Multi-purpose Black-Phosphorus/hBN heterostructures

Gabriel C. Constantinescu[*,†] and Nicholas D. M. Hine[‡]

†*Cavendish Laboratory, University of Cambridge, 19 JJ Thomson Avenue, Cambridge CB3 0HE, UK*

‡*Department of Physics, University of Warwick, Coventry CV4 7AL, UK*

E-mail: gcc32@cam.ac.uk

Abstract

Black phosphorus (BP) has recently emerged as a promising semiconducting two-dimensional material. However, its viability is threatened by its instability in ambient conditions, and by the significant decrease of its band gap in multilayers. We show that one could solve all the aforementioned problems by interfacing BP with hexagonal boron nitride (hBN). To this end, we simulate large, rotated hBN/BP interfaces using linear-scaling density functional theory (DFT). We predict that hBN-encapsulation preserves the main electronic properties of the BP monolayer, while hBN spacers can be used to counteract the band gap reduction in stacked BP. Finally, we propose a model for a tunneling field effect transistor (TFET) based on hBN-spaced BP bilayers. Such BP TFETs would sustain both low-power and fast-switching operations, including negative differential resistance behaviour with peak-to-valley ratios of the same order of magnitude as those encountered in transition metal dichalcogenide TFETs.

Keywords

2D heterostructures; tunneling transistor; linear-scaling DFT; electric fields

Since its inception, the two-dimensional (2D) material community has switched focus several times between different materials, aiming to satisfy different technological requirements. For instance, the lack of a band gap in graphene caused an increased interest in semiconducting transition metal dichalcogenides (TMDCs). While their many qualities have enabled their remarkable success in semiconductor electronics,[1,2] spintronics,[3] and optoelectronics,[2,4] the associated flaws of TMDCs are also of concern. Most notably, they only exhibit a direct band gap when in monolayer form,[5] the carrier mobilities are far smaller than those encountered in graphene, and their large band gap implies that the infrared (IR) spectrum cannot be harnessed by TMDC optoelectronics.

In this context, the spotlight has recently shifted to layered black phosphorus (BP),[6,7] a material which rectifies the most important shortcomings of TMDCs. The band gap of BP is direct even in stacked forms, and its value changes significantly with the number of layers: from 0.3 eV in bulk to 1.5-2.0 eV for the monolayer.[8,9] This range of values allows BP to exploit a range of spectrum previously unharnessed by graphene and TMDCs, corresponding to the mid- and near-IR. Moreover, the higher carrier mobilities ($1,000\ \mathrm{cm^2/Vs}$)[10] compared to TMDC monolayers, the strongly anisotropic electronic and optical properties,[11–15] as well as the robustness under elastic strain[16–18] have already ensured novel uses of BP in electronic,[9,10,12,19,20] photonic,[21–23] and thermoelectric[24,25] applications.

Unfortunately, alongside these highly desirable properties, BP also has drawbacks. Firstly, it is highly sensitive to molecules present in air[26] and it is structurally unstable in ambient conditions,[27,28] meaning that pristine BP flakes are unusable after only a couple of hours.[29] Such effects prevent the holes in monolayer BP from reaching their maximum predicted mobility of $10,000\ \mathrm{cm^2/Vs}$.[11] Secondly, the decrease of the band gap in BP stacks implies that

only the monolayer is usable for near-IR optoelectronic applications, limiting the external quantum efficiency of such devices. This band gap decrease cannot be avoided by simply misaligning the BP layers, as we have previously shown that stacked layers always interact at Γ in reciprocal-space,[30] exactly where the BP direct band-gap is. Lastly, due to the band gap being smaller than that of TMDCs, the current on/off ratio for BP transistors is lower; therefore, pristine BP channels are not at first sight as appealing as TMDCs in ultra-low power applications.

In this work, we propose that monolayers of an inert large-band gap insulator, such as hexagonal boron nitride (hBN), could solve or alleviate the aforementioned flaws of pristine BP. By simulating large rotated hBN/BP interfaces, we deduce three main conclusions: hBN encapsulation does not negatively influence any of the main electronic properties of BP, hBN spacers prevent the band gap reduction in stacked forms of BP, and BP/hBN/BP heterostructures are suitable as tunneling field-effect transistors (TFETs), an architecture that allows BP to be of use in both high-speed and low-power applications. The first point suggests that there is no electronic-structure downside to hBN-encapsulation, which is already known to protect BP from environmental interactions.[29,31,32] The second conclusion is especially relevant for optoelectronics, as it implies that hBN-spaced BP stacks have the same light-absorption properties as monolayer BP, with the essential difference that more photons are absorbed due to the increased thickness. Finally, since the operation of BP/hBN/BP as a TFET is based on quantum tunneling, instead of thermionic excitation of carriers, we encounter negative differential resistance (NDR) peaks[33] with peak-to-valley ratios (PVRs) comparable to those predicted for TMDC TFETs,[34] as well as subthreshold swings below the minimum theoretical limit for conventional field effect transistors.[35] We note that intricate heterostructure 2D stacks have already been experimentally obtained,[36] meaning that our proposition of hBN-encapsulated or hBN-spaced BP is realistic. Moreover, our first two conclusions do not require hBN to be in monolayer form, though of course the currents in the TFET application will decrease with the hBN-spacer thickness.

Since our hBN/BP heterostructures are too large for traditional DFT to be computationally feasable, we utilise linear-scaling (LS) DFT, as implemented in the ONETEP code.[37,38] This LS-DFT implementation is significantly more computationally efficient for large systems than plane-wave DFT methods, but with a tunable accuracy equivalent to the latter.[39] Localized orbitals, coined non-orthogonal generalised Wannier functions (NGWFs), are self-consistently optimized in-situ, eliminating the need for basis set superposition corrections[40] and keeping the basis set size required for good convergence small. Moreover, the density kernel is also self-consistently optimised, while constrained to be idempotent and normalised.[41]

We utilise the optB88-vdW functional for exchange and correlation, due to its proven track-record in describing both covalent and weak interactions in 2D materials.[30] The projector-augmented-wave (PAW) formalism[42] was employed for all the LS-DFT calculations, using the atomic datasets developed by Garrity *et al.*,[43] which have been validated for a wide range of covalent compounds.

Furthermore, we utilised a 800 eV cutoff energy, a 40.0 Å vertical supercell spacing, and NGWFs with a radius of 11.0 bohr for all involved atoms. Since the ONETEP calculations were performed at the Γ point, supercells must be chosen so as to converge with respect to the k-point sampling of the underlying primitive cell. We find that the BP/hBN interfaces must contain at least 12×12 hBN units and 11×8 BP units, in order to ensure sufficient convergence. Prior to the interface simulations, the structures of the individual hBN and BP layers were obtained by optimising the geometry of the primitive unit cells using the QuantumEspresso[44] plane-wave DFT code. The results match well to experimental lattice parameters (within $\approx 1\%$ for hBN,[45,46] and $< 3\%$ for BP[47]), as well as theory-derived ones (obtained through MP2 for hBN[48] and DFT for BP[49]). The full reasoning behind all our methodological choices is explained in the Supporting Information.

