Nanoimaging and Nanospectroscopy of Polaritons with Time Resolved s-SNOM

Ziheng Yao, Suheng Xu, Debo Hu, Xinzhuong Chen, Qing Dai,* and Mengkun Liu*

1. Introduction

Surface polaritons, a hybrid of electromagnetic wave and collective oscillation modes in materials,[1–4] have gained considerable attention for their potential applications in nanophotonic and quantum technologies. As a result, electronic and photonic techniques naturally merge and seek common ground. At the forefront of this technical trend is the field of polaritonics, where polaritons are half-light, half-matter quasiparticles that carry the properties of both “bare” photons and “bare” dipole-carrying excitations. The Janus-faced nature of polaritons renders the unique capability of operando control using photoexcitation or applied electric field. Here, state-of-the-art ultrafast polaritonic phenomena probed by scattering-type scanning near-field optical microscope (s-SNOM) techniques is reviewed. The ultrafast dynamical control and loss-reduction of the polariton propagation are discussed with special emphasis on the creation and probing of the tip or edge induced plasmon- and phonon–polaritons in low-dimensional systems. The detailed technical aspects of s-SNOM and its possible future development are also presented.

2. Origins and Classification of Surface Polaritons

2.1. Optical Systems Supporting Surface Polaritons

To understand why materials can support surface polaritons in specific frequency ranges, we have to review the electromagnetic theory of polaritons and optical polarization of materials. In this section, we mathematically derive the polariton modes from Maxwell equations. All materials involved are assumed to be nonmagnetic and therefore the relative permeabilities are equal to unity.[14] For systems with specific permittivity compositions, transverse magnetic (TM) polarized electromagnetic modes may exist. These modes are confined in one dimension and can propagate freely in the remaining two dimensions.
2.1.1. Interface between Two Media with Permittivities of Opposite Signs

Consider an interface between two materials. The interface lies in the plane labeled by $z = 0$. The relative permittivities of these two materials are $\varepsilon_1$ and $\varepsilon_2$ respectively (see Figure 1a for schematics). The wave vectors or field components in each area will also be labeled by 1 or 2 correspondingly.

For surface polaritons, the field is confined at the interface. The proper way of showing the evanescent nature of this field is to make the solution for Maxwell’s equations to take the following form

$$
\begin{align*}
E_x e^{-\alpha z} e^{i(qz - \omega t)} &= (E_{1x}, E_{1y}, E_{1z}) e^{-\alpha_1 z} e^{i(qz - \omega t)} \\
E_y e^{i\alpha z} e^{i(qz - \omega t)} &= (E_{2x}, E_{2y}, E_{2z}) e^{i\alpha_2 z} e^{i(qz - \omega t)}
\end{align*}
$$

(1)

where $E_1$ and $E_2$ are the electric fields, $\alpha_1$ and $\alpha_2$ are the decaying factors, and $q$ is the in-plane wavevector. Apply the Gauss’s law and the Faraday’s law to Equation (1). After imposing the interface conditions that the in-plane field components ($E_y$ and $H_x$ for transverse electric (TE) polarized mode, $H_y$ and $E_x$ for TM polarized mode) are continuous, we get

$$
\begin{align*}
E_{1y} &= E_{2y} \\
\alpha_1 E_{1x} &= -\alpha_2 E_{2x} \\
E_{1x} &= E_{2x} \\
\frac{\varepsilon_1}{\alpha_1} E_{1x} &= -\frac{\varepsilon_2}{\alpha_2} E_{2x}
\end{align*}
$$

(2)

which means $E_{1y} = E_{2y} = 0$ (no transverse electric field, TM mode only) and

$$
\frac{\alpha_1}{\alpha_2} = \frac{\varepsilon_1}{\varepsilon_2}
$$

(3)

Combining Equation (3) with the fact that

$$
\begin{align*}
\alpha_1 &= \sqrt{q^2 - k_0^2 \varepsilon_1} \\
\alpha_2 &= \sqrt{q^2 - k_0^2 \varepsilon_2}
\end{align*}
$$

(4)

we find

$$
q = k_0 \frac{E_1 E_2}{\varepsilon_1 + \varepsilon_2}
$$

(5)

Since the field strength decays exponentially with the distance away from the $z = 0$ interface, $\alpha_1$ and $\alpha_2$ are both positive. According to Equation (3), the real part of $\varepsilon_1$ and $\varepsilon_2$ must possess opposite signs. Here $\varepsilon_1$ stands for the permittivity of the topmost layer, which, in the near-field experiment, is usually air or dielectric cover of the dispersive material. So usually $\varepsilon_1$ is positive and has a flat frequency response. On the other hand, $\varepsilon_2 = \varepsilon_2(\omega)$ is dispersive and can be negative. Noble metals such as gold and silver can exhibit negative permittivities in the visible frequency range and support surface polaritons at their interfaces with dielectrics. Figure 1a shows the calculated field distribution of $H_y$ and the real space propagation of polaritons confined at the Au–Air interface at the wavelength of 532 nm. It can be observed that this polaritonic mode is tightly confined in the direction normal to the interface and can propagate freely along the interface, with certain damping.
2.1.2. Thin Film with Negative Permittivity

If the thickness \( d \) of the dispersive material is finite, polariton modes can be supported on both interfaces. Since \( \alpha \) denotes the inverse of decay length in the out-of-plane direction, for \( \alpha d << 1 \), the evanescent wave can reach the opposite interface before it dies off completely (see the field distribution in Au in Figure 1a), interacting with surface polariton at the other interface. In this situation, these two polaritons will combine into a symmetric and an antisymmetric mode. For each of them, the dispersion is modified from the case in a single interface.

The dispersions for such systems can be obtained from the characteristic equation for a planar waveguide due to structural similarities. Notice that only the fundamental and first order TM mode can be found\(^{[15]}\)

\[
\sqrt{\varepsilon_2 k_0^2 - q^2} = \tan^{-1}\left(\frac{q^2 - k_0^2 \varepsilon_2 \varepsilon_3}{\sqrt{\varepsilon_2 k_0^2 - q^2 \varepsilon_3}}\right) + \tan^{-1}\left(\frac{q^2 - k_0^2 \varepsilon_1 \varepsilon_3}{\sqrt{\varepsilon_1 k_0^2 - q^2 \varepsilon_3}}\right) + m\pi, \\
\varepsilon_1, \varepsilon_3 < 0, m = 0, 1
\]

(6)

where \( \varepsilon_1 \) and \( \varepsilon_3 \) are the permittivities of the upper and lower dielectric media, respectively.

If \( \varepsilon_1 = \varepsilon_3 \), the dispersion relation can be simplified to

\[
\sqrt{\varepsilon_2 k_0^2 - q^2} = 2\tan^{-1}\left(\frac{q^2 - k_0^2 \varepsilon_2 \varepsilon_3}{\sqrt{\varepsilon_2 k_0^2 - q^2 \varepsilon_3}}\right) + m\pi, m = 0, 1
\]

(7)

For the Air–Ag–Air structure shown in Figure 1b,c, there are two confined mode solutions for Equation (7), also at the wavelength of 532 nm. One of them exhibits a symmetric field distribution \((m = 0)\), a relatively lower field confinement, and a lower transmission loss, thus it is called long-range surface mode.\(^{[16,17]}\) The other one displays an antisymmetric field distribution \((m = 1)\), with higher field confinement as well as transmission loss, therefore it can be called short-range surface mode.

