Tunnel field-effect transistors for sensitive terahertz detection

I. Gayduchenko+1,2 S.G. Xu+3,4 G. Alymov+,1 M. Moskotin2,1 I. Tretyakov,5 T. Taniguchi,6 K. Watanabe,7 G. Goltsman,2,8 A.K. Geim,3,4 G. Fedorov,1,2 D. Svintsov,1 and D.A. Bandurin3,1

1Moscow Institute of Physics and Technology (National Research University), Dolgoprudny 141700, Russia
2Physics Department, Moscow Pedagogical State University, Moscow, 119335, Russia
3School of Physics, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom
4National Graphene Institute, University of Manchester, Manchester M13 9PL, United Kingdom
5Astro Space Center, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow 117997, Russia
6International Center for Materials Nanoarchitectonics, National Institute of Material Science, Tsukuba 305-0044, Japan
7Research Center for Functional Materials, National Institute of Material Science, Tsukuba 305-0044, Japan
8National Research University Higher School of Economics, Moscow, 101000, Russia

The rectification of high-frequency electromagnetic waves to direct currents is a crucial process for energy harvesting, beyond 5G wireless communications, ultra-fast science, and observational astronomy. As the radiation frequency is raised to the sub-terahertz (THz) domain, efficient ac-to-dc conversion by conventional electronics becomes increasingly challenging and requires alternative rectification protocols. Here we address this challenge by tunnel field-effect transistors made of dual-gated bilayer graphene (BLG). Taking advantage of BLG’s electrically tunable band structure, we create a lateral tunnel junction and couple it to a broadband antenna exposed to THz radiation. The incoming radiation is then down-converted by strongly non-linear interband tunneling mechanisms, resulting in exceptionally high-responsivity (exceeding 3 kV/W) and low-noise (0.2 pW/√Hz) detection at cryogenic temperatures. We demonstrate how the switching from intraband Ohmic to interband tunneling regime within a single detector can raise its responsivity by one order of magnitude, in agreement with the developed theory. Our work demonstrates an unexpected application of interband tunnel transistors for high-frequency detection and reveals bilayer graphene as one of the most promising platforms therefor.

Field effect transistors (FETs) recently found an unexpected application for the rectification of THz and sub-THz signals beyond their cutoff frequency [1, 2]. This technology paves the way for on-chip [3], low-noise [4], and sub-picosecond radiation detection [5, 6] offering the possibility of >10 Gb/s data transfer rates. Contrary to competing diode rectifiers, FETs offer the possibility of phase-sensitive detection [7, 8] vital for noise-immune communications with phase modulated signals. Recent innovations towards enhanced responsivity of FET-detectors include the use of novel materials [9–11], exotic nonlinearities [12–15], enhanced light-matter coupling [16] and plasmonic effects [17–19]. Despite the rich and complex physics of THz rectification, the responsibility of most FET-detectors is governed by the sensitivity of the channel conductivity $G_{ch}$ to the gate voltage $V_g$, parameterized by the normalized transconductance $F = -d \ln G_{ch}/dV_g$ [2, 20]. The transconductance in conventional FETs has a fundamental limit of $e/k_BT$ ($e$ is the elementary charge and $k_BT$ is the thermal energy) dictated by the leakages of thermal carriers over the gate-induced barrier, termed as ‘Boltzmann tyranny’. Although this process is well-recognized as a limiting factor for the minimal power dissipation of FETs in integrated circuits, it has been scarcely realised that it also imposes a bound on the responsivity of antenna-coupled FETs to THz fields.

One of the most promising routes to escape from the Boltzmann tyranny is the manipulation of interband tunneling instead of intraband thermionic currents. This idea is materialized in a tunnel field effect transistor (TFET) [21–23]. TFETs find their applications in low-voltage electrical and optical switching [24], accelerometry [25], chemical [26] and biological sensing [27, 28]. In spite of this variety, the use of TFETs for the rectification of high-frequency signals [29] has not been attempted so far. This is also surprising considering recent advances in the development of tunnelling high-frequency rectifiers and detectors based on quantum dots [30, 31], diodes [32–37] and superconducting tunnel junctions [38–40]. A possible reason is that low on-state current and cut-off frequency of TFETs stimulate the belief on their inapplicability in teratronics [41].

In this work, we show that the opposite is true and demonstrate the use of TFETs for highly-sensitive sub-THz and THz detection. Using bilayer graphene (BLG) as a convenient platform for this enquiry, we fabricate a dual-gated TFET and couple it to a broadband THz antenna. The received high-frequency signal is rectified by electrostatically-defined tunnel junction resulting in high-responsivity and low-noise detection. Our experimental results and the developed theory suggest that the origin of the exceptional responsivity in our detectors is not the large transconductance, but rather steep curvature...
of the tunneling $I - V$ characteristic [42]. Our findings point out that even TFETs without sub-$k_B T/e$ switching can act as efficient THz rectifiers preserving all the benefits of transistor-based detection technology.

**Device fabrication and characterization.** For the proof-of-principle demonstration, we constructed a TFET of a BLG taking advantage of its unique electronic properties. BLG is a narrow-band semiconductor characterized by a tunable band structure highly sensitive to the transverse electric field [43]. This ensures a steep ambipolar field effect and allows for an independent control of the band gap size and the carrier density, $n$ [44], providing a unique opportunity for a fully electrostatic engineering of the spatial band profile [45–47]. We employed this property to electrostatically define typical TFET configuration shown in Fig. 1a,b. In addition, BLG hosts a high-mobility electronic system, a crucial property for high-frequency applications. As we now proceed to show, these properties make BLG a convenient platform to demonstrate the drastic differences in performance of intraband field-effect-enabled detection and its interband tunneling counterpart within the same device.

We fabricated our detector by an encapsulation of BLG between two slabs of hexagonal boron nitride (hBN) using standard dry transfer technique described elsewhere [48] (See Methods). The BLG channel of length $L = 2.8 \mu m$ and width $W = 6.2 \mu m$ was assembled on top of a relatively thin ($\sim 10$ nm) graphite back gate which ensured efficient screening of remote charge impurities in Si/SiO$_2$ substrate [49]. The device was also equipped with a second (top) gate electrode deposited symmetrically between the source and drain contacts. Importantly, relatively short ($l < 100$ nm) regions near the contacts were not covered by the top gate and thus were affected by the bottom one only. This configuration allowed us to define a lateral tunnel junction between single- and double-gated regions when the top and bottom gate voltages ($V_{tg}$ and $V_{bg}$ respectively) had opposite polarities [45–47], as explained in Figs. 1b and 5. The device was coupled to the incident radiation via broadband bow-tie antenna. The rectified dc photovoltage, $V_{ph}$, was read out between the source and drain terminals as shown in Fig. 1a (See Methods).

Prior to photoresponse measurements, we characterized the transport properties of our device. Figure 1e shows the dependence of our detector’s two-terminal resistance, $r_{2pt}$, on $V_{tg}$ for two representative values of $V_{bg}$ measured at $T = 10$ K. At $V_{bg} = 0$ V, $r_{2pt}(V_{tg})$ exhibits familiar bell-like structure that peaks at the charge neutrality point (CNP) where $r_{2pt} \approx 0.4 \, \text{k}\Omega$ (inset of Fig. 1e). Application of $V_{bg} = 2$ V shifts the CNP to negative $V_{tg}$ and results in drastic increase of $r_{2pt}$ that reaches 20 k$\Omega$ already at $V_{tg} \approx -3.5$ V. This increase is a clear indicative of the electrically-induced band gap in BLG [43, 44].

**Tunneling-enabled detection.** Figure 2a shows the external responsivity of our detector, $R_v = V_{ph}/P_{in}$, as a function of $V_{tg}$ recorded in response to $f = 0.13$ THz radiation. Here $V_{ph}$ is the generated photovoltage and $P_{in}$ is the incident radiation power (See Methods for the details of responsivity determination). At $V_{bg} = 0$ V, $R_v(V_{tg})$ exhibits a standard antisymmetric sign-changing behaviour with $|R_v|$ reaching 200 V/W close to the CNP. The functional form of $R_v(V_{tg})$ follows that of the normalized transconductance $F = -(dG_{ch}/dV_{tg})/G_{ch}$ (Fig. 2b), where $G_{ch} = 1/r_{2pt}$, and is consistent with previous studies of graphene-based THz detectors [10, 18, 50]. This standard behaviour is routinely understood in terms of a combination of resistive self-mixing and photo-thermoelectric rectification, two predominant

---

**FIG. 1. Dual-gated bilayer graphene THz detector.** a, Schematic of an hBN encapsulated dual-gated BLG transistor. THz radiation is incident on a broadband antenna connected to the source and gate terminals yielding modulation of the gate-to-source voltage. The build-up photovoltage $V_{ph}$ is read out between the source and drain terminals. b, Band structure of the BLG at the interface between the $n$-doped bottom gate-sensitive region and dual-gated $p$-doped channel. c-d, Optical photographs of the fabricated dual-gated detector. The source and top-gate terminals are connected to a broadband bow-tie antenna. e, $r_{2pt}$ as a function of $V_{tg}$ for two representative $V_{bg}$ of 0 and 2 V. Inset: Zoomed-in $r_{2pt}(V_{tg})$ for $V_{bg} = 0$ V. $T = 10$ K.
mechanisms that govern THz detection in graphene-based FETs [51].

The response of our device changes drastically when a finite vertical electric field is applied perpendicular to the BLG channel. Figure 2a shows the $R_{0}(V_{bg})$ dependence for $V_{bg} = 1.5$ V and reveals a giant increase of $R_{0}$ exceeding 3 kV/W (red curve). A striking feature of the observed dependence is its strong asymmetry with respect to zero $V_{bg}$ behaviour: namely, $|R_{0}|$ is more than an order of magnitude larger for the $p$-doped channel (to the left from the CNP in Fig. 1e) as compared to the case of $n$-doping (to the right from the CNP in Fig. 1e). In addition, while the response decays rapidly with increasing $V_{bg}$ on the $n$-doped side, a finite $R_{0}$ is observed over the whole span of $V_{bg}$ at which the channel is $p$-doped. Importantly, the $F$-factor remains fairly symmetric for the $V_{bg}$/$V_{bg}$ combinations at which $R_{0}$ exhibits strong asymmetry. This observation suggests that the strong rectification of THz radiation in our device is not caused by the non-linearities in the dual-gated BLG channel. Moreover, the increase in $R_{0}$ cannot be explained by a trivial enhancement of channel resistance. To demonstrate this, in Fig. 1d and e we plot $R_{0}$ normalized to $r_{2pt}$, a quantity with the dimension of current responsivity. A symmetric $R_{0}/r_{2pt}(V_{bg})$ dependence measured at $V_{bg} = 0$ V is conceded with amplified and highly-asymmetric curve at finite $V_{bg}$, thereby excluding resistance-enabled $R_{0}$ enhancement.

