nature electronics

ARTICLES

https://doi.org/10.1038/s41928-019-0364-5

A transverse tunnelling field-effect transistor made from a van der Waals heterostructure

Xiong Xiong[1,5], Mingqiang Huang[1,4,5], Ben Hu[1], Xuefei Li[1], Fei Liu[2], Sichao Li[1], Mengchuan Tian[1], Tiaoyang Li[1], Jian Song[1] and Yanqing Wu[1,2,3*]

Semiconductor devices that rely on quantum tunnelling could be of use in logic, memory and radiofrequency applications. Tunnel devices that exhibit negative differential resistance typically follow an operating principle in which the tunnelling current contributes directly to the drive current. Here, we report a tunnelling field-effect transistor made from a black phosphorus/Al_2O_3/black phosphorus van der Waals heterostructure in which the tunnelling current is in the transverse direction with respect to the drive current. Through an electrostatic effect, this tunnelling current can induce a drastic change in the output current, leading to a tunable negative differential resistance with a peak-to-valley ratio of more than 100 at room temperature. Our device also exhibits abrupt switching, with a body factor (the relative change in gate voltage with respect to that of the surface potential) that is one-tenth of the Boltzmann limit for conventional transistors across a wide temperature range.

Devices that rely on a band-to-band tunnelling (BTBT) mechanism and exhibit negative differential resistance (NDR) are of potential use in radiofrequency oscillators and multi-value logic applications[1–6]. In such tunnel devices, the carriers tunnel through a barrier in the same direction as the total output current, and the tunnel current contributes directly to the overall current. High-quality interfaces and sharp energy band edges are desirable for the development of tunnel devices[7,8], and two-dimensional (2D) van der Waals heterostructures can thus provide unique opportunities[9–13]. In these structures, different 2D materials and thin insulators can be stacked on top of each other, without the lattice mismatch constraints typical of conventional heterostructures, due to the weak van der Waals interaction between layers. Moreover, the energy band at the interfaces in the structure can change abruptly with a sharp band edge, forming high-quality quantum wells or barriers with thickness down to a monolayer. The availability of numerous different 2D materials—with a variety of different band structures, from semi-metals to semiconductors to insulators—also makes it possible to assemble unique materials with well-designed band alignments[7,8,14,15]. A range of NDR and steep switching devices have been reported, including graphene/boron nitride (BN) resonant tunnel diodes[16–18], molecular electronic devices[19,20], amorphous-carbon quantum-well devices[21], transition metal dichalcogenide-based tunnelling devices[22–25] and hybrid 2D/3D tunnel devices based on MoS_2/Ge (ref. [26]). However, such devices are typically based on a conventional Esaki diode or resonant tunnel diode structure and have similar transport behaviour.

To achieve high tunnelling efficiency there are two key factors: (1) the ability to tune the density of states with Fermi-level alignment and (2) momentum conservation from the source to end in the momentum space, without involving phonons. In this regard, 2D black phosphorus (BP), which has a narrow bandgap of ~0.3 eV in multilayer thin films, has the potential to provide improved tunnel devices[23,25,27–29]. In particular, its direct bandgap can enhance the probabilities of band-to-band tunnelling efficiency because both the valence band maximum (VBM) and the conduction band minimum (CBM) are always located around the Γ point in the Brillouin zone, regardless of the layer thickness[30,31]. Moreover, few-layer BP shows satisfactory ambipolar electronic behaviour, with relatively high hole and electron mobilities, and its density of states can be controlled electrostatically[28–35].

In this Article, we report an efficient tunnelling device made from a 2D BP/Al_2O_3/BP sandwich structure. Unlike a conventional tunnel diode, the tunnelling carriers in this structure move in the transverse direction with respect to the drive current, and the electrostatic effect induced by the tunnelling carriers can significantly modulate the channel carrier density through the insulating layer, resulting in a giant NDR with a peak-to-valley ratio exceeding 100 at room temperature[16–18,22–25]. Furthermore, the tunnel structure enables an abrupt switching with a body factor (the relative change in gate voltage with respect to that of the surface potential) that is one-tenth of the thermionic limit in a wide temperature range (70–360 K).

Device structure and basic characterization

A schematic of the van der Waals structure device is shown in Fig. 1a. Layered BP flakes were first mechanically exfoliated onto a HfO_2/Si substrate. The 30 nm-thick HfO_2 layer was deposited by atomic layer deposition (ALD) at 300 °C on a p++ silicon wafer. The insulating region was patterned by electron-beam lithography, then a 2 nm aluminium seed layer was deposited by electron-beam evaporation (this is spontaneously oxidized into a thin Al_2O_3 layer on exposure to air). Next, another BP flake was deterministically transferred onto the top of the existing structure, with careful alignment. Considerable attention should be taken here to avoid any contact with the bottom BP layer (to avoid unwanted conduction paths in the device). Finally, 20/50 nm Ni/Au metal electrodes were deposited as contacts. A scanning transmission electron microscopy

[1]Wuhan National High Magnetic Field Center and School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan, China. [2]Institute of Microelectronics and Key Laboratory of Microelectronic Devices and Circuits (MoE), Peking University, Beijing, China. [3]Frontiers Science Center for Nano-optoelectronics, Peking University, Beijing, China. [4]Present address: Shenzhen Key Laboratory of Nanobiomechanics, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. [5]These authors contributed equally: Xiong Xiong, Mingqiang Huang. *e-mail: yqwu@pku.edu.cn

