Ultrafast and Nanoscale Plasmonic Phenomena in Exfoliated Graphene Revealed by Infrared Pump–Probe Nanoscopy

Martin Wagner,‡ Zhe Fei,‡ Alexander S. McLeod,‡ Aleksandr S. Rodin,‡,# Wenzhong Bao,‡,# Eric G. Iwinski,‡ Zeng Zhao,§ Michael Goldflam,† Mengkun Liu,‡ Gerardo Dominguez,¶ Mark Thiemens,‡ Michael M. Fogler,† Antonio H. Castro Neto,¶,# Chun Ning Lau,§ Sergiu Amarie,¶ Fritz Keilmann,¶ and D. N. Basov*‡,

†Department of Physics, University of California, San Diego, La Jolla, California 92093, United States
‡Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, Massachusetts 02215, United States
§Materials Research Science and Engineering Center, University of Maryland, College Park, Maryland 20742, United States
¶Department of Physics and Astronomy, University of California, Riverside, California 92521, United States
#Department of Physics, California State University, San Marcos, San Marcos, California 92096, United States
⊥Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, United States
¶Department of Physics, National University of Singapore, 117542, Singapore
□Neaspec GmbH, Bunsenstrasse S, 82152 Martinsried, München, Germany
Ludwig-Maximilians-Universität and Center for Nanoscience, 80539 München, Germany

Supporting Information

ABSTRACT: Pump–probe spectroscopy is central for exploring ultrafast dynamics of fundamental excitations, collective modes, and energy transfer processes. Typically carried out using conventional diffraction-limited optics, pump–probe experiments inherently average over local chemical, compositional, and electronic inhomogeneities. Here, we circumvent this deficiency and introduce pump–probe infrared spectroscopy with ~20 nm spatial resolution, far below the diffraction limit, which is accomplished using a scattering scanning near-field optical microscope (s-SNOM). This technique allows us to investigate exfoliated graphene single-layers on SiO2 at technologically significant mid-infrared (MIR) frequencies where the local optical conductivity becomes experimentally accessible through the excitation of surface plasmons via the s-SNOM tip. Optical pumping at near-infrared (NIR) frequencies prompts distinct changes in the plasmonic behavior on 200 fs time scales. The origin of the pump-induced, enhanced plasmonic response is identified as an increase in the effective electron temperature up to several thousand Kelvin, as deduced directly from the Drude weight associated with the plasmonic resonances.

KEYWORDS: Graphene, Dirac plasmon, pump–probe, time-resolved spectroscopy, near-field microscopy, s-SNOM

Graphene is emerging as a promising system for optoelectronic and plasmonic applications.1–4 This is in part because the optical, electronic, and plasmonic phenomena present in graphene are amenable to modifications by chemical doping and electrostatic gating. Specifically, these stimuli influence both the amplitude and the wavelength of Dirac plasmons, the surface charge oscillations in graphene.5,6 Tunable surface plasmons are of high interest in the context of confining and controlling long-wavelength electromagnetic radiation at the length scales typical of X-rays.7 The key novelty of the results presented here is that the plasmonic effects in graphene can be modified on ultrafast time scales with an efficiency rivaling that of electrostatic gating. This is achieved using a novel approach that enables pump–probe spectroscopy in the infrared (IR) region with nanoscale spatial resolution. Common optical pump–probe experiments have been extensively employed to investigate fundamental physical effects in condensed matter.8,9 However, the spatial resolution of these techniques is typically limited by diffraction restricting the minimum sampling area to about 1 μm2 for measurements in the visible and to 102–104 μm2 or more for IR and terahertz (THz) studies. For that reason, pump–probe experiments on graphene have been performed on large area samples10–13 that are prone to the formation of various defects or grain boundaries14,15 obscuring intrinsic electronic and optical phenomena. These irregularities are absent in exfoliated...
graphene with typical dimensions smaller than the wavelength of IR light: the type of specimens we investigated in our experiments.

