Optical Coulomb blockade lifting in plasmonic nanoparticle dimers

Dmitry Sivun and Thomas A. Klar

Institute of Applied Physics, Johannes Kepler University Linz, 4040 Linz, Austria

Abstract: If two metal nanoparticles are ultimately approached, a tunneling current prevents both an infinite redshift of the bonding dipolar plasmon and an infinite increase of the electric field in the hot spot between the nanoparticles. We argue that a Coulomb blockade suppresses the tunneling current and sustains a redshift even for sub-nanometer approach up to moderate fields. Only for stronger optical fields, the Coulomb blockade is lifted and a charge transfer plasmon is formed. Numerical simulations show that such scenarios are well in reach with manageable nanoparticle dimensions, even at room temperature. Applications may include ultrafast, optically driven switches, photo detectors operating at 500 THz, or highly nonlinear devices.

© 2020 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Classical electromagnetism (CEM) predicts an infinite redshift of the hybridized noble metal nanoparticle plasmon polariton (the bonding dipolar plasmon BDP) as well as an infinite increase of the field strength in the hotspot between two metallic nanoparticles in the limit that the surface-to-surface distance approaches zero. About ten years ago, it was proposed that quantum mechanical tunneling of electrons prevents both divergences [1–3]. Experimental proof was provided soon after [4–9]. In the case where vacuum or air fills the gap between the two nanoparticles, a tunneling current sets in if they are approached closer than approximately 0.5 nm surface-to-surface distance [4,5]. When a dielectric, specifically when comprising organic molecules, fills the gap, tunneling sets in at somewhat larger distances [7,8], relaxing a bit the technical difficulty of controlled sub-nanometric approach. The tunneling current, on the one hand, reduces the build-up of charge at the opposing surfaces and hence weakens the electric field strength inside the hot spot, and on the other hand leads to a smooth transition from the BDP, that redshifts upon approach, to the so called charge transfer plasmon (CTP) that blueshifts upon further approach. While fully quantum mechanical models are in excellent agreement with experimental results, it has been found that a quantum corrected model (QCM) based on solving Maxwell’s equations with finite elements, and introducing tunneling as a specifically adapted conductivity, leads to satisfactory results [10,11]. Including nonlocal effects adds some minor corrections, but the optical response of pairs of plasmonic nanoparticles in close approach is essentially covered by quantum tunneling [11].

To date, tunneling of electrons in quantum plasmonics was considered as a continuous effect where the conductivity between two nanoparticles increases gradually with decreasing distance [12]. Instead of single electrons, electron densities were discussed [13–15]. In contrast, we consider in this study that the electrical charge is quantized. Surprisingly, this has been almost ignored so far in plasmonics, although it should lead to decisive signatures in the absorption, scattering, and extinction spectra, which all should show discontinuities at certain applied optical powers. Similar non-monotonic behavior is expected for surface enhanced Raman scattering, the generation of higher harmonics and many other nonlinear effects.

We herewith propose that the Coulomb blockade of the tunneling current should prevent the formation of a CTP and the system should stay in the CEM regime until a certain threshold potential is overcome in the gap between two plasmonic nanoparticles. The transition from BDP
to CTP has usually been investigated as a function of nanoparticle distance, at constant impinging field strength [1–9], or at a fixed distance as a function of the (continuous) electron density in the gap [13–15]. In this study, we discuss spectra at a fixed sub-nanometric distance, as a function of illumination intensity and under the consideration of the discreteness of the elementary charge. Importantly, our considerations are conceptually different from two studies where the optical spectra of a pair of plasmonic nanoparticles has been switched either by a stepwise application of a potential difference from outside [16] or by a stepwise charging of a plasmonic nanoparticle as part of a single electron tunneling device [17]. In both cases, potential or charge was applied externally. In contrast, we are considering an electrically completely isolated pair of plasmonic nanoparticles, addressed only optically. Hence, we consider the Coulomb blockade as an intrinsic effect of quantum plasmonics and we discuss how it could be lifted by the optical field itself rather than by applying a more or less DC-potential from outside.

