


Graphene as a source of entangled plasmons

Zhiyuan Sun^{1,2,3}, D. N. Basov,² and M. M. Fogler¹

¹*Department of Physics, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA*

²*Department of Physics, Columbia University, 538 West 120th Street, New York, New York 10027, USA*

³*Department of Physics, Harvard University, 17 Oxford Street, Cambridge, Massachusetts 02138, USA*

 (Received 11 November 2021; accepted 28 April 2022; published 13 June 2022; corrected 21 November 2022)

We analyze nonlinear optics schemes for generating pairs of quantum entangled plasmons in the terahertz-infrared range in graphene. We predict that high plasmonic field concentration and strong optical nonlinearity of monolayer graphene enables pair-generation rates much higher than those of conventional photonic sources. The first scheme we study is spontaneous parametric down conversion in a graphene nanoribbon. In this second-order nonlinear process a plasmon excited by an external pump splits into a pair of plasmons, of half the original frequency each, emitted in opposite directions. The conversion is activated by applying a dc electric field that induces a density gradient or a current across the ribbon. Another scheme is degenerate four-wave mixing where the counter-propagating plasmons are emitted at the pump frequency. This third-order nonlinear process does not require a symmetry-breaking dc field. We suggest nano-optical experiments for measuring position-momentum entanglement of the emitted plasmon pairs. We estimate the critical pump fields at which the plasmon generation rates exceed their dissipation, leading to parametric instabilities.

DOI: [10.1103/PhysRevResearch.4.023208](https://doi.org/10.1103/PhysRevResearch.4.023208)

I. INTRODUCTION

Graphene plasmonics has emerged as a platform for realizing strong light-matter interaction on ultrasmall length scales [1,2]. Launching, detection, and manipulation of plasmons by near-field probes has been demonstrated at infrared/THz frequencies [3–14]. High concentration of electric field, a key factor for nonlinear optical phenomena, has been achieved by exciting plasmons of wavelength λ_p as small as few hundred nanometers, about two orders of magnitude shorter than the vacuum photon wavelength λ_v . Plasmonic quality factors up to $Q \sim 130$ have been reached [9] in encapsulated graphene structures [5,7], fulfilling another condition—low dissipation—necessary for prominent nonlinear effects.

Numerous theoretical [15–24] and a few experimental [25,26] papers have explored manifestations of nonlinear coupling of light to graphene in the context of conventional far-field optics. However, both nonlinear [23,27–30] and quantum [28,31,32] effects could be more pronounced in the near-field domain because of plasmonic field concentration. Plasmon interaction phenomena we study in this paper include the parametric down conversion (PDC) and four-wave mixing (FWM) [33]. Spontaneous PDC and FWM can generate entangled pairs of plasmons, similar to how the usual photonic spontaneous PDC produces entangled pairs of photons [34–41].

We propose that a graphene ribbon containing an electrostatically induced p - n junction [5,42–46], see Fig. 1(a), can be a highly efficient PDC source. The plasmon spectrum of such a system, which we will call the device of type A, consists of multiple continuously dispersing subbands. The lowest subband (labeled “0”) is gapless and the next one (labeled “1”) is gapped due to transverse confinement in the ribbon, see Fig. 1(b). In the simplest picture, the PDC process is splitting of a mode-1 plasmon of frequency ω_1 and momentum $q_1 = 0$ into two mode-0 plasmons of frequency $\omega_0 = \omega_1/2$ and momenta $\pm q_0$ that propagate away in opposite directions along the ribbon. The energy-momentum conservation (phase matching) in this conversion is guaranteed without any fine tuning of the device. This advantage of the counterpropagating PDC scheme has been previously exploited in photonic waveguides [36,39,47,48]. In a more precise description, frequencies ω_0 and so the momenta q of the outgoing plasmons have some uncertainty. The spontaneous PDC generates a superposition of states of different momentum:

$$|\psi\rangle = \int dq f(q) | -q \rangle | +q \rangle \quad (1)$$

where $| -q \rangle | +q \rangle$ means a two plasmon state with one plasmon at momentum $-q$ and the other at $+q$, see Sec. II D. This superposition is not factorizable, which implies that the pairs are position-momentum entangled, similar to particles in the original Einstein, Podolsky, and Rosen (EPR) paper [34]. The amplitude function $f(q)$ in Eq. (1) is peaked at q_0 and has a characteristic width $\delta q \sim \gamma/v_0$ where

$$\gamma = \omega_0/Q \quad (2)$$

is the plasmon damping rate [9] and $v_0 = d\omega_0/dq_0$ is the mode-0 plasmon group velocity.

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

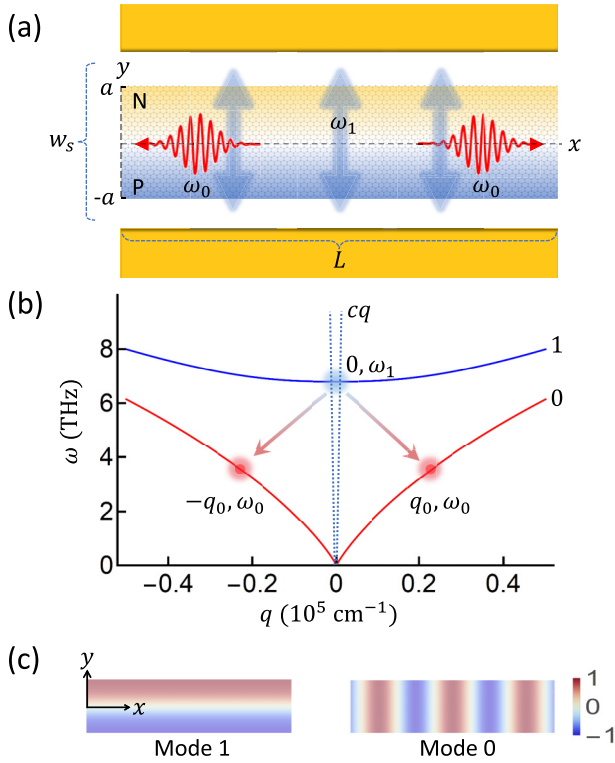


FIG. 1. (a) Type-A device: A graphene ribbon with a lateral p - n junction (blue/gold denoting electron/hole doping) induced by a split gate (golden bars). A pump field of frequency ω_1 applied in the \hat{y} direction (thick arrows) generates plasmon pairs of frequency ω_0 propagating in the $\pm\hat{x}$ direction (red-undulating arrows). (b) Plasmon dispersion in a ribbon of half-width $a = 200$ nm with the p - n junction on the midline and the carrier density $n = \pm 10^{12}$ cm^{-2} at the edges. The dotted lines form the “light cone” for the dispersion of vacuum photons. The blue dot represents the mode-1 plasmon pumped by the incident photon, which then decays into two mode-0 plasmons (the two red dots) as shown by the arrows. (c) The spatial distribution of the amplitude of the quasistatic electric potential of the two modes.

The role of the split-gate represented by the golden bars in Fig. 1(a) is twofold. First, it induces a gradient of carrier concentration in graphene when a dc voltage is applied between the two parts of the gate. The p - n junction is created when this voltage is high enough [44]. Such a system lacks inversion symmetry and therefore the PDC is allowed. Second, the split-gate serves as an optical antenna amplifying the incident pump field [44,49] that excites mode-1 plasmons. The coupling is facilitated by a large transverse dipole moment of the mode-1 plasmons and a strong field enhancement $\lambda_v/\lambda_p \gg 1$ of the antenna.

Note that the inversion symmetry can also be broken by a dc current across a uniformly doped ribbon, which we refer to as the device of type B. In contrast, the FWM is a third-order nonlinear effect, which does not require symmetry breaking fields. The degenerate FWM, in which the two pump photons of frequency ω are converted into a pair entangled plasmons of the same frequency, can take place in either a nanoribbon or a more typical larger-area graphene sheet, which are devices of type C.

Plasmonic effects similar to those we explore here have been considered in a few recent theoretical papers. For example, plasmonic sum-frequency generation in graphene nanoflakes [17] is based on the nonlinear mechanism similar to that underlying our plasmonic PDC. However, since the plasmon spectrum of a flake is discrete, only flakes of certain shapes can fulfill the energy-momentum conservation constraint for the sum (or difference) frequency generation. As mentioned above, in a long ribbon such kinematic constraints can be met without delicate fine tuning. Another related paper studied photon-plasmon difference frequency generation [26,50] and entanglement [51], which is a phenomenon intermediate between the usual all-photon PDC [34] and our fully plasmonic PDC.

The remainder of the paper is organized as follows. In Sec. II, we summarize our main results for the type-A device. In the last part of that section we discuss possible experiments that can probe the plasmon entanglement. In Sec. III we consider the third-order nonlinear effects. The FWM relevant for the operation of type-C devices is analyzed in Sec. III B and the current induced PDC in type-B devices is examined in Sec. III C. Section IV contains discussion and outlook. Finally, the summary of the notations and some details of the derivations are given in the Appendix.

II. PLASMONIC PDC IN A GRAPHENE RIBBON

To simplify the analysis of a type-A device, we assume that the edges of the ribbon $y = \pm a$ are not too close to those of the split-gate, so that the dc electric field created by the gate is approximately constant across the ribbon. The induced density response of graphene can be modeled [42] assuming the local chemical potential is linear in y :

$$\mu(y) = \frac{a+y}{2a}\mu_t + \frac{a-y}{2a}\mu_b. \quad (3)$$

If the chemical potentials μ_t and μ_b at the two edges of the ribbon are opposite in sign, a p - n junction forms. If $\mu_t = -\mu_b$, the junction is located on the midline $y = 0$ of the ribbon.

The plasmon frequency dispersion computed numerically as a function of momentum q along the strip is shown in Fig. 1(b) (For related analytical results, see Ref. [52]). The lowest-frequency branch, which we refer to as mode 0, is gapless:

$$\omega(q) = \xi_v \sqrt{\alpha_g a k_F} v_F q = \xi_v v_a q. \quad (4)$$

Here

$$v_a = \sqrt{(e^2/\hbar)v_F k_F a} \sim v_F \sqrt{k_F a}, \quad (5)$$

is a characteristic scale of the plasmon velocity, $v_F \approx c/300$ is the Fermi velocity, $\alpha_g = e^2/(\hbar v_F)$ is the dimensionless strength of the Coulomb interaction, and k_F denotes the Fermi momentum corresponding to the maximum absolute chemical potential $\max(|\mu_t|, |\mu_b|)$ in the ribbon. Dimensionless parameter $\xi_v \sim 1$ is a slow (logarithmic) function of q . It also depends on the doping profile of the ribbon. The dispersion law (4) is similar to that of one dimensional (1D) plasmons [6,10,13].

The first gapped mode, mode 1, can be viewed as a linear combination of 2D plasmons of momenta $q\hat{x} \pm q_y\hat{y}$ with

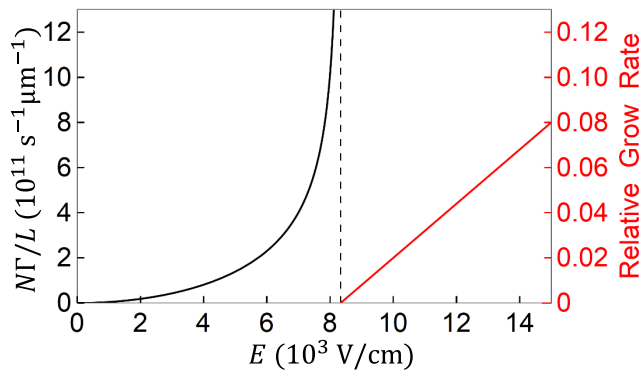


FIG. 2. Transition from stable to unstable regime of plasmon-pair generation in the type-A device as pump field is increased. The black curve is the rate of generation of plasmon pairs per unit length. The red curve is the relative growth rate $(\kappa - \gamma)/\omega_0$ in the unstable regime. The plasmon damping rate is assumed to be $\gamma = 0.1\omega_0$. Other parameters are from Fig. 1 and the temperature is $T = 300$ K.

$q_y \sim 1/a$. The frequency of this mode at $q = 0$ is

$$\omega_1 = \xi_1 \sqrt{2\alpha_g v_F^2 k_F / a}, \quad (6)$$

where $\xi_1 \sim 1$ is another dimensionless shape factor. The electric potential profile of the ω_1 mode [Fig. 1(c)] indicates that mode-1 excitations have a nonzero dipole moment in the y -direction. Thus, they can be resonantly excited by a y -polarized pump field [53].

In this paper we describe the mode-1 to mode-0 PDC in the quantum language. However, there is a complementary classical picture, which is as follows. Plasmonic oscillations of mode-1 modulate the total carrier density of the ribbon viewed as a one dimensional conductor. This in turn modulates the frequencies of mode-0 plasmons and causes their parametric excitation. In the standard theory of optical parametric amplification [33] (see also Appendix B), one of the generated mode-0 plasmon is referred to as the idler and the other as the signal. The idler is assumed to have a finite spectral power, e.g., due to equilibrium fluctuations in the beginning of the process. The signal is then produced from the pump and the idler by the difference frequency generation (DFG).

For weak pump fields E , the pair-generation rate grows linearly with the pump power, see Fig. 2. As E becomes stronger, a parametric instability is reached at some critical E . At even stronger pump field, the signal exhibits an exponential growth in the undepleted pump approximation. In Sec. II A below we evaluate the dependence of the pair-generation rate and the critical pump field on experimental parameters, such as the ribbon width and its doping profile.

A. PDC Hamiltonian

The second-order nonlinear conductivity [16,23,51,54] leads to interaction between the plasmons of modes 0 and 1. This interaction can be modeled by the

Hamiltonian $H = H_0 + H^{(2)}$:

$$H_0 = \hbar\omega_1(a_1^\dagger a_1 + 1/2) + \sum_q \hbar\omega(q)(a_q^\dagger a_q + 1/2),$$

$$H^{(2)} = \sum_q g_q^{(2)} a_1 a_q^\dagger a_{-q}^\dagger + \text{c.c.}, \quad (7)$$

where a_1 , a_q are the annihilation operators for mode-1 at zero momentum and mode-0 of momentum q , respectively. The interaction Hamiltonian $H^{(2)}$ leads to pair generation in the weak coupling regime ($g^{(2)} a_1 Q / \omega \ll 1$) and two-mode squeezing in the strong coupling regime ($g^{(2)} a_1 Q / \omega > 1$), as shown in Fig. 2. Since we study the PDC on resonance, $\omega(q_0) = \omega_0 = \omega_1/2$, we can treat the interaction strength $g^{(2)}$ as a constant, $g^{(2)} = g_{q_0}^{(2)}$. As shown in Appendix C,

$$g^{(2)} = \xi_g \frac{\pi^{1/2} e^{1/2} \hbar^{3/4} v_F^{3/4}}{2^{13/4} k_F^{5/4} a^{7/4} L^{1/2}}, \quad (8)$$

where L is the length of the gated part of the ribbon (Fig. 1) and dimensionless factor $0 \leq \xi_g \leq 1$ depends on the carrier density profile across the ribbon.

The Hamiltonian $H^{(2)}$ governs the spontaneous decay of the mode-1 plasmon into a pair of mode-0 plasmons with momenta q and $-q$. In the weak-coupling regime, the decay rate Γ can be calculated using Fermi's golden rule (see Appendix C5)

$$\Gamma = \frac{|g^{(2)}|^2 L}{\hbar^2 v_0} \left(n_0 + \frac{1}{2} \right), \quad (9)$$

where n_0 is the occupation number of mode-0 plasmons of frequency $\omega \simeq \omega_0$; in thermal equilibrium, $n_0 = 1/(e^{\hbar\omega_0/T} - 1)$ at temperature T . Normalizing Γ to the plasmon frequency ω_0 , we obtain the dimensionless decay rate of mode-1 plasmons:

$$\frac{\Gamma}{\omega_1} = \xi_{\text{df}} \frac{\Gamma_0}{\omega_1}, \quad \frac{\Gamma_0}{\omega_1} = \frac{\pi}{2^8} \frac{1}{\sqrt{\alpha_g}} \frac{1}{(k_F a)^{7/2}}, \quad (10)$$

where $\xi_{\text{df}} \sim 1$ is another dimensionless factor. The dependence of ξ_{df} on the carrier density profile across the ribbon is plotted in Fig. 3. For a ribbon with the carrier density $n = \pm 10^{12} \text{ cm}^{-2}$ at the edges and a half-width of $a = 200 \text{ nm}$, we find $\omega_1/(2\pi) \approx 5 \text{ THz}$ and $\Gamma_0/\omega_1 \sim 10^{-8}$. Therefore, due to PDC alone a mode-1 plasmon decays into a mode-0 plasmon pair every $\Gamma_0^{-1} \sim 10^{-5} \text{ s}$. If the plasmon Q factor due to other damping channels (phonon and impurity scattering [9]) is $Q = \omega_0/\gamma \sim 10$, the PDC efficiency is $\Gamma_0/\gamma \sim 10^{-7}$, which is much higher than in available photonic PDC devices [39,55]. Furthermore, this efficiency scales as a^{-4} , and so can be higher still in a narrower strip.

B. Pumping mode-1

In a uniform external ac electric field E in the y direction, the gapped mode-1, as a harmonic oscillator, is driven to a coherent state $|\beta\rangle = e^{\beta a_1^\dagger - (\beta^2/2)} |0\rangle$. Specifically, the magnitude of the charge density profile of the mode is related to the external pump field as

$$\rho = \frac{\frac{\sigma(\omega)}{\omega} \frac{\omega_1}{\sigma(\omega_1)}}{\left(1 - \frac{\sigma(\omega)}{\omega} \frac{\omega_1}{\sigma(\omega_1)}\right)} \frac{\alpha}{S} E \quad (11)$$

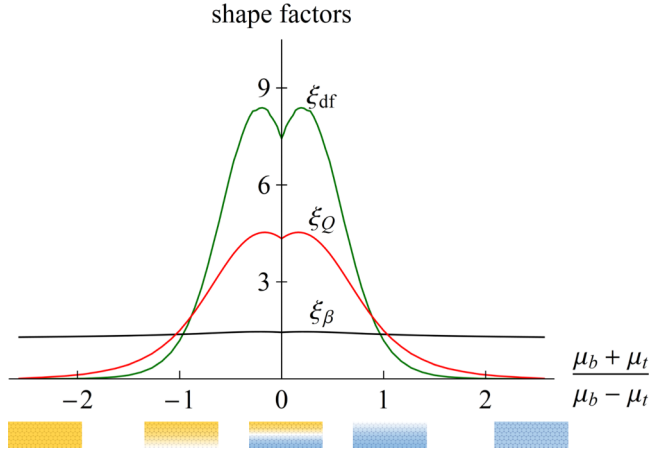


FIG. 3. The shape factors for type-A device as functions of the doping profile parameter $s_d = (\mu_b + \mu_t)/(\mu_b - \mu_t)$ at temperature $T = 300$ K. The green, black, and red curves are the decay rate shape factor ξ_{df} , the occupation number shape factor ξ_β and the shape factor ξ_Q for the relative subharmonic growth rate (defined in Appendix A). Each colored stripe below the horizontal axis illustrates the corresponding graphene ribbon with blue/gold representing electron/hole doping.

where $\sigma(\omega)$ is the two-dimensional optical conductivity and α/S is an $O(1)$ constant determined by the doping profile of the ribbon, see Appendix C1. On resonance, from its relation to ρ (see Appendices C2 and C4), β is found to be

$$\beta = \xi_\beta \beta_0, \quad \beta_0 = \frac{2^{\frac{3}{4}}}{\pi^{\frac{1}{2}}} Q \sqrt{\frac{L}{a}} \frac{eaE}{\left(\frac{e^2}{a}\right)^{\frac{3}{4}} \varepsilon_F^{\frac{1}{4}}} \quad (12)$$

where $\xi_\beta \sim 1$ is a shape factor shown in Fig. 3. The occupation of the gapped mode modifies the pair-generation rate to

$$R = N\Gamma \propto E^2 a^{-3/2} L, \quad N = \beta^2 \propto E^2 a^{5/2} L. \quad (13)$$

For $n = 10^{12} \text{ cm}^{-2}$, $a = 200 \text{ nm}$, $L = 1 \text{ }\mu\text{m}$, $E = 10^3 \text{ V/cm}$, and a plasmon quality factor of $Q \sim 10$ for ultraclean samples at room temperature [5,7,9,56], we have $\beta_0 = 15$, indicating an average occupation number of $N \sim 400$, thus boosting the generation rate to $R \sim 4 \times 10^9 \text{ s}^{-1}$. It means that in a nanosecond-long pulse, there are four pairs of subharmonic plasmons generated while there are $\sim 10^6$ photons incident onto the ribbon area. This rate is higher than conventional photonic [39] and quantum dot [38] devices.

