omega/c = 2\pi/\lambda_0$ is the free-space wavevector, and $\varepsilon_{\mathrm{d}} \equiv \varepsilon_{\mathrm{d}}(\omega)$ and $\varepsilon_{\mathrm{m}} \equiv \varepsilon_{\mathrm{m}}(\omega)$ are the relative permittivities (q-independent) of the dielectric and the metal, respectively.

Within the nonlocal hydrodynamic formalism[41,42,45], the nonlocal response introduces q-dependence in the material response (giving rise to a nonlocal response in the corresponding real-space representation) of the form[48]

$$\delta_{\mathrm{nl}}(q, \omega) = \frac{q^2}{\kappa_{\mathrm{nl}} \kappa_{\mathrm{m}}} \frac{\varepsilon_{\mathrm{m}} - \varepsilon_\infty}{\varepsilon_\infty}, \quad (2\mathrm{a})$$

$$\kappa_{\mathrm{nl}} = \sqrt{q^2 - \frac{1}{\xi^2} \frac{\varepsilon_{\mathrm{m}}}{\varepsilon_\infty}}. \quad (2\mathrm{b})$$

The nonlocal wavevector $\kappa_{\mathrm{nl}} \equiv \kappa_{\mathrm{nl}}(q, \omega)$ enters in the hydrodynamic model of plasmonics as an additional longitudinal wave[42], with the nonlocal length scale ξ given by[42,45].

$$\xi^2 = \frac{(3/5)v_{\mathrm{F}}^2}{\omega(\omega + i\gamma)} + \frac{\mathcal{D}}{i\omega}. \quad (2\mathrm{c})$$

Here, the first term originates from the Thomas–Fermi theory of metals[42], with v_{F} being the Fermi velocity. The second term is an addition from GNOR model, with $\mathcal{D}$ denoting the diffusion constant that embodies the nonlocal damping. Finally, $\varepsilon_\infty \equiv \varepsilon_\infty(\omega)$ is a heuristic frequency-dependent parameter originating from the Drude LRA model [$\varepsilon_{\mathrm{m}} \equiv \varepsilon_\infty - \omega_{\mathrm{p}}^2/(\omega^2 + i\gamma\omega)$] that takes into account polarization due to the presence of positively charged atomic ions and interband transitions in the background of the quasi-free electron gas.

While the quantum-corrected dispersion relation (1a) originates from a hydrodynamic treatment of the nonlocal response of metals (with $\mathrm{Re}\{\delta_{\mathrm{nl}}\} \propto v_{\mathrm{F}}$ and $\mathrm{Im}\{\delta_{\mathrm{nl}}\} \propto \sqrt{\mathcal{D}}$)[42], we emphasize that in a SRF[45,53], it can equally well be expressed in terms of the Feibelman parameter for the centroid of the induced charge[53], i.e., $\delta_{\mathrm{nl}} \propto d_\perp$ (see Supplementary Section S10). For convenience, we will use the effective-mode index $n_{\mathrm{GSP}} \equiv q/k_0 = \mathrm{Re}\{n_{\mathrm{GSP}}\} + i \cdot \mathrm{Im}\{n_{\mathrm{GSP}}\}$ and thus discuss all our results in terms of this dimensionless and complex-valued quantity. Naturally, the stronger the confinement of electromagnetic field is, the larger $\mathrm{Re}\{n_{\mathrm{GSP}}\}$. In the spirit of the above treatment, where $\delta_{\mathrm{nl}} \ll 1$, nonlocality is expected to make only a small correction, which can nevertheless become sizable for large GSP wavevectors

(herein promoted by small, nanometer-scale dielectric gap thicknesses, see schematic inset in Fig. 1), i.e., for large $\mathrm{Re}\{n_{\mathrm{GSP}}\}$. Intuitively, $\mathrm{Re}\{n_{\mathrm{GSP}}\} \gg \mathrm{Im}\{n_{\mathrm{GSP}}\}$ is required for rendering quantum nonlocal effects experimentally observable. To enter this regime, the surface-response function $\mathrm{Re}\{d_\perp\}$ should not be negligible in comparison to t_{d}, thus calling for nanofabrication techniques that can controllably realize MDM structures with sub-10-nanometer gaps. For sub-nanometric gaps with $t_{\mathrm{d}} \lesssim |\mathrm{Re}\{d_\perp\}|$, additional quantum mechanical effects (e.g., tunneling, electronic spill-out) may be needed to be incoporated as well[38,55].

Sample design and fabrication. We experimentally study the dependence of the GSP spectral features on the thickness of the dielectric gap t_{d}. To that end, we have fabricated five different planar MDM structures with varying t_{d} (nominally 2, 3, 5, 10, and 20 nm), along with tailored waveguide couplers. As previously mentioned, in order to explore potential signatures of nonlocal effects, it is essential to reduce as much as possible all losses of intrinsic origin (i.e., classical, bulk losses), but also those related to fabrication imperfections, such as surface roughness, contamination of metal and dielectric materials with impurities, etc. Therefore, we have utilized monocrystalline Au flakes as metal layers, whereas for the dielectric material in the core of the MDM structure we have employed plasma-assisted ALD of Al_2O_3 layers which allows controlled growth of homogeneous dielectric layers with approximately 2.2 Å precision (see Methods and Supplementary Fig. S1 for details).

Figure 2 shows optical (a–d) and scanning electron (e, f) micrographs of a fabricated sample with a 3 nm dielectric gap, also revealing the FIB-milled tapered waveguide coupler. This element of the sample design and fabrication is of particular importance, since due to the large wavevector mismatch between free-space light and confined GSP modes, excitation of the latter with a Gaussian laser beam is not efficient. As such, it is necessary to provide a compact and adequate coupling mechanism, which is critical to obtain sufficiently strong signal in SNOM measurements. Due to the short GSP propagation length ($L_{\mathrm{GSP}} \equiv [2\mathrm{Im}\{q\}]^{-1}$ is less than 0.5 μm for a 2-nm-thick dielectric gap), a typical grating coupler schemes[62] become unsuitable. The periodicity of the grating Λ required by the phase-matching condition $q = k_0 + k_{\mathrm{grating}} = k_0 \pm m2\pi/\Lambda$ (with m being an integer) even for a grating with just 3 periods, is comparable with the GSP propagation length. Therefore, in order to improve the coupling efficiency while maintaining a compact device, we exploited a tapered waveguide coupler design[63] and optimized its geometrical parameters for each dielectric gap thickness (see Methods and Supplementary Section S9 for details).

