

Single particle transport in two-dimensional heterojunction interlayer tunneling field effect transistor

Mingda (Oscar) Li,^{1,a)} David Esseni,² Gregory Snider,¹ Debdeep Jena,¹ and Huili Grace Xing^{1,b)}

¹University of Notre Dame, Notre Dame, Indiana 46556, USA

²University of Udine, Udine, Italy

(Received 9 December 2013; accepted 6 February 2014; published online 21 February 2014)

The single particle tunneling in a vertical stack consisting of monolayers of two-dimensional semiconductors is studied theoretically, and its application to a novel Two-dimensional Heterojunction Interlayer Tunneling Field Effect Transistor (Thin-TFET) is proposed and described. The tunneling current is calculated by using a formalism based on the Bardeen's transfer Hamiltonian, and including a semi-classical treatment of scattering and energy broadening effects. The misalignment between the two 2D materials is also studied and found to influence the magnitude of the tunneling current but have a modest impact on its gate voltage dependence. Our simulation results suggest that the Thin-TFETs can achieve very steep subthreshold swing, whose lower limit is ultimately set by the band tails in the energy gaps of the 2D materials produced by energy broadening. The Thin-TFET is thus very promising as a low voltage, low energy solid state electronic switch. © 2014 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4866076>]

I. INTRODUCTION

The electronic integrated circuits are the hardware backbone of today's information society, and the power dissipation has recently become the greatest challenge, affecting the lifetime of existing portable equipments, the sustainability of large and growing in number data centers, and the feasibility of energy autonomous systems for ambience intelligence,^{1,2} and of sensor networks for implanted monitoring and actuation medical devices.³ While the scaling of the supply voltage, V_{DD} , is recognized as the most effective measure to reduce switching power in digital circuits, the performance loss and increased device to device variability are a serious hindrance to the V_{DD} scaling down to 0.5 V or below.

The voltage scalability of VLSI systems may be significantly improved by resorting to innovations in the transistor technology and, in this regard, the ITRS has singled out Tunnel field effect transistors (FETs) as the most promising transistors to reduce the sub-threshold swing (SS) below the 60 mV/dec limit of MOSFETs (at room temperature), and, thus, to enable a further V_{DD} scaling.^{4,5} Several device architectures and materials are being investigated to develop Tunnel FETs offering both an attractive on current and a small SS, including III–V based transistors, possibly employing staggered or broken bandgap heterojunctions,^{6–9} or strain engineering.¹⁰ Even if encouraging experimental results have been reported for the on-current in III–V Tunnel FETs, to achieve a sub 60 mV/dec subthreshold swing is still a real challenge in these devices, probably due to the detrimental effects of interface states.^{6,11,12} Therefore, as of today, the investigation of new material systems and innovative device architectures for high performance Tunnel FETs is a timely

research field in both the applied physics and the electron device community.

In such a contest, two-dimensional (2D) crystals attract increasingly more attention primarily due to their scalability, step-like density of states, and absence of broken bonds at interface. They can be stacked to form a new class of tunneling transistors based on an interlayer tunneling occurring in the direction normal to the plane of the 2D materials. In fact, tunneling and resonant tunneling devices have been recently proposed,¹³ as well as experimentally demonstrated for graphene-based transistors.^{14,15} Furthermore, monolayers of group-VIB transition metal dichalcogenides MX_2 ($M = Mo, W$; $X = S, Se, Te$) have recently attracted remarkable attention for their electronic and optical properties.^{16,17} Monolayers of transition-metal dichalcogenides (TMDs) have a bandgap varying from almost zero to 2 eV with a sub-nanometer thickness, such that these materials can be considered approximately as two-dimensional crystals.¹⁸ The sub-nanometer thickness of TMDs can provide excellent electrostatic control in a vertically stacked heterojunction. Furthermore, the 2D nature of such materials make them essentially immune to the energy bandgap increase produced by the vertical quantization when conventional 3D semiconductors are thinned to a nanoscale thickness and, thus, immune to the corresponding degradation of the tunneling current density.¹⁹ Moreover, the lack of dangling bonds at the surface of TMDs may allow for the fabrication of material stacks with low densities of interface defects,¹⁹ which is another potential advantage of TMDs materials for Tunnel FETs applications.

In this paper, we propose a two-dimensional heterojunction interlayer tunneling field effect transistor (Thin-TFET) based on 2D semiconductors and develop a transport model based on the transfer-Hamiltonian method to describe the current voltage characteristics and discuss, in particular, the

^{a)}E-mail: mli7@nd.edu

^{b)}E-mail: hxing@nd.edu

subthreshold swing. In Sec. II, we first present the device concept and illustrate examples of the vertical electrostatic control, then we develop a formalism to calculate the tunneling current. Upon realizing that the subthreshold swing of the Thin-TFET is ultimately determined by the energy broadening, in Sec. II C, we show how this important physical factor has been included in our calculations. In Sec. II D, we address the effect of a possible misalignment between the two 2D semiconductor layers, while in Sec. II E, we derive some approximated, analytical expressions for the tunneling current density, which are useful to gain insight in the transistor operation and to guide the device design. In Sec. III, we present the results of numerically calculated current voltage characteristics for the Thin-TFET, and, finally, in Sec. IV, we draw some concluding remarks about the modeling approach developed in the paper and about the design perspectives for the Thin-TFET.

II. MODELING OF THE TUNNELING TRANSISTOR

A. Device concept and electrostatics

The device structure and the corresponding band diagram are sketched in Fig. 1, where the 2D materials are assumed to be semiconductors with sizable energy bandgap, for example, TMD semiconductors without losing generality.^{17,20} Both the top 2D and the bottom 2D material is a monolayer, and the thickness of the 2D layers is neglected in the modeling of the electrostatics.

The working principle of the tunneling transistor sketched in Fig. 1(a) can be explained as follows. When the conduction band edge E_{CT} of the top 2D layer is higher than the valence band edge E_{VB} of the bottom 2D layer (see Fig. 2(a)), there are no states in the top layer to which the electrons of the bottom layer can tunnel into. This corresponds to the off state of the device. When E_{CT} is pulled below E_{VB} (see Fig. 2(b)), a tunneling window is formed, and, consequently, an interlayer tunneling can flow from the bottom to the top 2D material. The crossing and uncrossing between the top layer conduction band and the bottom layer valence band are governed by the gate voltages, and they are described by the electrostatics of the device.

To calculate the band alignment along the vertical direction of the intrinsic device in Fig. 1, we write the Gauss law linking the sheet charge in the 2D materials to the electric fields in the surrounding insulating layers, which leads to

$$\begin{aligned} C_{TOX}V_{TOX} - C_{IOX}V_{IOX} &= e(p_T - n_T + N_D), \\ C_{BOX}V_{BOX} + C_{IOX}V_{IOX} &= e(p_B - n_B + N_A), \end{aligned} \quad (1)$$

where $C_{T(B)OX}$ is the capacitance per unit area of top oxide (interlayer, back oxide) and $V_{T(B)OX}$ is the potential drop across top oxide (interlayer, back oxide). The potential drop across the oxides can be written in terms of the external voltages V_{TG} , V_{BG} , V_{DS} , and of the energy $e\phi_{n,T} = E_{CT} - E_{FT}$ and $e\phi_{p,T} = E_{FB} - E_{VB}$ defined in Fig. 1(b) as

$$\begin{aligned} eV_{TOX} &= eV_{TG} + e\phi_{n,T} - eV_{DS} + \chi_{2D,T} - \Phi_{M,T}, \\ eV_{BOX} &= eV_{BG} - e\phi_{p,B} + E_{GB} + \chi_{2D,B} + \Phi_{M,B}, \\ eV_{IOX} &= eV_{DS} - e\phi_{p,B} - e\phi_{n,T} + E_{GB} + \chi_{2D,B} - \chi_{2D,T}, \end{aligned} \quad (2)$$

