Efficient All-Optical Plasmonic Modulators with Atomically Thin Van Der Waals Heterostructures

Xiangdong Guo, Ruina Liu, Debo Hu, Hai Hu, Zheng Wei, Rui Wang, Yunyun Dai, Yang Cheng, Ke Chen, Kaihui Liu, Guangyu Zhang, Xing Zhu, Zhipei Sun,* Xiaoxia Yang,* and Qing Dai*

DOI: 10.1002/adma.201907105

All-optical modulators are attracting significant attention due to their intrinsic perspective on high-speed, low-loss, and broadband performance, which are promising to replace their electrical counterparts for future information communication technology. However, high-power consumption and large footprint remain obstacles for the prevailing nonlinear optical methods due to the weak photon–photon interaction. Here, efficient all-optical mid-infrared plasmonic waveguide and free-space modulators in atomically thin graphene-MoS₂ heterostructures based on the ultrafast and efficient doping of graphene with the photogenerated carrier in the monolayer MoS₂ are reported. Plasmonic modulation of 44 cm⁻¹ is demonstrated by an LED with light intensity down to 0.15 mW cm⁻², which is four orders of magnitude smaller than the prevailing graphene nonlinear all-optical modulators (∼10³ mW cm⁻²). The ultrafast carrier transfer and recombination time of photogenerated carriers in the heterostructure may achieve ultrafast modulation of the graphene plasmon. The demonstration of the efficient all-optical mid-infrared plasmonic modulators, with chip-scale integrability and deep-sub wavelength light field confinement derived from the van der Waals heterostructures, may be an important step toward on-chip all-optical devices.

The increasingly widespread need to process and transmit massive amounts of data has motivated an enormous demand for high-speed and small-sized devices.[1,2] Since the increasing power consumption and heat generation in integrated electronic circuits are hampering further improvements in the speed and capacity of the information communication technology, it is natural to consider its optical alternative.[1,3] Indeed, all-optical logical circuitry capable of processing and transmitting data is believed to be superior to its electronic counterpart because of its much higher operating speed.[4–7] For this reason, many research efforts such as nonlinear optical dielectrics with optical Kerr effects,[8] photoinduced phase change,[4,9,10] and photogenerated carrier injection in semiconductors[11] have been devoted to the development of all-optical processing devices. However, due to the diffraction limit and the weak photon–photon interaction, all-optical devices are always exhausted by the large footprint and high-power consumption.[5,12]

Graphene plasmons with ultrahigh light confinement and wide tunability provide a promising platform to realize ultracompact all-optical devices.[13–18] For example, the ultrafast switching of nonequilibrium graphene plasmons has been demonstrated, which are originated from the generation and annihilation of hot electrons by controlling light-induced heating.[19] Unfortunately, the hot-electron plasmon...
approach is inefficient by the linear dispersion of graphene. Tuning of graphene plasmons has also been demonstrated with controllable desorption of oxygen by UV illuminations.\[30\] However, the physical adsorption modulation method as well as the phase change approach and electrical methods are typically slow.\[4,21\]

Constructing van der Waals (vdW) heterostructures of graphene and other 2D semiconductors provides a promising approach for achieving devices with new functions.\[22–25\] It has already been demonstrated that ultrafast and efficient charge transfer occurs in the atomically thin vdW heterostructures of graphene/TMD.\[26\] Compared to traditional bulk semiconductors based heterostructures, 2D heterostructures can significantly enhance the transfer of photogenerated carriers with high transfer speed and efficiency.\[27\] For example, in the monolayer graphene/WS\(_2\) vdW heterostructure, the photogenerated carrier transfer time is about 1.4 ps with the transfer efficiency of nearly 100\%.\[26\] Hence, vdW heterostructure can provide a new avenue for efficient and ultrafast optical modulation.