2. Coulomb blockade in plasmonic dimers

2.1. A rough estimate

To start with, we briefly recapitulate the Coulomb blockade: It has been observed and explained a long time ago in electrical transport measurements in metallic granular systems [18–21], followed by Coulomb staircases [22] and single electron tunneling [23,24]. Applications range from single-electron transistors to single electron pumps, to current standards, or infrared detectors, just to name a few [25]. Importantly, the Coulomb blockade is meanwhile easily observable even at room temperature, as the handling of metallic nanoparticles in the sub-5 nm range became feasible. For instance, room temperature Coulomb blockade and staircases have been observed using scanning tunneling microscopy [26] or electrochemistry with gold nanoparticles in the electrolyte [27].

To explain the basic concept, we consider a pair of spherical gold nanoparticles, each with radius \( R \) and a surface-to-surface distance \( d \) as shown in the inset of Fig. 1(a). If \( d \approx R \), \( C = 4\pi \varepsilon_r \varepsilon_0 R \) [28,29] (blue line in Fig. 1(a)) can be considered as a good approximation for the analytical expression for the capacitance of two non-intersecting spheres (red line in Fig. 1(a)), which was derived by Maxwell (based on image charges) [28]:

\[
C = 2\pi \varepsilon_0 \varepsilon_r R \sum_{n=1}^{\infty} \frac{\sinh \left( \ln \left( D + \sqrt{D^2 - 1} \right) \right)}{\sinh \left( n \ln \left( D + \sqrt{D^2 - 1} \right) \right)}
\]

(1)

where \( \varepsilon_0 \) is the permittivity of free space, \( \varepsilon_r \) is the relative permittivity of the medium in the gap, \( R \) is the radius of the spheres, and \( D = 1 + (d/2R) \) where \( d \) is the distance between the closest points on the surface of the spheres.

For example, \( R = 2.5 \) nm results in \( C = 0.278 \) aF using the linear approximation, or \( C = 0.358 \) aF using (1). The potential energy required to transfer a single electron from one to the other nanoparticle by tunneling is then given by the Coulomb blockade energy \( U_C = e^2/(2C) \) [20], where \( e \) is the elementary charge. Inserting \( C = 0.358 \) aF, we obtain \( U_C = 224 \) meV, which is roughly one order of magnitude above thermal noise at room temperature and hence should be easily in reach of experimental observation [27]. As a consequence, quantum tunneling should be suppressed and a CTP should not form.

Coulomb blockade, however, breaks down if the potential between the two nanoparticles surpasses \( U_C \). In contrast to previous considerations [16,17], we do not assume a potential or charge from outside, but we now consider that the optical power provided by an impinging light, driving the BDP, lifts the Coulomb blockade. Estimating that the field inside the hotspot is plasmonically enhanced compared to the incoming field by a factor \( f \approx 75 \) (at \( d = 0.2 \) nm,
obtained by numerical simulation), we end up with an intensity of

\[ I_C = \frac{|E|^2}{2Z_0} = \frac{(U_C/fe d)^2}{2Z_0} = 30 \text{ MW/cm}^2 \]  

(2)

Herein, \( E \) denotes the incident electric field and \( Z_0 = 376.73 \Omega \) is the vacuum impedance. Assuming the focal area, where the extinction cross-section of the dimer is measured in a microscope, being 1 \( \mu \text{m}^2 \) [30,31], the illumination power threshold necessary to lift the Coulomb blockade is only 0.30 W. If one applies less than this power, Coulomb blockade prevents tunneling and hence the CTP cannot form. Instead, the more intense and redshifted (compared to that of a single nanoparticle) extinction spectrum of the BDP will be observed. Once the applied power is increased beyond 0.30 W, the Coulomb blockade is lifted, the CTP forms, and the absorption spectrum blueshifts and weakens compared to the spectrum of the BDP. At a temperature of 0 K, the switch between the BDP and the CTP spectra will be sudden as soon as 0.30 W is surpassed. At room temperature \( T_R \), the change will smear out, but should still be very well observable because \( U_C \approx 9 \cdot k_B T_R \) at this specific example (\( k_B \) being Boltzmann’s constant).