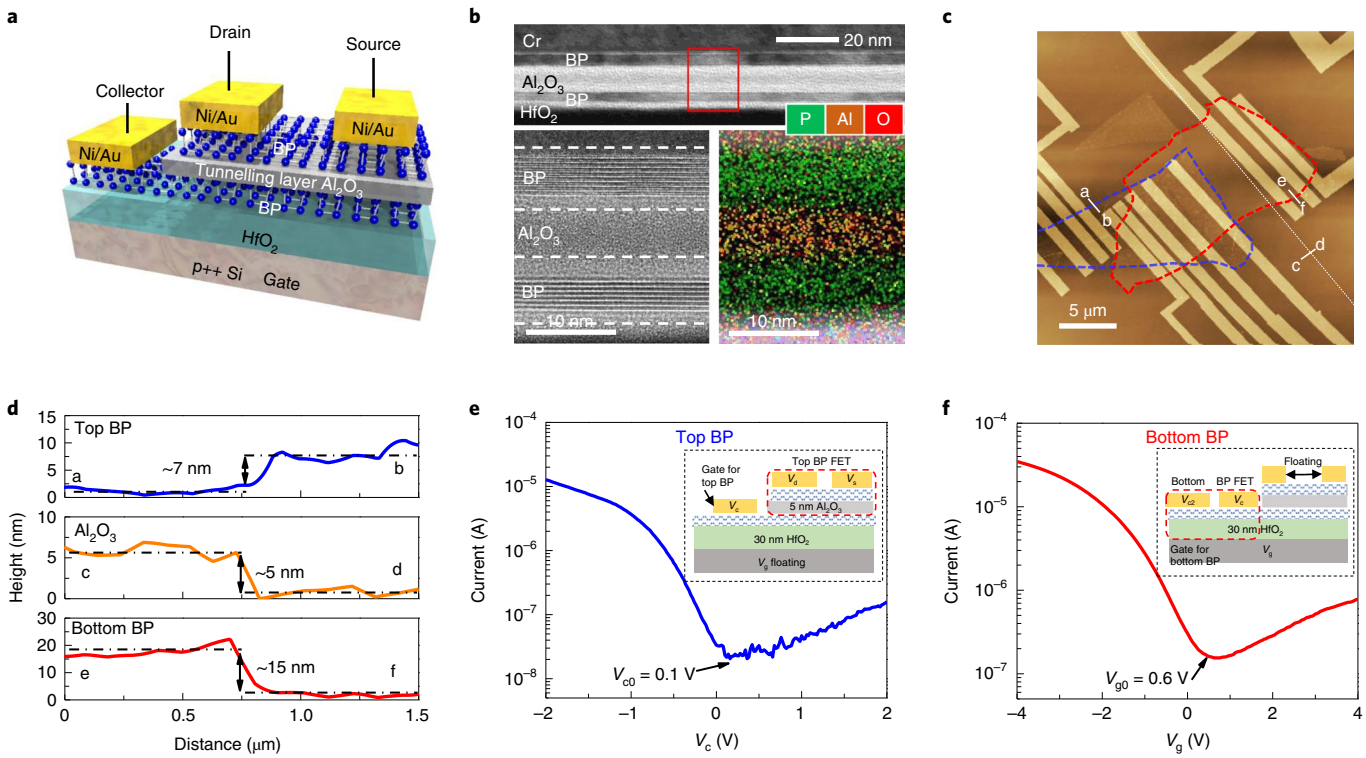

Fig. 1 | Device structure and characterization. a, Structure schematic view of the vertical van der Waals structure based on BP/alumina/BP. **b**, High-resolution STEM and EDX (inset) mappings of the device. **c**, Topography geometry of the device measured by AFM. The white dotted line represents the edge of the Al_2O_3 layer. **d**, Thickness of the top BP, alumina and bottom BP measured by AFM (7, 5 and 15 nm, respectively). **e**,**f**, Transfer characteristics and the operating schematics (insets) of the individual BP field-effect transistors at $V_{bias} = -0.05$ V, where the top BP FET (**e**) is modulated by V_c and the bottom BP FET (**f**) by V_g.

(STEM) image of the device is shown in Fig. 1b. The vertical sandwich structure contains two crystalline BP films separated by an amorphous thin alumina layer. The bottom right panel in Fig. 1b shows high-resolution energy-dispersive X-ray spectroscopy (EDX) mapping, where two sheets containing phosphorus element are clearly separated by the middle layer of Al. Figure 1c,d presents typical topography and layer thicknesses of the device, obtained using atomic force microscopy (AFM). The blue and red dashed frames represent the top and bottom BP flakes, respectively, and the white dotted line shows the edge of the sandwiching Al_2O_3 layer (left of this is the region capped by Al_2O_3). More characterizations are provided in Supplementary Fig. 2. Direct current (d.c.) transfer characteristics of the individual BP field-effect transistors of the top and bottom layer (and their operating schematic diagrams) are presented in Fig. 1e,f, respectively. The minimal current point V_{min} in the ambipolar transfer curves is located at $V_{c\_0} = 0.1$ V and $V_{g\_0} = 0.6$ V for the top and bottom BP, respectively (Supplementary Fig. 12). The difference between the two minimal points arises from the unintentional doping by the different dielectrics of the two layers; this creates a Fermi-level difference of ~30 meV, which plays an important role at the energy band edges, facilitating band bending for both BP layers, and is critical for the following tunnelling process.

NDR behaviour and transport mechanism

Figure 2a–d depicts the I_d–V_d output characteristics under four different V_g and V_c bias conditions, which create four different combinations of doping conditions for the two BP layers: PN, NN, PP and NP. When the bottom BP layer is p-doped under negative $V_g - V_{g0}$, as listed in Fig. 2a and Fig. 2c, both output curves show a monotonic increase with drain voltage regardless of the doping conditions of the top BP layer, with a current level on the order of 10 μA. The bottom BP layer is n-doped under positive $V_g - V_{g0}$, as listed in Fig. 2b and Fig. 2d. The drain current shows a monotonic increase at $V_c > V_{c\_0}$ in the range of a few μA when both layers are doped n-type in Fig. 2b. However, a striking difference emerges in Fig. 2d, where $V_c < V_{c\_0}$, meaning the top BP layer is p-doped while the bottom BP layer is n-doped. There is remarkable NDR behaviour where the drain current decreases abruptly from 19 μA to 0.7 μA within a narrow window of $\Delta V_d = 0.1$ V. In this NDR region, note that the drain current is more than 10^3 times larger than the gate leakage current I_g or collector current I_c, indicating that the NDR occurs in a much more magnified manner relative to the tunnelling current, which is fundamentally different from conventional tunnelling devices where the tunnelling current typically equals the change in drain current. The currents at other terminals of this device, under different bias conditions, are presented in Supplementary Fig. 3.