Here, we demonstrate highly efficient and ultrafast optical plasmonic waveguide and free-space modulators in vdW heterostructures of graphene and monolayer MoS\(_2\). The photogenerated electrons in MoS\(_2\) can dope graphene within around 7 ps, which may achieve ultrafast modulation of the graphene plasmon. In the waveguide modulators, flexible modulating of graphene plasmon propagation can be controlled by visible light, as measured by scattering-type scanning near-field optical microscope (s-SNOM). We also demonstrated the modulation efficiency can be quantitatively controlled by the wavelength and power density of illumination, for example, a 3.6 dB per atomic-layer relative modulation depth can be achieved by using a 440 nm laser with intensity of 180 mW cm\(^{-2}\). Our findings enable a versatile way for efficient control of graphene mid-infrared plasmons by using ordinary LEDs, which can be used for on-chip integration of all-optical light modulation and switching.

As shown schematically in Figure 1a, s-SNOM is used to detect the graphene plasmon controlled by turning on/off of visible light illumination (the orange beam). Here, graphene was p-type doped by the surface dangling bonds of the substrate\[28,29\] and more holes were introduced by doping with small amount of NO\(_2\) gas for the s-SNOM measurement.\[30,31\]

![Figure 1](https://www.advancedsciencenews.com/)

Figure 1. Optical modulation of graphene plasmons in graphene/MoS\(_2\) heterostructure. a) Schematic of the graphene plasmon waveguide modulator. Propagating graphene plasmons (concentric red ripples) are excited by focusing the infrared laser beam (left, red) on the s-SNOM tip. Meanwhile, a visible laser beam (right, orange) is focused on the sample. In the area illuminated by the visible laser, graphene plasmons switch off; however, switching off the visible light causes the graphene plasmons to recover. Here, 0 and 1 indicate the off and on states of the visible illumination, respectively. b) Up: photogenerated carriers are excited in the MoS\(_2\) layer under the visible light illumination. The electrons transfer into the p-type doped graphene, while the holes remain in the MoS\(_2\). Bottom: the photogenerated carriers recombine after the visible light is turned off. c) Schematic of ultrafast band-edge charge collection in the hole-doped graphene/MoS\(_2\) heterostructure under laser irradiation. d) PL spectra of monolayer MoS\(_2\) and the graphene/MoS\(_2\) heterostructure. e) Differential reflection (black circles) measured from the graphene/MoS\(_2\) heterostructure. The solid red line represents the corresponding fitting analysis. Here, the lasers employ a 600 nm pump and a 780 nm probe pulse.
The IR laser beam (red beam) incident on the metal-coated s-SNOM tip generates strong near-field confinement at the tip apex which results high momenta to excite graphene plasmons. The plasmons propagate along the graphene and reflect at the edge, which interfere with the incident light and form interference patterns as the tip scanning over the sample surface.\textsuperscript{[32,33]} The graphene plasmon in the heterostructure can be turned off and on by switching on and off the visible light (right part in Figure 1a). In contrast, the graphene plasmon has no response to the visible light in the individual graphene layer (left part in Figure 1a). Since the graphene plasmons are in the mid-infrared frequency range, the s-SNOM measurements were performed with mid-infrared laser to probe the graphene plasmon. The physical mechanism of the mid-infrared graphene plasmon modulation with visible light will be analyzed in Figure 1b–e by the PL spectroscopy and pump-probe spectroscopy.

The optical modulation of graphene plasmon in the heterostructure originates from the doping of graphene by the photogenerated electrons in MoS\textsubscript{2}. As shown in Figure 1b, the electrons in MoS\textsubscript{2} monolayer can absorb the visible light and jump from the valence band to the conduction band, generating electron-hole pairs. Photoelectrons in the MoS\textsubscript{2} then transfer to the p-type doped graphene film due its lower available energy levels for electrons, leaving the remaining holes trapped in the MoS\textsubscript{2} film (Figure 1b,c). The transferred photoelectrons recombine with the holes in the p-type doped graphene and thus erase the graphene plasmons, i.e., the resonant oscillation of conducting holes. When visible illumination is removed, the photogenerated carriers recombine and the Fermi level of graphene recover to the initial state, which revives the plasmon signal (bottom panel, Figure 1b).