Figure 3a-b details our observations further by showing maps of $R_{0}(V_{tg}, V_{bg})$ and $r_{2pt}(V_{tg}, V_{bg})$. Enhanced $R_{0}$ is observed in two distinct quadrants characterized by an anti-symmetric (with respect to the $V_{bg}$) sign pattern (see Supplementary Section 1 for the line cuts of the map in Fig. 3a). Outside these quadrants, $R_{0}$ was found negligibly small. Interestingly, $r_{2pt}$-map is fairly symmetric featuring gradual increase of resistance at the CNP with increasing vertical field as expected for BLG [43, 44]. We have also studied the performance of our detectors at higher $f$ and found consistent highly asymmetric response similar to that shown in Fig. 2b (Supplementary Section 2) highlighting broadband character of the rectification mechanism. Furthermore, using Johnson–Nyquist relation for the noise spectral density $S = 4k_{B}T r_{2pt}$, we estimate the noise equivalent power of our detector, $\text{NEP}=S/R_{0}$, to reach 0.2 pW/$\sqrt{\text{Hz}}$ at $T \approx 10$ K (Fig. 3c).

In order to understand deeper the peculiar detection mechanism of our cooled detector, we have studied the temperature dependence of its performance. Figure 3d shows the $R_{0}(V_{tg})$ dependencies measured at $T = 10$ and 77 K in response to $f = 0.13$ THz radiation. For $p$-doped channel, $R_{0}(V_{tg})$ drops by more than 2 orders of magnitude whereas a 10-fold decrease in $R_{0}(V_{tg})$ is observed for the $n$-doped side. Furthermore, in striking contrast to the behaviour observed at 10 K, $R_{0}(V_{tg})$ curves become more symmetric at liquid nitrogen $T$. To compare, $r_{2pt}$ at the CNP also drops with increasing $T$ (inset of Fig. 3d), demonstrating usual insulating behaviour of gapped BLG at zero doping. However, for the case of $n$-doped channel, $r_{2pt}$ increases with increasing $T$, a signature of phonon-limited transport, whereas on the $p$-doped side we observed a pronounced decrease of $r_{2pt}$ with increasing $T$. The giant increase of $R_{0}$ and insulating $T$-dependence of $r_{2pt}$, when $V_{tg}$ and $V_{bg}$ are
of opposite polarities, suggest that the behaviour of our BLG detector is governed by the interband tunneling as we now proceed to demonstrate.

**Modelling tunneling-enabled photoresponse.**

Our dual-gated BLG transistor can be modelled by an equivalent circuit shown in Fig. 5 (See Methods). It consists of the gate-controlled channel conductance $\tilde{G}_{\text{ch}}$ and tunnel junctions at the source and drain with conductances $G_S$ and $G_D$, respectively. The net responsivity $R_v$ of such a circuit is the sum of three ‘intrinsic’ responsivities (marked with subscript $i$) weighted with voltage division factors $\gamma = \frac{1 + G_S/\tilde{G}_{\text{ch}}}{2}$ and multiplied by a factor of $4Z_{\text{rad}}$ relating the mean square of the antenna’s output voltage with the incident power ($Z_{\text{rad}}$ is the radiation resistance of the antenna; exact expression for the prefactor is given in Methods):

$$R_v \approx 4Z_{\text{rad}}[R_{\text{TJ},i} |\gamma|^2 + R_{\text{TG},i} \Re \gamma + R_{\text{ch},i} |1 - \gamma|^2]. \quad (1)$$

The channel responsivity, $R_{\text{ch},i}$, is proportional to the transconductance [20] and appears due to resistive self-mixing effect, i.e. due to simultaneous modulation of carrier density by transverse gate field and their drag by longitudinal field. The tunnel junction responsivity $R_{\text{TJ},i}$ emerges due to non-linear dependence of tunneling current on junction voltage, $V_{\text{TJ}}$. Finally, the responsivity $R_{\text{TG},i}$ appears due to the simultaneous action of the gate voltage that modulates tunnel barrier and junction voltage that pulls the carriers. All three contributions can be calculated from the sensitivities of conductances $G_S$ and $\tilde{G}_{\text{ch}}$ to $V_{\text{bg}}$ and $V_{\text{tg}}$ (Methods and Supplementary Section 4).

Figure 4a plots the results of such calculations in a form of 2D map which shows $R_v$ dependence on $V_{\text{tg}}$ and $V_{\text{bg}}$ (see Methods and Supplementary Section 4). The map captures all the features of the experiment, in particular, the asymmetric gate voltage dependence of the responsivity and its giant increase when the voltage of top and bottom gates is of the opposite polarity. This is most clearly visible in Fig. 4c which compares $R_v(V_{\text{tg}})$ dependencies for the cases of zero and finite $V_{\text{bg}}$. Moreover, our model indicates a broadband character of the tunneling-assisted photoresponse (Supplementary Section 2) as well as provides a remarkable quantitative agreement with experiment provided that $Z_{\text{rad}} \approx 75 \, \Omega$, a typical value for the antenna of this type [18, 51].

The peculiar response of our detector can be understood with the band diagrams shown in Fig. 4b. The detector can operate in two regimes: the regime of intraband transport (white, grey, yellow and purple symbols on the map in Fig. 4a) and the regime of interband tunneling (green and blue symbols on the map in Fig. 4a), depending on the gate voltage configuration. At zero $V_{\text{bg}}$, BLG is practically gapless, so that the tunnel barrier between the source and the channel is almost absent (white and grey symbols on the map in Fig. 4a). In this regime, the device responsivity is controlled by $R_{\text{ch},d}$ which exhibits a symmetric dependence on $V_{\text{tg}}$ (cf Fig. 4c (black line) and Fig. 2a). On the contrary, when a finite bias is
FIG. 4. Modelling tunneling-assisted THz detection. a, Calculated $R_v(V_{tg}, V_{bg})$ map of our dual-gated BLG device in response to $f = 0.13$ THz radiation. b, Calculated band profiles for different $(V_{tg}, V_{bg})$ configurations indicated by the colored symbols in (a). White, grey, yellow and pink symbols point to the band diagrams of the FET mode whereas the green and blue symbols correspond to the regime of interband tunnelling. Red, blue and black lines illustrate conduction band minimum ($E_C$), valence band maximum ($E_V$), and the chemical potential ($\mu$), respectively. c, Line cuts of the map in (a) for $V_{bg} = 0$ V and $V_{bg} = 2.5$ V. $Z_{rad} \approx 75$ Ω was used for these calculations (Methods). Inset: The ratio between the tunnel junction, $R_{TJ}$, and the channel nonlinearity, $R_{ch}$, for $V_{bg} = 2.5$ V.

applied to the bottom gate, the tunnel junction is formed as illustrated in Fig. 4b (green and blue symbols). Its intrinsic rectifying capability $R_{TJ,i}$ exceeds that of the transistor channel $R_{ch,i}$, as shown in the inset to Fig. 4c by several orders of magnitude. This stems from an ultrastrong, exponential sensitivity of the tunnel conductance to the voltage at the junction, as compared to the smooth dependence of $\tilde{G}_{ch}$ on $V_{tg}$. Moreover, the ac voltage being rectified drops almost completely on a weakly conducting junction but not on the well-conducting channel in the tunneling regime ($|\gamma| \to 1$). This can be viewed as the ‘self-localization’ of the ac field in the tunneling rectifier, which contributes to the responsivity enhancement.

Our theory, which successfully describes the response of the BLG device, can also serve to demonstrate the prospects and fundamental limits of TFET-based THz detectors. In the current device, the tunneling is assisted by fluctuations of in-plane electric field induced by local groups of charged impurities [52]. In ideal devices, the responsivity would exceed 100 kV/W, according to the model calculations (Supplementary Section 5). It is also remarkable that the expected high transconductance of TFET concedes to even higher nonlinearity of the tunnel junction, thus $R_{TJ,i} \gg R_{TG,i}$ in the present and ideal devices. $R_{TG,i}$ can dominate in situations where the source and channel simultaneously possess large gap and remain undoped; thence electron tunneling occurs from a filled valence band of the source to an empty conduction band of the channel in the vicinity of band edges. Large density of states near the bottom of ‘Mexican hat’-like spectrum of BLG further increases TFET switching steepness [53]. Realization of such band alignment is possible with the application of the drain bias and/or with extra doping gates.

In conclusion, we have shown an opportunity to use TFETs as high-responsivity detectors of sub-THz and THz radiation. Constructing a prototypical device from a BLG dual-gated structure and coupling it to a broadband antenna allowed us to demonstrate the drastic difference between a conventional FET-based approach and TFET-enabled rectification. Furthermore, we have developed a full model enabling one to predict the performance of TFET detectors based on the details of their band structure. This model was applied to the case of BLG TFET detector and successfully captured all the experimentally observed features. As an outlook, we note that BLG is just a convenient platform to demonstrate exceptional performance of TFET-based THz detectors. This approach can be extended to larger-gap materials [54] enabling room-temperature operation, as well as to CMOS-compatible structures [55]. Furthermore, we envision that alternative transistor technologies enabling transconductance beyond Boltzmann limit (phase-change FETs [56], negative capacitance FETs [57]) would also demonstrate ultra-sensitive THz detection.

*Correspondence to: gefedorov@mail.ru, Svintcov.da@mipt.ru, bandurin@mit.edu.
METHODS

Device fabrication

To fabricate tunneling-enabled BLG photodetector we first encapsulated BLG between relatively thick hBN crystals using the standard dry-peel technique [48]. The thickness of hBN crystals was measured by atomic force microscopy. The stack was then transferred on top of a predefined back gate electrode made of graphite deposited onto a low-conductivity THz-transparent silicon wafer capped with a thin oxide layer (500 nm). The resulting van der Waals heterostructure was patterned using electron beam lithography to define contact regions. Reactive ion etching was then used to selectively remove the areas unprotected by a lithographic mask, resulting in trenches for depositing electrical leads. Metal contacts to BLG were made by evaporating 3 nm of chromium and 60 nm of gold. Afterwards, a second round of e-beam lithography was used to design the top gate. The graphene channel was finally defined by a third round of e-beam lithography, followed by reactive ion etching using Poly(methyl methacrylate) and gold top gate as the etching mask. Finally, a fourth round of e-beam lithography was used to pattern large bow-tie antenna connected to the source and the top-gate terminals, followed by evaporation of 3 nm of Cr and 400 nm of Au. Antennas were designed to operate at an experimentally relevant frequency range.

