Supplementary Files for

Spatially controlled electrostatic doping in graphene *p-i-n* junction for hybrid silicon photodiode

Tiantian Li[1], Dun Mao[1], Nick W. Petrone[2], Robert Grassi[3], Hao Hu[4], Yunhong Ding[4], Zhihong Huang[5], Guo-Qiang Lo[6], James C. Hone[2], Tony Low[3], Chee Wei Wong[7] and Tingyi Gu[1*]

[1]Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19711

[2]Department of Mechanical Engineering, Columbia University, New York, NY 10027

[3]Department of Electrical Engineering, University of Minnesota, Minneapolis, MN 55455

[4]DTU Fotonik, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[5]Hewlett-Packard Laboratories, 1501 Page Mill Rd., Palo Alto, CA 94304

[6]The Institute of Microelectronics, 11 Science Park Road, Singapore Science Park II, Singapore 117685

[7]Mesoscopic Optics and Quantum Electronics Laboratory, University of California Los Angeles, CA 90095

* Email: tingyigu@udel.edu

Supplementary Notes:

I Theoretical potential profiles of the graphene-silicon vertical heterojunctions and lateral homojunction

Let φ_g and χ be the work function of intrinsic graphene and the electron affinity of silicon, respectively, so that the work function difference Δ between the two materials, i.e., the difference between the Fermi levels of the two isolated materials, can be expressed as

$$\Delta = \Phi_g - \chi - \Phi_n , \qquad \text{[S–1]}$$

where $\varphi_n$is the difference between the conduction band edge and Fermi level in silicon (see Fig. S1a). The latter is determined by the doping concentration. In the non-degenerate limit:

$$\phi_n = \begin{cases} kT\ln(N_c/N) & \text{for n-type doping}\,, \\ E_g - kT\ln(-N_v/N) & \text{for p-type doping}\,, \end{cases} \qquad \text{[S–2]}$$

where k is Boltzmann's constant, T the temperature, E_g = 1.12 eV the energy band gap,

$N = N_D - N_A$ the net doping density, and N_c and N_v are the effective density of states in the conduction and valence band, respectively:

$$N_c = 2\left(\frac{2\pi m_n kT}{h^2}\right)^{3/2}, \qquad [S–3]$$

$$N_v = 2\left(\frac{2\pi m_p kT}{h^2}\right)^{3/2}. \qquad [S–4]$$

Here h is Planck's constant and m_n and m_p are the density-of-states effective mass for electron and holes, respectively. We use the values $m_n = 1.182 m_0$ and $m_p = 0.81 m_0$ (m_0 is the free electron mass). Also, we take $\Phi_g - \chi = 0.5$ eV. Since $\Phi_g - \chi \sim E_g/2$, it follows from [S–1]–[S–2] that, for large enough doping values $|N|$, $\Delta > 0$ ($\Delta < 0$) for n-type (p-type) doping.

When the two materials are put in contact, electrons or holes are transferred from silicon to graphene depending on the sign of Δ, resulting in a surface charge on the graphene layer and a band bending on the silicon side. The position of the Dirac point energy E_d with respect to the Fermi level E_F and the amount of band bending ψ_s (see Fig. S1b) are related to each other by:

$$E_F - E_d = \Delta - \psi_s \qquad [S–5]$$

With the above assumption on the sign of Δ with respect to the sign of N, the charge on the graphene layer is negative (positive) for n-type (p-type) silicon and a depletion region is formed on the silicon side of the junction. Using the full-depletion approximation, i.e., assuming that the charge density $\rho = q(p - n + N)$ in silicon is given by

$$\rho(z) = \begin{cases} 0 & \text{if } z < -z_d\,, \\ qN & \text{if } -z_d < z < 0\,, \end{cases} \qquad [S–6]$$

where q is the electronic charge, $N = N_D - N_A$ the net doping density, $z = 0$ the position of the silicon-graphene interface, and z_d the width of the depletion region, the solution of Poisson's equation $-d^2\phi/dz^2 = \rho/E_s$ is

$$\varphi(z) = \begin{cases} \varphi(-z_d) & \text{if } z < -z_d\,, \\ \varphi(-z_d) - \frac{qN(z+z_d)^2}{2\epsilon_s} & \text{if } -z_d < z < 0\,, \end{cases} \qquad [S–7]$$

d $2E_s$

with E_s being the dielectric constant of silicon, from which

$$\psi_s = -q[\varphi(0) - \varphi(-z_d)] = \frac{q^2 N z_d^2}{2\epsilon_s}\,. \tag{S–8}$$

Since the charges on graphene and silicon must compensate each other, i.e., $Nz_d = n_g$, where n_g is the net electron sheet density on the graphene layer, we get

$$\psi_s = \frac{q^2 n_g^2}{2\epsilon_s N}\,. \tag{S–9}$$

In turn, n_g is related to the energy difference E_F-E_d through the graphene density of states. Using the $T = 0$ approximation:

$$n_g = \mathrm{sgn}(E_F - E_d)\frac{(E_F - E_d)^2}{\pi(\hbar v_F)^2}\,. \tag{S–10}$$

Combining [S–5] and [S–9]–[S–10], we finally get:

$$\psi_s - \frac{q^2(\Delta - \psi_s)^4}{2\epsilon_s N\pi^2(\hbar v_F)^4} = 0\,, \tag{S–11}$$

which can be solved for ψ_s.

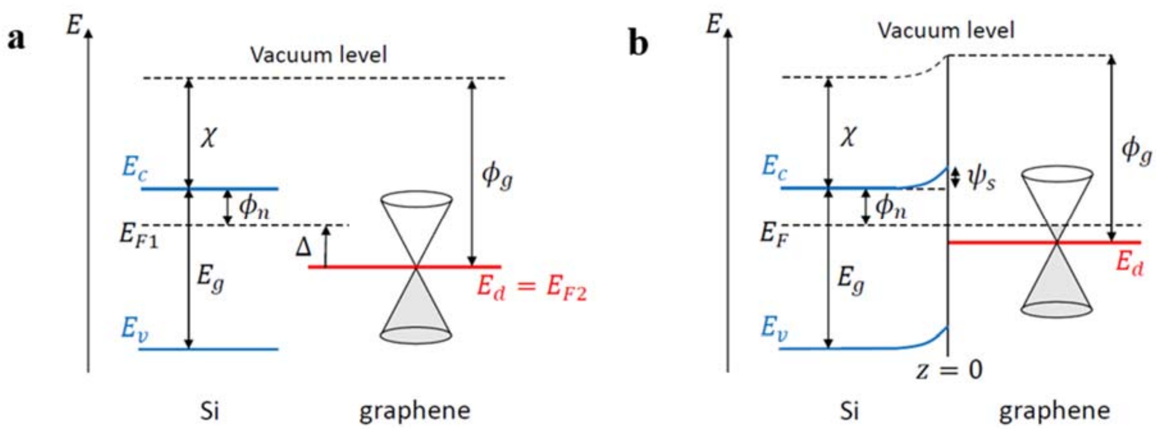

Figure S1. Schematic band diagram of doped silicon and intrinsic graphene. a, when the two materials are isolated and **b,** upon formation of the contact.