To test the operation principle, monolayer MoS\textsubscript{2} was grown directly onto a Si/SiO\textsubscript{2} substrate by chemical vapor deposition (CVD).\textsuperscript{[34,35]} Then exfoliated graphene was released on the MoS\textsubscript{2} layer by dry-transfer technique to form the graphene/MoS\textsubscript{2} heterostructure (details in the Experimental Section).\textsuperscript{[16]} Raman spectroscopy confirmed the high quality of monolayer graphene and MoS\textsubscript{2} (Figure S1, Supporting Information). The photoluminescence (PL) spectrum of the MoS\textsubscript{2} film exhibits a prominent peak around 670 nm in Figure 1d (black line), corresponding to the direct interband recombination of the photogenerated electron-hole pairs in the MoS\textsubscript{2}. In contrast, the PL of the MoS\textsubscript{2} is quenched to a magnitude nearly half that of the graphene/MoS\textsubscript{2} heterostructure, a direct reflection of the transfer of photogenerated electrons into the graphene and correspondingly reduced electron-hole recombination in the MoS\textsubscript{2}.

The transient carrier transfer process of carriers in the graphene/MoS\textsubscript{2} heterostructure was studied via the two-color pump-probe technique.\textsuperscript{[26,27]} We adapted a 600 nm pump laser (=100 fs width, 1.49 µJ µm\textsuperscript{-2} fluence) to excite MoS\textsubscript{2}, and a 780 nm probe laser (=100 fs width, 0.52 µJ µm\textsuperscript{-2} fluence) to avoid recording the optical response of the MoS\textsubscript{2} and substrate (Figure S2a,b, Supporting Information). Under this set-up, the photogenerated carriers can be attributed to the excitation of MoS\textsubscript{2} and the detection signal is emitted only from the graphene, which can reveal the charge transfer at the heterostructure interface,\textsuperscript{[35,37]} as indicated in Figure 1c.

The transient absorption signals of the graphene/MoS\textsubscript{2} heterostructure and pure graphene are shown in Figure 1e and Figure S2c (Supporting Information), respectively. The rising-up part of the transient absorption spectrum corresponds to the excitation, electron-electron interaction and transfer of photogenerated electrons. By deconvolution of the spectra with the laser cross correlation function (Figure S2d, Supporting Information), the calculated rising time of the heterostructure and individual graphene layer are about 1000 fs and 360 fs, respectively. The increase of the rising time in the heterostructure mainly originates from the charge transfer process of photogenerated electron from MoS\textsubscript{2} to the graphene layer, which is estimated to be in the level of 600 fs here. Theoretically, the charge transfer time could be reduced by an order of magnitude if conducted at a perfect interface. Although the dry-transfer method employed fabricates a clean interface, small amounts of impurities and air bubbles are inevitable.

Since the recombination of photogenerated electrons and holes in the heterostructure is much slower than the transfer and intraband equilibration, the former determines the modulation speed of the graphene plasmon modulators. The recombination time can be estimated by exponentially fitting the falling-down part of the transient absorption spectra. As shown in Figure 1e and Figure S2c (Supporting Information), the carrier recombination times of pure graphene and the heterostructure are ≈1.4 and 6.3 ps, respectively. Hence, assuming that the excitation time of plasmon is shorter than 50 fs,\textsuperscript{[19]} an ultrafast switching-on of graphene plasmon of ≈7 ps (i.e., <10 ps) could be expected by considering the transfer and recombination of photogenerated electrons. Nevertheless, if n-type doped graphene based heterostructure devices can be realized, where the switching off time of this optical modulation depends only on the photogenerated electrons transfer time, a response time around hundreds of femtoseconds can be expected.