Since hBN and BP belong to different crystal systems (hexagonal $P6_3/mmc$, and orthorhom-

bic D_{2h}^{18}, respectively) with significantly different lattice parameters, it is clear that their superposition produces Moiré patterns. To obtain the associated interface supercells, we have defined orthorhombic hBN unit cells from 2×1 primitive hBN units, and afterwards followed the instructions of Hermann[50] for the case of overlapping lattice-mismatched orthorhombic layers. We chose a rotation angle which reduces the supercell size as much as possible, while allowing a maximum strain of 1% in either direction of the BP layers. This value is smaller than the 2% strain which would be required to noticeably modify the BP electronic structure.[18] The resulting structures for the encapsulated BP monolayer (hBN/BP/hBN) and the spaced BP bilayer (BP/hBN/BP) are shown in Fig. 1.f and 1.g, respectively. Both interfaces have a rotation angle of 40.9° between the hBN and BP layers, and also obey our previously-discussed supercell-size constraints. Note that the 40.9° angle is not a prediction for an experimentally observable orientation, but merely a means of reducing the supercell size. However, based on our previous work on rotated TMDC interfaces,[30] we can expect that the rotation angle between the hBN spacer and the BP layers has no significant influence on the stability or the electronic structure near Γ for hBN/BP/hBN and BP/hBN/BP. The hBN/BP/hBN simulation cell has 1824 atoms, while BP/hBN/BP has 1632 atoms. In an effort to take into account structural modifications upon interfacing we first relaxed the atomic positions in hBN/BP/hBN and BP/hBN/BP, with force tolerances of to $2 \cdot 10^{-3}$ a.u./bohr. For hBN/BP/hBN, the average distance between the inner P subplanes and hBN was d=3.56 Å, which is within the range of previously predicted values for smaller aligned hBN/BP/hBN models.[51,52] We encounter a similar BP-hBN interlayer distance (d=3.49 Å) in BP/hBN/BP, the important distinction being that the hBN-corrugation was 5 times more pronounced in the case of BP/hBN/BP ($\Delta d < 0.24$ Å) than in hBN/BP/hBN ($\Delta d < 0.05$ Å).

One of the main goals of this work is to see if any significant alterations occur to the electronic-structure of BP when it is either encapsulated by hBN or when hBN is used as a spacer. As a measure of such effects, we utilise the one-particle Kohn-Sham spectral

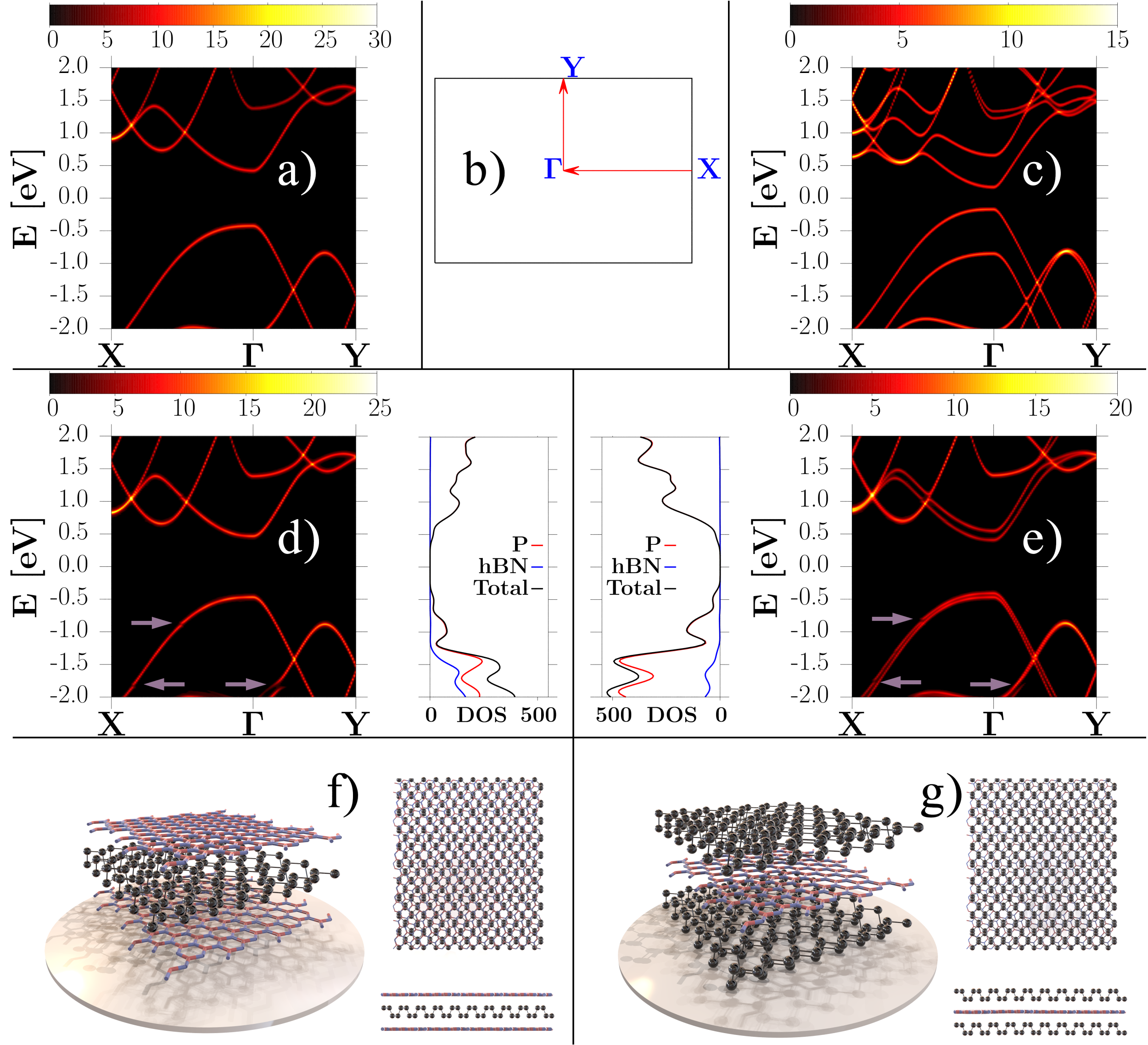

Figure 1: Unfolded spectral function for the BP monolayer (a), the top BP sheet in the BP AA-stacked bilayer (c), the BP layer in the hBN/BP/hBN heterostructure (d), and the top BP sheet in the BP/hBN/BP heterostructure (e). Primitive-cell Brillouin-zone sampling path depicted in (b). Local densities of states of BP and hBN are shown for hBN/BP/hBN (d) and BP/hBN/BP (e). All the energy scales are with respect to the mid-gap level. Violet arrows in (d),(e) highlight band discontinuities due to hBN-BP interactions. Artistic representation, and top/side views of the hBN/BP/hBN (f) and BP/hBN/BP (g) heterostructures – P atoms are black spheres.

function $A_{\mathbf{k}j,\mathbf{k}j}(\omega)$, which is equivalent to the probability that an electron of momentum $\mathbf{k}$ can be added/removed into/from band j, with an energy of ω. In this manner, we also gain means of direct comparison with angular-resolved photoemission spectroscopy experiments

(ARPES), as explained in previous works.[30,53] To access only the Brillouin zone of the BP primitive cell, we need to unfold the supercell spectral function by projecting the spectral function operator onto the monolayer of interest and afterwards changing the representation basis to the primitive-cell eigenstates of the monolayer in question:

$$A^{I}_{\mathbf{k}j,\mathbf{k}j} = \sum_{\substack{\mathbf{K},\\ \rho\in I}} \langle \Psi_{\mathbf{k}j} | \Psi_{\mathbf{K}J} \rangle \, A_{\mathbf{K}J,\mathbf{K}J} \, \langle \Psi_{\mathbf{K}J} | \phi^{\rho} \rangle \, \langle \phi_{\rho} | \Psi_{\mathbf{k}j} \rangle \,, \tag{1}$$

where $A^{I}_{\mathbf{k}j,\mathbf{k}j}$ is the spectral function projected on layer I in the representation of the primitive cell eigenstates $|\Psi_{\mathbf{k}j}\rangle$, $A_{\mathbf{K}J,\mathbf{K}J}$ is the heterostructure spectral function for the supercell eigenstates $|\Psi_{\mathbf{K}J}\rangle$ and $\sum_{\rho\in I} |\phi^{\rho}\rangle \langle \phi_{\rho}|$ projects onto the NGWFs of subsystem I.