Another important aspect of the design and fabrication of the MDM waveguides concerns the thickness of the upper gold flake, t_{u} (see sketch in Supplementary Fig. S13a). On the one hand, this upper Au layer needs to be sufficiently thin to make the NF of the ultraconfined GSP modes accessible to the s-SNOM tip, which scatters only weak evanescent tails of the mode penetrating through the top flake. On the other hand, t_{u} should be large enough to avoid significant modification of the GSP mode by undesired hybridization with the bare, single-interface SPP mode at the top air–gold interface. In both cases, the characteristic length-scale is the skin depth (~30 nm), making it a challenging task to fulfill both requirements. However, after thorough work on noise suppression in our SNOM, we managed to obtain near-field maps with sufficient signal-to-noise ratio for an upper-flake thickness of ~50 nm, for which the hybridization between the fundamental GSP and the single-interface SPP mode is negligible, while the assignment of a local thickness-independent ε_{m} is also justified[59,64]. Furthermore, atomic flatness of the monocrystalline
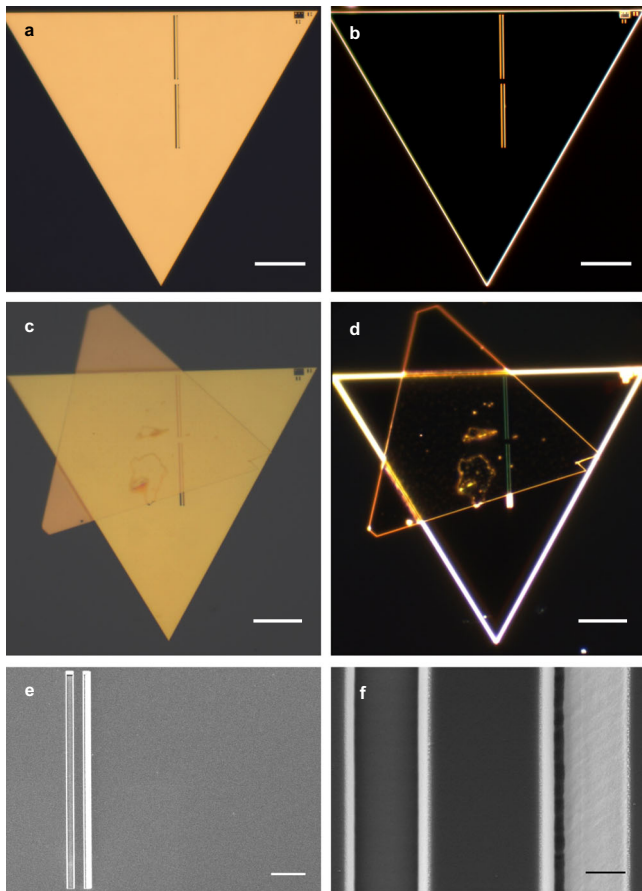
Fig. 2 Micrographs of the fabricated sample. a Bright- and (**b**) dark-field optical images of the sample with a 3 nm Al_2O_3 layer before the transfer of the top flake. **c** Bright- and (**d**) dark-field optical images of the sample after the transfer of the top flake. **e** SEM image of the flake during intermediate fabrication step and (**f**) close-up image of the FIB milled coupling element. Scale bars in panels (**a**–**d**) correspond to 10 µm, in panel (**e**) to 2 µm and in panel (**f**) to 50 nm.

Au flakes allows to reduce the noise induced by surface roughness in near-field measurement, which is an important aspect for such low-signal measurements.

Near-field measurements. Using the capabilities of our s-SNOM setup to measure both amplitude and phase, as illustrated in Fig. 3a (detailed descriptions of the setup can be found in the Methods section), we have obtained complex NF maps showing the propagating GSP modes in all the fabricated samples. As an example, Fig. 3b and c show pseudo-color images of the electric NF amplitude $|E_{NF}|$ and its real part $Re\{E_{NF}\}$, respectively, for the sample with $t_d = 3$ nm (NF maps of all other samples are provided in Supplementary Fig. S3). One-dimensional Fourier-transformation of the recorded NF maps along the GSP propagation direction ($x \rightarrow k_x$, as illustrated in Fig. 3d via the normalized absolute value of the transformed image) and averaging of the $E(k_x)$ spectrum along the y-axis direction, allows us to extract the real part of the GSP effective-mode index. As exhibited in Fig. 3e, the NF spectrum has peaks at two spatial frequencies, which correspond to two distinct propagating modes. The first one, with an effective-index of approximately unity, can be attributed to free-space light at an oblique incidence or to SPP modes in the topmost air–gold interface. The other contribution corresponds to the GSP mode and manifests as a peak at an effective-index slightly exceeding $Re\{n_{GSP}\} \approx 5.1$. We note the
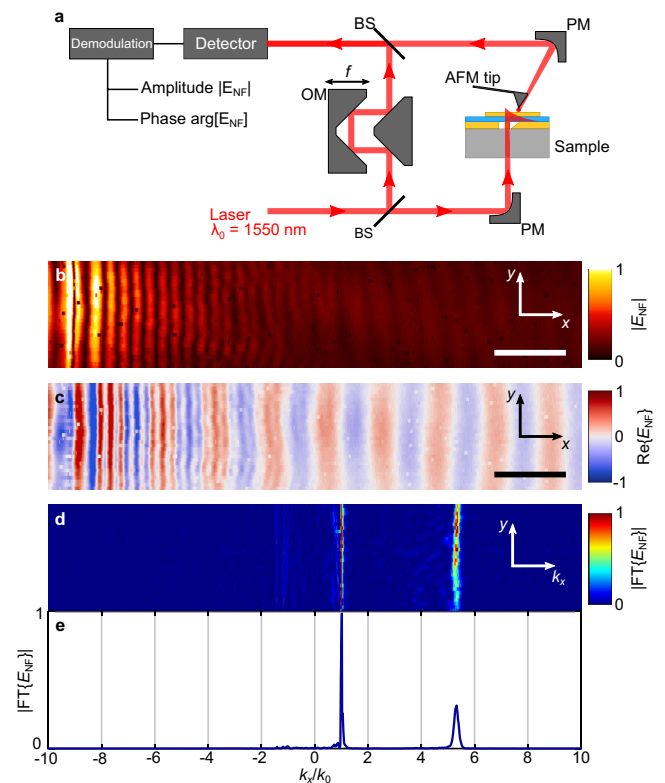