Figure 2a schematically exhibits the decrease of absolute value of the graphene Fermi level in the heterostructure under the illumination of visible light, according to the charge transfer process illustrated above. Figure 2b,c shows two typical s-SNOM images in the heterostructure at a frequency of 926 cm\textsuperscript{-1}. The dashed lines indicate the natural edge of the graphene layer supported on monolayer MoS\textsubscript{2} with the region to the left of the edge corresponding to the graphene/MoS\textsubscript{2} heterostructure and the region to the right corresponding to pure MoS\textsubscript{2}. In the absence of visible light illumination, two interference fringes of graphene plasmons appear along the graphene edge in the heterostructure region (Figure 2b). In contrast, when we turned on a 633 nm laser illumination, the graphene plasmon fringes in the same region nearly disappeared, as shown in Figure 2c. More evidences can be found in Figure S3 (Supporting Information). When visible laser was switched on and off, obvious boundaries (e.g., by line profiles) between the regions with and without plasmon fringes are shown, corroborating the efficient optical switching of the graphene plasmon in the heterostructure. For comparison, the plasmon imaging on pure graphene samples (i.e., without MoS\textsubscript{2}) with the visible laser on and off was also conducted, which revealed no change to the very stable images (Figure S3d, Supporting Information). This is direct evidence that the visible laser has no effect on pristine graphene plasmons.

To quantify the optical response of the graphene plasmon in the heterostructure, we extract the data on the cutting lines of
the plasmon images in Figure 2b,c for analysis, as plotted in Figure 2d. For the case in which the visible laser is switched off, the graphene plasmon interference fringes parallel to the edges are clearly detectable as peaks in the shaded area. These fringes result from the interference of the local field below the tip with the tip-launched and edge reflected fundamental plasmon mode. The fringe separation consequently corresponds to half of the graphene plasmon wavelength, and \( \lambda/2 = 100 \text{ nm} \) can be extracted from the red line. Most intriguingly, when the visible laser is switched on, the plasmon fringe information almost completely disappears, as evidenced by the lack of peaks in the shaded area (blue line), verifying the efficient response of the graphene plasmons to the MoS\(_2\) photogenerated carriers. And when the visible laser is switched off, the plasmon can recover to the initial state (i.e., the plasmonic peak positions and strengths), as demonstrated in Figure S3g (Supporting Information).

We also constructed free-space plasmonic modulators with the concept (Figure 3a). Large-area graphene/MoS\(_2\) heterostructures were fabricated with high quality CVD grown graphene and monolayer MoS\(_2\) on IR transparent MgF\(_2\)/Si substrate (Figure S5a, Supporting Information).\(^{42,43}\) The heterostructures were patterned into electrically continuous

---

**Figure 2.** The all-optical graphene plasmon waveguide modulation with a thickness of only several atomic layers. a) Schematic diagram of band alignment and the physical mechanism of photocarrier transfer in the hole-doped graphene/MoS\(_2\) heterostructure under visible light irradiation. b,c) Near-field images of a graphene/MoS\(_2\) heterostructure on SiO\(_2\) substrate in the b) absence and c) presence of 633 nm laser irradiation (6 mW cm\(^{-2}\)). Dashed lines indicate the graphene edge. d) The plasmon signals extracted from the cut-lines (red and blue lines) in (b) and (c), respectively.

