

All-Optical Generation of Surface Plasmons in Graphene

T. J. Constant,^{1,*} S. M. Hornett,¹ D. E. Chang,^{2,†} and E. Hendry¹

¹*Electromagnetic Materials Group, Department of Physics,
College of Engineering, Mathematics and Physical Sciences,
University of Exeter, Exeter, Devon, UK. EX4 4QL.*

²*ICFO - Institut de Ciències Fotòniques,
Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain*

(Dated: July 8, 2015)

Surface plasmons in graphene offer a compelling route to many useful photonic technologies [1–3]. As a plasmonic material, graphene offers several intriguing properties, such as excellent electro-optic tunability[4], crystalline stability, large optical nonlinearities[5] and extremely high electromagnetic field concentration[6]. As such, recent demonstrations of surface plasmon excitation in graphene using near-field scattering of infrared light[7, 8] have received intense interest. Here we present an all-optical plasmon coupling scheme which takes advantage of the intrinsic nonlinear optical response of graphene. Free-space, visible light pulses are used to generate surface plasmons in a planar graphene sheet using difference frequency wave mixing to match both the wavevector and energy of the surface wave. By carefully controlling the phase-matching conditions, we show that one can excite surface plasmons with a defined wavevector and direction across a large frequency range, with an estimated photon efficiency in our experiments approaching 10^{-5} .

Graphene has attracted significant interest in recent years as a unique optical material. In particular, it has been predicted and experimentally shown that graphene can support very highly confined surface plasmon modes[1, 9], with an electrically tunable dispersion[7, 8]. Despite these promising discoveries, the burgeoning field of graphene plasmonics has some serious obstacles to overcome if it is to progress from the proof-of-principle stage. Problems arise due to the small wavelength of the surface plasmons, two orders of magnitude smaller than light of the same frequency. This has led to the development of specialised measurement techniques, most of which use infrared light and geometries with scattering resonances [10–12] or near-field sources [7, 8] to excite graphene surface plasmons. However, the far-infrared regime, in which graphene plasmons are predicted to have long lifetimes, lacks developed sources and detectors compared to the visible regime. Alternative approaches, such as the manipulation of surface acoustic waves to couple to the graphene surface plasmons [13, 14], therefore hold promise.

Particularly desirable is the potential to excite a plasmon eigenstate with a singular energy, momentum and direction, vital for many future applications, including plasmonic circuits. In this respect, very recent progress has been made, with the development of carefully designed nano antennas which can locally excite and direct surface plasmons in graphene [11]. Here, the combination of infrared source frequency and nanoantenna dimen-

sions determine the frequency, wavevector and direction of the surface plasmons generated. In this article, we investigate a competing approach that embodies many of these desirable aspects of directivity without requiring careful nanofabrication of antennas. This all-optical approach can access a distinctly broad frequency range, even down to the far infrared. We coherently excite surface plasmons using two visible frequency free-space beams via difference frequency generation (DFG), an effect which we monitor through changes in reflectance, and can tune the frequency and wavevector of the surface plasmon through careful adjustment of incident light sources. This potential to excite and detect plasmons purely with free space optics, and at frequencies different than that of the plasmons themselves, has the potential to significantly expand the technological possibilities for graphene plasmonics.

The intrinsic nonlinear interactions of graphene with light are surprisingly large[5, 15–

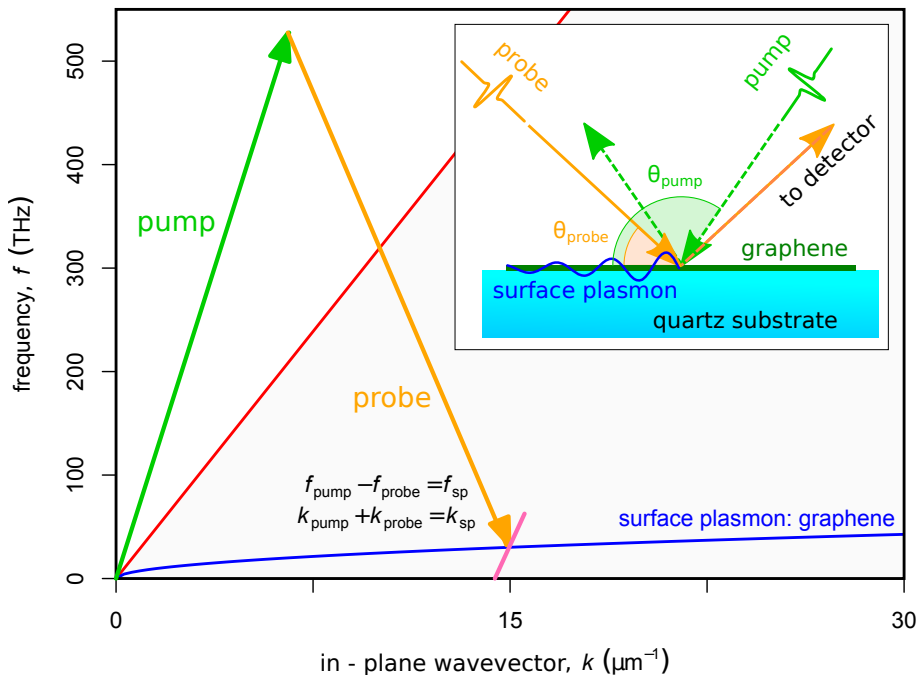


FIG. 1: The nonlinear coupling scheme illustrated on a dispersion diagram. The DFG of the pump (green arrow) and probe (orange arrow) allows access to wavevectors outside of the light line (red line). This permits phase-matching to the surface plasmon modes in graphene (blue line). The pink line illustrates a region that can be interrogated by altering the pump wavelength from 615 nm to 545 nm with the probe wavelength fixed at 615 nm. (Inset) The experimental arrangement used to excite surface plasmons on graphene.

18]. Moreover, large enhancements of nonlinear optical effects are predicted by the presence of highly confined plasmons in graphene[19, 20]. It seems intuitive, then, to attempt the converse: to use the nonlinear interaction between optical fields to resonantly drive surface plasmons. This kind of approach has been demonstrated for thin metallic films [21, 22], and has been recently proposed for graphene, with various coupling schemes for the difference frequency mixing of infra-red light in graphene clad waveguide structures suggested [23, 24]. Similar in concept, figure 1 shows our nonlinear coupling scheme illustrated on a dispersion diagram. By illuminating the graphene with two intense laser pulses with well-defined angles of incidence but different frequency, labeled here f_{pump} and f_{probe} , one can phase match both the frequency and wavevector, k , of the surface plasmon. This wave mixing process is a second order nonlinear effect, normally forbidden in centro-symmetric crystals[25], but possible in graphene because of the distinctively non-local, spatial character of the interaction [20]. The inset in fig. 1 shows the experimental arrangement used. An identical pair of optical parametric amplifiers (OPAs), pumped by an amplified femtosecond laser system, generate the 100 fs-pulses at a repetition rate of 1 kHz. The wavelengths of the two OPAs are selected independently, and the beams are directed to the sample. The experiments are carried out in a non-collinear geometry, using two beams incident on the samples at angles θ_{pump} and θ_{probe} . Oblique angles provide sufficient in-plane momentum to match to the surface plasmon, as illustrated in fig. 1. The incident beams are weakly focused on the sample using 30 cm focal length lenses, giving rise to a very small uncertainty in angle ~ 0.017 rad, and a similarly negligible uncertainty for the in-plane wavevectors. Sets of half-waveplates and polarizers determine both the average power and polarization, with the polarization set such that the electric vector of the light is in the plane of incidence (transverse magnetic polarized). The pump pulse fluence, Φ , used is typically in the range $\Phi \sim 0.1 - 0.2$ mJ/cm², with a pump spot size on the sample of ~ 300 μ m radius. This pump fluence is an order of magnitude less than the photo-modification threshold for graphene[26], and the probe fluence is typically two orders of magnitude smaller still. We measure the differential reflection of the probe beam defined as $\Delta R/R = (R - R_0)/R_0$ where R and R_0 are the reflections with and without the presence of the pump pulse, respectively. In order to isolate the nonlinear reflection signal, we vary the temporal overlap of the two pulses using a motorized delay stage.

For optical excitation pulses, one expects optical nonlinearity arising due to saturable

absorption caused by Pauli blocking of interband transitions[27]. A typical measurement of the temporal dynamics recorded for this process ($\lambda_{pump} = 547$ nm, $\lambda_{probe} = 615$ nm) is shown by the black curve in fig. 2. The asymmetric line shape of the signal is due to the relaxation dynamics of the excited electrons cooling[28, 29], with significant temporal broadening caused by the spatial overlap of the non-collinear beam spots. Note that there is no appreciable signal from the quartz substrate (See Supplementary Information (SI) fig. S1).

For non-degenerate pump and probe beams, in addition to (incoherent) saturable absorption effects, one can expect (coherent) wave mixing signals. This coherent contribution to the probe reflection is expected to be significantly enhanced when the difference frequency field generated by the pump and probe matches that of the graphene surface plasmons. This is analogous to that of a stimulated Raman process, corresponding to a transfer of energy from pump to probe pulses[25] via the generation of surface plasmons. An example of the recorded temporal dynamics under such a resonant condition is presented in fig. 2. Comparing the two curves in this figure, we see that the “non resonant” signal (i.e. when one is not phase matching to plasmon excitation) gives rise to an asymmetric lineshape representative of carrier cooling dynamics. Under “resonant” conditions (i.e. when phase matching conditions are satisfied) we observe a fast additional contribution to the signal, giving rise to a more symmetric lineshape, as one would expect for a coherent signal. For certain experimental geometries and excitation fluences, signal enhancements of up to approximately $\times 4$ can be observed (see supplementary figure S3). It should be noted that, depending on efficiencies, it may be possible to isolate the coherent signal using a heterodyne detection scheme[30], which could also allow detection of a plasmon in a different spatial position than generated.

