Nanotechnology

PAPER

Sub-5 nm monolayer black phosphorene tunneling transistors

To cite this article: Hong Li *et al* 2018 *Nanotechnology* **29** 485202

View the article online for updates and enhancements.

IOP ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

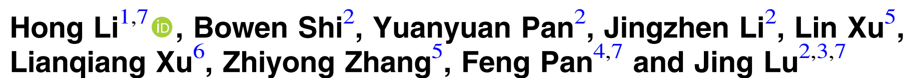
Start exploring the collection - download the first chapter of every title for free.

This content was downloaded from IP address 58.60.1.23 on 04/12/2018 at 07:01

IOP Publishing Nanotechnology

Nanotechnology **29** (2018) 485202 (10pp) https://doi.org/10.1088/1361-6528/aae0cb

Sub-5 nm monolayer black phosphorene tunneling transistors



Hong Li[1,7], Bowen Shi[2], Yuanyuan Pan[2], Jingzhen Li[2], Lin Xu[5], Lianqiang Xu[6], Zhiyong Zhang[5], Feng Pan[4,7] and Jing Lu[2,3,7]

[1] College of Mechanical and Material Engineering, North China University of Technology, Beijing 100144, People's Republic of China
[2] State Key Laboratory of Mesoscopic Physics and Department of Physics, Peking University, Beijing 100871, People's Republic of China
[3] Collaborative Innovation Center of Quantum Matter, Beijing 100871, People's Republic of China
[4] School of Advanced Materials, Peking University, Shenzhen Graduate School, Shenzhen 518055, People's Republic of China
[5] Key Laboratory for the Physics and Chemistry of Nanodevices and Department of Electronics, Peking University, Beijing 100871, People's Republic of China
[6] School of Physics and Electronic Information Engineering, Engineering Research Center of Nanostructure and Functional Materials, Ningxia Normal University, Guyuan, Ningxia 756000, People's Republic of China

E-mail: jinglu@pku.edu.cn, lihong@ncut.edu.cn and panfeng@pkusz.edu.cn

Received 25 August 2018
Accepted for publication 12 September 2018
Published 1 October 2018

Abstract

The successful fabrication of sub-5 nm 2D MoS_2 field-effect transistors (FETs) announces the approaching post-silicon era. It is possible for tunneling field-effect transistors (TFETs) based on monolayer black phosphorene (ML BP) to work well in the sub-5 nm region because of its moderate direct band gap, anisotropic electronic properties and high carrier mobility. We simulate the device performance limit of the ML BP TFETs at the sub-5 nm scale using *ab initio* quantum transport calculations. We predict that the on-state currents (I_{on}) of the sub-5 nm ML BP TFETs will exceed those of the ML WTe_2 TFETs, which possess the highest I_{on} among the transition-metal dichalcogenide family. In particular, the I_{on} of the ML BP TFETs can fulfill the 2028 requirements of the international technology roadmap for semiconductors (ITRS) for the high-performance (HP) devices until the gate length is scaled down to 4 nm, while the delay times and power dissipations always surpass the 2028 requirements of the ITRS HP devices significantly in the whole sub-5 nm region.

Supplementary material for this article is available online

Keywords: monolayer black phosphorene, sub-5 nm scale, tunneling transistor, device performance limit, *ab initio* quantum transport calculation

(Some figures may appear in colour only in the online journal)

1. Introduction

New-concept and post-silicon era field-effect transistors (FETs) have been intensively investigated as silicon complementary metal-oxide-semiconductor (CMOS) technology is believed to end at the 5 nm node with unacceptable low device performance and high power consumption [1–4]. The new-concept tunneling FET (TFET) architecture causes a lower subthreshold swing (SS) value than a thermionic FET and even breaks the SS limit of thermionic FETs (60 mV dec^{-1}) [5–8]. However, the on-state currents (I_{on}) are generally too small ($10^{-6} \sim 10^{-1}$ μA μm^{-1}) when using homogeneous bulk semiconductors as channel materials [6, 7]. The I_{on} can be improved to 400 and 142 μA μm^{-1} in

[7] Author to whom any correspondence should be addressed.

 © 2018 IOP Publishing Ltd Printed in the UK

the 20 nm-gate-long Ge-Si [9] and 5 nm-gate-long Ge/GaAs [10] thin heterojunction TFETs, respectively, but remain lower than those of the present high-performance (HP) logic device. Small I_{on} means slow switching speed in the logic device, which is one of the key stumbling blocks for the practical application of TFETs. 2D semiconducting materials are potential channel materials to extend Moore's law down to the sub-5 nm scale [7, 8, 11–15]. Their atomic thin bodies imply excellent gate controlling and geometric scaling capacities, while their smooth surfaces with fewer traps allow for efficient carrier transport [8, 16]. The inherent thin body and smooth surface of 2D materials provide a possible opportunity to elevate the on-state currents of TFETs.

Black phosphorene (BP), as a relatively new member of the 2D family, holds a moderate direct band gap, anisotropic electronic property and high carrier mobility. Since its first discovery in 2014, the thermionic BP FET has been experimentally fabricated and extensively studied [17–20]. A higher on-state current than the 2D transition-metal dichalcogenide channel has been found in the BP thermionic FET due to the anisotropic nature of monolayer (ML) BP [21, 22]. ML BP is also an extraordinary channel material for TFETs [23–27]. In particular, *ab initio* calculations have revealed that the ML BP TFETs with a gate length above 6.1 nm have on-state currents, delay times and power dissipations that meet the requirements of the international technology roadmap for semiconductors (ITRS) (2013 version) [28] for HP devices [25]. Very recently, 2D MoS_2 conventional FETs with a $1 \sim 9$ nm gate length [1, 29–31] have been fabricated successfully. Encouraged by the fast development of sub-5 nm technology, it is highly desirable to know whether the excellent device performance of ML BP TFETs can be maintained when the gate length is scaled down to below 5 nm.

In this paper, we simulate the ML BP TFETs with the gate length of $1 \sim 5$ nm along the armchair transport direction by using *ab initio* quantum transport simulations. The device performance of a 5 nm-gate-long ML WTe_2 TFET is calculated for comparison, as ML WTe_2 is predicted to be the best channel material for TFET in the transition-metal dichalcogenide family by a previous *ab initio* quantum transport simulation based on the flexible plane wave method [32]. The on-state current of the ML BP TFET with the 5 nm-long gate is more than four times greater than its WTe_2 counterpart, implying a faster switching speed of the former. The on-state currents of the ML BP TFET can even fulfill the 2028 requirements of the ITRS for HP devices (2013 version) [28] until the gate length is scaled down to 4 nm. Moreover, the delay times and power dissipations of the ML BP TFETs always surpass the 2028 requirements of the ITRS HP devices significantly in the whole sub-5 nm region.

2. Models and methods

A double-gated (DG) TFET model with planar *p-i-n* configuration based on ML BP is illustrated in figure 1(a). In this *p-i-n* configuration, the carrier transport mechanism is band-to-band tunneling (BTBT). The *p*- and *n*-type regions serve as the source and drain, respectively, and the intrinsic region forms the tunneling barrier, which can be tuned by the gate voltage to switch on and off. We take the gate lengths (L_g) as $1 \sim 5$ nm in steps of 1 nm. The device parameters, i.e. the equivalent oxide thickness (EOT), drain voltage (V_{ds}), and supply voltage (V_{dd}), are taken from the 2028 ITRS HP targets (2013 version) [28] with EOT = 0.41 nm and $V_{ds} = V_{dd} = 0.64$ V, respectively. The working voltage of 0.64 V is for a 5 nm-gate-long transistor for the ITRS 2028 target for HP application, which is a little bit smaller than that of $0.7 \sim 0.78$ V for a current $10 \sim 14$ nm-gate-long transistor for HP application (present CMOS) [28, 33, 34]. In the following, we are mainly concerned with the scale limit of the ML BP TFETs for HP application, so the device performance at a lower working voltage of $0.3 \sim 0.5$ V for low-power (LP) application is not studied. The connection between the dielectric and BP layer is ideal and has no trap density at the interface, so that the simulated device performance should be the upper limit. A good point is that the dangling-bond-free surface of the ML BP would depress the trap density in a real device configuration.

