Hot-Carrier Cooling in High-Quality Graphene Is Intrinsically Limited by Optical Phonons


Cite This: ACS Nano 2021, 15, 11285−11295

ABSTRACT: Many promising optoelectronic devices, such as broadband photodetectors, nonlinear frequency converters, and building blocks for data communication systems, exploit photoexcited charge carriers in graphene. For these systems, it is essential to understand the relaxation dynamics after photoexcitation. These dynamics contain a sub-100 fs thermalization phase, which occurs through carrier−carrier scattering and leads to a carrier distribution with an elevated temperature. This is followed by a picosecond cooling phase, where different phonon systems play a role: graphene acoustic and optical phonons, and substrate phonons. Here, we address the cooling pathway of two technologically relevant systems, both consisting of high-quality graphene with a mobility >10 000 cm² V⁻¹ s⁻¹ and environments that do not efficiently take up electronic heat from graphene: WSe₂-encapsulated graphene and suspended graphene. We study the cooling dynamics using ultrafast pump−probe spectroscopy at room temperature. Cooling via disorder-assisted acoustic phonon scattering and out-of-plane heat transfer to substrate phonons is relatively inefficient in these systems, suggesting a cooling time of tens of picoseconds. However, we observe much faster cooling, on a time scale of a few picoseconds. We attribute this to an intrinsic cooling mechanism, where carriers in the high-energy tail of the hot-carrier distribution emit optical phonons. This creates a permanent heat sink, as carriers efficiently rethermalize. We develop a macroscopic model that explains the observed dynamics, where cooling is eventually limited by optical-to-acoustic phonon coupling. These fundamental insights will guide the development of graphene-based optoelectronic devices.

KEYWORDS: graphene, cooling dynamics, hot electrons, transient absorption microscopy, optical phonons, phonon bottleneck

INTRODUCTION

The ultrafast dynamics of photoexcited charge carriers in graphene have received ample attention, initially driven by fundamental scientific interest in the intriguing electron−electron and electron−phonon interactions; cf. refs 1−5. More recently, interest has multiplied as a result of the emergence of highly promising technological applications that exploit these ultrafast dynamics. One example is ultrafast photodetection of visible (VIS) and infrared light,6,7 and even terahertz (THz) radiation.8−10 The ultrafast electronic response of graphene to incoming light has also led to the development of several concepts with relevance for data communication technologies, including modulators and receivers,11−13 and is demonstrated to be crucial for tailoring nonlinear photonics applications.14−17 Part of the aforementioned applications, for instance photodetectors and receivers, exploit the ultrafast photothermoelectric effect in graphene,18,19 where a longer hot-carrier cooling time leads to an increased photoresponse. Other applications, such as modulators, could instead benefit from a short cooling time, as this can lead to a higher modulation speed, i.e., to a broader bandwidth. Generally, these applications require high-quality graphene with a high mobility >10 000 cm² V⁻¹ s⁻¹.
electrical mobility. Clearly, it is crucial to properly understand the cooling dynamics of hot carriers in graphene, and in particular to identify the intrinsic mechanism that ultimately determines the cooling process in high-quality systems.

The decay of photoexcited charges in graphene occurs through a variety of dynamical processes that are the result of the specific properties of Dirac electrons and graphene phonons. Photoexcited charge carriers in graphene first undergo thermalization on a ~100 fs time scale, which occurs through carrier–carrier scattering and leads to a state with an elevated carrier temperature, i.e., a broadened Fermi–Dirac distribution.2−4,10 This photoinduced heating process is very efficient5,21 and, due to the small electronic heat capacity of graphene,6 the electron temperature $T_e$ can be increased significantly ($T_e \gg 500$ K). The hot-carrier state then relaxes back to the ground state with the electronic system at ambient temperature. This cooling process occurs through the interaction between charge carriers and graphene optical phonons, graphene acoustic phonons, and substrate phonons in nearby materials. These phonon-induced relaxation mechanisms have intricate dependencies on intrinsic (sample-dependent) parameters (such as disorder density and intrinsic doping), as well as extrinsic experimental parameters (such as photon energy, incident fluence, and ambient temperature). Thus, determining the dominant cooling channel(s) for excited graphene charge carriers has been challenging.

It is generally believed that charge carriers with enough kinetic energy can relax on an ultrashort time scale of a few hundred femtoseconds by interacting with strongly coupled optical phonons.22,23,24 These are the optical phonons at the $\Gamma$ and $K$ points with an energy of 0.2 and 0.16 eV, respectively.25 Thus, charge carriers with an excess energy $>0.16$ eV above the chemical potential can efficiently cool by optical phonon emission.26 Carriers with an excess energy $<0.16$ eV can only couple to acoustic phonons, generally resulting in very inefficient cooling with decay times up to the nanosecond range.27 In the case of graphene with significant disorder, however, coupling to acoustic phonons becomes much more efficient through disorder-assisted scattering. By scattering with defects, the large momentum mismatch between electrons and acoustic phonons is overcome.28 This disorder-assisted “supercollision” cooling process leads to typical cooling times of a few picoseconds at room temperature, as measured by electrical,29 optical,30,31 and optoelectronic32 techniques. Since disorder also limits the electrical mobility, this disorder-assisted cooling mechanism tends to play an increasingly important role for graphene with lower electrical mobility. Currently, there are contrasting views in the literature regarding the role of different cooling mechanisms. In particular, some studies, such as ref 26, question the role of supercollision cooling.