Fig. S2 shows the band diagram computed from the numerical solution of [S-11] for $N = -5 \times 10^{18}$, 5×10^{18} cm^{-3} at T = 300 K. In the presence of a residual chemical doping n_0 in the graphene layer, the considerations leading from [S-1] to [S-8] are still valid. However, charge neutrality now imposes $N z_d = n_g - n_0$, so that [S-9] is replaced by

$$\psi_s = \frac{q^2 (n_g - n_0)^2}{2\epsilon_s N} \qquad [S\text{–}12]$$

And [S-11] by

$$\psi_s - \frac{q^2}{2\epsilon_s N}\left[\mathrm{sgn}(\Delta - \psi_s)\frac{(\Delta - \psi_s)^2}{\pi(\hbar v_F)^2} - n_0\right]^2 = 0\,. \qquad [S\text{–}13]$$

Fig. S2 shows the band diagram computed with $n_0 = -1\times10^{12}$ cm^{-2} (p-type doping).

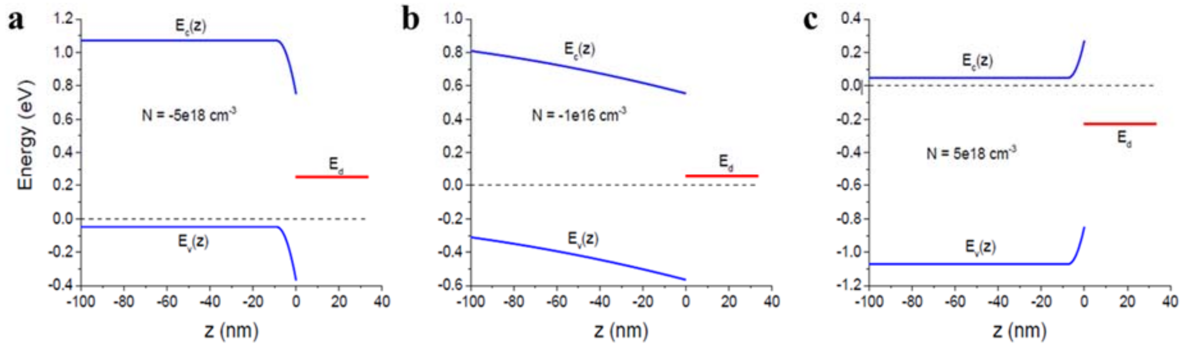

Figure S2. Band diagram of silicon-graphene junction for a, $N = -5 \times 10^{18}$, **b,** $N = -1 \times 10^{16}$, and **c,** $N = 5 \times 10^{18}$ cm^{-3}. E_F is taken as the energy reference.

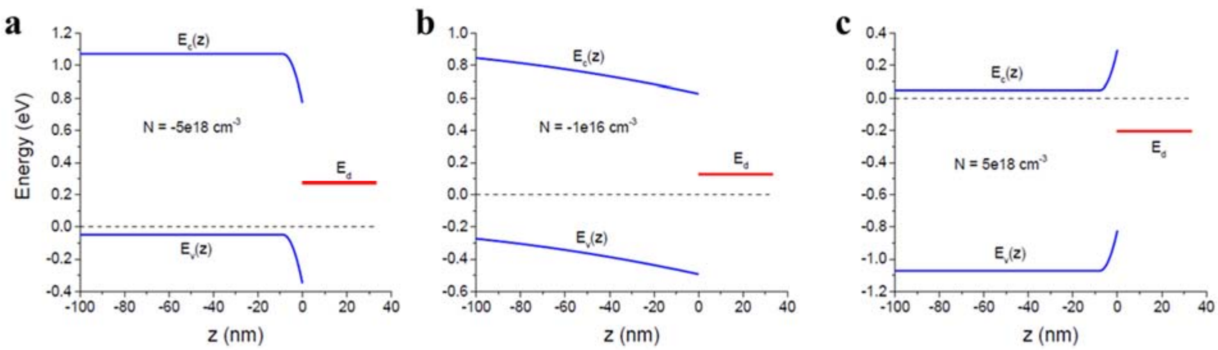

Figure S3. Same as in Fig. S2 but for a finite p-type residual doping of graphene of $\boldsymbol{n_0 = 1\times10^{12}}$ **cm^{-2}.** Compared to Fig. S2, the quantity E_F, E_d is slightly decreased (which means a more p-type character of graphene, as expected), an effect which is mostly evident for low $|N|$.

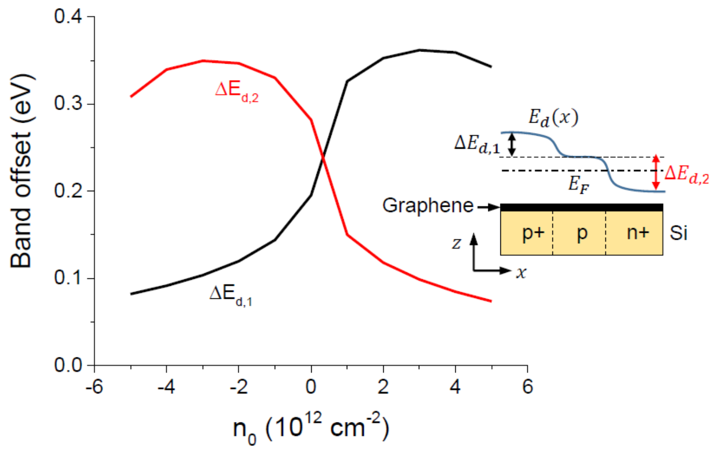

Figure S4. Graphene band offsets in a *p+–p–n+* silicon-graphene heterostructure. The doping values for silicon are the same as in Fig. S2–S3. The inset shows a schematic of the device structure and the definition of $\Delta E_{d,1}$ and $\Delta E_{d,2}$ with reference to the Dirac point energy profile $E_d(X)$.