Here, we implemented a nanospectroscopic IR local probe via S-SNOM16,17 under intense near-infrared (NIR) laser excitation. We employed this new technique to explore the ultrafast response of Dirac plasmons in graphene. Our key observation pertains to strong enhancement of plasmonic features by NIR pumping, which we directly monitor in our time-resolved spectra. We can fully account for this behavior by modeling NIR pumping as light-induced heating of the free-carrier subsystem of graphene. Specifically, we show that the intraband Drude weight \( D \) due to Dirac quasiparticles increases with their effective carrier temperature even though the density of electrons in the system, that is, the number of electrons in the conduction band minus the number of holes in the valence band, remains unchanged under NIR illumination. We stress that the Drude weight is directly accessible in our measurements through the analysis of the plasmonic features in the infrared spectra that we collect with the 200 fs temporal resolution. Furthermore, our data attest to the ability to optically modify Dirac plasmons on 20 nm length scales and 200 fs time scales. An important virtue of our approach is that the conversion of light to electron–hole pairs can now be investigated at length scales characteristic of these light-matter interaction processes and of prototypical photonics/plasmonics devices.1,2,4,7,18,19

Our setup (Figure 1a) integrates a commercial s-SNOM (Neaspec GmbH) with a mid-infrared (MIR) source (Lasnix) emitting ultrashort (<100 fs) pulses based on a 40 MHz NIR fiber laser (Toptica Photonics, Inc.).20–22 We note that owing to field enhancement by the s-SNOM tip23 an in-plane MIR peak field strength of the order of 107 kV/cm can be obtained with such a relatively low-cost, high repetition rate laser system (see Supporting Information). Ultrafast temporal resolution is achieved by operating the s-SNOM in pump–probe mode whereby the synchronized 100 fs NIR pulses of the fiber laser serve as 1.56 μm pump light with variable delay. Exfoliated graphene samples on 300 nm SiO2 were studied at ambient conditions. Raman spectroscopy confirmed an unintentional hole-doping ln = (2.6 ± 1) × 1012 cm–2, corresponding to a chemical potential \( \mu_0 = -(1500 ± 300) \) cm–1.

The near-field response of graphene on SiO2 in the MIR region is dominated by hybrid resonances originating from the coupling of plasmons to the surface phonon modes of the substrate.24 We first consider these plasmonic features in the near-field data in the absence of the NIR excitation. In Figure 2a, we plot \( s(\omega) \) spectra: the second harmonic of the near field scattering amplitude \( s_2 \) normalized to that of silicon. The \( s(\omega) \) spectrum of SiO2 alone (black curve in Figure 2a) is dominated by two surface phonons denoted as \( \alpha \) and \( \beta \), giving rise to a strong peak at \( \omega_\alpha = 1125 \) cm–1 and a weaker one at \( \omega_\beta = 785 \) cm–1.25 A single layer of graphene modifies the main features of the SiO2 substrate response (red curve). We register a 5 cm–1 blueshift of the \( \alpha \) resonance compared to SiO2 and a broadening of its high-frequency shoulder, effects rooted in

Figure 1. (a) Sketch of the NIR pump MIR probe near-field spectroscopy experiment on exfoliated graphene. (b) MIR spectroscopy data were acquired with probe pulses forming these two overlapping, broad-band spectra.

Figure 2. Graphene near-field spectra with and without NIR photexcitation. (a) Near-field amplitude MIR spectra \( s(\omega) \) (without NIR excitation) normalized to Si, for 300 nm SiO2 (black curve) and for single-layer graphene on SiO2 (red curve, probed with attenuated NIR intensity). The SiO2 phonon modes at \( \omega_\alpha = 1125 \) cm–1 and \( \omega_\beta = 785 \) cm–1 are labeled as \( \alpha \) and \( \beta \). (b) NIR pump-induced, spectral changes in the near-field amplitude \( \Delta s(\omega)/s(\omega) \) of graphene at zero delay for varying NIR average powers. The effects of electrostatic gating (without NIR pumping) is shown for comparison. The inset of panel (b) displays \( \Delta s(\omega)/s(\omega) \) for different graphene films at 10 mW pump power. (c) Theoretical effect of increasing the intraband Drude weight \( D \) above the initial nonequilibrium value \( 1.10D \) that serves as reference without NIR pump in intense MIR probe fields. While NIR pumping effects were modeled by an increase in carrier temperature, gating effects were simulated by adjusting the Fermi energy.
the interaction of the surface phonons of the SiO₂ substrate with the Dirac plasmons in graphene.²⁴ Note that the plasmon–phonon coupling is fundamentally a finite momentum effect (see dispersion plot in Figure 3a). The ability of near-field nanoscopy to investigate this effect stems from the AFM tip providing access to momenta of the order of \( \hbar q \approx \pi a \) (vertical dashed lines in Figure 3a,b) where \( a \) is the tip radius.²⁴ Thus, we are able to study surface plasmons in pristine, exfoliated graphene. We note that plasmonic coupling in far-field experiments requires nanostructuring of large-area CVD or epitaxial samples²⁵,²⁶ prone to defects inherent to both growth¹,¹⁴,¹⁵ and patterning²⁶ procedures.

