Plasmonic metal nanostructures with extremely small features: new effects, fabrication and applications

Huimin Shi, a Xupeng Zhu,b Shi Zhang,c Guilin Wen,*a Mengjie Zhengd and Huigao Duan*c

Surface plasmons in metals promise many fascinating properties and applications in optics, sensing, photonics and nonlinear fields. Plasmonic nanostructures with extremely small features especially demonstrate amazing new effects as the feature sizes scale down to the sub-nanometer scale, such as quantum size effects, quantum tunneling, spill-out of electrons and nonlocal states etc. The unusual physical, optical and photo-electronic properties observed in metallic structures with extreme feature sizes enable their unique applications in electromagnetic field focusing, spectra enhancing, quantum photonics, etc. In this review, we focus on the new effects, fabrication and applications of plasmonic metal nanostructures with extremely small features. For simplicity and consistency, we will focus our topic on the plasmonic metal nanostructures with feature sizes of sub-nanometers. Subsequently, we discussed four main and typical plasmonic metal nanostructures with extremely small features, including: (1) ultra-sharp plasmonic metal nanotips; (2) ultra-thin plasmonic metal films; (3) ultra-small plasmonic metal particles and (4) ultra-small plasmonic metal nanogaps. Additionally, the corresponding fascinating new effects (quantum nonlinear, non-locality, quantum size effect and quantum tunneling), applications (spectral enhancement, high-order harmonic wave generation, sensing and terahertz wave detection) and reliable fabrication methods will also be discussed. We end the discussion with a brief summary and outlook of the main challenges and possible breakthroughs in the
field. We hope our discussion can inspire the broader design, fabrication and application of plasmonic metal nanostructures with extremely small feature sizes in the future.

1. Introduction

The new fascinating effects caused by light–matter interaction in metal nanostructures enable lots of amazing applications. As one of the most typical and amazing effects, the coherent and collective electron oscillations in a dielectric and conductor interface (e.g., metal–dielectric interface), surface plasmons (SPs), attract wide and global attention in the past few decades. Generally, the electromagnetic field induced valence band electrons excitation in metals, named as the surface plasmon polaritons. Such electromagnetic surface waves arise through the coupling of the electromagnetic field to oscillation of the conductor’s electron plasma. Taking advantage of strong light–matter interactions and unique properties, SPs exhibit extensive applications and potential in light focusing, optical imaging, ultra-sensitive sensing, detection and medical therapy.2-4

Owing to the rapid development of SPs in the past three decades, the SP related new structures, materials, effects and applications experience explosive growth and flourishing. Therefore, it is of great importance to make a timely summary and outlook of these active issues. One of the most crucial issues is the amazing effects, as well as the potential terrific applications, enabled by plasmonic metal nanostructures with extremely small features. In this review, therefore, we focus on discussing the new effects, fabrication and applications of plasmonic metal nanostructures with extremely small features, as shown in Fig. 1. For simplicity and novelty, we mainly concentrate on plasmonic metal nanostructures with sub-10 nm feature sizes. Considering the fact that some excellent and outstanding reviews have demonstrated the pivotal role played by SPs in many well-known fields, such as in two dimensional materials, surface enhanced Raman spectroscopy, photo-electrochemistry etc., the general applications of SPs in these aforementioned fields will not be involved. We will start

Shi Zhang is currently a PhD student in the College of Mechanical and Vehicle Engineering at Hunan University. He received his B.Sc. degree (2015) from the School of Physics and Optoelectronics at Xiangtan University. Then he joined Prof. Huigao Duan’s group of Hunan University. His research focuses on fabrication of plasmonic nanostructures based on electron beam lithography and their optical applications.

Dr Guijin Wen received his M.S. and Ph.D. degrees in Mechanics from Southwest Jiaotong University, China, in 1998 and 2001, respectively. He is currently a full Professor with the School of Mechanical and Electrical Engineering, Guangzhou University, China. His research interests include nonlinear dynamics and intelligent control, materials and structure design, and vibration theory.

Dr Mengjie Zheng received her B.Sc. in Microelectronics from Lanzhou University, P. R. China, in 2014, and her PhD in Physics from Hunan University, P. R. China, in 2019. She joined the Research Laboratory of Electronics at MIT, USA and was a visiting student from 2017 to 2018. She is currently a research staff at Jihua Laboratory, P. R. China. Her research focuses on nanofabrication and nanostructured optical devices.

Dr Huigao Duan received his B.S. and Ph.D. in Physics from Lanzhou University in 2004 and 2010, respectively. From 2006 to 2008, he was working as a Researcher in the Institute of Electrical Engineering, Chinese Academy of Sciences. From 2008 to 2010, he was a visiting Researcher at Massachusetts Institute of Technology. He was working in IMRE, A*STAR, Singapore, as a Research Scientist from 2010 to 2012. He joined Hunan University, China, as a full professor in 2012 and is now a principal investigator in the College of Mechanical and Vehicle Engineering. His research interests include micro/nanomanufacturing, micro/nanosystems and their relevant applications.
the topic by giving a basic and brief introduction to the fundamentals of light–metal interaction. Then the typical plasmon effects, especially the effects induced by the sub-nanometer feature structures, generated in the crucial representation metal nanostructures with extremely small feature sizes will be presented sequentially. Four of the main and representative metal plasmonic nanostructures will be systematically discussed, which are (1) ultra-sharp plasmonic metal nanotips; (2) ultra-thin plasmonic metal films; (3) ultra-small plasmonic metal particles and (4) ultra-small plasmonic metal nanogap structures. The discussion will end up with a summary and outlook of the challenges and potential of plasmonic metal nanostructures with extremely small features in driving and exploiting future new applications and amazing discoveries.

We hope our discussion here can give a new insight to understand the distinctive applications induced by new plasmonic effects in metal nanostructures and inspire more exciting areas in metal plasmonics far beyond the current fields.

2. Fundamentals of light–metal interaction

As known, the unique response of metals to incident light is dominated by the collective behavior of free electrons in metals. In the early 20th century, Drude first explained the collective oscillation behavior of metal free electrons by using an incident frequency dependent complex permittivity, i.e. the Drude model. The core idea is that the response of metals to the incident light is mainly dominated by the free electrons, mainly derived from the valence electrons, in the metal. The ions maintain the neutrality of the entire metal and do not interact with free electrons, meaning that the situation is equivalent to a large number of free electrons in a uniform metal positive ion background field. Under the first-order approximation of the Drude model, the interaction between light and metal can be regarded as the forced oscillation of a damped harmonic oscillator with a natural frequency of zero under the incident electric field $E$. Through the kinematics solution of the damped harmonic oscillator and the introduction of the metal macroscopic polarization, the complex permittivity and complex refractive index of the metal can be obtained respectively.

The surface plasmon oscillation, the collective free electron oscillation behavior occurring at the interface between the metal and the dielectric, can propagate along the interface or be localized in the metal nanostructure, which generally includes the propagating surface plasmons (also called surface plasmon polaritons, SPP) and localized surface plasmons (LSP). The
collectedly oscillating electrons on the metal-dielectric interface form alternately distributed positive and negative charges along a certain direction of the interface, causing the collective oscillating energy of the electrons propagating forward (SPP). The oscillation frequency of SPP satisfies the dispersion relationship obtained by Maxwell’s electromagnetic equations and appropriate boundary conditions, which can be expressed as,

\[ \beta = k_0 \sqrt{\frac{\varepsilon_{d} \varepsilon_m}{\varepsilon_d + \varepsilon_m}} \]  

(2-1)

where \( \beta \) is the propagation constant of SPP, \( \varepsilon_d \) and \( \varepsilon_m \) are the permittivity of the dielectric and metal material, respectively. For the low frequency region, \( \varepsilon_m \rightarrow -\infty \),

\[ \beta = \lim_{\varepsilon_m \rightarrow -\infty} \left( \frac{\omega}{c} \sqrt{\frac{\varepsilon_d \varepsilon_m}{\varepsilon_d + \varepsilon_m}} \right) = \left( \frac{\omega}{c} \sqrt{\varepsilon_d} \right) \]  

(2-2)

SPP is a quasi-static surface bound wave; the coupling ability of incident light and free electrons and the binding ability of electric field near the interface are close to zero. As \( \omega = 0 \), there is no incident light hitting the metal. As \( \varepsilon_d \rightarrow -\varepsilon_m \),

\[ \beta = \frac{\omega}{c} \sqrt{\frac{\varepsilon_d \varepsilon_m}{\varepsilon_d + \varepsilon_m}} \rightarrow \infty \]  

and \( \omega = \omega_{\text{SP}} = \frac{\omega_{p}}{\sqrt{1 + \varepsilon_d}} \). \( \omega_{\text{SP}} \) is the limiting frequency of short wavelength, indicating that the incident light cannot be transmitted inside or on the surface of the metal and the metal exhibits extremely high reflectivity. Since the propagation constant \( \beta \rightarrow \infty \), the wavelength tends to zero, and the wave vector component in the normal direction of the interface tends to infinity, and the oscillating field decays rapidly in the normal direction of the interface and is greatly confined near the interface. In the small area near the interface, the effective dielectric constant of the metal-dielectric system \( \varepsilon_{eff} = \varepsilon_m + \varepsilon_d = 0 \). Therefore, the SPP is a quasi-static surface wave near the frequency \( \omega_{\text{SP}} \). Fig. 2 shows the metal full-frequency dispersion relationship. In fact, the electronic damping frequency of real metal is not equal to zero \( (\gamma \neq 0) \), so \( \beta \) does not completely approach infinity as \( \omega = \omega_{\text{SP}} \), a finite value allows a certain combination mode between \( \omega_{\text{SP}} \) and \( \omega_p \).

