Third-order nonlinear plasmonic materials: Enhancement and limitations

J. B. Khurgin
Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, Maryland 21218, USA

G. Sun
Department of Physics, University of Massachusetts Boston, Boston, Massachusetts 02125, USA

(Received 1 July 2013; published 25 November 2013)

We develop a rigorous and physically transparent theory of enhancement of third-order nonlinear optical processes achievable in plasmonic structures. The results show that the effective nonlinear index can be enhanced by many orders of magnitude, but, due to high metal losses the most relevant figure of merit, the amount of phase shift per one absorption length, remains very low. This makes nonlinear plasmonic materials a poor match for applications requiring high efficiency, such as all-optical switching and wavelength conversion, but does not preclude the applications where overall high efficiency is not required, such as sensing.

DOI: 10.1103/PhysRevA.88.053838 PACS number(s): 42.70.Nq, 42.65.Pc, 78.67.Pt, 42.65.Ky

I. INTRODUCTION

Various nonlinear optical phenomena have been attracting the interest of the scientific community ever since scientists gained access to intense optical fields with the invention of the laser in 1960 [1]. Very shortly after this invention practically all the major nonlinear optical phenomena of second and third order were successfully demonstrated [2–4]. Simultaneously the theory of nonlinear optics was developed by Bloembergen and many others [5–7]. Today a clear understanding of the nonlinear optical effects in various media exists and can be found in a number of excellent textbooks and monographs [8,9]. The fascinating promise of nonlinear optics has always been based on the fact that nonlinear optical phenomena allow one in principle to manipulate photons with other photons without relying on electronics. Hence a large number of all-optical devices that allow light manipulation based on either second- or third-order nonlinear effects, such as frequency conversion, switching, phase conjugation, and others, have been proposed and demonstrated in different materials and configurations [9]. However, while there have been some spectacular success stories that lead to practical products (such as frequency converters, optical parametric oscillators, frequency combs for measurements, and a few others), thus far the majority of nonlinear optical phenomena have not become competitive for practical applications, and not for the lack of trying.

The reason for this seeming incongruity is quite simple—all nonlinear optical phenomena can be divided into two broad classes: slow and ultrafast. The slow nonlinear phenomena are generally classified as such by the fact that optical fields do not interact directly, but through the various "intermediaries," such as electrons excited when the photons get absorbed, or through the temperature rise caused by the release of energy of the absorbed photons. For as long as these intermediaries exist, i.e., while electrons stay in the excited state or until the heat dissipates, their effect on the optical fields accumulates, hence these phenomena, such as saturable absorption, photorefractive effect, or thermal nonlinearity, can be quite strong, but this very fact makes them slow, as their temporal response is limited by the time constant associated with relaxation, recombination, or heat diffusion process. Furthermore, the slow nonlinearity always involves the so-called "real" process of photon absorption, and, once absorbed, these photons are never recovered, which means that "slow" nonlinearities are always associated with a significant loss. While there exists a legitimate niche for these slow nonlinearities (which may not be all that slow after all, as some saturable absorbers do show picosecond response), it is the other hand, the so-called virtual or ultrafast nonlinearities that have been the object of interest as they carry the promise of transforming the fields of information processing and communications.

The term "virtual," that is commonly associated with the ultrafast nonlinearity, implies that the nonlinear phenomenon does not involve excitation of the matter to the real excited states as there exist no transitions between the states that are resonant with the photon energy. When the non-energy-conserving virtual excitation does take place its duration is determined by the uncertainty principle, and thus can be as short as a few femtoseconds or even a fraction of a femtosecond which explicates the term "ultrafast." But it is precisely the fact that the excitation lasts such a short time interval that makes the ultrafast nonlinearities relatively weak. On a microscopic level one can explain this by the fact that for small electric fields the atoms and molecules always behave as essentially harmonic oscillators and only when the applied fields become appreciable relative to the intrinsic field E_i holding the electrons confined within the atom or the bond between the atoms does anharmonicity arise causing the nonlinear response. Overall, it can be loosely stated that the order of magnitude of the nonlinearity of the nth order can be determined as

$$\chi^{(n)} \sim \chi^{(1)} E_i^{n-1}, \tag{1}$$

where $\chi^{(1)} = n_d^2 - 1$ is the linear susceptibility and n_d is the index of refraction, typically anywhere between 1.5 and 3.5 for most solids in the visible to near-IR range. The typical value of the intrinsic field is on the order of the binding energy of a few eV divided by the bond length of 1–2 Å, i.e., 10^{10}–10^{11} V/m. Therefore, the second-order susceptibility $\chi^{(2)}$ in solids cannot exceed 1000 pm/V and is usually far less than that due to crystal symmetry, while

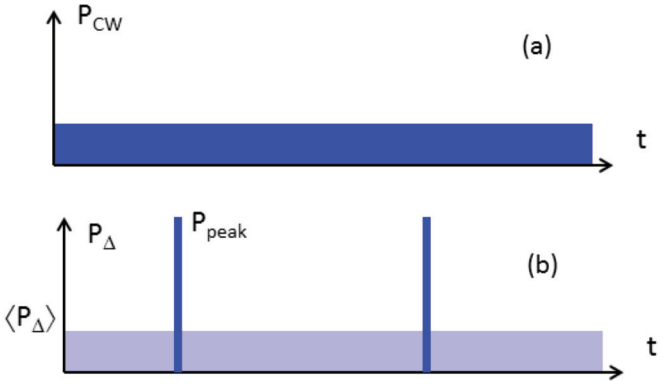

FIG. 1. (Color online) Comparison of cw (a) and mode-locked (b) laser outputs with equal average power.

the third-order susceptibility $\chi^{(3)}$ is expected to be less than 10^{-19} m^2/V^2. Through the relation between the third-order susceptibility and nonlinear refractive index, $n_2 = \chi^{(3)}\eta_0/n^2$ where the impedance of free space $\eta_0 = 377\ \Omega$, one can see that the latter is limited to a magnitude less than 10^{-13} cm^2/W. This is indeed the case, with a typical nonlinear index being strongly dependent on the width of the transparency region and ranging from $n_2 \sim 5 \times 10^{-16}$ cm^2/W for fused silica that is transparent all the way to UV, to perhaps $n_2 \sim 1 \times 10^{-13}$ cm^2/W for chalcogenide glasses transparent only in the IR range [10].

Clearly, very strong optical power density is required in order to produce appreciable ultrafast nonlinear optical phenomena. The average optical power available from a compact laser rarely exceeds a few hundred milliwatts; furthermore, if one wants to envision all-optical integrated circuits, the power dissipation requirements constrain the power to even much lower levels than that, possibly less than a milliwatt. Hence early on it was understood that to make nonlinear optical phenomena practical one must concentrate the power in both space and time. Concentration in space usually implies coupling the light into a tightly confining optical waveguide or a fiber. But the attainable concentration is limited to roughly a wavelength in the medium due to the diffraction limit. In addition, one may consider the resonant concentration of optical energy in microcavities [11,12], ring resonators [13], photonic band-gap structures [14], and slow-light devices [15,16], but all these resonant effects inevitably limit the bandwidth [17]. It is the concentration of optical power in the time domain provided by pulsed sources, particularly by the Q-switched [18] and mode-locked lasers [19,20], that has proven to be the winning technique in nonlinear optics.

The reasons for this can be easily grasped from the sketch in Fig. 1. Consider the light of a cw source with the power $P_{\rm cw}$ that propagates over a distance l in the medium with nonlinear susceptibility $\chi^{(n)}$. At the far end the nonlinear wave of power $P_{\rm cw}^{(n)} \propto l^2|\chi^{(n)}|^2 P_{\rm cw}^n$ will emerge [Fig. 1(a)]. If, on the other hand, one could use a periodically pulsed source with duty cycle Δ and the same average power $\langle P_\Delta\rangle = P_{\rm cw}$ [Fig. 1(b)], the peak power would obviously be $P_{\rm peak} = \Delta^{-1}P_{\rm cw}$ and the peak output nonlinear power would increase to $P_{\rm peak}^{(n)} \sim l^2|\chi^{(n)}|^2\Delta^{-n}P_{\rm cw}^n$ while the average nonlinear output would amount to $\langle P_\Delta^{(n)}\rangle \sim \Delta^{1-n}P_{\rm cw}^{(n)}$; i.e., the efficiency is increased by Δ^{1-n}. For a typical mode-locked laser with a picosecond-pulse duration and 100-MHz repetition rate $\Delta = 10^{-4}$, this indicates that the efficiency of the second-order nonlinear conversion can be boosted by 10^4 and for the third-order effect it is even higher.

Use of ultrashort low-duty-cycle laser pulses has become the ubiquitous method of obtaining excellent practical results for both the second-order (frequency conversion, parametric oscillation, and amplification) and the third-order (optical frequency comb and continuum generation) phenomena. And yet if one is thinking of applications in information processing, the switches are expected to operate at the same symbol rate and duty cycle as the data stream. In other words, if the signal itself is, say, a typical "no return to zero" (NRZ) stream of 5-ps FWHM pulses in 10-ps bit intervals, using 1-ps pulses at low duty cycle will not allow one to be able to fully switch each individual symbol. Then one should look at other methods of concentrating the energy and the attention is inevitably drawn back to the space domain and the question arises: Can one transfer the mode-locking techniques from time to space, i.e., to create a low-duty-cycle high-peak-power distribution of optical energy in space, rather than in time, and to use it to effectively enhance nonlinear optical effects.

Extending the time-space analogy, let us look at what limits the degree of energy concentration in time and space. In the time domain it is obviously the dispersion of group velocity, while in the space domain it is the diffraction. While there is obvious equivalence between the mathematical description of dispersion and diffraction, there is a stark difference—the group velocity dispersion can be minimized by a number of techniques because it can be either positive or negative, while the diffraction is always positive and there exists a hard diffraction limit to the optical confinement in an all-dielectric medium. Therefore, the closest space-domain analog to the mode-locking technique is a one-dimensional array of coupled resonators with either ring or photonic band-gap implementation, which does provide some enhancement of nonlinearity, but, as mentioned above, always at the expense of bandwidth [21,22].

It is important to realize, though, that the diffraction limit is applicable only to the all-dielectric structures with positive real parts of dielectric constants. In all-dielectric structures the energy oscillates between electric and magnetic fields, and if the volume in which one tries to confine the optical energy is much less than a wavelength the magnetic field essentially vanishes (so-called quasistatic limit) and without this energy "reservoir" for storage every alternative quarter cycle the energy simply radiates away. But if the structure contains a medium with negative real part of dielectric constant, i.e., free electrons, an alternative reservoir for energy opens up—the kinetic motion of these free carriers in metal or semiconductor, and the diffraction limit ceases being applicable. The optical energy can be then contained in the tightly confined sub-wavelength modes surrounding or filling the gap between the tiny metallic particles. These modes, combining electric field with charge oscillations, are called localized surface plasmons (LSPs), and in the last decade the entirely new closely related fields of plasmonics and metamaterials have arisen with

the ultimate goal of taking advantage of the unprecedented degree of optical energy concentration on the subwavelength scale [23].

Following these arguments, the researchers have observed enhancement of both linear (absorption, luminescence) [24,25] and nonlinear (Raman) phenomena [26–28] in the vicinity of small metal nanoparticles and their combinations. Enhancement of many orders of magnitude has been observed experimentally for surface-enhanced Raman scattering (SERS) [26–28], while the enhancement for luminescence and absorption was more modest. To address this issue we have developed a rigorous yet physically transparent theory explaining this enhancement provided by single [29] or coupled [30,31] nanoparticles in which we have traced the relatively weak enhancement of luminescence to the large absorption in the metal, which cannot be reduced in truly subwavelength mode in which the field is concentrated [32,33]. In that work [32] we have shown that the decay rate of the electric field in the subwavelength mode is always on the order of the scattering time in metal, i.e., 10–20 fs in noble metals. This is the natural consequence of the aforementioned fact that half of the time all the energy is contained in the kinetic motion of electrons in the metal where it dissipates with the scattering rate. As a result, a significant fraction of the LSP energy simply dissipates inside the metal rather than radiating away. The net result is that only very inefficient emitters [34] and also absorbers [35] can be enhanced by plasmonic effects such as, of course, the Raman process that is extremely inefficient [36], while the relatively efficient devices such as light-emitting diodes (LEDs) [37], solar cells [38], and detectors [39] do not exhibit any significant plasmonic enhancement relative to what can be obtained without the metal by purely dielectric means [40].