Responsivity measurements

To perform the photoresponse measurements we used variable temperature optical cryostat equipped with a polyethylene window that allowed us to couple the photodetector to incident THz radiation. A Zytex-106 infrared filter was mounted in the radiation shield of the cryostat to block the 300 K background radiation. The radiation was focused to the bow-tie antenna by a silicon hemispherical lens attached to the silicon side of the chip. The transparency of the silicon wafers to the incident THz radiation over the entire T— range was verified by transmission measurements using a THz spectrometer. Photovoltage was recorded using a home-made data acquisition system based on the PXI-e 6363 DAQ board.

The responsivity of our tunneling-enabled detector was calculated assuming that the full power delivered to the device antenna funnelled into the FET channel. The open-circuit voltage $V_{\text{ph}}$ determined by this way provides a lower bound for the detectors’ responsivity and is usually referred to as extrinsic. The calculation procedure comprised several steps. First the drain-to-source voltage was recorded as a function of the top gate voltage in the dark ($V_{\text{dark}}$). Then, the dependence of the drain-to-source voltage, $V_{DS}$, on $V_{tg}$ was obtained under the illumination with THz radiation. The latter was provided by a calibrated backward wave oscillator generating $= 0.13$ THz radiation with the output power $P_{\text{full}}$ accurately measured using Golay cell. The difference $V_{\text{ph}} = V_{DS} - V_{\text{dark}}$ formed the photovoltage. The responsivity was then calculated as $R_v = V_{\text{ph}}/P_{\text{in}}$, where $P_{\text{in}} \approx P_{\text{full}}/3.5$ is the power delivered to the antenna after taking into account the losses in the silicon lens and the cryostat optical window ($\approx 5.5 $ dB).

In order to study the photoresponse of our detectors at higher $f$, we used a quantum cascade continuous wave laser based on a GaAs/Al$_{0.1}$Ga$_{0.9}$As heterostructure emitting $f = 2.026$ THz radiation. Due to the low power of the QCL and non-optimized antenna design at this $f$, the calibration of the delivered to the device antenna power was rather challenging and therefore we only report tunneling-enabled operation of our detector in relative units.

![FIG. 5. Equivalent circuit of the BLG TFET detector.](image)

Antenna is modelled as an equivalent voltage source $V_{\text{ant}}$ that generates ac current $I_{THz}$ (red arrows) flowing into the source and escaping the FET channel through the gate capacitance. Rectification occurs mainly at the tunnel barrier between source and channel with voltage-dependent conductance $G_S$.

Rectification modelling

Our detector can be modelled by an equivalent circuit (shown in Fig. 5) comprising an effective voltage source $V_{\text{ant}}$ mimicking an antenna and two nonlinear junctions connected in series with transistor channel. The detector asymmetry, required to obtain a finite photovoltage at zero bias, is provided by the asymmetric connection of antenna between source and gate, and by zero-current condition at the drain. Calculation of detector voltage responsivity $R_v = V_{\text{ph}}/P_{\text{in}}$ includes three distinct steps:

1. Relating the non-linear $I(V)$ characteristics of circuit elements to the rectified voltage $V_{\text{ph}}$.
2. Relating the power incident on antenna with its open-circuit voltage $V_{\text{ant}}$.
3. Microscopic calculation of $I(V)$ characteristics for BLG channel and its tunnel contacts.
First, it is convenient to introduce “voltage-voltage” responsivity of the TFET, \( R_{\text{TFET}} = V_{\text{out}}/V_{\text{in}}^2 \). The responsivity of bare transistor channel coupled to antenna between source and drain is the log-derivative of the dc channel conductance \( G_{\text{ch}} \) with respect to the top gate voltage \( V_{\text{tg}} \) \cite{20}, up to a geometrical factor:

\[
R_{\text{ch},i} = -\frac{1}{2} \frac{d_b}{d_t} \frac{\partial \ln G_{\text{C,de}}}{\partial V_{\text{tg}}},
\]

The presence of a tunnel junction with conductance \( G_S \) (assumed frequency-independent) depending on the voltage at the junction \( V_{\text{TJ}} \) and the top gate voltage \( V_{\text{tg}} \) results in two extra contributions to \( R_{\text{TFET}} \), which also take the form of log-derivatives:

\[
R_{\text{TJ},i} = -\frac{1}{2} \frac{\partial \ln G_S}{\partial V_{\text{TJ}}}, \quad R_{\text{TG},i} = -\frac{\partial \ln G_S}{\partial V_{\text{tg}}},
\]

Summation law (1) for individual responsivities (2) and (3) follows directly from Kirchhoff’s circuit rules.

At the second stage, the experimentally measured “voltage-power” responsivity of the photodetector \( R_v \) is related to the “voltage-voltage” responsivity of the TFET as

\[
R_v = 4Z_{\text{rad}} \left| \frac{Z_{\text{GS}}}{Z_{\text{GS}} + Z_{\text{rad}}} \right|^2 R_{\text{TFET}},
\]

assuming the incident radiation is focused within the antenna’s effective aperture. The prefactor describes voltage division between the TFET impedance \( Z_{\text{GS}} \) between gate and source and antenna radiation resistance \( Z_{\text{rad}} \).

Finally, we calculate the \( I(V_{\text{in}},V_{\text{tg}}) \)-characteristics of circuit elements microscopically. The FET channel is described within drift-diffusive model with constant mobility \( \mu_{\text{BLG}} = 10^5 \text{ cm}^2/(\text{V} \cdot \text{s}) \), a value close to that found in the experiment. Short junctions are described within quantum ballistic model \cite{53}. Both the flux of carriers incident on tunnel barrier and its transparency depend on BLG band structure. This model results in an approximate relation for source junction conductance \( G_S \approx 2e^2/(\pi \hbar D_{\text{tun}} k_{\text{L-tun}} W) \), where \( D_{\text{tun}} \) is the barrier transparency for normal incidence, and \( k_{\text{L-tun}} \) is the characteristic transverse momentum of electrons participating in tunneling. To obtain vanishing junction resistance in the absence of tunnel barrier, \( D_{\text{tun}} \) was replaced by \( \sqrt{1 - D_{\text{tun}}} \) \cite{58}. The appearance of high-transparency regions across the tunnel barrier due to local electric potential fluctuations was modelled as an increase of the average field inside the tunnel barrier \( F_{\text{bar}} \) by a constant value \( F_{\text{fluct}} \). A value of \( F_{\text{fluct}} \approx 8 \text{ kV/cm} \) was extracted from the experimental resistance \( R_{2\text{pt}} \) in the tunneling regime of detector operation.

The calculation of TFET band structure in the double-gated and single-gated regions is based on a parallel-plate capacitor model supplemented with relations between charge densities on graphene layers, their electric potentials, and BLG bandstructure \cite{59}. The transient region with tunnel junction was modeled using an original approach, where screening by the charges in BLG was treated approximately by placing a fictitious conducting plane under BLG. The position and potential of this plane are chosen to yield the correct electric potential deep inside the source and channel regions of the BLG. This reduces our electrostatic problem to finding the fringing field of a capacitor, solved analytically by Maxwell \cite{60}.

**ACKNOWLEDGEMENTS**

This work was supported by the Russian Foundation for Basic Research within Grants No. 18-37-20058 and No. 18-29-20116. Experimental work of IG (photoreponsence measurements) was supported by the Russian Foundation for Basic Research (grant 19-32-80028). We acknowledge support of the Russian Science Foundation grant No. 19-72-10156 (NEP analyses) and grant No.17-72-30036 (transport measurements). The work of GA and DS (theory of THz detection) was supported by grant # 16-19-10557 of the Russian Scientific Foundation. K.W. and T.T. acknowledge support from the Elemental Strategy Initiative conducted by the MEXT, Japan, Grant Number JPMXP0112101001, JSPS KAKENHI Grant Number JP20H00354 and the CREST(JPMJCR15F3), JST. The authors thank A. Lisesanskas, W. Knap, A. I. Berdyugin and M.S. Shur for helpful discussions.

[1] M. Dyakonov and M. Shur, IEEE Trans. El. Dev. **43**, 380 (1996).
[2] W. Knap, M. Dyakonov, D. Coquillet, F. Teppe, N. Dyakonova, J. Lusakowski, K. Karpierz, M. Sakowicz, G. Valusis, D. Seliuta, et al., *Journal of Infrared, Millimeter, and Terahertz Waves* **30**, 1319 (2009).
[3] S. Boppel, A. Lisesanskas, M. Mundt, D. Seliuta, L. Minkevicius, I. Kasalynas, G. Valusis, M. Wittendorf, S. Winnerl, V. Krozer, and H. G. Roskos, *IEEE Transactions on Microwave Theory and Techniques* **60**, 3834 (2012).
[4] H. Hou, Z. Liu, J. Teng, T. Palacios, and S. Chua, *Scientific reports* **7**, 46664 (2017).
[5] L. Viti, A. R. Cadore, X. Yang, A. Vorobiev, J. E. Muench, K. Watanabe, T. Taniguchi, J. Stake, A. C. Ferrari, and M. S. Vitiello, *Nanophotonics* **1** (2020).
[6] V. M. Muravev, V. V. Solov’ev, A. A. Fortunatov, G. Tsydynzhapov, and I. V. Kukushkin, *JETP Letters* **103**, 792 (2016).
[7] S. Runyan, J. Liu, V. Kachorovskii, and M. Shur, *Applied Physics Letters* **111**, 121105 (2017).
[54] S. Kim, G. Myeong, W. Shin, H. Lim, B. Kim, T. Jin, S. Chang, K. Watanabe, T. Taniguchi, and S. Cho, Nat. Nanotechnol. 15, 203 (2020).
[55] R. Gandhi, Z. Chen, N. Singh, K. Banerjee, and S. Lee, IEEE Electron Device Letters 32, 437 (2011).
[56] N. Shukla, A. V. Thathachary, A. Agrawal, H. Paik, A. Aziz, D. G. Schlom, S. K. Gupta, R. Engel-Herbert, and S. Datta, Nature communications 6, 1 (2015).
[57] M. Si, C.-J. Su, C. Jiang, N. J. Conrad, H. Zhou, K. D. Maize, G. Qiu, C.-T. Wu, A. Shakouri, M. A. Alam, et al., Nature nanotechnology 13, 24 (2018).
[58] S. Datta, Electronic transport in mesoscopic systems (Cambridge university press, 1997) pp. 64–65.
[59] E. V. Castro, K. S. Novoselov, S. V. Morozov, N. Peres, J. L. Dos Santos, J. Nilsson, F. Guinea, A. Geim, and A. C. Neto, Journal of Physics: Condensed Matter 22, 175503 (2010).
[60] J. C. Maxwell, A treatise on electricity and magnetism, Vol. 1 (Clarendon press, 1873) pp. 246–248, art. 202.
SUPPLEMENTARY INFORMATION

for

Tunnel field-effect transistors for sensitive terahertz detection

I. Gayduchenko, S.G. Xu, G. Alymov, M. Moskotin, I. Tretyakov, T. Taniguchi, K. Watanabe, G. Goltsman, A.K. Geim, G. Fedorov, D. Svintsov, and D.A. Bandurin.