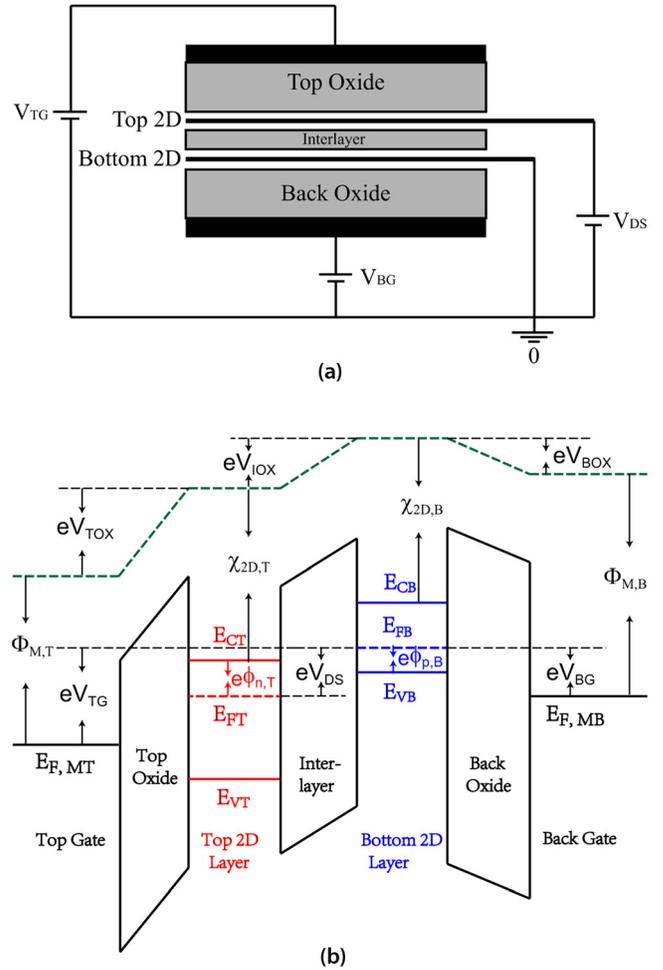


FIG. 1. (a) Schematic device structure for the Thin-TFET, where V_{TG} , V_{BG} , and V_{DS} are the top gate, back gate, and drain to source voltages; (b) sketch of the band diagram, where $\Phi_{M,T}$, $\Phi_{M,B}$ are the work-functions and $E_{F,MT}$, $E_{F,MB}$ the Fermi levels of the metal gates, while $\chi_{2D,T}$, $\chi_{2D,B}$ are the electron affinities, E_{FT} , E_{FB} the Fermi levels, E_{CT} , E_{CB} the conduction band edges, and E_{VT} , E_{VB} the valence band edges respectively in the top and bottom 2D layer. V_{TOX} , V_{IOX} and V_{BOX} are the potential drops respectively across the top oxide, interlayer and back oxide.

where E_{FT} , E_{FB} are fermi levels of majority carriers in the top and bottom layer. n_T , p_T are the electron and hole concentration in the top layer, n_B , p_B the concentrations in bottom layer, $\chi_{2D,T}$, $\chi_{2D,B}$ are the electron affinities of the 2D materials, Φ_T , Φ_B the workfunctions of the top and back gate, and E_{GB} is the energy gap in the bottom layer. Eq. (2) implicitly assumes that the majority carriers of the two 2D materials are at thermodynamic equilibrium with their Fermi levels, with the split of the Fermi levels set by the external voltages (i.e., $E_{FB} - E_{FT} = eV_{DS}$), and the electrostatic potential essentially constant in the 2D layers.

Since in our numerical calculations, we shall employ a parabolic effective mass approximation for the energy dispersion of the 2D materials, as discussed more thoroughly in Sec. III, the carrier densities can be readily expressed as an analytic function of $e\phi_{n,T}$ and $e\phi_{p,B}$ ²¹

$$n(p) = \frac{g_v m_c (m_v) k_B T}{\pi \hbar^2} \ln \left[\exp \left(-\frac{q\phi_{n,T}(\phi_{p,B})}{k_B T} \right) + 1 \right], \quad (3)$$

where g_v is the valley degeneracy.

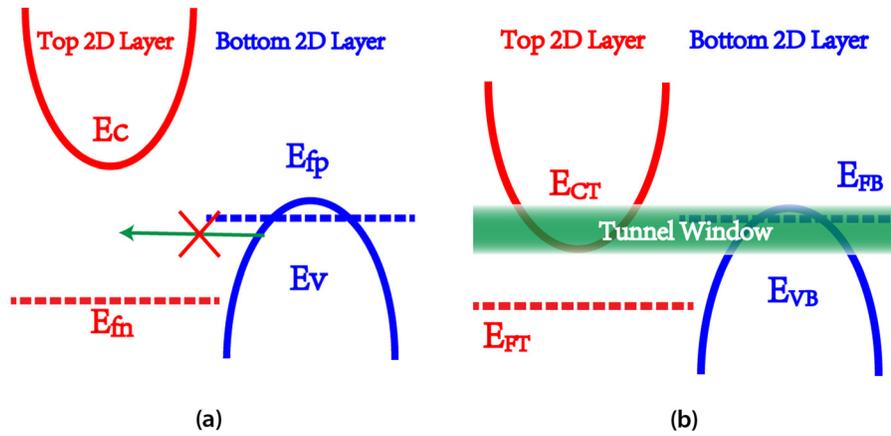


FIG. 2. Sketch of the band alignments in a Thin-TFET between the top and bottom 2D layer in (a) OFF state and (b) ON state.

When Eqs. (2) and (3) are inserted in Eq. (1), we obtain two algebraic equations for $\phi_{n,T}$ and $\phi_{p,B}$ that can be solved numerically and describe the electrostatics in a one dimensional section of the device.

B. Transport model

In this section, we develop a formalism to calculate the tunneling current based on the transfer-Hamiltonian method,^{22–24} as also revisited recently for resonant tunneling in graphene transistors.^{13,14,25} We start by writing the single particle elastic tunneling current as

$$I = g_v \frac{4\pi e}{\hbar} \sum_{\mathbf{k}_T, \mathbf{k}_B} |M(\mathbf{k}_T, \mathbf{k}_B)|^2 \delta(E_B(\mathbf{k}_B) - E_T(\mathbf{k}_T)) (f_B - f_T) \quad (4)$$

where e is the elementary charge, \mathbf{k}_B , \mathbf{k}_T are the wave-vectors, respectively, in the bottom and top 2D material, $E_B(\mathbf{k}_B)$ and $E_T(\mathbf{k}_T)$ denote the corresponding energies, f_B and f_T are the Fermi occupation functions in the bottom and top layer (depending respectively on E_{FB} and E_{FT} , see Fig. 1), and g_v is the valley degeneracy. The matrix element $M(\mathbf{k}_T, \mathbf{k}_B)$ expresses the transfer of electrons between the two 2D layers is given by¹⁴

$$M(\mathbf{k}_T, \mathbf{k}_B) = \int_A d\mathbf{r} \int dz \psi_{T, \mathbf{k}_T}^\dagger(\mathbf{r}, z) U_{sc}(\mathbf{r}, z) \psi_{B, \mathbf{k}_B}(\mathbf{r}, z), \quad (5)$$

where ψ_{B, \mathbf{k}_B} (ψ_{T, \mathbf{k}_T}) is the electron wave-function in the bottom (top) 2D layer and $U_{sc}(\mathbf{r}, z)$ is the perturbation potential in the interlayer region.

Eq. (5) acknowledges the fact that in real devices several physical mechanisms occurring in the interlayer region can result in a relaxed conservation of the in plane wave-vector \mathbf{k} in the tunneling process. We will return to the discussion of $U_{sc}(\mathbf{r}, z)$ in this section.

To proceed in the calculation of $M(\mathbf{k}_T, \mathbf{k}_B)$, we write the electron wave-function in the Bloch function form as

$$\psi_{\mathbf{k}}(\mathbf{r}, z) = \frac{1}{\sqrt{N_C}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}, z), \quad (6)$$

where $u_{\mathbf{k}}(\mathbf{r}, z)$ is a periodic function of \mathbf{r} and N_C is the number of unit cells in the overlapping area A of the two 2D materials. Eq. (6) assumes the following normalization condition:

$$\int_{\Omega_C} d\rho \int_z dz |u_{\mathbf{k}}(\rho, z)|^2 = 1, \quad (7)$$

where ρ is the in-plane abscissa in the unit cell area Ω_C and $A = N_C \Omega_C$.

The wave-function $\psi_{\mathbf{k}}(\mathbf{r}, z)$ is assumed to decay exponentially in the interlayer region with a decay constant κ ;^{13,14} such a z dependence is absorbed in $u_{\mathbf{k}}(\mathbf{r}, z)$, and we do not need to make it explicit in our derivations. It should be noticed that absorbing the exponential decay in $u_{\mathbf{k}}(\mathbf{r}, z)$ recognizes the fact that in the interlayer region the \mathbf{r} dependence of the wave-function changes with z . In fact, as already discussed,¹³ while the $u_{\mathbf{k}}(\mathbf{r}, z)$ are localized around the basis atoms in the two 2D layers, these functions are expected to spread out while they decay in the interlayer region, so that the \mathbf{r} dependence becomes weaker when moving farther from the 2D layers.

To continue in the calculation of $M(\mathbf{k}_T, \mathbf{k}_B)$, we let the scattering potential in the interlayer region be separable in the form¹⁴

$$U_{sc}(\mathbf{r}, z) = V_B(z) F_L(\mathbf{r}), \quad (8)$$

where $F_L(\mathbf{r})$ is the in-plane fluctuation of the scattering potential, which is essentially responsible for the relaxation of momentum conservation in the tunneling process.