In the case of graphene with a mobility above $10,000$ cm$^2$ V$^{-1}$ s$^{-1}$, which we here refer to as “high-quality graphene”, the disorder density is low enough that disorder-assisted cooling will not play an important role. Recent experiments on high-quality hBN-encapsulated graphene with a mobility above $30,000$ cm$^2$ V$^{-1}$ s$^{-1}$, however, did not show a strong increase in cooling time.33 This is because a parallel relaxation mechanism emerges in these heterostructures: hot carriers in graphene can decay through near-field interaction with hyperbolic phonon polaritons of the hBN.33–35 Hyperbolic phonon polaritons occur in spectral regions, where the in- and out-of-plane permittivities ($\epsilon_i$ and $\epsilon_d$, respectively) have opposite signs, i.e., the permittivity product $\epsilon_i\epsilon_d$ is negative.36 These hyperbolic spectral regions contain a high density of optical modes, which are accessible via near-field interaction with a large range of momentum vectors. As a result, out-of-plane super-Planckian cooling for hBN-encapsulated graphene occurs with a decay time of a few picoseconds at room temperature.33–35

Here, we aim to unveil the intrinsic physical mechanism that ultimately limits the cooling of hot carriers in high-quality graphene. We address this question using time-resolved optical measurements of the cooling dynamics in two specifically chosen material systems: WSe$_2$-encapsulated graphene and...
suspended graphene. Both systems contain “high-quality” graphene, according to our definition of having a mobility >10 000 cm² V⁻¹ s⁻¹, which is high enough to make “supercollision” cooling inefficient. Furthermore, they have basically nonhyperbolic environments, such that out-of-plane cooling to hyperbolic phonons does not play any role. We show that, despite eliminating these two relaxation channels, cooling still happens on a time scale of a few picoseconds. We attribute this to the high-energy tail of the hot-carrier distribution, where electrons with an energy >0.16 eV above the chemical potential reside. These electrons lose energy by coupling to optical phonons, which in turn couple to graphene acoustic phonons, while the electronic system continuously rethermalizes. Probing the transient optical properties of graphene in the VIS, near-infrared (NIR), and THz ranges, we will show that this cooling mechanism is consistent with the experimentally obtained cooling dynamics.

RESULTS AND DISCUSSION

Cooling Dynamics in High-Quality WSe₂-Encapsulated Graphene. The first material system we study is WSe₂-encapsulated graphene (see Figure 1a). Encapsulation with materials offering atomically planar surfaces via van der Waals stacking, is a successful route for obtaining high-quality graphene, suppressing rippling, preserving high carrier mobility and partially screening long-distance Coulomb scattering with substrate charges. Besides hBN, WSe₂ is arguably one of the most promising encapsulation materials for various technological applications of graphene.39 One of the main reasons is that it leads to a very high room-temperature mobility, which in a recent study reached up to 350 000 cm² V⁻¹ s⁻¹ for hBN/graphene/WSe₂.40 Our sample of WSe₂-encapsulated monolayer graphene was prepared using exfoliation and dry transfer. The WSe₂ flakes were obtained by adhesive tape exfoliation of WSe₂ crystals in the trigonal prismatic phase synthesized by chemical vapor transport, as detailed in the Methods. Bottom and top WSe₂ flakes are 63 and 61 nm thick, respectively, as measured by atomic force microscopy (see Figure 1b and the Supporting Information). The WSe₂/graphene/WSe₂ heterostructure was transferred onto a 280 μm thick CaF₂ substrate that is transparent for VIS and NIR light and allows for pump–probe measurements in transmission geometry.
We first characterize the WSe₂-encapsulated graphene using Raman spectroscopy with a 532 nm laser source (see Methods for details), in order to assess the quality of graphene and extract an estimation of its Fermi energy and charge mobility. In the regions where graphene is semiencapsulated, the Raman fingerprints of graphene, the G-peak and 2D-peak, are clearly visible (see Figure 1c). In the fully encapsulated region, the spectrum is dominated by a large background. Nevertheless, by extending the exposure time, we can identify the G-peak with center frequency of \( \omega_G \approx 1584 \text{ cm}^{-1} \) and width of \( \Gamma_G \approx 20 \text{ cm}^{-1} \) (spectra in the Supporting Information). From the G-peak position and width we estimate that \( E_G \leq 0.1 \text{ eV} \), in agreement with a previous observation for graphene supported on a backgate. We obtained an electron (hole) mobility of \( \sim 2 \times 10^4 \text{ cm}^2 \text{ V}^{-1} \text{s}^{-1} \) for the encapsulated graphene heterostructure with two metal contacts in the range 0.1 \text{ eV}, indicating that WSe₂-encapsulated graphene has a rather homogeneous strain distribution, with little intravalley scattering and high mobility. On the basis of the empirical correlation of ref 43, the measured \( \Gamma_{2D} \) corresponds to a mobility of \( \sim 80 \times 10^3 \text{ cm}^2 \text{ V}^{-1} \text{s}^{-1} \). In order to verify this estimate, we measured the transport properties at room temperature of a similar WSe₂-encapsulated graphene heterostructure with two metal contacts and a backgate. We obtained an electron (hole) mobility of \( \sim 39 \times 10^3 \text{ cm}^2 \text{ V}^{-1} \text{s}^{-1} \) (\( \sim 36 \times 10^3 \text{ cm}^2 \text{ V}^{-1} \text{s}^{-1} \), see Supporting Information). With such high mobility values, disorder-assisted “supercollision” cooling predicts a characteristic cooling time scale of \( >10 \text{ ps} \) (see Supporting Information).

