Control of semiconductor emitter frequency by increasing polariton momenta

Yaniv Kurman[1], Nicholas Rivera[2], Thomas Christensen[2], Shai Tsesses[1], Meir Orenstein[1], Marin Soljačić[2], John D. Joannopoulos[2] and Ido Kaminer[1,2]

[1]Department of Electrical Engineering, Technion, Israel Institute of Technology, 32000 Haifa, Israel

[2]Department of Physics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge 02139, Massachusetts, USA

Light emission and absorption is fundamentally a joint property of both an emitter and its optical environment. Nevertheless, because of the much smaller momenta of photons compared to electrons at similar energies, the optical environment typically modifies only the emission/absorption rates, leaving the emitter transition frequencies practically an intrinsic property. We show here that surface polaritons, exemplified by graphene plasmons, but also valid for other types of polaritons, enable substantial and tunable control of the transition frequencies of a nearby quantum well—demonstrating a sharp break with the emitter-centric view. Central to this result is the large momenta of surface polaritons that can approach the momenta of electrons, and impart a pronounced nonlocal behavior to the quantum well. This work facilitates non-vertical optical transitions in solids and empowers ongoing efforts to access such transitions in indirect bandgap materials, e.g. silicon, as well as enriches the study of nonlocality in photonics.

Our understanding of light-matter interactions has been instrumental to a range of scientific and technological breakthroughs. At the heart of the theory of light-matter interactions lies the quantized nature of electronic transitions in matter (e.g. atoms, molecules, solids, and quantum dots and wells). Indeed, this notion of quantization is critical to the understanding of both light emission and absorption, enabling key photonic technologies ranging from lasers [1] and LEDs [2] to CCD photodetectors [3] and solar cells [4].

By modifying the local optical environment of an emitter, the relative amplitude of the various transitions can be controlled with great fidelity through the Purcell effect [5-7]. The Purcell effect modifies the emission/absorption spectrum by enhancing the rate of certain frequencies and inhibiting others (i.e., it alters the local density of photonic states, e.g., [8,9]). The transition frequencies themselves, however, are generally perturbed only very weakly by the electromagnetic environment through effects like the Lamb shift when staying in weak coupling regimes [10-15], becoming, in effect, a fixed property of the emitter. Even in solids, where the electronic band structure contains a continuous range of electron energies, the absorption and emission frequencies are generally considered fixed by the electronic bandgap and the quasi Fermi levels. The fixed transition frequencies are due to the small momentum of the photons relative to that of the electrons, which effectively excludes all purely optical non-vertical/indirect transitions, although indirect processes such as phonon-mediated transitions can occur but are relatively inefficient (e.g., [4]). The impact of the small photon momentum reappears in diverse guises throughout the study of light-matter interaction [16]. A pertinent question, then, is whether this large momentum difference can be bridged efficiently by directly increasing the momentum of the optical component without requiring an intermediate process.

Here we show that polaritons in thin films or two-dimensional (2D) materials carry momentum large enough to approach the momenta of electrons in solids. Consequently, they enable tunable control over the absorption and emission frequencies in solids without altering

their intrinsic electronic properties. In particular, we analyze a graphene monolayer supporting plasmon polaritons adjacent to a GaAs/InGaAs/GaAs quantum well (QW) and consider the QW intersubband transitions (Figure 1a). We obtain tunable frequency shifts on the order of tens to hundreds of meV in the spectrum of the emitted and absorbed polaritons, as well as substantial rate enhancements.

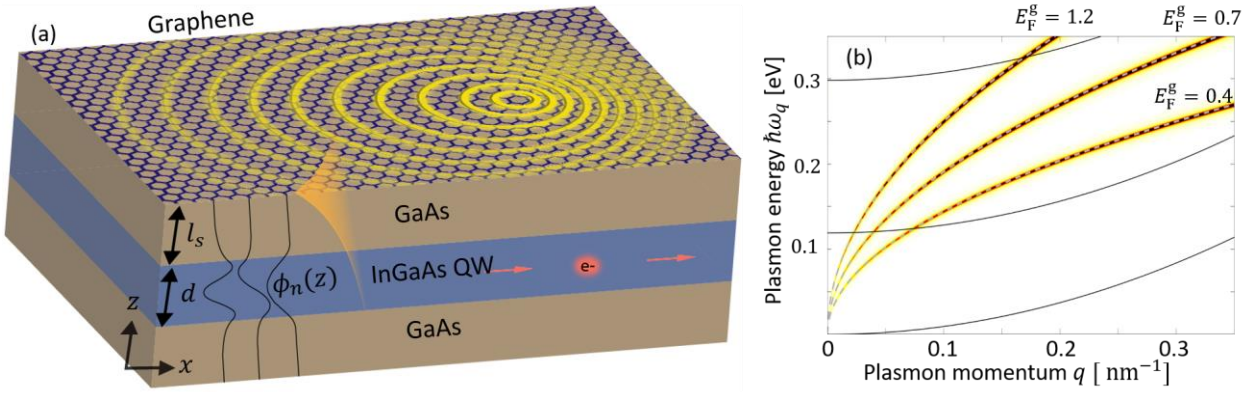

Figure 1: Illustration of the coupling of graphene plasmon polaritons to a quantum well (QW). (a) Scheme of the structure—a graphene layer situated above the QW. Plasmon polariton wavefronts supported by the graphene layer are illustrated by yellow waves. The z profiles of the QW eigenmodes $\phi_n(z)$ (black) and of the plasmon field (yellow) are illustrated on the left. (b) The plasmon dispersion for the structure in Figure 1a for graphene Fermi levels of $0.4, 0.7$ and 1.2 eV. The colormap presents the imaginary part of the reflection coefficient using the local RPA model [17], overlaid with the Drude model dispersion curves (dashed). The match between the curve and the colormap demonstrates the Drude model validity. In black, we plot the QW electronic subbands: the large plasmonic momenta enable non-vertical transitions between QW subbands. E.g., an electron initially at rest can undergo transitions corresponding to the intersections between the plasmonic and the QW dispersions.

Increasing the momentum of photons through polaritonic materials and approaching the range of electron momenta in solids is expected to break one of the most basic notions in optics – that the optical material response is *local*. Optical transitions facilitated by free space photons are nearly direct, owing to their small momentum: as a consequence, the associated energy change of the transitioning electron is fixed by the photon frequency—while its momentum change is negligible. As a result, the transition dynamics of the electron can be accurately captured by a *local* dielectric function that depends only on the transition

frequency (ω) and is independent of the photon momentum. In contrast to this, several studies have revealed the breakdown of the local paradigm, in deeply nanoscopic plasmonic structures [18-21], whose behavior necessitates an account of the material *nonlocal* response (generally described by the spatial dispersion of $\epsilon(\boldsymbol{r}, \boldsymbol{r}', \omega)$ that includes two space coordinates $\boldsymbol{r}, \boldsymbol{r}'$). In translationally invariant systems, where $\epsilon(\boldsymbol{r}, \boldsymbol{r}', \omega) = \epsilon(\boldsymbol{r} - \boldsymbol{r}', \omega)$, nonlocal response is equivalent to an explicit momentum ($\boldsymbol{q}$) dependence of the dielectric function, i.e., $\epsilon(\boldsymbol{q}, \omega)$ [22-26]. Recent experiments probing intrinsic nonlocalities in graphene plasmonics [27] and effective nonlocalities in metamaterials [28] further highlight the emerging primacy of nonlocality. An intriguing question related to these observations is whether it is possible to "induce" a nonlocal response or control the intrinsic nonlocality of a material by coupling it to an adjacent polaritonic material.