LSP is a non-propagating surface wave formed by the coupling of free electrons in metal nanostructures and incident light. The curved surface of the metal structures can provide an effective restoring force for the collectively oscillating electrons driven by the light field. Therefore, the maximum value of field enhancement can be generated near the surface of the metal nanostructures at a specific frequency, corresponding to the localized surface plasmon resonance (LSPR). LSP can be directly excited by the incident light because the scattering effect of the curved surface in metal nanostructure on the incident light field makes it a wide wave vector value distribution. The optical properties of metal nanostructures are dominated by LSPR, while the resonance peak position and full width at half maximum (FWHM) of the spectra depend on the geometric parameters of the structure, material composition and the surrounding environment.

In 1908, Gustav Mie applied Maxwell’s equations and appropriate boundary conditions to calculate the optical response of metal micro–nanospheres (Mie theory). Soon after, Richard Gans broke the limitations of Mie theory and extended it to Gans–Mie theory, being able to describe the optical response of ellipsoidal metal particles. Until then, the related theories are capable of describing most of the plasmonic metal nanostructures and corresponding effects in a reliable way.

3. Plasmonic metal nanotips (PMNTs) with ultra-small tip sizes

The investigation of the electromagnetic properties of metal tip structures can date back to the invention of the lightning rod in the early 18th-century, in which, the excellent charge conduction and localization abilities of the metal lightning rod perfectly demonstrate the most important and remarkable properties of metal tip structures. In contrast to the bulk metal lightning rod, metal nanotips demonstrate fascinating and far-reaching electromagnetic properties in optics, photonics and photoelectronics. The most remarkable plasmon effect of metal nanotips is LSPR-enabled extremely great electromagnetic field energy focus and charge localization at the tip. Such a plasmon induced high density of energy localization in an ultra-small area or a point gives rise to lots of significant applications in single molecule detection, field emission, nonlinear photonics, beyond classic light diffraction limit, high-resolution imaging energy/spectra enhancement etc. It is commonly known that the geometry and dielectric environment have a great influence on the SPR behavior of metal nanostructures. In most of the PMNT involved fields, the smoother surface, smaller diameter and sharper tip are always the holy grail since the electromagnetic enhancement is inversely proportional to the feature size of the structure. Although metallic tips and tapers have been fabricated via the metallization of fibers, electrochemical etching, electron-beam induced deposition, complex multistep processes, or standard metal deposition, the controllable and efficient fabrication of high-quality PMNTs is challenging. Electrochemical etching of metal nanowires is generally the simplest strategy to obtain high quality PMNTs with an apex of around 50–60 nm, even to sub-10 nm based on the anisotropic etching rate of metal.

---

**Fig. 2** The full frequency dispersion relation of metal materials.
materials. As shown in the inset in Fig. 3a, the tip apex is sharper than 50 nm. Noble metal gold (Au) and silver (Ag) nanotips can be obtained by electrochemical etching of gold and Ag nanowires, which demonstrate high spectral enhancement ability. Jens Steidtner and Bruno Pettinger demonstrated an ultrahigh vacuum tip enhanced Raman technology, achieving 15 nm resolution on single dye molecule imaging with Au nanotips with radii of 15–20 nm. Coating noble metal Au and Ag films on electrochemically etched W nanotips is another effective electrochemical etching process to obtain high quality PMNTs. Because the charge and electromagnetic field are generally confined on the surface of the plasmonic metal structures, the W core only imposes a weak influence on the SPR behavior of the noble metal coating layer. In such core–shell metal nanotip structures, electromagnetic etching of the tungsten core guarantees high quality tip topology and the noble metal coating shell makes sure efficient light coupling. Besides, followed by the subsequent focused ion beam (FIB) milling process, a kind of asymmetrical Au nanotip with an end diameter of 15 nm was fabricated, showing promising application in two photon excitation near-field fluorescence imaging. In the fields of tip enhanced spectroscopy, detection, field emission and imaging, the tips are integrated with an atomic force microscope (AFM) and scanning tunneling microscope (STM), becoming two powerful tools in scientific studies. Benefiting from the ultra-high SPR excitation, energy focusing and about $10^4$ field enhancement, tip-enhanced near-field spectroscopy is now able to reach ultra-high sensitivity and resolution at the single molecule/single layer molecule level.

The far-field coupling related nanofocusing, waveguide and nanolight source are another kind of vital fields for PMNTs. In contrast to tip-enhanced near spectroscopy, in far-field coupling systems, due to intrinsically short dephasing times and small absorption cross sections, high efficiency energy coupling,
transformation and emission in metal nanotips are desperately required. Mark I. Stockman predicted theoretically that SPPs propagating toward the tip of a tapered plasmonic waveguide will slow down and asymptotically stopped when they tend to the tip, causing accumulation of energy and giant local fields at the tip. Such a rapid 3D adiabatic nanofocusing lays the foundation of the high-efficiency coupling of the far-field radiation to the near-field zone, where the 3D energy concentration occurs at the tip of a smoothly tapered metal plasmonic waveguide. On the basis of this, for overcoming problems with background illumination in apertureless microscopy and higher coupling efficiency, grating coupling of surface plasma onto metal nanotips was proposed and being used as a nano-confined light source (Fig. 3b and c). By writing one-dimensional gratings onto the tip of 3D tapered conical nanotips through focused gallium ion beam sputtering, Ropers et al. demonstrated an efficient nonlocal optical excitation of the apex of a nanostructured metal taper based on SPP grating-coupling. The SPPs propagate over more than 10 μm along the tip shaft toward the apex of the tip, where they are reradiated into the far field. Raschke et al., subsequently, found that the breakdown of the adiabatic conditions occurred only within the last 10 s of nanometers. Despite the symmetry breaking side-on illumination and the resulting SPP excitation confined along the azimuth, the tip effectively acts as a mode filter, giving rise to a propagation-induced, nanoscopic SPP excitation with radial symmetry at the apex, and suggesting the robustness of conical SPP focusing with respect to asymmetries and imperfections in the SPP generation. The resulting dipolar nanoemitter with ~20 nm spatial confinement, that is, more than 1 order of magnitude beyond the diffraction limit, represents a novel light source. Soon after, by nanoscale field concentration via adiabatic SPP nanofocusing into monolithic Au tips, the same group demonstrated the independent nanometer spatial and femtosecond temporal optical waveform control, achieving a 10 nm spatially and few-femtosecond temporally confined excitation, as shown in Fig. 3d, showing great potential in scattering-type scanning near-field optical microscopy (s-SNOM). Becker and co-workers observed a coherent light scattering from a nanometer-sized metal tip-thin film gap region, finding a steep increase in scattering intensity in a near-field signal selected in k-space. As shown in Fig. 3e and f, this gap plasmon is different from conventional adiabatic nanofocusing determined apertureless scattering type SNOM, which allows confinement of light at the very apex of the taper to dimensions that are not limited by its apex diameter. Besides, the gap type localized plasmonic field intensity can exceed that of a bare apex mode by at least 2 orders of magnitude, revealing broader applications.

Despite the above conventional PMNTs, for specific purpose and scale-up fabricating requirements, some unconventional alien PMNTs also were purposely fabricated with various nano-fabricating techniques. Fig. 3g and h demonstrate laterally tapered planar Ag waveguide concentrated infrared surface SPPs. Taper/hole array microstructures are fabricated in a 140 ± 5 nm thick Ag film on the sample by electron beam lithography (EBL) followed by lift-off. By using erbium (Er) upconversion luminescence technology, the excitation, propagation, and concentration of SPPs in laterally tapered Ag waveguides are visualized. Thereafter, their study on tapered Au waveguide on a planar dielectric confirmed that SPP focusing is the dominant mechanism producing the spot of maximum intensity at the taper tip, and the nanofocusing is attributed to a guided SPP mode that has an asymmetric electric field distribution across the tapered waveguide. Park et al. fabricated vertically aligned funnel-shaped 3D PMNT arrays by directional photofluidization lithography (DPL), with which, the sharpness, shape, and orientation nanofunnel tips can be precisely controlled in a scalable and deterministic manner. A 3D plasmonic nanofocusing of light with patterned metallic pyramids is demonstrated in Fig. 3i. Gratings on the faces of the pyramids convert linearly polarized light into plasmons that propagate toward and converge at a ~10 nm apex. 3D Au and Ag pyramids plasmonic light nanofocusing tip structures with smooth surfaces were fabricated by a simple and impurity-isolated template stripping process. Moreover, the plasmonic properties in a combined structure with a metallic pyramid and an integrated C-shaped aperture, metallic thin layer and a dielectric pyramidal body structure composed of a tip and one-side metal-coated pyramidal cantilever tips were also investigated, exhibiting remarkable applications in non-linear photonics.

