Field-Effect Tunneling Transistor Based on Vertical Graphene Heterostructures

L. Britnell,[1] R. V. Gorbachev,[2] R. Jalil,[2] B. D. Belle,[2] F. Schedin,[2] A. Mishchenko,[1]
T. Georgiou,[1] M. I. Katsnelson,[3] L. Eaves,[4] S. V. Morozov,[5] N. M. R. Peres,[6,7] J. Leist,[8]
A. K. Geim,[1,2]* K. S. Novoselov,[1]* L. A. Ponomarenko[1]*

An obstacle to the use of graphene as an alternative to silicon electronics has been the absence of an energy gap between its conduction and valence bands, which makes it difficult to achieve low power dissipation in the OFF state. We report a bipolar field-effect transistor that exploits the low density of states in graphene and its one-atomic-layer thickness. Our prototype devices are graphene heterostructures with atomically thin boron nitride or molybdenum disulfide acting as a vertical transport barrier. They exhibit room-temperature switching ratios of ≈50 and ≈10,000, respectively. Such devices have potential for high-frequency operation and large-scale integration.

The performance of graphene-based field effect transistors (FETs) has been hampered by graphene's metallic conductivity at the neutrality point (NP) and the unimpeded electron transport through potential barriers caused by Klein tunneling, which limit the achie- vable ON-OFF switching ratios to ~10^3 and those achieved so far at room temperature to <10 (*1–7*). These low ratios are sufficient for individual high-frequency transistors and analog electronics (*4–7*), but they present a fundamental problem for any realistic prospect of graphene-based integrated circuits (*1–7*). A possible solution is to open a band gap in graphene—for example, by using bilayer

[1]School of Physics and Astronomy, University of Manchester, Manchester M13 9PL, UK. [2]Manchester Centre for Meso-science and Nanotechnology, University of Manchester, Manchester M13 9PL, UK. [3]Institute for Molecules and Materials, Radboud University of Nijmegen, 6525 AJ Nijmegen, Nether-lands. [4]School of Physics and Astronomy, University of Notting-ham, Nottingham NG7 2RD, UK. [5]Institute for Microelectronics Technology, 142432 Chernogolovka, Russia. [6]Departamento de Física, Universidade do Minho, P-4710-057, Braga, Portugal. [7]Graphene Research Centre and Department of Physics, National University of Singapore, 2 Science Drive 3, 117542 Singapore. [8]Momentive Performance Materials, 22557 West Lunn Road, Strongsville, OH 44070, USA.

*To whom correspondence should be addressed. E-mail: geim@manchester.ac.uk (A.K.G.); kostya@manchester.ac.uk (K.S.N.); leonid.ponomarenko@manchester.ac.uk (L.A.P.)

graphene (8, 9), nanoribbons (10, 11), quantum dots (11), or chemical derivatives (12)—but it has proven difficult to achieve high ON-OFF ratios without degrading graphene's electronic quality.

We report an alternative graphene transistor architecture—namely, a field-effect transistor based on quantum tunneling (13–17) from a graphene electrode through a thin insulating barrier [in our case, hexagonal boron nitride (hBN) or molybdenum disulfide of ~1 nm thickness]. The operation of the device relies on the voltage tunability of the tunneling density of states (DOS) in graphene and of the effective height Δ of the tunnel barrier adjacent to the graphene electrode. To illustrate the proposed concept, we concentrate on graphene-hBN-graphene devices [an alternative barrier material (MoS$_2$) is discussed in (18)].

The structure and operational principle of our FET are shown in Fig. 1. For convenience of characterization, both source and drain electrodes were made from graphene layers in the multiterminal Hall bar geometry (18). This device configuration allowed us to not only measure the tunnel current-voltage curves (I-V) but also characterize the response of the graphene electrodes, thus providing additional information about the transistor operation. The core graphene-hBN-graphene structure was encapsulated in hBN so as to allow higher quality of the graphene electrodes (19, 20). To fabricate the device shown in Fig. 1A, we first used the standard cleavage technique (21) to prepare relatively thick hBN crystals on top of an oxidized Si wafer (300 nm of SiO$_2$), which acted as a gate electrode (Fig. 1 and fig. S1). The crystals served as a high-quality atomically flat substrate and a bottom encapsulation layer (19). Monolayer graphene (Gr$_B$) was then transferred onto a selected hBN crystal (20 to 50 nm thick) using a dry transfer procedure (19, 22). After deposition of metal contacts (5 nm Ti/50 nm Au) and etching to form a multiterminal Hall bar mesa, the structure was annealed at 350°C in forming gas. A few-atom-thick hBN crystal was identified (23) and transferred on top of Gr$_B$ by using the same procedures. This hBN layer served as the tunnel barrier. The whole process of positioning, annealing, and defining a Hall bar was repeated to make the second (top) graphene electrode (Gr$_T$). Last, a thick hBN crystal encapsulated the entire multilayer structure (Fig. 1A and fig. S1). Further details of our multistep fabrication procedures can be found in (18, 22). We tested devices with tunnel barriers having thickness d from 1 to 30 hBN layers (18). To illustrate the basic principle of the tunneling FETs, we focus on the data obtained from four devices with a tunnel barrier made of 4 to 7 layers and discuss the changes observed for other d.

When a gate voltage V_g was applied between the Si substrate and the bottom graphene layer (Gr$_B$), the carrier concentrations n_B and n_T in both bottom and top electrodes increased because of the weak screening by monolayer graphene (24), as shown schematically in Fig. 1C. The increase of the Fermi energy E_F in the graphene

layers could lead to a reduction in Δ for electrons tunneling predominantly at this energy (18). In addition, the effective height also decreased relative to the NP because the electric field penetrating through Gr$_B$ altered the shape of the barrier (25, 26). Furthermore, the increase in the tunneling DoS as E_F moved away from the NP (24) led to an increase in the tunnel current I. Depending on parameters, any of the above three contributions could dominate changes in I with varying V_g. We emphasize that the use of graphene in this device architecture is critical

because this exploits graphene's low DOS, which for a given change in V_g led to a much greater increase in E_F as compared with a conventional two-dimensional gas with parabolic dispersion (13–17). This difference translated into much greater changes of both Δ and tunneling DOS.