As the pump field increases, Fermi's golden rule ceases to be valid. The pair-generation rate can be derived from the classical parametric oscillator theory (Appendix B)

$$R = \frac{N\Gamma}{\sqrt{1 - (\kappa/\gamma)}}, \quad \kappa = \frac{\beta g^{(2)}}{\hbar}. \quad (14)$$

The growth rate κ is to be discussed in the next section. The pair-generation rate diverges at the instability threshold $\kappa = \gamma$, as shown in Fig. 2.

C. Two mode squeezing and parametric amplification

The coupling Hamiltonian Eq. (7) in the interaction picture reads

$$H_I = \sum_q g_q^{(2)} e^{-i\delta t} a_1 a_q^\dagger a_{-q}^\dagger + \text{c.c.},$$

where $\delta = \omega_1 - 2\omega_q$ is the frequency mismatch. We focus on the pair of acoustic modes $\omega(q_1) = \omega_0 = \omega_1/2$ for which H_I is on resonance:

$$H_I = g^{(2)} a_1 a_q^\dagger a_{-q}^\dagger + \text{c.c.} \quad (15)$$

When mode-1 is pumped into a coherent state, one can replace a_1 by its classical value β . This Hamiltonian generates a two-mode-squeezing time evolution operator [57]

$$S(\beta g^{(2)} t/\hbar) = e^{-iH_I t} = e^{-i(g^{(2)} \beta a_q^\dagger a_{-q}^\dagger + \text{c.c.})t/\hbar}, \quad (16)$$

which squeezes the pair of mode-0 plasmons at momenta $(q, -q)$. Therefore, the plasmons are generated as two-mode squeezed states [34]. The squeezing operator leads to exponential growth of the amplitudes of the observables (e.g., the electric field of the mode) with the growth rate $\kappa = g^{(2)} \beta/\hbar$. We define the dimensionless relative growth rate

$$\frac{\kappa}{\omega_0} = \frac{1}{4} \xi_Q Q \frac{\zeta}{\alpha_g}, \quad \zeta = \frac{eE}{\varepsilon_F k_F} \quad (17)$$

that is made only of intensive quantities. Here ζ is the dimensionless small parameter that controls the second-order nonlinear effects of graphene plasmons. Note that the relative growth rate depends only on the maximum doping level, the pump electric field, and the dimensionless shape factor ξ_Q . Parameter $\xi_Q \sim 1$, plotted in Fig. 3, depends only on the scale invariant doping profile of the ribbon. The pump induced growth rate κ reduces the plasmon damping rate from γ to $\gamma - \kappa$, thereby further enhancing the pair-generation rate. The system reaches parametric instability at $\kappa = \gamma$, see Fig. 2. Beyond this threshold, the plasmons are unstable with exponential growth rate $\kappa - \gamma$. When amplifying a classical source as the seed, this effect is known as parametric amplification [33]. For $n = 10^{12} \text{ cm}^{-2}$ and $E = 10^4 \text{ V/cm}$, we have $\kappa = 0.12\omega_0$ in the p - n junction regime, high enough to compete with the damping rate γ .

The split-gate structure, as an antenna with width in y larger than the vacuum wave length λ_v of the incident light, could enhance the pump field E_v to the total field E by a factor of roughly $F = E/E_v = \lambda_v/(2\pi w_s)$ [44,49] where w_s is the distance separating the two halves of the gate, see Fig. 1. For the typical parameters in Fig. 1, one has $\lambda_v = 43 \text{ }\mu\text{m}$ at $\omega_1 = 7 \text{ THz}$ and w_s could be chosen as $0.5 \text{ }\mu\text{m}$, rendering the field enhancement factor $F \approx 14$. (To increase the working frequency to, e.g., $\omega_1 = 30 \text{ THz}$ corresponding to $\lambda_v = 10 \text{ }\mu\text{m}$, the ribbon width needs to be shrunk to $2a \sim 20 \text{ nm}$.) Therefore, to achieve the plasmon instability regime in Fig. 2, the actual incident field just needs to be of the order of $E_v = 10^2\text{--}10^3 \text{ V/cm}$. We note that the antenna also amplifies the radiative damping rate γ_R of the dipolar active modes (e.g., mode-1) of the ribbon by the same factor F . For ribbons of size much smaller than the vacuum wavelength, $a, L \ll \lambda_v$, the normalized radiative damping rate $Q_R^{-1} = \gamma_R/\omega_0$ is amplified by the antenna from $Q_R^{-1} \sim a^2 L/\lambda_v^3$ to aL/λ_v^2 , which is

still much smaller than $\gamma/\omega_0 \sim 0.1$ [5,7,9,56] for the parameters given above. Therefore, the intrinsic damping γ (e.g., due to phonons and the electronic system itself) should be the major plasmon dissipation pathway in the system and the effect of the radiative damping can be neglected.

D. Measuring the entanglement

In general, the plasmon pair state can be expressed as

$$|\psi\rangle = \int dp_1 dp_2 f(p_1, p_2) |p_1\rangle |p_2\rangle. \quad (18)$$

To quantify the standard deviations, we approximate the wave function by the Gaussian form

$$f(p_1, p_2) = Z e^{-\frac{(p_1+p_2)^2}{4\Delta_+^2} - \frac{(p_1-p_2-2p_0)^2}{4\Delta_-^2}} \quad (19)$$

where Z is a normalization factor. The EPR state in Eq. (1) corresponds to $\Delta_+ = 0$. After p_1 (or p_2) is measured, the standard deviation of p_2 (or p_1) becomes $\Delta p = (\frac{1}{\Delta_+^2} + \frac{1}{\Delta_-^2})^{-1/2}$. Similarly, after measuring either position, the standard deviation of the unmeasured position becomes $\Delta x = (\Delta_+^2 + \Delta_-^2)^{-1/2}$. Thus we arrive at

$$\Delta x \Delta p = \left(2 + \frac{\Delta_-^2}{\Delta_+^2} + \frac{\Delta_+^2}{\Delta_-^2}\right)^{-1/2} \leq \frac{1}{2}. \quad (20)$$

The apparent violation of the uncertainty principle happens whenever $\Delta_+ \neq \Delta_-$. In the extreme case of Eq. (1), we have $\Delta x \Delta p = 0$.

If momentum is conserved in the type-A device in Fig. 1, we have $p_1 = -p_2$ exactly satisfied. However, the length L of the device limits momentum conservation and leads to $\Delta_+ \sim 1/L$. The damping rate limits the plasmon line width and renders $\Delta_- \sim 1/l$ where $l = v_0/\gamma$ is the propagation length of the mode-0 plasmons. For device satisfying $L \ll l$, we have $\Delta x \Delta p \sim L/l \ll 1/2$. Note that in our case, entanglement is between continuous variables, and so the fidelity of the state can not be defined in the conventional way [58]. Instead, it is more natural to use Eq. (20) to quantify the entanglement. Alternatively, one could use entanglement witnesses or entanglement negativity [35], which are more technically involved. To detect the position-momentum entanglement of a plasmon pair, one can place grating structures on both sides of the device, which would convert the plasmons to far-field photons. Passing through the beam splitters, the photons are read either by single photon detectors [34,59], which have good time precision or those with good energy resolution, as shown in Fig. 4. The coincidence rate measured by the former as a function of path length difference gives a time uncertainty Δt , while the spectrum correlation of the energy detectors gives the energy uncertainty ΔE . The EPR entanglement is manifested in the relation $\Delta t \Delta E = \Delta x \Delta p < 1/2$ [34,60,61]. This “energy time entanglement” detection scheme is the same as that in Ref. [61] with the entangled photon source replaced by the plasmonic type-A device in Fig. 1 or type-B device in Fig. 8. Alternatively, homodyne detection experiments can measure the entanglement property of the squeezed state [34].

In addition, the Franson scheme [40,41,62] is also applicable as shown in Fig. 5. As mentioned in Sec. I, near-field technique based on scanning probes has been successful in

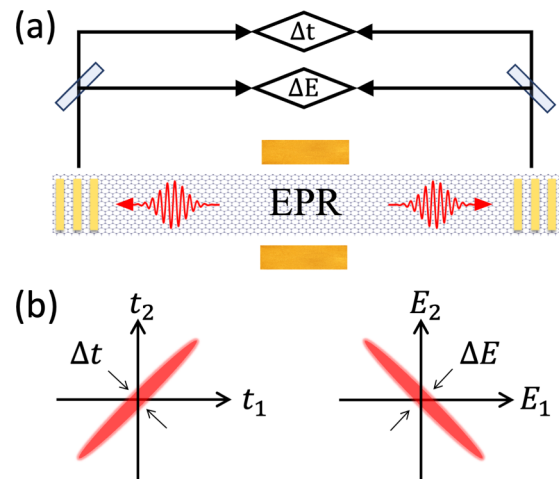


FIG. 4. (a) Schematic of the experiment setup for coincidence measurement of the energy-time entanglement of the generated EPR plasmon pairs. The plasmon pair generator can be either a type-A (Fig. 1) or type-B device (Fig. 8). The vertical yellow bars are the grating structures to convert plasmons into far field photons. (b) Illustration of the time (left) and energy (right) correlations. The widths of the ellipses are the Δt (left) and ΔE (right).

probing plasmons in graphene [2–13]. To implement the Franson scheme, one needs two scanning probes, as shown in Fig. 5(a). The plasmons can either travel to the scanning probes directly or after being reflected by the edges of the ribbon, which we call short and long paths respectively. The coincidence detection arises from interference between the short-short and long-long paths. If the coincidence rate between the signals from two near-field tips is measured as a function of $L_1 \approx L_2$, the distance between each tip and the edge, it will oscillate as $R_c \sim \cos^2((L_1 + L_2)/\lambda)$, like that in Fig. 5(b). However, the condition $L \ll L_i \ll l$ needs to be satisfied. The first inequality guarantees that there are no single photon (or “short-long” path) interference contributions, and the second inequality ensures that plasmon damping is

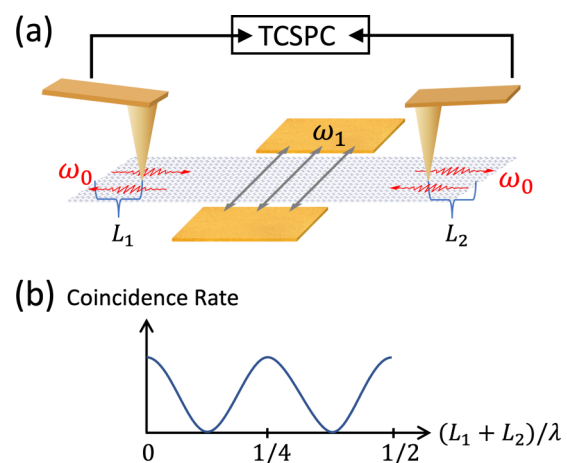


FIG. 5. (a) The Franson scheme of measuring EPR entanglement with two scanning probes as the two interferometers. The plasmon pair generator can be either a type-A (Fig. 1) or type-B device (Fig. 8). (b) The coincidence rate as a function of $(L_1 + L_2)/\lambda$.

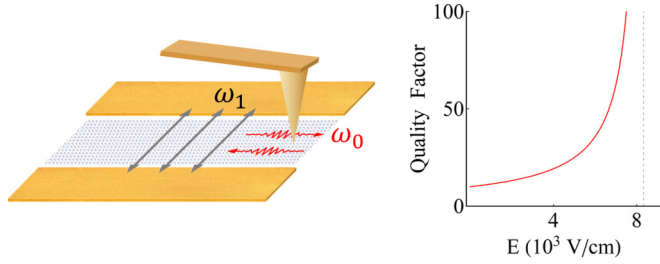


FIG. 6. Near-field imaging of the parametric amplification of the mode-0 plasmons. The right figure is plotted for a type-A device using the parameters in Fig. 1. The parametric plasmon amplifier can also be a type-B device in Fig. 8.

not important. Note that in a prior study on measuring plasmon entanglement [63], the interfering particles were actually photons in optical fibers. In the proposed scheme (Fig. 5) the interference occurs between plasmons. Additionally, our implementation involves highly confined terahertz-infrared plasmons and scanning probes operating on deeply subdiffractional length scales.

One can also probe the classical effects. If the pump field is comparable to the parametric instability threshold, the enhancement of plasmon lifetime can be detected. Upon pumping of the ω_1 mode of the type-A device in Fig. 1 (or type-B in Fig. 8), the ω_0 modes with momenta q and $-q$ are parametrically driven such that their effective damping rate is reduced to $\gamma - \kappa$ (see Appendix B), and their quality factor is boosted to $\omega_0/(\gamma - \kappa)$, as shown in Fig. 6. In a scanning near-field experiment, its manifestation is simply increased propagation length of the plasmons. One can also measure near-field DFG using classical interference between the signal and idler in these experiments, which is discussed in Appendix F.

III. PLASMON GENERATION BY THE THIRD-ORDER NONLINEARITY

In this section we study the effects of third-order nonlinearity [15,19,21,24,64,65] on plasmon interactions, which occurs already in the long wavelength limit. This nonlinearity originates from the linearity of the dispersion of Dirac quasi-particles in graphene, which breaks Galilean invariance [24]. At low frequency $\omega \ll \epsilon_F$, interband effects can be neglected and the third-order nonlinear conductivity $\sigma^{(3)}$ in graphene assumes the third-order ‘‘Drude’’ form

$$\sigma_{ilmn}^{(3)} = \frac{iD^{(3)}}{\omega_1\omega_2\omega_3} \Delta_{ilmn} \quad (21)$$

both in the kinetic [66] and hydrodynamic [24] regimes where ω_i are the frequencies of the three electric fields that generate the third-order current together. The symmetric tensor Δ_{ilmn} is the sum of all isotropic tensors $\Delta_{ilmn} = \delta_{il}\delta_{mn} + \delta_{im}\delta_{ln} + \delta_{in}\delta_{lm}$. Close to zero temperature ($T \ll \mu$), the third-order optical weight is predicted to be $D^{(3)} = \frac{1}{24\pi} \frac{e^4 v_F}{\hbar^3 k_F}$ in the kinetic regime and twice as large in the hydrodynamic regime [24]. In what follows, we assume the system is the kinetic regime since we are mostly interested in the mid infrared plasmons

whose frequencies are larger than the typical electron-electron scattering rate [23,24].

A. FWM Hamiltonian

We start with the plasmonic effective Lagrangian of an electronic system whose linear conductivity is in the dissipation-less Drude form: $\sigma_{jk} = \frac{iD/\pi}{\omega} \delta_{jk}$ where D is the Drude weight. In the near-field approximation (meaning there is only longitudinal electric field and no magnetic field) and in the gauge $A^\mu = (0, \mathbf{A})$ for the electric field where \mathbf{A} is the vector potential viewed as the plasmonic field, the Lagrangian is

$$\begin{aligned} L &= \int d^3r \left(\frac{1}{8\pi c^2} \dot{\mathbf{A}}^2 - \frac{D}{2\pi c^2} \mathbf{A}^2 \delta(z) \right) + L^{(3)} \\ &= \sum_q \left(\frac{1}{8\pi c^2} \frac{2}{|q|} \dot{A}_{-q} \dot{A}_q - \frac{D}{2\pi c^2} A_{-q} A_q \right) + L^{(3)}, \quad (22) \end{aligned}$$

where c is the speed of light. The first term is the electric field energy and the second term has the interpretation of the center of mass kinetic energy of the charge carriers. For two-dimensional (2D) Drude conductors modeled as the x - y plane embedded in 3D space, since the current is localized on the 2D plane, the second term of the top line in Eq. (22) is nonzero only on the plane. In the second line of Eq. (22), $\mathbf{A}_q = \hat{q} A_q$ are the Fourier components of \mathbf{A} evaluated on the 2D plane. The $2/|q|$ comes from integrating over the fields exponentially decaying into the three dimensional (3D) space. Note that the summation is over $\mathbf{q} = (q_x, q_y)$, and instead of three, there is only one plasmon mode for each \mathbf{q} since the field is constrained to be longitudinal. The resulting Euler-Lagrange equation of motion for A_q yields the 2D plasmon dispersion $\omega_q = \sqrt{2Dq}$ where the Drude weight is $D = v_F k_F e^2 / \hbar$ for graphene at zero temperature. Quantization of the plasmonic field

$$A_q = i\hat{q} A_{qu} (a_q + a_{-q}^\dagger), \quad A_{qu}/c = \sqrt{\pi \hbar \omega_q / (2SD)} \quad (23)$$

leads to the Hamiltonian

$$H = H_0 + H^{(3)}, \quad H_0 = \sum_q \hbar \omega_q \left(a_q^\dagger a_q + \frac{1}{2} \right) \quad (24)$$

where S is the area of the sample.