We find that the QW exhibits a pronounced nonlocal response when interacting with the large-momentum polariton (the optical field)—despite the fact that each system, considered separately, can be accurately described by a local optical response. Thereby, the coupled system enters a qualitatively new light-matter interaction regime beyond the local optical response and therefore also beyond the dipole approximation. We find two particularly interesting manifestations of this nonlocality: a Doppler effect for emission and absorption frequencies, and an intrinsic cutoff frequency for absorption. The platform proposed in this work enables several new opportunities in both fundamental and applied sciences: at a fundamental level, the controlled study of basic conceptual questions in nonlocality; at an applied level, opportunities range from improved solar cells, enabled by broader absorption spectra, to tunable solid-state LEDs and lasers, enabled by an effectively tunable bandgap.

Results

Theory of graphene-QW interactions

The system of interest, sketched in Figure 1a, is a QW of thickness d that may emit or absorb a plasmon, confined to an adjacent graphene layer, which, in turn, is separated from the QW active area by a spacer of thickness l_s =10 nm. The graphene sheet supports plasmons of momenta far greater than that of free space photons (see Figure 1b): consequently, these plasmons enable non-vertical transitions between electronic subbands of the QW (indicated as solid black lines in Figure 1b). We will now describe the physical properties of each of these individual components as well as the required characteristics for their interactions.

The graphene layer is the propagation medium of the graphene plasmons, whose electromagnetic field dominates the optical interaction with the QW. The electromagnetic field of a graphene plasmon is a highly confined evanescent wave supported by the sloshing of charge density in the graphene sheet. Graphene plasmons are especially attractive for our purpose thanks to the versatility of graphene's conductivity and Fermi level (adjustable via chemical or electrical doping in the graphene layer) [17,29,30], as well as its potentially low-loss extreme light confinement, i.e., wavelength shrinkage to lengths hundreds of times smaller than that of a free space photon [27,31-33]. The confinement factor, defined as $\eta_0 = \frac{c}{v_p} = \frac{\lambda_0}{\lambda_{pl}}$, represents in dimensionless terms the reduced plasmon phase velocity v_p relative to the speed of light in vacuum c—or, equivalently, the reduced plasmon wavelength λ_{pl} relative to the free space value λ_0.

In this work, we exploit a hallmark of graphene plasmonics, namely that η_0 can greatly exceed unity [27,33,34]. Consequently, an electric field mode $\boldsymbol{E}_q$ of a plasmon confined to the *xy* plane with wave vector $\boldsymbol{q}$ (momentum $\hbar\boldsymbol{q}$) can be simplified to:

$$\boldsymbol{E}_q(\boldsymbol{r}) \propto e^{-q|z|} e^{-i\boldsymbol{q}\cdot\boldsymbol{\rho}} (\hat{q} \pm i\hat{z}) \,, \tag{1}$$

where $\boldsymbol{\rho}$ is the position in the *xy* plane. The graphene plasmon dispersion relation can be expressed within the intraband approximation (i.e. by a Drude-like model, see methods) with intrinsic decay time of $\tau = 0.2$ ps [29]:

$$\omega_q \simeq \sqrt{\frac{4\alpha}{\epsilon_\mathrm{s}+1}\frac{E_\mathrm{F}^\mathrm{g}}{\hbar}cq} = v_\mathrm{p}(q)q \equiv \frac{cq}{\eta_0(q)}, \tag{2}$$

where ϵ_s is the spacer permittivity (taken 10.6 for GaAs), α is the fine structure constant, and $\hbar$ is the reduced Planck constant. E_F^g is the graphene Fermi level, which can be tuned by electrostatic gating or chemical doping to control the confinement factor and the plasmon dispersion.

Representing the dispersion through the (momentum–dependent) confinement factor $\eta_0(q)$ allows a direct extension of the key concepts discussed here to any surface polaritonic system (e.g. phonon polaritons [35-37], other families of polaritons [38,39], and their heterostructures [33,40]), all of which can be used to realize the results presented in this work. An important challenge pertinent to all of these high-confinement polaritons is the need to couple the polaritons in or out of the supporting material. Various methods facilitate such coupling: for example, plasmons hosted by the graphene sheet can be out-coupled by a SNOM tip [31] or by a periodic grating [41,42]—alternatively, they could be emitted or absorbed by any nearby quantum emitter, constrained only by the conservation of energy and momentum (Figure 1b).

We demonstrate the emission and absorption of plasmons with a 12 nm GaAs/InGaAs/GaAs QW, wide enough to support three electron subbands. The control over the energy distribution of the QW electrons can be accomplished by chemical doping, electrically, or optically. The distribution of the electrons is described by an effective

temperature T_{eff} and a QW quasi Fermi level $E_{\mathrm{F}}^{\mathrm{QW}}$ (see e.g., Ref. [43]), which are a function of the electronic excitation mechanism (e.g. the frequency and power of the driving laser in the case of optical excitation), and depend also on different charge relaxation rates. Under standard parameters, a simple single-particle treatment has the energies and the wave functions of the charge carriers as

$$E_{n\boldsymbol{k}} = E_n + \frac{\hbar^2 \boldsymbol{k}^2}{2m^*}\,, \tag{3a}$$

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{\rho}}\phi_n(z - z_0) \tag{3b}$$

with $n = 1, 2, 3$ for the three QW electron subbands. E_n and $\phi_n(z)$ represent the QW energy minima and out-of-plane wave function [44] (Figure 1a), z_0 is the location of the QW center, and m^* is the effective mass of charge carriers. $\boldsymbol{k}$ is the electron wave vector in the *xy* plane, in which the charge carriers behave like a 2D free electron gas, resulting in parabolic subbands.