Supplementary Section 1: Further examples of tunnel-enabled photoresponse

Figure S1 shows further examples of our tunnel detector responsivity \( R_v \) as a function of \( V_{tg} \) recorded in response to \( f = 0.13 \) THz radiation for varying \( V_{bg} \). For all \( V_{bg} \neq 0 \), \( R_v(V_{tg}) \) dependencies are highly asymmetric. With increasing \( V_{bg} \), \( R_v \) is increasing and for \( V_{bg} = 2.6 \) V reaches 4.5 kV/W overcoming zero \( V_{bg} \) value by more than an order of magnitude. A similar behaviour was observed if the polarity of \( V_{bg} \) is reversed (blue curve in Fig. S1). These observations highlights a drastic difference between the field-effect-enabled intraband (black curve) rectification and its interband tunneling counterpart (all other curves).

![Figure S1](image_url)

**FIG. S1.** Tunneling-enabled THz detection. \( R_v \) as a function of \( V_{tg} \) for given \( V_{bg} \) recorded in response to 0.13 THz radiation. \( T = 10 \) K.
Supplementary Section 2: Frequency dependence of tunnelling-enabled photoresponse

We have also studied the response of our detectors at higher frequency and found consistent tunnelling-enabled highly-asymmetric behaviour when the top and bottom gates are biased with opposite polarity. Examples of $R_v(V_{tg})$ are shown in Fig. S2a for two characteristic $f$ from sub-THz and THz domains. Note, due to the limitation of our measurements (See Methods) we only present a relative comparison between $R_v(V_{tg})$ recorded at $f = 0.13 \text{ THz}$ and $f = 2 \text{ THz}$. However, our modelling which provides remarkable agreement with experiment at $f = 0.13 \text{ THz}$ predicts that TFET detectors are expected to perform equivalently well at both sub-THz and THz frequencies as we show in Figs. S2b,c.

FIG. S2. Frequency dependence of tunnelling-assisted THz detection a, $R_v$ as a function of $V_{tg}$ for $V_{bg} = 1.2 \text{ V}$ obtained under illumination with THz radiation of given frequency. The data normalized to their maximum value. Peaks in $R_v$ correspond to the excitation of plasmon-resonances in the detector channel [S18]. b,c Theoretical $R_v(V_{tg})$ dependencies for given $f$: as-calculated (c) and normalized to their maximum value (b).
Supplementary Section 3: Comparison with existing technology

In Fig. S3 we compare the performance of our tunnel device with other THz detectors and rectifiers; some of them are available on the market (underlined labels). To this end, we plot their noise equivalent power (NEP) versus temperature, $T$, at which they operate. The comparison is made for the frequency range $0.1 – 2$ THz and for the NEP calculated via extrinsic responsivity, i.e. which takes into account the full power delivered to the device. The devices of different types are compared: cooled superconducting bolometers [S1, S2], cooled semiconducting bolometers [S3–S5], kinetic inductance sensors [S6, S7], cooled quantum dot devices [S8, S9], as well as transistor-based detectors [S10–S15] and Schottky diodes [S16]. One of the primary tasks for the next-generation THz technology is to produce low-NEP sensors operating at elevated temperatures as indicated by the yellow shaded area in Fig. S3. However, whereas the cooled devices feature exceptionally low NEP, room-$T$ devices are usually characterized by much higher NEP. Our BLG TFET offers a compromise to this enquiry: it features relatively low NEP and operates above liquid helium $T$. Furthermore, our model suggests that TFETs with optimized parameters can feature even lower NEP at room temperature (dark yellow star in Fig. S3) and thus offer a route to the next-generation THz technology. The details are given in Supplementary Section 5.

FIG. S3. Overview of THz detectors. NEP for THz detectors of various types plotted against the temperature at which they operate. Vertical error bars represent the spread of the detectors’ performance over the frequency range $0.1 – 2$ THz. Horizontal error bars show the temperature range at which the detectors operate. Underlined labels denote commercial technology.
Supplementary Section 4: Theoretical model of a BLG TFET photodetector

Supplementary Section 4.1: Modelling of tunneling-assisted THz detection

In this section, we derive a general expression for the responsivity of a TFET. The relevant circuit is shown in Fig. S4a. We will treat the TFET as if it was single-gated, since the bottom gate is held at a constant potential and its only function is to open a bandgap in BLG.

A TFET consists of two rectification units: a tunnel junction between the source and the channel, and the channel itself. (The drain tunnel junction is effectively excluded from the circuit by the zero drain current assumption, at least if the junction is too short to accommodate any spatial inhomogenities of the current.)

When a small ac voltage $V_{in}\cos(\omega t)$ is applied between the gate and source, it induces voltages and currents in different parts of the detector, having the general form $(\delta V, \delta I(t)) = Re\left((V, I)_{(1)}(1)e^{-i\omega t}\right) + (V, I)_{(2)} + \ldots e^{\pm 2i\omega t} + o(V_{in}^2)$, where we are interested in the first-order and dc second-order components.

We use a non-distributed model for the source junction, meaning current can flow through the junction only in presence of a nonzero voltage drop across the junction and not solely under the action of ac gate voltage. Keeping this in mind, the second-order expansion of its current-voltage characteristic $I_{s}\left(V_{sS}, V_{GS}\right)$ reads

$$I^{(1)}_{s} = G_s(\omega)V_{sS}^{(1)},$$

$$I^{(2)}_{s} = G_{s,dc}V_{sS}^{(2)} - R_{TJ,i}\left[V_{sS}^{(1)}\right]^2 - R_{TG,i}\operatorname{Re}\left(V_{GS}^{(1)}V_{sS}^{(1)*}\right),$$

where $G_{s,dc}$ and $G_s(\omega)$ are the dc and ac conductance of the junction, $R_{TJ,i}$ is the intrinsic tunnel junction responsivity, and $R_{TG,i}$ is the intrinsic “tunnel-gate” responsivity.

When writing similar expressions for the current $I_{s\leftarrow}flowing$ to the source from the channel, we make advantage of the linear dependence between $V_{ds}^{(1)}$ and $V_{GS}^{(1)}$ arising from zero drain current condition, and use only $V_{GS}^{(1)}$ and $V_{ds}^{(2)}$ as independent variables (remember that dc gate voltage $V_{GS}^{(2)}$ does not produce any current by itself):
\[ I_{sc}^{(1)} = \tilde{G}_{ch}(\omega)V_{ss}^{(1)}, \]
\[ I_{sc}^{(2)} = G_{ch,dc}\left[V_{ss}^{(2)} - R_{ch,i}\left|\frac{V_{ss}^{(1)}}{2}\right|^2\right], \]  

(S2)

where \( G_{ch,dc} \) is the dc channel conductance, \( \tilde{G}_{ch}(\omega) \equiv (\partial I_{sc}^{-1}(\omega)/\partial V_{GS}(\omega))|_{I_{sc}=0} \) is the ac channel conductance measured between source and gate, and \( R_{ch,i} \) is the intrinsic channel responsivity.

From continuity of current, \( I_{sc}^{(1)} = I_{sc}^{(1)} = I_{sc}^{(2)} = I_{sc}^{(2)} = 0, \) we find that the ac voltage \( V_{GS}^{(1)} \equiv V_{in} \) applied between the gate and source is divided into voltage \( V_{ss}^{(1)} \) at the source tunnel junction and voltage \( V_{ss}^{(1)} \) between the gate and the beginning of the channel:

\[ V_{ss}^{(1)} = \frac{\tilde{G}_{ch}(\omega)}{G_{S}(\omega) + \tilde{G}_{ch}(\omega)}V_{in}, \]
\[ V_{ss}^{(1)} = \frac{G_{S}(\omega)}{G_{S}(\omega) + \tilde{G}_{ch}(\omega)}V_{in}, \]  

(S3)

which are subsequently rectified by the tunnel junction and the channel:

\[ V_{ss}^{(2)} = R_{TJ,i}\left|\frac{V_{ss}^{(1)}}{2}\right|^2 + R_{TG,i}\left|\frac{V_{ss}^{(1)}}{2}\right|^2, \]
\[ V_{ss}^{(2)} = R_{ch,i}\left|\frac{V_{ss}^{(1)}}{2}\right|^2. \]  

(S4)

These rectified voltages sum together to yield the output voltage \( V_{DS}^{(2)} \equiv V_{out} \) of the photodetector (remember that the voltage at the drain junction \( V_{DD}^{(2)} = 0 \) because of zero drain current). Total responsivity of the TFET is given by the sum of tunnel junction responsivity, coming from the nonlinear current-voltage characteristic of the source tunnel junction, tunnel-gate responsivity, coming from resistive self-mixing in the gate-controlled source tunnel junction, and channel responsivity, coming from resistive self-mixing in the channel:

\[ R_{TFET} \equiv \frac{V_{out}}{V_{in}^2/2} \equiv \frac{V_{DS}^{(2)}}{\left(V_{ss}^{(1)}\right)^2/2} = R_{TJ} + R_{TG} + R_{ch}, \]
\[ R_{TJ} \equiv \left|\frac{\tilde{G}_{ch}(\omega)}{G_{S}(\omega) + \tilde{G}_{ch}(\omega)}\right|^2 R_{TJ,i}, \]
\[ R_{TG} \equiv \text{Re}\left|\frac{\tilde{G}_{ch}(\omega)}{G_{S}(\omega) + \tilde{G}_{ch}(\omega)}\right|^2 R_{TG,i}, \]
\[ R_{ch} \equiv \left|\frac{G_{S}(\omega)}{G_{S}(\omega) + \tilde{G}_{ch}(\omega)}\right|^2 R_{ch,i}. \]  

(S5)

We will neglect the frequency dependence of the tunnel junction current-voltage characteristic. With this assumption, intrinsic tunnel junction and tunnel-gate responsivities are given by the logarithmic derivatives of the junction conductance with respect to appropriate voltages:

\[ R_{TJ,i} = -\frac{1}{2}\left(\frac{\partial \ln G_{S}}{\partial V_{ss}}\right)_{V_{ss}}^{V_{GS}}, \]
\[ R_{TG,i} = -\left(\frac{\partial \ln G_{S}}{\partial V_{GS}}\right)_{V_{ss}}. \]  

(S6)
Due to the distributed nature of the channel, its current-voltage characteristics are inherently frequency-dependent. Nevertheless, the intrinsic channel responsivity can also be expressed in terms of the logarithmic derivative of its dc conductance, see Supplementary Section 4.6:

\[ R_{ch,i} \approx -\frac{1}{2} \frac{d_b}{d_t + d_b} \left( \frac{\partial \ln G_{ch,dc}}{\partial V_{ds}} \right)_{V_{ds}=0}. \]  

(S7)

A similar expression was originally derived in Ref. S17 for a single-gated FET. The extra prefactor represents the gate voltage division in a double-gated structure with top and bottom gate dielectrics of thicknesses \(d_t, d_b\).