Firstly, we investigate whether hBN encapsulation dilutes any of the desirable properties of monolayered BP, such as the direct band gap at Γ, or the carrier effective masses. This serves as a link with previous theoretical studies researching monolayer BP either encapsulated by hBN[51,52,54] or interfaced with other semiconducting materials,[55] while our large-scale approach has the added advantage of allowing the simulation of rotated interfaces. Therefore, our large, relatively unstrained, and rotated hBN/BP/hBN heterostructure provides a realistic representation of hBN-encapsulation. We compare the unfolded spectral function between the independent BP monolayer (Fig. 1.a) and the BP sheet in hBN/BP/hBN (Fig. 1.d). It is clear that no alterations occur to the DFT-predicted band gap (0.84 eV), as well as the shape of the bands near the Fermi level. We note that DFT is well known to significantly underestimate experimental band gap values,[8] but we are only interested in the relative band-gap changes in hBN/BP interfaces. The local density of states (LDOS) plot in Fig. 1.d clearly indicates that the hBN/BP/hBP interface has a type-I band-alignment, a conclusion which also applies to BP/hBN/BP (Fig. 1.e). For both hBN/BP/hBN and BP/hBN/BP (Fig. 1.d,e), the purple arrows highlight band discontinuities; these occur due to interactions between hBN and BP at energies where the hBN-associated DOS is non-zero, *i.e.* >0.5 eV below the valence band maximum at Γ. Therefore, we find that outer hBN sheets produce no

ill effects to the electronic structure of pristine monolayer BP. Moreover, it is experimentally known[32] that hBN substrates are usually free of charged impurities and provide shielding from unwanted scattering potentials, while hBN overlayers protect from interactions with air.[29,31,32] Consequently, we expect that hBN-encapsulation can have only beneficial implications for realistic BP layers, in agreement with the observed increased carrier mobilities,[56] durable and reliable transistor operation in ambient conditions,[31] or improved n-type charge transport.[29,32]

Secondly, we prove that BP/hBN/BP interfaces (Fig. 1.g) can be utilised as means of counteracting the near-IR to mid-IR light-absorption shift that occurs in multi-stacked BP sheets.[8] This effect is clearly observed as the transition from a 0.84 eV DFT-predicted band gap in monolayer BP (Fig. 1.a) to, for instance, a 0.34 eV gap in the AA-stacked bilayer BP (Fig. 1.c). To the best of our knowledge, the electronic-structure implications of such an arrangement have not been studied in any other theoretical work. We start by comparing the spectral function representation of the top BP sheet in the AA-stacked BP bilayer (Fig. 1.c) with the spectral function of the top BP sheet in the BP/hBN/BP heterostructure (Fig. 1.e).

Consequently, in BP/hBN/BP the interaction between the BP sheets is greatly reduced by the monolayer hBN spacer, as evident from the almost complete reversion to the monolayer bandstructure of BP. The spacer limits the degeneracy breaking between the states of two BP layers, significantly diminishing the splitting of the Γ-point conduction band (CB) valley and valence band (VB) peak in BP/hBN/BP compared to the BP AA-bilayer: 0.14 eV for the CB (0.49 eV for the bilayer) and 0.06 eV for the VB (0.68 for the bilayer). Most importantly, the band gap of the combined system is significantly increased, from 0.34 eV in the bilayer case, to 0.82 eV in the BP/hBN/BP heterostructure. This is explained by the presence of the hBN spacer, which significantly increases the distance between the BP layers, leading to a stringent reduction of their interlayer hybridisation. In essence, a BP bilayer is

transformed into a loosely interacting stack of BP monolayers. We thus demonstrate that single-layer BP and BP/hBN/BP interfaces share a similar band gap, a fact that could be used to greatly enhance the external quantum efficiency of BP optoelectronics in the near-IR spectrum, as the use of thicker hBN-spaced BP multilayers instead of thin BP monolayers would allow increased light absorption. We note that, even if the BP sheets of the bilayer and the BP/hBN/BP interface were chosen to be in the AA-stacking, our qualitative conclusions are generally applicable to any BP stacking in the case of hBN-spaced BP multilayers.

Furthermore, the remaining coupling between the BP layers in BP/hBN/BP, evident from the residual splitting of the valence band maximum (VBM) and conduction band minimum (CBM), can be effectively eliminated by using vertical electric fields. In Fig. 2 we plot the unfolded spectral function and density of states for BP/hBN/BP, but now we project on each BP monolayer separately (via Eq. 1) and analyse the results as a function of the applied vertical electric field.

In the zero-field state (Fig. 2.a), the band edges are clearly split into equally-weighted components in both the bottom and top layers, due to the symmetric degeneracy-breaking interactions. However, by applying even a relatively small vertical field of 0.7 V/nm (Fig. 2.b), the spectral weights of the intrusions of one layer into the other decrease dramatically. More specifically, in each split band-pair, the weight of one band decreases by a factor of $\approx$30, while the other doubles in weight, effects which are reversed in the other BP sheet, and which are more pronounced for larger fields (Fig. 2.c). In essence, the BP layers can be regarded as independent in the presence of vertical electric fields. In reality, substrate effects or impurities would possibly break the degeneracy between the BP films even without a field, achieving the same spectral weight redistribution effect. The hBN spacer vertically displaces the BP films, inducing a larger potential difference between them in a constant field. Thus, the band gap of the composite BP/hBN/BP system is highly tunable under electric fields, showing a relatively linear dependence with a slope of 0.29 eV nm/V; this tunability far