The coupling between plasmons and the QW charge carriers can be described by macroscopic quantum electrodynamics (originally discussed in the context of atomic systems, e.g., [34, 45-47]). The key steps are presented below, with additional technical details deferred to the SM Section 1. The electron–photon interaction is governed by the interaction Hamiltonian (in the Weyl gauge, i.e., with zero scalar potential)

$$H_{\mathrm{int}} = \frac{e}{2m^*}[\boldsymbol{p}\cdot\boldsymbol{A}(\boldsymbol{r}) + \boldsymbol{A}(\boldsymbol{r})\cdot\boldsymbol{p}], \tag{4a}$$

$$A_i(\boldsymbol{r}) = \sqrt{\frac{\hbar}{\pi\epsilon_0}}\int_0^\infty \mathrm{d}\omega\,\frac{\omega}{c^2}\int \mathrm{d}\boldsymbol{r}'\,\sqrt{\mathrm{Im}[\epsilon(\boldsymbol{r}',\omega)]}\,G_{ij}(\boldsymbol{r},\boldsymbol{r}',\omega)\hat{f}_j(\boldsymbol{r}',\omega) + \mathrm{h.c.}, \tag{4b}$$

where e is the QW carrier's charge, $\boldsymbol{p}$ its momentum operators, and $\boldsymbol{A}$ the vector potential operator of the electromagnetic field. The field quantization of the vector potential follows by expansion in annihilation (creation) operators $\hat{f}_i(\boldsymbol{r}',\omega)$ ($\hat{f}_i^{\dagger}(\boldsymbol{r}',\omega)$) that describe a dipole

excitation at position $\boldsymbol{r}'$ and with frequency ω. The vector potential operator incorporates the details of the optical environment (e.g., knowledge of the supported plasmon modes) through the imaginary part of the permittivity ϵ and the dyadic Green's function $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}', \omega)$. The dyadic Green's function components $G_{ij}(\boldsymbol{r}, \boldsymbol{r}', \omega)$ represent the ith component of the electromagnetic field at position $\boldsymbol{r}$ due to a time-harmonic point dipole at position $\boldsymbol{r}'$ oriented along the jth direction. An explicit expression for the Green's function of the system under consideration is given in the SM Eq. (S2).

The use of the dyadic Green's function allows a full account of the impact of intrinsic material losses. The impact of the losses is manifested in the spatial Fourier transformed $G(\boldsymbol{q}, \omega)$, whose amplitude is proportional to the imaginary part of the reflectance (plotted in Figure 1b), and therefore incorporates the plasmonic dispersion (Eq. (2)) and broadening. Furthermore, the Green's function also comprises the full spatial dependence of the mode $e^{-i\boldsymbol{q}\cdot\boldsymbol{\rho}-q|z|}$, which in our case includes the full momentum dependence $e^{-i\boldsymbol{q}\cdot\boldsymbol{\rho}}$ along the QW extended direction, and consequently goes beyond the dipole approximation that is unjustified in our settings. For this reason, we enter a regime that enables non-vertical transitions in the QW dispersion diagram, as shown in Figure 2a,b. We see that the full exploitation of the polaritonic momentum in electronic transitions is attainable in non-atomic, extended emitters, such as QWs.

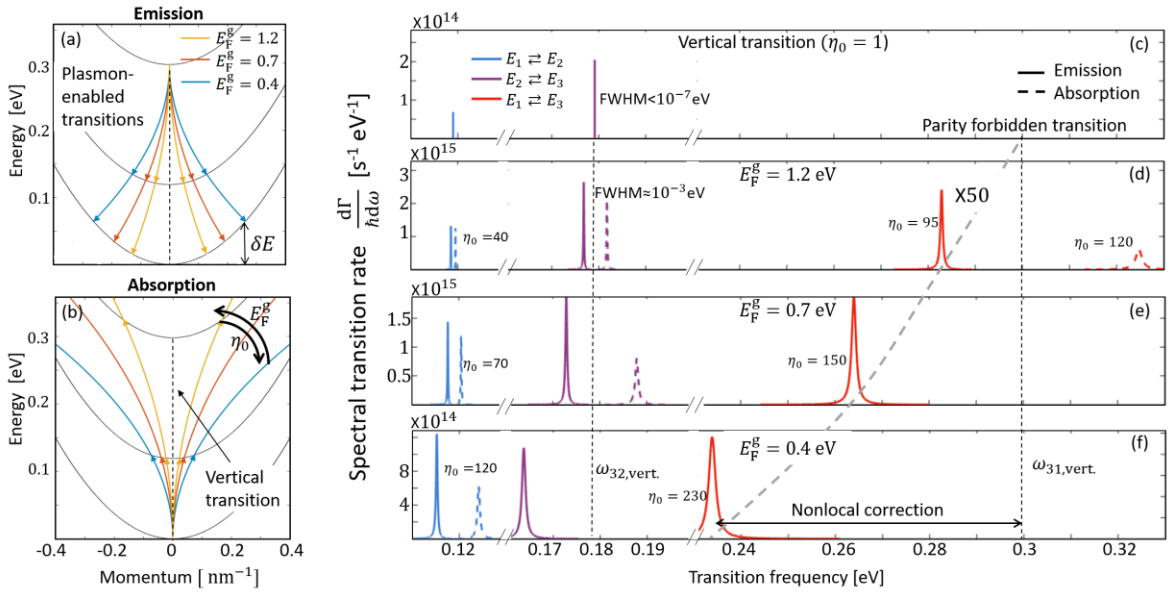

Figure 2: Control of the emission and absorption frequencies of a quantum well by tuning the graphene Fermi energy. (a,b) The concept of the controllable transition frequencies. Plasmonic emission (a) and absorption (b) between the QW subbands (black) for different graphene Fermi levels (blue $E_F^g = 0.4$, red 0.7, yellow 1.2 eV). The transition frequency shift is marked by δE. (c-f) Total rates of emission (solid) and absorption (dashed) per charge carrier of the different transitions (blue $E_1 \rightleftarrows E_2$, purple $E_2 \rightleftarrows E_3$, red $E_1 \rightleftarrows E_3$) for several values of the graphene Fermi level when the electron is initially at the bottom of the subband. Confinement factors of the corresponding transitions are indicated. Dashed black lines represent the local-response resonance frequencies (i.e. vertical transitions).

Transition frequency shift due to plasmon momentum

In this section, we analyze the transition spectra of plasmons emitted or absorbed by the QW and show how the large plasmon momentum alters the QW transition frequencies. The first step is to look at the transition rate between QW states due to the emission (em) or absorption (ab) of a plasmon in the considered structure. In this structure, the transition rates fall within the weak coupling paradigm; strong coupling will only become potentially relevant for sub-nanometre spacers. Hence, it is justified to obtain the rates directly from Fermi's golden rule [48]:

$$\Gamma_{\text{ab}}^{\text{em}} = \frac{2\pi}{\hbar} \sum_{\mathbf{k}_\text{i}, \boldsymbol{k}_\text{f}} f(\boldsymbol{k}_\text{i}) \left| \langle n_\text{f}, \boldsymbol{k}_\text{f}, n_{\boldsymbol{q}} \pm 1 | H_\text{int} | n_\text{i}, \boldsymbol{k}_\text{i}, n_{\boldsymbol{q}} \rangle \right|^2 \delta\left(E_{n_\text{i}\boldsymbol{k}_\text{i}} - E_{n_\text{f}\boldsymbol{k}_\text{f}} \mp \hbar\omega_q\right), \qquad (5)$$