TFET responsivity (S5) describes its response to the ac voltage at the gate, while the experimentally measured photodetector responsivity \(R_v\) describes response to the power \(P_{in}\) incident on the antenna. The relation between these responsivities can be obtained by considering the complete circuit of the photodetector, including the antenna radiation resistance \(Z_{rad}(\omega)\) (Fig. S4a). Assuming the incident radiation is focused within the antenna’s effective aperture, the incident power can be converted into the effective voltage \(V_{ant} = \sqrt{8Z_{rad}(\omega)P_{in}}\) [S18], which is divided between \(Z_{rad}(\omega)\) and the TFET gate-to-source ac impedance

\[ Z_{GS}(\omega) = G_s^{-1} + G_{ch}^{-1}(\omega), \]  

(S8)

yielding

\[ R_v \equiv \frac{V_{out}}{P_{in}} = 8Z_{rad}(\omega) \frac{V_{out}}{|V_{ant}|^2} = 4Z_{rad}(\omega) \left| \frac{Z_{GS}(\omega)}{Z_{GS}(\omega) + Z_{rad}(\omega)} \right|^2 R_{TFET}. \]  

(S9)

Supplementary Section 4.2: Bandstructure and charge density in bilayer graphene

BLG in external electric field is described by the Hamiltonian [S19, S20]

\[ \hat{H}(k) = \begin{pmatrix} \frac{-e\varphi_t}{\hbar} & \hbar v_0(\pm k_x - ik_y) & 0 & 0 \\ \hbar v_0(\pm k_x + ik_y) & \frac{-e\varphi_t}{\hbar} & 0 & 0 \\ 0 & 0 & -e\varphi_b & \hbar v_0(\pm k_x - ik_y) \\ 0 & 0 & \hbar v_0(\pm k_x + ik_y) & -e\varphi_b \end{pmatrix}. \]  

(S10)

in the vicinity of \(K, K'\) points of the Brillouin zone, where \(\varphi_t, \varphi_b\) are the electric potentials at top and bottom graphene layers, \(\gamma_1 = 0.38\ \text{eV}, v_0 = 10^6\ \text{m/s},\) and the signs depend on the valley.

The corresponding conduction and valence band dispersions are

\[ E_{c,v}(k) = -e\varphi_+ \pm E(k), \]  

\[ E(k) = \sqrt{\frac{E_g^2}{4} + \left( \frac{\gamma_1^2 - E_g^2}{4} + (\hbar v_0 k)^2 \right)^2} \]  

(S11)

with a bandgap

\[ E_g(\varphi_-) = \frac{\gamma_1}{\sqrt{\gamma_1^2 + e^2\varphi_-^2}} |e\varphi_-|. \]  

(S12)

where \(\varphi_+ \equiv (\varphi_t + \varphi_b)/2\) is the average potential of graphene layers, and \(\varphi_- \equiv \varphi_t - \varphi_b\) is the interlayer voltage. The bands have a “Mexican hat” shape with circular extrema around the corners of the Brillouin zone (Fig. S5).

The inverse dispersion relation is

\[ k_\pm(E - e\varphi_+) = \frac{1}{\hbar v_0} \sqrt{E^2 + \frac{e^2\varphi_-^2}{4} \pm \left( \frac{\gamma_1^2 + e^2\varphi_-^2}{4} \right) \left( E^2 - \frac{E_g^2}{4} \right)}. \]  

(S13)

It is double-valued within the “Mexican hat” region, \(E_g/2 < |E| \leq |e\varphi_-|,\) while only a single solution \(k_+\) remains above the hat, \(|E| > |e\varphi_-|\).

Given the dispersion relation (S11), we can express the charge density \(\rho_+ \equiv \rho_t + \rho_b\) in BLG at zero temperature through the chemical potential measured from the midgap, \(\mu \equiv \mu + e\varphi_+\), and vice versa:
FIG. S5. Bandstructure of biased bilayer graphene described by Hamiltonian (S10) (only the conduction and valence bands are shown). Circular band extrema are highlighted in yellow.

\[
\begin{align*}
\rho_+ (\tilde{\mu}) &= \begin{cases} 
0 & \text{if } |\tilde{\mu}| < \frac{E_g}{2}, \\
-\frac{e^2}{\pi} \frac{k_F^2}{\tilde{\mu}^2} \text{sgn } \tilde{\mu}, k_F = k_+ (\tilde{\mu}) & \text{if } |\tilde{\mu}| \geq \frac{|e\varphi_-|}{2} ,
\end{cases} \\
\tilde{\mu} (\rho_+) &= \begin{cases} 
-\frac{E}{2} \frac{\text{sgn } \rho_+, k_F = k_+ (\tilde{\mu})}{|\rho_+|^2} & \text{if } |\rho_+| \geq \rho_{\text{hat}}, \\
-\frac{E^2}{4} + \frac{1}{4(\gamma_1^2 + e^2 \varphi_+^2)} \left( \hbar v_0 \sqrt{\frac{\pi |\rho_+|^2}{e}} \right) \text{sgn } \rho_+ & \text{if } 0 < |\rho_+| < \rho_{\text{hat}}.
\end{cases}
\end{align*}
\]

(S14)

where \( \rho_{\text{hat}} = (e/\pi) (e\varphi_- / \hbar v_0)^2 \) is the charge density corresponding to \( \tilde{\mu} = \pm e\varphi_- / 2 \) (the Fermi level positioned at the tip of the “Mexican hat”). We have taken into account the double valley degeneracy in BLG.

**Supplementary Section 4.3: Electrostatics of double-gated bilayer graphene**

To calculate the band diagram of our TFET, we seek approximate analytical solution of electrostatic equations for double-gated BLG.

Let \( V_{tg}, V_{bg} \) be the potentials of the top and bottom gate, \( d_t, d_b \) the thicknesses of dielectric layers separating BLG from the gates, \( \kappa_t, \kappa_b \) the dielectric constants of these layers, and \( d \) the interlayer distance in BLG. Then, the total charge density \( \rho_+ \equiv \rho_t + \rho_b \) and interlayer charge transfer \( \rho_- \equiv (\rho_t - \rho_b)/2 \) are related to the electric potentials \( \varphi_t, \varphi_b \) of top and bottom graphene layers by

\[
\begin{align*}
\rho_+ &= -C_t (V_{tg} - \varphi_t) - C_b (V_{bg} - \varphi_b), \\
\rho_- &= -\frac{C_t (V_{tg} - \varphi_t)}{2} + \frac{C_b (V_{bg} - \varphi_b)}{2} + C_{\text{cl}} (\varphi_t - \varphi_b),
\end{align*}
\]

(S15)

where we introduced capacitances per unit area \( C_{\text{cl}} \equiv \epsilon_0 / d, C_t \equiv \kappa_t \epsilon_0 / d_t, C_b \equiv \kappa_b \epsilon_0 / d_b \).

The potentials of graphene layers stay close to the Fermi level (compared to the gate voltages), and we can substitute \( V_{tg} - \varphi_t / h \to V_{tg} + \mu / e \) in the second equation. We cannot do the same in the first equation, otherwise it would not work properly in the undoped case. Instead, in the first equation we approximate \( \varphi_t \approx \varphi_b \approx \varphi_+ \) to decouple \( \rho_+, \varphi_+ \) and \( \rho_-, \varphi_- \).
\(\varphi_{-} = 10 \text{ meV}\)
\(\varphi_{-} = 30 \text{ meV}\)
\(\varphi_{-} = 100 \text{ meV}\)
\(\varphi_{-} = 10 \text{ meV}\)
\(\varphi_{-} = 30 \text{ meV}\)
\(\varphi_{-} = 100 \text{ meV}\)

\(T = 0 \text{ K}\)
\(T = 77 \text{ K}\)

**FIG. S6.** Ratio between the interlayer charge transfer calculated using a constant interlayer quantum capacitance \(C^q_{-} = \frac{3e^2 \gamma_1}{4\pi\hbar^2 v_0^2}\) and the exact interlayer charge transfer \(\rho_{-}\) calculated from Hamiltonian (S10) as described in [S19]. Left panel: zero temperature, right panel: \(T = 77 \text{ K}\).

\[
\rho_+ \approx -C_t V_{tg} - C_b V_{bg} + (C_t + C_b) \varphi_+,
\]
\[
\rho_- = -\frac{C_t}{2} \left( V_{tg} + \frac{\mu}{e} \right) + \frac{C_b}{2} \left( V_{bg} + \frac{\mu}{e} \right) + C^q_{-} \varphi_-.
\]

The introduced approximations essentially amount to a minor shift of gate voltages, by the order of magnitude equal to \(\varphi_t, \varphi_b\).

These equations have to be supplemented with explicit expressions for \(\rho_{+}(\tilde{\mu}, \varphi_-)\) in BLG (\(\tilde{\mu} \equiv \mu + e\varphi_+\) is the chemical potential with respect to the midgap). To facilitate analytical treatment, we use zero-temperature expression for the total charge density (S14) and a constant quantum capacitance model for the interlayer charge transfer:

\[
\rho_+ = \begin{cases} 0 & \text{if } |\tilde{\mu}| \leq \frac{E_g(\varphi_-)}{2}, \\ \rho_+(\tilde{\mu}) & \text{if } |\tilde{\mu}| > \frac{E_g(\varphi_-)}{2}, \end{cases}
\]

\[
\rho_- \approx -C^q_{-} \varphi_-.
\]

The constant interlayer quantum capacitance \(C^q_{-} = \frac{3e^2 \gamma_1}{4\pi\hbar^2 v_0^2}\) approximates the interlayer charge transfer in BLG over a wide range of bangaps and doping levels within 50% accuracy (see Fig. S6).

