Ultrafast optical switching of infrared plasmon polaritons in high-mobility graphene

G. X. Ni[1,2†], L. Wang[3†], M. D. Goldflam[2], M. Wagner[2], Z. Fei[2], A. S. McLeod[2], M. K. Liu[2,4], F. Keilmann[5], B. Özyilmaz[1,6], A. H. Castro Neto[1], J. Hone[3], M. M. Fogler[2] and D. N. Basov[2,7★]

The success of metal-based plasmonics for manipulating light at the nanoscale has been empowered by imaginative designs and advanced nano-fabrication. However, the fundamental optical and electronic properties of elemental metals, the prevailing plasmonic media, are difficult to alter using external stimuli. This limitation is particularly restrictive in applications that require modification of the plasmonic response at sub-picosecond timescales. This handicap has prompted the search for alternative plasmonic media[1-3], with graphene emerging as one of the most capable candidates for infrared wavelengths. Here we visualize and elucidate the properties of non-equilibrium photo-induced plasmons in a high-mobility graphene monolayer[4]. We activate plasmons with femtosecond optical pulses in a specimen of graphene that otherwise lacks infrared plasmonic response at equilibrium. In combination with static nano-imaging results on plasmon propagation, our infrared pump–probe nano-spectroscopy investigation reveals new aspects of carrier relaxation in heterostructures based on high-purity graphene.

Graphene plasmonics[5-7] has progressed rapidly, propelled by the electrical tunability, high field confinement[8,9], potentially long lifetimes[10,11] of plasmons and the strong light–matter interactions[12-15] in graphene. An earlier spectroscopic study has reported photo-induced alteration of the plasmonic response of graphene on optical pumping[16]. In this work, we harnessed ultrafast optical pulses to generate mid-infrared (mid-IR) plasmons in a sample that lacks a plasmonic response at equilibrium. We examined the real-space aspects of non-equilibrium plasmon–polariton generation and propagation under femtosecond (fs) photo-excitation using a new ultrafast nano-infrared (IR) technique that fuses real-space plasmon imaging with spectroscopy. We applied this method to investigate high-quality graphene specimens encapsulated in hexagonal boron nitride: hBN/G/hBN[4].

We performed time-resolved broadband nano-IR experiments using antenna-based near-field nanoscopy (see Methods). This set-up (Fig. 1a,b) combines exceptional spatial, spectral and temporal resolution[16-18], allowing an experimental probe of the dispersion of graphene plasmons under photo-excitation—a feat previously considered technologically infeasible. In our measurements, the metalized tip of an atomic force microscope (AFM) was illuminated by a focused IR probe beam, generating strong evanescent electric fields beneath the tip. These fields possess a wide range of in-plane momenta q and therefore facilitate efficient coupling to graphene plasmons[19]. Such evanescent fields extend ~20 nm beneath the top surface of our structures, which is sufficient to

launch and detect surface plasmons in a graphene microcrystal protected by a thin (10 nm) encapsulating layer of hBN[10]. The tip of the nanoscope acts as a launcher for surface plasmons of wavelength (λ_p) that propagate radially outwards from the tip. On reflection from the sample edge, these plasmons form standing waves between the tip and the reflector. While scanning the tip towards a graphene edge, one can collect the near-field signal, which displays oscillations with a period of $\lambda_p/2$. Recent nano-IR imaging experiments with encapsulated high-mobility graphene have registered oscillations with both $\lambda_p/2$ and λ_p periodicity; the latter were assigned to plasmons launched directly from the sample edge by incident light[10].

We begin with the principal result of this work: the ultrafast dynamics of plasmons in encapsulated graphene revealed by means of nano-IR pump–probe spectroscopy (Fig. 1c,d). The broadband mid-IR probe allows for visualization of the frequency–momentum dispersion of graphene plasmons in the course of a single line-scan across the sample surface[20]. We investigated the photo-induced changes in near-field scattering amplitude $s(\omega,x)$ collected from sequential 20×20 nm^2 spatial pixels that together constitute a hyperspectral line-scan. Acquired at varying pump–probe delay times, hyperspectral scans reveal a rich spatiotemporal plasmonic response, which arises and then decays according to the dynamics of photo-excited carriers. It is instructive to present hyperspectral line-scans in the form of two-dimensional frequency-position maps $s(\omega,x)$ plotted in Fig. 1c,d (and see Supplementary Information (SI)). This representation highlights the novelty of our experimental approach, which combines non-equilibrium spectroscopy with imaging of plasmonic standing waves.

The strongest photo-induced signal is recorded at zero time delay between pump and probe pulses (Fig. 1c): we observe a set of three dispersing peaks in the $s(\omega,x)$ hyperspectral map. The dashed lines shown in Fig. 1c trace peaks of the signal determined from the spatial derivative of the raw data, $ds(\omega,x)/dx = 0$ (Fig. 1e). The spatial period of oscillations, as well as the relative separation between peaks and troughs, systematically increases with decreasing ω. These findings are consistent with expectations for plasmonic modes dispersing with positive group velocity.

The photo-induced plasmonic signal decays as the probe pulse is temporally delayed with respect to the pump pulse (Fig. 1d,f). After $t = 2$ ps pump–probe delay, we observe only a single peak in close proximity to the sample edge. Plasmonic features completely vanish after 5 ps (Supplementary Section 6): a timescale consistent with earlier diffraction-limited measurements[21-23]. Figure 1c,d

[1]Centre for Advanced 2D Materials and Graphene Research Centre, National University of Singapore, Singapore 117546, Singapore. [2]Department of Physics, University of California, San Diego, La Jolla, California 92093, USA. [3]Department of Mechanical Engineering, Columbia University, New York, New York 10027, USA. [4]Department of Physics, Stony Brook University, Stony Brook, New York 11794, USA. [5]Ludwig-Maximilians-Universität and Center for Nanoscience, 80539 München, Germany. [6]Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542, Singapore. [7]Department of Physics, Columbia University, New York, New York 10027, USA. [†]These authors contributed equally to this work. *e-mail: dbasov@physics.ucsd.edu
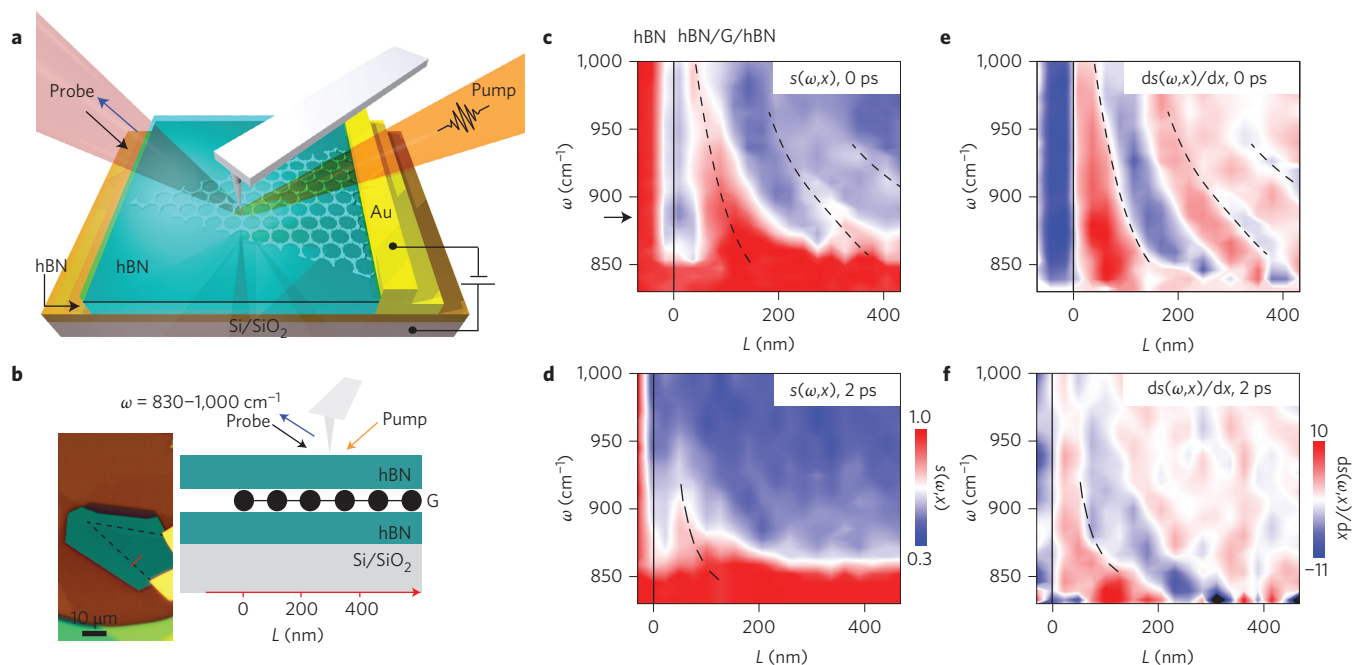
Figure 1 | Experimental configuration and ultrafast pump–probe plasmonic control. a,b, The pump–probe nano-IR set-up and optical image of the hBN/hBN device. **b**, The black dotted lines mark the graphene layer that is covered by a thin hBN layer. The red solid line denotes the nano-IR line-scan position. The probe beam spans frequencies from 830–1,000 cm^{-1}. Under pumping conditions, plasmonics polariton waves are launched by the AFM tip and reflected at the graphene edge at $L = 0$. **c,d**, Pump–probe $s(\omega,x)$ data revealing the dispersion of photo-induced plasmons for different pump–probe time delays. These data are displayed in the form of two-dimensional hyperspectral maps of the scattering amplitude $s(\omega,x)$. The black dashed lines mark the peaks of the dispersing traces obtained from $ds(\omega,x)/dx = 0$, as shown in **e,f**, and detailed in Supplementary Section 6. **e,f**, Derivatives of the pump–probe $s(\omega,x)$ data in **c,d**, respectively. The vertical solid line marks the edge of the graphene.

depicts the first direct observation of plasmon interference patterns produced by photo-excited Dirac quasiparticles.