By substituting Eqs. (6) and (8) in Eq. (5) and writing $\mathbf{r} = \mathbf{r}_j + \rho$, where \mathbf{r}_j is a direct lattice vector and ρ is the in-plane position inside each unit cell, we obtain

$$M(\mathbf{k}_T, \mathbf{k}_B) = \frac{1}{N_C} \sum_{j=1}^{N_C} e^{i(\mathbf{k}_B - \mathbf{k}_T) \cdot \mathbf{r}_j} \int_{\Omega_C} d\rho \int dz e^{i(\mathbf{k}_B - \mathbf{k}_T) \cdot \rho} \times u_{T, \mathbf{k}_T}^\dagger(\mathbf{r}_j + \rho, z) F_L(\mathbf{r}_j + \rho) V_B(z) u_{B, \mathbf{k}_B}(\mathbf{r}_j + \rho, z). \quad (9)$$

We now assume that $F_L(\mathbf{r})$ corresponds to relatively long range fluctuations, so that it can be taken as approximately constant inside a unit cell, and that, furthermore, the top and bottom 2D layer have the same lattice constant, hence the Bloch functions u_{T, \mathbf{k}_T} and u_{B, \mathbf{k}_B} have the same periodicity in the \mathbf{r} plane. Moreover, for the time being, we consider that the conduction band minimum in the top layer and the valence band maximum in the bottom layer are at the same point of the 2D Brillouin zone, so that $\mathbf{q} = \mathbf{k}_B - \mathbf{k}_T$ is small

compared to the size of the Brillouin zone and $e^{iq\rho}$ is approximately 1.0 inside a unit cell. These considerations and approximations allow us to rewrite Eq. (9) as

$$M(\mathbf{k}_T, \mathbf{k}_B) \simeq \frac{1}{N_C} \sum_{j=1}^{N_C} e^{iq\mathbf{r}_j} F_L(\mathbf{r}_j) \int_{\Omega_C} d\rho \int dz u_{T, \mathbf{k}_T}^\dagger(\rho, z) \times V_B(z) u_{B, \mathbf{k}_B}(\rho, z), \quad (10)$$

where the integral in the unit cell has been written for $\mathbf{r}_j = \mathbf{0}$ because it is independent of the unit cell.

Consistent with the assumption that \mathbf{k}_B and \mathbf{k}_T are small compared to the size of the Brillouin zone, in Eq. (10), we neglect the \mathbf{k}_B (\mathbf{k}_T) dependence of u_{B, \mathbf{k}_B} (u_{T, \mathbf{k}_T}) and simply set $u_{T, \mathbf{k}_T}(\rho, z) \approx u_{0T}(\rho, z)$, $u_{B, \mathbf{k}_B}(\rho, z) \approx u_{0B}(\rho, z)$, where $u_{0T}(\rho, z)$ and $u_{0B}(\rho, z)$ are the periodic parts of the Bloch function at the band edges, which is the simplification typically employed in the effective mass approximation approach.²¹ By recalling that the u_{0B} and u_{0T} retain the exponential decay of the wave-functions in the interlayer region with a decay constant κ , we now write

$$\int_{\Omega_C} d\rho \int dz u_{0T}^\dagger(\rho, z) V_B(z) u_{0B}(\rho, z) \simeq M_{B0} e^{-\kappa T_{IL}}, \quad (11)$$

where T_{IL} is the interlayer thickness and M_{B0} is a \mathbf{k} independent matrix element that will remain a prefactor in the final expression for the tunneling current. Since $F_L(\mathbf{r})$ has been assumed a slowly varying function over a unit cell, then the sum over the unit cells in Eq. (10) can be rewritten as a normalized integral over the tunneling area

$$\frac{1}{\Omega_C N_C} \sum_{j=1}^{N_C} \int_{\Omega_C} e^{iq\mathbf{r}_j} F_L(\mathbf{r}_j) \simeq \frac{1}{A} \int_A e^{iq\mathbf{r}} F_L(\mathbf{r}) d\mathbf{r}. \quad (12)$$

By introducing Eqs. (11) and (12) in Eq. (10), we can finally express the squared matrix element as

$$|M(\mathbf{k}_T, \mathbf{k}_B)|^2 \simeq \frac{|M_{B0}|^2 S_F(\mathbf{q})}{A} e^{-2\kappa T_{IL}}, \quad (13)$$

where $\mathbf{q} = \mathbf{k}_B - \mathbf{k}_T$ and $S_F(\mathbf{q})$ is the power spectrum of the random fluctuation described by $F_L(\mathbf{r})$, which is defined as²¹

$$S_F(\mathbf{q}) = \frac{1}{A} \left| \int_A e^{iq\mathbf{r}} F_L(\mathbf{r}) d\mathbf{r} \right|^2. \quad (14)$$

By substituting Eq. (13) in Eq. (4) and then converting the sums over \mathbf{k}_B and \mathbf{k}_T to integrals, we obtain

$$I = \frac{g_v e |M_{B0}|^2 A}{4\pi^3 \hbar} e^{-2\kappa T_{IL}} \int_{\mathbf{k}_T} \int_{\mathbf{k}_B} d\mathbf{k}_T d\mathbf{k}_B S_F(\mathbf{q}) \delta(E_B(\mathbf{k}_B) - E_T(\mathbf{k}_T)) (f_B - f_T). \quad (15)$$

Before we proceed with some important integrations of the basic model that will be discussed in Secs. II C and II D, a few comments about the results obtained so far are in order below.

According to Eq. (15), the current is proportional to the squared matrix element $|M_{B0}|^2$ defined in Eq. (11) and

decreases exponentially with the thickness interlayer T_{IL} according to the decay constant κ of the wave-functions. Attempting to derive a quantitative expression for M_{B0} is admittedly very difficult, in fact, it is difficult to determine how the periodic functions $u_{0T}(\rho, z)$ and $u_{0B}(\rho, z)$ spread out when they decay in the barrier region and, furthermore, it is not even perfectly clear what potential energy or Hamiltonian should be used to describe the barrier region itself, which is an issue already recognized and thoroughly discussed in the literature, since a long time.²⁴ Our model essentially circumvents these difficulties by resorting to the semi-empirical formulation of the matrix element given by Eq. (11), where M_{B0} is left as a parameter to be determined and discussed by comparing to experiments.

It is also worth noting that in our calculations, we have not explicitly discussed the effect of spin-orbit interaction in the bandstructure of 2D materials, even if giant spin-orbit couplings have been reported in 2D transition-metal dichalcogenides.²⁶ If the energy separations between the spin-up and spin-down bands are large, then the spin degeneracy in current calculations should be one instead of two, which would affect the current magnitude but not its dependence on the gate bias. Our calculations neglected also the possible modifications of band structure in the TMD materials produced by the vertical electrical field, in fact, we believe that in our device, the electrical field in the 2D layers is not strong enough to make such effects significant.²⁷

The decay constant κ in the interlayer region may be estimated from the electron affinity difference between the 2D layers and the interlayer material.¹³ Moreover, according to Eq. (15) the constant κ determines the dependence of the current on T_{IL} , so that κ may be extracted by comparing to experiments discussing such a dependence, which, for example, have been recently reported for the interlayer tunneling current in a graphene-*h*BN system.¹⁵

As for the spectrum $S_F(\mathbf{q})$ of the scattering potential, in our calculations, we utilize

$$S_F(\mathbf{q}) = \frac{\pi L_C^2}{(1 + q^2 L_C^2 / 2)^{3/2}}, \quad (16)$$

where $q = |\mathbf{q}|$ and L_C is the correlation length, which in our derivations has been assumed large compared to the size of a unit cell. Eq. (16) is consistent with an exponential form for the autocorrelation function of $F_L(\mathbf{r})$,²¹ and a similar \mathbf{q} dependence has been recently employed to reproduce the experimentally observed line-width of the resonance region in graphene interlayer tunneling transistors.¹⁴ Such a functional form can be representative of phonon assisted tunneling, short-range disorder,²⁸ charged impurities,²⁹ or Moiré patterns that have been observed, for instance, at the graphene-*h*BN interface.³⁰⁻³² We will see in Sec. II E that the L_C has an influence on the gate voltage dependent current, which has a neat physical interpretation, hence a comparison to experimental data will be very informative for an estimate of L_C .

C. Effects of energy broadening

According to Eqs. (4) and (15), the tunneling current is simply zero when there is no energy overlap between the

conduction band in the top layer and the valence band in the bottom layer, that is, for $E_{CT} > E_{VB}$. In a real device, however, the 2D materials will inevitably have phonons, disorder, host impurities in the 2D layer and be affected by the background impurities in the surrounding materials, so that a finite broadening of the energy levels is expected to occur because of the statistical potential fluctuations superimposed to the ideal crystal structure.³³ The energy broadening in 3D semiconductors is known to lead to a tail of the density of states (DoS) in the gap region, that has been also observed in optical absorption measurements and denoted Urbach tail.^{34,35} It is thus expected that the finite energy broadening will be a fundamental limit to the abruptness of the turn on characteristic attainable with the devices of this work, hence it is important to include this effect in our model.