Now, the equation for \(\rho_-\) becomes trivial to solve, while the equation for \(\rho_+\) requires some additional simplifications to allow analytical solution. We consider two opposite cases: (1) Fermi level lies within the bandgap, (2) Fermi level lies outside the gap. In the first case, \(\rho_+ = 0\) at zero temperature, and \(\varphi_+\) is readily obtained from (S16). In the second case, we can pick some initial guess for \(\varphi_+\), find \(\rho_+\) from (S16), and find a better approximation for \(\varphi_+\) from (S17). Since the quantum capacitance \(C^q_{+} \sim \epsilon_0/d\) is much larger than \(C_t + C_b\), the precise value of the initial guess is unimportant, and we initially assume the Fermi level is pinned at the band edge, \(\tilde{\mu} = \pm E_g/2\) (this choice avoids spurious discontinuities in \(\varphi_+(V_{tg}, V_{bg})\)).

The overall procedure is summarized in the following equations:

\[
\varphi_- \approx \frac{C_t (V_{tg} + \mu/e) - C_b (V_{bg} + \mu/e)}{2C_-}, \quad C_- \equiv C^q_{-} + C^q_{+},
\]
\[
E_g = E_g(\varphi_-),
\]
\[
\tilde{\mu}_0 = e \frac{C_t (V_{tg} + \mu/e) + C_b (V_{bg} + \mu/e)}{C_t + C_b},
\]
\[
\tilde{\mu} \approx \begin{cases} \tilde{\mu}_0 & \text{if } |\tilde{\mu}_0| \leq \frac{E_g}{2}, \\ \tilde{\mu}_0 \pm C_t + C_b & \text{if } |\tilde{\mu}_0| > \frac{E_g}{2}, \end{cases}
\]
where \( E_q(\varphi_-), \tilde{\mu}(\rho_+) \) are given by (S12) and (S14).

In our calculations, we used \( d_t = 80 \) nm, \( d_b = 50 \) nm, \( d = 0.335 \) nm, and \( \kappa_t = \kappa_b = 3.76 \) (out-of-plane static dielectric constant of hexagonal boron nitride [S21]).

We use this parallel-plate capacitor model to find the electric potentials \( \varphi_{+S}, \varphi_{+C} \) and interlayer voltages \( \varphi_{-S}, \varphi_{-C} \) in the source region and in the channel in absence of ac signal, and also to calculate the channel response to an ac signal, see Supplementary Section 4.6. In the source region, there is only the bottom gate, while the role of a top gate is played by infinity, held at zero potential. This means \( C_t = 0 \), and the top gate disappears from the equations.

**Supplementary Section 4.4: Tunneling field**

In a TFET based on double-gated BLG, a tunnel junction is formed under the top gate edge, where the parallel-plate capacitor model of Supplementary Section 4.3 cannot be applied, and an accurate calculation of the tunneling field requires solving a two-dimensional electrostatic problem. This problem can be solved analytically in the absence of BLG [S22], and the answer is

\[
\tilde{E}_x = \frac{1 - \tilde{\varphi}}{\left[1 + \tilde{y} (1 - \tilde{\varphi}) \cot \tilde{\varphi}\right]^2 + \tilde{y}^2 \left(1 - \tilde{\varphi}\right)^2},
\]

where

\[
\tilde{E}_x \equiv \frac{E_x d_{tb}}{\pi \left[\varphi(x = -\infty) - \varphi(x = +\infty)\right]}, \quad \tilde{y} \equiv \frac{\pi y}{d_{tb}}, \quad \tilde{\varphi} \equiv \frac{\varphi - \varphi(x = -\infty)}{\varphi(x = +\infty) - \varphi(x = -\infty)}
\]

are the dimensionless field in the plane of BLG, dimensionless position of BLG with respect to the gates (\( \tilde{y} = 0 \) at the bottom gate and \( \pi \) at the top gate), and dimensionless electric potential at the point where the field is calculated.

\( d_{tb} = d_t + d_b + d \) is the distance between gates, \( \varphi(x = -\infty) = V_{bg} \) and \( \varphi(x = +\infty) = V_{bg} + \frac{y}{d_{tb}}(V_tg - V_{bg}) \) are the electric potential in the source region and in the channel. The top and bottom dielectrics are assumed to be the same, as in our experiment.

Across a wide range of \( \tilde{y} \), \( \tilde{E}_x \) is close to its low-\( \tilde{y} \) limit

---

**FIG. S7.** (a) Color map showing the distribution of electric potential near the source-channel junction. Black lines: equipotential lines and field lines. Dashed red line: fictitious conductor introduced to obtain the correct potential in the source and the channel without explicitly considering screening by BLG. Potential above the fictitious conductor was calculated as prescribed in Ref. S22, potential below the fictitious conductor was calculated in the parallel-plate capacitor model. (b) Electric potential \( \varphi_{+}(x) \) inside BLG.
\[
\tilde{E}_x \approx \tilde{\varphi}^2 (1 - \tilde{\varphi}).
\] (S21)

In the presence of BLG, exact calculation of the tunneling field would require solving the two-dimensional electrostatic problem numerically. To avoid this, we notice that adding BLG into the system reduces \( \varphi(x = +\infty) - \varphi(x = -\infty) \) from several volts to tens or hundreds of millivolts. This suggests to approximate the screening by BLG via introducing a fictitious perfect conductor placed very close to the BLG. The potential of this conductor and its distance from the BLG are chosen so as to reproduce the correct potentials in the source and channel regions of BLG.

The resulting electric potential distribution in the system is shown in Fig. S7. Introducing the fictitious conductor allows us to keep Eq. (S21) for the electric field in BLG, if \( d_0 \) is replaced with \( d_t \) in the definition of (Eq. (S20)), and \( \varphi(x = \pm\infty) \) are calculated in the parallel-plate capacitor model described in Supplementary Section 4.3.

Knowing the distribution of electric potential in BLG, we can calculate the tunnel current through the source-channel junction. Before we actually do this, we introduce two additional simplifications. First, we neglect field variations inside the barrier and assume tunneling through uniform field. This field is calculated at the point where the tunneling electron crosses the midgap \( (E + e\varphi = 0, \text{where } E \text{ is the electron energy}) \). Second, instead of using different values of the tunneling field for electrons of different energies, we use a single value calculated for energy \( E = (E_{\text{tun,min}} + E_{\text{tun,max}})/2 \). \( E_{\text{tun,min}} \) and \( E_{\text{tun,max}} \) are the boundaries of the energy region where tunneling is possible. Assuming zero temperature and both quasi-Fermi levels \( \mu_S, \mu_C \) in the source and the channel (near its beginning) lying within the band overlap region, we can write \( E_{\text{tun,min}} = \min \{\mu_S, \mu_C\} \) and \( E_{\text{tun,max}} = \max \{\mu_S, \mu_C\} \). (Remember that we are interested in the small-signal case, when the quasi-Fermi levels are close to each other and either lie both inside the band overlap region, or both outside. In the latter case, tunneling is impossible.)

To summarize, we use the following expression for the tunneling field:

\[
F_{\text{tun}} \approx \frac{\pi |\varphi+ - \varphi_{+\text{tun}}|}{d_t} \left( \frac{\varphi_{+\text{tun}} - \varphi_S}{\varphi_C - \varphi_S} \right)^2,
\] (S22)

where \( \varphi_S, \varphi_C \) are calculated as described in Supplementary Section 4.3, and \( -e\varphi_{+\text{tun}} = (\mu_S + \mu_C)/2 \).

**Supplementary Section 4.5: Responsivity of the source-channel junction**

A zero-temperature ballistic expression for the tunnel current through the source-channel junction is

\[
I_{\leftarrow s,\text{tun}} = 8eW \int_{\mu_C}^{\mu_S} \frac{dE}{2\pi\hbar} \int_{-k_{\text{max}}(E)}^{k_{\text{max}}(E)} \frac{dk_\perp}{2\pi} D(E, k_\perp)
\] (S23)

if \( \mu_S > \mu_C \) (the opposite case is treated similarly). Here, \( W = 6.2 \mu m \) is the channel width, \( D(E, k_\perp) \) is the barrier transparency, the wavevector integral is taken up to the maximum possible transverse wavevector \( k_{\text{max}}(E) \) that an electron with energy \( E \) can have both in the source and in the channel, and the factor of 8 results from two spin projections, two valleys, and two tunneling paths in the imaginary \( k \)-space (interference between them [S23] is neglected).

An analytical approximation can be derived by expanding the WKB barrier transparency in powers of \( k_\perp \) up to second order and extending the wavevector integration up to infinity [S24]:

\[
I_{\leftarrow s,\text{tun}} \approx \frac{2e}{\pi^{3/2}\hbar} D_{\text{tun}} k_{\text{tun}} W (\mu_S - \mu_C),
\]

\[
D_{\text{tun}} \approx \exp \left( \frac{\pi}{4\hbar v_0 e F_{\text{tun}}} \right),
\]

\[
k_{\text{tun}} \approx \sqrt{\frac{4}{\pi}} \frac{\gamma_1}{E_{g,\text{tun}}} e F_{\text{tun}}.
\]

We assume that the transition between the source and the channel has the same shape for both the interlayer voltage and the electric potential and, therefore, the tunnel current flows through the bandgap \( E_{g,\text{tun}} \approx |\varphi_{-\text{tun}}| \), where

\[
\frac{\varphi_{-\text{tun}} - \varphi_S}{\varphi_C - \varphi_S} = \left( \frac{\varphi_{+\text{tun}} - \varphi_S}{\varphi_C - \varphi_S} \right) + C,
\] (S25)
At experimental conditions, the bandgap does not exceed 60 meV, so we use $E_g \approx |\varphi_ -|$ instead of a more accurate expression (S12).

