Plasmonic silicon Schottky photodetectors: The physics behind graphene enhanced internal photoemission (F)

Uriel Levy, Meir Grajower, P. A. D. Gonçalves (iD), N. Asger Mortensen (iD), and Jacob B. Khurgin (iD)

COLLECTIONS

F This paper was selected as Featured

View Online   Export Citation   CrossMark

ARTICLES YOU MAY BE INTERESTED IN

The physics and chemistry of the Schottky barrier height
Applied Physics Reviews **1**, 011304 (2014); https://doi.org/10.1063/1.4858400

Metal-semiconductor-metal photodetectors based on graphene/p-type silicon Schottky junctions
Applied Physics Letters **102**, 013110 (2013); https://doi.org/10.1063/1.4773992

Schottky contact surface-plasmon detector integrated with an asymmetric metal stripe waveguide
Applied Physics Letters **95**, 021104 (2009); https://doi.org/10.1063/1.3171937

Plasmonic silicon Schottky photodetectors: The physics behind graphene enhanced internal photoemission

Uriel Levy,[1] Meir Grajower,[1] P. A. D. Gonçalves,[2,3] N. Asger Mortensen,[2,3] and Jacob B. Khurgin[4,a]

[1]*Department of Applied Physics, The Benin School of Engineering, The Center for Nanoscience and Nanotechnology, The Hebrew University of Jerusalem, Jerusalem, Israel*
[2]*Department of Photonics Engineering, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark*
[3]*Center for Nanostructured Graphene, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark*
[4]*Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, Maryland 21218, USA*

Recent experiments have shown that the plasmonic assisted internal photoemission from a metal to silicon can be significantly enhanced by introducing a monolayer of graphene between the two media. This is despite the limited absorption in a monolayer of undoped graphene ($\sim\pi\alpha = 2.3\%$). Here we propose a physical model where surface plasmon polaritons enhance the absorption in a single-layer graphene by enhancing the field along the interface. The relatively long relaxation time in graphene allows for multiple attempts for the carrier to overcome the Schottky barrier and penetrate into the semiconductor. Interface disorder is crucial to overcome the momentum mismatch in the internal photoemission process. Our results show that quantum efficiencies in the range of few tens of percent are obtainable under reasonable experimental assumptions. This insight may pave the way for the implementation of compact, high efficiency silicon based detectors for the telecom range and beyond. © *2017 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).* [http://dx.doi.org/10.1063/1.4973537]

In recent years, plasmonic enhanced silicon detectors based on internal photoemission (IPE) across the metal-silicon Schottky barrier have been demonstrated and shown to be useful in the detection of telecom band (1300-1600 nm)—beyond the detection capability of conventional silicon photodiodes and photoconductors. Both guided wave and free space configurations were implemented.[1–13] While such demonstrations open new opportunities for CMOS compatible nanophotonics and nanoplasmonics, the quantum efficiency of such detectors still falls well short off the commercially available InGaAs detectors. This is primarily due to the low transmission probability of electrons traversing the metal-silicon interface, combined with fast thermalization ($\sim$10 fs) of photo-excited hot electrons in the metal.[14]

Recently, it has been shown that the insertion of a graphene monolayer in between the metal and the semiconductor greatly enhances the internal quantum efficiency to about 7%.[15] While a large variety of photodetectors based on two dimensional materials have been demonstrated,[16] most of them are based on mechanisms which are different from the internal photoemission effect and to date, the graphene-metal-silicon photodetector seems to be a leading candidate for CMOS compatible internal photoemission detection in the telecom window. The observed enhancement in quantum efficiency makes photodetectors based on internal photoemission highly relevant for a variety of applications, e.g., chip scale optical communications. The reasons for such a significant enhancement

---

[a]Author to whom correspondence should be addressed. Electronic mail: jakek@jhu.edu

are not yet fully understood, and in general the transport of hot carriers (electrons and holes) between two-dimensional and three-dimensional materials is still an active topic of research.

In this work, we address this issue by developing a phenomenological model which explains the physics behind graphene-silicon internal photoemission based detectors. The model not only captures the latest findings but also provides guidelines for obtaining efficient internal photoemission processes, in terms of surface roughness, field confinement, and momentum matching.

The structure that is considered in this paper is that of a graphene monolayer, sandwiched between a metal and a semiconductor, as depicted in Fig. 1.