Energy broadening in the 2D systems can stem from the interaction with randomly distributed impurities and disorder in the 2D layer or in the surrounding materials,^{33,36,37} by scattering events induced by the interfaces,³⁸ as well as by other scattering sources. We recognize the fact that a detailed description of the energy broadening is exceedingly complicated due to the many-body and statistical fluctuation effects,³⁹ and thus resort to a relatively simple semi-classical treatment.^{33,36} We start by recalling that the density of states $\rho_0(E)$ for a 2D layer with no energy broadening is

$$\rho_0(E) = \frac{g_s g_v}{4\pi^2} \int_{\mathbf{k}} d\mathbf{k} \delta[E - E(\mathbf{k})], \quad (17)$$

where $E(\mathbf{k})$ denotes the energy relation with no broadening and g_s, g_v are spin and valley degeneracy. In the presence of a randomly fluctuating potential $V(\mathbf{r})$, instead, the DoS can be written as^{33,36}

$$\begin{aligned} \rho(E) &= \int_0^\infty dv \rho_0(v) P_v(E - v) \\ &= \frac{g_s g_v}{4\pi^2} \int_{\mathbf{k}} d\mathbf{k} \left[\int_0^\infty dv \delta[v - E(\mathbf{k})] P_v(E - v) \right] \\ &= \frac{g_s g_v}{4\pi^2} \int_{\mathbf{k}} d\mathbf{k} P_v[E - E(\mathbf{k})], \end{aligned} \quad (18)$$

where $P_v(v)$ is the distribution function for $V(\mathbf{r})$ (to be further discussed below), and we have used the $\rho_0(E)$ definition in Eq. (17) to go from the first to the second equality.

Comparing Eq. (18) to Eq. (17), we see that the $\rho(E)$ of the system in the presence of broadening can be calculated by substituting the Dirac function in Eq. (17) with a finite width function $P_v(v)$, which is the distribution function of $V(\mathbf{r})$, and it is thus normalized to one.

In order to include the energy broadening in our current calculations, we rewrite the tunneling rate in Eq. (4) as

$$\begin{aligned} \frac{1}{\tau_{\mathbf{k}_T, \mathbf{k}_B}} &= \frac{2\pi}{\hbar} |M(\mathbf{k}_T, \mathbf{k}_B)|^2 \delta[E_T(\mathbf{k}_T) - E_B(\mathbf{k}_B)] \\ &= \frac{2\pi}{\hbar} |M(\mathbf{k}_T, \mathbf{k}_B)|^2 \int_{-\infty}^{\infty} dE \delta[E - E_T(\mathbf{k}_T)] \delta[E - E_B(\mathbf{k}_B)], \end{aligned} \quad (19)$$

and note that, consistent with Eq. (18), the energy broadening can be included in the current calculation by substituting

$\delta[E - E(\mathbf{k})]$ with $P_v[E - E(\mathbf{k})]$. By doing so the tunneling rate becomes

$$\frac{1}{\tau_{\mathbf{k}_T, \mathbf{k}_B}} \simeq \frac{2\pi}{\hbar} |M(\mathbf{k}_T, \mathbf{k}_B)|^2 S_E(E_T(\mathbf{k}_T) - E_B(\mathbf{k}_B)), \quad (20)$$

where we have introduced an energy broadening spectrum S_E that is defined as

$$\begin{aligned} S_E(E_T(\mathbf{k}_T) - E_B(\mathbf{k}_B)) &= \int_{-\infty}^{\infty} dE P_{vT}[E - E_T(\mathbf{k}_T)] \\ &\quad \times P_{vB}[E - E_B(\mathbf{k}_B)], \end{aligned} \quad (21)$$

where P_{vT} and P_{vB} are the potential distribution functions due to the presence of randomly fluctuating potential $V(\mathbf{r})$ in the top and the bottom layer, respectively.

On the basis of Eq. (20), in our model for the tunneling current, we accounted for the energy broadening by using in all numerical calculations the broadening spectrum $S_E(E_T(\mathbf{k}_T) - E_B(\mathbf{k}_B))$ defined in Eq. (21) in place of $\delta[E_T(\mathbf{k}_T) - E_B(\mathbf{k}_B)]$. More precisely, we used a Gaussian potential distribution for both the top and the bottom layer

$$P_v(E - E_{k0}) = \frac{1}{\sqrt{\pi}\sigma} e^{-(E - E_{k0})^2/\sigma^2}, \quad (22)$$

which has been derived by Evan O. Kane for a broadening induced by randomly distributed impurities,³⁶ in which case σ can be expressed in terms of the average impurity concentration.

Quite interestingly, for the Gaussian spectrum in Eq. (22), the overall broadening spectrum S_E defined in Eq. (21) can be calculated analytically and reads

$$S_E(E_T(\mathbf{k}_T) - E_B(\mathbf{k}_B)) = \frac{1}{\sqrt{\pi}(\sigma_T^2 + \sigma_B^2)} e^{-(E_T(\mathbf{k}_T) - E_B(\mathbf{k}_B))^2/\sigma^2}. \quad (23)$$

Hence also S_E has a Gaussian spectrum, where σ_T and σ_B are the broadening energies for the top and bottom 2D layer, respectively.

D. Rotational misalignment and tunneling between inequivalent extrema

The derivations in Sec. II B assumed that there is a perfect rotational alignment between the top and the bottom layer and that the tunneling occurs between equivalent extrema in the Brillouin zone, that is tunneling from a K to a K extremum (or from K' to K' extremum). We now denote by θ the angle expressing a possible rotational misalignment between the two 2D layers (see Fig. 3), and still assume that the top 2D crystal has the same lattice constant a_0 as the bottom 2D crystal. The principal coordinate system is taken as the crystal coordinate system in the bottom layer, and we denote with \mathbf{r}' , \mathbf{k}' the position and wave vectors in the crystal coordinate system of the top layer (with \mathbf{r} , \mathbf{k} being the vectors in the principal coordinate system). The wave-function in the top layer has the form given in Eq. (6) in terms of \mathbf{r}' , \mathbf{k}' , hence, in order to calculate the matrix element in the

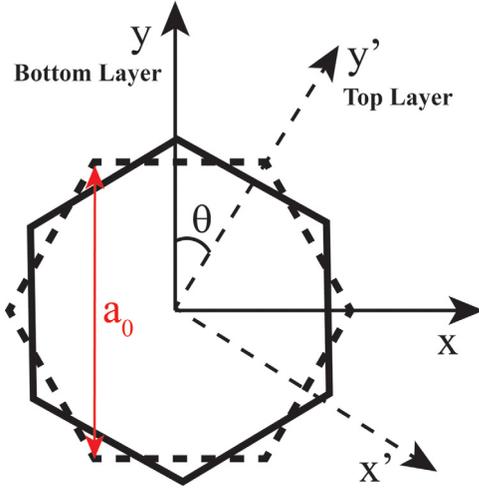


FIG. 3. Sketch of a possible rotational misalignment between the top and bottom 2D layer, x - y is the reference coordinate for the bottom 2D layer and x' - y' is the reference coordinate for the top 2D layer. θ is the rotational misalignment angle. We assume the top layer and the bottom layer have the same lattice constant a_0 .

principal coordinate system, we start by writing $\mathbf{r}' = \hat{R}_{B \rightarrow T} \mathbf{r}$, $\mathbf{k}' = \hat{R}_{B \rightarrow T} \mathbf{k}$, where $\hat{R}_{B \rightarrow T}$ is the rotation matrix from the bottom to the top coordinate system, with $\hat{R}_{T \rightarrow B} = [\hat{R}_{B \rightarrow T}]^T$ being the matrix going from the top to the bottom coordinate system and M^T denoting the transpose of the matrix M . The rotation matrix can be written as

$$\hat{R}_{T \rightarrow B} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad (24)$$

in terms of the rotational misalignment angle θ .

Consistent with Sec. II B, we set $u_{T, \mathbf{k}_T}(\mathbf{r}', z) \approx u_{0T}(\mathbf{r}', z)$, $u_{B, \mathbf{k}_B}(\mathbf{r}, z) \approx u_{0B}(\mathbf{r}, z)$, where $u_{0T}(\mathbf{r}', z)$, $u_{0B}(\mathbf{r}, z)$ are the periodic part of the Bloch function respectively at the band edge in the top and bottom layer. We then denote with \mathbf{K}_{0T} the wave-vector at the conduction band edge in the top layer (expressed in the top layer coordinate system) and with \mathbf{K}_{0B} the wave-vector at the valence band edge in the bottom layer (expressed in the principal coordinate system); the derivations in this section account for the fact that \mathbf{K}_{0T} and \mathbf{K}_{0B} may be inequivalent extrema (i.e., $\mathbf{K}_{0T} \neq \mathbf{K}_{0B}$).

