Thermionic graphene/silicon Schottky infrared photodetectors

S. Doukas,1 P. Mensz,2 N. Myoung,3 A. C. Ferrari,‡ I. Goykhman,∗ and E. Lidorikis†,1,5
1Department of Materials Science and Engineering, University of Ioannina, 45110, Ioannina, Greece
2Micro Nanoelectronics Research Center, Technion, Haifa 320000, Israel
3Department of Physics Education, Chosun University, Gwangju, 61452, Republic of Korea
4Cambridge Graphene Centre, University of Cambridge, Cambridge CB3 0FA, United Kingdom
5University Research Center of Ioannina (URCI), Institute of Materials Science and Computing, 45110 Ioannina, Greece

(Received 12 September 2021; revised 17 December 2021; accepted 19 January 2022; published 16 March 2022)

Optical communications, imaging, and biomedicine require efficient detection of infrared radiation. Growing demand pushes for the integration of such detectors on chips. It is a challenge for conventional semiconductor devices to meet these specs due to spectral limitations arising from their finite band gap, as well as material incompatibilities. Single layer graphene (SLG) is compatible with complementary metal-oxide-semiconductor (CMOS) Si technology, while its broadband (UV to THz) absorption makes the SLG/Si junction a promising platform for photodetection. Here we model the thermionic operation of SLG/Si Schottky photodetectors, considering SLG’s absorption, heat capacity, and carrier cooling dependence on temperature and carrier density. We self-consistently solve coupled rate equations involving electronic and lattice temperatures, and nonequilibrium carrier density under light illumination. We use as an example the infrared photon energy of 0.4 eV, below the threshold for direct photoemission over the Schottky barrier, to study the photothermionic response as a function of voltage bias, input power, pulse width, electronic injection, and relaxation rates. We find that device and operation parameters can be optimized to reach responsivities competitive with the state of the art for any light frequency, unlike conventional semiconductor-based devices. Our results prove that the SLG/Si junction is a broadband photodetection platform.

DOI: 10.1103/PhysRevB.105.115417

I. INTRODUCTION

Efficient light detection in the infrared (IR) is key to a variety of applications ranging from optical communications [1,2], imaging [3], biomedicine [4,5], tomography [6], and more. In the midwave IR (MWIR), corresponding to the wavelength range $\lambda \sim 3 - 8 \mu m$, current photodetector (PD) technology utilizes low ($<0.4$ eV) band-gap semiconductors such as HgCdTe [7] or III-V compounds (e.g., InSb, InAs) [8] with detectivities (i.e., the inverse of the minimal signal detected limited by the noise [9]) almost reaching the limits of ideal PD operation (e.g., $>10^{10}$ cmHz$^{0.5}$/W when cryogenically cooled [7]). The ever-growing demand for performance improvement and miniaturization, however, requires room-temperature (RT) PD cointegrated with supporting circuitry (e.g., drivers, amplifiers, and processors) on the same Si chip, utilizing complementary metal-oxide-semiconductor (CMOS) technology. III-V materials are not compatible with standard CMOS processes because of cross contamination [10] and dopant redistribution effects [10], requiring separate manufacturing [10], followed by complex bonding and assembling processes [11] and their spectral response is limited to photon energies above their band gap. A simple, yet promising, broadband CMOS-compatible approach is the single layer graphene (SLG)/Si Schottky junction [12–15], whereby sub-band-gap photons (i.e., of energy lower than the Si band gap of 1.1 eV [9]) are absorbed in SLG due to its gapless nature [12,16] and injected over the barrier in an internal photoemission (IPE) process [9,17–19].

SLG/semiconductor (SLG/SC) Schottky diodes [12,14,16,20–25] with rectifying electrical characteristics have been explored for PDs using free-space [26–47] and guided-mode configurations [15,48–50]. A SLG/SC Schottky PD has 3 modes of operation for photocurrent generation: (a) photovoltaic (PV) [26–35], where photons with energy $E_{ph} > E_g$ (SC band gap) are absorbed by the SC creating electron-hole (e-h) pairs, which then split under the internal field in the depletion region, (b) IPE [36–40], in which sub-band-gap $E_{ph} < E_g$ photons are absorbed in SLG to yield photoexcited carriers with sufficient energy to overcome the Schottky barrier at the SLG/SC interface and emit into the SC, and (c) photothermionic (PTh) [41–45], where low-energy photons (below both SC band gap and Schottky barrier) are absorbed in SLG yield photoexcited carriers, without sufficient energy to overcome the Schottky barrier directly. Instead, they multiply [51–53] and thermalize due to e-e scattering [53–57], resulting in an increased SLG electronic temperature, $T_e$.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

S. Doukas1 P. Mensz2 N. Myoung3 A. C. Ferrari4 I. Goykhman2,* and E. Lidorikis1,5,†
1Department of Materials Science and Engineering, University of Ioannina, 45110, Ioannina, Greece
2Micro Nanoelectronics Research Center, Technion, Haifa 320000, Israel
3Department of Physics Education, Chosun University, Gwangju, 61452, Republic of Korea
4Cambridge Graphene Centre, University of Cambridge, Cambridge CB3 0FA, United Kingdom
5University Research Center of Ioannina (URCI), Institute of Materials Science and Computing, 45110 Ioannina, Greece

(Received 12 September 2021; revised 17 December 2021; accepted 19 January 2022; published 16 March 2022)

Optical communications, imaging, and biomedicine require efficient detection of infrared radiation. Growing demand pushes for the integration of such detectors on chips. It is a challenge for conventional semiconductor devices to meet these specs due to spectral limitations arising from their finite band gap, as well as material incompatibilities. Single layer graphene (SLG) is compatible with complementary metal-oxide-semiconductor (CMOS) Si technology, while its broadband (UV to THz) absorption makes the SLG/Si junction a promising platform for photodetection. Here we model the thermionic operation of SLG/Si Schottky photodetectors, considering SLG’s absorption, heat capacity, and carrier cooling dependence on temperature and carrier density. We self-consistently solve coupled rate equations involving electronic and lattice temperatures, and nonequilibrium carrier density under light illumination. We use as an example the infrared photon energy of 0.4 eV, below the threshold for direct photoemission over the Schottky barrier, to study the photothermionic response as a function of voltage bias, input power, pulse width, electronic injection, and relaxation rates. We find that device and operation parameters can be optimized to reach responsivities competitive with the state of the art for any light frequency, unlike conventional semiconductor-based devices. Our results prove that the SLG/Si junction is a broadband photodetection platform.

DOI: 10.1103/PhysRevB.105.115417

I. INTRODUCTION

Efficient light detection in the infrared (IR) is key to a variety of applications ranging from optical communications [1,2], imaging [3], biomedicine [4,5], tomography [6], and more. In the midwave IR (MWIR), corresponding to the wavelength range $\lambda \sim 3 - 8 \mu m$, current photodetector (PD) technology utilizes low ($<0.4$ eV) band-gap semiconductors such as HgCdTe [7] or III-V compounds (e.g., InSb, InAs) [8] with detectivities (i.e., the inverse of the minimal signal detected limited by the noise [9]) almost reaching the limits of ideal PD operation (e.g., $>10^{10}$ cmHz$^{0.5}$/W when cryogenically cooled [7]). The ever-growing demand for performance improvement and miniaturization, however, requires room-temperature (RT) PD cointegrated with supporting circuitry (e.g., drivers, amplifiers, and processors) on the same Si chip, utilizing complementary metal-oxide-semiconductor (CMOS) technology. III-V materials are not compatible with standard CMOS processes because of cross contamination [10] and dopant redistribution effects [10], requiring separate manufacturing [10], followed by complex bonding and assembling processes [11] and their spectral response is limited to photon energies above their band gap. A simple, yet promising, broadband CMOS-compatible approach is the single layer graphene (SLG)/Si Schottky junction [12–15], whereby sub-band-gap photons (i.e., of energy lower than the Si band gap of 1.1 eV [9]) are absorbed in SLG due to its gapless nature [12,16] and injected over the barrier in an internal photoemission (IPE) process [9,17–19].

SLG/semiconductor (SLG/SC) Schottky diodes [12,14,16,20–25] with rectifying electrical characteristics have been explored for PDs using free-space [26–47] and guided-mode configurations [15,48–50]. A SLG/SC Schottky PD has 3 modes of operation for photocurrent generation: (a) photovoltaic (PV) [26–35], where photons with energy $E_{ph} > E_g$ (SC band gap) are absorbed by the SC creating electron-hole (e-h) pairs, which then split under the internal field in the depletion region, (b) IPE [36–40], in which sub-band-gap $E_{ph} < E_g$ photons are absorbed in SLG to yield photoexcited carriers with sufficient energy to overcome the Schottky barrier at the SLG/SC interface and emit into the SC, and (c) photothermionic (PTh) [41–45], where low-energy photons (below both SC band gap and Schottky barrier) are absorbed in SLG yield photoexcited carriers, without sufficient energy to overcome the Schottky barrier directly. Instead, they multiply [51–53] and thermalize due to e-e scattering [53–57], resulting in an increased SLG electronic temperature, $T_e$.
These thermalized hot carriers in SLG are long-lived, due to their slow cooling (~few ps) through phonon modes [41,57,60–67], raising the prospects for achieving high enough \( T_e \) to gain thermal energy for thermionic emission [41] (photocurrent) over the Schottky barrier.

In contrast to PV and IPE, where the spectral response of SLG/SC Schottky PDs is limited by \( E_g \) and Schottky barrier height (SBH, \( \Phi_B \)) [26–40], PTH enables low-energy \( (E_{ph} < 2\Phi_B) \) multispectral photodetection [41,43,45] owning to SLG broadband absorption [68,69]. The lower density of states in SLG [70] compared to bulk metals leads to SLG Fermi level (\( E_F \)) shifts upon doping [71], enabling the control of the SBH beyond the classical Schottky effect [9,72] of barrier lowering [9,14], by application of an external electric field via gating [73], and/or a reverse voltage bias [74,75]. This electrical tuning of SBH provides an extra degree of freedom for optimizing the PTH response in SLG/SC PDs and also allows one to reach ~100% optical absorption in SLG when the device is integrated into an optical cavity [76–80] by adjusting SLG absorption (i.e., optical loss) toward the critical coupling condition (loss rate equal to cavity decay rate) [76,77]. Thus, SLG/SC Schottky PDs are a promising material platform for photodetection across multiple MWIR spectral bands.

Reference [41] reported SLG/SC Schottky PD junctions in the PTh regime. In SLG/WSe\(_2\)/SLG Schottky PDs (with WSe\(_2\) thicknesses from ~2 to 40 nm [41]), the transition between PV and PTH was demonstrated, with a strong reduction in generated photocurrent, as well as a superlinear dependence on input power, the latter being a manifestation of thermal activation of carriers over the SBH. While it is common [32,35,41] that the photocurrent is strongly reduced in the PTH regime (compared to IPE or PV), it is not clear if this reduction is due to fundamental PTH limitations, or to improper device architecture and operation conditions. A theoretical understanding of photocurrents in the SLG/SC Schottky PTh domain is thus needed if one is to explore and optimize the PD operation parameters and performance.

The photocurrent in the PTh regime can be calculated within the framework of the Landauer transport formalism [14,22,41]. This simplifies to a modified Richardson equation [9] for low \( T_e \) (\( k_B T_e \ll \Phi_B \)) in SLG [41] (see Appendix D). The high \( T_e \) limit (\( T_e > 300 \) K), when the thermionic current dominates cooling, has also been considered [45]. There is, however, little understanding of what to expect in the intermediate regime, i.e., as \( T_e \) raises enough to change SLG’s properties (chemical potential, absorptivity, and heat capacity) [64] and PD operation. This is a nonlinear problem requiring self-consistent solutions: \( T_e \) is determined by light absorption, thermal capacity, and cooling rates [41,64,65], which, in turn, are \( T_e \) dependent [64].