To qualitatively analyse the NDR behaviour, Fig. 2e–j shows the energy band diagrams of the device. Note that the voltage drop on the contacting BP area and related series resistance results in less efficient gate coupling of V_c compared with V_g. When the bottom BP layer is electrostatically n-doped by gate voltage V_g and the top BP layer is electrostatically p-doped by collector voltage V_c, the energy band alignment in this vertical structure forms a type III broken gap where the VBM of the p-type channel is very close to the CBM of the n-type layer. The extra band bending from the original Fermi-level difference will further shift energy bands in opposite directions and create a broken-gap structure while introducing heavily doped regions near the vicinity of the Al_2O_3 tunnel barrier, meeting the band alignment requirement for efficient band-to-band tunnelling (Supplementary Fig. 5), while the efficiency can

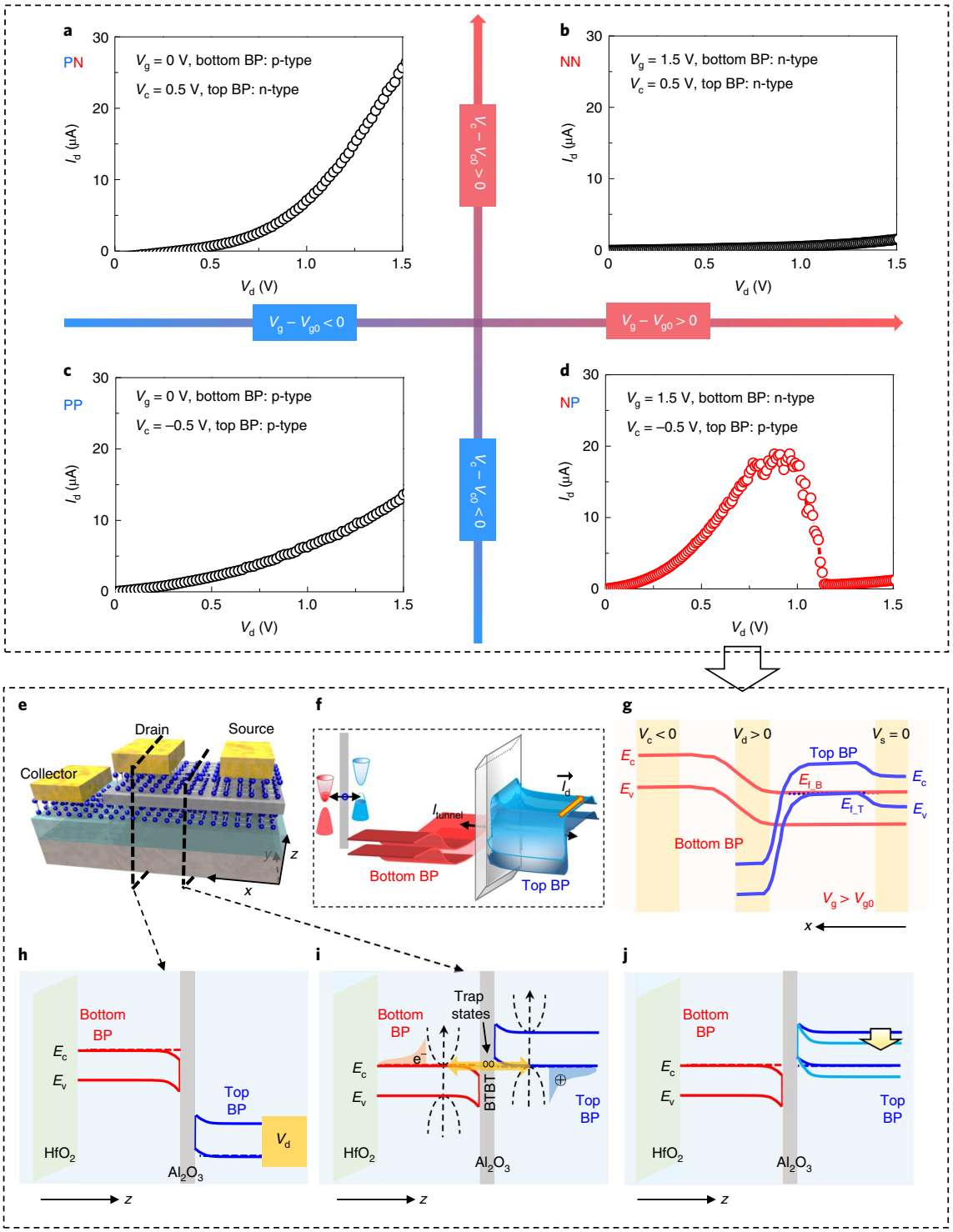

Fig. 2 | Four-terminal tunnel device with NDR. a–d, I_d–V_d curves under different V_g and V_c bias conditions (shown in **a–d**). NDR appears only when the bottom BP is n-doped and top BP is p-doped. **e–i**, Structure and energy band diagrams of the device under bias conditions for NDR, showing band alignment conditions for band-to-band tunnelling: structure of the devices (**e**); three-dimensional (3D) energy band diagram (**f**); 2D energy band diagram (**g**); energy band diagram slices of the drain region (**h**) and channel region (**i**). **j**, The positive virtual gating effect changes the channel potential causing NDR.

be increased with assisting traps from the interface and inside the Al_2O_3 (Supplementary Fig. 9). On the other hand, instead of contributing directly to the output current, transverse carrier tunnelling will act as a positive virtual gate to change the potential of the top channel, which further reduces the p-doping level of the top BP, as shown in Fig. 2j (for more details see Supplementary Fig. 16). The current of the top BP channel will be reduced dramatically via this tunnelling process.