Fig. 3 Scanning near-field optical microscopy measurements. a Schematic illustration of the s-SNOM setup (see "Methods" for details). Pseudo-color images of (**b**) amplitude and (**c**) real-part of the detected near-field (NF) signal for the sample with 3 nm dielectric gap at $\lambda_0 = 1550$ nm excitation wavelength (scale bars: 2 µm). **d** Amplitude of the Fourier transformed NF map along the propagation coordinate ($x \rightarrow k_x$) and (**e**) its profile (averaged along the y-axis).

absence of any significant contribution at a corresponding negative value ($Re\{n_{GSP}\} \approx -5.1$), which could indicate GSP back-scattering ($k_x \rightarrow -k_x$) due to surface roughness, potentially competing with nonlocal effects[65]. Moderate roughness-induced scattering, as well as inhomogeneity of the effective dieielectric constant, may result in the broadening of the forward-scattering peak ($k_x \rightarrow k_x + \Delta k_x, \Delta k_x \ll k_x$), which would effectively manifest in a slightly increased imaginary part of the wavevector.

Further data post-processing, namely filtering of the NF maps and selecting only spatial frequencies in the vicinity of n_{GSP} in the Fourier domain, allows us to clean up interference with other near-fields as well as to reduce the noise, and retrieve the spatial evolution of the pure GSP mode along the propagation direction. Figure 4a illustrates the results of that procedure, showing how the GSP wavelength is shortened ($Re\{n_{GSP}\}$ increased) and exhibits a faster decay (i.e., increasing $Im\{n_{GSP}\}$) as t_d is reduced from $t_d = 20$ nm down to 2 nm. The propagation length of the GSP modes can be estimated by fitting an exponential envelope to the $|E_{NF}|$ field. Further details about the NF map processing can be found in Supplementary Section S4.

Fit of the experimental results to GNOR model. The parametric plot in Fig. 4b summarizes the experimental results (square data-points with error bars) and contrasts them against classical LRA calculations (dashed curve) as well as to those based on the GNOR model (solid curve). In both calculations, we used experimentally obtained material parameters at $\lambda_0 = 1550$ nm from literature, namely, $\varepsilon_m = -106.62 + 6.1257i$ for
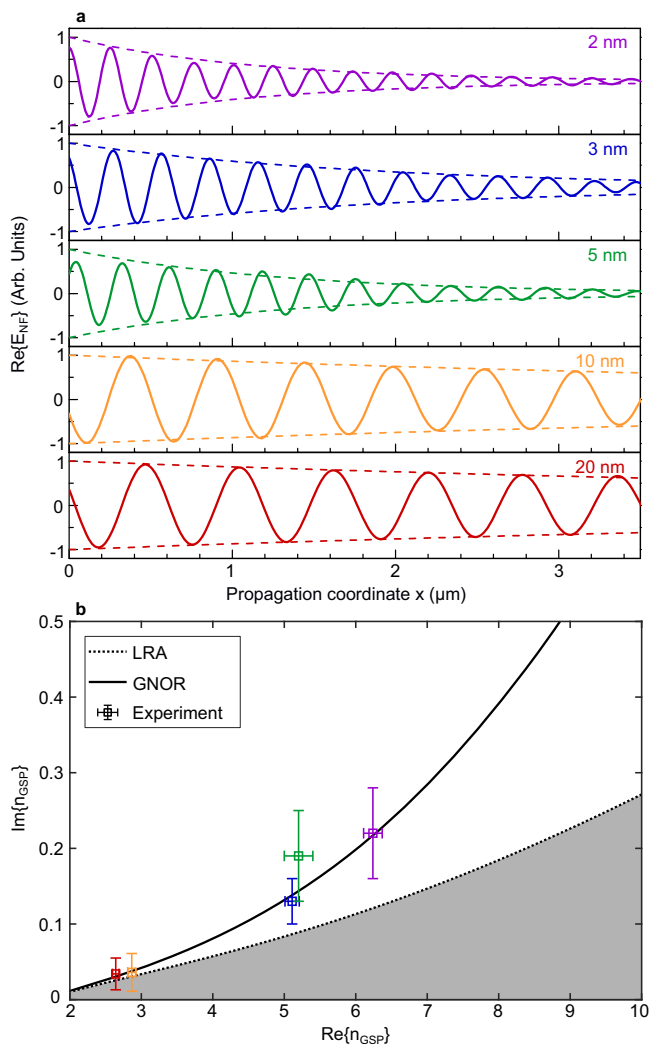
Fig. 4 Analysis of the experimental and theoretical results. a Real part (solid curves) and exponential envelopes (dashed curves) of the GSP mode profiles extracted from s-SNOM measurements for the five samples with different gap thickness (see labels). **b** Parametric plot of the effective-mode index n_{GSP} (at the excitation wavelength $\lambda_0 = 1550$ nm) for varying dielectric gap thickness: calculated using LRA (dashed curve), GNOR model (solid curve) and experimentally obtained data (squares with error bars that represent standard deviation of the measurement). Colors of the indicated points on the curves and experimental data points correspond to 2, 3, 5, 10, and 20 nm gap thicknesses, as in panel (a).