We will now examine the physical mechanism underlying the ultrafast plasmonic effects in Fig. 1. By way of a photo-induced increase in the Drude weight (D), this transient plasmonic response originates from a strongly enhanced electronic temperature (T_{el}) in photo-excited graphene[16,23]. Specifically, at timescales longer than ~30 fs, it is expected that the optical excitation of graphene preserves the electron density, but results in an elevated T_{el} (ref. 24, for details see Supplementary Section 5). At $t = 0$, T_{el}, $D(T_{el})$, and $\lambda_p(\omega)$ are the largest. The reduction in $\lambda_p(\omega)$ shown in Fig. 1d at 2 ps following the pump pulse can be explained by a decrease of $D(T_{el})$ as electrons begin to cool. Thus, the plasmonic response observed experimentally within the first several ps following photo-excitation is consistent with the expected dynamical evolution in $D(T_{el})$.

We now proceed to survey nano-imaging data acquired for the same device using a continuous-wave (cw) laser at $\omega = 890$ cm^{-1} (Fig. 2b–d). With zero applied back-gate voltage $V_g = 0$, the near-field signal from the entire structure is nearly uniform with no optical contrast from graphene (Fig. 2b). This is because our encapsulated graphene remains nearly charge neutral even at ambient conditions ($V_g = V_{CN} = 0$ V, where V_{CN} is the charge neutrality voltage)[4]. As V_g increases, plasmon fringes start to develop, with the most prominent plasmon fringe observed closest to the sample edges. Figure 2c shows the near-field image obtained at $V_g = 60$ V. Compared with the $V_g = 0$ case, the triangular graphene shape can be clearly discerned. Moreover, multiple plasmonic fringes along the graphene edges can be detected. Increasing V_g to 115 V improves the visibility of the graphene plasmon fringes, yielding at least 12 fringes within a micrometre of the sample edge (Fig. 2d). The round-trip distance travelled by plasmons in Fig. 2 is therefore approaching 2 µm,

similar to ref. 10, and attests to a reduction of plasmonic losses in high-mobility samples.

Alternating bright and dark fringes in Fig. 2 encode information about λ_p and the dimensionless damping constant γ_p[8–11]. We recall that $\gamma_p = (\sigma_1/\sigma_2) + (\kappa_2/\kappa_1)$ is related to the ratios of the real and imaginary parts of the conductivity $\sigma(\omega) = \sigma_1(\omega) + i\sigma_2(\omega)$, and the effective permittivity $\kappa(\omega) = \kappa_1(\omega) + i\kappa_2(\omega)$ of the graphene environment[25,26] (see Supplementary Section 4). For a quantitative analysis, we plot imaging data in the form of line-profiles (Fig. 2e, f) acquired along the lines marked in Fig. 2c,d. Multiple plasmon fringes are clearly resolved in the graphene interior. These fringes appear in pairs, indicating the coexistence of two types of oscillations: one with period $\lambda_p/2$ and the other with period λ_p (Supplementary Section 3). The plasmonic line-profiles can be simulated numerically by approximating the tip as an elongated spheroid (Supplementary Section 4). The best fit, shown by the green curve in Fig. 2f, yields an estimate of $\gamma_p = 0.02$ and $\lambda_p = 150$ nm.

Our analysis of plasmonic images acquired under near-equilibrium conditions confirms the relevance of both λ_p- and $\lambda_p/2$-periodic plasmonic interference patterns to the infrared response of the high-mobility device. Nevertheless, these two distinct periodicities are not apparent in our dynamic data. This is not unexpected. Pump–probe traces show substantially broadened plasmonic fringes: an anticipated outcome of increased damping in the regime of elevated T_{el}. In Fig. 3a we display an attempt to model the influence of scattering associated with an increase in T_{el}. We start with parameters that most accurately describe our equilibrium nano-IR data (red trace), and then examine the effect of increasing γ_p. Enhanced scattering rapidly washes out weaker fringes with period $\lambda_p/2$ whereas stronger oscillations of period λ_p persist over a broader range of γ_p. Based on this analysis, we assume that separations between the dashed lines in Fig. 1c,e are equal to λ_p (Supplementary Section 3).
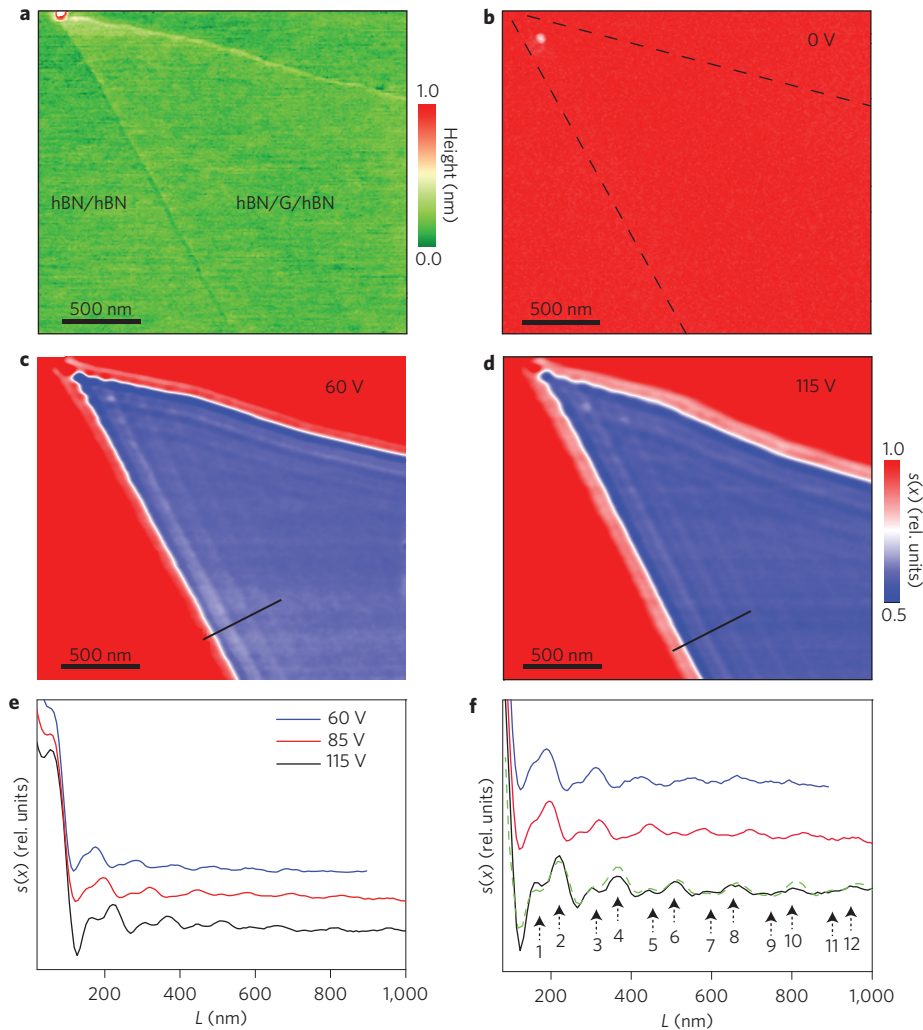
Figure 2 | AFM and static nano-IR imaging and line-profiles at different gate voltages. a, AFM topography image of the hBN/G/hBN encapsulated device. The triangular outline highlights the edges of graphene. **b–d,** Images of the normalized scattering amplitude $s(x)$ for the hBN/G/hBN structure acquired at $\omega = 890$ cm^{-1} at gate voltages of 0 (**b**), 60 (**c**) and 115 V (**d**), respectively (see Methods for details of data acquisition). In **b** the dashed lines mark the graphene position. **e,** Line-profiles across the graphene edge acquired for different gate voltages at the locations marked with black solid lines in **c** and **d**. **f,** Zoom-in of the plasmon fringes in **e**. Periodic double peaks are clearly visible. Round-trip plasmon travelling lengths of ~ 2 μm and at least 12 plasmonic oscillations are resolved in these traces. The green dashed line is the line profile obtained within a model described in Supplementary Section 3.
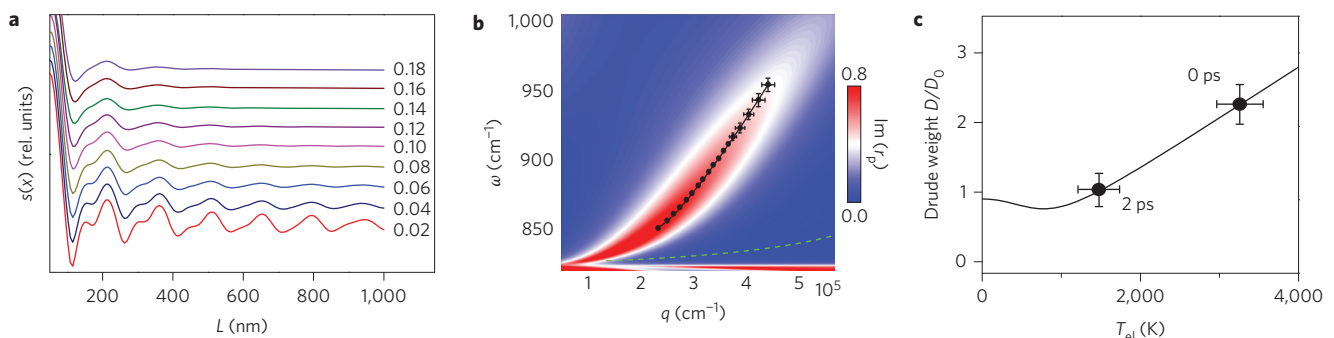