where $f(\boldsymbol{k}_\mathrm{i}) = \left(1 + \exp\left[\left(\frac{\hbar^2 k_\mathrm{i}^2}{2m^*} - E_\mathrm{F}^\mathrm{QW}\right)/k_\mathrm{B}T_\mathrm{eff}\right]\right)^{-1}$ is the Fermi–Dirac distribution as a function of an electronic momentum $\hbar\boldsymbol{k}_\mathrm{i}$ and the final state is assumed to be vacant. Subscripts i and f denote the initial and final states of the electron, while $n_{\boldsymbol{q}}$ is the number of plasmon quanta (increased/decreased for emission/absorption). Although the calculation can be done for any $n_{\boldsymbol{q}}$, for simplicity we focus on the cases where $n_{\boldsymbol{q}} = 0$ when considering spontaneous plasmon emission and $n_{\boldsymbol{q}} = 1$ when considering plasmon absorption. The results presented are normalized to yield transition rates *per charge carrier*, i.e. such that $1 = \frac{A}{(2\pi)^2}\int d\boldsymbol{k}_\mathrm{i} f(\boldsymbol{k}_\mathrm{i})$ for a slab of area A.

The matrix element in Eq. (5) that specifies the interaction amplitude between graphene plasmons and the QW is approximately given by (see SM Section 1):

$$\left\langle n_\mathrm{f}, \boldsymbol{k}_\mathrm{f}, n_{\boldsymbol{q}} \pm 1 \middle| H_\mathrm{int} \middle| n_\mathrm{i}, \boldsymbol{k}_\mathrm{i}, n_{\boldsymbol{q}} \right\rangle \propto \overbrace{\langle n_\mathrm{f}|\mathrm{e}^{-qz}\partial_z|n_\mathrm{i}\rangle}^{\text{(i)}} \overbrace{\left\langle \boldsymbol{k}_\mathrm{f} \middle| \mathrm{e}^{\pm i\boldsymbol{q}\cdot\boldsymbol{\rho}} \middle| \boldsymbol{k}_\mathrm{i} \right\rangle}^{\text{(ii)}}. \tag{6}$$

The first term (i) gives the *z*-dependent interaction amplitude due to overlap between the evanescent tail of the plasmon and the charge carrier state inside the QW (Figure 1a), where the ∂_z derivative arises from the momentum operator. The term (ii) ensures momentum conservation in the *xy* plane, forcing $\boldsymbol{k}_\mathrm{i} \pm \boldsymbol{q} = \boldsymbol{k}_\mathrm{f}$, which appears as a delta function in the lossless case or a Lorentzian when including plasmonic losses. Meanwhile, the frequency-dependent delta function, $\delta\left(E_{n_i \boldsymbol{k}_i} - E_{n_\mathrm{f}\boldsymbol{k}_\mathrm{f}} \mp \hbar\omega_q\right)$ in Eq. (5), enforces energy conservation. The allowed transition frequencies associated with the emission or absorption of a plasmon are consistent with the combined energy-momentum conservation, and can be determined by substituting the dispersion relations of both the plasmon and the QW electron, Eqs. (2) and (3a), into these conservation laws. Figure 2a,b illustrates the tunable transition frequencies that

follow from these conservation laws (for an electron initially at rest), with the tunability courtesy of the gate tunable plasmon dispersion.

Figure 2c-f displays the tunability in the transition frequencies when increasing the momentum of the plasmon, as expected from Figures 2a,b. Specifically, the absorption and emission frequencies show blue and red shifts respectively, breaking the symmetry between them with a split as wide as $100\ \mathrm{meV}$. Comparing panels c-f illustrates how the frequency shifts increase as the Fermi level of graphene decreases, which corresponds to increasing the plasmon momentum, i.e., increasing the confinement factor. In addition to a frequency shift, the QW transitions also exhibit spectral broadening due to plasmonics losses, with an increase in the FWHM of the spectral lines of up to $10\ \mathrm{meV}$, which will be further broadened as discussed below by a Doppler broadening effect. The broadening of the FWHM could be understood through Figure 1b from the width of the plasmonic dispersion and its intersection with the electronic dispersion line.

In line with previous findings [49, 50], Figure 2 also shows that the total transition rates (the integrated spectra) are significantly enhanced for interaction with plasmons that have large confinement factors (η_0). As a result of the large Purcell factor, we find that the radiative plasmon transitions dominate all other radiative processes (see SM Figure S5), such as QW emission into free space photons. In contrast with previous findings, Figure 2 shows that there is a difference between the absorption rate and the emission rate. This change in spectral response between absorption and emission, and the large change compared to the local-response (i.e. vertical transitions) plotted in Figure 2a,b underscore the impact of the induced nonlocality in the QW response. Moreover, we find that the spacer thickness affects the enhancement of all transitions, and limits the range of confinement factors that enhance the transitions. The reason lies in the fact that large confinements shorten the plasmon tail in the z

direction to a characteristic length of $\frac{c/\omega_q}{\eta_0}$, as seen in Eq. (6) term (i) (more details in SM Figure S8).

Figure 3 summarizes the large dependence of the absorption frequencies, emission frequencies, and transition rates of distinct intersubband transitions on the plasmon confinement factor. In this representation, the gradual change from the free space photon transition rates and spectra ($\eta_0 = 1$) to the high momentum plasmon transition rates and spectra (i.e., high values of η_0) is apparent. The first notable feature in Figure 3 is the strong enhancement of transition $E_1 \rightleftarrows E_3$, which due the parity symmetry of the wavefunctions, has a vanishing dipole moment but a non-vanishing quadrupole moment in the z direction. In the absence of coupling to plasmons, this transition is much slower than the $E_1 \rightleftarrows E_2$ transition on account of the weak variation of the electromagnetic field that practically forbids any beyond-dipole transition in the quantum well. However, in the presence of plasmons, the $E_1 \rightleftarrows E_3$ transition can be enhanced by up to seven orders of magnitude (here achieved for confinement factor of 85). This transforms the $E_1 \rightleftarrows E_3$ transition from effectively forbidden to one with a rate exceeding even the conventional, non-forbidden transitions (e.g., $E_1 \rightleftarrows E_2$) of the QW. Interestingly, we find that the rates of both kinds of transitions are not merely influenced by the out-of-plane fields and field gradients; they are also strongly influenced by the in-plane plasmon momentum, which emphasizes the importance of the complete nonlocal description. The solid black lines in Figure 3a presents a comparison with the conventional local calculation of the transition rates, where we neglect the finite plasmon momentum by approximating $e^{i\boldsymbol{q}\cdot\boldsymbol{\rho}} \simeq 1$ in Eq. (6) that yields a conventional dipole $\boldsymbol{d}\cdot\boldsymbol{E}$ interaction term. The deviation between the curves in Figure 3a shows that the local calculation fails for the $E_1 \rightleftarrows E_3$ transition even at relatively low confinements, further underscoring the importance of going beyond the dipolar approximation (further details in SM Section 6).

