



OPEN Anatomy of nanomagnetic switching at a 3D topological insulator PN junction

Yunkun Xie¹, Hamed Vakili^{2✉}, Samiran Ganguly^{1,3} & Avik W. Ghosh^{1,2}

A P-N junction engineered within a Dirac cone system acts as a gate tunable angular filter based on Klein tunneling. For a 3D topological insulator with a substantial bandgap, such a filter can produce a charge-to-spin conversion due to the dual effects of spin-momentum locking and momentum filtering. We analyze how spins filtered at an in-plane topological insulator PN junction (TIPNJ) interact with a nanomagnet, and argue that the intrinsic charge-to-spin conversion does not translate to an external gain if the nanomagnet also acts as the source contact. Regardless of the nanomagnet's position, the spin torque generated on the TIPNJ is limited by its surface current density, which in turn is limited by the bulk bandgap. Using quantum kinetic models, we calculated the spatially varying spin potential and quantified the localization of the current versus the applied bias. Additionally, with the magnetodynamic simulation of a soft magnet, we show that the PN junction can offer a critical gate tunability in the switching probability of the nanomagnet, with potential applications in probabilistic neuromorphic computing.

Three-dimensional topological insulator surfaces enjoy gapless topological surface states (TSS) protected by spatial inversion and time-reversal symmetry. In many ways, they are similar to low-energy graphene bands, except the Dirac cones are labeled by spins rather than pseudospins, with a similar Chern number^{1–3}. Numerous novel physical phenomena have been explored with TSS across graphene PN junctions. In particular, we can draw inspiration from graphene PN junctions that have served as a convenient laboratory for Klein tunnel devices, with potentially important contributions to low-power digital as well as high-speed analog switches^{4–10}.

The intrinsic spin-momentum locking of TSS in a TI material allows conversion of a charge current into a spin current at a higher efficiency than heavy metal underlayers in common spin-orbit torque devices^{11–13}. In fact, the spin-momentum locking feature of TSS imposes spin conservation rules across a PN junction, suggesting a charge-to-spin conversion ratio (scaled dimensionally like the Spin Hall Angle except between longitudinal current densities) can exceed unity and be as high as 20¹⁴. This number is much higher than the Spin Hall Angle measured in heavy metals such as Pt (0.07)¹⁵, β -Ta (0.12–0.15)¹⁶, and in Pt-doped Au (0.12)¹⁷. However, besides an idealized assumption about the presence of a large non-leaky bandgap (perhaps more suitable to TIs like SmB₆ than Bi₂Se₃), it is also worth emphasizing that the spin ‘amplification’ effect was studied on a homogeneous TIPNJ, where the source/drain contacts are assumed to be extensions of the TI surface as well. While it is easy to over-generalize the high charge-to-spin conversion efficiency to the TI-ferromagnet interface, it is not a straightforward equivalence to the homogeneous TI surface and deserves proper analysis.

In this paper, we argue that the aforementioned large spin amplification is an intrinsic effect that does not automatically induce a large spin current across the TI/nanomagnetic contact interface. Figure 1 shows some possible setups for a nanomagnet on TIPNJ. Our focus is on configuration (a) and we will use Non-Equilibrium Green's Function (NEGF) simulations on the TSS to help clarify the difference between intrinsic versus extrinsic charge-to-spin gain in TIPNJ. We also show a NEGF treatment of a full 3D slab that allows us to compare the surface versus bulk shunt resistances. In practice, it can be a challenge to isolate the surface contribution from the bulk in common binary TI compounds such as Bi₂Se₃, Bi₂Te₃^{18,19}. One possible solution is to alloy the binary TIs into ternary compounds like Bi₂Te₂Se tuned to have a low carrier density in the bulk²⁰. Another way to improve the TSS portion of the total current is by inserting a ferromagnetic layer with a perpendicular anisotropy to one surface (Fig. 5b,c), which we will cover later.

¹School of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA, 22903, USA. ²Department of Physics, University of Virginia, Charlottesville, VA 22903, USA. ³Department of Electrical and Computer Engineering, Virginia Commonwealth University, Richmond, VA 23284, USA. ✉email: hv8rf@virginia.edu

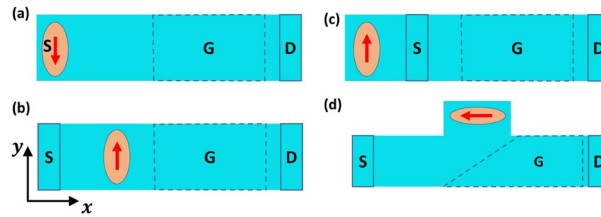


Figure 1. Top view of different configurations of nanomagnet on a TIPNJ. (a) Nanomagnet as the source contact. (b) Floating nanomagnet located between source and drain contacts. (c) Floating nanomagnet outside the source-drain path. (d) A variation of the out-of-location floating nanomagnet. The gate (PN junction interface) is at 45° degree from the source-drain direction. The angled PN junction can reflect electrons from the source to the nanomagnet.

A PN junction provides a gate tunable filter on the spins. For sufficient current densities it can flip a low barrier in-plane magnet unidirectionally. The latter can be flipped back with an oppositely directed current in a symmetric set-up. The low barrier magnet is useful for three-terminal stochastic computing^{21,22}, where the PN junction provides a gate tunability of the average magnetization that follows a neuron-like nonlinear activation function.

Computational method

Near the Dirac point, the 3DTI Hamiltonian can be described as³:

$$\epsilon_0(k) + \begin{pmatrix} \mathcal{M}(k) & A_1 k_z & 0 & A_2 k_- \\ A_1 k_z & -\mathcal{M}(k) & A_2 k_- & 0 \\ 0 & A_2 k_+ & \mathcal{M}(k) & -A_1 k_z \\ A_2 k_+ & 0 & -A_1 k_z & -\mathcal{M}(k) \end{pmatrix} \quad (1)$$

with $k_{\pm} = k_x \pm ik_y$, $\epsilon_0(k) = C + D_1 k_z^2 + D_2 (k_x^2 + k_y^2)$ and $\mathcal{M}(k) = M - B_1 k_z^2 - B_2 (k_x^2 + k_y^2)$. For Bi₂Se₃, the parameters used are: $M = -0.28$ eV, $A_1 = 2.3$ eVÅ, $A_2 = 3.3$ eVÅ, $B_1 = 6.8$ eVÅ², $B_2 = 44.5$ eVÅ², $C = -0.0068$ eV, $D_1 = 5.74$ eVÅ², $D_2 = 30$ eVÅ². These parameters yield a bulk bandgap of ≈ 300 meV. We used the 3D Hamiltonian to study the surface versus bulk current. By discretizing the 3DTI Hamiltonian we get the lattice Hamiltonian:

$$H_{3DTI} = \sum_i c_i^\dagger \epsilon_{3DTI} c_i + \sum_i \left(c_{i,i}^\dagger \frac{T_x}{2} c_{i+1,i} + c_{i,i}^\dagger \frac{T_y}{2} c_{i,i+1} + c_{i,i}^\dagger \frac{T_z}{2} c_{i,i+1} + \text{H.C.} \right) \quad (2)$$

Where $\epsilon_{3DTI} = (C - 2D_1/a^2 - 4D_2/a^2)I_{4 \times 4} + (M - 2B_1/a^2 - 4B_2/a^2)I_{2 \times 2} \otimes \tau_z$, $T_{x,y} = B_2/a^2 I_{2 \times 2} \otimes \tau_z + D_2/a^2 I_{4 \times 4} - iA_2/2a \sigma_{x,y} \otimes \tau_x$, $T_z = B_1/a^2 I_{2 \times 2} \otimes \tau_z + D_1/a^2 I_{4 \times 4} - iA_1/2a \sigma_z \otimes \tau_x$, and H.C. is the Hermitian conjugate. The z-direction is set perpendicular to the quantum layers. While the 3DTI Hamiltonian was used to study the bulk-surface current distributions and current shunting, we need to adopt a simpler TSS model for simulating the charge/spin transport in the combined system of a TIPNJ and a nanomagnet (Fig. 2): $H_{TSS} = v_F \hat{z} \cdot (\boldsymbol{\sigma} \times \mathbf{p})$. To discretize this Hamiltonian, a Wilson mass term needs to be added to avoid the fermion doubling problem²³:

$$H_{TSS} = v_F \hat{z} \cdot (\boldsymbol{\sigma} \times \mathbf{p}) + \gamma \hbar v_F \sigma_z (k_x^2 + k_y^2) \\ = \sum_i c_i^\dagger \epsilon_{TI} c_i + \sum_i \left(c_{i,i}^\dagger t_x c_{i,i+1} + \text{H.C.} \right) \\ + \sum_j \left(c_{j,j}^\dagger t_y c_{j,j+1} + \text{H.C.} \right) \quad (3)$$

where $\epsilon_{TI} = -4\hbar v_F \alpha \sigma^z / a$, $t_x = \hbar v_F [i\sigma^y / 2 + \alpha \sigma^z] / a$, $t_y = \hbar v_F [-i\sigma^x / 2 + \alpha \sigma^z] / a$. $a = 5$ Å is the grid spacing, $v_F = 0.5 \times 10^6$ m/s, and $\alpha = \gamma / a$ is a fitting parameter set to $\alpha = 1$ to generate a bandstructure that reproduces the ideal linear bands within a energy window of 0.5 eV¹⁴. $\hat{z}$ is the normal vector to the surface and v_F is the speed of electrons near the Dirac point. $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices.

A generic ferromagnetic (FM) nanomagnet is modeled by a tight-binding Hamiltonian in a cubic lattice with a single orbital per site:

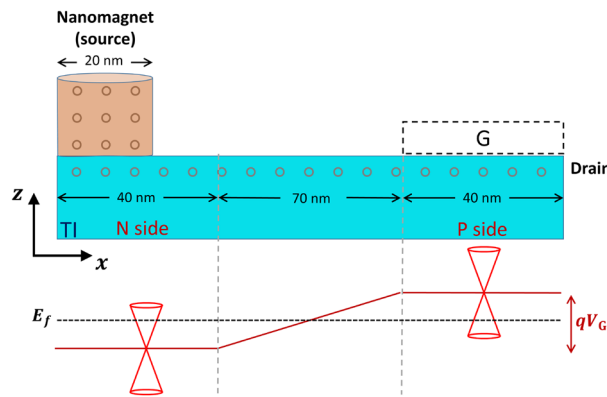


Figure 2. Schematics of the discretization of a TIPNJ with a nanomagnet as the source contact. The ferromagnetic contact has dimension 20 nm along the x-direction. The TI surface is 150 nm long with 40 nm on the P and N side and 70 nm transitioning from N to P side. The gate on the drain side can swing the local TI surface to P-type. The red line represents the electrostatic potential of the TIPNJ.

$$H_{\text{FM}} = \sum_{i,\sigma,\sigma'} c_{i\sigma}^\dagger \left(\varepsilon_{\text{FM}} \delta_{\sigma\sigma'} - \frac{\Delta}{2} \cdot \sigma \right) c_{i\sigma'} - t \sum_{i,i',\sigma} \left(c_{i\sigma}^\dagger c_{i'\sigma} + \text{H.C.} \right) \quad (4)$$

where ε_{FM} is the onsite energy. $\Delta = \Delta \mathbf{m}$ is the exchange energy split along the direction of the magnetization $\mathbf{m}$. It is important that $\mathbf{m}$ be in-plane as an out-plane magnetic orientation can open up a gap in the TI and lower the efficiency for the spin current generation. i' goes through all neighbors of site i . $t = \hbar^2/2m^*a^2$ is the electron hopping energy with effective electron mass m^* . The hopping term between the FM and TI surface is tuned to $t_{\text{FM-TI}} = 2.3t$. This value of hopping term is chosen to minimize the contact resistance. The FM parameters $m^* = 0.5m_e$, $\varepsilon_{\text{FM}} = 1.3 \text{ eV}$, $\Delta = 0.8 \text{ eV}$ result in a spin polarization $\eta = (D_\uparrow - D_\downarrow)/(D_\uparrow + D_\downarrow) \approx 0.57$ around the Fermi energy with density of states $D_\uparrow = 1.34 \times 10^{46} \text{ J}^{-1} \text{ m}^{-3}$, $D_\downarrow = 0.357 \times 10^{46} \text{ J}^{-1} \text{ m}^{-3}$ for the spin up ($-y$)/down ($+y$) channels. These numbers are a bit lower than the nanomagnet modeled in²⁴, mostly due to the reduced size of our simulated magnet. Figure 2 shows the discretization of the coupled system including the ferromagnetic contact and the TI surface. The TI surface is assumed to be doped N-type with a single gate controlling the drain side. The periodic boundary condition is adopted in the y direction and characterized by the transverse quasi-momentum $\mathbf{k}_\perp$. Part of the FM contact is included in the channel Hamiltonian while the rest is assumed to be infinite in the $+z$ direction and modeled by a self-energy Σ_S using recursive Green's functions. For simplicity, the drain is just an extended part of the TI surface and included as a surface self-energy term Σ_D that can be obtained with an iterative approach²⁵. The retarded Green's function G^r and electron correlation matrix G^n can be calculated through the Non-equilibrium Green's function method²⁶

$$G^r(E, \mathbf{k}_\perp) = [EI - H(\mathbf{k}_\perp) - \Sigma_S(E, \mathbf{k}_\perp) - \Sigma_D(E, \mathbf{k}_\perp)]^{-1} \quad (5)$$

$$G^n(E, \mathbf{k}_\perp) = G^r(f_S \Gamma_S + f_D \Gamma_D) G^{r\dagger}$$

$$\Gamma_{S,D} = i \left(\Sigma_{S,D} - \Sigma_{S,D}^\dagger \right) \quad (6)$$

with $f_{S,D}$ the Fermi–Dirac distribution function on the source/drain sides, and I the identity matrix. The charge current is conserved throughout the structure and can be evaluated as:

$$I_q(E) = \frac{q}{h} \sum_{\mathbf{k}_\perp} \text{Tr} [\Gamma_S G^r \Gamma_D G^{r\dagger}] (f_S - f_D) \quad (7)$$

while the charge and spin current from site i to j are calculated as:

$$J_s^{i \rightarrow j}(E) = \frac{q}{ih} \sum_{\mathbf{k}_\perp} \text{Tr} \left[\sigma \left(H_{ij} G_{ji}^n - G_{ji}^n H_{ij} \right) \right] \quad (8)$$

$$J_q^{i \rightarrow j}(E) = \frac{q}{ih} \sum_{\mathbf{k}_\perp} \text{Tr} \left[\left(H_{ij} G_{ji}^n - G_{ji}^n H_{ij} \right) \right]$$

The bias induced carrier density $n = n_{\text{neq}} - n_{\text{eq}}$ where $n_{\text{neq}}, n_{\text{eq}}$ are obtained from:

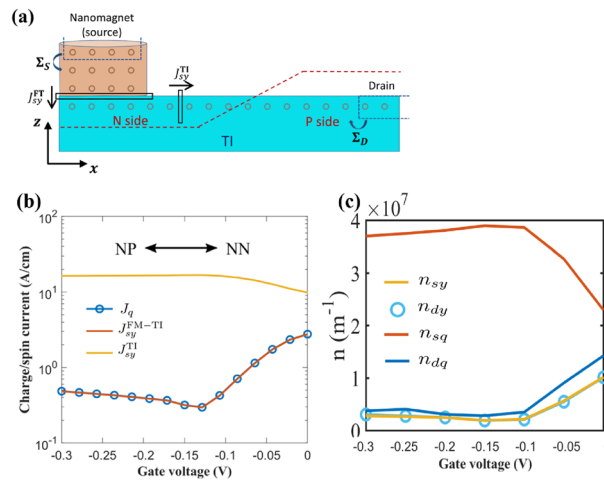


Figure 3. I–V characteristic of TIPNJ with a ferromagnetic source contact. **(a)** TIPNJ setup with a nanomagnet as the source. The magnetization of the FM contact is oriented to the $-y$ direction. The gate contact is present but not visualized in the schematic. **(b)** Charge and spin current densities calculated at different locations. J_q is the charge current density (conserved throughout the system). J_{sy}^{TI} is the spin current (polarized along $-y$ direction) calculated on the TI surface on the source side (N side). J_{sy}^{FM-TI} is the spin current density at the FM-TI interface. The source-drain bias V_{sd} is fixed at 0.1 V and the magnetization of the nanomagnet is aligned in the $-y$ direction. The scattering of charge current at the FM-TI interface suppress the charge current density J_q compared to the spin current J_{sy}^{TI} which explains why at the NN limit, the ratio of spin current and charge current is larger than $\pi/2$ shown in our previous work¹⁴. **(c)** Biased density spin y density is the same in the source (n_{sy}) and drain side (n_{dy}) of TI. This indicates that the spin current amplification in the source region of TI can not be used for magnet switching. n_{sq} and n_{dq} are the charge carrier density in the source and drain, respectively.

$$n_{eq} = -\frac{1}{\pi} \text{Tr} \left(\sum_{\mathbf{k}_{\perp}} \int dE \text{Im}(G^r(E, \mathbf{k}_{\perp})) \right) f_0(E) \quad (9)$$

$$n_{neq} = \frac{1}{2\pi i} \text{Tr} \left(\sum_{\mathbf{k}_{\perp}} \int dE G^n(E, \mathbf{k}_{\perp}) \right)$$

where f_0 is the equilibrium Fermi–Dirac distribution.

Results: charge-to-spin conversion in TIPNJ

Our previous work¹⁴ showed that the TIPNJ acts like a spin collimator that increases the non-equilibrium spin current while reducing the charge current at the same time. The corresponding spin-to-charge ratio (the longitudinal equivalent of the Spin Hall Angle θ_{SH}) at the source contact can go up to as high as 20¹⁴. While it is easy to assume that this conversion would dramatically improve the switching efficiency if we replace the ‘TI source contact’ with a ferromagnetic contact, we show here that the impressive gain is limited to the TI surface (we refer to this as the ‘intrinsic’ gain). To see the difference between the intrinsic and the external gain, we calculated the spin current at two locations along the transport direction: one between the FM contact and the TI surface, the other one on the N-type TI surface, as indicated in Fig. 3a. Figure 3b compares the charge current with the spin currents at the above-mentioned locations as a function of the gate voltage. As the gate voltage sweeps from -0.3 V to 0 V, the TI surface transitions from a PN junction to a homogeneous N-type surface. The behavior of the charge current (independent of where it is being calculated) and TI surface spin current J_{sy}^{TI} , with a maximum ratio of $J_{sy}^{TI}/J_q > 20$ resembles the results from¹⁴, where the PN junction effectively suppresses the charge current while amplifying the nonequilibrium spin current. However, the spin current across the nanomagnetic contact J_{sy}^{FM-TI} does not follow the in-plane surface spin current but instead follows the charge current. While the ratio J_{sy}^{FM-TI}/J_q is very close to 100%, J_{sy}^{FM-TI} never exceeds the charge current regardless of which regime the TI surface is in. Looking at the bias-induced charge and spin carrier density, we can see the picture more clearly (Fig. 3c). The spin density shows no change on the source and drain sides of TI. However, the charge densities show significant variations in the source and drain regions (opposite of currents’ behavior). This further solidifies the above explanation, as the charge densities from the incoming and reflected currents add up in the source region but are subtracted for spin density.

To better understand this discrepancy, we look at the individual Fermi levels f_{μ} in each spin channel (where spin-up corresponds to $-y$ while spin-down corresponds to $+y$) on the TI surface (Fig. 4). The Fermi levels

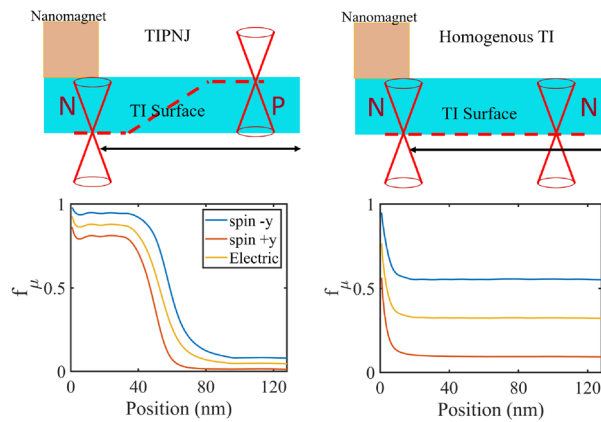


Figure 4. Fermi levels f_μ for different spin channels on the TI surface in different regimes (TIPNJ versus homogeneous TI surface). μ_s (μ_d) is the Fermi level of the source (drain) contact. As can be seen from the comparison, PN junction raises the Fermi level of all spin channels on the source side due to the reflection of the junction potential barrier. The biggest potential drop happens near the junction interface. In the case of homogeneous TI, the Fermi level remains uniform throughout the TI surface from the source to the drain.

are calculated by simulating a magnetic probe at each lattice point which draws net zero currents from the channel with negligible deformation of the channel Hamiltonian. The calculated quantity is $f_\mu = \frac{\text{Tr}(\Gamma_{\text{FM}} G^{\mu})}{\text{Tr}(\Gamma_{\text{FM}} A)}$ where $\Gamma_{\text{FM}}(\pm y) = \gamma_m(I \pm \sigma_y)$ is the coupling between FM probe and the TI surface with $\gamma_m \ll \hbar v_f/a$ being the coupling strength. Due to the spin-momentum locking of the TI surface states, electrons in the spin-up channel can only move in the $+x$ direction while electrons in the spin-down channel move in the opposite direction, roughly along two one-way streets (for 3D TSS, the angular transition is continuous)²⁷. This behavior results in the ‘intrinsic’ spin amplification observed on the TI surface: non-equilibrium electrons moving in opposite directions due to electron reflection at the PN junction reduce the net charge current while increasing the net nonequilibrium spin current. However, when the nanomagnet contact exchanges electrons with the TI surface, the electrons flow bidirectionally regardless of their spin orientations. Ultimately, the rate of inflow and outflow of spins from the nanomagnet is controlled by the difference in the electrochemical potentials between the magnet and the TI surface. Being the source contact, the magnet has the highest electrochemical potential in both spin channels. Therefore one expects a net outflow of electrons from the magnet to the TI surface in both spin channels. The net spin current is just the difference between the charge currents in different spin channels, which is always lower than the sum of them. The extrinsic spin amplification in this configuration, with the magnet doubling as the current source, cannot exceed unity.

The limit on external charge-to-spin gain is imposed by having the nanomagnet as the source contact. A more common setup is to have a floating nanomagnet on top of a TI (see Fig. 1b). There are several research papers that discuss this configuration (without a PN junction) in detail, such as the proximity effect induced by the magnetic field from the magnet²⁸, and the shunting of the charge current from the TI surface to a conductive magnetic layer²⁴. With this configuration, the electrochemical potential of the nanomagnet is free to adjust in order to draw a net zero charge current.

The charge-to-spin conversion rate depends on the splitting of the electrochemical potentials of the opposite spin channels. With TIPNJ, Fig. 4 shows that the splitting narrows compared to the homogeneous TI surface, which means the PN junction effectively turns off the spin torque from the TI surface.

Because of the current shunting to the bulk of the 3DTI, the surface spin current calculated from TSS Hamiltonian Eq. (3) is overestimated. To study the surface-bulk current distribution, the 3DTI Hamiltonian as described in Eq. (2) is employed. With Eq. (8), the current density for each layer is calculated and separated into bulk and surface currents. To get a simple resistance model, we use two parallel resistances for bulk R_b and surface R_s . Based on the current distributions, the ratio of resistances is calculated. Here we define the surface current as the current in the top (bottom) layers: starting from the top (bottom) surface layer to the layer where the current density is peaked as shown in Fig. 5c. TSS has a practical limitation on how much current increases with applied bias. For electrons with higher energies that are only accessible with a higher bias, the conductive state will no longer be strongly localized at the surface and will shunt into the bulk (Fig. 5a). Such a shunting²⁹ implies that a significant portion of the current will be diverted away from the surface, reducing the carrier exchange between the FM and TI and degrading the switching efficiency. Increasing the bias above the bulk bandgap will therefore have diminishing returns. For example, for Bi_2Se_3 the localized current at the surface (about 2nm thick, Fig. 5b) carries only 30% of the total current. One way to alleviate this is by finding a 3DTI material with a larger bulk bandgap, as can be seen in Fig 5a,b. Alternatively, inserting or doping the TI surface in contact with the substrate (Fig. 5b,d) the 3DTI grown on with a PMA magnet can also increase θ_{SH} . This is due to the fact that out-of-plane Zeeman energy^{8,30–32} (Added as $m_z \sigma_z \otimes I_{2 \times 2}$ to TI onsite energy for the bottom surface) opens up a gap. Since this effect is localized only at one surface, the opposite surface is nearly unaffected. However, the total current is now smaller as the total resistance of the channel is increased.

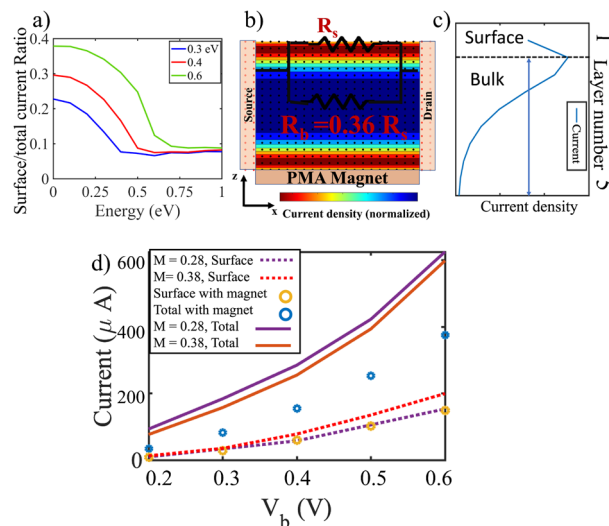


Figure 5. (a) The ratio of TSS-to-Bulk current contributions as a function of the bulk band gap for various carrier energies. We simulate different bulk bandgaps by varying M 0.28, 0.38 and 0.55 eV. For electron energies inside the bulk bandgap, the TSS dominate. We see that for energies above the band gap the TSS is weakly localized. (b) Current distribution and a corresponding simple resistance model for Bi_2Se_3 (with a bulk gap of 0.3 eV) at 0.2 V applied bias. The magnetic layer with perpendicular anisotropy (PMA) inserted between the bottom TI surface and the substrate (not shown) shows a possible way to lower the shunting effect. (c) The surface current is defined as the current in the top layers; from the top layer to the layer with peaked current density. (d) Two ways to improve the surface-to-bulk current ratio: with a larger bulk band gap, the surface current increases while the total current decreases. Alternatively, by inserting a magnetic layer with perpendicular anisotropy at the bottom surface, the total current is reduced while the current at the opposite surface is almost unaffected, translating to a higher contribution for the top surface. V_s and V_d are $\pm V_b/2$.

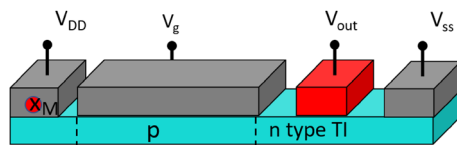


Figure 6. Figure shows the structure of a 1T-1MTJ memory based on TIPNJ. The TIPNJ acts as the transistor element of the memory. V_g is the gate voltage which controls the type of TI (P or N), which then can tune the TSS accordingly. The V_{out} is the output voltage which depends on the magnetization direction.

TIPNJ as a RAM

For the nanomagnet-as-source-contact configuration sans external gain, the gate controls the TIPNJ resistance and thus can be used to tune the total spin current and the spin torque applied on the nanomagnet. The structure of such a device is shown in Fig. 6, where the TIPNJ would have a transistor-like behavior. However, the maximum bias window is limited by the bandgap of the bulk TI. While we have discussed ways to alleviate the shunting issue, they might be challenging to achieve experimentally. Having said that, we looked at a possible 1 transistor-1 magnetic tunnel junction (1T1MTJ)³³ device based on TIPNJ. The switching of a magnet parallel or anti-parallel to the injected spin current polarization is known as the anti-damping switching. The required critical spin current is given by (ignoring the effects of the rotation of spins from interfacial Zeeman field)³⁴:

$$I_s \approx \frac{4q\alpha k_B T \Delta}{\hbar} \quad (10)$$

where Δ is the dimensionless energy barrier and α is the Gilbert damping. In the NEGF calculations, we take the spin current (as calculated from Eq. 8) in the two top layers to be the spin current interacting with the ferromagnetic. We justify it by looking at the current vs layer distribution (Fig. 5c), which shows the bulk effects come into effect after the first two layers. From the calculations, we can see that at low biases, spin currents of 100 μA at a width of 40 nm is achievable. This would be enough to switch a magnet with $\Delta = 40$, (needed for a reliable memory at room temperature¹⁶), $\alpha = 0.1$ at a write error rate (WER) of 10^{-9} and a switching time of 5 ns. The switching roughly translates to an energy consumption of ≈ 600 fJ/bit. It has been reported that 3DTI can be more energy efficient than heavy metals^{12,35}. However, by improving the interfacial quality of FM and heavy

metal, a significant increase in efficiency of heavy metal has also been reported^{9,36}. This suggests that although 3DTI might not provide higher energy efficiency compared to heavy metals, TIPNJ can operate at low voltages suitable for more energy efficient applications such as RAM. For instance, the tuning of the TI surface state can be achieved by small applied voltages ($V_d \leq 0.1 - 0.5$ V as compared to 0.9 V for CMOS transistors).

From a practical perspective, the required on/off current ratio β from reliable spin torque switching can be deduced from the error rate^{37,38}:

$$\beta = \log \left[\frac{\pi^2 \Delta}{4P_e} \right] / \log \left[\frac{\pi^2 \Delta}{4(1 - P_e)} \right] \quad (11)$$

For a typical write error rate $P_e = 10^{-9}$, the minimum β is about 5.5. In comparison, the $\beta \approx 8$ calculated from NEGF simulations shows that the gating mechanism works well enough for a TI based 1T-1MTJ RAM.

Volatile memory

Although it is possible to enhance the TIPNJ current through material engineering to achieve the required current to drive traditional RAM applications, we also explored what is possible under conditions where leakage current and interfacial imperfections lower the spin current generated on a TIPNJ. To quantify the driveability of the TIPNJ, we have self-consistently simulated the magnetodynamics of a nanomagnet under the TIPNJ current with the phenomenological Landau–Lifshitz–Gilbert (LLG) equation with a stochastic Langevin correction:

$$\begin{aligned} \frac{d\mathbf{m}}{dt} = & -\frac{\mu_0\gamma}{1+\alpha^2} \mathbf{m} \times \mathbf{h} - \frac{\alpha\mu_0\gamma}{(1+\alpha^2)} \mathbf{m} \times (\mathbf{m} \times \mathbf{h}) \\ & - \frac{\mu_B}{q\Omega M_s(1+\alpha^2)} \mathbf{m} \times (\mathbf{m} \times \mathbf{I}_s) \\ \mathbf{h} = & \mathbf{H}_{\text{eff}} + \boldsymbol{\tau}_s, \quad \boldsymbol{\tau}_s = \sqrt{\frac{2\alpha k_B T}{\mu_0\gamma\Omega M_s}} \mathbf{G} \end{aligned} \quad (12)$$

where $\mathbf{m}$ is the normalized magnetic moment; $\mathbf{H}_{\text{eff}}$ is the effective field from the intrinsic anisotropy of the magnet; M_s is the saturation magnetization; Ω is the volume of nanomagnet; $\mathbf{I}_s$ is the spin current with the vector pointing at the spin direction calculated using Eq. (8) as a function of the magnetic direction $\mathbf{m}$. $\boldsymbol{\tau}_s$ is a stochastic thermal torque term representing the thermal noise. $\mathbf{G}$ a three dimensional Gaussian white noise with mean $\langle G(t) \rangle = 0$ and standard deviation of $\langle G^2(t) \rangle = 1$. The spin current is calculated self-consistently from Eq. (8). Note that we use a macrospin model for the nanomagnet. The possible higher-order contributions from Hubbard energy and finite-size effects of the nanomagnet are ignored here. First, we propose a configuration (Fig. 7a) for an experimental verification of the switching of a nanomagnet from a TIPNJ. A symmetric structure is needed for flipping the magnetization back-and-forth. By applying a periodic voltage pulse (Fig. 7b) a periodic signal is generated (Fig. 7c) which can be read out by stacking an MTJ or spin valve structure on top of the switching magnet. Our simulation shows that the TIPNJ can reliably drive a small magnet with an adequate anisotropy energy

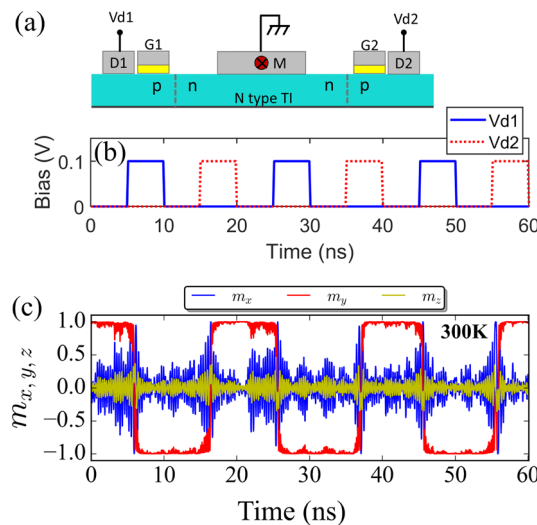


Figure 7. (a) Proposed experimental set-up to test gate tunable up to down (1 to 0) switching and vice-versa. Dual TIPNJ setup for switching the nanomagnet back and forth. (b) Voltage pulse profile. (c) Stochastic Landau Lifshitz Gilbert (LLG) simulation of magnetodynamics of an in-plane nanomagnet ($40 \times 20 \times 4$ nm) with $M_s = 400$ emu/cc, $\alpha = 0.1$ with thermal stability $\Delta \approx 5$ at 300 K. When the drain V_{d1} is activated with a positive bias relative to the source, the spin torque flips the magnet's in-plane magnetization ($m_y : +1 \rightarrow -1$), at which point further spin injection is blocked by spin-momentum coupling and the magnetization stays pinned. Subsequently activating the other TIPNJ with V_{d2} flips the magnet back to the $-y$ orientation.

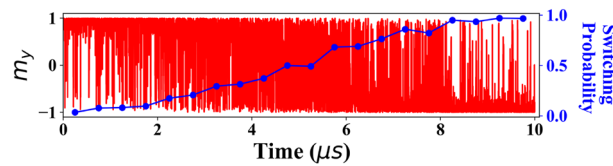


Figure 8. Programming a magnet based binary stochastic neuron using gate tunable PN junction barrier. Low-energy barrier magnet operating in the stochastic regime. A nanomagnet of dimension $20 \text{ nm} \times 18 \text{ nm} \times 5 \text{ nm}$ with approximately $\sim 1k_B T$ energy barrier. The current is applied as a linear ramp from $+5 \mu\text{A}$ to $-5 \mu\text{A}$ where negative current means the current is flowing through the other PN junction, which can be achieved by tuning the drain or gate bias.

barrier ($5k_B T$) to reliably switch back and forth with a current (proportional to the longitudinal dimension of the magnet) as low as $I_s = 10 \mu\text{A}$. When we further lower the anisotropy barrier, we entered a regime where the tunability of the TIPNJ is favored over its driveability. A nanomagnet with very small energy barrier ($\sim 1k_B T$) can operate in the stochastic switching regime. At room temperature, those magnets flip their magnetic moments frequently and stochastically under the influence of thermal noise. Room temperature simulations have shown a decrease in spin current in a factor of two in Bi_2Se_3 ²⁹. A small current can tip the balance of the two energy minimum orientations along the easy axis and affect the probability of the magnetic moment pointing at the direction decided by the current direction, with a binary stochastic neuron like tanh behavior on average. Figure 8 shows the tuning of switching probability under a small spin current $I_s = 5 \mu\text{A}$ generated from the TIPNJ for $\mathbf{m} = \pm \hat{y}$. The state (probability) of the nanomagnet can be extracted through a time-average measurement on the magnetic moment. This kind of setup can be used in generating random numbers³⁹, or stochastic signal processing⁴⁰. The energy cost for this is calculated to be $\approx 20 \text{ fJ/bit}$. The nanomagnet-TI contact resistance and the shunting does decrease the efficiency of the TI to generate the spin current, which can be countered by using a wider TI and magnet, which lowers the requirement for the spin current density. This would mean a trade-off between size density and energy consumption needs to be considered in an actual device. The hopping term between FM and TI where chosen here to minimize the contact resistance. However, with $t_{\text{FM-TI}} = t$, a reduction of around 20–25% in the spin current was seen. However, by taking smaller effective electron masses, this reduction can be decreased. As we take $m^* = 0.5$, there is still room for decreasing the effective mass.

Conclusion

In this paper, we have taken a closer look at the tuning of the TSS spin current in a TIPNJ as was originally proposed in¹⁴. We have analyzed the intricate nature of the charge-to-spin conversion on the TI surface with a nanomagnetic source contact. More importantly, we have clarified an important difference between the intrinsic charge-to-spin gain observed on a homogeneous TI surface, and the external gain useful for switching a nanomagnet. Additionally, we have looked at the limitations on the magnitude of the surface current, which is mostly determined by the bulk bandgap, and discussed two ways to improve it. While those issues seem to suggest the TIPNJ might not be able to offer a superior solution for common memory/logic applications, it can still work well as a RAM because of its gate tunability and low bias operation. Furthermore, the spatial programmability introduced by the PN junction may find use in novel applications such as probabilistic or neuromorphic applications. An interesting direction for this work is to look at magnets both as the source and drain to see how the parallel- anti-parallel orientation of the two magnets would affect the spin current generation in the TI.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 25 October 2022; Accepted: 21 May 2023

Published online: 10 June 2023

References

- Chen, Y. *et al.* Experimental realization of a three-dimensional topological insulator, Bi_2Te_3 . *Science* **325**, 178–181 (2009).
- Xia, Y. *et al.* Observation of a large-gap topological-insulator class with a single Dirac cone on the surface. *Nat. Phys.* **5**, 398 (2009).
- Zhang, H. *et al.* Topological insulators in Bi_2Se_3 , Bi_2Te_3 and Sb_2Te_3 with a single Dirac cone on the surface. *Nat. Phys.* **5**, 438 (2009).
- Beenakker, C. Colloquium: Andreev reflection and Klein tunneling in graphene. *Rev. Mod. Phys.* **80**, 1337 (2008).
- Stander, N., Huard, B. & Goldhaber-Gordon, D. Evidence for Klein tunneling in graphene p–n junctions. *Phys. Rev. Lett.* **102**, 026807 (2009).
- Sajjad, R. N. & Ghosh, A. W. Manipulating chiral transmission by gate geometry: Switching in graphene with transmission gaps. *ACS Nano* **7**, 9808–9813 (2013).
- Reza, A. K., Fong, X., Azim, Z. A. & Roy, K. Modeling and evaluation of topological insulator/ferromagnet heterostructure-based memory. *IEEE Trans. Electron Devices* **63**, 1359–1367. <https://doi.org/10.1109/TED.2016.2520941> (2016).
- Chiba, T., Takahashi, S. & Bauer, G. E. W. Magnetic-proximity-induced magnetoresistance on topological insulators. *Phys. Rev. B* **95**, 094428. <https://doi.org/10.1103/PhysRevB.95.094428> (2017).
- Rojas-Sánchez, J.-C. & Fert, A. Compared efficiencies of conversions between charge and spin current by spin-orbit interactions in two- and three-dimensional systems. *Phys. Rev. Appl.* **11**, 054049. <https://doi.org/10.1103/PhysRevApplied.11.054049> (2019).

10. Ghosh, S. & Manchon, A. Spin-orbit torque in a three-dimensional topological insulator-ferromagnet heterostructure: Crossover between bulk and surface transport. *Phys. Rev. B* **97**, 134402. <https://doi.org/10.1103/PhysRevB.97.134402> (2018).
11. Fan, Y. *et al.* Magnetization switching through giant spin-orbit torque in a magnetically doped topological insulator heterostructure. *Nat. Mater.* **13**, 699–704. <https://doi.org/10.1038/nmat3973> (2014).
12. Han, J. *et al.* Room-temperature spin-orbit torque switching induced by a topological insulator. *Phys. Rev. Lett.* **119**, 77702. <https://doi.org/10.1103/PhysRevLett.119.077702> (2017).
13. Ding, J. *et al.* Switching of a magnet by spin-orbit torque from a topological dirac semimetal. *Adv. Mater.* **33**, 2005909. <https://doi.org/10.1002/adma.202005909> (2021).
14. Habib, K. M., Sajjad, R. N. & Ghosh, A. W. Chiral tunneling of topological states: Towards the efficient generation of spin current using spin-momentum locking. *Phys. Rev. Lett.* **114**, 176801 (2015).
15. Liu, L., Moriyama, T., Ralph, D. & Buhrman, R. Spin-torque ferromagnetic resonance induced by the spin hall effect. *Phys. Rev. Lett.* **106**, 036601 (2011).
16. Liu, L. *et al.* Spin-torque switching with the giant spin hall effect of tantalum. *Science* **336**, 555–558 (2012).
17. Gu, B. *et al.* Surface-assisted spin hall effect in au films with pt impurities. *Phys. Rev. Lett.* **105**, 216401 (2010).
18. Taskin, A. & Ando, Y. Quantum oscillations in a topological insulator $\text{Bi}_{1-x}\text{Sb}_x$. *Phys. Rev. B* **80**, 085303 (2009).
19. Butch, N. P. *et al.* Strong surface scattering in ultrahigh-mobility bi_2se_3 topological insulator crystals. *Phys. Rev. B* **81**, 241301 (2010).
20. Jia, S. *et al.* Low-carrier-concentration crystals of the topological insulator $\text{bi}_2\text{te}_2\text{se}$. *Phys. Rev. B* **84**, 235206 (2011).
21. Ganguly, S., Camsari, K. Y. & Ghosh, A. W. Analog signal processing using stochastic magnets. *IEEE Access* **9**, 92640–92650. <https://doi.org/10.1109/ACCESS.2021.3075839> (2021).
22. Camsari, K. Y., Faria, R., Sutton, B. M. & Datta, S. Stochastic p -bits for invertible logic. *Phys. Rev. X* **7**, 031014. <https://doi.org/10.1103/PhysRevX.7.031014> (2017).
23. Hong, S., Diep, V., Datta, S. & Chen, Y. P. Modeling potentiometric measurements in topological insulators including parallel channels. *Phys. Rev. B* **86**, 085131 (2012).
24. Roy, U. *et al.* Magnetization switching of a metallic nanomagnet via current-induced surface spin-polarization of an underlying topological insulator. *J. Appl. Phys.* **117**, 163906 (2015).
25. Sancho, M. L., Sancho, J. L., Sancho, J. L. & Rubio, J. Highly convergent schemes for the calculation of bulk and surface green functions. *J. Phys. F* **15**, 851 (1985).
26. Datta, S. *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, 1997).
27. Xie, Y., Tan, Y. & Ghosh, A. W. Spintronic signatures of Klein tunneling in topological insulators. *Phys. Rev. B* **96**, 205151. <https://doi.org/10.1103/PhysRevB.96.205151> (2017).
28. Duan, X., Li, X.-L., Semenov, Y. G. & Kim, K. W. Nonlinear magnetic dynamics in a nanomagnet-topological insulator heterostructure. *Phys. Rev. B* **92**, 115429 (2015).
29. Gupta, G., Lin, H., Bansil, A., Jalil, M. B. A. & Liang, G. Role of acoustic phonons in bi_2se_3 topological insulator slabs: A quantum transport investigation. *Phys. Rev. B* **89**, 24519. <https://doi.org/10.1103/PhysRevB.89.24519> (2014).
30. Hasan, M. Z. & Kane, C. L. Colloquium: Topological insulators. *Rev. Mod. Phys.* **82**, 3045–3067. <https://doi.org/10.1103/RevModPhys.82.3045> (2010).
31. Mahfouzi, F., Nagaosa, N. & Nikolić, B. K. Spin-to-charge conversion in lateral and vertical topological-insulator/ferromagnet heterostructures with microwave-driven precessing magnetization. *Phys. Rev. B* **90**, 115432. <https://doi.org/10.1103/PhysRevB.90.115432> (2014).
32. Tokura, Y., Yasuda, K. & Tsukazaki, A. Magnetic topological insulators. *Nat. Rev. Phys.* **1**, 126–143. <https://doi.org/10.1038/s42254-018-0011-5> (2019).
33. Wang, Z. *et al.* High-density nand-like spin transfer torque memory with spin orbit torque erase operation. *IEEE Electron Device Lett.* **39**, 343–346. <https://doi.org/10.1109/LED.2018.2795039> (2018).
34. Sun, J. Z. Spin-current interaction with a monodomain magnetic body: A model study. *Phys. Rev. B* **62**, 570–578. <https://doi.org/10.1103/PhysRevB.62.570> (2000).
35. Han, J. & Liu, L. Topological insulators for efficient spin-orbit torques. *APL Mater.* **9**, 060901. <https://doi.org/10.1063/5.0048619> (2021).
36. Zhu, L., Zhu, L. & Buhrman, R. A. Fully spin-transparent magnetic interfaces enabled by the insertion of a thin paramagnetic nio layer. *Phys. Rev. Lett.* **126**, 107204. <https://doi.org/10.1103/PhysRevLett.126.107204> (2021).
37. Butler, W. H. *et al.* Switching distributions for perpendicular spin-torque devices within the macrospin approximation. *IEEE Trans. Magn.* **48**, 4684–4700. <https://doi.org/10.1109/TMAG.2012.2209122> (2012).
38. Xie, Y., Behin-Aein, B. & Ghosh, A. Numerical fokker-planck simulation of stochastic write error in spin torque switching with thermal noise. In *2016 74th Annual Device Research Conference (DRC)* 1–2. <https://doi.org/10.1109/DRC.2016.7548462> (2016).
39. Fukushima, A. *et al.* Spin dice: A scalable truly random number generator based on spintronics. *Appl. Phys. Express* **7**, 083001 (2014).
40. Ganguly, S., Camsari, K. Y. & Ghosh, A. W. Analog signal processing using stochastic magnets. arXiv preprint [arXiv:1812.08273](https://arxiv.org/abs/1812.08273) (2018).

Acknowledgements

This work is supported by the Army Research Lab (ARL) and in part by the NSF I/UCRC on Multi-functional Integrated System Technology (MIST) Center; IIP-1439644, IIP-1439680, IIP-1738752, IIP-1939009, IIP-1939050, and IIP-1939012. We acknowledge helpful discussions with Mahesh Neupane and George De Coster at ARL.

Author contributions

Y.X. and H.V. wrote the main manuscript text. Y.X. and H.V. prepared the figures. Y.X. and A.G. conceived of the project. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.V.

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

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023