Figure 3 | Plasmonic line-profiles, dispersion and Drude weight calculations. a, Plasmon line-profiles calculated for $\lambda_p = 150$ nm at different γ_p. As γ_p increases, the fringes of period λ_p persist whereas those of period $\lambda_p/2$ diminish and become unobservable for $\gamma_p > 0.14$. **b,** The dispersion of photo-induced plasmons represented by the imaginary part of the complex reflectance Im $r_p(q,\omega)$ of our multilayered structure (Supplementary Section 5) calculated for an electron temperature of $\sim 3,200$ K at $\gamma_p = 0.15$. The black dots connected by the solid line are the experimental data obtained from Fig. 1c,e. The green dashed curve is the calculated dispersion at chemical potential $\mu = 1,330$ cm^{-1} corresponding to a gate voltage of $V_g = +30$ V. **c,** Drude weight as a function of electronic temperature calculated using equation (S3). The data point at $t = 2$ ps corresponding to $T_{el} \sim 1,700$ K was obtained from Im $r_p(q,\omega)$ calculations (Supplementary Sections 5 and 6). In both **b** and **c**, the error bars represent the 95% confidence intervals.

The dispersion of non-equilibrium photo-induced plasmons is presented in Fig. 3b. The dots were obtained from the fringe separations extracted from derivative plots in Fig. 1c,e. To model the plasmonic response we followed a standard procedure as detailed in the Supplementary Sections 2 and 5. The plasmon dispersion in Fig. 3b is visualized via a false colour map of the imaginary part of the reflectivity $r_p = r_p(q,\omega)$[16]. By varying T_{el}, we obtain a best match to the experimental data for $T_{el} = 3,200$ K. The corresponding D versus T_{el} relationship is plotted in Fig. 3c, which is similar to previous results[16]. With all parameters determined, we are at the position to plot the output of the model in the form of hyperspectral maps (Supplementary Section 5) that reproduce the gross features of the raw data in Fig. 1c. This overall agreement establishes the plasmonic origin for the real-space features observed in pump-probe data.

An important unresolved issue is the origin of the tenfold increase in the linewidth of the non-equilibrium plasmon dispersion compared with the equilibrium one. One possible candidate for the enhanced plasmon decay is the emission of acoustic phonons[10,27]. Another possibility is electron–hole scattering, for which recent relativistic hydrodynamics theories[28,29] predict that $\gamma_{ee} = (2\omega\tau)^{-1} \sim \alpha^2 T_{el}/\omega$. For $\alpha \sim 1$ and $T_{el} = 3,200$ K the estimate yields $\gamma_{ee} \sim 2.5$, which is more than an order of magnitude higher than the measured γ_p. Two possible reasons for this discrepancy are that the theory assumes undoped graphene, whereas in our case the doping is non-negligible, and the assumption of weak interactions ($\alpha \ll 1$) that may not describe the realistic case of $\alpha \sim 1$. On the other hand, the lower than expected intrinsic damping is good news for the implementation of practical ultrafast plasmonic devices using graphene.

Methods

Methods and any associated references are available in the online version of the paper.

References

1. Boltasseva, A. & Shalaev, V. M. All that glitters need not be gold. *Science* **8**, 1086–1101 (2014).
2. MacDonald, K. F., Samson, Z. L., Stockman, M. I. & Zheludev, N. I. Ultrafast active plasmonics. *Nature Photon.* **3**, 55–58 (2009).
3. Atwater, H. A. The promise of plasmonics. *Sci. Am.* **296**, 56–62 (2007).
4. Wang, L. *et al.* One-dimensional electrical contact to a two-dimensional material. *Science* **342**, 641–617 (2013).
5. Grigorenko, A. N., Polini, M. & Novoselov, K. S. Graphene plasmonics. *Nature Photon.* **6**, 749–758 (2012).
6. Basov, D. N., Fogler, M. M., Lanzara, A., Wang, F. & Zhang, Y. Colloquium: graphene spectroscopy. *Rev. Mod. Phys.* **86**, 959–993 (2014).
7. Javier García de Abajo, F. *et al.* Graphene plasmonics: challenges and opportunities. *ACS Photon.* **1**, 135–152 (2014).
8. Fei, Z. *et al.* Gate-tuning of graphene plasmons revealed by infrared nano-imaging. *Nature* **487**, 82–85 (2012).
9. Chen, J. *et al.* Optical nano-imaging of gate-tunable graphene plasmons. *Nature* **487**, 77–81 (2012).
10. Woessner, A. *et al.* Highly confined low-loss plasmons in graphene–boron nitride heterostructures. *Nature Mater.* **14**, 421–425 (2014).
11. Principi, A. *et al.* Plasmon losses due to electron–phonon scattering: the case of graphene encapsulated in hexagonal boron nitride. *Phys. Rev. B* **90**, 165408 (2014).
12. Echtermeyer, T. J. *et al.* Strong plasmonic enhancement of photovoltage in graphene. *Nature Commun.* **2**, 458 (2011).
13. Koppens, F. H. L. *et al.* Photodetectors based on graphene, other two-dimensional materials and hybrid systems. *Nature Nanotech.* **9**, 780–793 (2014).
14. Mak, K. F., Ju, L., Wang, F. & Heinz, T. F. Optical spectroscopy of graphene: from the far infrared to the ultraviolet. *Solid State Comm.* **152**, 1341–1349 (2012).
15. Brida, D. *et al.* Ultrafast collinear scattering and carrier multiplication in graphene. *Nature Commun.* **4**, 1987 (2014).
16. Wagner, M. *et al.* Ultrafast and nanoscale plasmonic phenomena in exfoliated graphene revealed by infrared pump-probe nanoscopy. *Nano Lett.* **14**, 894–900 (2014).
17. Wagner, M. *et al.* Ultrafast dynamics of surface plasmons in InAs by time-resolved infrared nanospectroscopy. *Nano Lett.* **14**, 4529–4534 (2014).
18. Eisele, L. *et al.* Ultrafast multi-terahertz nano-spectroscopy with sub-cycle temporal resolution. *Nature Photon.* **8**, 841–845 (2014).
19. Fei, Z. *et al.* Infrared nanoscopy of Dirac plasmons at the graphene-SiO$_2$ interface. *Nano Lett.* **11**, 4701–4705 (2011).
20. Dai, S. *et al.* Tunable phonon polaritons in atomically thin van der Waals crystals of boron nitride. *Science* **343**, 1125–1129 (2014).
21. Lui, C. H., Mak, K. F., Shan, J. & Heinz, T. F. Ultrafast photoluminescence from graphene. *Phys. Rev. Lett.* **105**, 127404 (2010).
22. Ulbricht, R., Hendry, E., Shan, J., Heinz, T. F. & Bonn, M. Carrier dynamics in semiconductors studied with time-resolved terahertz spectroscopy. *Rev. Mod. Phys.* **83**, 543–586 (2011).
23. Frenzel, A. J., Lui, C. H., Shin, Y. C., Kong, J. & Gedik, N. Semiconducting-to-metallic photoconductivity crossover and temperature-dependent Drude weight in graphene. *Phys. Rev. Lett.* **113**, 056602 (2014).
24. Gierz, I. *et al.* Snapshots of non-equilibrium Dirac carrier distributions in graphene. *Nature Mater.* **12**, 1119–1124 (2013).
25. Caldwell, J. D. *et al.* Sub-diffractional, volume-confined polaritons in a natural hyperbolic material: hexagonal boron nitride. *Nature Commun.* **5**, 5221 (2014).
26. Ni, G. X. *et al.* Plasmons in graphene moiré superlattices. *Nature Mater.* **14**, 1217–1222 (2015).
27. Winnerl, S. *et al.* Carrier relaxation in epitaxial graphene photoexcited near the Dirac point. *Phys. Rev. Lett.* **107**, 237401 (2011).
28. Kashuba, A. B. Conductivity of defectless graphene. *Phys. Rev. B* **78**, 085415 (2008).
29. Briskot, U. *et al.* Collision-dominated nonlinear hydrodynamics in graphene. *Phys. Rev. B* **92**, 115426 (2015).