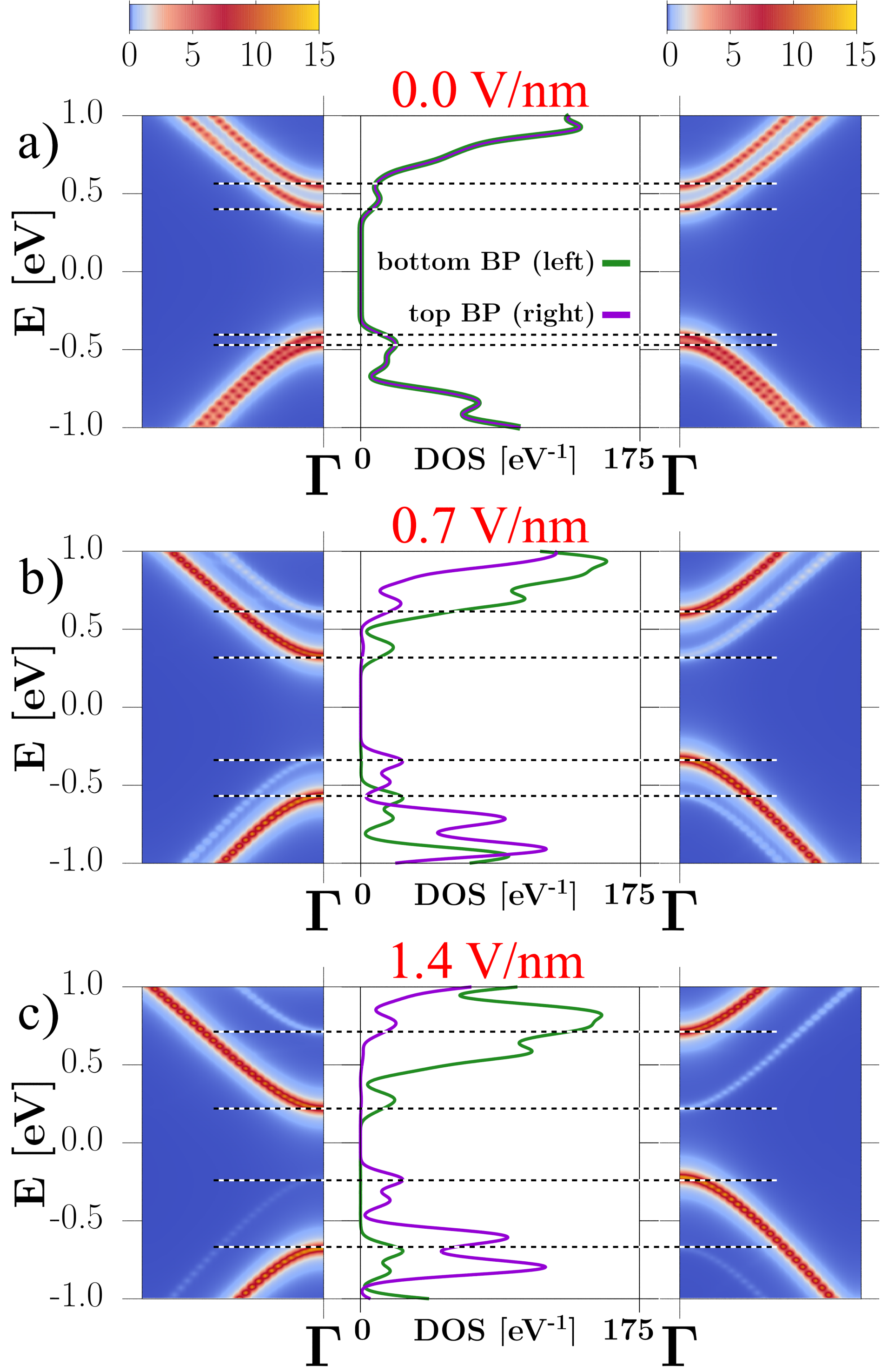

Figure 2: Spectral function representations along the Γ-Y direction for the bottom (left insets) and top (right insets) BP layers in BP/hBN/BP, under constant vertical electric fields: 0.0 V/nm (a), 0.7 V/nm (b), and 1.4 V/nm (c); white bands have significantly lower spectral weight than red bands. Center insets show the projected DOS on either BP layer. Dashed lines highlight the split-band edges.

exceeds what was previously encountered in pristine BP.[49] The field-induced rigid shifts in the projected density of states of each layer (Fig. 2 - center insets) clearly indicate a spatial separation of electrons and holes, with the former being on the bottom BP layer, while the latter are on the top BP layer.

This enhanced degree of control exerted by electric fields on the energy levels of the component BP layers warrants consideration of the BP/hBN/BP interface as a potential candidate for Tunneling Field Effect Transistor (TFET) architectures. A model arrangement is shown in Fig. 3.a, where the BP/hBN/BP heterostructure is sandwiched between a 4-layer hBN dielectric and a graphene gate on each side; gate voltages (V_{BG}, V_{TG}) control the carrier concentrations on each BP layer, while a bias voltage (V_{DS}) is also maintained between the BP films. More details can be found in the Supporting Information.

In conventional Metal-Oxide-Semiconductor Field Effect Transistors (MOSFETs), TMDC channels are preferred over BP channels, due to the significantly lower on/off current ratios identified in the latter,[6,10] owing to the smaller band gap of the multilayer BP currently employed in such devices. A smaller band gap implies that thermal excitations can more easily inject electrons from one layer into the other in the off-state, leading to a larger off-state current and implicitly worse switching capabilities. However, in a TFET arrangement (Fig. 3.a), the hBN barrier (Fig. 3.b) in between the BP layers strongly inhibits interlayer thermal carrier injection, leaving only quantum tunneling as a current-generation mechanism. Furthermore, the TFET concept is experiencing a rebirth in the context of 2D materials, as these allow the possibility of cleaner interfaces, without dislocations or trap states, the main causes for slower switching and increased background current in 3D TFETs. Consequently, our proposal of a BP/hBN/BP TFET shows promise due to the superior quality of 2D interfaces, as well as the possible reduction in thermally-induced currents.

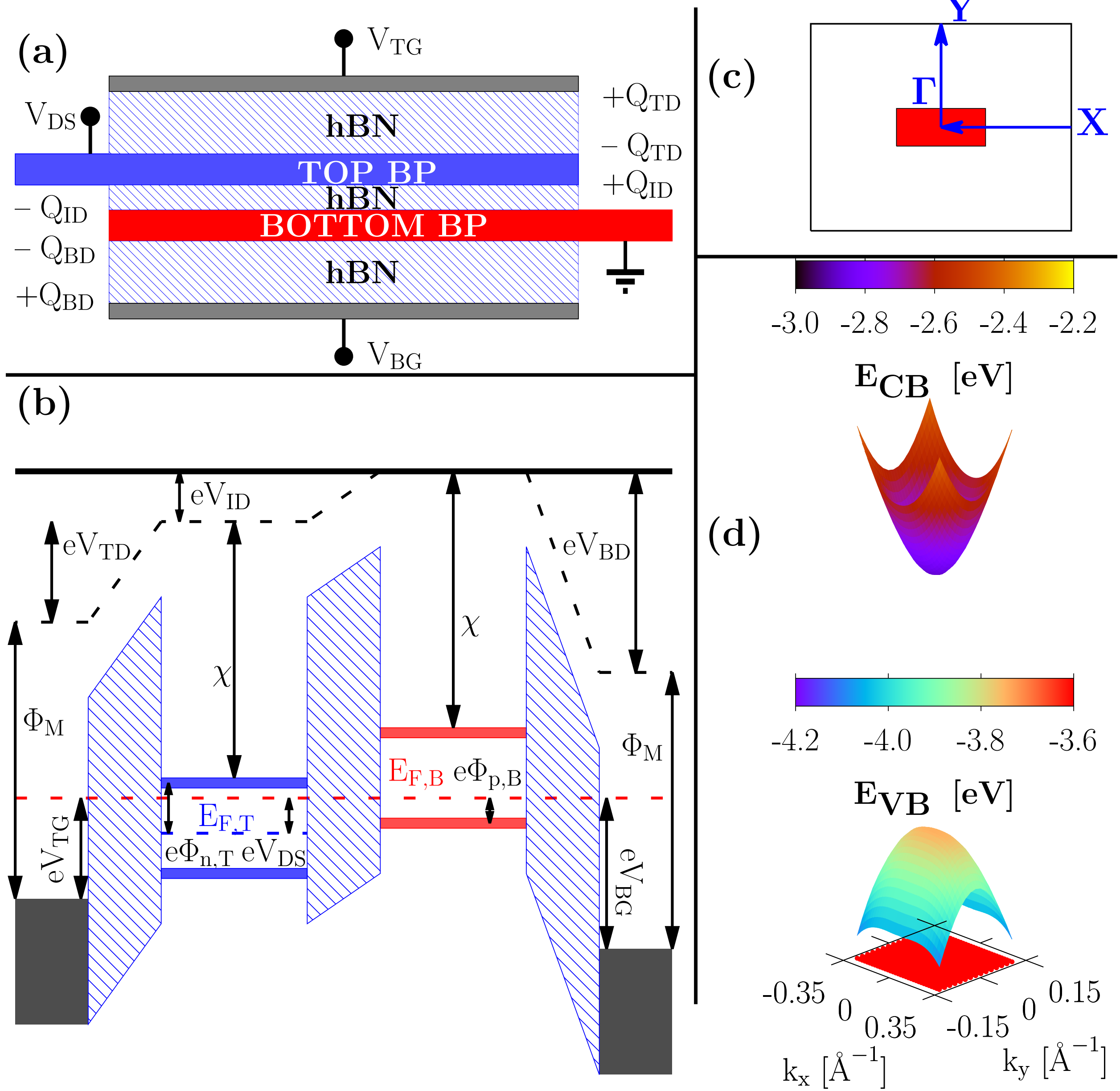

Figure 3: (a) 2D schematic of BP/hBN/BP transistor, also showing induced charges in the dielectrics; associated charges in BP have opposite sign. (b) Band diagram, with ϕ_M being the gate work-function, χ the electron affinity of BP, and $E_{F,T}$ and $E_{F,B}$ the Fermi levels to the top and bottom BP layers, respectively; $E_{F,B}$ taken as energy reference. (c) Region (red) in the BP primitive-cell Brillouin-zone that was sampled in calculating the energy dispersion relation, as well as the coupling matrix terms. (d) The BP energy levels used in our model (highest valence band and lowest conduction), as seen in the sampled reciprocal-space region.