The $H^{(3)} = -L^{(3)}$ contains products of four plasmonic creation and annihilation operators. It describes FWM, which is caused by the third-order optical conductivity $\sigma^{(3)}$. Writing the third-order current $j^{(3)}$ in terms of vector potentials, Eq. (21) indicates

$$j_i^{(3)} = \Pi_{ilmn} A^l A^m A^n / c^3$$

and the interaction Hamiltonian

$$\begin{aligned} H^{(3)} &= -\frac{1}{4c^4} \int d^2r \Pi_{ilmn} A^i A^l A^m A^n \\ &= -\frac{D^{(3)}}{4c^4} \sum \Delta_{ilmn} A_{q_1}^i A_{q_2}^l A_{q_3}^m A_{q_4}^n, \quad (25) \end{aligned}$$

where the kernel $\Pi_{ilmn} = D^{(3)} \Delta_{ilmn}$ is perfectly local in space and time, and the summation runs over all q_i constrained by $q_1 + q_2 + q_3 + q_4 = 0$ due to momentum conservation. Note that although Eq. (25) is a negative ϕ^4 term in this bosonic

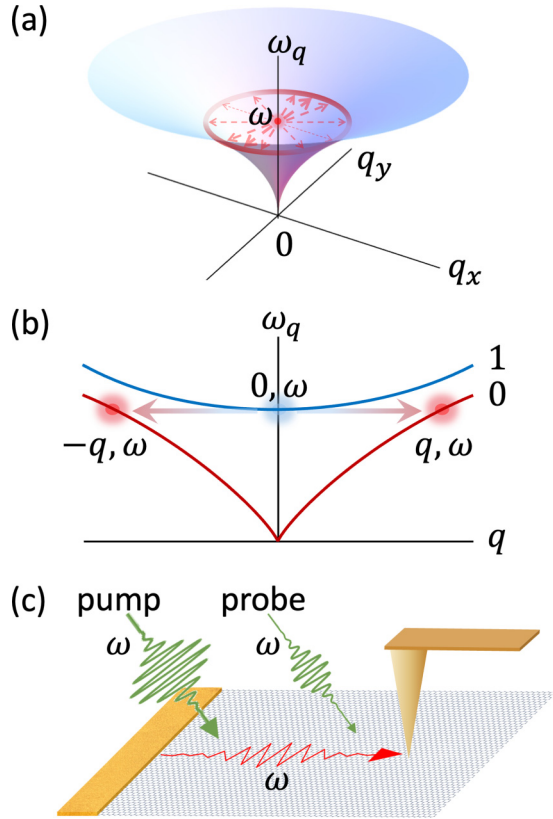


FIG. 7. (a) Generation of entangled plasmon pairs due to spontaneous four wave mixing. The red dot at the center represents the pump. The red circle on the plasmon dispersion curve represents the generated plasmons at the same frequency as the pump but nonzero momenta. The opacity of the circle and thickness of the arrows show the angular dependence of the pair generation intensity. (b) The decay of two mode-1 plasmons into a pair of entangled mode-0 plasmons on a graphene ribbon (same as type-A device in Fig. 1 but without doping gradient) due to spontaneous FWM. (c) The schematic of a near-field measurement of the enhanced plasmon lifetime on a type-C device.

field theory, the system is stable due to higher-order nonlinear couplings.

B. Spontaneous four wave mixing

This subsection discusses spontaneous FWM shown in Fig. 7 in uniformly doped graphene. Equation (25) implies that entangled plasmon pairs with frequency ω can be generated by incident light of the same frequency. We model far field photon incident on the graphene plane by a uniform ac electric field with amplitude $E_0 = \omega A_0$ in \hat{y} direction, and represent it by the vector potential

$$\mathbf{A}_0 = \hat{y}A_0(e^{-i\omega t} + e^{i\omega t})$$

where we have assumed A_0 to be real without loss of generality. When two fields in the product are replaced by the source \mathbf{A}_0 , Eq. (25) becomes

$$H^{(3)} = - \sum_{q_x > 0} g_q^{(3)}(t)(a_q + a_{-q}^\dagger)(a_{-q} + a_q^\dagger) \quad (26)$$

where the pair generation strength is

$$g_q^{(3)}(t) = \frac{1}{8} \hbar \omega_q \xi_\omega^2 (1 + 2 \sin^2 \theta_q) (2 + e^{-2i\omega t} + e^{2i\omega t}) \quad (27)$$

and θ_q is the angle of the plasmon momentum q relative to the \hat{x} axis. We have defined the dimensionless small parameter

$$\xi_\omega = \frac{eE_0/\omega}{\hbar k_F}, \quad (28)$$

which controls the strength of FWM and other third-order effects in graphene. In Eqs. (24) and (26), it is enough to consider terms close to resonance in the interaction picture. The number-conserving terms in Eq. (26) like $a_q a_q^\dagger$ imply the field induced Kerr shift [33] of the plasmon frequency, which are red shifts in graphene due to its negative third-order conductivity [24]. The terms like $e^{-2i\omega t} a_q^\dagger a_{-q}^\dagger$ lead to pair generation in the weak-interaction regime ($g^{(3)}Q/\omega \ll 1$ where $Q = \omega/\gamma$ is the plasmon quality factor) and two mode squeezing (instability) in the strong-interaction regime ($g^{(3)}Q/\omega > 1$).

In the weak-interaction regime, the plasmons with exactly the same frequency as ω are generated as entangled pairs and with an angular distribution of $(1 + 2 \sin^2 \theta_q)^2$, as shown in Fig. 7(a). The pair-generation rate is determined by Fermi's golden rule [analogous to Eq. (9)]:

$$\Gamma = \frac{2\pi}{\hbar} \langle |g^{(3)}|^2 \rangle_\theta \cdot \frac{S}{16\pi} \frac{\omega^3}{D^2} \quad (29)$$

where $\langle \rangle_\theta$ means angular average, the second part comes from the 2D plasmon density of states and we have assumed $\omega \gg T$ so the relevant plasmons are not occupied in equilibrium. The resulting dimensionless generation rate is

$$\frac{\Gamma}{\omega} = \frac{9}{2^{10}} \xi^4 S \frac{\omega^4}{\alpha_g^2 v_F^2 \omega_F^2}. \quad (30)$$

For $n = 10^{12} \text{ cm}^{-2}$, $E_0 = 10^3 \text{ V/cm}$, and $S = 1 \mu\text{m}^2$, we have $\Gamma/\omega = 3.2 \times 10^{-8}$. This leads to superior generation rates compared with conventional FWM sources, as we will discuss in Sec. IV. Since there is no need to break the inversion symmetry, this effect could be observed simply using uniformly doped graphene, which we call the type-C device. One may also use the ribbon in Fig. 1(a) but a lateral p - n junction is no longer necessary. The spontaneous FWM process in a transversely pumped ribbon is illustrated in Fig. 7(b). To distinguish the generated entangled plasmon pairs from the linearly excited plasmons, the key experimental signature would be satisfaction of the EPR criterion [34] in the quadrature measurement in Fig. 4 or the nonlocal interference phenomenon in the Franson experiment in Fig. 5.

In the strong-interaction regime, the plasmons are squeezed in pairs and experience parametric amplification. The resulting relative exponential growth rate is [similar to Eq. (17)]

$$Q_g^{-1} = \frac{|g_q^{(3)}|}{\hbar \omega} = \frac{1}{8} \xi^2 (1 + 2 \sin^2 \theta_q) \quad (31)$$

for the plasmons propagating in θ_q direction, which agrees with the classical result (see Appendix E). In a typical near-field experiment shown in Fig. 7(b), this reduces the plasmon damping rate from γ to $\gamma - g_q^{(3)}/\hbar$. Therefore, to overcome the typical damping rate $\gamma \sim 0.1\omega$ of graphene plasmons, one needs $\xi \sim 0.3$, which requires a pump field of $E_0 \sim 10^5 \text{ V/cm}$ for $n = 10^{12} \text{ cm}^{-2}$ and $\omega = 30 \text{ THz}$.

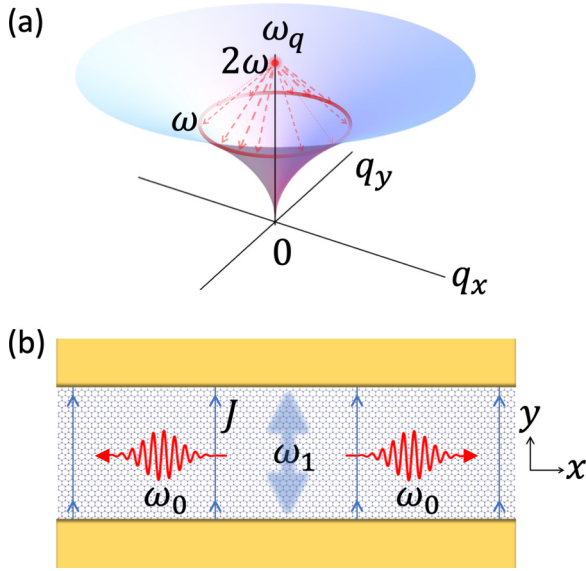


FIG. 8. (a) Parametric down conversion of the pump of frequency 2ω into pairs of subharmonic plasmons of frequency ω in current carrying graphene. (b) Realization in a ribbon (type-B device), which is similar to the type-A device in Fig. 1 except that the density gradient is replaced by a dc current J flowing between the contacts.

C. Spontaneous PDC in current carrying graphene

The third-order nonlinearity can also lead to PDC in a current-carrying graphene, as shown in Fig. 8. This effect can also be understood as FWM with one “pump wave” in Eq. (25) replaced by a dc current flow ($\omega = 0$). If the other pump source has frequency 2ω , then the process looks like a PDC: $2\omega \rightarrow \omega + \omega$. Another point of view is that the dc current with flow velocity u breaks the inversion symmetry, resulting in a nonzero second-order nonlinear conductivity $\sigma^{(2)} \sim \frac{u}{v_F} \frac{e^3 v_F}{\hbar^2} \frac{1}{\omega^2}$. To leading order in the dc current $J = enu$, the equations from Sec. III B can be directly applied to the present case. We assume that the pump field is polarized parallel to the dc current, such that the PDC Hamiltonian is the same as Eqs. (26) and (27) but with ξ_ω^2 replaced by $\xi_{2\omega}\xi_0$ where $\xi_0 = eE_{dc}/(\gamma_{dc}\hbar k_F) \approx u/v_F$ is the dimensionless parameter due to the dc electric field E_{dc} that drives the dc current.

In the weak pump regime, entangled subharmonic plasmon pairs at frequency ω are generated with the angular distribution of $(1 + 2\sin^2\theta_q)^2$ in an infinite graphene sheet, same as that of spontaneous FWM. Summed over all angles, the total relative generation rate is the same as Eq. (30) but with ξ^4 replaced by $\xi_{2\omega}^2\xi_0^2$. To measure the entanglement property of the emitted pairs, one can use the graphene ribbon in Fig. 8(b) as a plasmon pair source to perform either the energy-time entanglement measurement in Fig. 4 or the Franson scheme in Fig. 5. In the ribbon, the generated subharmonic pairs are not in every direction, but can only propagate along \hat{x} [similar to Fig. 1(b)] with the generation rate described by roughly the same formula as that for an infinite sheet. For $n = 10^{12} \text{ cm}^{-2}$, $\omega_1 = 2\omega = 30 \text{ THz}$, $E_0 = 10^3 \text{ V/cm}$, a ribbon size of $S = 0.2 \mu\text{m} \times 5 \mu\text{m}$ and a dc current of $J = 0.16 \text{ mA}/\mu\text{m}$ corre-

sponding to $u = 0.1v_F$, one has $\xi_0 \approx 0.1$ and a normalized generation rate of $\Gamma/\omega \approx 0.9 \times 10^{-6}$ from this type-B source.

In the strong pump regime, the plasmon lifetime is enhanced. The pump induced relative growth rate of plasmons of frequency ω is

$$Q_g^{-1} = \frac{1}{16}\xi_{2\omega}\xi_0\Delta_{ilmn}\hat{q}_i\hat{q}_l\hat{E}_m\hat{J}_n, \quad (32)$$

similar to Eq. (31). Here \hat{X} means the unit vector along X and \hat{q} is the unit vector of the momentum of the subharmonic plasmon. For a current of $J = 0.5 \text{ mA}/\mu\text{m}$ at the doping level of $n = 10^{12} \text{ cm}^{-2}$ and under a pump field of $3 \times 10^5 \text{ V/cm}$ at $\omega_1 = 2\omega = 30 \text{ THz}$ polarized along \hat{y} , the subharmonic plasmon running along \hat{y} has a relative growth rate of $Q_g^{-1} \approx 0.03$, while those along \hat{x} has a growth rate of $Q_g^{-1} \approx 0.01$, comparable to the plasmon loss at low temperatures [9].

D. Field enhancement in a ribbon

From previous discussion, high field is essential to make the proposed devices practical. By design, these devices naturally enhance the incident field by two mechanisms, helping to achieve this goal. First, as discussed in Sec. II C(a), the split gate for type-A device in Fig. 1, or the contacts for type-B device in Fig. 8(b), act as antennas. They enhance the pump field by a factor of $F = E/E_v = \lambda_v/(2\pi w_s) \approx 10$ [49] for the channel width $w_s = 200 \text{ nm}$ and the vacuum wave length of the pump light $\lambda_v \approx 10 \mu\text{m}$ at $\omega_1 = 30 \text{ THz}$. Second, close to the resonance at ω_1 with the gapped mode of the ribbon [see, e.g., Fig. 1(b) and Fig. 7(b)], there is another field enhancement factor of $F_Q = \xi_f \frac{\omega_f^2}{(\omega - \omega_1 + i\gamma)(\omega + \omega_1 + i\gamma)}$ where ξ_f is an $O(1)$ shape factor, see Eqs. (11) and (C1). On resonance, this leads to a field enhancement factor of $F_Q \sim Q \sim 10$.

The antenna effect together with the resonance field enhancement gives a net enhancement factor of $QF \sim 100$. Therefore, the spontaneous PDC rate in the type-B device (Fig. 8) is enhanced by a factor of $(QF)^2$ while the spontaneous FWM rate in a ribbon is enhanced by $(QF)^4$. Similarly, to reach a pump field of 10^5 V/cm for parametric instability, the external incident field just needs to be of the order of 10^3 V/cm .

Note that the resonant field enhancement also occurs in photonic cavities such as microring resonators [40,67–70]. In our devices, the vacuum photons play the role of the waveguide photons in the bus waveguide therein, and the graphene ribbon plays the role of the ring resonator. Unlike the plasmon modes in the nanoribbon whose line width is dominated by intrinsic damping, the field enhancement on resonance in the micro-rings scales as \sqrt{Q} instead of Q [68] because the line width γ_p of the photon modes in the micro-ring resonator is dominated by radiation leakage into the bus waveguide. As a result, the enhancement of the FWM pair-generation rate due to resonant pumping scales as Q^2 , see, e.g., Eq. (4) of Ref. [67]. Note also that in both our devices (short devices with length $L_0 \ll l$) and photonic cavities, the resonance conditions with the generated modes contributes another factor of Q to the enhancement of the generation rate.

IV. DISCUSSION AND EXPERIMENTAL OUTLOOK

We proposed several graphene spontaneous PDC and FWM devices, which have promising quantum information

TABLE I. Comparison of graphene with conventional platforms for entangled photon generation following that in Ref. [40]. Although plasmons in these graphene devices can work in a broad frequency range from terahertz to infrared, we picked specific frequencies for estimations in the table. For graphene type-A device, the plasmon frequency is shown in Fig. 1(b) while for type-B device (with a current flow velocity of $u = 0.1v_F$) and FWM, we used a pump frequency of 30 THz. We used $a = 200$ nm and the quality factor of $Q = 10$ for all the graphene devices. Since we are using the Gaussian unit system while the majority of nonlinear optics literature uses the SI unit, we translated from the latter in the literature to the former in this table. The refraction indices used for conversion of $\chi^{(3)}$ are AlGaAs: $N = 3.2$ at $\lambda = 1.55 \mu\text{m}$, Silicon: $N = 3.5$ at $\lambda = 1.55 \mu\text{m}$, InP: $N = 3.4$ at $\lambda = 1.49 \mu\text{m}$, Si_3N_4 : $N = 2.0$ at $\lambda = 1.55 \mu\text{m}$.

PDC devices	Frequency	Q	$\chi^{(2)}$ (cm/V)	Generation rate ($\text{s}^{-1}\text{mW}^{-1}$)
Graphene type-A device [Fig. 1(a)]	4 THz	10–100	2×10^{-6}	9×10^{11}
Graphene type-B device [Fig. 8(b)]	15 THz	10–100	3×10^{-6}	2×10^{12}
Semiconductor coupled quantum wells [47,71,72]	1–50 THz		$\sim 10^{-6}$	
LiNbO ₃ [73]	0.79 eV		2×10^{-10}	2×10^7
AlN [74]	0.80 eV	1×10^5	4×10^{-11}	6×10^6
FWM devices	Frequency	Q	$\chi^{(3)}$ (cm^2/V^2)	Generation rate ($\text{s}^{-1}\text{mW}^{-2}$)
Graphene type-B device with no dc current	30 THz	10–100	4×10^{-14}	6×10^{14}
AlGaAs on insulator [40]	0.80 eV	1×10^6	7×10^{-17}	2×10^{10}
Silicon on insulator [75,76]	0.80 eV	$\sim 10^5$	4×10^{-17}	2×10^8
InP [77]	0.83 eV	4×10^4	2×10^{-16}	1×10^8
Si_3N_4 [78]	0.80 eV	2×10^6	4×10^{-19}	4×10^6

applications [79] as efficient sources of entangled plasmon pairs, photon-plasmon converters and plasmon amplifiers [80–84], etc., for future quantum plasmonic/polaritonic circuits. These plasmons are in the THz-infrared spectra range, which is an important energy range in condensed matter physics where superconductivity, magnetism, spin liquid, fractional quantum Hall effect, and other correlated phenomena are relevant. The quantum sources proposed here may enable quantum imaging [85] and sensing [86] of these novel states of matter.

Let us compare our devices with conventional photonic PDC sources. In photonic PDC, the nonlinearity is usually characterized by a nonlinear susceptibility $\chi^{(2)}$, which has the unit of inverse electric field. We can convert the 2D conductivity $\sigma^{(2)}$ to an effective 3D nonlinear susceptibility $\chi^{(2)}$ as follows. The 2D nonlinear susceptibility (in the Gaussian unit system) is [87] $\chi_{2D}^{(2)} = i\sigma^{(2)}/\omega_0$. For the type-A device, which is a ribbon with half width a , the extent of the electric field of the plasmon in the out-of-plane direction is $1/q_0 \sim a$. Hence, the effective 3D nonlinear susceptibility is $\chi^{(2)} \sim \chi_{2D}^{(2)}/a \sim 1/(ek_F^2)$. For $k_F = \sqrt{\pi} \times 10^6 \text{ cm}^{-1}$ corresponding to carrier density $n = 10^{12} \text{ cm}^{-2}$, we find $\chi^{(2)} \sim 2 \times 10^{-6} \text{ cm/V}$, which is much larger than that of conventional photonic media, e.g., LiNbO₃ [41,73], but is of the same order as $\chi^{(2)}$ in sophisticated photonic waveguides made of coupled quantum wells [47,71,72], see Table I. Nevertheless, a more important figure of merit is the dimensionless decay factor

$$\frac{\Gamma}{\omega_0} \sim |\chi^{(2)}|^2 \frac{\hbar\omega_0}{a^3} \quad (33)$$

[same as Eq. (10) noting that $g^{(2)} \sim \chi^{(2)}\omega_0^{3/2}a^{-1}L^{-1/2}$] of a waveguide mode into subharmonic modes due to spontaneous PDC. Due to strong mode confinement ($a \ll c/\omega_0$), this figure of merit is much larger than those of photonic devices. Physically, this is because electric field of a single plasmon in our device is much larger than that of a single photon.