The device performances are calculated with the Atomistix ToolKit (ATK) 2016 package [35–37] based on density functional theory (DFT) coupled with the nonequilibrium Green's function method. We use the generalized gradient approximation (GGA) of Perdew–Burke–Ernzerhof (PBE) [38] form as the exchange-correlation functional, Hartwigsen–Goedecker–Hutter type as norm-conserving pseudopotentials and Tier 1 as the basis set. The density mesh cutoff is set to 100 Ha and the electron temperature is set to 300 K. The Monkhorst–Pack *k*-point mesh [39] is set to $29 \times 1 \times 100$ for the electronic self-consistent calculations. The current at a given gate voltage V_g and bias voltage V_{ds} is an integration of the transmission coefficient $T\,(E,\,V_{ds},\,V_g)$ using the Landauer–Büttiker formula [40]:

$$I(V_{ds},\,V_g) = \frac{2e}{h}\int_{-\infty}^{+\infty}\{T(E,\,V_{ds},\,V_g)[f_L(E-\mu_L) - f_R(E-\mu_R)]\}dE,$$

where $f_{L/R}$ is the Fermi–Dirac distribution function and μ_L/μ_R the electrochemical potential of the L/R electrode. The gate effect is calculated by solving the Poisson and Kohn–Sham equations self-consistently. $T\,(E)$ is an average of $T\,(E,\,k_{x'})$ over different $k_{x'}$ ($k_{x'} = 57$) in the irreducible Brillouin zone, where $T\,(E,\,k_{x'})$ is calculated by

$$T(E,\,k_{x'}) = Tr[G^r(E,\,k_{x'})\cdot\Gamma_S(E,\,k_{x'})\cdot G^a(E,\,k_{x'}) \cdot\Gamma_D(E,\,k_{x'})].$$

Here, $G^{r/a}\,(E,\,k_{x'})$ and $\Gamma_{S/D}(E,\,k_{x'}) = i(\sum_{S/D}^{r}(E,\,k_{x'}) - \sum_{S/D}^{a}(E,\,k_{x'}))$ are the retarded/advanced Green's function and the level broadening due to the left and right electrodes, respectively. $\Gamma_{S/D}\,(E,\,k_{x'})$ is expressed in terms of the electrode self-energies $\sum_{S/D}^{r/a}(E,\,k_{x'})$, which reflects the influence of the electrodes on the scattering region [41].

DFT-GGA is a good approximation for the transport calculation in an FET configuration [42–46]. For an intrinsic

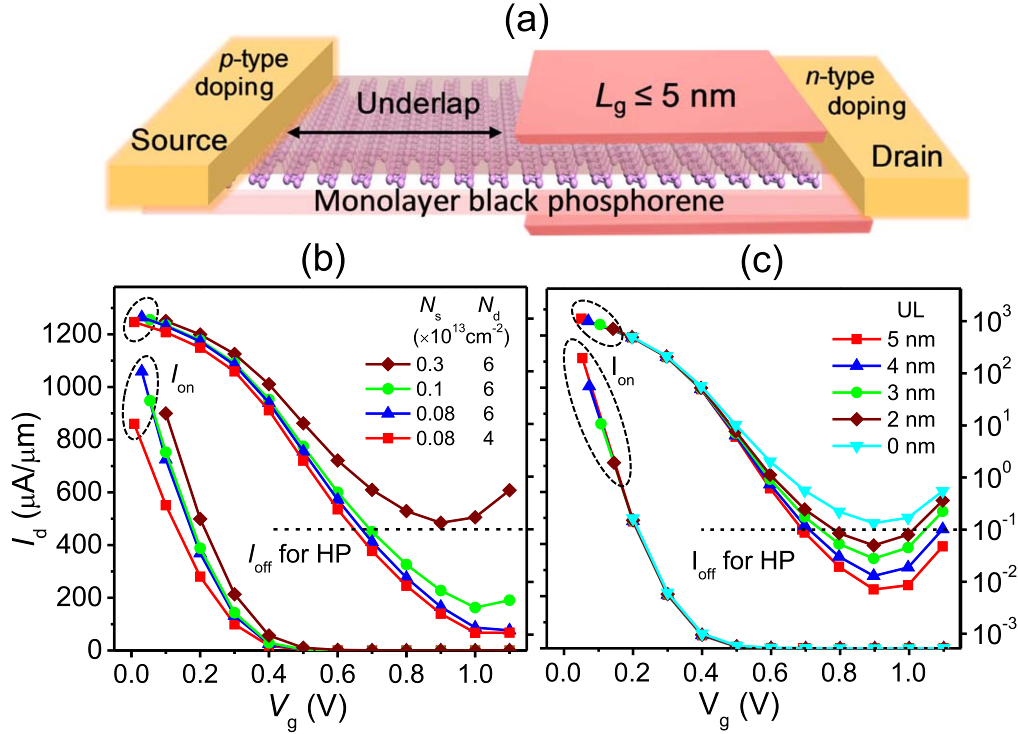


Figure 1. (a) Schematic view of the sub-5 nm DG ML BP TFET along the armchair transport direction. (b), (c) Transfer characteristics for the 5 nm-gate-long DG ML BP TFETs with (b) source/drain doping concentrations N_s/N_d ($UL_s = 0$ nm) and (c) source underlap length UL_s ($N_s/N_d = 0.3/6 \times 10^{13}$ cm^{-2}) at $V_{bias} = V_{dd} = 0.64$ V. We mark the on-state current for each device as the largest current in each curve. Purple ball, P.

(undoped) semiconductor, DFT-GGA takes the single-electron approximation and tends to underestimate the band gap, while the quasi-particle approach considers the many-body effect and is able to give the exact band gap. However, the many-body effect is strongly depressed in a degenerately doped semiconductor, and the quasi-particle band gap becomes close to the DFT-GGA band gap. For example, the band gap of ML BP of 1.1 eV is calculated at DFT-GGA level, which is independent of the doping concentration. The quasi-particle band gap of the ML BP decreases from 2.2 eV in the undoped state to 1.4 eV at a degenerately doped level [47], and the latter is close to the DFT-GGA value. In an FET, the channel semiconductor is in a degenerately doped state, and hence, the DFT-GGA band gap approximates the quasi-particle band gap in an FET. The reliability of *ab initio* quantum transport simulation in describing the sub-10 nm FET is partially validated from the general agreement of the observed and simulated transfer characteristic for a 1 nm-gate-long thermionic MoS_2 FET [22]. In particular, the calculated SS value is 66 mV dec^{-1} [22], which is nearly equal to the experimental one of 65 mV dec^{-1} [1].

