Title
Ultraconfined Plasmonic Hotspots Inside Graphene Nanobubbles

Permalink
https://escholarship.org/uc/item/8h60c1xg

Journal
Nano Letters, 16(12)

ISSN
1530-6984

Authors
Fei, Z
Foley, JJ
Gannett, W
et al.

Publication Date
2016-12-14

DOI
10.1021/acs.nanolett.6b04076

Peer reviewed
Ultraconfined Plasmonic Hotspots Inside Graphene Nanobubbles

Z. Fei,†,‡,§ J. J. Foley, IV, †, , W. Gannett, , M. K. Liu, †, Il S. Dai, † G. X. Ni, † A. Zettl, †,§ M. M. Fogler, † G. P. Wiederrecht, ‡ S. K. Gray, ‡ and D. N. Basov†, ‡

†Department of Physics, University of California, San Diego, La Jolla, California 92093, United States
‡Center for Nanoscale Materials, Argonne National Laboratory, Argonne, Illinois 60439, United States
§Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, United States
IlDepartment of Chemistry, William Paterson University, Wayne, New Jersey 07470, United States
¶Department of Physics, University of California at Berkeley, Berkeley, California 94720, United States
fiMaterials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States
†Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, United States
fiDepartment of Physics, Columbia University, New York, New York 10027, United States

Supporting Information

ABSTRACT: We report on a nanoinfrared (IR) imaging study of ultraconfined plasmonic hotspots inside graphene nanobubbles formed in graphene/hexagonal boron nitride (hBN) heterostructures. The volume of these plasmonic hotspots is more than one-million-times smaller than what could be achieved by free-space IR photons, and their real-space distributions are controlled by the sizes and shapes of the nanobubbles. Theoretical analysis indicates that the observed plasmonic hotspots are formed due to a significant increase of the local plasmon wavelength in the nanobubbles. Such an increase is attributed to the high sensitivity of graphene plasmons to its dielectric environment. Our work presents a novel scheme for plasmonic hotspot formation and sheds light on future applications of graphene nanobubbles for plasmon-enhanced IR spectroscopy.

KEYWORDS: Graphene nanobubbles, heterostructures, nanoinfrared imaging, plasmon hotspots, FDTD simulation

Plasmonic hotspots are specific surface regions of metallic nanostructures where the plasmon field is highly confined in real space and thus becomes strongly enhanced in intensity. They attracted broad research interest in the context of both fundamental physics related to high-field electrodynamics as well as technological applications in biosensing and surface enhanced Raman/infrared (IR) spectroscopy. So far, studies of plasmonic hotspots are focused on noble metal based nanostructures where plasmons are normally in the visible or near-ultraviolet frequency range. Here we report on real-space nanoinaging of plasmonic hotspots in graphene, a novel plasmonic material that supports IR plasmons with high confinement, long lifetimes, electrostatic tunability, and propensity toward forming hybrid modes with other polaritons in two-dimensional van der Waals heterostructures. The observed hotspots reside in graphene nanobubbles that are mass-produced by transferring chemical-vapor-deposited graphene films onto the hexagonal boron nitride (hBN) substrates following a standard thermal annealing process. As introduced in the Supporting Information, the nanobubbles could possibly encapsulate a mixture of air, water, or hydrocarbons that originate from the graphene transfer and annealing processes. Prior to our research, these unique nanobubbles have served as platforms for studying the strain effects of graphene as well as the plasmon—phonon coupling phenomena of the graphene/hBN heterostructure. The IR nanoscope we utilized for nano-IR imaging is illustrated in Figure 1, panel a, where the IR light (solid arrow) from a continuous-wave laser is focused at a metalized tip of an atomic force microscope (AFM). The IR-illuminated sharp AFM tip acts as both a launcher and a detector of surface plasmon polaritons. The backscattered light (dashed arrow) off the tip—graphene system contains essential information about plasmons underneath the tip. The nanoscope collects simultaneously the AFM topography and near-field scattering amplitude (s). As demonstrated in previous studies, s is a reliable measure of the out-of-plane electric field amplitude (E_z) underneath the tip. In Figure 1, panel a, we plot an image of three-dimensional topography for a typical sample, where two graphene nanobubbles (labeled as “A” and...
"B") can be clearly visualized. Both nanobubbles are pyramid-shaped with a triangular base. The lengths of the bottom edges \(L\) of the bubble A are about 250 nm with a height \(H\) of less than 20 nm. Bubble B has a smaller size in all dimensions compared to A. For both A and B, the aspect ratios \(H/L\) are below 0.1, which indicate relatively shallow topographic curvature of these bubbles. Such small aspect ratios apply to all other nanobubbles we found on our samples. As an example, we show a large-area scan with tens of nanobubbles densely distributed in the sample area (Figure 1b). Among all these bubbles, the majority reveals the pyramidal shape. In addition to these bubbles, we can also see a number of wrinkles that connect one bubble to another throughout the map. The nano-IR image simultaneously taken with Figure 1, panel b is plotted in Figure 1, panel c where one can see many bright hotspots within the sample area.