In order to observe the presence of the coherent signal, we vary the difference frequency in order to isolate any resonant, coherent conditions. In experiment, the pump wavelength is varied from 615 nm to 545 nm, with the probe wavelength set at 615 nm, allowing difference frequencies ranging from 0 to 60 THz. In this way, it is possible to interrogate a section of the surface plasmon dispersion, for example the region illustrated by the pink line in fig. 1. Note that we normalize the signal by pump fluence in order to remove artefacts due to power variation[28]. By altering the experimental geometry, we investigate here three different regions of the dispersion diagram corresponding to ($\theta_{pump} = 55^\circ, \theta_{probe} = 45^\circ$),

($\theta_{pump} = 50^\circ, \theta_{probe} = 70^\circ$) and ($\theta_{pump} = 15^\circ, \theta_{probe} = 125^\circ$).

Figure 3 shows the results of these three measurement geometries, superimposed on the surface plasmon dispersion (black line). The dispersion was calculated according to the model outlined in ref. [31], with the SiO₂ substrate phonon frequencies as given and a Fermi energy of $E_f = 0.5$ eV. This Fermi energy is larger than the expected intrinsic doping of our graphene samples (see Methods for sample details), which we attribute to a significantly raised electron temperature expected under illumination by ultrafast pulses (see SI, fig. S2). Hybridization with the substrate phonons leads to four branches[31, 32]. The overlaid colour

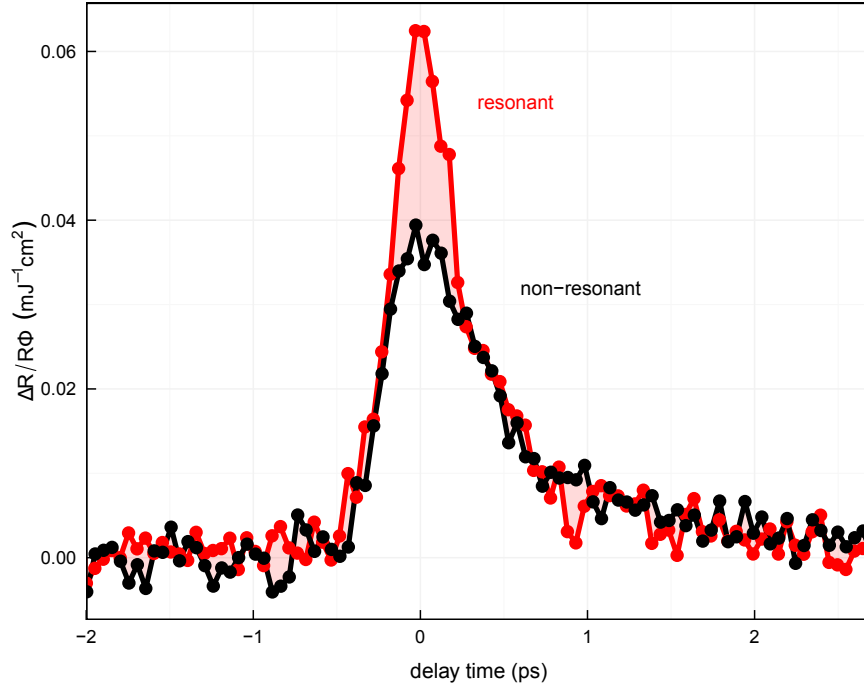


FIG. 2: Differential reflection normalized to average pump fluence as a function of temporal overlap for the geometry $\theta_{pump} = 15^\circ, \theta_{probe} = 125^\circ$. At zero delay time, both the pump and probe pulses arrive simultaneously, leading to a nonlinear change in the probe reflection. Two curves are shown: The black curve labelled “non-resonant” shows a typical time asymmetric measurement when the difference frequency produced by the pump and probe (61.2 THz) does not coincide with a surface plasmon energy state. The red curve shows an additional fast symmetric contribution to the recorded reflection signal when the difference frequency matches the energy of a graphene surface plasmon (23.8 THz).

plots are placed on the diagram so that the maximum differential reflection signal achieved in each delay-scan corresponds to the difference frequency and wavevector of the dataset. The grey shading around the plasmon dispersion curve indicates the expected spectral broadening of the signals (~ 7.5 THz) due to the finite bandwidth of ~ 100 fs-pulses.

Near the regions defined by the surface plasmon dispersion in graphene, we observe clear enhancement in the differential reflection. The assignment of the spectral features to surface plasmon excitation is further supported by the polarization dependence of the signal

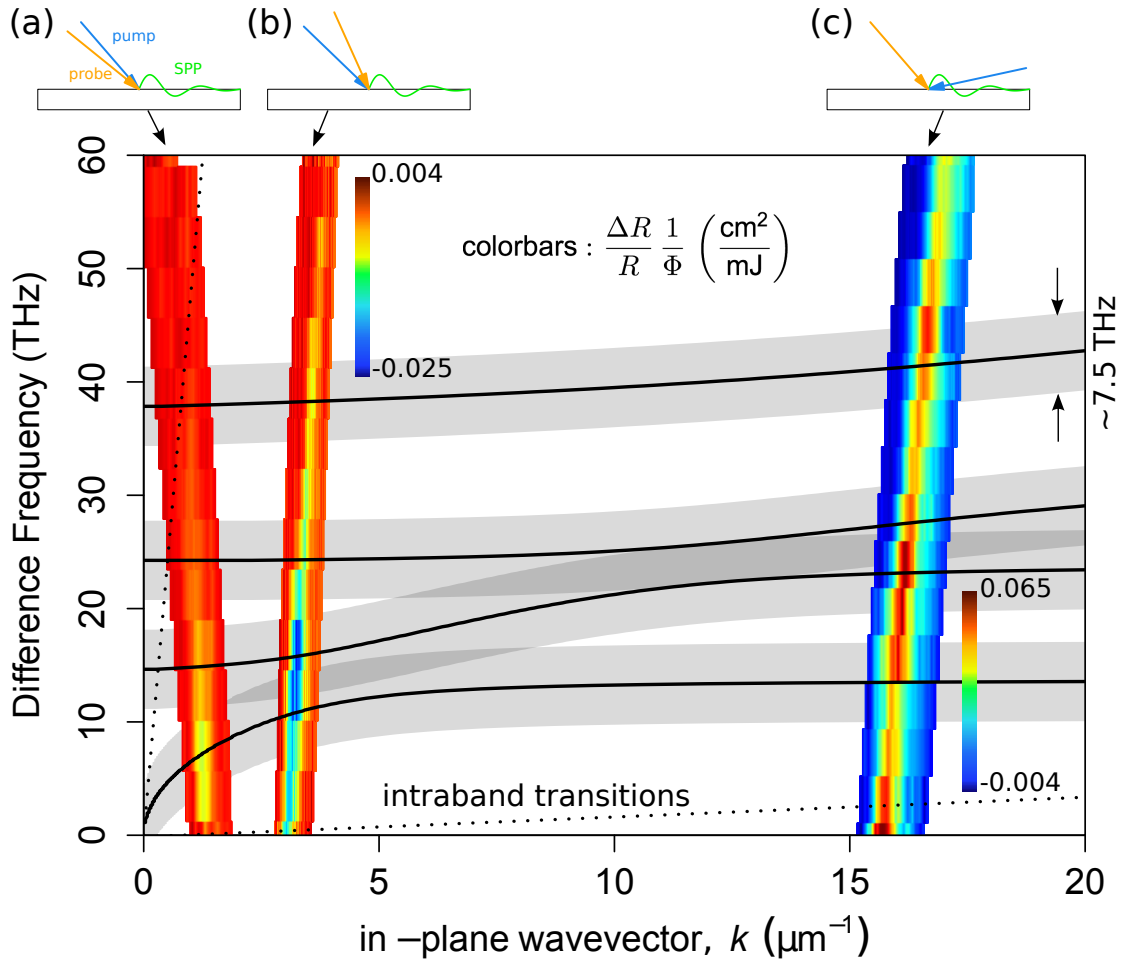


FIG. 3: Plots of differential reflection normalized to pump fluence for the three different experimental geometries, superimposed on the graphene surface plasmon-phonon dispersion. (black lines). Angles used are: (a) $\theta_{pump} = 55^\circ, \theta_{probe} = 45^\circ$, (b) $\theta_{pump} = 50^\circ, \theta_{probe} = 70^\circ$ and (c) $\theta_{pump} = 15^\circ, \theta_{probe} = 125^\circ$. The intraband transition threshold and light line (dotted lines) are labelled on the diagram. The grey regions indicate the spectral resolution of our experiment.

(see fig. S5). The observation of these resonant features over the incoherent background is strongly dependent on the magnitudes of both pump and probe intensities (see SI figures S2 and S3). For larger difference frequencies, up to 150 THz, we do not observe any further resonance features in our spectra (see fig. S4).

The lower branch of the plasmon dispersion relation gives rise to the largest mixing signals for the low wavevector phase-matching (angles (a) and (b) in fig. 3), while the upper frequency branches give rise to largest signals for the high wavevector (angle (c) in fig. 3) region. Whilst we observe clear resonance features in all three of these experimental geometries, we also observe a change in sign of the signal between low (figure 3(a) and (b)) and high (figure 3(c)) wavevector regions. The absolute differential reflectivity signal size also increases with increasing wavevector.