3. Results and Discussions

3.1. Device optimization

ML BP forms a puckered honeycomb layer with an orthorhombic unit cell with the optimized lattice constants of 3.32 and 4.41 Å. The band structure is highly anisotropic with a direct band gap of 1.09 eV at Γ point at PBE level. The electron and hole effective masses are very small ($m_h/m_e = -0.135/0.141\ m_0$) and heavy ($m_h/m_e = -2.445/1.210\ m_0$) along the armchair and zigzag directions, respectively. The lattice constants, band gap and effective masses are all consistent with a previous theoretical study [24]. The device performance should be distinct along the armchair and zigzag directions due to the highly anisotropic electronic structure. We only investigate the armchair transport direction due to the eight orders of magnitude larger current than that along the zigzag direction from previous semi-empirical calculations [23, 24]. We use light and heavy doping concentrations for the source and drain regions (N_s/N_d), respectively, to obtain a lower leakage current and a higher on-state current, as illustrated in our previous work [25]. We first test N_s/N_d and the length of the source underlap region (UL_s, intrinsic region neither biased nor gated) of the ML BP TFET at $L_g = 5$ nm, for the sake of a high on-state current (I_{on}), which is desired to maximize the switching speed. The transfer characteristics of these tested BP TFETs as functions of N_s/N_d (fixed $UL_s = 0$ nm) and UL_s (fixed $N_s/N_d = 3 \times 10^{12}/{-6} \times 10^{13}$ cm^{-2}) are given in figures 1(b) and (c), respectively. Under optimal schemes, the leakage current (I_{leak}, the smallest current, not the off-state current) can meet the ITRS requirements of 0.1 μA μm^{-1} for HP devices, but still higher than the requirements of 5×10^{-5} μA μm^{-1} for the ITRS LP devices, as can be seen in figures 1(b) and (c). We get the on-state current I_{on} at the on-state gate voltage V_g (on), where V_g (on) $= V_g$ (off)$-V_{dd} = V_g$ (off)-0.64 V. The off-state gate voltage V_g (off) is taken at the point where the off-state current I_{off} is 0.1 μA μm^{-1} according to the ITRS HP goal. We marked the

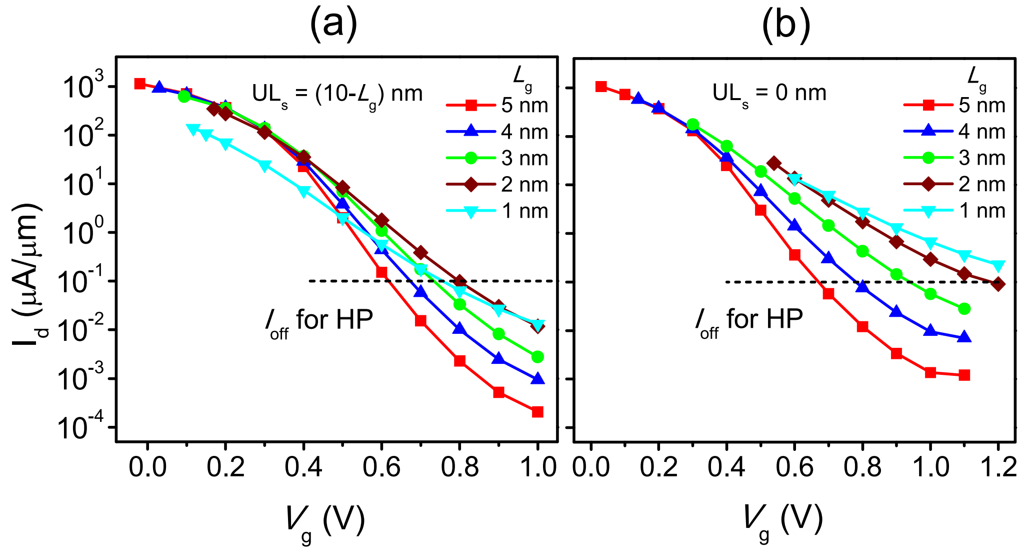


Figure 2. Transfer characteristics for the sub-5 nm DG ML BP TFETs with source underlap length (a) $UL_s = (10 - L_g)$ nm and (b) $UL_s = 0$ nm. We mark the on-state current for each device as the largest current in each curve.

on-state currents as the highest currents marked in figures 1(b) and (c). The lowest $N_s = 8 \times 10^{11}$ cm^{-2} and higher $N_d = -6 \times 10^{13}$ cm^{-2} are the best N_s/N_d for highest I_{on}, as shown in figure 1(b). And the longest $UL_s = 5$ nm is the best choice for the highest I_{on}, as shown in figure 1(c).

We also take ML WTe_2 as channel material to construct a TFET with a 5 nm-long gate for the sake of comparison as it is the best channel material for TFET in the transition-metal dichalcogenide family according to a previous *ab initio* quantum transport simulation based on the flexible plane wave method [32]. The transfer characteristics of these ML WTe_2 TFETs with $L_g = UL_s = 5$ nm as functions of source/drain doping concentrations are given in figure S1(a), which is available online at stacks.iop.org/NANO/29/485202/mmedia in the supporting material. All the I_{leak} are low enough to meet the ITRS LP goals of 5×10^{-5} μA μm^{-1}. We take I_{off} (HP) and I_{off} (LP) at 0.1 and 5×10^{-5} μA μm^{-1} according to the ITRS HP and LP standard, respectively, and the corresponding gate voltage is V_g (off/HP) and V_g (off/LP) for HP and LP application, respectively. Then I_{on} (HP) and I_{on} (LP) are obtained at V_g (on/HP) = V_g (off/HP)−0.64 V and V_g (on/LP) = V_g (off/LP)−0.64 V, respectively. The benchmark I_{on} of ML WTe_2 TFET against the ITRS 2028 requirements for HP and LP devices are given in figures S1(b) and (c), respectively. Reducing N_s has the effect of improving both I_{on}(HP) and I_{on}(LP). But the highest I_{on} (HP) and I_{on} (LP) are only 253 and 77 μA μm^{-1}, respectively, which still cannot meet the ITRS HP and LP 2028 goals.

3.2. Device performance

We take the optimal $N_s/N_d = 8 \times 10^{11}/-6 \times 10^{13}$ cm^{-2} and optimal $UL_s = (10 - L_g)$ nm for the ML BP TFETs with $L_g = 1 \sim 5$ nm in steps of 1 nm. The optimal transfer characteristics are presented in figure 2(a) and those without UL_s are given for comparison in figure 2(b). I_{leak} are in the range of $2.06 \times 10^{-4} \sim 1.32 \times 10^{-2}$ and $1.21 \times 10^{-3} \sim 2.26 \times 10^{-1}$ μA μm^{-1} for the sub-5 nm ML BP TFETs with $UL_s = 10 - L_g$ and $UL_s = 0$ nm, respectively. All the leakage currents except those of the ML BP TFET with $L_g = 1$ nm and $UL_s = 0$ nm can meet the ITRS HP goals of 0.1 μA μm^{-1}, but are still higher than the requirements of 5×10^{-5} μA μm^{-1} for ITRS LP devices. We benchmark the key figures of merit of the sub-5 nm ML BP TFETs against those of the ITRS 2028 requirements for HP transistors and those of the ML WTe_2 TFET with $L_g = 5$ nm in table 1 and figure 3.

I_{on} is one of the critical parameters for a transistor to denote the switching speed of logic transitions. High I_{on} means fast switching, which is beneficial for efficient applications such as HP servers. From figure 3(a), I_{on} monotonously decreases rapidly with the decreased L_g, irrespective of the underlap. I_{on} of the sub-5 nm ML BP TFETs ($1.39 \times 10^2 \sim 1.12 \times 10^3$ μA μm^{-1}) with $UL_s = (10 - L_g)$ nm are always larger than those ($2.50 \times 10^1 \sim 1.06 \times 10^3$ μA μm^{-1}) with $UL_s = 0$ nm, given the same L_g. The improvement of I_{on} by UL_s is generally more effective for shorter L_g. The I_{on} of the ML BP TFETs with $L_g = 5$ nm are 1.12×10^3 and 1.06×10^3 μA μm^{-1} with $UL_s = 5$ and 0 nm, respectively, significantly higher than the value of 9×10^2 μA μm^{-1} for the ITRS 2028 requirements for HP transistors by 24.4% and 17.8%, respectively. The I_{on} of that with $L_g = 4$ nm and $UL_s = 6$ nm is 9.02×10^2 μA μm^{-1}, which can still meet the ITRS 2028 requirements for HP transistors. Remarkably, the I_{on} value of 1.12×10^3 μA μm^{-1} for the ML BP TFETs with $L_g = UL_s = 5$ nm is four times larger than that of 2.53×10^2 μA μm^{-1} for the examined ML WTe_2 counterpart with the same L_g and UL_s at the same calculation level. The I_{on} of the ML WTe_2 TFET with $L_g = 7$ nm is 1890 μA μm^{-1} taken at $V_{ds}/V_{dd} = 0.5/0.7$ V with $I_{off} = 7.5 \times 10^{-2}$ μA μm^{-1} from a previous *ab initio* quantum transport simulation based on the flexible plane wave method [32]. The several times higher I_{on} of the