Systematic studies on the tunability of the dependence of NDR on the back gate voltage V_g and collector voltage V_c were carried out as shown in Fig. 3a,b. A sweep rate of ~35 ms per point (201 points per single curve sweep) was mainly used for the measurements. Figure 3a

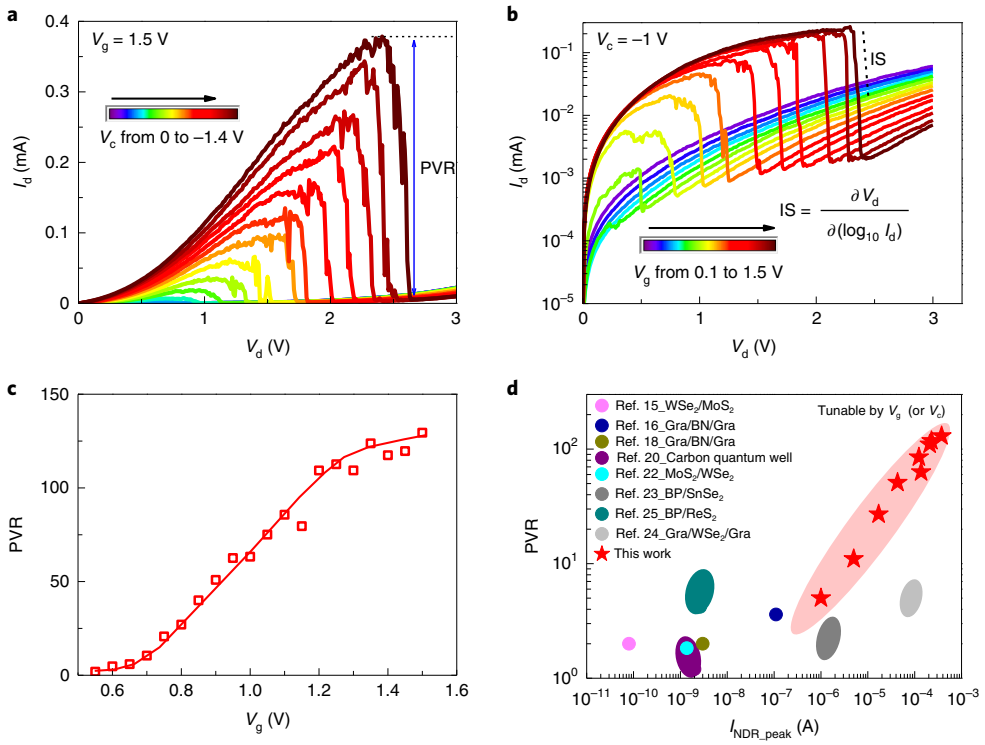

Fig. 3 | Tunable NDR behaviour. a, I_d–V_d curves with V_c changing from 0 V to −1.4 V at V_g = 1.5 V. **b**, I_d–V_d curves with V_g changing from 0.1 V to 1.5 V at V_c = −1 V. The inverse slope (IS) is defined as a change in drain voltage that produces an order of magnitude change in current in the NDR region. **c**, The PVR at different V_g, extracted from **b**. **d**, Benchmark of PVR versus peak drain current I_d for various previously reported 2D material NDR devices.

shows the output characteristics, I_d–V_d, of the tunnel device, with V_c changing from 0 V to −1.4 V at fixed V_g = 1.5 V. The collector voltage can effectively modulate the channel Fermi level such that the drain current increases with V_c when NDR occurs. In the NDR region, the position of peak voltage V_p increases with larger negative V_c, and this trend persists over the entire voltage sweep range. This can be attributed to the larger negative V_c further pulling up the energy band of the top BP channel; under this condition, the critical condition for band-to-band tunnelling requires larger positive V_d to pull down the band edge near the junction to meet the Fermi-level alignment. Figure 3b presents the output characteristics of the same device with V_g increasing from 0.1 V to 1.5 V at a fixed V_c of −1 V. The drain current shows a steep drop of about two orders of magnitude within 70 mV of drain voltage at high current densities, which indicates that a vital parameter, the inverse slope (IS, defined in a similar way as subthreshold slope, SS), is very low. This small IS value can be used to cause a very small SS below the thermionic limit, together with other parameters that are unique to this device structure. Similarly, for the back gate voltage, the peak voltage V_p also shows a right shift as V_g increases. This is because, under larger positive V_g, the energy band of the bottom BP will be pulled down. To meet the critical condition of Fermi-level alignment, more positive V_d is needed to pull down the band edge of the top BP layer. The currents of other terminals of this device under different bias conditions are presented in Supplementary Fig. 13.

For technologically relevant applications, the peak-to-valley current ratio (PVR) and the peak current magnitude are typically considered the two most important parameters. As summarized in Fig. 3c, the PVR can be tuned continuously by the gate voltage over a wide range, up to a huge PVR of 130 at room temperature. Another important parameter, the peak current, was extracted from Fig. 3b and is plotted versus PVR along with values for other similar devices in Fig. 3d. The peak current can be modulated over a wide range (more than two decades) and can reach hundreds of μA (for a device channel width of ~10 μm), much higher than in previous 2D devices. The PVR of 130 at room temperature is also higher than in any 2D-based tunnel device investigated so far, with almost 20 times improvement over previous results[16–18,22–25]. This large PVR benefits from the large peak current from the top BP layer before tunnelling and the low valley current, which is dominated by the effective modulation of carrier density after tunnelling from the electrostatic control of the virtual gating effect. Measurements from four other representative devices with NDR are presented in Supplementary Fig. 14. All devices, with different BP channel layer thicknesses, are presented and summarized in Supplementary Figs. 10, 11 and 14 and Supplementary Table 1.