To understand the origin of these coupling behaviours, we have developed a simple theoretical model that captures the salient features of this nonlinear reflection and generation of plasmons, and we briefly summarize the ideas here (details of the model are presented in the supplementary material). In general, equations for the electromagnetic boundary conditions at the air-graphene-substrate interface relate the wavevector and frequency dependent reflection and transmission coefficients $r(k, \omega)$, $t(k, \omega)$ to the graphene current density $J(k, \omega)$. The current density, on the other hand, can be written in terms of the electric field via conductivity functions, which allows the equations to be solved in terms of fields alone. Nonlinear contributions imply that $J(k, \omega)$ depends on fields at other wavevectors and frequencies, which couple the various reflection and transmission coefficients together. For a second-order conductivity $\sigma^{(2)}$, we find that the probe transmission depends on the pump via

$$t_{probe} = \frac{t_{probe}^{(L)}}{1 - A (\sigma^{(2)})^2 |t_{pump}|^2 I_{pump}}, \quad (1)$$

with an analogous equation for the pump transmission (expressions for r are more involved but are directly related to t , see supplementary materials). Here $t_{probe}^{(L)}$ is the linear transmission coefficient, A is a function of linear optical properties and beam angles, and for notational simplicity, dependencies on k, ω are implicit. In Fig. 4, we plot the numerical solution for the differential probe reflectance, normalized by fluence, for the simplified case of continuous plane-wave pump and probe beams. While this simple model ignores the non-equilibrium nature of the excitation, as observed in experiment, we show below that

it is sufficient to describe some of the salient features of our results. Similar to Fig. 3, the differential reflectance is plotted versus difference frequency and in-plane wavevector, whose values are scanned by continuously varying the probe incidence angle and pump wavelength. The probe wavelength and pump angles are fixed at 615 nm and 50° , respectively, and the pump and probe intensities are chosen to be 10 and $0.1 \text{ W}/\mu\text{m}^2$, to closely correspond to the configuration in Fig. 3(b). It can be seen that the simulation qualitatively produces the main features of Fig. 3. In particular, the change in the sign of differential reflectance at the Brewster angle is clearly observed, as is the enhancement of the signal when the difference frequency and wavevector align with the plasmon dispersion relation.

The model also reproduces some of the main features arising from different coupling efficiencies to different bands (fig. 4). Generally, the highest coupling efficiency occurs for the dispersion regions which are most ‘plasmon-like’ in origin (see additionally fig. S8). This is most obvious comparing the data in fig. 3(c) and fig. 4(c), where the coupling to the

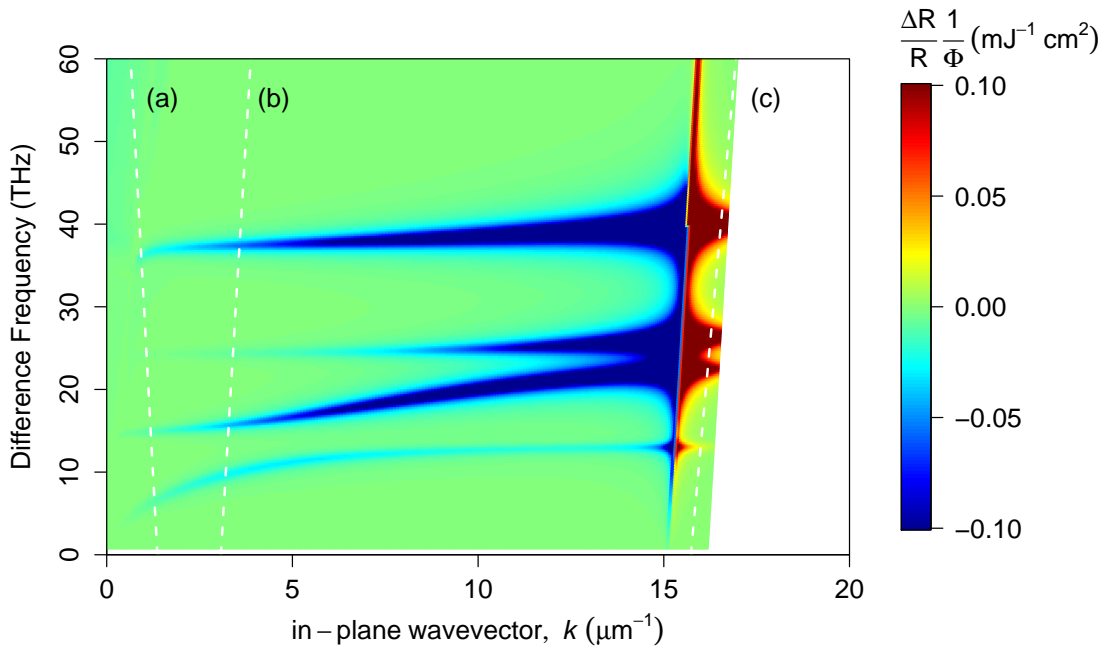


FIG. 4: The numerical solution for the differential probe reflectance normalized by fluence, calculated using the model outlined in the SI. The white dotted lines indicate the region of the dispersion relation probed by the experimental geometries shown in fig. 3.

upper bands is much stronger than the lowest band in both model and experiment. For lower wavevector cases, the coupling to the highest band is overestimated in the model compared to the experiment. This could possibly be caused by frequency-dependant losses in the graphene sheet unaccounted for in our simple model. The model additionally reproduces the increasing absolute signal strength with increasing wavevector observed in experiment, which is a consequence of both larger changes in the reflection coefficient for a corresponding change in absorption for higher angles, and due to spatial dispersion[20] in the signal. Indeed, it can be shown that the magnitude of enhancement is proportional to the square of the plasmon quality factor, Q^2 (see SI), in agreement with predictions from ref. [23].

In addition to the surface plasmon resonance conditions, for the highest wavevector region in fig. 3(c) there is an additional resonant enhancement found at low frequencies < 3 THz measured in experiment that is not reproduced in our model (fig. 4(c)). The position of this peak lies within the expected region of intraband transitions in graphene, indicated by the dotted line in fig. 3. This feature is also largely polarization independent, unlike the enhancements we attribute to surface plasmon coupling (see SI, fig. S5).

In principle, Eq. 1 can be inverted to allow an experimental determination of the nonlinear conductivity $\sigma^{(2)}$, given transmission or reflection data. This is difficult in the present setup, in part given the broad bandwidth of the pulses, uncertainty over some system parameters, and difficulty of investigating a large number of angles to quantify possible wavevector and frequency dependence of $\sigma^{(2)}$. However, as an estimate, we take the simplest possible model, in which the effective nonlinear susceptibility $\chi^{(2)}$ is frequency and wavevector-independent. This corresponds to a nonlinear conductivity function obeying $\sigma^{(2)}(\omega) = i|\sigma^{(2)}(\omega_{probe})|(\omega/\omega_{probe})$, where the value at the (fixed) probe frequency represents a single fitting parameter. We find that a value of $|\sigma^{(2)}(\omega_{probe})| \approx 2.4 \times 10^{-12} \text{ A} \cdot \text{m}/\text{V}^2$ produces the same peak signal as observed in Fig. 3(b). It should be emphasized that this represents a rather conservative estimate of $\sigma^{(2)}(\omega)$. In particular, the mobility of $\mu \approx 2000 \text{ cm}^2/\text{Vs}$ corresponds to a plasmon linewidth of $\gamma = 2\pi \times 1.6 \text{ THz}$ that is narrower than the measurement bandwidth, indicating that only a fraction of the pulse can efficiently excite plasmons. Reducing measurement bandwidth could therefore give rise to greater coupling to the surface plasmons, while also reducing the effects of non-equilibrium carriers on the measurements. While a comparison to a bulk nonlinear crystal is not directly meaningful, it is nonetheless interesting to note that a bulk nonlinear crystal with

the thickness $t \approx 0.3$ nm of a graphene layer would require a nonlinear susceptibility of $\chi^{(2)} \sim |\sigma^{(2)}(\omega_{probe})|/(\epsilon_0\omega_{probe}t) \sim 3 \times 10^{-7}$ m/V to produce the equivalent in-plane nonlinear currents. This value is approximately 3 orders of magnitude larger than in GaAs.

Finally, from the inferred value of $\sigma^{(2)}$ and the input beam parameters, our model enables us to estimate the conversion efficiency η of pump photons to plasmons (see SI). We find a value of $\eta \approx 6 \times 10^{-6}$, while noting that the actual conversion could be significantly higher with narrow pulses, again as the estimated value of $\sigma^{(2)}$ does not account for the large pulse bandwidth. We note that this experimentally obtained value of η is of the same order as predicted in ref. [23], once adjusted for our experimental parameters.

In summary, surface plasmon generation in a planar single layer graphene crystal has been demonstrated using nonlinear wave mixing of visible frequency light. We observe enhancements in the nonlinear, differential reflection of light for various frequencies and phase-matching conditions which agree well with the surface plasmon-phonon dispersion for graphene on quartz substrates. In contrast to near-field infrared scattering[7, 8], our approach may be used to access an distinctly broad frequency range, including the ordinarily hard to reach THz “gap” [33]. Moreover, by careful manipulating the phase-matching conditions, we show that one can generate surface plasmons with a defined wavevector, with an efficiency approaching 10^{-5} . This efficiency by no means represents a fundamental limit, and we believe that it could in principle be pushed towards a 10^{-2} level with future adjustments, such as increasing the surface plasmon Q factor from ~ 5 to ~ 30 with lattice-matched hBN substrates[34], equalizing the intensities of the pump and probe beams (see SI, fig. S3), or the use of narrower bandwidth pulses. Moreover, in principle, our approach could be extended to higher or lower frequencies, regions that are generally hard to access using current approaches[2].