The behavior of in-plane resistivity ρ for the Gr$_B$ and Gr$_T$ layers as a function of V_g is shown in Fig. 2A. The curves indicate little residual doping for encapsulated graphene (≈ 0 and $<10^{11}$ cm^{-2} for Gr$_B$ and Gr$_T$, respectively). In both layers, ρ strongly depended on V_g, showing that Gr$_B$ did not screen out the electric field induced by the Si-gate electrode. The screening efficiency was quantified by Hall effect measurements (Fig. 2, B to D), which showed that the gate induced approximately the same amount of charge in both layers at low concentrations—that is, there was little screening if n_B was small. As the concentration in Gr$_B$ increased, the $n_B(V_g)$ and $n_T(V_g)$ dependences became super- and sublinear, respectively (Fig. 2, B and C), because of the increase in n_B, which led to an increasingly greater
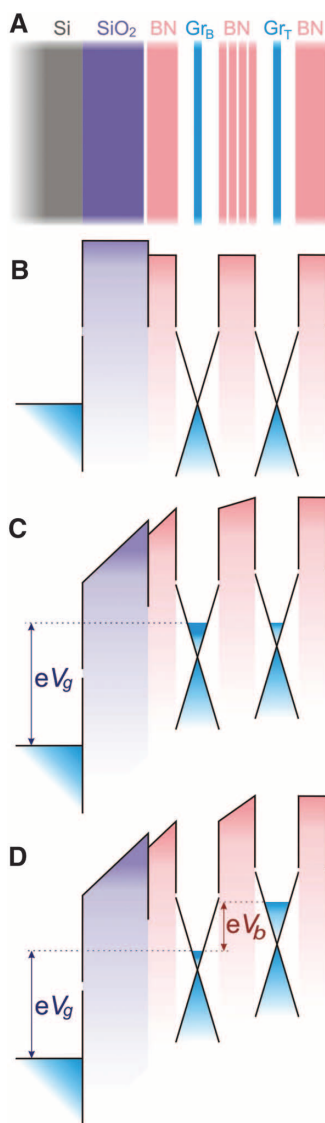
Fig. 1. Graphene field-effect tunneling transistor. (**A**) Schematic structure of our experimental devices. In the most basic version of the FET, only one graphene electrode (Gr$_B$) is essential, and the outside electrode can be made from a metal. (**B**) The corresponding band structure with no gate voltage applied. (**C**) The same band structure for a finite gate voltage V_g and zero bias V_b. (**D**) Both V_g and V_b are finite. The cones illustrate graphene's Dirac-like spectrum and, for simplicity, we consider the tunnel barrier for electrons.
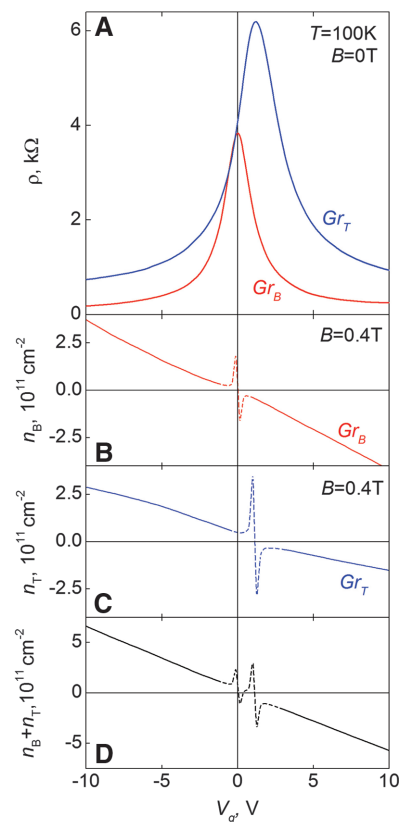
Fig. 2. Graphene as a tunneling electrode. (**A**) Resistivities of the source and drain graphene layers as a function of V_g. (**B** to **D**) Carrier concentrations in the two layers induced by gate voltage, which were calculated from the measured Hall resistivities ρ_{xy} by using the standard expression $n = B/e\rho_{xy}$, where B is the magnetic field and e is the electron charge. Close to the NP, the spikes appear (shown by dotted curves) because the above expression is not valid in the inhomogeneous regime of electron-hole puddles. The shown device has a 4-layer hBN barrier.

fraction of the gate-induced electric field being screened out by Gr_B (*18*). Hence, more electrons accumulated in the bottom graphene electrode, and fewer reached the top electrode. The total charge accumulated in both layers varied linearly in V_g (Fig. 2D), as expected. We could describe the observed redistribution of the charge between the two graphene layers in terms of the corresponding sequential circuit including the quantum capacitance (*13, 27*) of the graphene layers (fig. S2). For a parabolic band, the ratio between n_B and n_t would be independent on V_g, and therefore, the electric field penetrating into the tunnel barrier would be substantially reduced even in the limit of zero n_B (*13*).
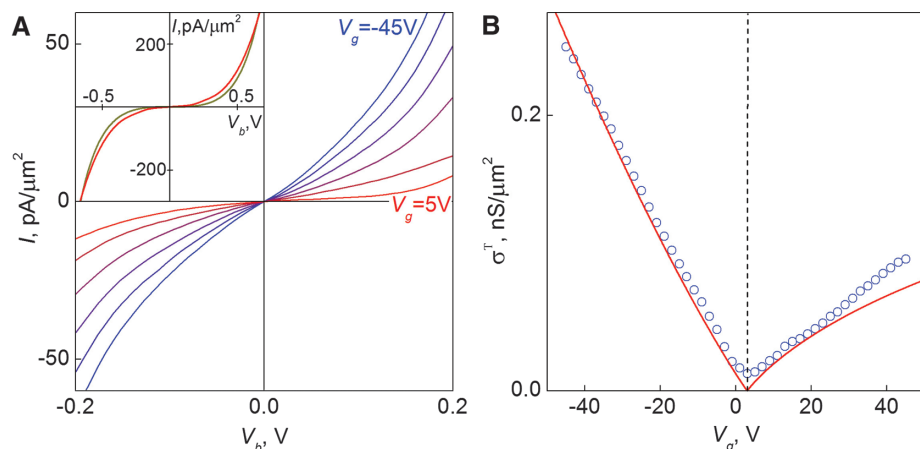
A bias voltage V_b applied between Gr_B and Gr_T gave rise to a tunnel current through the thin hBN barrier that scaled with device area. I-V characteristics for one of our devices at various V_g are shown in Fig. 3A. First, we consider the case of zero V_g. At low V_b, I was linear in bias, yielding a tunnel resistivity $\rho^T = V_b/I \approx 100$ gigohms μm^2 for this hBN thickness. At higher voltages (V_b above ~0.1 V), I grew more rapidly. The I-V curves could be described (Fig. 3A, inset, and fig. S3) by the standard quantum-tunneling formulae (*25, 26*), assuming energy conservation but no momentum conservation at the mismatched graphene-hBN interface (*28*).