Having established the high quality of our WSe₂-encapsulated graphene sample, we now study its hot-carrier cooling dynamics. We perform time-resolved transient absorption (TA) measurements in the NIR range using a high-sensitivity microscope (see Methods for details on the experimental setup), which effectively serves as an ultrafast electronic thermometer. All experiments are performed under ambient conditions and at room temperature, monitoring the time-dependent differential transmission \( \Delta T/T(t) \). We use pump pulses at 1550 nm (0.8 eV) with a 40 MHz repetition rate and probe pulses at 1700 nm (0.729 eV). These photon energies are chosen well below the band gap of WSe₂ (\( \sim 1.35 \text{ eV} \) for multilayer WSe₂), to avoid photoexcitation of charges in the encapsulant (see reflectance measurements in the Supporting Information). Thus, the role of the WSe₂ is solely that of a dielectric environment. The pump pulses are absorbed in graphene, where they induce carrier heating, \( i.e., \) a broadening of the Fermi–Dirac distribution of the electronic system. The probe pulses are sensitive to this because the broadening leads to Pauli blocking of the interband transitions at the probe photon energy. Therefore, the pump-induced increase in probe transmission is a measure of the electronic temperature. Accordingly, the \( \Delta T/T \) signal in Figure 2a is positive, indicating pump-induced transmission increase, \( i.e., \) absorption bleaching. We can distinguish between fully encapsulated and semiencapsulated graphene regions in Figure 2a, as the former gives rise to a lower transient signal (at a time delay of \( \sim 100 \text{ fs} \)), which is likely the result of strong Fresnel reflections at the air–WSe₂ and WSe₂–graphene interfaces.

Figure 2b shows the relaxation dynamics of the \( \Delta T/T \) signal of WSe₂-encapsulated graphene and compares it to the dynamics for hBN-encapsulated graphene. The decay in WSe₂-encapsulated graphene is significantly slower: the signal decays to 10% of the initial value (at time zero) after a time delay of \( \sim 5.5 \text{ ps} \) for WSe₂-encapsulated graphene, instead of a delay of \( \sim 2 \text{ ps} \) for the hBN-encapsulated graphene. As shown recently, in high-quality, hBN-encapsulated graphene, hot-carrier cooling is dominated by out-of-plane coupling to hyperbolic phonons in hBN. Since the hyperbolic nature of hBN is crucial for making this cooling channel efficient, \( i.e., \) we study the hyperbolicity of WSe₂ using \( ab \text{ initio} \) density functional theory calculations. The comparison of in-plane and out-of-plane permittivities of WSe₂ and hBN, reported in the Supporting Information, shows that WSe₂ is much less hyperbolic than hBN. Therefore, out-of-plane cooling to the phonons in the encapsulant will be relatively inefficient in WSe₂-encapsulated graphene, in qualitative agreement with the slower cooling that we observe. Quantitatively, however, both disorder-assisted “supercollision” cooling and out-of-plane cooling to the encapsulant would predict longer cooling times (see Supporting Information), well above 10 ps for the WSe₂-encapsulated graphene sample, which is not what we observe.