To reveal the details of these hotspots, we show in Figure 1, panel d a zoomed-in nano-IR image that is in the same sample area (Figure 1b). Among all these bubbles, the majority reveals the pyramidal shape. In addition to these bubbles, we can also see a number of wrinkles that connect one bubble to another throughout the map. The nano-IR image simultaneously taken with Figure 1, panel b is plotted in Figure 1, panel c where one can see many bright hotspots within the sample area.

To reveal the details of these hotspots, we show in Figure 1, panel d a zoomed-in nano-IR image that is in the same sample area (Figure 1b). Among all these bubbles, the majority reveals the pyramidal shape. In addition to these bubbles, we can also see a number of wrinkles that connect one bubble to another throughout the map. The nano-IR image simultaneously taken with Figure 1, panel b is plotted in Figure 1, panel c where one can see many bright hotspots within the sample area.

To reveal the details of these hotspots, we show in Figure 1, panel d a zoomed-in nano-IR image that is in the same sample area (Figure 1b). Among all these bubbles, the majority reveals the pyramidal shape. In addition to these bubbles, we can also see a number of wrinkles that connect one bubble to another throughout the map. The nano-IR image simultaneously taken with Figure 1, panel b is plotted in Figure 1, panel c where one can see many bright hotspots within the sample area.

To reveal the details of these hotspots, we show in Figure 1, panel d a zoomed-in nano-IR image that is in the same sample area (Figure 1b). Among all these bubbles, the majority reveals the pyramidal shape. In addition to these bubbles, we can also see a number of wrinkles that connect one bubble to another throughout the map. The nano-IR image simultaneously taken with Figure 1, panel b is plotted in Figure 1, panel c where one can see many bright hotspots within the sample area.

To reveal the details of these hotspots, we show in Figure 1, panel d a zoomed-in nano-IR image that is in the same sample area (Figure 1b). Among all these bubbles, the majority reveals the pyramidal shape. In addition to these bubbles, we can also see a number of wrinkles that connect one bubble to another throughout the map. The nano-IR image simultaneously taken with Figure 1, panel b is plotted in Figure 1, panel c where one can see many bright hotspots within the sample area.
area as Figure 1, panel a. The excitation laser frequency was set to $\omega = 910$ cm$^{-1}$ corresponding to a photon energy of 113 meV. Here multiple hotspots are clearly located inside these bubbles. We first discuss the bubble A at the center of the image, where three hotspots are distributed close to the corners of the pyramid. In Figure 1, panel e, we plot the line profiles for both topography (black curve) and IR amplitude (red curve) across one of the hotspots along the dashed line (1) in Figure 1, panel d. From Figure 1, panel f, one can see that the IR amplitude nearly doubles at the center of the hotspot despite a minute topographic variation of less than 10 nm. Moreover, the IR amplitude peak of the hotspot has a full-width at half-maximum (FWHM) of $\sim 50$ nm, which indicates that the hotspot is highly localized in real space compared to the excitation IR wavelength ($\lambda_{\text{IR}} \approx 11$ μm). Consequently, the mode volume ($\sim W^3$) is more than one-million-times smaller than what could be achieved by free-space IR photons ($\sim \lambda_{\text{IR}}^3$). In Figure 1, panel e, we plot the nano-IR image taken at $\omega = 950$ cm$^{-1}$ in the same sample area as Figure 1, panel d. Compared to Figure 1, panel d, we find that the three hotspots of bubble A in Figure 1, panel e move toward the corners of the bubble. As a result, the separations between hotspots increase at $\omega = 950$ cm$^{-1}$. This can be seen more clearly in Figure 1, panel g, where we plot the extracted line profiles along dashed lines (2) that connect two hotspots (Figure 1d,e). The observed frequency dependence is consistent with the plasmonic origin of these hotspots (Figure S6).