Therefore, it would only be natural to look into what plasmonic enhancement can do for the inherently weak nonlinear processes, and, although the first works along this direction are over 30 years old [41–46] the interest has peaked significantly in the last decade [47]. There are a number of ways where nonlinear optical effects can be enhanced by the surface plasmons (SPs). One is the coupling of the excitation field to form the much stronger localized field near the surface of the metal structure that leads to the enhancement of optical processes [48]. Such a strong near-field effect is responsible for the experimental observations of significant Raman enhancement that has resulted in single-molecule detection [26–28,49] and SP-enhanced wave mixing such as second-harmonic generation (SHG) on random [50–52] and defined plasmonic structures [53–59], as well as the enhancement of linear processes such as optical absorption and luminescence [24,25]. Another is the fact that SP resonance is ultrasensitive to the dielectric properties of the metal and its surrounding medium—a minor modification in the refractive index around the metal surface can lead to a large shift of plasmonic resonance [60]. Such a phenomenon brings about the prospect of controlling light with another light where the latter induces optical property change in the plasmonic structure which in turn modifies the propagation of the original light. Motivated by this promise, researchers around the world have been pursuing the goal of practical all-optical modulation or switching based on Kerr nonlinearities in either unconfined plasmonic materials [61–64] or waveguides [65–69], which has remained elusive up to this date.

At this point it is important to differentiate between the sources of nonlinearity in these works, because both metals and dielectrics possess nonlinearity. The nonlinear susceptibility of metal can be due to either free carriers or to band-to-band transitions. The nonlinearity of band-to-band transition (typically involving d bands in noble metals) is no different from the interband nonlinearity of dielectrics and semiconductors, except it always occurs in the region of large absorption due to free carriers, and in addition, the nonlinearity is strongest in the blue region of the spectrum, while we prefer to concentrate on the telecommunication region of 1300–1500 nm. As far as nonlinearity of free electrons is concerned, it is extremely weak because LSPs (at least when there are only a few of them per nanoparticle) are nearly perfect harmonic oscillators. That can be easily understood from the following back-of-the envelope calculation. To maintain just a single LSP with, say, $\hbar\omega_0 = 1.25$ eV in a subwavelength mode that, as we mentioned above, decays with a time constant of 10^{-14} fs, it would mean power dissipation of 20 μW in a small volume on the order of, say, 2×10^{-17} cm^{-3}, i.e., very high power density of 10^{12} W/cm^3 and temperature rise on the order of about 10 K per picosecond. Clearly, one cannot expect to find more than a few SPs per mode before the catastrophic meltdown. But then, as we have mentioned before, the energy of SP for half the time is contained in the form of kinetic energy of electrons, hence one can write

$$N_e V m_0 \omega_0^2 x_0^2/2 = \hbar\omega_0, \tag{2}$$

where $N_e \sim 8 \times 10^{22}$ cm^{-3} is the electron density, V is the metal volume, and x_0 is the classical amplitude of each individual electron. From Eq. (2) we immediately find $x_0 \simeq 0.002$ Å and with such a tiny amplitude of motion the free electron cannot "see" any anharmonicity of the potential.

Therefore, we shall consider the structure in which the metal nanoparticles are embedded into the nonlinear material with large nonlinearity and low loss. We shall limit our consideration to the third-order nonlinearity because it leads to optical switching and other interesting phenomena without phase matching, and furthermore, we shall limit ourselves to the nonlinear modulation of the refractive index (real part of susceptibility) rather than absorption (imaginary part). One reason for it is that for the amplitude modulation it is desirable to maintain the "zero" bit level as close to real zero as possible, which can only be done by the interference (as in, for instance, Mach Zehnder interferometer). Another reason is that by modulating the index one can take advantage of advanced phase-modulation formats, such as quadrature phase-shift keying (QPSK), quadrature amplitude modulation (QAM), etc. Index modulation is typically broadband and, in addition to simple modulation and switching, can be used for frequency conversion, while absorption modulation is an inherently resonant phenomenon. Finally, changes in absorption are usually associated with real excitations; hence they are not truly ultrafast.

In Sec. III E we shall briefly consider the implications of using metal nonlinearity as well as modulation of the

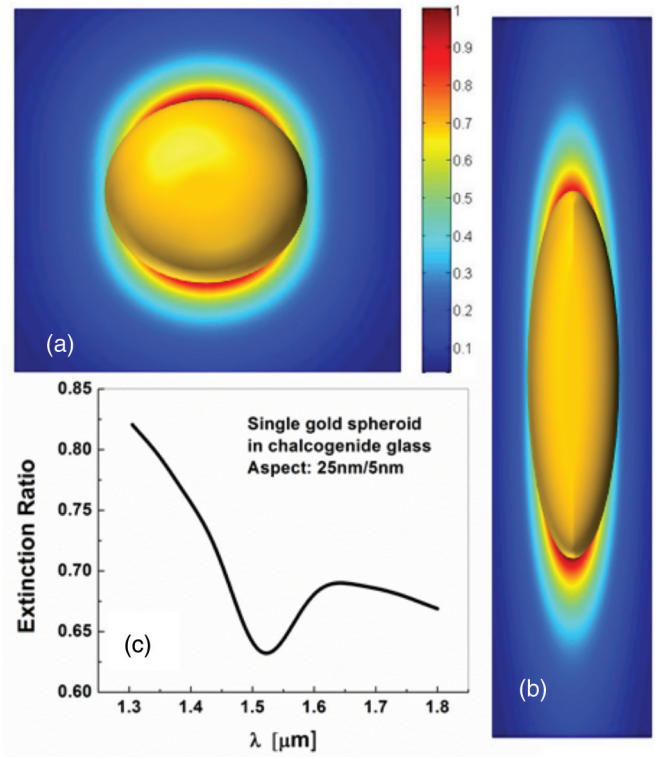

FIG. 2. (Color online) (a) Gold spherical nanoparticle ($a =$ 20 nm) with the electric field distribution. (b) Gold elliptical nanoparticle [long (short) axes: 25 nm (5 nm)] with resonance at 1550 nm and associated electric field distribution. (c) Extinction spectrum of the elliptical nanoparticle.

absorption coefficient and show that essentially the same figures of merit will apply to those plasmonically enhanced nonlinear schemes as to the one which we shall consider at length—the structure shown in Figs. 2(a) and 2(b) in which metal nanoparticles are surrounded by a nonlinear dielectric medium. The goal of our treatment is to evaluate the enhancement of the third-order nonlinear polarizability of this metamaterial, or one can use the term "artificial dielectric" consisting of metal nanoparticles that enhance local field. In the course of this work we shall introduce figures of merit relevant to practical applications and see how the plasmonically enhanced nonlinear materials stack up against the conventional ones. To make our treatment both general and physically transparent we shall fully rely on analytical derivations which, of course, would require certain simplifications that are justified for as long as one is looking just for the order of magnitude of enhancement. For instance, we shall consider spherical and elliptical (or spheroidal) nanoparticles, single and coupled, but we shall indicate how the treatment can be expanded to other shapes of nanoparticles, including nanoshells [70], that can be defined by just three parameters: resonant SP frequency ω_0, quality factor Q, and effective SP mode volume V_{eff}. For this purpose, in Fig. 2(a) we show spherical nanoparticles and in Fig. 2(b) we show elliptical nanoparticles with resonance at the telecommunication wavelength of 1550 nm with actual field distribution calculated numerically. Also shown in Fig. 2(c) is the extinction spectrum of the elliptical particle obtained numerically where the resonance can be observed.

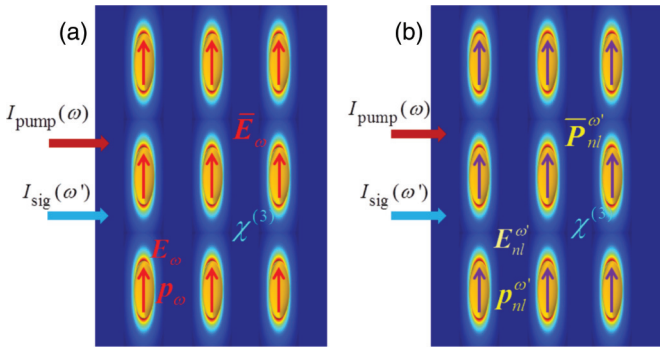

FIG. 3. (Color online) Fields and polarizations in the plasmonically enhanced nonlinear metamaterial: (a) average $\bar{\boldsymbol{E}}_\omega$ and local $\boldsymbol{E}_\omega$ electric fields and dipole $\boldsymbol{p}_\omega$ at the pump frequency; (b) local nonlinear field $\boldsymbol{E}_{\omega'}$, dipole moment $\boldsymbol{p}_{nl}^{\omega'}$, and average nonlinear polarization $\bar{\boldsymbol{P}}_{nl}^{\omega'}$.

II. LINEAR OPTICAL PROPERTIES OF METAL NANOPARTICLES EMBEDDED IN A DIELECTRIC

A. Polarizability and local field enhancement

Consider a rather general scheme for plasmonically enhanced nonlinearity shown in Fig. 3(a) consisting of nanospheres of radius a surrounded by the nonlinear dielectric with relative permittivity ε_d and nonlinear susceptibility tensor $\chi^{(3)}$. The density of these spheres is N_s. In the most general case $\chi^{(3)}$ implies four wave interactions, with some of the waves being the pumps (of switching signals) and some being the nonlinear output signals. In many practical cases, such as cross- and self-phase modulation, degeneracy reduces the number of interacting waves. In Fig. 3(a) we show just one pump (or switching) wave of frequency ω and one signal wave of frequency ω'.