Here, we develop a self-consistent framework to investigate the performance metrics, such as responsivity and detectivity, of SLG/Si Schottky PDs under IR illumination in the PTh regime, and optimize them by exploring different device architectures (cavity integration) and operation parameters (reverse bias, input power, carrier injection time, etc.). We consider PDs where SLG is integrated inside a dielectric Bragg cavity [76–80], and show that, by tuning the cavity to the critical coupling condition, almost total light absorption in SLG can be achieved, even for SLG nanoribbons of reduced surface coverage. This corresponds to SLG absorption density >100%, i.e., an absorption cross section larger than the geometrical cross section, resulting in a higher \( T_e \) compared to that for full SLG coverage at 100% absorption.

We self-consistently solve the coupled rate equations of \( T_e \), lattice temperature, \( T_i \), and nonequilibrium carrier density, to obtain, using the Landauer formalism [14,22], the PD PTh response as a function of input optical power, optical pulse duration, applied bias, and charge injection/relaxation rates, taking into account the phonon-mediated [81–83], disorder-assisted [62,84–86], and thermionic-injection [45] cooling processes. We implement realistic considerations about all device aspects, such as mobility, input power density, reverse bias, and doping. We show that, under appropriate combinations of the above, the PD external responsivity and specific detectivity can, in principle, reach ~1 A/W and ~10\(^7\) cmHz\(^2\)/W, comparable with graphene-based state-of-the-art performances reported in MWIR [12,16].

When combined with the SLG broadband absorption and fast electron dynamics, our results imply similarly sensitive and fast photodetection in other spectral bands, such as long wavelength IR (8–15 \( \mu \)m) and THz (>15 \( \mu \)m). Our methods and conclusions are applicable to any PTh SLG/SC Schottky PD configuration, providing a comprehensive framework for the SLG/SC Schottky junction as a photodetection platform throughout the IR. Our approach can be implemented in commercial simulators to create key software modules for designing graphene-based free-space photonic devices and integrated circuits. We provide in the Supplemental Material a detailed table of acronyms and symbols definition, which is useful for the rest of this paper.

II. BACKGROUND

A. Schottky junction formation

We consider a SLG with electron affinity [energy difference between SLG charge neutrality point (CNP) and vacuum level] \( \chi_{SLG} = 4.5 \) eV [87] (theoretical value, experiments put it between 4.2 eV and 5 eV, depending on metal contacts and fabrication [88]) and initial (i.e. before contact) p-doping at Fermi level \( E_{F,SLG} < 0 \). Henceforth, all quantities referring to SLG will be indicated with a SLG subscript, while those referring to Si with Si. Quantities without a subscript refer to whole device parameters. Table S1 in Supplemental Material [89] summarizes all acronyms and symbols used in this paper.

We place this SLG in contact with n-type Si (Fig. 1). The Si work function is \( \Phi_{Si} = \chi_{Si} + \phi_{Si} \), where \( \chi_{Si} = 4.05 \) eV is the electron affinity in Si [14], \( \phi_{Si} = E_{C, Si} - E_{F, Si} \) is the energy difference between the conduction band, \( E_{C, Si} \), and the Fermi level, \( E_{F, Si} \), in Si, and \( e \) is the electron charge. \( \phi_{Si} = k_B T_Si \ln \left( \frac{N_{C, Si}}{N_{C, Si}} \right) \), where \( N_{C, Si} = 2.8 \times 10^{14} \) cm\(^{-3}\) is the effective density of states in the conduction band of Si [9], \( N_{d, Si} \) cm\(^{-3}\) is the Si doping concentration, \( k_B \) is the Boltzmann constant, and \( E_{C, Si} \) is the electronic temperature in Si. Throughout this paper, we set \( T_{Si} = 300 \) K (due to the larger heat capacitance of Si compared to SLG [90]) and, unless stated otherwise, assume a commonly used dopant concentration in SLG Schottky devices [14] \( N_{d, Si} = 10^{16} \) cm\(^{-3}\).
The work function of SLG (before contact with Si) is $\Phi_{\text{SLG}} = \chi_{\text{SLG}} - \mu_{\text{SLG}}$, where $\mu_{\text{SLG}}$ is the equilibrium chemical potential (i.e., the Fermi energy at $T_{\text{SLG}} > 0$ [64]) of the suspended SLG. Upon contact [Fig. 1(b)], because $\Phi_{\text{Si}} < \Phi_{\text{SLG}}$, electrons will flow from Si to SLG elevating its Fermi level to $E_{F,\text{SLG}}$ (thermodynamic equilibrium, i.e., $E_{F,\text{SLG}}$ constant across the junction [14,74,91]) and work function to $\Phi_{\text{SLG}}$. At finite $T_{\text{e,SLG}}$, this corresponds to a SLG chemical potential $\mu_{\text{SLG}}(\Delta n_{\text{SLG}} < 0, \rho \text{ doped}; \mu_{\text{SLG}} > 0, \text{ n doped})$. Assuming no $E_{F,\text{SLG}}$ pinning [72] and neglecting surface-states effects [14,22], the energy balance in thermodynamic equilibrium gives a built-in voltage $V_0$ [14]:

$$eV_0 = \Phi_{\text{SLG}} - \Phi_{\text{Si}},$$

and SBH [14]:

$$\Phi_{R0} = eV_0 + e\phi_{\text{Si}}.$$  

In SLG Schottky junctions, the SBH depends on $\mu_{\text{SLG}}$ [14]. Therefore, if the device is reverse biased with external voltage $V_R$ (positive potential to Si), $e$ will flow into SLG, changing its Fermi level to $E_{F,\text{SLG}} = E_{F0,\text{SLG}} + \Delta E_{F,\text{SLG}}$, where $\Delta E_{F,\text{SLG}}$ is the induced Fermi level shift. $V_R$ will be distributed between SLG and Si, i.e., $|V_R| = |\Delta E_{F,\text{SLG}}|/e + |V'_{R}|$, where $V'_{R}$ is the reverse bias voltage drop in the Si depletion region resulting in energy band bending $|e(V_0 + V'_R)|$ [24,74,91]. To calculate $E_{F,\text{SLG}}, \mu_{\text{SLG}}$, and the reduced SBH $\Phi_{R0}$ when the device is reverse biased, we consider that the electrical charge transferred across the junction results in a change of depletion region charge, $\Delta Q_{D,\text{Si}} > 0$ due to $V'_R$ [14,24,74,92]

$$\Delta Q_{D,\text{Si}} = \sqrt{2\epsilon_{\text{Si}}N_d,\text{Si}} (V_0 + V'_{R} - \frac{k_B T_{e,\text{Si}}}{e}) - \sqrt{2\epsilon_{\text{Si}}N_d,\text{Si}} (V_0 - \frac{k_B T_{e,\text{Si}}}{e}),$$

where $\epsilon_{\text{Si}} = 11.7 \epsilon_0$ is the dielectric permittivity of Si [9] and $\epsilon_0$ the vacuum permittivity. The same charge is accumulated in SLG ($\Delta Q_{\text{SLG}} = -\Delta Q_{D,\text{Si}} < 0$) and causes $\Delta E_{F,\text{SLG}}$, i.e., $\Delta Q_{\text{SLG}} = e(n_{\text{SLG}} - n_{0,\text{SLG}})$, where $n_{\text{SLG}} = E_{F,\text{SLG}}^2/(\pi \hbar^2 v_{F,\text{SLG}}^2)$ is the e concentration associated with $E_{F,\text{SLG}}(\mu_{\text{SLG}})$ [69], $n_{0,\text{SLG}}$ is the initial e concentration in SLG associated with $E_{F0,\text{SLG}}(\mu_{0,\text{SLG}})$, $\hbar$ is the reduced Planck’s constant, and $v_{F,\text{SLG}} = 10^6 \text{m/s}$ is the SLG Fermi velocity [14]. Combining the above, we get a transcendental equation [24,74,91]

$$s_f E_{F,\text{SLG}}^2 - s_{F0} E_{F0,\text{SLG}}^2 = \sqrt{\Gamma(\chi_{\text{SLG}} - \Phi_{\text{Si}} - \mu_{\text{SLG}} + eV'_R - k_B T_{e,\text{Si}})}$$

$$- \sqrt{\Gamma(\chi_{\text{SLG}} - \Phi_{\text{Si}} - \mu_{0,\text{SLG}} - k_B T_{e,\text{Si}})}.$$  

where $\Gamma = 2\epsilon_{\text{Si}} N_d,\text{Si} \pi^2 \hbar^4 v_{F,\text{SLG}}^2 e^2$, $s_{F0} = \text{sign}(E_{F,\text{SLG}}, E_{F0,\text{SLG}})$, and $\mu_{\text{SLG}} = E_{F,\text{SLG}}$ for $T_{e,\text{SLG}} \to 0$, in which case Eq. (4) is solved algebraically. Several additional effects can also modify the SBH, such as image force barrier lowering [14,93] (i.e., the Schottky effect) and Schottky barrier inhomogeneity [14,93] but, for simplicity, we neglect them (also see Appendix A).

### B. Light absorption and carrier dynamics

We assume photon energies in the thermionic regime [41] $\hbar\omega/2 < \Phi_{\text{B}}^{\text{CNP}}$, where $\Phi_{\text{B}}^{\text{CNP}} = \chi_{\text{SLG}} - \chi_{\text{Si}}$ is the SBH with reference to the CNP, but with allowed interband transitions in SLG, i.e., $\hbar\omega/2 > E_{F,\text{SLG}}$. These cases involve photon energies below the Si band gap ($\hbar\omega < 1.1 \text{ eV}$ [9]), so absorption occurs only in SLG. Figure 2 presents the calculated charge density distributions [detailed derivation in Eqs. (18) and (19)] for SLG following IR absorption at $\hbar\omega = 0.4 \text{ eV}$. To demonstrate this concept, we assume, without loss of generality, $E_{F,\text{SLG}} = 0.15 \text{ eV}$ (a wider range of $E_{F,\text{SLG}}$ will be discussed later). The interband transitions create hot $e-e$ collisions [41,53,57,60,64], into an equilibrium $e$ bath at some elevated $T_{e,\text{SLG}}$ [Fig. 2(d)]. We assume $T_{e,\text{SLG}} = 900 \text{ K}$ at equilibrium, consistent with values found in our simulations. As a result,
the tail of the Fermi-Dirac distribution extending above the SBH gets enlarged, enabling thermionic emission. In pulsed optical operation, the $e^{-}$ bath eventually cools down to RT by electron-phonon ($e$-ph) scattering within a picosecond (ps) timescale $\tau_{e-ph}$ [53,57,62,66,67] [Fig. 2(e)].