In order to explain the relatively fast cooling dynamics observed in the WSe₂-encapsulated graphene, we therefore consider an intrinsic cooling mechanism, based on the coupling of electrons with energy \( >0.16 \text{ eV} \) above the chemical potential to optical phonons. The optical phonons, at the \( \Gamma \) and \( K \) points, in turn, are anharmonically coupled to acoustic phonons. An important ingredient for this cooling channel is the continuous rerethermalization of the electronic system. Microscopically, this means that once the electrons with high enough energy (more than 0.16 eV above the chemical potential) have relaxed by coupling to optical phonons, the remaining electrons will rerethermalize through carrier–carrier scattering. This means that some electrons will end up with an energy that is high enough to emit optical phonons. Even at an electron temperature of 300 K, there is a significant fraction of electrons that can emit optical phonons. As a result, this is a rather efficient cooling channel for graphene at room temperature. We will describe the analytical model of this cooling mechanism in more detail in the following discussion and in the Supporting Information.

We calculate the cooling dynamics and compare the results to the experimental data. In the calculation, we also take into account the superlinear relation connecting the \( \Delta T/T \) signal to the change in carrier temperature. As shown in Figure 2c, the experimental data and calculated signal are in good agreement when we use an optical-to-acoustic phonon decay time of 2 ps and a Fermi energy of 0.1 eV. This optical phonon lifetime lies within the range 1.2–2.55 ps, previously reported for graphene. Figure 2d shows the temperature dynamics corresponding to the calculated transient signals in Figure 2c, which are clearly nonexponential. The fast initial decay corresponds to efficient coupling to optical phonons, whereas the slower subsequent decay is the result of the hot-phonon bottleneck, where optical phonons are cooling to acoustic phonons.

**Cooling Dynamics in High-Quality Suspended Graphene.** In order to determine if this cooling mechanism is intrinsic to graphene, we study a second technologically relevant material system in which any thermal exchange with the environment is inherently excluded: high-quality suspended graphene. This sample contains large-area graphene grown by chemical vapor deposition (CVD), which was then transferred onto a transmission electron microscopy (TEM)
grid; see Figure 3a. This was achieved through a polymer-free approach, using the method described in ref 49. As a TEM grid we used holey carbon, which is convenient as it has locations with relatively large hole sizes and high yield. (c) TEM image of the suspended graphene, showing the absence of disorder. (d) Plot of 2D peak frequency $\omega_{2D}$ as a function of G peak frequency $\omega_{G}$. Dashed lines represent the expected dependence in the strain-free (red dot-dot-dashed line) and doping-free (blue dot-dashed line) material. (e) Fermi energy $E_F$ distribution extracted from vector decomposition of peaks positions in panel d, and centered at $|E_F| \approx 0.18$ eV. (f) Obtained distribution of 2D-peak widths, indicating a mobility of $>17 000 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$, following calculations of ref 43.

Figure 4. Hot-carrier cooling dynamics of suspended graphene. (a) Transient transmission $\Delta T/T$ map at time zero, showing individual holes where graphene is suspended, acquired with pump tuned at 3.1 eV (400 nm) and probe at 1.55 eV (800 nm). (b) $\Delta T/T$ dynamics (colored dots) for suspended graphene at five different pump fluences from 3 to 38 $\mu$J/cm$^2$. The positive $\Delta T/T$ results from the pump-induced carrier heating that leads to decreased interband conductivity due to Pauli blocking, and thus decreased absorption of the NIR probe light. The colored lines are the calculated $\Delta T/T$ dynamics, based on the intrinsic cooling mechanism where hot electrons cool via the combination of emission of optical phonons, continuous rethermalization of the electron distribution, and coupling of optical to acoustic phonons. (c) Dynamical evolution of the electron temperature $T_e$, corresponding to the calculated $\Delta T/T$ dynamics in panel b. Here, a fast decay due to electron-optical-phonon coupling is followed by a slower decay due to optical-to-acoustic phonon coupling.

We characterize our suspended graphene sample using various microscopic techniques; see Figure 3. First, we show an image taken using scanning electron microscopy, evidencing excellent graphene coverage of the holes and providing an indication of hole sizes (see Figure 3b). TEM measurements furthermore show an atomically perfect lattice (see Figure 3c). We then perform Raman spectroscopy, in order to estimate the Fermi energy and charge mobility of the suspended graphene. Performing a strain-doping analysis of the G-peak and 2D-peak, obtained from several positions on the sample (see Figure 3d), we obtain a distribution of carrier densities (see Figure 3e), which corresponds to an average Fermi energy of $\sim 0.18$ eV. Using the measured width of the 2D-peaks in Figure 3e and the empirical correlation of ref 43, we extract a mobility of $>17 000 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. This result is in agreement with electrical measurements performed on a sample of suspended graphene, prepared in an identical way; see ref 50. The extracted mobility confirms the high quality of the sample, for which disorder-assisted “supercollision” cooling predicts a cooling time $>10$ ps (see Supporting Information).
Figure 5. Hot-carrier cooling dynamics in high-quality graphene. (a) Process of electronic cooling explained through schematics of the carrier density $N$ as a function of carrier energy for three electron temperatures (1200, 800, and 400 K). Cooling occurs through a combination of (i) optical phonon emission by electrons with energy $>0.16$ eV above the chemical potential (red-shaded area), (ii) rethermalization of the electronic distribution, and (iii) anharmonic coupling of optical phonons to acoustic phonons. (b) Calculated “instantaneous” cooling time for a given initial electron temperature for three different Fermi energies. (c) Calculated cooling time as a function of lattice temperature $T_L$ for a very small $\Delta T_e$ and $E_F = 0.03$ eV. For panels b and c, cooling occurs through optical phonon emission and continuous rethermalization of the electronic system. The hot-phonon bottleneck is not included.