2.1.3. vdW Materials with Negative Permittivity

Due to their layered structure, vdW materials are intrinsically optical anisotropic.\(^{[18,19]}\) Together with the precondition for the existence of polaritons which requires a negative permittivity, vdW materials supporting polaritons may be hyperbolically anisotropic.\(^{[20–26]}\) Take hexagonal boron nitride (h-BN) as an example, let \( \varepsilon_\perp \) and \( \varepsilon_\parallel \) be the relative permittivities perpendicular and parallel to the optic axis, which is the \( z \) axis, respectively.\(^{[27]}\) For a specific frequency, the iso-frequency surface in the \( k \)-space is given by\(^{[22]}\)

\[
\varepsilon_\perp k_x^2 + \varepsilon_\parallel k_y^2 - \left(\frac{2\pi \omega}{c}\right)^2 = 0
\]

(8)

When \( \varepsilon_\perp \) and \( \varepsilon_\parallel \) take opposite signs, the iso-frequency surface is a hyperboloid, which spans to the region where \( k_x \rightarrow \infty \), \( k_y \rightarrow \infty \), meaning the bulk can support a mode with very large in-plane momentum.

When the in-plane momentum, say \( k_y \), is large, the hyperboloid can be approximated by a cone. The ratio between \( k_x \) and \( k_z \) converges to a fixed value for all large \( k_x \). The direction of group velocity is perpendicular to such cone or hyperboloid. The angle of the group velocity asymptotically approaches\(^{[22]}\)

\[
\tan \theta(\omega) = \frac{\varepsilon_\perp(\omega)}{\varepsilon_\parallel(\omega)}
\]

(9)

Which means, the polaritons with high in-plane momentum, will propagate along the same direction. This ray-like propagation behavior of polaritons gives rise to multiple applications such as superlensing.\(^{[22,23]}\) Such polaritons can reflect back and forth between the two surfaces of an h-BN slab, making h-BN...
an effective waveguide for the hyperbolic phonon polaritons. Taking the out-of-plane anisotropy into account, Equation (6) should be modified as

\[
\frac{\varepsilon_\perp}{\varepsilon_\parallel} \sqrt{k_\perp^2 - q^2} d = \tan^{-1} \left( \frac{\sqrt{q^2 - k_\perp^2} \varepsilon_\parallel}{\varepsilon_\perp} \right) + \tan^{-1} \left( \frac{\sqrt{q^2 - k_\perp^2} \varepsilon_\parallel}{\varepsilon_\perp} \right) + n\pi
\]

(10)

The in-plane wavevector of hyperbolic polaritons can be very large, and the higher the mode order, the larger the wavevector. The available modes can be different in different frequency ranges: in the spectral intervals where the in-plane permittivity \( \varepsilon_\perp \) is negative and the out-of-plane one \( \varepsilon_\parallel \) is positive, the in-plane wavevector is positive (Re\( \eta > 0 \)) and the fundamental mode is absent (Re\( \eta > 0 \)). In the spectral intervals where the signs of the in-plane and the out-of-plane permittivities are reversed, the in-plane wavevector is negative (Re\( \eta < 0 \)) and the fundamental mode (n = 0) exists (note that by convention in waveguide optics, the fundamental TM mode refers to the mode whose principal field component, i.e., the magnetic field, has no zero in the guiding layer. This definition is superior to that in ref. [30] where the mode order has no definite physical meanings). In Figure 1d we show the second order (n = 2) phonon polariton mode of a 50 nm thick h-BN slab in the case of Re\( \eta < 0 \)). Obviously, the field confinement is much higher than those of polaritons in noble metals.

2.2. Origins of Negative Permittivity and Categories of Surface Polaritons

Permittivity encodes the optical response of a material. The larger the permittivity, the larger the material’s polarizing capability. In fact, the word “polariton” is coined from the polarization field “particles,” analogous to photons. Generally, intrinsic collective modes of material interacting strongly with photons facilitate the origin of negative permittivity, because the Kramers–Kronig relationships usually demand a steep dispersion of the real part of the permittivity, in the vicinity of the collective mode resonance frequency. A negative permittivity means the external electric field is screened. Such screening confines the field at the interface and forms surface polariton. Based on the types of the collective modes associated, polaritons can be classified into surface plasmon polaritons (SPPs), surface phonon polaritons (SPhPs), exciton polaritons (EPs), Cooper pair polaritons, and magnon polaritons, etc. Here we briefly introduce the first three, which have been successfully investigated in the ultrafast-near field experiments.

2.2.1. Plasma Resonance and SPPs

For materials with free charge carriers, the carrier density oscillations can resonate with external field at the plasma frequency \( \omega_p \). Without loss of generality, the permittivity of plasmonic systems can be expressed in the Drude form

\[
\varepsilon(\omega) = \varepsilon_\infty \left(1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega} \right)
\]

(11)

where \( \varepsilon_\infty \) is the high-frequency limit of permittivity and \( \gamma \) is the damping rate. The plasma frequency is given by \( \omega_p = \sqrt{4\pi n e^2/(m^* \varepsilon)} \). According to Equation (11), negative permittivity can be achieved below the plasma frequency, supporting SPPs. The theories, realizations and applications of SPPs can be found in many books and reviews.

2.2.2. Phonons and SPhPs

Polar materials are usually of permanent dipoles which interact resonantly with mid-infrared light. This resonance constitutes another mechanism of negative permittivity. The frequency-dependent permittivity for polar dielectrics can be expressed as

\[
\varepsilon(\omega) = \varepsilon_\infty \left(1 + \frac{\omega_{\text{TO}} - \omega_{\text{LO}}}{\omega_{\text{TO}} - \omega - i\gamma\omega} \right)
\]

(12)

where \( \omega_{\text{TO}} \) and \( \omega_{\text{LO}} \) refer to the transverse optical phonon (TO) and longitudinal optical phonon (LO) frequency respectively. According to Equation (12), negative permittivity can be realized in the frequency gap between the LO branch and the TO branch, which is known as the Reststrahlen band. The band location and width are determined by the oscillator’s effective mass and the interatomic field strength. As observed in SiC and h-BN using s-SNOM, SPhPs show considerable long lifetime (low damping rate).