---

**Figure 3.** Graphene plasmon free-space modulator. a) A schematic of FTIR measurement of plasmons in graphene/MoS\(_2\) nanoribbon arrays. Graphene plasmons are excited by an incident infrared beam (the red shaded pillar) and can be measured by in situ background subtraction through the MgF\(_2\)/Si back gating. A visible light beam (orange shaded pillar) incident to the heterostructure is used to optically excite MoS\(_2\). b,c) Graphene plasmon extinction spectra (colored lines) of the graphene/MoS\(_2\) heterostructure under irradiation of varied wavelengths and densities. The laser power density in (b) is fixed at 80 mW cm\(^{-2}\) while the wavelength in (c) is fixed at 440 nm. The extinction spectrum in the absence of laser irradiation is also shown (gray line) at \( \Delta_{\text{CNP}} = -12 \text{ V} \). The circles represent the experimental data, which can be fitted with the Lorentz function (solid lines). d) Frequency shifts and modulation depth as a function of the optical power illuminating the graphene/MoS\(_2\) device.
nanoribbon arrays which were measured by a far-field Fourier transform infrared microscope (FTIR), as shown in Figure 3a. The measured plasmonic extinction spectra can be described as $1 - T_{\text{VR}}/T_{\text{CNP}}$, where $T_{\text{VR}}$ and $T_{\text{CNP}}$ represent the transmittance measured with an applied gate voltage $V_g$ and at the charge neutral points (CNPs) of the graphene nanoribbons, respectively. Collectively, these data allow us to quantitatively study the optical tunability of the plasmon from the aspects of plasmon frequency and strength in the heterostructures.

We systematically studied the plasmonic modulator under illumination with varied wavelengths. Three typical laser wavelengths, i.e., 440 nm, 520 nm and 650 nm with power density fixed at 80 mW cm$^{-2}$ were used and the measured graphene plasmon extinction spectra are presented in Figure 3b. The plasmon resonance frequencies are extracted to be 996 cm$^{-1}$ (dark), 981 cm$^{-1}$ (650 nm laser), 956 cm$^{-1}$ (520 nm laser), and 892 cm$^{-1}$ (440 nm laser), respectively by fitting the extinction spectra with Lorentz function (solid lines). Therefore, shorter wavelength laser irradiation prompts a greater redshift of the plasmon frequency and decreased plasmon intensity.

The electrical transfer curves were monitored while measuring the plasmons (Figure S5b, Supporting Information). The graphene CNP points exhibit a gradually decreasing trend from 7 V (dark) to 6.1 V (650 nm laser), 5.6 V (520 nm laser), and 5 V (440 nm laser). This indicates that graphene is heavily hole-doped throughout the experiment, and that the injection of photogenerated electrons decreases the number of holes and drives the Fermi level upward. Illumination of a shorter wavelength causes more photoinduced carriers to be injected, as is consistent with the change observed in the graphene plasmon extinction spectra. The physical mechanism proceeds as follows: Photons with higher energy will transit electrons to a higher conduction band level in MoS$_2$ to enable more electrons to transfer to graphene layer. Therefore, there is a greater decrease to the hole density in graphene, which makes the hole plasmons shift to red further and prompts a larger decrease in the resonance. Unfortunately, simple pursuit of lasers with shorter wavelength is not an effective strategy for improving efficiency because once wavelengths are decreased to the ultraviolet range, undesirable phenomena such as the ultraviolet light-induced O$_2$ doping of graphene occur.[20,44] To restrict the focus of this study to an optically tunable method, the shortest laser wavelength used herein was 440 nm.

The optical modulation of graphene plasmon can also be controlled by the laser intensity. Figure 3c displays the plasmonic extinction spectra under 440 nm laser illumination with varied intensity. As shown, the plasmon frequency and intensity obviously decrease with increasing laser power densities. This is because more photogenerated electrons are transferred from MoS$_2$ into the graphene layer, which can be confirmed from corresponding electrical transfer curves of the device (Figure S6, Supporting Information). To allow for quantitative analysis, the plasmon frequency shifts are calculated and plotted as a function of the power density (Figure 3d). The frequency shift and the relative modulation depth exhibit a sublinear dependence on power intensity, and under a laser intensity of 180 mW cm$^{-2}$ can reach values of approximately 140 cm$^{-1}$ and 56% (3.6 dB per atomic-layer), respectively. The curves can be fitted by the equation: $\Delta \omega_i = P \beta_i$, ($i = 1, 2$), where $P$ is the power of the irradiation light, $\Delta \omega_i$ and $\Delta \beta_i$ represent the wavenumber shifts and the modulation depth, respectively. $\beta_1$ and $\beta_2$ at around 0.28 and 0.32 can give the best fitting. This curve implies that the all-optical modulation effect gradually becomes saturated at high laser power. The phenomenon is similar to the photoresponsivity of MoS$_2$ on illumination intensity,[43] which could be explained in terms of trap states present either in MoS$_2$ or at the interface between the MoS$_2$ and the underlying MgF$_2$ layer or the graphene layer. This is aggravated by the high surface-to-volume ratio of the MoS$_2$. Under higher illumination intensities the density of available states is reduced, resulting in saturation of the photoresponsivity.