In the context of photodetectors based on internal photoemission from a metal to a semiconductor, graphene may offer the advantage of high absorption the 2D layer provides, while at the same time it can be placed right at the metal-semiconductor interface. This high absorption is attributed to the direct interband transition. This is in contrast to typical metals, in which absorption is based on free carriers, and must involve a mediator such as phonons or defects to provide momentum conservation. As a result, the graphene monolayer is capable of absorbing the light within a propagation length of few tens of microns and can be considered as a significant absorption mechanism, co-existing with the Ohmic damping in the metal. This can be shown by the following expression for the dispersion relation of the plasmonic mode with and without (by making $\sigma \to 0$) the presence of the intermediate graphene layer (see, e.g., Refs. 17 and 18 and references therein)

$$\frac{\varepsilon_{si}}{\sqrt{q^2 - \varepsilon_{si}\omega^2/c^2}} + \frac{\varepsilon_M}{\sqrt{q^2 - \varepsilon_M\omega^2/c^2}} = \frac{\sigma}{i\varepsilon_0\omega}, \tag{1}$$

where $\varepsilon_M = 1 - \frac{\omega_p^2}{\omega*(\omega+i\gamma)}$ is the Drude-like dielectric function of the metal, $\varepsilon_{si} = 12$ denotes the relative permittivity of silicon, ω is the angular frequency, q is the in-plane wave vector, and σ is the optical conductivity of graphene. Neglecting doping, this conductivity is entirely due to direct interband transitions, i.e., $\sigma = e^2/4\hbar = c\pi\alpha\varepsilon_0$. Due to the small magnitude of the fine structure constant, $\alpha \sim 1/137$, we perturbatively treat the effect of graphene on the SPP dispersion relation: to the lowest order, we find the complex valued shift in the wavevector to be $\Delta q \propto \alpha$. In the next step we calculate the enhanced absorption $\Gamma = \frac{\mathrm{Im}(\Delta q)}{\mathrm{Im}(q_0)}$ of the surface plasmon polariton (SPP) due to the presence of graphene. For low frequencies, the result is given by $\Gamma \simeq 2\pi\alpha\frac{\omega^2}{\gamma\omega_p}$, where γ and ω_p are the damping rate and the plasma frequency of the metal, respectively. By substituting typical values for gold at an operation frequency of 0.8 eV (telecom), one obtains up to 20% enhancement, depending on the specific choice of damping factor. Therefore, the relative absorption in graphene with respect to the total absorption is given by $\Gamma / (1 + \Gamma)$ and the SPP propagation length is now $L = \frac{L_0}{1+\Gamma}$, where L_0 is the SPP propagation length in the absence of graphene. With the above estimate, about 17% of the light is being absorbed in the graphene layer over a length of $\sim 30~\mu$m.

It should be noted that operating closer to the surface plasmon resonance, there is even a regime with slow light enhancement (not captured by the perturbative expansion). Another alternative for
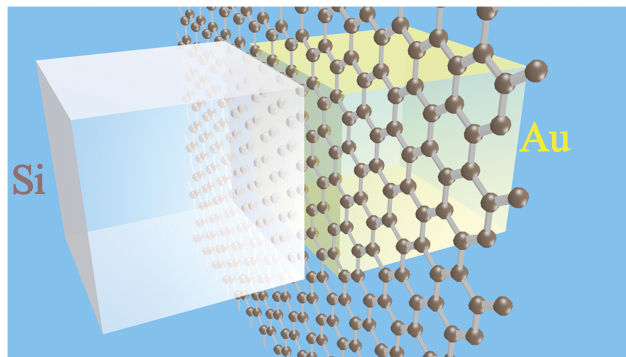

FIG. 1. Schematic diagram illustrating the geometry of the studied structure.

enhancing the absorption in graphene is by patterning the metallic layer on top with subwavelength structures, as discussed, e.g., in Ref. 19.

Since the transition in graphene is vertical, the energy of all the hot carriers is always $\hbar\omega/2$ above the Dirac point (Fig. 2(a)) in clear contrast to the absorption in the metal where photo-generated carriers have broad energy distribution. Assuming this value is above the Schottky barrier, the electron has sufficient energy and will cross the barrier provided that lateral momentum is preserved (Fig. 2(b)). We further note that in graphene, being a 2D material, any hot carrier is *at the boundary* with silicon.