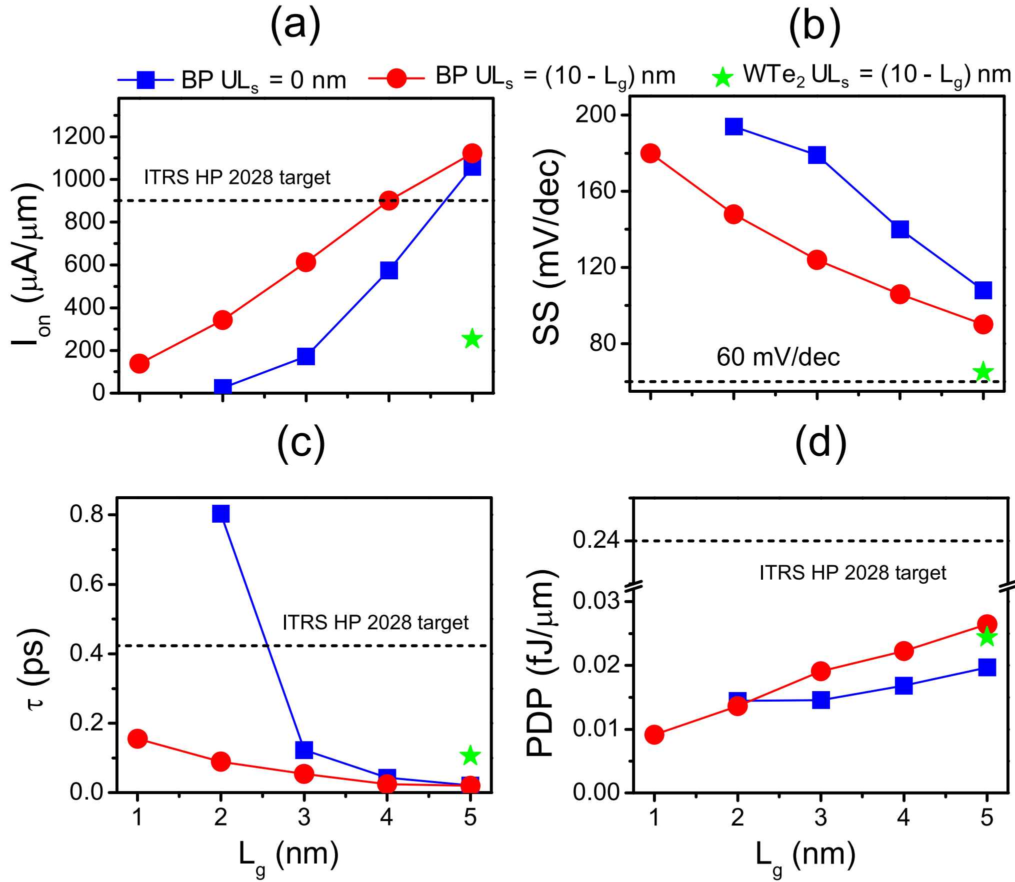


Figure 3. Benchmark of the (a) I_{on}, (b) SS, (c) delay time (τ) and (d) PDP of the sub-5 nm DG ML BP TFETs against the ITRS 2028 requirements for HP devices and the DG ML WTe_2 TFET.

Table 1. Benchmark of the ballistic device performances of the sub-5 nm DG ML BP TFETs against the ITRS 2028 requirements for HP transistors and those of the DG ML WTe_2 TFET with a 5 nm-long gate. N_s/N_d: source/drain doping concentrations; UL_s: source underlap length; EOT: equivalent oxide thickness; V_{dd}: supply voltage; SS: subthreshold swing; C_g: intrinsic gate capacitance; τ: delay time; and PDP: power dissipation. We take the armchair direction as the transport direction and set the EOT = 0.41 nm, $V_{bias} = V_{dd} = 0.64$ V and $I_{off} = 0.1$ μA μm^{-1} for all the calculated devices.

	L_g (nm)	N_s/N_d (10^{13} cm^{-2})	UL_s (nm)	I_{on} (μA/μm)	SS (mV/dec)	C_g (fF/μm)	τ (ps)	PDP (fJ/μm)
ML BP HP	5	0.3/6	5	1113	101	0.036	0.021	0.021
ML BP HP	5	0.08/6	0	1059	108	0.035	0.021	0.020
ML BP HP	5	0.08/6	5	1123	90	0.034	0.020	0.026
ML WTe_2 HP	5	0.08/6	5	253	65	0.042	0.106	0.024
ITRS HP 2028	5.1	—	—	900	—	0.4	0.423	0.24
ML BP HP	4	0.08/6	6	902	106	0.034	0.024	0.022
ML BP HP	3	0.08/6	7	613	124	0.033	0.054	0.019
ML BP HP	2	0.08/6	8	342	148	0.031	0.089	0.014
ML BP HP	1	0.08/6	9	139	180	0.022	0.155	0.009

previous study is due to the plane wave method, where we use an atom orbital basis set.

SS is a key factor to represent the gate control ability, where a smaller value infers a better gate control at the sub-threshold region. The definition is $SS = \frac{\partial V_g}{\partial \lg I_d}$, which indicates the needed V_g to change the drain current I_d by one order of magnitude. From figure 3(b), the calculated SS of the ML BP TFETs increase obviously with the decreased L_g, irrespective of UL_s. The SS of the sub-5 nm BP TFETs (90 ~ 180 mV dec^{-1}) with $UL_s = (10 - L_g)$ nm are always smaller than those (108 ~ 194 mV dec^{-1}) without UL_s, given the same L_g. The ultrashort L_g is responsible for the relatively larger SS, compared to 53 mV dec^{-1} for the one with $L_g = 10$ nm [25]. The minimum SS of 90 mV dec^{-1} of the ML BP TFETs with $L_g = 5$ nm and $UL_s = (10 - L_g)$ nm is not as low as a real TFET (sub-60 mV dec^{-1}). However, this does not affect the device performance, especially the on-state

current. This is because the SS value of the ML BP TFET can be held in a wide drain current scope until the drain current reaches ~100 μA μm^{-1}, which in turn leads to a high on-state current of 1.12 × 10^3 μA μm^{-1}. On the other hand, a smaller minimum SS of 65 μA μm^{-1} is obtained in the examined ML WTe_2 counterpart, but its I_{on} is only 2.53 × 10^2 μA μm^{-1}, which is only one fourth of that of the ML BP one. This is because the SS value of the ML WTe_2 one can be kept until the drain current is ~1 μA μm^{-1}. We compared the SS of the sub-5 nm TFETs without UL_s with those of their metal-oxide-semiconductor field-effect transistor (MOSFET) counterparts in figure S2 in the supporting material. Notably smaller SS of 108 ~ 194 mV dec^{-1} is obtained in the TFET architecture than those of 129 ~ 372 mV dec^{-1} in the MOSFET architecture under the same L_g.