For a p^+-p-n^+ silicon-graphene heterostructure, as the one shown in the inset of Fig. S4, the shape of the band diagram along the vertical direction z at fixed longitudinal position x is similar to the plots in Fig. S2–S3. The different doping values of silicon give rise to band offsets $\Delta E_{d,1}$, $\Delta E_{d,2}$ in the graphene layer (see inset of Fig. S4 for symbol definition). Fig. S4 shows the dependence of $\Delta E_{d,1}$ and $\Delta E_{d,2}$ on the residual doping n_0.

II. Scanning photocurrent microscopy and integrated waveguide photocurrent measurement

Scanning photocurrent microscopy (SPCM) is used to investigate the optoelectronic properties of the hybrid G-Si structure in the visible band. SPCM provides the spatially resolved information about the photocurrent generation and transport process. In this study, the normal incident laser is focused down to a sub-micrometer full-width at half-maximum spot through a ×10 objective, with spot position controlled by a two-axis scanning mirror (Scheme 1 in Fig. S5) [S1]. The wavelength of the laser source can be tuned from visible to mid-infrared, but the operating wavelength is limited to be in the visible band by the mirrors, attenuators, and objectives.

The measured SPCM image, superimposed on the reflection image, is shown in the inset of Fig. S5. The area covered by the graphene is marked by the dashed line (illustrated in detail in Fig. S5 inset), which partially cover the silicon waveguide (solid line). Excitation light from lasers (continuous laser, tunable from 1320 to 1620nm, and subpicosecond pulsed laser, centered at

1550nm) is coupled on and off the chip through inverse tapers and lensed fiber. The photocurrent is collected by contacting standard ground-signal-ground (GSG) probe to the device with 100 μm pitch-to-pitch distance.

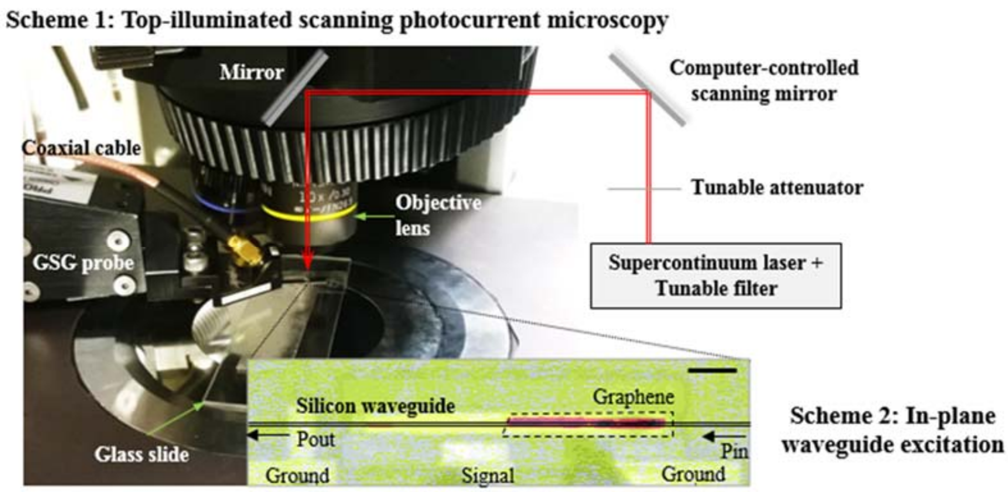

Figure S5. Scanning photocurrent microscopy set up. Optical image showing the top illumination and the electrical current readout, with the schematic of the optical path. Inset: the SPCM image (532nm) of the graphene covered the right part of the *p-i-n* junction. The photocurrent (red) and the reflection (grey) images are superimposed and show photocurrent generation from the graphene covered part of intrinsic region. Scale bar: 30μm. The position of the silicon waveguide is indicated on the Scanning photocurrent microscopy image, with in-plane excitation (P_{in}) and output (P_{out}).

III. Characteristics of graphene *p-i-n* junction

Lateral potential gradient along the graphene plane and low carrier densities are important for studying the built-in electric field driven photoresponse in graphene. The unsymmetrical doping in silicon leads to a built-in potential in the silicon substrate, which is directly contacted to the graphene: $V_0 = (k_BT/e)ln(N_aN_d/n_i^2\text{-}\Delta E_F/e)$ where k_BT/e = 0.0259 V at room temperature, with Boltzmann constant k_B, room temperature T and single electron charge e. The doping densities on both p and n sides are $N_a = N_d = 5\times10^{18}$ cm^{-3}. At the intrinsic region, the net carrier density in Si is 4.55×10^{11} cm^{-3}, with an electrostatic potential (V_0) of 0.84 eV in dark. Incident pump photons are focused to a sub-1 μm^2 area and normally incident to the Si photonic membrane, exciting electron-hole pairs in the intrinsic region. The ultrafast carrier transfer on vertical G-Si heterojunction significantly suppress the carrier loss channels through local recombination in Si. The transferred photocarriers would be drained through the lateral built-in electric field.

In absence of external bias, the built-in electric field in graphene can be introduced by asymmetric source and drain contacts. Here the asymmetric source and drain contacts are replaced by the Si with different doping types, rather than gate activation. This configuration is illustrated with the equivalent circuit model shown in Fig. S6a. The two diode symbols correspond to *p-i* and *i-n* silicon junctions. Graphene (golden part) is on the top of silicon. Two tunneling junctions to the highly-doped Si on the sides (dash lines on the two sides) are formed and one Schottky contact is formed with intrinsic silicon part. In the photocarrier separation region (Graphene on intrinsic silicon of PhC WG), the Schottky barrier is calculated to be around 0.6 eV (Supplementary file S1) for near infrared photocarrier generation. The lateral built-in electric field formed in graphene (Figure S4) drives the hole towards the graphene-*p*-Si contact region, where the Schottky barrier is less than 0.05 eV. Driven by the strong vertical built-in electric field, holes travel back into the *p*-Si through field emission.