As discussed in Sec. III C, a comparable $\chi^{(2)}$ can also be achieved in a type-B device where inversion symmetry is broken by applying a dc current flow to an uniformly doped ribbon if the flow velocity is moderately fast: $u > 0.1v_F$.

We also consider the FWM [33] in Sec. III B, a third-order nonlinear effect. Such a process does not require inversion-symmetry breaking and can take place in either a ribbon or a plain graphene sheet (type-C device in Fig. 7). This plasmon instability requires a relatively strong incident field controlled by the small parameter $\chi^{(3)}$, but it may be easier to realize experimentally since there is no specific requirement for the nanostructure. Compared with other nonlinear materials, $\chi^{(3)} = i\sigma^{(3)}/(\omega a)$ in graphene is much stronger, as shown in Table I.

A brief discussion of units conversion is in order. The Gaussian unit system defines the 3D nonlinear susceptibilities in terms of the effective dielectric function $\epsilon = 1 + 4\pi[\chi + \chi^{(2)}E + \chi^{(3)}E^2 + \dots]$ while the SI system defines them as $\epsilon/\epsilon_0 = 1 + \chi_{\text{SI}} + \chi_{\text{SI}}^{(2)}E + \chi_{\text{SI}}^{(3)}E^2 + \dots$ [33] where ϵ_0 is the vacuum permittivity. Therefore, besides converting the units of electric fields, there is an additional 4π factor: $\chi^{(n)} = \chi_{\text{SI}}^{(n)}/(4\pi)$. The unit m^2/W is also used in the literature [40,75] for third-order nonlinearity N_3 defined as $\delta N = \frac{1}{2N}\chi_{\text{SI}}^{(3)}E^2 = \frac{4\pi}{2N}\chi^{(3)}E^2 \equiv N_3P$ where $N = \sqrt{1 + 4\pi\chi}$ is the linear refraction index, δN is the change of effective refraction index due to third-order nonlinearity and the pump field, and P is the power flux of the pump field. Therefore, besides converting $\text{m}^2/\text{W} \rightarrow 0.00132806 \text{ m}^2/\text{V}^2$, one also needs the conversion $\chi^{(3)} = \frac{N}{2\pi}N_3$ to obtain the $\chi^{(3)}$ in Gaussian unit.

We note that the properties of the type-A device may have substantial temperature dependence arising from the middle of the ribbon [Fig. 1(a)] where the system is close to charge neutrality. The gapped mode is an antisymmetric charge oscillation between the upper and lower half of the strip, and thus strongly depends on the conductivity at the middle of the junction where the carriers are thermally excited

electrons and holes. Therefore, the frequency of the gapped mode (the shape factor ξ_1 in Sec. II) increases with temperature. In our derivation, we included only the intraband contributions [20,23,51,54] to the second-order nonlinear conductivity [Eq. (C20)]. Close to the middle of the junction where $\mu \lesssim \hbar\omega$, interband contributions may be important. There the total $\sigma^{(2)}$ is suppressed by a factor of $(\mu/\hbar\omega)^M$ with $M \geq 2$ compared to the intraband expressions used in this paper, but $\sigma^{(2)}$ may diverge at the narrow region where $\mu = \hbar\omega$ at the interband threshold for Pauli blocking [20,51,54,87–89]. This divergence is rounded by nonzero temperature, and would thus lead to substantial temperature dependence if included. We note that even in homogeneously doped graphene, there are 3 ~ 4 orders of magnitude inconsistencies [90] between theories (e.g., Refs. [20,51,54,87,88]) and experiments [25,26]. At this stage, it is premature to include such details in the model, and instead, we provide order of magnitude scaling results for the PDC. Incidentally, experiment and theory seem to agree for $\sigma^{(3)}$ [65].

As mentioned in Sec. I, generation of plasmons by PDC in graphene has also been considered by Ref. [51]. In that study, the pump and idler are far field photons. In contrast, here all three modes are plasmons such that frequencies are much lower and momenta are much larger, making this process much more efficient. Moreover, our proposed device involves an antenna that further increases the coupling efficiency to far field pump. As a result, the threshold power to achieve parametric instability is about four orders of magnitude lower than Ref. [51].

Since plasmons do not have the polarization degree of freedom (they are always longitudinally polarized), entanglement of polarization frequently discussed in quantum optics does not apply here. In plasmonics one focuses on the position-momentum (or energy-time) entanglement. For example, it has been shown this entanglement survives photon-plasmon conversion processes [31,63,91]. Here we proposed a scheme to directly generate and measure the entanglement of plasmon pairs on a chip using the modern tools of near-field optics.

The following estimate shows that graphene will not be damaged by the strong electric field required. The incident light heats up the electron gas in graphene with the heating power $P_{\text{heat}} = 2\text{Re}[\sigma] |E|^2$ and electron phonon scattering provides its cooling channel. The cooling power is approximated by $P_{\text{cool}} = C_v \gamma_c (T_e - T_l)$ where $T_e(T_l)$ is the electron(lattice) temperature, $C_v \sim g(\varepsilon_F)T$ is the heat capacity of the electron gas and $\gamma_c \sim 1 \text{ ps}^{-1}$ is the cooling rate revealed by previous experiment [7]. The lattice is a good heat conductor and is assumed to stay at the same temperature as the environment. The balancing of heating and cooling $P_{\text{heat}} = P_{\text{cool}}$ determines the stationary electron temperature $T_e - T_l \sim \frac{1}{4\pi^2} \frac{\gamma_c \varepsilon_F^2}{\gamma_c T} \xi^2$ where we have set the Boltzmann constant to one. At the typical doping level and frequency scale considered in this paper, for $T_l = 300 \text{ K}$ and $E = 10^4 \text{ V/cm}$, we conclude that $T_e - T_l \sim 6 \text{ K}$, which is quite small compared to either room temperature or the Fermi energy.

We worked in the kinetic regime where plasmon frequency is much larger than the electron-electron scattering rate $\omega \gg \Gamma_{ee}$ [23]. In the low frequency hydrodynamic regime $\omega \ll \Gamma_{ee}$ [23,81,92–98], the nonlinear conductivities are similar in

magnitude as long as the temperature is not much larger than chemical potential [23,24,99]. Therefore, our estimate of the generation rates should apply as well to the collective modes in the hydrodynamic regime such as the charged “demons” [23,81,92–94,96,97]. However, in the case of $T \gg \mu$, the demons become almost charge neutral, driven mainly by kinematic pressure, and the physics could be quite different. It may be interesting to study nonlinear and quantum effects for these collective modes [24,100].

Since the ribbon width is much larger than the lattice constant, the Dirac dispersion approximation for the electrons in graphene are valid. We used the “local” approximation to compute the plasmonic modes, which neglects dependence of the optical conductivity $\sigma(\omega, q)$ on the field wave vector q , see Appendix C1. The nonzero q effects enter $\sigma(\omega, q)$ as a power expansion in the dimensionless small number $v_F^2 q^2 / \omega^2$ [8,11,23], see, e.g., the supporting information of Refs. [11,23]. For the relevant modes of the ribbons such as mode-1 in Fig. 1, one has $q \approx 2\pi/(4a)$ and $v_F q / \omega \approx 0.36$, meaning that the devices are indeed in the local regime. As an estimate, the leading-order nonlocal correction $\delta\omega_1$ to the plasmon frequency ω_1 is just $\frac{\delta\omega_1}{\omega_1} = \frac{1}{2} \frac{3}{4} \left(\frac{v_F q}{\omega}\right)^2 \approx 0.05$ in the kinetic regime we are considering. Therefore, the Drude (or “local”) approximation is well justified.

Extension of these nonlinear effects to polaritonic modes [101–103] beyond those in graphene is a meaningful future direction. For example, monolayer hexagonal boron nitride [104,105] should exhibit strong second-order optical nonlinearity due to broken inversion symmetry of the crystal lattice, and would be a natural platform for generating entanglement between the long lived hyperbolic phonon polaritons [106–115]. Similar nonlinear processes exist for optical phonons in SiC [116], Josephson plasmons in layered superconductors [84,117–119] and the collective modes in excitonic insulators [120–124].

ACKNOWLEDGMENTS

M.M.F and Z.S. were supported by the Office of Naval Research under Grant No. N00014-18-1-2722. Z.S. acknowledges support from the Harvard Quantum Initiative through the Postdoctoral Fellowship in Quantum Science and Engineering. Research at Columbia University was supported by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), under Award No. DE-SC0018426. D.N.B. is an investigator in Quantum Materials funded by the Gordon and Betty Moore Foundation’s EPiQS Initiative through Grant No. GBMF4533. We thank Y. Shao for bringing to our attention the current carrying devices, and thank I. L. Aleiner, A. J. Sternbach, S. Moore, R. Jing, Y. Dong, L. Xiong, and B. S. Kim for helpful discussions.

APPENDIX A: SUMMARY OF NOTATIONS FOR THE RIBBON CASE

The main results of this paper are written in terms of various “shape” parameters and physical quantities such as size (a, L), Fermi momentum (ε_F), Fermi velocity (v_F), electric field (E), and frequency (ω). Most of the shape parameters are dimensionless numbers of the order of unity. They are S'_0 ,

S'_1 , α' , η' , ξ_0 , ξ_1 , ξ_g , ξ_{df} , ξ_β , and ξ_Q . Their definitions for the case of a ribbon are as follows:

$$\begin{aligned} S'_0 &= \int_{-1}^1 dy \tilde{\sigma}(y) |\tilde{\mathbf{E}}_0|^2, \\ S'_1 &= \int_{-1}^1 dy \tilde{\sigma}(y) |\tilde{\mathbf{E}}_1|^2, \\ \alpha' &= \int_{-1}^1 dy \tilde{\sigma}(y) \hat{\mathbf{e}} \cdot \tilde{\mathbf{E}}_1, \\ \eta' &= q_x^2 \int_{-1}^1 dy \partial_y^2 \varphi_1 \varphi_0^2 \tilde{g}(y), \\ \xi_v(q) &= \frac{\omega(q)}{\sqrt{\alpha_g(a k_F) v_F q}} = \frac{\omega(q)}{v_a q}, \\ \xi_0 &= \xi_v(q_0), \\ \xi_1 &= \frac{\hbar \omega_1}{\sqrt{2e^2 E_F / a}}, \\ \xi_g &= \xi_1^{1/2} \frac{\eta'}{S_1^{1/2} S'_0}, \\ \xi_{df} &= \frac{v_a}{v_l} \frac{\eta'}{S_1 S'_0}, \\ \xi_\beta &= \frac{1}{\xi_1^{3/2}} \frac{\alpha'}{S_1^{1/2}}, \\ \xi_Q &= \frac{\eta'}{\xi_1^2} \frac{\alpha'}{S'_0 S_1}. \end{aligned}$$

The remaining ones, S , α and $|\tilde{\rho}|^2$, are just the corresponding dimensionless ones multiplied by aL , one half of the total area of the ribbon.

APPENDIX B: PARAMETRIC OSCILLATOR

The spontaneous PDC process in Sec. II is similar to what happens in a parametric oscillator (PO) except that the former leads to two-mode squeezing while the latter corresponds to single-mode squeezing [57]. Optical POs are widely used to generate subharmonic light and to create entangled photon pairs. To make the paper self-contained, we review the basic theory of the PO in this Appendix.

A PO can be modeled with the equation of motion

$$\hat{L}x(t) = [\partial_t^2 + \gamma \partial_t + (1 + 2\delta \cos \Omega t) \omega_0^2] x(t) = F(t), \quad (\text{B1})$$

where Ω is the frequency of the pump modulating the natural frequency ω_0 of the oscillator by a relative amount $\sim \delta$. The difference

$$\Delta = \Omega - 2\omega_0 \quad (\text{B2})$$

of Ω from the primary parametric resonance frequency $2\omega_0$ is referred to as the detuning. Given the initial conditions $x(0)$, $\dot{x}(0) \equiv \partial_t x(t)|_{t=0}$, the solution of Eq. (B1) can be written in terms of the retarded Green's function $G = G(t, t')$:

$$\begin{aligned} x(t) &= \{[\dot{x}(0) + \gamma x(0)]G - x(0) \partial_{t'} G\}|_{t'=0} \\ &+ \int G(t, t - \tau) \Theta(t - \tau) F(t - \tau) d\tau, \quad (\text{B3}) \end{aligned}$$

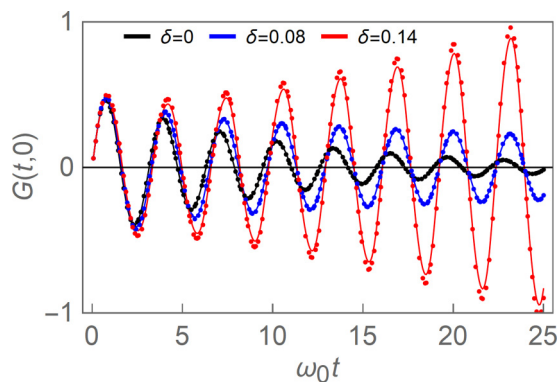


FIG. 9. Behavior of the Green's function $G(t, 0)$ of Eq. (B1) at several values of the modulation parameter δ at resonance $\Omega = 2\omega_0$. The solid lines show the exact Green's function [Eq. (B4)], the dots are the approximations constructed from Eq. (B18). The damping rate is $\gamma = 0.1\omega_0$ corresponding to $\delta_c \approx 0.1$.

where $\Theta(t)$ is the Heaviside step-function. One simple example is $x(0) = 0$, $\dot{x}(0) = 1$, $F(t) \equiv 0$, in which case $x(t) = G(t, 0)$.

As usual, the Green's function is constructed from an arbitrary pair of linearly independent solutions $x_1(t)$, $x_2(t)$ of the homogeneous equation $\hat{L}x(t) = 0$:

$$G(t, t') = \Theta(t - t') \frac{x_1(t')x_2(t) - x_2(t')x_1(t)}{W(t')}, \quad (\text{B4})$$

$$W(t) = x_1(t) \partial_t x_2(t) - x_2(t) \partial_t x_1(t). \quad (\text{B5})$$

By virtue of the Abel formula, the Wronskian determinant [Eq. (B5)] has the following t dependence:

$$W(t) = W(0) e^{-\gamma t}. \quad (\text{B6})$$

One possible choice of $x_1(t)$ and $x_2(t)$ is

$$x_1(t) = e^{-\frac{1}{2}\gamma t} \text{C}\left(\frac{4\omega_0^2 - \gamma^2}{\Omega^2}, -\frac{4\delta\omega_0^2}{\Omega^2}, \frac{1}{2}\Omega t\right), \quad (\text{B7})$$

$$x_2(t) = e^{-\frac{1}{2}\gamma t} \text{S}\left(\frac{4\omega_0^2 - \gamma^2}{\Omega^2}, -\frac{4\delta\omega_0^2}{\Omega^2}, \frac{1}{2}\Omega t\right), \quad (\text{B8})$$

where $\text{C}(a, b, z)$ and $\text{S}(a, b, z)$ are the even and the odd Mathieu functions, respectively, as defined in Mathematica [125] (Mathematica notations are $\text{MathieuC}[a, b, z]$ and $\text{MathieuS}[a, b, z]$). These functions are normalized as $\text{C}(a, 0, z) = \cos(\sqrt{az})$, $\text{S}(a, 0, z) = \sin(\sqrt{az})$ for $b = 0$.

From Eqs. (B4)–(B8), we conclude that

$$G(t, 0) = \frac{x_2(t)}{\dot{x}_2(0)}, \quad (\text{B9})$$

so that an instability, i.e., the exponential growth of $G(t, 0)$ is possible if $\text{S}(a, b, z)$ is increasing faster than $e^{\frac{1}{2}\gamma t}$. This occurs if $|\delta|$ exceeds some threshold value $\delta_c = \delta_c(\gamma, \Delta)$. Conversely, the response function $G(t, 0)$ exhibits exponential decay with time if $|\delta| < \delta_c$. Examples of such stable and unstable behaviors are shown in Fig. 9.

Approximate analytical calculation of δ_c can be done for low damping $\gamma \ll \omega_0$ and small detuning $\Delta \ll \omega_0$. Thus, one

can show that

$$\delta_c \simeq \frac{\sqrt{\Delta^2 + \gamma^2}}{\omega_0}, \quad (\text{B10})$$

in agreement with Fig. 9.

The derivation goes as follows. Let us seek a solution of Eq. (B1) in the form

$$x(t) = A(t)e^{i\omega_0 t} + B(t)e^{-i\omega_0 t}, \quad (\text{B11})$$

where $A(t), B(t)$ are slowly varying. Matching the coefficients for the rapidly oscillating phase factors of Eq. (B1), we obtain

$$\begin{pmatrix} 2i\partial_t + i\gamma & \omega_0\delta e^{i\Delta t} \\ \omega_0\delta e^{-i\Delta t} & -2i\partial_t - i\gamma \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (\text{B12})$$

To eliminate the factors $e^{\pm i\Delta t}$, we change variables to

$$A(t) = e^{\frac{i}{2}\Delta t} a(t), \quad B(t) = e^{-\frac{i}{2}\Delta t} b(t), \quad (\text{B13})$$

arriving at

$$\begin{pmatrix} 2i\partial_t + i\gamma - \Delta & \omega_0\delta \\ \omega_0\delta & -2i\partial_t - i\gamma - \Delta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (\text{B14})$$

The general solution of this equation is

$$\begin{pmatrix} a \\ b \end{pmatrix} = e^{\lambda_1 t} \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} + e^{\lambda_2 t} \begin{pmatrix} u_2 \\ v_2 \end{pmatrix}, \quad (\text{B15})$$

where

$$\lambda_{1,2} = -\frac{\gamma}{2} \pm \frac{i}{2} \sqrt{\Delta^2 - (\omega_0\delta)^2} \quad (\text{B16})$$

are the growth rates and

$$\begin{pmatrix} u_j \\ v_j \end{pmatrix} = \begin{pmatrix} \omega_0\delta \\ \Delta - i\gamma - i\lambda_j \end{pmatrix}, \quad j = 1, 2 \quad (\text{B17})$$

are the corresponding eigenvectors. Two linearly independent solutions for $x(t)$ could be chosen as

$$x_j = e^{\lambda_j t} (u_j e^{i\frac{\Omega}{2}t} + v_j e^{-i\frac{\Omega}{2}t}), \quad j = 1, 2, \quad (\text{B18})$$

and the Green's function can be constructed per Eq. (B4). Assuming $\text{Re}[\lambda_2] \geq \text{Re}[\lambda_1]$, the instability threshold is determined by the condition

$$\text{Re}[\lambda_2] = 0, \quad (\text{B19})$$

which leads to Eq. (B10). A comparison with the exact Green's function confirms the validity of this approximation for small damping and detuning, such as $\gamma = 0.1\omega_0$ and $\Delta = 0$ in Fig. 9.