Author contributions

All authors were involved in designing the research performing the research and writing the paper.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.N.B.

Competing financial interests

F.K. is one of the cofounders of Neaspec, producer of the s-SNOM apparatus used in this study.

Methods

Devices. Our hBN/G/hBN high-quality devices were fabricated using a polymer-free multilevel stacking dry-transfer process to form the encapsulated graphene structure (Supplementary Section 1). This technique is essential for maintaining graphene flakes free of contamination, bubbles and wrinkles. Consequently, the charge neutrality point of graphene in these samples is very close to the zero gate voltage V_g at ambient conditions, as confirmed from our near-field gating experiments. After multi-stacking transfer, the metal contacts were fabricated at the exposed graphene edge[4]. Note that graphene is entirely encapsulated within the hBN flakes.

At ambient conditions, the carrier mobility near the Dirac point exceeds 140,000 cm^2 V^{-1} s^{-1}. For the carrier densities relevant to the plasmonic regime $n > 3 \times 10^{12}$ cm^{-2}, the mobility is reduced considerably. However, it still exceeds by more than an order of magnitude that of the first-generation graphene devices and approaches the room-temperature limit set by the acoustic phonon scattering[4].

Figures 1b and 2a show optical microscope and AFM images of the hBN/G/hBN device. The region of a tapered graphene microcrystal encapsulated between the two hBN layers can be clearly seen in the AFM topography. The graphene edges in our devices were preserved in their pristine state without any exposure to oxygen plasma etching, a process that is likely to introduce extrinsic edge doping and/or scattering. The underlying Si/SiO$_2$ (285 nm) substrate serves as the back gate to which we apply voltage V_g (up to 120 V at ambient conditions) to tune the carrier density in graphene[30]. Equation (S4) is based on the assumption[19] that the sheet optical conductivity $\sigma(\omega)$ of graphene is dominated by the intraband component $\sigma_{intra}(\omega) = D(T_{el})/[\pi(\Gamma - i\omega)]$.

Infrared nano-imaging experiments. The infrared nano-imaging experiments were performed using a scattering-type scanning near-field optical microscope (s-SNOM). Our s-SNOM (http://www.neaspec.com) is equipped with continuous wave mid-IR quantum cascade lasers (http://www.daylightsolutions.com) and CO$_2$ lasers (http://www.accesslaser.com). The s-SNOM is based on an AFM with a tip apex radius of 25 nm operating in a tapping frequency around 270 kHz.

A pseudo-heterodyne interferometric detection module is implemented in our s-SNOM to extract both scattering amplitude s and phase of the near-field signal. In the current work, we discuss only the former. To subtract the background signal, we demodulated the near-field signal at the third harmonic of the tapping frequency. All of the infrared nano-imaging experiments were performed under ambient conditions.

Nano-Fourier transform infrared (nano-FTIR) experiments. The ultrafast nano-FTIR set-up is based on s-SNOM. The s-SNOM is integrated with a mid-IR source (Lasnix) that emits ultrashort (<40 fs) pulses produced via an 80 MHz near-IR Er-doped fibre laser (Toptica Photonics). Pulses of mid-IR probe light are generated by difference-frequency mixing of near-IR pulses in a 2 mm thick, z-cut GaSe crystal. Ultrafast temporal resolution (~ 200 fs) is achieved by operating the s-SNOM in a pump–probe mode whereby the synchronized 40 fs near-IR pulses of the fibre laser ($\lambda = 1.56$ µm) serve as a pump with a variable pump–probe time delay.

An important technical detail of our pump–probe study is that application of 30 V back-gate voltages was necessary to achieve the plasmonic features shown in Fig. 1c,d. The static Drude weight produced by this gate voltage alone is insufficient to support steady-state plasmonic oscillations within our mid-IR probing window without photo-excitation (Fig. 3b). Previously, we carried out an analogous spatio-temporal imaging protocol on lower mobility graphene/SiO$_2$ structures[16]. Although time-resolved nano-spectroscopy revealed a pump-induced modification of the plasmonic response[16], we were unable to observe real-space plasmonic oscillations, with or without an applied gate voltage. We also wish to point out that the probe pulse in our apparatus is insufficient to launch plasmons or trigger plasmon interference patterns (Supplementary Section 6).

References

30. Das, A. *et al.* Monitoring dopants by Raman scattering in an electrochemically top-gated graphene transistor. *Nature Nanotech.* **3,** 210–215 (2008).

Ultrafast optical switching of infrared plasmon polaritons in high-mobility graphene

G. X. Ni[1,2*], L. Wang[3*], M. D. Goldflam[2], M. Wagner[2], Z. Fei[2], A. S. McLeod[2], M. K. Liu[2,4], F. Keilmann[5], B. Özyilmaz[1], A. H. Castro Neto[1], J. Hone[3], M. M. Fogler[2], D. N. Basov[2,6†]

[1]Centre for Advanced 2D Materials and Graphene Research Centre, National University of Singapore, Singapore, 117546.

[2]Department of Physics, University of California, San Diego, La Jolla, California 92093, USA.

[3]Department of Mechanical Engineering, Columbia University, New York, NY 10027, USA.

[4]Department of Physics, Stony Brook University, Stony Brook, NY 11794, USA.

[5]Ludwig-Maximilians-Universität and Center for Nanoscience, 80539 München, Germany.

[6]Department of Physics, Columbia University, New York, USA.

[*]These authors contribute equally.

[†]Corresponding author: dbasov@physics.ucsd.edu

1. **BN/G/BN van der Waals structures fabrication and characterization**
2. **Graphene plasmon dispersion**
3. **Plasmon fringes in cw nano-imaging data**
4. **On the origin of plasmon damping**
5. **Calculations of non-equilibrium plasmon dispersion**
6. **Derivative analysis of non-equilibrium plasmon dispersion**

1. BN/G/BN van der Waals structures fabrication and characterization

We obtained microcrystals of both graphene (G) and hexagonal boron nitride (BN) using mechanical exfoliation. The starting point was exfoliating BN micro-crystals on top of polypropylene carbonate (PPC) polymer. The stack is assembled using the same method as in our previous work [S1]. In short, the BN/graphene/BN stack is made using the top BN micro-crystals to pick up single layer graphene and the bottom BN micro-crystals in sequence. Bubble- and wrinkle-free stacks were then transferred on to SiO₂/Si substrates. For the etching process, only a small portion of the graphene edge was exposed for edge-contact deposition (Fig. 1b). Thus most of the graphene specimen remains intact. Finally, the edge-contact was made by depositing 1/10/60 nm of Cr/Pd/Au layers. A complete BN/G/BN device optical image is shown in Fig. 1b of the main text.

2. Graphene plasmon dispersion

Figure S1 represents the frequency (ω) – momentum (q) dispersion diagram of the polaritonic modes at the air/BN/G/BN/SiO$_2$/Si interface visualized via a false color map of the imaginary part of the reflectivity $r_p = r_p(q, \omega)$ [S2]. The graphene chemical potential is E$_F$ ~ 0.32 eV, which corresponds to the carrier density $n = 7.5 \times 10^{12}$ cm^{-2}. This choice of E$_F$ gives the plasmon dispersion in a good agreement with the data points at $\omega = 890$ cm^{-1} and 935 cm^{-1} (the black dots in Fig. S1), which are obtained from the cw images presented in Fig. 2 of the main text. The additional dispersion lines seen in Fig. S1 are due to phonon-polariton guided waves in BN slabs [S3, S4]. These latter modes are confined to the frequency bands $\omega = 746$–819 cm^{-1} and 1370–1610 cm^{-1}. Both of these bands are outside the frequency range in the current study, and so we do not discuss them further.