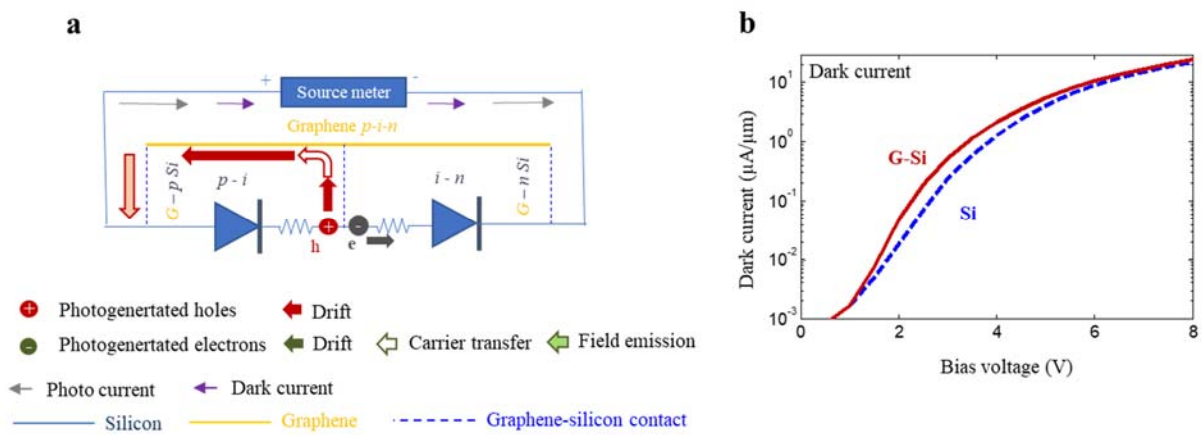

Figure S6. Characteristics of the hybrid device. a, Small signal model for the multi-junction formation in the hybrid *p-i-n* junction. **b,** Measured current-voltage characteristics for graphene integrated (red solid curve) and monolithic Si *p-i-n* junctions (blue dashed curve).

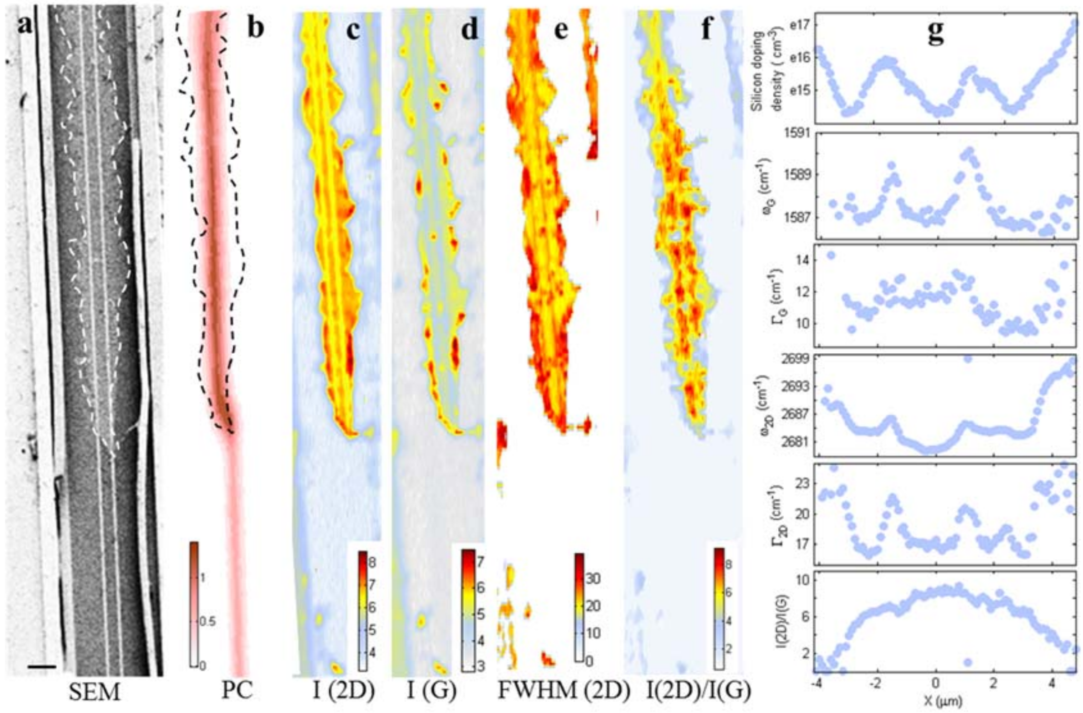

Figure S7. Spatially-resolved Raman mapping of the graphene on silicon *p-i-n* junction. a, SEM image of the active region covered by graphene (dashed line). **b,** Photocurrent mapping. **c,** Raman mapping of *2D* peak intensity. **d,** *G* peak intensity. **e,** Full-width-half-maximum (FWHM) of the *2D* peak. **f,** *2D* versus *G* peak ratio for the highlighted section in **a. g,** Doping profile of the Si substrate measured by EFM, *G* peak wavenumber (ω_G), full-width half-maximum (FWHM) of *G* peak (Γ_G), *2D* peak wavenumber (ω_{2D}), FWHM of *2D* peak (Γ_{2D}), and the intensity ratio of the *2D*-to-*G* peak near the intrinsic region. The patterns are repeatable at different positions along the junction.

For the device shown in Fig. 1d, the corresponding characteristics by SEM, spatially-resolved photocurrent and Raman mapping are shown in Fig. S7a-f. The spatially dependent Raman characteristics of graphene on the Si *p-i-n* junction are compared to the electrical force measurements (EFM) in Fig. S7g. The substrate screening effect from the double photonic crystal waveguide is reflected by the twin peaks in EFM mapping and the Raman *G* peak wavenumber, representing the line defect on photonic crystal plane. The substrate doping influences the wavenumber, the full-width half-maximum of the Raman *G* peak, and the intensity ratio of the *2D*-to-*G* peak.