As shown below, we could distinguish experimentally between electron and hole tunneling and found that the tunneling was due to holes. This result is in agreement with a recent theory for the graphene-hBN interface (*29*), which reports a separation between the Dirac point in graphene and the top of the hBN valance band of ≈1.5 eV, whereas the conduction band is >4 eV away from

the Dirac point. The fit to our data with Δ = 1.5 eV yielded a tunneling mass $m \approx 0.5 m_0$ (m_0 is the free electron mass), which is in agreement with the effective mass for holes in hBN (*30*). Furthermore, our analysis indicated that I varied mainly with the change in the tunneling DOS, whereas the change in tunneling probability with applied bias was a secondary (albeit important) effect (*18*). For our atomically thin hBN barriers with relatively low ρ^T, we were not in a regime of exponential sensitivity to changes in $\Delta[E_F(V_b)]$.

We demonstrate transistor operation in Fig. 3A, which plots the influence of gate voltage on I. V_g substantially enhanced the tunnel current, and the changes were strongest at low bias. The field effect was rather gradual for all gate voltages up to ±50 V, a limit set by the electrical breakdown of our SiO_2 gate dielectric (typically ≈60 V). This response is quantified in Fig. 3B, which plots the low-bias tunneling conductivity $\sigma^T = I/V_b$ as a function of V_g. The influence of V_g was highly asymmetric: σ^T changed by a factor of ≈20 for negative V_g (holes) and by a factor of 6 for positive V_g (electrons). We observed changes up to ≈50 for hole tunneling in other devices and always the same asymmetry (fig. S4) (*18*). Also, the ON-OFF ratios showed little change between room and liquid-helium temperatures, as expected for Δ >> thermal energy.

To analyze the observed behavior of $\sigma^T(V_g)$, we modeled the zero-bias conductivity by using the relation $\sigma^T \propto DOS_B(V_g) \times DOS_T(V_g) \times T(V_g)$, where the indices refer to the two graphene layers and $T(V_g)$ is the transmission coefficient through the hBN barrier (*25, 26*). The resulting curve shown in Fig. 3B accounts qualitatively for the main features in the measured data, using

self-consistently the same tunneling parameters m and Δ given above. At V_g near zero, corresponding to tunneling from states near the NP, the tunneling DOS in both graphene layers was small and nonzero and was the result of residual doping, disorder, and temperature effects (*18*). The application of a gate voltage of either polarity led to a higher DOS and, therefore, higher σ^T. The gradual increase in $\sigma^T(V_g)$ for both polarities in Fig. 3B was therefore caused by the increasing DOS. However, V_g also affected the transmission coefficient. Because of the shift of E_F with changing V_g, the effective barrier height Δ decreased for one sign of charge carriers and increased for the other (Fig. 1B), which explains the asymmetry in both experimental and calculated $\sigma^T(V_g)$ in Fig. 3B in terms of the change in $T(V_g)$. For our devices, the effect of V_g on $T(V_g)$ was relatively weak (nonexponential) and comparable with the effect caused by changes in the tunneling DOS. The sign of the asymmetry infers that the hBN barrier height was lower for holes than for electrons, which is in agreement with the graphene-hBN band structure calculations (*29*). The weaker dependence of I on V_g at high bias can also be understood in terms of the more gradual increase in the tunneling DOS and in E_F at high doping ($V_b = 0.5V$ correspond to $n_B \approx 10^{13}$ cm^{-2}).

Our results and analysis suggest that higher ON-OFF ratios could be achieved by using either higher V_g or making devices with larger d, so that the tunneling depends exponentially on bias and is controlled by the barrier height rather than the DOS. The former route is limited by the electrical breakdown of dielectrics at ~1 V/nm ($V_g \approx 300V$ for our SiO_2 thickness). By extrapolating the analysis shown in Fig. 3B to such voltages, we found that ON-OFF ratios >10^4 would be possible for our 4-to-7-layer devices if SiO_2 of highest quality were used. However, it would still require unrealistically large V_g to enter the regime where E_F becomes comparable with Δ and changes in $\sigma^T(V_g)$ are exponentially fast. Therefore, we explored the alternative option and investigated devices with both thinner and thicker hBN barriers. For 1- to 3-hBN layers, zero-bias σ^T increased exponentially with decreasing number of layers, which is consistent with quantum tunneling, and we observed a weaker influence of V_g on I, as expected for the more conductive regime. On the other hand, the thicker hBN barriers were prone to electrical breakdown. Nonetheless, for a few devices with d ≈ 6 to 9 nm, we were able to measure a tunnel current without breakdown. A current >10 pA appeared at biases of several volts and increased exponentially with V_b. The thicker devices' I-V characteristics could be fitted by using the same hole-tunneling parameters used above, thus indicating quantum tunneling rather than an onset of electrical breakdown. Unfortunately, no substantial changes (exceeding 50%) in the tunnel current could be induced by V_g. This insensitivity to gate voltage remains to be understood but was

Fig. 3. Tunneling characteristics for a graphene-hBN device with 6 ± 1 layers of hBN as the tunnel barrier. (**A**) I-Vs for different V_g (in 10-V steps). Because of finite doping, the minimum tunneling conductivity is achieved at $V_g \approx 3V$. The inset compares the experimental I-V at $V_g = 5$ V (red curve) with theory (dark), which takes into account the linear DOS in the two graphene layers and assumes no momentum conservation. Further examples of experimental curves and their fitting can be found in (*18*). (**B**) Zero-bias conductivity as a function of V_g. The symbols are experimental data, and the solid curve is our modeling. The curve is slightly shifted with respect to zero V_g because of remnant chemical doping. In all the calculations, we assumed the hole tunneling with $m = 0.5 m_0$ and $\Delta \approx 1.5$ eV (*29, 30*) and used d as measured by atomic force microscopy. Both I and σ are normalized per tunnel area, which was typically 10 to 100 μm^2 for the studied devices. Temperature, 240 K.

probably caused by charge traps that screened the influence of the gate.

An alternative method to achieve an exponential dependence of the tunneling current on gate voltage would be to use a barrier dielectric with a smaller Δ, which would be comparable with typical E_F realizable in graphene. One of such candidate materials is MoS_2, which has a band gap of about 1.3 eV and can be obtained in a mono- or few-layers state similar to hBN and graphene (21). Our first graphene-MoS_2–based devices demonstrate ON-OFF ratio close to 10.000 (fig. S5), which is sufficient for certain types of logic circuits.