To better understand the mechanisms of this unique tunnelling structure, we compared the classical Esaki diode with those operated based on this principle. The p+n+ junctions were created by heavily doping in opposite polarities and forming type III homo- or heterojunctions, where the CBM in the n-side meets or becomes lower than the VBM in the p-side, whereby sufficient band-to-band tunnelling can occur under reverse bias or small forward biases. As the forward bias increases, the energy band in the p-side will be pulled down and the overlap region, which determines the available states for tunnelling, becomes smaller and eventually diminishes once reaching the critical voltage point where the current reaches the valley point. In essence, the band-to-band tunnelling dominates the whole region before the valley point, and NDR occurs where tunnelling probability starts to decrease because of the band alignment, and thermal current dominated by high energy tail states starts to take over in the higher bias region. The apparent difference in our structure, which we call transverse tunnelling (TT), lies in the fact that, before the critical point of NDR, the drive current in the channel operates like a conventional field-effect transistor, following the diffusion-drift model, without any tunnelling process. When Fermi-level and band-edge alignment modulated by the gate voltages meets the band-to-band tunnelling

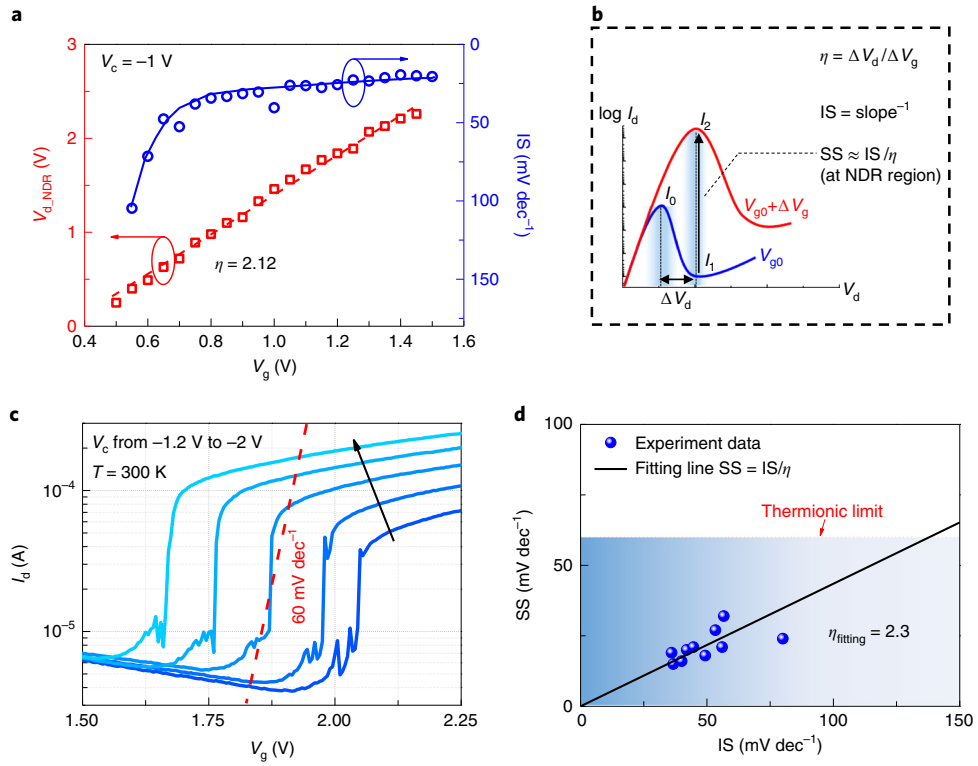

Fig. 4 | Abrupt switching in TT-FETs. a, The V_d of the NDR current peak and IS from the NDR region for different gate voltages extracted from Fig. 3b, in which $\eta = d(V_{d\_NDR})/d(V_g)$ is the gate-to-drain control efficiency of the device. **b**, Schematic of the prediction of abrupt switching in the NDR region. **c**, Abrupt switching, much less than 60 mV dec^{-1}, in the I_d–V_g curves at different V_c with V_d = 1.5 V at 300 K. **d**, Fitting of the experimental data from **c** using SS = IS/η. The obtained η value is ~2.3, very close to the experimental data of η = 2.4 shown in Supplementary Fig. 6.

condition, efficient tunnelling occurs, and the carriers tunnel through the insulator between the two BP layers at the peak position. These positively charged carriers in the bottom BP layer also act as a virtual gate to reduce the hole doping density in the top channel, which in turn drastically reduces the drive current electrostatically at the valley position.

Conventional real space transfer devices demonstrate NDR phenomena via thermal emission transport over the top of the barrier from channel to collector[36]. In such transistors, the onset of NDR is triggered by dynamical channel depletion by hot carriers (as described in detail for negative resistance field-effect transistor (NERFET) and charge injection transistor (CHINT) devices)[37,38]. In tunnel FETs and the Esaki diode, the output current typically equals the tunnelling current before the valley position of NDR behaviour, until the BTBT current diminishes and thermal current takes over[39]. It should also be noted that, unlike conventional tunnel FETs where tunnelling current contributes directly to the output current, the carrier tunnel direction in the TT-FET is transverse to the output current flow and serves as a virtual gating effect, and the gate voltage can efficiently control energy band alignment for the onset of carrier tunnelling (Supplementary Fig. 8).

Abrupt switching properties

We also analysed the dependence of the position of peak voltage on gate voltage in the NDR behaviour (Fig. 4a, left axis). The linear fitting line shows the gate-to-drain control efficiency (η = 2.12). It should be noted that this positive larger-than-unity efficiency indicates that the gate control can effectively shift the critical condition of band-to-band tunnelling where a larger drain voltage needs to be applied to compensate for the shift. This linear relationship shows the direct link between the occurrence of NDR and the doping level in the bottom BP layer, and the extra doping effect offsetting the back gate electrostatic doping can only originate from BTBT carriers via the virtual positive gating effect. Another critical factor is the IS, corresponding to the steepness of the NDR region, where current undergoes a sharp decrease. Figure 4a (right axis) also shows the IS exhibiting a decreasing trend from ~100 mV dec^{-1} then quickly reaching a value as low as ~20 mV dec^{-1} as V_g increases. These small IS values under various gate voltages show the ultra-efficient control of the transverse electric field and current conduction in the lateral channel. As a result, on combining these two parameters together, there is a region where current can rise quickly from a valley to the next peak with a small increase in gate voltage, as shown in the neighbouring two curves under different gate voltage illustrated in Fig. 4b, and this rise can be steep enough to beat the thermionic limit if the ratio of IS to η is small enough (for details see Supplementary Fig. 7).