Besides heterodyne absorption spectroscopy, which allows for detecting plasmonic nanoparticles down to diameters of 1.4 nm [31], a variety of other methods exists to optically detect gold nanoparticles of sub-ten nanometer size, for instance using heterodyne interference [32] or position modulation [33,34]. Hence, we conclude that a Coulomb blockade and its lifting by increasing the illumination power should be readily observable with well-established setups, even at room temperature.

2.2. Numerical simulations

When \( R \) is much larger than \( d \) (\( R >> d \)), one has to use either the analytical expression (1), which often needs hundreds of sum terms to converge, or a numerical solver. A numerical solver also allows using arbitrary shapes of particles. Figure 1(b) shows a comparison of the full analytical solution (1) \( (n_{\max}=1000) \) with the numerical simulation for the dimer with fixed radii \( R = 2.5 \) nm and as a function of the distance between the spheres. All numerical simulations are carried out using the finite element (FEM) solver COMSOL Multiphysics under the assumption of electrostatic conditions. The entire surface of the first sphere is grounded and the second sphere is under constant potential. The spheres were placed in vacuum \( (\varepsilon_r = 1) \) in these simulations.

The numerically calculated capacitance for a pair of gold nanoparticles (Fig. 2(a)) is shown in Fig. 2(b) whereby the (equal) radius of both is varied from 0.5 to 2.5 nm and the distance is varied

Fig. 1. (a) Capacitance of a dimer as a function of the spheres’ radii \( R \) with a fixed interparticle distance \( d = 0.2 \) nm. The linear approximation \( C = 4\pi \varepsilon_0 R \) (blue line, \( \varepsilon_r = 1 \)) is well valid for radii up to 1.0 nm and starts to distinctly deviate from Maxwell’s formula (Eq. (1), red line) for radii \( R > 1.5 \) nm. (b) Comparison of the capacitance calculated using Maxwell’s Eq. (1) (red solid line) or COMSOL simulations (blue dots) for a dimer, consisting of spheres with equal radii \( R = 2.5 \) nm as a function of the interparticle distance.
from contact (0 nm) to 1.0 nm. For distances above 0.1 nm, the capacitance indeed depends only weakly on distance as argued above, but strongly on the radius. The numerically determined Coulomb charging energies are displayed in Fig. 2(c). At a distance of 0.2 nm, Coulomb blockade energies of 9 or 40 $k_B T_R$ are reached for diameters of 5 or 1.4 nm (Au$_{55}$), respectively.

Fig. 2. (a) Sketch of a dimer of Au nanospheres. (b) Capacitance of the dimer as a function of $R$ and $d$. (c) Energy of the Coulomb blockade (transfer of a single electron), normalized to room temperature thermal energy. (d) Extinction spectra of a Au$_{55}$ ($R = 0.7$ nm) dimer with separation $d = 0.2$ nm. Black solid line: CEM (potential between the dimers is not sufficient to exceed the Coulomb blockade). Red dash-dotted line: QCM (Coulomb blockade is lifted). (e) The total extinction power as a function of the illumination intensity. Assuming $T = 0$ K, a sudden drop of the extinction is observed and the slope is reduced afterwards. Calculations carried out at 524 nm (dotted line in (d)).