2.2.3. Excitons

Exciton, a typical quasiparticle in excited semiconductors, can also couple with photon in proper energy range and forms EPs. Exciton can be regarded as a hydrogen-like system. When the Coulomb interaction between electron and hole is weak, the Wannier excitons with Bohr radius larger than lattice spacing are delocalized and screen out the external electric field. One unique property of exciton polaritons is the spatial dispersion effect because of the non-negligible ratio of energy transported by excitons. A wavevector-dependent term is included in the dielectric function

\[
\varepsilon(\omega,k) = \varepsilon_\infty \left(1 + \frac{\omega \beta k^2 - i\omega \gamma}{\omega^2 - \omega_{\text{exc}}^2 - i\gamma\omega} \right)
\]

(13)

where \( \omega_{\text{exc}} \) is the resonance frequencies of longitudinal exciton and transverse exciton. \( \beta = (\hbar \omega_{\text{exc}}/M^*) \) where \( M^* \) is the effective mass of exciton. The \( k^2 \)-square term introduces more tunable parameters into the dielectric function and results in many unique properties of the EPs. For example, recent studies have experimentally verified that strong binding exciton can remain stable even in room temperature in Group VI transition-metal dichalcogenides (TMDs) with chemical formula \( \text{MX}_2 \) (\( M = \text{Mo, W}; X = \text{S, Se}) \).
3. Experimental Technique and Polariton Detection

3.1. Near Field Technique

The capability of probing beyond diffraction limit and accessing evanescent waves make s-SNOM a promising technique for investigating polaritons.\[52\] Here we briefly introduce the basic concepts in near-field optics and the s-SNOM technique. For a more detailed discussion of modern near-field optical microscopy, readers are directed to several recent reviews.\[53–56\]

When electromagnetic wave is scattered by an object, the scattered wave always contains both propagating wave and evanescent wave.\[57\] For evanescent wave, there will be at least one direction, in which the wave vector component is larger than that in free space. A larger wave vector, or a shorter wavelength, can carry information of spatial variations finer than the wavelength of the incident light. On the other hand, the evanescent wave fades away quickly in a distance in the order of wavelengths away from the object surface, therefore undetectable in the far-field.

The detection of evanescent wave is the key to achieve optical imaging beyond the diffraction limit. One method is to put a small probe near the sample surface.\[58,59\] within the extension of the evanescent field. The probe can act as an antenna and reradiate the evanescent wave to far-field.\[59\] With the advent of atomic force microscopy (AFM), this idea was developed into different techniques, one of which is known as s-SNOM. The AFM probe, or the “tip,” is placed tens of nanometers above the sample surface and polarized by the incident light. The local electric-field beneath the tip apex is greatly enhanced within a range comparable to its radius of curvature $a$. When the tip is brought close to the sample surface, this localized field would interact with the sample strongly. Therefore, the scattered light due to this interaction contains information about the sample volume right beneath the tip. A common practice is to assume the strength of near-field signal as proportional to the localized field between the tip and sample surface, especially in the case of imaging polaritons using s-SNOM.

To obtain the genuine near-field signal, the tip-sample separation is modulated by tapping the AFM tip with an amplitude of $\approx 50$ nm and a frequency of $\Omega$. The exact tapping amplitude depends on the wavelength of light and the specific setup. Since the near-field signal exhibits strong nonlinearity with the tip-sample distance, it can be effectively extracted by demodulating the output of the detector using a lock-in amplifier at high harmonics of $\Omega$. Due to the limitation of the signal-to-noise ratio, $S_2$ or $S_3$ (the second or the third harmonic near-field signal) are usually chosen.

3.2. Role of s-SNOM in Studying Polaritons

Since polaritons are modes confined at sample surface and have an imaginary out-of-plane wave vector, their in-plane momenta are larger than the momentum of light in free space (from Equation (5), when $\varepsilon_1$ and $\varepsilon_2$ are of opposite signs, $|q|>|k_0|$). As a result, direct free space excitation at smooth surfaces cannot yield polaritonic mode. Various methods have been applied to provide extra in-plane momentum. For example, attenuated total reflection (ATR) with dielectric materials is widely used to excite and detect surface polaritons in the far-field experiments.\[60–65\] The nonlinear excitation\[66,67\] provides momentum from the nonlinear susceptibility of the material. Gratings\[55,68,69\] can provide multiples of $g = 2\pi/d$ for the in-plane momentum, where $d$ is the spacing of gratings. An artificial metasurface with gradient reflection phase\[70\] breaks the translational invariance at the interface, thus providing an intrinsic in-plane momentum component at the surface.

Roughly speaking, the tip scattered evanescent wave in a prototypical s-SNOM setup spans in a wide momentum range in the order of $1/a$,\[71\] providing sufficient in-plane momentum for launching polaritons.\[72\] The tip-launched polaritons spread circularly (as shown in Figure 2a) and can be reflected back to the tip by edges or defects of the sample. The total electric field of the incident and reflected polaritons can be picked up by the tip and converted into far-field signal. By scanning the tip across the sample surface, the interference pattern (Figure 2b) between the incident and the reflected polaritons can be registered. For CW laser excitation, the interference pattern comes from a standing wave with alternative nodes and antinodes, where the field amplitude reaches local maxima or minima. The distance between successive antinodes (or nodes) of the standing wave is approximately half of the polariton wavelength.\[70,71\] In the tip-launched case, the polaritons propagate as cylindrical wave. To comply with the law of conservation of energy, the field strength does not only decay due to the imaginary part of the wave vector, but also with a geometrical factor $1/\sqrt{r}$.

In some cases, the sample edges or other defects can also act as polariton launchers.\[74\] The spacing between adjacent signal maxima is approximately one wavelength instead of half wavelength.\[75\] For simplicity, a straight edge can be regarded as a plane wave source without geometric decay. In this case, the decay of edge-launched polariton is purely attributed to the imaginary part of the wavevector. Considering that the tip-launched polaritons travel twice as much as the edge-launched ones, it is easy to understand that, in a system where both integer wavelength and half wavelength are observed, the half-wavelength pattern usually decays faster.\[75,76\]

For s-SNOM imaging using ultrafast laser pulses, the imaging mechanism can be slightly different. To illustrate one of the many possible scenarios, we divide the imaging process into two steps. The first is shown in Figure 2c. The ultrafast pulse is focused onto the tip apex and induces polaritons. The electric field of the polaritons is picked up by the tip and scattered as a far-field signal into the detector. This far-field signal is labeled as the pink “emitted signal” $E_i(t)$ in Figure 2c. As the emitted signal propagates in the free space, the polaritons propagate as cylindrical wave pulse along the sample surface, denoted as “ultrafast polariton” in Figure 2c.

The second step is shown in Figure 2d. The ultrafast polaritons reach the sample edge, then reflected back. As the reflected polaritons reach the tip, the electric field is picked up by the tip again, converted to another set of far-field signal, and scattered into the same detector at a later time. Here it is shown as the black “reflected light” in Figure 2d, labeled as $E_r(t)$.

A toy model considers $E_i(t) = E_i(t + \Delta t)$ can be made to illustrate the signal detection here, where $\Delta t$ is the time for...
polariton travel from the tip, to the edge, and finally back to the tip. Here we assume tip-launched-edge-reflected polaritons instead of edge launched ones. In this toy model, we ignore all the dispersion, tip-sample coupling, dissipation, or phase changes by reflection and capture simply the mechanisms of fringe formation.