4. Plasmonic ultra-thin metal films

Ultra-thin metal films (UTMFs) are another interesting plasmonic structures with a long history. In ancient China, an ultrathin gold film was generally stuck to great Buddha to create a powerful and solemn atmosphere. After thousand times of repeated forging, a piece of thin Au plate turns into a hundred nanometer ultrathin Au film. The excellent mechanical ductility and chemical stability make sure the uniformity and corrosion resistance of the sticky Au film, enabling thousand years’ shininess, holiness and resplendence of the Buddha. Currently, with the thickness of UTMFs down to the nanometer or atomic level, unique properties, such as new quantum optical effects, peculiar dielectric properties and strong plasmon field enhancement abilities lead to more exciting applications in optical sensing, imaging and flexible displays. Magnetron sputtering, thermal evaporation, electron beam evaporation, atomic layer deposition (ALD) and molecular-beam epitaxy (MBE) are traditional methods to fabricate UTMFs. However, large surface roughness and discontinuity of metal films are common disadvantages of sputtering and evaporation processes. Although ALD and MBE are theoretically able to control atomic level thickness and good continuity during the metal film growth, the low yield of ALD and rigorous substrate matching of MBE are bottlenecks for these two strategies.

As schematically illustrated in Fig. 4, the growth of metal films is generally governed by three known models, that is, Volmer–Weber (bulk island), Frank–van der Merwe (layer by layer) and Stranski–Krastanov (layer-plus-bulk). During the growth of metal films, the presence of step and kink Ehrlich–Schwoebel effects (SESE and KESE) leads to arbitrary
orientation of surface morphology. The diffusion of atoms along the steps not only affects the morphological stability of the steps, but also can induce up-step mass current leading to mound formation. Consequently, reducing the Ehrlich–Schwoebel carrier become an impactful way to improve the metal film quality. Introducing hybrid materials (Al, Cu, Ti, Ni etc.), acting as the adsorption center and seeds to promote the nucleation of host metals, is an effective way to obtain UTMFs with small roughness.

The inset SEM images (without (left) and with a Cu seed layer (right), the scale bars are 500 nm) and the plot in Fig. 5a demonstrate that after 1 nm of Cu seed layer deposition, atomically smooth and stable Ag films with 0.5 nm roughness on a silica substrate were successfully fabricated. The introduced Cu promotes nucleation, thus reducing the grain size. The same strategy was demonstrated by Al-doped Ag films, in which, 15 nm of 4% Al atom doped Ag films maintained smooth surface morphology even after 300 °C annealing in N₂, exhibiting excellent thermal stability. The Al atoms enhance the nuclei density of films on an oxide surface due to the much larger bond strength of Al–O bonds than Ag–O bonds, and a capping layer spontaneously formed over the Al-doped Ag films restrains the surface diffusion and mass transportation of Ag atoms, strengthening the thermal stability of Ag films. The semiconductor or metal buffer layers are also crucial in fabricating sub-10 nm UTMFs. The surface defects of buffer layer anchor the metal atoms, decreasing the atom diffusion, and strong bonding between metal atoms and semiconductor non-metal atoms inhibits the grain growth, reducing the surface roughness. Recently, Volkov et al., depending on the monolayer MoS₂ buffer layer, achieved 3–4 nm thickness ultra-thin Au films with continuous smooth morphology and Drude plasmonic response on a SiO₂ substrate. Besides, by introducing surfactants and additives (metallic surfactants, molecular monolayer surfactants, polymer surfactants, metal additives and gas additives), seed layer-free growth, self-assembled organosilane monolayer linkers polymers etc. were also used for preparing sub-10 nm UTMFs. The substitution of metal atoms with polymers in the buffer layer decreases the cost and eliminates the subsequent deterioration of the optical and electrical properties of the metal thin films.

For a long time, one of the most important applications of UTMFs is the transparent electrodes in flexible devices. The strong SPR, high electrical conductivity and low loss in visible and near-infrared of Ag and Au make them perfect substitutions for traditional indium tin oxide film (ITO) film materials. The good malleability and flexibility of Au and Ag UTMFs bridge the gap between rigid ITO films and flexible photoelectrical devices. More importantly, the general UTMF fabrication methods are compatible with standard CMOS (complementary metal-oxide-semiconductor transistor) and MEMS (micro-electromechanical systems) nano-manufacturing, promising better compatibility and integrated performance. Yun's group demonstrated a nitrogen-doped copper film with an average transmittance of 84% over a spectral range of 380–1000 nm and a sheet resistance lower than 20 Ω sq⁻¹. The corresponding flexible organic solar cell achieves a power conversion efficiency of 7.1%, exceeding that (6.6%) of solar cells utilizing the conventional ITO film (Fig. 5b). In the flexible photoelectrical device field, Liu et al. demonstrated giant Kerr response of ultrathin Au films from quantum size effects in metal quantum well (MQW) structures (Fig. 5c–e). The measured nonlinear Kerr susceptibility reaches 2.06 × 10⁻¹² m² V⁻², showing a four orders of magnitude enhancement compared with the intrinsic value of bulk Au. Compared with the LSPR based nonlinear effects, the MQWs have the advantages of broad-band operation and free of scattering, promising high compatibility with existing nano-photonic waveguides for integrated nonlinear optics. Moreover, although the SPR on metal films is generally localized on the surface of the film,
new SPR effects occur as the thickness scales down to several atoms of UTMFs. Compared to the 5 nm thick Au film, the 3 nm thick Au film fabricated by a cryogenic temperature deposition process not only displays improved transmittance, but also degraded relative electric permittivity, revealing decreased plasmonic adsorption. El-Fattah et al. demonstrated wafer scale 8 atomic monolayer (~1.9 nm) crystalline Ag films with excellent plasmonic and electrical properties, as shown in Fig. 5f–h.\(^74\) Despite the plasmonic resonance shift, they also obtained sharp near-infrared optical spectra with a quality factor of \(~4\) and strong plasmon localization in such an atomic UTMF, suggesting great potential for electro-optical modulation and optical sensing applications. Recently, Pruneri and co-workers demonstrated plasmons in few-nanometer-thick Au UTMFs, with clear evidence of new dispersion regimes and large electrical tunability.\(^77,78\) Resonance peaks at wavelengths of 1.5–5 \(\mu\)m are shifted by hundreds of nanometres and amplitude-modulated by tens of percent through gating using relatively low voltages. Combined with a large-scale fabrication approach, the fabricated films can find applications in transparent conductors, plasmon-enhanced spectroscopy, optical biosensing and electrochromic devices.

Beyond the new plasmonic effects and applications mentioned above, UTMFs imply broader future applications in the fields of optoelectronics, nonlinear optics, plasmon-enhanced detection and sensing, thermal manipulation, two-photon emission and plasmonic color printing.\(^79,81\)
5. Quantum plasmonics in ultra-small metal particles

Ultra-small metal particles (USMPs) are one of the simplest but extraordinarily vital plasmonic structures. With the decease of particle size, the plasmonic behaviors of USMPs evolve from classical electrodynamic theory to quantum theory. The quantum size effect, nonlocal effect, electron density spill-out and tunneling effect become prominent as the particle size scales down to the sub-nanometer scale.82–88

The plasmonic spectrum of a single noble metal spherical nanoparticle experiences a blue-shift with the gradual decrease of the particle size, and such a dependence can be well-described by classical electrodynamic theory. However for USMPs with size smaller than ~2 nm, things seem to go in a confused way, both blue-shift and red-shift of plasmonic spectra appeared.89,90 The underlying reason is that as the particle size scales down to the sub-nanometer scale, the quantum effects cannot be neglected. In certain circumstances, the quantum effects even play a dominating role in determining the spectral behaviors of the USMPs. In 1987, S. Das Sarma investigated the surface-plasmon excitation in simple metals using a hydrodynamical model based on an electron gas-based nonlocal theory.82 The infrared divergence inherent in local theories is suppressed by nonlocal corrections, and as the velocity of the probe particle vanishes, the excitation probability goes to zero. He also showed that for low energy external protons interacting with the metal surface, nonlocal effects reduce the excitation probability compared with the local theory. For a long period, such nonlocal theories prevail in many quantum plasmonic systems. In 2014, Mortensen et al. put forward a comparatively simple semiclassical generalized non-local optical response theory that unifies quantum pressure convection effects and induced charge diffusion kinetics for plasmonic nanostructures.91 The theory described both the