As the pump wave propagates through the material with the average electric field of $\bar{\boldsymbol{E}}_\omega$, the nanospheres become polarized by this field, and acquire the dipole moment [71]

$$\boldsymbol{p}_\omega = \frac{\varepsilon_m - \varepsilon_d}{\varepsilon_m + 2\varepsilon_d} 4\pi\varepsilon_0\varepsilon_d a^3 \bar{\boldsymbol{E}}_\omega = 3\varepsilon_0 V \frac{\varepsilon_m - \varepsilon_d}{\varepsilon_m + 2\varepsilon_d}\varepsilon_d \bar{\boldsymbol{E}}_\omega, \tag{3}$$

as shown in Fig. 3(a). Using the Drude model for the dielectric constant of metal $\varepsilon_m = 1 - \omega_p^2/(\omega^2 + j\omega\gamma)$ with plasma frequency ω_p and scattering rate γ we can write

$$\begin{aligned}\frac{\varepsilon_m - \varepsilon_d}{\varepsilon_m + 2\varepsilon_d} &= \frac{1 - \frac{\omega_p^2}{\omega^2 + i\omega\gamma} - \varepsilon_d}{1 - \frac{\omega_p^2}{\omega^2 + i\omega\gamma} + 2\varepsilon_d} \\ &= \frac{(1-\varepsilon_d)\,\omega^2 - \omega_p^2 + i\omega\gamma(1-\varepsilon_d)}{(1+2\varepsilon_d)\,\omega^2 - \omega_p^2 + i\omega\gamma(1+2\varepsilon_d)} \\ &\approx \frac{\omega_0^2 + \omega^2\frac{\varepsilon_d - 1}{2\varepsilon_d + 1}}{\omega_0^2 - \omega^2 - i\omega\gamma},\end{aligned} \tag{4}$$

where $\omega_0 = \omega_p/\sqrt{1 + 2\varepsilon_d}$ is the LSP resonant frequency [34]. Not far from the resonance we obtain

$$\boldsymbol{p}_\omega = \frac{\alpha\omega_0^2 \bar{\boldsymbol{E}}_\omega}{\omega_0^2 - \omega^2 - i\omega\gamma} \approx \alpha\frac{Q}{L(\omega)}\bar{\boldsymbol{E}}_\omega, \tag{5}$$

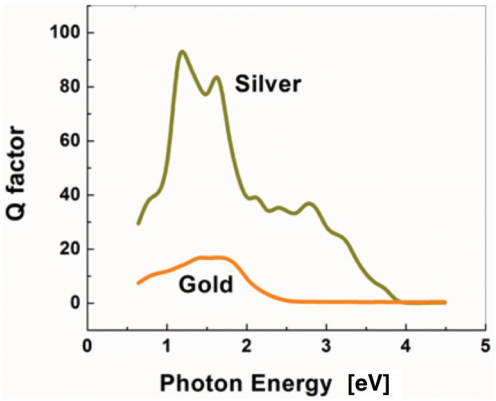

FIG. 4. (Color online) Dispersions of Q factors for gold and silver nanoparticles.

where we have introduced the Q factor of the mode as

$$Q = \omega_0/\gamma, \tag{6}$$

resonant Lorentzian denominator,

$$L(\omega) = Q\left(1 - \omega^2/\omega_0^2\right) - i, \tag{7}$$

and polarizability of the nanoparticle as

$$\alpha \approx 3\varepsilon_0\varepsilon_d V\left(1 + \frac{\varepsilon_d - 1}{2\varepsilon_d + 1}\right) = 3\varepsilon_0\varepsilon_d V\frac{3\varepsilon_d}{2\varepsilon_d + 1} = 3\varepsilon_0\varepsilon_d V\beta. \tag{8}$$

Here, for the spherical nanoparticle, the factor $\beta = 3\varepsilon_d/(2\varepsilon_d + 1)$, and for particles of different shapes it will be somewhat different and polarization dependent, yet still within the same order of magnitude. Similarly, the value of resonant frequency will change; however, since we are interested only in the order of magnitude results in this work, all the conclusions obtained here for spherical particles and their combinations will hold for the particles of different shapes. It should be noted that the Q factor for a particular shape depends only on its resonant frequency ω_0 since the decay rate γ does not depend on the shape (or exact dimensions) as long as particles are much smaller than the wavelength (which is of course required to avoid scattering and diffraction effects).

The Q factor for the gold and silver, the two least-lossy plasmonic materials are shown in Fig. 4 as functions of frequency. Near 1550 nm the Q factor of bulk gold is about 12 and for bulk silver it is closer to 30, according to Johnson and Christy [72], although for silver nanoparticles the interface scattering usually decreases the Q factor by a factor of a few. Also, gold is easier to work with than silver, as it does not get oxidized, so in the subsequent discussion we use gold as the material of choice, although in the end changing Q by a factor of a few will not affect any of our conclusions, since, we stress once again, our results are all just an order of magnitude estimates.

Equation (5) can be construed as the solution of the equation of motion of the harmonic oscillator, or the LSP mode characterized by the dipole moment $\boldsymbol{p}$,

$$\frac{d^2\boldsymbol{p}}{dt^2} + \gamma\frac{d\boldsymbol{p}}{dt} = -\omega_0^2\boldsymbol{p} + \omega_0^2\alpha\bar{\boldsymbol{E}}, \tag{9}$$

and consisting of coupled oscillations of the free electron current inside the nanoparticle, and the electric field inside and outside the spherical nanoparticle [34],

$$\boldsymbol{E}_\omega(\boldsymbol{r}) = \begin{cases} -\frac{\boldsymbol{p}}{4\pi\varepsilon_0\varepsilon_d a^3} & r < a \\ \frac{1}{4\pi\varepsilon_0\varepsilon_d r^3}[3(\boldsymbol{p}\cdot\hat{\boldsymbol{r}})\hat{\boldsymbol{r}} - \boldsymbol{p}] & r > a \end{cases}, \tag{10}$$

with the maximum field near the surface of the nanoparticle equal to

$$\boldsymbol{E}_{\max,\omega}(a) = \frac{1}{4\pi\varepsilon_0\varepsilon_d}\frac{2\boldsymbol{p}_\omega}{a^3} = \frac{2\boldsymbol{p}_\omega}{3\varepsilon_0\varepsilon_d V}, \tag{11}$$

where $V = 4\pi a^3/3$ is the volume of the nanosphere.

The total maximum field near the surface is then a sum of a dipole field and the original average field,

$$\begin{aligned}\boldsymbol{E}_{\max,\omega} &= \bar{\boldsymbol{E}}_\omega + \frac{2}{3\varepsilon_0\varepsilon_d V}\alpha\frac{Q}{L(\omega)}\bar{\boldsymbol{E}}_\omega \\ &= \bar{\boldsymbol{E}}_\omega\left[1 + \frac{2\beta Q}{L(\omega)}\right] \approx \frac{2\beta Q}{L(\omega)}\bar{\boldsymbol{E}}_\omega.\end{aligned} \tag{12}$$

Hence near the resonance the local field is enhanced roughly by a factor of $2Q$ relative to the average field.

The dipole LSP mode equation (10) is the lowest order ($l = 1$) among the many orthogonal modes defined by the angular momentum number l. The electric field of the lth mode can be written as

$$\boldsymbol{E}_l(\boldsymbol{r}) = E_{\max,l}\boldsymbol{F}_l(\boldsymbol{r}), \tag{13}$$

where the normalized mode shape function $\boldsymbol{F}_l(\boldsymbol{r})$ has a maximum value of unity and one can introduce the effective volume $V_{\text{eff},l}$ in such a way,

$$\int F_l(\boldsymbol{r})\frac{\partial[\omega\varepsilon_r'(\boldsymbol{r})]}{\partial\omega}F_l(\boldsymbol{r})d^3\boldsymbol{r} = \varepsilon_d V_{\text{eff},l}, \tag{14}$$

that the total mode energy can be found as

$$U_l = \tfrac{1}{2}\varepsilon_d E_{\max,l}^2 V_{\text{eff},l}. \tag{15}$$

The effective volume of the lth-order mode [30],

$$V_{\text{eff}1} = \frac{\pi a^3}{\varepsilon_d} = \frac{3}{4\varepsilon_d}V, \tag{16}$$

is commensurate with the volume of the nanoparticle itself.

B. Effective index and absorption

If the nanoparticles are much smaller than the wavelength of light in the dielectric, one can apply a classical polarizability theory in which each nanoparticle is treated as a polarizable atom. The *effective dielectric constant* of the composite medium (or a metamaterial if one wants to use a more modern, *de rigueur* terminology) can be found as the sum of the original dielectric constant and the susceptibility of the nanoparticles with a density N_s,

$$\begin{aligned}\varepsilon_{\text{eff}} &= \varepsilon_d + \frac{N_s\alpha}{\varepsilon_0}\frac{Q}{L(\omega)} = \varepsilon_d + \frac{N_s 3\varepsilon_0\varepsilon_d V\beta}{\varepsilon_0}\frac{Q}{L(\omega)} \\ &= \varepsilon_d\left[1 + 3f\beta\frac{Q}{L(\omega)}\right],\end{aligned} \tag{17}$$

where we have introduced the *effective filling factor*,

$$f = N_s V \ll Q^{-1}. \tag{18}$$

The latter condition is practically always satisfied in the medium with $Q \sim 10$ and is required to avoid taking into account local field (collective) effects that would change the LSP resonant frequency according to the Lorentz-Lorenz formula. But, once again, even for a very dense medium the frequency renormalization is not going to change the main conclusions of this work.

In this approximation we can find the *effective index of refraction* as

$$n_{\text{eff}} = \varepsilon_{\text{eff}}^{1/2} \approx n_d \left[1 + \frac{3f\beta}{2}\frac{Q^2\left(1-\omega^2/\omega_0^2\right)}{|L(\omega)|^2} + \frac{3f\beta}{2}\frac{iQ}{|L(\omega)|^2}\right], \tag{19}$$

where $n_d = \sqrt{\varepsilon_d}$. Obviously, the *effective absorption coefficient*,

$$\alpha_a = \frac{2\pi n_d}{\lambda}\frac{3f\beta Q}{|L(\omega)|^2}, \tag{20}$$

also gets resonantly enhanced by the Q factor.

Next we shall mention what happens to energy propagation described by the *Poynting vector* given as

$$\bar{S}_\omega = \frac{\bar{E}_\omega^2}{2\eta_0} n_{\text{eff},r} = \frac{\bar{E}_\omega^2 n_d}{2\eta_0}\left[1 + \frac{3f\beta}{2}\frac{Q^2\left(1-\omega^2/\omega_0^2\right)}{|L(\omega)|^2}\right]. \tag{21}$$

At the same time the average *energy density* inside is

$$\begin{aligned} U_\omega &= \frac{\varepsilon_0\varepsilon_d\bar{E}_\omega^2}{2} + \frac{N_s\varepsilon_0\varepsilon_d E_{\max,\omega}^2 V_{\text{eff}}}{2} \\ &= \frac{\varepsilon_0\varepsilon_d\bar{E}_\omega^2}{2}\left[1 + N_s\frac{3}{4\varepsilon_d}V\frac{4\beta^2Q^2}{|L(\omega)|^2}\right] \\ &= \frac{\varepsilon_0\varepsilon_d\bar{E}_\omega^2}{2}\left[1 + \frac{3f\beta^2}{\varepsilon_d}\frac{Q^2}{|L(\omega)|^2}\right]. \end{aligned} \tag{22}$$

Now we can define the *energy velocity* as

$$v_{E,\omega} = \frac{\bar{S}_\omega}{U_\omega} = \frac{c}{fn_d}\frac{1 + \frac{3f\beta}{2}\frac{Q^2(1-\omega^2/\omega_0^2)}{|L(\omega)|^2}}{1 + \frac{3f\beta}{\varepsilon_d}\frac{Q^2}{|L(\omega)|^2}}. \tag{23}$$

Note that away from resonance the energy propagation velocity is identical to the group velocity, but at resonance we prefer to use energy propagation velocity since the group velocity loses physical meaning (for instance, by becoming negative). The slowing down factor, or relative group index, at resonance is

$$n_g = \frac{c}{n_d v_E} = 1 + \frac{3f\beta^2Q^2}{\varepsilon_d}, \tag{24}$$

which is the consequence of the fact that a large fraction of energy being "stored" in the nonpropagating LSP modes. This result is important because if we now compare Eq. (24) with Eq. (12) we can see that local intensity enhancement at resonance, $4\beta^2Q^2$, can be written roughly as $(n_g - 1)/f$. This indicates the connection between the plasmonics and slow-light structures [73] in which local intensity also gets enhanced, but note that for $Q \sim 10$ and $f < 0.01$ most of the enhancement for plasmonic structures comes not from the slowing down of light but rather redistribution of energy on the subwavelength scale.