To describe the $T_{e,SLG}$ dynamics, we assume a two-temperature model [41]:

\[
\begin{align*}
\tau_{e,SLG} \frac{dT_{e,SLG}}{dt} &= \alpha_{SLG} P_{in}(t) - J_{e-ph} - J_{th}, \quad (5) \\
\tau_{e,SLG} \frac{dT_{l,SLG}}{dt} &= J_{e-ph} - \Gamma_{SLG-Si}(T_{l,SLG} - T_{Si}), \quad (6)
\end{align*}
\]

where $T_{e,SLG}$ does generally differ from the SLG lattice temperature $T_{l,SLG}$ [41], $\alpha_{SLG} \equiv \alpha_{SLG}(\mu_{e,SLG}, \mu_{v,SLG}, T_{e,SLG})$ is the optical absorptance in SLG [64], $P_{in}$ is the input optical power density, $c_{e,SLG} \equiv c_{e,SLG}(\mu_{e,SLG}, \mu_{v,SLG}, T_{e,SLG})$ the electronic heat capacity, $c_{l,SLG} \equiv c_{l,SLG}(T_{l,SLG})$ the lattice heat capacity, $J_{th} \equiv J_{th}(\mu_{e,SLG}, \Phi_{B0}, T_{e,SLG}, T_{Si})$ the thermal current density carried by thermionic emission across the junction [45], $J_{e-ph} \equiv J_{e-ph}(\mu_{e,SLG}, \mu_{v,SLG}, T_{e,SLG}, T_{l,SLG})$ the electronic thermal cooling current into the phonon bath, $\mu_{e,SLG}, \mu_{v,SLG}$ the nonequilibrium chemical potentials due to photoexcited carriers, $\Gamma_{SLG-Si} \sim 20 \text{ MWm}^{-2}\text{K}^{-1}$ the cooling rate of the SLG lattice into Si [90] (see Appendix B), and $T_{Si}$ the Si lattice temperature (assumed fixed at 300 K). Both $\alpha_{SLG}$ and $c_{e,SLG}$ are $T_{e,SLG}$ dependent [64], so we must seek self-consistent solutions to Eqs. (5) and (6). They also depend on the nonequilibrium chemical potentials $\mu_{e,SLG}, \mu_{v,SLG}$, which are themselves $T_{e,SLG}$ dependent [64]. To calculate $\mu_{e,SLG}, \mu_{v,SLG}$ [derived in Eqs. (16)–(20)], we note that photoexcited hot carriers stem from the interband absorption creating hot $e^{-}h$ pairs [Fig. 2(c)]. These will relax into a Fermi-Dirac thermal equilibrium distribution by $e^{-}e$ scattering within the timescale $\tau_{e-e}$ as for [64]:

\[
\tau_{e-e} \frac{d\delta_{\text{e,SLG}}}{dt} = \frac{P_{in}(t)}{h\omega} \alpha_{\text{inter,SLG}}(\mu_{e,SLG}, \mu_{v,SLG}, T_{e,SLG}) - \delta_{\text{e,SLG}} \tau_{e-e}, \quad (7)
\]

where $\delta_{\text{e,SLG}}$ is the nonequilibrium carrier density and $\alpha_{\text{inter,SLG}}$ the interband SLG absorption. Throughout this work, we set $\tau_{e-e} = 20 \text{ fs}$ [53,57].

To calculate $J_{e-ph}$, we account for two dominant processes, $e$-ph scattering with optical phonons [53,57,66,94] and with acoustic phonons via disorder-assisted supercollisions [45,62,84–86]:

\[
\begin{align*}
J_{e-ph} &\equiv J_{SC}(\mu_{e,SLG}, \mu_{v,SLG}, T_{e,SLG}, T_{l,SLG}) \\
&\quad + J_{op}(\mu_{SLG}, T_{e,SLG}, T_{l,SLG}). \quad (8)
\end{align*}
\]

Normal (momentum-conserving) acoustic phonon scattering is suppressed due to SLG’s small Fermi surface [84,95] and we neglect its contribution [45,62,82,83,85,86]. The thermal current due to optical phonon scattering is [82,96,97]

\[
\begin{align*}
J_{op}(\mu_{SLG}, T_{e,SLG}, T_{l,SLG}) &= \sum_{\omega} \left\{ \frac{9\hbar v_{\omega}^2}{\pi (\hbar v_{\omega}^{EF,SLG})^4} \rho_{\text{SLG}} \\
&\quad \times \left[ N\left(\frac{\Omega_{\omega,SLG}}{k_{B}T_{e,SLG}}\right) - N\left(\frac{\Omega_{\omega,SLG}}{k_{B}T_{l,SLG}}\right) \right] \right\} F(\mu_{SLG}, T_{e,SLG}). \quad (9)
\end{align*}
\]

where

\[
F(\mu_{SLG}, T_{e,SLG}) = \int_{-\infty}^{\infty} |\chi(1-x)| [f(\Omega(x-1)) - f(\Omega,x)] dx. \quad (10)
\]

We account for two optical phonon branches, with $\Omega_{\text{K}} = 161 \text{ meV}$, corresponding to the Raman D-peak at the $K$ point of Brillouin zone [94,98,99] and $\Omega_{\Gamma} = 196 \text{ meV}$, corresponding to the Raman G-peak at the $\Gamma$ point of the Brillouin zone [94,98,99]. In Eq. (9), $\rho_{\text{SLG}} = 7.6 \times 10^{-7} \text{ Kg/m}^2$ is the SLG mass density [81,95,100–102], $\gamma' \sim 40 \text{ eV/nm}$ the derivative of the nearest-neighbor coupling amplitude (hopping integral) $\gamma$ with respect to SLG bond length [96], $N(x) = [e^x - 1]^{-1}$ the Bose distribution [96], and $f_{FD}(\epsilon; \mu_{SLG}, T_{e,SLG}) = \frac{1}{e^{(\epsilon - \mu_{SLG})/k_BT_{e,SLG}} + 1}$.
{\exp[(\varepsilon - \mu_{\text{SLG}})/(k_BT_{\text{SLG}})] + 1}^{-1} \text{ is the Fermi-Dirac distribution.}

An additional cooling pathway of SLG carriers is the disorder-mediated (momentum nonconserving) emission of acoustic phonons, labeled as supercollisions [62,84–86], which consist in three-body collisions between carriers, phonons, and impurities [84–86]. With normal (momentum conserving) collisions constrained by SLG’s small Fermi surface, and with optical phonon energies (\Omega_r = 196 eV, \Omega_{\alpha} = 161 eV) \gg k_BT_{\text{SLG}} \text{ for } T_{\text{SLG}} \text{ below a few hundred K, which makes optical phonon scattering insignificant, disorder-mediated scattering can potentially dominate } J_{e-ph} \text{ by using the entire thermal distribution of phonons [86] and leading to energy dissipation } \sim \text{ per scattering event [85].}

This cooling channel is modeled by assuming disordered short-range scatterers with a mean-free path \( l \). It is an open debate whether \( J_{e-ph} \) is dominated by \( J_{op} \) or \( J_{SC} \) [62,85,86,97]. For simplicity, here we assume \( J_{op} \) and \( J_{SC} \) to be roughly at the same level. This can be achieved by setting \( l = 200 \text{ nm.} \)

The thermal current due to disorder-assisted supercollisions is [62,84–86]

\[ J_{SC} = \gamma_{SC}(T_{e,\text{SLG}}^3 - T_{l,\text{SLG}}^3), \]

where

\[ \gamma_{SC} = \frac{9.62 D_{\text{SLG}}^2 k_B^4}{2 \rho_{\text{SLG}}^2 k_{\text{F,SLG}}^2 h^2} \left( 4(n_e,\text{SLG} + n_h,\text{SLG} + 2\delta_{n,\text{SLG}}) \right) \pi (h v_{F,\text{SLG}})^3, \]

(12)

with \( n_e,\text{SLG} \) (\( n_h,\text{SLG} \)) = \( (e)(h) \) concentration, \( \delta_{n,\text{SLG}} \) the nonequilibrium carrier density, and \( k_{\text{F,SLG}} \) the disorder parameter [62,84–86], defined as the product of SLG Fermi wave vector \( k_{\text{F,SLG}} \) and mean-free path \( l \) for short-range scatterers. The latter is inversely proportional to the concentration of impurities [86]. For \( l = 200 \text{ nm}, k_{\text{F,SLG}} \sim 50 \text{ (for } E_{F,\text{SLG}} = 0.15 \text{ eV), we obtain an e-ph relaxation time } \tau_{e-ph} \sim 1.5 \text{ ps}, \text{ consistent with previous reports [41,53,57,60–62].} \)

\( D_{\text{SLG}} \) is the deformation potential for supercollision scattering, i.e., the effective potential associated with lattice deformation upon disorder-mediated acoustic phonon emission [84–86,102,103]. Literature values are between 10 and 30 eV [62,85,95,102,104]. Here, we set it at an intermediate value of 20 eV [45,62,84,85]. \( s_{\text{SLG}} \) in Eq. (12) is the sound velocity in SLG, with theoretical [81,83,84,100,105] and experimental [62,86,95,102] values ranging from 10^4 m/s to 2 \times 10^4 m/s. Here, we assume \( s_{\text{SLG}} = 1.5 \times 10^4 \text{ m/s, in agreement with theoretical predictions accounting for both transverse and longitudinal acoustic phonons [105].} \)

Equation (12) for \( \gamma_{SC} \) takes into account all available carriers for supercollision scattering, i.e., \( e, h \), and \( \delta_{n,\text{SLG}} \). For completeness, in Appendix B we compare the results derived with the cooling pathways of Eqs. (8)–(12) (adopted throughout this work) with another approach which averages all the e-ph scattering channels using a single temperature-independent relaxation rate [64,65].

Figure 13 in Appendix B shows that both approaches lead to similar results for e-ph cooling, in agreement with previous theoretical [84,85] and experimental [41,53,57,60,62] studies.

When \( \tau_{\text{pulse}} \gg \tau_{e-e} \), Eq. (7) can be solved for the quasi-cw case:

\[ \delta_{n,\text{SLG}} = \frac{\tau_{e-e}}{\alpha_{\text{inter,SLG}}(\mu_{\text{e,SLG}}, \mu_{v,\text{SLG}}, T_{e,\text{SLG}})P_{in}}. \]

(13)

Equation (13) needs to be solved self-consistently, since \( \mu_{\text{e,SLG}} \) and \( \mu_{v,\text{SLG}} \) depend on \( \delta_{n,\text{SLG}} \). For \( \tau_{\text{pulse}} \gg \tau_{ph} \), Eqs. (5) and (6) can also be solved in the quasi-cw case to get

\[ \alpha_{\text{SLG}}P_{in} = J_{e-ph} + J_{th} \]

(14)

and

\[ \tau_{l,\text{SLG}} = \tau_{\text{Si}} + \tau_{\text{SLG}}^{-1} J_{e-ph}. \]

(15)

Since \( \alpha_{\text{SLG}}, J_{e-ph} \) and \( J_{th} \) depend on \( T_{e,\text{SLG}} \), Eqs. (14) and (15) must also be solved self-consistently.

C. Equilibrium, nonequilibrium carriers, and heat capacity

\( E_{F,\text{SLG}} \) sets the net SLG charge density. Let us initially assume n-doped SLG. At \( T_{e,\text{SLG}} = 0 \):

\[ n_{e,\text{SLG}}(E_{F,\text{SLG}}, 0) = \frac{E_{F,\text{SLG}}^2}{\pi \hbar^2 v_{F,\text{SLG}}^2}, n_{h,\text{SLG}}(E_{F,\text{SLG}}, 0) = 0. \]

(16)

At \( T_{e,\text{SLG}} > 0 \) and equilibrium, the individual charge carrier concentrations increase, but the net charge remains constant:

\[ n_{e,\text{SLG}}(\mu_{\text{SLG}}, T_{e,\text{SLG}}) - n_{h,\text{SLG}}(\mu_{\text{SLG}}, T_{e,\text{SLG}}) = \frac{E_{F,\text{SLG}}^2}{\pi \hbar^2 v_{F,\text{SLG}}^2}, \]

(17)

where [64]

\[ n_{e,\text{SLG}}(\mu_{\text{SLG}}, T_{e,\text{SLG}}) = \int_0^{\infty} \nu(\epsilon) f_{FD}(\epsilon; \mu_{\text{SLG}}, T_{e,\text{SLG}}) \ d\epsilon, \]

(18)

\[ n_{h,\text{SLG}}(\mu_{\text{SLG}}, T_{e,\text{SLG}}) = \int_0^{\infty} \nu(\epsilon) f_{FD}(\epsilon; -\mu_{\text{SLG}}, T_{e,\text{SLG}}) \ d\epsilon, \]