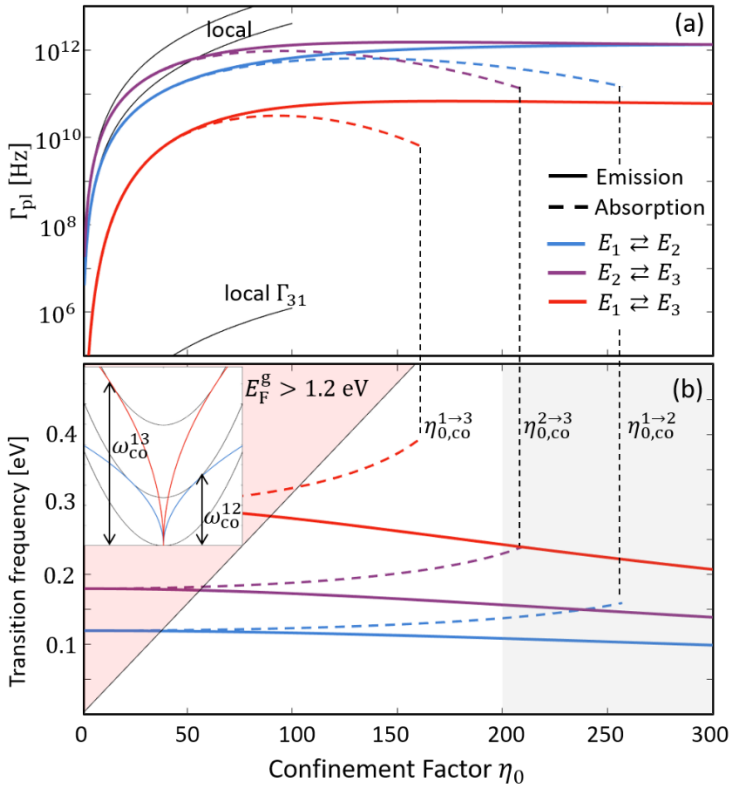

Figure 3: Confinement factor effect on plasmonic transitions. Transition rates (a) and frequencies (b) as a function of plasmon confinement factor (dashed lines, absorption; solid lines, emission) for the three possible transitions (blue $E_1 \rightleftarrows E_2$, purple $E_2 \rightleftarrows E_3$, red $E_1 \rightleftarrows E_3$). Furthermore, the corresponding local calculation is shown in black (equal for absorption and emission). In (b) the red shaded area marks graphene Fermi level above 1.2 eV, and the gray shaded area marks confinements where the Drude model fails. Inset: the QW subbands overlaid by the plasmon dispersion for two absorption transitions, with the Fermi levels having the cutoff values $\mathrm{E}_\mathrm{F}^\mathrm{g} = \mathrm{E}_{\mathrm{F,co}}^\mathrm{g}$ for the second and third subbands. All curves are for the plasmonic lossless case, with a single electron initially at rest ($\boldsymbol{k}_\mathrm{i} = 0$). For the same plots as a function of E_F^g see SM Figure S7.

Another interesting effect revealed by rigorously accounting for the large plasmon momenta is the appearance of a cutoff confinement factor $\eta_{0,\mathrm{co}}$—or, equivalently, a cutoff Fermi level—for the QW absorption (Figure 3b). In particular, above this confinement there are no absorptive transitions that satisfy the conservation of energy and momentum laws previously discussed. The reason is straightforward and immediately appreciable from Fig. 2b: at sufficiently low Fermi levels, the plasmon dispersion bends so rapidly, that it "misses" its intersection with the next subband (inset in Figure 3b). For an electron initial at rest, we obtain

a closed-form expression for the cutoff graphene Fermi level, $E_{\mathrm{F,co}}^{\mathrm{g}} = \frac{\epsilon_s+1}{\alpha}\frac{\left(\frac{2}{3}\Delta E\right)^{3/2}}{\sqrt{m^* c^2}}$ where $\Delta E = E_{n_\mathrm{f},\mathbf{0}} - E_{n_\mathrm{i},\mathbf{0}}$ is the QW energy level difference. Interestingly, the cutoff Fermi level corresponds to a maximum confinement factor $\eta_{0,\mathrm{co}} = \frac{3}{8}\frac{\hbar\omega_q}{\Delta E}\sqrt{\frac{3m^* c^2}{2\Delta E}}$, which is found within experimental range thanks to a combination of the electron's small effective mass and the graphene's large confinement factor. When $E_{\mathrm{F}}^{\mathrm{g}} < E_{\mathrm{F,co}}^{\mathrm{g}}$, absorption between the two relevant subbands is forbidden in any frequency, and when $E_{\mathrm{F}}^{\mathrm{g}} > E_{\mathrm{F,co}}^{\mathrm{g}}$, the absorption channels open (where the higher frequency absorption channel is significantly slower – see SM Figure S5). In the same spirit, additional absorption channels are likely to exist at higher frequencies due to other intersubband transitions beyond the parabolic band approximation.

To further investigate the frequency shift due to the plasmon's momentum, we examine the transitions of a charge carrier *not* initially at rest ($\boldsymbol{k}_\mathrm{i} \neq 0$) and reveal a Doppler effect. We denote the angle θ between a plasmon at direction $\widehat{\boldsymbol{q}}$ and the direction of the initial momentum of the charge carrier $\widehat{\boldsymbol{k}}_\mathrm{i}$, so that $\theta = \arccos(\widehat{\boldsymbol{k}}_\mathrm{i} \cdot \widehat{\boldsymbol{q}})$. Figure 4a shows that for the case of plasmon emission, when θ varies from 0 to π, the plasmon frequency can vary by tens of meV. Consequently, the distribution of charge carriers in the QW (i.e., distribution in $\boldsymbol{k}_\mathrm{i}$) alters the frequency bandwidth of each QW transition. We now explain this behavior analytically as a manifestation of a Doppler shift.

According to the Doppler shift, a moving charge observes waves with a shifted frequency that depends on the relative angle θ (as defined above). Importantly, the conventional Doppler effect carries the hidden assumption that the charge momentum is unchanged by the interaction with the wave. The Doppler shift in our case is different from the conventional one since the momentum of the charge can change significantly during the interaction with the plasmon (inset of Figure 4a), i.e., the recoil of the charge carrier alters the

conventional Doppler formula. As expected, the recoil is negligible for low confinement factors (and low plasmon momentum), so that the transition rates are angle-independent, and the transition frequencies follow the conventional Doppler formula:

$$\omega_q(\theta) = \omega_q(\pi/2)\left(1 - \frac{v_\mathrm{i}\cos(\theta)}{v_\mathrm{p}(q)}\right)^{-1}, \tag{8}$$

where $\omega_q(\theta)$ represents the frequency of the plasmon emission\absorption for a specific angle of emission or absorption. The charge carrier velocity is $v_\mathrm{i}(k_\mathrm{i})$ and the plasmon phase velocity is $v_\mathrm{p}(q)$. Equation (8) reveals not only the angle-dependence of frequency but also the graphene Fermi energy dependence since it controls the plasmon phase velocity. While Eq. (8) takes the form of a conventional Doppler shift, we find that for high confinement factors, the recoil of the charge carrier can cause substantial deviations from the conventional Doppler shift, leading to a recoil dependent generalized Doppler shift alongside with highly angle-dependent transition rates (see SM Figure S9).