III. NONLINEAR PROPERTIES OF METAMATERIAL WITH ISOLATED NANOPARTCLES

A. Nonlinear polarizability

Let us now turn our attention to Fig. 3(b) where the local nonlinear microscopic polarization at the frequency ω',

$$\boldsymbol{P}_{nl}(\boldsymbol{r},t) = P_{\max,nl}^{\omega'}\boldsymbol{G}(\boldsymbol{r})e^{-i\omega' t}, \tag{25}$$

is established near the nanoparticle due to the presence of a strong local pump field. As mentioned above ω' could be the same as or different from the pump frequency ω that drives the nonlinear polarization. The maximum nonlinear polarization is $|\boldsymbol{P}_{nl}(\boldsymbol{r}_{\max})| = P_{\max,nl}^{\omega'}$, occurring usually at the same location where the local pump field reaches maximum, and $\boldsymbol{G}(\boldsymbol{r})$ is the normalized shape of nonlinear polarization. The nonlinear polarization can now drive the LSP oscillations at the same frequency ω' according to the wave equation for the electric field of the LSP mode,

$$\nabla^2\boldsymbol{E}(\boldsymbol{r},t) - \frac{\varepsilon_r(\boldsymbol{r})}{c^2}\frac{\partial^2}{\partial t^2}\boldsymbol{E}(\boldsymbol{r},t) = \frac{1}{\varepsilon_0 c^2}\frac{\partial^2}{\partial t^2}\boldsymbol{P}_{nl}(\boldsymbol{r},t) \tag{26}$$

where the relative dielectric constant can be written inside of the metal as

$$\varepsilon_r(\boldsymbol{r}) = \varepsilon_r'(\boldsymbol{r}) + i\varepsilon_r''(\boldsymbol{r}) \approx \varepsilon_r'(\boldsymbol{r})\left[1 + i\gamma(\boldsymbol{r})/\omega\right]. \tag{27}$$

We look for the solution of the form

$$\boldsymbol{E}(\boldsymbol{r},t) = \sum_l E_{\max,l}^{\omega'}\boldsymbol{F}_l(\boldsymbol{r})e^{-i\omega' t}, \tag{28}$$

where $\boldsymbol{F}_l(\boldsymbol{r})$ is the normalized electric field of the lth LSP eigenmode with $l = 1$ being the dipole mode described by Eq. (10), whose amplitude $E_{\max,1}$ we are trying to determine. Each eigenmode is a solution of the homogeneous wave equation

$$\nabla^2\boldsymbol{F}_l(\boldsymbol{r}) = -\omega_{0,l}^2\frac{\varepsilon_r'(\boldsymbol{r})}{c^2}\boldsymbol{F}_l(\boldsymbol{r}), \tag{29}$$

and the modes are orthogonal and normalized to the effective volume $V_{\text{eff},l}$ of the lth mode as defined in Eq. (14),

$$\int F_m(\boldsymbol{r})\frac{\partial[\omega\varepsilon_r'(\boldsymbol{r})]}{\partial\omega}F_l(\boldsymbol{r})d^3\boldsymbol{r} = \varepsilon_d V_{\text{eff},l}\delta_{lm}. \tag{30}$$

Substituting Eq. (28) into Eq. (26) and using Eq. (29), we obtain

$$\begin{aligned} \sum_l \Big\{&\left[-\omega_{0,l}^2 + \omega'^2 + i\omega'\gamma(\boldsymbol{r})\right]E_{\max,l}(t) \\ &+ 2i\omega'\frac{\partial E_{\max,l}(t)}{\partial t}\Big\}\boldsymbol{F}_l(\boldsymbol{r}) \\ &= -\frac{\omega'^2}{\varepsilon_0\varepsilon_r'(\boldsymbol{r})}P_{\max,nl}^{\omega'}\boldsymbol{G}(\boldsymbol{r}). \end{aligned} \tag{31}$$

If we now multiply Eq. (31) by $\frac{\partial(\varepsilon_r'\omega)}{\partial\omega}\boldsymbol{F}_1(r)$, integrate over the volume, and take advantage of the orthogonality condition equation (30) to obtain the steady-state amplitude of the $l = 1$ dipole mode driven by the nonlinear polarization at frequency ω',

$$E_{\max}^{\omega'} = \frac{P_{\max,nl}^{\omega}\kappa}{\varepsilon_0\varepsilon_d}\frac{Q}{L(\omega')}, \tag{32}$$

where the overlap coefficient, assuming that the dielectric is nondispersive and nonlossy, is

$$\kappa = \frac{\varepsilon_d \int \boldsymbol{F}_1(\boldsymbol{r}) \cdot \boldsymbol{G}(\boldsymbol{r}) dV}{\int \frac{\partial(\varepsilon_r' \omega)}{\partial \omega} F_1^2(r) dV} = \frac{\varepsilon_d \int \boldsymbol{F}_1(\boldsymbol{r}) \cdot \boldsymbol{G}(\boldsymbol{r}) dV}{V_{\text{eff},1}}. \quad (33)$$

Now, according to Eq. (11) we can find the amplitude of the dipole mode,

$$p_{nl}^{\omega'} = \frac{3}{2} V \varepsilon_0 \varepsilon_d E_{\max,nl}^{\omega'} = \frac{3}{2} V \kappa \frac{Q}{L(\omega)} P_{\max,nl}^{\omega'}. \quad (34)$$

And overall effective nonlinear polarization of the metamaterial is

$$\bar{\boldsymbol{P}}_{nl}^{\omega'} = \frac{3}{2} f \kappa \frac{Q}{L(\omega)} \boldsymbol{P}_{\max,nl}^{\omega'}, \quad (35)$$

shown in Fig. 3(b). As one can see, local nonlinear polarization gets enhanced by being at resonance with the nanoparticle dipole mode and enhancement is once again proportional to the Q factor of the resonance.

Before continuing we shall recap the chain of events that leads to the establishment of enhanced nonlinear polarization as shown in Fig. 3:

(i) The average pumping field $\bar{\boldsymbol{E}}_\omega$ polarizes nanoparticles engendering linear dipole moment $\boldsymbol{p}_\omega$ in each of them;

(ii) Dipole oscillations are coupled with the linear local field $\boldsymbol{E}_\omega(\boldsymbol{r})$ in the vicinity of each nanoparticle. This field is resonantly enhanced by a factor of the order of Q relative to $\bar{\boldsymbol{E}}_\omega$;

(iii) A local nonlinear polarization $\boldsymbol{P}_{nl}^{\omega'}(\boldsymbol{r})$ is established in the vicinity of each nanoparticle. Since this polarization is proportional to the third order of the electric field, it is enhanced roughly by a factor of Q^3;

(iv) This polarization resonantly couples into the dipole LSP mode of the nanoparticle thus establishing the local nonlinear field $\boldsymbol{E}_{\omega'}(\boldsymbol{r})$ and the dipole moment $\boldsymbol{p}_{nl}^{\omega'}$. Resonance causes enhancement by another Q factor;

(v) Finally the localized dipoles $\boldsymbol{p}_{nl}^{\omega'}$ combine to establish the average nonlinear polarization $\bar{\boldsymbol{P}}_{nl}^{\omega'}$.

Of course all the steps outlined above occur simultaneously and instantly, but in our view tracing the process step by step provides the clarity of a physical picture. Let us now turn our attention to specific third-order processes.

B. Effective third-order nonlinearity

Consider now the third-order nonlinearity in which interaction of electromagnetic waves at three different frequencies is described by the general local third-order susceptibility,

$$\boldsymbol{P}_{nl}^{\omega_1-\omega_2+\omega_3}(\boldsymbol{r}) = \varepsilon_0 \chi^{(3)}(\omega_3, -\omega_2, \omega_1) \boldsymbol{E}_{\omega_1}(\boldsymbol{r}) \boldsymbol{E}_{\omega_2}^*(\boldsymbol{r}) \boldsymbol{E}_{\omega_3}(\boldsymbol{r}). \quad (36)$$

In general, when all four frequencies, ω_1, ω_2, ω_3, and $\omega_4 = \omega_1 - \omega_2 + \omega_3$ are different (but typically close to each other) the nonlinear process described by Eq. (36) is four-wave mixing (FWM), when $\omega_3 = \omega_1$ $\omega_4 = 2\omega_1 - \omega_2$ Eq. (36) describes optical parametric generation (OPG), when $\omega_1 = \omega_2$ and $\omega_3 = \omega_4$ it describes cross-phase modulation (XPM), and for the case when all frequencies are equal Eq. (36) describes self-phase modulation (SPM). FWM and OPG are both of great interest in wavelength conversion while both XPM and SPM are important for optical switching.

In a composite medium (metamaterial) the local fields $\boldsymbol{E}_{\omega_k}(\boldsymbol{r})$ in Eq. (36) in the vicinity of the nanoparticle are all locally enhanced relative to the mean fields $\bar{E}_{\omega_k}$ according to Eq. (12), i.e.,

$$\boldsymbol{E}_{\omega_k}(\boldsymbol{r}) = \frac{2\beta Q}{L(\omega_k)} \bar{E}_{\omega_k} \boldsymbol{F}_1(\boldsymbol{r}). \quad (37)$$

Hence the local third-order nonlinear polarization is

$$\begin{aligned} \boldsymbol{P}_{nl}^{\omega_1-\omega_2+\omega_3}(\boldsymbol{r}) &= \varepsilon_0 \chi^{(3)}(\omega_3, -\omega_2, \omega_1) \frac{(2\beta Q)^3}{L(\omega_1) L^*(\omega_2) L(\omega_3)} \\ &\quad \times \boldsymbol{F}_1(\boldsymbol{r}) \boldsymbol{F}_1(\boldsymbol{r}) \boldsymbol{F}_1(\boldsymbol{r}) \bar{E}_{\omega_1} \bar{E}_{\omega_2}^* \bar{E}_{\omega_3} \\ &= P_{\max,nl}^{\omega_1-\omega_2+\omega_3} \boldsymbol{G}(\boldsymbol{r}), \end{aligned} \quad (38)$$

where its amplitude is

$$\begin{aligned} P_{\max,nl}^{\omega_1-\omega_2+\omega_3} &= \varepsilon_0 |\chi^{(3)}(\omega_3, -\omega_2, \omega_1)| \frac{(2\beta Q)^3}{L(\omega_1) L^*(\omega_2) L(\omega_3)} \\ &\quad \times \bar{E}_{\omega_1} \bar{E}_{\omega_2}^* \bar{E}_{\omega_3}, \end{aligned} \quad (39)$$

the shape function is

$$\boldsymbol{G}(\boldsymbol{r}) = \frac{\chi^{(3)}}{|\chi^{(3)}|} \boldsymbol{F}_1(\boldsymbol{r}) \boldsymbol{F}_1(\boldsymbol{r}) \boldsymbol{F}_1(\boldsymbol{r}), \quad (40)$$

and $\frac{\chi^{(3)}}{|\chi^{(3)}|}$ is the normalized fourth-order nonlinear susceptibility tensor. Substituting Eq. (40) into Eq. (35) we obtain

$$\begin{aligned} \bar{\boldsymbol{P}}_{nl}^{\omega_4} &= \frac{3}{2} f \kappa_3 \frac{(2\beta)^3 Q^4}{L(\omega_1) L^*(\omega_2) L(\omega_3) L(\omega_4)} \varepsilon_0 \chi^{(3)}(\omega_3, -\omega_2, \omega_1) \\ &\quad \times \bar{\boldsymbol{E}}_{\omega_1} \bar{\boldsymbol{E}}_{\omega_2}^* \bar{\boldsymbol{E}}_{\omega_3} \\ &\equiv \varepsilon_0 \chi_{\text{eff}}^{(3)}(\omega_3, -\omega_2, \omega_1) \bar{\boldsymbol{E}}_{\omega_1} \bar{\boldsymbol{E}}_{\omega_2}^* \bar{\boldsymbol{E}}_{\omega_3}, \end{aligned} \quad (41)$$

where the coupling coefficient for the third-order nonlinearity, according to Eq. (33), is

$$\kappa_3 = \frac{\varepsilon_d}{V_{\text{eff},1}} \int_{r>a} \boldsymbol{F}_1(\boldsymbol{r}) \cdot \frac{\chi^{(3)}}{|\chi^{(3)}|} \boldsymbol{F}_1(\boldsymbol{r}) \boldsymbol{F}_1(\boldsymbol{r}) \boldsymbol{F}_1(\boldsymbol{r}) d^3\boldsymbol{r}. \quad (42)$$

Since typically all the frequencies are close to each other, we can see that the *effective nonlinear susceptibility* gets enhanced as

$$\chi_{\text{eff}}^{(3)} \simeq \frac{3}{2} f \kappa_3 (2\beta)^3 \frac{Q^4}{L^2(\omega) |L(\omega)|^2} \chi^{(3)}, \quad (43)$$

i.e., by a factor proportional to Q^4. This is an outstanding result, exciting enough to attract the attention of both the plasmonic and nonlinear optics communities to this topic, which has witnessed a surge of research efforts and publications as reviewed in Sec. I. Indeed, even with filling factor $f \sim 0.01$ one can expect 100-fold enhancement of susceptibility. It means 100-fold enhancement of the nonlinear refractive index and indicates that one can achieve the same efficiency of nonlinear phase modulation at about $1/100$ of the length, and, more dramatically, the same efficiency of the wavelength conversion in only $1/10\,000$ of the length. These are precisely the results that prompted many scientists to enter the race for the largest enhancement. It would be nice if we could end our discussion right here on this optimistic note, but

one needs to maintain caution when it comes to reporting these giant plasmonic enhancements. Our prior research of plasmonic enhancement of various emission processes including photoluminescence [29], electroluminescence [34], and Raman scattering [36] has shown that large enhancements are feasible only for the processes that have very low original efficiencies (such as Raman scattering) but are far more modest for those efficient processes such as fluorescence and electroluminescence. It is therefore reasonable to expect that there must exist an upper limit of the nonlinear plasmonic enhancement.