2.3. Plasmon quantization

So far, we have considered the quantization of the electron, but not yet the quantization of the plasmon itself. It turns out that this might be a pothole on the way to observing the optical Coulomb blockade in a dimer of spherical metallic nanoparticles, unless care is taken to find a suitable geometry. The quantization of a bosonic plasmon-polariton of angular frequency $\omega_{Pl}$ tells that the minimal energy stored on a dimer must be $\hbar \cdot \omega_{Pl}$. Assuming that about 5% of this energy is stored in the gap between the nanoparticles and most of it is stored inside the metal (see Appendix A), and assuming that the volume of the hotspot is well estimated by $V = \pi R d^2$ [4,35], further setting $R = 2.5$ nm, $d = 0.2$ nm and $\hbar \cdot \omega_{Pl} = 2.38$ eV, we end up with an electric field of $E = \sqrt{0.1 \cdot \hbar \omega_{Pl} / (\epsilon_0 V)} \approx 3.7 \cdot 10^9$ V/m. This turns into a potential energy drop between the two surfaces of about $U_{Pl} = 0.74$ eV. A more elaborate numerical evaluation of $U_{Pl}$ yields $U_{Pl} = 0.62$ eV. Hereby in the simulations, we first adjusted the amplitude of the incident electric field in such a way, that the totally stored energy is equal to the energy of a single quantum of plasmonic oscillations $\hbar \omega_{Pl}$. Then the integral over the absolute value of the electric field along the shortest distance between the two apices was taken. The potential difference between the two spheres $U_{Pl}$, induced already by a single quantum of a plasmonic oscillation, exceeds the Coulomb blockade $U_C$ and hence, the latter cannot be observed for this initial set of dimensions of a nanoparticle dimer.
If however, one chooses $R = 0.7 \text{nm}$ (Au$_{55}$) and $d = 0.2 \text{nm}$, the numerical results are $U_C = 1027 \text{meV}$ and $U_{pl} = 816 \text{meV}$. In the case where the incoming intensity is adjusted such that there is not more than one plasmonic excitation on the Au$_{55}$ dimer at a time, the potential between the two spheres does not reach the Coulomb blockade, not even with the help of thermal noise, since $U_C - U_{pl} = 8k_BT_R$. For more intense excitation, the potential exceeds the Coulomb blockade. In the first case, the dimer must be simulated in the CEM model, in the latter case, the QCM needs to be applied. Details of our implementation of the QCM can be found in the Appendix B. Figure 2(d) shows the extinction cross-section $\sigma_{ext}$ (for such small nanoparticles dominated by absorption). Setting $T = 0 \text{K}$, a sharp transition is observed between the case where two electrically isolated spheres form a BDP (black solid line in Fig. 2(d)) and the case where two spheres are connected by a tunneling contact, forming a CTP (red dash-dotted line in Fig. 2(d)).

The Coulomb blockade is lifted when the Au$_{55}$ dimer is illuminated with $19.1 \text{GW/cm}^2$. If one measures the total extinction power $P_{ext} = I_{inc} \cdot \sigma_{ext}$ as a function of the applied intensity $I_{inc}$ (Fig. 2(e)), one observes a sudden drop in extinction at this intensity. Further, the slope is less above the lifting of the Coulomb blockade. We note, that $19.1 \text{GW/cm}^2$ is tolerable for Au$_{55}$ dimers if this intensity is applied in ultrashort pulses such that the thermal budget stays small. Further down, we will also comment on the possibility of a “thermal” breakdown of the Coulomb blockade.

**Fig. 3.** (a) Sketch of a nanosphere in proximity to a flat nanorod with dimensions $W = 15 \text{nm}$, $L = 40 \text{nm}$, and $t = 2R$. (b) Capacitance of the sphere/rod dimer as a function of $R$ and $d$. (c) Energy of the Coulomb blockade (transfer of a single electron), normalized to room temperature thermal energy. (d) Extinction spectra of the sphere/rod dimer with a separation $d = 0.2 \text{nm}$ and $R = 1 \text{nm}$. Black solid line: CEM (potential between the dimers is not sufficient to exceed the Coulomb blockade). Red dash-dotted line: QCM (Coulomb blockade is lifted). (e) The total extinction power as a function of the illumination intensity at $905 \text{nm}$ (dashed line in (d)). Assuming $T = 0 \text{K}$, a sudden drop of the extinction is observed and the slope is reduced afterwards.
2.4. A realistic example

As a second example, we want to introduce a more practicable sample, namely dimers consisting of a sphere of variable radius \( R \), and a flat rod, with thickness \( t = 2R \), a width \( W = 15 \text{ nm} \) and a length \( L = 40 \text{ nm} \) (Fig. 3(a)). The basic idea is that the small spherical nanoparticle provides the small capacitance, but the rod acts as an antenna. Figure 3(b) shows the capacitance as a function of the sphere radius \( R \) and the distance \( d \) between the sphere and the rod. Figure 3(c) shows the charging energy as a function of \( R \) and \( d \). Coulomb energies of \( 20 k_B T \) can be reached easily, rendering also this dimer capable of showing Coulomb blockade at room temperature.