In addition to the nanopore A, hotspots are also seen inside the nanobubble B (Figure 1d,e). There are also three hotspots in B, but they are much closer toward the center, barely distinguishable from each other. We provide additional nanoimaging data of various nanobubbles in Figure 2. In general, for large pyramid-shaped bubbles, the hotspots normally reside at the corners, while for small bubbles they tend to appear at the center. In all cases, the hotspots are highly confined in space with strong field enhancement. Note that hotspots were also seen frequently inside wrinkles with sizable widths and heights (Figure 1d,e and Figure 2). For example, the wrinkle at location C in Figure 1, panel a with a width of $\sim 100$ nm and a height of $\sim 6$ nm has a hotspot at the center.

We assert that the observed hotspots are formed due to the localization of tip-excited graphene plasmons inside the graphene nanobubbles or wide wrinkles. We are able to reproduce the hotspot patterns as well as their dependences with the bubble size or the plasmon wavelength through real-space simulations assuming mode localization effects (Figure S6). Plasmonic hotspots were also observed previously in patterned graphene nanostructures, where localization of the plasmonic energy is due to the termination of graphene at the edges. In the case of graphene nanobubbles, the plasmon localization occurs due to the impedance mismatch at the boundaries between graphene nanobubbles and flat graphene region (blue dashed curves in Figures 1 and 2).

We now elaborate on two possible causes of impedance mismatch that are responsible for the plasmon localization and hence the hotspot formation inside the nanobubbles. One is the variation of topography, and the other is the change of the local plasmon wavevector ($q_p$). Under the Drude approximation, $q_p$ can be written as

$$q_p = \frac{\hbar \kappa(\omega)}{2e^2 \sqrt{\pi n F}} \sqrt{\omega(\omega + i\Gamma)}$$

where $\kappa(\omega)$, $\nu_F$, $n$, and $\Gamma$ are the effective dielectric function of the environment, Fermi velocity, density, and charge scattering rate of graphene, respectively. On the basis of eq 1, it can be seen that $\kappa(\omega)$, $\nu_F$, $n$, and $\Gamma$ are the main parameters
that control $q$. It is well-known that hBN is an ideal substrate that contributes little doping or scattering to graphene. Therefore, it is safe to rule out the possibility that $n$ and $\Gamma$ on top of the graphene bubbles are consistently and largely different from the flat graphene/hBN region. The Fermi velocity ($v_F$) scales with graphene strain. The strain effects of graphene nanobubbles on hBN have been studied in an earlier work, where it was found that thermal annealing induces isotropic compressive strain in the flat graphene/hBN region and the strain is released in the bubble region. The induced compressive strain, which scales with the annealing temperature, is estimated to be about 0.06% in our samples. Such a tiny strain will have negligible effects on $v_F$ and hence the plasmonic properties of graphene.

To evaluate the effects of topography and dielectric environment on graphene plasmons, we performed two control simulations with the finite difference time domain (FDTD) method. In one case, we study suspended graphene with a nanobubble at the center (Figure S5) to evaluate the effects of topographic curvature on graphene plasmons. For simplicity, we consider a Gaussian-shaped bubble in our simulation. The radius of curvature ($R_{\text{bubble}} \approx 600 \text{ nm}$) of the modeled bubble is comparable to that of the bubbles in our experiment (Figure 1). As shown in Figure S5, graphene plasmons can propagate almost freely off the bubble with negligible energy loss despite the curve. Such an observation is attributed to the high spatial confinement of graphene plasmons that leads to little radiation loss on curved graphene. Indeed, the plasmon wavelength ($\lambda_p$) is in the range of 100–200 nm (see discussions below) that is far smaller than $R_{\text{bubble}}$. Therefore, we are safe to conclude that the minute topographic variations of our nanobubbles are not responsible for the hotspots formation inside the bubbles.

We also investigate flat graphene on an hBN substrate where a cylindrical vacancy sits at the center of the substrate (Model 1, Figure 3a). More detailed sketches about the geometric shape, orientation, and aspect ratio of the vacancy are shown in Figure S2, where we plot the 3D side view (Figure S2a), the $x$–$y$ plane top view (Figure S2b), and the $x$–$z$ plane cross-sectional view (Figure S2c) of the hBN substrate together with the cylindrical vacancy. With this simplified model, we are able to simulate the dielectric environment of the nanobubbles in our experiments without introducing topographic variations. As an example, we consider a homogeneous distribution of a dielectric medium with a permittivity of $\varepsilon_{\text{vac}} = 1.0$ to 2.4 (see Supporting Information for discussions about the filling materials inside the bubble). Inset plots the modeled heterostructure for the calculation. Calculated frequency ($\omega$)–momentum ($q$) dispersion diagrams for all polariton modes in the heterostructure system considering $\varepsilon_{\text{vac}} = 1$. The white arrows in panels b–d mark the two reststrahlen bands 746–819 cm$^{-1}$ and 1370–1610 cm$^{-1}$ of hBN where graphene plasmons couple strongly with hyperbolic phonon polaritons in hBN. The red marks the plasmon dispersion without consideration of phonon polaritons in hBN. The horizontal and vertical dashed lines in panels b–d mark the excitation frequency ($\omega = 910 \text{ cm}^{-1}$) and corresponding plasmon wave vectors of graphene.