By expressing \mathbf{r}' and \mathbf{k}' in the principal coordinate system, we can essentially follow the derivations in Sec. II B and write the matrix element as

$$M(\mathbf{k}_T, \mathbf{k}_B) \simeq \frac{1}{N_C} \sum_{j=1}^{N_C} e^{i(\mathbf{q} + \mathbf{Q}_D) \cdot \mathbf{r}_j} F_L(\mathbf{r}_j) \times \int_{\Omega_C} d\mathbf{r} \int dz u_{0T}^\dagger \times (\hat{R}_{B \rightarrow T}(\mathbf{r}_j + \boldsymbol{\rho}), z) V_B(z) u_{0B}(\mathbf{r}_j + \boldsymbol{\rho}, z), \quad (25)$$

where $\mathbf{q} = (\mathbf{k}_B - \mathbf{k}_T)$, and we have introduced the vector

$$\mathbf{Q}_D = \mathbf{K}_{0B} - \hat{R}_{T \rightarrow B} \mathbf{K}_{0T}. \quad (26)$$

Eq. (25) is an extension of Eq. (10) that accounts for a possible rotational misalignment between the 2D layers and describes also the tunneling between inequivalent extrema. The vector \mathbf{Q}_D is zero only for tunneling between equivalent

extrema (i.e., $\mathbf{K}_{0B} = \mathbf{K}_{0T}$) and for a perfect rotational alignment (i.e., $\theta = 0$). Considering a case where all extrema are at the \mathbf{K} point, we have $|\mathbf{K}_{0B}| = |\mathbf{K}_{0T}| = 4\pi/3a_0$, then for $\mathbf{K}_{0B} = \mathbf{K}_{0T}$, the magnitude of \mathbf{Q}_D is simply given by $Q_D = (8\pi/3a_0)\sin(\theta/2)$.¹³

One significant difference in Eq. (25) compared to Eq. (10) is that, in the presence of rotational misalignment, the top layer Bloch function $u_{0T}(\hat{R}_{B \rightarrow T} \mathbf{r}, z)$ has a different periodicity in the principal coordinate system from the bottom layer $u_{0B}(\mathbf{r}, z)$. Consequently, the integral over the unit cells of the bottom 2D layer is not the same in all unit cells, so that the derivations going from Eq. (10) to Eq. (15) should be rewritten accounting for a matrix element $M_{B0,j}$ depending on the unit cell j . Such an $M_{B0,j}$ could be formally included in the calculations by defining a new scattering spectrum that includes not only the inherently random fluctuations of the potential $F_L(\mathbf{r})$ but also the cell to cell variations of the matrix element $M_{B0,j}$. A second important difference of Eq. (25) compared to Eq. (10) lies in the presence of \mathbf{Q}_D in the exponential term multiplying $F_L(\mathbf{r}_j)$.

For the case of tunneling between inequivalent extrema and with a negligible rotational misalignment (i.e., $\theta \simeq 0$), Eq. (26) gives $\mathbf{Q}_D = \mathbf{K}_{0B} - \mathbf{K}_{0T}$, and the current can be expressed as in Eq. (15) but with the scattering spectrum evaluated at $|\mathbf{q} + \mathbf{Q}_D|$. Since, in this case, the magnitude of \mathbf{Q}_D is comparable to the size of the Brillouin zone, the tunneling between inequivalent extrema is expected to be substantially suppressed if the correlation length L_c of the scattering spectrum $S_F(\mathbf{q})$ is much larger than the lattice constant, as it has been assumed in all the derivations.

Quite interestingly, the derivations in this section suggest that a possible rotational misalignment is expected to affect the absolute value of the tunneling current, but not to change significantly its dependence on the terminal voltages.

From a technological viewpoint, if the stack of the 2D materials is obtained using a dry transfer method the rotational misalignment appears inevitable.^{14,40} Experimental results have shown that, when the stack of 2D materials is obtained by growing the one material on top of the other, the top 2D and bottom 2D layer can have a fairly good angular alignment.^{41,42}

E. An analytical approximation for the tunneling current

The numerical calculations for the tunneling current obtained with the model derived in Secs. II B and II C will be presented in Sec. III while in this section, we discuss an analytical, approximated expression for the tunneling current, which is mainly useful to gain an insight about the main physical and material parameters affecting the current versus voltage characteristic of the Thin-TFET. In order to derive an analytical current expression, we start by assuming a parabolic energy relation and write

$$E_{VB}(\mathbf{k}_B) = E_{VB} - \frac{\hbar^2 k_B^2}{2m_v} \quad E_{CT}(\mathbf{k}_T) = E_{CT} + \frac{\hbar^2 k_T^2}{2m_c}, \quad (27)$$

where $E_{VB}(\mathbf{k}_B)$, $E_{CT}(\mathbf{k}_T)$ are the energy relation, respectively, in the bottom layer valence band and top layer conduction band and m_v , m_c the corresponding effective masses.

In the analytical derivations, we neglect the energy broadening and start from Eq. (15), so that the model is essentially valid only in the on-state of the device, that is for $E_{CT} < E_{VB}$.

We now focus on the integral over \mathbf{k}_B and \mathbf{k}_T in Eq. (15) and first introduce the polar coordinates $\mathbf{k}_B = (k_B, \theta_B)$, $\mathbf{k}_T = (k_T, \theta_T)$, and then use Eq. (27) to convert the integrals over k_B, k_T to integrals over respectively E_B, E_T , which leads to

$$I \propto \int_{\mathbf{k}_T} \int_{\mathbf{k}_B} d\mathbf{k}_T d\mathbf{k}_B S_F(q) \delta(E_B(\mathbf{k}_B) - E_T(\mathbf{k}_T)) (f_B - f_T) \\ = \frac{m_c m_v}{\hbar^4} \int_0^{2\pi} d\theta_B \int_0^{2\pi} d\theta_T \int_{E_{CT}}^{\infty} dE_T \int_{-\infty}^{E_{VB}} dE_B S_F(q) \\ \times \delta(E_B - E_T) (f_B - f_T), \quad (28)$$

where the spectrum $S_F(q)$ is given by Eq. (16) and thus depends only on the magnitude q of $\mathbf{q} = \mathbf{k}_B - \mathbf{k}_T$. Assuming $E_{CT} < E_{VB}$, the Dirac function reduces one of the integrals over the energies and sets $E = E_B = E_T$, furthermore the magnitude of $\mathbf{q} = \mathbf{k}_B - \mathbf{k}_T$ depends only on the angle $\theta = \theta_B - \theta_T$, so that Eq. (28) simplifies to

$$I \propto \frac{m_c m_v (2\pi)}{\hbar^4} \int_0^{2\pi} d\theta \int_{E_{CT}}^{E_{VB}} dE S_F(q) (f_B - f_T). \quad (29)$$

In the on-state condition (i.e., for $E_{CT} < E_{VB}$), the zero Kelvin approximation for the Fermi–Dirac occupation functions f_B, f_T can be introduced to further simplify Eq. (29) to

$$I \propto \frac{m_c m_v (2\pi)}{\hbar^4} \int_0^{2\pi} d\theta \int_{E_{min}}^{E_{max}} dE S_F(q), \quad (30)$$

where $E_{min} = \max\{E_{CT}, E_{FT}\}$ and $E_{max} = \min\{E_{VB}, E_{FB}\}$ define the tunneling window $[E_{max} - E_{min}]$.

The evaluation of Eq. (30) requires to express q as a function of the energy E inside the tunneling window and of the angle θ between \mathbf{k}_B and \mathbf{k}_T . By recalling $q^2 = k_B^2 + k_T^2 - 2k_B k_T \cos(\theta)$, we can use Eq. (27) to write

$$q^2 = \frac{2m_v}{\hbar^2} (E_{VB} - E) + \frac{2m_c}{\hbar^2} (E - E_{CT}) \\ - \frac{4\sqrt{m_c m_v}}{\hbar^2} \sqrt{(E_{VB} - E)(E - E_{CT})} \cos(\theta), \quad (31)$$

with $E = E_B = E_T$. When Eq. (31) is substituted in the spectrum $S_F(q)$, the resulting integrals over E and θ in Eq. (30) cannot be evaluated analytically. Therefore, to proceed further, we now examine the maximum value taken by q^2 . The θ value leading to the largest q^2 is $\theta = \pi$, and the resulting q^2 expression can be further maximized with respect to the energy E varying in the tunneling window. The energy leading to maximum q^2 is

$$E_M = \frac{E_{CT} + (m_c/m_v)E_{VB}}{1 + (m_c/m_v)}, \quad (32)$$

and the corresponding q_M^2 is

$$q_M^2 = \frac{2(m_c + m_v)(E_{VB} - E_{CT})}{\hbar^2}. \quad (33)$$

When neither the top nor the bottom layer are degenerately doped, the tunneling window is given by $E_{min} = E_{CT}$ and $E_{max} = E_{VB}$, in which case the E_M defined in Eq. (32) belongs to the tunneling window, and the maximum value of q^2 is given by Eq. (33). If either the top or the bottom layer is degenerately doped, the Fermi levels become the edges of the tunneling window, and the maximum value of q^2 may be smaller than in Eq. (33).