We conclude that our tunneling devices offer a viable route for high-speed graphene-based analog electronics. The ON-OFF ratios already exceed those demonstrated for planar graphene FETs at room temperature by a factor of 10 (3–7). The transit time for the tunneling electrons through the nanometer-thick barriers is expected to be extremely fast (a few femtoseconds) (13–17) and exceeds the electron transit time in submicrometer planar FETs. It should also be possible to decrease the lateral size of the tunneling FETs down to the 10 nm scale, a requirement for integrated circuits. Furthermore, there appears to be no fundamental limitation

to further enhancement of the ON-OFF ratios by optimizing the architecture and by using higher-quality gate dielectrics and, in particular, lower tunnel barriers (Δ < maximum achievable E_F).

References and Notes

1. P. Avouris, Z. H. Chen, V. Perebeinos, *Nat. Nanotechnol.* **2**, 605 (2007).
2. A. K. Geim, *Science* **324**, 1530 (2009).
3. F. Schweirz, *Nat. Nanotechnol.* **5**, 487 (2010).
4. Y. Wu *et al.*, *Nature* **472**, 74 (2011).
5. L. Liao *et al.*, *Nature* **467**, 305 (2010).
6. S. J. Han *et al.*, *Nano Lett.* **11**, 3690 (2011).
7. Y. M. Lin *et al.*, *Science* **332**, 1294 (2011).
8. E. V. Castro *et al.*, *Phys. Rev. Lett.* **99**, 216802 (2007).
9. J. B. Oostinga, H. B. Heersche, X. Liu, A. F. Morpurgo, L. M. K. Vandersypen, *Nat. Mater.* **7**, 151 (2008).
10. M. Y. Han, B. Ozyilmaz, Y. B. Zhang, P. Kim, *Phys. Rev. Lett.* **98**, 206805 (2007).
11. C. Stampfer *et al.*, *Front. Phys.* **6**, 271 (2011).
12. D. C. Elias *et al.*, *Science* **323**, 610 (2009).
13. S. Luryi, *Appl. Phys. Lett.* **52**, 501 (1988).
14. M. Heiblum, M. V. Fischetti, *IBM J. Res. Develop.* **34**, 530 (1990).
15. J. A. Simmons *et al.*, *J. Appl. Phys.* **84**, 5626 (1998).
16. A. Zaslavsky *et al.*, *Appl. Phys. Lett.* **83**, 1653 (2003).
17. A. Sciambi *et al.*, *Phys. Rev. B* **84**, 085301 (2011).
18. Materials and methods are available as supporting material on *Science* Online.
19. C. R. Dean *et al.*, *Nat. Nanotechnol.* **5**, 722 (2010).
20. A. S. Mayorov *et al.*, *Nano Lett.* **11**, 2396 (2011).
21. K. S. Novoselov *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10451 (2005).
22. L. A. Ponomarenko *et al.*, *Nat. Phys.* **7**, 958 (2011).
23. R. V. Gorbachev *et al.*, *Small* **7**, 465 (2011).
24. A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, A. K. Geim, *Rev. Mod. Phys.* **81**, 109 (2009).
25. J. G. Simmons, *J. Appl. Phys.* **34**, 1793 (1963).
26. E. L. Wolf, *Principles of Electron Tunneling Spectroscopy* (Oxford Univ. Press, Oxford, 1985).
27. L. A. Ponomarenko *et al.*, *Phys. Rev. Lett.* **105**, 136801 (2010).
28. J. M. Xue *et al.*, *Nat. Mater.* **10**, 282 (2011).
29. N. Kharche, S. K. Nayak, *Nano Lett.* **11**, 5274 (2011).
30. Y. N. Xu, W. Y. Ching, *Phys. Rev. B* **44**, 7787 (1991).

Supporting Online Material
www.sciencemag.org/cgi/content/full/science.1218461/DC1
Materials and Methods
SOM Text
Figs. S1 to S5
References (31–35)

Field-effect tunneling transistor based on vertical graphene heterostructures

L. Britnell *et al*

#1 *Experimental structures*

Our devices contain two graphene Hall bars placed on top of each other with a thin layer of hBN in between. Figure S1 shows one of the studied devices. The turquoise area in Fig. S1A is a thick hBN crystal on top of an oxidized Si wafer (brown-purple). The hBN layer served as a substrate to ensure the quality of the bottom graphene electrode. The actual graphene-hBN-graphene-hBN sandwich is highly transparent and practically invisible on this image taken in an optical microscope (Fig. S1A). Nonetheless, one may discern a mesa structure in the central area between the Au leads. The multilayer Hall bar geometry is illustrated in Fig. S1B. This is an electron micrograph of the same device but before depositing Au contacts. The colored image of various layers was used at a design stage for the last round of electron-beam lithography. The Au leads (deposited later) are shown in violet, and two graphene mesas in orange and green. The hBN crystal used as the tunnel barrier can be seen as a light grey patch of irregular shape. Its thickness was determined using atomic force microscopy, Raman microscopy and optical contrast (*26*).
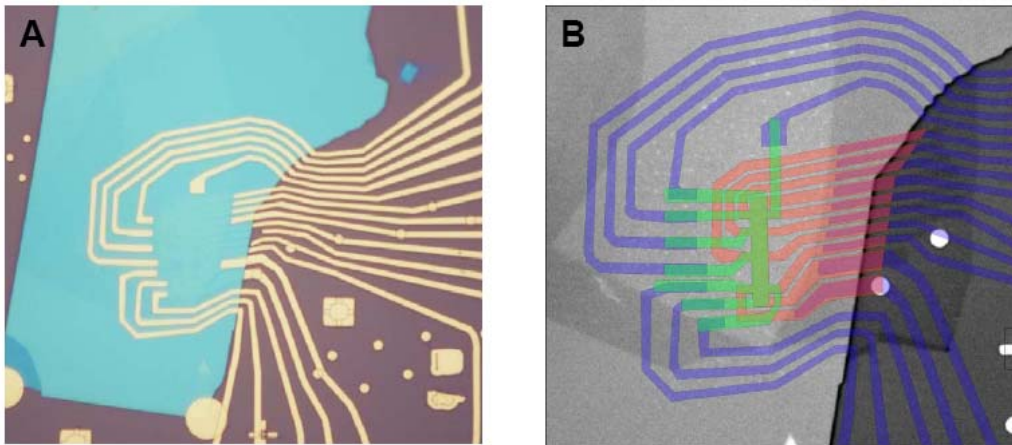

Figure S1. One of our hBN-graphene-hBN-graphene-hBN devices. (A) Optical image of the final device. (B) Electron micrograph of the same device at the final design stage before evaporating Au leads. Two 10-terminal Hall bars made from graphene are shown in green and orange. The spatial scale is given by the width of the Hall bar, which was 2 µm for this device. Fabrication required 4 dry transfers and alignments of the graphene and hBN crystals, 4 nonconsecutive rounds of electron-beam lithography, 3 rounds of plasma etching and two separate metal depositions.