C. Effective nonlinear index and maximum phase shift

To understand the limitations of the enhancement we shall first consider XPM (or SPM) for which nonlinear polarization in Eq. (36) can be written as

$$P_{nl}^{\omega_2}(\boldsymbol{r}) = 2\varepsilon_0 n_d n_2 I_{\omega_1}(\boldsymbol{r}) E_{\omega_2}(\boldsymbol{r}), \tag{44}$$

where $n_2 = \chi^{(3)}(\omega_1, -\omega_1, \omega_2)\eta_0/\varepsilon_d$ is the nonlinear index of the dielectric, and $I_{\omega_1}(\boldsymbol{r}) = |E_{\omega_1}(\boldsymbol{r})|^2 n_d/2\eta_0$ is the local intensity (of course the energy in the near field does not propagate and this definition of local intensity is just a convenient way to relate the susceptibility with the nonlinear index). The local change of refractive index is then

$$\Delta n(\boldsymbol{r}) = n_2 I_{\omega_1}(\boldsymbol{r}) = n_2 \bar{I}_{\omega_1} (2\beta)^2 \frac{Q^2}{|L(\omega_1)|^2} F_1^2(\boldsymbol{r}), \tag{45}$$

where the average intensity $\bar{I}_{\omega_1} = n_d |\overline{E}_{\omega_1}|^2/2\eta_0$.

Similarly, we now introduce the *effective nonlinear index* as $n_{2,\text{eff}} = \chi_{\text{eff}}^{(3)}(\omega_1, -\omega_1, \omega_2)\eta_0/\varepsilon_d$ and rewrite Eq. (41) as

$$\bar{P}_{nl}^{\omega_2} = 2\varepsilon_0 n_d n_{2,\text{eff}} \bar{I}_{\omega_1} \bar{E}_{\omega_2}. \tag{46}$$

According to Eq. (43) the effective nonlinear index gets enhanced by the same giant factor proportional to Q^4,

$$n_{2,\text{eff}} \simeq \frac{3}{2} f \kappa_3 (2\beta)^3 \frac{Q^4}{L^2(\omega_2)\,|L(\omega_1)|^2} n_2. \tag{47}$$

Since the pump intensity decreases due to absorption, the nonlinear phase shift can be found as

$$\Delta\Phi(z) = \frac{2\pi}{\lambda} n_{2,\text{eff}} \int_0^z \bar{I}_{\omega_1}(z) dz = \frac{2\pi}{\lambda\alpha_a} n_{2,\text{eff}} \bar{I}_0 (1 - e^{-\alpha_a z}), \tag{48}$$

where $\bar{I}_0$ is the input pump intensity and α_a is the absorption coefficient defined in Eq. (20). This means that the maximum phase shift obtained after about one absorption length is

$$\Delta\Phi_{\max} = \frac{|L(\omega_1)|^2}{3 f \beta Q} \frac{n_{2,\text{eff}}}{n_d} \bar{I}_0 \approx \kappa_3 (2\beta)^2 \frac{Q^3}{L^2(\omega_2)} \frac{n_2}{n_d} \bar{I}_0. \tag{49}$$

Achieving the π-phase shift required to get photonic switching would then require at resonance $\bar{I}_\pi \sim \pi n_d [\kappa_3 n_2 (2\beta)^2 Q^3]^{-1}$. If we assume $\beta \sim 1.35$ (estimated numerically for the actual ellipsoid resonant at 1550 nm), $Q \sim 12$, and large nonlinear index characteristic of chalcogenide glass $n_2 = 10^{-13}$ cm^2/W, we obtain the required switching intensity of the order of $\bar{I}_\pi \sim 6 \times 10^9$ W/cm^2.

This result indicates that the giant nonlinear index enhancement equation (47) can only be used to reduce the length of the device, while the switching intensity remains very high—requiring peak powers of about 60 W into a 1-μm^2 waveguide. But the situation is actually even less optimistic; according to Eq. (12) the maximum local intensity near the nanoparticle surface is

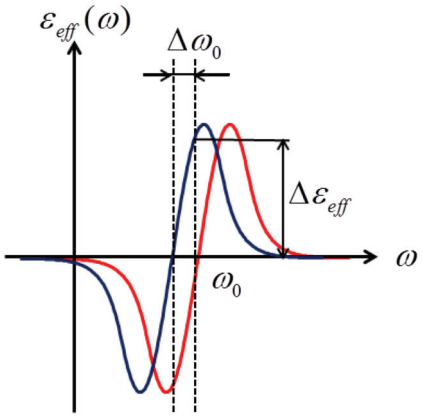

FIG. 5. (Color online) Change of dielectric constant ε_{eff} caused by the shift of the resonance frequency of SP ω_0.

$$I_{\max} = (2\beta Q)^2 \bar{I}_\pi \approx 6 \times 10^{12} \text{ W/cm}^2 \tag{50}$$

(or maximum local field is of 5×10^7 V/cm), which is significantly higher than the damage threshold of the material, and besides as simple calculations may show, it leads to the local temperature rise on the scale of 10^4 K/ps! In fact, if one searches through all the nonlinear materials, it is difficult to find one that is capable of achieving ultrafast refractive index change larger than 0.1%. In addition to limitation due to overheating and optical damage, at high power the nonlinearities of higher than the third order, i.e., $\chi^{(5)}$ and $\chi^{(7)}$, become important, and they often have their signs opposite to $\chi^{(3)}$ [74] which leads to actual decrease in the nonlinear index change at high intensities.

Therefore, let us define the maximum local nonlinear index change attainable in a given material as $\Delta n_{\max}$. Then from Eqs. (45) and (47) we obtain

$$\Delta n_{\text{eff, max}} \approx 3 f \kappa_3 \beta \frac{Q^2}{L^2(\omega_2)} \Delta n_{\max}. \tag{51}$$

As we can see now the enhancement is only proportional to Q^2. This result makes perfect sense if we recognize that local change of dielectric constant $\Delta\varepsilon_{d,\max} = 2n_d \Delta n_{\max}$ simply causes the shift of the LSP resonant frequency $\omega_0 = \omega_p/\sqrt{1+2\varepsilon_d}$, as shown in Fig. 5, which in turn changes the effective dielectric constant of the metamaterial $\varepsilon_{\text{eff, max}}$ according to Eq. (17) as

$$\Delta\varepsilon_{\text{eff, max}} \approx \frac{\partial\varepsilon_{\text{eff}}}{\partial\varepsilon_d} \Delta\varepsilon_{d,\max} = \frac{\partial\varepsilon_{\text{eff}}}{\partial L(\omega_2)} \frac{\partial L(\omega_2)}{\partial\omega_0^2} \frac{\partial\omega_0^2}{\partial\varepsilon_d} \Delta\varepsilon_{d,\max} \sim \frac{6\varepsilon_d \beta f}{1+2\varepsilon_d} \frac{Q^2}{L^2(\omega_2)} \Delta\varepsilon_{d,\max}, \tag{52}$$

where we have kept only the largest term proportional to Q^2 and disregarded smaller terms proportional to Q.

What is most important though is that the *maximum obtainable phase shift* equation (49) becomes

$$\Delta\Phi_{\max} = \frac{2\pi}{\lambda\alpha_a}\Delta n_{\text{eff, max}} = \kappa_3 Q \frac{|L(\omega_1)|^2}{L^2(\omega_2)}\frac{\Delta n_{\max}}{n_d}. \quad (53)$$

The simple meaning of Eq. (53) is that, even if we assume enormous local nonlinear index change of 1% (local intensity of 10^{11} W/cm^2), we cannot expect to get phase shift higher than 0.1, almost two orders of magnitude less than what is required for π-phase shift switching.

D. Frequency conversion using FWM

It is easy to see that small maximum phase shift for XPM or SPM corresponds to even smaller efficiency of the frequency conversion for FWM or OPG. Indeed the growth of the idler $\bar{E}_{\omega_3}(z)$ in the presence of pump $\bar{I}_{\omega_1}(z) = \bar{I}_0 e^{-\alpha_a z}$ and signal $E^*_{\omega_2}(z) = E_s e^{-\frac{\alpha_a}{2}z}$ can be described by

$$\frac{d\bar{E}_{\omega_3}(z)}{dz} = \frac{2\pi}{\lambda} n_{2,\text{eff}} \bar{I}_{\omega_1}(z) E^*_{\omega_2}(z) - \frac{\alpha_a}{2} E_{\omega_3}(z) \quad (54)$$

with the solution

$$\bar{E}_{\omega_3}(z) = \frac{2\pi}{\lambda\alpha_a} n_{2,\text{eff}} \bar{I}_0 E_s [1 - e^{-\alpha_a z}] e^{-\alpha_a z/2} \quad (55)$$

that reaches a maximum near $z = \alpha_a^{-1} \ln 3$ equal to

$$\bar{E}_{i,\max} = \frac{2\pi}{\lambda\alpha_a}\frac{2}{3^{3/2}} n_{2,\text{eff}} \bar{I}_0 E_s = \frac{2}{3^{3/2}}\Delta\Phi_{\max} E_s. \quad (56)$$

Therefore, maximum conversion efficiency from signal to idler is

$$\frac{I_i}{I_s} = \frac{1}{3}\left(\frac{2}{3}\Delta\Phi_{\max}\right)^2, \quad (57)$$

and under no conceivable conditions can it exceed -30 dB.

E. Absorption modulation

Although we have mentioned that absorption modulators are not nearly as versatile as the phase modulators, we can simply find the maximum change of the imaginary part of the refractive index from Eq. (51) which will occur when $\omega \approx \omega_0(1 \pm 2Q^{-2})$ and is equal to

$$(\text{Im}\Delta n_{\text{eff}})_{\max} \approx 2 f \kappa_3 \beta \frac{Q^2}{L^2(\omega_2)}\Delta n_{\max}, \quad (58)$$

which upon comparison with Eq. (19) translates to the maximum absorption coefficient modulation,

$$\frac{\Delta\alpha_a}{\alpha_a} \sim \frac{4}{3}\kappa_3 Q \frac{\Delta n_{\max}}{n_d}. \quad (59)$$

Thus the change in absorption per unit length does not appear to increase far beyond a few percent. Since the absorption process is exponential, one can, of course, achieve deeper modulation than that of the total transmission by propagating over many absorption lengths, but the insertion loss would be overwhelming. Therefore, if we use the figure of merit, the change of transmission per one absorption length, i.e., $1 - \exp(\Delta\alpha_a/\alpha_a) \approx \Delta\alpha_a/\alpha_a$, we can see that one cannot achieve significant absorption modulation without incurring enormous loss.