Having established the high quality of our suspended graphene sample, we now study its hot-carrier cooling dynamics using time-resolved transient absorption microscopy (see Methods for details on the experimental setup). We use a pump tuned at 400 nm (3.1 eV) and a probe at 800 nm (1.55 eV). Similar to the NIR TA measurements, we observe photoinduced bleaching ($\Delta T/T > 0$) of the probe, due to heating-induced Pauli blocking. The $\Delta T/T$ map at zero time delay (overlapping pump and probe pulses) in Figure 4a shows that we can resolve the individual holes spatially and that a clearly distinctive transient signal comes from inside—from the suspended graphene. For this experiment we use relatively small spot sizes (see Methods), in order to resolve individual suspended regions.

The experimental TA dynamics of our high-quality suspended graphene sample are shown in Figure 4b for five different fluences. Similar transient optical responses were observed in suspended graphene prepared by exfoliation, in the VIS (400 nm pump, 800 nm probe), and in the NIR (830 nm pump, 1100–1400 nm probe). We describe the dynamics with the same cooling mechanism as for WSe$_2$-encapsulated graphene, with the same optical-to-acoustic phonon coupling time of 2 ps (details in the Supporting Information). For this sample we use an $E_F = 0.15$ eV, close to the experimentally determined average value. Figure 4c reports the temperature dynamics corresponding to the $\Delta T/T$ dynamics in Figure 4b, again showing an initial fast decay due to coupling between electrons and optical phonons, followed by slower decay due to optical phonons coupling to acoustic phonons.

As a final experimental test, we complement our TA measurements on suspended graphene, which probe interband transitions, with optical-pump THz probe (OPTP) measurements, which probe intraband transitions; see Figure S3 of Supporting Information. Despite the different optical transitions that are probed, both techniques essentially function as ultrafast electrical thermometers for graphene. This is because in both cases the probe absorption is affected by pump-induced changes in the carrier temperature. The observed OPTP dynamics, reported in the Supporting Information, can be described with the same model as before, confirming the validity of this intrinsic cooling mechanism.

We now discuss in more detail the cooling mechanism that we have used to describe the experimentally obtained cooling dynamics (schematically shown in Figure 5a). First, we note that in many early time-resolved studies on graphene the dynamics were explained using a qualitatively similar mechanism involving electrons decaying to optical phonons, and a hot-phonon bottleneck; cf. refs 24, 46, 52, and 53. Some of these studies applied a phenomenological two- (or three-) temperature model for electron cooling via optical phonons (and acoustic phonons), which (implicitly) based on (re)thermalization of the electronic system. Also using electrical transport measurements, signatures of cooling via optical phonons were identified (for bilayer graphene).55 This cooling mechanism, however, was often thought to only mediate cooling for carriers with high enough energy to couple directly to optical phonons55 and was typically used for experiments with relatively high fluence and therefore relatively high electron temperatures. For the rest of the carriers in the hot-carrier distribution, and for cooling of systems with relatively low electron temperature, alternative cooling channels were considered. In particular, for cooling below a carrier temperature of $\sim 1500$ K, disorder-assisted “supercollision cooling” to graphene acoustic phonons18–30,32 and out-of-plane cooling to (hyperbolic) substrate phonons,33–35 which can both give rise to picosecond cooling at room temperature, were thought to be the dominant mechanisms. In
2016, a microscopic, numerical, simulation of the cooling dynamics of hot carriers in graphene was presented, based on electron-to-optical phonon coupling, and including rerethermalization of the carrier system. The calculated cooling times were used to explain qualitative trends in decay times measured by OToP in samples of multilayer epitaxial graphene on SiC and monolayer graphene grown by CVD. These results motivated us to (re)consider cooling via optical phonons as the intrinsic cooling pathway for high-quality graphene, where disorder-assisted cooling and out-of-plane cooling are inefficient. The cooling mechanism is schematically explained in Figure 5a. We consider two phonon baths, $\Gamma$ and $K$ phonons, and we capture qualitatively the effect of temperature smearing of the electron Fermi surface in the phonon density of states. The latter is important, because we are dealing with elevated temperatures that lead to a smeared-out Fermi surface.