We calculate the tunneling current in the spirit of the Bardeen formalism:[57]

$$\mathrm{I} = \frac{4\pi e}{\hbar} \sum_{i,j} \sum_{\mathbf{k}_B,\mathbf{k}_T} |\mathrm{O}_{ij}(\mathbf{k}_\mathcal{B},\mathbf{k}_\mathcal{T})|^2 \; \delta\Big[\mathrm{E}_{i\mathcal{B}}(\mathbf{k}_\mathcal{B}) - \mathrm{E}_{j\mathcal{T}}(\mathbf{k}_\mathcal{T})\Big] \; \Big(\mathrm{f}_{i\mathcal{B}} - \mathrm{f}_{j\mathcal{T}}\Big) \;, \tag{2}$$

where $\mathbf{k}_\mathcal{B}$ and $\mathbf{k}_\mathcal{T}$ are the in-plane momentum vectors in the bottom ($\mathcal{B}$) and top ($\mathcal{T}$) layers, and $\mathrm{E}_{i\mathcal{B}}(\mathbf{k}_\mathcal{B})$ are eigenvalues of the bottom layer eigenvectors i. The sum over eigenvectors only contains the highest valence band and the lowest conduction band for each individual BP layer (Fig. 3.d), and their description is restricted to the reciprocal-space region where there is only one peak (valley) in the valence (conduction) band, *i.e.* the red shaded area in Fig. 3.c. We ensure this constraint by choosing our drain-source and gate voltages such that all generated electrons and holes, as well as the BP Fermi levels are guaranteed to be at most 0.15 eV below the valence band maximum (VBM) and at most 0.30 eV above the conduction band minimum (CBM). This boundary is only valid when $|\mathrm{V_{ID}}|$-$\mathrm{E_G}$ (Fig. 3) is smaller than the minimum between the aforementioned VB depth and CB height. Moreover, having previously proven that electric fields fully decouple the BP layers in BP/hBN/BP (Fig. 3), we can indeed utilise the bandstructure of the independent BP monolayer. Previous works have employed parabolic dispersions, but here, with access to the full E-$\mathbf{k}$ relation via Kohn-Sham eigenvalues, we instead sample it explicitly in the relevant region of reciprocal space. The δ-function enforces tunneling between bands with the same energies; however, we emulate a realistic behaviour by using a Gaussian broadening σ_E. The $\mathrm{f}_{i\mathcal{B}}$ and $\mathrm{f}_{j\mathcal{T}}$ terms are bottom and top-layer Fermi factors corresponding to levels i and j, with energies $\mathrm{E}_{i\mathcal{B}}(\mathbf{k}_\mathcal{B})$ and $\mathrm{E}_{j\mathcal{T}}(\mathbf{k}_\mathcal{T})$, respectively; their role is to prevent transitions between similarly-occupied levels.

We note that both the shifts in energy and Fermi levels are implicitly dependent on the bias and gate voltages. Their positions can be exactly determined by solving a system of non-linear equations for each set of gate and bias voltages, as explained in detail in the Supporting Information. The coupling matrix term $\mathrm{O}_{ij}(\mathbf{k}_\mathcal{B},\mathbf{k}_\mathcal{T})$ in the Bardeen formalism[57]

represents the probability that a carrier can tunnel through a barrier from one momentum state into another:

$$\mathrm{O}_{ij}(\mathbf{k}_{\mathcal{B}},\mathbf{k}_{\mathcal{T}}) = \langle \tilde{\psi}_{i\mathcal{B}}(\mathbf{k}_{\mathcal{B}}) \mid \mathrm{V} \mid \tilde{\psi}_{j\mathcal{T}}(\mathbf{k}_{\mathcal{T}}) \rangle \ , \tag{3}$$

where V is the scattering barrier potential of hBN and $\tilde{\psi}_{j\mathcal{T}}(\mathbf{k}_{\mathcal{T}})/\tilde{\psi}_{i\mathcal{B}}(\mathbf{k}_{\mathcal{B}})$ are primitive-cell eigenvectors from the top and bottom BP layers, respectively. Similar to the expression for the spectral function representation, Eq. 3 was evaluated through an unfolding procedure, explained in the Supporting Information (SI). In the SI derivations, one can observe that Eq. 3 implicitly enforces strict in-plane momentum conservation. This condition can be relaxed through a Gaussian broadening of effective width $\sigma_{\mathbf{q}}$, the inverse of which ($\sigma_{\mathbf{q}}^{-1}$) is essentially a real-space coherence length. According to this definition, large coherence lengths are equivalent to small momentum-space Gaussian broadenings, which would induce sharper and larger current peaks upon resonant tunneling.

Previous studies on TFETs[34,58,59] have used simple analytical expressions for the barrier V. In this work, because the full system has been explicitly simulated via LS-DFT, we can utilise a more realistic form, obtained from the local (pseudo, Hartree and exchange-correlation) potentials. Therefore, we implicitly take into consideration the interlayer distances and atom corrugation patterns, both of which are accurately described by our chosen ab-initio nonlocal van-der-Walls functional. For these reasons, our coupling matrix terms, derived entirely from DFT-calculated quantities, can be expected to be more realistic than purely analytical expressions.

To make our simulations as relevant as possible to future experiments, we employ rather conservative parameters, indicative of an imperfect system. The presence of random impurities was implied by setting the band-energy broadening (σ_E) to 20 meV, room temperature was assumed unless specified, and the coherence length was set to a modest value of $\sigma_{\mathbf{q}}^{-1}$=25.4 nm, well within the range of coherence lengths encountered in other studies.[34] Surprisingly, even in these relatively suboptimal conditions, our simulated BP/hBN/BP TFET shows promise

in each of the three main operating regimes. For comparison purposes, in the Supporting Information we have also considered the cases where $\sigma_{\mathbf{q}}^{-1}$=6.4 nm and σ_E = 40 meV.

In the first mode (Fig. 4.a), we operate our TFET as a Zener diode, in reverse bias. The operating principle can be understood by observing the relative band alignments of the BP films, as shown in the miniature band diagrams in Fig. 4; the energy levels that can contribute to tunneling in the T=0 K limit are highlighted by darker shades of red (for the bottom BP) or blue (for the top BP). In this particular case, the current is generated by valence electrons from the top layer tunneling to unoccupied conduction states in the bottom layer, with both layers being p-doped. In this regime, the subthreshold swing (SS), defined as the change in gate voltage required to change the current by an order of magnitude, expresses the on/off switching capabilities. In MOSFETs, the thermal carrier injection mechanism limits SS to a minimum value of 60 mV/dec.[35] By contrast, the TFET architecture is not constrained by any lower bound in switching speed, as predicted for TMDC TFETs.[58] In fact, by calculating SS for different back gate voltages (right inset of Fig. 4.a), we predict a SS lower than the theoretical limit for MOSFETs, over a range of six orders of magnitude of current. This essentially implies that a BP/hBN/BP TFET could switch between the on and off state faster than any MOSFET in existence, thus showing outstanding potential for high speed electronics, even under our very conservative assumptions regarding the BP purity and lack of in-plane momentum conservation for tunneling. In our model, the SS is lower than 60 mV/dec only for currents as large as 10^{-5} A/μm^2, when the transistor is not fully on.