Fig. 6 (a) Extinction cross-section for the dipole resonance in a metal sphere.91 (b) Collection of normalized, deconvoluted EELS data from Ag particles. (c) Plot of the SPR energy versus particle diameter, with the inset depicting bulk resonance energies.93 (d) Schematic illustration of the "SPL" approach.94 (e) PL of Au nanodisks with different sizes95 (reproduced with permission from ref. 91 Copyright 2014, ref. 93 Copyright 2012 Springer Nature, ref. 94 Copyright 2016 and ref. 95 Copyright 2012, the American Chemical Society).
frequency shifts and size-dependent damping in individual metallic nanoparticles as well as the broadening of the cross-over regime from bonding dipole plasmons to charge-transfer plasmons in plasmonic nanoparticle dimers (Fig. 6a). However, with hard-wall boundaries, the standard hydrodynamic Drude model is capable of giving out accurate quantitative predictions for the optical response of plasmonic metal nanoparticles. For the nanosystem where the surface effect caused by the electron density spill-out in free space cannot be neglected, the standard hydrodynamic Drude model is less accurate. Toscano et al. studied the resonance shifts and spill-out effects in self-consistent hydrodynamic nanoplasmamics. They constructed a self-consistent hydrodynamic model to explain the size-dependent surface resonance shifts of Na and Ag nanowires and nanospheres. The theory provides a good attempt to develop a versatile hydrodynamic theory and numerical method that account for both retardation and electronic spill-out.

Yet with lots of challenges, the experimental investigation on USMPs also demonstrates exciting advances. The theoretical descriptions and experimental results of USMPs with size larger than 10 nm are well matched under the classical electromagnetic framework, and will not be the main points here. The USMPs or clusters are generally synthesized by bottom-up wet chemical processes, and in some occasions, ligands are indispensable in precise control of the cluster size and atom numbers. The addition of ligands may weaken the optical response of metal USMPs and clusters due to metal–ligand interaction. In 2012, Scholl et al. synthesized individual ligand-free silver nanoparticles and investigated the plasmon resonances by aberration-corrected TEM imaging and monochromated scanning TEM electron energy-loss spectroscopy (EELS). They found that as the diameter of nanoparticle decreases from 20 nm to less than 2 nm, the plasmon resonance blue-shifts to higher energy by 0.5 eV, which is a substantial deviation from classical predictions, indicating a transition to the quantum-influenced regime (Fig. 6b and c). Soon after, Haberland pointed out that the blueshift of LSPR with the decrease of particle size in Scholl’s work is flawed, because they neglected the influence of the carbon substrate and possibly the residual ligand. A convincing way is taking both the small and large size limit into account in the study of size dependent plasmon resonance in metal nanoparticles. Therefore, what is the small-particle limit of the surface plasmon becomes an unavoidable and fundamental question to be answered.

In 2016, Zhou et al. carefully studied the evolution from the plasmon to exciton state in ligand-protected atomically precise gold nanoparticles. They investigated the transition from the metallic (or plasmonic) to molecular state in metal nanoparticles by performing ultrafast spectroscopic studies based on atomically precise thiolate-protected Au25, Au32, Au144, Au123, Au–520 and Au–940 nanoparticles. Three distinct states: metallic (size larger than Au133 and larger than 2.3 nm), transition regime (between Au133 and Au144, 2.3–1.7 nm) and non-metallic or excitonic state (smaller than Au144, smaller than 1.7 nm) are clearly demonstrated. The results of CO oxidation and electrocatalytic oxidation of alcohol proved that the transition state apparently impacts the catalytic ability of the nanoparticles, providing a solid guide to the synthesis of atomic metal catalysts in the catalysis process. Furthermore, Higaki and coworkers observed sharp transition from nonmetallic Au246 to metallic Au279 with nascent SPR using thiolate-protected Au nanoparticles. By femtosecond transient absorption spectroscopic analysis and steady-state absorption spectroscopy, they found that the Au279 nanocluster shows the metallic state of the particle while Au246 displays a nonmetallic electronic structure of the nanocluster. The findings are surprising that only 33 gold atoms lesser of Au246 than that of Au279, but the optical properties and electron dynamics between Au246 and Au279 are dramatically distinguished from each other. Recently, Wang’s group investigated the plasmonic evolution of atomically size-selected Au clusters by electron energy loss spectroscopy (EELS). They systematically investigated the plasmonic response of Au clusters with atom number (N) from 100 to 70 000. The bulk plasmon remains unchanged above Au887 and disappears as N decreases to 887. The SPR firstly redshifts from 2.4 to 2.3 eV above Au887 and blueshifts towards 2.6 eV down to Au100 and finally splits into 3 fine features. Naturally, an evolution picture of plasmon physics with atomically-precise was revealed and divided to 3 regimes, i.e. classical plasmon (Au887 to Au70 000), quantum confinement corrected plasmon (Au100 to Au887) and molecule related plasmon (particle size less than Au100). The observations from solid-state plasmon classical electrodynamics, quantum confinement corrected plasmon physics, and molecular plasmon theory pave the way for novel applications of plasmonic USMPs. Note that the changes in crystal structures during the size decrement are also the vital phenomena that should be considered in investigating the quantum size effects of USMPs.

Compared with ultra-small clusters synthesized by bottom-up wet chemistry, the artificial metal atoms/molecules fabricated by top-down techniques, such as EBL, FIB, chemical etching RIE etc., possess much larger size. However, benefiting from the better size and morphology control ability of top-down nanofabrication approaches, the investigation on the plasmonic optical properties and applications of artificial monomer atoms is always of great importance and desirable.

To overcome the low efficiency and high time-consumption of transitional top-down approaches, recently, our group developed a “Sketch and Peel” lithography (SPL) approach for large scale, high efficiency, multiscale and high precision metal nanostructure fabrication (Fig. 6d). By taking the method analogous to paper-cutting, in the SPL approach, only the outside outlines of nanostructures are needed to be exposed, which greatly improves the fabrication efficiency of EBL. Moreover, the effective inhibition of the proximity effect enables ultra-high precision and shape control ability. We estimate that the fabrication efficiency is improved by hundreds of times. By extending the idea to the FIB process, we found that the SPL process is compatible to state-of-the-art helium ion beam milling technology and both the fabrication efficiency and precision are greatly improved. Furthermore, by introducing a self-assembled monolayer to engineer the surface energy of the substrate, the adhesiveness of the Al film outside the outline template is significantly decreased to implement the selective
peeling process, and various plasmonics Al nanostructures with sub-20 nm feature sizes were successfully achieved by the SPL approach, suggesting the tremendous potential and fabricating ability of the technology. With regard to optical properties and applications of artificial plasmonic metal atoms, the hybridization model for the plasmon response of complex nanostructures provides a simple effective guideline. As shown in Fig. 6e, Hu et al. performed a systematic study on the photoluminescence (PL) and scattering spectra of EBL defined individual Au nanodisk structures. They found that the PL spectra and scattering spectra of individual Au nanodisks showed the same dependencies on shape, size, and plasmon coupling as the particle plasmon resonances, providing conclusive evidence that the PL in Au nanostructures indeed occurs via radiative damping of plasmon resonances driven by excited electrons in the metal itself and the PL peak position can be adjusted by their plasmon resonances. Subsequently, Chen et al. found that the Cr adhesion layer between the SiO₂ substrate and Au nanostructures not only weakens the scattering intensity of the structures, but also leads to the red-shift of the spectral position. Lawrence J. Tauzin et al. explored the relationship between plasmon damping and PL in Au nanorods. The damping resulting from poor crystallinity in nanostructures and adhesion layers can be quantitatively linked to the decrease in PL quantum yield, and the annealing process can improve the emission intensity of lithographically prepared Au nanorods.

6. Plasmon coupling in metal nanogaps

Plasmonic metal nanogaps, have a unique and vital position in light–matter interactions. Sub-nanometer metal nanogaps are able to focus the incident light into the region beyond the classic diffraction limit, and have wide applications in subwavelength waveguides, molecular detection, plasmonic antennas, medical treatment, and nano-optics. As the distance of nanogaps scales down to sub-nanometer, some quantum effects, such as nonlocal electromagnetic effects, optical rectification, inelastic electron tunneling, rings interference generally emerged (Fig. 7). In this small section, we will briefly give a comprehensive discussion on the typical advances in theories, fabrication and applications of sub-nanometer nanogaps.

**Plasmonic metal split rings (PMSRs) with sub-nanometer split gaps**

The PMSR nanostructure with a sub-nanometer split gap is a crucial and unique gap structure, in which, generally, the distance of split gaps can be well-controlled by top-down fabrication processes. For large metal split rings, the main attention is focused on their excellent applications in THz applications. For PMSRs with sub-nanometer split gaps, top-down EBL and FIB are two of the most vital and popular techniques. Besides, new strategies and methods are also demonstrated. For example Gwinner et al. proposed a fast and cheap, large-area, high-coverage shadow nanosphere lithography technique to fabricate periodic metallic split-ring resonator structures. In this fabrication process, the inner- and outer-ring diameters, gap angles, as well as thickness and periodicity of the metallic split-ring resonator structures can be well controlled.