(19)

with \( \nu(\epsilon) = 2|\epsilon|/[(\pi \hbar^2)^2 v_{F,\text{SLG}}^2] \) the carrier density of states [70]. We account for the nonequilibrium photoexcited carrier density \( \delta_{n,\text{SLG}} \) [from the solution of Eq. (7)] and use Eqs. (18) and (19) to solve for \( \mu_{e,\text{SLG}} \), \( \mu_{v,\text{SLG}} \) according to [64]

\[ n_{e,\text{SLG}}(\mu_{\text{e,SLG}}, T_{e,\text{SLG}}) = n_{e,\text{SLG}}(\mu_{\text{SLG}}, T_{e,\text{SLG}}) + \delta_{n,\text{SLG}}, \]

(20a)

\[ n_{h,\text{SLG}}(\mu_{\text{v,SLG}}, T_{e,\text{SLG}}) = n_{h,\text{SLG}}(\mu_{\text{SLG}}, T_{e,\text{SLG}}) + \delta_{n,\text{SLG}}. \]

(20b)

Once \( \mu_{e,\text{SLG}} \) and \( \mu_{v,\text{SLG}} \) are known, the electronic heat capacity can be calculated as [64]

\[ c_{e,\text{SLG}}(\mu_{\text{e,SLG}}, \mu_{\text{v,SLG}}, T_{e,\text{SLG}}) = \frac{\partial}{\partial T_{e,\text{SLG}}} \int_0^{\infty} \nu(\epsilon) \left[ f_{FD}(\epsilon; \mu_{\text{e,SLG}}, T_{e,\text{SLG}}) + f_{FD}(\epsilon; -\mu_{\text{v,SLG}}, T_{e,\text{SLG}}) \right] \ d\epsilon \]

(21)
D. SLG optical properties

Light absorption in SLG is governed by inter- and intraband transitions, described by the optical conductivities $\sigma_{\text{opt,SLG}}^\text{inter}$ and $\sigma_{\text{opt,SLG}}^\text{intra}$ using the Kubo formula [69,106]:

$$
\sigma_{\text{opt,SLG}}^\text{opt} (\omega; \mu_{c,\text{SLG}}, \mu_{v,\text{SLG}}, T_{e,\text{SLG}}) \\
= \sigma_{\text{opt,SLG}}^\text{inter} (\omega; \mu_{c,\text{SLG}}, \mu_{v,\text{SLG}}, T_{e,\text{SLG}}) \\
+ \sigma_{\text{opt,SLG}}^\text{intra} (\omega; \mu_{c,\text{SLG}}, \mu_{v,\text{SLG}}, T_{e,\text{SLG}}),
$$

(22)

with

$$
\sigma_{\text{opt,SLG}}^\text{intra} = -\frac{i e^2}{\pi \hbar^2} \int_0^\infty \left[ \frac{\partial f_{FD} (-\epsilon; \mu_{c,\text{SLG}}, T_{e,\text{SLG}})}{\partial \epsilon} - \frac{\partial f_{FD} (\epsilon; \mu_{c,\text{SLG}}, T_{e,\text{SLG}})}{\partial \epsilon} \right] d\epsilon,
$$

(23)

$$
\sigma_{\text{opt,SLG}}^\text{inter} = \frac{i e^2 \Omega}{2 \pi \hbar^2} \int_0^\infty \left[ f_{FD} (-\epsilon; \mu_{c,\text{SLG}}, T_{e,\text{SLG}}) - f_{FD} (\epsilon; \mu_{c,\text{SLG}}, T_{e,\text{SLG}}) \right] d\epsilon,
$$

(24)

where $\Omega = \omega + i \tau_{\text{opt}}^{-1}$ and $\tau_{\text{opt}}$ is the free electron relaxation time related to charge carrier mobility in SLG [107]. The latter depends on both SLG quality (i.e., contaminants, defects, wrinkles, and nonuniformities) as well as the substrate on which SLG is placed and the interface quality [107]. Depending on all above parameters, $\tau_{\text{opt}}$ can have values from 10 fs up to >300 fs [107–109]. For 0.4 eV photon energy the SLG absorption is dominated by intraband transitions because $2|E_F| < \hbar \omega$. For these, the $\tau_{\text{opt}}$ effect is small (it mainly affects intraband absorption [107]). For the $E_{F,\text{SLG}}$ values used in this work, we estimate that the change in $\alpha_{\text{SLG}}$ when varying $\tau_{\text{opt}}$ between 100 and 300 fs is $<$1% at RT. Thus, for simplicity, we set $\tau_{\text{opt}} = 200$ fs.

The dielectric function of SLG is calculated as [110,111]:

$$
\epsilon_{\text{SLG}} (\omega; \mu_{c,\text{SLG}}, \mu_{v,\text{SLG}}, T_{e,\text{SLG}}) \\
= \epsilon_{\infty,\text{SLG}} + \frac{i \sigma_{\text{SLG}}^\text{opt} (\omega; \mu_{c,\text{SLG}}, \mu_{v,\text{SLG}}, T_{e,\text{SLG}})}{\epsilon_0 \omega d_{\text{SLG}}},
$$

(25)

where we assume a SLG thickness $d_{\text{SLG}} = 0.335$ nm [112] and $\epsilon_{\infty,\text{SLG}} = 5.7$ is the high-frequency limit of the SLG dielectric permittivity [110]. Equation (25) can be used in a 1D transfer matrix model [113] to extract $\alpha_{\text{SLG}}$ in a multilayer resonator structure. Since $\alpha_{\text{SLG}}$ is proportional to $\Im [\epsilon_{\text{SLG}}]$ [110] i.e., to $Re [\sigma_{\text{SLG}}^\text{opt}]$, according to Eq. (25), the interband contribution to optical absorption is determined by

$$
\alpha_{\text{inter,SLG}} = \alpha_{\text{SLG}} \frac{Re [\sigma_{\text{opt,SLG}}^\text{opt}]}{Re [\sigma_{\text{opt,SLG}}^\text{intra} + \sigma_{\text{opt,SLG}}^\text{inter}]}.
$$

(26)

E. Photothermionic current

For given $T_{e,\text{SLG}}$, $T_{e,\text{Si}}$, and $\Theta_B$ the thermionic current across the SLG/Si Schottky junction can be calculated using the Landauer formalism [14,22]:

$$
J_{et} = -\frac{e}{\tau_{\text{inj}}} \int_{-\infty}^{\infty} v(\epsilon) \tau(\epsilon) D f(\epsilon) d\epsilon,
$$

(27)

where $\tau_{\text{inj}}^{-1}$ is the effective injection rate of charge carriers from SLG to Si [14,22,41,114] (which may be limited by the injection rate from contacts to Si [22,114]), $D f(\epsilon) = f_{FD}(\epsilon; \mu_{c,\text{SLG}}, T_{e,\text{SLG}}) - f_{FD}(\epsilon; \mu_{c,\text{Si}}, T_{e,\text{Si}})$, and $\tau(\epsilon)$ is the transmission probability of charge carriers from SLG to Si over $\Phi_B$. For simplicity, we assume a Heaviside step function for the latter, $\tau(\epsilon) = \Theta(\epsilon - \Phi_B^{\text{CNP}})$ [14,22]. There is also a corresponding thermionic thermal current ($J_B$, since the transporting electrons carry energy $\epsilon - \mu_{c,\text{SLG}}$ across the junction [45]):

$$
J_B = \frac{1}{\tau_{\text{inj}}} \int_{-\infty}^{\infty} v(\epsilon)(\epsilon - \mu_{c,\text{SLG}}) \tau(\epsilon) D f(\epsilon) d\epsilon.
$$

(28)

This current will contribute to cooling as noted in Eq. (5).

Alternative transport mechanisms, e.g., thermionic field emission (TFE) [14,115] and field emission (FE) [14,115], which include tunneling through the barrier, may also contribute to the photocurrent across the junction. For $N_{\text{Si}}/\text{Si} \lesssim 3 \times 10^{17} \text{ cm}^{-3}$ and $T_{e,\text{SLG}} > RT$, examined in this paper, TE is the dominating mechanism [14], prevailing over TFE and FE as well as over additional mechanisms such as electron diffusion in the depletion region [14,116]. Indeed,Refs. [14,22,41] showed that Landauer transport for TE can accurately describe the photocurrent across a SLG/Si Schottky junction. We also note that due to the absence of $T_{e,\text{SLG}}$ gradients in SLG in the proposed platform, we do not expect photo thermoelectric effects to contribute to the photocurrent.

F. Carrier injection time

$t_{\text{inj}} [\text{ps}]$ is a phenomenological parameter characterizing the rate at which charge carriers are injected from SLG into Si across the SLG/Si interface [14,22]. Its value depends on various fundamental and/or device aspects, such as SLG quality [14,22], Schottky interface quality [23,41,74,92], momentum conservation/relaxation of charge carriers upon injection over the Schottky barrier [114,117,118], and Ohmic contact interfaces [14,22,119], as well as $E_{F,\text{SLG}}$ and $T_{e,\text{SLG}}$ [114,117,118].

It remains an open question of how low $t_{\text{inj}}$ can be achieved at the SLG/Si interface, e.g., $t_{\text{inj}} \sim 0.11$ ps was extracted for SLG/nSi Schottky junctions in Ref. [114] and ~50 ps for SLG/nSi Schottky junctions in Ref. [22] using Cr/Au contacts, the latter dropping to ~0.13 ps if Pd metal contacts are used instead [22,119]. A theoretical estimation of a lower bound for $t_{\text{inj}}$ [117] using the Fermi liquid theory for doped SLG [117] $(|E_{F,\text{SLG}}| \gg k_B T_{e,\text{SLG}})$ and the Dirac liquid theory for intrinsic SLG $(|E_{F,\text{SLG}}| \ll k_B T_{e,\text{SLG}})$ [117], derived minimum values $t_{\text{inj}} \sim \hbar |E_{F,\text{SLG}}|/(k_B T_{e,\text{SLG}})^2$ for the former and $t_{\text{inj}} \sim \hbar/(k_B T_{e,\text{SLG}}) \sim 0.025$ ps at RT for the latter. We thus consider our device performance as a function of $t_{\text{inj}}$ from a minimum of 0.13 ps to a maximum of 30 ps, with the upper limit accounting for $t_{\text{inj}}$ limited by contact resistance [22,41].
Figure 6 also considers the fs regime to account for possible future experimental improvements.

For a qualitative comparison to the bulk case, we note that for metal/nSi Schottky junctions, the Richardson constant is $A_{\text{eff}} = 112 \text{cm}^{-2}\text{K}^{-2}$ [14]. In the Landauer transport formalism for 2d systems, the effective Richardson constant is $A_{\text{eff}}^2 \sim 2e k_B^2/(\tau_{\text{eq}} \pi \hbar^2 v_F^{2,\text{SLG}})(\Phi_B^{\text{CNP}}/k_B T)$ (see Appendix E). For $\Phi_B^{\text{CNP}} = 0.45$ eV and RT, $A_{\text{eff}}^2$ reaches the $A_{\text{eff}}^3$ value for $\tau_{\text{eq}} \sim 0.03$ ps. This is close to the lower limits of $\tau_{\text{eq}}$ estimated in Ref. [117].