IV. Absorption saturation in monolithic structure

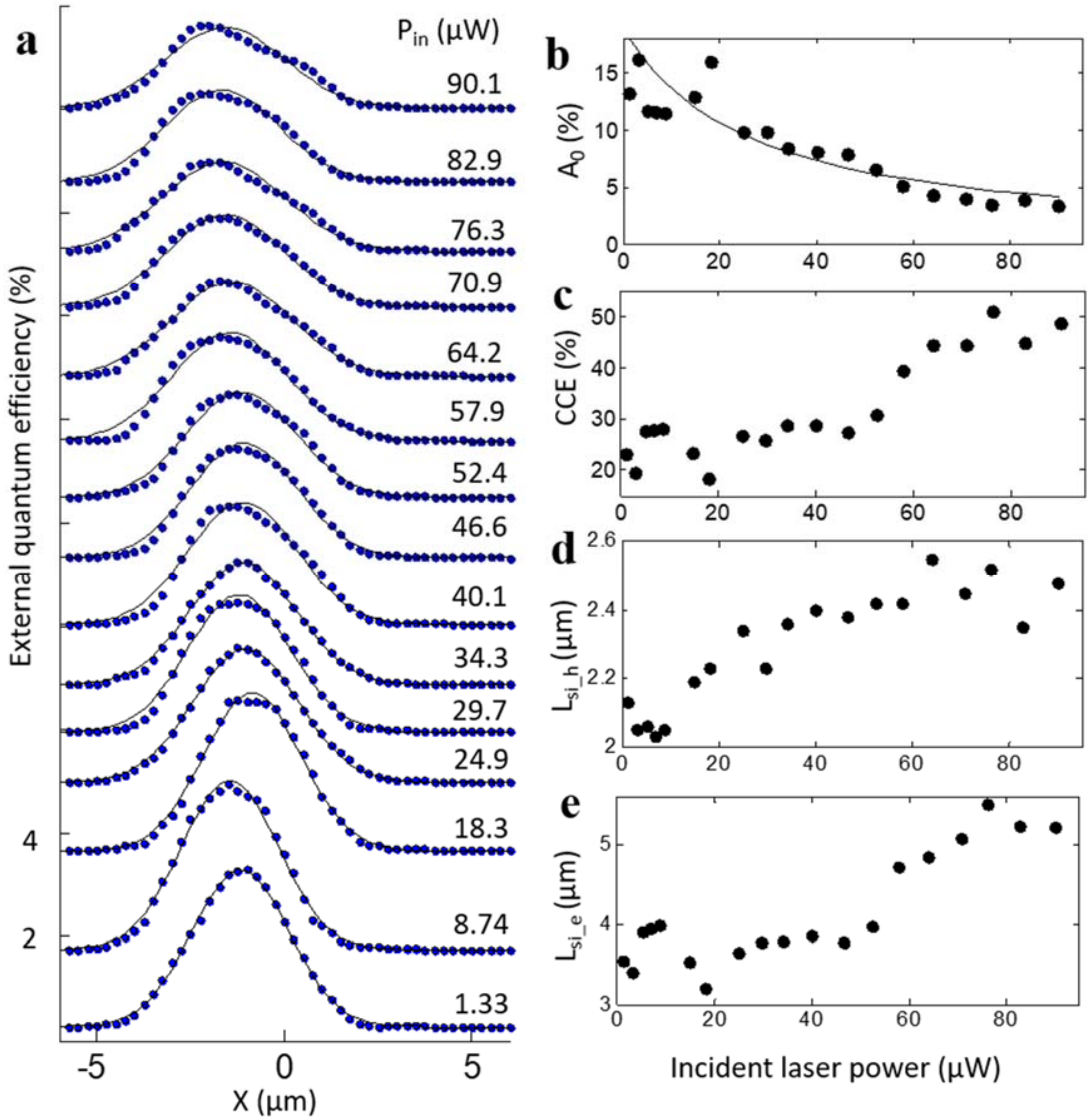

Figure S8. Photocurrent mapping across the monolithic silicon *p-i-n* junction. a, EQE mapping as the 532 nm green laser spot moves across the *p-i-n* junction, with plotting offset with the laser power as marked in the figure for clarity. The blue circles are experimental data and the solid black curves are corresponding fits with equations S-16. The intrinsic region is defined from -2.5 μm to 2.5 μm. The laser spot diameter is 0.6 μm **b,** Absorption coefficient versus laser power. The circles are experimental data and the curve is the fit by absorption saturation model ($A=A_0/(1+P/P_0)$). The amplitude A_0 is fitted to be 0.19, and saturation power P_0 is 25 μW. **c,** Peak charge collection efficiency of the lateral *p-i-n* junction. **d,** Mean free path length of holes and **e,**

electrons in the intrinsic region of the nanostructured Si as the temperature rises with high optical injection.

The photocarrier density dynamics can be described by the master equation:

$$\frac{dn}{dt}=\frac{I\alpha}{\hbar\omega}-\frac{n}{\tau_t} \qquad [S-14]$$

where n is the electron density. The first term on the right represents local carrier generation rate ($\Gamma=I\alpha/\hbar\omega$). The absorption coefficient α is determined by the Si photonic crystals in the visible band and graphene in the infrared range. The carrier lifetime depends on carrier loss rate through local recombination ($1/\tau_{rec}$) and carrier extraction rate ($1/\tau_{transport}$): $1/\tau_t = 1/\tau_{rec} + 1/\tau_{transport}$. By solving the equation [S-14] in steady state, the density of the total carrier is $n=I\alpha\tau_t/\hbar\omega$.

At higher optical injection intensities, the local carrier density increases and reduces the absorption coefficient, both in Si and graphene:

$$\alpha(n)=\frac{\alpha_0}{1+n/n_s}=\frac{\alpha_0}{1+I/I_s} \qquad [S-15]$$

where α_0 is the absorption coefficient at low light intensity, n_s is the saturation carrier density and I_s is the saturation light intensity.

We measured the photocurrent profile for monolithic devices at different optical injection levels, fitted with the equation (Fig. S8a):

$$EQE_{Si}(X)=A_{Si}(I)\times e^{-(X-X_e)^2/L_e^2-(X-X_h)^2/L_h^2} \qquad [S-16]$$

The efficient carrier extraction (EQE) leads to higher saturation threshold of light intensity. $L_{e/h}$ is the mean free path for electrons/holes of majority carriers in intrinsic graphene on intrinsic Si substrate. X is the spatial location of the laser, as $X_{e/h}$ is defined as the border of the intrinsic region to *p/n* doped regions. I_s is fitted with the light intensity dependent absorption coefficient, to be 25 μW/ μm^2 (Fig. S8b). The charge collection efficiency of lateral *p-i-n* junction (η_{pin} = $e^{-(X-X_e)^2/L_e^2-(X-X_h)^2/L_h^2}$) is fitted as in Fig. S8c. The asymmetric mean free path of majority carriers is given in Fig. S8d-e.

V. Current saturation in graphene

The drift current from the electron in G-Si ($I_{G\text{-}Si}$) can be expressed as [S2, S3]:

$$I_{G-Si} = \frac{W}{L}\int_0^L (n_G \mu_{G\_e} \nabla E_D + (n_{Si} - n_G)\mu_{Si\_e} \nabla E_C)dx \qquad \text{[S-17]}$$

where the drift current is the integral of electron and hole densities along the channel (L). The cross-section of the local channel is 5 μm wide (W) and 250 nm thick (d). The potential gradient for conduction/valance bands ($\nabla E_{c/v} = F_{Si}$) is 0.17 V/μm at low optical injection region, determined by lateral band offset (Fig. 1b). The lateral built-in electric field is shared by the atomic thin graphene layer ($\nabla E_D = F_G$), forming the potential gradient of the Dirac point (∇E_D). In Si, the mobilities for electrons and holes are μ_{Si_e}=1400 and μ_{Si_h}=450 $cm^2V^{-1}s^{-1}$. The drift carrier velocity in Si saturates as the built-in electric filed goes beyond 0.3V. This threshold voltage becomes lower (0.22 to 0.26 V) for the hybrid device. F_G and F_{Si} are the built-in electric field for graphene plane and Si *p-i-n* junction respectively. In the hybrid structure, the electron/hole numbers per second in Si ($N_{G\text{-}Si\_Si}/P_{G\text{-}Si\_Si}$) and graphene ($N_{G\text{-}Si\_G}/P_{G\text{-}Si\_G}$) is determined by vertical potential gradient on the G-Si heterojunction (Fig.1b). The photocurrent in the hybrid device ($I_{G\text{-}Si}$) can be expressed as:

$$I_{G-Si} = N_G(\mu_{G\_e} + \mu_{G\_h}) \times F_G + (N_{Si} - N_G) \times (\mu_{Si\_e} + \mu_{Si\_h}) \times F_{Si} \qquad \text{[S-18]}$$