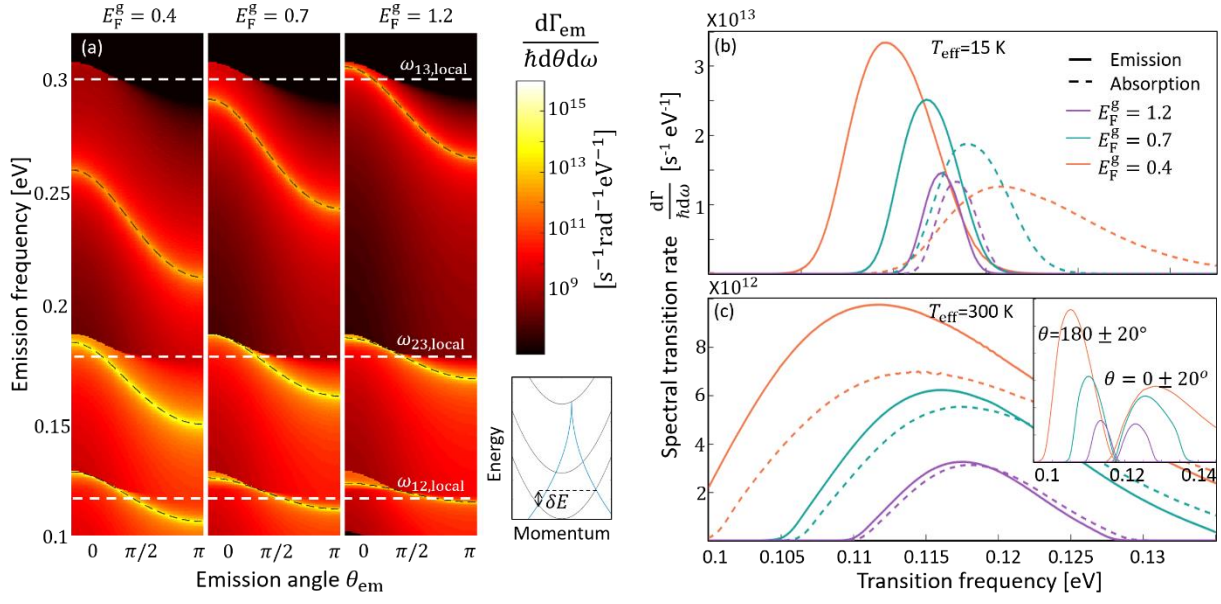

Figure 4: Doppler effect in plasmonic transitions. (a) Angular spectral emission rate for an electron with initial momentum of $k_\mathrm{i} = 0.1\ \mathrm{nm}^{-1}$ and three graphene Fermi level energies $\mathrm{E_F^g} = 0.4$,0.7 and 1.2 eV, illustrating the sharp break with the local description (dashed white). The dashed black lines present the emission frequencies according to our lossless formula that

generalizes the conventional Doppler shift, showing a good match with full Green's function simulations. Inset: QW band-diagram and possible transitions to opposite directions for a nonzero initial electron momentum $\boldsymbol{k}_{\mathrm{i}}$, showing the direction-dependence Doppler frequency shift marked by δE. (b,c) Spectral emission (solid) and absorption (dashed) rate in several graphene Fermi level values for the $E_1 \rightleftarrows E_2$ transition when the effective temperature of the electrons in the QW is (b) $T_{\mathrm{eff}} = 15$ K and (c) $T_{\mathrm{eff}} = 300$ K where the quasi Fermi level is 1.5 meV above the bottom of the initial subband. Inset: Spectral emission rates forward in angles $0 \pm 20°$ and backward in angles $180 \pm 20°$.

Figures 4b,c presents a Doppler broadening of the emission and absorption spectra for transitions between the first and second electronic subbands, for three different graphene Fermi levels $\mathrm{E_F^g}$ and two effective temperatures, calculated by integrating over k_{i}, θ (Eq. 5).Although the known Doppler broadening expression agrees with the results for low confinement factors (see SM Section 3), in high confinement factors the inhomogeneous transition rates in the Doppler effect cause the overall peak to be redshifted, as demonstrated in Figure 4c and its inset. Despite the broadening, at low temperatures (Figure 4b) the separation between the peaks of emission and absorption is observable, with a split exceeding 10 meV for $\mathrm{E_F^g} < 0.7$ eV.

We conclude our discussion of the Doppler shift by noting that the Doppler effect in the system is itself a manifestation of nonlocality, as previously noted by e.g., Landau and Lifschitz [51]: the frequency and rate of plasmon absorption is explicitly dependent on the plasmon wavevector through the angle of absorption θ. Since the absorption rate is directly proportional to the imaginary part of the dielectric function, $\Gamma \propto \mathrm{Im}\, \epsilon(\boldsymbol{q}, \omega)$, it follows that the QW permittivity is wave vector dependent and therefore spatially nonlocal.

Section IV - Discussion

To summarize, we have demonstrated that the large momenta of surface polaritons can be used to control the frequencies of polaritons emitted or absorbed by electrons in a solid-state

emitter. Specifically, the combination of a QW and a plasmon supporting monolayer of graphene enables substantial control over the spectral properties of the system, through graphene's tunable Fermi level. The full transition dynamics manifests new effects such as an absorption cutoff and pronounced Doppler shifts that suggest intriguing possibilities, e.g., bound electron acceleration (see SM Section 7). This control of the transition spectrum also enables tuning the intrinsic optical response and effective permittivity of the system (since $\Gamma \propto \mathrm{Im}\,\epsilon(\boldsymbol{q},\omega)$), spanning over twentyfold when the wave vector and the frequency fulfill the plasmonic dispersion relations (see SM Section 5 and Figure S2). As a result, our work suggests a framework for designing new, tunable, nonlocal metamaterials from local constituents.