METHODS

Sample Preparation

Samples for our experiments are fabricated from commercially grown CVD graphene on copper foil (graphene supermarket). Transfer to quartz substrates was performed in house via a standard metal etching and float technique using Ammonium persulfate to etch the cop-

per and PMMA as a support structure. Combined resistance and Raman spectroscopy[35] give an estimated mobility of the samples of around $2000 \text{ cm}^2/\text{Vs}$ and a natural Fermi energy of $\sim 300 \text{ meV}$. Raman imaging indicates that the graphene is nominally single layer, with $\sim 80\%$ coverage of the substrate.

ACKNOWLEDGEMENTS

This research has been supported by the European Commission (FP7-ICT-2013-613024-GRASP) and EPSRC fellowship (EP/K041215/1). DEC also acknowledges support from Fundacio Privada Cellex Barcelona and the Ramon y Cajal program. The authors would like to thank Javier García de Abajo and Marinko Jablan for useful discussions, and Nick Cole for technical assistance.

AUTHOR CONTRIBUTIONS

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

FIGURE CAPTIONS

* Electronic address: t.j.constant@ex.ac.uk

† Electronic address: darrick.chang@icfo.es

- [1] A. N. Grigorenko, M. Polini, and K. S. Novoselov, *Nat. Photonics* **6**, 749 (2012), ISSN 1749-4885, 1301.4241, URL <http://dx.doi.org/10.1038/nphoton.2012.262>.
- [2] F. G. de Abajo, *ACS Photonics* **1**, 135 (2014), arXiv:1402.1969v2, URL <http://pubs.acs.org/doi/abs/10.1021/ph400147y>.
- [3] F. H. L. Koppens, D. E. Chang, and F. J. de Abajo, *Nano Letters* **11**, 3370 (2011), URL <http://dx.doi.org/10.1021/nl201771h>.

- [4] M. F. Craciun, S. Russo, M. Yamamoto, and S. Tarucha, *Nano Today* **6**, 42 (2011), ISSN 17480132, URL <http://linkinghub.elsevier.com/retrieve/pii/S1748013210001623>.
- [5] E. Hendry, P. J. Hale, J. Moger, A. K. Savchenko, and S. A. Mikhailov, *Physical Review Letters* **105**, 097401 (2010), ISSN 0031-9007, URL <http://link.aps.org/doi/10.1103/PhysRevLett.105.097401>.
- [6] V. W. Brar, M. S. Jang, M. Sherrott, J. J. Lopez, and H. a. Atwater, *Nano Letters* **13**, 2541 (2013), ISSN 15306984.
- [7] 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., *Nature* **487**, 82 (2012), ISSN 1476-4687, URL <http://www.ncbi.nlm.nih.gov/pubmed/22722866>.
- [8] J. Chen, M. Badioli, P. Alonso-González, S. Thongrattanasiri, F. Huth, J. Osmond, M. Spasenović, A. Centeno, A. Pesquera, P. Godignon, et al., *Nature* **487**, 77 (2012), ISSN 1476-4687, URL <http://www.ncbi.nlm.nih.gov/pubmed/22722861>.
- [9] Y. V. Bludov, A. Ferreira, N. M. R. Peres, and M. I. Vasilevskiy, *International Journal of Modern Physics B* **27**, 1341001 (2013), ISSN 0217-9792, URL <http://www.worldscientific.com/doi/abs/10.1142/S0217979213410014>.
- [10] W. Gao, J. Shu, C. Qiu, and Q. Xu, *ACS Nano* **6**, 7806 (2012), ISSN 19360851.
- [11] P. Alonso-Gonzalez, A. Y. Nikitin, F. Golmar, A. Centeno, A. Pesquera, S. Velez, J. Chen, G. Navickaite, F. Koppens, A. Zurutuza, et al., *Science* **344**, 1369 (2014), ISSN 0036-8075, URL <http://www.sciencemag.org/cgi/doi/10.1126/science.1253202>.
- [12] W. Gao, G. Shi, Z. Jin, J. Shu, Q. Zhang, R. Vajtai, P. M. Ajayan, J. Kono, and Q. Xu, *Nano letters* **13**, 3698 (2013), ISSN 1530-6992, URL <http://www.ncbi.nlm.nih.gov/pubmed/23895501>.
- [13] M. Farhat, S. Guenneau, and H. Bağcı, *Physical Review Letters* **111**, 237404 (2013), ISSN 0031-9007, URL <http://link.aps.org/doi/10.1103/PhysRevLett.111.237404>.
- [14] J. Schiefele, J. Pedrós, F. Sols, F. Calle, and F. Guinea, *Physical Review Letters* **111**, 237405 (2013), ISSN 0031-9007, arXiv:1309.0767v1, URL <http://link.aps.org/doi/10.1103/PhysRevLett.111.237405>.
- [15] M. T. Manzoni, F. J. Garc, D. Abajo, and D. E. Chang, Preprint at <http://arxiv.org/abs/1406.4360> (2014), 1406.4360v1.
- [16] J. J. Dean and H. M. van Driel, *Physical Review B* **82**, 125411 (2010), ISSN 1098-0121, URL

- <http://link.aps.org/doi/10.1103/PhysRevB.82.125411>.
- [17] S.-Y. Hong, J. I. Dadap, N. Petrone, P.-C. Yeh, J. Hone, and R. M. Osgood, *Physical Review X* **3**, 021014 (2013), ISSN 2160-3308, URL <http://link.aps.org/doi/10.1103/PhysRevX.3.021014>.
- [18] T. Gu, N. Petrone, J. F. McMillan, A. van der Zande, M. Yu, G. Q. Lo, D. L. Kwong, J. Hone, and C. W. Wong, *Nature Photonics* **6**, 554 (2012), ISSN 1749-4885, 1205.4333, URL <http://www.nature.com/doi/10.1038/nphoton.2012.147>.
- [19] M. Gullans, D. E. Chang, F. H. L. Koppens, F. J. G. de Abajo, and M. D. Lukin, *Physical Review Letters* **111**, 1 (2013), ISSN 00319007, arXiv:1309.2651v1.
- [20] S. A. Mikhailov, *Phys. Rev. B* **84**, 045432 (2011), ISSN 10980121.
- [21] J. Renger, R. Quidant, N. van Hulst, S. Palomba, and L. Novotny, *Physical Review Letters* **103**, 266802 (2009), ISSN 0031-9007, URL <http://link.aps.org/doi/10.1103/PhysRevLett.103.266802>.
- [22] S. Palomba and L. Novotny, *Physical Review Letters* **101**, 056802 (2008), ISSN 0031-9007, URL <http://link.aps.org/doi/10.1103/PhysRevLett.101.056802>.
- [23] X. Yao, M. Tokman, and A. Belyanin, *Physical Review Letters* **112**, 055501 (2014), ISSN 00319007, arXiv:1308.2005v1, URL <http://arxiv.org/abs/1308.2005><http://link.aps.org/doi/10.1103/PhysRevLett.112.055501>.
- [24] Y. Sun, G. Qiao, and G. Sun, *Optics Express* **22**, 27880 (2014), ISSN 1094-4087, arXiv:1405.3320, URL <http://www.opticsinfobase.org/abstract.cfm?URI=oe-22-23-27880>.
- [25] R. W. Boyd, *Nonlinear Optics* (Elsevier, 2008), third edit ed., ISBN 978-0-12-369470-6.
- [26] E. Alexeev, J. Moger, and E. Hendry, *Applied Physics Letters* **103**, 2011 (2013), ISSN 00036951, 1310.0970.
- [27] D. Brida, A. Tomadin, C. Manzoni, Y. J. Kim, A. Lombardo, S. Milana, R. R. Nair, K. S. Novoselov, A. C. Ferrari, G. Cerullo, et al., *Nature communications* **4**, 1987 (2013), ISSN 2041-1723, URL <http://www.ncbi.nlm.nih.gov/pubmed/23770933>.
- [28] J. M. Dawlaty, S. Shivaraman, M. Chandrashekhara, F. Rana, and M. G. Spencer, *Applied Physics Letters* **92**, 21 (2008), ISSN 00036951.
- [29] P. J. Hale, S. M. Hornett, J. Moger, D. W. Horsell, and E. Hendry, *Phys. Rev. B* **83**, 121404 (2011), ISSN 10980121, 1012.3927.

- [30] P. Borri, F. Romstad, W. Langbein, a. Kelly, J. Mork, and J. Hvam, *Optics express* **7**, 107 (2000), ISSN 1094-4087.
- [31] I. J. Luxmoore, C. H. Gan, P. Q. Liu, F. Valmorra, P. Li, J. Faist, and G. R. Nash, *ACS Photonics* **1**, 1151 (2014), ISSN 2330-4022, URL <http://pubs.acs.org/doi/abs/10.1021/ph500233s>.
- [32] H. Yan, T. Low, W. Zhu, Y. Wu, M. Freitag, X. Li, F. Guinea, P. Avouris, and F. Xia, *Nature Photonics* **7**, 394 (2013), ISSN 1749-4885, URL <http://www.nature.com/doi/abs/10.1038/nphoton.2013.57>.
- [33] M. Tonouchi, *Nature Photonics* **1**, 97 (2007), ISSN 1749-4885.
- [34] A. Woessner, M. B. Lundberg, Y. Gao, A. Principi, P. Alonso-González, M. Carrega, K. Watanabe, T. Taniguchi, G. Vignale, M. Polini, et al., *Nature Materials* **14**, 421 (2014), ISSN 1476-1122, URL <http://www.nature.com/doi/abs/10.1038/nmat4169>.
- [35] J. E. Lee, G. Ahn, J. Shim, Y. S. Lee, and S. Ryu, *Nature Communications* **3**, 1024 (2012), ISSN 2041-1723, URL <http://dx.doi.org/10.1038/ncomms2022>.