A drastic simplification in the evaluation of Eq. (30) is obtained for $q_M^2 \ll 1/L_c^2$, in which case Eq. (16) returns to $S_F(q) \approx \pi L_c^2$, so that by substituting $S_F(q)$ in Eq. (29) and then in Eq. (15) the expression for the current simplifies to

$$I \simeq \frac{eg_v A(m_c m_v)}{\hbar^5} |M_{B0}|^2 e^{-2\kappa T_{IL}} L_c^2 (E_{max} - E_{min}), \quad (34)$$

where we recall that $E_{min} = \max\{E_{CT}, E_{FT}\}$, $E_{max} = \min\{E_{VB}, E_{FB}\}$ define the tunneling window.

It should be noticed that Eq. (34) is consistent with a complete loss of momentum conservation, so that the current is simply proportional to the integral over the tunneling window of the product of the density of states in the two 2D layers. Since for a parabolic effective mass approximation the density of states is energy independent, the current turns out to be simply proportional to the width of the tunneling window. In physical terms, Eq. (34) corresponds to a situation where the scattering produces a complete momentum randomization during the tunneling process.

As can be seen, as long as the top layer is not degenerately doped, we have $E_{min} = E_{CT}$ and the tunneling window widens with the increase of the top gate voltage $V_{T,G}$, hence, according to Eq. (34), the current is expected to increase linearly with $V_{T,G}$. However, when the tunneling window increases to such an extent that q_M^2 becomes comparable to or larger than $1/L_c^2$, then part of the q values in the integration of Eq. (30) belong to the tail of the spectrum $S_F(q)$ defined in Eq. (16), and so their contribution to the current becomes progressively vanishing. The corresponding physical picture is that, while the tunneling window increases, the magnitude of the wave-vectors in the two 2D layers also increases, and, consequently, the scattering can no longer provide momentum randomization for all the possible wave-vectors involved in the tunneling process. Under these circumstances, the current is expected to first increase sub-linearly with $V_{T,G}$ and eventually saturate for large enough $V_{T,G}$ values.

III. NUMERICAL RESULTS FOR THE TUNNELING CURRENT

The 2D materials used for the tunneling current calculations reported in this paper are the hexagonal monolayer MoS_2 and WTe_2 . The band structure for MoS_2 and WTe_2 has been calculated by using a density functional theory (DFT) approach,^{18,43} showing that these materials have a direct bandgap with the band edges for both the valence and the conduction band residing at the K point in the 2D Brillouin zone. Fig. 4 shows that in a range of about 0.4 eV from the

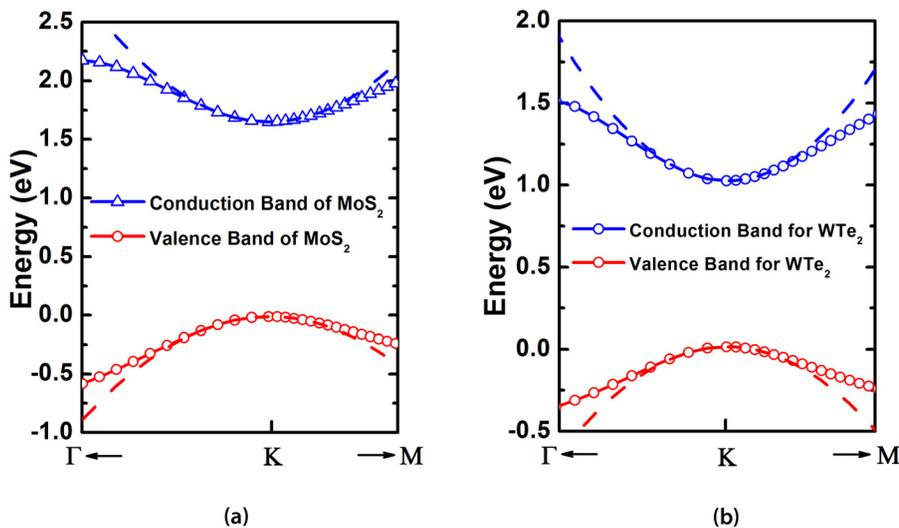


FIG. 4. (a) Band structure for hexagonal monolayer MoS₂ and (b) hexagonal monolayer WTe₂ as obtained using DFT method described in the paper of Gong *et al.*¹⁸ The dashed lines represent the analytical approximation obtained with a parabolic effective mass model.

band edges, the DFT results can be fitted fairly well by using an energy relation based on a simple parabolic effective mass approximation (dashed lines). Hence, the parabolic effective mass approximation appears adequate for the purposes of this work, which is focussed on a device concept for extremely small supply voltages (< 0.5 V). The values for the effective masses inferred from the fitting of the DFT calculations are tabulated in Table I together with some other material parameters relevant for the tunneling current calculations.

In all current calculations, we assume a top gate work function of 4.17 eV (Aluminium) and back gate work function of 5.17 eV (p++ Silicon), and the top and bottom oxides have an effective oxide thickness (EOT) of 1 nm (see Fig. 1). The top 2D layer consists of hexagonal monolayer MoS₂ while the bottom 2D layer is hexagonal monolayer WTe₂. An *n*-type and *p*-type doping density of 10^{12} cm^{-2} by impurities and full ionization are assumed respectively in the top and bottom 2D layer, and the relative dielectric constant of the interlayer material is set to 4.2 (e.g., boron nitride). The voltage V_{DS} between the drain and the source is set to 0.3 V, and the back gate is grounded for all calculations, unless otherwise stated.

As already pointed out in Sec. II B, it is very difficult to derive a quantitative expression for the tunneling matrix element $M_{B,0}$. However, the value of $M_{B,0}$ could be inferred from the experimental data. In the lack of experimental data for a vertical transistor consisting of transition-metal dichalcogenides, we have set the value of $M_{B,0}$ to be 0.01 eV in our calculations, so that the resultant current density is in the same order of magnitude with the experimental value reported in the graphene/hBN system.⁴⁴

TABLE I. The band gaps, electron affinities, and effective masses used for MoS₂ and WTe₂.

	Bandgap (eV)	Electron affinity (χ)	Conduction band effective mass (m_0)	Valence band effective mass (m_0)
MoS ₂	1.8	4.30	0.378	0.461
WTe ₂	0.9	3.65	0.235	0.319

In Fig. 5, the results of numerical calculations are shown for the band alignment and the current density versus the top gate voltage V_{TG} . Figure 5(a) shows that the top gate voltage can effectively govern the band alignment in the device and, in particular, the crossing and uncrossing between the conduction band minimum E_{CT} in the top layer and the valence band maximum E_{VB} in the bottom layer, which discriminates between the on and off state of the transistor.

The I_{DS} versus V_{TG} characteristic in Fig. 5(b) can be roughly divided into three different regions: sub-threshold region, linear region, and saturation region. The sub-threshold region corresponds to the condition $E_{CT} > E_{VB}$ (see also Fig. 5(a)), where the very steep current dependence on V_{TG} is illustrated better in Fig. 6 and will be discussed below.

In the second region, I_{DS} exhibits an approximately linear dependence on V_{TG} , in fact, the current is roughly proportional to the energy tunneling window, as discussed in Sec. II E and predicted by Eq. (34), because the tunneling window is small enough that the condition $q_M^2 \ll 1/L_c^2$ is fulfilled. In this region I_{DS} is proportional to the long-wavelength part of the scattering spectrum $S_F(q)$ (i.e., small q values), hence the current increases with L_c , as expected from Eq. (34). The super-linear behavior of I_{DS} at small V_{TG} values observed in Fig. 5(b) is due to the tail of the Fermi occupation function in the top layer. When V_{TG} is increased above approximately 0.5 V, the current in Fig. 5(b) enters the saturation region, where I_{DS} increasing with V_{TG} slows down because of the decay of the scattering spectrum $S_F(q)$ for q values larger than $1/L_c$ (see Eq. (16)).

In Fig. 6, we analyze the I-V curves for different interlayer thicknesses T_{IL} and broadening energies σ ; in all cases, an average inverse sub-threshold slope is extracted in the I_{DS} range from 10^{-5} and $10^{-2} \mu\text{A}/\mu\text{m}^2$. Figure 6(a) shows that the tunneling current increases exponentially by decreasing T_{IL} , and the decay constant $\kappa = 3.8 \text{ nm}^{-1}$ employed in our calculations results in a dependence on T_{IL} that is roughly consistent with the dependence experimentally reported in graphene based interlayer tunneling devices.¹⁵ The threshold voltages are also shifted to lower values by increasing T_{IL} . It can be seen that the T_{IL} impact on SS is overall weak, and a very steep sub-threshold region is obtained for all the T_{IL}