The number of carriers generated per second in the graphene Si hybrid structure is same as the one in the monolithic device (N_{Si}). Graphene carries part of the photocurrent by contacting to the Si substrate (N_G), determined by the vertical contact on the G-Si heterojunction (Fig. 1c). The carrier mobility in graphene is carrier density dependent [S4, S5]:

$$\mu_{G\_e/h}(n) = \frac{\mu_0}{1 + (n_G / n_0)^{\alpha}} \qquad \text{[S-19]}$$

where μ_0 = 4650 cm^2/Vs, n_0=1.1×10^{13} cm^{-2}, and α=2.2 at room temperature. At high optical injection region, the photocurrent profile can be fitted by:

$$I_{G-Si}(X) = I_G \times e^{-(X - X_{G\_e})^2/L_{G\_e}{}^2 - (X - X_{G\_h})^2/L_{G\_h}{}^2} + (I_{Si} - I_G) \times e^{-(X - X_{Si\_e})^2/L_{Si\_e}{}^2 - (X - X_{Si\_h})^2/L_{Si\_h}{}^2} \qquad \text{[S-20]}$$

where the current in graphene (I_G) saturates at higher optical powers, and the Si membrane carries the majority of the current (I_{Si}-I_G). By curve-fitting, the model to experimentally measured photocurrent profile across the *p-i-n* junction (Fig. S9a), the photocurrent carried by Si and graphene can be separated (Fig. S9b). Graphene carries most of the photocurrent until it saturates at high incident laser powers (40 μW). The corresponding charge collection efficiency of the lateral *p-i-n* junction and mean free path of holes/electrons in graphene are shown in Fig. S9c,d.

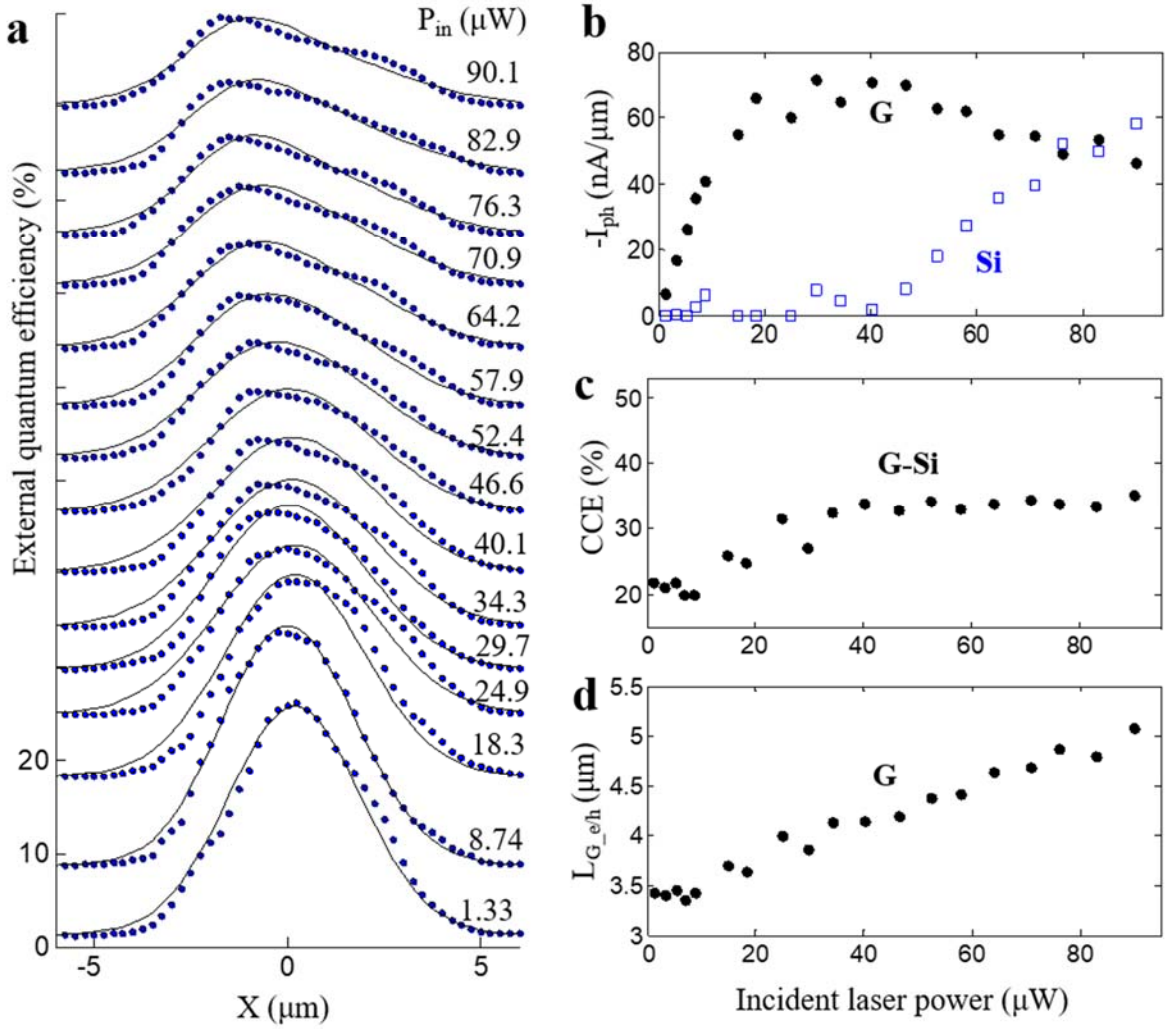

Figure S9. Photocurrent mapping across the hybrid *p-i-n* junction. a, External quantum efficiency (*EQE)* mapping as the 532nm green laser is spot moved across the *p-i-n* junction, with an offset of laser power (marked in the figure) for clarity. The blue circles are experimental data and the solid black curves are corresponding fits via equation S-18. The intrinsic region is defined from -2.5 μm to 2.5 μm. The laser spot diameter is 0.6 μm. **b,** Photocurrent carried by graphene (black dots) and Si (blue squares) versus laser power. **c,** Peak charge collection efficiency of the lateral *p-i-n* junction. **d,** Mean free path of holes/electrons in graphene.