Figure 4. Theoretical calculations of the local plasmon wavelength ($\lambda_p$) and dispersion diagrams of a modeled heterostructure. (a) Calculated $\lambda_p$ versus gap size $t_{\text{gap}}$. We consider a homogeneous gap medium with a dielectric constant of $\varepsilon_{\text{vac}} = 1.0$ to 2.4 (see Supporting Information) for discussions about the filling materials inside the bubble. Inset plots the modeled heterostructure for the calculation. (b–d) Calculated frequency ($\omega$)–momentum ($q$) dispersion diagrams for all polariton modes in the heterostructure system considering $\varepsilon_{\text{vac}} = 1$. The white arrows in panels b–d mark the reststrahlen bands 746–819 cm$^{-1}$ and 1370–1610 cm$^{-1}$ of hBN where graphene plasmons couple strongly with hyperbolic phonon polaritons in hBN. The red marks the plasmon dispersion without consideration of phonon polaritons in hBN. The horizontal and vertical dashed lines in panels b–d mark the excitation frequency ($\omega = 910 \text{ cm}^{-1}$) and corresponding plasmon wave vectors of graphene.
The FDTD simulations discussed above suggest that the variation of dielectric environment is a likely cause for the impedance mismatch between nanobubbles and flat graphene on hBN. The physical origin behind the impedance mismatch, as detailed below, is a significant increase of the local plasmon wavelength ($\lambda_p$) on top the graphene nanobubbles. To determine $\lambda_p$, we first calculate the plasmon dispersion considering the entire graphene/gap/substrate system (inset of Figure 4a). The added gap region here is to simulate the dielectric medium inside the graphene nanobubbles. Details of the calculations are introduced in the Supporting Information. The gap layer here between graphene and hBN is filled with a uniform medium with a dielectric constant of $\varepsilon_m$. Here in the dispersion color plots (Figure 4b-d), $\varepsilon_m$ is set to be 1, and the thickness of the gap layer ($t_{gap}$) varies from 0 to 20 nm. The bright curves in the dispersion color plots represent various surface modes of our system. One can see that the intrinsic graphene plasmon mode (with $q^{-2}$ scaling, marked with red curves) coupled strongly with the hBN hyperbolic phonon waveguide modes in the two reststrahlen bands 746–819 cm$^{-1}$ and 1370–1610 cm$^{-1}$ (within the ranges defined by the white arrows).36,35,39 To avoid the strong coupling regions, we chose excitation frequencies away from the two reststrahlen bands. For example, the horizontal dashed lines in Figure 4, panels b–d mark the frequency of $\omega = 910$ cm$^{-1}$, where the response is predominantly plasmonic. The cross-points between the horizontal dashed lines and the graphene plasmon mode determine the plasmon wavevector $q_p$ (vertical dashed line); hence, the plasmon wavelength $\lambda_p \approx 2\pi/q_p$ at this frequency. With this method, we are able to plot the complete $t_{gap}$-$\lambda_p$ dependence curve at $\omega = 910$ cm$^{-1}$ assuming $\varepsilon_m = 1$ (Figure 4a, black curve). In Figure 4, panel a, we also plot the calculated dependence curves considering $\varepsilon_m = 1.3, 1.8$, and 2.4 corresponding to other possible filling materials (Supporting Information). In all cases, the plasmon wavelength increases with $t_{gap}$ which is due to the relatively lower values of $\varepsilon_m$ compared to the effective dielectric constant of hBN ($\varepsilon_{eff} \approx 4.0$) at $\omega = 910$ cm$^{-1}$ (Supporting Information). The increase of the local plasmon wavelength can also be seen by comparing FDTD simulations of Model 1 and Model 2 (see Figure 3c,d) and their zoomed-in snapshots in Figure S4b,d). These simulated field maps reveal a larger mode period of graphene plasmons above the vacancy region compared to that on the flat hBN substrate, which is a direct evidence of the strong sensitivity of graphene plasmons to its dielectric environment.