Delay time (τ) is the drain current response time to the gate voltage, and a small τ directly reflects a rapid operation. It is defined as $\tau = \frac{C_g V_{dd}}{I_{on}}$. Here, C_g is the intrinsic gate capacitance, defined as $C_g = \frac{\partial Q_{ch}}{\partial V_g}$, where Q_{ch} is the total charge of the gate region. The calculated τ of the sub-5 nm ML BP TFETs with and without UL_s is given in figure 3(c). The addition of UL_s has the effect of decreasing the delay time when given the same L_g due to the effective enhanced I_{on}. The calculated τ for the 5 nm-gate-long ML BP TFETs with UL_s = 5 nm is 0.020 ps, and it increases almost linearly to 0.156 ps as L_g decreases to 1 nm. Whereas, its counterpart with L_g = 5 nm and no UL_s has a comparable τ value of 0.021 ps, but it first increases to 0.123 ps at L_g = 3 nm and then jumps to 0.804 ps at L_g = 2 nm. The delay times of the ML BP TFETs with $UL_s = (10 - L_g)$ nm are always smaller than the ITRS 2028 requirement of 0.423 ps for an HP device. In particular, the delay times of these TFETs with longer L_g = 4 ~ 5 nm are more than one order of magnitude smaller than that of the ITRS HP 2028 requirement. τ of 0.020 ps for the ML BP TFET with $L_g = UL_s$ = 5 nm is five times smaller than that of the 0.106 ps of its ML WTe_2 counterpart. The quite small τ indicates the fast switching ability of ML BP TFETs even at sub-5 nm scale.

Another major concern for FETs is power consumption, especially when considering integration application at very large scale. We indicate the device power consumption using power dissipation (PDP) per width, which is defined as $\mathrm{PDP} = \frac{(Q_{on} - Q_{off}) \cdot V_{dd}}{W}$. Here, Q_{on}/Q_{off} are the total charges of the gate region under the on-/off-states and W is the channel width. From figure 3(d), the calculated PDPs decrease with the decreased L_g. The PDPs are 0.014 ~ 0.026 fJ μm^{-1} for the ML BP TFETs with $UL_s = (10 - L_g)$ nm, which are generally larger than those of 0.014 ~ 0.020 fJ μm^{-1} with no UL_s, given the same L_g. This is because UL_s has increased the electron transport distance between the two electrodes. The calculated PDPs are one order of magnitude smaller than the ITRS HP 2028 requirement of 0.24 fJ μm^{-1}. The PDP of 0.26 fJ μm^{-1} for the ML BP TFET with $L_g = UL_s$ = 5 nm is slightly larger than that of the 0.024 fJ μm^{-1} of its ML WTe_2 counterpart. It is inspiring that the ML BP TFETs would cost one order of magnitude less switching energy during the one order of magnitude faster switching speed compared to the ITRS HP 2028 target when L_g > 3 nm.

3.3. Discussions

The I_{leak} and I_{on} of the ML BP TFET are three orders of magnitude and four times larger than those of its ML WTe_2 counterpart, respectively. In the two devices, device technical parameters (i.e. EOT, V_{ds} and V_{dd}), gate length, source underlap region length, source/drain doping concentration and calculation parameters are all the same. That is, we take $L_g = UL_s$ = 5 nm and $N_s/N_d = 8 \times 10^{11}/-6 \times 10^{13}$ cm^{-2} in the ML BP and WTe_2 TFETs with EOT = 0.41 nm and $V_{ds} = V_{dd}$ = 0.64 V using the same GGA-PBE functional and Tier 1 basis set. Thus, the distinct values of the I_{leak} and I_{on} of the two TFETs must arise from the intrinsic factors of the two materials. To explore the intrinsic reasons, we compare the projected density of states (PDOS) of the source/drain region, local device density of states (LDDOS) and transmission spectra ($T(E)$) under the off-/on-states for the ML BP and WTe_2 TFETs.

In a TFET, the current comes from BTBT carriers in the bias window due to the overlap of the source valence and drain conduction bands. The hole/electron densities of the source/drain (i.e. PDOS(source/drain)), the band gap E_g of the channel material and the hole/electron effective masses of the source/drain m_h/m_e along the transport direction are three factors that affect the BTBT transmission spectrum, that is $T(E) \propto \mathrm{PDOS(source)} \times \mathrm{PDOS(drain)} \times e^{-\sqrt{E_g}\cdot\sqrt[4]{m_h m_e}}$. We plot the PDOS (source/drain) of the ML BP and WTe_2 TFETs in figures 4(a) and (d), respectively. The PDOS (source/drain) of the ML WTe_2 are larger, which is advantageous for a larger transmission spectrum. The channel E_g determines the BTBT barrier height, and it can be reflected from the LDDOS of the ML BP and WTe_2 TFETs, as shown in figures 4(b) and (e), respectively. As E_g of ML BP and WTe_2 are very similar, having the values of 1.09 and 1.15 eV, respectively, the effect of E_g can be ignored. The effective masses in ML BP and WTe_2 are $m_h/m_e = -0.135/0.141$ m_0 and $-0.514/0.344$ m_0, respectively. The product of m_h and m_e in ML BP is 0.019 m_0^2, nearly one order of magnitude smaller than that of 0.18 m_0^2 in ML WTe_2, which would cause a notably higher BTBT transmission spectrum in ML BP.

The finial BTBT transmission spectrum is a comprehensive effect of PDOS(source) × PDOS(drain), E_g, and $m_h \cdot m_e$ in the transport direction, and $m_h \cdot m_e$ in the transport direction is the dominant factor. Ultimately, we get a larger on-state transmission spectrum in the ML BP TFET than its WTe_2 counterpart (see figures 4(c) and (f)), which leads to a larger on-state current in the ML BP TFET. For HP application, the I_{on} of ML BP and WTe_2 TFETs are 1123 and 253 μA μm^{-1}, respectively, with I_{off} = 0.1 μA μm^{-1} according to the ITRS HP requirement. The choice of ML BP channel can meet the HP required of 900 μA μm^{-1}. On the other hand, the higher effective masses are more beneficial for a lower leakage current [48, 49], and actually the I_{leak} are 1.29 × 10^{-7} and