Consider now the response of the PO to a periodic driving source

$$F(t) = f_\omega e^{-i\omega t} + \text{c.c.} \quad (\text{B20})$$

In the stable case, $|\delta| < \delta_c$, where the effect of the initial conditions disappears at long times, x_ω is the linear response to F :

$$x_\omega(t) = f_\omega e^{-i\omega t} \int_0^\infty G(t, t - \tau) e^{i\omega\tau} d\tau + \text{c.c.} \quad (\text{B21})$$

Without the pump, $G(t, t - \tau)$ is a function of τ only, and so the integral on the right-hand side reduces to a constant. Under pumping, $G(t, t - \tau)$ is periodic in t with the period

$2\pi/\Omega$. Accordingly, the Fourier spectrum of $x_\omega(t)$ becomes a frequency comb of Floquet harmonics $\omega + m\Omega$ where m is an arbitrary integer. For small γ and Δ , however, it is easier to compute $x_\omega(t)$ not from Eq. (B21) but from Eq. (B14) supplemented with the appropriate right-hand side. The result is

$$x_\omega(t) = \frac{f_\omega e^{-i\omega t}}{(\omega - \nu_1)(\omega - \nu_2)} \left(\frac{\Omega - 2\omega - i\gamma + \Delta}{4\omega_0} - \frac{\delta}{4} e^{i\Omega t} \right) + \text{c.c.} \quad (\text{B22})$$

Apparently, this approximation accounts only for the dominant Floquet harmonics: $\omega - \Omega$, ω , and $\omega + \Omega$. In the complex plane of ω , function x_ω (or the linear response kernel) has two poles, at ν_1 and ν_2 , given by

$$\nu_{1,2} = \frac{\Omega}{2} + i\lambda_{1,2} = \omega_0 + \frac{\Delta}{2} - \frac{i\gamma}{2} \mp \frac{1}{2} \sqrt{\Delta^2 - (\omega_0\delta)^2}. \quad (\text{B23})$$

The pole ν_2 has a larger imaginary part, which grows with $|\delta|$. At $|\delta| = \delta_c$, ν_2 crosses over into the upper half-plane of ω , signaling the instability.

Instead of a periodic driving source, we can consider a stochastic one, namely, a Langevin source defined by the two-point correlator

$$\langle F(t)F(0) \rangle = \frac{\gamma T}{\pi} \delta(t). \quad (\text{B24})$$

In this case the total driving force $F(t)$ contains a continuum of Fourier harmonics, and the response $x(t)$ is obtained by integrating expressions like Eq. (B22) over ω . The stochastic drive leads to random fluctuations of x , causing a nonzero average power injection by the parametric pump into the system:

$$P = 2\omega_0^2 \delta \langle \cos \Omega t x \partial_t x \rangle. \quad (\text{B25})$$

A straightforward calculation based on Eqs. (B22) and (B25) yields

$$P = \frac{(\omega_0\delta)^2 \gamma T}{\gamma^2 + \Delta^2 - (\omega_0\delta)^2}. \quad (\text{B26})$$

Finally, we generalize this result to the quantum case. We do not show a formal derivation, but sketch a more physical approach. The essential point is that the power injection P by the pump is enabled by the nonzero quantum/thermal fluctuations of x . The starting point is a Hamiltonian of the total system (PO plus environment) in the form

$$H(t) = H_{\text{PO}} - F(t)x + H_{\text{env}}, \quad (\text{B27})$$

$$H_{\text{PO}} = \frac{p^2}{2} + \frac{1}{2}(1 + 2\delta \cos \Omega t)\omega_0^2 x^2. \quad (\text{B28})$$

The purpose of H_{env} is to describe the Langevin noise source and the dissipation effects. A standard device to achieve this is to use the Caldeira-Leggett model of a bath of harmonic oscillators. We surmise that the final answer for P is similar to the classical Eq. (B26) except the thermal energy T is replaced by $\hbar\omega_0(n_b + \frac{1}{2})$ where $n_b = 1/(e^{\hbar\omega_0/T} - 1)$ is the boson occupation number. Since the injected power increases the energy of the PO, the pump must be generating quanta of

TABLE II. Physical meanings of the dimensionless profile functions whose typical local values are order unity.

Symbol	Physical quantity
$\tilde{\sigma}(\mathbf{r})$	conductivity
$\tilde{g}(\mathbf{r})$	2nd-order conductivity $\sigma^{(2)}$
$\tilde{\rho}(\mathbf{r})$	charge density of a mode
$\tilde{\mathbf{E}}(\mathbf{r})$	electric field of a mode

the PO motion. Dividing P by the energy $2\hbar\omega_0$ of two such quanta, we obtain the pair-generation rate

$$R = \frac{P}{2\hbar\omega_0} = \frac{\gamma}{\gamma^2 + \Delta^2 - (\omega_0\delta)^2} \frac{(\omega_0\delta)^2}{2} \left(n_b + \frac{1}{2} \right). \quad (\text{B29})$$

Note that formally, in the quantum mechanical language, this nonperturbative result in the parametric pump δ can be obtained by computing the linear response to δ (bubble diagram) [120], with the boson propagators replace by the renormalized ones by δ [see Eq. (B22)]. As shown by Eq. (B29), at weak pumping, R scales quadratically with δ , consistent with Fermi's Golden Rule. The precise match is obtained in the limit $\gamma \rightarrow 0$, $\omega_0\delta/\gamma \rightarrow 0$ where R is nonzero only at the resonance, $R \propto \delta(\Delta)$. As δ increases, deviations from Fermi's Golden rule appear. As δ approaches the critical value δ_c [Eq. (B10)], the pair-generation rate diverges.

Lastly, expressing δ in terms of the coupling constant $g^{(2)}$ in Sec. II and integrating over the acoustic modes in the vicinity of ω_0 we obtain Eq. (14).

APPENDIX C: DERIVATION FOR THE GRAPHENE RIBBON (TYPE-A DEVICE)

1. Plasmon modes on a generic nano structure

Unlike the uniform two-dimensional (2D) electron gas, a generic conducting nanostructure [17,126,127] breaks the translational symmetry in at least one direction, and the plasmon modes can not be classified by 2D momenta. In this section, we define necessary profile functions describing the graphene nanoisland (e.g., a ribbon) such that the quantization could be described semi-analytically. Meanings of the dimensionless profile functions are summarized in Table II.

The linear optical conductivity is assumed local and in the Drude form: $\mathbf{j}(\mathbf{r}) = \hat{\sigma}(\mathbf{r})\mathbf{E}(\mathbf{r}) = \frac{i}{\pi(\omega+i\gamma)}D(\mathbf{r})\mathbf{E}(\mathbf{r})$ where $D(\mathbf{r})$ is the local Drude weight. This holds for $\omega \gg v_F q$ where q is the characteristic momentum corresponding to the plasmon modes and size of the nanostructure. In the case of a graphene ribbon with a being its half width, this momentum is roughly $q \approx \pi/(2a)$. We define the dimensionless conductivity profile function $\tilde{\sigma}$ by $\hat{\sigma}(\mathbf{r}) = \hat{\sigma}\tilde{\sigma}(\mathbf{r})$ (or $D(\mathbf{r}) = D\tilde{\sigma}(\mathbf{r})$) where D is the maximum local Drude weight. Assume a certain plasmon mode has the charge density profile $\rho(\mathbf{r}) = \rho\tilde{\rho}(\mathbf{r})$, its electric field profile is

$$\mathbf{E}(\mathbf{r}) = -\nabla\hat{V}\rho(\mathbf{r}) = \rho\tilde{\mathbf{E}}(\mathbf{r}), \quad \tilde{\mathbf{E}}(\mathbf{r}) \equiv -\nabla\hat{V}\tilde{\rho}(\mathbf{r}) \quad (\text{C1})$$

where the Coulomb kernel $\hat{V}(\mathbf{r}, \mathbf{r}')$ is defined as an operator such that $\hat{V}f(\mathbf{r}) = \int d\mathbf{r}'^2 \frac{1}{|\mathbf{r}-\mathbf{r}'|} f(\mathbf{r}')$. Note that $\tilde{\sigma}$, $\tilde{\rho}$ and $\tilde{\mathbf{E}}$ are all dimensionless profile functions whose values are order one.

From the charge continuity equation (CCE)

$$\partial_t \rho(\mathbf{r}) + \nabla \cdot \mathbf{j} = 0 \quad (\text{C2})$$

with only linear conductivity, the mode $\rho(\mathbf{r}) = \rho\tilde{\rho}(\mathbf{r})$ being an eigenmode with frequency ω_1 implies

$$-i\omega_1\tilde{\rho}(\mathbf{r}) - \sigma(\omega_1)\nabla\tilde{\sigma}(\mathbf{r})\nabla\hat{V}\tilde{\rho}(\mathbf{r}) = 0. \quad (\text{C3})$$

Define the linear operator $\hat{L} = \nabla\tilde{\sigma}(\mathbf{r})\nabla\hat{V}$, the above equation simplifies to

$$\hat{L}\tilde{\rho}(\mathbf{r}) = \frac{-i\omega_1}{\sigma(\omega_1)}\tilde{\rho}(\mathbf{r}), \quad (\text{C4})$$

i.e., the shape function $\tilde{\rho}(\mathbf{r})$ is an eigenvector of \hat{L} with eigenvalue $-i\omega_1/\sigma(\omega_1)$. We define inner product to be $\langle f|g \rangle \equiv \int f^*(r)\hat{V}(r, r')g(r')d^2rd^2r'$ such that \hat{L} is Hermitian. Although the plasmon frequency is complex with a small imaginary part $\text{Im}[\omega_1] = \gamma/2$, the eigenvalue $-i\omega_1/\sigma(\omega_1)$ is real since \hat{L} is a Hermitian operator.

If an external uniform electric field $\mathbf{E}_{\text{ext}} = E_{\text{ext}}\mathbf{e}$ with frequency ω is applied, the CCE becomes

$$(-i\omega - \sigma(\omega)\hat{L})\tilde{\rho}(\mathbf{r})\rho + \sigma(\omega)\nabla\tilde{\sigma}(\mathbf{r})\mathbf{e}E_{\text{ext}} = 0. \quad (\text{C5})$$

Taking inner product with $\tilde{\rho}(\mathbf{r})$, we arrive at

$$\left(-i\omega - \sigma(\omega)\frac{-i\omega_1}{\sigma(\omega_1)} \right) |\tilde{\rho}|^2 \rho + \sigma(\omega)\alpha E_{\text{ext}} = 0, \quad (\text{C6})$$

where

$$\begin{aligned} |\tilde{\rho}|^2 &= \langle \tilde{\rho}|\tilde{\rho} \rangle = \int \tilde{\rho}(r)\hat{V}(r, r')\tilde{\rho}(r')d^2rd^2r' \\ &= \frac{\sigma(\omega_1)}{i\omega_1} \int \tilde{\sigma}(r)|\tilde{\mathbf{E}}(\mathbf{r})|^2 d^2r = \frac{\sigma(\omega_1)}{i\omega_1} S, \end{aligned} \quad (\text{C7})$$

and the dipole factor is

$$\begin{aligned} \alpha &= \langle \tilde{\rho}|\mathbf{e} \cdot \nabla\tilde{\sigma} \rangle = \mathbf{e} \cdot \int \tilde{\rho}(r)\hat{V}(r, r')\nabla'\tilde{\sigma}(r')d^2rd^2r' \\ &= \mathbf{e} \int \tilde{\sigma}(r)\tilde{\mathbf{E}}(r)d^2r. \end{aligned} \quad (\text{C8})$$

Two new shape quantities are introduced: S has the interpretation of the effective area [17] of this mode, α is the effective dipole moment of the mode along \mathbf{e} . Therefore, the charge density amplitude induced by the external driving field is Eq. (11). In this response function, there is a simple pole at resonance $\omega = \text{Re}[\omega_1]$, leading to divergence. However, this pole is broadened by the plasmon damping rate γ , and the enhancement factor at resonance will be

$$\rho = (Q_1/i)(\alpha/S)E_{\text{ext}}, \quad (\text{C9})$$

where $Q_1 = \omega_1/\gamma$ is the quality factor of this plasmon mode.

2. Quantization

For each plasmon mode as a harmonic oscillator, we choose charge density ρ as the generalized coordinate and write it in terms of plasmon creation and annihilation operators as

$$\rho = \rho_u(a + a^\dagger). \quad (\text{C10})$$

Equivalently [Eq. (C1)] the electric field is

$$\mathbf{E}(\mathbf{r}) = \rho_u(a + a^\dagger)\tilde{\mathbf{E}}(\mathbf{r}). \quad (\text{C11})$$

Working in the gauge $A = (0, \mathbf{A})$ for the vector potential, from the relation $\mathbf{E} = -\partial_t \mathbf{A}/c$, we can write the vector potential as

$$\mathbf{A}(\mathbf{r})/c = -i \frac{\rho_u}{\omega_1} (a - a^\dagger)\tilde{\mathbf{E}}(\mathbf{r}), \quad (\text{C12})$$

where the operators are time dependent ones from the interaction picture $a(t) = ae^{-i\omega_1 t}$.

Since the potential energy of a harmonic oscillator is half of its total energy, the quantum unit of density ρ_u can be determined by

$$\langle V_{\text{Coulomb}} \rangle = \frac{1}{2} \rho_u^2 |\tilde{\rho}|^2 = \frac{1}{4} \hbar \omega_1, \quad (\text{C13})$$

which yields

$$\rho_u = \sqrt{\hbar \omega_1 / (2 |\tilde{\rho}|^2)}. \quad (\text{C14})$$

We note that due to nonzero damping, the modes here should be understood as quasi-normal modes [128–130]. Nevertheless, due to the nice properties of Drude response for plasmons, we were able to cast the formalism into a Hermitian one with a simpler approach [Eq. (C4)] such that the basis functions are automatically orthonormal. The quantization procedure assumes no dissipation, which is then added phenomenologically and should be understood as coming implicitly from a bath.

3. Three mode coupling due to second-order nonlinearity

In general, the current generated in response to electric field contains a second-order term

$$j_i^{(2)}(\mathbf{r}, t) = \hat{\sigma}_{ilm}^{(2)} [E_l(\mathbf{r}, t) E_m(\mathbf{r}, t)], \quad (\text{C15})$$

where $\hat{\sigma}_{ilm}^{(2)}$ is a tensor-valued operator nonlocal in space and time. In a uniform system, this operator is diagonal in the frequency-momentum space, so that

$$j_i^{(2)}(-\mathbf{q}_3, -\omega_3) = \sigma_{ilm}^{(2)} E_l(\mathbf{q}_1, \omega_1) E_m(\mathbf{q}_2, \omega_2). \quad (\text{C16})$$

where $\omega_3 = -(\omega_1 + \omega_2)$, $\mathbf{q}_3 = -(\mathbf{q}_1 + \mathbf{q}_2)$, and

$$\sigma_{ilm}^{(2)} = \sigma_{ilm}^{(2)}(\mathbf{q}_1, \omega_1; \mathbf{q}_2, \omega_2). \quad (\text{C17})$$

By convention, the nonlinear second-order conductivity $\sigma_{ilm}^{(2)}$ is symmetrized, i.e., invariant under the interchange ($1 \leftrightarrow 2, l \leftrightarrow m$). It is convenient to define another tensor Π_{ilm} , which describes the current response to the vector potential \mathbf{A} . In the temporal gauge, $\mathbf{E}(\mathbf{q}, \omega) = (i\omega/c)\mathbf{A}(\mathbf{q}, \omega)$, it is given by

$$\hat{\Pi}_{ilm}(\omega_3, \mathbf{q}_3; \omega_1, \mathbf{q}_1, \omega_2, \mathbf{q}_2) \equiv -\omega_1 \omega_2 \sigma_{ilm}^{(2)}, \quad (\text{C18})$$

so that

$$j_i^{(2)}(-\mathbf{q}_3, -\omega_3) = \frac{1}{c^2} \hat{\Pi}_{ilm} A_l(\mathbf{q}_1, \omega_1) A_m(\mathbf{q}_2, \omega_2). \quad (\text{C19})$$

Due to inversion symmetry of graphene, $\sigma_{ilm}^{(2)}$ vanishes at $q_1 = q_2 = 0$. It grows linearly with q_1, q_2 when these momenta are small. The same properties are inherited by the tensor Π_{ilm} . In particular, in the kinetic regime of graphene

[16,23,51,54] and neglecting dissipation, Π_{ilm} can be written as

$$\hat{\Pi}_{ilm} = -\frac{D^{(2)}}{\omega_1 \omega_2 \omega_3} [(\omega_2^2 q_1 + \omega_1^2 q_2 + 2\omega_1 \omega_2 q_3)_l \delta_{lm} + \text{perm}], \quad (\text{C20})$$

[When comparing to other expressions in the literature, one should remember the constraint $\sum_k \omega_k = \sum_k \mathbf{q}_k = 0$.] The notation ‘‘perm’’ stands for two additional terms obtained from the first one by the permutations

$$\begin{aligned} (3, i, 1, l; 2, m) &\rightarrow (1, l; 3, i; 2, m), \\ (3, i, 1, l; 2, m) &\rightarrow (2, m; 1, l; 3, i). \end{aligned} \quad (\text{C21})$$

Accordingly, $\hat{\Pi}_{ilm}$ is symmetric under arbitrary permutations of the triads $\{\mathbf{q}_k, \omega_k, \mu_k\}$, where $k = 1, 2, 3$ and $\mu_k = i, l, m$. This symmetry can be understood as a consequence of energy conservation.

The tensor Π_{ilm} defines the operator $\hat{\Pi}_{ilm}$ nonlocal in space and time. The three-mode coupling Hamiltonian can be constructed as

$$H_c = -\frac{1}{3c^3} \int d^2 r \hat{\Pi}_{ilm} A_i A_l A_m. \quad (\text{C22})$$

This Hamiltonian obeys the requirement $\partial H_c / \partial \mathbf{A} = -\mathbf{j}^{(2)}/c$. Assume any three plasmon modes 1, 2, and 3, Eq. (C22) together with Eq. (C12) lead to their coupling Hamiltonian

$$\begin{aligned} H_c = -2i \frac{\rho_{1u} \rho_{2u} \rho_{3u}}{\omega_0 \omega_2 \omega_3} \int d^2 r \hat{\Pi}_{ilm} \tilde{E}_{1i} \tilde{E}_{2l} \tilde{E}_{3m} \\ \cdot (a_1 - a_1^\dagger)(a_2 - a_2^\dagger)(a_3 - a_3^\dagger). \end{aligned} \quad (\text{C23})$$

In $\hat{\Pi}$, the momenta should be replaced by spatial gradients acting on the corresponding field profiles, and the frequency arguments should match those of the creation/annihilation operators [e.g., for terms like $a_1 a_2^\dagger a_3^\dagger$, it should be $\hat{\Pi}(-\omega_0, \omega_2, \omega_3)$].

If the nanostructure has inversion symmetry, the plasmon modes can have even or odd parity. In order for the interaction Eq. (C23) not to vanish, the product of the three fields must have even parity. In the case of subharmonic decay, the mode 0 should couple to light whose electric field is nearly uniform, and is thus an odd mode. The other two modes should be nearly identical whose product is thus even, leading the product of the three modes to be odd. Therefore, the desired subharmonic decay does not happen in inversion symmetric nanostructures. We investigate the approach to break the inversion symmetry in the next section.