Here we should also briefly mention that one could use modulation of the refractive index of the metal itself, but it is difficult to see how one can change the index of metal by more than 1% unless one operates near the interband transitions where the Q factor is greatly reduced, which defeats the whole purpose of plasmonic enhancement.

IV. METAMATERIALS WITH DIMERS OR NANOLENSES

A. Local field enhancement

Now we have concluded that while nonlinear susceptibility and nonlinear index of refraction do get enhanced significantly in the simple nanostructures, the strong absorption makes maximum attainable phase shift less than desired. From the previous work of ourselves [75,76] as well as others [77,78], we have established that local fields can be enhanced even further in more complicated nanoparticle structures. In the gap between two identical nanoparticles (dimer) [75] or in the vicinity of a smaller nanoparticle coupled to a larger nanoparticle of the same shape (nanolens) [76], the maximum field enhancement was proportional to Q^2 rather than Q for a single nanoparticle; hence much larger "cascaded" enhancements of absorption, Raman scattering, and in some cases photoluminescence could be achieved in these "hot spots." Therefore, it is tempting to evaluate the possibility of using the hot spots to enhance nonlinearity. Since we have shown that in either dimer or nanolens the field enhancement

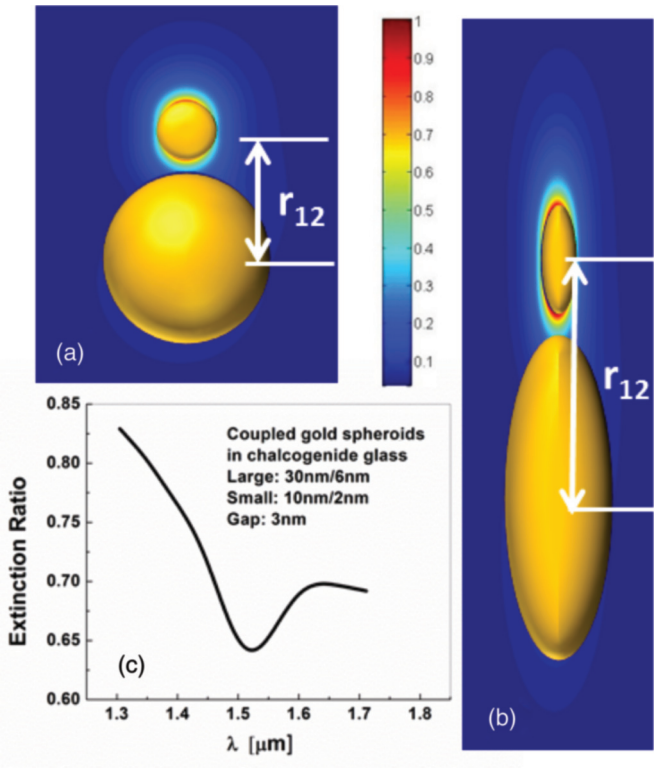

FIG. 6. (Color online) (a) Gold spherical nanoparticle dimer ($a_1 = 20$ nm, $a_2 = 10$ nm, gap $g = 4$ nm, $r_{12} = a_1 + a_2 + d$) with the electric field distribution. (b) Gold elliptical nanoparticle dimer [long (short) axes: 30 nm (6 nm) and 10 nm (2 nm); 5-nm gap] with resonance at 1550 nm and associated electric field distribution. (c) Extinction spectrum of the elliptical dimer.

is similar, we shall limit our analysis to the case of nanolens only, as it is easier to describe analytically.

Consider two spherical nanoparticles of radii a_1 and a_2 separated by a vector $\boldsymbol{r}_{12}$ as shown in Fig. 6(a). The dipole oscillation equation (9) is augmented by the dipole-dipole interaction between the two dipoles associated with the two coupled nanoparticles,

$$\frac{d^2\boldsymbol{p}_{1(2)}}{dt^2}+\gamma\frac{d\boldsymbol{p}_{1(2)}}{dt} = -\omega_0^2\boldsymbol{p}_{1(2)}+\omega_0^2\alpha_{1(2)}\bar{\boldsymbol{E}}_\omega+\omega_0^2\alpha_{1(2)}\frac{2\boldsymbol{p}_{2(1)}}{4\pi\varepsilon_0\varepsilon_d r_{12}^3}. \quad (60)$$

Using the definition of polarizability, Eq. (8), we obtain for the harmonic field $\bar{\boldsymbol{E}}_\omega$ parallel to the direction of $\boldsymbol{r}_{12}$,

$$\begin{aligned}\left(\omega_0^2-\omega^2-j\omega\gamma\right)p_1^\omega-2\omega_0^2\beta\frac{a_1^3}{r_{12}^3}p_2^\omega&=\omega_0^2\alpha_1\bar{E}_\omega,\\ \left(\omega_0^2-\omega^2-j\omega\gamma\right)p_2^\omega-2\omega_0^2\beta\frac{a_2^3}{r_{12}^3}p_1^\omega&=\omega_0^2\alpha_2\bar{E}_\omega,\end{aligned} \quad (61)$$

or simply

$$\begin{aligned}L(\omega)Q^{-1}p_1^\omega-2\beta\left(\frac{a_1}{r_{12}}\right)^3p_2^\omega&=\alpha_1\bar{E}_\omega\\ -2\beta\left(\frac{a_2}{r_{12}}\right)^3p_1^\omega+L(\omega)Q^{-1}p_2^\omega&=\alpha_2\bar{E}_\omega,\end{aligned} \quad (62)$$

which leads to the solution

$$\begin{aligned}p_1^\omega&=Q\frac{L(\omega)\alpha_1+2\beta Q\left(\frac{a_1}{r_{12}}\right)^3\alpha_2}{L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3}\bar{E}_\omega\\ &=4\pi\varepsilon_0\varepsilon_d\beta\frac{L(\omega)a_1^3+2\beta Q\left(\frac{a_1}{r_{12}}\right)^3a_2^3}{L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3}\bar{E}_\omega,\\ p_2^\omega&=Q\frac{L(\omega)\alpha_2+2\beta Q\left(\frac{a_2}{r_{12}}\right)^3\alpha_1}{L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3}\bar{E}_\omega\\ &=4\pi\varepsilon_0\varepsilon_d\beta\frac{L(\omega)a_2^3+2\beta Q\left(\frac{a_2}{r_{12}}\right)^3a_1^3}{L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3}\bar{E}_\omega.\end{aligned} \quad (63)$$

For the maximum electric fields we obtain

$$\begin{aligned}E^\omega_{\max,1}&=\frac{1}{4\pi\varepsilon_0\varepsilon_d}\frac{2p_1^\omega}{a_1^3}=2Q\frac{L(\omega)\beta+2\beta^2Q\left(\frac{a_2}{r_{12}}\right)^3}{L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3}\bar{E}_\omega,\\ E^\omega_{\max,2}&=\frac{1}{4\pi\varepsilon_0\varepsilon_d}\frac{2p_2^\omega}{a_2^3}=2Q\frac{L(\omega)\beta+2\beta^2Q\left(\frac{a_1}{r_{12}}\right)^3}{L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3}\bar{E}_\omega.\end{aligned} \quad (64)$$

Obviously the maximum enhancement will take place if one can get this condition,

$$4\beta^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3\ll Q^{-2},\quad 2\beta\left(\frac{a_1}{r_{12}}\right)^3\gg Q^{-1}. \quad (65)$$

Essentially, we are going to the limit of $a_2\approx 0$ $a_1\approx r_{12}$, which brings us to

$$E^\omega_{\max,1}\approx\frac{2\beta Q}{L(\omega)}\bar{E}_\omega,\quad E^\omega_{\max,2}\approx\left[\frac{2\beta Q}{L(\omega)}\right]^2\bar{E}_\omega. \quad (66)$$

As one can see in Fig. 6(a) the field is greatly enhanced in the vicinity of the smaller particle. In our prior work [76], using more precise calculations we have shown that the simple analytical results of Eq. (66) can be used as an upper bound on the field enhancement in the nanolens, or as a matter of fact, in the nanogap between two particles. In Fig. 6(b) we show the dimer that resonates at the wavelength of choice of 1550 nm, as well as its extinction spectrum in Fig. 6(c). These results have been obtained using precise numerical calculations. Therefore we can get the enhancement on the order of Q^2 for the smaller nanoparticle and the prospect seems to look bright.

B. Nonlinear polarizability

The high field in the vicinity of the smaller nanoparticle will cause nonlinear polarization, Eq. (25),

$$\boldsymbol{P}^\omega_{nl,2}(\boldsymbol{r},t)=P^\omega_{\max,2}\boldsymbol{G}_2(\boldsymbol{r})e^{-i\omega t}, \quad (67)$$

where $\boldsymbol{G}_2(\boldsymbol{r})$ is the normalized distribution of nonlinear polarization near the smaller particle. Then according to Eqs. (31)–(34) this polarization will induce the nonlinear dipoles of two particles via the new driving term on the right-hand side of Eq. (62),

$$\begin{aligned}L(\omega)Q^{-1}p^\omega_{nl,1}-2\beta\left(\frac{a_1}{r_{12}}\right)^3p^\omega_{nl,2}&=0,\\ -2\beta\left(\frac{a_2}{r_{12}}\right)^3p^\omega_{nl,1}+L(\omega)Q^{-1}p^\omega_{nl,2}&=\frac{3}{2}V_2\kappa P^\omega_{\max,2}.\end{aligned} \quad (68)$$

We then arrive at

$$\begin{aligned}p^\omega_{nl,1}&=2\pi a_2^3Q^2\frac{2\beta\left(\frac{a_1}{r_{12}}\right)^3\kappa P^\omega_{\max,2}}{L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3},\\ p^\omega_{nl,2}&=2\pi a_2^3Q\frac{L(\omega)\kappa P^\omega_{\max,2}}{L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3}.\end{aligned} \quad (69)$$

As one can see from comparison to Eq. (34), the nonlinear dipole of the larger nanoparticle 1 experiences additional enhancement relative to the dipole of the smaller nanoparticle 2. But note that now the volume of the smaller nanoparticle is present in the numerator of Eq. (69), hence the situation that is optimum for the external field enhancement in the nanolens equation (65), i.e., the limit of $a_2\approx 0$ $a_1\approx r_{12}$, is far from being optimal for the enhancement of nonlinear polarization.

C. Effective third-order nonlinearity of the nanolens medium

Let us now estimate the effective nonlinear susceptibility of the nanolens. According to Eqs. (39) and (64) the maximum

nonlinear polarization near the smaller nanoparticle is

$$P^{\omega_1-\omega_2+\omega_3}_{\max,2} = \varepsilon_0|\chi^{(3)}(\omega_3,-\omega_2,\omega_1)|(2Q)^3\frac{\left[L(\omega)+2\beta Q\left(\frac{a_1}{r_{12}}\right)^3\right]\left|L(\omega)+2\beta Q\left(\frac{a_1}{r_{12}}\right)^3\right|^2}{\left[L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3\right]\left|L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3\right|^2}\bar{E}_{\omega_1}\bar{E}^*_{\omega_2}\bar{E}_{\omega_3}, \tag{70}$$

and substituting it into Eq. (69) we obtain the nonlinear dipole of the larger particle 1 being equal to

$$p^{\omega}_{nl,1} = 16\kappa_3\varepsilon_0\chi^{(3)}\beta Q^5\frac{3}{2}V_1\left(\frac{a_2}{r_{12}}\right)^3\frac{\left[L(\omega)+2\beta Q\left(\frac{a_1}{r_{12}}\right)^3\right]\left|L(\omega)+2\beta Q\left(\frac{a_1}{r_{12}}\right)^3\right|^2}{\left[L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3\right]^2\left|L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3\right|^2}\bar{E}_{\omega_1}\bar{E}^*_{\omega_2}\bar{E}_{\omega_3}, \tag{71}$$

and the effective susceptibility becomes

$$\chi^{(3)}_{\text{eff}} = 24f\chi^{(3)}\kappa_3\beta Q^5\frac{\left[L(\omega)\left(\frac{a_2}{r_{12}}\right)^3+2\beta Q\left(\frac{a_1a_2}{r_{12}^2}\right)^3\right]\left|L(\omega)+2\beta Q\left(\frac{a_1}{r_{12}}\right)^3\right|^2}{\left[L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3\right]^2\left|L^2(\omega)-4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3\right|^2}. \tag{72}$$

So what is the maximum attainable nonlinearity enhancement? According to Eq. (66), the local field gets enhanced by a factor proportional to Q^2 instead of Q for a single nanoparticle. For Raman scattering, which is also a third-order nonlinear process, the enhancement with the nanolens system can be Q^8 instead of Q^4 for a single nanoparticle, a tremendous improvement. Can we expect similar improvement for the FWM and other third-order nonlinear processes? The answer is no, because according to Eq. (65) the largest enhancement of local fields is always attained when the volume of the smaller particle becomes negligibly small. But the key characteristic of Eq. (72), already noted above, is the presence of the volume of the smaller nanoparticle in the numerator; hence the optimum condition for the maximum effective $\chi^{(3)}$ will not coincide with the condition for maximum local field enhancement and overall enhancement will be less than Q^8.