monocrystalline Au from Olmon et al.[66] and $\varepsilon_d = 2.657$ for Al_2O_3 from Boidin et al.[67]. The Drude model parameters which enter the nonlocal correction factor, δ_{nl}, were obtained by fitting the experimental data from Olmon et al. in the 1000–2000 nm wavelength range, yielding $\hbar\omega_p = 8.29$ eV, $\hbar\gamma = 47.72$ meV and $\varepsilon_\infty = 2.61$; and a tabulated value for gold's Fermi velocity $v_F = 1.4 \times 10^6$ ms^{-1} was used[68]. Even though the slope of the LRA curve is sensitive to the exact value of ε_m (see Supplementary Fig. S7), we deliberately invoke data from Olmon et al.[66], being appropriate for descriptions of optical response of monocrystalline Au flakes[23]. The diffusion constant $\mathcal{D}$ is used as the GNOR model's fitting parameter (see details in Supplementary Section S5). By fitting the dispersion relation [Eq. (1a)] to the the imaginary part of the experimental data points (at the corresponding values of Re{n_{GSP}}), we obtained an estimate

$\mathcal{D} \simeq 8 \times 10^{-4}$ m^2 s^{-1}. We emphasize that this procedure does not strictly rely on particular assumptions about the parameters of the dielectric gap, thus allowing a robust extraction of $\mathcal{D}$ even with significant uncertainties in the thickness t_d and effective permittivity ε_d. Further justification of this point is discussed in the following section.

As the LRA represents a minimal model for the light–matter interactions (disregarding nonclassical damping mechanisms), the gray-shaded area of this parametric regime in Fig. 4b would be "prohibited". Within the measurement accuracy of our s-SNOM experiment, all the data points indeed fall outside this regime, except for points corresponding to the large gaps, where quantum nonlocal contributions become negligible and the solid line asymptotically approaches the LRA curve. However, for the smaller gaps the deviations of the experimental data from the LRA curve suggest evidence for including gap-dependent broadening mechanisms in theoretical treatments of the ultraconfined GSP modes, e.g., nonlocal surface-scattering corrections.

Discussion

Nonlocal corrections to the classical, local-response electrodynamics have traditionally been extracted from optical measurements by confronting experimental data with classical predictions based on the LRA that invoke precise information about morphology of the sample. Typically, this information is obtained via structural characterization techniques, e.g., scanning electron microscopy (SEM), revealing particle or gap sizes and shapes (see, for instance, refs. [69–75]). The success of these approaches to determine nonlocal corrections thus relies on the ability to correlate spectral shifts of resonances with accurate structural information on the morphology, which is no simple task even with state-of-the-art electron microscopy. Other approaches utilized well-controlled spacer materials, which are challenging to work with, such as gaps formed by varying numbers of stacked graphene layers[76,77], while graphene plasmons may also be used on their own to unveil nonlocal quantum effects in nearby metals and to probe surface-response functions[58].

As shown in Fig. 4, our analysis of GSP propagation partially eliminates the above-noted challenge since the experimentally measured n_{GSP} data can be considered in a parametric way (where t_d is varied through different, but otherwise identical devices) and can be confronted with the theory in the same way, i.e., by plotting solutions of Eq. (1a) for varying t_d. As such, the dashed curve in Fig. 4b depends only on ε_m and ε_d, while the solid curve has an additional dependence on δ_{nl}. The solid curve is swept away from the dashed counterpart when one increases δ_{nl} (effectively by varying complex-valued ξ, see inset in Fig. 1). Furthermore, as shown in Supplementary Fig. S8, the slopes of both the LRA and GNOR dispersion curves in this parametric space are nearly independent of ε_d, and insensitive to addition of a thin air void in the gap too (Supplementary Fig. S11). This alleviates the need for extremely accurate quantification of the possible deviations of ε_d and t_d (as fabricated) from its nominal values (as intended in the initial design).

While we note that the real part of n_{GSP} obtained in experiments is consistently smaller than the predictions of both the LRA and calculations incorporating nonlocal corrections (see Supplementary Fig. S4), this fact does not deny nor contradict our analysis in the parametric way, as explained above. A possible explanation for smaller n_{GSP} could be presence of an air void or an additional dielectric layer between the Au flakes and the ALD film. The former could be present due to poor adhesion or other fabrication imperfection (for example, contamination with small particles that suspend the upper Au flakes slightly above the ALD

layer), while the latter might appear as residuals of the organic molecules from the Au crystal growth solution or deposited conduction layer for FIB milling. We refer to the additional calculations provided in Section S8 of Supplementary Information, which support the plausibility of this explanation, as well as its compatibility with our analysis.

The diffusion constant $\mathcal{D}$, which embodies nonlocal damping, was estimated from a fit to the experimental data (see Section S5 of the Supplementary Information). The obtained value $\mathcal{D} \simeq 8 \times 10^{-4}\,\mathrm{m^2\,s^{-1}}$ is slightly smaller than previously reported experimental value $\mathcal{D} \simeq 8.8 \times 10^{-4}\,\mathrm{m^2\,s^{-1}}$ [75]. In the initial GNOR theory[41], it was consolidated that the diffusion is effectively linked to the surface-scattering properties[78] and that the diffusion constant for air-gold interface is anticipated to be even smaller, on the order of $\mathcal{D} \sim 3 \times 10^{-4}\,\mathrm{m^2\,s^{-1}}$ [79], while dielectric surroundings ($\varepsilon_d > 1$) further promotes Landau damping[80,81]. Eventually, the higher diffusion constant revealed by our experiments (as well as ref. [75]) could also indicate some additional surface scattering in our samples, despite our dedicated efforts to eliminate surface roughness through the use of monocrystalline gold and ALD films. Overview and comparison of the previously reported values of $\mathcal{D}$ is provided in Supplementary Table S1 of the Supplementary Information.