4. The graphene ribbon

For the graphene ribbon shown in Fig. 1, there is mirror symmetry of $x \rightarrow -x$ while the static transverse electric field breaks the mirror symmetry of $y \rightarrow -y$. The gapped mode 1 and the acoustic mode 0 with momentum q along x have the electric potential

$$\varphi_1 = \varphi_1(y), \quad \varphi_0 = \varphi_0(y) e^{iqx}. \quad (\text{C24})$$

Fully extracting the dependence on length scales, the effective areas can be written as

$$\begin{aligned} S_0 &= aLS'_0 = aL \int_{-1}^1 dy \bar{\sigma}(y) |\tilde{\mathbf{E}}_0|^2, \\ S_1 &= aLS'_1 = aL \int_{-1}^1 dy \bar{\sigma}(y) |\tilde{\mathbf{E}}_1|^2, \end{aligned} \quad (\text{C25})$$

where S'_0 and S'_1 are dimensionless factors of order one. The dipole factor for the gapped mode is

$$\alpha = aL\alpha' = aL \int_{-1}^1 dy \bar{\sigma}(y) \hat{\mathbf{e}} \cdot \tilde{\mathbf{E}}_1. \quad (\text{C26})$$

where α' is order one.

In Sec. C2, we have assumed that the density shape functions $\tilde{\rho}(\mathbf{r})$, as transformation coefficients for the “generalized coordinates” ρ , are all real functions. Due to the translational invariance along x , it is convenient to define the “complex” modes ρ_q with well defined momenta. In the conventional quantization rule $\rho_q = \rho_u(a_q + a_{-q}^\dagger)$, the creation operator a_q^\dagger generates plasmon with momentum q , i.e., its action adds $\hbar q$ to the total momenta. This set of creation and annihilation operators are related to those of Eq. (C10) by a canonical transformation.

The interaction strength—If the chemical potential on this ribbon varies slowly enough: $1/\lambda_\mu \ll \omega/v_F$ where λ_μ is the characteristic spatial scale of chemical potential variation, one can make the local approximation, using Eq. (C20) as the local second-order nonlinear coupling strength. The second-order optical weight can be written as $D^{(2)} = D_k^{(2)} \tilde{g}(y)$ where $D_k^{(2)}$ is a typical value of $D^{(2)}$ on the strip and $\tilde{g}(y)$ is an order one-dimensional shape function. Given fixed chemical potential, the value of $D_k^{(2)}$ is temperature dependent [23]. In the high-frequency kinetic regime, $D_k^{(2)}$ goes to a constant value

$$D_0^{(2)} = -\frac{1}{8\pi} \frac{e^3 v^2}{\hbar^2} \text{sgn } \mu \quad (\text{C27})$$

at $T \ll \mu$ and vanishes as $T \gg \mu$. Since $\mu \gg T$ at the edge of the strip, we define $D_k^{(2)}$ to be its zero temperature value $D_0^{(2)}$.

With Eqs. (C20), (C25), and (C14), the coupling Hamiltonian Eq. (C23) applied to modes 0, q and $-q$ becomes

$$H_c = g_q^{(2)} a_1 a_q^\dagger a_{-q}^\dagger + \text{c.c.} \quad (\text{C28})$$

where we have kept only the near resonant terms. Therefore, we have derived the interaction term in Eq. (7) where the interaction strength g is given by Eq. (8) with $\xi_g = \xi_1^{1/2} \eta' / (S_1^{1/2} S_0')$ and the dimensionless integral η' defined as

$$\eta' = q_x^2 \int_{-1}^1 dy \partial_y^2 \varphi_1 \varphi_0^2 \tilde{g}(y). \quad (\text{C29})$$

In the above expression for η' , we have dropped the terms involving $\partial_y \varphi_0$. This approximation is good if $\partial_x \varphi_0 \gg \partial_y \varphi_0$, which is true for the subharmonic plasmons in Fig. 1(b). The dimensionless momentum $q'_x = q_x a$ is order one.

In Fig. 3, we plot the resulting shape factors for the strip as the linear doping profile varies. There we have assumed the temperature and doping dependence of the Drude weight and the second-order optical

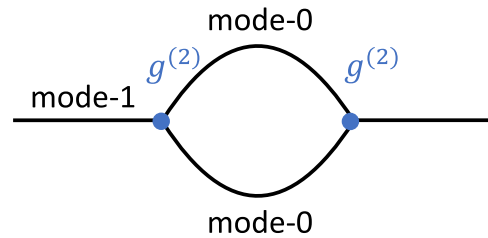


FIG. 10. Feynman diagram for the self energy of mode-1 plasmon due to its parametric coupling to two mode-0 plasmons. Solid lines are the plasmon propagators.

weight are $D(\mu, T) = 2 \frac{e^2}{\hbar^2} T \ln[2 \cosh(\frac{\mu}{2T})]$ and $D_k^{(2)}(\mu, T) = D_0^{(2)} \tanh(\frac{\mu}{2T})$, so that their dimensionless profile functions are $\bar{\sigma}(y) = \frac{2T}{\mu(y)} \ln[2 \cosh(\frac{\mu(y)}{2T})]$ and $\tilde{g}(y) = \tanh(\frac{\mu(y)}{2T})$.

The occupation number—Under the ac electric field of the incident light, the gapped mode ω_1 is driven to a coherent state $|\beta(t)\rangle$ where $\beta(t) = \beta e^{-i\omega_1 t}$. From Eq. (C10), the average value of ρ is

$$\langle \rho \rangle = \rho_u \langle \beta(t) | (a + a^\dagger) | \beta(t) \rangle = \rho_u (\beta e^{-i\omega_1 t} + \text{c.c.}), \quad (\text{C30})$$

which combined with Eq. (C9) yields

$$\beta = (Q/i)(\alpha/S) E_{\text{ext}} / \rho_u. \quad (\text{C31})$$

The above equation and Eq. (C14), (6) yields Eq. (12).

5. The decay rate

In this subsection, we derive Eq. (9), the decay rate Γ of the mode-1 plasmon into two mode-0 plasmons through the interaction Hamiltonian Eq. (7). This could be done by calculating the self energy (Fig. 10) of the mode-1 plasmon. It is a bubble diagram formed by two mode-0 propagators:

$$\begin{aligned} \Sigma(i\omega) &= \left| \frac{g^{(2)}}{\hbar} \right|^2 \sum_{q, i\Omega_n} \frac{\omega(q)}{(i\Omega_n)^2 - \omega(q)^2} \frac{\omega(q)}{(i\omega + i\Omega_n)^2 - \omega(q)^2} \\ &= \left| \frac{g^{(2)}}{\hbar} \right|^2 \sum_q \frac{[2n(\omega(q)) + 1]4\omega(q)}{(i\omega)^2 - 4\omega(q)^2}, \end{aligned} \quad (\text{C32})$$

where $n(x) = 1/(e^{\hbar x/T} - 1)$ is the boson occupation number. Its imaginary part at $i\omega = \omega_1$ would simply give the Fermi's golden rule:

$$\begin{aligned} \text{Im}[\Sigma(\omega_1)] &= \left| \frac{g^{(2)}}{\hbar} \right|^2 \sum_q [2n(\omega(q)) + 1] \pi \delta(\omega_1 - 2\omega(q)) \\ &= \left| \frac{g^{(2)}}{\hbar} \right|^2 \frac{L}{v_0} \left(n(\omega_1/2) + \frac{1}{2} \right), \end{aligned} \quad (\text{C33})$$

which is just Eq. (9).

APPENDIX D: CLASSICAL PARAMETRIC AMPLIFICATION VIA SECOND-ORDER NONLINEARITY

1. Infinite plane

General—For the subharmonic generation to happen in inversion symmetric systems, there needs to be a strong pump

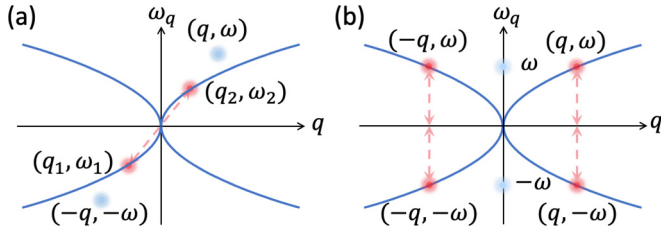


FIG. 11. (a) Schematic illustration of 2D plasmons parametrically driven by the pump through second-order nonlinearity. The blue curve is the plasmon frequency momentum dispersion. The red dots are the two plasmon modes coupled by the pump field at (q, ω) and its complex conjugate at $(-q, -\omega)$. (b) Degenerate FWM of plasmons induced by third-order nonlinearity. The dots at zero momentum represent the uniform pump field with frequency ω and its complex conjugate. The dots at the same frequency but nonzero momentum represent the plasmon modes coupled by this field through third-order nonlinearity.

field with frequency ω and appreciable wave vector \mathbf{q} whose electric potential is

$$\Phi = \phi e^{i(\mathbf{q}\mathbf{r} - \omega t)} + \phi^* e^{-i(\mathbf{q}\mathbf{r} - \omega t)}. \quad (\text{D1})$$

The equation of motion (EOM) of the plasmons is conveniently described by the second-order nonlinear conductivity

$$j_\mu(\omega, \mathbf{q}) = \int d\omega' d\mathbf{q}' \sigma_{\mu\alpha\beta}^{(2)}(\omega - \omega', \mathbf{q} - \mathbf{q}', \omega', \mathbf{q}') \times E_\alpha(\omega - \omega', \mathbf{q} - \mathbf{q}') E_\beta(\omega', \mathbf{q}'). \quad (\text{D2})$$

Since $\sigma^{(2)}$ is a real function in space-time domain, its frequency momentum representation has the general property

$$\sigma_{\mu\alpha\beta}^{(2)}(-\omega_1, -\mathbf{q}_1, -\omega_2, -\mathbf{q}_2) = \sigma_{\mu\alpha\beta}^{(2)*}(\omega_1, \mathbf{q}_1, \omega_2, \mathbf{q}_2). \quad (\text{D3})$$

And if the system has inversion symmetry, then

$$\sigma_{\mu\alpha\beta}^{(2)}(\omega_1, -\mathbf{q}_1, \omega_2, -\mathbf{q}_2) = -\sigma_{\mu\alpha\beta}^{(2)}(\omega_1, \mathbf{q}_1, \omega_2, \mathbf{q}_2), \quad (\text{D4})$$

which means $\sigma_{\mu\alpha\beta}^{(2)}(\omega_1, 0, \omega_2, 0) = 0$. However, it can be nonzero at finite momentum. By convention, $\sigma^{(2)}$ is defined to be symmetric: $\sigma_{\mu\alpha\beta}^{(2)}(\omega_1, \mathbf{q}_1, \omega_2, \mathbf{q}_2) = \sigma_{\mu\beta\alpha}^{(2)}(\omega_2, \mathbf{q}_2, \omega_1, \mathbf{q}_1)$.

Due to $\sigma^{(2)}$ and the driving field, the equation of motion for plasmons is modified. As illustrated in Fig. 11(a), any two modes $(\mathbf{q}_1, \mathbf{q}_2)$ with $\mathbf{q}_1 + \mathbf{q} = \mathbf{q}_2$ are coupled by the driving field through $\sigma^{(2)}$. The charge continuity equation for the

$$\sigma_{ilm}^{(2)}(\omega_1, \mathbf{q}_1, \omega_2, \mathbf{q}_2) = \frac{D_k^{(2)}}{\omega_1 \omega_2 \omega_3} \left[\left(q_3 - \frac{\omega_1}{2\omega_2} q_2 - \frac{\omega_2}{2\omega_1} q_1 \right)_i \delta_{lm} + \left(2q_2 - 2q_1 - \frac{\omega_2}{\omega_1} q_1 + \frac{3\omega_1}{\omega_2} q_2 \right)_m \delta_{il} \right] + \begin{pmatrix} 1 & \leftrightarrow & 2 \\ l & \leftrightarrow & m \end{pmatrix}, \quad (\text{D13})$$

where $(\omega_3, \mathbf{q}_3) = (\omega_1 + \omega_2, \mathbf{q}_1 + \mathbf{q}_2)$ and the second-order optical weight is $D_k^{(2)} = D_0^{(2)} \tanh(\mu/(2T))$ [23] whose zero-temperature limit is $D_0^{(2)}$. Applied to Eq. (D10) in the ideal

mode with momentum \mathbf{q}_2 is

$$\partial_t \rho_{\mathbf{q}_2} + \nabla \cdot \mathbf{j}_{\mathbf{q}_2} = 0 \quad (\text{D5})$$

and the current $\mathbf{j}_{\mathbf{q}_2}$ can be expanded to second order in electric field

$$j_\mu(\omega_2, \mathbf{q}_2) = \sigma_{\mu\nu}(\omega_2, \mathbf{q}_2) E^\nu(\omega_2, \mathbf{q}_2) + \sigma_{\mu\alpha\beta}^{(2)}(\omega_1, \mathbf{q}_1, \omega, \mathbf{q}) \times E^\alpha(\omega_1, \mathbf{q}_1) E^\beta(\omega, \mathbf{q}) + \sigma_{\mu\alpha\beta}^{(2)}(\omega, \mathbf{q}, \omega_1, \mathbf{q}_1) \times E^\alpha(\omega, \mathbf{q}) E^\beta(\omega_1, \mathbf{q}_1). \quad (\text{D6})$$

The electric field $E(\omega_2, \mathbf{q}_2)$ is related to the charge density through the Coulomb kernel

$$\mathbf{E}(\omega_2, \mathbf{q}_2) = -i\mathbf{q}_2 V_{q_2} \rho(\omega_2, \mathbf{q}_2). \quad (\text{D7})$$

Similar set of equations apply to the pairing mode $\rho_{\mathbf{q}_1}$. Separating the densities into products of amplitude and phase as

$$\rho_{\mathbf{q}_1} = A_1(t) e^{i(\mathbf{q}_1 \mathbf{r} - \omega_1 t)}, \quad \rho_{\mathbf{q}_2} = A_2(t) e^{i(\mathbf{q}_2 \mathbf{r} - \omega_2 t)}, \quad (\text{D8})$$

and plugging them into the continuity equations [Eq. (D5) *et al.*], we get the equations for time evolution equation of the amplitudes:

$$\begin{pmatrix} \partial_t + \gamma & \kappa_2 e^{i\Delta t} \\ \kappa_1 e^{-i\Delta t} & \partial_t + \gamma \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = 0. \quad (\text{D9})$$

The parameters κ_1, κ_2 are

$$\begin{aligned} \kappa_1 &= iq_{2\mu} \sigma_{\mu\alpha\beta}^{(2)}(\omega_1, \mathbf{q}_1, \omega, \mathbf{q}) (-iq_{1\alpha} V_{q_1}) (-iq_\beta \phi) \\ &\quad + iq_{2\mu} \sigma_{\mu\alpha\beta}^{(2)}(\omega, \mathbf{q}, \omega_1, \mathbf{q}_1) (-iq_\alpha \phi) (-iq_{1\beta} V_{q_1}) \\ &= -i2\phi V_{q_1} \sigma_{\mu\alpha\beta}^{(2)}(\omega_1, \mathbf{q}_1, \omega, \mathbf{q}) q_{2\mu} q_{1\alpha} q_\beta, \\ \kappa_2 &= i2\phi^* V_{q_2} \sigma_{\mu\alpha\beta}^{(2)}(\omega_2, \mathbf{q}_2, -\omega, -\mathbf{q}) q_{1\mu} q_{2\alpha} q_\beta. \end{aligned} \quad (\text{D10})$$

Similar to Eq. (B12), the solution to Eq. (D9) has an exponential factor $e^{\lambda t}$. For there to be nonzero solutions, we require the determinant to be zero, and obtain

$$\lambda = -\gamma \pm \sqrt{\kappa_1 \kappa_2 - \Delta^2/2} \quad (\text{D11})$$

where $\Delta = \omega - (\omega_2 - \omega_1)$. Therefore, the criterion for there to be an exponentially growing solution is

$$\text{Re}[\sqrt{\kappa_1 \kappa_2 - \Delta^2/2}] > \gamma. \quad (\text{D12})$$

Let us temporarily assume that there is no damping $\gamma = 0$. In the ideal case, $\mathbf{q}_1 = -\mathbf{q}_2$, and $\Delta = 0$, we have $\kappa_2 = \kappa_1^*$ from the property Eq. (D3). Therefore, Eq. (D12) is guaranteed to be satisfied in any dissipationless system.

Applied to Graphene—In the kinetic regime of graphene, expanding to linear order in q , the second-order nonlinear conductivity reads [20,23]

case $\mathbf{q}_1 = -\mathbf{q}_2$, and $\Delta = 0$, we obtain the growth rate

$$\sqrt{\kappa_1 \kappa_2} = \frac{3}{2} \frac{e^3 v_F^2}{\hbar^2} \frac{q^3}{\omega^3} |\phi|. \quad (\text{D14})$$

Taking into account the plasmon dispersion $\omega_1 = \sqrt{2Dq_1}$, we arrive at the dimensionless relative growth factor

$$Q_g^{-1} = \frac{3}{16\alpha_g} \zeta. \quad (\text{D15})$$

2. Graphene ribbon

Assume the ω_1 mode is excited by an external source tuned exactly at ω_1 , the density is $\rho(t) = \rho_0 e^{-i\omega_1 t} + \text{c.c.}$. Through second-order nonlinearity, the strong field of mode ω_1 is going to modify the CCE of the ω_0 mode. Assuming the charge density of the subharmonic mode is $\rho_1 = A(t) e^{-i\omega_0 t} + \text{c.c.}$, plugging it into the CCE of this mode and taking inner product with $\tilde{\rho}_1$, one obtains equation for the amplitude:

$$\partial_t A + \kappa A^* = 0, \quad (\text{D16})$$

where $|\tilde{\rho}_1|^2$ is defined in Eq. (C7) and $|\kappa| = |\beta g / \hbar|$. Equation (D16) leads to exponentially growing solution with the relative growth rate

$$Q_g^{-1} = \frac{|\kappa|}{\omega_0} = \zeta \frac{Q}{8\alpha_g} \frac{\eta' \alpha'}{\xi_0^2 S_1' S_0'}, \quad (\text{D17})$$

which agrees with Eq. (17), the quantum mechanical result.