The unique geometry structures of PMSRs with sub-nanometer split gaps enable them with many attractive nonlinear effects, such as pronounced Fano resonance, high-order harmonic wave generation and light twisting. For example, Cao et al. designed gradient-rotation split-ring
antenna metasurfaces to simultaneously generate and separate pure optical vortices in a broad wavelength range, which has the potential in miniaturized on-chip orbital angular momentum-multiplexers. As shown in Fig. 8a and b, by hybridizing with nanodisks, the split-ring resonators/disk nanocavities with strong Fano resonance are capable of generating high-order magnetic modes, indicating great promise in refraction index sensors. Besides, multiple-split rings with sub-nanometer feature sizes can confine the LSPR and the associated field enhancements to specific areas, meaning higher flexibility in manipulating the parameters of field hot spots, thus leading to higher capability and availability for sensing. More interestingly, by decreasing the inner radius of split rings to zero, split nanodisks are obtained. Fang et al., introduced a single metallic nanodisk with a missing wedge shaped slice to support a Fano resonance, in which, the coupling between a hybridized plasmon resonance of the disk and a narrower quadrupolar mode supported by the edge of the missing wedge slice supports the Fano line. Very recently, we demonstrated the pronounced Fano resonance in the near-infrared region using Au split nanodisk structures fabricated by the SPL process. The ultra-small split gaps of 15 nm enable strong plasmon coupling and field enhancement in the gap region. The Fano resonance behavior can be easily controlled by changing the disk diameter and split length. We found that the Fano resonances observed in the split nanodisks are induced by mode interference between the bright antibonding dipole mode of split disks and the subradiant mode supported by the narrow split gap. Enabled by the giant near-field enhancement, strong wavelength-dependent second harmonic generation was observed under near-infrared excitation in split nanodisks (Fig. 8c and d). Recently, a tunable dual-split-disk resonator composed of top and bottom Au split-disk resonators in the THz wave was theoretically demonstrated, which exhibits electromagnetically induced transparency characteristics and a high merit factor, showing potential in sensing, programming, filtering, and attenuating devices.

**Artificial plasmonic metal nanoparticle pairs (APMNP)s with sub-nanometer nanogaps**

Two closely positioned APMNPs with sub-nanometer nanogaps are the simplest and extremely significant nanostructures in the plasmonic coupling field. The high controllability in fabricating artificial plasmonic molecules by top-down micro/nano-manufacturing methods makes APMNPs an ideal candidate to explore new effects in sub-nanometer gap size structures. Generally, the gap regions, the so-called electromagnetic energy “hot spots”, enable huge field density and large intensity enhancement, thus leading to great potential in sensing and detection. The gap distance between two nanoparticles can be precisely controlled. With the gap size of APMNPs down to sub-nanometer, the associated optical properties also experienced a transition from classic electrodynamics to quantum dynamics, that is, the quantum nonlinear effects, such as size effect, nonlocal effect, electron spill-out effect, Coulomb block effect and quantum tunneling effect will gradually become pronounced.

In general, top-down and bottom-up nanofabrication techniques are all capable of fabricating paired metal nanoparticles with sub-nanometer gaps. Compared to highly controlled top-
down approaches, bottom-up wet chemical synthesis processes are more superior in obtaining ultra-small nanogaps. However, both techniques have shortcomings in fabricating APMNPs with sub-nanometer gaps, so optimizations on high-efficiency, reliability and time/economy consuming fabrication approaches are always desirable and indispensable. In top-down methods, the proximity effect in lithography is the biggest challenge. In 2002, Liu et al. developed an automated proximity effect correction algorithm based on the self-consistent exposure equalization technique to fabricate electrode pairs. Koh et al. patterned triangle dimers with sub-1 nm gaps on electron-transparent 30 nm thick silicon nitride transmission electron microscope (TEM) windows based on the fact that the substrate with lower density and atomic number could greatly reduce the range and intensity of the proximity effect during the lithography process. In 2010, Theiss et al. demonstrated an angle evaporation technique that can fabricate arrays of plasmonic nanoparticles with the gap size of 1 nm (Fig. 9a). As shown in Fig. 9b, a two-step EBL process to decrease overlap was proposed by Zhu and Crozier, and metallic dimers with angstrom-scale gaps were achieved. Lately, Hoogenboom’s group developed an electron-beam induced deposition etch mask method, achieving Au dimers with sub-10 nm spacing distance. For the FIB process, Ga-ion beams generally only can obtain structures with gap size larger than 10 nm due to the low resolution. Carl Zeiss Company reported a 4 nm nanogap on a suspended Au film obtained using He-ion lithography. Recently, by combing EBL, reaction ion etching, HF etching and Ga⁺ FIB milling processes, Cui et al. developed a method to produce ~2 nm Au gap electrodes. Our group also developed a series of EBL-based approaches to reliably fabricate nanostructures with sub-10 nm gaps (Fig. 9c). Besides, other methods such as capillary-force self-assembling of patterned nanostructures, tip-based approaches, controlled metal deposition and breaking or crack strategies are also able to achieve sub-10 nm gap APMNPs, which recently have been reviewed by Yang et al., and will not be discussed in this paper.

To better understand and illustrate the optical response transition of APMNPs from classic to quantum effects, innovative advances and developments are presented uninterruptedly in the past few decades. Early studies on the quantum tunneling exposed that when the “hot spot” between two Ag nanospheres is small, its local field is almost identical to a replaced plate. As the gap size is smaller than 0.6 nm, the quantum tunneling will dramatically reduce the enhancing ability of the interstice between nanoparticles. In 2011, Esteban et al. proposed a quantum-corrected model that incorporates quantum-mechanical effects within a classical electrodynamical framework. Tserkezis et al. carefully studied the underlying reasons for nonlocal damping induced plasmon-enhanced fluorescence reduction in ultra-narrow gaps. Recently, by incorporating Feibelman d-parameters, as displayed in Fig. 9d, Gontalves and co-workers demonstrated that the neglect of quantum effects in sub-nanometer gap structures can be well-remedied in a unified theoretical treatment of mesoscopic electrodynamics, which bridged the gap between the purely quantum and classical domains. Along with theoretical advancements, experimental investigations on the quantum effects of APMNPs also showed enormous progress. Baumberg’s group revealed the quantum regime of tunneling plasmonics by measuring the electrical and optical properties of Au nanostructures with controllable sub-nanometer separation at the same time. Combining with the quantum models of plasmonic systems, the singularities predicted by the classical theories were eliminated. Their experimental results suggest that the quantum tunneling constructs a quantum limit for plasmonic field confinement (~10⁻⁵) for visible light. In 2013, by employing STEM and EELS, Scholl et al. demonstrated the observation of quantum tunneling between two coupled plasmonic Ag nanoparticles as the gap size is reduced to atomic dimensions. Their observations demonstrate a gradual transition from classical calculations to quantum theories, accompanied by the transition of optical spectra from the redshift of the dominant dipolar peak to electron tunneling induced intensity reduction as the gap size is gradually reduced from 7 nm to smaller than ~0.5 nm (Fig. 9e and f). Furthermore, the bonding dipolar mode is substituted by a dipolar charge transfer mode in the overlapped particle systems. Nijhuis et al. presented another interesting quantum plasmonic gap system, a molecular tunnel junction, composed of self-assembled monolayer bridged plasmonic resonators at length scales in the range of 0.4 to 1.3 nm. The tunnel barrier width and height can be controlled by the properties of the molecules. Depending on EELS, they directly observed the showing that the potential plasmonic induced tunneling can be observed in reconciling molecular electronics.