To find this condition we consider the resonant case $L^2(\omega\sim\omega_0)=-1$ and assume $2\beta Q(a_1/r_{12})\gg 1$; then we need to optimize,

$$\chi^{(3)}_{\text{eff}} \approx 48f\chi^{(3)}\kappa_3\beta^2Q^6\left(\frac{a_1}{r_{12}}\right)^6\frac{4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3}{\left[1+4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3\right]^4}. \tag{73}$$

This enhancement reaches its maximum when

$$4\beta^2Q^2\left(\frac{a_1a_2}{r_{12}^2}\right)^3 = \frac{1}{3}, \tag{74}$$

and equals

$$\chi^{(3)}_{\text{eff}} \approx 5f\chi^{(3)}\kappa_3\beta^2Q^6. \tag{75}$$

Well, as one can see, the enhancement of $\chi^{(3)}$ and nonlinear index n_2 provided by the nanolens system is only proportional to the Q^6. This is rather easy to interpret. The local intensity in the nanolens gets enhanced by a factor proportional to Q^4 and then, according to Eq. (52) the nonlinear refractive index change gets enhanced by the same additional factor Q^2, whether it is a single particle, nanolens, dimer, or nanoantenna. The additional enhancement provided by the coupled particles composite equation (75), compared to the isolate nanoparticle composite equation (43), is about

$$\frac{\chi^{(3)}_{\text{eff},2}}{\chi^{(3)}_{\text{eff},1}} \approx \frac{5}{12}\frac{Q^2}{\beta}, \tag{76}$$

i.e., a factor on the order of 100. Overall enhancement for the previously considered case of $\beta\sim 1.35$, $Q\sim 12$, and $f=0.01$ in chalcogenide glass can be as high as 1×10^5, but the relevant question is what it means in terms of maximum phase shift that can be obtained.

D. Maximum attainable phase shift in nanolens

This maximum shift can be obtained in a way similar to Eq. (49),

$$\Delta\Phi_{\max} \approx 1.7\kappa_3\beta Q^5\frac{n_2}{n_d}\bar{I}_0. \tag{77}$$

Therefore the pump optical intensity required to achieve π-phase shift is $\bar{I}_\pi\sim 8\times 10^7$ W/cm^2, i.e., less than 1 W of peak power into a 1-μm^2 waveguide. This appears to be a reasonable power, but, of course the problem is that the local intensity is enhanced according to Eqs. (64) and (74) roughly by

$$\frac{I_{\max}}{\bar{I}_\pi} = \left|\frac{E^{\omega}_{\max,2}}{\bar{E}_\omega}\right|^2 \approx 9\beta^4Q^4 \approx 6\times 10^5, \tag{78}$$

where we have used $L^2(\omega)\approx -1$ near the resonance, indicating that the local intensity is on the scale of $I_{\max}\approx 5\times 10^{13}$ W/cm^2 which is way beyond the optical damage value. If we introduce once again the maximum local nonlinear index as $\Delta n_{\max}=n_2\bar{I}_{\max}=9\beta^4Q^4n_2\bar{I}_0$, Eq. (77) can be rewritten as

$$\Delta\Phi_{\max} \approx 0.2\kappa_3\frac{Q}{\beta^3}\frac{\Delta n_{\max}}{n_d}. \tag{79}$$

This result for the nanolens is even worse (by a factor of about 5) than the result in Eq. (53) for the isolated nanoparticles. Clearly, the dependence $\kappa_4 Q$ is common to any type of nanostructure, monomer, dimer, trimer, or nanoantenna. The maximum achievable index of refraction $\Delta n_{\max}$ changes the resonant frequency according to Eq. (52) which provides enhancement by the factor of Q^2, but then the absorption coefficient also gets enhanced by the factor of Q so only a

single factor of Q survives in the end. The factor in front of $\kappa_3 Q$ is reduced in dimers and more complicated structures relative to the monomers simply because a smaller fraction of the mode energy is contained in the region where the index change is maximal. Hence one should not expect any improvement in the maximum obtainable nonlinear phase shift $\Delta\Phi_{\max}$ beyond a single factor of Q in more complicated structures such as trimers, bowtie antennae, and so on, even if the effective nonlinear index can be enhanced beyond the already huge enhancement in Eq. (75). Giant enhancement of nonlinearity will only mean that the nonlinear phase shift will saturate at much shorter distance but at essentially the same value of Eq. (53) or less, indicating that, to the best of our knowledge on existing materials, it is impossible to achieve true all-optical switching using plasmonic enhancement.

V. DISCUSSION

Everything said above can be summarized in a simple figure which tells it all (Fig. 7). We consider a 1-μm^2 cross-section waveguide made of chalcogenide glass with nonlinear index $n_2 = 10^{-13}$ cm^2/W. We consider the third-order nonlinear process of XPM (although the case of SPM is no different). The peak pump power is $P_0 = 1.6$ mW so that the intensity inside the waveguide free of metal nanoparticles is $I_0 = 1.6 \times 10^5$ W/cm^2 and the nonlinear index change is $\Delta n_0 = 1.6 \times 10^{-8}$. If the waveguide is filled with solitary Au nanoparticles, the local intensity is enhanced by a factor of $(2\beta Q)^2$, to $I_s = 1.8 \times 10^8$ W/cm^2, and the local nonlinear index change is $\Delta n_s = 1.8 \times 10^{-5}$. When the waveguide is filled with Au dimers optimized according to Eq. (74), the local intensity is enhanced by Eq. (78) to $I_d = 1.0 \times 10^{11}$ W/cm^2 which causes

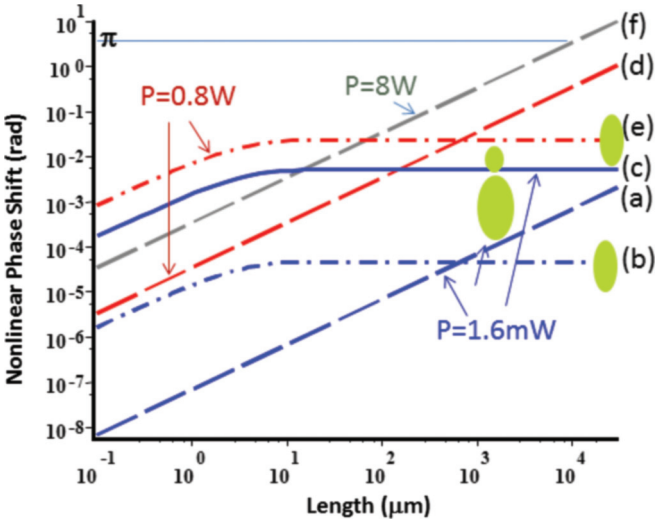

FIG. 7. (Color online) Nonlinear phase shift in the 1-μm^2 chalcogenide waveguide: (a) input power of 1.6 mW, no plasmonic enhancement; (b) input power of 1.6 mW, nonlinearity enhanced by the elliptical Au nanoparticles with the filling factor of 0.001; (c) input power of 1.6 mW, nonlinearity enhanced by the dimers of elliptical Au nanoparticles with the filling factor of 0.001; (d) input power of 0.8 W, no plasmonic enhancement; (e) input power of 0.8 W, nonlinearity enhanced by the elliptical Au nanoparticles with the filling factor of 0.001; (f) input power of 8 W, no plasmonic enhancement.

local index shift $\Delta n_d = 0.01$—a rather nonrealistic value in view of potential optical damage and saturation, but we shall consider it to be an upper limit.

The nonlinear phase shift versus distance curves are shown in Figs. 7(a)–7(c) for the three aforementioned cases. In the waveguide without nanoparticles [curve (a)] the nonlinear phase shift increases gradually and reaches 2×10^{-3} rad at length $z = 3$ cm. For the waveguide impregnated with Au monomers with the volume fraction $f = 0.001$ [curve (b)] the nonlinear shift is much larger but at $z \sim 5$ μm the phase shift saturates at a rather small value of 2×10^{-5} rad. In the waveguide impregnated with Au dimers with the same volume fraction $f = 0.001$ [curve (c)] the enhancement is stronger and it also saturates at 5×10^{-3} rad. In all three cases the required 180° shift is attainable at this low power.

If we increase the input power by a factor of 500 to $P_0 = 800$ mW the nonlinear shift in the waveguide without nanoparticles [curve (d)] would now reach a value of 1 rad at $z = 3$ cm. For the waveguide with Au monomers [curve (e)] the phase shift will now saturate at 0.02 rad and this is the maximum phase shift attainable because local intensity now approaches the damage threshold of 10^{11} W/cm^2. There is no curve for the waveguide with dimers because the local intensity there would be nearly 10^{14} W/cm^2 which is way beyond optical damage. Once again, no "full optical switching" can be achieved over a reasonably small distance.

Finally we can further increase input power by another factor of 10–8 W. Now one cannot use waveguides with plasmonic enhancement because local intensities would be beyond optical damage threshold for both monomers and dimers. So the only curve (f) is that for the waveguide without nanoparticles and as one can see the phase shift indeed achieves π-phase shift at a length of about 9 mm.

VI. CONCLUSIONS

Thus we arrive at a rather dichotomous conclusion. On one hand, using waveguides impregnated with metallic monomers, dimers, and other constructions (one may call them plasmonic metamaterials) allows one to achieve huge enhancement of effective nonlinear index, up to the order of 10^5 and more due to the high degree of field concentration in the hot spots On the other hand, strong absorption in the metal causes saturation of the nonlinear phase shift for SPM and XPM or frequency conversion efficiency in the case of FWM and OPG at very short distances. Given the fact that maximum local index change is limited, generously, to about 1% due to optical damage, the nonlinear phase shift saturates at a very small value of a few tens of milliradians—which is insufficient for any photonic switching operation. Similarly, conversion efficiency saturates at values less than -30 dB, making use of plasmonic nonlinear metamaterials for this purpose highly inefficient. It is also clear that changes in Q by a factor of 2–3 that might be attainable in silver (although not demonstrated to date due to oxidation and surface scattering) will not change the results in any substantial way, and only assure earlier saturation of the nonlinear conversion. The one and only advantage of nonlinear plasmonic metamaterial is that nonlinear effects may be observable at very small propagation distances of a few micrometers with reasonable (but not low!)

optical powers—but there is a giant chasm between being observable and being practical, and at this point, with the existing metals and nonlinear materials, one cannot see how the nonlinear plasmonic metamaterials can bridge this chasm.