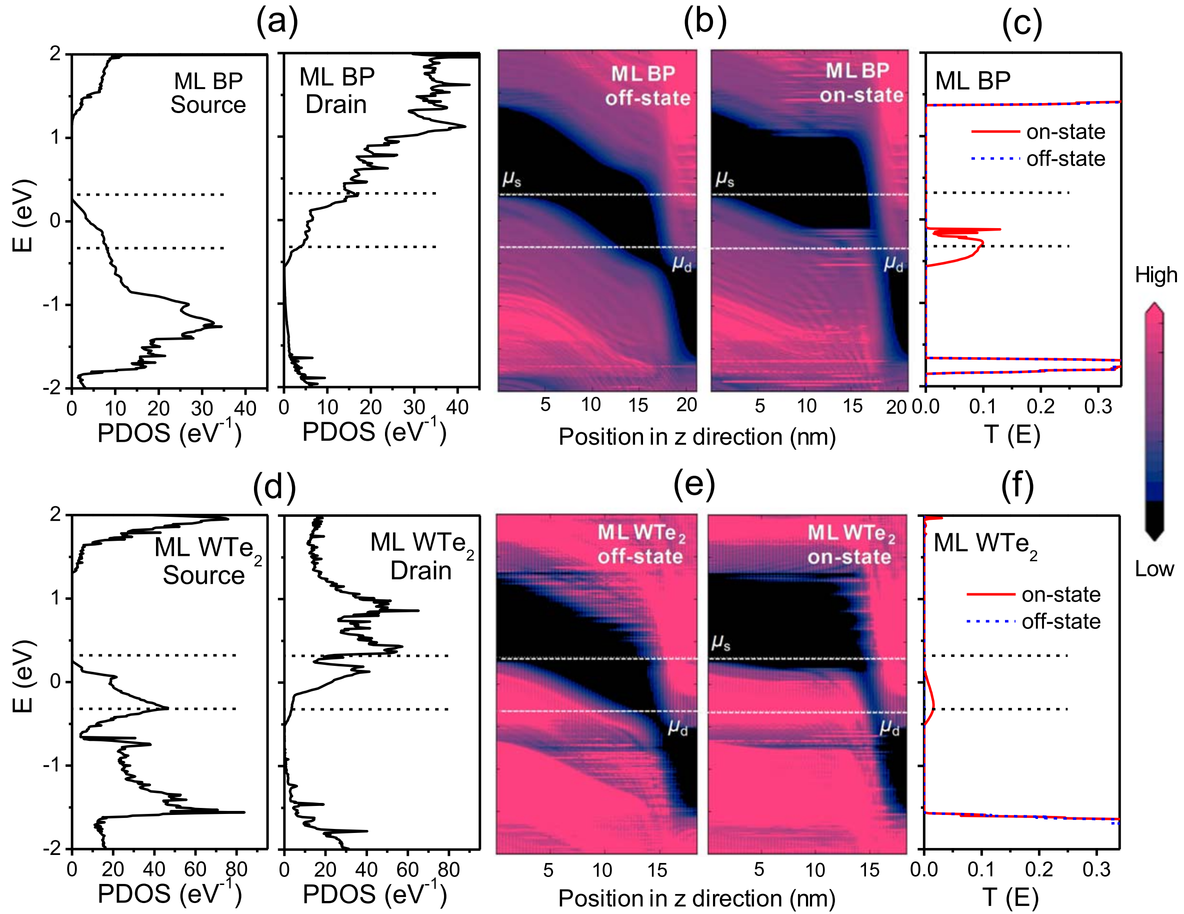


Figure 4. (a), (d) PDOS of the source and drain; (b), (e) LDDOS under off-/on-states; and (c), (f) transmission spectra (T(E)) for the ML BP (a)–(c) and WTe_2 TFETs (d)–(f). In these TFETs, $L_g = UL_s = 5$ nm and $N_s/N_d = 8 \times 10^{11}/-6 \times 10^{13}$ cm^{-2}. The black/white dashed lines indicate the source to drain tunneling window of $V_{ds} = 0.64$ V.

$2.06 \times 10^{-4}\ \mu A\ \mu m^{-1}$ for the ML WTe_2 and BP channels, respectively. Taking $I_{off} = 5 \times 10^{-5} \mu A\ \mu m^{-1}$ according to the ITRS LP requirement, the I_{on} of ML WTe_2 in planar *p-i-n* TFET architecture is 77 $\mu A\ \mu m^{-1}$, which, however, is lower than the LP required of 295 $\mu A\ \mu m^{-1}$. We suggest two possible schemes to explore a TFET for effective LP application. One is finding a desired channel material, probably with a comparable band gap of around 1.0 eV and a product of m_h and m_e between 0.019 m_0^2 (ML BP) and 0.18 m_0^2 (ML WTe_2). The other one is exploring vertical heterojunction architecture instead of planar homojunction architecture.

The transport ability of carriers can be reflected from the transmission eigenvalues and eigenstates at $E = -0.32$ eV under the off-/on-states for the ML BP and WTe_2 TFETs taken at (0, 0) and (0.33, 0) *k*-points, respectively. The two transmission eigenvalues under off-states are almost the same with the values of 4.81×10^{-5} and 6.61×10^{-5} for the ML BP and WTe_2 TFETs, respectively. Whereas the transmission eigenvalue of the ML BP TFET under on-state increases significantly to 1.31×10^{-1}, much larger than that of 2.70×10^{-2} for its ML WTe_2 counterpart. We plot the corresponding off-/on-state eigenstates for the ML BP and WTe_2 TFETs in figures S3(a) and (b), respectively. The incoming electron wavefunctions from the source regions are forbidden to reach the drain regions under the off-states from the upper images. Whereas, the incoming electron wavefunctions from the source region pass through the whole channels to reach the drain regions, and more electron wavefunction is found in the drain region of the ML BP TFET, which represents a larger I_{on} in the ML BP TFET.

We have noted that the source underlap region UL_s has an activity effect that improves the I_{on} of ML BP TFETs: the longer the UL_s, the higher the I_{on}. To explore the function of UL_s, we compare the LDDOS of the ML BP TFETs with $L_g = 2$ nm and different UL_s values in figure 5. For the off-state LDDOS in figure 5(a), the potentials of the gate regions are lifted with the increased UL_s. Correspondingly, the potentials of the gate regions under the on-states are then lifted with increased UL_s as the applied voltages are the same (see figure 5(b)). Therefore, the BTBT barrier heights are reduced with the increased UL_s, which results in a larger I_{on} in the ML BP TFET with longer UL_s, given the same L_g.

3.4. Prospective

The ML BP TFET is a prospective candidate to extend Moore's law down to sub-5 nm scale for HP application with rapid operation. Economical synthesizing methods of ML BP, patterning techniques to 1 nm resolution and feasible doping ways to the electrodes are three key points for the realization of

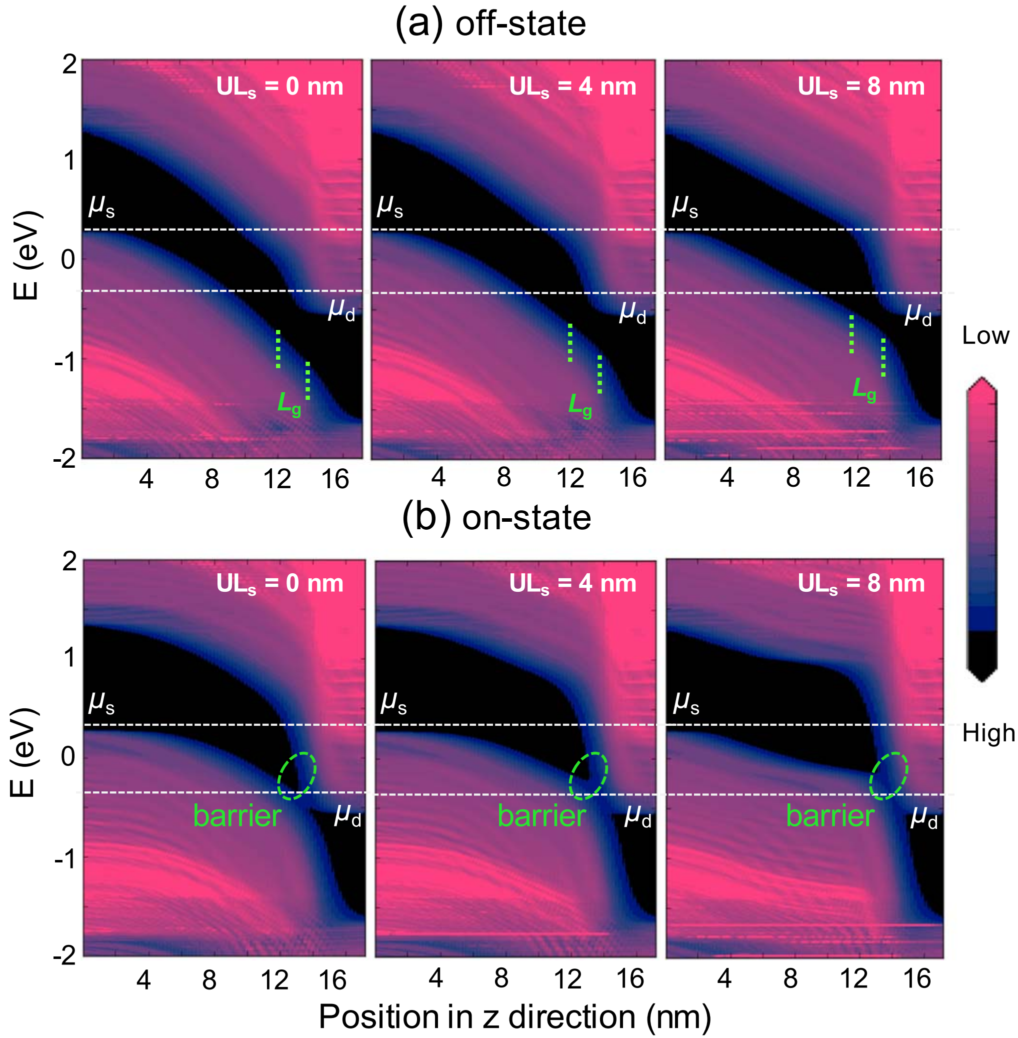


Figure 5. LDDOS of the 2 nm-gate-long DG ML BP TFETs with $UL_s = 0$, 4 and 8 nm under (a) off- and (b) on-states. We marked the gate region and the width of the barrier with dashed green lines and circles, respectively. White dashed lines indicate the source to drain tunneling window of $V_{ds} = 0.64$ V.