In metals, the above-mentioned scenario is different for two major reasons: (a)—the plasmonic mode decays about 10 nm into the metal and thus the probability of the hot carrier to be generated at the boundary is finite. As a result, some of the hot carriers will experience thermalization and will not make it to the barrier. (b)—the hot carriers have a distribution of energies. Some of the carriers have sufficient energy to cross the barrier, while others do not.

Another important issue relates to the relaxation dynamics of hot electrons. Thermalization time in graphene, which is determined by electron-electron scattering, is on the time scale of 100 fs, while typical values for metals are significantly shorter, in the 10 fs range. This allows for multiple attempts for a hot electron to be transported across the Schottky barrier into the silicon as they are "bounced" back and forth in the atomically thin layer of graphene. This process further enhances the probability of internal photoemission.

Before providing a more detailed analysis, let us consider the impact of having all the absorption occurring in a monolayer. We conceptually represent the monolayer as a quantum well with a thickness $a \sim 3$ Å and assume that a photoexcited carrier is travelling back and forth in the well, with a velocity of $v = \pi\hbar/m_0 a$. Therefore, the round trip time is $\tau_{rt} = 2a/v = 2a^2 m/\pi\hbar = \hbar\,(\pi\hbar^2/2ma^2)^{-1}$ which is roughly 0.5 fs.

Assigning a transmission probability T to the transition from graphene to either silicon or metal, the effective emission time becomes $\tau_{em} = \tau_{rt}/(2T)$. At the same time the thermalization time of hot carriers is $\tau_{th} \sim 100$ fs (for recent ultrafast spectroscopy experiments of hot electrons in graphene see, e.g., Ref. 20). Then, the quantum efficiency of the emission into the silicon is given by $\eta = \frac{1}{2}\frac{\tau_{th}}{\tau_{th}+\tau_{em}}$ $= \frac{1}{2}\frac{\tau_{th}}{\tau_{th}+\tau_{rt}/2T}$. Assuming a relatively low transmission value of 0.1%, which is a characteristic value of metal-semiconductor interfaces,[14] one obtains a quantum efficiency in excess of 10%! This huge enhancement is entirely due the multiple attempts for transmission. Clearly, the large absorption in a single-layer of graphene plays a pivotal role in such an enhancement, as every photo-excited carrier is always not further than $\sim$3 Å from the graphene-silicon interface. It should be stated that the relatively long thermalization time of free carriers in graphene ($\sim$100 fs, as opposed to $\sim$10 fs
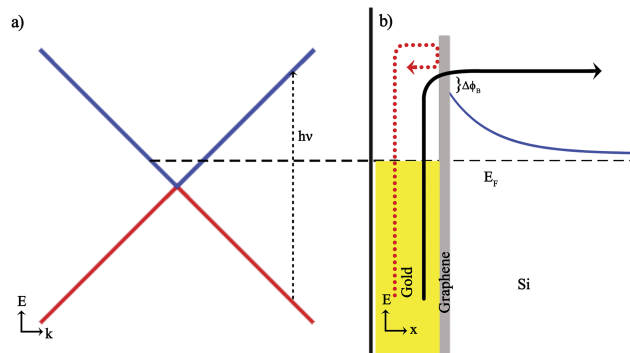
FIG. 2. (a) Electron energy band diagram of graphene, near the Dirac point. The horizontal dashed line is associated with the Fermi level under weak doping conditions (not to scale). For generality, we have slightly shifted the Fermi energy level from the Dirac point. (b) Schematic illustration of the internal photoemission process from the metal through the graphene and to the semiconductor. Photon energy is assumed to be larger than the Schottky barrier. The black solid line represents a hot electron that was transported to silicon through the barrier, while the red dashed line represents a hot electron that was reflected from the barrier. The blue line denotes the conduction band level in the silicon.

in metal), combined with the use of a waveguide configuration which allows the efficient absorption of light over relatively short propagation distance absorption, is pivotal in obtaining such high efficiencies.