The scheme introduced in this paper can be generalized to a host of surface polariton systems interfaced with various solid-state emitters simply by substituting material properties of the polaritons and the electrons, i.e., their dispersions. Other 2D materials and polar dielectric structures supporting phonon polaritons, such as hBN, offer several interesting opportunities in this regard. E.g., the highly frequency-selective nature of phonon polaritons [35-37] may enable additional spectral control over the QW intersubband transitions in addition to the strong enhancements of light-matter interactions offered by such materials [49]. The formalism can also be generalized to intersubband transitions in other QW structures such as in few-layer transition metal dichalcogenides [52] and narrow bandgap materials [53], as well to interband transitions in these materials and in others. Further still, the formalism could be extended to cascade processes in multi-level systems as in [54], thus capturing the complete dynamics of the QW electrons when experiencing multiple consecutive transitions. The limit of our formalism is the strong coupling regime that might occur when a spacer of few atomic layers separates the electronic states and the plasmons, leading to the formation of a new type of polariton composed by the QW intersubband polariton and the graphene plasmon (potentially accompanied by a Lamb-shift-type frequency shift [55]).

The ability to access very large electromagnetic field's momenta could also affect the optical response of indirect bandgap bulk semiconductors [56,57], which are typically constrained in their inability to interact with light near its indirect bandgap due to the large electronic momentum mismatch between valence maximum and conduction minimum. Coupling such materials to a polariton–supporting 2D layer could conceivably provide the momentum necessary to overcome this mismatch—even without altering its electronic band structure (e.g., by introducing defects). A proof-of-concept can be implemented using multilayer $MoSe_2$ as an emitter, which was found to change from direct to indirect depending on the number of layers [58, 59], making this material an ideal candidate for such an experimental proof-of-concept (potentially in a heterostructure with adjacent polariton supporting 2D materials). A particularly exciting realization of this idea would be inducing indirect bandgap transitions in silicon, a material of extreme technological importance. To cross the Brillouin zone with a high momentum polariton, one would typically still want to modify the electronic structure (e.g., with porous silicon [60]). With a combination of band-structure engineering and large momenta polaritons, one could envision designing more efficient silicon solar cells, silicon photodetectors and silicon-based light sources. Such devices, which could be easily integrated on a chip, would greatly contribute to creating simple, robust hybrid electronic-photonic systems.

Methods

The considerations of plasmonic losses

The plasmon dispersion is presented in Figure 1b via the imaginary part of the reflectance (colormap), computed from the full local conductivity of graphene (i.e. including both intraband and interband dispersion) calculated from the random phase approximation (RPA) [17,49]. The dashed lines give the associated intraband approximation, Eq. (2), neglecting losses. The excellent agreement between the two conductivity models illustrates that for the range of parameters considered in this work, the intraband (or Drude) treatment is completely sufficient. This result can be used whenever the Fermi level is larger than the plasmon energy and the influence of both intrinsic graphene nonlocality and interband corrections are negligible [17,29,30]. The intrinsic graphene losses considered in Figure 1b and in our results below reflect an intrinsic decay time of $\tau = 0.2$ ps. This is a reasonable estimate of the decay time attainable in current experimental setups [61], with even higher decay times attainable in under cryogenic conditions or via hexagonal boron nitride encapsulation [33]. For corrections due to a change in graphene losses (different Drude decay times τ), see SM figure S6. In practice, it translates to plasmon propagation lengths on the order of a micron for the Fermi levels and frequencies considered here. More about the plasmonic losses in the Supplementary Material (SM) Section 8.

References

[1] Yariv, A. *Quantum Electronics* 3rd edn (Wiley, New York,1989).

[2] Schubert, E.F. *Light emitting diodes* (Cambridge Univ. Press, 2006).

[3] Theuwissen, A.J.P. *Solid-State Imaging with Charge-Coupled Devices* Vol. **1** (Kluwer Academic, Boston, 1995).

[4] Green, M.A. *Solar Cells: Operating Principles, Technology and System Applications* (Univ. New South Wales, 1998).

[5] Purcell, E.M. Spontaneous emission probabilities at radio frequencies. *Phys. Rev.* **69**, 681–681 (1946).

[6] Kleppner, D. Inhibited spontaneous emission. *Phys. Rev. Lett.* **47**, 233–236 (1981).

[7] Pelton, M. Modified spontaneous emission in nanophotonic structures. *Nature Photon.* **9**, 427–435 (2015).

[8] Vats, N., John, S. & Busch, K. Theory of fluorescence in photonic crystals. *Phys. Rev. A* **65**, 043808 (2002).

[9] Neogi, A., *et al*. Enhancement of spontaneous recombination rate in a quantum well by resonant surface plasmon coupling. *Phys. Rev. B* **66**, 153305 (2002).

[10] Lamb, W.E. & Retherford, R.C. Fine structure of the hydrogen atom by a microwave method. *Phys. Rev.* **72**, 241–243 (1947).

[11] Bethe, H.A. The Electromagnetic Shift of Energy Levels. *Phys. Rev.* **72**, 339 (1947).

[12] Wylie, J.M. & Sipe, J. E. Quantum electrodynamics near an interface. *Phys. Rev. A* **30**, 1185–1193 (1984).

[13] Khitrova, G., Gibbs, H.M., Kira, M.S., Koch, W. & Scherer, A. Vacuum Rabi splitting in semiconductors. *Nature Phys.* **2**, 81-90 (2006).

[14] Zhang, Y. *et al*. Sub-nanometre control of the coherent interaction between a single molecule and a plasmonic nanocavity. *Nature Commun.* **8**, 15225 (2017).

[15] Chang, C.H., Rivera, N., Joannopoulos, J.D., Soljačić, M. & Kaminer, I. Constructing "designer atoms" via resonant graphene-induced lamb shifts. *ACS Photonics* **4**, 3098-3105 (2017).

[16] Cohen-Tannoudji, C., Dupont-Roc, J. & Grynberg, G.*Photons and Atoms: Introduction to Quantum Electrodynamics* (Wiley, 1989).

[17] Garcia de Abajo, F.J. Graphene plasmonics: Challenges and opportunities. *ACS Photonics* **1**, 135-152 (2014).

[18] Scholl, J. A., Koh, A.L. & Dionne, J.A. Quantum plasmon resonances of individual metallic nanoparticles. *Nature* **483**, 421–427 (2012).

[19] Raza, S. *et al*. Blueshift of the surface plasmon resonance in silver nanoparticles studied with EELS. *Nanophotonics* **2**, 131–138 (2013).

[20] Jin, D. *et al*. Quantum-Spillover-Enhanced Surface-Plasmonic Absorption at the Interface of Silver and High-Index Dielectrics. *Phys. Rev. Lett.* **115**, 193901 (2015).

[21] Raza, S. *et al*. Multipole plasmons and their disappearance in few-nanometre silver nanoparticles. *Nature Commun.* **6**, 8788 (2015).

[22] Raza, S., Bozhevolnyi, S.I., Wubs, M. & Mortensen, N.A. Nonlocal optical response in metallic nanostructures. *Journal of Physics: Condensed Matter* **27**, 183204 (2014).

[23] Christensen, T. *From Classical to Quantum Plasmonics in Three and Two Dimensions* (Springer, 2017).