the sub-5 nm BP TFETs. ML BP was achieved successfully by mechanical exfoliation in 2014 [17, 19]. Very recently, single crystals of orthorhombic BP samples, several millimeters in size, have been realized with high crystal quality and purity from red phosphorus with Sn/SnI_4 [50], making the large-area preparation of ML BP possible. Till now, the fabricated 2D BP Schottky barrier field-effect transistors (SBFETs) have already been scaled down to 20 nm [20]. Electron beam lithography [51], utilizing the natural dimension of a single-walled carbon nanotube as the gate [1], hydrogen plasma etching [29], directed self-assembly of block copolymer [30] and corrosion cracking along with cleavage plane [31] are all feasible ways for 1 nm patterning resolution. It is inspiring that sub-5 nm 2D SBFETs have been experimentally achieved using the ML MoS_2 channel very recently [1, 29], and scaling the ML BP channel to this patterning resolution is foreseeable in the near future following a similar fabrication technique.

Finding a feasible way to dope ML BP is another key to fabricating ML BP TFETs. Doping the 2D electrodes is distinguished from bulk materials, since the conventional state-of-the-art substitution doping strategy is not suitable for 2D materials. Also, the contact of high/low work function bulk metal is improper because the BP electronic structure usually cannot be kept intact after such metal contacts [52–54]. It is inspiring that several doping strategy techniques, i.e. electrostatic doping [55, 56], surface charge transfer doping [57, 58] and bulk doping [59, 60] have been successfully used to dope BP in labs. Among these, the electrostatic gating technique is the most feasible way because controllable doping levels can be achieved by injecting electrons or holes into the respective bands. Since the techniques for synthesizing ML BP, patterning to sub-5 nm resolution, and doping 2D electrodes, are all practicable, the experimental fabrication of sub-5 nm ML BP TFET is feasible.

4. Conclusions

To summarize, we simulate the device performance limit of the ML BP TFETs at sub-5 nm scale by using *ab initio* quantum transport calculations. With optimal doping concentration and underlap configuration, we predict that the

on-state current of the ML BP TFET at a gate length of 5 nm is four times larger than that of its ML WTe_2 counterpart, which possesses the highest I_{on} among the transition-metal dichalcogenide family. In particular, the I_{on} of the ML BP TFETs can even fulfill the 2028 target of the ITRS for HP devices until the gate length is scaled down to 4 nm. The large on-state current implies a fast switching speed of the sub-5 nm ML BP TFETs. More encouragingly, the delay times and power dissipations surpass the 2028 requirements of the ITRS for HP devices significantly even at an ultimate gate length of 1 nm. The predicted remarkable device performance makes ML BP a competitive candidate for TFETs in the sub-5 nm nodes.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. 11704008 and 11674005), the Basic Scientific Research Foundation of Beijing Municipal Education Commission, the Youth Innovation Foundation of North China University of Technology, the Project of Talent Youth Foundation of North China University of Technology, the National Materials Genome Project (No. 2016YFB0700600) and the National Basic Research Program of China (No. 2013CB932604).

ORCID iDs

Hong Li https://orcid.org/0000-0003-3737-0833

References

[1] Desai S B, Madhvapathy S R, Sachid A B, Llinas J P, Wang Q, Ahn G H, Pitner G, Kim M J, Bokor J and Hu C 2016 MoS_2 transistors with 1-nanometer gate lengths *Science* **354** 99
[2] Ieong M, Doris B, Kedzierski J, Rim K and Yang M 2004 Silicon device scaling to the sub-10 nm regime *Science* **306** 2057
[3] Cavin R K, Lugli P and Zhirnov V V 2012 Science and engineering beyond Moore's law *Proc. IEEE* **100** 1720
[4] Waldrop M M 2016 The chips are down for Moore's law *Nature* **530** 144
[5] Ionescu A M and Riel H 2011 Tunnel field-effect transistors as energy-efficient electronic switches *Nature* **479** 329
[6] Lu H and Seabaugh A 2014 Tunnel field-effect transistors: state-of-the-art *IEEE J Electron Dev. Soc.* **2** 44
[7] Fiori G, Bonaccorso F, Iannaccone G, Palacios T, Neumaier D, Seabaugh A, Banerjee S K and Colombo L 2014 Electronics based on two-dimensional materials *Nat. Nanotechnol.* **9** 768
[8] Chhowalla M, Jena D and Zhang H 2016 Two-dimensional semiconductors for transistors *Nat. Rev. Mater.* **1** 16052
[9] Seabaugh A C and Zhang Q 2010 Low-voltage tunnel transistors for beyond CMOS logic *Proc. IEEE* **98** 2095
[10] Yoon Y J, Seo J H, Cho S, Kwon H I, Lee J H and Kang I M 2016 Sub-10 nm Ge/GaAs heterojunction-based tunneling field-effect transistor with vertical tunneling operation for ultra-low-power applications *J. Semicond. Tech. Sci.* **16** 172
[11] Léonard F and Talin A A 2011 Electrical contacts to one- and two-dimensional nanomaterials *Nat. Nanotechnol.* **6** 773
[12] Allain A, Kang J, Banerjee K and Kis A 2015 Electrical contacts to two-dimensional semiconductors *Nat. Mater.* **14** 1195
[13] Akinwande D, Petrone N and Hone J 2014 Two-dimensional flexible nanoelectronics *Nat. Commun.* **5** 5678
[14] Schwierz F, Pezoldt J and Granzner R 2015 Two-dimensional materials and their prospects in transistor electronics *Nanoscale* **7** 8261
[15] Cao W, Kang J, Sarkar D and Liu W 2015 2D semiconductor FETs—projections and design for sub-10 nm VLSI *IEEE Trans. Electron Dev.* **62** 3459
[16] Kang J, Liu W, Sarkar D, Jena D and Banerjee K 2014 Computational study of metal contacts to monolayer transition-metal dichalcogenide semiconductors *Phys. Rev.* X **4** 031005
[17] Liu H, Neal A T, Zhu Z, Luo Z, Xu X, Tománek D and Ye P D 2014 Phosphorene: an unexplored 2D semiconductor with a high hole mobility *ACS Nano* **8** 4033
[18] Das S, Zhang W, Demarteau M, Hoffmann A, Dubey M and Roelofs A 2014 Tunable transport gap in phosphorene *Nano Lett.* **14** 5733
[19] Li L, Yu Y, Ye G J, Ge Q, Ou X, Wu H, Feng D, Chen X H and Zhang Y 2014 Black phosphorus field-effect transistors *Nat. Nanotechnol.* **9** 372
[20] Miao J, Zhang S, Cai L, Scherr M and Wang C 2015 Ultrashort channel length black phosphorus field-effect transistors *ACS Nano* **9** 9236
[21] Quhe R, Peng X, Pan Y, Meng Y, Wang Y, Han Z, Feng S, Zhang Q, Shi J and Yang J 2017 Can a black phosphorus Schottky-barrier transistor be good enough? *ACS Appl. Mater. Inter.* **9** 3959
[22] Quhe R *et al* 2018 Simulations of quantum transport in sub-5 nm monolayer phosphorene transistors *Phys. Rev. Appl.* **10** 024022
[23] Liu F, Shi Q, Wang J and Guo H 2015 Device performance simulations of multilayer black phosphorus tunneling transistors *Appl. Phys. Lett.* **107** 203501
[24] Chang J and Hobbs C 2015 Theoretical study of phosphorene tunneling field effect transistors *Appl. Phys. Lett.* **106** 2095
[25] Li H *et al* 2018 High-performance sub-10 nm monolayer black phosphorene tunneling transistors *Nano Res.* **11** 2658
[26] Lu S C, Mohamed M and Zhu W 2016 Novel vertical hetero- and homo-junction tunnel field-effect transistors based on multi-layer 2D crystals *2D Mater.* **3** 011010
[27] Feng S, Zhang Q, Yang J, Lei M and Quhe R 2017 Tunneling field effect transistors based on in-plane and vertical layered phosphorus heterostructures *Chin. Phys.* B **26** 421
[28] International Technology Roadmap for Semiconductors (ITRS), 2013, (http://itrs2.net/)
[29] Xie L, Liao M, Wang S, Yu H, Du L, Tang J, Zhao J, Zhang J, Chen P and Lu X 2017 Graphene-contacted ultrashort channel monolayer MoS_2 transistors *Adv. Mater.* **29** 1702522
[30] Nourbakhsh A, Zubair A, Sajjad R N, Amir T K G, Wei C, Fang S, Xi L, Jing K, Dresselhaus M S and Kaxiras E 2016 MoS_2 field-effect transistor with sub-10 nm channel length *Nano Lett.* **26** 7798
[31] Xu K, Chen D, Yang F, Wang Z, Yin L, Wang F, Cheng R, Liu K, Xiong J and Liu Q 2017 Sub-10 nm nanopatterns architecture for 2D materials field-effect transistors *Nano Lett.* **17** 1065
[32] Jiang X W, Luo J W, Li S S and Wang L W 2015 How good is mono-layer transition-metal dichalcogenide tunnel field-effect transistors in sub-10 nm?—an *ab initio* simulation study *Proc. of 2015 IEEE Int. Electron Devices Meeting (IEDM) (Washington, DC, USA)* (https://doi.org/10.1109/IEDM.2015.7409683)