At a specific frequency, a Fourier transform of the electric field scattered by the tip shows as

\[ E_i(t) \rightarrow E_i(\omega) \]  

(14)

\[ E_i(t) = E_i(t + \Delta t) \rightarrow E_i(\omega) \exp(i\omega\Delta t) \]  

In the frequency domain, they add up as

\[ E_{tot}(\omega) = E_i(\omega) + E_i(\omega) = E_i(\omega)\left[1 + \exp(i\omega\Delta t)\right] \]  

(16)

Again, \( \Delta t \) is the time for polariton to have a round trip between tip and edge. As the distance between the tip and sample edge changes, the time \( \Delta t \) will change accordingly:

\[ \Delta t = 2x/v_\gamma = 2x/\left(\beta_\gamma\right) = 4\pi x/(\omega\lambda_\gamma) \]  

(17)

Here \( v_\gamma, \lambda_\gamma \) are velocity and wavelength of the polariton. \( f = \omega/2\pi \) is the frequency. The relation \( v_\gamma = \beta_\gamma \) is used to substitute the velocity. The 2x appears due to the total travel distance of the polariton is twice the tip-edge distance. Plug it into the total field

\[ E_{tot}(\omega) = E_i(\omega)\left[1 + \exp(i\omega\Delta t)\right] = E_i(\omega)\left[1 + \exp\left(i4\pi x/\omega\lambda_\gamma\right)\right] \]  

(18)

For the phase in the last term to change \( 2\pi \), the tip-edge distance should change

\[ 2\pi = 4\pi x/\lambda_\gamma \Rightarrow x = \lambda_\gamma/2 \]  

(19)

Hence, the spatial periodicity in the detected electric field is half of the polariton wavelength, the same result as in the CW wave experiment. Here, we did not require \( E_i(t) \) to be nonzero throughout the time of detection. \( E_i(t) \) being an ultrafast pulse or a CW laser does not affect our result. The key for the fringe formation here is that, in a Fourier transform, two pulses with changing intervals between them will result in a changing magnitude at a certain frequency.

For the case of edge-launched polaritons, the travel distance of polaritons is the same as the tip-edge distance, so the time delay \( \Delta t = x/v_\gamma \), leading to a result of \( x = \lambda_\gamma \).

From the discussions above, we can see that, in the ultrafast detection the polariton “pulses” do not have to interfere with themselves to form observable fringes, the distance between adjacent fringes are the same as the CW experiment under the same conditions. This promises that the detection of ultrafast polaritons can share the same modeling as polaritons generated by CW laser.

A realistic experimental setup for near-field detection of CW laser-induced polaritons is shown in Figure 3. The whole setup can be separated into a light source, an asymmetric Michelson interferometer, a focusing optics, and an AFM. The light source produces CW laser, which propagates through the beam splitter and is split into a detection arm and a reference arm. The detection arm contains the focusing optics, the AFM, and the sample. For a single-wavelength laser, the focusing optics can be a lens. Otherwise, an off-axis parabolic mirror is often used.[56] As the AFM tip raster scans the surface of the sample, the spatial variations in the near-field image suggests the wave...
vector $q$ in the corresponding laser frequency $\omega$, settling a data point on the dispersion. With tunable laser or variable CW light sources, multiple mappings can be done to yield the dispersion relation. The light from reference arm interfere with the scattered light from the detection arm at the detector, amplifying the near-field signal significantly and eliminating the far-field background. Two commonly used detection methods involving a reference mirror are homodyne detection (HD) and pseudoheterodyne detection (PHD).\cite{77-79} In HD, the reference mirror stays at fixed positions. In PHD, the reference mirror oscillates at frequency $\omega_m$ and the signal is demodulated at sidebands of the frequency $n\Omega + m\omega_m$ ($n$ and $m$ are integers with $n > 1$). This allows the simultaneous acquisition of the near-field amplitude and phase.

The setup for broadband or time-domain measurements is shown in Figure 4. The probe beam is an ultrafast laser. By using a long delay stage, the reference beam can map out the time-profile of the near-field signal. This is similar to the setup of Fourier-transform infrared spectroscopy (FTIR). With a broadband probe light source, such as synchrotron light,\cite{80} high temperature plasma light source,\cite{81} or laser driven plasma,\cite{82} such setup, so-called nano-FTIR,\cite{83} can measure IR spectra with a spatial resolution of $\approx 20$ nm.\cite{84} The technique has produced flourish results in phase transition materials,\cite{85-88} polariton dispersion,\cite{89,90} catalytic chemistry,\cite{91} biology,\cite{92} and geoscience, among other fields.\cite{80,94}

Another idea for ultrafast measurement is to combine pump-probe setup with s-SNOM. A pump light, usually with

Figure 3. Multicolor mapping of plasmon polariton dispersion using tunable CW lasers. a) Top panel: dispersion relation of plasmon polariton of graphene single-layer on SiO$_2$. Adapted with permission.\cite{72} Copyright 2011, American Chemical Society. The positions where the vertical and horizontal dashed blue and red lines cross refer to experimentally extracted data points at two different frequencies. Bottom panel: the interference patterns of tip-launched polaritons corresponding to different frequencies. b) Schematics of the near-field experimental setup, capable of both HD and PHD detection.

Figure 4. Phonon plasmon polariton dispersion obtained by hyperspectral line scan in an ultrafast broadband near-field setup. a) Top panel: schematics of the dispersion relation of hybridized phonon–plasmon polaritons of h-BN/Graphene/h-BN heterostructure. Bottom panel: the hyperspectral maps of hybridized phonon–plasmon polaritons in a h-BN/graphene/h-BN heterostructure with pump-probe delay at 0 fs (Adopted with permission.\cite{75} Copyright 2016, Springer Nature). b) Schematics of the ultrafast near-field experimental setup.
wavelength shorter than the probe light, is focused onto the tip with the same or different focusing optics. A delay stage is added to the path of the pump beam to vary the time delay between these two. As the system is pumped away from equilibrium, the dynamics of the photo-excited electrons and their recovery to thermal equilibrium will be detected by the probe beam arriving at various time delays after pumping. The setup has been carried out successfully in the infrared, high-THz, and THz frequency regions. This is also the basic scheme for the detection of ultrafast polaritons using s-SNOM. The pseudo-heterodyne detection is also determined to be compatible with the near-field pump probe measurements and yield artifact free results. However, this method has not been carried out in ultrafast polaritons experiments using s-SNOM.

3.3. Imaging CW Polaritons with s-SNOM

Before near-field experiments, graphene plasmons have been widely investigated in the far-field. However, the large momentum mismatch between free-space photons and plasmons confined in 2D results in low light-plasmon coupling efficiency in launching and probing of the graphene SPPs. The direct visualization of propagating SPPs in graphene was first achieved by two near-field studies using s-SNOM. These two works demonstrate that the tapered graphene ribbon on 6H-SiC and graphene/SiO\(_2\)/Si back-gate structure can effectively support the propagation of SPPs with highly reduced wavelength (\(\lambda_0 = \lambda_0 / 40\)). The 2D maps of the near-field amplitude present periodic fringes close to the graphene edges or defects. These fringes are interpreted as interference between the incident and reflected SPPs, as described in the previous section.