G. Geometrical focusing and giant absorption

The interband absorption in suspended SLG is 2.3% [68] and it decreases to 0.46% when SLG is transferred on Si due to the Fresnel reflections at the air-substrate interface [110]. This is a severe limitation for any SLG PD. An improvement is needed to enforce the optical power to be fully absorbed in SLG. Integrating a SLG/Si thermionic PD in an optical cavity can result in $\sim 100\%$ absorption of incident light by SLG [76–80]. If we consider a one-port asymmetric cavity [78], i.e., a Bragg cavity with a metal mirror on the backside, the total absorption of SLG, nSi, and back mirror is given by [76–78]

$$\alpha = \frac{4\gamma_a \gamma_d}{(\omega - \omega_0)^2 + (\gamma_a + \gamma_d)^2},$$  \hspace{1cm} (29)$$

where $\gamma_a$ [s$^{-1}$] and $\gamma_d$ [s$^{-1}$] are the total cavity absorption and decay rates, and $\omega_0$ is the resonant frequency. Under the critical coupling condition $\gamma_a = \gamma_d$, 100% absorption can be achieved when the cavity is at resonance ($\omega = \omega_0$). Assuming a fixed cavity geometry with predefined (fixed) $\gamma_d$, we achieve critical coupling by tuning $E_{F,\text{SLG}}$ (thus SLG absorption, i.e., $\gamma_a$) using reverse bias across the junction. This allows for perfect absorption even in nanostructured (e.g., nanoribbons of subwavelength width) SLG partially covering the PD illuminated area. Perfect absorption in a partially covered SLG area means that the absorption per unit SLG area is inversely proportional to its coverage fraction. We quantify this by defining the SLG inverse surface coverage ratio, $F = S_0/S$, as the ratio of the diffraction limited area $S_0 = \lambda^2/\pi$ over the SLG geometrical cross section $S$ (within $S_0$). In this configuration, the absorption density, defined as the ratio of absorption cross section to SLG geometrical cross section, can be up to 100% $\times F$. This is equivalent to light focusing within a spot below the diffraction limit. We term this geometrical focusing. The increased absorption density will lead to increased $T_{e,\text{SLG}}$, and thus increased thermionic emission (see Appendix C).

H. Performance metrics

To evaluate PD performance, we use the following figures of merit: (a) external responsivity [9,37] $R_{\text{ext}} = J_d/P_{\text{in}}A/[\text{W}], where J_d = J_d - J_i$ is the photocurrent density, $J_d$ the total reverse current density under illumination, and $J_i$ the saturation current density at dark; (b) specific detectivity [9,37] $D^* = \sqrt{S \Delta f}/\text{NEP} [\text{cm}$,$\sqrt{\text{Hz/W}}$, where $S$ is the device area, $\Delta f$ is the bandwidth, and NEP is the noise equivalent power, i.e., the minimum detectable power when the signal-to-noise ratio SNR = 1 [9]. The total noise $i_n$ comes from the combined contributions of shot (quantum) and Johnson (thermal) noise $i_n = i_s + i_j$, where $i_s$, $i_j$ are the noise currents, normalized to the 1 Hz spectral band [9,37]: $i_s = [2e(i_d + I_{ph})]^{0.5}$ and $i_j = (4k_B T_e R_{eq})^{0.5}$, with $i_d$ the dark current, $I_{ph}$ the photocurrent, and $R_{eq} = dV/dI$ the equivalent resistance at reverse bias in the dark.

III. RESULTS AND DISCUSSION

We fist consider SLG on a n-Si substrate illuminated by a 1.5 ps pulse at $\lambda = 3.1 \mu\text{m}$, where $h\nu = 0.4 \text{eV}$ at peak power $P_{\text{in}} = 10^8 \text{W/cm}^2$ (fluence $\sim 150 \mu\text{J/cm}^2$, enough to create observable thermal effects, Fig. 3). Equations (5)–(7) are time integrated using a fourth-order Runge-Kutta method [120,121], self-consistently including all definitions in Eqs. (1)–(28).
The structure and layers thicknesses are optimized to cavity schematic) to allow duration and $\tau$ as in Fig. 3. For $F = 1$, $R_{\text{ext}}$ as a function of $P_{\text{in}}$. $R_{\text{ext}}$ is also plotted for $F = 4, 20, 100$. Inset: Asymmetric one-port Bragg cavity configuration.

We assume $\tau_{e-e} = 20$ fs, $\tau_{\text{opt}} = 200$ fs, $\tau_{\text{inj}} = 30$ ps, $l = 200$ nm, resulting in $k_{\text{F-SLG}} l \sim 50$ and $E_{F,\text{SLG}} = 0.15$ eV before contact, giving rise to $\Phi_{B0} = 0.30$ eV at zero voltage bias, a value commonly found in SLG/Si heterojunctions [14,122]. Figure 3 plots all the dynamics within and after the pulse duration. In Fig. 3(a), $T_{\text{e-SLG}}$ reaches $\sim 1200$ K and decays slowly to RT within a ps timescale, dictated by $e$-ph scattering. In Fig. 3(b), the fractional changes of SLG absorption, thermal capacity, and chemical potential are shown, with baseline values $\alpha_{\text{SLG}} = 0.0041, c_{\text{e-SLG}} = 2.34 \times 10^{-11}$ Jcm$^{-2}$K$^{-1}$, and $\mu_{\text{SLG}} = 0.15$ eV. A critical parameter is $c_{\text{e-SLG}}$, whose sharp increase with temperature (about one order of magnitude at $T_{\text{e-SLG}} \sim 1200$ K) limits the achievable $T_{\text{e-SLG}}$. By comparison, the SLG absorption reduction of $\sim 20\%$ at peak $T_{\text{e-SLG}} \sim 1200$ K is less severe. Figure 3(c) shows the thermal currents created. In Fig. 3, cooling is dominated by $J_{\text{e-ph}}$ while $J_{\text{th}}$ [Eq. (28)] is negligible. This is due to the imbalance between cooling and injection times: $\tau_{e-ph} \sim 1.3$ ps and $\tau_{\text{inj}} \sim 30$ ps, which leads to ineffective hot carrier injection across the Schottky barrier and $J_{\text{e-ph}} \gg J_{\text{th}}$. We will explore different carrier injection times (1 fs $\leq \tau_{\text{inj}} \leq 30$ ps) and the crossover between the two cooling regimes (i.e., $J_{\text{e-ph}}$ versus $J_{\text{th}}$).

To test the effect of pulse duration on PD peak responsivity ($R_{\text{peak}}$), defined as the peak $R_{\text{ext}}$ during the pulse duration, and to define a quasi-cw approximation limit, Fig. 4 plots the normalized responsivity $R_{\text{peak}}/R_{\text{cw}}$ (where $R_{\text{cw}}$ is the external responsivity under the quasi-cw approximation) at fixed power $P_{\text{in}} = 10^8$ W/cm$^2$ as a function of pulse duration for different $\tau_{\text{inj}}$. The other simulation parameters are as in Fig. 3. For $\tau_{\text{inj}} > 1$ ps, the cooling is dominated by $e$-ph scattering, and the quasi-cw approximation is valid only for pulse durations $> 10$ ps. On the other hand, the sub-ps pulses cannot be described by a quasi-cw approximation for any $\tau_{\text{inj}}$.

We next consider a SLG/$n$-Si Schottky PD integrated into an asymmetric Bragg cavity at critical coupling (Fig. 5 inset schematic) to allow $\sim 100\%$ SLG optical absorption [76–80]. The structure and layers thicknesses are optimized to cavity resonance at $\lambda = 3.1$ $\mu$m ($\hbar \omega = 0.4$ eV), i.e., $d_{\text{SiO}_2} = 530.7$ nm, $d_{\text{n-Si}} = 226.6$ nm, $c_{\text{SiO}_2}^{\text{cavity}} = 1061.5$ nm, and $d_{\text{n-Si}}^{\text{cavity}} = 453$ nm. The $n$-Si layer ($N_{\text{d-Si}} = 10^{16}$ cm$^{-3}$) in contact with SLG forms a Schottky junction, and the combined contribution to optical absorption is taken into account (see Appendix C). All other Si layers in contact with $\text{SiO}_2$ to realize the Bragg mirrors are assumed undoped. We consider the PD area equal to the diffraction limit $\lambda^2/\pi \sim 10^{-7}$ cm$^2$ fully covered by SLG.

Figure 5 plots the calculated $J_{\text{el}}$ as a function of $P_{\text{in}}$ for quasi-cw illumination ($\tau_{e-e} = 20$ fs, $\tau_{\text{opt}} = 200$ fs, $\tau_{\text{inj}} = 30$ ps, $\Phi_{B} = 0.3$ eV, and $k_{\text{F-SLG}} l \sim 50$). At low input powers ($P_{\text{in}} < 10^7$ W/cm$^2$ down to the NEP level), $J_{\text{el}}$ is linear with $P_{\text{in}}$, so the PD exhibits a linear dynamical range (i.e., the power range of constant $R_{\text{ext}}$) over five decades for $10^{-3}$ W/cm$^2 < P_{\text{in}} < 10^5$ W/cm$^2$. At higher $P_{\text{in}} > 10^2$ W/cm$^2$, the PD response becomes nonlinear and $R_{\text{ext}}$ increases exponentially by almost three orders of magnitude. In this range, a substantial increase in $T_{\text{e-SLG}}$ leads to exponentially enhanced thermionic emission across the Schottky barrier, as also evident from the Richardson equation [14], i.e., $J_{\text{el}} \propto \exp[-\Phi_{B}/(k_{B}T_{\text{e-SLG}})]$ (see also Appendix E). $R_{\text{ext}}$ reaches its maximum at a certain $P_{\text{in}}$, beyond which it declines. This decline is understood as a combination of several factors occurring as $T_{\text{e-SLG}}$ increases: (i) the exponential term saturates (i.e., when $k_{B}T_{\text{e-SLG}}$ becomes comparable to $\Phi_{B}$), (ii) the SLG absorption drops, and (iii) the SLG thermal capacitance increases significantly [see Fig. 3(b)], requiring increasingly larger input power to achieve the same $T_{\text{e-SLG}}$ increase. Therefore, $T_{\text{e-SLG}}$ and $J_{\text{el}}$ continue to rise with increased $P_{\text{in}}$, but at a progressively slower sublinear rate, resulting in the $R_{\text{ext}}$ drop. In our discussion, $P_{\text{in}}$ is still below the threshold where SLG saturable absorption [123,124] starts to dominate and becomes a limiting factor to the PD performance.

If the SLG electrode is nanostructured (e.g., ribbons narrower than $\lambda/2$), its spatially averaged absorptivity will decrease (assuming no resonant phenomena, such as plasmons [125], become active), but critical coupling conditions can still be met if the cavity decay rate is reduced to match the new total absorption rate. This is done by adding Si/$\text{SiO}_2$ layers on either side of the cavity [78] to increase the reflectance of Bragg mirrors. Figure 5 plots the resulting $R_{\text{ext}}$ for SLG.
reduced below the diffraction limit for \( F = 4, 20, 100 \). As more multilayer periods are added to achieve critical coupling, the corresponding absorption densities increase by a factor \( F \), giving rise to higher \( T_{\text{SLG}} \) (see Appendix C). At the same time, the SLG/Si contact area is proportionally smaller, limiting the overall current (both bright and dark) across the junction. As a result, the peak \( R_{\text{ext}} \) remains the same, but is observed at a smaller input power and with a smaller dark current, both reduced by a factor \( F \).

Next, we consider the effect of \( \tau_{\text{inj}} \) on PD performance. Figure 6(a) plots the normalized \( J_{\text{ph}} \) and \( J_{\text{th}} \) as a function of \( \tau_{\text{inj}} \) for different \( F \). For simplicity, we assume \( F = 1 \) and 100% SLG absorption. While \( \tau_{\text{inj}} \approx 0.13 \, \text{ps} \) is the lowest value reported to date [22,114,119], we anticipate that future studies may do better, so we extend our work into the fs regime. The crossover between phonon-dominated \( (J_{\text{ph}} > J_{\text{th}}) \) and thermionic-dominated \( (J_{\text{ph}} < J_{\text{th}}) \) cooling regimes is observed in all cases [Fig. 6(a)] but at a different \( \tau_{\text{inj}} \), depending on \( P_{\text{in}} \). For larger \( P_{\text{in}} \), \( T_{\text{SLG}} \) increases, leading to enhanced thermionic cooling, and the crossover point is shifted to the higher \( \tau_{\text{inj}} \). The corresponding \( T_{\text{SLG}} \) and \( R_{\text{ext}} \) are shown in Fig. 6(b). \( R_{\text{ext}} \approx 1 \, \text{A/W} \) is achievable for \( \tau_{\text{inj}} \) in the sub-ps regime. A smaller \( \tau_{\text{inj}} \) results in lower NEP and higher \( D^* \) [Fig. 6(c)]. In the limit of \( \tau_{\text{inj}} \approx 1 \, \text{fs}, D^* \) approaches \( 10^8 \, \text{cmHz}^{-0.5} / \text{W} \), a value typical in SLG-based IPE PDs [36,37], thus indicative of the potential competitiveness of the PTh configuration.