All-Optical Generation of Surface Plasmons in Graphene: Supplementary Information

T. J. Constant,^{1,*} S. M. Hornett,¹ D. E. Chang^{1,†} and E. Hendry¹

¹*Electromagnetic Materials Group, Department of Physics, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, Devon, UK. EX4 4QL.*

²*ICFO - Institut de Ciències Fotòniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain*

SUBSTRATE RESPONSE

Differential reflection was recorded as a function of delay time for both the graphene on quartz and for the bare quartz substrate to ascertain any contribution to the nonlinear response from the substrate. The experimental parameters for these data were set to the resonant condition of figure 3(b), with $\theta_{pump} = 50^\circ$, $\theta_{probe} = 70^\circ$ and the difference frequency set to 10 THz. The measurements are compared in figure S1, and negligible signal, compared to the resonant measurement from graphene, is observed.

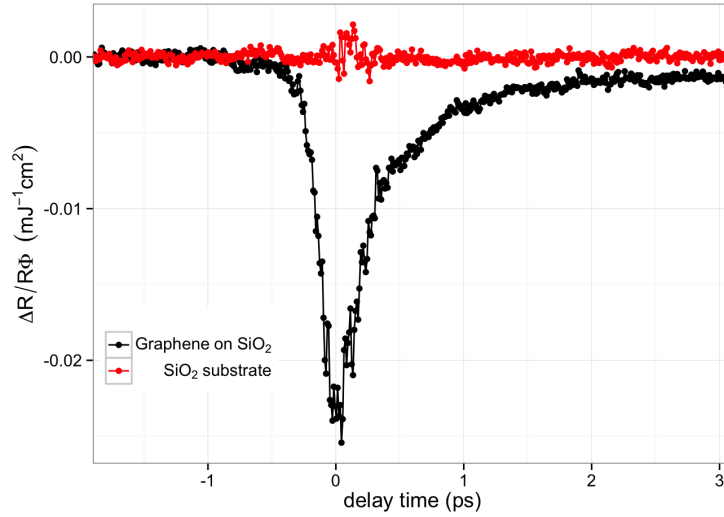


FIG. S1: Differential reflection normalized to fluence as a function of temporal overlap for: Graphene on quartz (black) and the bare quartz substrate (red) at $\theta_{pump} = 50^\circ$, $\theta_{probe} = 70^\circ$.

INCREASING PUMP FLUENCE

For pulsed fluence $\sim 0.1 \text{ mJ/cm}^2$ we expect to generate an electron temperature of $\sim 1000 \text{ K}$ [S1]. This means that we are probing a very non-equilibrium electron distribution. In order to investigate the effect of this electron heating, we increased the pump fluence used in the experiment to $\sim 1.1 \text{ mJ/cm}^2$. These results are shown in fig. S2. We find that a higher fluence significantly suppresses the surface plasmon resonance features with respect to the background, off-resonance signal. We believe two factors contribute to this effect: firstly, an increased electron temperature will increase saturable absorption, the effect primarily responsible for the incoherent background signal. Secondly, due to the negative photoconductivity usually exhibited by graphene for pulsed femtosecond excitation[S2], one can expect increased losses and quenching of the surface plasmon, leading to broadening of the spectral features associated with their excitation.

Evidence that the increased electron temperature also raises the effective Fermi energy of the sample can be inferred by comparing fig. S2(a) and fig. S2(b), where the resonant regions of differential reflectivity shift to higher frequencies in the high-fluence case, as would be expected for the graphene surface plasmon dispersion for a higher doping level.

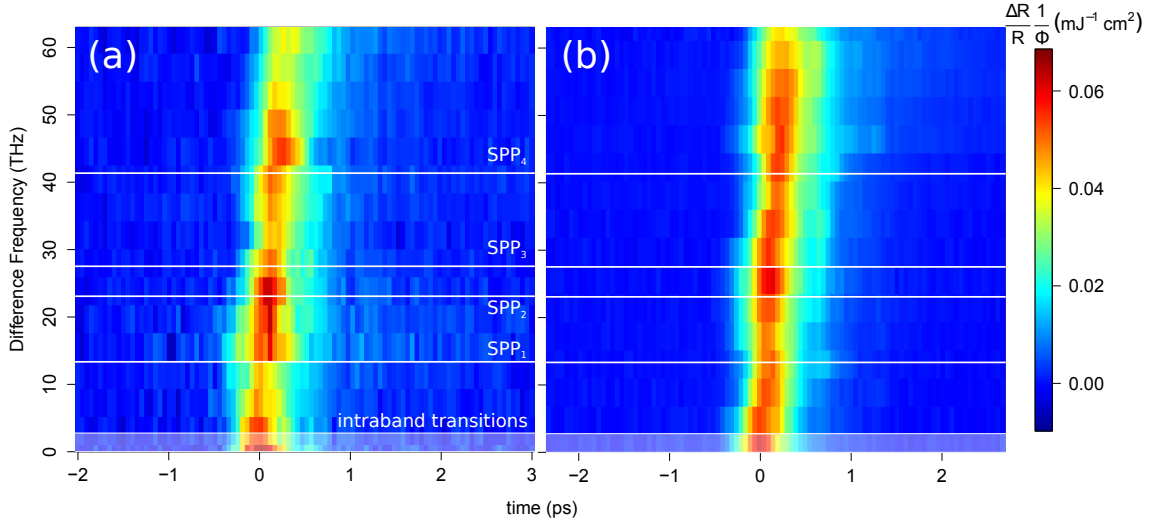


FIG. S2: Differential reflection normalized to fluence as a function of temporal overlap for: low (left) and high (right) pump fluence at $\theta_{pump} = 15^\circ$ and $\theta_{probe} = 125^\circ$.

REDUCING PUMP FLUENCE

Since electron heating clearly has a negative impact in the experiment, one wants to minimize the pump intensity. However, reducing the light intensity also reduces our difference frequency coupling efficiency, since surface plasmon generation here is a nonlinear process. However, a possible route to better isolating coherent signals is to reduce the pump beam intensity and increase the probe beam intensity, illuminating the sample with similar fluences for both beams. This reduces the Pauli blocking of the probe induced by the pump beam, decreasing the signal due to saturable absorption decreases, while maintaining the efficiency of the difference frequency mixing process.

Figures S3(a) and S3(b) show differential reflection normalized to pump fluence for two difference frequencies, 0 THz ($\lambda_{pump} = 615$ nm) and 12 THz ($\lambda_{pump} = 600$ nm), measured for the angles $\theta_{pump} = 50^\circ$ and $\theta_{probe} = 70^\circ$. In this geometry we expect a resonant enhancement for a difference frequency ~ 12 THz due to plasmon excitation and no enhancement for 0 THz, as previously measured in fig. 3(b). In figure S3 (a) we show a typical measurement

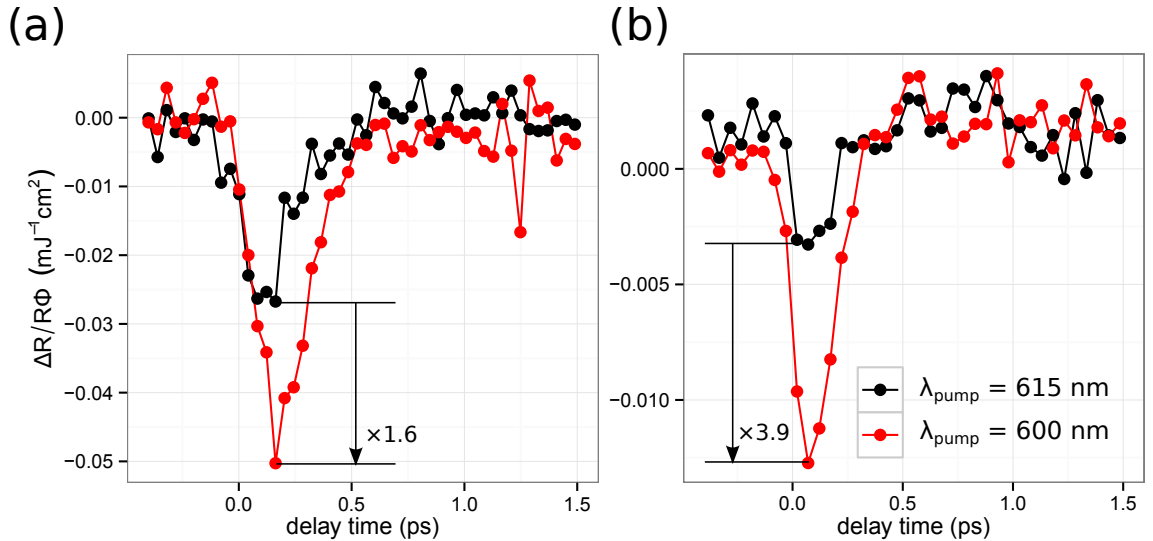


FIG. S3: Differential reflection normalized to pump fluence as a function of temporal overlap of the pulses for: (a) high-fluence pump and low-fluence probe, and (b) comparable fluence in both pump and probe. The black lines indicate the case for the difference frequency of 0 THz (no resonant plasmon coupling) and the red lines are for a difference frequency of 12 THz (resonant plasmon coupling). $\theta_{pump} = 55^\circ$, $\theta_{probe} = 70^\circ$, $\lambda_{probe} = 615$ nm.

for a high-power pump beam (0.26 mJ/cm^2) and a low power-probe beam (0.0028 mJ/cm^2). In this case, when the difference frequency matches that of the surface plasmon, a resonant change to the reflectivity by a factor of 1.6 occurs. For equal pump and probe fluences ($\sim 0.07 \text{ mJ/cm}^2$), as shown in fig. S3(b), we observe a significant suppression of the background, non-resonant signal. The enhancement then measured from a non-resonant condition to the resonant surface plasmon excitation is increased to a factor of 3.9.