We developed an analytical model to describe the hot-carrier cooling dynamics in graphene. The details of the derivation are shown in the Supporting Information. Briefly, we solve the following rate equations for the electron temperature $T_e(t)$, and phonon temperature $T_\alpha(t)$:

$$
C_\alpha (T_e(t)) \frac{dT_e(t)}{dt} = -h \sum_\alpha \omega_\alpha R_\alpha (T_e(t), T_\alpha(t))
$$

$$
D(\omega_\alpha, T_e(t)) \frac{dT_\alpha(t)}{dt} = \frac{R_\alpha (T_e(t), T_\alpha(t))}{M_\alpha (T_e(t))}
$$

$$
- \gamma_\alpha n_\alpha (T_\alpha(t)) - n_\alpha (T_\alpha^{(0)})]
$$

The left-hand side of the first rate equation contains the electronic heat capacity $C_\alpha (T_e(t))$ and the temporal derivative of the electron temperature. For each electronic temperature, the chemical potential is obtained by imposing that the total carrier density is fixed. The description of electrons in terms of a single temperature and chemical potential implicitly assumes that electron–electron interband scattering processes occur on a much shorter time scale than the electron–phonon dynamics. The dependence of the chemical potential on carrier temperature that we use (see Supporting Information) correspond to those that follow from conservation of total carrier density, where both intraband and interband carrier–carrier scattering are taken into account. The right-hand side of eq 1 describes the emission of optical phonons, where the sum is over the two optical phonon modes (labeled by $\alpha$), at the $\Gamma$ and $K$ point, $\omega_\alpha$ is the frequency of mode $\alpha$, and $R_\alpha (T_e, T_\alpha)$ is the rate of $\alpha$-phonon emission. We calculate this rate analytically using a Boltzmann-equation approach, which contains one important input parameter, namely, the electron–phonon coupling strength. For this, we use the value 11.4 eV/Å, obtained by density functional theory calculations, and verified by experiments. The left-hand side of the second rate equation contains $D(\omega_\alpha, T_e) = \partial n_\alpha (T_e) / \partial T_e$, where $n_\alpha (T_e)$ is the phonon occupation function. Its right-hand side contains a first term due to the emission of optical phonons by electrons, where the parameter $M_\alpha (T_e)$ measures the size of the portion (an annulus) of the phonon Brillouin zone that is heated in the electron-cooling process. This parameter depends, assuming the phonon dispersion to be flat, on the maximum momentum that can be exchanged between electrons and phonons, and thus on the electron temperature $T_e$. The second term describes the decay of optical phonons to acoustic phonons, where $T_\alpha^{(0)}$ is the equilibrium phonon temperature, i.e., the lattice temperature $T_L$. For the optical phonon decay term, we use the parameter $\gamma_\alpha$ as a phenomenological damping rate. There are essentially two adjustable parameters in our calculations: the optical phonon lifetime $\tau_\alpha^{-1}$, and the parameter $\nu$ that governs the temperature dependence of the phonon number density (see Supporting Information). We note that this thermodynamic model, characterized by a single electronic temperature and chemical potential, is based on underlying microscopic carrier–carrier interactions that lead to carrier (re)thermalization, as well as carrier–phonon and phonon–phonon interactions that lead to carrier cooling.

We have seen that this analytical model is able to accurately describe the experimentally obtained cooling dynamics. The hot-phonon bottleneck, occurring when the density of emitted optical phonons is so high that they cannot completely decay into acoustic phonons and part of their energy is scattered back to the electronic system, becomes more and more important with increasing initial $T_e$, and it leads to an overall slower cooling of the graphene hot carriers. For applications operating with a small heating $\Delta T_e$, however, cooling is ultimately determined by electron-optical-phonon cooling. This regime is likely relevant for applications that require low power consumption with low light intensities. Therefore, we analytically study the cooling in the absence of the hot-phonon bottleneck, i.e., when $\tau_\alpha \to \infty$. We calculate that in this case, cooling at room temperature takes about $\sim 500$ fs, for $E_F \sim 0.3$ eV, and $\sim 1.4$ ps for 0.1 eV. Cooling to optical phonons will quickly become less efficient upon decreasing the lattice temperature $T_L$. Around 200 K, we find a cooling time around 5 ps, whereas this increases to $\sim 4$ ns at 100 K (see Figure 5c). We note that when $\Delta T_e$ is not small, the effect of increased cooling time with decreased lattice temperature is much weaker. Thus, obtaining a longer intrinsic cooling time requires a reduction of both $T_L$ and $\Delta T_e$.