VI. Charge transfer efficiency in graphene-silicon contact

Charge separation and recombination in nanostructured *p-i-n* junction occur near-place instantaneously after the carrier generation [S4]. For photoabsorption in the visible band, the charge absorption and recommendation take place in Si. Charge transfer rate ($1/\tau_{transfer}$) and recombination rate ($1/\tau_{rec}$) determines the internal quantum efficiency of the semiconductor device: $IQE_{G\text{-}Si}=G_0(1/\tau_{tranfer})/(1/\tau_{transfer}+1/\tau_{rec})$. Built-in electric field assisted charge transfer on the Van der Waals interface improves $IQE_{G\text{-}Si}$ in two ways: (1) introduce vertical carrier transfer channel with high charge transfer rate; (2) suppress $1/\tau_{rec}$ through reducing local carrier density. Recombination in monolithic nanostructured Si significantly reduces the quantum efficiency. The recombination current has both bulk (R_b) and surface contributions (R_S). Considering the recombination rate in bulk Si is much lower than the measured minority carrier lifetime in Si photonic crystal structures, the recombination of photogenerated carriers in the nanostructured intrinsic region is dominated by surface recombination, normalized by the surface (S) to volume ratio (V) [S5-S12]:

$$R_{Rec} = S/V \times R_S \qquad \text{[S-21]}$$

The rate of surface recombination (R_S) is:

$$R_S = \frac{(n+n_s)(p+p_s)-n_i^2}{(p+p_s)/S_n+(n+n_s)/S_p} \qquad \text{[S-22]}$$

where n/p is the local electron/hole densities; S_n and S_p are surface recombination velocities for electrons/ holes respectively and enhanced with graphene cladding. p_s and n_s are the carrier density on the Si surface (in contact with graphene). In Si, the photogenerated hot carriers relax to the band edge in 300 fs. The surface recombination rate for Si nanostructure is about 2×10^4 cm/s. Given the photonic crystal structure with a lattice constant of 415 nm, hole radius of 124 nm and membrane thickness of 250 nm, the surface to volume ratio (S/V) is derived to be 9×10^5 /cm. As the product of surface recombination rate and surface to volume ratio, the local recombination lifetime $1/\tau_{rec}$ is estimated to be 2×10^{10} /s (20 GHz). In the monolithic device, the local IQE is measured to be only 12% (blue squares in Fig. 3c), which corresponds to the $1/\tau_{transport}$ of 3×10^9 /s in lateral Si *p-i-n* junction. The hybrid structure allows much more efficient carrier transport through the vertical G-Si heterojunction, with charge transfer efficiency of near 95% (red crosses in Fig. 3c). The vertical carrier transfer rate from Si to graphene is then derived to be faster than 10^{11}/s (100 GHz).

VII. Hot carrier separation and avalanche gain

VII.1 Hot carrier response on G-Si junction

In Fig. 4b, the photocurrent on the hybrid and monolithic region to near-infrared light is measured through top illumination on the same device. As shown in Fig. 1d, a 120 μm long photonic crystal waveguide is partially covered by 60 μm long graphene. Single mode fiber with a spot size of 10 μm is placed on top of the graphene covered region. The XYZ position of the fiber is finely adjusted for the maximum photocurrent output. Continuous and pulsed (duration of sub-picosecond) optical signal with a center wavelength of 1550 nm is illuminated on the intrinsic region of G-Si *p-i-n* junction through the single mode fiber. The tip of the fiber is then moved to the intrinsic region for collecting the photocurrent under the same optical excitation conditions (Fig. S10). The voltage-dependent photocurrent is then derived by extracting dark current from the total current under illumination (Fig. 4b). The optical power levels under two different excitations are adjusted for generating the same photocurrent in Si *p-i-n* junction, to ensure the same average power level for pulsed and continuous wave excitation coupled to the Si photonic crystal membrane, as the hot carrier contribution to photocurrent is minimal in Si devices.

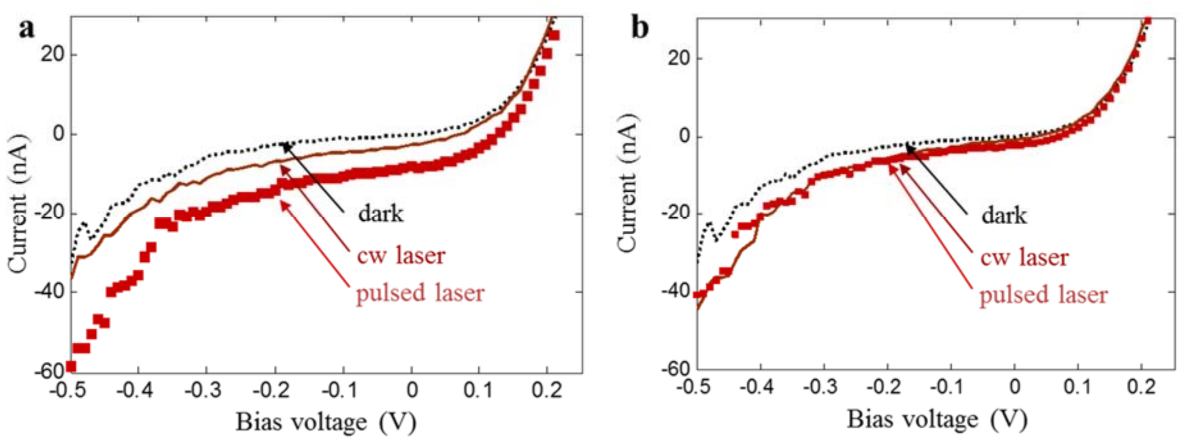

Figure S10. The IV characteristics of the *p-i-n* junction in dark, continuous wave laser and sub-picosecond pulsed laser illumination. a, G-Si hybrid structure. **b,** Monolithic Si structure.