Our nano-IR imaging experiments establish graphene nanobubbles on hBN substrates as an effective plasmonic cavity for trapping plasmons with subhundred-nanometer spatial confinement. These highly confined hotspots are ideal for plasmon-enhanced IR spectroscopy of nanoscale objects. Of particular interest are biomolecules whose IR vibration modes could be significantly enhanced by electrostatically tunable graphene plasmons.60,61 Moreover, graphene nanobubbles are perfect for encapsulating biomolecules in an aqueous environment.62–64 Therefore, one can in principle monitor structural changes of molecules upon growth inside these nanobubbles. Our study sheds light on future applications of graphene nanobubbles for plasmon-enhanced IR spectroscopy of biomolecules.
Supporting Information of
“Ultra-confined plasmonic hotspots inside graphene nanobubbles”

Z. Fei¹,²,³*, J. J. Foley IV²,⁴, W. Gannett⁵,⁶, M. K. Liu¹,⁷, S. Dai¹, G. X. Ni¹, A. Zettl⁵,⁶, M. M. Fogler¹, G. P. Wiederrecht², S. K. Gray², D. N. Basov¹,⁸

¹Department of Physics, University of California, San Diego, La Jolla, 92093, USA
²Center for Nanoscale Materials, Argonne National Laboratory, Argonne, Illinois 60439, USA
³Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA
⁴Department of Chemistry, William Paterson University, Wayne, New Jersey, 07470, USA
⁵Department of Physics, University of California at Berkeley, Berkeley, California 94720, USA
⁶Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA
⁷Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York, 11794.
⁸Department of Physics, Columbia University, New York, NY 10027, USA.

*E-mail: (Z.F.) zfei@iastate.edu

List of contents

1. Nano-infrared imaging experiments
2. Sample preparation and characterization
3. Graphene nanobubbles
4. Calculation of the local plasmon wavelength
5. FDTD simulation
6. Real-space simulation of hotspot patterns in nanobubbles
1. Nano-infrared imaging experiments

The nano-infrared (IR) imaging experiments introduced in the main text were performed at UCSD by using a scattering-type scanning near-field optical microscope (s-SNOM). Our s-SNOM is a commercial system (neaspec.com) equipped with continuous wave mid-IR quantum cascade lasers (daylightsolutions.com). The unit for the IR frequency used in the work is wavenumber (cm\(^{-1}\)). The typical laser frequency used in the work is \(\omega = 910 \text{ cm}^{-1}\), corresponding to a photon energy of 113 meV. The s-SNOM is based on an atomic force microscope (AFM) operating in the tapping mode with a tapping frequency of \(\approx 270 \text{ kHz}\) and tapping amplitude of \(\approx 50 \text{ nm}\). A pseudo-heterodyne interferometric detection module is implemented in our s-SNOM to extract both the scattering amplitude \(s\) and phase \(\varphi\) of the near-field signal. In the current work, we discuss mainly the amplitude part of the signal that is sufficient to describe the plasmon hotspots. In order to subtract the background signal, we demodulated the near-field signal at the \(n\)th harmonics of the tapping frequency \((n = 3\) in the current work). In all the displayed near-field images, we plotted the near-field amplitude normalized to that taken off the nanobubbles on the flat graphene/hBN region. Our nano-IR imaging experiments were performed at ambient conditions.

2. Sample preparation and characterization

Our graphene films were fabricated using a two-step low pressure chemical vapor deposition (CVD) process.\(^1\) We then transferred graphene films to exfoliated hexagonal boron nitride (hBN) crystals on standard SiO\(_2\)/Si substrates using a sacrificial polymethyl methacrylate (PMMA) layer.\(^2,3\) After transferring, we cleaned the PMMA residuals by acetone following by a thermal annealing at 340 °C with H\(_2\)/Ar. The thermal annealing could potentially induce a tiny compressive strain of about 0.06% on the flat graphene on hBN region away from the nanobubbles.\(^4\) Such a small strain has negligible effects on both the electronic and plasmonic properties of graphene.\(^5,6\) Optical microscopy and Raman spectroscopy were used to determine the graphene thickness and to locate graphene/hBN (GBN) and graphene/SiO\(_2\) (GOS) regions (Figure S1). All our data images about graphene nanobubbles were taken in the graphene/hBN regions.