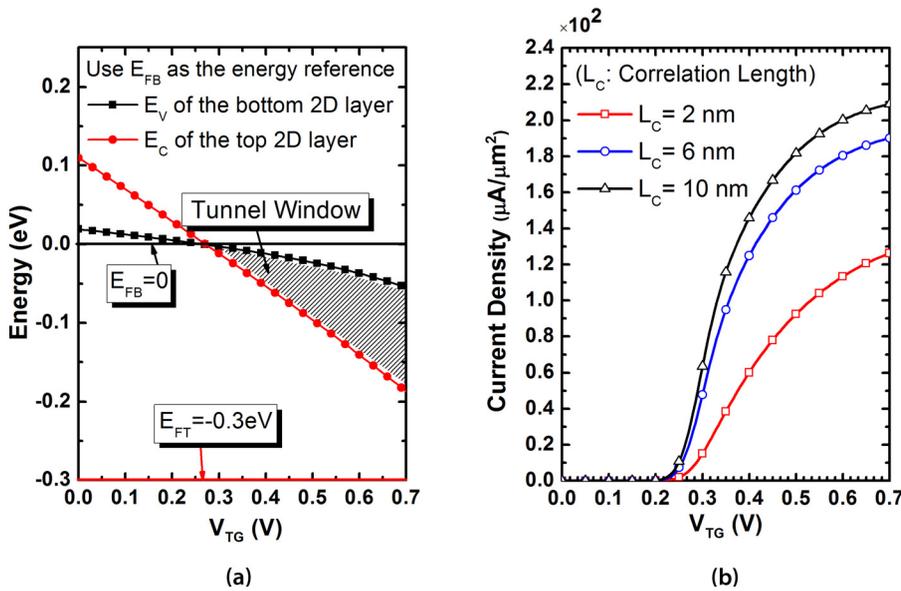


FIG. 5. Numerical results of (a) band alignment versus the top gate voltage V_{TG} and (b) tunnel current density versus the top gate voltage V_{TG} for different values of the correlation length L_C . The parameters used in (b) are as follows: Matrix element is $M_{B0} = 0.01\text{ eV}$; decay constant of wave-function in the interlayer is $\kappa = 3.8\text{ nm}^{-1}$; energy broadening is $\sigma = 10\text{ meV}$; and interlayer thickness is $T_{IL} = 0.6\text{ nm}$ (e.g., 2 atomic layers of BN). $V_{BG} = 0$ and $V_{DS} = 0.3\text{ V}$ in both (a) and (b).

values examined in Fig. 6(a). This is because, in order for the Thin-TFET to obtain a small SS, it is absolutely necessary that V_{TG} has a tight control on the electrostatic potential in the top semiconductor layer, but has a negligible influence on the potential of the bottom semiconductor layer. The SS is thus insensitive to T_{IL} as long as T_{IL} does not change the control of V_{TG} on such potentials. In short, for Thin-TFETs, a larger interlayer thickness reduces substantially the current density, but does not deteriorate SS.

Figure 6(b) shows that according to the model employed in our calculations SS is mainly governed by the parameter σ of the energy broadening Eq. (22). This result is expected, as already mentioned in Sec. II C, since, in our model, the energy broadening is the physical factor setting the minimum value for SS, and the I_{DS} versus V_{TG} approaches a step-like curve when σ is zero due to the step-like DoS of these 2D semiconductors.⁴⁵ More specifically, Fig. 6(b) shows that according to our calculations, the Thin-TFET may be able to provide an SS below the 60 mV/dec (i.e., the limit of conventional MOSFETs at room temperature), even for fairly large broadening energies up to about 40 meV. It is here worth

noting that energy broadening and band tails have been already recognized as a fundamental limit to the SS of band-to-band tunneling transistors,⁴⁶ and are not a specific concern of the Thin-TFET. As already mentioned in Sec. II C, the band tails in 3D semiconductors have been investigated by using thermal measurements and are described in terms of the so called Urbach parameter E_0 .^{34,35} Values for E_0 comparable to the room temperature thermal energy, $k_B T \approx 26\text{ meV}$, have been reported for GaAs and InP.^{47,48} Our results suggest that energy broadening and band tails in 2D materials play a critical role in the minimum SS attainable by Thin-TFETs, and, at the time of writing, we are not aware of experimental data reported for band tails in monolayers of transition-metal dichalcogenides.

IV. DISCUSSION AND CONCLUSIONS

This paper proposed a new steep slope transistor based on the interlayer tunneling between two 2D semiconductor materials and presented a detailed model to discuss the physical mechanisms governing the device operation and to gain

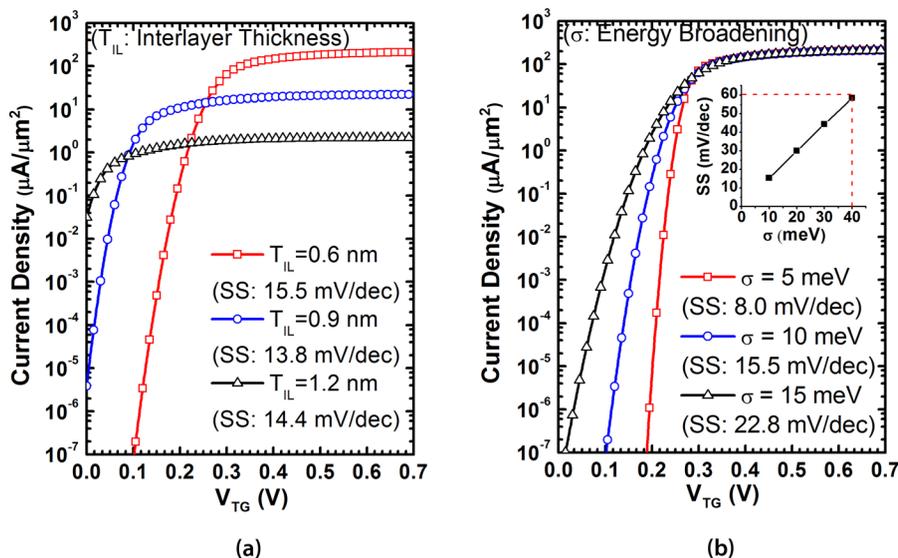


FIG. 6. Numerical calculations for: (a) current density versus V_{TG} with several interlayer thicknesses; (b) current density versus V_{TG} with different values of energy broadening σ . The inset shows that SS increases with σ , and a SS value of 60 mV/dec corresponds to a energy broadening as high as 40 meV. The matrix element is $M_{B0} = 0.01\text{ eV}$; the decay constant of wave-function in the interlayer is $\kappa = 3.8\text{ nm}^{-1}$. In (a), the energy broadening is $\sigma = 10\text{ meV}$. In (b), the interlayer thickness is $T_{IL} = 0.6\text{ nm}$ (e.g., 2 atomic layers of BN). $V_{BG} = 0$ and $V_{TG} = 0.3\text{ V}$ in both (a) and (b).

an insight about the tradeoffs implied in the design of the transistor.

The tunnel transistor based on 2D semiconductors has the potential for a very steep subthreshold region, and the subthreshold swing is ultimately limited by the energy broadening in the two 2D materials. The energy broadening can have different physical origins, such as disorder, charged impurities in the 2D layers or in the surrounding materials,^{37,39} phonon scattering⁴⁹ and microscopic roughness at interfaces.³⁸ In our calculations, we accounted for the energy broadening by assuming a simple Gaussian energy spectrum with no explicit reference to a specific physical mechanism. However, a more detailed and quantitative description of the energy broadening is instrumental in physical modeling of the device and its design.

Quite interestingly, our analysis suggests that, while a possible rotational misalignment between the two 2D layers can affect the absolute value of the tunneling current, the misalignment is not expected to significantly degrade the steep subthreshold slope, which is the crucial figure of merit for a steep slope transistor.

An optimal operation of the device demands a good electrostatic control of the top gate voltage V_{TG} on the band alignments in the material stack, as shown for example in Fig. 5(a), which may become problematic if the electric field in the interlayer is effectively screened by the high electron concentration in the top 2D layer. Consequently, since high carrier concentrations in the 2D layers are essential to reduce the layer resistivities, a tradeoff exists between the gate control and layer resistivities; as a result, doping concentrations in these 2D layers are important design parameters in addition to tuning the threshold voltage. In this respect, chemical doping of TMD materials have been recently demonstrated,^{50,51} however these doping technologies are still far less mature than they are for 3D semiconductors, and improvements in *in-situ* doping will be very important for optimization of the device performance. Since our model does not include the lateral transport in the 2D materials, an exploration of the above design tradeoffs goes beyond the scope of the present paper and demands the development of more complete transport models.

The transport model proposed in this work does not account for possible traps or defects assisted tunneling, which have been recently recognized as a serious hindrance to the experimental realization of Tunnel-FETs exhibiting a subthreshold swing better than 60 mV/dec.^{11,12} A large density of states in the gap of the 2D materials may even lead to a Fermi level pinning that would drastically degrade the gate control on the band alignment and undermine the overall device operation. In this respect, from a fundamental viewpoint, the 2D crystals may offer advantages over their 3D counterparts because they are inherently free of broken/dangling bonds at the interfaces.¹⁹ However, the fabrication technologies for 2D crystals are still in an embryonal stage compared to technologies for conventional semiconductors, hence, the control of defects in the 2D materials will be a challenge for the development of the proposed tunneling transistor.