To further analyze the optically tunable plasmon in the graphene/MoS$_2$ heterostructure, we perform FEM simulation of the experimentally measured extinction spectra.[43,46] The graphene Fermi energy in the absence of laser illumination was measured to be 0.216 eV according to the absorption spectrum (Figure S7 and Note S3, Supporting Information),[47] which was used to simulate the experimental extinction spectrum (the gray line in Figure 3c). Exposure to the visible laser illumination only changes the Fermi energy. By simulating the colored lines in Figure 3c, we can extract the corresponding graphene Fermi energy, which ranges from 0.216 to 0.145 eV at the different tested laser intensities (details in Figure S8 and Note S4, Supporting Information). All the experimental and simulated plasmon frequencies are plotted as a function of Fermi energy with different laser intensities and wavelengths in Figure S9 (Supporting Information). These data satisfy the graphene plasmon dispersion relationship $\omega \propto |E_F|^{1/2}$, implying that the optically tunable plasmon phenomenon depends mainly on the carrier concentration and Fermi energy in graphene, rendering it very consistent with the traditional physical mechanism of graphene plasmons.

Because the relationship between plasmon modulation and the laser energy density implies that this plasmon optical modulation should be very efficient, we studied the modulation in response to an LED light. Figure 4a is the real-time drain-source current ($I_{ds}$) of the heterostructure device which was recorded simultaneously to the switching of the LED illumination. This weak LED light with a power density of 0.15 mW cm$^{-2}$ can cause about 15% decrease in the current intensity, and the CNP is reduced from 7 to 6 V (Figure S10a, Supporting Information), further corroborating the sensitive of the all-optical modulation in the heterostructure. However, since the smallest sampling interval of the electrical measurement is 50 ms which is much slower than the doping of photogenerated electrons, it cannot directly reflect the ultrafast process. This can be demonstrated that about 50% current change happens within the 50 ms as shown in Figure 4b. Most importantly, in electrical experiments, the current change is not only affected by visible light modulation, but also limited by other factors such as the persistent photoconductivity effect.[21] Figure 4c depicts the extinction spectra of the graphene/MoS$_2$ nanoribbon arrays with the LED illumination turning off (gray line), on (red line) and off (blue line) successively at a fixed gate voltage $V_g = -12$ V. When the LED is on, the plasmon frequency can be modulated by approximately 43 cm$^{-1}$. After switching off the LED, the plasmon can achieve nearly 100% recovery. The results of the simulations (Figure S10b, Supporting Information) are consistent with
the experiment. If we consider the modulating time as 7 ps (Figure 1e), the corresponding modulation energy consumption is on the order of 0.1 attojoule, which is the lowest energy consumption that can realize optical modulation. In addition, a series test of graphene plasmon to LED light response were performed for pure graphene and heterostructure under different gate voltage conditions (Figure S11, Supporting Information). The results demonstrated a stable and significant optical-modulated plasmon phenomenon in the heterostructure while the pure graphene devices nearly have no visible light response.