These results impose an upper limit for thermionic emission in SLG/Si Schottky PDs for a given SBH. The latter depends on the initial \( E_{\text{FSLG}} \) upon contact and the applied reverse bias \( V_R \), which both affect \( \Phi_B \) and the leakage (dark) current. For \( V_R = 0, \Phi_B \) depends on SLG Fermi level before contact \( E_{\text{FSLG}} \) and \( n\text{-Si} \) dopant concentration \( N_{d,\text{Si}} \), having a lower limit \( \Phi_B \approx 0.2 \, \text{eV} \) [Eq. (2)], depending on \( N_{d,\text{Si}} \) (assuming the ideal case of no charges and Fermi level pinning at the interface).

Figure 7 plots the linear regime responsivity \( R_{\text{lin}} \) (at low input power, \( P_{\text{in}} < 10^2 \, \text{W/cm}^2 \)) and peak responsivity \( R_{\text{peak}} \) (i.e., derived at \( P_{\text{in}} > 10^2 \, \text{W/cm}^2 \) for \( F = 1 \)) as a function of \( \Phi_B \). We consider two cases for \( \tau_{\text{inj}} \approx 30 \, \text{ps} \) and the lower bound of 0.13 ps. The span of possible SBH achievable for different \( N_{d,\text{Si}} \) is denoted in Fig. 7 by a vertical dashed line (i.e., \( \Phi_B = 235 \, \text{meV} \) and \( 150 \, \text{meV} \) for \( N_{d,\text{Si}} = 10^{16} \) and \( 3 \times 10^{17} \, \text{cm}^{-3} \), respectively). \( R_{\text{lin}} > 0.1 \, \text{A/W} \) is reached for \( \tau_{\text{inj}} = 0.13 \, \text{ps} \) at \( \Phi_B < 0.25 \, \text{eV} \) even in the linear (low power) regime. \( R_{\text{lin}} \) and \( R_{\text{peak}} \) plateau at lower SBH for both \( \tau_{\text{inj}} \). This is due to more efficient cooling at higher \( E_{\text{FSLG}} \), more pronounced in the linear regime \( (T_{\text{SLG}} < 310 \, \text{K}) \), where a peak at \( \Phi_B \approx 0.2 \, \text{eV} \) is obtained for \( \tau_{\text{inj}} = 0.13 \, \text{ps} \). Such behavior is expected when the thermionic cooling becomes dominant in combination with Pauli blocking at \( E_{\text{FSLG}} > 0.2 \, \text{eV} \) (\( \Phi_B < 0.25 \, \text{eV} \)). The latter contrasts with the high \( T_{\text{SLG}} > 800 \, \text{K} \) case, where Pauli blocking is suppressed by the thermally...
excited carriers and decreasing $\mu_{c,SLG}$ [see Fig. 3(b)]. For $\tau_{inj} = 30 \text{ ps}$, cooling is phonon dominated and does not result in a performance drop by further increasing $E_{F0,SLG}$.

SBH can also be controlled by $V_R$. The dependence of SBH on $V_R$ is plotted in Fig. 12(b) of Appendix A. Figure 8 plots $R_{peak}$ as a function of $V_R$ for different $\Phi_{B0}$. We consider the two cases $\tau_{inj} = 30 \text{ ps}$ and $0.13 \text{ ps}$, assuming, for simplicity, $\tau_{inj}$ independent of $V_R$. For $N_{d,SLG} = 10^{16} \text{ cm}^{-3}$, avalanche breakdown is expected at $V_R \sim 50 \text{ V}$ (see Appendix A), so $V_R < 25 \text{ V}$ is considered to be safely applied. The operation near the breakdown regime, including possible avalanche multiplication effects [126] from impact ionization in Si, is not taken into account. For $\tau_{inj} = 30 \text{ ps}$, $R_{peak}$ is limited to the mA/W range. For $\tau_{inj} = 0.13 \text{ ps}$, $R_{peak}$ is limited to 1 A/W can be reached. Higher $E_{F0,SLG}$ and/or $V_R$ both contribute to SBH lowering, thus higher $R_{peak}$. In Fig. 8, the slope changes at $V_R = 4 \text{ V}$ for $\Phi_{B0} = 0.52 \text{ eV}$ and at $V_R = 21 \text{ V}$ for $\Phi_{B0} = 0.58 \text{ eV}$. This is attributed to SLG transitioning from $p$ to $n$ as $V_R$ increases (for $\Phi_{B0} \leq 0.45 \text{ eV}$ SLG is already $n$ doped at $V_R = 0$) and $E_{F,SLG}$ shifts from valence to conduction band via the CNP. During this transition, the rate of $\mu_{SLG}$ variation with $V_R$ changes due to the smaller density of states around the CNP (see also Fig. 12 in Appendix A).

Another important factor affecting $R_{peak}$ is $k_{F,SLG} l$ [85,86], related to supercollision cooling $J_{SC}$ [Eqs. (11) and (12)], with $k_{F,SLG} = \sqrt{\pi n_{F0,SLG}^2}$, where $n_{F0,SLG}^2$ is the free carrier concentration associated with SLG doping. Figure 9 shows the $R_{peak}$ dependence on $k_{F,SLG} l$ for $\Phi_{B0} = 0.3 \text{ eV}$ (by varying $l$) for two representative $\tau_{inj}$ and $V_R$. In the slow $\tau_{inj}$ regime (i.e., $30 \text{ ps}$), where cooling is dominated by $e$-$\text{ph}$ scattering, as $k_{F,SLG} l \rightarrow 0$, supercollision cooling is enhanced and becomes dominant (surpassing $e$-optical phonon scattering and thermionic cooling). This leads to a drop in $R_{peak}$. As $k_{F,SLG} l$ increases, for example, for a lower density of defects, supercollision scattering is suppressed. At a certain point, the $e$-optical phonon scattering becomes dominant, and $R_{peak}$ plateaus, as $k_{F,SLG} l$ increases further. For the fast $\tau_{inj}$ (i.e., $0.13 \text{ ps}$), cooling is dominated by the thermionic heat flow ($J_{th}$) from SLG to Si, resulting in large $R_{peak}$, but this quickly deteriorates as $k_{F,SLG} l \rightarrow 0$.

Figure 10 plots $D^*$ as a function of $V_R$ for different $\Phi_{B0}$. $D^*$ strongly depends on the PD noise level. In our calculations, we consider shot and thermal (Johnson) noise contributions. Shot noise is proportional to the PD current [9], while thermal noise is inversely proportional to the diode resistance [9]. At low $V_R$, the PD dark (leakage) current is reduced and the Schottky junction resistance is increased, so the total noise figure decreases, and $D^*$ is expected to grow for $V_R \rightarrow 0$. Figure 10(a) shows that this is valid only for $n$-doped SLG (e.g., $\Phi_{B0} = 0.3,0.35,0.4 \text{ eV}$). In contrast, when the SLG/Si Schottky contact results in $p$-doped SLG (e.g., $\Phi_{B0} = 0.52,0.58 \text{ eV}$), $D^*$ increases with $V_R$. This is understood as follows: for $p$-doped SLG, a larger $V_R$ increases both photo- and dark currents (thus noise levels), but at the same time shifts $E_{F,SLG}$ toward the CNP, weakening both $e$-optical phonon and supercollision-assisted heat dissipation channels [Eqs. (9)–(12)], making the SBH lowering effect more profound, resulting in enhanced $R_{eq}$ (see Appendix F). These counteracting phenomena lead to increased $D^*$ with $V_R$. These effects are more pronounced for $\tau_{inj} = 30 \text{ ps}$ [Fig. 10(a)], where $J_{\text{ph}}$ cooling dominates, and less for $\tau_{inj} = 0.13 \text{ ps}$ [Fig. 10(b)], where thermionic cooling is important. Overall, the effect of $V_R$ on $D^*$ depends on $E_{F,SLG}$ relative to Si: it will
enhance $D'$ when the initial $E_{F,SLG}$ and $E_{F,SlG}$ are of opposite signs, but it will reduce it in the opposite case. $R_{ext}$, on the other hand, always increases with $VR$, as for Figs. 8 and 9.

Finally, we estimate the limiting factors defining the time response. These are (1) $\tau_r$, the transit time of charge carriers across the depletion zone; (2) $\tau_{RC}$, the charge/discharge resistance-capacitance (RC) time constant of the diode/circuit combination; and (3) $\tau_{ph}$, the photon lifetime inside the optical cavity. The overall time response is thus limited by [27]:

$$\tau = \sqrt{\tau_r^2 + \tau_{RC}^2 + \tau_{ph}^2}. \quad (30)$$

$\tau_r$ can be estimated as $\tau_r = \chi_d/v_{sat,SlG}$, where $\chi_d$ is the depletion region width and $v_{sat,SlG} = 10^7 cm/s$ the carrier saturation velocity in Si [9,37]. The depletion region varies with $VR$ and initial SLG doping (i.e., $\chi_d \propto (V_0 + V_F^{(0.5)})$). $\chi_d$ ranges from 150 nm (at $VR = 0$) up to the fully depleted nSi layer thickness $d_{SlG}$, i.e., $\chi_d = d_{SlG} = 453$ nm (see Appendix G). These values bring $\tau_r$ between 1.5 and 4.5 ps, with the corresponding transit-time limited cutoff frequency $f_{tr} = (2\pi \tau_r)^{-1} \sim 35 - 100$ GHz.

On the other hand, the RC time constant enforces a limiting frequency $f_{RC} = (2\pi R_{SlG} C_J)^{-1}$, where $R_{SlG} = R_s + R_e$ is the sum of series and contact resistances, $C_J = \varepsilon_0 \varepsilon_r \varepsilon_r/\chi_d$ is the junction capacitance, $S_c \sim 10^{-7} cm^2$ is the device area, $\varepsilon_0$ is the vacuum permittivity, and $\varepsilon_r$ is the Si dielectric constant (quantum capacitance does not have a significant contribution, see Appendix G). For $R_e$, we estimate the Si contribution to be $R_S_{SlG} = (d_{SlG} N_d \mu_{SlG} e_{SlG})^{-1} \sim 10 k\Omega/\square$ and the SLG contribution $R_{SlG} = \sigma_{SlG}/\delta_{SlG}$ between 0.1 and 1 k\Omega, where $\sigma_{SlG}$ is the dc SLG conductivity (see Appendix G). Assuming $R_e \sim 1 k\Omega \mu m [109,119,127]$, we get $R_e \ll R_s$. Depending on configuration (i.e., initial SLG doping, $V_F$) $f_{RC}$ is estimated in the range 1–10 GHz, but can be increased by reducing the SLG footprint (thus capacitance), following the geometrical focusing approach. Another way to overcome the frequency limit imposed by $f_{RC}$ would be by increasing Si doping and lowering $R_{SlG}$. Bringing $R_{SlG}$ to the same value as $R_{SlG}$ would lead to an almost sixfold increase in $f_{RC}$.

The optical cavity limited frequency response is given by [78] $f_{opt} = c/(n2LQ)$, where $c/n$ is the speed of light inside the cavity, $L$ is the cavity length, and $Q$ the cavity quality factor (see Appendix G). Depending on the chosen configuration (i.e., geometrical focusing), $f_{opt}$ ranges from 10 GHz (for $F = 100$) up to 1 THz (for $F = 1$). Consequently, the time response limitations are mainly imposed by $f_{RC}$, bringing overall operation frequencies in the range of 1–100 GHz, depending on configuration.