It is worth noting that artificial plasmonic metal dimers are also crucial parts in revealing the quantum size effects of APMNPs with sub-nanometer gaps. Artificial metal dimer molecules are the simplest inter-coupling structures that are composed of two “atoms”. The inter-coupling of two “atoms” makes metal dimers having more complex LSPR modes. Zuloaga et al. proposed a fully quantum mechanical description of the plasmon resonances in a metal dimer by using time-dependent density functional theory, showing that for gap sizes below 1 nm quantum mechanical effects will greatly shape the optical response and drastically decrease the field enhancements relative to classical predictions. In 2012, we fabricated nanoprisms with varying gap and bridge widths. By systematic reduction of the gap to 0.5 nm size and narrowing the bridge to 3 nm width (Fig. 9g), we observed, for the first time, the gradual appearance of the charge-transfer plasmon mode and identify higher-order antenna modes, shedding light on the evolution of plasmon modes in the transition regime from a disappearing gap to a conductive bridge between two metal nanoprisms (Fig. 9h and i). One year later, we developed a three-dimensional nonlocal hydrodynamic solution of Maxwell’s equations and investigated the electron-energy loss probing of nonlocal effects in metal structures below the nanoscale. For connected bowties, we showed that it is the dipole–dipole mode that is most affected by nonlocality, resulting in blueshifts as large as 0.7 eV for a dimer made up of prisms with a side length of 5 nm. The experimental EELS for Au
nanoprisms connected by bridges as narrow as 1.6 nm indicated that nonlocal effects do not appear to be pronounced. In such nanoprisms, quantum confinement and granularity at the nanometer scale need to be taken into consideration. In addition to classic boundary limit investigation, nonlinear effects and corresponding application of plasmonic metal dimers are also inspiring. Capasso et al. reported a bowtie plasmonic quantum cascade laser antenna that can confine coherent mid-infrared radiation well below the diffraction limit, showing great potential in spatial-resolved high-resolution chemical and biological imaging and spectroscopy. Soon after, nanogap detectors for fast DNA analysis, high-harmonic generation
devices, and single-molecule sensing systems based on artificial plasmonic metal dimers with sub-nanometer gaps are also demonstrated. Moreover, the single-particle level sensor composed of two different materials can be achieved by a named hetero-plasmonic dimer. Wang’s group systematically investigated the coupling behaviors of 2D homo- and hetero-plasmonic dimers with universal sub-10 nm gaps. They found that compared with homo-dimers of Au–Au and Ag–Ag,

---

**Fig. 10** (a) Schematic illustration of elevated Au bowties on top of Si posts, the calculated spatial distribution of the E field intensity, SEM image and SERS spectra of bowtie substrates. (b and c) The fabrication process (b) and SEM image (c) of AgAu 3D nanostar dimer. (d) The schematic illustration of atomic layer lithography. (e) The top-view SEM of a 5 nm-wide annular gap in a 200 nm–thick Ag film. (f) Schematic flow and corresponding SEM images in each stage of the fabrication process used for the coaxial nanocavity array using glancing-angle ion polishing. (g) Conceptual schematics of the self-folding of nanocylinders with plasmonic nanogaps for bimolecular sensing (reproduced with permission from ref. 161 Copyright 2010, ref. 162 Copyright 2014, ref. 164 Copyright 2016, ref. 165 Copyright 2020, the American Chemical Society, and ref. 163 Copyright 2013, Nature Publishing Group).
the primitive Au and Ag plasmon modes differ in energy, leading to much weaker hybridization.\textsuperscript{160}

3D plasmonic metal nanostructures (3D-PMNs) with sub-nanometer gaps

3D-PMNs are more effective than those supported by in-plane coupling since the coupling between plasmonic structures and the supported substrates weakens the inter-coupling of plasmonic metal structures. With added dimensions, decoupled 3D-PMNs enable higher flexibility in geometric structures, coupling modes and possible applications. 3D-PMNs are capable of many amazing properties and applications, including geometric and coupling mode variabilities, decreased substrate-induced loss and enhanced plasmon coupling efficiency. Besides, high-aspect ratio nanostructures are potentially of great importance in biological sensing and X-ray diffraction.

Hatab et al. described plasmonic interactions in suspended Au bowtie nanoantenna with sub-10 nm gap distance by EBL and following metal deposition processes, achieving up to 2 orders of magnitude additional enhancement compared with that of bowtie structures directly attached on substrates (Fig. 10a).\textsuperscript{161} The improved SERS enhancement is attributed to a plasmonic nanocavity effect occurring when the plasmonic interactions enter a strongly coupled regime. Wang’s group described a 3D plasmonic Au mushroom array refractive index sensing with a Wood’s anomaly and the LSPR interference induced figure of merits of $\sim 108$, which is comparable to the theoretically predicted upper limit for standard SPP resonance sensors.\textsuperscript{162} In 2014, Chirumamilla and co-workers introduced novel complex 3D star-shaped Au nanostructures with inter-particle spacing below 10 nm by EBL and the subsequent reaction-ion etching approach. Endowed with multiple tips that are decoupled from the substrate, the 3D Au nanostar dimer exhibits the ability for single-/few-molecule detection by SERS.\textsuperscript{167} Soon after, the same group fabricated bimetallic 3D nanostar dimers in ring cavities, achieving the recyclable and robust SERS substrates for signal detection from few molecules (Fig. 10b and c).\textsuperscript{168} Very recently, Wu et al. fabricated high aspect-ratio collapsible nanofinger-like plasmonic structures with precisely controllable nanogaps of sub-10 nm by combining nanoprinting lithography, reaction-ion etching, metal deposition and ALD processes. Depending on this approach, they simultaneously observed fluorescence and the Raman signal in the same spectra, providing an opportunity to analyze plasmon enhanced fluorescence with \textit{in situ} local field enhancement monitoring.\textsuperscript{168}

Another interesting 3D PMNs are vertically oriented gap structures with the gap size less than 1 nm, which are mainly demonstrated by Kim and Oh \textit{et al.} in the University of Minnesota.\textsuperscript{163,164,169–175} In 2013, as shown in Fig. 10d and e, they introduced an atomic layer lithography process and successfully fabricated wafer-scale vertically oriented gaps in opaque metal films along the entire contour of a millimetre-sized pattern, with gap widths as narrow as 9.9 Å.\textsuperscript{163} With such reliable ultra-small nanogaps, they not only observed resonant transmission of near-infrared waves through 1.1 nm-wide gaps (l/1295, an effective refractive index of 17.8), but also a resonant transmission of millimetre waves through 1.1 nm-wide gaps (l/4 000 000, inferring a field enhancement factor of 25 000). The method is capable of producing sub-nanometer gaps over millimetre-scale contours, enabling resonantly enhanced transmission of THz waves. Soon after, based on the same approach, they experimentally demonstrate that even the refractive index change produced by only a 1 nm thick dielectric overlayer (Al\textsubscript{2}O\textsubscript{3} film) can be detected by the THz waves confined in sub-10 nm metallic gaps.\textsuperscript{176} Besides, the method is also extended to the high-throughput fabrication of resonant metamaterials with ultra-small coaxial apertures by combining with glancing angle ion polishing (Fig. 10f). The realized ultra-compact, high aspect-ratio coaxial nanocavities with 2 nm gaps and 100 nm diameters exhibit higher order resonances at optical and the zeroth-order F–P mode frequencies.\textsuperscript{184} Recently, they further demonstrated the broader applications of these vertically oriented gap structures with sub-10 nm gaps in low-power optical trapping of nanoparticles and proteins, high-contrast infrared absorption spectroscopy, and nanogap electrodes, as well as mid-infrared nonlocality and quantum plasmonics.\textsuperscript{169,173,175–178} The work on vertically oriented gap structures with sub-1 nm structures greatly strengthens the power and ability of extreme small gap structures, which also promise more possibility and potential in metal plasmonic nonlinear optics and plasmonic quantum photonics.

Beyond top-down micro/nanofabricating processes, Dai \textit{et al.} demonstrated a self-assembled 3D nano-split ring originating from a 2D nanobelt for plasmon-enhanced optofluidic sensing (Fig. 10g). The realized segmented nanocylinders were achieved by electron irradiation triggered self-assembly. The optical response of the 3D plasmonic nanostructures shows an intensity enhancement over 22 times.\textsuperscript{165} Although, 3D PMNs generally enable higher coupling efficiency, the difficulties in precisely and large scale fabricating of 3D-PMNs inhibit the practical applications, the advances and breakthroughs in the more efficient and economic approaches will have heavy impact on the fields.

7. Summary and outlook

In summary, plasmonic metal nanostructures with extremely small feature sizes play crucial roles in the fields of nanophotonics, sensing, nonlinear optical effects, spectral filtering, high-resolution imaging, light focusing and quantum optics, especially in the areas that are beyond the ability of conventional devices, such as quantum nonlocality, quantum tunneling and high-order harmonic wave generation. Nowadays, motivated by the pursuit of scientific limit and technological innovation, plasmonic metal nanostructures with extremely small features have demonstrated amazing and exciting ability in light manipulation that is far beyond the capability of conventional optical devices. Specifically, plasmonic metal metamaterials play important role in quantum and nonlinear photonics, as well as related applications. Combined with dielectric and ultrathin 2D layer nanomaterials, plasmonic metal nanostructures with extremely small feature
sizes are capable of opening up a new era for planar nano-focusing, ultra-high resolution sensing, ultra-high sensitive detection and energy conversion devices and systems.\textsuperscript{179–187} Moreover, the advancement in tiny metal nanoholes also promises valuable and promising applications.\textsuperscript{188} However, the inherent high loss of metals, high cost and low manufacturing efficiency of traditional nanofabrication techniques are the main obstacles that restrict the wide and large-scale application of plasmonic metal nanostructures with extremely small feature sizes. Therefore, the advances in nanofabrication technologies, more efficient structural design and material hybrids will speed up the development of the field. Specifically, by hybridizing with functional materials, the sensitivity, precision, range and flexibility of light manipulation will be further improved, thus leading to more exciting and versatile functionalities for broader applications, as well as continuous innovation of novel concepts and effects.