Expressions (S23), (S24) require that the chemical potentials $\mu_ S, \mu_ C$ are taken at the points where the deviations of the carrier distributions from the Fermi-Dirac form are negligible. Since we consider the tunnel junction connected in series with the channel, we need an expression for the tunnel current in terms of the voltage directly at the junction, otherwise a certain part of the channel would be counted twice. This can be achieved by introducing a $1 + \mathcal{D}_\text{tun}$ correction in the denominator:

$$I_ {\text{es}, \text{tun}} \approx \frac{2e}{\pi^{3/2} \hbar} \frac{\mathcal{D}_\text{tun}}{1 - \mathcal{D}_\text{tun}} k_ {\perp \text{tun}} W (\mu_ S - \mu_ C),$$

(S26)

similarly to the one-dimensional Landauer formula containing $\mathcal{D}/(1 - \mathcal{D})$ [S25, S26].

The idealistic model that led to Eq. (S26) gives very small barrier transparency and huge tunnel resistance, orders of magnitude larger than in our experiment. This suggests there is some mechanism affecting the junction resistance, most likely electron-hole puddles, that create field fluctuations and may increase the average tunneling field. We take this effect into account phenomenologically, introducing a single fitting parameter $F_{\text{fluct}}$, which represents the average fluctuating field and is added to the tunneling field (S22) calculated without disorder:

$$I_ {\text{es}, \text{tun}} \approx \frac{2e}{\pi^{3/2} \hbar} \frac{\mathcal{D}_\text{tun}}{1 - \mathcal{D}_\text{tun}} k_ {\perp \text{tun}} W (\mu_ S - \mu_ C),$$

(S27)

$$\mathcal{D}_\text{tun} \approx \exp \left( -\frac{\pi \sqrt{\gamma_1 E_{g, \text{tun}}}}{4\hbar v_0 e (F_{\text{tun}} + F_{\text{fluct}})} \right),$$

$$k_ {\perp \text{tun}} \approx \sqrt{\frac{4}{\pi} \frac{\gamma_1 e (F_{\text{tun}} + F_{\text{fluct}})}{E_{g, \text{tun}} \hbar v_0}}.$$  

This is the final expression for the tunnel current that we used in our calculations. The value $F_{\text{fluct}} = 8$ kV/cm was found by fitting the experimental resistance in the tunnel regime and simultaneously gave responsivity in reasonable agreement with the experiment.

Assuming grounded source, $\mu_ S = 0$, we identify $\mu_ C$ with $-eV_ {S} \mu$ and $V_ \mu$ with $V_ {GS}$ from Supplementary Section 4.1. Now, we can calculate the junction conductance as $G_ S = -e\partial I_ {\text{es}, \text{tun}}/\partial \mu_ C$ and the intrinsic tunnel junction and tunnel-gate responsivities through (S6). When the doping types of source and channel are the same, or channel is undoped, there is no tunnel junction. In this case, we set the junction conductance to infinity and tunnel junction and tunnel-gate responsivities to zero.

**Supplementary Section 4.6: Responsivity of a long double-gated channel**

In this section, we consider resistive self-mixing in a long [S27] double-gated channel and find its responsivity. Our derivation closely follows that of Ref. S17, but extends it by (1) allowing the carrier density to depend separately on the top gate voltage and the Fermi level (because $\rho_+ = \rho_+ (\mu + eV_ \mu)$ is no longer true in the presence of a bottom gate), (2) using frequency-dependent channel conductivity.

The basic assumptions of our model are that the dc channel conductivity $\sigma_\text{dc}(x, t)$ is instantaneously related to the local charge density $\rho_+(x, t)$, which, in turn, is related (also locally and instantaneously) to the top gate voltage $V_ \mu(t)$ and the Fermi level $\mu(x, t)$. Response to ac perturbations is described within the Drude model. Together with the charge conservation, we get a system of four equations:

$$\rho_+(x, t) = \rho_+ (V_ \mu(t), \mu(x, t)), \quad \frac{\partial J_+ (x, t)}{\partial t} = -\frac{1}{e} \frac{\sigma_\text{dc}(x, t)}{\tau} \mu(x, t) = -\frac{J_+ (x, t)}{\tau},$$

(S28)

$$\sigma_\text{dc}(x, t) = \sigma_\text{dc} [\rho_+(x, t)], \quad \frac{\partial \rho_+(x, t)}{\partial t} = \frac{\partial J_+ (x, t)}{\partial x},$$

with the boundary conditions of grounded source and zero drain current:

$$\mu(0, t) = 0, J_+ (+\infty, t) = 0.$$  

(S29)
The top gate voltage consists of a constant bias and an ac signal, \( V_{tg}(t) = V_{tg}^{(0)} + V_{in} \cos(\omega t) \). (Hereafter, quantities in the absence of the ac signal will be denoted by the \(^{0}\) superscript, while next orders in \( V_{in} \) will be denoted by \(^{1}\) and \(^{2}\), as in Supplementary Section 4.1.)

To the first order in \( V_{in} \), we obtain

\[
\rho_{+}^{(1)} = \frac{\partial \rho_{+}}{\partial V_{tg}} V_{in} + \frac{\partial \rho_{+}}{\partial \mu} \mu^{(1)},
\]

\[
J_{\downarrow}^{(1)} = -\frac{1}{e} \frac{\sigma_{dc}^{(0)}}{1 - i \omega \tau} \frac{\partial \mu^{(1)}}{\partial x},
\]

\[
\sigma_{dc}^{(1)} = \frac{d \sigma_{dc}^{(1)}}{d \rho_{+}^{(1)}},
\]

\[
-i \omega \rho_{+}^{(1)} = \frac{\partial J_{\downarrow}^{(1)}}{\partial x}.
\]

(S30)

Using the boundary conditions (S29), we get the following solution

\[
\mu^{(1)}(x) = -\left( \frac{\partial \rho_{+}}{\partial \mu} \right)^{-1} \frac{\partial \rho_{+}}{\partial V_{tg}} V_{in} \left[ 1 - e^{i q_{pl} x} \right],
\]

\[
\rho_{+}^{(1)}(x) = \frac{\partial \rho_{+}}{\partial V_{tg}} V_{in} e^{i q_{pl} x},
\]

\[
J_{\downarrow}^{(1)}(x) = -\frac{1}{e} \frac{\sigma_{dc}^{(0)}}{1 - i \omega \tau} i q_{pl} \left( \frac{\partial \rho_{+}}{\partial \mu} \right)^{-1} \frac{\partial \rho_{+}}{\partial V_{tg}} V_{in} e^{i q_{pl} x},
\]

\[
q_{pl} \equiv \sqrt{\frac{i \omega (1 - i \omega \tau)}{\sigma_{dc}^{(0)}} \frac{\partial}{\partial (-\mu/e)}},
\]

(S31)

Having found the first-order current, we can write the channel “source-gate” conductance (\( W \) is the channel width):

\[
\tilde{G}_{ch}(\omega) \equiv \frac{J_{\downarrow}^{(1)}(0) W}{V_{in}} = -\frac{\sigma_{dc}^{(0)}}{1 - i \omega \tau} i q_{pl} W \left( \frac{\partial (-\mu/e)}{\partial V_{tg}} \right) \rho_{+}.
\]

(S32)

(Note that we use \( \exp(-i \omega t) \) for the time dependence of harmonic signals instead of \( \exp(+j \omega t) \) convention prevalent in electrical engineering, resulting in reactances having unconventional signs.)

The equation on the second order dc current results from the zero dc drain current condition,

\[
J_{\downarrow}^{(2)} = -\frac{1}{e} \sigma_{dc}^{(0)} \frac{\partial \mu^{(2)}}{\partial x} - \frac{1}{2 e} \text{Re} \left( \sigma_{dc}^{(1)} \frac{\partial \mu^{(1)*}}{\partial x} \right) = 0,
\]

(S33)

yielding the intrinsic channel responsivity

\[
R_{ch,i} \equiv \frac{V_{out}}{|V_{in}|^2/2} = \frac{[\mu^{(2)}(\infty) - \mu^{(2)}(0)] / (-\mu)}{|V_{in}|^2/2}
\]

\[
= -\frac{1}{2} \left( \frac{\partial \ln \sigma_{dc}}{\partial V_{tg}} \right)_{\mu} \left( \frac{\partial (-\mu/e)}{\partial V_{tg}} \right) \rho_{+}.
\]

(S34)

The expressions (S32), (S34), and the definition of \( q_{pl} \) (S31) differ from the results of Ref. S17 by two extra factors. The first factor \( (\partial (-\mu/e) / \partial V_{tg})_{\rho_{+}} \) is unity in a single-gated FET and reduces to approximately \( C_{t} / (C_{t} + C_{b}) = \frac{\tilde{d}_{b}}{\tilde{d}_{t} + \tilde{d}_{b}} \) in the presence of a bottom gate. The second factor \( 1/(1 - i \omega \tau) \) appears due to the frequency dependence of conductivity.
Calculations show that the difference between \( \partial (-\mu/e) / \partial V_{tg} \) and \( \tilde{d}_b / (\tilde{d}_t + \tilde{d}_b) \) is minor and can be neglected within the accuracy of our model, so we used the following expressions for the channel “source-gate” conductance and intrinsic channel responsivity:

\[
\begin{align*}
\tilde{G}_{\text{ch}}(\omega) &= -\frac{\sigma_{dc}^{(0)}}{1-i\omega\tau} iq_{pl} W \frac{\tilde{d}_b}{\tilde{d}_t + \tilde{d}_b}, \\
R_{\text{ch},i} &= -\frac{1}{2} \left( \frac{\partial \ln \sigma_{dc}}{\partial V_{tg}} \right) \frac{d_b}{d_t + d_b}, \\
q_{pl} &= \sqrt{i\omega(1-i\omega\tau) \frac{\partial \rho_+}{\partial (-\mu/e)}}.
\end{align*}
\] (S35)

The derivatives in \( (S35) \) were evaluated with the help of the approximate electrostatic model presented in Supplementary Section 4.3 and the constant-mobility approximation for the channel dc conductivity:

\[
\sigma_{dc}^{(0)} = |\rho_+| \mu_{\text{BLG}} + \sigma_{\text{residual}}(V_{bg}),
\] (S36)

where we take \( \mu_{\text{BLG}} = 10^5 \text{ cm}^2/(\text{V} \cdot \text{s}) \) (according to measurements performed on similar devices [S14]). The transport relaxation time \( \tau \) was taken to be 2 ps according to the relation \( \mu_{\text{BLG}} = e\tau/m^* \), where \( m^* = \gamma_1/2v_0^2 \) (this is the carrier effective mass in the band extrema of gapless BLG; in gapped BLG band dispersion is similar to the gapless case except in close vicinity of the band edges, so we neglect the bandgap dependence of \( m^* \)).