IV. CONCLUSIONS

We assessed the behavior and performance limits of SLG/Si Schottky IR PDs operating in the thermionic regime. We described the dynamics under realistic conditions in both biased and unbiased cases, and extracted the performance figures of merit. In the proposed Bragg cavity configuration, $R_{ext}$ up to $\sim 1 A/W$ is achievable in the nonlinear (high-power) regime and $D'$ peaks at $\sim 10^7 cmHz^{-0.5}/W$, with a seven-decade linear dynamic range at low ($< 10^2 W/cm^2$) input powers. Our results and conclusions are general and applicable throughout the IR, if the architecture length scale is appropriately adjusted to the corresponding wavelength. This PD platform is promising for multispectral detection at RT using CMOS compatible hybrid SLG-Si technology.

ACKNOWLEDGMENTS

We acknowledge funding from the EU Graphene and Quantum Flagships, the Israel Science Foundation (Grant No. 1732/18), the Israel Innovation Authority (Grant No. 63350), ERC grant Hetero2D, EPSRC Grants No. EP/L016087/1, No. EP/K01711X/1, No. EP/K017144/1, No. EP/N010345/1, No. EP/V000055/1, and DSTL.

APPENDIX A: SCHOTTKY JUNCTION FORMATION

Equations (1)–(4) are self-consistently solved to get $E_{F,SLG}$, $\mu_{SLG}$ and $\Phi_B$ as a function of $E_{F,SLG}$, $T_{SlG}$ and $VR$. To validate our calculations, we reproduce the results of Ref. [74] and show in Fig. 11 the calculated change in SBH due to the $E_{F,SlG}$ shift, $\delta \Phi_B$, as a function of $VR$.

In Ref. [74] SLG was contacted with pSi ($N_{a,SlG} = 3 \times 10^{19} cm^{-3}$) and initially p doped with h concentration $n_{0,SlG} = 3.5 \times 10^{12} cm^{-3}$. Thus, $VR$ led to reduced $E_{F,SlG}$.
and SBH. This is equivalent to our nSLG–nSi configuration, except that Eq. (1) needs to be changed to:

$$eV_0 = \chi_{\text{Si}} + E_{g,\text{Si}} - \chi_{\text{SLG}} - |E_{F,0,\text{SLG}} - \delta E_{F,\text{SLG}}| - e\phi_{\text{Si}}$$  \hspace{1cm} (A1)

with $E_{g,\text{Si}} = 1.1$ eV the Si band gap [9], $E_{F,0,\text{SLG}} = -0.25$ eV and $\delta E_{F,\text{SLG}}$ the change in $E_{F,\text{SLG}}$ induced by the contact with Si (in both biased and unbiased cases). In the pSi case, $e\phi_{\text{Si}}$ is given by [9]:

$$e\phi_{\text{Si}} = k_B T_{c,\text{Si}} \ln \left( \frac{N_{\text{d,Si}}}{N_{\text{a,si}}} \right),$$ \hspace{1cm} (A2)

with $N_{\text{d,Si}}$ and $N_{\text{a,si}}$ the acceptor concentration and the effective density of states at the valence band edge, respectively. Apart from these changes, the core of our algorithm remains unchanged, reproducing Ref. [74]. Both in Ref. [74] and here, additional mechanisms of SBH modulation, such as image force lowering [9,14] and SBH inhomogeneity [14], are neglected since their contribution is small [14].

In the main text, we assumed $V_R$ up to 25 V. In our configuration (contact with nSi, $N_{\text{d,si}} = 10^{16}$ cm$^{-3}$), reverse bias breakdown occurs at $V_R \sim 51$ V, as calculated from [9]:

$$V_R^{\text{BD}} = \frac{\epsilon_0 E_{c,\text{cr}}}{2eN_{\text{d,si}}},$$ \hspace{1cm} (A3)

where $E_{c,\text{cr}}$ is the critical breakdown field [9]:

$$E_{c,\text{cr}} = \frac{4 \times 10^5}{1 - \frac{1}{3} \log_{10} \left( \frac{N_{\text{d,si}}}{10^{10}} \right)} \text{ [V/cm].}$$ \hspace{1cm} (A4)

We assume that the device safely operates up to half $V_R^{\text{BD}}$. Our approach also handles the SLG transition from $p$ to $n$ by $V_R$. Figure 12 shows the results for SBH, $\mu_{\text{SLG}}$ (assuming no photoexcited carriers, i.e., $\mu_{\text{SLG}} = \mu_{\text{e,SLG}} = \mu_{\text{c,SLG}}$), and dark current density across the junction for different $E'_{F,\text{SLG}}$ (i.e., the initial SLG Fermi level before Si contact).

**APPENDIX B: ELECTRON-PHONON COOLING**

Heat dissipation through e–ph interactions is described using a two-temperature model in Eqs. (5) and (6) [41]. We account for both optical phonon and supercollision scattering using Eqs. (8)–(12), with Eq. (12) considering the total number of carriers to describe supercollision scattering in the non–degenerate limit. An alternative approach is to average all e–ph interactions (i.e., supercollision scattering, optical, and acoustic phonons) using a single mean e–ph relaxation time. This can be done by modifying Eq. (8) to:

$$J_{e,\text{ph}} = \frac{C_{e,\text{SLG}}}{\tau_{e,\text{ph}}} \left( T_{c,\text{SLG}} - T_{l,\text{SLG}} \right)$$ \hspace{1cm} (B1)

with $\tau_{e,\text{ph}}$ the average relaxation time. The latter can be treated as free parameter, mainly depending on SLG quality (i.e., defects, wrinkles, and contaminants), similarly to the case of $l$ for short-range scatterers [Eq. (12)]. Larger $l$ (or longer $\tau_{e,\text{ph}}$) will lead to reduced e–ph heat dissipation and better PD performance, Fig. 9. This behavior is more profound when the carrier injection time is large (e.g., for the $\tau_{\text{inj}} = 30$ ps case).

FIG. 11. SBH shift due to $V_R$ in a SLG/Si junction, presented in Ref. [74] (green squares) and our calculations (red solid line).

FIG. 12. (a) $\mu_{\text{SLG}}$, (b) SBH, (c) dark current density as a function of $V_R$ and $E'_{F,\text{SLG}}$. 
Je the calculations using Tl lines using Eq. (B1) (average relaxation time). Although the estimation (underestimation) of R using the explicit phonon channels, Eqs. (8)–(12), and dashed carrier cooling is dominated by e-ph scattering. Figure 13 compares the two approaches (i.e., explicit phonon channels versus single average relaxation time) using the same configuration as in Fig. 9. We adjust τ_e-ph to have the same overall e-ph cooling rate at RT as the explicit phonon approach. We find that for l = 200 nm, the corresponding τ_e-ph is ∼ 1.3 ps, matching the values reported in previous works [62,67,128]. The resulting R_{ext} versus input power density and e-ph thermal current density versus T_{e,SLG}, assuming E_{F,SLG} = 0.15 eV, are shown in Fig. 13. Solid lines correspond to calculations using the explicit phonon channels, Eqs. (8)–(12), and dashed lines using Eq. (B1) (average relaxation time). Although the single mean e-ph relaxation time approach leads to an over-estimation (underestimation) of R_{ext} (J_e-ph), both approaches lead to qualitatively similar results, including peak values and performance saturation.

\[ c_{i,SLG}(T_{i,SLG}) \] is calculated from Ref. [90]. In the range of interest for our work (T_{i,SLG} = 300 – 500 K), c_{i,SLG} scales linearly with T_{i,SLG}. We use the following parametrization [90]:

\[
    c_{i,SLG} = -1.4 \times 10^{-5} + 1.9 \times 10^{-6} T_{i,SLG} \quad [\text{J/m}^2].
\]  

The lattice cooling rate into Si, \( \Gamma_{SLG-Si} \sim 20 \text{ MW/m}^2\text{K} \), is estimated using finite element heat transfer simulations and is in good agreement with Refs. [41,90].

**APPENDIX C: CRITICAL COUPLING AND GEOMETRICAL FOCUSING**

The optical absorption of SLG can be enhanced with the use of Bragg cavities [78–80]. We assume an asymmetric Bragg cavity, as presented in the inset of Fig. 5, consisting of two periods of Si/SiO_2 layers on either side, with an Au mirror on the back. SLG is sandwiched between a SiO_2 and a n-doped Si layer of double thickness compared to the Bragg layers. The Bragg layer thicknesses are d_Si = 226.6 nm and d_{SiO2} = 530.7 nm, following the quarter-wave rule for the target \( h_0 = 0.4 \text{ eV} \), using the refractive indexes n_Si = 3.42 [9] and n_{SiO2} = 1.46 [9]. The Au back mirror has d_Au = 100 nm and n_Au = 1.73 + 19.2i at \( h_0 = 0.4 \text{ eV} \) [129]. Only the Si layer beneath SLG is doped, while the other Si layers are assumed intrinsic. For \( F = 1 \), SLG covers an area equal to the diffraction limit (\( S_0 \sim 10^{-7} \text{ cm}^2 \)) and its permittivity is modeled using the Kubo formula, as for Eqs. (22)–(24). The highly doped Si layer contribution to the optical absorption is also considered, using a Drude model for its dielectric function [78]. We get the SLG optical absorption under normal incidence by employing the Fresnel relations using the transfer matrix method [78]. In Fig. 14, we plot the absorption in each element as a function of \( E_{F,SLG} \) at RT. Our device is effectively a single port system, since transmission is suppressed by the Au back mirror. By tuning the SLG absorptivity, one can achieve critical coupling [76–80] and absorb all incoming radiation. Figure 14 shows that \( \sim 100\% \) incoming light is absorbed. Most absorption occurs in SLG, which increases \( T_{e,SLG} \), enabling thermionic emission through the Schottky junction.

For critical coupling, the decay rate of the resonator must match the total absorption rate of its elements [Eq. (29)]. By patterning SLG into ribbons, the absorption rate is reduced. A straightforward way to reduce the decay rate of the cavity is adding more Bragg layers. For \( F = 1 \), we use 2 Bragg periods, while for \( F = 4, 20, 100 \) we use 3, 4, 5, respectively. Absorption in SLG in these cases can be up to 95% (Fig. 14), while the absorption density is enhanced by a factor \( F \). Conversely, the electronic current (both dark and bright) through the junction will be reduced, due to the smaller contact area. Overall, these will lead to the same \( R_{ext} \), but achieved at a reduced input power by a factor \( F \), as presented in Fig. 5. This allows us to achieve optimum performance, requiring less power input.

**APPENDIX D: WAVELENGTH DEPENDENCE ON PD RESPONSE**

We now investigate the wavelength dependence of our thermionic PD platform. Figure 15 plots the SLG absorption in the range of \( h\omega = 0.38–0.42 \text{ eV} \). The Bragg cavities (composed by Bragg mirrors with \( N = 2 \) Si/SiO_2 repetitions...
(each) are scaled for each target wavelength, based on the methods described in Appendix C. For simplicity, we restrict our analysis to $F = 1$ and $E_{F,\text{SLG}} = 0.15$ eV. Almost total light absorption can be achieved, as can be seen from the continuous lines in Fig. 15. The dashed lines correspond to the calculated SLG absorption using 3 Bragg bilayers for each mirror. In the latter case, the decay rate of the cavity is reduced, critical coupling conditions are not met, and SLG absorption decreases. In either case, however, one can tune $E_{F,\text{SLG}}$ and adjust SLG absorptivity to reach critical coupling, thus perfect absorption.