It is noticeable that the line defects play an important role in the formation of the plasmonic interference patterns in real space. Conversely, nanoimaging is a powerful tool to discover other type of defects with plasmonic response, for instance, the grain boundaries in graphene prepared by chemical vapor deposition (CVD). Such boundaries are hard to reveal by AFM but can be characterized by twin fringes in near-field nanoimaging.

Using the Drude-like conductivity model for graphene, the plasmon wave vector can be written as

\[
q_p = \frac{\hbar}{2eE_f} a_0 \left( \omega + i\frac{\tau}{\gamma} \right)
\]  

where \(\kappa(\omega)\) refers to the effective dielectric function considering the interaction of the substrate. \(E_f\) is the fermi energy and \(\tau\) is the relaxation time. The plasmon wavelength is determined by the real part of the plasmon wave vector \(\lambda_p = \frac{2\pi}{Re(q_p)}\), and the plasmon damping rate is defined as the ratio between the real part and the imaginary part of the wave vector, \(\gamma_p = \frac{\operatorname{Re}(q_p)}{\operatorname{Im}(q_p)}\).

According to the linear dispersion at two Dirac points in the Brillouin zone, the Fermi momentum can be expressed by the free carrier density in a rather simple form \(k_F = \sqrt{\frac{\pi n}{\lambda}}\). By fitting to the line profiles of the interference pattern in experiments, the carrier density and damping rate of plasmon polaritons can be obtained simultaneously.

The electric tuning of graphene plasmons is originated from the \(E_f\) term in Equation (20). As the gate voltage in a p-doped graphene/SiO\(_2\)/Si back-gated system decreases, the plasmon wavelength increases and the strength of the interference pattern is enhanced due to increased free-carrier density. Besides the tunability, the quality factor, defined as the inverse of the damping rate \(Q_p = \frac{1}{\gamma_p}\), is also one of the most important indices for polaritonic devices. Compared with plasmons of noble metals with high loss due to absorption, graphene plasmons provide promising low-energy loss and relatively long propagation length. In the first near-field experiments, the quality factor measured in graphene/SiO\(_2\)/Si structure is unexpectedly low, around 7.4, indicating a rather strong damping rate caused by surface irregularity, enhanced electronic relaxation rate and many-body effects.

The later CW imaging of high-quality graphene encapsulated in h-BN yields a much higher quality factor valued 50. The damping can also be significantly suppressed in low temperature environment. An unprecedented high quality factor for encapsulated graphene reaches 130 in cryogenic temperature (60 K).

A system with strong SPhPs which has been investigated extensively in the near-field regime is h-BN, in the form of both nano flakes and nanotubes. Fourier transform IR nanospectroscopy (nano-FTIR) was performed to investigate the dispersion of SPhPs in h-BN. Compared with the data obtained by the CW imaging, the broadband nanospectroscopy provides an efficient way to map out the dispersion in a single line scan.

The near-field studies on EPs are realized in WSe\(_2\) using aperture-SNOM and in MoSe\(_2\) using s-SNOM, respectively. Complex line profiles formed by the interferences of various waveguide modes with different in-plane momenta are observed. The waveguide modes are highly damped when the incident photon energy is larger than the exciton energy. In the MoSe\(_2\) waveguides, a long propagation distance of waveguide exciton polariton up to 12 \(\mu\)m was realized.

In a multilayer structure where different types of polaritons coexist, hybridization happens in the vicinity of dispersion crossover. One prototype of the hybrid modes is the phonon–plasmon polariton. A strong enhancement of the near-field strength intensity was observed in monolayer graphene on SiO\(_2\)/Si substrate in 2011. The gate voltage–dependent spectral enhancement around the bare surface optical phonon frequency of SiO\(_2\) is explained by the phonon–plasmon coupling at the graphene–SiO\(_2\) interface. The effect of hybridization strongly relies on the excitation frequency. The hyperbolic phonon polaritons in h-BN effectively couple with plasmon polaritons in graphene, giving rise to the hyperbolic plasmon–phonon polaritons (HPPPs). The HPPPs can be tuned by varying gate voltage but still inherit the low damping rate from hyperbolic phonon polaritons. The propagation length is 1.5–2 times larger than the unhybridized hyperbolic phonon polaritons in h-BN. Therefore, HPPPs show unprecedented features that provide alternative routes to the design and optimization of the polaritonic devices.
3.4. Imaging Ultrafast Polaritons with s-SNOM

3.4.1. Spectral Measurements

The first near-field imaging of graphene plasmons in 2012 has ignited a worldwide interest in investigating polaritons using s-SNOM.[73,106] Up to now, the near-field study of polaritons continued to spur innovation and galvanize thought in a variety of photonics systems that extend to many different research fields.[118–121] Later on, the graphene nanoimaging, and graphene plasmon nanoimaging, have become the touchstone of near-field imaging systems, showing the capabilities of, for example, ultrafast s-SNOM,[75,95,122] THz s-SNOM,[123] and cryogenic s-SNOM.[109]

Inspired by the successes of imaging time-independent plasmons excited by CW lasers, investigating the plasmon dynamics in the ultrafast regime becomes a natural follow-up. Two years after the original work, infrared pump-probe experiment beyond diffraction limit was first realized in the graphene/SiO2/Si system (Figure 5) with a temporal resolution of ≈200 fs.[95] Owing to the fine spatial resolution, especially with the artifact-free methods,[102] the authors in ref. [95] managed to measure the ultrafast electron and lattice dynamics free from interferences from the defects and grain boundaries.

The experiment setup is shown in Figure 5a. The ultrafast (<100 fs) near-infrared (NIR) pulse (λ = 1.56 µm) was generated by a 40 MHz Er-doped fiber laser. The MIR probe was generated by difference-frequency mixing of NIR pulse and supercontinuum (SCIR) pulse at λ = 1.8 µm in GaSe. The NIR pulse also serves as the optical pump to photoexcite the plasmonic responses, which interact with two phonon modes of SiO2, α (1125 cm⁻¹) and β (785 cm⁻¹), marked by the two horizontal lines in Figure 5c. The coupling between the surface phonon in SiO2 and the plasmons in graphene gives rise to the hybrid phonon–plasmon mode, shown in the calculated dispersion in Figure 5b.