We also note that in the low-power pump case (fig. S3(b)), the lineshape of the recorded temporal dynamics is far more symmetric than in the high-power case (fig. S3(a)), which clearly exhibits the typical asymmetric lineshape indicative of incoherent carrier cooling dynamics.

HIGHER DIFFERENCE FREQUENCIES

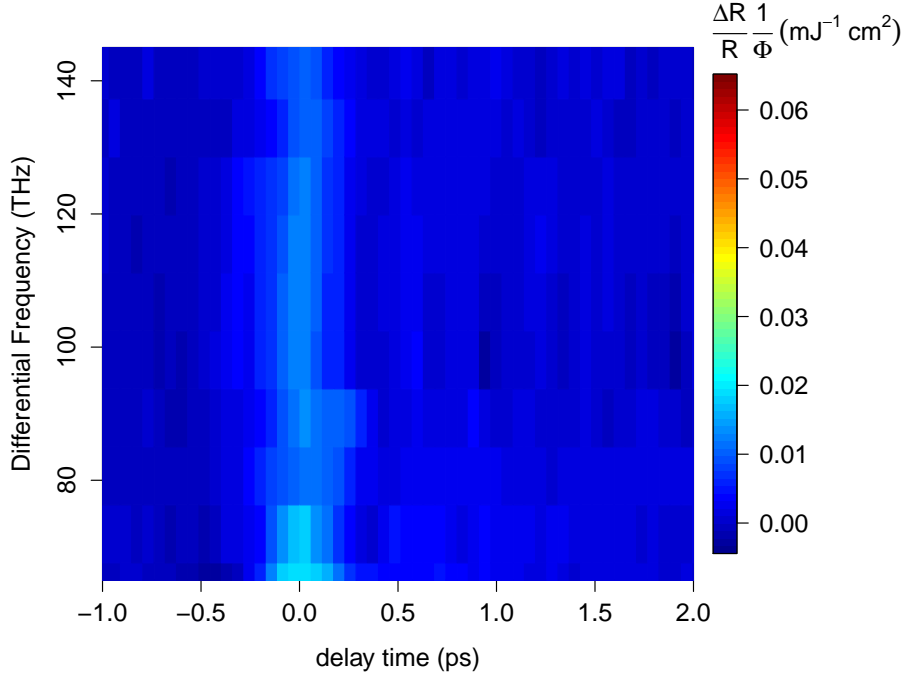


FIG. S4: Plot of differential reflection normalized to pump fluence as a function of temporal overlap for $\theta_{pump} = 15^\circ$ and $\theta_{probe} = 125^\circ$ at higher frequencies. The colorscale has been scaled to fig. 3(c) for ease of comparison.

We have preformed several measurements for a wider range of difference frequencies. In

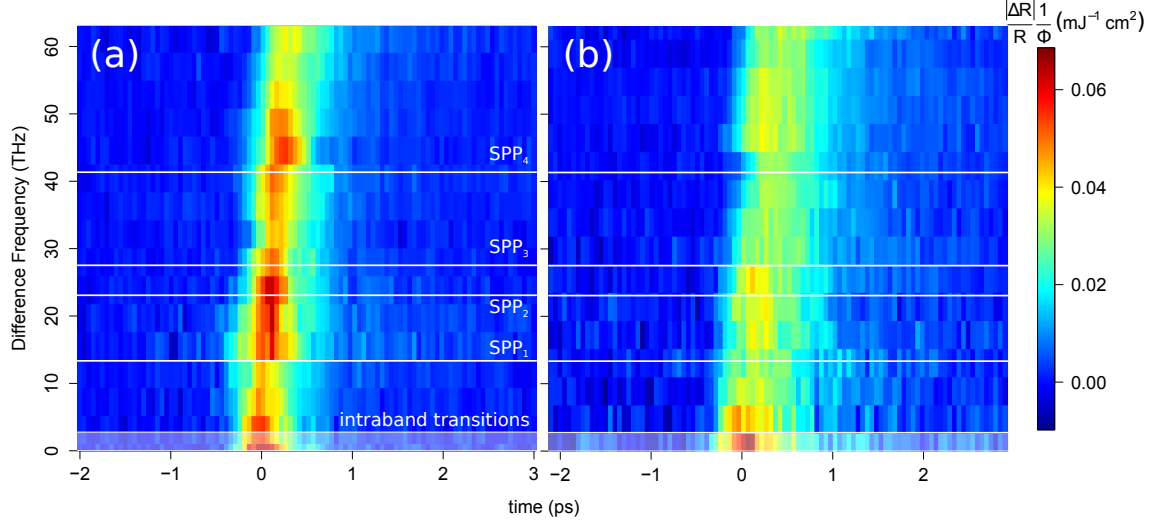


FIG. S5: Differential reflection normalized to pump fluence as a function of temporal overlap for transverse magnetic (a) and transverse electric (b) polarizations. The resonances for the predicted surface plasmon frequencies are clearly suppressed (labelled), while the intraband resonance is largely unaffected.

figure S4, we present measurements taken for $\theta_{pump} = 15^\circ$ and $\theta_{probe} = 125^\circ$, where the pump wavelength was varied from 540 nm to 475 nm, with the probe wavelength fixed at 615 nm. This gives a difference frequency range from 70 THz to 140 THz. For these larger difference frequencies, we observe no resonance features above 70 THz, indicating there is no coherent coupling to higher frequency modes.

POLARIZATION DEPENDENCE

The experiment shown in fig. 3(c) was repeated with the pump and probe both polarized with the electric vector parallel to the graphene surface (transverse electric, TE polarized). Under these circumstances we expect surface plasmon excitation to be suppressed. The results are shown in figure S5. When illuminating with TE polarized light, there is a clear decrease in the nonlinear enhancement to the reflectivity for the 23-27 THz peak and the ~ 45 THz peak. The peak at ~ 0 THz is of the same order in both polarization cases, suggesting that the nonlinearity arising due to the intraband transitions are less sensitive to the polarization of the light. This is attributed to how carriers respond to the EM field

in the plane of the surface. Since, for the intraband response the conductivity functions are in-plane, they make no distinction as to the out of plane component (which distinguishes TE and TM polarization).

While the higher-frequency resonances are strongly suppressed, they are not absent completely: we attribute this to imperfect polarization of the beams and inhomogeneity in the graphene sample.

THEORETICAL MODEL

We describe a theoretical model of two continuous-wave, free-space beams of frequencies $\omega_{1,2}$ (without loss of generality, assume that $\omega_1 > \omega_2$) interacting with graphene via a difference frequency generation process. The convention in our calculations to define the field polarizations and beam angles is illustrated in Fig. S7. The beams are taken to be incident from air (refractive index $n \approx 1$). Important to modeling this experiment is the inclusion of a frequency-dependent and complex refractive index of the substrate at low frequencies, in order to capture the lattice vibrations in silica and the resulting surface optical phonons. To do this, we take a simple dielectric response model based upon three transverse optical (TO) phonon modes [S3],

$$n^2(\omega) = \epsilon_\infty + \sum_{j=1}^3 \frac{f_j \omega_{TO,j}^2}{\omega_{TO,j}^2 - \omega^2 - i\omega\gamma_{TO,j}}. \quad (\text{S1})$$

From Ref. [S3], the high-frequency dielectric constant is taken to be $\epsilon_\infty = 2.4$, while the TO phonon frequencies and oscillator weights are $\omega_{TO} = 2\pi \times (13.44, 23.75, 33.84)$ THz and $f = (0.7514, 0.1503, 0.6011)$, respectively. The damping rates are taken to be $\gamma_{TO} = 2\pi \times (0.80, 1.27, 1.27)$ THz. The resulting real and imaginary parts of the refractive index, plotted in Fig. S6, approximately correspond to experimentally measured values [S4]. In practice, this refractive index function is only relevant for the substrate response at the low difference frequency of $\omega_3 = \omega_1 - \omega_2$, while for the high frequencies $\omega_{1,2}$ the response is nearly frequency-independent, $n \approx \sqrt{\epsilon_\infty}$.

In general, one can obtain equations relating the reflection and transmission coefficients to each other by enforcing electromagnetic boundary conditions (continuity of the normal electric displacement and tangential electric field) at the graphene interface. The solution

for the transmission coefficient of field i ($i = 1, 2$) is readily found to be

$$t_i = \frac{2 \sin \theta_i}{n_i \sin \theta_i + \sin \phi_i} - \frac{(\rho_{is}/\epsilon_0) \sec \phi_i \sin \theta_i}{E_{Ii}(n_i^2 \sin \theta_i + n_i \sin \phi_i)}, \quad (\text{S2})$$

where $n_i = n(\omega_i)$ denotes the substrate refractive index at the field frequency, while the reflection coefficient is related by $r_i = 1 - t_i \sin \phi_i \csc \theta_i$. Here E_{Ii} are the incident field amplitudes, and θ_i and ϕ_i are the angles of the fields on the vacuum and substrate sides, respectively. $\rho_{is} = \rho(\omega_i, k_{ix})$ is the graphene surface charge density at frequency ω_i and in-plane wavevector $k_{ix} = (\omega_i/c) \cos \theta_i$. Note that the first term on the right-hand side of Eq. (S2) reproduces the standard Fresnel coefficient in the absence of a graphene layer ($\rho_{is} = 0$). The angle of the transmitted field is related to the incident by Snell's Law, $\cos \theta_i = n_i \cos \phi_i$.