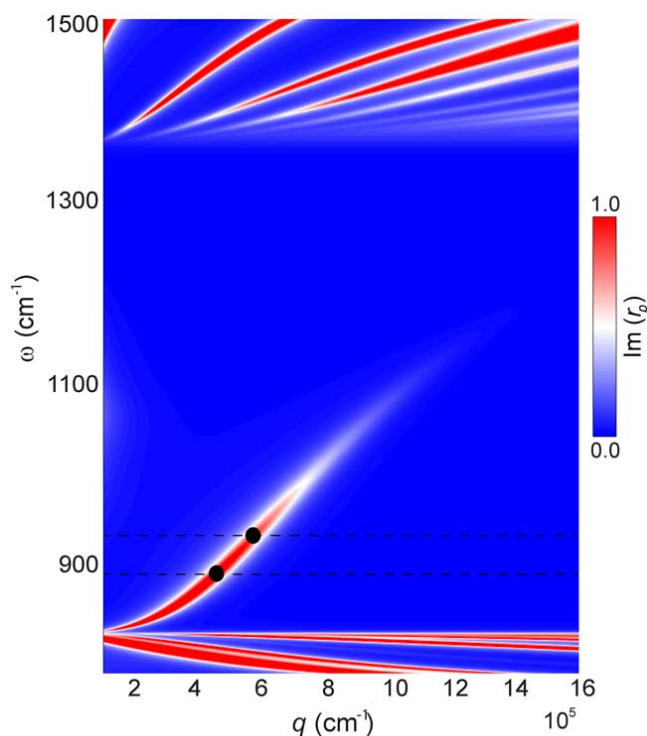**Figure S1 | Graphene dispersion diagram.** Calculated frequency (ω) – momentum (q) dispersion diagram for polaritonic modes of BN/G/BN on SiO$_2$/Si substrates at E$_F$ ~ 0.32 eV. The black dots are the experimental data derived from cw nano-imaging.

3. Plasmon fringes in cw nano-imaging data

For a general theoretical analysis of the plasmon fringes observed in the near-field experiments we imagine formally splitting the total electric field underneath the tip into two contributions: the field that would exist in an infinite graphene sheet and the perturbation caused by the presence of the graphene edge. The fringes must result

from (constructive or destructive) interference of these two contributions. In the previous work [S5, S6], it was effectively assumed that the perturbation caused by the edge can be described by the method of images. In other words, the effect of the edge located at distance L from the tip was taken to be equivalent to having an "image" tip the distance $2L$ away from the physical tip. The amplitude of the plasmons launched by the image tip at the position of the actual tip carries an extra factor equal approximately to $e^{2iqL}L^{-1/2}$, where the second factor describes the algebraic decay of a wave spreading radially in two dimensions. This leads to the $\lambda_p/2$-periodic variation of the total signal with L. This physical picture has been successful in describing plasmonic fringes in relatively low-mobility graphene. However, our new experimental data presented in the main text (as well as data reported in [S7]) indicate that for the high-mobility graphene this picture is incomplete. These data reveal that the plasmon fringes appear as double peaks, implying that two distinct periodicities exist, $\lambda_p/2$ and λ_p.
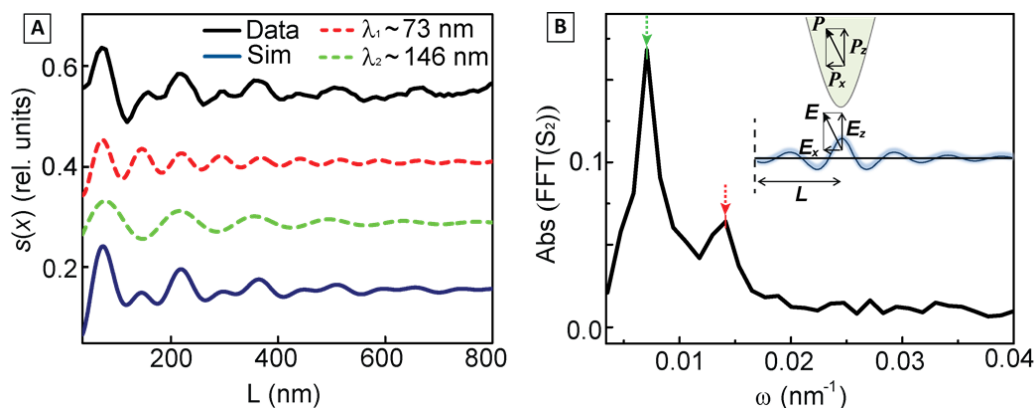

Figure S2 | Double periodicity of plasmon fringes. A) The black solid line is the plasmonic trace acquired at gate voltage 115 V; the red and green dashed lines correspond to $\frac{\exp(2iqL)}{\sqrt{L+R1}}$ and $\frac{\exp(iqL)}{L+R2}$, respectively; the blue solid line displays their sum, which reproduced the key features of the black trace, except very near the graphene edge. The curves are offset for clarity. B) The Fourier transform amplitude of the black trace in panel A; the two peaks correspond to the λ_1 and λ_2 periods. Inset: a schematic illustrating a possible origin of the λ_p-periodic fringes (see text).

In order to quantify the two periods, we utilized Fourier transform, which we found quite adequate and simpler to implement than Hankel transform employed in ref. S7. The Fourier transform of the line profile acquired at the gate voltage of 115 V is shown in Fig. S2A&B. The two prominent peaks correspond to the periods $\lambda_1 = 73$ nm and $\lambda_2 = 146$ nm. Following [S7], we also considered a data fit in the form $A\frac{\exp(2iqL)}{\sqrt{L+R1}} + B\frac{\exp(iqL)}{L+R2}$ with adjustable parameters $A = -0.25, B = 1.74, R_1 = -9.34$ nm, $R_2 = -28.36$ nm. This phenomenological expression reproduces the measured double-peak plasmon fringes except for the very first one, see Fig. S3A. The fit has an unphysical divergence at small $L = -R_2$.
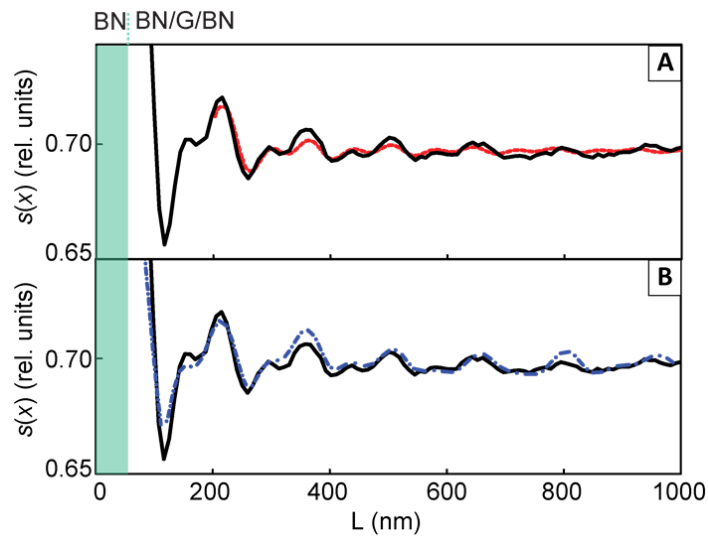
Figure S3 | Plasmon line profile fitting. A) The black solid line is the raw data from Fig. 2d of the main text; the red dotted line is the fit using the phenomenological formula of [S7] (see text). B) Numerical simulation of the line profile using the spheroidal tip model (the blue dotted line).

In accordance with the above discussion, the shorter period is $\lambda_p/2$ and its appearance is understood from the method of images. The physical origin of the λ_p periodicity requires further explanation. One possibility, raised in [S7], is that the light beam incident on the graphene edge can also launch plasmons. Such plasmons would have the requisite qL phase shift with respect to those launched by the tip. The excitation of the plasmons by the edge can occur only due to the component of the incident electric field normal to the edge [S9]. If this mechanism were operational, the amplitude of the longer-period oscillations would show a cosine-like dependence on the polarization angle θ of the incident light. Therefore the plasmon fringes launched by the edge should be absent when the incident electric field is parallel to the edge ($\theta = 0$) and should be the strongest when it is perpendicular to it ($\theta = \pi/2$). No such dependence has been observed in the experiment. Figure S4A shows the raw data obtained from two different orientations of graphene edges with respect to the incident radiation. These patterns are nearly identical. Clearly, the λ_p periodicity of plasmon fringes is found at both polarizations. This finding indicates that direct launching of the plasmons by incident field scattered off the graphene edge is a small effect.
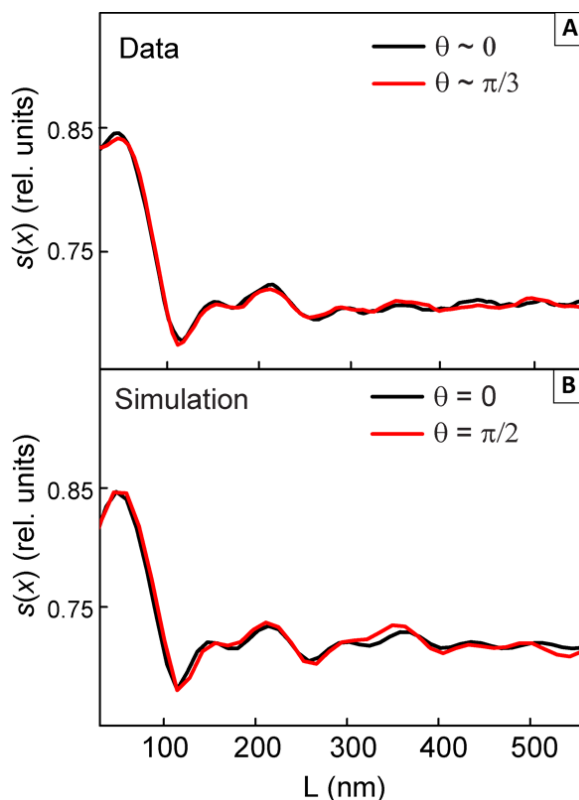
Figure S4 | Plasmonic fringes near the edge of graphene. A) Raw data for plasmon fringes across two different graphene edges: one is oriented at $\theta \sim 0$ with respect to the direction of the incident beam and the other one is at $\theta \sim \pi/3$. B) Using the spheroidal model, we calculated the plasmon line profiles both for $\theta = 0$ and $\theta = \pi/2$. The agreement between data and modeling results indicates that the contribution from direct launching of the plasmons by graphene edge alone is small.