The residual conductivity \( \sigma_{\text{residual}} \) due to potential fluctuations in the channel was obtained by fitting the following formula to the experimental dc resistance at the channel neutrality point:

\[
\frac{L}{W} \sigma_{\text{residual}}^{-1}(V_{bg}) = \frac{r_{\infty} V_{bg}^2 + r_0 V_0^2}{V_{bg}^2 + V_0^2}.
\] (S37)

The fitting procedure yielded \( r_0 = 200 \Omega, r_{\infty} = 150 \text{ k}\Omega, V_0 = 5.5 \text{ V} \).

Using the intrinsic channel responsivity and channel “source-gate” conductance, together with the intrinsic tunnel junction and tunnel-gate responsivities and the tunnel junction conductance found in Supplementary Section 4.5, we can obtain the total responsivity of our transistor through Eq. (S5) and convert it to the photodetector responsivity through Eq. (S9).
Supplementary Section 5: Performance limits of BLG TFET photodetectors

The theory described in the previous Supplementary Sections was used to calculate the theoretical responsivity of our photodetector, which is shown in Fig. 4 of the main text. The detector responsivity in our theory is limited by the electric potential fluctuations and could be substantially improved in devices with reduced density of charged impurities. Figure S8 shows the theoretical responsivity of our photodetector in absence of potential fluctuations (that is, with $F_{\text{fluct}} = 0$), which reaches hundreds kV/W.

Another way to increase detector responsivity is to exploit large nonlinearity of the tunnel junction at small values of band overlap (when the tunnel current is about to be switched off). This requires that the conduction band edge in the source region is simultaneously aligned with the valence band edge in the channel (or vice versa) and with the Fermi level. Such kind of band alignment can be realized by introducing an additional gate above the source region and could potentially result in infinite responsivity in the idealized model (no potential fluctuations, no leakage currents). In practice, the maximum achievable responsivity will be limited by potential fluctuations and leakage currents. Thermionic leakage hinders the performance of our detector at non-cryogenic temperatures because of the small bandgap ($<60$ meV) realized in our TFET, but this problem can be mitigated by increasing the bandgap, either by applying a larger vertical field to BLG, or by using larger-gap materials, such as black phosphorus. Electric potential fluctuations present a more fundamental issue and limit the logarithmic derivatives of the tunnel conductance to $\sim 1/V_{\text{fluct}}$, where $V_{\text{fluct}}$ is the magnitude of these fluctuations.

Assuming the total responsivity is dominated by $R_{\text{TJ}}$ (as in our photodetector) and using equations (S5), (S6), and (S9), we can estimate the achievable room-temperature noise equivalent power as

$$\text{NEP}_{\text{min}} = \frac{\sqrt{4r_{\text{pt}}k_BT}}{|R_v|} = \frac{\sqrt{4r_{\text{pt}}k_BT}}{4Z_{\text{rad}} \left( Z_{GSS} + Z_{\text{rad}} \right)} \approx \frac{\sqrt{4r_{\text{pt}}k_BT}}{4Z_{\text{rad}} \left( \frac{r_S}{r_S + Z_{\text{rad}}} \right)^2} \frac{1}{2} \left| \frac{\partial \ln G_S}{\partial V_S} \right|$$

where $r_S = G_S^{-1}$ is the resistance of the source tunnel junction. To minimize the noise equivalent power, we assumed $r_S = 3Z_{\text{rad}}$, the drain junction is absent, and the channel resistance is negligible.

Taking $Z_{\text{rad}} = 75$ $\Omega$ and $V_{\text{fluct}} = 1$ mV (an experimentally achievable value [S28]), we estimate that the room-temperature noise equivalent power in TFET-based photodetectors can be made as low as 0.02 pW/$\sqrt{\text{Hz}}$ (shown in Fig. S3).
SUPPLEMENTARY REFERENCES

S1. Boston Electronics Corporation, https://www.boselec.com.
S2. SCONTEL, http://www.scontel.ru.
S3. Infrared Laboratories, Inc., https://www.infraredlaboratories.com.
S4. D. C. Alsop, C. Inman, A. E. Lange, and T. Wilbanks, “Design and construction of high-sensitivity, infrared bolometers for operation at 300 mK,” Applied Optics 31, 6610–6615 (1992).
S5. S. T. Tanaka, A. Clapp, M. J. Devlin, M. L. Fischer, C. Hagmann, A. E. Lange, and P. L. Richards, “100-mK bolometric receiver for low-background astronomy,” in Infrared Detectors and Instrumentation, Vol. 1946 (International Society for Optics and Photonics, 1993) pp. 110–115.
S6. J. Hubmayer, J. Beall, D. Becker, H.-M. Cho, M. Devlin, B. Dober, C. Groppi, G. C. Hilton, K. D. Irwin, D. Li, et al., “Photon-noise limited sensitivity in titanium nitride kinetic inductance detectors,” Applied Physics Letters 106, 073505 (2015).
S7. A. Monfardini, A. Benoit, A. Bideaud, L. Swenson, A. Cruciani, P. Camus, Hoffmann, F. X. Désert, S. Doyle, P. Ade, et al., “A dual-band millimeter-wave kinetic inductance camera for the IRAM 30 m telescope,” The Astrophysical Journal Supplement Series 194, 24 (2011).
S8. P. Kleinschmidt, S. P. Giblin, V. Antonov, H. Hashiba, L. Kulik, A. Tzalenchuk, and S. Koniyma, “A highly sensitive detector for radiation in the terahertz region,” IEEE Transactions On Instrumentation and Measurement 56, 463–467 (2007).
S9. J.-H. Dai, J.-H. Lee, Y.-L. Lin, and S.-C. Lee, “In(Ga)As quantum rings for terahertz detectors,” Japanese Journal of Applied Physics 47, 2924 (2008).
S10. Terasense Group, Inc., https://terasense.com.
S11. L. Vicarelli, M. S. Vitiello, D. Coquillat, A. Lombardo, A. C. Ferrari, W. Knap, M. Polini, V. Pellegrini, and A. Tredicucci, “Graphene field-effect transistors as room-temperature terahertz detectors,” Nature Materials 11, 865–871 (2012).
S12. A. A. Generalov, M. A. Andersson, X. Yang, A. Vorobiev, and J. Stake, “A 400-GHz graphene FET detector,” IEEE Transactions on Terahertz Science and Technology 7, 614–616 (2017).
S13. D. A. Bandurin, I. Gayduchenko, Y. Cao, M. Moskotin, A. Principi, I. V. Grigorieva, G. Goltsman, G. Fedorov, and D. Svintsov, “Dual origin of room temperature sub-terahertz photoresponse in graphene field effect transistors,” Appl. Phys. Lett. 112, 141101 (2018).
S14. D. A. Bandurin, D. Svintsov, I. Gayduchenko, S. G. Xu, A. Principi, M. Moskotin, I. Tretyakov, D. Yagodkin, S. Zhukov, T. Tamiguchi, et al., “Resonant terahertz detection using graphene plasmons,” Nature Communications 9, 1–8 (2018).
S15. A. El Fatimy, F. Teppe, N. Dyakonova, W. Knap, D. Seliuta, G. Valuˇsis, A. Shchepetov, Y. Roelens, S. Bollaert, A. Cappy, et al., “Resonant and voltage-tunable terahertz detection in InGaAs/InP nanometer transistors,” Applied Physics Letters 89, 131926 (2006).
S16. TOPTICA Photonics AG, https://www.toptica.com.
S17. M. Sakowicz, M. B. Lifshits, O. A. Klimenko, F. Schuster, D. Coquillat, F. Teppe, and W. Knap, “Terahertz responsivity of field effect transistors versus their static channel conductivity and loading effects,” Journal of Applied Physics 110, 054512 (2011).
S18. A. Sanchez, C. F. Davis Jr, K. C. Liu, and A. Javan, “The MOM tunneling diode: Theoretical estimate of its performance at microwave and infrared frequencies,” Journal of Applied Physics 49, 5270–5277 (1978).
S19. E. V. Castro, K. S. Novoselov, S. V. Morozov, N. M. R. Peres, J. M. B. Lopes Dos Santos, J. Nilsson, F. Guinea, A. K. Geim, and A. H. Castro Neto, “Electronic properties of a biased graphene bilayer,” Journal of Physics: Condensed Matter 22, 175503 (2010).
S20. E. McCann and M. Koshino, “The electronic properties of bilayer graphene,” Reports on Progress in Physics 76, 056503 (2013).
S21. A. Laturia, M. L. Van de Put, and W. G. Vandenberghe, “Dielectric properties of hexagonal boron nitride and transition metal dichalcogenides: from monolayer to bulk,” npj 2D Materials and Applications 2, 1–7 (2018).
S22. J. C. Maxwell, A treatise on electricity and magnetism, Vol. 1 (Clarendon press, 1873) pp. 246–248, art. 202.
S23. R. Nandkishore and L. Levitov, “Common-path interference and oscillatory Zener tunneling in bilayer graphene p-n junctions,” Proceedings of the National Academy of Sciences 108, 14021–14025 (2011).
S24. G. Alymov, V. Vyurkov, V. Ryzhii, and D. Svintsov, “Abrupt current switching in graphene bilayer tunnel transistors enabled by van Hove singularities,” Scientific Reports 6, 24654 (2016).
S25. S. Datta, Electronic transport in mesoscopic systems (Cambridge university press, 1997) pp. 64–65.
S26. Strictly speaking, a proper averaging of $D(E, k_z)$ is required in the two-dimensional case. This is not an easy task, especially because we need not only linearity, but also quadratic response, and the result will depend on the precise form of scattering that drives the carrier distributions toward equilibrium. This would anyway exceed the accuracy of our model, so we resort to this simple substitution. Its only role in our model is to provide smooth interpolation between the gapless case (unit transparency, there is essentially no tunnel junction, its conductance should become infinite), and the gapped case (small transparency, ($S24$) is applicable).
S27. Our use of the long-channel approximation is motivated by the fact that the experimentally obtained plasmon resonances in responsivity are not as prominent as they are in the finite-channel theory. This suggests there are additional plasmon damping mechanisms not taken into account in our model, like interband absorption or effects of electron viscosity. Instead of complicating the model by considering them explicitly, we take into account the extra plasmon damping approximately.
not by using a shorter decay length, but by setting the channel length to infinity.

S28. T. Uwanno, T. Taniguchi, K. Watanabe, and K. Nagashio, “Electrically inert h-BN/bilayer graphene interface in all-two-dimensional heterostructure field effect transistors,” *ACS Applied Materials & Interfaces* **10**, 28780–28788 (2018).