As a quantitative analysis on the effect of the coherence length and energy broadening for this mode, in Fig. S2 from the SI we show a comparison between different combinations of coherence lengths (σ_q^{-1}=25.4 nm and 6.4 nm) and energy broadenings (σ_E=20 meV and 40 meV). We find that the coherence length has no effect on the shape of the I-V characteristics or the SS, except for a reduction of the current by a factor of 10. However, a change from σ_E=20 meV to σ_E=40 meV has more significant effects: while the current values are

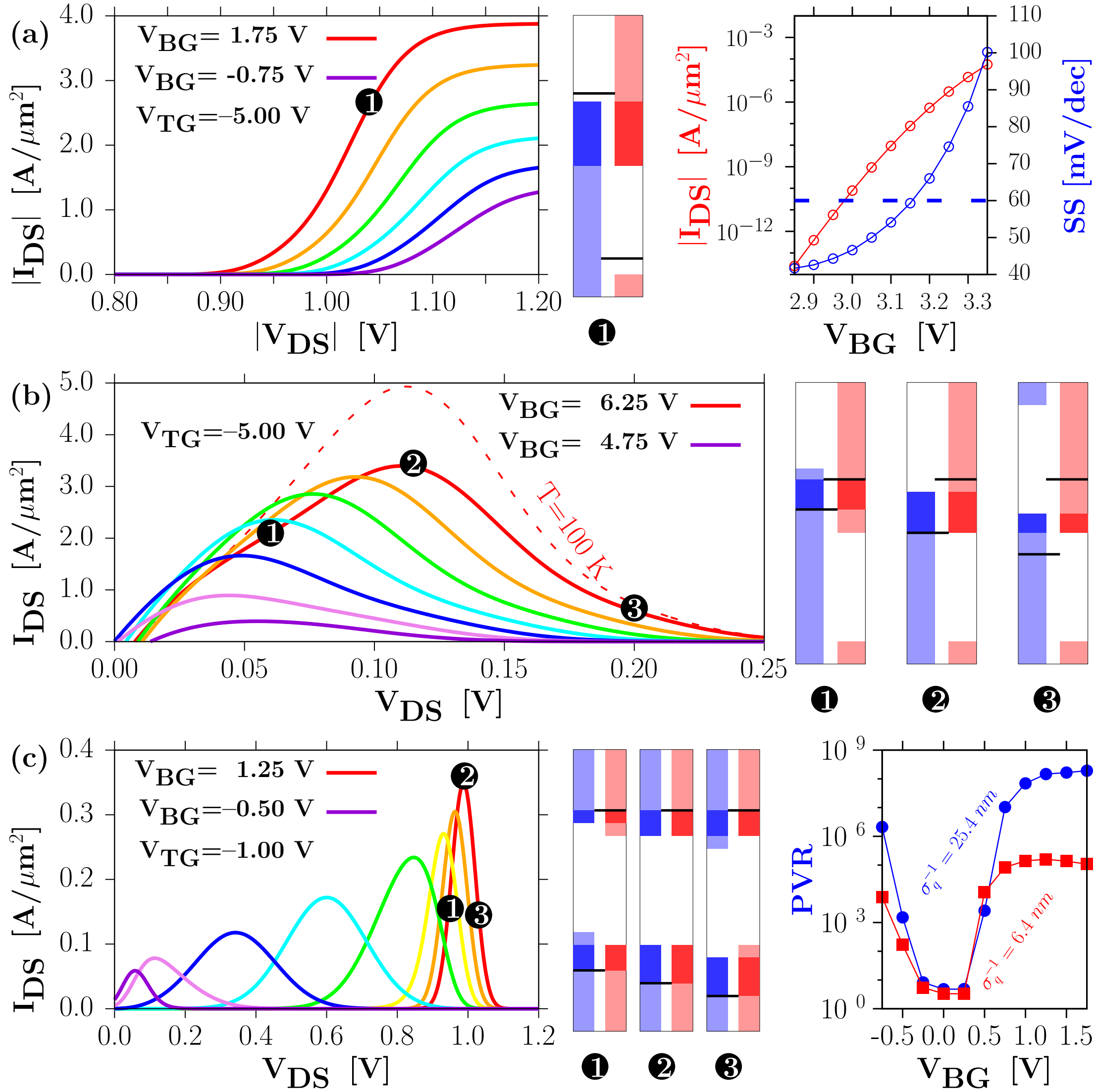

Figure 4: Operation modes in the BP/hBN/BP TFET. The top-gate voltage (V_{TG}) is fixed for each case, back-gate voltage (V_{BG}) is varied in steps of 0.50 V for (a) and 0.25 V for (b),(c). Band diagrams show energy levels of the top (blue) and bottom (red) BP layers – white suggests forbidden-energy gaps, darker red/blue highlights energy levels involved in tunneling at 0 K, black lines are the Fermi levels in each layer. (a) Left inset: current dependence on bias voltage for the reverse-bias/Zener regime. Right inset: current (red) and subthreshold swing (blue) dependence on the back-gate voltage, at fixed bias voltage V_{DS}=-0.6 V; dotted blue line shows the theoretical limit for SS in MOSFETs. (b) Left inset: negative differential resistance (NDR) peak for broken-gap arrangement, in forward bias. All solid I-V curves are at a temperature of 300 K, while the dotted one is at 100 K, with the same settings as the solid red line. (c) Left inset: NDR peak for the aligned-gap arrangement, at low gate voltages. Right inset: peak-to-valley ratio (PVR) for different back-gate voltages, for coherence lengths σ_q^{-1}=25.4 nm and σ_q^{-1}=6.4 nm; the valley current is measured at a bias voltage 0.2 V larger the NDR peak position, while σ_E = 20 meV throughout.

largely maintained, the I-V curves in Fig. S2 from the SI are broader, with less defined features and peaks. More importantly, the SS is much larger, with its minimum value being 144.4 mV/dec for $\sigma_E = 40$ meV, instead of 41.7 mV/dec as for $\sigma_E = 20$ meV. Thus, the switching speed is highly dependent on the impurity concentration in BP (linked to σ_E), but rather independent of the BP/hBN interface quality (linked to σ_q^{-1}).

The second operational mode (Fig. 4.b), is defined by a negative differential resistance (NDR) peak.[33] Such NDR peaks have applications in oscillatory circuits, memory devices, and even multi-valued logic.[60] This mode is accessible by changing the bias voltage in the previously discussed regime (Fig. 4.a) from reverse bias to forward bias. As before, the band diagrams (Fig. 4.b - right insets) indicate a broken gap arrangement, but now both BP films require degenerate doping for the NDR peak to occur, with the bottom (top) layer being of n (p) type. The tunneling carriers are clearly the conduction electrons in the bottom layer (dark red shade) and holes in the top layer (dark blue shade). For increasing bias, the energy levels of the top layer are essentially shifted down, until eventually one obtains a maximum current (point 2 in Fig. 4.b) when the top-BP Fermi level aligns with the bottom-BP conduction band minimum. For even higher bias, due to the forbidden energy gaps, the number of states which can tunnel is cut off, leading to a decrease in current (point 3 in Fig. 4.b).

Without the hBN interlayer barrier, carriers from the n-side conduction band would eventually start to be thermally injected into the p-side empty states. This effect would lead to a valley in the current, followed by an increase; this is a behaviour generally observed in Esaki diodes, such as a recent realisation of a 2D BP/$SnSe_2$ diode.[61] Such thermal excitations lead to a lower peak-to-valley ratio in devices exhibiting NDR. However, the hBN barrier shifts the thermally-driven current increase to higher forward-bias, thus extending the bias-voltage range over which the tunneling current can decrease, and increasing the obtainable peak-to-valley ratio. Note that if the temperature is lowered to 100 K, the current peak is more pronounced, due to the steepening of the Fermi factors in Eq. 2. Such a behaviour is in

perfect agreement with the experimentally-observed temperature dependence of NDR peaks in MoS_2/WSe_2 TFETs.[62] We note that this operation mode exhibits the same dependencies on the coherence length and energy broadening (see Fig. S3 in the SI) as described in the previous TFET regime. On the other hand, the drawback of this operation mode is the fact that large gate voltages are needed to enable the required degenerate doping of the BP layers and ensure a broken-gap arrangement, meaning that energy efficiency is not ensured.