Because our inquiry into the ultrafast plasmon dynamics relies on monitoring the hybrid plasmon–phonon resonances, we discuss the dispersion relationships unraveling the physics of this coupling to set the stage for later analysis of the pump–probe data. The dispersion of the Dirac plasmon alone (red curves in Figure 3a,b) and of the hybrid plasmon–phonon modes (color plots in Figure 3a,b) is calculated²⁸ based on the standard sheet conductivity of graphene intra- and interband components as

\[
\sigma(\omega) = \sigma_{\text{intra}}(\omega) + \sigma_{\text{inter}}(\omega), \quad \sigma_{\text{intra}}(\omega) = \frac{i D}{\pi \omega + i\gamma}, \quad \sigma_{\text{inter}}(\omega) = \frac{i D}{\pi \omega + i\gamma} \tag{1}
\]

The plasmonic response of graphene is dominated by the intraband part that is defined by two parameters: the Drude weight \( D \) and the scattering rate \( \gamma.²⁷ \) Ignoring for a moment the impact of the substrate, the dispersion of the plasmon (red curves in Figure 3a,b) follows the familiar square-root behavior \( \omega \sim (k_F q)^{1/2} \) for \( q \) values below the Fermi wavevector \( k_F \). The horizontal lines in Figure 3a,b denote the SiO₂ surface phonons, defined by \( \text{Re} \kappa(\omega) = 0 \), where \( \kappa(\omega) = (1 + \varepsilon_{\text{SiO}_2}(\omega))^{-1} \) is the arithmetic mean of the dielectric functions of air and SiO₂. Once coupling between the plasmon and SiO₂ phonons is taken into account, the dispersion is dominated by avoided crossings. An immediate consequence of the coupling is the blueshift of the \( \alpha \) resonance that is consistent with the data in Figure 2a.

Next, we examine the influence of NIR pumping on the graphene response at zero time delay between pump and probe pulses. We present pump-induced changes of the scattering amplitude \( \Delta s(\omega)/s(\omega) \) (Figure 2b), collected from a \( \sim 10–20 \) nm spot that is defined by the radius of the tip. The spectra show an increase of the scattering amplitude, which is particularly pronounced at frequencies just above the \( \alpha \) and \( \beta \) modes, amounting to as much as 20% around \( \omega = 1200 \text{ cm}^{-1} \). The upper frequency cutoff in Figure 2b is governed by the signal-to-noise of our data. At \( \omega = 900 \text{ cm}^{-1} \), a weaker and broader peak appears whereas the \( \beta \) mode is only marginally amplified. Multilayer graphene shows a stronger photoinduced response (see inset of Figure 2b) that grows nearly linearly with the number of layers.

Diffraction-limited pump–probe experiments on graphene have established several facts relevant to our nanoscale observations.¹²,¹³,²⁸,²⁹ Immediately after photoexcitation graphene reveals distinct electron and hole chemical potentials that last only 10–30 fs after the pump pulse arrives. Interband scattering of electrons and holes occurs via Auger and carrier multiplication processes that are kinematically allowed and efficient in graphene. The net effect is that 30 fs after the pump pulse is incident on graphene, electrons and holes are characterized by a single Fermi-Dirac distribution with the same effective temperature \( T \) and chemical potential \( \mu(T) \).¹²,²⁸ The 200 fs temporal resolution of our apparatus only enables access to the latter regime of “hot electrons” whereas the effects of rapid equilibration are beyond our range.