To verify the above theory, transfer characteristic I_d–V_g curves were measured at different V_c, as shown in Fig. 4c for gate voltage steps of 5 mV (to obtain reliable data). All the curves show abrupt switching of less than 25 mV dec^{-1}, with a step change of current from 10 μA to 100 μA. The simple equation SS = IS/η, as predicted in Fig. 4b, was used to fit the experimental data from Fig. 4c, and the best fitting parameter, η = 2.3, was obtained as shown in Fig. 4d. This is very close to the experimental value of η = 2.4 extracted from the drain voltage change at peak current in the NDR behaviour (Supplementary Fig. 6). This further confirms that the SS below the thermal limit can only be obtained from a combination of small IS and larger gate-to-drain control efficiency η for these NDR devices. Furthermore, the steep switching in I_d–V_d or I_d–V_g plots shows little dependence on sweep direction and hysteresis loops (Supplementary Fig. 4). With a faster sweep rate during measurements, the knee point of NDR shows a slight shift to the right, resulting in a higher peak current and valley current (Supplementary Fig. 15). However,

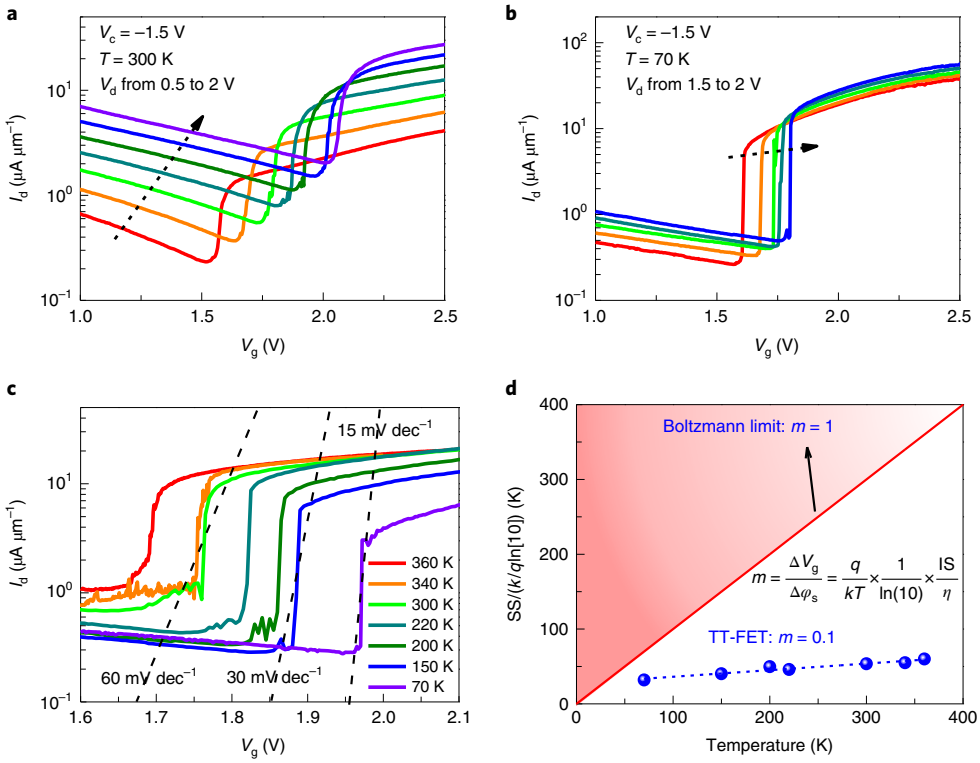

Fig. 5 | Abrupt switching at various temperatures. a,**b**, I_d–V_g characteristics with different V_d at room temperature (**a**) and 70 K (**b**). The V_g step during measurements was 5 mV. The negative DIBL effect can be observed with a steep slope. **c**, I_d–V_g curves at various temperatures with fixed V_c = −1.8 V and V_d = 1.5 V. Three classic thermionic limits with 300, 150 and 75 K guidelines have been plotted. **d**, A body factor of 0.1 ($m = \partial V_g/\partial \varphi_s$) is extracted from the experimental data using SS = mkT/qln(10) from 360 K to 70 K, far below the Boltzmann limit of m = 1.

the average IS and the η factor are almost the same with different sweep rates, proving that the prediction of SS in Fig. 4d is reliable.

Negative drain-induced barrier-lowering (DIBL) effects, usually associated with NDR, have been predicted and demonstrated in negative-capacitance FETs due to the drain-coupled negative capacitance effect[40,41]. Similarly, the I_d–V_g characteristics at different V_d of TT-FETs at 300 K and 70 K are presented in Fig. 5a,b; negative DIBL with a steep slope is shown at both room temperature and low temperature, which is a signature effect of NDR and fundamentally different from the behaviour of previous tunnelling FETs. To further investigate the transport mechanism, the temperature-dependent current–voltage characteristics of these devices were investigated. I_d–V_g curves under the same voltage bias V_d = 1.5 V and V_c = −1.8 V at various temperatures are shown in Fig. 5c. The devices exhibit very consistent abrupt switching at various temperatures, with the SS below the Boltzmann limit for a wide temperature range from 360 K to 70 K. Body factor analysis is shown in Fig. 5d; this corresponds to the relative change in gate voltage with respect to the change in surface potential (Supplementary Note 1). The lower limit of body factor for conventional thermionic transport is m = 1, corresponding to the limit of carrier transport at the high energy tail in Boltzmann statistics from the Fermi–Dirac distribution. Body factor values less than 1 have become the primary goal for tunnel FETs[39,42,43] and negative-capacitance transistors[44,45]. In our TT-FETs, an extremely small body factor of m = 0.1 is achieved across a wide temperature range from 360 K to 70 K, demonstrating the ultra-efficient electrostatic control of this tunnelling dominant transport, surpassing previous tunnelling transistor structures.