To investigate this issue further, we studied the electrodynamic response of the system (tip plus sample) by numerical simulations. Our modeling assumes the spheroidal shape of the tip. As the basic elements of this model have already been described in detail in the previous papers [S5, S8], we mention only new aspects of this numerical code. Previously, the scattering amplitude $s(\omega)$ (before demodulation) was ascribed to the component p_z of the total radiating dipole of the tip normal to the surface. In the current work, the in-plane component p_x is also taken into account. The model also includes the in-plane dipole moment of the graphene sheet. Thus, the described above hypothesis of launching of the plasmons by the edge should be automatically incorporated in the model as well. As shown in Fig. S3B, our numerical simulations reproduce the doubly-periodic line form of the observed signal. Additionally, its polarization dependence (see Fig. S4B) is almost negligible, in agreement with the experiment shown in Fig. S4A but in contrast with edge-launching hypothesis. Our modeling therefore suggests that the λ_p-periodic part of the signal could be due to the p_x polarization of the tip, which is included in the new simulation code. The effect of this component of the radiating dipole is shown schematically in the inset of Fig. S2B.

The physical explanation of this effect must have to do with the fact that the method of images is only an approximation. It would be strictly valid only if the plasmons obeyed the 2D wave equation. Instead, the plasmons in graphene are governed by a more complicated integral-differential equation [S5, S8]. We speculate that because of this distinction, there should be a correction to the method of images, in the form of a local charge accumulation near the edge. Because of the long-range nature of the Coulomb force, this extra charge can produce the in-plane electric field E_x and hence the dipole polarization p_x of the tip directly, without the intermediary of a plasmon wave, see Fig. S2B. The net result is that the plasmon has to propagate only the distance L not $2L$, leading to the evident λ_p-periodicity of the near-field signal. Although this scenario gives a qualitative explanation of the observed effect, the quantitative analysis requires further theoretical work, which we will report elsewhere.

4. On the origin of plasmon damping

In the following paragraphs, we discuss the origin of graphene plasmon losses in the cw nano-IR images that we examined as a function of the gate-induced carrier density. As with previous studies [S7], it is instructive to represent these data in a form of two-dimensional plots mapping of near-field $s(x)$ signal as a function of gate voltage across the graphene edge within a single experiment (in Fig. S5A). These data were acquired by consecutive scanning along the same line across graphene edge at increasing gate voltage. Information on the variation of both λ_p and γ_p with respect to electrostatic doping is encoded in data in Fig. S5A. The analysis reported below allows us to extract the plasmon lifetime as well as the optical conductivity of encapsulated graphene. As described in the main text, the plasmon damping is defined as [S5]:

$$\gamma_p(\omega) \approx \frac{\sigma_1(\omega)}{\sigma_2(\omega)} + \frac{\kappa_2(\omega)}{\kappa_1(\omega)} \tag{S1}$$

Here, $\sigma_1(\omega)$ and $\sigma_2(\omega)$ are the real and imaginary parts of optical conductivity of graphene; $\kappa_1(\omega)$ and $\kappa_2(\omega)$ are the real and imaginary parts of the effective dielectric function of the substrate environment. The first term in Eq. (S1) accounts for plasmonic losses associated with intrinsic electronic properties of graphene whereas the second term is solely determined by the optical phonon scattering of the surrounding materials.

Using the spheroidal tip model [S8], the value of γ_p can be obtained through a fitting protocol matching the experimental data against a theoretical fringe profile. From the best fit, we obtained $\gamma_p \sim 0.02$, independent of carrier density in the range of 4×10^{12} to 7.5×10^{12} cm^{-2}. In contrast, the carrier mobility scales inversely with carrier density as $\mu \propto 1/n$. At n $\sim 7.5\times10^{12}$ cm^{-2}, the intrinsic room-temperature carrier mobility which is only limited by acoustic phonon scattering was predicted to be $\sim$ 16,000 cm^2/Vs (see Fig. S6) [S7, S10, S11]. The hBN optical constants used here were obtained from recent measurements [S3, S4], from which we estimate that plasmon dissipation

induced by scattering of optical phonons in hBN is $\frac{\kappa_2(\omega)}{\kappa_1(\omega)} \approx 0.005$. Taking the value of $\frac{\sigma_1(\omega)}{\sigma_2(\omega)} \approx 0.015$, we obtain a dissipation rate of graphene plasmons $\tau^{-1} = \frac{\sigma_1(\omega)}{\sigma_2(\omega)} \times \omega \approx 13$ cm^{-1}, corresponding to a DC mobility of $\sim 14,000$ cm^2/Vs at $\sim 7.5 \times 10^{12}$ cm^{-2} doping level, which agrees well with the theoretically calculated phonon scattering limitation. We therefore conclude that our results approach the ultimate limit of graphene plasmon propagation at ambient conditions, in which the dominant dissipation mechanism arises from intrinsic scattering from the graphene acoustic phonon. Our results also agree well with the recently reported experimental and theoretical study of intrinsic plasmon damping in BN/G/BN systems [S7, S10].

The low plasmon damping is also reflected in the optical conductivity, as shown in Fig. S5B. Here the real part of the optical conductivity is less than $0.2\pi e^2/2h$ at a doping level of 7.5×10^{12} cm^{-2}, consistent with the theoretical estimate of σ_1 obtained from non-interacting Dirac quasiparticle model. To this end, we also calculated the plasmon lifetime τ as a function of carrier densities. It is found that τ is ~ 1 ps, which is remarkably long and confirms that our results approaching the ultimate limitation set by phonon scattering. [S10, S12].
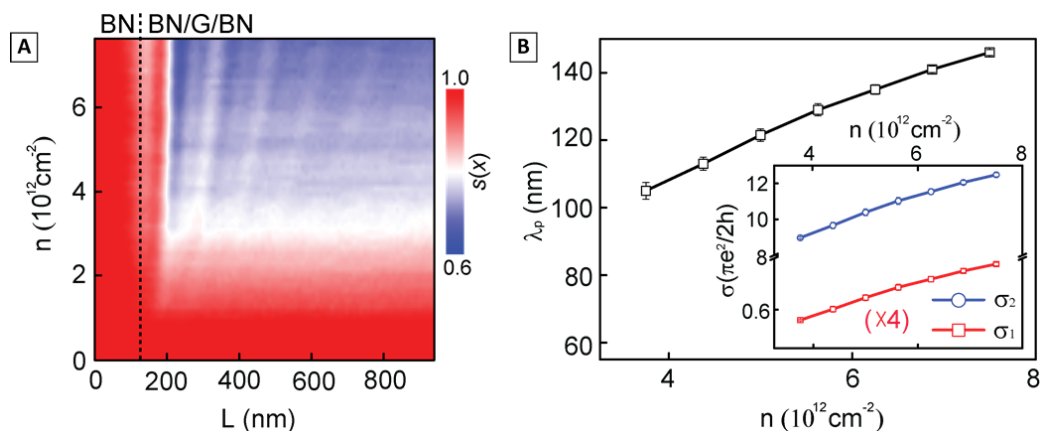

Figure S5 | Analysis the origin of plasmon damping. A) 2D diagram of *s(x)* as a function of gate voltage in real space for an infrared wavelength of 890 cm^{-1}. Multiple plasmon fringes with monotonically increasing plasmon wavelength as function of gate voltage are clearly seen. B) Extracted plasmon wavelength versus carrier density. Inset shows the optical conductivity as function of doping.

In Figure S6, we also plotted the carrier mobility μ as a function of the carrier density n, at doping levels relevant to this work. Here, the intrinsic acoustic phonon scattering at room-temperature was treated as the limiting factor [S7, S10, S11]. At $n \sim 7.5 \times 10^{12}$ cm^{-2}, the intrinsic room-temperature carrier mobility was predicted to be $\sim 16,000$ cm^2/Vs.
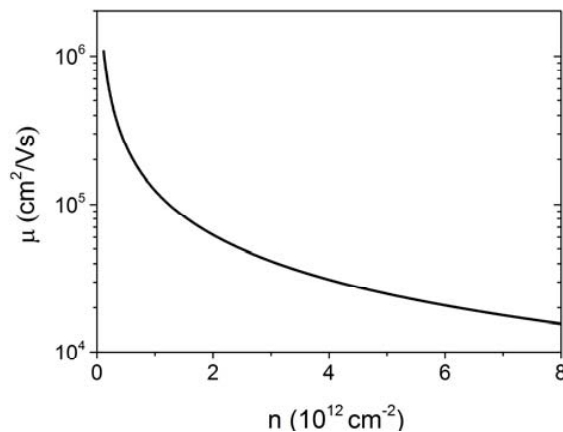
Figure S6 | Carrier mobility versus doping. Calculated DC mobility limited by acoustic phonon scattering with the deformation-potential coupling constant D = 19 eV at 300 K for different densities n = 1.0×10^{12} cm^{-2} and 8.0×10^{12} cm^{-2}.