**Author contributions**

H. Shi and X. Zhu prepared the original draft, S. Zhang and M. Zheng collected and classified the references and figures, H. Shi, G. Wen and H. Duan discussed the topic and keypoints of the manuscript, and G. Wen and H. Duan provided funding support. All of the authors discussed the content of the manuscript and reviewed the draft.

**Conflicts of interest**

The authors declare no conflicts of interest.

**Acknowledgements**

This work was supported by the Key Program of National Natural Science Foundation of China (No. 11832009), the College Science Foundation of Bureau of Education of Guangzhou Municipality (No. 202032830), Guangzhou Municipal Science and Technology Bureau Basic and Applied Basic Research Foundation (Grant No. 202102020962), the National Natural Science Foundation of China (No. 11672104, No. 12004150), and the Chair Professor of Lotus Scholars Program in Hunan Province (No. XJT2015408). M. Z. was supported by the Guangdong Basic and Applied Basic Research Foundation (Grant No. 2020A1515110971) and Jihua Laboratory (X201321XQ200).

**References**

1 N. J. Halas, S. Lal, W. S. Chang, S. Link and P. Nordlander, *Chem. Rev.*, 2011, **111**, 3913–3961.
2 S. M. Morton, D. W. Silverstein and L. Jensen, *Chem. Rev.*, 2011, **111**, 3962–3994.
3 A. Kristensen, J. K. W. Yang, S. I. Bozhevolnyi, S. Link, P. Nordlander, N. J. Halas and N. A. Mortensen, *Nat. Rev. Mater.*, 2016, **2**, 16088.
4 W. A. Murray and W. L. Barnes, *Adv. Mater.*, 2007, **19**, 3771–3782.
5 G. V. Hartland, *Chem. Rev.*, 2011, **111**, 3858–3887.
6 K. M. Mayer and J. H. Hafner, *Chem. Rev.*, 2011, **111**, 3828–3857.
7 M. Li, S. K. Cushing and N. Wu, *Analyt.,* 2015, **140**, 386–406.
8 H. Yu, Y. Peng, Y. Yang and Z.-Y. Li, *npj Comput. Mater.*, 2019, **45**, 1–14.
9 S. A. Maier, S. R. Andrews, L. Martin-Moreno and F. J. Garcia-Vidal, *Phys. Rev. Lett.*, 2006, **97**, 176805.
10 J. Christensen, A. Manjavacas, S. Thongrattanasiri, F. H. L. Koppens and F. J. García de Abajo, *ACS Nano*, 2012, **6**, 431–440.
11 R. Zhang, Y. Zhang, Z. C. Dong, S. Jiang, C. Zhang, L. G. Chen, L. Zhang, Y. Liao, J. Aizpurua, Y. Luo, J. L. Yang and J. G. Hou, *Nature*, 2013, **498**, 82–86.
12 S. A. Maier, *Plasmonics: Fundamentals and Applications*, Springer Science & Business Media, New York, 2007.
13 P. Drude, *Ann. Phys.*, 1900, **306**, 566–613.
14 K. A. Willets and R. P. Van Duyne, *Annu. Rev. Phys. Chem.*, 2007, **58**, 267–297.
15 X. P. Zhu, H. M. Shi, S. Zhang, Z. Q. Chen, M. J. Zheng, Y. S. Wang, S. W. Xue, J. Zhang and H. G. Duan, *Acta Phys. Sin.*, 2019, **68**(14), 147304.
16 G. Mie, *Ann. Phys.*, 1908, **330**, 377–445.
17 M. Meier and A. Wokaun, *Opt. Lett.*, 1983, **8**, 581–583.
18 R. Gans, *Ann. Phys.*, 1912, **342**, 881–900.
19 S.-W. Hla, L. Bartels, G. Meyer and K.-H. Rieder, *Phys. Rev. Lett.*, 2000, **85**, 2777–2780.
20 A. Bouhelier, M. Beverluis, A. Hartshuh and L. Novotny, *Phys. Rev. Lett.*, 2003, **90**, 013903.
21 H. G. Frey, S. Witt, K. Felderer and R. Guckenberger, *Phys. Rev. Lett.*, 2004, **93**, 208001.
22 J. M. Gerton, L. A. Wade, G. A. Lessard, Z. Ma and S. R. Quake, *Phys. Rev. Lett.*, 2004, **93**, 180801.
23 T. Ichimura, N. Hayazawa, M. Hashimoto, Y. Inouye and S. Kawata, *Phys. Rev. Lett.*, 2004, **92**, 220801.
24 M. I. Stockman, *Phys. Rev. Lett.*, 2004, **93**, 137404.
25 T. Schmid, L. Opilik, C. Blum and R. Zenobi, *Angew. Chem., Int. Ed.*, 2013, **52**, 5940–5954.
26 F. Lu, M. Jin and M. A. Belkin, *Nat. Photonics*, 2014, **8**, 307–312.
27 B. N. Tugchin, N. Janunts, A. E. Klein, M. Steinert, S. Fasold, S. Diziaîn, M. Sison, E.-B. Kley, A. Tünnemann and T. Pertsch, *ACS Photonics*, 2015, **2**, 1468–1475.
28 B. Xue, D. Wang, J. Zuo, X. Kong, Y. Zhang, X. Liu, L. Tu, Y. Chang, C. Li, F. Wu, Q. Zeng, H. Zhao, H. Zhao and H. Zhang, *Nanoscale*, 2015, **7**, 8048–8057.
29 L. Meng and Z. Yang, *Nanophotonics*, 2018, **7**, 1325–1332.
30 F. Walla, M. M. Wiecha, N. Mecklenbeck, S. Beldi, F. Keilmann, M. D. Thomson and H. G. Roskos, *Nanophotonics*, 2018, **7**, 269–276.
31 M. Yang, M. S. Mattei, C. R. Cherqui, X. Chen, R. P. Van Duyne and G. C. Schatz, *Nano Lett.*, 2019, **19**, 7309–7316.
32 H. Shi, S. Zhang, X. Zhu, Y. Liu, T. Wang, T. Jiang, G. Zhang and H. Duan, *ACS Appl. Mater. Interfaces*, 2017, **9**, 36907–36916.
33 H. Shi, X. Wang, M. Zheng, X. Wu and Y. Chen, *Adv. Mater. Interfaces*, 2016, **3**, 1600588.
Nanoscale Advances