**Figure S1.** (a) Typical optical image of CVD graphene transferred to exfoliated hBN microcrystals on SiO\(_2\)/Si wafers. Here the graphene on hBN and graphene on SiO\(_2\) regions are marked with ‘GBN’ and ‘GOS’ respectively. (b) Raman spectra of graphene on hBN and graphene on SiO\(_2\).

3. Graphene nanobubbles

The nanobubbles are formed in the graphene on hBN regions after the transferring and annealing processes. The formation of these nanobubbles is believed to be related to the negative thermal expansion coefficient of graphene\(^7\) as well as residual materials at the graphene/hBN interface.\(^5,9\) These residual materials will aggregate together as the filling materials inside the
bubbles. They are originated from interfacing materials of graphene during the transfer process, for example air, water, acetone and PMMA. The follow-up thermal annealing could cause thermal degradation of PMMA into monomer MMA that is normally in the liquid form at ambient temperature. Indeed, previous works about similar nanobubbles have seen evidence of hydrocarbons inside the bubbles. It was also observed that these residual materials are mobile and easy to escape or evaporate when nanobubbles were broken. Moreover, our real-space simulations in Figure S6 suggest that a homogeneous distribution of the filling materials is necessary to reproduce the hotspots patterns observed in our experiments. In our analysis, we consider all possible forms of residual materials. For air or hydrocarbon gas, the permittivity is \( \varepsilon_m \approx 1.0 \). For water, acetone, and solid PMMA, \( \varepsilon_m \) at \( \omega \sim 910 \text{ cm}^{-1} \) is about 1.3, 1.8 and 2.4 at \( \omega \sim 910 \text{ cm}^{-1} \), respectively. For other hydrocarbon liquids, though optical constants vary depending on the exact chemical composition, \( \varepsilon_m \) at \( \omega \sim 910 \text{ cm}^{-1} \) is commonly far below 4.0 that is the effective permittivity of hBN (Figure S3). The relative smaller permittivity of the bubble medium compared to that of hBN is the major cause of the plasmonic impedance mismatch between graphene nanobubbles and the flat graphene regions.

4. **Graphene doping**

As shown in Eq. 1 in the main text, carrier density \( n \) is one essential parameter that determines the plasmon wavevector and hence the plasmon wavelength of graphene. Nevertheless, we believe the doping inhomogeneity \( \delta n \) of our graphene samples is a minor effect compared to variations of dielectric environment as discussed in detail in the main text. Compared to detached graphene in the bubble region, the flat support graphene is coupled closely with the substrate, which could possibly cause additional doping to graphene. This is certainly an issue if using SiO\(_2\) as the substrate that could introduce strong doping to graphene. Nevertheless, our graphene sample is sitting on hBN that is a much better substrate with negligible doping effect on graphene, as confirmed by the transport study and very recently a Raman study. In the latter Raman work, the authors investigated graphene on various substrates including hBN as well as suspended graphene over trenches, and their data suggest that the doping level of graphene on hBN is roughly the same as suspended graphene. Therefore, we believe that the carrier density of supported graphene on hBN will not be consistently and largely different from the detached graphene bubble region. Moreover, we wish to emphasize that our graphene samples are highly doped at ambient conditions with a carrier density of \( n \approx 3 \times 10^{12} \text{ cm}^{-2} \) according to our simulations, so small doping inhomogeneities \( \delta n \) will not affect much the plasmon wavelength \( \lambda_p \) of graphene since \( \lambda_p \sim (n + \delta n)^{1/2} \approx n^{1/2} (1 + 0.5 \delta n/n) \). For example, a doping inhomogeneity as big as \( \delta n = 0.5 \times 10^{12} \text{ cm}^{-2} \) only causes a change of 8% to the plasmon wavelength. This is certainly a small effect compared to the effect due to variation of dielectric environment (~70%) discussed in the main text.

5. **Calculation of the local plasmon wavelength**

In order to determine the local plasmon wavelength, we first calculate the plasmon dispersion by evaluating numerically the imaginary part of the reflection coefficient \( \text{Im}(r_p) \) for the entire graphene/gap/hBN/SiO\(_2\) heterostructure system by using the transfer matrix method (Figure 4 in the main text). Here the gap layer is added to simulate the dielectric medium inside the graphene nanobubble. Based on the calculated dispersion diagrams, we can determine directly the local plasmon wavevector \( q_p \) hence the local plasmon wavelength \( \lambda_p = 2\pi/q_p \) of graphene. With this method, we produced the gap size \( t_{\text{gap}} \) dependent plasmon wavelengths for graphene as shown in Figure 4a in the main text. We set the hBN thickness to be 150 nm according to our AFM
measurement, but the calculation results do not show noticeable changes for graphene plasmons when the thickness of hBN is above 100 nm.