5. Calculations of non-equilibrium plasmon dispersion

In this section, we discuss the pump-induced plasmon dispersion calculations. In the main text (Fig. 3b), we have simulated the ω–q maps for the surface modes at the air/BN/G/BN/SiO$_2$/Si interface with the electronic temperature as one of the fitting parameters. This is different from the graphene plasmon dispersion calculation described in section 2, where the carrier density serves as the variable parameter. The difference stems from non-equilibrium state of graphene at time scales relevant to our experiments: 30 fs – 5 ps [S13]. At these time scales, optical excitation in graphene preserves the electron density, but results in an elevated carrier temperature [S13, S14].

In order to calculate the temperature dependent plasmon dispersion we employed the standard sheet conductivity of graphene consisting of inter- and intraband contributions:

$$\sigma(\omega) = \sigma_{\text{intra}}(\omega) + \sigma_{\text{inter}}(\omega), \qquad \sigma_{\text{intra}}(\omega) = \frac{i}{\pi}\frac{D(T_{\text{el}})}{\omega + i\Gamma}. \qquad (S2)$$

The Drude-like intraband component that has been shown to dominate for our excitation conditions [S14] contains the scattering rate Γ and the temperature dependent Drude weight (Eq. S3). In Figure S7 we plot the calculated plasmon dispersion for carrier temperatures from 1500 to 3200 K. At T_{el} = 3200 K, a best match to experimental data was obtained. This good agreement further confirms the plasmonic origin for the real-space features observed in our pump-probe data.

The Drude spectral weight governing the plasmonic response is given by the equation

$$D(T_{el}) = \left(\frac{8\pi^2 e^2}{h^2}\right) k_B T_{el} \ln\left[2\cosh\frac{\mu}{2k_B T_{el}}\right], \tag{S3}$$

where μ is the chemical potential, e is the elementary charge, and h is Planck's constant [S14]. At zero temperature, Eq. (S3) predicts $D = (2\pi e/h)^2 E_F$, where $E_F = \mu(T_{el} = 0)$ is the Fermi energy. At high T_{el}, $k_B T_{el} \gg |\mu|$, Eq. (S3) implies a linear increase of D with T_{el}. The plasmon wavelength λ_p is directly proportional to $D(T_{el})$ and inversely proportional to the effective permittivity $\kappa(\omega)$ (the geometric mean of the in- and out-of-plane permittivities) of the encapsulating hBN:

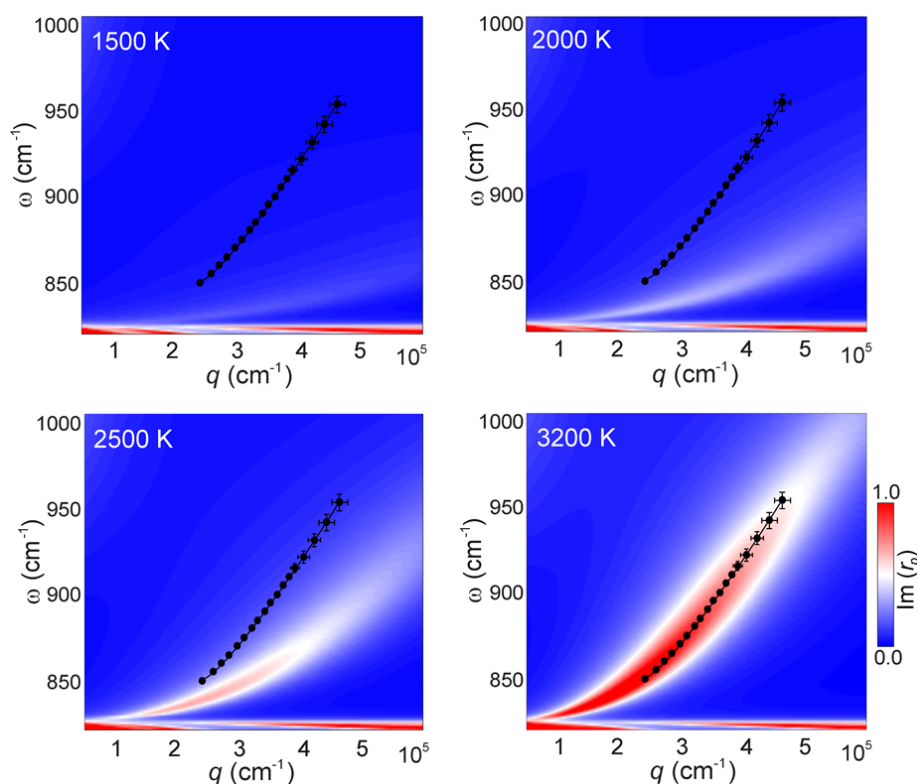$$\lambda_p(\omega) = \frac{4\pi}{\kappa(\omega)\omega^2} D(T_{el}) . \tag{S4}$$

Figure S7 | Plasmon dispersion at different temperatures. Calculated plasmon dispersion at temperatures from 1500 K to 3200 K (false color). The black dots with solid line are the experimental data as discussed in the main text.

6. Derivative analysis of non-equilibrium plasmon dispersion

As described in the main text, we have utilized the derivative of the raw data at $t = 0$ ps delay time to determine the peak positions in our scattering amplitude data, as shown in Fig. 1c (reproduced here as Fig. S8A). This derivative $\frac{ds(\omega,x)}{dx} = 0$ plot is

presented in Fig. S8B. The first prominent aspect of this two-dimensional derivative map is the alternating blue and red regimes which are separated by the white lines. The blue corresponds to negative values; the red corresponds to positive values of the derivative and the white lines correspond to the zeros. Thus, the white lines correspond to the extrema of the signal, in particular, the 1st, 3rd and 5th (when apparent) lines are the maxima. The dashed lines in Fig. S8A depict these maxima overlaid on the raw data map. The vertical solid line marks the edge of graphene inside our encapsulated structures. In Fig. S8C, we show an attempt to model the hyper-spectral data obtained in Fig. S8A. Using the spheroidal tip model, the calculated ω-q map does reproduce the gross features of the raw data shown in Fig. S8A. This confirms the correlation of plasmon scattering associated with an increase in T_{el}, as shown in Fig. 3a.

Figure S8D&E shows the raw data and its derivative of s(ω,x) at 2 ps time delay. The peak position can be extracted using the approach discussed above. To determine the electronic temperature at t = 2 ps, we first computed the imaginary part of the complex reflectance $r_p(q, \omega)$, see Fig. S8F. The momentum at which Im $r_p(q, \omega)$ has the maximum is identified with q_p. This value serves as the input into the spheroidal model simulations, which produce the ω-x diagram. The numerically computed fringe positions can then be compared with those found from the derivative of the measured signal. The best match of these two was achieved at T_{el} = 1700 K.
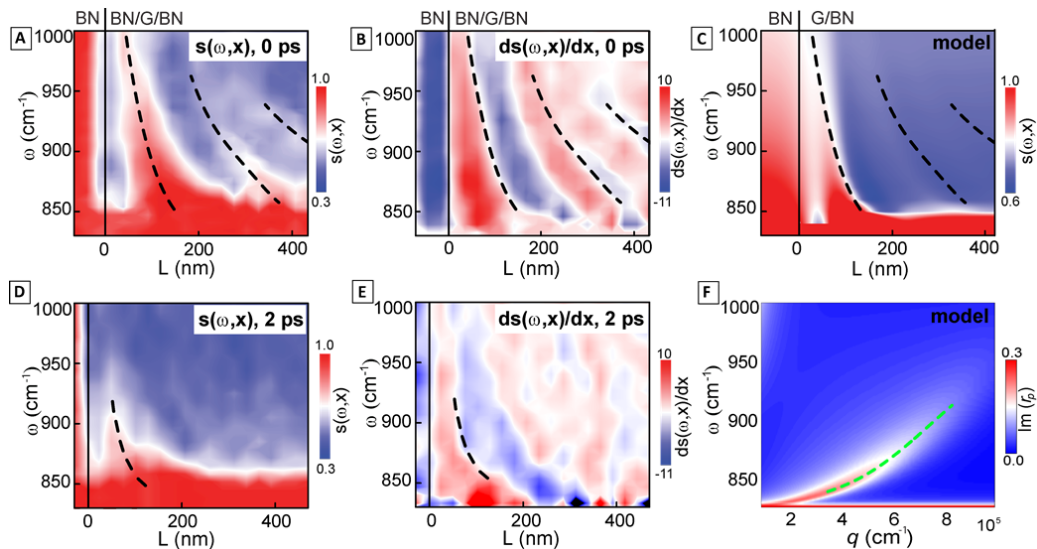