#2 *Penetration of electric field through the graphene electrode*

Consider the geometry shown in Fig. 1A of the main text. The external electric field between the Si and bottom graphene electrodes, which are separated by distance D, is $F_g = V_g/D$ (dielectric constants for both SiO$_2$ and hBN are similar and, for simplicity, we assume them both equal to ε). The electric field F_b between Gr$_B$ and Gr$_T$ and the induced carrier densities in the graphene plates n_T and n_B are related by the equations

$$\varepsilon(F_b - F_g) = 4\pi n_B e$$

$$-\varepsilon F_b = 4\pi n_T e$$

A bias voltage V_b between the two graphene electrodes is given by

$$eV_b = eF_b d - \mu(n_T) + \mu(n_B)$$

where d is the hBN thickness and $\mu(n)$ are the chemical potentials in the corresponding graphene layers. For simplicity, we assume that graphene electrodes are chemically undoped and, therefore, $n_T = n_B = 0$ in the absence of applied voltages.

Taking into account the electron-hole symmetry $\mu(-n) = -\mu(n)$, we obtain the following equation

$$\frac{4\pi e^2 d}{\varepsilon} n_T + \mu(n_T) + \mu\left(n_T + \frac{\varepsilon F_g}{4\pi e} \right) + eV_b = 0 \tag{S1}$$

which allows us to determine n_T induced by the field effect in Gr$_T$ for a given V_g. For a conventional two-dimensional (2D) electron gas, $\mu(n) \propto n$ and the first term in eq. (S1), which describes the classical capacitance of the tunnel barrier, is dominant for any realistic d, larger than interatomic distances. In graphene with its low DoS and Dirac-like spectrum, $\mu(n) \propto \sqrt{n}$ and this leads to a qualitatively different behavior, which can be described in terms of quantum capacitance (*27*) (also note the discussion of doping of graphene through an hBN spacer in ref. *31*).

The above expressions were employed to find n_T and n_B as a function of bias V_b and gate voltage V_g and the results were then used to model the I-V characteristics (see the theory curves in Fig. 3 of the main text). To illustrate the agreement between the experiment and theory at the intermediate stage of determining n_T and n_B, Figure S2 shows the same experimental data for carrier concentrations in the top and bottom graphene layer $n(V_g)$ as in Fig. 2B,C and compares them with the behavior expected from solving eq. (S1).
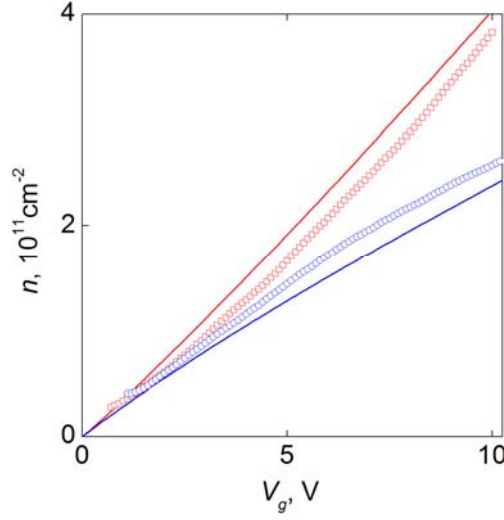
Figure S2. Nonlinear dependence of charge carrier concentrations in the two graphene electrodes as a function of gate voltage. The symbols are experimental data (red symbols for the bottom graphene layer; blue for the top). The solid curves in the corresponding colors are our modeling. No fitting parameters are used.

#3 *Modeling of our device operation*

I-V curves for a tunnel junction are generally described by (*23*)

$$I(V) \propto \int dE DoS_{\mathrm{B}}(E) DoS_{\mathrm{T}}(E - eV) T(E)[f(E - eV) - f(E)] \tag{S2}$$

where $f(E)$ is the Fermi distribution function. At low temperatures the difference of the Fermi functions restricts the relevant energy E integral to $\mu < E < \mu + eV$ where μ is the chemical potential and, to be specific, we consider the case $eV > 0$. The above formula assumes that there is no in-plane momentum conservation, which is most likely to be the case of realistic graphene-hBN interfaces. There are several possible mechanisms for elastic scattering at the interface and, in particular, unavoidable fluctuations of the mass term due to the lattice mismatch (*32*). Note that elastic tunneling is forbidden between two 2D systems if in-plane momentum is conserved.

If the tunneling conductance per channel is much smaller than the conductivity quantum e^2/h (as in our case) the transmission probability T is exponentially small and depends strongly on the energy E of tunneling electrons,

$$T(E) = A(E) \exp[-W(E)] \tag{S3}$$

where *A* is a smooth function that depends on details of the wave-function matching at the interface. In our modeling, we assume *A=const.*

Let us now discuss some functional forms of *W(E)*. For the case of an isotropic barrier, we need to solve the dispersion equation $E = \varepsilon_n(k_x, k_y, k_z)$ for each band of the barrier material, where *E* is the energy of electrons tunneling in the *z* direction. No real solution for k_z is possible inside the energy gap, and the minimal Imk_z $\mathrm{Im}\,k_z$ for a given *E* and arbitrary k_x and k_y, which dominates the tunneling probability, is given by

$$W(E) = 2d\,\mathrm{Im}\,k_z$$

For the case of parabolic bands, $\mathrm{Im}\,k_z = \dfrac{\sqrt{2m\Delta}}{\hbar}$ where Δ is the barrier height (in our case, the distance to the valence band) and *m* is the effective mass (*22,23,33*).

In the case of layered crystals, their band structure can be described in the simplest approximation as

$$\varepsilon(k_x, k_y, k_z) = \tau(k_z) + \varepsilon_1(k_x, k_y) \tag{S4}$$

where $\tau(k_z) = 2t_\perp \cos(k_z l)$; $t_\perp$ describes the interlayer coupling and *l* is the interlayer distance (for the case of hBN, $l \approx 3.4\text{Å}$). By solving the corresponding tunneling equation, we find k_z within the gap to be

$$k_z = \frac{i}{l} \ln\left(\left| \frac{E - \varepsilon_1}{2t_\perp} \right| + \sqrt{\left(\frac{E - \varepsilon_1}{2t_\perp} \right)^2 - 1} \right)$$