Conclusions

We have shown that a vertical van der Waals tunnel structure, in which the tunnelling carriers are in the transverse direction with respect to the drive current, can offer a tunable NDR with a PVR of up to 130. In these TT-FETs, the critical transport conditions can be tuned by an electric field and the transverse tunnelling carriers exhibit efficient electrostatic control over the drive current. The high efficiency in gate control can lead to an abrupt switching with a body factor that is only one-tenth of the Boltzmann limit, across a broad temperature range of 70–360 K.

Methods

Device fabrication. Few-layer BP flakes were first mechanically exfoliated from commercially available bulk BP crystal (Smart Elements) and transferred onto 30 nm HfO_2 by atomic layer deposition on a Si substrate in a glove box, with the oxygen and water contents kept below 0.1 ppm. $Hf[N(CH_3)(C_2H_5)]_4$ and ozone (O_3) were used as precursors for Hf and oxygen sources. The estimated equivalent oxide thickness was ~7 nm. The insulating region was patterned by electron-beam lithography and a 2 nm aluminium seed layer was deposited by electron-beam evaporation, which was spontaneously oxidized into a thin Al_2O_3 layer on exposure to air. Next, another BP flake was transferred on top of the existing structure with careful alignment on top of the insulating layer, without any contact of the bottom BP layer to avoid potential in-series or in-parallel conduction paths in the device. Finally, 20/50 nm Ni/Au metal electrodes were separately deposited as the ohmic contacts on the top channel and floating gate. The entire device fabrication process is shown in Supplementary Fig. 1.

STEM images. A cross-sectional specimen of the device was prepared by using a lift-out process in a dual-beam FEI Quanta 3D FEG focus ion beam (FIB) system. The STEM images of the sample were then obtained using a Titan G2 60–300 high-resolution transmission electron microscope.

D.c. characterizations. The device was placed inside a Lakeshore cryogenic probe station in vacuum (<2×10^{-4} mbar) during all electrical characterizations. Electrical measurements were carried out using an Agilent B1500A semiconductor parameter analyser.

Data availability

The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

Received: 7 October 2018; Accepted: 30 December 2019;
Published online: 03 February 2020