We emphasize that we have refrained from any attempt to fit the real part of ξ, related to the Fermi velocity v_F, to force non-local curve to fit better the experimental data as the value of $1.4 \times 10^6\,\mathrm{m\,s^{-1}}$ is well-established and was obtained independently in specific experiments. Instead, we argue that our results offer a way to infer the value of $d_\perp$ from experiment, admittedly being a determination of the surface-response function for this particular wavelength only and for this particular interface between Au and $\mathrm{Al_2O_3}$ [i.e., $d_\perp^{\mathrm{Al_2O_3-Au}}(\omega_0)$]. In principle, this procedure amounts to a fitting-extraction of the dimensionless quantity $d_\perp/t_d$, so that any inevitable uncertainties in t_d (fabrication-induced deviations from nominal values) would in practice limit the accuracy by which we could in turn estimate $d_\perp$. Since theory accounts, including ab initio calculations, suggest that $d_\perp$ is to be found in the ångström range[45,59], we would need to experimentally determine t_d with atomic-scale accuracy, which remains a significant challenge as mentioned above. The fitting of the data with Eq. (1a) gives a value of $\delta_{nl} = -0.0822 + 0.0303\,i \pm (0.007 + 0.02i)$ for the smallest studied dielectric gap ($t_d = 2$ nm), which indeed meets our initial expectation that $|\delta_{nl}| \ll 1$. While here we suggest a nonlocal interpretation for the observed gap-dependent broadening, we acknowledge that in principle this could also be qualitatively explained within the LRA by invoking other more lossy material-response models (see Supplementary Fig. S7), which phenomenologically mimic additional effects of roughness, grain boundaries, etc. While our experiments cannot totally dismiss such alternative explanations, they appear less obvious for our samples fabricated using high-quality materials. Indeed, SNOM measurements on single Au flakes[33,34]—where the bare surface plasmons exhibit the local-response dynamics, since there are no small spatial sample features to promote significant effects from the large wavevector response of the plasmons—appear well-explained by the Olmon data for crystalline Au[66], while more recent experiments conclude even lower bulk damping[82,83] in better accordance with the McPeak data[22].

In summary, we have performed an experimental study of record-high GSP mode-index in MDM structures comprised of high-quality monocrystalline Au flakes and ultrathin ALD-deposited $\mathrm{Al_2O_3}$ films. Analysis of s-SNOM measurements from samples with different dielectric gap thicknesses, reveals signatures of gap-dependent nonclassical broadening, which we found to be well-described by the generalized hydrodynamic model of plasmonics[41]. Such broadening becomes progressively

more significant for smaller dielectric gap thicknesses that give rise to higher effective-mode indices. Our results suggest that quantum nonlocal corrections should be taken into account when treating extremely confined gap plasmon modes supported by MDM structures with sub-10-nanometer dielectric gaps.

Methods

Sample fabrication. The fabrication recipe flowchart diagram can be found in Supplementary Fig. S1. Below we provide details about particular fabrication steps and used equipment.

Synthesis of monocrystalline Au flakes. Monocrystalline Au flake samples were prepared using a recipe adopted from reference[24]. In short, thin and flat Au crystals were synthesized on BK-7 glass substrates via endothermic reduction of chlorooauric acid ($\mathrm{HAuCl_4 \cdot 3H_2O}$) in ethylene glycol ($\mathrm{C_2H_6O_2}$). All reagents were purchased from Sigma-Aldrich. The substrates were put into a vial with the solution and kept in an oven at 90 °C for 24 h. Afterwards, samples were cleaned in acetone, isopropyl alcohol (IPA) and distilled water and dried with a nitrogen blow. Each substrate carried a large number of Au flakes with diverse sizes and shapes, and suitable samples were selected by visual inspection using optical microscopy.

FIB milling of waveguide couplers. Milling of tapered waveguide coupler structures in monocrystalline Au flakes was performed using a gallium (Ga) ion FIB instrument (Helios NanoLab G3 UC, FEI company), with beam oriented perpendicularly to the sample plane an accelerating voltage of 30 kV and an ion current of 7.7 pA. In a preparatory step, substrates carrying Au flake samples that were selected in the previous fabrication step were coated with a thin (≈ 6 nm) conductive carbon layer with a Leica Sputter Coater LEICA EM ACE600 in order to discharge the sample directly on the glass substrate, subsequently removed by mild plasma etching during the next fabrication step.

ALD of dielectric gaps. After FIB patterning, the samples were coated with 2, 3, 5, 10, and 20 nm–thick layers of $\mathrm{Al_2O_3}$ using Oxford Plasma Technology OPAL ALD system (Bristol, UK) equipped with an inductively coupled plasma source. To avoid degradations of the samples, the aluminum oxide layer was deposited at low substrate temperatures (30 °C). As precursors trimethylaluminum (TMA) and oxygen plasma were used, with dose times of 30 ms and 5 s, respectively. Between each dose step, the reaction chamber was purged with inert argon and nitrogen for 5 s. The ALD process is defined by two half-cycles of self-terminating single surface reactions of each precursor, resulting in highly homogeneous and conformal $\mathrm{Al_2O_3}$ coatings with nearly perfect thickness control[84]. The growth per cycle was 2.2 Å[85]. Hence, the desired film thickness was adjusted by the number of applied ALD-cycles.

Transfer and assembly of Au flakes. Upper Au flakes were transferred from the substrates at which they have been synthesized using customized 2D material transfer system (HQ+ graphene). Polydimethylsiloxane (PDMS) stamps (WF X4 Gel-Film from Gel-Pak) were used as a carrier substrates. Transfer was performed at elevated temperature of 130 °C to promote adhesion of the Au flake to the target substrate.

Numerical electrodynamics simulations. Numerical simulations and optimization of the tapered coupler geometric parameters were carried out using a commercially available finite-element method (FEM) solver (COMSOL Multiphysics 5.4, Wave Optics module). Since a planar waveguide was considered, modeling domain was two-dimensional (i.e., no spatial variation along y-axis, see Supplementary Fig. S13). Experimentally measured values of material optical parameters at 1550 nm wavelength were used: $\varepsilon_m = -106.62 + 6.1257i$ for monocrystalline Au from Olmon et al.[66]; $\varepsilon_d = 2.657$ for $\mathrm{Al_2O_3}$ from Boidin et al.[67]; and $\varepsilon = 2.1$ for the BK-7 glass.

Simulations were performed in a wavelength domain at $\lambda_0 = 1550$ nm using a Gaussian beam source in a scattered field formulation. Value of 0.15 was used as a numerical aperture of the Gaussian beam to mimic the experimental excitation conditions in the s-SNOM setup. Geometrical parameters of the tapered waveguide coupler (w_t, w_a, w_1 and w_2, see Supplementary Section S9) were optimized using the Levenberg–Marquardt algorithm (available from the COMSOL Optimization module), with the coupling efficiency being the objective function.