6. FDTD simulation

We performed rigorous electrodynamics simulations to further investigate the impact of bubble topography and dielectric environment on the plasmon propagation and subsequent hot-spot formation. A commercial-grade simulator based on the finite-difference time-domain (FDTD) method was used to perform the calculations. We considered four different models in our simulations to test the effects of pure dielectric environment (with Model 1 & 2) and pure topography (with Model 3 & 4) on graphene plasmons:

(1) Graphene on an hBN substrate with a cylindrical vacancy to simulate the pure dielectric environment of graphene nanobubbles without introducing topographic features. Detailed sketches about the cylindrical vacancy are plotted in Figure S2. The depth and the diameter of the cylindrical vacancy are set to be 10 nm and 250 nm to match the aspect ratio of the nanobubble ‘A’ in Figure 1 of the main text. Note that the average height of the nanobubble ‘A’ is roughly 10 nm.

(2) Graphene on a flat and intact hBN substrate.

(3) Suspended graphene with a Gaussian-shaped bubble at the center to simulate the pure topographic effects of our graphene bubbles without introducing variations of dielectric environment. The height and the full width at half maximum of the Gaussian-shaped graphene bubble are set to be 20 nm and 125 nm to match the aspect ratio of the nanobubble ‘A’ in Figure 1 of the main text.

(4) Suspended and flat graphene.

Figure S2. Detailed sketches about Model 1 for FDTD simulation. These sketches include the 3D side view (a), x-y plane top view (b), and x-z plane cross-sectional view (c) about the model. Note that graphene (not shown in these sketches) will be placed right above the hBN surface and vacancy.

In our modeling, the Fermi energy of graphene ($E_F$) is set to be $\sim 0.2$ eV corresponding to a 2D optical conductivity $\sigma_{2D} \approx (0.2 + 2i)G_0$ at $\omega = 910$ cm$^{-1}$, where $G_0 = \pi e^2/4h \approx 6.07 \times 10^{-5}$ $\Omega^{-1}$ is the universal optical conductivity of graphene. The effective thickness of graphene is set to be 5 nm, which is two orders of magnitude smaller than the plasmon wavelength. We employ a multi-
resolution grid with a grid spacing of 1 nm in the graphene layer that gradually increases to 10 nm away from graphene. Excitation of graphene plasmons is achieved using a point dipole emitter source with IR frequency around 910 cm\(^{-1}\). The dipole emitter is placed 2 nm above the graphene plane and is polarized along the z-axis (normal to the graphene film). For hBN, we modeled it as an isotropic medium with an effective dielectric constant of \(\varepsilon_{\text{eff}} \approx 4.0\) at \(\omega = 910\) cm\(^{-1}\). This is validated by the calculated dispersion diagrams in Figure S3 with both effective and realistic dielectric constants (\(\varepsilon_{\text{ab}} \approx 8.1+0.01i\) and \(\varepsilon_c \approx 1.7+0.02i\)).\(^{18}\) Here we found that the plasmon wavevector \(q_p\) at \(\omega = 910\) cm\(^{-1}\) is roughly the same for both dispersion diagrams.

**Figure S3.** Calculated frequency (\(\omega\)) – momentum (\(q\)) dispersion diagrams of graphene on hBN considering realistic (a) and effective (b) dielectric constants of hBN. Here the realistic dielectric constants of hBN is adopted from previous literature\(^{18}\) and the effective dielectric constant is set to be 4.0 at \(\omega = 910\) cm\(^{-1}\). The arrows in panel a mark the two reststrahlen bands 746-819 cm\(^{-1}\) and 1370-1610 cm\(^{-1}\) of hBN where graphene plasmons couple strongly with hyperbolic waveguide polaritons in hBN. The horizontal and vertical dashed lines mark the excitation frequency (\(\omega = 910\) cm\(^{-1}\)) and corresponding plasmon wavevectors of graphene.

**Figure S4.** Zoomed-in view of the FDTD simulations (Model 1 and Model 2) shown in Figure 3 of the main text. The zoomed-in regions are marked with dashed squares in panels a and c. The plasmon field maps in these zoomed-in regions are plotted in panels b and d.
The FDTD simulation results of Model 1 and Model 2 are shown in Figure 3 in the main text. In Figure S4, we plot the zoomed-in snapshots of the plasmon field ($E_z$) maps close to the emitter. As shown in these field maps, the plasmon wavelength ($\lambda_p$) in the zoomed-in regions of Model 1 (graphene/air/hBN) is about 70% larger than that in Model 2 (graphene/hBN), consistent with the dispersion calculations in Figure 4 of the main text. As discussed in the main text, the increase of the plasmon wavelength above the bubble/vacancy region is the origin of the plasmon confinement inside the bubbles.