References

1. Maekawa, T. et al. Frequency increase in terahertz oscillation of resonant tunnelling diode up to 1.55 THz by reduced slot-antenna length. *Electron. Lett.* **50**, 1214–1216 (2014).
2. Feiginov, M. et al. Resonant-tunnelling-diode oscillators operating at frequencies above 1.1 THz. *Appl. Phys. Lett.* **99**, 233506 (2011).
3. Koyama, Y. et al. Oscillations up to 1.40 THz from resonant-tunneling-diode-based oscillators with integrated patch antennas. *Appl. Phys. Express* **6**, 064102 (2013).
4. Jin, N. et al. Tri-state logic using vertically integrated Si-SiGe resonant interband tunneling diodes with double NDR. *IEEE Electron Device Lett.* **25**, 646–648 (2004).
5. Gan, K.-J., Tsai, C.-S., Chen, Y.-W. & Yeh, W.-K. Voltage-controlled multiple-valued logic design using negative differential resistance devices. *Solid-State Electron.* **54**, 1637–1640 (2010).
6. Berezowski, K. S. & Vrudhula, S. B. Multiple-valued logic circuits design using negative differential resistance devices. In *37th Int. Symposium on Multiple-Valued Logic* (*ISMVL'07*) (IEEE, 2007).
7. Gong, C., Zhang, H., Wang, W. & Colombo, L. Band alignment of two-dimensional transition metal dichalcogenides: application in tunnel field effect transistors. *Appl. Phys. Lett.* **103**, 053513 (2013).
8. Shewchun, J. & Temple, V. Theoretical tunneling current characteristics of the SIS (semiconductor–insulator–semiconductor) diode. *J. Appl. Phys.* **43**, 5051–5061 (1972).
9. Chhowalla, M., Jena, D. & Zhang, H. Two-dimensional semiconductors for transistors. *Nat. Rev. Mater.* **1**, 16052 (2016).
10. Huang, M. et al. Multifunctional high-performance van der Waals heterostructures. *Nat. Nanotechnol.* **12**, 1148–1154 (2017).
11. Mou, X., Register, L. F., MacDonald, A. H. & Banerjee, S. K. Bilayer pseudospin junction transistor (BiSJT) for 'beyond-CMOS' logic. *IEEE Trans. Electron Devices* **64**, 4759–4762 (2017).
12. Geim, A. K. & Grigorieva, I. V. Van der Waals heterostructures. *Nature* **499**, 419–425 (2013).
13. Yan, R. et al. GaN/NbN epitaxial semiconductor/superconductor heterostructures. *Nature* **555**, 183–189 (2018).
14. Zhang, C. et al. Systematic study of electronic structure and band alignment of monolayer transition metal dichalcogenides in van der Waals heterostructures. *2D Mater.* **4**, 015026 (2017).
15. Lin, Y.-C. et al. Atomically thin resonant tunnel diodes built from synthetic van der Waals heterostructures. *Nat. Commun.* **6**, 7311 (2015).
16. Britnell, L. et al. Resonant tunnelling and negative differential conductance in graphene transistors. *Nat. Commun.* **4**, 1794 (2013).
17. Mishchenko, A. et al. Twist-controlled resonant tunnelling in graphene/boron nitride/graphene heterostructures. *Nat. Nanotechnol.* **9**, 808–813 (2014).
18. Fallahazad, B. et al. Gate-tunable resonant tunneling in double bilayer graphene heterostructures. *Nano Lett.* **15**, 428–433 (2015).
19. Chen, J., Reed, M., Rawlett, A. & Tour, J. Large on–off ratios and negative differential resistance in a molecular electronic device. *Science* **286**, 1550–1552 (1999).
20. Perrin, M. L. et al. Large negative differential conductance in single-molecule break junctions. *Nat. Nanotechnol.* **9**, 830–834 (2014).
21. Bhattacharyya, S. et al. Resonant tunnelling and fast switching in amorphous-carbon quantum-well structures. *Nat. Mater.* **5**, 19–22 (2006).
22. Roy, T. et al. Dual-gated MoS_2/WSe_2 van der Waals tunnel diodes and transistors. *ACS Nano* **9**, 2071–2079 (2015).
23. Yan, R. et al. Esaki diodes in van der Waals heterojunctions with broken-gap energy band alignment. *Nano Lett.* **15**, 5791–5798 (2015).
24. Burg, G. W. et al. Coherent interlayer tunneling and negative differential resistance with high current density in double bilayer graphene–WSe_2 heterostructures. *Nano Lett.* **17**, 3919–3925 (2017).
25. Shim, J. et al. Phosphorene/rhenium disulfide heterojunction-based negative differential resistance device for multi-valued logic. *Nat. Commun.* **7**, 13413 (2016).
26. Sarkar, D. et al. A subthermionic tunnel field-effect transistor with an atomically thin channel. *Nature* **526**, 91–95 (2015).
27. Li, L. et al. Black phosphorus field-effect transistors. *Nat. Nanotechnol.* **9**, 372–377 (2014).
28. Li, T. et al. High field transport of high performance black phosphorus transistors. *Appl. Phys. Lett.* **110**, 372 (2017).
29. Xiong, X. et al. High performance black phosphorus electronic and photonic devices with HfLaO dielectric. *IEEE Electron Device Lett.* **39**, 127–130 (2017).
30. Castellanos-Gomez, A. Black phosphorus: narrow gap, wide applications. *J. Phys. Chem. Lett.* **6**, 4280–4291 (2015).
31. Cai, Y., Zhang, G. & Zhang, Y.-W. Layer-dependent band alignment and work function of few-layer phosphorene. *Sci. Rep.* **4**, 6677 (2014).
32. Asahina, H. & Morita, A. Band structure and optical properties of black phosphorus. *J. Phys. C* **17**, 1839–1852 (2000).
33. Xia, F., Wang, H. & Jia, Y. Rediscovering black phosphorus as an anisotropic layered material for optoelectronics and electronics. *Nat. Commun.* **5**, 4458 (2014).
34. Du, Y., Liu, H., Deng, Y. & Ye, P. D. Device perspective for black phosphorus field-effect transistors: contact resistance, ambipolar behavior and scaling. *ACS Nano* **8**, 10035–10042 (2014).
35. Li, X. et al. Mechanisms of current fluctuation in ambipolar black phosphorus field-effect transistors. *Nanoscale* **8**, 3572–3578 (2016).
36. Sze, S. M. & Ng, K. K. *Physics of Semiconductor Devices* (Wiley, 2006).
37. Kastalsky, A. & Luryi, S. Novel real-space hot-electron transfer devices. *IEEE Electron Device Lett.* **4**, 334–336 (1983).
38. Kastalsky, A. et al. A field-effect transistor with a negative differential resistance. *IEEE Electron Device Lett.* **5**, 57–60 (1984).
39. Lu, H. & Seabaugh, A. Tunnel field-effect transistors: state-of-the-art. *IEEE J. Electron Devices Soc.* **2**, 44–49 (2014).
40. Ota, H. et al. Fully coupled 3-D device simulation of negative capacitance FinFETs for sub 10 nm integration. In *IEEE Int. Electron Devices Meeting* 318–321 (IEEE, 2016).
41. Si, M. et al. Steep-slope hysteresis-free negative capacitance MoS_2 transistors. *Nat. Nanotechnol.* **13**, 24–28 (2018).
42. Ganjipour, B., Wallentin, J., Borgström, M. T., Samuelson, L. & Thelander, C. Tunnel field-effect transistors based on InP-GaAs heterostructure nanowires. *ACS Nano* **6**, 3109–3113 (2012).
43. Dewey, G. et al. Fabrication, characterization, and physics of III–V heterojunction tunneling field effect transistors (H-TFET) for steep sub-threshold swing. In *IEEE Int. Electron Devices Meeting* 785–788 (IEEE, 2011).
44. Salahuddin, S. & Datta, S. Use of negative capacitance to provide voltage amplification for low power nanoscale devices. *Nano Lett.* **8**, 405–410 (2008).
45. Then, H. W. et al. Experimental observation and physics of 'negative' capacitance and steeper than 40 mV/decade subthreshold swing in $Al_{0.83}In_{0.17}N$/AlN/GaN MOS-HEMT on SiC substrate. In *IEEE Int. Electron Devices Meeting* 691–694 (IEEE, 2013).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grants 91964106 and 61874162), the 111 Project (B18001) and the Strategic Priority Research Program of the Chinese Academy of Sciences (grant XDB30000000). We thank the staff at the Center of Micro-fabrication and Characterization of Wuhan National Laboratory for Optoelectronics and Huazhong University of Science and Technology Analytical and Testing Center for support with electron-beam lithography, electron-beam evaporation and transmission electron microscopy.

Author contributions

Y.W. proposed and supervised the project. Y.W., X.X. and M.H. designed the experiment. X.X. and M.H. performed device fabrication and characterization. B.H. performed the simulations. X.L., F.L., S.L., M.T., T.L. and J.S. assisted with device fabrication and discussions. X.X., M.H. and Y.W. analysed the data and co-wrote the manuscript. All authors discussed the results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41928-019-0364-5.

Correspondence and requests for materials should be addressed to Y.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.