[24] Varas, A., Garcia-Gonzalez, P., Feist, J., Garcia-Vidal, F.J. & Rubio, A. Quantum plasmonics: from jellium models to ab initio calculations. *Nanophotonics* **5**, 409–426 (2016).

[25] Zhu, W. *et al*. Quantum mechanical effects in plasmonic structures with subnanometre gaps. *Nature Commun.* **7**, 11495 (2016).

[26] Fitzgerald, J.M., Narang, P., Craster, R.V., Maier, S.A. & Giannini, V. Quantum Plasmonics. *Proceedings of the IEEE* **104**, 2307–2322 (2016).

[27] Lundeberg, M. *et al*. Tuning quantum non-local effects in graphene plasmonics. *Science*, eaan2735 (2017).

[28] Ginzburg, P. *et al*. Spontaneous emission in non-local materials. *Light Sci. Appl.* **6**, e16273 (2017).

[29] Jablan, M., Buljan, H. & Soljacic, M. Plasmonics in graphene at infrared frequencies. *Phys. Rev. B* **80**, 1–7 (2009).

[30] Jablan, M., Soljacic, M. & Buljan, H. Plasmons in Graphene: Fundamental Properties and Potential Applications. *Proc. IEEE* **101**, 1689–1704 (2013).

[31] Chen, J., *et al*. Optical nano-imaging of gate-tunable graphene plasmons. *Nature* **487**, 77-81 (2012).

[32] Fei, Z. *et al*. Gate-tuning of graphene plasmons revealed by infrared nano-imaging. *Nature* **487**, 82-85 (2012).

[33] Woessner, A. *et al*. Highly confined low-loss plasmons in graphene-boron nitride heterostructures. *Nature Mat.* **14**, 421–425 (2014).

[34] Rivera, N., Kaminer, I., Zhen, B., Joannopoulos, J.D. & Soljačić, M. Shrinking light to allow forbidden transitions on the atomic scale. *Science* **353**, 263-269 (2016).

[35] Hillenbrand, R., Taubner, T. & Keilmann, F. Phonon-enhanced light-matter interaction at the nanometre scale. *Nature* **418**, 159-162 (2002).

[36] Taubner, T., Korobkin, D., Urzhumov, Y., Shvets, G. & Hillenbrand, R. Near-field microscopy through a SiC superlens. *Science* **313**, 1595 (2006).

[37] Dai, S. *et al*. Tunable Phonon Polaritons in Atomically Thin van der Waals Crystals of Boron Nitride. *Science* **343**, 1125–1129 (2014).

[38] Basov, D.N., Fogler, M.M. & Garcia de Abajo, F.J. Polaritons in van der Waals materials. *Science* **354**, aag1992 (2016).

[39] Low, T. *et al*. Polaritons in layered two-dimensional materials. *Nature Mat.* **16**, 182-194 (2017).

[40] Lin, X., *et al*., All-angle negative refraction of highly squeezed plasmon and phonon polaritons in graphene–boron nitride heterostructures. *PNAS* **114**, 6717-6721 (2017).

[41] Thongrattanasiri, S., Koppens, F.H.L., & Garcia de Abajo, F.J. Complete optical absorption in periodically patterned graphene. *Phys. Rev. Lett.* **108,** 047401 (2012).

[42] Brar, V.W., Jang, M.S., Sherrott, M., Lopez, J.J., & Atwater, H.A. Highly confined tunable mid-infrared plasmonics in graphene nanoresonators. *Nano Lett.* **13**, 2541-2547, (2013).

[43] Coldren, L.A. & Corzine, S.W. *Diode Lasers and Photonic Integrated Circuits*. (Wiley, New York, 2012).

[44] Griffiths, D. *Introduction to Quantum Mechanics* (Pearson Prentice Hall 1995).

[45] Glauber, R.J., & Lewenstein, M. Quantum optics of dielectric media. *Phys. Rev. A* **43**, 467–491 (1991).

[46] Scheel, S. & Buhmann, S. Y. Macroscopic QED - concepts and applications. *Acta Physica Slovaca* **58**, 675-809 (2008).

[47] Bitan, G. *Energy Distribution of Single and Two Photon Emitters in Plasmonic Environments*. MSc thesis, Technion IIT (Department of Electrical Engineering, 2014).

[48] Kaminer, I., *et al*. Efficient plasmonic emission by the quantum Čerenkov effect from hot carriers in graphene. *Nature commun.* **7**, 11880 (2016).

[49] Koppens, F.H.L., Chang, D.E. & Garcia de Abajo, F.J. Graphene plasmonics: A platform for strong light-matter interactions. *Nano Lett.* **11**, 3370–3377 (2011).

[50] Rivera, N., Rosolen, G., Joannopoulos, J.D., Kaminer, I., & Soljačić, M. Making two-photon processes dominate one-photon processes using mid-IR phonon polaritons. *PNAS* **114**, 13607 (2017).

[51] Landau, L.D., Lifshitz, E.M. & Pitaevskii, L.P. *Electrodynamics of Continuous Media* Vol. **8** (Elsevier, 2008).

[52] Wang, Q.H., Kalantar-Zadeh, K., Kis, A., Coleman, J.N. & Strano, M.S. Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. *Nature nanotech.* **7**, 699-712 (2012).

[53] McGill, T.C., & Collins, D.A. Prospects for the future of narrow bandgap materials. *Semiconductor Science and Technology* **8**, S1 (1993).

[54] Sloan, J., Rivera, N., Soljačić, M. & Kaminer, I. Tunable UV-emitters through Graphene Plasmonics, *Nano Lett.* **18**, 308-313 (2017).

[55] Törmä, P., & Barnes, W.L. Strong coupling between surface plasmon polaritons and emitters: a review. *Reports on Progress in Physics* **78**, 013901 (2014).

[56] Jung, J., *et al*. Dyadic Green's functions of thin films: Applications within plasmonic solar cells. *Phys. Rev. B* **83**, 085419 (2011).

[57] Trolle, M.L. & Pedersen, T.G. Indirect optical absorption in silicon via thin-film surface plasmon. *Journal of Applied Physics* **112**, 043103 (2012).

[58] Tongay, S. *et al.* Thermally driven crossover from indirect toward direct bandgap in 2D semiconductors: $MoSe_2$ versus MoS_2. *Nano Lett.* **12**, 5576–5580 (2012).

[59] Zhang, Y. *et al*. Direct observation of the transition from indirect to direct bandgap in atomically thin epitaxial MoSe2. *Nature Nanotech*. **9**, 111 (2014).

[60] Malý, P. *et al*. Picosecond and millisecond dynamics of photoexcited carriers in porous silicon. *Phys. Rev. B* **54**, 7929–7936 (1996).

[61] Tan, Y.W. *et al*. Measurement of scattering rate and minimum conductivity in graphene. *Phys. Rev. Lett*. **99**, 246803 (2007).

Acknowledgements

We thank Moriel Hoffman for illustrating Figure 1a.

Statement of author contribution

All authors made significant contributions to the manuscript.