APPENDIX E: CLASSICAL PARAMETRIC AMPLIFICATION VIA THIRD-ORDER NONLINEARITY

Assume a strong uniform electric field $\mathbf{E} = \mathbf{E}_0 e^{-i\omega t} + \text{c.c.}$ which, due to momentum mismatch, cannot excite plasmons through linear response. However, through third-order nonlinear effect, this uniform \mathbf{E} tends to make the plasmons grow in amplitude. The CCE of the plasmons with momentum q is

$$\partial_t \rho_{\mathbf{q}} + \nabla \cdot \mathbf{j}_{\mathbf{q}} = 0, \quad (\text{E1})$$

and the currents are

$$\begin{aligned} j(\mathbf{q}, \omega)_i &= \sigma_{il}(\omega) E(\mathbf{q}, \omega)_l + 3\sigma_{ilmn}^{(3)}(-\omega, \omega, \omega) \\ &\quad \times E(\mathbf{q}, -\omega)_l E_{0m} E_{0n}, \\ j(\mathbf{q}, -\omega)_i &= \sigma_{il}(-\omega) E(\mathbf{q}, -\omega)_l + 3\sigma_{ilmn}^{(3)}(\omega, -\omega, -\omega) \\ &\quad \times E(\mathbf{q}, \omega)_l E_{0m}^* E_{0n}^*, \end{aligned} \quad (\text{E2})$$

where $E(\mathbf{q})_l = -iq_l v_q \rho_{\mathbf{q}}$. Thus the CCEs of the plasmon mode q at frequencies ω and $-\omega$ are coupled by the pump field E_0 , as shown in Fig. 11(b). Separating amplitude and phase of the charge density as

$$\rho_{\mathbf{q}}(t) = A(t) e^{-i\omega t} + B(t) e^{i\omega t}, \quad (\text{E3})$$

we arrive at the equation for the amplitudes

$$\begin{pmatrix} 2\partial_t & \lambda \\ \lambda' & 2\partial_t \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = 0, \quad (\text{E4})$$

where

$$\begin{aligned} \lambda &= 3\sigma_{ilmn}^{(3)}(-\omega, \omega, \omega) q_i q_l v_q E_{0m} E_{0n}, \\ \lambda' &= 3\sigma_{ilmn}^{(3)}(\omega, -\omega, -\omega) q_i q_l v_q E_{0m}^* E_{0n}^*. \end{aligned} \quad (\text{E5})$$

Due to the fact that $\sigma_{ilmn}^{(3)}(\omega, -\omega, -\omega) = \sigma_{ilmn}^{(3)*}(-\omega, \omega, \omega)$, we have $\lambda' = \lambda^*$. Therefore, Eq. (E4) has the solution

$$\begin{aligned} A(t) &= A_+ e^{\kappa t} + A_- e^{-\kappa t}, \\ B(t) &= B_+ e^{\kappa t} + B_- e^{-\kappa t}, \end{aligned} \quad (\text{E6})$$

where $\kappa = |\lambda|$ is the growth/decay rate. Taking the ‘‘Drude’’ form in Eq. (21) for $\sigma^{(3)}$ of graphene in the kinetic regime, plugging in the square root dispersion of plasmons $\omega = \sqrt{2Dq}$, the dimensionless relative growth rate $Q_g^{-1} = \kappa/\omega$ simplifies to

$$Q_g^{-1} = \frac{3}{8} \frac{e^2 E_0^2 / \omega^2}{(\hbar k_F)^2} = \frac{3}{8} \xi^2 \quad (\text{E7})$$

for the plasmons propagating parallel to the direction of the pump field. The dimensionless small number $\xi = \frac{eE_0/\omega}{\hbar k_F} = \delta p / p_F$ is simple to memorize: δp is just the change in electron momentum caused by the electric field during one half cycle of the oscillations, $\delta t \sim \pi/\omega$ [24]. Note that due to third-order nonlinearity, the plasmon frequency itself is also renormalized by this strong uniform field [24], and the frequency ω throughout this section should be considered as the renormalized one.

Although a result of third-order nonlinearity, this phenomenon is different from modulational instability, which is ubiquitous in nonlinear optics and fluid mechanics [131], e.g., in surface gravity waves. In modulational instability, the strong pump is a finite momentum ‘‘wave train’’ on resonance, and the exponentially growing waves have momentums different from the pump by a small amount δk . If the criterion of instability is satisfied, the growth rate scales linearly with δk . For a negative Kerr nonlinearity as for graphene plasmons [24], the criterion requires $\partial_k^2 \omega > 0$, not satisfied by the square root dispersion. However, this criterion assumes a non dispersive $\sigma^{(3)}$, which is not true in graphene. Therefore, whether modulational instability happens in graphene needs further investigation and is a weaker effect anyway. Indeed, in a recent work on nonlinear plasmons [132], it was found that second-order nonlinearity could lead to growth of side bands in the presence of a wave train, similar to modulational instability.

APPENDIX F: NEAR-FIELD PROBE OF DFG

In this section, we discuss a scanning near-field experiment that could measure plasmonic DFG by classical interference. The setup is similar to the left part of Fig. 6. Upon pumping of the ω_1 mode of the device Fig. 1(a), if the ω_0 mode with momentum q is launched by a classical source combined with an ‘‘antenna’’, e.g., the left edge of the ribbon, the counter propagating mode with momentum $-q$ will be generated by DFG from the pump and the q mode. The q mode can be described by a coherent state $|a_0\rangle$, which satisfies $a_q |a_0\rangle = a_0 |a_0\rangle$. The Hamiltonian (7) leads to equation of motion for the $-q$ mode:

$$(\partial_t + \gamma) \langle a_{-q} \rangle = -i\omega_0 \langle a_{-q} \rangle - ig^{(2)} \beta a_0^* e^{-i\omega_0 t} / \hbar, \quad (\text{F1})$$

where the phenomenological damping rate γ has been added. After the driving term has been turned on for a time long enough, the steady state solution is

$$\langle a_{-q} \rangle = i \frac{g^{(2)}\beta}{\hbar\gamma} a_0^* e^{-i\omega_0 t} = i \frac{\kappa}{\gamma} a_0^* e^{-i\omega_0 t}. \quad (\text{F2})$$

Therefore, if the “reflection coefficient” κ/γ is at the order of one, the two waves would interfere to form fringes with period $\lambda_0/2$ where $\lambda_0 = 2\pi/q$ is the wavelength of the subharmonic plasmon. These fringes could be picked up by the near-field scanning probe.

-
- [1] F. H. L. Koppens, D. E. Chang, and F. J. García de Abajo, Graphene plasmonics: A platform for strong light–matter interactions, *Nano Lett.* **11**, 3370 (2011).
- [2] D. N. Basov, M. M. Fogler, A. Lanzara, F. Wang, and Y. Zhang, Colloquium: Graphene spectroscopy, *Rev. Mod. Phys.* **86**, 959 (2014).
- [3] Z. Fei, A. S. Rodin, G. O. Andreev, W. Bao, A. S. McLeod, M. Wagner, L. M. Zhang, Z. Zhao, M. Thiemens, G. Dominguez *et al.*, Gate-tuning of graphene plasmons revealed by infrared nano-imaging, *Nature (London)* **487**, 82 (2012).
- [4] J. Chen, M. Badioli, P. Alonso-González, S. Thongrattanasiri, F. Huth, J. Osmond, M. Spasenović, A. Centeno, A. Pesquera, P. Godignon *et al.*, Optical nano-imaging of gate-tunable graphene plasmons, *Nature (London)* **487**, 77 (2012).
- [5] A. Woessner, M. B. Lundberg, Y. Gao, A. Principi, P. Alonso-Gonzalez, M. Carrega, K. Watanabe, T. Taniguchi, G. Vignale, M. Polini *et al.*, Highly confined low-loss plasmons in graphene-boron nitride heterostructures, *Nat. Mater.* **14**, 421 (2015).
- [6] Z. Shi, X. Hong, H. A. Bechtel, B. Zeng, M. C. Martin, K. Watanabe, T. Taniguchi, Y.-R. Shen, and F. Wang, Observation of a Luttinger-liquid plasmon in metallic single-walled carbon nanotubes, *Nat. Photon.* **9**, 515 (2015).
- [7] G. X. Ni, L. Wang, M. D. Goldflam, M. Wagner, Z. Fei, A. S. McLeod, M. K. Liu, F. Keilmann, B. Özyilmaz, A. H. Castro Neto *et al.*, Ultrafast optical switching of infrared plasmon polaritons in high-mobility graphene, *Nat. Photon.* **10**, 244 (2016).
- [8] M. B. Lundberg, Y. Gao, R. Asgari, C. Tan, B. Van Duppen, M. Autore, P. Alonso-González, A. Woessner, K. Watanabe, T. Taniguchi *et al.*, Tuning quantum nonlocal effects in graphene plasmonics, *Science* **357**, 187 (2017).
- [9] G. X. Ni, A. S. McLeod, Z. Sun, L. Wang, L. Xiong, K. W. Post, S. S. Sunku, B.-Y. Jiang, J. Hone, C. R. Dean *et al.*, Fundamental limits to graphene plasmonics, *Nature (London)* **557**, 530 (2018).
- [10] S. Wang, S. Zhao, Z. Shi, F. Wu, Z. Zhao, L. Jiang, K. Watanabe, T. Taniguchi, A. Zettl, C. Zhou, and F. Wang, Nonlinear Luttinger liquid plasmons in semiconducting single-walled carbon nanotubes, *Nat. Mater.* **19**, 986 (2020).
- [11] Y. Dong, L. Xiong, I. Y. Phinney, Z. Sun, R. Jing, A. S. McLeod, S. Zhang, S. Liu, F. L. Ruta, H. Gao *et al.*, Fizeau drag in graphene plasmonics, *Nature (London)* **594**, 513 (2021).
- [12] W. Zhao, S. Zhao, H. Li, S. Wang, S. Wang, M. I. B. Utama, S. Kahn, Y. Jiang, X. Xiao, S. Yoo *et al.*, Efficient Fizeau drag from dirac electrons in monolayer graphene, *Nature (London)* **594**, 517 (2021).
- [13] S. Wang, S. Yoo, S. Zhao, W. Zhao, S. Kahn, D. Cui, F. Wu, L. Jiang, M. I. B. Utama, H. Li *et al.*, Gate-tunable plasmons in mixed-dimensional van der Waals heterostructures, *Nat. Commun.* **12**, 5039 (2021).
- [14] H. Hu, R. Yu, H. Teng, D. Hu, N. Chen, Y. Qu, X. Yang, X. Chen, A. S. McLeod, P. Alonso-González, X. Guo *et al.*, Active control of micrometer plasmon propagation in suspended graphene, *Nat. Commun.* **13**, 1465 (2022).
- [15] S. A. Mikhailov, Non-linear electromagnetic response of graphene, *Europhys. Lett.* **79**, 27002 (2007).
- [16] S. A. Mikhailov, Theory of the giant plasmon-enhanced second-harmonic generation in graphene and semiconductor two-dimensional electron systems, *Phys. Rev. B* **84**, 045432 (2011).
- [17] M. T. Manzoni, I. Silveiro, F. J. G. D. Abajo, and D. E. Chang, Second-order quantum nonlinear optical processes in single graphene nanostructures and arrays, *New J. Phys.* **17**, 83031 (2015).
- [18] M. Glazov and S. Ganichev, High frequency electric field induced nonlinear effects in graphene, *Phys. Rep.* **535**, 101 (2014).
- [19] J. L. Cheng, N. Vermeulen, and J. E. Sipe, Third order optical nonlinearity of graphene, *New J. Phys.* **16**, 053014 (2014).
- [20] Y. Wang, M. Tokman, and A. Belyanin, Second-order nonlinear optical response of graphene, *Phys. Rev. B* **94**, 195442 (2016).
- [21] S. A. Mikhailov, Influence of optical nonlinearities on plasma waves in graphene, *ACS Photon.* **4**, 3018 (2017).
- [22] S. A. Mikhailov, Nonperturbative quasiclassical theory of the nonlinear electrodynamic response of graphene, *Phys. Rev. B* **95**, 085432 (2017).
- [23] Z. Sun, D. N. Basov, and M. M. Fogler, Universal linear and nonlinear electrodynamic dynamics of a dirac fluid, *Proc. Natl. Acad. Sci. USA* **115**, 3285 (2018).
- [24] Z. Sun, D. N. Basov, and M. M. Fogler, Third-order optical conductivity of an electron fluid, *Phys. Rev. B* **97**, 075432 (2018).
- [25] T. J. Constant, S. M. Hornett, D. E. Chang, and E. Hendry, All-optical generation of surface plasmons in graphene, *Nat. Phys.* **12**, 124 (2015).
- [26] B. Yao, Y. Liu, S.-W. Huang, C. Choi, Z. Xie, J. Flor Flores, Y. Wu, M. Yu, D.-L. Kwong, Y. Huang *et al.*, Broadband gate-tunable terahertz plasmons in graphene heterostructures, *Nat. Photon.* **12**, 22 (2018).
- [27] N. Kroo, S. Varró, G. Farkas, P. Dombi, D. Oszetzky, A. Nagy, and A. Czitrovsky, Nonlinear plasmonics, *J. Mod. Opt.* **55**, 3203 (2008).
- [28] M. Gullans, D. E. Chang, F. H. L. Koppens, F. J. Garcia de Abajo, and M. D. Lukin, Single-Photon Nonlinear Optics with Graphene Plasmons, *Phys. Rev. Lett.* **111**, 247401 (2013).
- [29] M. Kauranen and A. V. Zayats, Nonlinear plasmonics, *Nat. Photon.* **6**, 737 (2012).

- [30] J. D. Cox, A. Marini, and F. J. G. de Abajo, Plasmon-assisted high-harmonic generation in graphene, *Nat. Commun.* **8**, 14380 (2017).
- [31] M. S. Tame, K. R. McEnery, S. K. Özdemir, J. Lee, S. A. Maier, and M. S. Kim, Quantum plasmonics, *Nat. Phys.* **9**, 329 (2013).
- [32] A. Reserbat-Plantey, I. Epstein, I. Torre, A. T. Costa, P. A. D. Gonçalves, N. A. Mortensen, M. Polini, J. C. W. Song, N. M. R. Peres, and F. H. L. Koppens, Quantum nanophotonics in two-dimensional materials, *ACS Photon.* **8**, 85 (2021).
- [33] R. W. Boyd, *Nonlinear Optics* (Elsevier Science, Amsterdam, 2020).
- [34] M. D. Reid, P. D. Drummond, W. P. Bowen, E. G. Cavalcanti, P. K. Lam, H. A. Bachor, U. L. Andersen, and G. Leuchs, Colloquium: The Einstein-Podolsky-Rosen paradox: From concepts to applications, *Rev. Mod. Phys.* **81**, 1727 (2009).
- [35] J.-W. Pan, Z.-B. Chen, C.-Y. Lu, H. Weinfurter, A. Zeilinger, and M. Żukowski, Multiphoton entanglement and interferometry, *Rev. Mod. Phys.* **84**, 777 (2012).
- [36] A. Orioux, M. A. M. Versteegh, K. D. Jöns, and S. Ducci, Semiconductor devices for entangled photon pair generation: A review, *Rep. Prog. Phys.* **80**, 076001 (2017).
- [37] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih, New High-Intensity Source of Polarization-Entangled Photon Pairs, *Phys. Rev. Lett.* **75**, 4337 (1995).
- [38] H. Wang, H. Hu, T.-H. Chung, J. Qin, X. Yang, J.-P. Li, R.-Z. Liu, H.-S. Zhong, Y.-M. He, X. Ding, Y.-H. Deng, Q. Dai, Y.-H. Huo, S. Höfling, C.-Y. Lu, and J.-W. Pan, On-Demand Semiconductor Source of Entangled Photons Which Simultaneously Has High Fidelity, Efficiency, and Indistinguishability, *Phys. Rev. Lett.* **122**, 113602 (2019).
- [39] A. De Rossi and V. Berger, Counterpropagating Twin Photons by Parametric Fluorescence, *Phys. Rev. Lett.* **88**, 043901 (2002).
- [40] T. J. Steiner, J. E. Castro, L. Chang, Q. Dang, W. Xie, J. Norman, J. E. Bowers, and G. Moody, Ultrabright entangled-photon-pair generation from an AlGaAs-on-insulator microring resonator, *PRX Quantum* **2**, 010337 (2021).
- [41] U. A. Javid, J. Ling, J. Staffa, M. Li, Y. He, and Q. Lin, Ultrabroadband Entangled Photons on a Nanophotonic Chip, *Phys. Rev. Lett.* **127**, 183601 (2021).
- [42] L. M. Zhang and M. M. Fogler, Nonlinear Screening and Ballistic Transport in a Graphene p - n Junction, *Phys. Rev. Lett.* **100**, 116804 (2008).
- [43] L. Vicarelli, M. S. Vitiello, D. Coquillat, A. Lombardo, A. C. Ferrari, W. Knap, M. Polini, V. Pellegrini, and A. Tredicucci, Graphene field-effect transistors as room-temperature terahertz detectors, *Nat. Mater.* **11**, 865 (2012).
- [44] A. Woessner, R. Parret, D. Davydovskaya, Y. Gao, J.-S. Wu, M. B. Lundeberg, S. Nanot, P. Alonso-Gonzalez, K. Watanabe, T. Taniguchi *et al.*, Electrical detection of hyperbolic phonon-polaritons in heterostructures of graphene and boron nitride, *npj 2D Mater. Appl.* **1**, 25 (2017).
- [45] P. Alonso-González, A. Y. Nikitin, Y. Gao, A. Woessner, M. B. Lundeberg, A. Principi, N. Forcellini, W. Yan, S. Vélez, A. J. Huber *et al.*, Acoustic terahertz graphene plasmons revealed by photocurrent nanoscopy, *Nat. Nanotechnol.* **12**, 31 (2017).
- [46] S. Castilla, B. Terrés, M. Autore, L. Viti, J. Li, A. Y. Nikitin, I. Vangelidis, K. Watanabe, T. Taniguchi, E. Lidorikis *et al.*, Fast and sensitive terahertz detection using an antenna-integrated graphene pn junction, *Nano Lett.* **19**, 2765 (2019).
- [47] Y. J. Ding, S. J. Lee, and J. B. Khurgin, Transversely Pumped Counterpropagating Optical Parametric Oscillation and Amplification, *Phys. Rev. Lett.* **75**, 429 (1995).
- [48] S. Saravi, T. Pertsch, and F. Setzpfandt, Generation of Counterpropagating Path-Entangled Photon Pairs in a Single Periodic Waveguide, *Phys. Rev. Lett.* **118**, 183603 (2017).
- [49] B.-Y. Jiang, E. J. Mele, and M. M. Fogler, Theory of plasmon reflection by a 1d junction, *Opt. Express* **26**, 17209 (2018).
- [50] C. Wolff and N. A. Mortensen, Stimulated plasmon polariton scattering, *Nat. Commun.* **11**, 4039 (2020).
- [51] M. Tokman, Y. Wang, I. Oladyshkin, A. R. Kutayiah, and A. Belyanin, Laser-driven parametric instability and generation of entangled photon-plasmon states in graphene, *Phys. Rev. B* **93**, 235422 (2016).
- [52] E. G. Mishchenko, A. V. Shytov, and P. G. Silvestrov, Guided Plasmons in Graphene p - n Junctions, *Phys. Rev. Lett.* **104**, 156806 (2010).
- [53] F. Hu, Y. Luan, Z. Fei, I. Z. Palubski, M. D. Goldflam, S. Dai, J. S. Wu, K. W. Post, G. C. A. M. Janssen, M. M. Fogler, and D. N. Basov, Imaging the localized plasmon resonance modes in graphene nanoribbons, *Nano Lett.* **17**, 5423 (2017).
- [54] J. L. Cheng, N. Vermeulen, and J. E. Sipe, Second order optical nonlinearity of graphene due to electric quadrupole and magnetic dipole effects, *Sci. Rep.* **7**, 43843 (2017).
- [55] A. Orioux, X. Caillet, A. Lemaître, P. Filloux, I. Favero, G. Leo, and S. Ducci, Efficient parametric generation of counterpropagating two-photon states, *J. Opt. Soc. Am. B* **28**, 45 (2011).
- [56] H. Yan, T. Low, W. Zhu, Y. Wu, M. Freitag, X. Li, F. Guinea, P. Avouris, and F. Xia, Damping pathways of mid-infrared plasmons in graphene nanostructures, *Nat. Photon.* **7**, 394 (2013).
- [57] C. M. Caves and B. L. Schumaker, New formalism for two-photon quantum optics. I. Quadrature phases and squeezed states, *Phys. Rev. A* **31**, 3068 (1985).
- [58] A. Martin, T. Guerreiro, A. Tiranov, S. Designolle, F. Fröwis, N. Brunner, M. Huber, and N. Gisin, Quantifying Photonic High-Dimensional Entanglement, *Phys. Rev. Lett.* **118**, 110501 (2017).
- [59] G.-H. Lee, D. K. Efetov, W. Jung, L. Ranzani, E. D. Walsh, T. A. Ohki, T. Taniguchi, K. Watanabe, P. Kim, D. Englund, and K. C. Fong, Graphene-based Josephson junction microwave bolometer, *Nature (London)* **586**, 42 (2020).
- [60] J. C. Howell, R. S. Bennink, S. J. Bentley, and R. W. Boyd, Realization of the Einstein-Podolsky-Rosen Paradox Using Momentum- and Position-Entangled Photons from Spontaneous Parametric Down Conversion, *Phys. Rev. Lett.* **92**, 210403 (2004).
- [61] J.-P. W. MacLean, J. M. Donohue, and K. J. Resch, Direct Characterization of Ultrafast Energy-Time Entangled Photon Pairs, *Phys. Rev. Lett.* **120**, 053601 (2018).
- [62] J. D. Franson, Bell Inequality for Position and Time, *Phys. Rev. Lett.* **62**, 2205 (1989).
- [63] S. Fasel, F. Robin, E. Moreno, D. Erni, N. Gisin, and H. Zbinden, Energy-Time Entanglement Preservation in