Figure 16(a) plots the SLG absorption as a function of photon energy, from quasi-cw calculations as described in the main text. We assume $E_{F,\text{SLG}} = 0.15$ eV, $F = 1$, $\tau_{e-e} = 20$ fs, $\tau_{\text{opt}} = 200$ fs, $\tau_{\text{nj}} = 0.13$ ps, $\Phi_B = 0.3$ eV, $k_{F,\text{SLG}} l \sim 50$ for both $N = 2$ and $N = 3$ Bragg bilayers. Dashed lines correspond to the absorption in the linear regime (low input power, $T_{e,\text{SLG}} \sim 301$ K) and solid lines for input power corresponding to the peak of $R_{\text{ext}}$ (high input power, $T_{e,\text{SLG}} \sim 1200$ K). We observe that, at lower $\hbar \omega$, SLG absorption is reduced in both cases, due to Pauli blocking ($\hbar \omega < 0.3$ eV). Nonetheless, we can compensate for the reduced SLG absorption rate by adding a third Bragg bilayer, thus reducing the cavity decay rate. In this case, critical coupling conditions are restored in these lower photon energies. Figure 16(b) plots $R_{\text{ext}}$ as a function of $\hbar \omega$ in the linear regime (dashed lines) and at its peak (solid lines).
As expected, the $R_{ext}$ trend follows that of SLG absorption. The black lines, labeled with $R_{ext}^{(100\%)}$, show the limiting case of 100% SLG absorption at all photon wavelengths. In the linear regime, $R_{ext}^{(100\%)}$ is flat, as expected, while in the nonlinear one we observe a slight decrease with increasing photon energy. This is due to the excited nonequilibrium carriers: for the same input power a higher number of photons are absorbed at lower $\hbar\omega_0$. As a result, a higher nonequilibrium carrier density is involved [Eq. (13)], resulting in a higher $\mu_c$ [Eq. 20(a)]. Since SLG is contacted with $n\text{Si}$, the higher $\mu_c$ will lead to reduced SBH $\Phi_B$, thus a slightly increased thermionic emission for lower $\hbar\omega_0$.

These results show that our methods and conclusions are applicable throughout the IR spectrum. By exploiting the critical coupling mechanism it is possible to design PDs with the lowest input power, i.e., at NEP level, under bias. For both cases, $V_R$ leads to an increased $E_{F,\text{SLG}}$, thus lower SBH $\Phi_B$, and materials used in Bragg cavities.

**APPENDIX E: LANDAUER TRANSPORT FORMALISM**

The electronic current density across the SLG–Si junction is calculated using the Landauer transport formalism [Eq. (27)]. At low $T_{e,\text{SLG}}$ ($k_BT_{e,\text{SLG}} \ll \Phi_B$), Eq. (27) can be solved analytically, resulting in the ideal diode equation [22,41]:

$$J_{el} = A'[\mathcal{R}(T_{e,\text{Si}})e^{\nu_0/k_BT_{e,\text{Si}}} - \mathcal{R}(T_{e,\text{SLG}})].$$

(E1)

where $A' = 2\epsilon k_B^2/(\tau_{inj}\pi h^2v_{F,\text{SLG}}^2)$ and:

$$\mathcal{R}(T) = e^{-\nu_0/k_BT}T^2\left(\frac{\Phi_{\text{CNP}}}{k_BT} + 1\right)e^{-\nu_0/k_BT}.$$  

(E2)

provided that $E - \mu_{c,\text{SLG}} \approx E - \mu_{\text{Si}} \gg k_BT$. For $T_{e,\text{SLG}} = T_{e,\text{Si}} = T$ this simplifies to the Richardson equation [9]:

$$J_{el} = J_0(e^{\nu_0/k_BT} - 1).$$

(E3)

with saturation current:

$$J_0 = A'T^2\left(\frac{\Phi_{\text{CNP}}}{k_BT} + 1\right) e^{-\nu_0/k_BT}.$$  

(E4)

A similar analytical expression can be derived for the thermal current density across the junction [corresponding to Eq. (28)]:

$$J_{th} = A_{th}'[\mathcal{R}_{th}(T_{e,\text{SLG}}) - \mathcal{R}_{th}(T_{e,\text{Si}}) e^{\nu_0/k_BT_{e,\text{Si}}}] - \mathcal{R}_{th}'(T_{e,\text{SLG}}) + \mathcal{R}_{th}'(T_{e,\text{Si}}) e^{\nu_0/k_BT_{e,\text{Si}}}],$$

(E5)

where $A_{th}' = 2/(\tau_{inj}\pi h^2v_{F,\text{SLG}}^2)$ and:

$$\mathcal{R}_th(T) = e^{-\nu_0/k_BT}k_BT^3\left[\left(\frac{\Phi_{\text{CNP}}}{k_BT}\right)^2 + \frac{2\Phi_{\text{CNP}}}{k_BT} + 2\left(1 - e^{-\nu_0/k_BT}\right).$$

(E6)

$$\mathcal{R}_{th}'(T) = e^{-\nu_0/k_BT}k_BT^2\left(\mu_{c,\text{SLG}}\frac{\Phi_{\text{CNP}}}{k_BT} + 1\right)e^{-\nu_0/k_BT}. $$

(E7)

Figure 17 compares the current densities calculated using Landauer transport [Eqs. (27) and (28)] and extracted analytical relations [Eqs. (E1) and (E5)] as a function of $T_{e,\text{SLG}}$. The calculations use the same parameters as in Fig. 5 for $F = 1$, assuming $E_{F,\text{SLG}} = 0.15 \text{ eV}$ and $\tau_{inj} = 30 \text{ ps}$. The two curves slightly diverge for $T_{e,\text{SLG}} > 2000 \text{ K}$. The overall agreement validates the simplified diode equation for low ($k_BT_{e,\text{SLG}} \ll \Phi_B$) temperatures.

**APPENDIX F: RESPONSE AT LOW INPUT POWER**

We consider the response of $n\text{SLG}$ and $p\text{SLG}$ PDs at the lowest input power, i.e., at NEP level, under bias. For both cases, $V_R$ leads to an increased $E_{F,\text{SLG}}$, thus lower SBH. This is more profound in the $p\text{SLG}$ case [see also Fig. 12(b)] because
of the reduced charge density as one approaches the Dirac point. Also, in the nSLG case, the total number of carriers available for SC cooling as well as the ability for e-optical phonon scattering (ΩK = 161 meV and ΩF = 196 meV) [94] is increased, while in pSLG the additional e transfer from Si to SLG will lead (initially) to a reduced number of carriers and scattering due to optical phonon, thus reduced cooling. Both these effects lead to increased Ret with V R in the pSLG case, and reduction in the nSLG case, as shown in the D* graphs of Fig. 10. After D* peaks in the pSLG case, further increase in V R leads to a reduction, for the same reasons it reduces in nSLG. A critical factor determining the peak position is the thermionic cooling mechanism becoming dominant. Figure 18 compares the two cooling channels (e-ph and thermionic) as a function of V R for two E′ F,SLG, assuming input power at the NEP level. We verify that in nSLG, thermionic cooling is always dominant (except close to V R = 0), while in pSLG a crossover is found for V R ≃ 22 V. Thus, the combined counteracting phenomena described above lead to an initial increase in D* with V R in the pSLG/nSi. For nSLG/nSi, all V R values lead to lower D*. This is more profound in the high injection time scenario, since for low τ inj = 0.13 ps, thermionic cooling is more important than e-ph cooling.

APPENDIX G: TEMPORAL RESPONSE LIMITATIONS

The temporal response limitations for the proposed configuration are given in Eq. (30) and have 3 contributions:

1. Transition time of carriers through the depletion region: \( \tau_d = \chi_d / \nu_{sat,Si} \), which varies depending on the value of the depletion region width [9]:

\[
\chi_d = \frac{2e_{Si}}{\epsilon N_{d,Slg}} \left( V_0 + V_R - k_B T_{e,Slg} \right).
\]

(\( G1 \))

\( \chi_d \) will vary mainly due to the different \( V_R \) from a minimum \( \sim 150 \) nm up to the whole thickness of the \( nSi \) layer (\( d_{nSi} = 453 \) nm). Given that the carrier saturation velocity in Si is \( \nu_{sat,Si} = 10^7 \) cm/s [9], the transition time through the depletion region lies in the range \( \sim 1.5 - 4.5 \) ps, with the corresponding limiting frequency in the range \( f_\text{c}(2\pi \tau_d)^{-1} \sim 35 - 100 \) GHz.

2. Electronics limited frequency is given by \( f_B = (2\pi R_C C_j)^{-1} \). \( R_C = R_s + R_e \) is the sum of the series and contact resistances. \( R_s \) has two contributions, one due to nSi and one due to SLG. The Si resistance is \( R_S = (d_{nSi} N_{d,Slg} e \mu_{e,Slg})^{-1} \sim 10 \) kΩ/cm, assuming \( N_{d,Slg} = 10^{16} \) cm\(^{-3} \) with the electron mobility in Si \( \mu_{e,Slg} = 1350 \) cm\(^2\)/Vs [9]. For SLG, we use \( R_{SLG} = (\sigma_{dc,SLG})^{-1} \), where [109]:

\[
\sigma_{dc,Slg} = n_{min,Slg} e \mu_{q,Slg} \left( 1 + \frac{n_e R_e + n_h R_h}{n_{min,Slg}} \right)
\]

(\( G2 \))

is the dc conductivity of SLG [109], assuming a minimum charge paddle concentration \( n_{min,Slg} \sim 10^{12} \) cm\(^{-2} \) and \( \mu_{q,Slg} \sim 10^4 \) cm\(^2\)/Vs for the SLG mobility. \( R_{SLG} \) is estimated to be \( \sim 0.1 - 1 \) kΩ, depending on \( E_{F,SLG} \), \( V_R \), and \( T_{e,SLG} \). We assume \( R_e \sim 1 \) kΩ \( \mu \)m [109,119,127]. The total capacitance is \( C_1 = C_j^{-1} + C_{q,slg}^{-1} \), with \( C_j = \varepsilon_0 \varepsilon_r / \chi_d \) the junction capacitance [37] and \( C_{q,slg} \) the SLG quantum capacitance [71,130]. The latter, given the assumed \( n_{min,Slg} \), is estimated in the order of \( \sim 1 \mu F \) cm\(^2 \) [130], almost two orders of magnitude above \( C_j \). It is thus safe to take \( C_j \sim C_j \). Considering all the above ranges, we find \( f_B = (2\pi R_C)^{-1} \sim 1 - 10 \) GHz. This can be increased using a higher Si doping and/or reducing \( C_j \) by geometrical focusing. However, \( F > 1 \) will eventually impose an upper limit in the temporal response due to the required high-Q optical cavities.

3. The optics-limited frequency can be written as [78] \( f_{opt} = c / (\pi 2L Q) \). 2L is the cavity roundtrip length (the middle \( nSi \) and \( SiO_2 \) layers) and \( c / \nu \) the speed of light inside it. Since the latter consists of two different materials we get:

\[
L = \frac{c}{\left( \frac{1}{n_{slg} d_{slg}} + \frac{1}{n_{Si} d_{Si}} \right)},
\]

(\( G3 \))

where \( d_{slg} \sim 1.06 \mu m \) and \( d_{Si} \sim 0.45 \mu m \) are the thicknesses of the two middle layers. The quality factor, \( Q \), is defined as [78]:

\[
Q = \frac{\omega}{\Delta \omega},
\]

(\( G4 \))
with $\omega_0$ the cavity resonant frequency and $\Delta \Omega$ the full-width at half-maximum (FWHM) of the resonance. We find $Q \sim 200$ for $F = 1$ and $Q \sim 20\,000$ for $F = 100$. These two extremes give $f_{\text{opt}} \sim 10 - 1000 \text{GHz}$.

We thus conclude that, with careful selection of device parameters and use of geometrical focusing, we could achieve operating frequencies up to $\sim 100 \text{GHz}$.
[106] G. W. Hanson, J. Appl. Phys. 103, 064302 (2008).
[125] A. N. Grigorenko, M. Polini, and K. S. Novoselov, Nat. Photonics 6, 749 (2012).