The overall energy consumption required for information processing and communications is one of the most pressing issues in current technological progress. The target energy consumption demanded for an optical device, as well as complementary metal–oxide–semiconductor devices, is on the order of tens of attojoules. Previous work examining the all-optical modulation of graphene has mainly exploited third-order nonlinear effects, such as saturation absorption and Kerr effect, which require high energy or intensity. For example, since saturated absorption requires strong field laser injection into graphene, the corresponding laser intensity is \( \approx 10^7 \text{ mW cm}^{-2} \) or greater. Here, the construction of atomicallyvdW heterostructures of graphene and a direct band TMD monolayer allows us to tune graphene plasmons via extremely low energy consumption for optical modulation. By using weak LED light down to 0.15 mW cm\(^{-2}\), the plasmon frequency can be tuned by \( \approx 43 \text{ cm}^{-1} \).

The response time of this optical modulation in the heterostructure may reach the picosecond timescale, which is faster than the electrical-modulation plasmonic method. In our design, the optically modulated time mainly includes not only the transfer time of photogenerated electrons (at the level of 600 fs), but also photogenerated electron-hole recombination time (\( \approx 6.3 \text{ ps} \)) due to the modulation of the p-type doped graphene plasmon inherent to the device as a result of its preparation. If n-type graphene can be effectively obtained for the processing of future devices, the optically modulated time will only depend on the transfer time of photogenerated electrons, and therefore improve to the order of \( \approx 100 \text{ fs} \).

To summarize, we have demonstrated efficient all-optical waveguide and free-space graphene plasmonic modulators based on atomically thick graphene/MoS\(_2\) heterostructures. The ultra-high photosensitivity of single layer MoS\(_2\) to visible light can prompt excitation of the photogenerated carriers in MoS\(_2\) which are transferred to graphene where they tune the graphene plasmons. We experimentally demonstrated that weak LED light (density 0.15 mW cm\(^{-2}\)) can modulate the graphene plasmons by 43 cm\(^{-1}\), which is about 4 orders of magnitude weaker than the nonlinear all-optical graphene devices. And the modulation depth can be further improved with shorter the laser wavelength and the stronger the laser power. The time response of optical tunable graphene plasmons is mainly determined by transfer time of photogenerated electrons (\( \approx 600 \text{ fs} \)) and the recombination time of photogenerated electron-hole pairs (\( \approx 6.3 \text{ ps} \)), which may enable ultrafast optical-modulation graphene plasmons in the future. Therefore, our heterostructure system design achieves efficient all optical modulation, which lays the foundation for the realization of ultracompact, low energy consumption all-optical devices for future information processing and communication.

**Experimental Section**

**Fabrication and Characterization of Heterostructures:** Monolayer MoS\(_2\) was grown directly on the Si/SiO\(_2\) substrate by CVD. The monolayer graphene was mechanically exfoliated onto Si/SiO\(_2\) substrates and then transferred onto the MoS\(_2\) using a common dry method. The polydimethylsiloxane/graphene films were clamped by a manipulator equipped with a stepper motor to assist both their peeling-off from SiO\(_2\) substrates and stamping onto receiving MoS\(_2\)/SiO\(_2\) substrates. Graphene was doped with NO\(_2\) gas by insertion in a gas chamber for half an hour. PL and Raman characterizations were performed using a Horiba Jobin Yvon LabRAM HR-Evolution Raman system with 514 nm laser excitation.

**Transient Absorption Spectrum:** The laser system that was used was based on Ti: sapphire oscillator and optical parametric oscillator (OPO) (76 MHz, 100 fs pulse duration in the output, Coherent, Inc.). The diameters of the focused pump and probe pulses were about 2 and 1 \( \mu \text{m} \), respectively. The transient absorption signal, \( \Delta R/R = (R_{\text{with pump}} - R_{\text{without pump}}) / R_{\text{without pump}} \), was recorded by a PMT and lock-in amplifier with reflective geometry. In addition, the pump and probe fluence were quite low to prevent the introduction of nonlinear processes.