[33] Nikonov D E and Young I A 2015 Benchmarking of beyond-CMOS exploratory devices for logic integrated circuits *IEEE J. Explor. Solid-State Comput. Dev. Circuits* **1** 3

[34] Natarajan S, Agostinelli M, Akbar S and Bost M 2015 A 14 nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 μm 2 SRAM cell size *2014 IEEE Int. Electron Devices Meeting* (https://doi.org/10.1109/IEDM.2014.7046976)

[35] Atomistix ToolKit version 2016.3, QuantumWise A/S (www.quantumwise.com): Copenhagen, Denmark

[36] Brandbyge M 2002 Density-functional method for nonequilibrium electron transport *Phys. Rev.* B **65** 5401

[37] Soler J M, Artacho E, Gale J D, García A, Junquera J, Ordejón P and Sánchez-Portal D 2002 The SIESTA method for ab initio order-N materials simulation *J. Phys. Condens. Matter.* **14** 2745

[38] Perdew J P, Burke K and Ernzerhof M 1996 Generalized gradient approximation made simple *Phys. Rev. Lett.* **78** 3865

[39] Monkhorst H J 1976 Special points for Brillouin-zone integrations *Phys. Rev.* B **13** 5188

[40] Datta S 1995 *Electronic Transport in Mesoscopic Systems* (Cambridge: Cambridge University Press)

[41] Çakır D and Peeters F M 2014 Dependence of the electronic and transport properties of metal-$MoSe_2$ interfaces on contact structures *Phys. Rev.* B **89** 245403

[42] Wang Y, Yang R X, Quhe R, Zhong H, Cong L, Ye M, Ni Z, Song Z, Yang J and Shi J 2015 Does p-type ohmic contact exist in WSe_2-metal interfaces? *Nanoscale* **8** 1179

[43] Zhong H, Quhe R, Wang Y, Ni Z, Meng Y, Song Z, Pan Y, Yang J, Li Y and Ming L 2016 Interfacial properties of monolayer and bilayer MoS_2 contacts with metals: beyond the energy band calculations *Sci. Rep.* **6** 21786

[44] Wang Y, Fei R, Quhe R, Li J, Zhang H, Zhang X, Shi B, Xiao L, Song Z and Yang J 2018 Many-body effect and device performance limit of monolayer InSe *ACS Appl. Mater. Inter.* **10** 23344

[45] Wang Y, Huang P, Ye M, Quhe R, Pan Y, Zhang H, Zhong H, Shi J and Lu J 2017 Many-body effect, carrier mobility, and device performance of hexagonal arsenene and antimonene *Chem. Mater.* **29** 2191

[46] Sun X, Song Z, Liu S, Wang Y, Li Y, Wang W and Lu J 2018 Sub-5 nm monolayer arsenene and antimonene transistors *ACS Appl. Mater. Interfaces* **10** 22363

[47] Gao S and Yang L 2017 Renormalization of the quasiparticle band gap in doped two-dimensional materials from many-body calculations *Phys. Rev.* B **96** 155410

[48] Sylvia S S, Park H H, Khayer M A, Alam K, Klimeck G and Lake R K 2012 Material selection for minimizing direct tunneling in nanowire transistors *IEEE Trans. Electron Dev.* **59** 2064

[49] Cao W, Sarkar D, Khatami Y, Kang J and Banerjee K 2014 Subthreshold-swing physics of tunnel field-effect transistors *AIP Adv.* **4** 133116

[50] Akhtar M, Anderson G, Zhao R, Alruqi A, Mroczkowska J E, Sumanasekera G and Jasinski J B 2017 Recent advances in synthesis, properties, and applications of phosphorene *npj 2D Mater. Appl.* **1** 5

[51] Manfrinato V R, Stein A, Zhang L, Nam C Y, Yager K G, Stach E A and Black C T 2017 Aberration-corrected electron beam lithography at the one nanometer length scale *Nano Lett.* **17** 4562

[52] Pan Y, Wang Y, Ye M, Quhe R, Zhong H, Song Z, Peng X, Yu D, Yang J and Shi J 2016 Monolayer phosphorene–metal contacts *Chem. Mater.* **28** 2100

[53] Pan Y, Yang D, Wang Y, Meng Y, Han Z, Quhe R, Zhang X, Li J, Guo W and Li Y 2017 Schottky barriers in bilayer phosphorene transistors *ACS Appl. Mater. Interfaces* **9** 12694

[54] Zhang X *et al* 2018 Three-layer phosphorene-metal interfaces *Nano Res.* **11** 707

[55] Buscema M, Groenendijk D J, Steele G A, Hs V D Z and Castellanosgomez A 2014 Photovoltaic effect in few-layer black phosphorus p-n junctions defined by local electrostatic gating *Nat. Commun.* **5** 4651

[56] Robbins M C and Koester S J 2017 Black phosphorus p- and n-MOSFETs With electrostatically doped contacts *IEEE Electron Dev. Lett.* **38** 285

[57] Xiang D, Han C, Wu J, Zhong S, Liu Y, Lin J, Zhang X A, Ping H W, Barbaros Ö and Neto A H 2015 Surface transfer doping induced effective modulation on ambipolar characteristics of few-layer black phosphorus *Nat. Commun.* **6** 6485

[58] Koenig S P, Doganov R A, Seixas L, Carvalho A, Tan J Y, Watanabe K, Taniguchi T, Yakovlev N, Neto A H C and Özyilmaz B 2016 Electron doping of ultrathin black phosphorus with Cu adatoms *Nano Lett.* **16** 2145

[59] Yang B, Wan B, Zhou Q, Wang Y, Hu W, Lv W, Chen Q, Zeng Z, Wen F and Xiang J 2016 Te-doped black phosphorus field-effect transistors *Adv. Mater.* **28** 9408

[60] Xu Y, Yuan J, Fei L, Wang X, Bao Q, Wang Y, Zhang K and Zhang Y 2016 Selenium-doped black phosphorus for high-responsivity 2D photodetectors *Small* **12** 5000