The top of the valence band corresponds to $E_{\max} = \max \varepsilon_1(k_x, k_y) + 2t_\perp$ (to be specific, we choose $t_\perp > 0$), and the optimal value for the tunneling wavevector is then

$$\mathrm{Im}\,k_z = \frac{1}{l} \ln\left(\left(\frac{\Delta}{2t_\perp} + 1 \right) + \sqrt{\left(\frac{\Delta}{2t_\perp} + 1 \right)^2 - 1} \right) \tag{S5}$$

where $\Delta = E - E_{\max}$. If $\Delta \gg 2t_\perp$, this expression can be simplified as $k_z = \dfrac{i}{l} \ln\left(\dfrac{\Delta}{t_\perp} \right)$ and yields the tunneling probability $T \propto (t_\perp / \Delta)^{2n}$ where *n* =*d/l* is the number of atomic layers in the tunnel barrier. In the opposite limit $\Delta \ll 2t_\perp$, we obtain $k_z = \dfrac{i}{l} \sqrt{\dfrac{\Delta}{t_\perp}} = i \dfrac{\sqrt{2m^*\Delta}}{\hbar}$ where $m^* = \dfrac{\hbar^2}{2t_\perp l^2}$ is the effective mass in the tunneling direction. This shows that the standard isotropic model is applicable to layered crystals, provided the tunneling occurs not too far from the band-gap edge.

Eq. (S4) is a simplified version of the real band structure of hBN, which depends on stacking order. hBN crystals usually have AA′ stacking (*34*). In the next approximation that allows an analytical solution by neglecting the mixing of π and σ bands (*29,30*), we obtain the following dispersion relation [S4]

$$\varepsilon^2\left(k_x, k_y, k_z\right) = \frac{E_g^2}{4} + \tau^2\left(k_z\right) + \varepsilon_1^2\left(k_x, k_y\right) \pm 2\tau\left(k_z\right)\varepsilon_1\left(k_x, k_y\right) \tag{S6}$$

where E_g is the energy difference between boron and nitrogen sites (*34*). In this case, we find

$$\mathrm{Im}\, k_z = \frac{1}{l}\ln\left(\frac{\Phi}{2t_\perp} + \sqrt{\left(\frac{\Phi}{2t_\perp}\right)^2 - 1}\right) \tag{S7}$$

where $\Phi = \sqrt{E^2 - \dfrac{E_g^2}{4}} - \left|\varepsilon_1\left(k_x, k_y\right)\right|$. Eq. (S7) differs from (S5) by replacement $E \to \sqrt{E^2 - E_g^2/4}$, which indicates the general validity of equation $\mathrm{Im}\, k_z \propto \ln(\Delta)$ for describing vertical tunneling through strongly layered materials. (S5) and (S7) fit our experimental data equally well. It is worth noting that the tunneling exponent through layered crystals depends on E only weakly (logarithmically) in comparison with isotopic crystals that exhibit the standard square-root energy dependence. For small changes in Δ, this difference is unimportant (see below).

Finally, in the case of a strong electric field such that it changes the rectangular shape of the tunnel barrier (Fig. 1D) the above expressions for *W* can be generalized within the WKB approximation (*33*) as

$$W = 2\int_0^d dx\, \mathrm{Im}\, k_z\left(\Delta \to \Delta(x)\right).$$

#4 *Layered versus isotropic barrier*

In the main text, we have chosen for the sake of brevity to ignore the fact that our tunnel barriers are made from a strongly layered material. This simplification allowed us to refer to the standard tunneling theory. However, the assumption can be justified further by the fact that, for our device parameters, we have found no difference between the I-V characteristics calculated for the layered and isotropic materials and, therefore, we cannot distinguish between the two cases. To illustrate the indifference to the layered structure of our tunnel barrier, Figure S3 shows experimental I-V curves for two devices and compares them with the behavior expected for layered and isotropic cases. No major difference can be

seen, except at low bias in Fig. S3A. The exact shape of experimental curves at low bias varies from sample to sample (cf. Fig. S3A & B) and, hence, we do not discuss the difference.
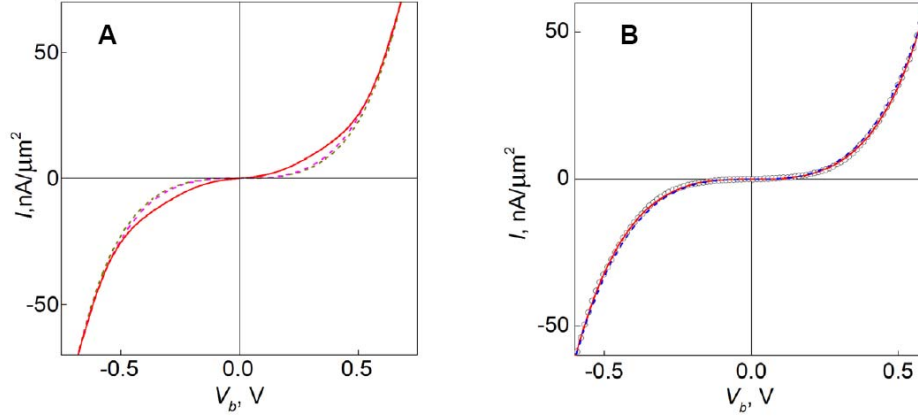

Figure S3. Tunneling I-V characteristics for two different 4-hBN-layer devices at zero gate voltage and their comparison with theory. (A) The red solid curve is the experimental data from Fig. 3. The two dashed curves are our modeling for an isotropic barrier (Δ and m as in the main text) and for a layered barrier of the same height and $t_\perp$ =0.6eV, by using formulae from the above section. Note that $t_\perp \approx$0.6eV corresponds to m =0.5m_0. (B) Nominally similar device (for clarity, the experimental data are shown by symbols). The curves are again the layered and isotropic versions of the tunneling theory. The fitting parameter is the constant A in eq. (S3), which determines the absolute value of I. The close agreement between functional forms of the theoretical curves validates the use of the conventional tunneling formulae in the main text.