Obviously, this rough prediction depends upon the value of T which reflects the character of atomic wavefunctions in the semiconductor and in the graphene, and also relies critically on the in-plane momentum conservation as depicted in Fig. 3 where the projection of the cubic Brillouin zone of Si onto the (k_x, k_y) plane is shown superimposed onto the hexagonal Brillouin zone of graphene, with the principal axis rotated by an arbitrary angle which cannot be predicted during fabrication. The key issue is that the photo-excited carriers (both holes and electrons) in graphene are located in the vicinity of the K points of graphene. The magnitude of the lateral wavevector of these carriers is about $K = 1.71 \, \text{Å}^{-1}$, and their energy is comparable to either the top of the valence band of a p-type silicon, located at the center of the Brillouin zone ($k_v \sim 0$), or the bottom of the conduction band, located at one of the six valleys along the 100 directions. Two of these valleys (pink circles) have an in-plane wavevector ($k_c \sim 0$) while the other four (pink ellipsoids) are located at ($k_c \sim 1 \, \text{Å}^{-1}$).

Therefore, the momentum mismatch involved in the transmission of a carrier from the graphene to the silicon is fairly large. If the photo-excited carrier is a hole going into the Γ valley of Si, then the mismatch is on the scale of $\delta k_{K\Gamma} \sim 1.7 \, \text{Å}^{-1}$, while for the photo-excited electrons going into the X valley of Si the mismatch is about $\delta k_{KX} \sim 0.9 \, \text{Å}^{-1}$.

However, this dire situation can be relaxed if one considers the possibility of the Umklapp process,[21] where momentum conservation, shown in Fig. 3, is written as $\vec{K} = \vec{k}_{c/v} + \vec{G}$, where $\vec{G}$ is one of the reciprocal lattice vectors. For example, $\vec{G} = \left[\frac{2\pi}{a}, \frac{2\pi}{a}, \frac{2\pi}{a} \right]$ in silicon has an absolute value of its projection onto the xy plane (shown in Fig. 3 as an orange arrow) of $\overrightarrow{|G_{xy}|} = 1.63 \, \text{Å}^{-1}$, which is close to the K point in graphene. As one can see from Fig. 3 with the help of the Umklapp process, the overall mismatch between the states in graphene and silicon $\Delta k_{||}$ is significantly reduced, yet it is still not zero, especially when one considers the inevitable randomness of the relative orientation of the

FIG. 3. Brillouin zones of Si (red) and graphene (blue) in the xy-plane. Because of the violation of momentum conservation, the hot electrons (holes) photo-generated in the vicinity of K-points of graphene (light blue circles) cannot be transferred to the minima (maxima) of the conduction (valence) band located near X (Γ) points and shown as pink ellipsoids (circles). They can, however, be transferred to the energy minimum (maximum) of the second Brillouin zone of Si-circle Γ2Si via the Umklapp process enabled by the lattice vector G. For generality, we have slightly rotated the graphene lattice with respect to the silicon lattice. In fact, while this rotation can take any value in the range of $0°-360°$, owing to the six-fold symmetry of the hexagonal lattice, it is sufficient to describe the angle between $0°$ and $30°$.

lattices. However, this small momentum mismatch can be mitigated by the disorder and roughness always present at the interface, as was observed, e.g., in the internal photoemission from the metal to the silicon in a plasmonic waveguide Schottky photodetector.[8] We now discuss this scenario in greater detail.

The coupling between graphene and silicon can be described by the Hamiltonian H_{coup} $= \langle \Psi_G | H | \Psi_{si} \rangle$, corresponding to the scale of few eV. This Hamiltonian has a constant value H_0 augmented by the disorder term $V_d(x, y)$, given by $V_d(x, y) = \iint F(k_x, k_y) e^{i(k_x x + k_y y)} dk_x dk_y$ which obviously is a random function but with a well-defined power spectrum $|F(k_x, k_y)|^2 \equiv |F(k_\parallel)|^2$ $= \pi^{-1} \Lambda V_0^2 e^{-k_\parallel^2 \Lambda^2}$, where V_0^2 is the power of the coupling integral variations and Λ is the correlation length.

According to the Fermi's golden rule, we can now write the rate of transfer from a given state in graphene with energy E_{gr} above the Schottky barrier to a state in the silicon with an energy E_{si} separated by the wavevector $\Delta \boldsymbol{k}_\parallel$ as

$$\mathfrak{R}(\Delta k_\parallel) = \frac{2\pi}{\hbar} |F(\Delta k_\parallel)|^2 \frac{a}{2L} \delta \left(E_{gr} - E_{si}(\Delta k_\parallel) \right), \tag{2}$$

where L is the extent of the Si wavefunctions in the direction normal to the plane used for normalization of these wavefunctions. This term will be cancelled once the integration over k_z takes place in Eq. (3). Next we need to perform the summation over all the states in the silicon having the same $k_{||}$. Using the effective mass approximation, the delta-function in Eq. (2) becomes

$$\delta \left(E - \frac{\hbar^2 k_\parallel^2}{2m^*} - \frac{\hbar^2 k_z^2}{2m^*} \right) = \frac{m^*}{\hbar^2 k_z} \delta \left(k_z - \sqrt{\frac{2m^* E}{\hbar^2} - k_\parallel^2} \right). \tag{3}$$