However, the third operational mode (Fig. 4.c), also showing NDR behaviour, is highly energy efficient, as proven by the lower required gate voltages. Low-power operation can be achieved because the band-gaps of the same-material (BP) films now need to be aligned, not staggered, reducing the need for large bias or gate voltages. While the currents are an order of magnitude lower than in the previous regime, the position of the NDR peak is significantly more tunable under changes in gate voltage, a useful prospect for memory devices. The band diagrams of Fig. 4.c show that the current generation is two-fold, between the bottom conduction electrons and the top unoccupied conduction states, and from the bottom valence electrons to the top holes. By increasing the bias to the point where the bands are perfectly aligned (point 2 in Fig. 4.c), a maximum number of carriers which can tunnel is obtained, hence a maximum current. For even larger bias (point 3 in Fig. 4.c), however, while the number of tunneling-capable carriers is maintained, the current drops due to the momentum-mismatch caused by the band misalignment, *i.e.* decreasing coupling matrix terms $|\mathrm{O}_{ij}(\mathbf{k}_{\mathcal{B}}, \mathbf{k}_{\mathcal{T}})|$ between the top ($\mathbf{k}_{\mathcal{T}}$) and bottom ($\mathbf{k}_{\mathcal{B}}$) momentum vectors. Clearly, this momentum-mismatch induced current decrease would be most significant in TFETs with same-material films, such as our BP/hBN/BP case or previously studied TMDC/hBN/TMDC architectures,[34] and is highly dependent on the coherence length $\sigma_{\mathbf{q}}^{-1}$ (see SI). We do not compute any thermally-induced current in our model, meaning we cannot precisely pinpoint the value and position of the valley current associated with the current-voltage curves in Fig. 4.c. However, in the spirit of Campbell *et al.*[34] we define the valley current at a fixed offset of 0.2 V from the NDR peak; using this, we can obtain a measure of the peak-to-valley ratio (PVR) dependence on

the back-gate voltage.

As shown in the right inset of Fig. 4.c, there are two distinct means for large PVR to occur: either at low bias (V_{DS}), for back-gate voltages (V_{BG}) of the same polarity as the fixed top-gate voltage (V_{TG}=-1.00 V), or at higher bias for back-gate voltages of opposite polarity to the top-gate voltage. In the former case, one has the advantage of low required bias, implying an energy-efficient behaviour; however, the generated peak currents are also low. In the latter case, the peak currents, as well as the PVRs increase dramatically, but at the cost of larger gate bias voltages. Therefore, one has a choice between low-power/low-current and high-power/large-PVR operation, depending on the technological application.

An important observation is that our obtained PVRs are of the same order of magnitude ($\approx 2 \cdot 10^8$) as those predicted by Campbell *et al.*[34] (10^9) in TMDC/hBN/TMDC TFETs. These results imply that BP can be as useful as TMDCs in TFET applications, as opposed to the case of MOSFETs where BP layers exhibit on/off current ratios[10] three orders of magnitude smaller than TMDC films.[63] Furthermore, BP multilayers have the advantage of significantly smaller band gaps than TMDCs, meaning that only small voltage changes are needed to switch between the different operational modes presented in our work (Fig. 4.a,b,c).

As a last note, a change in coherence length produces profound changes in this operation mode, as shown in Fig. S4 from the SI. Firstly, if the coherence length is four times smaller (i.e. from 25.4 nm to 6.4 nm), the current decreases by a factor of 500, while the NDR peak current becomes more sensitive to the back-gate voltage. Such enhanced effects are explained by the increased momentum-coupling between alike bands (as in Fig. 4.c), as opposed to bands of different character (as in Fig. 4.a,b). Not only the current, but also the peak-to-valley ratios (PVRs) are highly dependent on the coherence length. For instance, upon a fourfold reduction in coherence length, the maximum PVR is reduced by three orders of magnitude, from $1.92 \cdot 10^8$ to $1.59 \cdot 10^5$. If subsequently the energy broadening is also increased from 20 meV to 40 meV, the maximum PVR is further reduced by another three

orders of magnitude, to $1.22 \cdot 10^2$. Therefore, both the current values and the peak-to-valley ratios encountered in the aligned-band NDR regime are highly sensitive to both the BP/hBN interface quality and the BP intrinsic purity.

In conclusion, we have shown that hBN encapsulation of BP does not worsen the desirable electronic properties compared to pristine monolayer BP. This offers the prospect of stable passivated flakes, allowing BP devices to function in ambient conditions.[29,31] Moreover, hBN spacers can be used to counteract the band gap reduction in stacked BP, thus improving the efficiency of BP optoelectronic devices in the near-IR range, and allowing more control over light absorption frequencies. Furthermore, based on our finding that the band gap of the BP/hBN/BP heterostructure is highly tunable under electric fields, we propose and simulate a tunneling field effect transistor built around such spaced BP bilayers. Lastly, different operational modes were identified for this transistor, with applications ranging from ultra-fast switching to low-power operation, or enormous peak-to-valley ratios as part of negative differential resistance regimes.

Acknowledgement

The authors acknowledge the support of the Winton Programme for the Physics of Sustainability. Computing resources were provided by the Darwin Supercomputer of the University of Cambridge High Performance Computing Service. GCC acknowledges the support of the Cambridge Trust European Scholarship. The underlying research data can be accessed via the following persistent URL: https://www.repository.cam.ac.uk/handle/1810/254019. The authors declare no competing financial interests.

Supporting Information Available

Convergence tests and specification of the computational methodology. Mathematical derivations for the coupling matrix terms, where the effective potential predicted by LS-DFT is used as scattering barrier. The implicit dependence of the energy and Fermi levels on the gate and bias voltages: obtaining the system of non-linear equations. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Yoon, Y.; Ganapathi, K.; Salahuddin, S. *Nano Lett.* **2011**, *11*, 3768–3773.

(2) Wang, Q. H.; Kalantar-Zadeh, K.; Kis, A.; Coleman, J. N.; Strano, M. S. *Nat. Nanotechnol.* **2012**, *7*, 699–712.

(3) Mak, K. F.; He, K.; Shan, J.; Heinz, T. F. *Nat. Nanotechnol.* **2012**, *7*, 494–498.

(4) Splendiani, A.; Sun, L.; Zhang, Y.; Li, T.; Kim, J.; Chim, C.-Y.; Galli, G.; Wang, F. *Nano Lett.* **2010**, *10*, 1271–1275.

(5) Kuc, A.; Zibouche, N.; Heine, T. *Phys. Rev. B* **2011**, *83*, 245213.

(6) Ling, X.; Wang, H.; Huang, S.; Xia, F.; Dresselhaus, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 4523–4530.

(7) Du, Y.; Ouyang, C.; Shi, S.; Lei, M. *J. Appl. Phys.* **2010**, *107*.

(8) Tran, V.; Soklaski, R.; Liang, Y.; Yang, L. *Phys. Rev. B* **2014**, *89*, 235319.

(9) Liu, H.; Neal, A. T.; Zhu, Z.; Luo, Z.; Xu, X.; Tománek, D.; Ye, P. D. *ACS Nano* **2014**, *8*, 4033–4041.

(10) Li, L.; Yu, Y.; Ye, G. J.; Ge, Q.; Ou, X.; Wu, H.; Feng, D.; Chen, X. H.; Zhang, Y. *Nat. Nanotechnol.* **2014**, *9*, 372–377.

(11) Qiao, J.; Kong, X.; Hu, Z.-X.; Yang, F.; Ji, W. *Nat. Commun.* **2014**, *5*.

(12) Xia, F.; Wang, H.; Jia, Y. *Nat. Commun.* **2014**, *5*.

(13) Hong, T.; Chamlagain, B.; Lin, W.; Chuang, H.-J.; Pan, M.; Zhou, Z.; Xu, Y.-Q. *Nanoscale* **2014**, *6*, 8978–8983.

(14) Wu, J.; Mao, N.; Xie, L.; Xu, H.; Zhang, J. *Angew. Chem. Int. Ed.* **2015**, *54*, 2366–2369.

(15) Baba, M.; Izumida, F.; Morita, A.; Koike, Y.; Fukase, T. *Jpn. J. Appl. Phys.* **1991**, *30*, 1753.

(16) Fei, .; Yang, L. *Nano Lett.* **2014**, *14*, 2884–2889.

(17) Padilha, J.; Fazzio, A.; da Silva, A. J. *Phys. Rev. Lett.* **2015**, *114*, 066803.

(18) Peng, X.; Wei, Q.; Copple, A. *Phys. Rev. B* **2014**, *90*, 085402.

(19) Das, S.; Demarteau, M.; Roelofs, A. *ACS Nano* **2014**, *8*, 11730–11738.