Review

34 P. K. Jain, K. S. Lee, I. H. El-Sayed and M. A. El-Sayed, J. Phys. Chem. B, 2006, 110, 7238–7248.
35 M. Schenk, M. Kruger and P. Hommelhoff, Phys. Rev. Lett., 2010, 105, 257601.
36 Y. Lin, Y. Zou, Y. Mo, J. Guo and R. G. Lindquist, Sensors, 2010, 10, 9397–9406.
37 S. Kharintsev, A. Alekseev and J. Loos, Spectrochim. Acta, Part A, 2017, 171, 139–143.
38 M. Kruger, M. Schenk and P. Hommelhoff, Nature, 2011, 475, 78–81.
39 W. Zhang, B. S. Yeo, T. Schmid and R. Zenobi, J. Phys. Chem. C, 2007, 111, 1733–1738.
40 A. B. Zrimsek, N. Chiang, M. Mattei, S. Zaleski, M. O. McNally, C. T. Chapman, A. I. Henry, G. C. Schatz and R. P. Van Duyne, Chem. Rev., 2017, 117, 7583–7613.
41 J. Steidtner and B. Pettinger, Phys. Rev. Lett., 2008, 100, 236101.
42 C. C. Neacsu, G. A. Reider and M. B. Raschke, Phys. Rev. B: Condens. Matter Mater. Phys., 2005, 71, 204102.
43 E. J. Sánchez, L. Novotny and X. S. Xie, Phys. Rev. Lett., 1999, 82, 4014–4017.
44 E. M. van Schroevenstein Lantman, T. Deckert-Gaudig, A. J. Mank, V. Deckert and B. M. Weckhuyzen, Nat. Nanotechnol., 2012, 7, 583–586.
45 C. Ropers, C. C. Neacsu, T. Elsaesser, M. Albrecht, M. B. Raschke and C. Lienau, Nano Lett., 2007, 7, 2784–2788.
46 S. Berweger, J. M. Atkin, X. G. Xu, R. L. Olmon and M. B. Raschke, Nano Lett., 2011, 11, 4309–4313.
47 S. F. Becker, M. Essmann, K. Yoo, P. Gross, R. Vogelgesang, N. Park and C. Lienau, ACS Photonics, 2016, 3, 223–232.
48 E. Verhagen, L. Kuipers and A. Polman, Nano Lett., 2007, 7, 334–337.
49 N. C. Lindquist, P. Nagpal, A. Lesuffleur, D. J. Norris and S. H. Oh, Nano Lett., 2010, 10, 1369–1373.
50 F. De Angelis, G. Das, P. Candeloro, M. Patrini, M. Galli, A. Bek, M. Lazzerino, I. Maksymov, C. Liberale, L. C. Andreani and E. Di Fabrizio, Nat. Nanotechnol., 2010, 5, 67–72.
51 C. Ropers, C. C. Neacsu, M. B. Raschke, M. Albrecht, C. Lienau and T. Elsaesser, Jpn. J. Appl. Phys., 2008, 47, 6051–6054.
52 S. Berweger, J. M. Atkin, R. L. Olmon and M. B. Raschke, J. Phys. Chem. Lett., 2012, 3, 945–952.
53 S. Lee, J. Shin, Y.-H. Lee and J.-K. Park, ACS Nano, 2010, 4, 7175–7184.
54 T. Umakoshi, Y. Saito and P. Verma, Nanoscale, 2016, 8, 5634–5640.
55 R. Kato, K. Taguchi, R. Yadav, T. Umakoshi and P. Verma, Nanotechnology, 2020, 31, 335207.
56 X. Chen, N. C. Lindquist, D. J. Klemme, P. Nagpal, D. J. Norris and S. H. Oh, Nano Lett., 2016, 16, 7849–7856.
57 Y.-J. Lu, J. Kim, H.-Y. Chen, C. Wu, N. Dabidian, C. E. Sanders, C.-Y. Wang, M.-Y. Lu, B.-H. Li, X. Qiu, W.-H. Chang, L.-J. Chen, G. Shvets, C.-K. Shih and S. Gwo, Science, 2012, 337, 450.
58 L. Britnell, R. M. Ribeiro, A. Eckmann, R. Jalil, B. D. Belle, A. Mishchenko, Y. J. Kim, R. V. Gorbachev, T. Georgiou, S. V. Morozov, A. N. Grigorenko, A. K. Geim, C. Casiraghi, A. H. C. Neto and K. S. Novoselov, Science, 2013, 340, 1311.
59 T. Stefaniuk, P. Wrobel, E. Gorecka and T. Szoplik, Nanoscale Res. Lett., 2014, 9, 153.
60 U. Diebold, J.-M. Pan and T. E. Madey, Surf. Sci., 1995, 331–333, 845–854.
61 A. Kossyo, V. Mek, D. Simakov, K. Leossan, S. Kéna-Cohen and S. A. Maier, Adv. Opt. Mater., 2015, 3, 71–77.
62 J. H. Park, P. Ambwani, M. Manno, N. C. Lindquist, P. Nagpal, S. H. Oh, C. Leighton and D. J. Norris, Adv. Mater., 2012, 24, 3988–3992.
63 O. Pierre-Louis, M. R. D’Orsogna and T. L. Einstein, Phys. Rev. Lett., 1999, 82, 3661–3664.
64 P. Politi and J. Villain, Phys. Rev. B: Condens. Matter Mater. Phys., 1996, 54, 5114–5129.
65 S. A. Chambers, T. Droubay, D. R. Jennison and T. R. Mattsson, Science, 2002, 297, 827.
66 S. Conoci, S. Petralia, P. Samori, F. M. Raymo, S. Di Bella and S. Sortino, Adv. Funct. Mater., 2006, 16, 1425–1432.
67 A. Anders, E. Byron, D.-H. Kim, K. Fukuda and S. H. N. Lim, Solid State Commun., 2006, 140, 225–229.
68 D. Gu, C. Zhang, Y.-K. Wu and L. J. Guo, ACS Nano, 2014, 8, 10343–10351.
69 N. Formica, D. S. Ghosh, A. Carrilero, T. L. Chen, R. E. Simpson and V. Pruneri, ACS Appl. Mater. Interfaces, 2013, 5, 3048–3053.
70 M. Todeschini, A. Bastos da Silva Fanta, F. Jensen, J. B. Wagner and A. Han, ACS Appl. Mater. Interfaces, 2017, 9, 37374–37385.
71 D. I. Yakubovsky, Y. V. Stebunov, R. V. Kirtaev, G. A. Ermolaev, M. S. Mironov, S. M. Novikov, A. V. Arsenin and V. S. Volkov, Adv. Mater. Interfaces, 2019, 6, 1900196.
72 G. Zhao, W. Shen, E. Jeong, S. G. Lee, S. M. Yu, T. S. Bae, G. H. Lee, S. Z. Han, J. Tang, E. A. Choi and J. Yun, ACS Appl. Mater. Interfaces, 2018, 10, 27510–27520.
73 H. Qian, Y. Xiao and Z. Liu, Nat. Commun., 2016, 7, 13153.
74 Z. M. Abd El-Fattah, V. Mkhitaryan, J. Brede, L. Fernandez, C. Li, Q. Guo, A. Ghosh, A. R. Echarri, D. Naveh, F. Xia, J. E. Ortega and F. J. Garcia de Abajo, ACS Nano, 2019, 13, 7771–7779.
75 D. S. Ghosh, T. L. Chen and V. Pruneri, Appl. Phys. Lett., 2010, 96, 041109.
76 J. Yun, Adv. Funct. Mater., 2017, 27, 1606641.
77 R. Yu, P. Mazumder, N. F. Borrelli, A. Carrilero, D. S. Ghosh, R. A. Maniyara, D. Baker, F. J. Garcia de Abajo and V. Pruneri, ACS Photonics, 2016, 3, 1194–1201.
78 R. Yu, V. Pruneri and F. J. Garcia de Abajo, Sci. Rep., 2016, 6, 32144.
79 N. Luhmann, D. Hoj, M. Piller, H. Kahler, M. H. Chien, R. G. West, U. L. Andersen and S. Schmid, Nat. Commun., 2020, 11, 2161.
80 Y. Muniz, A. Manjavacas, C. Farina, D. A. R. Dalvit and W. J. M. Kort-Kamp, Phys. Rev. Lett., 2020, 125, 033601.
132 A. Neutens, P. Van Dorpe, I. De Vlaminck, L. Lagae and C. Ciraci, N. Yu, E. Cubukcu, L. Diehl, D. Bour, S. Corzine, J. Zhu, G. Höfler, K. B. Crozier and F. Capasso, Opt. Express, 2007, 15, 13272–13281.

133 N. Y. Cubukcu, L. Diehl, D. Bour, S. Corzine, J. Zhu, G. Höfler, K. B. Crozier and F. Capasso, Opt. Express, 2007, 15, 13272–13281.

134 X. Liang and S. Y. Chou, Nano Lett., 2008, 8, 1472–1476.

135 M. K. Kim, H. Sim, S. J. Yoon, S. H. Gong, C. W. Ahn, Y. H. Cho and Y. H. Lee, Nano Lett., 2015, 15, 4102–4107.

136 S. S. Acimović, M. P. Kreuzer, M. U. González and R. Quidant, ACS Nano, 2009, 3, 1231–1237.

137 Z. Fang, L. Fan, C. Lin, D. Zhang, A. J. Meixner and X. Zhu, Nano Lett., 2011, 11, 1676–1680.

138 G. C. Li, Y. L. Zhang, J. Jiang, Y. Luo and D. Y. Lei, ACS Nano, 2017, 11, 3067–3080.
178 A. Barik, X. Chen and S.-H. Oh, *Nano Lett.*, 2016, **16**, 6317–6324.

179 I. Kim, G. Yoon, J. Jang, P. Genevet, K. T. Nam and J. Rho, *ACS Photonics*, 2018, **5**, 3876–3895.

180 K. Huang, Z. Dong, S. Mei, L. Zhang, Y. Liu, H. Liu, H. Zhu, J. Teng, B. Luk’yanchuk, J. K. W. Yang and C.-W. Qiu, *Laser Photonics Rev.*, 2016, **10**, 500–509.

181 G. Hu, X. Hong, K. Wang, J. Wu, H.-X. Xu, W. Zhao, W. Liu, S. Zhang, F. Garcia-Vidal, B. Wang, P. Lu and C.-W. Qiu, *Nat. Photonics*, 2019, **13**, 467–472.

182 A. Schirato, M. Maiuri, A. Toma, S. Fugattini, R. Proietti Zaccaria, P. Laporta, P. Nordlander, G. Cerullo, A. Alabastri and G. Della Valle, *Nat. Photonics*, 2020, **14**, 723–727.

183 H. Sroor, Y.-W. Huang, B. Sephton, D. Naidoo, A. Vallés, V. Ginis, C.-W. Qiu, A. Ambrosio, F. Capasso and A. Forbes, *Nat. Photonics*, 2020, **14**, 498–503.

184 S. Divitt, W. Zhu, C. Zhang, H. J. Lezec and A. Agrawal, *Science*, 2019, **364**, 890.

185 M. Khorasaninejad, W. T. Chen, R. C. Devlin, J. Oh, A. Y. Zhu and F. Capasso, *Science*, 2016, **352**, 1190.

186 A. M. Shaltout, V. M. Shalaev and M. L. Brongersma, *Science*, 2019, **364**, eaat3110.

187 M. Khorasaninejad, A. Ambrosio, P. Kanhaiya and F. Capasso, *Sci. Adv.*, 2016, **2**, e1501258.

188 C. Genet and T. W. Ebbesen, *Nature*, 2007, **445**, 39–46.