As a final discussion, we mention another physical mechanism that can play a significant role in the cooling process of hot carriers in high-mobility graphene (see Supporting Information for details). This mechanism is called “diffusive cooling”, where the electron system cools down by in-plane heat diffusion out of the initially heated spot. Until now, this mechanism was mainly considered in the case of micron-sized channels of ultrahigh-quality graphene ($\gg 10,000$ cm$^2$/(V s)) and at cryogenic temperatures. Cooling via diffusion depends on the electronic heat diffusivity $D$, which increases with charge mobility and Fermi energy. In the case of pump–probe measurements, the observed evolution of the electron temperature increase due to lateral heat diffusion depends on both pump and probe spot sizes ($\sigma_p$ and $\sigma_e^2$ respectively), according to (see Supporting Information):

$$
\Delta T_e(t) \propto \frac{\sigma_p^2 \sigma_e^{2}}{\sigma_p^2 + \sigma_e^{2} + 2 D t}
$$

Using this equation, we calculate that, for spot sizes below 1 mm, “diffusive cooling” can become the dominant cooling channel in high-mobility graphene samples, even at room temperature. In the measurements we performed on WSe$_2$-encapsulated graphene, “diffusive cooling” does not play a large role, due to the relatively large spot sizes that we used (see Figure S12 of the Supporting Information). However, for the measurements on suspended graphene, with smaller spot sizes,
it is very likely that “diffusive cooling” contributes to the probed cooling dynamics, in particular in the tail of the dynamics. It is also conceivable that this leads to an overestimation of the optical phonon decay rate $\gamma_r$. We do note that this “diffusive cooling” is not an intrinsic cooling mechanism, since its contribution becomes negligible for sufficiently large spot sizes.

CONCLUSION

Using three different time-resolved measurement techniques (NIR TA, VIS TA, and OPTP) and two different high-quality, technologically relevant, graphene systems (WSe$_2$-encapsulated and suspended CVD graphene), we have shown that hot carriers decay through an intrinsic mechanism involving optical phonon emission and constant rethermalization of the electronic system. Electrons with an energy $>$0.16 eV above the chemical potential couple to optical phonons, which in turn decay to acoustic phonons while the electronic system continuously rethermalizes. The electron-to-optical-phonon cooling component gives rise to subpicosecond cooling. Due to the hot-phonon bottleneck governed by the anharmonic coupling of optical to acoustic phonons, an approximately biexponential cooling occurs, where the second decay component has a characteristic time scale of a few picoseconds. The overall decay becomes slower for increasing initial electron temperature (higher incident fluence) due to the hot-phonon bottleneck. Whereas previous studies considered a similar mechanism for cooling of high-energy carriers, we here show that this is the intrinsically limiting cooling mechanism for high-mobility graphene, also under relatively low incident fluences and therefore lower electron temperatures. Our analytical model suggests that this mechanism will quickly become less efficient upon decreasing the ambient temperature $T_0$, provided that also the amount of heating is small $\Delta T_e < T_0$. Thus, operating graphene with low incident fluence and at reduced ambient temperatures is likely a promising approach to optimize optoelectronic applications exploiting hot carriers in graphene.

METHODS

High-Sensitivity Transient Absorption Microscopy. The transient absorption microscope for measurements in the NIR is custom-built starting from an Er-doped fiber laser (Toptica-Femto fiber pro) generating 300 mW, 150 fs pulses centered at 1550 nm with 40 MHz repetition rate. A portion of the output of the laser is used as the pump pulse, and it is modulated with an acousto-optic modulator operating at 1 MHz. The NIR probe pulse is obtained by a supercontinuum (SC) generation focusing part of the laser fundamental in a highly nonlinear fiber. The high-energy component of the SC and the fundamental frequency are filtered out with a long-pass filter cutting at 1600 nm, and the component at 1700 nm is selected with a double-grating monochromator with 5 nm spectral resolution. The pump and the probe are collinearly focused on the sample with an objective (Olympus-UCPLN-IR with magnification 100X and NA = 0.85) over a spot size of 2.8 (3.0) $\mu$m for the probe (pump) pulse. The probe transmitted by the sample is collected with an achromatic doublet with a 8 mm focal length and detected by an InGaAs balanced amplified photodiode with 4 MHz bandwidth. The component of the probe at the modulation frequency is measured with a lock-in amplifier (HFLI, Zurich Instruments) using 300 ms effective time constant resulting in a $\Delta T/T$ sensitivity below $10^{-6}$. Transient transmission dynamics is monitored by changing the pump–probe time delay with an optical delay line, while the pump–probe maps (images) at fixed time-delay are acquired by moving the sample with a motorized three-axis piezo-stack linear stage (Newport NPXYZ100). Image size of 120 × 120 pixels is used in the experiment. The width of the cross-correlation between pump and probe pulses at the sample is $\sim 210$ fs.

The transient absorption measurements in the VIS on suspended graphene are performed using 400 nm pump and 800 nm probe pulses with 150 fs pulses at 76 MHz repetition rate. Both pump and probe pulses are carefully overlapped and focused at the same focal plane with a 40X/0.6 NA objective to focal spot sizes of 1.35 and 0.9 $\mu$m, respectively. The pump beam is modulated with an optical chopper at 6.4 kHz. The probe is delayed temporally with a mechanical delay line and detected in transmission on a balanced photodiode via lock-in detection. Further details of the setup are described in ref 59.