#5 *Additional examples of our device operation*

We have studied 6 multiterminal devices such as shown in Fig. S1 and >10 simpler tunneling FETs with only one or two Ohmic contacts attached to each graphene electrode. The latter type does not provide much information about the properties of the graphene electrodes but even one contact is sufficient to study their tunneling I-V characteristics. The devices with the same hBN thickness have exhibited qualitatively similar behavior, as discussed in the main text. To illustrate the degree of reproducibility for different samples, Figure S4 plots the behavior observed in another device with the tunnel barrier consisting of 4 hBN layers. One can see that the nonlinear I-V characteristics are qualitatively similar to those presented in the main text, and their response to gate voltage is also similar.
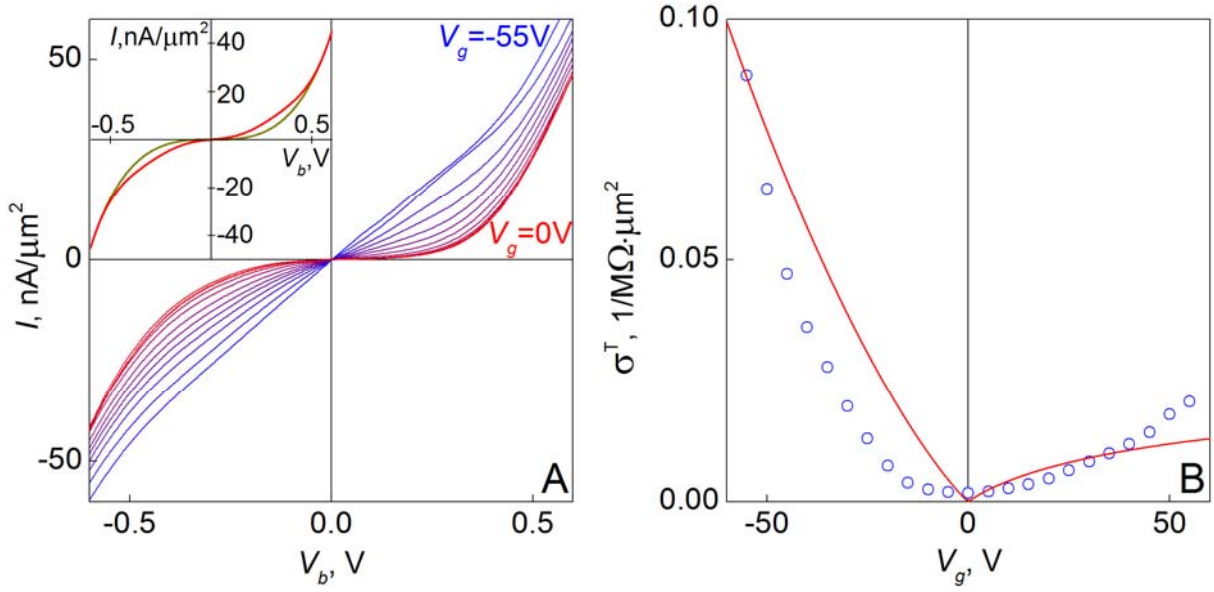
Figure S4. Another hBN-graphene-hBN-graphene-hBN field-effect device. (A) Tunneling I-Vs and their response to gate voltage (in 5V steps, cf. Fig. 3 of the main text). The inset compares the experimental I-V at zero gate voltage (red curve) with theory (dark) which takes into account the linear DoS in the two graphene layers and assumes no momentum conservation.Temperature: 300 K. (B) Changes in low-bias tunneling (symbols) and the theory fit for 4 hBN layers (solid curve). The main difference with respect to the device in the main text is a weak response at low gate voltages, which is probably due to stronger disorder and chemical doping that smears the gate influence. The electron-hole asymmetry again implies the hole tunneling as discussed in the main text.

The only consistent difference that we noticed for a number of devices with 4 or more atomic layers of hBN was the absolute value of σ^T which could vary by a factor of 100 for nominally the same d. Although this can be attributed to possible errors in determining the number of layers in thicker hBN (26), more careful analysis of the devices' response to bias and gate voltages reveals that the reason for these variations is more likely to be inhomogeneous thickness of hBN. We believe that in some devices one or two layers can be missing locally (in submicron scale patches) so that the tunnel current then concentrates within these thinner areas. Graphite is known to cleave leaving occasional stripes of smaller thickness for few-layer graphene crystals and, whereas it is possible to see missing graphene patches in an optical microscope, hBN does not allow the required resolution (26).

#6 Vertical transistors using few-layer MoS₂ as a barrier

One of the possible routes to increase the ON-OFF ratio is to use materials with smaller Δ. In this way it would be viable to use thicker barriers but shift E_F closer to the barrier edge, which should allow exponential dependence of the tunnel current on gate voltage. One of the candidate materials is MoS_2. It is a layered semiconductor, which can be cleaved down to a monolayer (*24*). It has an indirect band gap of $\approx$1.3eV (*35*), significantly lower than that in hBN. We have exploited MoS_2 to prepare graphene-MoS_2 devices by using the same procedures as described in the main text and in #1 of SOM.

I-V characteristics for a transistor with a 6-layer MoS_2 barrier are presented in Figure S5A. Measurements of its conductivity at a fixed small bias as a function of gate voltage are plotted in Fig. S5B. The dependence is clearly exponential and in this device an ON-OFF ratio of ~10,000 has been achieved. Further work is needed to improve the observed ratios further and to verify whether the mechanism responsible for vertical transport through MoS_2 is indeed tunneling.
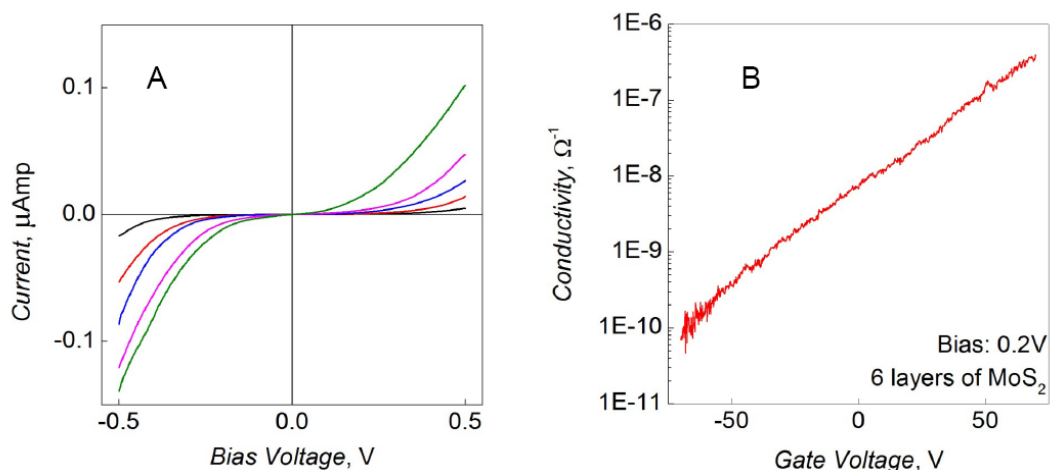

Figure S5. I-V characteristics of a graphene MoS_2 device. Thickness of MoS_2 is 6 layers. (A) - Different curves correspond to various gate voltages applied. Black: -40V; red: -20V; blue: 0V; purple: +20; green: +40V. (B) - Conductivity measured at a bias voltage of 0.2V as a function of gate voltage. The ON/OFF ratio of $>7\times10^3$ is observed even for the relatively limited range of gate voltages.