Near-field measurements. NF measurements were performed with the aid of s-SNOM[61], using the transmission module of a customized commercially available setup (Neaspec). Pseudo-heterodyne demodulation allows to simultaneously obtain information about amplitude and phase of the NF signal. Figure 3a shows a schematic diagram of the setup: the laser beam from a CW telecom laser (with wavelength $\lambda_0 = 1550$ nm) is split into two interferometric arms using a beam splitter (BS). In the reference arm, the signal is modulated using an oscillating mirror (OM) driven at the frequency $f \approx 300$ Hz. In the other arm, the laser beam is focused onto the waveguide coupler of the sample using a parabolic mirror (PM), with the full-width-half-max (FWHM) of the spot size ~3 μm. An atomic-force

microscope (AFM) tip (Arrow NCPt from NanoWorld) raster-scans the surface of the sample (which simultaneously allows to obtain topography of the sample) in a tapping mode (at the frequency $f_d \approx 250$ kHz and amplitude ~50 nm) and scatters the optical NF that is collected by another PM. Finally, reference and measurement arms are combined with a BS to allow interferometric detection and subsequent pseudo-heterodyne demodulation. In order to suppress background (bulk scattering from the tip and sample), the signal is demodulated at high harmonics of tip's oscillation frequency, mf_d, with $m = 3$ for the presented results. The recorded data for the propagation constant is reproducible using different, but nominally identical tips, while inevitable tip-to-tip variations would manifest locally in the intensity of scattered light and in the resolution of spatial dynamics below the tip radius. Details on the subsequent analysis of the recorded near-field maps is provided in Supplementary Section S4.

Data availability

The data that underlie the findings of this study are available from the corresponding authors upon reasonable request.