- Plasmon-Assisted Light Transmission, *Phys. Rev. Lett.* **94**, 110501 (2005).
- [64] G. Soavi, G. Wang, H. Rostami, D. G. Purdie, D. De Fazio, T. Ma, B. Luo, J. Wang, A. K. Ott, D. Yoon *et al.*, Broadband, electrically tunable third-harmonic generation in graphene, *Nat. Nanotechnol.* **13**, 583 (2018).
- [65] T. Jiang, D. Huang, J. Cheng, X. Fan, Z. Zhang, Y. Shan, Y. Yi, Y. Dai, L. Shi, K. Liu *et al.*, Gate-tunable third-order nonlinear optical response of massless dirac fermions in graphene, *Nat. Photon.* **12**, 430 (2018).
- [66] S. A. Mikhailov, Quantum theory of the third-order nonlinear electrodynamic effects of graphene, *Phys. Rev. B* **93**, 085403 (2016).
- [67] P. P. Absil, J. V. Hryniewicz, B. E. Little, P. S. Cho, R. A. Wilson, L. G. Joneckis, and P.-T. Ho, Wavelength conversion in GaAs micro-ring resonators, *Opt. Lett.* **25**, 554 (2000).
- [68] L. G. Helt, Z. Yang, M. Liscidini, and J. E. Sipe, Spontaneous four-wave mixing in microring resonators, *Opt. Lett.* **35**, 3006 (2010).
- [69] M. Ferrera, L. Razzari, D. Duchesne, R. Morandotti, Z. Yang, M. Liscidini, J. E. Sipe, S. Chu, B. E. Little, and D. J. Moss, Low-power continuous-wave nonlinear optics in doped silica glass integrated waveguide structures, *Nat. Photon.* **2**, 737 (2008).
- [70] J. S. Levy, A. Gondarenko, M. A. Foster, A. C. Turner-Foster, A. L. Gaeta, and M. Lipson, CMOS-compatible multiple-wavelength oscillator for on-chip optical interconnects, *Nat. Photon.* **4**, 37 (2010).
- [71] C. Sirtori, F. Capasso, J. Faist, L. N. Pfeiffer, and K. W. West, Far-infrared generation by doubly resonant difference frequency mixing in a coupled quantum well two-dimensional electron gas system, *Appl. Phys. Lett.* **65**, 445 (1994).
- [72] M. A. Belkin, F. Capasso, A. Belyanin, D. L. Sivco, A. Y. Cho, D. C. Oakley, C. J. Vineis, and G. W. Turner, Terahertz quantum-cascade-laser source based on intracavity difference-frequency generation, *Nat. Photon.* **1**, 288 (2007).
- [73] J. Zhao, C. Ma, M. Rüsing, and S. Mookherjea, High Quality Entangled Photon Pair Generation in Periodically Poled Thin-Film Lithium Niobate Waveguides, *Phys. Rev. Lett.* **124**, 163603 (2020).
- [74] X. Guo, C.-I. Zou, C. Schuck, H. Jung, R. Cheng, and H. X. Tang, Parametric down-conversion photon-pair source on a nanophotonic chip, *Light Sci. Appl.* **6**, e16249 (2017).
- [75] C. Koos, L. Jacome, C. Poulton, J. Leuthold, and W. Freude, Nonlinear silicon-on-insulator waveguides for all-optical signal processing, *Opt. Express* **15**, 5976 (2007).
- [76] C. Ma, X. Wang, V. Anant, A. D. Beyer, M. D. Shaw, and S. Mookherjea, Silicon photonic entangled photon-pair and heralded single photon generation with $\text{CAR} > 12,000$ and $g^{(2)}(0) < 0.006$, *Opt. Express* **25**, 32995 (2017).
- [77] R. R. Kumar, M. Raevskaia, V. Pogoretskii, Y. Jiao, and H. K. Tsang, Entangled photon pair generation from an InP membrane micro-ring resonator, *Appl. Phys. Lett.* **114**, 021104 (2019).
- [78] S. Ramelow, A. Farsi, S. Clemmen, D. Orquiza, K. Luke, M. Lipson, and A. L. Gaeta, Silicon-nitride platform for narrow-band entangled photon generation, [arXiv:1508.04358](https://arxiv.org/abs/1508.04358).
- [79] I. Alonso Calafell, J. D. Cox, M. Radonjić, J. R. M. Saavedra, F. J. García de Abajo, L. A. Rozema, and P. Walther, Quantum computing with graphene plasmons, *npj Quantum Inf.* **5**, 37 (2019).
- [80] A. F. Page, F. Ballout, O. Hess, and J. M. Hamm, Nonequilibrium plasmons with gain in graphene, *Phys. Rev. B* **91**, 075404 (2015).
- [81] Z. Sun, D. N. Basov, and M. M. Fogler, Adiabatic Amplification of Plasmons and Demons in 2D Systems, *Phys. Rev. Lett.* **117**, 076805 (2016).
- [82] S. de Vega and F. J. García de Abajo, Plasmon generation through electron tunneling in graphene, *ACS Photon.* **4**, 2367 (2017).
- [83] J. Wilson, F. Santosa, M. Min, and T. Low, Temporal control of graphene plasmons, *Phys. Rev. B* **98**, 081411(R) (2018).
- [84] S. Rajasekaran, E. Casandru, Y. Laplace, D. Nicoletti, G. D. Gu, S. R. Clark, D. Jaksch, and A. Cavalleri, Parametric amplification of a superconducting plasma wave, *Nat. Phys.* **12**, 1012 (2016).
- [85] P.-A. Moreau, E. Toninelli, T. Gregory, and M. J. Padgett, Imaging with quantum states of light, *Nat. Rev. Phys.* **1**, 367 (2019).
- [86] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, *Nat. Photon.* **12**, 724 (2018).
- [87] X. Yao, M. Tokman, and A. Belyanin, Efficient Nonlinear Generation of THz Plasmons in Graphene and Topological Insulators, *Phys. Rev. Lett.* **112**, 055501 (2014).
- [88] H. Rostami, M. I. Katsnelson, and M. Polini, Theory of plasmonic effects in nonlinear optics: The case of graphene, *Phys. Rev. B* **95**, 035416 (2017).
- [89] C. Wolff, C. Tserkezis, and N. A. Mortensen, Enhanced ponderomotive force in graphene due to interband resonance, *New J. Phys.* **21**, 073046 (2019).
- [90] C. J. Tollerton, J. Bohn, T. J. Constant, S. A. R. Horsley, D. E. Chang, E. Hendry, and D. Z. Li, Origins of all-optical generation of plasmons in graphene, *Sci. Rep.* **9**, 3267 (2019).
- [91] E. Altewischer, M. P. van Exter, and J. P. Woerdman, Plasmon-assisted transmission of entangled photons, *Nature (London)* **418**, 304 (2002).
- [92] T. V. Phan, J. C. W. Song, and L. S. Levitov, Ballistic heat transfer and energy waves in an electron system, [arXiv:1306.4972](https://arxiv.org/abs/1306.4972).
- [93] D. Svintsov, Hydrodynamic-to-ballistic crossover in dirac materials, *Phys. Rev. B* **97**, 121405(R) (2018).
- [94] A. Lucas and S. Das Sarma, Electronic sound modes and plasmons in hydrodynamic two-dimensional metals, *Phys. Rev. B* **97**, 115449 (2018).
- [95] G. Patrick, Y. Chan-Shan, L. Tairu, T. Fanglin, K. Rai, Z. Hai, W. Kenji, T. Takashi, and W. Feng, Quantum-critical conductivity of the Dirac fluid in graphene, *Science* **364**, 158 (2019).
- [96] V. Andreeva, D. A. Bandurin, M. Luskun, and D. Margetis, Dipole excitation of collective modes in viscous two-dimensional electron systems, *Phys. Rev. B* **102**, 205411 (2020).
- [97] B. N. Narozhny, I. V. Gornyi, and M. Titov, Hydrodynamic collective modes in graphene, *Phys. Rev. B* **103**, 115402 (2021).

- [98] A. Levchenko and J. Schmalian, Transport properties of strongly coupled electron–phonon liquids, *Ann. Phys.* **419**, 168218 (2020).
- [99] A. Principi, D. Bandurin, H. Rostami, and M. Polini, Pseudo-Euler equations from nonlinear optics: Plasmon-assisted photodetection beyond hydrodynamics, *Phys. Rev. B* **99**, 075410 (2019).
- [100] D. Svintsov, V. Vyurkov, V. Ryzhii, and T. Otsuji, Hydrodynamic electron transport and nonlinear waves in graphene, *Phys. Rev. B* **88**, 245444 (2013).
- [101] D. N. Basov, M. M. Fogler, and F. J. García de Abajo, Polaritons in van der Waals materials, *Science* **354**, 195 (2016).
- [102] T. Low, A. Chaves, J. D. Caldwell, A. Kumar, N. X. Fang, P. Avouris, T. F. Heinz, F. Guinea, L. Martin-Moreno, and F. Koppens, Polaritons in layered two-dimensional materials, *Nat. Mater.* **16**, 182 (2017).
- [103] D. N. Basov, A. Asenjo-Garcia, P. J. Schuck, X. Zhu, and A. Rubio, Polariton panorama, *Nanophotonics* **10**, 549 (2021).
- [104] S. Dai, W. Fang, N. Rivera, Y. Stehle, B.-Y. Jiang, J. Shen, R. Y. Tay, C. J. Ciccarino, Q. Ma, D. Rodan-Legrain *et al.*, Phonon polaritons in monolayers of hexagonal boron nitride, *Adv. Mater.* **31**, 1806603 (2019).
- [105] N. Li, X. Guo, X. Yang, R. Qi, T. Qiao, Y. Li, R. Shi, Y. Li, K. Liu, Z. Xu *et al.*, Direct observation of highly confined phonon polaritons in suspended monolayer hexagonal boron nitride, *Nat. Mater.* **20**, 43 (2021).
- [106] S. Dai, Z. Fei, Q. Ma, A. S. Rodin, M. Wagner, A. S. McLeod, M. K. Liu, W. Gannett, W. Regan, K. Watanabe *et al.*, Tunable phonon polaritons in atomically thin van der Waals crystals of boron nitride, *Science* **343**, 1125 (2014).
- [107] E. Yoxall, M. Schnell, A. Y. Nikitin, O. Txoperena, A. Woessner, M. B. Lundeberg, F. Casanova, L. E. Hueso, F. H. L. Koppens, and R. Hillenbrand, Direct observation of ultraslow hyperbolic polariton propagation with negative phase velocity, *Nat. Photon.* **9**, 674 (2015).
- [108] P. Li, M. Lewin, A. V. Kretinin, J. D. Caldwell, K. S. Novoselov, T. Taniguchi, K. Watanabe, F. Gaussmann, and T. Taubner, Hyperbolic phonon-polaritons in boron nitride for near-field optical imaging and focusing, *Nat. Commun.* **6**, 7507 (2015).
- [109] Z. Sun, Á. Gutiérrez-Rubio, D. N. Basov, and M. M. Fogler, Hamiltonian optics of hyperbolic polaritons in nanogranules, *Nano Lett.* **15**, 4455 (2015).
- [110] N. Rivera, G. Rosolen, J. D. Joannopoulos, I. Kaminer, and M. Soljačić, Making two-photon processes dominate one-photon processes using mid-IR phonon polaritons, *Proc. Natl. Acad. Sci. USA* **114**, 13607 (2017).
- [111] A. J. Giles, S. Dai, I. Vurgaftman, T. Hoffman, S. Liu, L. Lindsay, C. T. Ellis, N. Assefa, I. Chatzakis, T. L. Reinecke *et al.*, Ultralow-loss polaritons in isotopically pure boron nitride, *Nat. Mater.* **17**, 134 (2018).
- [112] L. V. Brown, M. Davanco, Z. Sun, A. Kretinin, Y. Chen, J. R. Matson, I. Vurgaftman, N. Sharac, A. J. Giles, M. M. Fogler *et al.*, Nanoscale mapping and spectroscopy of nonradiative hyperbolic modes in hexagonal boron nitride nanostructures, *Nano Lett.* **18**, 1628 (2018).
- [113] I.-H. Lee, M. He, X. Zhang, Y. Luo, S. Liu, J. H. Edgar, K. Wang, P. Avouris, T. Low, J. D. Caldwell, and S.-H. Oh, Image polaritons in boron nitride for extreme polariton confinement with low losses, *Nat. Commun.* **11**, 3649 (2020).
- [114] G. Ni, A. S. McLeod, Z. Sun, J. R. Matson, C. F. B. Lo, D. A. Rhodes, F. L. Ruta, S. L. Moore, R. A. Vitalone, R. Cusco *et al.*, Long-lived phonon polaritons in hyperbolic materials, *Nano Lett.* **21**, 5767 (2021).
- [115] S. L. Moore, C. J. Ciccarino, D. Halbertal, L. J. McGilly, N. R. Finney, K. Yao, Y. Shao, G. Ni, A. Sternbach, E. J. Telford *et al.*, Nanoscale lattice dynamics in hexagonal boron nitride moiré superlattices, *Nat. Commun.* **12**, 5741 (2021).
- [116] A. Cartella, T. F. Nova, M. Fechner, R. Merlin, and A. Cavalleri, Parametric amplification of optical phonons, *Proc. Natl. Acad. Sci. USA* **115**, 12148 (2018).
- [117] M. H. Michael, A. von Hoegen, M. Fechner, M. Först, A. Cavalleri, and E. Demler, Parametric resonance of Josephson plasma waves: A theory for optically amplified interlayer superconductivity in $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$, *Phys. Rev. B* **102**, 174505 (2020).
- [118] Z. Sun, M. M. Fogler, D. N. Basov, and A. J. Millis, Collective modes and terahertz near-field response of superconductors, *Phys. Rev. Research* **2**, 023413 (2020).
- [119] P. E. Dolgirev, A. Zong, M. H. Michael, J. B. Curtis, D. Podolsky, A. Cavalleri, and E. Demler, Periodic dynamics in superconductors induced by an impulsive optical quench, [arXiv:2104.07181](https://arxiv.org/abs/2104.07181).
- [120] D. Golež, Z. Sun, Y. Murakami, A. Georges, and A. J. Millis, Nonlinear Spectroscopy of Collective Modes in an Excitonic Insulator, *Phys. Rev. Lett.* **125**, 257601 (2020).
- [121] Z. Sun and A. J. Millis, Bardasis-schrieffer polaritons in excitonic insulators, *Phys. Rev. B* **102**, 041110(R) (2020).
- [122] Y. Murakami, D. Golež, T. Kaneko, A. Koga, A. J. Millis, and P. Werner, Collective modes in excitonic insulators: Effects of electron-phonon coupling and signatures in the optical response, *Phys. Rev. B* **101**, 195118 (2020).
- [123] T. Kaneko, Z. Sun, Y. Murakami, D. Golež, and A. J. Millis, Bulk Photovoltaic Effect Driven by Collective Excitations in a Correlated Insulator, *Phys. Rev. Lett.* **127**, 127402 (2021).
- [124] Z. Sun, T. Kaneko, D. Golež, and A. J. Millis, Second-Order Josephson Effect in Excitonic Insulators, *Phys. Rev. Lett.* **127**, 127702 (2021).
- [125] E. W. Weisstein, Mathieu function, from MathWorld—A Wolfram Web Resource, <https://mathworld.wolfram.com/MathieuFunction.html>.
- [126] F. J. García De Abajo, Colloquium: Light scattering by particle and hole arrays, *Rev. Mod. Phys.* **79**, 1267 (2007).
- [127] W. Wang, T. Christensen, A.-P. Jauho, K. S. Thygesen, M. Wubs, and N. A. Mortensen, Plasmonic eigenmodes in individual and bow-tie graphene nanotriangles, *Sci. Rep.* **5**, 9535 (2015).
- [128] S. Franke, S. Hughes, M. K. Dezfouli, P. T. Kristensen, K. Busch, A. Knorr, and M. Richter, Quantization of Quasinormal Modes for Open Cavities and Plasmonic Cavity Quantum Electrodynamics, *Phys. Rev. Lett.* **122**, 213901 (2019).
- [129] J. Ren, S. Franke, and S. Hughes, Quasinormal Modes, Local Density of States, and Classical Purcell Factors for Coupled Loss-Gain Resonators, *Phys. Rev. X* **11**, 041020 (2021).

- [130] C. Sauvan, T. Wu, R. Zarouf, E. A. Muljarov, and P. Lalanne, Normalization, orthogonality, and completeness of quasinormal modes of open systems: The case of electromagnetism, *Opt. Express* **30**, 6846 (2022).
- [131] G. Falkovich, *Fluid Mechanics: A Short Course for Physicists* (Cambridge University Press, Cambridge, England, 2011).
- [132] B. Eliasson and C. S. Liu, Nonlinear plasmonics in a two-dimensional plasma layer, *New J. Phys.* **18**, 053007 (2016).

Correction: The previously published Figure 2 contained an incorrect unit on the abscissa and has been replaced to reflect the correct unit, V/cm. The same unit has been fixed in the last paragraph of Sec. III C and in the second paragraph of Sec. III D.