(20) Kamalakar, M. V.; Madhushankar, B. N.; Dankert, A.; Dash, S. P. *Small* **2015**, *11*, 2209–2216.

(21) Deng, Y.; Luo, Z.; Conrad, N. J.; Liu, H.; Gong, Y.; Najmaei, S.; Ajayan, P. M.; Lou, J.; Xu, X.; Ye, P. D. *ACS Nano* **2014**, *8*, 8292–8299.

(22) Buscema, M.; Groenendijk, D. J.; Blanter, S. I.; Steele, G. A.; van der Zant, H. S. J.; Castellanos-Gomez, A. *Nano Lett.* **2014**, *14*, 3347–3352.

(23) Engel, M.; Steiner, M.; Avouris, P. *Nano Lett.* **2014**, *14*, 6414–6417.

(24) Qin, G.; Yan, Q.-B.; Qin, Z.; Yue, S.-Y.; Cui, H.-J.; Zheng, Q.-R.; Su, G. *Sci. Rep.* **2014**, *4*.

(25) Fei, R.; Faghaninia, A.; Soklaski, R.; Yan, J.-A.; Lo, C.; Yang, L. *Nano Lett.* **2014**, *14*, 6393–6399.

(26) Cai, Y.; Ke, Q.; Zhang, G.; Zhang, Y.-W. *J. Phys. Chem. C* **2015**, *119*, 3102–3110.

(27) Liu, X.; Wood, J. D.; Chen, K.-S.; Cho, E.; Hersam, M. C. *J. Phys. Chem. Lett.* **2015**, *6*, 773–778.

(28) Island, O., J; Steele, G. A.; van-der Zant, H. S. J.; Castellanos-Gomez, A. *2D Mater.* **2015**, *2*, 011002.

(29) Doganov, R. A.; O'Farrell, E. C. T.; Koenig, S. P.; Yeo, Y.; Ziletti, A.; Carvalho, A.; Campbell, D. K.; Coker, D. F.; Watanabe, K.; Taniguchi, T.; Castro-Neto, A. H.; Özyilmaz, B. *Nat. Commun.* **2015**, *6*.

(30) Constantinescu, G. C.; Hine, N. D. M. *Phys. Rev. B* **2015**, *91*, 195416.

(31) Avsar, A.; Vera-Marun, I. J.; Tan, J. Y.; Watanabe, K.; Taniguchi, T.; Castro Neto, A. H.; Özyilmaz, B. *ACS Nano* **2015**, *9*, 4138–4145.

(32) Doganov, R. A.; Koenig, S. P.; Yeo, Y.; Watanabe, K.; Taniguchi, T.; Özyilmaz, B. *Appl. Phys. Lett.* **2015**, *106*, 083505.

(33) Esaki, L.; Tsu, R. *IBM J. Res. Dev.* **1970**, *14*, 61–65.

(34) Campbell, P. M.; Tarasov, A.; Joiner, C. A.; Ready, W. J.; Vogel, E. M. *ACS Nano* **2015**, *9*, 5000–5008.

(35) Ionescu, A. M.; Riel, H. *Nature* **2011**, *479*, 329–337.

(36) Withers, F.; Del Pozo-Zamudio, O.; Mishchenko, A.; Rooney, A. P.; Gholinia, A.; Watanabe, K.; Taniguchi, T.; Haigh, S. J.; Geim, A. K.; Tartakovskii, A. I.; Novoselov, K. S. *Nat. Mater.* **2015**, *14*, 301–306.

(37) Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. *J. Chem. Phys.* **2005**, *122*, 084119.

(38) Wilkinson, K. A.; Hine, N. D. M.; Skylaris, C.-K. *J. Chem. Theory Comput.* **2014**, *10*, 4782.

(39) Skylaris, C.-K.; Haynes, P. D. *J. Chem. Phys.* **2007**, *127*, 164712.

(40) Haynes, P. D.; Skylaris, C.-K.; Mostofi, A. A.; Payne, M. C. *Phys. Status Solidi B* **2006**, *243*, 2489–2499.

(41) Haynes, P. D.; Skylaris, C. K.; Mostofi, A. A.; Payne, M. C. *J. Phys.: Condens. Matter* **2008**, *20*, 294207.

(42) Blöchl, P. E. *Phys. Rev. B* **1994**, *50*, 17953–17979.

(43) Garrity, K. F.; Bennett, J. W.; Rabe, K. M.; Vanderbilt, D. *Comput. Mater. Sci.* **2014**, *81*, 446–452.

(44) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; et al., *J. Phys.: Condens. Matter* **2009**, *21*, 395502.

(45) Pease, R. S. *Nature* **1950**, *165*, 722–723.

(46) Warner, J. H.; Rümmeli, M. H.; Bachmatiuk, A.; Büchner, B. *ACS Nano* **2010**, *4*, 1299–1304.

(47) Brown, A.; Rundqvist, S. *Acta Crystallogr.* **1965**, *19*, 684–685.

(48) Constantinescu, G.; Kuc, A.; Heine, T. *Phys. Rev. Lett.* **2013**, *111*, 036104.

(49) Dai, J.; Zeng, X. C. *J. Phys. Chem. Lett.* **2014**, *5*, 1289–1293.

(50) Hermann, K. *J. Phys.: Condens. Matter* **2012**, *24*, 314210.

(51) Cai, Y.; Zhang, G.; Zhang, Y.-W. *J. Phys. Chem. C* **2015**, *119*, 13929–13936.

(52) Rivero, P.; Horvath, C. M.; Zhu, Z.; Guan, J.; Tománek, D.; Barraza-Lopez, S. *Phys. Rev. B* **2015**, *91*, 115413.

(53) Lee, C.-C.; Yukiko, Y.-T.; Taisuke, O. *J. Phys.: Condens. Matter* **2013**, *25*.

(54) Hu, T.; Hong, J. *ACS Appl. Mater. Interfaces* **2015**, *7*, 23489–23495.

(55) Guo, H.; Lu, N.; Dai, J.; Wu, X.; Zeng, X. C. *J. Phys. Chem. C* **2014**, *118*, 14051–14059.

(56) Gillgren, N.; Wickramaratne, D.; Shi, Y.; Espiritu, T.; Yang, J.; Hu, J.; Wei, J.; Liu, X.; Mao, Z.; Watanabe, K.; Taniguchi, T.; Bockrath, M.; Barlas, Y.; Lake, R. K.; Lau, C. N. *2D Mater.* **2015**, *2*, 011001.

(57) Bardeen, J. *Phys. Rev. Lett.* **1961**, *6*, 57–59.

(58) Li, M.; Esseni, D.; Snider, G.; Jena, D.; Grace Xing, H. *J. Appl. Phys.* **2014**, *115*, 074508.

(59) Britnell, L.; Gorbachev, R. V.; Jalil, R.; Belle, B. D.; Schedin, F.; Mishchenko, A.; Georgiou, T.; Katsnelson, M. I.; Eaves, L.; Morozov, S. V.; Peres, N. M. R.; Leist, J.; Geim, A. K.; Novoselov, K. S.; Ponomarenko, L. A. *Science* **2012**, *335*, 947–950.

(60) Chan, H.; Mohan, S.; Mazumder, P.; Haddad, G. *IEEE J. Solid-State Circuits* **1996**, *31*, 1151–1156.

(61) Rusen, Y.; Sara, F.; Yimo, H.; Bo, S.; Shudong, X.; Mingda, L.; Nan, M.; Vladimir, P.; David, A. M.; Debdeep, J.; Huili, G. X. *Nano Lett.* **2015**, *15*, 5791–5798.

(62) Tania, R.; Mahmut, T.; Xi, C.; Hui, F.; Der-Hsien, L.; Peida, Z.; Yu-Ze, C.; Yu-Lun, C.; Jing, G.; Ali, J. *ACS Nano* **2015**, *9*, 2071–2079.

(63) Wu, W.; De, D.; Chang, S.-C.; Wang, Y.; Peng, H.; Bao, J.; Pei, S.-S. *Appl. Phys. Lett.* **2013**, *102*.

Graphical TOC Entry