Figure S8 | Non-equilibrium plasmon dispersion. A-C) plasmon dispersion at 0 ps. A) Raw data of s(ω,x) at 0 ps as presented in the main text of Fig. 1c. B) Map of the derivative $\frac{ds(\omega,x)}{dx} = 0$ of our hyper-spectral map at 0 ps. The black dashed lines mark the maxima. C) Calculated two-dimensional hyper-spectral map of s(ω,x) using the spheroidal model. Here the plasmon damping was set to be γ_p = 0.15. The solid black line represents the graphene edge location. D-F) plasmon dispersion at 2 ps. D) Raw data of s(ω,x) at 2ps as presented in the main text of Fig. 1E. E) Two-dimensional hyper-spectral map of the derivative $\frac{ds(\omega,x)}{dx} = 0$ at 2ps. The black dashed line marks

the maximum. F) The dispersion of photo-induced plasmons represented by the imaginary part of the complex reflectance Im $r_p(q,\omega)$ calculated for an electronic temperature of ~ 1700 K at γ_p= 0.15. We obtained Im $r_p(q,\omega)$ through calculations that take into account the multilayer structure of our air/BN(10 nm)/G/BN(30 nm)/SiO$_2$/Si sandwiched devices.
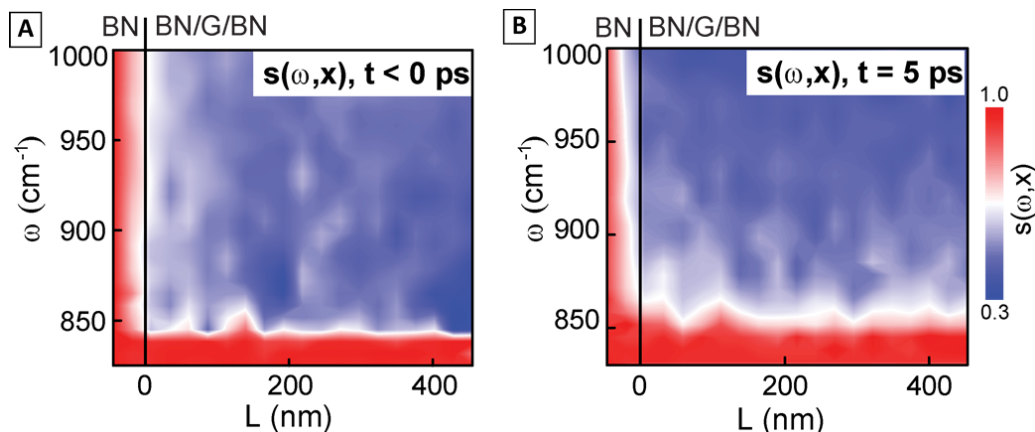

Figure S9 | Ultrafast pump-probe data at different time delays. A) Hyper-spectral map of the near-field scattering amplitude s(ω,x) before pump overlap with probe beam. B) Hyper-spectral map of the near-field scattering amplitude s(ω,x) at a pump-probe delay time of 5 ps, where no dispersion is observed. The black solid line marks the edge between BN/BN and BN/G/BN.

In Figure S9, we show the pump-probe s(ω,x) data obtained at time delay of t = 5 ps and another map obtained before the pump beam overlaps with probe beam (t < 0 ps). In both cases, no plasmon dispersion was observed. This confirms that the observed plasmonic patterns within several picoseconds of photo-excitation is due to the strongly enhanced electronic temperature associated with this dynamical variation of the Drude weight. It also indicates that although the probe beam alone can heat the electrons [S14], it is insufficient to trigger plasmon interference patterns.
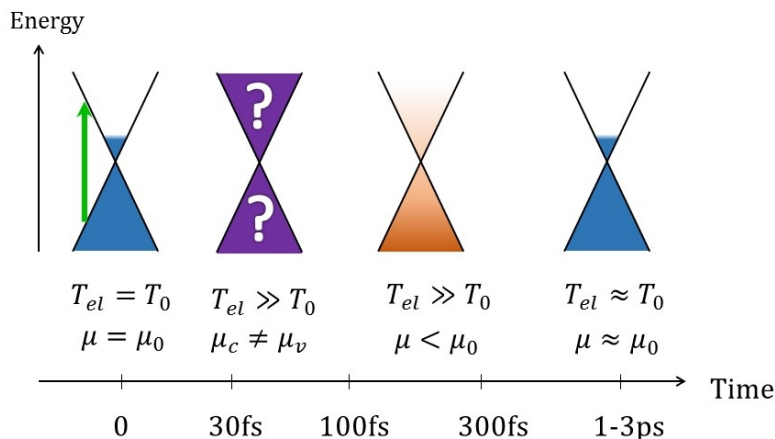

Figure S10 | Schematic evolution of the carrier population dynamics in graphene. Before pumping, electrons and holes have equal temperature T$_0$ and chemical

potential μ_0. Immediately after pumping, the carrier distribution is complicated (non-Fermi-Dirac). Within the first 30-50 fs the interaction of electrons with one another and with optical phonons establish two separate Fermi-Dirac distributions for electrons and holes with nearly equal and very high temperature T_{el} but unequal chemical potentials μ_c, μ_v. After about 100 fs the two chemical potentials equilibrate to the common value μ. Finally, the electron temperature gradually approaches the lattice temperature T_0 mainly due to emission of acoustic phonons. This regime, lasting for several picoseconds, is probed in the present experiments.

Finally, we remark on the possible decay mechanisms for non-equilibrium plasmons. Under the conditions of our experiments, the most likely decay channel appears to be the emission of acoustic phonons, as suggested recently in [S7, S10]. On the other hand, the optical phonons, which have energy of ~0.2 eV may play a role during the first 30-50 fs after the pump pulse when the electron temperature is still higher than 0.2 eV. This and other stages of the relaxation of the electrons are illustrated in Fig. S10, which is constructed based on the literature devoted to highly photo-excited graphene [S13, S15-S18].

References

S1.   Wang, L. *et al*. One-Dimensional Electrical Contact to a Two-Dimensional Material. *Science* **342**, 641-617, (2013).

S2.   Fei, Z. *et al*. Infrared Nanoscopy of Dirac Plasmons at the Graphene-SiO_2 interface. *Nano Lett.* **11**, 4701-4705 (2011).

S3.   Dai, S. *et al*. Tunable Phonon Polaritons in Atomically Thin van der Waals Crystals of Boron Nitride. *Science* **343**, 1125-1129, (2014).

S4.   Caldwell, J. D. *et al*. Sub-diffractional, volume-confined polaritons in a natural hyperbolic material: hexagonal boron nitride. *Nature Commun*, **5**, 5221 (2014).

S5.   Fei, Z. *et al*. Gate-tuning of graphene plasmons revealed by infrared nano-imaging. *Nature* **487**, 82-85 (2012).

S6.   Chen, J. *et al*. Optical nano-imaging of gate-tunable graphene plasmons. *Nature* **487**, 77-81 (2012).

S7.   Woessner, A. *et al*. Highly confined low-loss plasmons in graphene–boron nitride heterostructures. *Nature Mater*. **14**, 421-425 (2014).

S8.   Fei, Z. *et al*. Electronic and plasmonic phenomena at graphene grain boundaries. *Nature Nanotech*. **8**, 821-825 (2013).

S9.   Yan, H. *et al*. Damping pathways of mid-infrared plasmons in graphene nanostructures, *Nature Photonics*, **7**, 394 (2013).

S10. Principi, A. *et al*. Plasmon losses due to electron-phonon scattering: the case of graphene encapsulated in hexagonal Boron Nitride. *Phys. Rev. B* **90**, 165408 (2014).

S11. Hwang, E., Das Sarma, S. *Phys. Rev. B* **77**, 115449 (2008).

S12. Javier Garcia de Abajo, F. Graphene Plasmonics: Challenges and Opportunities, *ACS Photonics*, **1**, 135-152 (2014).

S13. Gierz, I. *et al*. Snapshots of non-equilibrium Dirac carrier distributions in graphene. *Nature Mater.* **12**, 1119–1124, (2013).

S14. Martin, W. *et al*. Ultrafast and Nanoscale Plasmonic Phenomena in Exfoliated Graphene Revealed by Infrared Pump–Probe Nanoscopy. *Nano Lett.* **14**, 894-900 (2014).

S15. Brida, D. *et al*. Ultrafast collinear scattering and carrier multiplication in graphene. *Nature Commun.* **4**, 1987 (2013).

S16. Tierooij, K. J. *et al*. Photoexcitation cascade and multiple hot-carrier generation in graphene. *Nature Phys.* **9**, 248–252 (2013).

S17. Gierz, I. *et al*. Tracking Primary Thermalization Events in Graphene with Photoemission at Extreme Time Scales. *Phys. Rev. Lett.* **115**, 086803 (2015).

S18. Dani, K. M. *et al*. Intraband conductivity response in graphene observed using ultrafast infrared-pump visible-probe spectroscopy. *Phys. Rev. B* **86** 125403 (2013).