VII.2 Discussion on carrier avalanche mechanisms

We analyzed the origin of bias dependent avalanche gain in the hybrid structure. The avalanche gain could be attributed to two junctions: (1) carrier multiplication along the biased graphene across the *p-i-n* junction. (2) avalanche on graphene-doped Si interface. Both processes might contribute to the overall reverse bias dependent photocurrent gain, as both models can be adjusted to fit the data:

Avalanche along biased graphene: The near-infrared light absorption in graphene generates hot carriers with the energy of half of the photon energy. The hot carriers in graphene will quickly thermalize and lose their energy through carrier-carrier scattering (in the time scale of sub 50fs [S13]) and carrier-phonon interactions (several ps) until their carrier temperature (T_e) reaches equilibrium with the lattice temperature (T_{ph}). Since only the hot carrier with energy above the Schottky barrier would be able to be collected, carrier population with higher T_e leads to higher emission probability. The transient carrier temperature after sub-picosecond laser excitation is much higher than the carrier temperature in equilibrium with phonon bath (under continuous wave excitation), and thus leads to the 3.2× photocurrent enhancement through more efficient charge emission through the Schottky barrier.

As a graphene sheet is placed in an electric field (F) along the graphene, the charge generation, transport, and recombination follow the 1D continuity equations and can be numerically solved [S14]. The charge generation rate through impact ionization can be approximated by the following expression:

$$U \propto T_e^{3/2} F^{\alpha} \exp[-(n/n_0)] \qquad \text{[S-23]}$$

T_e is the electronic temperature, which quadratically depends on the electric field: $T_e=T_L[1+(F/F_{CT})^2]$. T_L is the lattice temperature. F_{CT} is a critical field for the onset of hot carrier effects, limited by various scattering mechanisms. α is a numerical factor of 1.7. n is local carrier density. n_0 is around 1×10^{12} cm^{-2}, and slightly increases with T_e. The multiplication factor can be derived as $M=(n+U)/n$. The clean graphene Si interface with asymmetric semiconductor contact lowers the FCT to be only 1.1kV/cm, as fitted to the experimental data in Fig. 4b.

Avalanche on graphene-Si interface: The voltage dependence of photocurrent has contributions from both thermionic emission (TE) and avalanche gain on G-Si interface. As the intrinsic region of Si is moderately doped (10^{16} cm^{-3}), thermionic emission domivnates the carrier transport process. The voltage dependence of TE can be expressed as $\sqrt{V_R+\phi_B/E_0}\exp(qV_R/\varepsilon')$, where E_0 and ε' are two constants determined by the Si doping and operation temperature, and V_R is the reverse bias. The avalanche multiplication factor follows the empirical model $M=1/(1-(V_R/V_{BD})^k)$ [29]. The bias dependence of photocurrent is proportional to the product of TE and M. The breakdown voltage (V_{BD}) and the power coefficient k are fitted to be -0.63 V and 3.2 respectively. M = 4.18 is achieved as V_R set at -0.5V bias. Photocurrent ratio between the pulsed laser and continuous wave excitation is weakly dependent on the reverse bias, indicating the independence of the hot carrier generation and amplification processes.

VII.3 Dark current analysis by Landauer transport model

As shown in Fig. S10, both the photocurrent and dark current depend on the reverse bias. The limited density of states in graphene leads unique carrier transport behavior on G-Si junction. Landauer transport formalism can be used for predicting the thermionic emission current of carrier injection on direct contact between G-Si interface [S15-S16]:

$$J_0 = \frac{2}{\pi} R_{Injection} q_0 (\frac{k_B T}{\hbar v_F})^2 (\frac{\phi_0}{k_B T} + 1) \exp(\frac{-\phi_0}{k_B T}) \qquad \text{[S-24]}$$

where $R_{Injection}$ is the charge injection rate, q_0 is the elementary charge, k_B is the Boltzmann's constant, $\hbar$ is the reduced Planck's constant, v_F is the graphene Fermi velocity, ϕ_B is the Schottky barrier at zero bias, and V_R is the reverse bias. Laudauer transport model can be applied to dark current analysis on Graphene-Si contact.

Supplementary References:

S1. C.-H. Lee, et al. Atomically thin p-n junctions with van der Waals heterointerfaces. Nat. Nanotechnol. **9**, 676-681 (2014).

S2. I. Meric, et al. Shepard, Current saturation in zero-bandgap, top-gated graphene field-effect transistors. Nat. Nanotechnol. **3**, 654-659 (2008).

S3. V. E. Dorgan, M.-H. Bae, and E. Pop, Mobility and saturation velocity in graphene on SiO2. Appl. Phys. Lett. **97**, 082112 (2010).

S4. M. M. Gabriel, et al. Imaging charge separation and carrier recombination in nanowire *p-i-n* junctions using ultrafast microscopy. *Nano Lett.* **14**, 3079-3087 (2014).

S5. S. L. Chuang, *Physics of Optoelectronic Devices* (Wiley, New York, 1995).

S6. T.-T. D. Tran, et al. High brightness InP micropillars grown on silicon with Fermi level splitting larger than 1 eV. *Nano Lett.* **14**, 3235-3240 (2014).

S7. E. M. Grumstrup, et al. Ultrafast carrier dynamics of silicon nanowire ensembles: the impact of geometrical heterogeneity on charge carrier lifetime. *J. Phys. Chem. C* **118**, 8626 (2014).

S8. T. Gu, M. A. El-Emawy, K. Yang, A. Stintz, and L. F. Lester. Resistance to edge recombination in GaAs-based dots-in-a-well solar cells. *Appl. Phys. Lett.* **95**, 261106 (2009).

S9. O. Breitenstein, P. Altermatt, K. Ramspeck, and A. Schenk. The origin of ideality factors $n >$ 2 of shunts and surfaces in the dark IV curves of Si solar cells. In *Proceedings of the 21st European Photovoltaic Solar Energy Conference*, 625-628 (2006).

S10. W. Shockley, and W. T. Read, Statistics of the recombinations of holes and electrons. *Phys. Rev.* **87**, 835 (1952).

S11. J. M. Liu, *Photonic Devices* (Cambridge University, 2005).

S12. J. A. Brews, Charge-sheet model of the MOSFET. *Solid State Electron.* **21**, 345 (1978).

S13. K.J. Tielrooij, et al. Generation of photovoltage in graphene on a femtosecond timescale through efficient carrier heating. *Nat. Nanotechnol.* **10**, 437-443(2015).

S14. L. Pirro, A. Girdhar, Y. Lebleici, and J.-P. Leburton, Impact ionization and carrier multiplication in graphene. *J. of Appl. Phys.* **112**, 093707 (2012).

S15. S. Datta, *Quantum transport: atom to transistor* (Cambridge University Press, 2005).

S16. S. Dhiraj, and J. U. Lee, Ideal graphene/silicon Schottky junction diodes. *Nano Lett.* **14**, 4660-4664 (2014).