The FDTD simulation results of Model 3 and Model 4 are shown in Figure S5, where we plot the plasmon field ($E_z$) map of suspended graphene with a Gaussian-shaped graphene bubble at the center (Figure S5c) and that of the completely-flat suspended graphene (Figure S5d), respectively. Similar to Figure 3 in the main text, we plot the zoomed-in views of the region far away from the emitter (marked with rectangles in Figure S5c,d) in Figure S5e,f and the corresponding $E_z$ profiles right above graphene in Figure S5g. In this zoomed-in region, plasmons are the dominant source of field. By comparing Model 3 and Model 4 through these field maps and profiles, one can see that plasmon field of the two models are generally the same indicating that graphene plasmons can propagate freely off the bubble with negligible energy loss despite the curvature. Similar results were also obtained in previous simulation work, where the authors conclude that the high spatial confinement of graphene plasmons is responsible for the negligible radiation loss on slightly-curved graphene. Indeed, the plasmon wavelength ($\lambda_p \sim 100-200$ nm) is far smaller than the radius of curvature of the bubble ($\sim 600$ nm). The simulation results in Figure S5 prove that pure topographic variations of our nanobubbles are not responsible for the plasmon confinement and hence the hotspots formation inside the bubbles.

**Figure S5.** The FDTD simulations of surface plasmon polaritons launched by a point emitter above suspended graphene. (a) Model 3: graphene with a Gaussian-shaped bubble (height = 20 nm; full width at half maximum = 125 nm) at the center right beneath the emitter. (b) Model 4: flat suspended graphene. (c,d) Simulated $E_z$ maps of model 3 and model 4, respectively. (e,f) Zoomed-in $E_z$ maps of the regions defined by rectangles in panels c and d, respectively. (g) Horizontal line
profiles of $E_z$ taken right above the surface of graphene in panels e and f. Scale bars represent 100 nm in panels c and d, and represent 50 nm in panels e and f.

7. **Real-space simulation of hotspot patterns in nanobubbles**

In order to reproduce the real-space patterns of hotspots distribution, we construct a simplified triangular cavity (Figure S6a). Considering that the nanobubbles that we investigated are relatively flat (height $\ll$ side length), the triangular cavity is a good approximation of the pyramid-shaped bubbles (Figure 1 in the main text). In Figure S6a, we plot the simulated field distribution map $|E_\alpha(x, y)|$ that resembles what we measure in our nano-IR experiments. Here $(x, y)$ are the coordinates of the plasmon source that is scanning over the cavity and $E_\alpha(x, y)$ is the total plasmon field at a given position $(x, y)$. The purpose of the simulation is for qualitatively understanding the hotspots pattern formation and their dependence with both the plasmon wavelength and bubble size. The tip-launched plasmon field distribution takes the standard circular (cylinder) waveform obtained by solving the Helmholtz equation. We considered mode localization effects by assuming plasmon reflections off the bubble edges. Similar method has been used in an earlier work, where fringes patterns of phonon polaritons of hBN were simulated.

The simulation results are given in Figure S6b-f, where the varying parameter is the ratio between the length of the cavity edge ($L$) and the graphene plasmon wavelength ($\lambda_p$). One can see that, as $L/\lambda_p$ varies from 2.78 to 0.55 (Figure S6b-f), the $E_{\alpha}(x, y)$ maps show systematic evolutions. The dominant features of these simulations are the bright hotspots at the corners of the triangular cavity. As $L$ decreases or $\lambda_p$ increases (e.g. when excitation frequency decreases), the hotspots merge towards the center of the cavity and eventually merge into one single hotspot at the center (Figure S6f). All these features agree well with our imaging data of graphene nanobubbles with different bubble sizes or excitation frequencies (Figure 1 and Figure 2 in the main text).

![Figure S6](image)

**Figure S6.** Real-space simulations of plasmonic hotspots inside a triangular cavity. (a) Illustration of the triangle cavity for the purpose of real-space simulation. The length of the side is $L$. (b-f) Real-space simulations about $|E_z|$ field amplitude of an equilateral triangle cavity with the ratio $L/\lambda_p$ from 2.78 to 0.55.

**References for the Supporting Information**