The relative change in near-field amplitude Δs(ω)/s(ω) increases with the increasing pump power (Figure 5c top). At 10 mW pump power, the pumping effect is similar with a −70 V DC gating. Compared with the peak at high-frequency range (ω > 1100 cm⁻¹), the one at the β mode (ωβ = 785 cm⁻¹) is wider and more subtle, corresponding to a higher damping and a steeper slope of the plasmon dispersion (Figure 5b). The effect of pumping is explained as an increase of electron temperature. In a previous FTIR spectroscopy measurement of the transmission spectrum of graphene, a Drude-like frequency dependence of conductivity was observed.[124] The conductivity of graphene can be modeled by the following formulas

$$\sigma(\omega) = \sigma_{\text{intr}}(\omega) + \sigma_{\text{inter}}(\omega), \quad \sigma_{\text{intr}}(\omega) = \frac{i D}{\pi \omega + i \gamma}$$

(21)

where D is the Drude weight and γ is the scattering rate. The intraband Drude weight

Figure 5. Ultrafast nanospectroscopy of exfoliated graphene. a) Schematic of the near-field pump-probe setup. b) Dispersion relation of phonon–plasmon polariton with different Drude weights D and damping rates γ (left: D = D0, γ = 300 cm⁻¹, right: D = 1.46 D0, γ = 300 cm⁻¹). c) Experimental (top) and theoretical (bottom) results of the relative changes in near-field amplitude induced by optical pumping with different power (solid curves) and electrostatic gate (black dots). Two vertical lines β and α correspond to the SiO2 phonon modes at 785 and 1125 cm⁻¹, respectively. Comparison of the spectral features between pump- and gate-induced data is the first attempt to render insight into the ultrafast plasmon. d) Time resolved changes of spectrally integrated near-field amplitude Δs/s for single-layer, bilayer and trilayer graphene. e) Δs/s on different layers of graphene at various time delays. The number of graphene layers are shown at the left panel. Area: 6 µm × 6 µm. Adopted with permission.[95] Copyright 2014, ACS.
Dk

\[
\eta = 2e \ln \cosh \left( \frac{2B}{\mu} \right) 
\]

(22)
is temperature dependent. \cite{125,126} As \( k_B T \gg \mu \), the Drude weight approximately linearly depends on electron temperature. Higher Drude weight increases the slope of plasmon dispersion in graphene, therefore increases the spectral strength of the upper branch (Figure 5b). Meanwhile, larger scattering rate causes the broadening of the dispersion (Figure 5b right). By tuning the Drude weight and scattering rate, the near-field response of the system is modeled in the bottom of Figure 5c, which shows good agreement with the experiments (Figure 5c top). It is noteworthy that the probe pulse with high in-plane field strength also increases the Drude weight and scattering rate. The plasmonic features in equilibrium state obtained from ultrafast nanospectroscopy is, therefore, inevitably perturbed by the tip-enhanced probe pulse. As a result, the probe beam heats the electrons to above \( \approx 1270 \, \text{K} \) \((D = 1.10 \, \text{D}_0 \text{ with MIR probe only})\); therefore the real-space polariton features are totally damped out.

The first realization of ultrafast near-field response in a narrow bandgap semiconductor using s-SNOM is in InAs. \cite{96} The experimental setup is identical to that in Figure 5a. The energy of NIR pump (1560 nm, or 0.794 eV) is larger than the direct bandgap in InAs (0.35 eV). The InAs sample is weakly n-doped so that in equilibrium it does not show any distinct plasmonic feature. The NIR beam pumps electrons into the \( \Gamma \) valley (Figure 6d), leading to a higher carrier density before electron–hole recombination. Figure 6a shows near-field amplitude under increasing pump power at 0.8 ps time delay, where the most pronounced changes occurred. A peak appears at the lower frequency end, moves toward higher frequencies and gets broadened. The changes in peak frequency and width indicate the increase of plasma frequency and scattering rate with increasing pump power. As the pump power increases, the carrier concentration increases accordingly, shifting the plasma frequency into the frequency region available for MIR probe. Figure 6b shows the simulated plot of near-field amplitude with plasma frequency \( \nu_p \) and scattering rate \( \gamma \) in Drude model.
The experimentally retrieved $\nu_p$ and $\gamma$'s are shown in inset of Figure 6b.

The changes in plasma frequency and scattering rate can be mapped out in real time by measuring spectra at different time delays with a fixed pump power (18 mW in Figure 6c). The plasmonic peak appears at $\approx$ 0.2 ps and reaches the highest frequency at 1 ps. The spectral changes in the experiment are well reproduced by Drude model considering the change of plasma frequency $\nu_p$ and scattering rate $\gamma$. The retrieved dynamics of $\nu_p$ and $\gamma$ show $\approx$ 1 ps raising time, $\approx$ 1 ps fast decay time and $\approx$ 60 ps slow decay time, which can be attributed to relaxation of pumped electrons in the $\Gamma$ valley, carrier-lattice cooling and defect mediated recombination, respectively.[127–129] The evolution of plasma frequency $\nu_p$ is shown in Figure 6e.

No real-space interference patterns were observed in InAs regardless of the very large time constant ($>$ 60 ps) at the plasma frequency. A very important reason is the near-zero group velocity. As shown in Figure 6f, the dispersion for plasmon polariton in InAs is nearly flat at the tip-accessible momentum range (maximum of the yellow curve in Figure 6f), implying a near-zero group velocity of plasmon polariton, preventing the polariton propagation. Since the real-space imaging requires interference between incident and reflected polaritons as stated in Section 3.2, no interference pattern could be imaged with group velocity $= 0$.

### 3.4.2. Real Space Imaging of Ultrafast Polaritons

To real-space propagation dynamics of plasmon polaritons is observed in high-mobility graphene encapsulated in hexagonal boron nitride (h-BN).[75] The dissipation rate of graphene was reduced to $\approx$ 130 cm$^{-1}$ when being pumped, around one third of the graphene/SiO$_2$ system. The carrier mobility near the Dirac point exceeds 14 000 cm$^2$ Vs$^{-1}$, larger than that in the graphene/SiO$_2$ sample with an optimistic estimation around 10 000 cm$^2$ Vs$^{-1}$.

The experimental setup is shown in Figure 7a. High mobility graphene encapsulated in h-BN is transferred onto SiO$_2$ substrate and back-gated. When gated with a voltage of 30 V and pumped with a NIR (1560 nm) ultrafast laser, interference pattern can be observed at the graphene boundary with broadband MIR probe. The plasmon polaritons dispersion can be revealed in a single line scan. In Figure 7c,d, the hyperspectral maps at two different time delays show the increase of plasmon wavelength with increasing frequency. The plasmonic response is the strongest at the zero-time delay, showing three observable interference fringes. At the 2 ps time delay, only 1 fringe is observable. The wavelength at the same frequency $\omega$ also decreases.

By comparing the ultrafast results (Figure 7c,d) with interference pattern under 890 cm$^{-1}$ CW laser, the spacings between the adjacent fringes are determined as $\lambda_p$ instead of $\lambda_p/2$.

By counting $\lambda_p$, a set of frequency-wavelength relations can be obtained. The effective electron temperature and Drude weight can be determined via model fitting. The relation between plasmon wavelength and Drude weight is given by

$$\lambda_p = 4\pi D(T_e)/\kappa(\omega)\omega^2$$

where $D(T)$ is the Drude weight in graphene and $\kappa$ is the permittivity of h-BN. As $D$ decreases, the plasmon wavelength decreases, which is clearly observed by comparing Figure 7c to 7d. In Figure 7c, the effective electron temperature $T_{el}$ is 3200 K at zero-time delay. As the electron temperature drops with increasing time delay, the Drude weight decreases and leads to a smaller wavelength at the same frequency, as shown in Figure 7d, where $T_{el} = 1700$ K.
The authors suggested a transient process in which carriers follow a non-Fermi–Dirac distribution immediately after the pumping, followed by two distinct Fermi–Dirac distributions for electrons and holes with near-equal temperature $T_{el}$ but different chemical potentials. The chemical potentials for equations and holes equilibrate at a later stage and $T_{el}$ gradually approaches lattice temperature. However, due to insufficient temporal resolution, these dynamics, especially the first two transient processes, were not directly probed by this work. The real-space plasmon polariton propagation as a function of time is also not clearly resolved.