References and Notes

1. P. Avouris, Z. H. Chen, V. Perebeinos, Carbon-based electronics. *Nat. Nanotechnol.* **2**, 605 (2007). doi:10.1038/nnano.2007.300 Medline

2. A. K. Geim, Graphene: Status and prospects. *Science* **324**, 1530 (2009). doi:10.1126/science.1158877 Medline

3. F. Schweirz, Graphene transistors. *Nat. Nanotechnol.* **5**, 487 (2010). doi:10.1038/nnano.2010.89

4. Y. Wu *et al*., High-frequency, scaled graphene transistors on diamond-like carbon. *Nature* **472**, 74 (2011). doi:10.1038/nature09979 Medline

5. L. Liao *et al*., High-speed graphene transistors with a self-aligned nanowire gate. *Nature* **467**, 305 (2010). doi:10.1038/nature09405 Medline

6. S. J. Han *et al*., High-frequency graphene voltage amplifier. *Nano Lett.* **11**, 3690 (2011). doi:10.1021/nl2016637 Medline

7. Y. M. Lin *et al*., Wafer-scale graphene integrated circuit. *Science* **332**, 1294 (2011). doi:10.1126/science.1204428 Medline

8. E. V. Castro *et al*., Biased bilayer graphene: Semiconductor with a gap tunable by the electric field effect. *Phys. Rev. Lett.* **99**, 216802 (2007). doi:10.1103/PhysRevLett.99.216802 Medline

9. J. B. Oostinga, H. B. Heersche, X. Liu, A. F. Morpurgo, L. M. K. Vandersypen, Gate-induced insulating state in bilayer graphene devices. *Nat. Mater.* **7**, 151 (2008). doi:10.1038/nmat2082 Medline

10. M. Y. Han, B. Ozyilmaz, Y. B. Zhang, P. Kim, Energy band-gap engineering of graphene nanoribbons. *Phys. Rev. Lett.* **98**, 206805 (2007). doi:10.1103/PhysRevLett.98.206805 Medline

11. C. Stampfer *et al*., Transport in graphene nanostructures. *Front. Phys.* **6**, 271 (2011).

12. D. C. Elias *et al*., Control of graphene's properties by reversible hydrogenation: Evidence for graphane. *Science* **323**, 610 (2009). doi:10.1126/science.1167130 Medline

13. S. Luryi, Quantum capacitance devices. *Appl. Phys. Lett.* **52**, 501 (1988). doi:10.1063/1.99649

14. M. Heiblum, M. V. Fischetti, Ballistic hot-electron transistors. *IBM J. Res. Develop.* **34**, 530 (1990). doi:10.1147/rd.344.0530

15. J. A. Simmons *et al*., Planar quantum transistor based on 2D–2D tunneling in double quantum well heterostructures. *J. Appl. Phys.* **84**, 5626 (1998). doi:10.1063/1.368610

16. A. Zaslavsky *et al*., Ultrathin silicon-on-insulator vertical tunneling transistor. *Appl. Phys. Lett.* **83**, 1653 (2003). doi:10.1063/1.1600832

17. A. Sciambi *et al*., Vertical field-effect transistor based on wave-function extension. *Phys. Rev. B* **84**, 085301 (2011). doi:10.1103/PhysRevB.84.085301

18. Materials and methods are available as supporting material on *Science* Online.

19. C. R. Dean *et al*., Boron nitride substrates for high-quality graphene electronics. *Nat. Nanotechnol.* **5**, 722 (2010). doi:10.1038/nnano.2010.172 Medline

20. A. S. Mayorov *et al*., Micrometer-scale ballistic transport in encapsulated graphene at room temperature. *Nano Lett.* **11**, 2396 (2011). doi:10.1021/nl200758b Medline

21. K. S. Novoselov *et al*., Two-dimensional atomic crystals. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10451 (2005). doi:10.1073/pnas.0502848102 Medline

22. L. A. Ponomarenko *et al*., Tunable metal–insulator transition in double-layer graphene heterostructures. *Nat. Phys.* **7**, 958 (2011). doi:10.1038/nphys2114

23. R. V. Gorbachev *et al*., Hunting for monolayer boron nitride: Optical and Raman signatures. *Small* **7**, 465 (2011). doi:10.1002/smll.201001628 Medline

24. A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, A. K. Geim, The electronic properties of graphene. *Rev. Mod. Phys.* **81**, 109 (2009). doi:10.1103/RevModPhys.81.109

25. J. G. Simmons, Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film. *J. Appl. Phys.* **34**, 1793 (1963). doi:10.1063/1.1702682

26. E. L. Wolf, Principles of electron tunneling spectroscopy. *Oxford University Press* (1985).

27. L. A. Ponomarenko *et al*., Density of states and zero Landau Level probed through capacitance of graphene. *Phys. Rev. Lett.* **105**, 136801 (2010). doi:10.1103/PhysRevLett.105.136801 Medline

28. J. M. Xue *et al*., Scanning tunnelling microscopy and spectroscopy of ultra-flat graphene on hexagonal boron nitride. *Nat. Mater.* **10**, 282 (2011). doi:10.1038/nmat2968 Medline

29. N. Kharche, S. K. Nayak, Quasiparticle band gap engineering of graphene and graphone on hexagonal boron nitride substrate. *Nano Lett.* **11**, 5274 (2011). doi:10.1021/nl202725w Medline

30. Y. N. Xu, W. Y. Ching, Calculation of ground-state and optical properties of boron nitrides in the hexagonal, cubic, and wurtzite structures. *Phys. Rev. B* **44**, 7787 (1991). doi:10.1103/PhysRevB.44.7787 Medline

31. M. Bokdam, P. A. Khomyakov, G. Brocks, Z. Zhong, P. J. Kelly, Electrostatic doping of graphene through ultrathin hexagonal boron nitride films. *Nano Lett.* **11**, 4631 (2011). doi:10.1021/nl202131q Medline

32. B. Sachs, T. O. Wehling, M. I. Katsnelson, A. I. Lichtenstein, Adhesion and electronic structure of graphene on hexagonal boron nitride substrates. *Phys. Rev. B* **84**, 195414 (2011). doi:10.1103/PhysRevB.84.195414

33. L. D. Landau, E. M. Lifshitz, *Quantum Mechanics*. Oxford: Pergamon (1977).

34. R. M. Ribeiro, N. M. R. Peres, Stability of boron nitride bilayers: Ground-state energies, interlayer distances, and tight-binding description. *Phys. Rev. B* **83**, 235312 (2011). doi:10.1103/PhysRevB.83.235312

35. Gmelin, Gmelin handbook of inorganic and organometallic chemistry (SpringerVerlag, Berlin, 1995).