Raman Spectroscopy. For Raman characterization, we used an inVia confocal Raman spectrometer from Renishaw plc, equipped with 473 and 532 nm CW laser sources. The laser beam was focused onto the sample through a 100X objective lens, with 0.89 NA. The nominal fwhm of the Gaussian beam at the focus is estimated to be $\sim 1$ $\mu$m, and the step size for the maps is set to 4 $\mu$m. The laser power hitting the sample in the selected configurations for the measurements were 1.6 mW for the 473 nm laser and 0.7 mW for the 532 nm. With an exposure time per pixel of 60 s, the fluence was 122 and 53 mJ/$\mu$m$^2$, respectively. Depending on the measurement, a 1800 lines/mm grating or a denser 2400 lines/mm were used. The spectra are calibrated with respect to the Si peak at 520 cm$^{-1}$.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsnano.0c10864. Hyperbolic cooling model, cooling dynamics probed with terahertz pulses, topography of encapsulated graphene, electron mobility of WSe$_2$-encapsulated graphene, Raman characterization of WSe$_2$-encapsulated graphene, fully encapsulated vs semiconfined graphene, differential reflectance of WSe$_2$-encapsulated graphene, cooling due to lateral heat diffusion, cooling via disorder-assisted acoustic phonon scattering, cooling via optical phonons (PDF)

AUTHOR INFORMATION

Corresponding Author
Klaas-Jan Tielrooij — Catalan Institute of Nanoscience and Nanotechnology (ICN2), BIST & CSIC, Barcelona 08193, Spain; orcid.org/0000-0002-0055-6231; Email: klaas.tielrooij@icn2.cat

Authors
Eva A. A. Pogna — NEST, Institute Nanoscienze-CNR and Scuola Normale Superiore, 56127 Pisa, Italy; Department of Physics, Politecnico di Milano, 20133 Milan, Italy; orcid.org/0000-0003-4779-3549
Xiaoyu Jia — Max-Planck-Institut für Polymerforschung, 55128 Mainz, Germany
Alessandro Principi — School of Physics and Astronomy, University of Manchester, M13 9PL Manchester, U.K.
Alexander Block — Catalan Institute of Nanoscience and Nanotechnology (ICN2), BIST & CSIC, Barcelona 08193, Spain
Luca Banszerus — JARA-FIT and second Institute of Physics, RWTH Aachen University, 52074 Aachen, Germany; orcid.org/0000-0002-1855-1287
Jincan Zhang — Center for Nanochemistry, College of Chemistry and Molecular Engineering, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing
ACKNOWLEDGMENTS

We would like to thank Andrea Tomadin for discussions. The authors acknowledge funding from the European Union Horizon 2020 Programme under Grant Agreement No. 881603 Graphene Core 3. ICN2 was supported by the Severo Ochoa program from Spanish MINECO (Grant No. SEV-2017-0706). A.P. acknowledges support from the European Commission under the EU Horizon 2020 MSCA-RISE-2019 programme (project 873028 HYDROTRONICS) and from the Leverhulme Trust under grant RPG-2019-363. K.J.T. acknowledges funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement No. 804349 (ERC StG CUHL), RyC fellowship No. RYC-2017-22330, and IAE project PID2019-111673GB-I00 and financial support through the MAINZ Visiting Professorship. X.J. acknowledges the support from the Max Planck Graduate Center with the Johannes Gutenberg-Universität Mainz (MPGC). J.Z. acknowledges the support from National Natural Science Foundation of China (No. S2072042). Z.L. acknowledges support from the University of Liege under Special Funds for Research, IPD-STEMA Programme. M.J.V. gratefully acknowledges funding from the Belgian Fonds National de la Recherche Scientifique (FNRS) under PDR grant T.0103.19-ALPS. Computational resources were provided by CERI (FRS-FNRS G.A. 2.5020.11) and the Zenobe Tier-1 supercomputer (Gouvernement Wallon G.A. 1117545) and by a PRACE-3IP DECI grant 2DSPin and Pylight on Beskow (G.A. 653838 of H2020). ICFO was supported by the Severo Ochoa program for Centers of Excellence in R&D (CEX2019-000910-S), Fundació Privada Cellex, Fundació Privada Mir-Puig, and the Generalitat de Catalunya through the CERCA program. N.v.H. acknowledges funding from the European Commission (ERC AdG 670949-LightNet), the Spanish Plan Nacional (PGC2018-096875-B100), and the Catalan AGAUR (2017SGR1369).

REFERENCES

(7) Koppen, F. H. L.; Mueller, T.; Avouris, P.; Ferrari, A. C.; Vitiello, M. S.; Polini, M. Photodetectors Based on Graphene, Other

Notes

The authors declare no competing financial interest.