In retrospect, our rather unenthusiastic conclusion about the prospects of using plasmonic resonances to enhance nonlinearity does not appear to be surprising at all. Numerous resonant schemes for enhancement nonlinearity have been proposed and investigated at length [79,80]. Some of the schemes rely upon intrinsic material resonances; others try to take advantage of photonic resonant structures, such as microresonators and photonic crystals. The Q factor of the resonances ranges from a few hundreds to tens of thousands, and yet in the end, none of the resonant schemes has found practical applications to this day, due to the fact that resonance is always associated with excessive absorption and dispersion. To this day optical fiber remains the nonlinear medium of choice in which low nonlinear coefficients are more than compensated by the long propagation length and high degree of confinement. All kinds of all-optical switching and frequency conversion techniques have been successfully demonstrated in fiber [81]. The only other media in which all-optical switching has been consistently demonstrated is semiconductor optical amplifier (SOA) in which the loss simply does not exist due to optical gain. Neither fiber nor SOA relies upon any resonance despite its apparent appeal—one always tries to avoid loss and excessive dispersion.

So if the numerous relatively high-Q resonant schemes for enhancing optical nonlinearity have failed to achieve practicality, it would have been naive to expect plasmonic resonance in metal nanoparticles with Q barely of the order of 10 to succeed where so many have failed. Thus in retrospect this work only confirms the obvious. And yet this obvious fact has not been universally accepted by the community, and we hope that our effort has been useful as it has revealed the nature and limitations of the plasmonic enhancement of $\chi^{(3)}$ in great detail and without reliance on excessive numerical modeling.

We emphasize that it has not been our purpose to make broad predictions of where the research in nonlinear plasmonics research may go in the future; our modest goal was to provide a set of simple expressions and relevant numbers for others so they can ascertain the prospects for using nonlinear plasmonic metamaterials in their own applications. Still we may attempt to make a very general statement, that plasmonically enhanced structures in nonlinear optics might not find too many applications requiring decent efficiency, such as switching, wavelength conversion, etc., but may be of use in such applications where efficiency is not much of an issue such as sensing, as well as in fundamental studies of optical properties of different materials under extremely high fields.

[1] T. H. Maiman, Nature **187**, 493 (1960).
[2] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, Phys. Rev. Lett. **7**, 118 (1961).
[3] P. D. Maker, R. W. Terhune, M. Nisenhoff, and C. M. Savage, Phys. Rev. Lett. **8**, 21 (1962).
[4] J. A. Giordmaine, Phys. Rev. Lett. **8**, 19 (1962).
[5] J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, Phys. Rev. **127**, 1918 (1962).
[6] N. Bloembergen and P. S. Pershan, Phys. Rev. **128**, 606 (1962).
[7] N. Bloembergen and Y. R. Shen, Phys. Rev. **133**, A37 (1964).
[8] Y. R. Shen, *Principles of Nonlinear Optics* (Wiley, New York, 1984).
[9] R. W. Boyd, *Nonlinear Optics* (Academic Press, New York, 1992).
[10] L. B. Fu, M. Rochette, V. G. Ta'eed, D. J. Moss, and B. J. Eggleton, Opt. Express **13**, 7637 (2005).
[11] S. X. Qian, J. B. Snow, H. M. Tzeng, and R. K. Chang, Science **231**, 486 (1986).
[12] H. B. Lin and A. J. Campillo, Phys. Rev. Lett. **73**, 2440 (1994).
[13] J. E. Heebner and R. W. Boyd, Opt. Lett. **24**, 847 (1999).
[14] V. Berger, Phys. Rev. Lett. **81**, 4136 (1998).
[15] M. Bajcsy, S. Hofferberth, V. Balic, T. Peyronel, M. Hafezi, A. S. Zibrov, V. Vuletic, and M. D. Lukin, Phys. Rev. Lett. **102**, 203902 (2009).
[16] C. Monat, M. de Sterke, and B. J. Eggleton, J. Opt. **12**, 104003 (2010).
[17] J. B. Khuhin, J. Opt. Soc. Am. B **22**, 1062 (2005).
[18] R. W. Hellwarth, in *Advances in Quantum Electronics*, edited by R. Singer (Columbia University Press, New York, 1961), p. 334.
[19] L. Hargrove, R. L. Fork, and R. L. Pollock, Appl. Phys. Lett. **5**, 4 (1964).
[20] A. J. De Maria, D. A. Stetson, and H. Heyma, Appl. Phys. Lett. **8**, 22 (1966).
[21] Y. Hamachi, S. Kubo, and T. Baba, Opt. Lett. **34**, 1072 (2009).
[22] J. Khurgin, Opt. Lett. **30**, 643 (2005).
[23] M. I. Stockman, Opt. Express **19**, 22029 (2011).
[24] S. Kühn, U. Håkanson, L. Rogobete, and V. Sandoghdar, Phys. Rev. Lett. **97**, 017402 (2006).
[25] P. Bharadwaj and L. Novotny, Opt. Express **15**, 14266 (2007).
[26] M. Moskovits, L. Tay, J. Yang, and T. Haslett, Top. Appl. Phys. **82**, 215 (2002).
[27] K. Kneipp, Y. Wang, H. Kneipp, L. T. Perelman, I. Itzkan, R. R. Dasari, and M. S. Feld, Phys. Rev. Lett. **78**, 1667 (1997).
[28] S. Nie and S. R. Emory, Science **275**, 1102 (1997).
[29] G. Sun, J. B. Khurgin, and R. A. Soref, Appl. Phys. Lett. **94**, 101103 (2009).
[30] G. Sun, J. B. Khurgin, and A. Bratkovsky, Phys. Rev. B **84**, 045415 (2011).
[31] G. Sun and J. B. Khurgin, Appl. Phys. Lett. **98**, 113116 (2011).
[32] J. B. Khurgin and G. Sun, Appl. Phys. Lett. **99**, 211106 (2011).
[33] J. B. Khurgin and G. Sun, Appl. Phys. Lett. **100**, 011105 (2012).
[34] J. B. Khurgin, G. Sun, and R. A. Soref, Appl. Phys. Lett. **93**, 021120 (2008).
[35] J. B. Khurgin, G. Sun, and R. A. Soref, Appl. Phys. Lett. **94**, 071103 (2009).
[36] G. Sun and J. B. Khurgin, Phys. Rev. A **85**, 063410 (2012).
[37] K. Okamoto, I. Niki, and A. Scherer, Appl. Phys. Lett. **87**, 071102 (2005).

[38] S. Pillai, K. R. Catchpole, T. Trupke, and M. A. Green, J. Appl. Phys. **101**, 093105 (2007).
[39] S. C. Lee, S. Krishna, and S. R. J. Brueck, Opt. Express **17**, 23160 (2009).
[40] M. B. Dühring, N. A. Mortensen, and O. Sigmund, Appl. Phys. Lett. **100**, 211914 (2012).
[41] M. Fleischmann, P. J. Hendra, and A. J. McQuillan, Chem. Phys. Lett. **26**, 163 (1974).
[42] D. A. Weitz, S. Garoff, J. I. Gersten, and A. Nitzan, J. Chem. Phys. **78**, 5324 (1983).
[43] M. Moskovits, Rev. Mod. Phys. **57**, 783 (1985).
[44] S. I. Anisimov, B. L. Kapeliovich, and T. L. Perelman, Sov. Phys. JETP **39**, 375 (1974).
[45] C. K. Chen, A. R. B. de Castro, and Y. R. Shen, Phys. Rev. Lett. **46**, 145 (1981).
[46] A. Wokaun, J. G. Bergman, J. P. Heritage, A. M. Glass, P. F. Liao, and D. H. Olson, Phys. Rev. B **24**, 849 (1981).
[47] M. Kauranen and A. V. Zayats, Nat. Photonics **6**, 737 (2012).
[48] A. V. Zayats, I. I. Smolyaninov, and A. A. Maradudin, Phys. Rep. **408**, 131 (2005).
[49] B. Sharma, R. R. Frontiera, A. Henry, E. Ringe, and R. P. van Duyne, Mater. Today **15**, 16 (2012).
[50] I. I. Smolyaninov, A. V. Zayats, and C. C. Davis, Phys. Rev. B **56**, 9290 (1997).
[51] S. I. Bozhevolnyi, J. Beermann, and V. Coello, Phys. Rev. Lett. **90**, 197403 (2003).
[52] C. Anceau, S. Brasselet, J. Zyss, and P. Gadenne, Opt. Lett. **28**, 713 (2003).
[53] J. L. Coutaz, M. Nevière, E. Pic, and R. Reinisch, Phys. Rev. B **32**, 2227 (1985).
[54] S. Linden, F. B. P. Niesler, J. Förstner, Y. Grynko, T. Meier, and M. Wegener, Phys. Rev. Lett. **109**, 015502 (2012).
[55] M. W. Klein, C. Enkrich, M. Wegener, and S. Linden, Science **313**, 502 (2006).
[56] N. Feth, S. Linden, M. W. Klein, M. Decker, F. B. P. Niesler, Y. Zeng, W. Hoyer, J. Liu, S. W. Koch, J. V. Moloney, and M. Wegener, Opt. Lett. **33**, 1975 (2008).
[57] M. D. McMahon, R. Lopez, R. F. Haglund, Jr., E. A. Ray, and P. H. Bunton, Phys. Rev. B **73**, 041401(R) (2006).
[58] T. Xu, X. Jiao, G. P. Zhang, and S. Blair, Opt. Express **15**, 13894 (2007).
[59] A. Lesuffleur, L. K. S. Kumar, and R. Gordon, Appl. Phys. Lett. **88**, 261104 (2006).
[60] J. Homola, Chem. Rev. **108**, 462 (2008).
[61] S. Link and M. A. El-Sayed, J. Phys. Chem. B **103**, 8410 (1999).
[62] H. Baida, D. Mongin, D. Christofilos, G. Bachelier, A. Crut, P. Maioli, N. Del Fatti, and F. Vallée, Phys. Rev. Lett. **107**, 057402 (2011).
[63] M. Abb, P. Albella, J. Aizpurua, and O. L. Muskens, Nano Lett. **11**, 2457 (2011).
[64] I. I. Smolyaninov, A. V. Zayats, A. Gungor, and C. C. Davis, Phys. Rev. Lett. **88**, 187402 (2002).
[65] A. V. Krasavin and N. I. Zheludev, Appl. Phys. Lett. **84**, 1416 (2004).
[66] D. Pacifici, H. J. Lezec, and H. A. Atwater, Nat. Photonics **1**, 402 (2007).
[67] A. V. Krasavin, T. P. Vo, W. Dickson, P. M. Bolger, and A. V. Zayats, Nano Lett. **11**, 2231 (2011).
[68] K. F. MacDonald, Z. L. Samson, M. I. Stockman, and M. I. Zheludev, Nat. Photonics **3**, 55 (2009).
[69] A. V. Krasavin, S. Randhawa, J.-S. Bouillard, J. Renger, R. Quidant, and A. V. Zayats, Opt. Express **19**, 25222 (2011).
[70] E. Prodan, C. Radloff, N. J. Halas, and P. Nordlander, Science **302**, 419 (2003).
[71] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999), p. 158.
[72] P. B. Johnson and R. W. Christy, Phys. Rev. B **6**, 4370 (1972).
[73] J. Khurgin, Adv. Opt. Photon. **2**, 287 (2010).
[74] S. Wu, X. C. Zhang, and R. L. Fork, Appl. Phys. Lett. **61**, 919 (1992).
[75] G. Sun and J. B. Khurgin, Appl. Phys. Lett. **97**, 263110 (2010).
[76] G. Sun and J. B. Khurgin, Appl. Phys. Lett. **98**, 153115 (2011).
[77] S. Toroghi and P. G. Kik, Appl. Phys. Lett. **101**, 013116 (2012).
[78] S. Toroghi and P. G. Kik, Phys. Rev. B **85**, 045432 (2012).
[79] J. Renger, R. Quidant, N. van Hulst, and L. Novotny, Phys. Rev. Lett. **104**, 046803 (2010).
[80] S. Kim, J. H. Jin, Y. J. Kim, I. Y. Park, Y. Kim, and S. W. Kim, Nature **453**, 757 (2008).
[81] G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed. (Academic Press, New York, 2001).