The surface charge density can be related to the current density J in the graphene layer via the continuity equation, which in the Fourier domain reads

$$\rho_s(\omega, k_x) = (k_x/\omega)J_x(\omega, k_x). \quad (\text{S3})$$

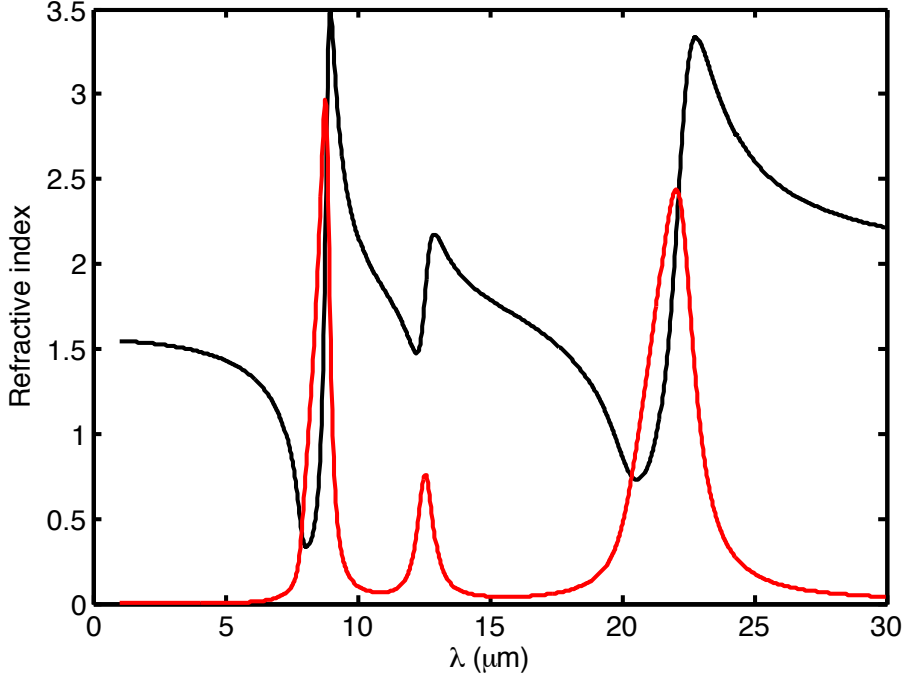


FIG. S6: Real (black) and imaginary (red) parts of the refractive index of silica versus free-space wavelength ($\lambda = 2\pi c/\omega$), based upon the model of Eq. (S1).

At the same time, the current density can be related to the electric fields via conductivity functions. We are particularly interested in the case of difference frequency generation, where the field produced at the difference frequency and wavevector $\omega_3 = \omega_1 - \omega_2$ and $k_{3x} = k_{1x} - k_{2x}$ is resonantly enhanced by aligning them with the plasmon dispersion of graphene $\omega_p(k_x)$. This motivates a truncated model in which we include only the linear and second-order conductivities, and frequencies $\omega_{1,2}$ and ω_3 (thus left out is sum frequency generation and the generation of even higher harmonics). Then, the current density for field 1 is given by

$$J_x(\omega_1, k_{1x}) = \sigma^{(1)}(\omega_1, k_{1x})E_x(\omega_1, k_{1x}) + \sigma^{(2)}(\omega_1, k_{1x}; \omega_2, k_{2x}, \omega_3, k_{3x})E_x(\omega_2, k_{2x})E_x(\omega_3, k_{3x}). \quad (\text{S4})$$

Here, $E_x(\omega_1, k_{1x}) = t_1 E_{1I} \sin \phi_1$ is the total parallel field for $i = 1$ at the graphene layer, which we have written in terms of the incident field and transmission coefficient. $\sigma^{(1)}$ is the linear conductivity function, while $\sigma^{(2)}(\omega_1, k_{1x}; \omega_2, k_{2x}, \omega_3, k_{3x})$ is the second-order nonlinear conductivity functions relating the current density generated at ω_1, k_{1x} given fields at ω_2, k_{2x}

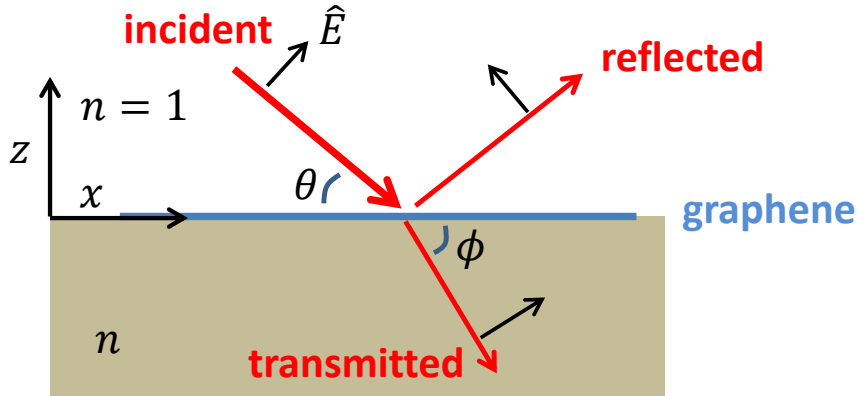


FIG. S7: Illustration of p -polarized electromagnetic fields propagating in the x - z plane and interacting with graphene. The fields $i = 1, 2$ consist of incident, reflected, and transmitted components, with the directions of propagation and polarizations indicated by the red and black arrows, respectively. The angles of incidence and transmission are θ and ϕ . The incident field is assumed to propagate in vacuum (refractive index $n = 1$), while the graphene sits on top of a substrate with index n (possibly frequency dependent).

and ω_3, k_{3x} . Similar expressions as Eq. (S4) can be written down for the current density at ω_i, k_{ix} ($i = 2, 3$), and we use an analogous set of conventions to indicate the fields and conductivities at other frequencies and wavevectors. In what follows, we will also adopt the more compact notation $\sigma^{(2)}(\omega_1) = \sigma^{(2)}(\omega_1, k_{1x}; \omega_2, k_{2x}, \omega_3, k_{3x})$, where the dependence on wavevectors and input frequencies is understood.

The substitution of Eqs. (S3) and (S4) into Eq. (S2) (along with analogous equations for the other fields $i = 2, 3$) yields a set of nonlinear equations relating the transmission coefficients and incident fields,

$$t_1 = t_1^{(L)} \left[1 - \frac{|t_2 E_{2I}|^2}{(2c\epsilon_0)^2} t_1^{(L)} t_3^{(L)} \sigma^{(2)}(\omega_1) \sigma^{(2)}(\omega_3) \sin \phi_1 \sin^2 \phi_2 \sin \phi_3 \right]^{-1}, \quad (\text{S5})$$

$$t_2 = t_2^{(L)} \left[1 - \frac{|t_1 E_{1I}|^2}{(2c\epsilon_0)^2} t_2^{(L)} t_3^{(L)*} \sigma^{(2)}(\omega_2) \sigma^{(2)*}(\omega_3) \sin^2 \phi_1 \sin \phi_2 \sin^* \phi_3 \right]^{-1}. \quad (\text{S6})$$

Here $t^{(L)} = \frac{2 \sin \theta}{n \sin \theta + \sin \phi + (\sigma^{(1)}/c\epsilon_0) \sin \theta \sin \phi}$ is the linear transmission coefficient, and the angle of the generated field is defined via $k_{3x} \equiv \frac{n_3 \omega_3}{c} \cos \phi_3$. The complex in-plane field amplitude generated at the difference frequency and at the position of the graphene layer $z = 0$ is given by

$$E_{3x} = -\frac{t_1 t_2^* t_3^{(L)}}{2c\epsilon_0} E_{1I} E_{2I}^* \sigma^{(2)}(\omega_3) \sin \phi_1 \sin \phi_2 \sin \phi_3. \quad (\text{S7})$$

While Eqs. (S5) and (S6) may appear somewhat complicated, here we note their main features. First, we note that the input field amplitudes, the beam angles, and the linear optical properties of the system are generally known. Thus, on one hand, given a theoretical model of the nonlinear conductivity $\sigma^{(2)}$, these equations can be solved to obtain the predicted changes in transmission and reflection of the input beams, due to the generation of plasmons at ω_3, k_{3x} . The plasmon field amplitude itself can be found from Eq. (S7). On the other hand, even absent a theoretical model, if changes in transmission or reflection of the incident fields are experimentally measured, one can attempt to invert these equations in order to obtain an experimentally inferred value of $\sigma^{(2)}$.