Experiments of encapsulated graphene revealed a mechanism of edge-launched polaritons and the capability of observing ultrafast polariton dynamics using near-field setup. However, the real space interference pattern between the tip launched polariton and the backscattered polariton from edge or impurity has not been realized using ultrafast lasers until 2016 by Huber et al., who utilize the low damping nature of hybridized polariton in $\text{SiO}_2$/black phosphorus (BP)/$\text{SiO}_2$ heterostructure.$^{[76]}$

Encapsulated BP shows relatively low near-field signal compared with the $\text{SiO}_2$ substrate without pumping (Figure 8a top). When pumped by a 1560 nm, 40 fs NIR beam, the near-field signal increased substantially, and shows alternative bright and dark fringes parallel to the sample edge (Figure 8a bottom), which is similar to the aforementioned interference pattern of SPPs or SPhPs observed in CW based near-field experiments.

To determine the origin of the near-field pattern, the polariton dispersions of BP/$\text{SiO}_2$ heterostructure are derived under unpumped and pumped conditions (Figure 8c). The SPhPs of $\text{SiO}_2$ splits into two branches due to the coupling between two interfaces of cover layer $\text{SiO}_2$. The lower frequency longitudinal optical mode is labeled as LO-. Similarly, the SPP mode of BP splits due to the coupling between two BP/$\text{SiO}_2$ interfaces, into a symmetric (asymmetric) mode with lower (higher) energy, SPP- (SPP+). Without pump, LO- and SPP+ stay uncoupled (Figure 8c left). When pumped by NIR pulse, electron–hole pairs will be excited, raising the carrier density, as well as the plasma frequency $\nu_p$. When $\nu_p$ is raised to the Reststrahlen band (between two horizontal white dashed lines in Figure 8c), SPP+ and LO- will intersect, forming hybrid modes. The right panel of Figure 8c shows the hybrid mode of SPP+ and LO- in the white circle when $\nu_p = 38$ THz, which resides in the Reststrahlen band of $\text{SiO}_2$.

A line scan across the $\text{SiO}_2$/BP heterostructure boundary was done to obtain the dispersion relation for comparison with theoretical predictions. The hyperspectral map of this line scan is shown in Figure 8d. The polariton wavelength does not change much with the frequency. By imposing Fourier transform for each frequency on Figure 8d, the dispersion of the hybrid polaritons can be obtained, as shown in Figure 8e. The polariton response

---

**Figure 8.** Femtosecond photo-switching of interface polaritons in black phosphorus heterostructures. a) Near-field amplitude $I_{4}$ before (top) and 250 fs after (bottom) pumping; $\text{SiO}_2$/BP/ $\text{SiO}_2$ heterostructure sample is at left and $\text{SiO}_2$ substrate at right. The interference fringes parallel to the sample boundary are clearly seen. Scale bar: 2 $\mu$m. b) Left panel: near-field line scans at different time delays across the sample boundary. Right panel: Dependence of average interference amplitude on time delay. c) Calculated dispersion of BP polaritons with plasma frequency at 10 THz (left) and 38 THz (right), corresponding to unpumped and pumped status respectively. $\text{SiO}_2$ Reststrahlen band sits between the two white dashed lines. d) Hyperspectral line scan in the direction perpendicular to the stripes. e) The dispersion of hybrid phonon–plasmon polaritons, obtained by Fourier transform in $x$-direction of Figure 8d. f) Theoretical dispersion with plasma frequency at 36.6 THz, showing the same frequency-momentum region as in (e). (Adopted with permission.$^{[76]}$ Copyright 2016, Springer Nature).
is confined in the hybrid region of SPP+ of BP and LO- of SiO2. Two vertical dashed lines show the minimum momentum resolution limited by the total length of the line scan. The theoretical dispersion in the same frequency–momentum region is shown in Figure 8f with \( \nu_p = 36.6 \text{ THz} \) for comparison.

To investigate the time evolution of plasmon–phonon polaritons, line scans at different time delays are performed. The results are shown Figure 8b left panel. Fitting sine function to each line scan produces the evolution of polariton strength as shown in Figure 8b (right panel). Thus, the rise time for the polariton response to reach half of its maximum is determined to be within 90 fs. It can also be observed that, the wavelength does not change much with time delay, meaning the wavelength of hybrid polariton only weakly depends on the carrier density or plasma frequency. The hybrid polariton response is relatively decoupled from pump photon energy or carrier density, occupying a region selected by the intersection in frequency–momentum space during the entire \( \approx 5 \text{ ps} \) lifetime of the hybrid polaritons. The line profiles in experiment also show good agreement with damping-free propagation theoretical calculation, indicating low damping rate for the hybrid polariton.

### 3.4.3. Spatiotemporal Imaging

The real-space imaging of polaritons directly revealed the electric field profiles in a subwavelength scale, utilizing the high spatial resolution capability of s-SNOM. However, due to the limited temporal resolution and high damping, one is unable to obtain useful information such as the group or phase velocity. Recently, in the visible–near-infrared (VIS–NIR) frequency range, the temporal and spatial resolution of near-field pump-probe system has achieved 43 fs and 50 nm respectively, allowing the first direct spatiotemporal imaging of EP in a WSe2 slab, and revealing its group velocity to be \( v_g \approx 0.017 c \).\[130\]

The experiment setup is shown in Figure 9a. A sub-10 fs broadband Ti:Sapphire laser beam is split into pump (650–700 nm) and probe (700–1050 nm) beams by a dichroic mirror (DM). The temporal resolution is determined by the convolution between the pump and probe to be \( \approx 43 \text{ fs} \). The near-field images are obtained at the wavelength of 760 nm, close to the A exciton transition in WSe2 (inset of Figure 9a) thus showing strong near-field signal contrast.

The snapshots of interference patterns at different time delays are shown in Figure 9b. The number of interference fringes is increasing with time. Figure 9c shows the result of spatiotemporal imaging, where the line profile of the polariton fringes is plotted versus time (inset shows the real-space image and the linescan). With increasing time delay, the first stripe moves inward (to the left in Figure 9c) and new stripes appear at the right boundary. The right triangle shows the distance \( \Delta x \) the first stripe has moved within a time interval \( \Delta t \), yielding an estimation of the group velocity \( v_g \approx 4.7 \pm 0.5 \times 10^6 \text{ m s}^{-1} \).