We now fix \( R = 1 \text{ nm} \) and \( d = 0.2 \text{ nm} \), and simulate the extinction slightly below the Coulomb blockade (CEM calculation, black solid line in Fig. 3(d) and slightly above (QCM, red dash-dotted line in Fig. 3(d)). Plotting the extinction versus the illumination intensity (Fig. 3(e)) reveals that the Coulomb blockade is lifted at \( 4.21 \text{ MW/cm}^2 \). This is four orders of magnitude less than in case of the \( \text{Au}_{55} \) dimer because the rod acts as an effective antenna. Noteworthy, in the sphere/rod example, \( U_C = 412 \text{ meV} \gg k_B T \) and \( U_{Pl} = 155 \text{ mV} < U_C \). Hence, a single quantum of a plasmon will not reach the Coulomb blockade, not even with the help of thermal noise.

3. Critical discussion of the concept

When introducing a new concept, a critical discussion is mandatory. The first question that might arise is whether sub-nm spacings are too difficult to achieve and might even be closed by moving atoms [36]. The latter would also lead to discontinuities in the optical spectra which could easily be mistaken to be Coulomb blockade effects. In fact, except for the very first ones [4,5], experiments in tunneling quantum plasmonics are usually carried out in systems where the sub-nm spacings are defined by a dielectric substance [37], in particular by organic molecules [7–9]. This leads to substantially enlarged tunneling distances up to 1 nm and beyond. If organic molecules are used as spacers, resonant tunneling should be taken into account [38,39]. It would be interesting (but far beyond the scope of this initial discussion) to consider Coulomb blockade with resonant tunneling and the possibility that the optical field driving the plasmons leads to the equilibration of the Fermi level of one of the plasmonic nanoparticles with one of the molecular levels of the spacer.

A further concern could be that the intensity reaches the GW/cm\(^2\) range at \( U_C \) and at sub-nanometric distances, and that dielectric breakdown might hence be an issue. If this was true, then basically all literature on quantum tunneling plasmonics should be revised. As experimental evidence of the formation of the CTP is out of discussion [4–9], we argue that dielectric breakdown does not cause problems at plasmonic timescales. Nevertheless, high electric field strengths lead to field emission (Fowler-Nordheim tunneling) [13–15,40], which leads to nonlinearities in quantum plasmonics. It is worth noting that such nonlinearities are expected to occur at excitation powers of \( 1–10 \text{ GW/cm}^2 \) [13,40], which is reached for the \( \text{Au}_{55} \) dimer, but not in case of the sphere/rod dimer. It was further pointed out that space charge needs to be considered at separations of 0.2 nm [13–15], however, we note that in the currently considered situation there is either no tunneling electron or at most one electron. The question, whether a single electron, while tunneling inside the gap, may give rise to Coulomb-blocking the conductivity of the gap [41], and hence preventing the formation of a CTP, is interesting, but totally beyond the scope of this initial discussion.

Although rarely discussed in tunneling plasmonics [12], time constants are of particular importance when discussing Coulomb blockade effects. Field induced tunneling times are on the order of \( T_t \approx 0.5 \text{ fs} \) [42]. In order to yield an effective tunneling contact, the half-cycle \( T_{1/2} \) of a plasmon resonance should exceed \( T_t \). While this is easily fulfilled in the infrared, CTPs at visible frequencies have \( T_{1/2} \approx 1 \text{ fs} \), just twice the tunneling time. Further, the spectral width of a CTP implies a plasmon lifetime on the order of 4 fs [43], meaning that within its lifetime, the CTP undergoes only four half-cycles. Further, the short CTP lifetime implies an energetic uncertainty
of roughly $\Delta U = 90$ meV. $U_C$ should exceed this value, say, threefold. It actually turns out that this requirement is more strict than the requirement $U_C > k_B T_R$. However, it is achieved for the Au$_{55}$ dimer and the sphere/rod dimer discussed above. Using the time-energy uncertainty $U_C \tau \geq \hbar/2$ with $\tau = R_\text{el} C$, one ends up with a gap resistance $R_\text{el} \geq 4.1$ kΩ. In QCM, $R_\text{el}$ depends on the surface-to-surface distance. One finds that $d \geq 0.15$ nm is required, which is the case for all examples discussed above. This argument is, by the way, what is usually termed the “third condition” of the orthodox theory of Coulomb blockade [25]. The first condition states that size-quantization effects do not play a role (which is a good assumption for gold nanoparticles from $2R = 1.4$ nm (Au$_{55}$) onwards) and the second condition is that $T_t$ is much smaller than all other time scales, a condition which is just about to be fulfilled, because $T_t \approx 0.5 \cdot T_1^{1/2}$.

Next, we would like to discuss how hot electrons of temperature $T_e$ might affect the Coulomb blockade. (Heating of the lattice is 100 fold less than heating of the electron sea and can hence be ignored). Clearly, the Coulomb blockade breaks down as soon as $U_C \approx k_B T_e$ and one would simply see a thermal lifting of the Coulomb blockade (though optically induced) rather than a genuine optical lifting, in which the plasmonically enhanced field lifts the blockade directly. As long as there is only one quantum of plasmon oscillation excited at a time on an otherwise “cold” (i.e. $T_R$) dimer, no problems should occur, because only after some femtoseconds, the plasmons decay into single electron and hole excitations, which thermalize within some 100 fs. In the case of the Au$_{55}$ dimer, a single plasmon quantum will heat $T_e$ by roughly 1000 K, while in case of the sphere/rod dimer (Fig. 3(a)) $T_e$ heats up by only 15 K due to the 140 times larger volume. In both cases, the Coulomb blockade is optically lifted if two (Au$_{55}$) or roughly three (sphere/rod) plasmons are excited at a time. Hence, in case of the Au$_{55}$, the second plasmon might be absorbed by an electronically hot Au$_{55}$ dimer and it is impossible to distinguish thermal from optical lifting of the Coulomb blockade. In contrast, a $\sim 15$ K increase of $T_e$ will not lead to a thermal lifting of the Coulomb blockade, a further advantage of the sphere/rod system.

4. Conclusions

In summary, we have taken the quantization of the electronic charge and of the plasmon seriously and argue that the Coulomb blockade should prevent the establishment of a charge transfer plasmon as long as the applied light intensity does not lead to potential differences between the surfaces of a plasmonic dimer that lifts this blockade. We propose a dimer comprising a nanosphere in intimate neighborhood to a gold rod as a facile test object where Coulomb blockade shall be observable at room temperature using state of the art experimental setups. Hot electron effects do not thermally lift the Coulomb blockade in this arrangement. The rod could be, for instance, milled out of a single crystal gold flake using an ion beam [44] and a small gold nanoparticle could be approached via AFM nanomanipulation [45,46]. A natural next step is time domain density functional simulations that include exchange-correlation potentials and respect charge quantization. Applications may include ultrafast optical switches operating at IR or visible light frequencies (500 THz) or highly nonlinear optical devices such as high frequency generators; further, plasmonic biosensors based on quantum concepts, or current standards directly linked to optical frequencies. All-optical Coulomb staircases should also be observable.

Appendix A: Energy density calculations

In case of a dispersive medium, the average energy density assumes the form [47]:

$$\bar{W} = \frac{1}{4} \left( \frac{d(\omega \epsilon_\text{e} \epsilon_0)}{d\omega} |E|^2 + \frac{d(\omega \mu_\text{m} \mu_0)}{d\omega} |H|^2 \right)$$

(3)

However, this equation is only applicable when the imaginary parts of $\epsilon(\omega)$ and $\mu(\omega)$ are very small in comparison with their real parts. Evaluation of this formula for plasmonic nanostructures
near resonance frequencies results in negative total energy densities of the system, which is unphysical. The origin of such an anomaly is the energy density inside the metal, where \( \varepsilon(\omega) = \varepsilon'(\omega) + i\varepsilon''(\omega) \) is not only dispersive but its real part is negative and the absolute values of the real and imaginary parts are on the same order. The correct expression for the energy density in highly absorbing materials was first derived by Loudon [48] and later generalized to a highly dispersive and absorbing material by Ruppin [49]:

\[
\bar{W} = \frac{\varepsilon_0}{4} \left( \varepsilon' + \frac{2\omega\varepsilon''}{\Gamma_E} \right) |E|^2 + \frac{\mu_0}{4} \left( \mu' + \frac{2\omega\mu''}{\Gamma_H} \right) |H|^2
\]

(4)

where \( \Gamma_E \) and \( \Gamma_H \) are damping frequencies.

In case of no magnetic dispersion \( (\mu' = 1, \mu'' = 0) \), Eq. (4) reduces to:

\[
\bar{W} = \frac{\varepsilon_0}{4} \left( \varepsilon' + \frac{2\omega\varepsilon''}{\Gamma_E} \right) |E|^2 + \frac{\mu_0}{4} |H|^2
\]

(5)

Using expression (5) inside and outside the metal with the respective dielectric constants, one can estimate the distribution of energy in the metal and in the dielectric around. Figure 4 shows how much of the stored energy is inside the NPs and how much is stored in the dielectric (essentially inside the hot spot). Data are shown in Figs. 4(a)-(c) for the three geometries discussed in the main paper.

**Fig. 4.** Black lines: total energy of the plasmonic dimers; red lines: energy inside the metal and blue lines: energy in the hot spot between the nanoparticles. Three arrangements are calculated: (a) Dimer consisting of two equal spheres with radii 0.7 nm and distance \( d = 0.2 \) nm. (b) Dimer consisting of two equal spheres with radii 2.5 nm and \( d = 0.2 \) nm. (c) Dimer consisting of a sphere with radius 1 nm and a flat nanorod (dimensions as given in Fig. 3(a)).

**Appendix B: Implementation of the quantum corrected model (QCM) in the numerical simulations**

The main idea behind the QCM is to locally replace the permittivity in the gap between two metallic NPs with an effective permittivity which will allow charge transfer across the gap [10]. This effective permittivity mimics the tunnelling conductivity across the gap and depends exponentially on the gap size. It can be easily implemented in any classical Maxwell’s equation solver to get a quantum corrected optical and electrical response of a given nanostructure.

In order to cover all separation distances, an effective permittivity is required with a smooth transition from the dielectric properties of the gap medium at large separation distances, to a fully metallic behavior at the touching regime. We used the following expression for the permittivity
of an effective medium in the gap [11],

\[ \varepsilon_{\text{gap}}(\omega, d) = \varepsilon_{\text{medium}}(\omega) + \left( \varepsilon_{\text{JC}}(\omega) + \frac{\omega_p^2}{\omega^2 + i\omega\gamma_p} - \varepsilon_{\text{medium}}(\omega) \right) \exp(-d/l_d) - \frac{\omega_p^2}{\omega^2 + i\omega\gamma_p \cdot \exp(d/l_c)} \]  

(6)

where \( \varepsilon_{\text{medium}}(\omega) \) is the relative permittivity of the surrounding medium, \( \varepsilon_{\text{JC}}(\omega) \) is the empirical data for the dielectric function of gold [50], \( \omega_p \) and \( \gamma_p \) are the plasma and the damping frequencies of the conduction electrons in the Drude model, \( d \) is the distance between the metallic surfaces, \( l_c \) is the characteristic length where the tunneling of the conduction band electrons is switched on with decreasing \( d \) and \( l_d \) is the characteristic length where the dielectric constant of the medium is changed to the background dielectric constant of the metal (which is \( \varepsilon_{\text{JC}} \) without the Drude contribution). For a gap filled with air, the characteristic length \( l_c \) is taken to be 0.04 nm [11]. The exact value of \( l_d \) can be set equal to the radial decay of the 5d orbitals, which is 0.079 nm for gold [11].