**Near-Field Optical Microscopy Measurements:** Near-field imaging was conducted using a commercially available s-SNOM (Neaspec GmbH), equipped with infrared lasers. P-polarized infrared light from the monochromatic quantum cascade lasers was focused via a parabolic mirror onto both the tip and sample at an angle of 60\(^\circ\) to the surface normal. The probes were metallized atomic force microscope (AFM) probes with an apex radius of \( \approx 25 \text{ nm} \) (Nanoworld). Using this
technique, both the amplitude and phase of $E(x,y)$ were recorded. All figures showed the magnitude of $E(x,y)$.

**Nanofabrication and Electrical Characterization of Graphene/MoS$_2$ Plasmon Free-Space Modulators:** Graphene and monolayer MoS$_2$ sheets were grown on copper foil and single-crystalline sapphire wafers, respectively, by chemical vapor deposition. The sheets were then transferred onto a MgF$_2$ (600 nm)/Si substrate using the poly(methyl methacrylate) (PMMA)-assisted method. Nanoribbon arrays were patterned onto the graphene surface using electron beam lithography (Vistec 5000 +ES, Germany) in PMMA followed by oxygen plasma etching. Two Ti (10 nm)/Au (60 nm) electrode patterns were fabricated using a second electron-beam lithography cycle combined with electron beam evaporation (OHMIKER-SOBT, Taiwan). The morphologies of the fabricated graphene/MoS$_2$ nanoribbons were characterized by scanning electron microscopy (Hitachi S-4800). The electrical measurements were carried out with a semiconductor parameter analyzer (Agilent 4294A).

**Fourier Transform Infrared Microscopy Measurements:** Infrared transmission measurements were performed via FTIR microscopy (Thermo Fisher Nicolet iN10). The graphene/MoS$_2$ nanoribbon array transmission spectra at CNP ($T_{CNP}$) were used to generate a background spectrum for each measurement. The plasmon extinction spectra were given by the normalized transmittance spectra $T$ ($V_{bg} = -5$ V corresponds to $A_{bg} = -12$ V) relative to the transmittance at the CNP as $1 - T(V_{bg})/T_{CNP}$. Thus, the extinction spectra were purely plasmonic signals initiated by incident light excitation. Each measurement was repeated several times to confirm the extinction spectra.

**Simulations:** The numerical calculations were performed using the finite element method. The graphene was modeled as a homogeneous 2D conducting layer with the conductivity given by the Kubo formula. A dielectric constant of monolayer 3.7 was adopted for the MoS$_2$. To balance the 3D FEM calculations and the mesh quality, the smallest mesh size of graphene was 0.5 nm and the mesh size gradually increased outside the graphene/MoS$_2$ layer. The near-field optical image simulation of s-SNOM and the far-field infrared spectroscopy simulation of FTIR were presented in previous works.

**Supporting Information**

Supporting Information is available from the Wiley Online Library or from the author.

**Acknowledgements**

This work was supported by the National Basic Key Research Program of China (Grant No. 2015CB932400), the National Key Research and Development Program of China (Grant No. 2016YFA0201600), the National Natural Science Foundation of China (Grant Nos. 11674073, 51925203, 51902065, and 11704085), the key program of the Bureau of Frontier Sciences and Education, Chinese Academy of Sciences (Grant No. QYZDB-SSW-SLH021), the Key Research Program of the Chinese Academy of Sciences (Grant No. ZDBS-SSW-JSC002), Youth Innovation Promotion Association CAS, and CAS Interdisciplinary Innovation Team (Grant No. JCTD-2018-03), Aalto Centre for Quantum Engineering, Business Finland (A-Photonics), Academy of Finland (Grant Nos. 276376, 284548, 295777, 304666, 312297, 312551, and 314810), Academy of Finland Flagship Programme (Grant No. 320167, PREIN), the European Union’s Horizon 2020 research and innovation program (Grant No. 820423, S2QUIP), ERC (Grant No. 834742), and Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB50020200, XDPB06).

**Conflict of Interest**

The authors declare no conflict of interest.

**Keywords**

all-optical plasmonic modulators, graphene plasmon, 2D materials, van der Waals heterostructures

Received: October 30, 2019
Revised: December 5, 2019
Published online:


