doi:10.1038/nature15387

S1. ADVANTAGES OF 2D MATERIALS for ULTRA SHORT-CHANNEL FIELD-EFFECT TRANSISTORS

In a field-effect-transistor (**Fig. S1(a)**), current flows through a semiconducting channel between the source and drain electrodes. The current flow is controlled by a third electrode called gate, which is capacitively connected to the channel and hence, the gate can turn the device ON or OFF. However, as the channel length is scaled to deep sub-micron regime, electric fields from the source and drain begin to influence the potential of the channel and the gate can no longer effectively control the channel potential, resulting in degradation of device electrostatics. The electric field from the drain lowers the barrier in the channel close to the source-channel junction and this phenomenon is called Drain Induced Barrier Lowering (DIBL) (**Fig. S1(b)**). DIBL can lead to significant current flow even when the device is supposed to be OFF. This increase in OFF-current (or leakage) leads to increase in undesirable static power dissipation, which is a burning issue for the IT Industry. Hence, it is of absolute necessity to improve the gate control over the channel.

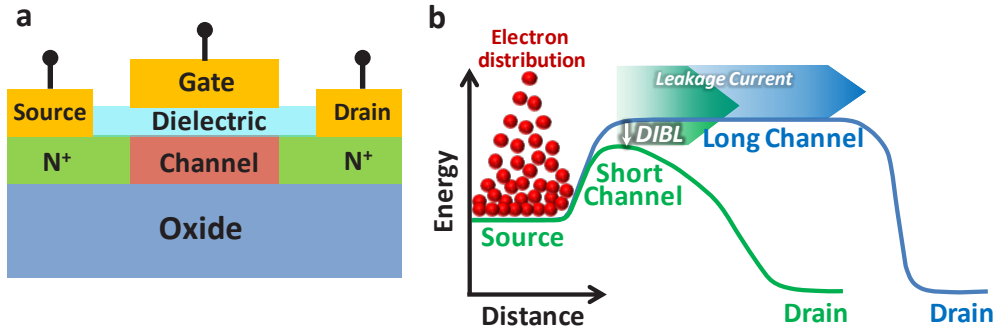


Fig. S1. (**a**) The schematic diagram of a conventional FET (CFET). (**b**) Conduction band profile of a conventional FET for both long and short channel lengths.

A parameter that indicates how efficiently the gate controls the channel is the natural length scale (λ),[1] which depends on the gate capacitance as well as the permittivity and thickness of the semiconducting channel. A rule of thumb is that in order for the gate to maintain efficient electrostatics, λ should be 5-10 times smaller than the channel length (L_g).[2] The expressions for λ for Silicon-ON-Insulator (SOI) MOSFET structure is given by

$$\lambda_{SOI} = \sqrt{\frac{\varepsilon_s}{C_g} t_s} \qquad \text{(s1)}$$

where t_s and ε_s are the thickness and permittivity of the semiconducting channel respectively, C_g is the gate capacitance per unit area. It is evident from eqn s1, that reducing the channel thickness can help in reducing the λ and hence, improve the device electrostatics. 2D layered materials are highly promising in this respect.[3] Since the adjacent layers of the 2D materials have van der Waal's bonding, it is relatively easy to obtain as thin as a single atomic layer of these materials. Moreover, they are characterized by pristine, dangling-bond free interfaces. In addition, they offer easy patternability due to their planar nature[4] compared to 1D structures such as nanowires and nanotubes. Moreover, thinning down conventional 3D materials to nanometer-scale dimensions is not only process-wise challenging; such structures suffer from interfaces with trap states and worse (w.r.t 2D materials) mobility degradation (**Fig. S2**). Hence, the 2D semiconducting materials are considered potential candidate as channel material for beyond Si scaling era.[5]

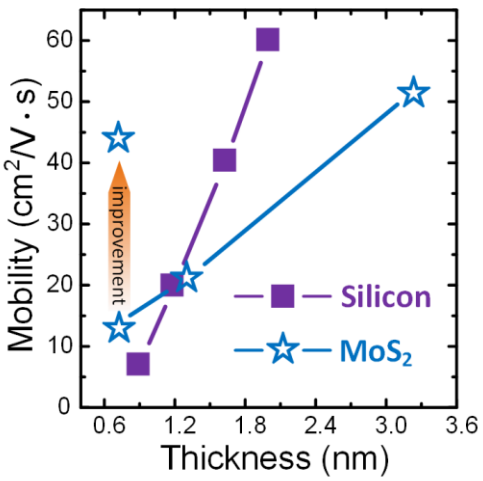


Fig. S2. This figure shows how mobility degrades with material thickness. The mobility of Si[6] degrades much faster than that of MoS_2[7]. At 0.9 nm, it is only around 6-7 $cm^2/V{\cdot}s$. While the 0.65 nm thick 1L MoS_2 has a mobility of around 12 $cm^2/V{\cdot}s$. Recently, this value has been boosted to 44 $cm^2/V{\cdot}s$ by W. Liu et al.[8] In general, 2D materials pay much less for being thin compared to bulk materials.

While graphene is the most widely known 2D material, the lack of an intrinsic bandgap and the extreme difficulties in engineering the same, makes it unsuitable for digital applications. In this respect, another class of 2D layered materials in the form of Transition-Metal-Dichalcogenides (TMDs), which possess finite bandgaps are highly attractive.

S2. FUNDAMENTAL SUBTHRESHOLD SWING LIMITATION IN CONVENTIONAL FETs - IMPLICATIONS FOR SUPPLY VOLTAGE SCALING AND POWER DISSIPATION

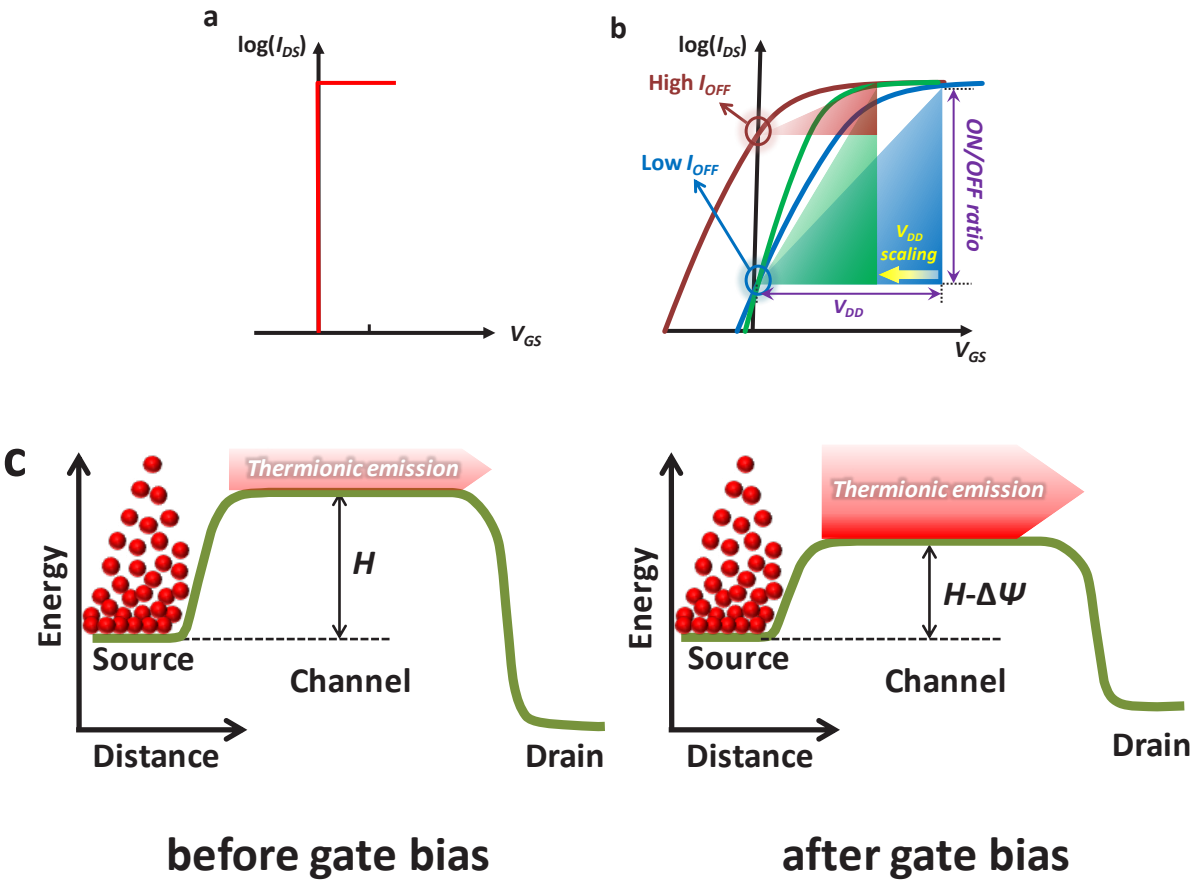


Fig. S3. (**a**) Ideal step-like log(I_{DS})-V_{GS} characteristics. (**b**) In reality, I_{DS}-V_{GS} characteristics have finite slopes and look like the blue curve. If we try to scale down the supply voltage (V_{DD}) and hence, make the device turn on with a lower V_{GS}, the characteristics will look like the brown curve leading to significantly higher leakage current and low ON-OFF current ratio. In order to reduce V_{DD}, and at the same time achieve high ON-OFF current ratio, it is necessary to make the slope of the I_{DS}-V_{GS} curve steeper or the inverse of the slope (defined as the Subthreshold Swing) smaller, as illustrated by the green curve. (**c**) Band diagrams before and after application of gate potential illustrating the transport mechanism in conventional FETs. Only electrons with energy higher than the barrier height can transport to the drain. H is the height of the barrier before application of gate potential and $\Delta\psi$ is change in surface potential of the semiconductor channel after application of gate potential.

Lowering the power in electronic components is of critical importance for energy-efficiency leading to reduction in Green House Gases, increasing battery life of portable electronics and implanted medical devices as well as solving heat dissipation issues in mobile and non-portable systems. The most effective way to control the power density is to scale down the supply voltage (V_{DD}) as the dynamic power displays a quadratic dependence on V_{DD}. Ideally we would prefer a transistor with very abrupt step like transfer characteristics (**Fig. S3a**) such that with the application of an infinitesimally small gate voltage, the device can be turned ON from the OFF state. However, in conventional FETs (CFETs) there is a fundamental limitation of the abruptness of the turn-ON characteristics (**Fig. S3b**) and they need the gate voltage to be changed by at least 60 mV to cause a corresponding change of the drain current by one decade (or 10X). This implies that in order to achieve an ON-OFF ratio of 4 decades, supply voltage of at least 240 mV (4*60 mV) is required. Hence, in case of CFETs, it is not possible to reduce the supply voltage and still maintain reasonable ON-OFF ratio.

The reason behind this fundamental limitation is discussed below. The abruptness of the turn-ON behavior of the FET is characterized by a parameter called the Subthreshold Swing (*SS*), which is defined as the inverse of the subthreshold slope and is given by $SS = (d\log_{10} I_{DS} / dV_{GS})^{-1}$ where I_{DS} : drain current, V_{GS}: gate to source voltage. The *SS* is dependent on two factors as shown by the two terms in eqn. s2.

$$SS = \frac{\partial V_{GS}}{\partial \psi} \quad \frac{\partial \psi}{\partial (\log_{10} I_{DS})} \qquad \text{(s2)}$$

First term is determined by the device electrostatics, or in other words, how well the surface potential of the semiconductor (ψ) at the interface of the semiconductor and gate-dielectric is modulated by the gate voltage and is given by the ratio of the change in V_{GS} to the change in ψ. Even in the case of perfect electrostatics, the lowest possible value of the first factor is 1, which means a particular change in gate voltage (ΔV_{GS}) leads to almost similar change in surface potential of the semiconductor channel ($\Delta\psi$) such that $\Delta\psi = \Delta V_{GS}$. The second term affecting *SS*, is determined by how efficiently the current is modulated by the change in the surface potential. The lowest possible value of the second factor is 60 mV/dec at room temperature for conventional FETs as explained below. The transport mechanism (**Fig. S3c**) of the conventional FETs is based on thermionic emission over the barrier, which implies only electrons having energy higher than the barrier height can transport to the drain and contribute to current. Hence, current can be increased (or decreased) by decreasing (or increasing) the barrier height *H*. Since the electrons in the source is distributed according to the Fermi-Dirac distribution, which can be approximated as

Boltzmann distribution for energies much higher than the Fermi level, the occupied density of states, or equivalently the electron density per unit energy (n) decreases exponentially with increase in energy (E) ($n \propto e^{-E/k_BT}$). Thus, with a decrease in surface potential by $\Delta\psi$, and hence, decrease in barrier by the same amount, current will increase following the trend $I_{DS} \propto e^{\Delta\psi/k_BT}$. Hence, the minimum value for the second factor, can be calculated as [$K_BT/q\ ln(10)$], which has the value of 60 mV/dec at room temperature. Thus, combining the minimum values for the first and second term, the limitation of *SS* in case of conventional FETs comes out to be 60 mV/decade. Note that this minimum value of 60 mV/dec for *SS*, is obtained even in presence of perfect electrostatics. Degradation of device electrostatics will degrade (increase) the *SS* further.

S3. WHY SINGLE CARRIER TUNNELING BARRIER CANNOT LEAD TO SUBTHERMIONIC (< 60 mV/decade at Room Temperature) SUBTHRESHOLD SWING

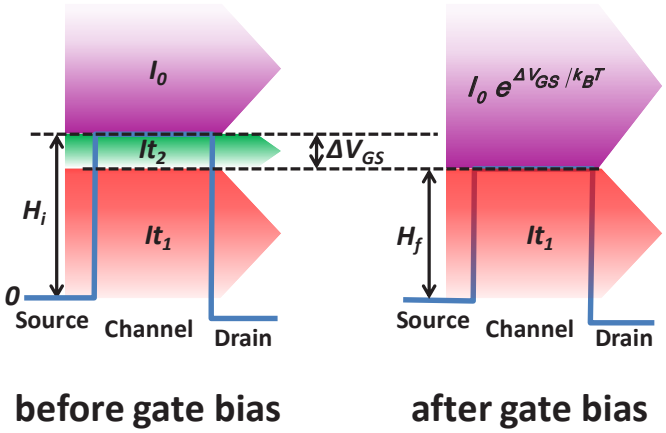


Fig. S4. Band diagrams of a barrier in which carriers can transport not only through thermionic emission above the barrier but they can also tunnel through the barrier, before and after application of gate voltage (V_{GS}).

In the previous section, we discussed that for transport based on thermionic emission over the barrier, the minimum *SS* achievable is $K_BT/q\ ln(10)$. Here, we show that, this is true also for any barriers, which involve combination of single carrier tunneling (but not Band-to-Band-tunneling) and thermionic emission. By single carrier tunneling, we mean situations where either electrons or holes are involved in tunneling but not both. On the other hand, Band-to-Band- tunneling involves both electrons and holes[9]. A single carrier tunneling barrier is shown in **Fig. S4,** which, for example, can be achieved by vertically stacking a 2D material (with non-zero bandgap) between two layers of graphene. This barrier can be thought to be similar to that shown in **Fig. S3c**, but the only difference is that

the width of the barrier is small enough so that apart from transmission over the barrier through thermionic emission, electrons with energy lower than the barrier height, can also tunnel directly from source to drain. Initially, before application of gate voltage, let the barrier height be H_i, and after increasing the gate voltage by ΔV_{GS}, the height is reduced to H_f (where $H_f = H_i - \Delta V_{GS}$ in the best case assuming perfect electrostatics). Let us divide the initial current (I_i) into 3 components: let the current above the barrier through thermionic emission be I_0, tunneling current through the barrier from energy *0* to H_f be It_1 and that from energy H_f to H_i be It_2 (**Fig. S4**). After the application of gate voltage, the final current (I_f) can be thought to be comprised of two components, current above the barrier, which is increased compared to the initial thermionic emission current and is given by $I_0 e^{\Delta V_{GS}/k_B T}$ (as explained in previous section) and the tunneling current through the barrier given by It_1. For small change in gate voltage, the *SS* can be written as:

$$SS = \frac{\Delta V_{GS}}{\log_{10}\left(\frac{I_f}{I_i}\right)} = \frac{\ln(10)\,\Delta V_{GS}}{\ln\left(\frac{I_f}{I_i}\right)} \qquad \text{(s3)}$$

Now, inserting the current components for I_i and I_f in eqn (s3), we get,

$$SS = \frac{\ln(10)\Delta V_{GS}}{\ln\left(\frac{I_0 e^{\Delta V_{GS}/k_B T} + I_{t1}}{I_0 + I_{t1} + I_{t2}}\right)} \qquad \text{(s4)}$$

$$\text{Now, since } \frac{I_0 e^{\Delta V_{GS}/k_B T} + I_{t1}}{I_0 + I_{t1} + I_{t2}} < \frac{I_0 e^{\Delta V_{GS}/k_B T}}{I_0} \qquad \text{(s5)}$$

$$\text{we get, } \frac{\ln(10)\Delta V_{GS}}{\ln\left(\frac{I_0 e^{q\Delta V_{GS}/k_B T} + I_{t1}}{I_0 + I_{t1} + I_{t2}}\right)} > \frac{\ln(10)\Delta V_{GS}}{\ln\left(\frac{I_0 e^{q\Delta V_{GS}/k_B T}}{I_0}\right)} = \ln(10)\frac{k_B T}{q} \qquad \text{(s6)}$$

This implies that, in the case when single carrier tunneling is present along with thermionic emission, not only it is fundamentally impossible to achieve *SS* lower than $K_B T/q\ ln(10)$, but in fact, the *SS* is further degraded from this minimum value.

Even, in case of a Schottky barrier (**Fig. S5**), it is not possible to achieve *SS* lower than the fundamental limit of MOSFET. Solomon has showed using simple mapping technique, that it is not possible to achieve lower *SS* than

$K_BT/q\ ln(10)$ when single carrier tunneling is involved. His methodology is valid for both Schottky barriers as well as direct source-drain tunneling. Readers are referred to his paper[10] for further details of the methodology.

Note, that in case of band-to-band tunneling[11], where carriers tunnel from the valence to the conduction band, it is possible to achieve *SS* below $K_BT/q\ ln(10)$ as discussed in the main manuscript.

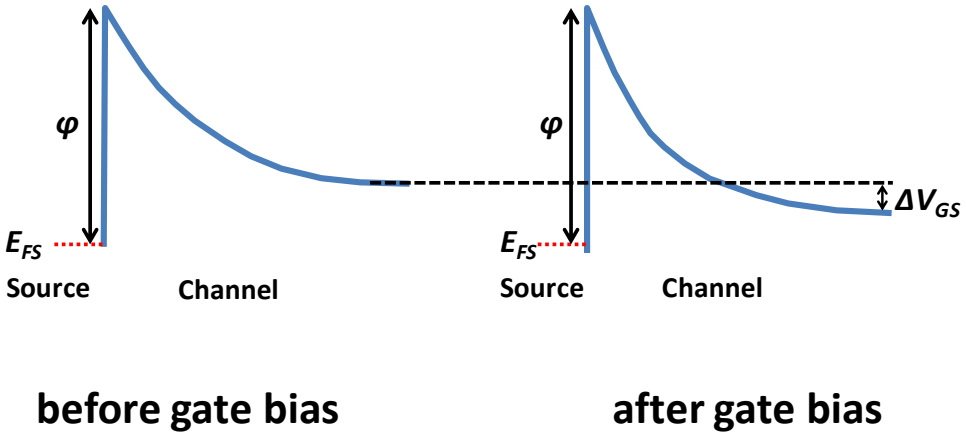


Fig. S5. Band diagrams of Schottky barrier at the source-channel junction, before and after application of gate potential. φ denotes the Schottky barrier height and E_{FS} denotes the Fermi level of the source.

S4. MoS_2 SYNTHESIS METHOD

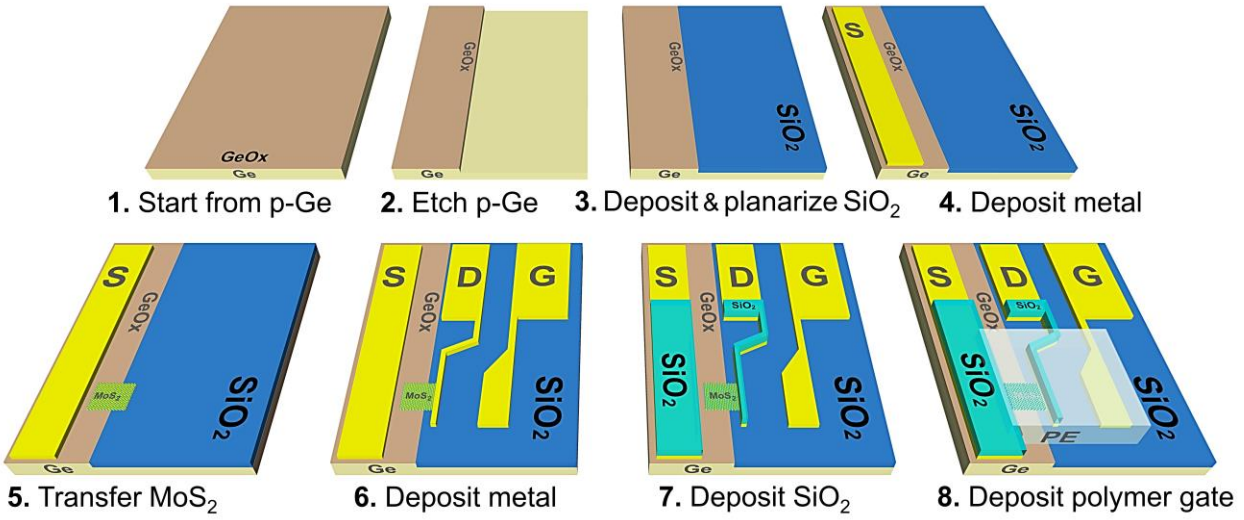
We used chemical vapor deposition (CVD) method to synthesize the MoS_2 layers instead of using mechanical exfoliation, in order to enable a scalable technology. Sulfur (S) and Molybdenum Oxide (MoO_3) powder were used as the S and Mo precursor, respectively.[12] MoO_3 powder contained in a boat was inserted in a fused quartz tube and placed at the center of the CVD furnace. S powder was placed at the upstream region of the furnace, which had lower temperature. A clean Si wafer with thermally grown SiO_2 (275 nm) on top was used as the substrate for MoS_2 growth and it was put above the MoO_3 powder with the SiO_2 layer facing down. The furnace temperature was increased to 850 $^{\circ}$C in 20 min, and was held at that temperature for 10 min, leading to the formation of bilayer MoS_2. The temperature of sulfur was kept at about 200 $^{\circ}$C. Throughout the process, 50 sccm of argon was used as the carrier gas and the growth was carried out under atmospheric pressure. The high resolution Scanning Transmission Electron Microscopy (STEM) images were recorded at 300 KV to investigate the MoS_2 samples.

S5. DETAILS OF DEVICE FABRICATION

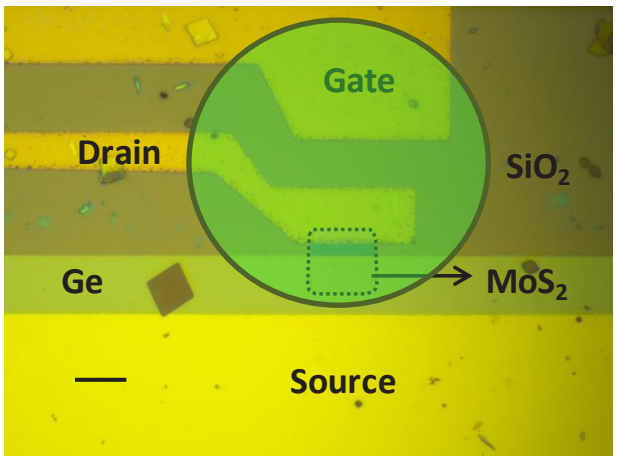
The fabrication process flow for ATLAS-TFET is shown in **Fig. S6(a)**. We started with a commercially available Ge wafer, highly doped (p-type) with Ga and having a resistivity of 0.0011 ohm.cm. The wafer was patterned using photolithography and 300 nm trenches were etched using deep reactive ion etching. Subsequently, without removing the photoresist, about 350 nm of SiO_2 was grown using ebeam deposition. Next, the photoresist was removed using acetone upon which oxide only remained in the trenches and was washed away (with the resist) from the other regions. After that, very dilute HF was used to slowly etch out the SiO_2 in the trenches, to planarize it with respect to the surrounding Ge. This made the substrate ready for device fabrication. MoS_2 synthesized using CVD was then transferred to the engineered substrate using PMMA transfer process and then the PMMA was removed using acetone. MoS_2 was etched from the regions where devices were not intended to be built using Inductively Coupled Plasma (ICP) etching based on BCl_3 (15 sccm) and Ar (60 sccm) at pressure of 0.6 Pa, RF source power of 100 W and RF bias power of 50 W[13] for 1 min. Drain contacts were formed using ebeam lithography and subsequent metallization using Yttrium/Gold (Yttrium was chosen due to its low workfunction of 3.1 eV and presence of d-orbitals which can lead to efficient contacts[14–16]). As the drain region of MoS_2 is undoped, a parasitic contact resistance occurs in series in the current path. In future, reducing the parasitic resistance through doping of the drain region, can lead to further performance improvement. In our work, we have used solid polymer electrolyte (PE) gating. The drain electrode is covered with SiO_2 to prevent the drain voltage from influencing the electric double layer[17] formed in the polymer. The gate electrode is formed using ebeam lithography and metallization of 20 nm/ 50 nm Yttrium/Gold using ebeam deposition. For polymer electrolyte gating, the gate electrode does not need to be directly above the channel and is formed on the side as shown in **Fig. S6.** The PE is formed by mixing poly(ethylene oxide) (PEO) and lithium perchlorate (LiClO4*)* in 8:1 ratio in methanol and it can be drop casted or even lithographically patterned[18]. The dielectric constant for the PEO matrix has been reported in literature as 5.[19] A fabricated prototype device is shown in **Fig. S6(b)**.

Note that a p-type ATLAS-TFET can also be achieved on the same substrate using highly n-doped Ge source and WSe_2 with low electron affinity as the channel. Moreover, it is possible to build large scale circuits with complimentary ATLAS-TFETs in a planar platform. By starting with a undoped Ge wafer and selectively doping certain portions of it as p-type and other portions as n-type and then transferring in-plane heterostructures of MoS_2 and WSe_2 to it, it is possible to build both p and n-type ATLAS-TFETs for integrated circuits. The recent

experimental demonstration[20] of CVD synthesis of in-plane heterostructures of MoS_2 and WSe_2, indicate the feasibility of the growth of such structures. It is possible to build the complimentary circuits even without the use of the in-plane heterostructures by employing a 2-step procedure in which large scale MoS_2 can be first transferred to the substrate and patterned followed by the transfer and patterning of large scale WSe_2.



(**a**)



(**b**)

Fig. S6 (**a**) Schematic diagram showing fabrication process flow of ATLAS-TFET. (**b**) The ATLAS-TFET device showing the source, drain and gate contacts. The circled region represents the solid polymer electrolyte. The covering of drain electrode with oxide has not been shown in this figure for clarity. Scale bar: 10 μm.

S6. DETAILS OF X-RAY PHOTOELECRON SPECTROSCOPY

X-ray Photoelectron Spectroscopy (XPS) can be used to investigate the electronic structure of the materials. It involves knocking out electrons from the surface of a material using X-rays (with energy = $E_{X\text{-}ray}$) and measuring the kinetic energy ($E_{kinetic}$) of the emitted electrons. Then the binding energy (BE) of the electrons can be calculated as:

$$BE = E_{X\text{-}ray} - (E_{kinetic} + \phi) \qquad \text{(s7)}$$

where ϕ is the work function. In case of a junction formed of two different materials, XPS can indicate the presence of band bending through the shift of the BE. Moreover, the direction of the shift of BE can tell whether the band bending is due to positive (shift towards higher BE) or negative (shift towards lower BE) charge.

In case the sample for XPS is conducting, the photoelectrons emitted from the material are compensated by the electrons from the sample holder. However, if the sample is insulating, charge compensation cannot occur and there can be build-up of positive charges due to knocking out of electrons, which can shift the binding energies to higher values than actual. This is called charging effect. In the XPS setting, a charge neutralizer, which provides extra electrons to the sample, can be used. In the case of insulating substrate, use of charge neutralizer (CN), shifts the binding energies to lower values than actual, due to the presence of negative charges due to extra electrons. By doing the XPS measurements with and without the CN, it can be found out whether there is any charging effect in the sample. If the BE does not change, with and without the CN, then it can be confirmed that charging is absent while if the BE changes, then charging effect is present. In our case, the MoS_2 on Ge did not show any charging effect as both Ge and MoS_2 are conducting as confirmed by measurements with and without CN. However, the MoS_2 on SiO_2/Si substrate suffers from charging effect due to the presence of 280 nm thick insulating SiO_2. To mitigate this effect, and find the actual BE, we used the procedure shown by Alay et. al.[21] First, XPS measurements were performed with and without the CN and the shift in BE for both Oxygen (O) and Molybdenum core levels were derived and found to be similar. This implies that charging effect leads to similar shift in MoS_2 and SiO_2 core levels. Next, Si with ultra-thin 1 nm SiO_2 was measured with and without CN, and no charging effect was found. The actual BE of O core level can be found from this experiment as charging is absent. The error in BE due to charging was calculated by subtracting this actual BE of O from the BE of O obtained from XPS of 280 nm SiO_2 on Si without CN. The core levels of MoS_2 measured in the case of MoS_2 on SiO_2/Si substrate without CN, were corrected for this error to obtain the actual BE.

In **Fig. 2i** of the main manuscript, we have shown the comparison of Ge 3d core level for just Ge and for Ge with MoS_2 but have omitted the deconvolution of the level for the sake of clarity. Here, the deconvolution of the Ge 3d core level is shown (**Fig. S7**).

XPS data were measured using a Kratos Axis Ultra DLD system (monochromated Al k-alpha radiation 1486 eV). While survey scans were measured using pass energy of 80 eV, high-resolution scans were run using 20 eV.

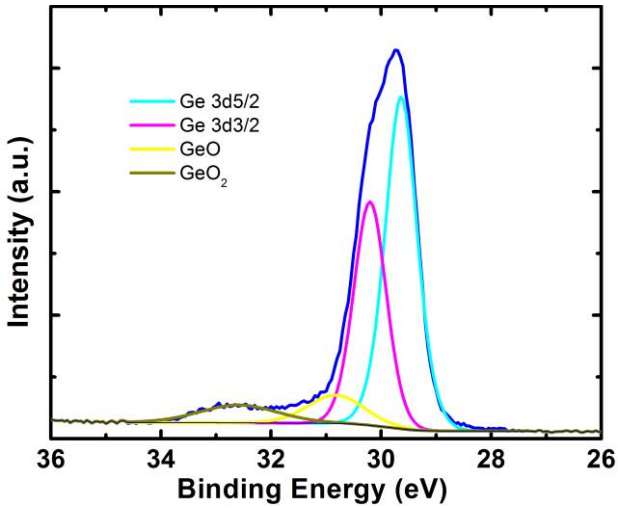


Fig. S7. The Ge core level (blue) deconvoluted to show Ge 3d5/2 (cyan), Ge 3d3/2 (magenta), GeO (yellow) and GeO_2 (dark yellow).

S7. NDR AS A SIGNATURE OF BAND-TO-BAND-TUNNELING (BTBT)

Here, we explain why we see a trend towards Negative Differential Resistance (NDR) in the forward bias characteristics of the p-n junction measurements. For that, we first explain the case where we can observe prominent NDR[22]. Prominent NDR can be observed in a p-n junction when both the p and n regions are highly doped. The current in forward bias condition of such a p-n junction as a function of voltage applied to the p-region with respect to the n-region is shown in **Fig. S8**. The band diagrams corresponding to different points in the curve marked with numbers from 1 to 5 are also shown. When the applied voltage is zero, the Fermi levels in p and n regions align and there is no current flow (point 1, **Fig. S8**). When a positive bias is applied to the p region, the electrons from the filled states of n region can tunnel into the empty states of p-region leading to current flow (point 2, **Fig. S8**).

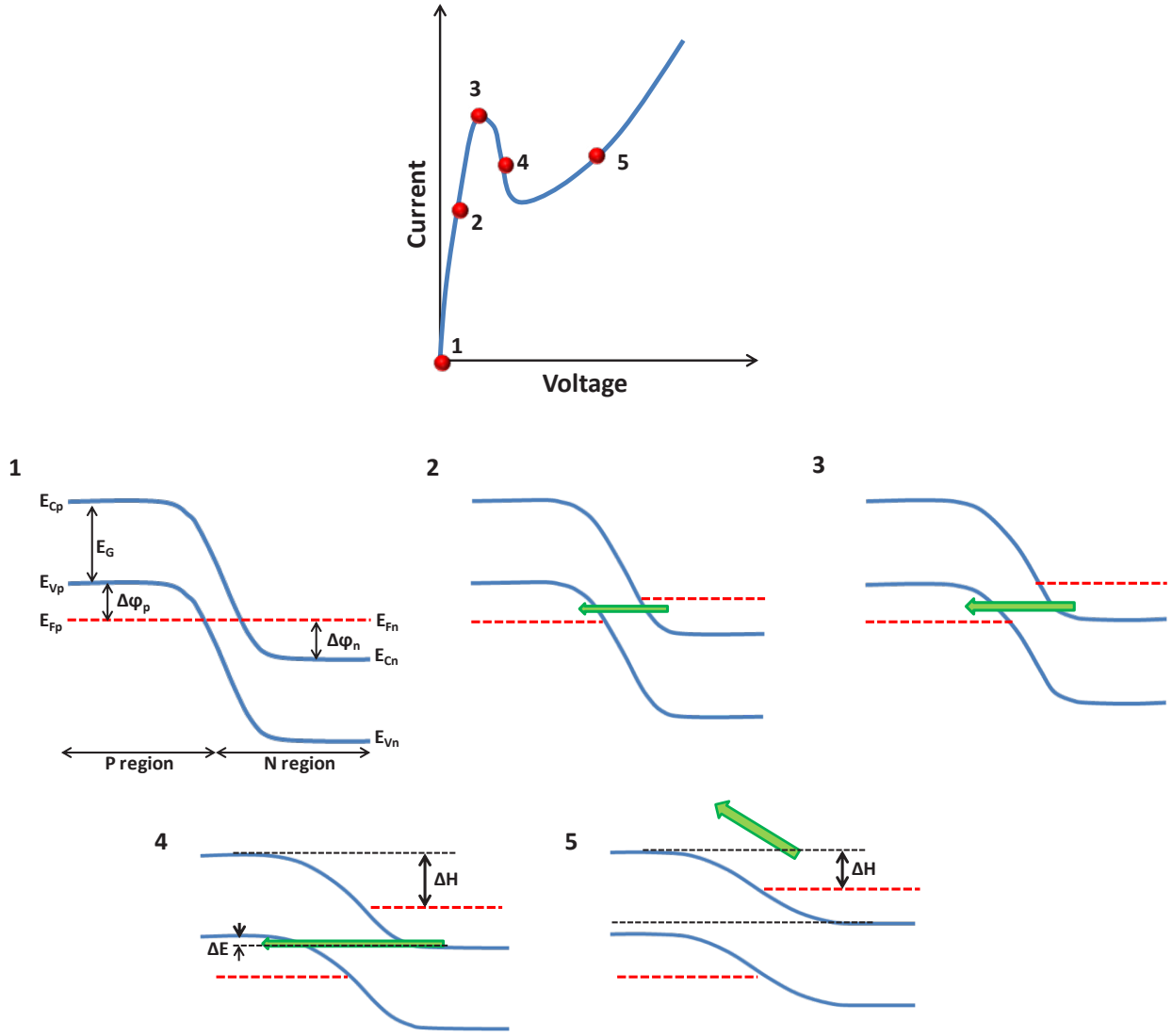


Fig. S8. The forward bias characteristics of a highly doped p-n junction along with band diagrams representing different points in the current-voltage curve.

When the bias is increased and the energy of most of the electrons in n-region gets aligned with that of the free states in p-region, maximum current is achieved (point 3, **Fig. S8**). With further increase in bias, the number of electrons in n-region with energy aligned to the empty states in p-region decreases, resulting in decrease in current and hence, NDR (point 4, **Fig. S8**). As the bias is increased yet further, the barrier height (shown by ΔH) for electrons in the n-region to be thermionically emitted to the p-region decreases, and hence current begins to increase again due to thermionic emission (point 5, **Fig. S8**). As is clear from the current-voltage curve, prominent NDR

effect can be observed here as the tunneling current at point 3 is much higher than the thermionic emission at that bias. However, in case either p or n-region is lightly doped, prominent NDR is not observed (**Fig. S9**). Supposing the n region is lightly doped and the Fermi level is very near or below the conduction band in that region, then there are very few electrons available that can tunnel to p region, leading to lower tunneling current (point 2, **Fig. S9**). Thus, the tunneling current in forward bias is not substantially higher than the thermionic current leading to a smoother transition from tunneling (point 2, **Fig. S9**) to thermionic emission (point 3, **Fig. S9**) showing a trend towards NDR.

In our case, the p-region (Ge) is highly doped while n-region (MoS_2) is lightly doped. Hence, we also observe a trend toward NDR.

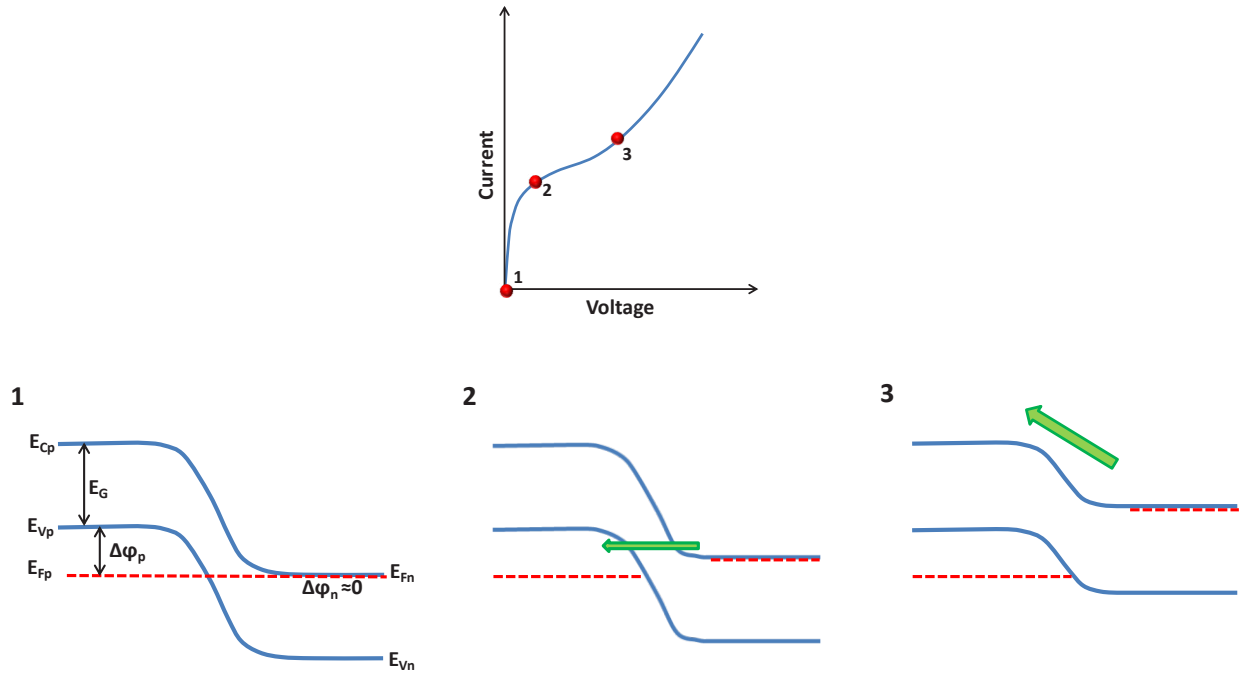


Fig. S9. The forward bias characteristics of a p-n junction with highly doped p region and lightly doped n region, along with band diagrams representing different points in the current-voltage curve.

S8. TEMPERATURE DEPENDENT BEHAVIOR OF BAND-TO-BAND-TUNNELING

In **Fig S10,** the p-n junction current of our device measured in a two terminal configuration (where source pad contacts the p-type Ge and drain pad contacts the naturally n-type MoS_2) at different temperatures is shown. We observe negligible temperature dependence in the reverse bias characteristics. Negligible temperature dependence of

the current is a typical signature that the current is dominated by BTBT rather than thermionic emission, since BTBT is relatively temperature independent. Temperature independence of the reverse bias characteristics also indicates that the effect of trap states is negligible.

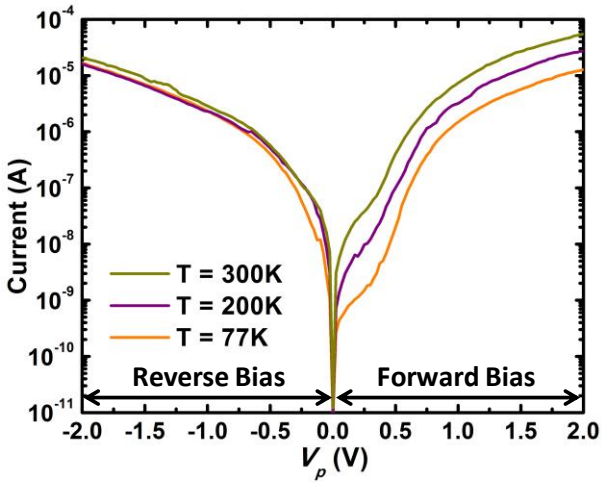
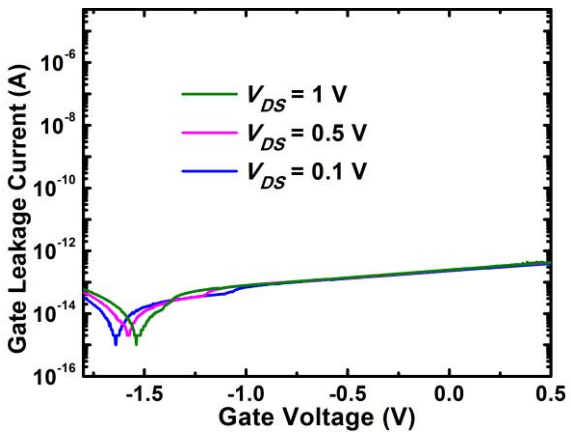


Fig. S10. The forward and reverse bias p-n junction characteristics shown for 3 different temperatures of 77K, 200K and 300K. The reverse bias current is dominated by band-to-band tunneling and shows negligible temperature dependence. The forward bias current involves thermionic emission as well as recombination and hence, is temperature dependent.

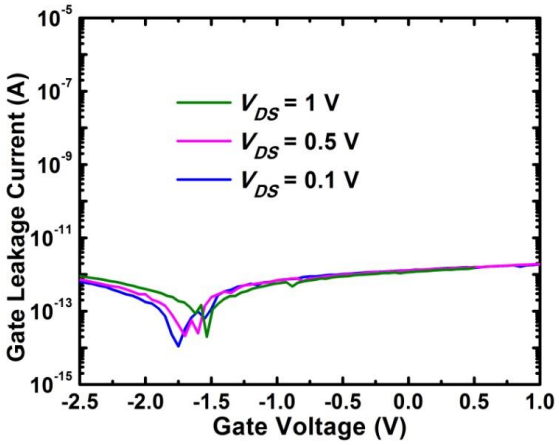
S9. GATE LEAKAGE CHARACTERISTICS AND SOURCE CURRENT OF ATLAS-TFET

The gate leakage (I_G-V_{GS}) characteristics of the ATLAS-TFET are illustrated in **Fig. S11(a)**. As observed from this figure, absolute values of I_G first decreases with increase in gate voltage, reaches a minimum point (around or below -1.5 V) and then increases again. This is because I_G is positive when electrons flow into the gate and is negative when electrons flow out of it and the zero crossing point between the positive and negative values leads to the minimum point. The effect of V_{DS} on zero crossing point can be explained as follows. For a particular V_{DS} say 0.1 V, let us consider any V_{GS} (let us call it V_{GS1}) just to the right of the zero crossing where I_G is positive, meaning electrons are flowing into the gate. Now, if V_{DS} is increased, making the potential at the source-channel junction more positive, electrons will have a tendency to flow out of the gate making the I_G negative at $V_{GS} = V_{GS1}$. This will lead to the shift in the zero crossing point to the right with increasing V_{DS} as observed in **Fig. S11(a)**. This effect is not unique to the polymer gate. We had also fabricated devices having conventional gating with HfO_2 as gate dielectric and similar effect can be observed there as well, as shown in **Fig. S11(b)**.

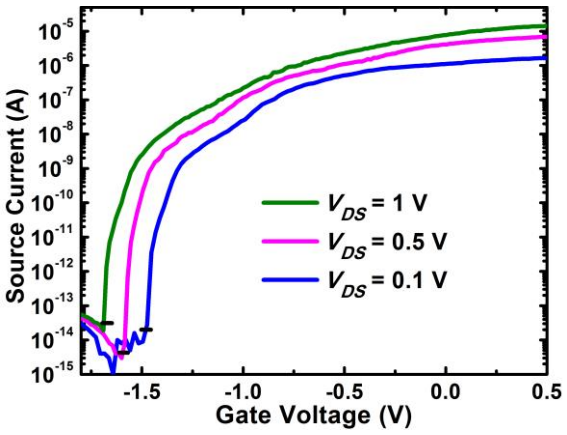
Below the cross points (where source-drain current become similar to the gate leakage current level) current from the gate mainly goes to the source due to the higher overlap between Ge and the polymer complex gate. The plot of source current as a function of gate voltage is shown in **Fig. S11(c)**, which clearly shows that the source current follows the trend of gate leakage current below cross points, thus confirming that gate current flows through the source.



(a)



(b)

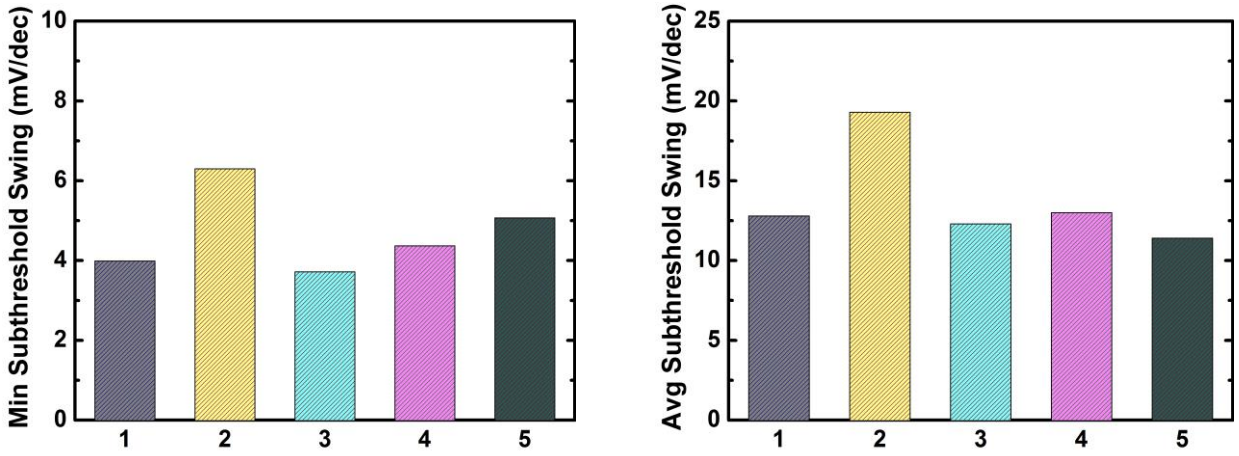


(c)

Fig. S11 (**a**) Gate leakage current as a function of gate voltage for 3 different drain voltages of 0.1V, 0.5V and 1 V for ATLAS-TFET with polymer gating. (**b**) Gate leakage current as a function of gate voltage for 3 different drain voltages of 0.1V, 0.5V and 1 V in the case of conventional gating with HfO_2 as gate dielectric. HfO_2 is deposited using Atomic Layer Deposition and its thickness is around 50 nm. (**c**) Source current as a function of gate voltage for 3 different drain voltages of 0.1V, 0.5V and 1 V for ATLAS-TFET with polymer gating. The cross point, where source current become similar to the gate leakage current level is shown by a short black dash on each curve.

S10. SUMMARY OF DIFFERENT DEVICES MEASURED

The minimum as well as average subthreshold swings (over two decades of current) of 5 devices are shown in **Fig. S12(a)** and **(b)**, respectively. Devices 1-3 were fabricated on the same substrate and were processed and measured together. Devices 4 and 5 were on a different substrate and were processed and measured a month later. This shows that the achievement of sub-60 mV/dec *SS* in ATLAS-TFETs can be repeated.



(a) (b)

Fig. S12. (a) Minimum and **(b)** average subthreshold swing (over two decades of current) of 5 devices.

S11. NEGLIGIBLE HYSTERESIS IN TRANSFER CHARACTERISTICS

The transfer characteristics of the device show negligible hysteresis as evident from **Fig. S13** where it is shown that the I_{DS}-V_{GS} curve measured from negative to positive gate voltages (blue) coincides with that measured from positive to negative voltages (red).

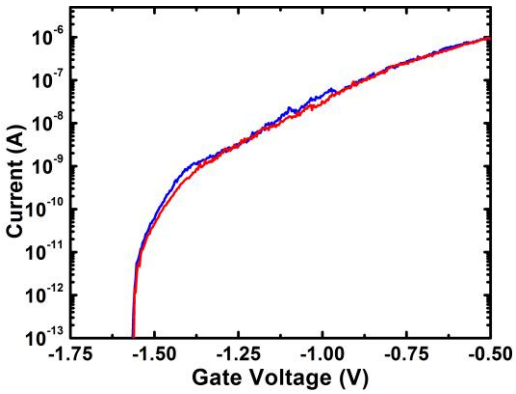


Fig. S13. The I_{DS}-V_{GS} characteristics of ATLAS-TFET showing negligible hysteresis.

S12. OUTPUT CHARACTERISTICS OF ATLAS-TFET

The output (I_{DS}-V_{DS}) characteristics of the device are illustrated in **Fig. S14**. TFETs generally have high saturation voltage for drain current [23,24] and we also do not observe saturation till V_{DS} of 1V in our ATLAS-TFET. This is because of the high value of drain to channel capacitance.[25] This arises from the fact that while there is a tunneling barrier between the source and channel, no such barrier exists between channel drain and hence, electrons from drain can easily transport to the channel compared to that from source, leading to higher drain to channel capacitance compared to source to channel capacitance. Thus, it is difficult to bring the conduction band of channel below that of the drain, as electrons from drain can easily populate the channel and bring the conduction band up. Due to this effect, at very small V_{DS}, it becomes difficult to bring the conduction band of channel below the valence band of source for turning the device ON. This effect occurs in all TFETs, in general. In ATLAS-TFET, this effect is reduced to some extent due to the low tunneling barrier, high junction electric field and higher tunneling area, which increases the source to channel coupling and hence, the source to channel capacitance. This effect can be further reduced by incorporation of proper drain engineering techniques such as drain underlap or heterogeneous dielectrics.[26,27] The curves in Fig. 4b (main article) shift to the left with increasing V_{DS}, due to this higher drain capacitance effect. Drain voltage can also influence the *SS*. At very small drain voltages, the *SS* degrades due to the drain capacitance effect.[23,25] *SS* first improves with increase in drain voltage but, at higher drain biases, degrades again due to drain induced barrier thinning. The increasing trend in *SS* for the ATLAS-TFET with drain current (or gate voltage) (shown in Fig. 4c of main article) is common to all TFETs and the lowest *SS* is obtained when the conduction band of the channel is lowered just below the valence band of the source.

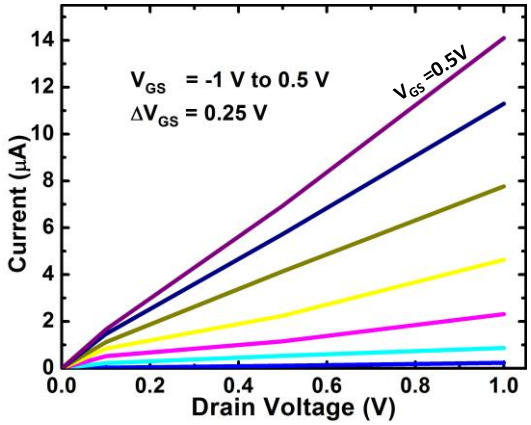


Fig. S14. Drain current as a function of drain voltage for different gate voltages starting from -1V to 0.5 V. The data is obtained from measured I_{DS}-V_{GS} characteristics at different drain voltages.

S13. TRANSFER CHARACTERISTICS OF CONVENTIONAL MoS_2 FET (CFET)

The schematic diagram of a conventional FET is shown in **Fig. S15**. Here, the source and drain metallic contacts are both on MoS_2 and current flows from the source to the drain laterally. The substrate consists of 280 nm thermally grown SiO_2 on Si and is used to host the ultra-thin MoS_2 conventional FET. The measured transfer and output characteristics of the conventional FET (CFET) is shown in **Fig. S16**. Note, that the same polymer gating technique has been used for this CFET, as has been used for ATLAS-TFET. The *SS* obtained for CFET, is above 60 mV/dec for all the drain voltages. This control experiment with CFET, using the same gate dielectric and measured under similar conditions and with same tools as that for the ATLAS-TFET, proves that the sub-60 mV/dec *SS* in case of ATLAS-TFET is not the outcome of polymer gating or instrumentation used in measurements. Moreover, the ideal *SS* obtained in conventional FET with polymer gate indicates that the gating effect indeed occurs through electrostatic control following the electrostatic equations (eqn. (s2)) and not through any chemical reaction in the polymer. Note that, *SS* increases at the lower end of drain current due to gate leakage and at the upper end due to device operation towards the superthreshold region (as shown explicitly in Fig. 4c of the main article).

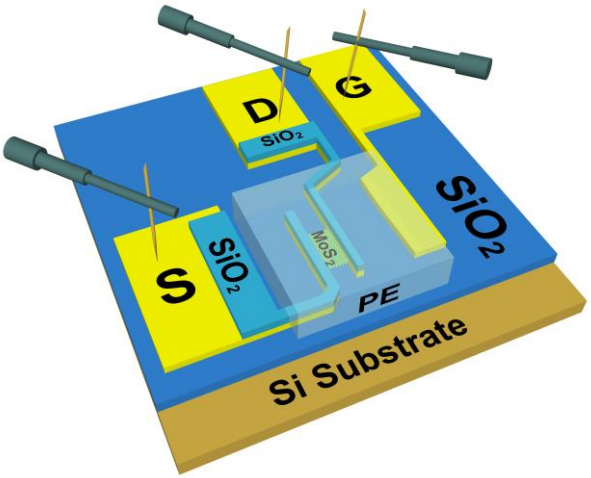
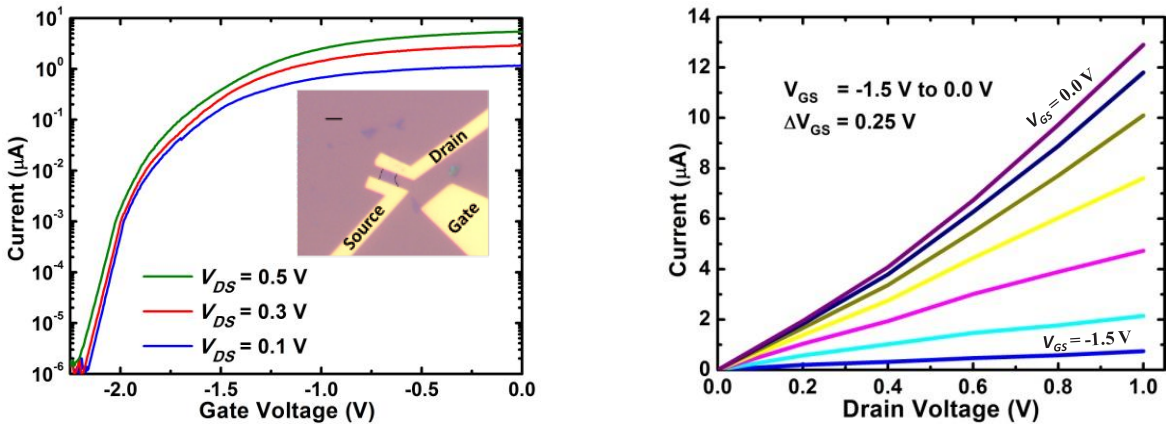


Fig. S15. Schematic diagram of a conventional FET with bilayer MoS_2 as the channel material.



(a) (b)

Fig. S16. (**a**) The I_{DS}-V_{GS} characteristics of the conventional-FET based on bilayer MoS_2 (as shown in Fig. S15) for different V_{DS}. The inset figure shows the optical image of the device. Scale bar: 5 μm. Length and average width of the device are 3.6 μm and 5.8 μm, respectively. (**b**) The I_{DS}-V_{DS} characteristics of the same device measured in (**a**). The data is obtained from measured I_{DS}-V_{GS} characteristics at different drain voltages.

S14. COMPARISON WITH EXPERIMENTAL TFETs WITH SUB-THERMIONIC MINIMUM-SS

Table-I U: unknown (not mentioned in that particular paper), N: Did not achieve Sub-thermal SS_{avg} over 4 decades

References	Channel Material	Structure [Channel Dimension] (nm)	$\|V_{DS}\|$ (V)	SS_{min} (mV/dec)	SS_{avg} over 3 decades (mV/dec)	Sub-thermal SS_{avg} over 4 decades (mV/dec)	I_{on} (μA/μm)	I_{on}/I_{off} Ratio
This work	*MoS_2*	*Planar [1.3]*	*1*	*4.3*	*27*	*34.3*	*1*	*1.6×10^{8}*
			0.5	*3.9*	*22*	*31*	*0.5*	*8.3×10^{7}*
			0.1	*3.8*	*28*	*36.5*	*0.11*	*1.8×10^{7}*
Tomioka et. al. [28] (Hokkaido Univ./JST)	Si/III-V Hetero-J	NW [30]	1	12	21	21	1	2×10^{6}
			0.1	U	21	N	0.001	1×10^{6}
Gandhi et. al. [29,30] (NUS/IME/UCSB)	Si	NW [30-40]	0.1	30	50	N	0.003	1×10^{4}
		NW [18]	0.1	30	50	N	0.44	1.6×10^{5}
Jeon et. al.[31] (SEMATECH/Berkeley/Texas State Univ.)	Si	SOI [40]	1	32	~55	N	1.2	6×10^{6}
Kim et. al.[32] (UC. Berkeley)	Si	SOI [70]	0.5	38	~50	N	0.42	3.5×10^{6}
Mayer et. al.[33] (CEA-LETI)	Si	SOI [20]	0.8	42	>60	N	0.02	4×10^{4}
			0.1	42	>60	N	4×10^{-4}	4×10^{4}
Choi et. al.[34] (UC. Berkeley)	Si	SOI [70]	1	52.8	>60	N	12.1	2×10^{3}
			0.1	~52.8	>60	N	1	1×10^{4}
Appenzeller et. al.[35] (IBM)	CNT	SOI [U]	0.5	40	>60	N	0.01	1×10^{3}
Knoll et. al.[36] (Peter Grunberg Inst.)	Si	NW [7×45]	0.1	30	>60	N	2	2×10^{7}
			0.5	>30	>60	N	10	2×10^{6}
Krishnamohan et. al.[37] (Stanford Univ.)	Ge	DG [10]	0.5	50	>60	N	1	2×10^{6}
Huang et. al.[38] (Peking Univ.)	Si	SOI [100]	0.6	29	70	N	20	1×10^{8}
			0.05	~29	~70	N	1.5	7.5×10^{6}
Villalon et. al.[39] (CEA-LETI)	Ge_xSi_{1-x}	SOI [7-11]	0.1	33	>60	N	1	1×10^{7}
Ganjipour et. al.[40] (Lund Univ.)	InP/GaAs Hetero-J	NW [85-100]	0.75	50	150	N	2.2	1.1×10^{7}
Kim et. al.[41] (U. Tokyo/JST-CREST)	Si	SOI [10-13]	0.05	28	70	N	0.1	1×10^{7}
Dewey et. al.[42] (Intel)	III-V Hetero-J	bulk	0.05	~58	>60	N	0.3	3×10^{4}

The above Table I summarizes the key features of all the experimental TFETs in literature, to date, reporting minimum *SS* of sub 60 mV/dec at room temperature. It is evident from the table that out of all such previously reported TFETs, only those reported by the four groups (as highlighted in the 2nd to 5th rows in bold) achieved an average *SS* below 60 mV/dec (at least over 3 decades). ITRS has prescribed attainment of average sub-thermionic *SS* over at least 4 decades such that current ON-OFF ratio of 4 orders of magnitude can be achieved under application of low bias. The TFETs achieving the same are highlighted in purple in the table. It can be observed that, except, Tomioka et. al., no other previous groups have been able to achieve this. Tomioka et. al., have reported average sub-thermionic *SS* over 4 decades at V_{DS}=1V. The ATLAS-TFET is the first TFET in a planar platform to satisfy this ITRS prescription. Also, it is the only TFET to be able to do so up to a low V_{DS} of 0.1V. Our ON currents are comparable to those of previously reported TFETs with average sub-thermal *SS* (at least over 3 decades). Moreover, an excellent ON-OFF ratio spanning over 7-8 decades of current has been achieved in our ATLAS-TFET.

S15. TUNNELING RESISTANCE OF INTRINSIC GERMANIUM OXIDE

Due to the presence of the intrinsic Germanium Oxide between the Ge and MoS_2 (**Fig. S17**), an extra tunneling resistance is added in series in the tunneling path from the Ge to the MoS_2. Here, we derive the tunneling probability

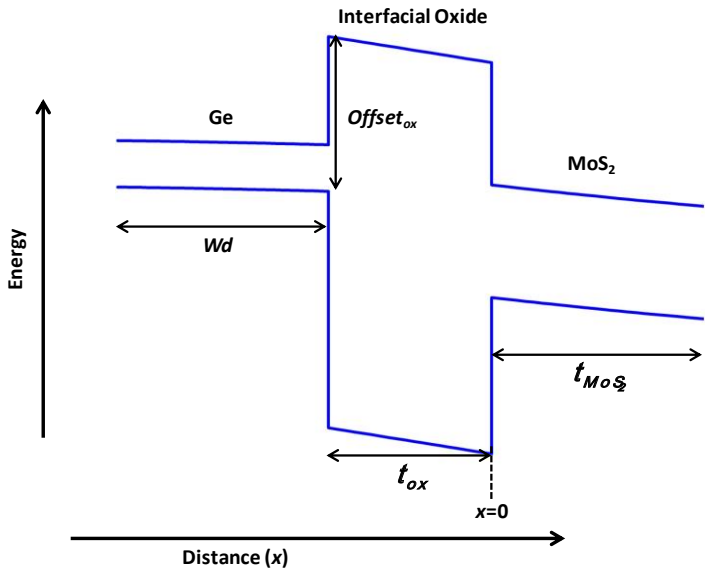


Fig. S17. Band diagram of the tunneling junction between Ge and MoS_2 with the interfacial oxide in-between. While in **Fig. 1c** in the manuscript the bands for the two layers of MoS_2 along with the van der Waal's gap have been shown separately for better visual interpretation, for simulation purpose, it is much more convenient to use the effective band-structure of the bilayer MoS_2 as is shown in this figure.

through this oxide as a function of the voltage applied to the MoS_2 with respect to the Ge, since inverse of this tunneling probability gives indication of the tunneling resistance. For the above derivation, first the band bending in the device needs to be determined, for which, Poisson equations are solved in the device. The Poisson equations in the 3 different regions namely the Ge, interfacial oxide and MoS_2 (TMD) is given below:

$$\frac{d^2\psi_{Ge}(x)}{dx^2} = \frac{q \times dop_{Ge}}{\varepsilon_{Ge}} \qquad \text{(s8a)}$$

$$\frac{d^2\psi_{ox}(x)}{dx^2} = 0 \qquad \text{(s8b)}$$

$$\frac{d^2\psi_{MoS_2}(x)}{dx^2} = \frac{q \times dop_{MoS_2}}{\varepsilon_{MoS_2}} \qquad \text{(s8c)}$$

where ψ_{Ge}, ψ_{ox} and ψ_{MoS2} are the potential profile of valence band of Ge, interfacial oxide and MoS_2 respectively, $\varepsilon_{Ge,}$ and ε_{MoS2} are the dielectric constants of Ge and MoS_2 respectively while, $dop_{Ge,}$ and dop_{MoS2} are the doping concentrations in Ge and MoS_2, respectively. x is in the direction from Ge to MoS_2 and we have defined x=0 at the interface of the oxide (with a thickness of t_{ox}) and MoS_2 (with a thickness of t_{MoS2}), q is the electronic charge.

Eqns (s8a) and (s8c) are first solved with the following four boundary conditions given by:

(i) $\frac{d\psi_{Ge}(-Wd - t_{ox})}{dx} = 0$ as electric field is zero in Ge beyond the depletion region (Wd) in it.

(ii) $-\frac{d\psi_{Ge}(-t_{ox})}{dx} \times \varepsilon_{Ge} = F_{ox} \times \varepsilon_{ox} = -\frac{d\psi_{MoS_2}(0)}{dx} \times \varepsilon_{MoS_2}$ due to the continuity of electric displacement field where $\varepsilon_{Ge,}$ ε_{ox} and ε_{MoS2} are the dielectric constants of Ge, interfacial oxide and MoS_2 respectively and F_{ox} is the electric field in the oxide.

(iii) $\psi_{MoS_2}(0) = \psi_{Ge}(-t_{ox}) + Ev_{offset} - V_{ox}$ where Ev_{offset} is the valence band offset of MoS_2 with respect to Ge and V_{ox} is the voltage drop in the interfacial oxide and can be written as: $-\frac{d\psi_{MoS_2}(0)}{dx} \times \frac{\varepsilon_{MoS_2}}{\varepsilon_{ox}} \times t_{ox}$

(iv) $\psi_{MoS_2}(t_{MoS_2}) = U_{MoS_2}$ where U_{MoS2} is the valence band position of MoS_2 at x=$t_{MoS2.}$

By solving eqns (s8a) and (s8c) with the above boundary conditions we obtain the potential profile in the Ge and the MoS_2 with the depletion region in Ge (Wd) being the unknown factor in them as given below:

$$\psi_{Ge}(x)=\frac{1}{2}\frac{q\,dop_{Ge}\,x^2}{\varepsilon_{Ge}}+\frac{q\,dop_{Ge}\,(Wd+t_{ox})\,x}{\varepsilon_{Ge}}+\frac{1}{2}\frac{1}{\varepsilon_{MoS_2}\varepsilon_{ox}\varepsilon_{Ge}}\{-2\varepsilon_{MoS_2}\varepsilon_{Ge}t_{ox}q\,dop_{Ge}\,Wd$$
$$-2\varepsilon_{ox}\varepsilon_{Ge}t_{MoS_2}q\,dop_{Ge}\,Wd+2\varepsilon_{MoS_2}\varepsilon_{ox}t_{ox}q\,dop_{Ge}\,Wd-2\varepsilon_{MoS_2}\varepsilon_{Ge}\varepsilon_{ox}\,Ev_{offset}$$
$$+\varepsilon_{MoS_2}\varepsilon_{ox}q\,dop_{Ge}\,t_{ox}{}^2-\varepsilon_{Ge}\varepsilon_{ox}q\,dop_{MoS_2}\,t_{MoS_2}{}^2+2\varepsilon_{MoS_2}\varepsilon_{Ge}\varepsilon_{ox}U_{MoS_2}\} \quad \text{(s9a)}$$

$$\psi_{MoS_2}(x)=\frac{1}{2}\frac{q\,dop_{MoS_2}\,x^2}{\varepsilon_{MoS_2}}+\frac{q\,dop_{Ge}\,Wd\,x}{\varepsilon_{MoS_2}}-\frac{1}{2}\frac{q\,dop_{MoS_2}\,t_{MoS_2}{}^2-2U_{MoS_2}\varepsilon_{MoS_2}+2t_{MoS_2}q\,dop_{Ge}\,Wd}{\varepsilon_{MoS_2}} \quad \text{(s9b)}$$

Now, without any loss of generality, $\psi_{Ge}(-Wd-tox)$ is set to zero such that all other potentials are calculated with respect to it. Hence, by solving $\psi_{Ge}(-Wd-tox)=0$ using eqn (s9a), we can derive Wd as:

$$Wd=-\frac{1}{q\varepsilon_{MoS_2}\varepsilon_{ox}dop_{Ge}}\{\varepsilon_{Ge}\varepsilon_{MoS_2}t_{ox}q\,dop_{Ge}+\varepsilon_{Ge}\varepsilon_{ox}t_{MoS_2}q\,dop_{Ge}+[(\varepsilon_{Ge}\varepsilon_{MoS_2}t_{ox}q\,dop_{Ge})^2$$
$$+2\varepsilon_{Ge}{}^2\varepsilon_{MoS_2}t_{ox}q^2\,dop_{Ge}{}^2\varepsilon_{ox}t_{MoS_2}+(\varepsilon_{Ge}\varepsilon_{ox}t_{MoS_2}q\,dop_{Ge})^2$$
$$+2\varepsilon_{MoS_2}{}^2\varepsilon_{ox}{}^2q\,dop_{Ge}\varepsilon_{Ge}\left(U_{MoS_2}-Ev_{offset}\right)-\varepsilon_{MoS_2}\varepsilon_{ox}{}^2q^2dop_{Ge}\varepsilon_{Ge}dop_{MoS_2}t_{MoS_2}{}^2]^{1/2}\} \quad \text{(s10)}$$

The voltage drop in Ge, V_{Ge} can be derived by subtracting the potential of Ge at the interface of Ge and interfacial oxide ($\psi_{Ge}(-t_{ox})$) from the potential at the end of the depletion region ($\psi_{Ge}(-Wd-t_{ox})$ which has been set to zero). Therefore, $V_{Ge}=0-\psi_{Ge}(-t_{ox})$ and from eqn s9a, we can derive,

$$V_{Ge}=\frac{1}{2}\frac{1}{\varepsilon_{MoS_2}\varepsilon_{ox}\varepsilon_{Ge}}(-2\varepsilon_{MoS_2}\varepsilon_{Ge}t_{ox}q\,dop_{Ge}Wd-2\varepsilon_{Ge}\varepsilon_{ox}t_{MoS_2}q\,dop_{Ge}Wd$$
$$-2\varepsilon_{MoS_2}\varepsilon_{Ge}\varepsilon_{ox}Ev_{offset}-\varepsilon_{Ge}\varepsilon_{ox}q\,dop_{MoS_2}t_{MoS_2}{}^2$$
$$+2\varepsilon_{MoS_2}\varepsilon_{Ge}\varepsilon_{ox}U_{MoS_2} \quad \text{(s11)}$$

Now, F_{ox} can be derived from eqn (s9b) using $-\frac{d\psi_{MoS_2}(0)}{dx}\times\frac{\varepsilon_{MoS_2}}{\varepsilon_{ox}}$ as:

$$F_{ox}=\frac{q\,dop_{Ge}\,Wd}{\varepsilon_{ox}} \quad \text{(s12)}$$

With the electric field in the interfacial oxide being derived, we can now derive the tunneling probability (P_{ox}) through it. Using the WKB approximation, the tunneling probability at the energy of the valence band edge of Ge at x=-t_{ox} can be written as:

$$P_{ox} = e^{-2\sqrt{\frac{2mq}{\hbar^2}}\int_0^{t_{ox}}\sqrt{Offset_{ox}-F_{ox}x}\,dx} \quad \text{(s13)}$$

where $Offset_{ox}$ is the height of the oxide conduction band with respect to the valence band of Ge, $\hbar = h/2\pi$ with h being the Planck's constant, m is the mass of the carrier.

Solving the integral in eqn (s13), the tunneling probability can be derived as:

$$P_{ox} = e^{-\frac{4}{3}\sqrt{\frac{2mq}{\hbar^2}}\frac{(Offset_{ox}^{3/2}-(Offset_{ox}-V_{ox})^{3/2})}{F_{ox}}} \quad \text{(s14)}$$

In **Fig. S18**, the tunneling probability through the oxide is plotted as a function of the voltage on MoS_2 with respect to Ge and it is observed that the probability is extremely low and is relatively independent of the applied bias. This implies that even by increasing the bias, it is not possible to increase the tunneling probability (and hence, decrease the resistance posed by the interfacial oxide) significantly.

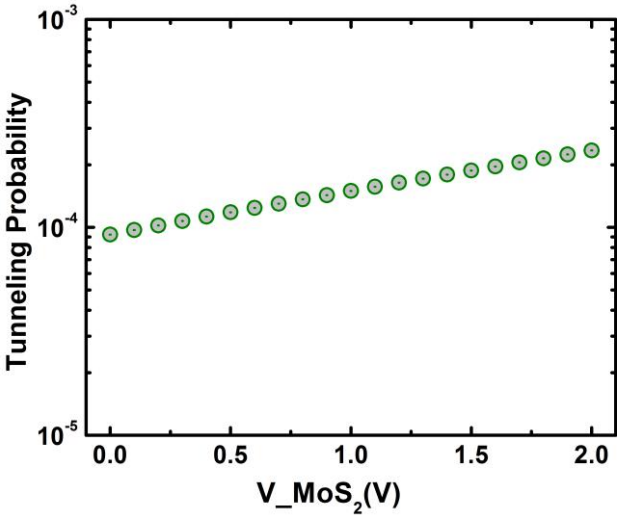


Fig. S18. Tunneling probability through the interfacial oxide as a function of the voltage applied to MoS_2.

S16. ATLAS-TFET: POTENTIAL FOR HIGH PERFORMANCE

As explained in the previous section, the presence of intrinsic germanium oxide leads to additional tunneling resistance. Drain current of the ATLAS-TFET can be improved by removing the intrinsic oxide. Intrinsically, ATLAS-TFET is advantageous for obtaining high ON current. Not only it leads to larger area for tunneling, but it is also easier to scale down the tunneling barrier width (W_t), which basically depends on the thickness of the channel.

Note that in case of a point-TFET where the gate is perpendicular to the source-channel junction (**Fig. S19** (**a**)), it is very difficult to scale down the W_t. This is because in case of point-TFET, W_t is primarily determined by the device electrostatics. Thus, a point-TFET would lead to much higher W_t compared to that of line-TFET, even if bilayer MoS_2 and similar gate capacitance are used. For elucidating the above argument, we derive the analytical expression for band profile in case of point-TFET. Assuming the source to be highly doped and band bending in it to be negligible, the Poisson's equation for conduction band profile in the channel ($\varphi_p(x)$) for point-TFET can be written as[1]

$$\frac{d^2\varphi_p(x)}{dx^2} = \eta(\varphi_p(x) - EG_{ov} + V_{GSeff}) \qquad \text{(s15)}$$

EG_{ov} is the band overlap between source and channel, V_{GSeff} is the effective gate to source voltage and given by V_{GSeff} $=V_{GS}$ - V_{FB} where V_{GS} is the applied gate to source voltage and V_{FB} is the flat-band voltage and η is the inverse of the square of natural length scale, which for the single gated SOI structure can be given by

$$\eta = \frac{\varepsilon_{Di}}{\varepsilon_{Ch} t_{Ch} t_{Di}}$$

where, t_{Di} and ε_{Di} are the thickness and dielectric constant of the gate dielectric respectively, and t_{Ch} and ε_{Ch} represent those of the channel. Setting x=0 at the source-channel junction and E=0 at the valence band of the source, the boundary conditions can be written as:

$$\varphi_p(0) = EG_{ov}$$

$$\varphi_p(L) = Ec_D$$

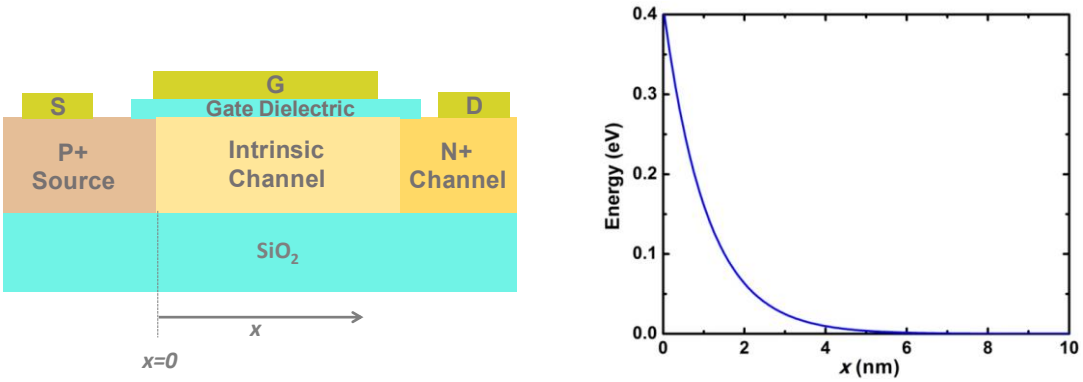
where, L is the channel length and Ec_D is the channel conduction band potential at the drain end.

Solving eqn (s15) using the above boundary conditions, we can derive the channel potential profile as:

$$\varphi_p(x) = e^{(x-L)/\sqrt{\eta}}(Ec_D + V_{GSeff} - EG_{ov}) + e^{-x/\sqrt{\eta}} V_{GSeff} - V_{GSeff} + EG_{ov} \qquad \text{(s16)}$$

Using eqn (s16), the channel conduction band profile for point-TFET near the source-channel junction is plotted in **Fig. S19** (**b**), assuming all material properties remaining same as ATLAS-TFET (such as band overlap at source-

channel junction, solid polymer electrolyte gating and bilayer MoS_2 as channel). It is observed that the W_t where the conduction band of the channel lowers down to the valence band of the source (E=0), is around 5 nm which is much higher than W_t of 1.3 nm in case of line-TFET. Thus, for the same length and width, ATLAS-TFET can lead to higher current due to the lower W_t as well as higher area of tunneling. Note that, while with the scaling down of channel length the benefit due to larger area of tunneling decreases, advantage arising from ultra-small W_t and hence, higher electric field can still lead to high current.



(**a**) (**b**)

Fig. S19 (**a**) Schematic diagram of a point-TFET. x=0 represents the source-channel junction. (**b**) Channel conduction band profile for point-TFET near the source-channel junction as a function of distance (x) along the channel.

It is also noteworthy that, to date, there has not been any successful experimental demonstration of line-TFET with sub-thermionic *SS*. This undermines the importance of atomically thin MoS_2 with dangling-bond free interfaces, which has helped in the first successful realization of a line-TFET with sub-thermionic *SS*.

The vertical band-to-band tunneling (BTBT) is the major factor determining the drain current in ATLAS-TFET as the tunneling resistance is higher than the lateral drift-diffusion resistance due to presence of a tunnel barrier. The region where BTBT takes place is the area of overlap between the MoS_2 and the Ge**.** There are two ways of normalizing the current. One is to normalize with respect to the width, which is the conventional method for MOSFETs and can allow comparison with MOSFETs as well as point TFETs. Another way is to normalize with respect to the overlap area as the current scales with the area. Both methods have been used in the literature. To avoid any ambiguity, we have plotted the current in absolute sense but have mentioned both the width and area of overlap in the caption of Fig. 3 (main Letter), so that comparison can be done as per requirement.

S17. DISCUSSION ON MILLER CAPACITANCE

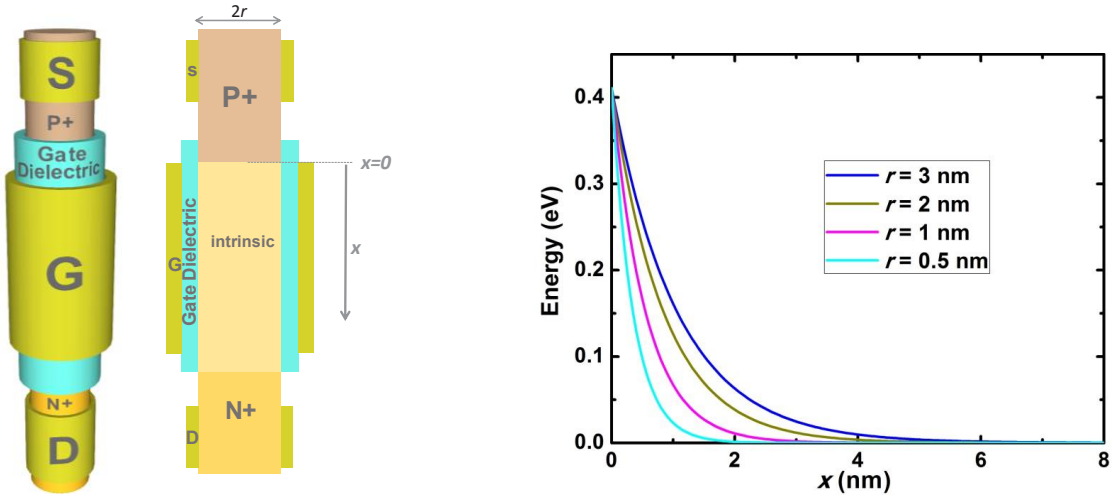
The gate capacitance (C_G) of a transistor consists of 2 parts namely the gate-to-source capacitance (C_{GS}) and gate-to-drain capacitance (C_{GD}). In TFETs, C_{GS} is small due to the presence of tunnel barrier at the source-channel junction and hence, the difficulty in carrier injection by source into the channel. C_{GS} is improved in ATLAS-TFET due to the smaller tunnel barrier width. Also, in TFETs, there exists very small potential drop between the drain and the channel. Thus, carriers from the drain can be easily injected into the channel leading to a high C_{GD} which, results in Miller effect and large voltage overshoot or undershoot in the large-signal circuit switching characteristics. If a gate underlap region is introduced in the ATLAS-TFET, then most of drain voltage will appear in the underlap region. This will reduce the injection of carriers from the drain and thus, reducing the C_{GD} and thereby, the Miller effect. [20]

S18. NOTE ON 1D TFETs AND 2D-2D TUNNEL JUNCTIONS

The ATLAS-TFET will be advantageous compared to 1D TFETs. When compared to 1D point-TFET (**Fig. S20** (**a**)), the current of ATLAS-TFET can be much higher due to 2 reasons. First, the ATLAS-TFET presents larger area for tunneling. Second, it will be difficult to achieve tunneling width as small as in the ATLAS-TFET with a 1D point-TFET. The channel conduction band profile for 1D point-tunnel FET can be given by eqn (s16), except that the equation for η will be different and can be written as:

$$\eta_{1D} = \frac{\varepsilon_{ch}\, r^2 \ln(1 + t_{Di}/r)}{2\varepsilon_{Di}}$$

where r is the radius of the 1D channel. The channel conduction band profile for point-TFET is plotted in **Fig. S20** (**b**) for different radii of 1D structure. It is observed that, all other material properties remaining same (such as band overlap at source-channel junction, solid polymer electrolyte gating as well as channel dielectric constant), a 1D point-TFET will require a radius smaller than 1 nm to match the W_t = 1.3 nm of 2D line-TFET based on bilayer MoS_2. It will be extremely challenging to achieve such small radius with nanowires. Though nanotubes can have small radius, they suffer from chirality issues. Overall, processing of 1D structures is much more difficult compared to planar architectures.



(**a**) (**b**)

Fig. S20 (**a**) Schematic diagram of a 1D point-TFET showing both the 3D view (left) and cross-sectional view (right). x=0 represents the source-channel junction. (**b**) Channel conduction band profile for 1D point-TFET as a function of distance (*x*) along the channel for different radii of the 1D structure.

Moreover, fabrication of a 1D line-TFET, will be even more challenging. A schematic of a 1D line-TFET is shown in **Fig. S21**. Realization of such a structure will require fabrication of concentric cylinders of P+ source and N/intrinsic channel. To make the structure work like an 1D architecture, the radii of the cylinders should be very small. Fabrication of such a structure is extremely daunting.

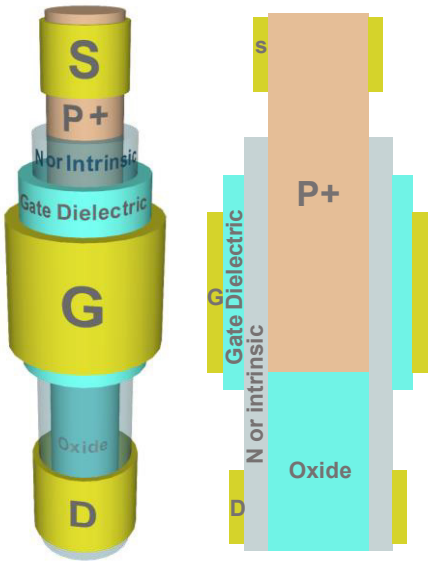
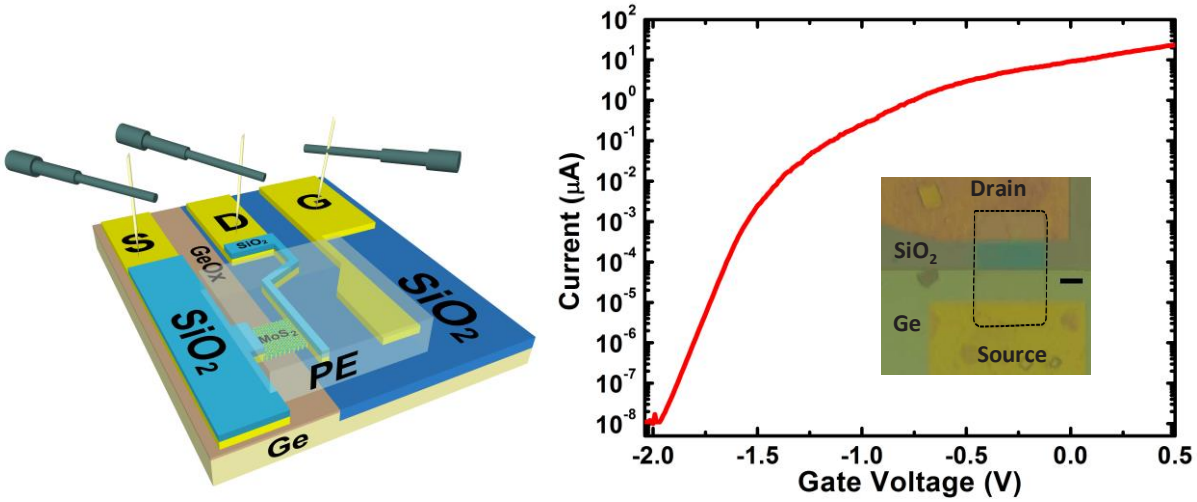


Fig. S21. Schematic diagram of a 1D line-TFET showing both the 3D view (left) and cross-sectional view (right).

It is noteworthy, that 2D-2D tunnel junctions have been theorized to lead to steep *SS*, due to the unique transverse momentum conservation effect[43]. In practical TFETs, it is necessary to form tunnel junctions between two different materials leading to heterostructures. In such heterostructures, it is difficult to obtain a single wave-vector connecting the valence and conduction bands, which makes involvement of phonons necessary for tunneling. Along with phonon scattering, surface scattering will result in loss of momentum conservation, and thereby, loss of the advantage due to momentum conservation in 2D-2D junctions.

S19. CONVENTIONAL MoS_2 FET WITH GERMANIUM SUBSTRATE

To confirm that the sub-thermionic *SS* in case of ATLAS-TFET is not due to the substrate effect, we have fabricated a MoS_2 FET with p-Ge as part of the substrate. The FET device is similar in structure to that shown in Fig. 2 and with source metal contacting the MoS_2. The polymer complex gate overlapped the source region as shown in **Fig. S22** (**a**). In this case, since the source metal is directly contacting the MoS_2, electrons from the source metal can directly transport to the MoS_2 channel and hence, the current is mainly dominated by the thermionic emission. Thus, the *SS* of this device is above the fundamental limitation of 60mV/dec (**Fig. S22** (**b**)).



(**a**) (**b**)

Fig. S22 (a) Schematic diagram of a FET based on bilayer MoS_2 with p-Ge as part of the substrate. The structure is similar to that of ATLAS-TFET and the source metal is connected directly to the MoS_2. **(b)** Drain current as a function of gate voltage at V_{DS}=0.5V. The *SS* obtained is around 78 mV/dec. The inset figure shows the optical image of the device. Scale bar: 5 μm. Length and width of the device are 11.8 μm and 14.7 μm, respectively. MoS_2 is shown by the dotted rectangular region.

REFERENCES

1. Yan, R., Ourmazd, A. & Lee, K. F. Scaling the Si MOSFET : From bulk to SO1 to bulk. *IEEE Trans. Electron Devices* **39,** 704–710 (1992).

2. Colinge, J. Multiple-gate SOI MOSFETs. *Solid State Electron.* **48,** 897–905 (2004).

3. Sarkar, D. *et al.* MoS_2 field-effect transistor for next-generation label- free biosensors. *ACS Nano* **8,** 3992–4003 (2014).

4. Kang, J., Sarkar, D., Khatami, Y. & Banerjee, K. Proposal for all-graphene monolithic logic circuits. *Appl. Phys. Lett.* **103,** 083113 (2013).

5. Cao, W., Kang, J., Sarkar, D., Liu, W. & Banerjee, K. Performance evaluation and design considerations of 2D semiconductor based FETs for sub-10 nm VLSI. In *IEEE International Electron Devices Meeting*, 30.5.1–30.5.4 (IEEE, 2014).

6. Schmidt, M. *et al.* Mobility extraction in SOI MOSFETs with sub 1nm body thickness. *Solid. State. Electron.* **53,** 1246–1251 (2009).

7. Liu, W. *et al.* High-performance few-layer-MoS_2 field-effect-transistor with record low contact-resistance. In *IEEE International Electron Devices Meeting*, 499–502 (IEEE, 2013).

8. Liu, W., Sarkar, D., Kang, J., Cao, W. & Banerjee, K. Impact of contact on the operation and performance of back-gated monolayer MoS_2 field-effect-transistors. *ACS Nano* **8**, 7904–7912 (2015).

9. Sarkar, D., Krall, M. & Banerjee, K. Electron-hole duality during band-to-band tunneling process in graphene-nanoribbon tunnel-field-effect-transistors. *Appl. Phys. Lett.* **97,** 263109 (2010).

10. Solomon, P. M. Inability of single carrier tunneling barriers to give subthermal subthreshold swings in MOSFETs. *IEEE Electron Device Lett.* **31,** 618–620 (2010).

11. Cao, W., Sarkar, D., Khatami, Y., Kang, J. & Banerjee, K. Subthreshold-swing physics of tunnel field-effect transistors. *AIP Adv.* **4,** 067141 (2014).

12. Najmaei, S. *et al.* Vapour phase growth and grain boundary structure of molybdenum disulphide atomic layers. *Nat. Mater.* **12,** 754–759 (2013).

13. Liu, H., Gu, J. & Ye, P. D. MoS_2 nanoribbon transistors: Transition from depletion mode to enhancement mode by channel-width trimming. *IEEE Electron Device Lett.* **33,** 1273–1275 (2012).

14. Kang, J., Liu, W., Sarkar, D., Jena, D. & Banerjee, K. Computational study of metal contacts to monolayer transition-metal dichalcogenide semiconductors. *Phys. Rev. X* **4,** 031005 (2014).

15. Kang, J., Sarkar, D., Liu, W., Jena, D. & Banerjee, K. A computational study of metal-contacts to beyond-graphene 2D semiconductor materials. In *IEEE International Electron Devices Meeting*, 17.4.1–17.4.4 (IEEE, 2012).

16. Kang, J., Liu, W. & Banerjee, K. High-performance MoS_2 transistors with low-resistance molybdenum contacts. *Appl. Phys. Lett.* **104,** 093106 (2014).

17. Lin, M.-W. *et al.* Mobility enhancement and highly efficient gating of monolayer MoS_2 transistors with polymer electrolyte. *J. Phys. D. Appl. Phys.* **45,** 345102 (2012).

18. Carrad, D. J. *et al.* Electron-beam patterning of polymer electrolyte films to make multiple nanoscale gates for nanowire transistors. *Nano Lett.* **14,** 94–100 (2014).

19. Boyd, R. H. The dielectric constant of lamellar semicrystalline polymers. *J. Polym. Sci. Polym. Phys. Ed.* **21,** 505–514 (1983).

20. Mos, W. S. *et al.* Vertical and in-plane heterostructures from WS_2/MoS_2 monolayers. *Nat. Mater.* **13,** 1135–1142 (2014).

21. Alay, J. L. *et al.* Determination of valence band alignment at ultrathin SiO_2/Si interfaces by high-resolution x-ray photoelectron spectroscopy. *Jpn. J. Appl. Phys.* **34,** 653–656 (1995).

22. Khatami, Y., Kang, J. & Banerjee, K. Graphene nanoribbon based negative resistance device for ultra-low voltage digital logic applications. *Appl. Phys. Lett.* **102,** 043114 (2013).

23. Mallik, A. & Chattopadhyay, A. Drain-dependence of tunnel field-effect transistor characteristics: The role of the channel. *IEEE J. Electron Devices Soc.* **58,** 4250–4257 (2011).

24. Boucart, K. & Ionescu, A. M. A new definition of threshold voltage in tunnel FETs. *Solid. State. Electron.* **52,** 1318–1323 (2008).

25. Mookerjea, S., Krishnan, R., Datta, S. & Narayanan, V. Effective capacitance and drive current for tunnel FET (TFET) CV/I estimation. *IEEE Trans. Electron Devices* **56,** 2092–2098 (2009).

26. Verhulst, A. S., Vandenberghe, W. G., Maex, K. & Groeseneken, G. Tunnel field-effect transistor without gate-drain overlap. *Appl. Phys. Lett.* **91,** 053102 (2007).

27. Narang, R., Saxena, M., Gupta, R. S. & Gupta, M. Device and circuit level performance comparison of tunnel FET architectures and impact of heterogeneous gate dielectric. *J. Semicond. Technol. Sci.* **13,** 224–236 (2013).

28. Tomioka, K., Yoshimura, M. & Fukui, T. Steep-slope tunnel field-effect transistors using III-V nanowire / Si heterojunction. In *Symposium on VLSI Technology*, 47–48 (IEEE, 2012).

29. Gandhi, R. *et al.* Vertical Si-nanowire n -type tunneling FETs with low subthreshold swing (≤ 50 mV/decade) at room temperature. *IEEE Electron Device Lett.* **32,** 437–439 (2011).

30. Gandhi, R. *et al.* CMOS-compatible vertical-Silicon-nanowire gate-all-around p-type tunneling FETs with ≤ 50 mV/decade subthreshold swing. *IEEE Electron Device Lett.* **32,** 1504–1506 (2011).

31. Jeon, K. *et al.* Si tunnel transistors with a novel silicided source and 46 mV/dec swing. In *Symposium on VLSI Technology*, 121-122 (IEEE, 2010).

32. Kim, S. H., Kam, H., Hu, C. & Liu, T. K. Germanium-source tunnel field effect transistors with record high I_{ON} / I_{OFF}. In *Symposium on VLSI Technology*, 178–179 (2009).

33. Mayer, F. *et al.* Impact of SOI, $Si_{1-x}Ge_xOI$ and GeOI substrates on CMOS compatible tunnel FET performance. In *2008 IEEE International Electron Devices Meeting*, 1–5 (IEEE, 2008).

34. Choi, W. Y., Park, B., Lee, J. D. & Liu, T. K. Tunneling field-effect transistors (TFETs) with subthreshold swing (SS) less than 60 mV/dec. *IEEE Electron Device Lett.* **28,** 743–745 (2007).

35. Appenzeller, J., Lin, Y.-M., Knoch, J. & Avouris, P. Band-to-band tunneling in carbon nanotube field-effect transistors. *Phys. Rev. Lett.* **93,** 196805 (2004).

36. Knoll, L. *et al.* Inverters with strained Si nanowire complementary tunnel field-effect transistors. *IEEE Electron Device Lett.* **34,** 813–815 (2013).

37. Krishnamohan, T., Kim, D., Raghunathan, S. & Saraswat, K. Double-gate strained-Ge heterostructure tunneling FET (TFET) with record high drive currents and <60mV/dec subthreshold slope. In *IEEE International Electron Devices Meeting*, 1-3 (IEEE, 2008).

38. Huang, Q. *et al.* Comprehensive performance reassessment of TFETs with a novel design by gate and source engineering from device/circuit perspective. In *IEEE International Electron Devices Meeting*, 335–338 (IEEE, 2014).

39. Villalon, A. *et al.* Strained Tunnel FETs with record I_{ON}: First demonstration of ETSOI TFETs with SiGe channel and RSD. In *Symposium on VLSI Technology,* 49–50 (IEEE, 2012).

40. Ganjipour, B., Wallentin, J., Borgstro, M. T., Samuelson, L. & Thelander, C. Tunnel field-effect transistors based on InP-GaAs heterostructure nanowires. *ACS Nano* **6,** 3109–3113 (2012).

41. Kim, M. *et al.* High I_{on} / I_{off} Ge-source ultrathin body strained-SOI tunnel FETs - Impact of channel strain , MOS interfaces and back gate on the electrical properties. In *IEEE International Electron Devices Meeting*, 331–334 (IEEE, 2014).

42. Dewey, G. *et al.* Fabrication, characterization, and physics of III–V heterojunction tunneling field effect transistors (H-TFET) for steep sub-threshold swing. In *IEEE International Electron Devices Meeting*, 33.6.1–33.6.4 (IEEE, 2011).

43. Agarwal, S. & Yablonovitch, E. Pronounced effect of pn-junction dimensionality on tunnel switch sharpness. *arXiv:1109.0096* (2012).