And thus

$$R(K) = \frac{2\pi}{\hbar} \frac{a}{L} \int_0^{\sqrt{\frac{2m^* E^2}{\hbar^2}}} 2\pi k_\parallel dk_\parallel |F(\boldsymbol{K} - \boldsymbol{k}_\parallel)|^2 \int_0^\infty \frac{m^*}{\hbar^2 k_z} \delta \left(k_z - \sqrt{\frac{2m^* E}{\hbar^2} - k_\parallel^2} \right) dk_z \times \frac{L}{2\pi}, \tag{4}$$

where K is wavevector of the state in graphene, roughly equal to the K-point. By integrating over k_z, one obtains

$$R(K) = \frac{2\pi}{\hbar} \frac{m^* a}{\hbar^2} \frac{1}{2} \int_0^{\frac{2m^* E}{\hbar^2}} \frac{dk_\parallel^2}{\sqrt{\frac{2m^* E}{\hbar^2} - k_\parallel^2}} |F(\boldsymbol{K} - \boldsymbol{k}_\parallel)|^2 \approx \left| F \left(\langle \Delta k_\parallel^2 \rangle \right) \right|^2 \frac{\pi}{\hbar} \frac{m^* a}{\hbar^2} \sqrt{\frac{2m^* E}{\hbar^2}} \int_0^1 \frac{dx}{\sqrt{1-x}}$$

$$= \frac{V_0^2}{\hbar} \frac{m^* \Lambda^2}{\hbar^2} \sqrt{\frac{2m^* a^2 E}{\hbar^2}} e^{-\langle \Delta k_\parallel^2 \rangle \Lambda^2}. \tag{5}$$

To further proceed, we should now estimate the transition rate using a reasonable parameter. The first energy level of our fictitious quantum well is given by $\hbar^2 / 2m^* a^2 \sim 3$ eV whereas a typical energy for the electron in the silicon is estimated to be $E = 1$ eV (this value is obtained by assuming a reasonable Schottky Barrier of 0.3 eV and using the photon energy at the wavelength of interest around 1.5 μm, corresponding to $\hbar \omega / 2 = 0.4$ eV). In this way, we obtain $\sqrt{\frac{2m^* a^2 E}{\hbar^2}} \simeq 0.2$.

Next, we make $V_0 = 0.1$ eV and $\Lambda \approx 10$ Å, so that $\frac{m^* \Lambda^2}{\hbar^2} = 6.5$ eV^{-1} and thereby arriving at

$$R(K) \approx 2 \times 10^{13} e^{-100 \langle \Delta k_\parallel^2 \rangle} \text{ s}^{-1}, \tag{6}$$

where $\Delta k_\parallel$ is given in units of Å^{-1}. The quantum efficiency can be estimated as simply the ratio $\eta \sim R \tau_{th} / (R \tau_{th} + R' \tau_{th} + 1)$, where R' is the rate of transfer of carriers from graphene to the metal. This rate is small because the Fermi wavevector in gold is $k_F \sim 1.2$ Å^{-1} $< K$. Hence once again the process of transfer from K states in graphene into the metal has Umklapp character. Therefore, for 10% efficiency assuming thermalization time of 100 fs, we need $R \sim 10^{12}$ s^{-1} which requires a very reasonable momentum mismatch of about $\Delta k_\parallel = 0.17$ Å^{-1} (this corresponds, e.g., to rotation by about 9.7° between the projected Brillouin zones of graphene and silicon; see Fig. 3).

In summary, we have presented a heuristic model which provides a possible explanation for the recently observed enhancement of internal photoemission efficiency from a metal to silicon.