The description above generally holds regardless of the values of ω_3, k_{3x} . It is particularly interesting, however, to focus on the case where they align with the plasmon dispersion relation. In the small wavevector limit and for frequencies smaller than twice the Fermi frequency, $\omega \lesssim 2\omega_F$, the linear conductivity is well-approximated by the Drude model [S5],

$$\sigma(\omega) \approx \frac{ie^2}{\pi \hbar} \frac{\omega_F}{\omega + i\gamma}. \quad (\text{S8})$$

Here we have included a phenomenological damping term γ . The plasmon dispersion relation can be found by solving for the pole of the linear transmission coefficient, $t^{(L)}$. For simplicity, we will momentarily consider the case of a substrate with frequency-independent refractive index, so that the role of plasmon damping can be more clearly identified. In the absence of losses, the dispersion relation is found to be

$$k_{x,p} = \frac{(1+n^2)\omega_p^2}{4\alpha c\omega_F}, \quad (\text{S9})$$

where $\alpha \approx 1/137$ is the fine-structure constant. In the presence of losses, choosing ω_3, k_{3x} to lie on the plasmon dispersion relation yields a linear transmission amplitude of $|t_3^{(L)}| \approx \frac{2nQ}{1+n^2}$, where $Q = \omega_3/\gamma$ is the plasmon quality factor. Under these conditions, one thus sees from Eq. (S7) that the field intensity experiences a resonant enhancement of $|E_3|^2 \propto Q^2$. A similar resonant effect appears in the transmission and reflection coefficients of the incident fields.

While the Eqs. (S5) and (S6) can in principle be inverted to infer $\sigma^{(2)}$ given experimental data for reflection or transmission coefficients, in the present experimental setup this procedure can only be done semi-quantitatively due to a number of unknowns. First, the signal lies significantly above the noise floor only near the plasmon dispersion relation, and only a limited number of beam angles are investigated. This makes it difficult to infer a specific wavevector and frequency dependence of the nonlinear conductivity (fundamentally, there must be a dependence on wavevector, as otherwise $\sigma^{(2)} = 0$ for a centrosymmetric material). Furthermore, the experiment employs pulses whose bandwidths are significantly larger than the plasmon linewidth. Given that, here we aim to reach a conservative estimate for the strength of $\sigma^{(2)}$, while we anticipate that future improved experiments (such as with longer pulses and nano-structures) and theoretical models will enable more detailed comparisons.

The full conductivity function at zero temperature is given by [S5]

$$\sigma(\omega) = \frac{ie^2}{\pi\hbar} \frac{\omega_F}{\omega + i\gamma} + \frac{e^2}{4\hbar} \left[\Theta(\omega - 2\omega_F) + \frac{i}{\pi} \log \left| \frac{\omega - 2\omega_F}{\omega + 2\omega_F} \right| \right], \quad (\text{S10})$$

where $\Theta(x)$ is the Heaviside step function. The Fermi energy $\hbar\omega_F \approx 0.5$ eV of graphene is significantly lower than the pump and probe photon energies of ~ 3 eV. At these frequencies, the linear conductivity of graphene is nearly frequency independent and real, $\sigma^{(1)}(\omega)/(c\epsilon_0) \approx \pi\alpha$, which we use to obtain the pump and probe linear reflection coefficients. Furthermore, we take the simplest possible function for the nonlinear conductivity, $\sigma^{(2)}(\omega) = i|\sigma^{(2)}(\omega_2)|(\omega/\omega_2)$, where $|\sigma^{(2)}(\omega_2)|$ is a single fitting parameter (the probe frequency ω_2 is fixed in the experiment). With this choice of function, graphene would be

equivalent to a nonlinear material with a frequency-independent bulk nonlinear susceptibility of $\chi^{(2)} = -i\sigma^{(2)}(\omega)/(\omega\epsilon_0 t) = |\sigma^{(2)}(\omega_2)|/(\omega_2\epsilon_0 t)$, where t is the effective thickness of graphene. Inserting this nonlinear conductivity into Eqs. (S5), (S6), and (S7), we find that a value of $|\sigma^{(2)}(\omega_2)| \approx 2.4 \times 10^{-12} \text{ A}\cdot\text{m}/\text{V}^2$ produces a good qualitative fit to the experimental data.

We now discuss how to obtain the conversion efficiency of pump photons to plasmons. The number of photons dissipated per unit area and time by the field at the difference frequency consists of two terms, $\Gamma_d = \Gamma_{d,g} + \Gamma_{d,s}$. The first term consists of damping from the graphene layer due to the real part of its conductivity, and is given by $\Gamma_{d,g} = (\text{Re } \sigma^{(1)}(\omega_3)) |E_{3x}|^2 / (2\hbar\omega_3)$. The second term is due to damping from the substrate, due to the fact that at low frequencies its refractive index is complex. This contribution is given by $\Gamma_{d,s} = \frac{\epsilon_0}{4\hbar|\text{Im}(k_{3x} \tan \phi_3)|} (\text{Im } n^2(\omega_3)) |E_{3x}|^2 (1 + |\cot \phi_3|^2)$. On the other hand, the incident photon flux in the pump field is $\Gamma_{\text{in}} = I_1 \sin \theta_1 / (\hbar\omega_1)$, where I_1 is the pump intensity.

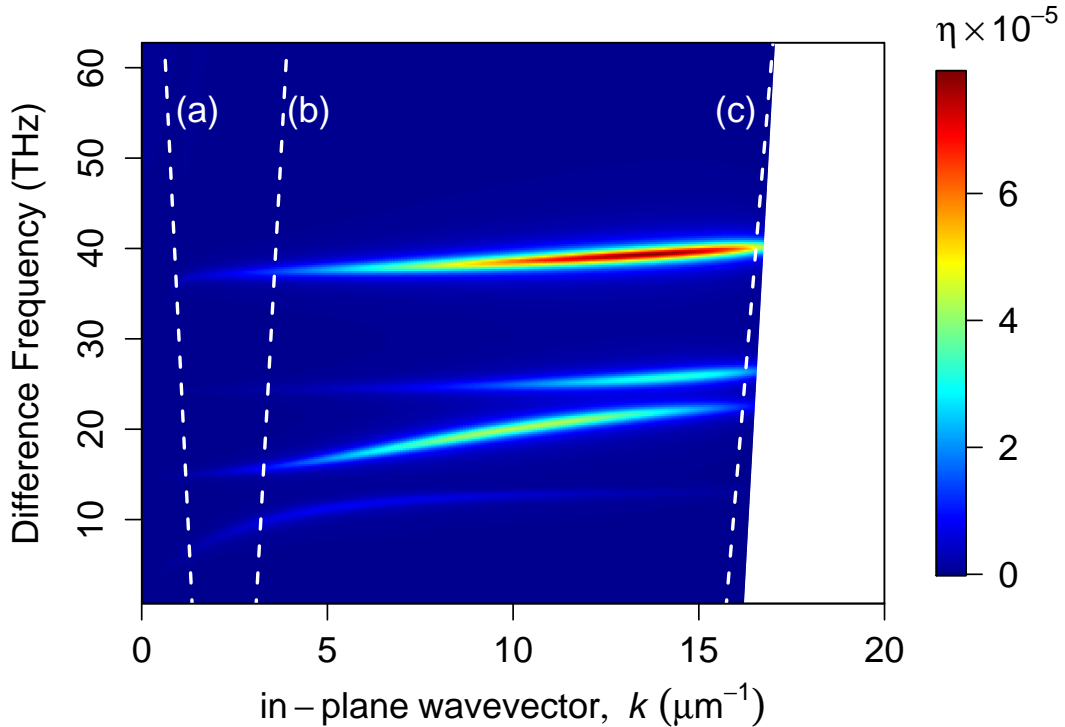


FIG. S8: The numerical solution for the pump conversion efficiency, η , as a function of wavevector and difference frequency.

Generally, the amplitude of the generated plasmon field rate will depend on both the pump and probe intensities. However, at the level of individual photons, the process is that an incoming pump photon gets converted to a plasmon (assisted by stimulated emission of a photon into the probe). Thus, we define the conversion efficiency relative to the pump alone. In steady state, the rates of photons dissipated and generated at the difference frequency are equal, and thus the overall conversion efficiency of pump photons to plasmons is $\eta = \Gamma_d/\Gamma_{\text{in}}$. This efficiency is shown in fig. S8 for the range of frequencies and wavevectors used in our experiments. Using the estimated value of $|\sigma^{(2)}(\omega_2)|$, we find that the conversion efficiency for the experimental arrangement shown in fig. 3(b) at the point of maximum signal is approximately $\eta \approx 6 \times 10^{-6}$.

* Electronic address: t.j.constant@ex.ac.uk

† Electronic address: darrick.chang@icfo.es

- [S1] Hale, P. J., Hornett, S. M., Moger, J., Horsell, D. W. & Hendry, E. Hot phonon decay in supported and suspended exfoliated graphene. *Phys. Rev. B* **83**, 121404 (2011). 1012.3927.
- [S2] Tielrooij, K. J. *et al.* Photoexcitation cascade and multiple hot carrier generation in graphene. *Nature Physics* **9**, 248–252 (2013). URL <http://dx.doi.org/10.1038/nphys2564>. arXiv:1210.1205v2.
- [S3] Luxmoore, I. J. *et al.* Strong Coupling in the Far-Infrared between Graphene Plasmons and the Surface Optical Phonons of Silicon Dioxide. *ACS Photonics* **1**, 1151–1155 (2014). URL <http://pubs.acs.org/doi/abs/10.1021/ph500233s>.
- [S4] Kitamura, R., Pilon, L. & Jonasz, M. Optical constants of silica glass from extreme ultraviolet to far infrared at near room temperature. *Applied optics* **46**, 8118–8133 (2007).
- [S5] Wunsch, B., Stauber, T., Sols, F. & Guinea, F. Dynamical polarization of graphene at finite doping. *New Journal of Physics* **8**, 318 (2006). 0610630.