References

1. Barnes, W. L., Dereux, A. & Ebbesen, T. W. Surface plasmon subwavelength optics. *Nature* **424**, 824–830 (2003).
2. Gramotnev, D. K. & Bozhevolnyi, S. I. Plasmonics beyond the diffraction limit. *Nat. Photon.* **4**, 83–91 (2010).
3. Gramotnev, D. K. & Bozhevolnyi, S. I. Nanofocusing of electromagnetic radiation. *Nat. Photon.* **8**, 14–23 (2014).
4. Fernández-Domínguez, A. I., García-Vidal, F. J. & Martín-Moreno, L. Unrelenting plasmons. *Nat. Photon.* **11**, 8–10 (2017).
5. Economou, E. N. Surface plasmons in thin films. *Phys. Rev.* **182**, 539–554 (1969).
6. Sorger, V. J., Oulton, R. F., Ma, R.-M. & Zhang, X. Toward integrated plasmonic circuits. *MRS Bull.* **37**, 728–738 (2012).
7. Ding, F., Pors, A. & Bozhevolnyi, S. I. Gradient metasurfaces: a review of fundamentals and applications. *Rep. Prog. Phys.* **81**, 026401 (2018).
8. Smith, C. L. C., Stenger, N., Kristensen, A., Mortensen, N. A. & Bozhevolnyi, S. I. Gap and channeled plasmons in tapered grooves: a review. *Nanoscale* **7**, 9355–9386 (2015).
9. Baumberg, J. J., Aizpurua, J., Mikkelsen, M. H. & Smith, D. R. Extreme nanophotonics from ultrathin metallic gaps. *Nat. Mater.* **18**, 668–678 (2019).
10. Langer, J. et al. Present and future of surface-enhanced raman scattering. *ACS Nano* **14**, 28–117 (2020).
11. Bozhevolnyi, S. I. & Khurgin, J. B. The case for quantum plasmonics. *Nat. Photon.* **11**, 398–400 (2017).
12. Fernández-Domínguez, A. I., Bozhevolnyi, S. I. & Mortensen, N. A. Plasmon-enhanced generation of nonclassical light. *ACS Photonics* **5**, 3447–3451 (2018).
13. Dionne, J. A., Sweatlock, L. A., Atwater, H. A. & Polman, A. Plasmon slot waveguides: Towards chip-scale propagation with subwavelength-scale localization. *Phys. Rev. B* **73**, 035407 (2006).
14. Bozhevolnyi, S. I. & Jung, J. Scaling for gap plasmon based waveguides. *Opt. Express* **16**, 2676–2684 (2008).
15. Dionne, J. A., Lezec, H. J. & Atwater, H. A. Highly confined photon transport in subwavelength metallic slot waveguides. *Nano Lett.* **6**, 1928–1932 (2006).
16. Preiner, M. J., Shimizu, K. T., White, J. S. & Melosh, N. A. Efficient optical coupling into metal-insulator-metal plasmon modes with subwavelength diffraction gratings. *Appl. Phys. Lett.* **92**, 113109 (2008).
17. Verhagen, E., Dionne, J. A., Kuipers, L. K., Atwater, H. A. & Polman, A. Near-field visualization of strongly confined surface plasmon polaritons in metal-insulator-metal waveguides. *Nano Lett.* **8**, 2925–2929 (2008).
18. Deshpande, R., Zenin, V. A., Ding, F., Mortensen, N. A. & Bozhevolnyi, S. I. Direct characterization of near-field coupling in gap plasmon-based metasurfaces. *Nano Lett.* **18**, 6265–6270 (2018).
19. Khurgin, J. B. & Sun, G. In search of the elusive lossless metal. *Appl. Phys. Lett.* **96**, 181102 (2010).
20. Khurgin, J. B. How to deal with the loss in plasmonics and metamaterials. *Nat. Nanotechnol.* **10**, 2–6 (2015).
21. Naik, G. V., Shalaev, V. M. & Boltasseva, A. Alternative plasmonic materials: beyond gold and silver. *Adv. Mater.* **25**, 3264–3294 (2013).
22. McPeak, K. M. et al. Plasmonic films can easily be better: Rules and recipes. *ACS Photonics* **2**, 326–333 (2015).
23. Hoffmann, B. et al. New insights into colloidal gold flakes: structural investigation, micro-ellipsometry and thinning procedure towards ultrathin monocrystalline layers. *Nanoscale* **8**, 4529–4536 (2016).
24. Krauss, E. et al. Controlled growth of high-aspect-ratio single-crystalline gold platelets. *Cryst. Growth Des.* **18**, 1297–1302 (2018).
25. Huang, J.-S. et al. Atomically flat single-crystalline gold nanostructures for plasmonic nanocircuitry. *Nat. Commun.* **1**, 150 (2010).
26. Kumar, S. et al. Efficient coupling of single organic molecules to channel plasmon polaritons supported by v-grooves in monocrystalline gold. *ACS Photonics* **7**, 2211–2218 (2020).
27. Méjard, R. et al. Advanced engineering of single-crystal gold nanoantennas. *Opt. Mater. Express* **7**, 1157–1168 (2017).
28. Kejík, L., Horák, M., Šikola, T. & Křápek, V. Structural and optical properties of monocrystalline and polycrystalline gold plasmonic nanorods. *Opt. Express* **28**, 34960–34972 (2020).
29. Viarbitskaya, S. et al. Tailoring and imaging the plasmonic local density of states in crystalline nanoprisms. *Nat. Mater.* **12**, 426–432 (2013).
30. Frank, B. et al. Short-range surface plasmonics: Localized electron emission dynamics from a 60-nm spot on an atomically flat single-crystalline gold surface. *Sci. Adv.* **3**, e1700721 (2017).
31. Spektor, G. et al. Revealing the subfemtosecond dynamics of orbital angular momentum in nanoplasmonic vortices. *Science* **355**, 1187–1191 (2017).
32. Großmann, S. et al. Nonclassical optical properties of mesoscopic gold. *Phys. Rev. Lett.* **122**, 246802 (2019).
33. Kaltenecker, K. J. et al. Mono-crystalline gold platelets: A high quality platform for surface plasmon polaritons. *Nanophotonics* **9**, 509 (2020).
34. Kaltenecker, K. J. et al. Near-infrared nanospectroscopy using a low-noise supercontinuum source. *APL Photonics* **6**, 066106 (2021).
35. Boroviks, S. et al. Anisotropic second-harmonic generation from monocrystalline gold flakes. *Opt. Lett.* **46**, 833–836 (2021).
36. Munkhbat, B., Canales, A., Küçüköz, B., Baranov, D. G. & Shegai, T. O. Tunable self-assembled Casimir microcavities and polaritons. *Nature* **597**, 214–219 (2021).
37. Tame, M. S. et al. Quantum plasmonics. *Nat. Phys.* **9**, 329–340 (2013).
38. Zhu, W. et al. Quantum mechanical effects in plasmonic structures with subnanometre gaps. *Nat. Commun.* **7**, 11495 (2016).
39. Bozhevolnyi, S. I. & Mortensen, N. A. Plasmonics for emerging quantum technologies. *Nanophotonics* **6**, 1185 (2017).
40. Christensen, T. et al. Nonlocal response of metallic nanospheres probed by light, electrons, and atoms. *ACS Nano* **8**, 1745–1758 (2014).
41. Mortensen, N. A., Raza, S., Wubs, M., Søndergaard, T. & Bozhevolnyi, S. I. A generalized non-local optical response theory for plasmonic nanostructures. *Nat. Commun.* **5**, 3809 (2014).
42. Raza, S., Bozhevolnyi, S. I., Wubs, M. & Mortensen, N. A. Nonlocal optical response in metallic nanostructures. *J. Phys.: Cond. Matter* **27**, 183204 (2015).
43. Gonçalves, P. A. D. *Plasmonics and Light–Matter Interactions in Two-Dimensional Materials and in Metal Nanostructures: Classical and Quantum Considerations* (Springer Nature, 2020).
44. Gonçalves, P. A. D. et al. Plasmon–emitter interactions at the nanoscale. *Nat. Commun.* **11**, 366 (2020).
45. Mortensen, N. A. Mesoscopic electrodynamics at metal surfaces — from quantum-corrected hydrodynamics to microscopic surface-response formalism. *Nanophotonics* **10**, 2563–2616 (2021).
46. Zhou, Q., Zhang, P. & Chen, X.-W. General framework of canonical quasinormal mode analysis for extreme nano-optics. *Phys. Rev. Lett.* **127**, 267401 (2021).
47. Baghramyan, H. M., Della Sala, F. & Ciracì, C. Laplacian-level quantum hydrodynamic theory for plasmonics. *Phys. Rev. X* **11**, 011049 (2021).
48. Raza, S., Christensen, T., Wubs, M., Bozhevolnyi, S. I. & Mortensen, N. A. Nonlocal response in thin-film waveguides: Loss versus nonlocality and breaking of complementarity. *Phys. Rev. B* **88**, 115401 (2013).
49. Karanikolas, V. et al. Quantum surface effects in strong coupling dynamics. *Phys. Rev. B* **104**, L201405 (2021).
50. Yan, W., Wubs, M. & Mortensen, N. A. Hyperbolic metamaterials: Nonlocal response regularizes broadband supersingularity. *Phys. Rev. B* **86**, 205429 (2012).
51. Khurgin, J. B. Ultimate limit of field confinement by surface plasmon polaritons. *Faraday Disc.* **178**, 109–122 (2015).
52. Varas, A., García-González, P., Feist, J., García-Vidal, F. J. & Rubio, A. Quantum plasmonics: from jellium models to ab initio calculations. *Nanophotonics* **5**, 409–426 (2016).
53. Feibelman, P. J. Surface electromagnetic-fields. *Prog. Surf. Sci.* **12**, 287–407 (1982).
54. Teperik, T. V., Nordlander, P., Aizpurua, J. & Borisov, A. G. Robust subnanometric plasmon ruler by rescaling of the nonlocal optical response. *Phys. Rev. Lett.* **110**, 263901 (2013).
55. Yan, W., Wubs, M. & Mortensen, N. A. Projected dipole model for quantum plasmonics. *Phys. Rev. Lett.* **115**, 137403 (2015).
56. Christensen, T., Yan, W., Jauho, A.-P., Soljačić, M. & Mortensen, N. A. Quantum corrections in nanoplasmonics: Shape, scale, and material. *Phys. Rev. Lett.* **118**, 157402 (2017).