Based on our model, the photo-generated hot electrons generated at the graphene layer can have multiple attempts of penetrating into the silicon due to the long thermalization time with respect to the round trip time of carriers in the monolayer, and thus having higher probability of escaping into the semiconductor and be collected as a useful photocurrent. We argue that the momentum mismatch between graphene and silicon can be partially overcom with the aid of Umklapp processes. This, combined with inevitable interface roughness, can provide efficiencies in excess of 10%, which makes graphene a viable candidate for enhancing Schottky barrier sub-bandgap photodetector. As a final note, it is interesting to observe the role of SPP in enhancing the interaction of light with graphene to such an extent that the low absorption of graphene monolayer is still sufficient for achieving decent photocurrents, owing to the field enhancement at the interface and the use of guided surface modes.

[1] A. Akbari and P. Berini, "Schottky contact surface-plasmon detector integrated with an asymmetric metal stripe waveguide," Appl. Phys. Lett. **95**, 21104 (2009).

[2] P. Berini, "Surface plasmon photodetectors and their applications," Laser Photonics Rev. **8**, 197–220 (2014).

[3] M. L. Brongersma, N. J. Halas, and P. Nordlander, "Plasmon-induced hot carrier science and technology," Nat. Nanotechnol. **10**, 25–34 (2015).

[4] H. Chalabi, D. Schoen, and M. L. Brongersma, "Hot-electron photodetection with a plasmonic nanostripe antenna," Nano Lett. **14**, 1374–1380 (2014).

[5] M. W. Knight *et al.*, "Embedding plasmonic nanostructure diodes enhances hot electron emission," Nano Lett. **13**, 1687–1692 (2013).

[6] M. W. Knight, H. Sobhani, P. Nordlander, and N. J. Halas, "Photodetection with active optical antennas," Science **332**, 702–704 (2011).

[7] B. Desiatov *et al.*, "Plasmonic enhanced silicon pyramids for internal photoemission Schottky detectors in the near-infrared regime," Optica **2**, 335–338 (2015).

[8] I. Goykhman, B. Desiatov, J. Khurgin, J. Shappir, and U. Levy, "Waveguide based compact silicon Schottky photodetector with enhanced responsivity in the telecom spectral band," Opt. Express **20**, 28594–28602 (2012).

[9] I. Goykhman, B. Desiatov, J. Khurgin, J. Shappir, and U. Levy, "Locally oxidized silicon surface-plasmon Schottky detector for telecom regime," Nano Lett. **11**, 2219–2224 (2011).

[10] M. Casalino, G. Coppola, M. Iodice, I. Rendina, and L. Sirleto, "Critically coupled silicon Fabry-Perot photodetectors based on the internal photoemission effect at 1550 nm," Opt. Express **20**, 12599 (2012).

[11] W. Li and J. Valentine, "Metamaterial perfect absorber based hot electron photodetection," Nano Lett. **14**, 3510–3514 (2014).

[12] P. Berini, A. Olivieri, and C. Chen, "Thin Au surface plasmon waveguide Schottky detectors on p-Si," Nanotechnology **23**, 444011 (2012).

[13] Z. Han *et al.*, "On-Chip detection of radiation guided by dielectric-loaded plasmonic waveguides," Nano Lett. **15**, 476–480 (2015).

[14] I. Goykhman, B. Desiatov, and, J. Shappir, "Model for quantum efficiency of guided mode plasmonic enhanced silicon Schottky detectors," preprint arXiv:1401.2624 (2014).

[15] I. Goykhman *et al.*, "On-chip integrated, silicon–graphene plasmonic Schottky photodetector with high responsivity and avalanche photogain," Nano Lett. **16**, 3005–3013 (2016).

[16] F. H. L. Koppens *et al.*, "Photodetectors based on graphene, other two-dimensional materials and hybrid systems," Nat. Nanotechnol. **9**, 780 (2014).

[17] P. A. D. Gonçalves and N. M. R. Peres, *An Introduction to Graphene Plasmonics* (World Scientific, 2016).

[18] S. Xiao, X. Zhu, B.-H. Li, and N. A. Mortensen, "Graphene-plasmon polaritons: From fundamental properties to potential applications," Front. Phys. **11**, 117801 (2016).

[19] M. Hashemi, M. H. Farzad, N.A. Mortensen, and S. Xiao, "Enhanced absorption of graphene in the visible region by use of plasmonic nanostructures," J. Opt. **15**, 55003 (2013).

[20] Z. Mics *et al.*, "Thermodynamic picture of ultrafast charge transport in graphene," Nat. Commun. **6**, 7655 (2015).

[21] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).