The simulation results reported in this paper indicate that the newly proposed transistor based on interlayer

tunneling between two 2D materials has the potential for a very steep turn-on characteristic, because the vertical stack of 2D materials having an energy gap is probably the device structure that allows for the most effective, gate controlled crossing and uncrossing between the edges of the bands involved in the tunneling process. Our modeling approach based on the Bardeen's transfer Hamiltonian is, by no means, a complete device model but, instead, a starting point to gain insight about its working principle and its design. At the present time, an experimental demonstration of the device appears of crucial importance, first of all to validate the device concept and then to help estimate the numerical value of a few parameters in the transport model that can be determined only by comparing to experiments.

ACKNOWLEDGMENTS

This work was supported in part by the Center for Low Energy Systems Technology (LEAST), one of six SRC STARnet Centers, sponsored by MARCO and DARPA, by the Air Force Office of Scientific Research (FA9550-12-1-0257), and by a Fulbright Fellowship for D. Esseni. The authors are also grateful for the helpful discussions with Profs. K. J. Cho, R. Feenstra, and A. Seabaugh. The authors are especially thankful to Dr. Cheng Gong in Dr. K. J. Cho's Group for providing the calculated band structure data shown in Fig. 4.

¹J. M. Rabaey, J. Ammer, T. Karalar, S. Li, B. Otis, M. Sheets, and T. Tuan, in *Solid-State Circuits Conference, 2002. Digest of Technical Papers. ISSCC. 2002 IEEE International* (IEEE, 2002), vol. 1, pp. 200–201.

²R. Amirtharajah and A. P. Chandrakasan, *IEEE J. Solid-State Circuits* **33**, 687 (1998).

³R. G. Dreslinski, M. Wieckowski, D. Blaauw, D. Sylvester, and T. Mudge, *Proc. IEEE* **98**, 253 (2010).

⁴I. W. Group *et al.*, see <http://www.itrs.net> for International Technology Roadmap for Semiconductors (ITRS) [Online]. Available: <http://www.itrs.net/Links/2011ITRS/Home2011.htm>.

⁵A. C. Seabaugh and Q. Zhang, *Proc. IEEE* **98**, 2095 (2010).

⁶G. Zhou, R. Li, T. Vasen, M. Qi, S. Chae, Y. Lu, Q. Zhang, H. Zhu, J.-M. Kuo, T. Kosel *et al.*, *Tech. Dig – Int Electron Devices meet* **2012**, 32.6.1–32.6.4.

⁷Y. Lu, G. Zhou, R. Li, Q. Liu, Q. Zhang, T. Vasen, S. D. Chae, T. Kosel, M. Wistey, H. Xing *et al.*, *IEEE Electron Device Lett.* **33**, 655 (2012).

⁸L. Knoll, Q.-T. Zhao, A. Nichau, S. Trellenkamp, S. Richter, A. Schafer, D. Esseni, L. Selmi, K. K. Bourdelle, and S. Mantl, *IEEE Electron Device Lett.* **34**, 813 (2013).

⁹G. Zhou, Y. Lu, R. Li, Q. Zhang, Q. Liu, T. Vasen, H. Zhu, J.-M. Kuo, T. Kosel, M. Wistey *et al.*, *IEEE Electron Device Lett.* **33**, 782 (2012).

¹⁰F. Conzatti, M. Pala, D. Esseni, E. Bano, and L. Selmi, *IEEE Trans. Electron Devices* **59**, 2085 (2012).

¹¹M. Pala and D. Esseni, *IEEE Trans. Electron Devices* **60**, 2795 (2013).

¹²D. Esseni and M. G. Pala, *IEEE Trans. Electron Devices* **60**, 2802 (2013).

¹³R. M. Feenstra, D. Jena, and G. Gu, *J. Appl. Phys.* **111**, 043711 (2012).

¹⁴L. Britnell, R. Gorbachev, A. Geim, L. Ponomarenko, A. Mishchenko, M. Greenaway, T. Fromhold, K. Novoselov, and L. Eaves, *Nat. Commun.* **4**, 1794 (2013).

¹⁵L. Britnell, R. V. Gorbachev, R. Jalil, B. D. Belle, F. Schedin, M. I. Katsnelson, L. Eaves, S. V. Morozov, A. S. Mayorov, N. M. Peres *et al.*, *Nano Lett.* **12**, 1707 (2012).

¹⁶B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, *Nat. Nanotechnol.* **6**, 147 (2011).

¹⁷Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, *Nat. Nanotechnol.* **7**, 699 (2012).

¹⁸C. Gong, H. Zhang, W. Wang, L. Colombo, R. M. Wallace, and K. Cho, *Appl. Phys. Lett.* **103**, 053513 (2013).

- ¹⁹D. Jena, *Proc. IEEE* **101**, 1585 (2013).
- ²⁰K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, *Phys. Rev. Lett.* **105**, 136805 (2010).
- ²¹D. Esseni, P. Palestri, and L. Selmi, *Nanoscale MOS Transistors: Semi-classical Transport and Applications* (Cambridge University Press, 2011).
- ²²J. Bardeen, *Phys. Rev. Lett.* **6**, 57 (1961).
- ²³W. A. Harrison, *Phys. Rev.* **123**, 85 (1961).
- ²⁴C. B. Duke, *Tunneling in Solids* (Academic Press, New York, 1969), Vol. 1999.
- ²⁵P. Zhao, R. Feenstra, G. Gu, and D. Jena, *IEEE Trans. Electron Devices* **60**, 951 (2013).
- ²⁶Z. Zhu, Y. Cheng, and U. Schwingenschlögl, *Phys. Rev. B* **84**, 153402 (2011).
- ²⁷A. Ramasubramaniam, D. Naveh, and E. Towe, *Phys. Rev. B* **84**, 205325 (2011).
- ²⁸Q. Li, E. Hwang, E. Rossi, and S. D. Sarma, *Phys. Rev. Lett.* **107**, 156601 (2011).
- ²⁹J. Yan and M. S. Fuhrer, *Phys. Rev. Lett.* **107**, 206601 (2011).
- ³⁰M. Yankowitz, J. Xue, D. Cormode, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, P. Jacquod, and B. J. LeRoy, *Nat. Phys.* **8**, 382 (2012).
- ³¹J. Xue, J. Sanchez-Yamagishi, D. Bulmash, P. Jacquod, A. Deshpande, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, and B. J. LeRoy, *Nature Mater.* **10**, 282 (2011).
- ³²R. Decker, Y. Wang, V. W. Brar, W. Regan, H.-Z. Tsai, Q. Wu, W. Gannett, A. Zettl, and M. F. Crommie, *Nano Lett.* **11**, 2291 (2011).
- ³³P. Van Mieghem, G. Borghs, and R. Mertens, *Phys. Rev. B* **44**, 12822 (1991).
- ³⁴F. Urbach, *Phys. Rev.* **92**, 1324 (1953).
- ³⁵G. Cody, *J. Non-Cryst. Solids* **141**, 3 (1992).
- ³⁶E. O. Kane, *Phys. Rev.* **131**, 79 (1963).
- ³⁷S. D. Sarma and B. Vinter, *Surf. Sci.* **113**, 176 (1982).
- ³⁸A. Knabchen, *J. Phys.: Condens. Matter* **7**, 5209 (1995).
- ³⁹A. Ghazali and J. Serre, *Solid-State Electron.* **28**, 145 (1985).
- ⁴⁰L. Britnell, R. Gorbachev, R. Jalil, B. Belle, F. Schedin, A. Mishchenko, T. Georgiou, M. Katsnelson, L. Eaves, S. Morozov *et al.*, *Science* **335**, 947 (2012).
- ⁴¹S. Tiefenbacher, C. Pettenkofer, and W. Jaegermann, *Surf. Sci.* **450**, 181 (2000).
- ⁴²A. Koma, *J. Cryst. Growth* **201**, 236 (1999).
- ⁴³G.-B. Liu, W.-Y. Shan, Y. Yao, W. Yao, and D. Xiao, preprint arXiv:1305.6089 (2013).
- ⁴⁴L. Britnell, R. V. Gorbachev, R. Jalil, B. D. Belle, F. Schedin, M. I. Katsnelson, L. Eaves, S. V. Morozov, A. S. Mayorov, N. M. Peres *et al.*, preprint arXiv:1202.0735 (2012).
- ⁴⁵S. Agarwal and E. Yablonovitch, preprint arXiv:1109.0096 (2011).
- ⁴⁶M. A. Khayer and R. K. Lake, *J. Appl. Phys.* **110**, 074508 (2011).
- ⁴⁷J. Pankove, *Phys. Rev.* **140**, A2059 (1965).
- ⁴⁸A. V. Subashiev, O. Semyonov, Z. Chen, and S. Luryi, *Appl. Phys. Lett.* **97**, 181914 (2010).
- ⁴⁹U. Bockelmann and G. Bastard, *Phys. Rev. B* **42**, 8947 (1990).
- ⁵⁰H. Fang, M. Tosun, G. Seol, T. C. Chang, K. Takei, J. Guo, and A. Javey, *Nano Lett.* **13**, 1991 (2013).
- ⁵¹H. Fang, S. Chuang, T. C. Chang, K. Takei, T. Takahashi, and A. Javey, *Nano Lett.* **12**, 3788 (2012).