The observed pattern was determined to be the “constructive interferences of the EP wave packets propagating inside WSe2 slab waveguide.” The loss of EP wave packet was estimated to be \( \approx 2.6 \mu \text{m}^{-1} \), which is likely caused by collisions between EP and photo-injected free electrons. In this work, the modification of the dielectric function introduced by intense laser around the exciton resonance,\[131\] so called “renormalization,”

---

**Figure 9.** Ultrafast exciton polaritons in a WSe2 flake. a) Schematics of the ultrafast near-field setup. The laser beam was split into broadband pump (500 to 700 nm) and probe (700 to 1050 nm). The spectra of the 10 fs Ti:Sapphire beam, the pump beam and the probe beam are shown. Inset: absorbance spectra of WSe2 slab. Two exciton resonance A and B are shown together with the wavelength of pump and probe beam. b) Snapshots of near field signal intensity \( |S|^2 \) with different time delays on WSe2 wedge. c) Line profile of exciton polariton (horizontal axis) with different delay time (vertical axis). The onset of the exciton fringe is shown as the white dashed line. The group velocity can be estimated by the right triangle, which illustrates the distance \( \Delta x \) the fringe has moved in time \( \Delta t \). (Adopted with permission.[130] Copyright 2019, AAAS).
was realized in WSe2. The renormalization was a 40–80 meV splitting between longitudinal exciton and the corresponding transverse exciton, which is large comparing to the typical value of 2 to 3 meV caused by exciton-light coupling in bulk semiconductors.\[132\] The large renormalization, being the cause of slow EP group velocity, is a manifestation of the growing strong Coulomb interaction in TMD-layered materials as they are thinned down.

The ultrafast near-field measurements of phonon polaritons in h-BN by Yoxall et al.\[30\] reveals a propagation mode with negative phase velocity and a group velocity down to 0.002 c using \(\approx 100 \) fs pulses. In a homodyne scheme (Figure 10a), \( E_{\text{ref}}E_{\text{nl}} \), i.e., the multiplication of reference with the near-field signal is measured. The reference beam amplifies the near-field signal and provides a time-resolved measurement, yielding the time profile of the electric field of polaritons. In this setup, polaritons are launched by the edge of gold film, propagate in a h-BN slab, and picked up by the AFM tip. By a line scan across the gold edge, a 2D mapping of signal intensity with respect to position and delay time \( \tau \) can be obtained (Figure 10c). When \( \tau \) is close to 0, the bright fringes show the light scattered directly from the tip \( E_{\text{dir}} \). As \( \tau \) increases, the hyperbolic polaritons that reach the tip can be observed \( E_{\text{HP}} \). A 2D Fourier transform (2DFT) of Figure 10c produces the frequency–momentum relation. Figure 10d is the 2DFT of one of the position–time delay mappings, in which the fundamental hyperbolic polaritonic mode \( (M_0) \) are shown with other modes. From the dispersion, it can be observed that the group velocity, \( \partial \omega / \partial |k| \), is negative. Figure 10d is the real part of the inverse 2DFT of the modified dispersion with only \( M_0 \) preserved, which is done by applying a Hanning window to filter out all other modes. A clear wave packet propagation from gold edge is shown. The inset is a zoom-in of the wave packet. While the wave packet travels from top-left to bottom-right in Figure 10d, the equal-phase line moves from bottom-left to top-right. That is to say, the phase velocity and the group velocity travel in opposite directions. This opposite motion can further be corroborated by the right panel of Figure 10d, which are several succeeding time slices in the left panel.

It is worth noting that the slow group velocity down to 0.002c and the negative phase velocity only appear in one of the two hyperbolic frequency regions of h-BN. The one with negative phase velocity, HPI, is found between 760 and 825 cm\(^{-1}\). There is also HPII, found in 1370 and 1610 cm\(^{-1}\), has group velocity \( \approx 0.027c \) and positive phase velocity. Figure 10c is a position–time mapping of HPII.

4. Outlook

In this report, we have reviewed recent progress on ultrafast near-field characterization of polaritonic materials. Still a nascent field, ultrafast near-field nano-optics has already demonstrated exceptional ability to investigate time-resolved phenomenon at fundamental time and length scales of polaritonic excitations. Here we underscore important current trends that may lead to the advances of future ultrafast plasmonic, or more generally, polaritonic circuits. We also conclude the important challenges that remain to be addressed in the field of s-SNOM.
1. The realization of the “lab-on-chip” experiments of polaritonic circuits is within reach.\cite{133} For example, a plasmonic circuit\cite{134} where sources, modulators, and detectors are integrated into a single chip will yield a compact experimental layout. Tip launched or detected polaritonic waves can serve as a key component for exciting or probing the local element to facilitate the design phase of such experiments. This idea is schematically illustrated in Figure 11.

2. Concept of photonic crystals,\cite{135} metamaterials,\cite{115} meta-surfaces,\cite{136} and recently proposed moiré photonic architectures\cite{137} can be incorporated into the polaritonic chip. Polariton waves modulated via photonic bandgap engineering or subdiffractional metallic composites can enable many intriguing functionalities such as extraordinary optical transmission and light trapping in functional devices such as thin film solar cells.

3. Using artificially induced TM-TE mode splitting as effective spin-orbit coupling,\cite{138} one-way propagating edge modes can be constructed. Combined with new chiral phenomenon which predicted to be nondissipative, the low dissipation chiral polaritons\cite{120,139} can yield important quantum phenomenon such spin-dependent polariton–polariton interaction.

4. The loss in polaritronics can in principle be compensated by parametric amplification,\cite{140,141} photo-active plasmonic cavities, or optical gain media.\cite{142,143} For example, integrated quantum wells with gain-enhanced antenna structures enables sustained polaritonic on-chip devices.

5. Grating structures, or subwavelength antennas\cite{144} which direct the polaritons and control the re-emission will serve as important modulation methods for ultrafast electronics. For example, fast scanning of the laser beam using electrical means can be utilized to redirect or reshape of the beam profile, where the control of group velocity and phase front play a pivotal role.

6. Nonlinear media\cite{145} or mode conversion can modify group velocities in the low range, which are required for sensitive spectroscopy and nonlinear light conversion.

In conclusion, ultrafast s-SNOM provides a versatile platform for the study of polaritons, especially their spatiotemporal behaviors under photoexcitation. New approaches to and applications of light manipulations via tunable polaritonic modes can be advanced with the further development of near-field techniques such as tip-engineering and ultrafast pulse shaping. Being a nascent field, ultrafast near-field polaritons show great potentials in exploring the nano world.

Acknowledgements

Z.H.Y., S.H.X., and D.B.H. contributed equally to this work. The authors thank Alexander S. McLeod, Siyuan Dai, Wenjie Wang (Ithatron Optics), Andrea Huber (Neaspec, Inc.), G. Lawrence Carr, Hans A. Bechtel, Michael Martin, Haidan Wen, S. T. Chui, and Gregory O. Andreev for their helpful discussions. Z.H.Y. thanks the support of the Advanced Light Source Doctoral Fellowship in Residence Program, funded by DOE Office of Science under contract no. DE-AC02-05CH11231. Q.D. and D.H. thank the support from the National Basic Key Research Program of China (No. 2015CB9324000) and the National Natural Science Foundation of China (No. 11704085).

Conflict of Interest

The authors declare no conflict of interest.