It is easily seen that this expression (6) converges to the medium dielectric constant for large \( d \) and it converges to the metal dielectric constant for very small \( d \). Since the gap size \( d \) is changing in radial direction (c.f. Figure 5), the gap permittivity \( \varepsilon_{\text{gap}}(\omega, d(r)) \) should be inhomogeneous in the radial direction. In the original publication of Esteban et. al. [10], the authors describe the gap region by a finite number of coaxial cylinder barrels with a respective homogeneous permittivity \( \varepsilon_{\text{gap}}^i(\omega) = \varepsilon_{\text{gap}}^i(\omega, d_i) \) within each of the \( i \) cylinders (see Fig. 5(a)). In this work, instead, we used a smooth variation of the permittivity (Fig. 5(b)) as described by the analytical notation for the distance between two equal spheres:

\[ d(r) = d + 2 \cdot (r_0 - r_0 \cos(\alpha(r))) \]  

(7)

where \( d \) is the closest possible distance between the two spheres (at the axis of the dimer), \( \alpha = \arcsin(r/r_0) \), and \( r \) is the radial distance in cylindrical coordinates with respect to the symmetry axis (c.f. Figure 5(c)). Finally, all numerical simulations of the optical response of the dimers where carried out in frequency domain using COMSOL Multiphysics with the Wave Optics package. In both CEM and QCM models, the dielectric function of gold was taken from Johnson and Christy [50]. Table 1 shows the parameters for the quantum corrected gap permittivity model.

**Fig. 5.** Real part of relative permittivity \( \varepsilon_{\text{gap}}(\omega, d) \) in the gap region in case of a gap distance \( d = 0.1 \) nm at 520 nm wavelength. (a) Model built with 6 coaxial cylinders of discrete radii \( r \), as used by Esteban et.al. [10]. (b) Analytically described permittivity with a continuous dependence on \( r \), as used in this work. (c) Sketch of the metallic dimer structure with an artificial gap material in the middle.
Table 1. Parameters for the quantum corrected gap permittivity model.

<table>
<thead>
<tr>
<th>$l_c$</th>
<th>$l_d$</th>
<th>$\omega_p$</th>
<th>$\gamma_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.04 nm</td>
<td>0.079 nm</td>
<td>9.06 eV</td>
<td>70.76 meV</td>
</tr>
</tbody>
</table>

Appendix C: Thermal smearing

At zero temperature, a rapid jump from the classical to the quantum corrected extinction cross-section should occur as soon as the potential between the two particles surpasses $U_C$. However, for temperatures larger than zero, the jump should be smeared out. We describe such thermal smearing by a modified logistic function:

$$\sigma_{\text{ext}}(U_{\text{inc}}) = \sigma_{\text{ext}}^{\text{QCM}} + \frac{\sigma_{\text{ext}}^{\text{CEM}} - \sigma_{\text{ext}}^{\text{QCM}}}{1 + \exp \left[ \frac{U - U_C}{k_B T} \right]}$$  \hspace{1cm} (8)

where $\sigma_{\text{ext}}^{\text{CEM}}$ and $\sigma_{\text{ext}}^{\text{QCM}}$ are the extinction cross-sections for CEM and QCM regimes, respectively. $U$ is the applied potential between the NPs, $U_C$ is the Coulomb blockade energy, $k_B$ is Boltzmann’s constant, and $T$ is the temperature.

Figure 6 shows the extinction cross-section as a function of the potential $U$ between the two nanoparticles for temperatures of 0 and 300 K for the dimer of two spherical nanoparticles with radii $R = 2.5$ nm and the distance between them $d = 0.2$ nm. $U_C = 224$ meV was obtained from COMSOL simulations (see Fig. 2(c)).

![Fig. 6. Extinction cross-section as a function of the potential between the two nanoparticles with radii $R = 2.5$ nm and the distance between them $d = 0.2$ nm. Black solid and blue dashed lines correspond to temperatures of 0 and 300 K, respectively.](image)

Acknowledgments

The authors acknowledge fruitful discussions with Javier Aizpurua and Ruben Esteban about QCM and Vitaly Golovach about single electron transistors and Coulomb staircases during TAK’s sabbatical at Donostia International Physics Center (DIPC) and discussions with Calin Hrelescu. TAK thanks DIPC for a travel allowance. Article processing charges have been provided by Johannes Kepler Open Access Publishing Fund.

Disclosures

The authors declare no conflicts of interest.
References