Remarkably, the spectral changes associated with pump-induced heating of electrons in single-layer graphene closely resemble those originating from electrostatic gating. As shown in Figure 2b (dots), similar \( \Delta s(\omega)/s(\omega) \) spectra are obtained by applying a gate voltage of \( V_g = -70 \text{ V} \) and thus increasing the sample’s hole-doping. To the best of our knowledge, data in Figure 2b is the first attempt to explicitly compare spectral features induced in graphene by gating and ultrafast photoexcitation. We were able to account for this peculiar similarity of the response to the two distinct stimuli within our intraband Drude model of plasmonic response introduced in the following section.

We now proceed to our quantitative description of the plasmonic response in the regime of pump-induced heating of the electronic system based on eq 1 and introduce a temperature dependent intraband Drude weight (Figure 3c) in the following form:²⁰.
theoretical spectra. The simulation correctly replicates a peak in slight mismatch between the details of the experimental and scattering rate constant at spectra capture the essential aspects of our data even if we keep parameters is presented in Table 1. We found that the model concentration in the conduction band minus the hole fi dominates in our NIR pumping results. Speci\(\text{fi}\) is due to the 3-fold increase of the scattering rate \(\gamma\) results in an increase of the \(\alpha\) of graphene. Owing to small signal-to-noise ratios the full MIR spectral features associated with both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts for the \(\text{meas}\) features of both the transient and the equilibrium data in Figure 2. At equilibrium (Figure 2a), this model accounts For bi- and trilayer graphene, similar behavior showing enhanced intraband Drude weight with temperature is expected to be similar to those originating from electrostatic gating despite the assumption of constant electron density in our modeling. Note that intense in-plane dc fields have recently been predicted to increase the slope of the plasmon dispersion. Another effect that is not expected to play a significant role in our case. That is because in our experiments the ac electric fields average to zero both in time and due to the radial field distribution directly under the tip also in space.

We stress that even though \(\mu(T)\) is suppressed at elevated temperatures (Figure 3c), the Drude weight \(D\) is enhanced above its equilibrium value \(D_0\) for \(T > 0.60 |\mu_0/k_0| = 1270 K\) (black square). The sensitivity of the Drude weight to the electron temperature enables ultrafast optical control of plasmonic effects in graphene at femtosecond time scales. Since \(D\) defines both the amplitude and wavelength of surface plasmons in graphene, both of these fundamental characteristics will be impacted by photoexcitation. For bi- and trilayer graphene, similar behavior showing enhanced intraband Drude weight with temperature is expected when the electron temperature \(k_0 T\) approaches (or exceeds) the interlayer tunneling energy \(\gamma_T = 0.4\) eV, a situation close to our experimental parameters. An approximately linear dispersion in that region should mimic the temperature dependence of \(D\) in single-layer graphene. Data in the inset of Figure
2b and comparable time constants in Figure 4 qualitatively support this interpretation, but quantitative comparison warrants further study.

We now analyze the temporal evolution of the near-field plasmonic response by measuring the spectrally integrated scattering amplitude. In these latter experiments, the probe beam was centered at a frequency close to 1200 cm\(^{-1}\) (see Figure 1b). Thus the scattering amplitude is dominated by the high-frequency flank of the \(\alpha\) plasmon–phonon mode where pump-induced changes are the most distinct, according to Figure 2b. We have carried out these experiments for a sample with terraces of single-, bi- and trilayer graphene (Figure 4c,d). Our assignment of the number of layers in each terrace is based on optical contrast and is confirmed by Raman microprobe spectroscopy. In Figure 4a, we plot the photoinduced changes of the scattering signal \(\Delta s/s\) for various pump–probe time delays \(\Delta t\) collected from an area as small as \(\sim 500 \text{ nm}^2\). We found that the temporal traces show a biexponential time dependence\(^{10,37}\) and extracted the two time constants (Figure 4b). Time-resolved imaging data are presented in Figure 4e. Here we display images of NIR photoinduced changes \(\Delta s/s\) for different pump–probe time delays. A significant increase of the scattering amplitude occurs at temporal overlap (\(\Delta t = 0\)). Naturally, the signal in the region of the SiO\(_2\) substrate is not impacted by the NIR pump. Few-layer graphene exhibits stronger pump-induced changes in scattering amplitude compared to single-layer terraces. This is expected since the absorption cross-section of the pump light scales with the number of layers. For increasing time delay \(\Delta t\), the sequence in Figure 4e displays a decrease in the pump-induced \(\Delta s/s\) signal. Note that our time-resolved imaging does not reveal significant spatial variations in the observed decay constants.