57. Yang, Y. et al. A general theoretical and experimental framework for nanoscale electromagnetism. *Nature* **576**, 248–252 (2019).

58. Gonçalves, P. A. D. et al. Quantum surface-response of metals revealed by acoustic graphene plasmons. *Nat. Commun.* **12**, 3271 (2021).

59. Echarri, A. R. et al. Optical response of noble metal nanostructures: Quantum surface effects in crystallographic facets. *Optica* **8**, 710–721 (2021).

60. Kern, J. et al. Atomic-scale confinement of resonant optical fields. *Nano Lett.* **12**, 5504–5509 (2012).

61. Keilmann, F. & Hillenbrand, R. Near-field microscopy by elastic light scattering from a tip. *Philos. Trans. A* **362**, 787–805 (2004).

62. Maier, S. A. *Plasmonics: Fundamentals and Applications* (Springer-Verlag, 2007).

63. Pile, D. F. P. & Gramotnev, D. K. Adiabatic and nonadiabatic nanofocusing of plasmons by tapered gap plasmon waveguides. *Appl. Phys. Lett.* **89**, 041111 (2006).

64. Echarri, A. R., Cox, J. D. & García de Abajo, F. J. Quantum effects in the acoustic plasmons of atomically thin heterostructures. *Optica* **6**, 630–641 (2019).

65. Hajisalem, G., Min, Q., Gelfand, R. & Gordon, R. Effect of surface roughness on self-assembled monolayer plasmonic ruler in nonlocal regime. *Opt. Express* **22**, 9604–9610 (2014).

66. Olmon, R. L. et al. Optical dielectric function of gold. *Phys. Rev. B* **86**, 235147 (2012).

67. Boidin, R., Halenkovič, T., Nazabal, V., Beneš, L. & Němec, P. Pulsed laser deposited alumina thin films. *Ceram. Int.* **42**, 1177–1182 (2016).

68. Ashcroft, N. D. & Mermin, N. D. *Solid State Physics* (Saunders College Publishing, 1976).

69. Miyazaki, H. T. & Kurokawa, Y. Squeezing visible light waves into a 3-nm-thick and 55-nm-long plasmon cavity. *Phys. Rev. Lett.* **96**, 097401 (2006).

70. Duan, H., Fernández-Domínguez, A. I., Bosman, M., Maier, S. A. & Yang, J. K. W. Nanoplasmonics: classical down to the nanometer scale. *Nano Lett.* **12**, 1683–1689 (2012).

71. Ciracì, C. et al. Probing the ultimate limits of plasmonic enhancement. *Science* **337**, 1072 (2012).

72. Scholl, J. A., Koh, A. L. & Dionne, J. A. Quantum plasmon resonances of individual metallic nanoparticles. *Nature* **483**, 421 (2012).

73. Raza, S. et al. Blueshift of the surface plasmon resonance in silver nanoparticles studied with EELS. *Nanophotonics* **2**, 131–138 (2013).

74. Raza, S. et al. Multipole plasmons and their disappearance in few-nanometer silver nanoparticles. *Nat. Commun.* **6**, 8788 (2015).

75. Shen, H. et al. Optical observation of plasmonic nonlocal effects in a 2D superlattice of ultrasmall gold nanoparticles. *Nano Lett.* **17**, 2234–2239 (2017).

76. Mertens, J. et al. Controlling subnanometer gaps in plasmonic dimers using graphene. *Nano Lett.* **13**, 5033–5038 (2013).

77. Iranzo, D. A. et al. Probing the ultimate plasmon confinement limits with a van der Waals heterostructure. *Science* **360**, 291–295 (2018).

78. Svendsen, M. K., Wolff, C., Jauho, A.-P., Mortensen, N. A. & Tserkezis, C. Role of diffusive surface scattering in nonlocal plasmonics. *J. Phys.: Cond. Matter* **32**, 395702 (2020).

79. Raza, S., Wubs, M., Bozhevolnyi, S. I. & Mortensen, N. A. Nonlocal study of ultimate plasmon hybridization. *Opt. Lett.* **40**, 839–842 (2015).

80. Jin, D. et al. Quantum-spillover-enhanced surface-plasmonic absorption at the interface of silver and high-index dielectrics. *Phys. Rev. Lett.* **115**, 193901 (2015).

81. Mortensen, N. A. et al. Surface-response functions obtained from equilibrium electron-density profiles. *Nanophotonics* **10**, 3647–3657 (2021).

82. Lebsir, Y., Boroviks, S., Thomaschewski, M., Bozhevolnyi, S. I. & Zenin, V. A. Ultimate limit for optical losses in gold, revealed by quantitative near-field microscopy. https://arxiv.org/abs/2203.00754 (2022).

83. Casses, L. N., Kaltenecker, K. J., Xiao, S., Wubs, M. & Stenger, N. Quantitative near-field characterization of surface plasmon polaritons on monocrystalline gold platelets. *Opt. Express* **30**, 11181–11191 (2022).

84. Jang, B. et al. Fine-tuning of the optical properties of hollow-core light cages using dielectric nanofilms. *Opt. Lett.* **45**, 196–199 (2020).

85. Yüksel, S. et al. Background-free bottom-up plasmonic arrays with increased sensitivity, specificity and shelf life for sers detection schemes. *J. Phys. Chem. C* **119**, 13791–13798 (2015).

Author contributions

S.B., J.-S.H., and N.A.M. conceived and planned the research. S.B. designed the MDM coupling structure, drawing also on expertise in MDM waveguides from V.A.Z. and S.I.B. P.A.D.G. and S.B. derived the dispersion relation within surface-response formalism. Samples were fabricated by S.B. and Z.-H.L., while S.B. synthesized the Au flakes and M.Z. performed the ALD. The FIB milling and SEM characterization was performed by S.B., Z.-H.L., and A.D. SNOM measurements were performed by V.A.Z. and S.B., while S.B., V.A.Z., and C.W. performed the data analysis. All authors contributed to the discussion and interpretation of the results, and the writing of the manuscript was done in a joint effort.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-30737-2.

Correspondence and requests for materials should be addressed to N. Asger Mortensen.

Peer review information *Nature Communications* thanks Luis Perez and the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.