Below, we briefly outline the key features revealed by the temporal profiles of our pump–probe data. The rise time in Figure 4a is limited by the temporal resolution of the MIR probe pulses (200 fs; see Supporting Information). The positive sign of the pump-induced signal is consistent with the idea of increased intraband Drude weight \(D\) that we have established based on the analysis of the pump–probe spectra in Figure 2b. We assign the faster time scale \(\tau_1 \approx 200\) fs to cooling of the hot carriers via optical phonon emission.\(^{12,26,37}\) Analysis of the spectra informed us that this cooling process occurs in the regime of constant electron density even though the Drude weight is suppressed with decreasing carrier temperature. The faster time constant is quite similar for the heavily doped single- and few-layer graphene. The longer time constant \(\tau_2\) is assigned to energy relaxation involving acoustic phonons.\(^{13,37}\) Bi- and trilayer samples show larger \(\tau_2\) values: a plausible consequence of energy diffusion to the substrate.\(^7\)

In conclusion, we reported near-field pump–probe spectroscopy based on s-SNOM combining exceptional spatial, spectral, and temporal resolution. Experimental capabilities of our ultrafast IR nanoscope significantly extend previously reported results.\(^{20,38–42}\) Well-established fiber-based near-field methods\(^{43,44}\) may be easier to implement but are unfortunately limited to visible and NIR frequencies. In contrast, ultrafast s-SNOM described here is capable of probing a broad spectral region from visible\(^{45}\) to far-infrared energies.\(^{46}\) First pump–probe spectroscopy data collected with this apparatus revealed ultrafast optical modulation of the infrared plasmonic response of graphene. This finding is a precondition for ultrafast graphene-based plasmonic devices. In terms of its efficiency, optical control of plasmons in graphene is on par with conventional electrostatic gating. Remarkably, the pulse energies needed to modify the infrared plasmonic response are 2 orders of magnitude smaller than what is typically necessary for comparable ultrafast switching times in metal-based plasmonic structures at NIR frequencies.\(^{47}\) Notably, ultrafast plasmonic tuning in graphene can be readily accomplished using moderately priced fiber lasers, as first demonstrated here. A relevant figure of merit in this context is \(R_{\text{mod}}\); defined by the ratio of the probe intensity modulation depth (\%) to the necessary pump fluence (\(\text{mJ/cm}^2\)). We estimated \(R_{\text{mod}}\) \(\sim 100\) for graphene, an unprecedentedly high value that exceeds \(R_{\text{mod}}\) \(\sim 6\) for ultrafast schemes in metal structures\(^{48}\) or \(R_{\text{mod}}\) \(\sim 40\) for much slower (~100 ps) Si-based devices.\(^{47}\) We remark that this value has been achieved close to the strong 1125 cm\(^{-1}\) SiO\(_2\) surface phonon where the plasmon–phonon–polariton behaves predominantly phonon-like with long polariton lifetimes.\(^{28}\) Real-space plasmon propagation potentially relevant for plasmonic devices has been observed at lower frequencies\(^{49}\) where still \(R_{\text{mod}}\) \(\sim 25\). Significant improvement could be expected from higher pump pulse energies and more efficient NIR light concentration using resonant tips.\(^{50}\) Optimized tips would be beneficial since in our study an area \(\sim 10^6\) times larger than the probed area under the...
tip was illuminated and the total amount of NIR field enhancement is unclear (see Supporting Information). Nevertheless, this first infrared pump–probe experiment beyond the diffraction limit demonstrates the capability of this technique and paves the way to the exploration of a wide range of problems in condensed matter physics, biology, and chemistry.

ASSOCIATED CONTENT

Supporting Information
 Supporting experimental data and theory details. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
 *E-mail: dbasov@physics.ucsd.edu.

Notes

1 The authors declare the following competing financial interest(s): F. K. is cofounder of Neaspec and Lasnix, producers and paves the way to the exploration of a wide range of

REFERENCES

(2) Bao, Q.; Loh, K. P. ACS Nano 2012, 6, 3677–3694.
(4) Grigorenko, A. N.; Polini, M.; Novoselov, K. S. Nat. Photonics 2012, 6, 749–758.