Classical Computing with Topological States: Coping with a post-Moore World

Avik Ghosh
ECE
Physics

UNIVERSITY of VIRGINIA

Virginia **N**an**O** computing **(ViNO)**

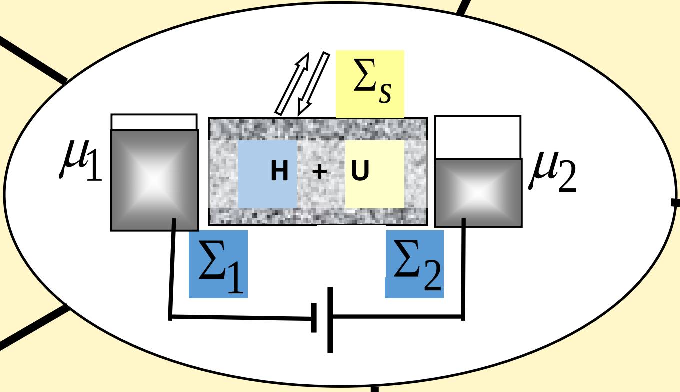
Director of NSF site on **M**ultifunctional **I**ntegrated **S**ystems **T**echnology (**MIST**)

$$ NSF, NASA, SRC, DARPA, ORNL

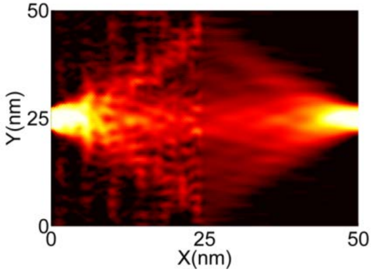
Purdue U, June 10, 2021

S. Ganguly, Y. Tan, H. Vakili, G. Morshed, N. Sakib
R. Sajjad, M. Habib, M. Elahi

ViNO group Grand Challenges

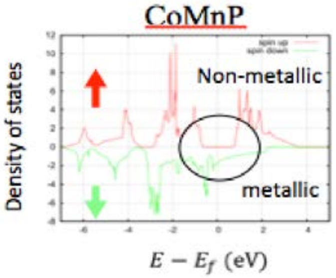


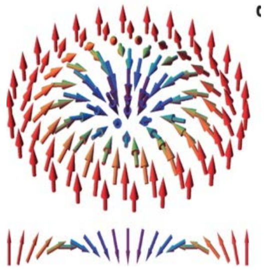
1. 2D-MATERIALS



What can we do ***Uniquely*** with 2-D materials?

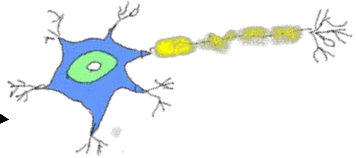
2. MAGNETS





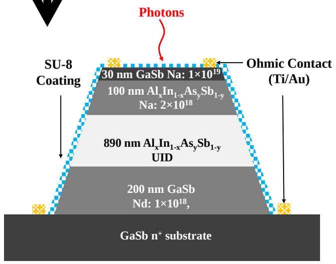
What limits Ultimately scaled memory?

3. NEUROMORPHIC COMPUTING



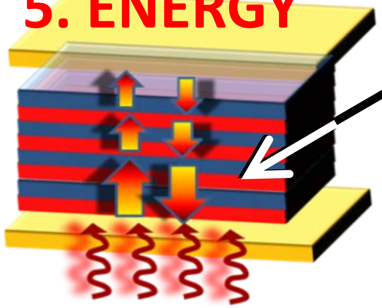
Can we do real-time learning & inferencing?

4. SINGLE PHOTON DETECTION

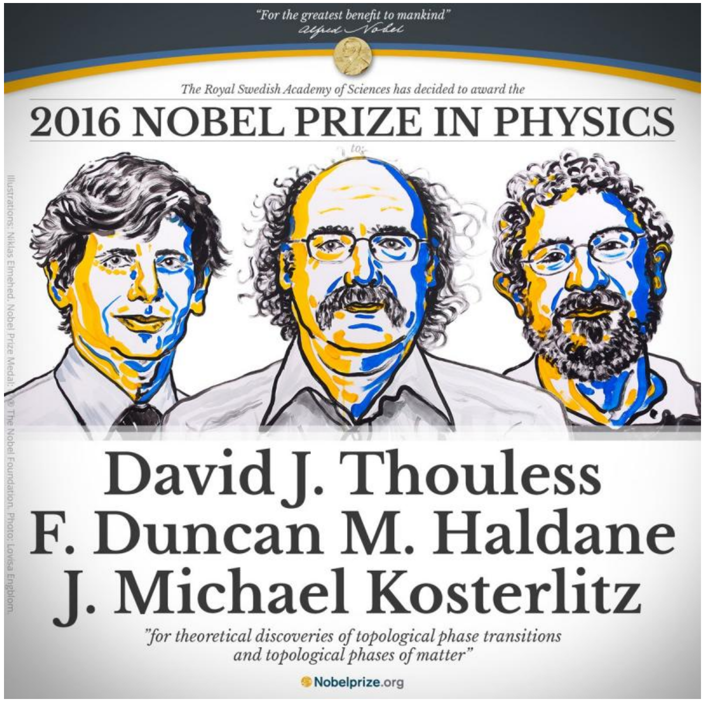


What limits single photon sensing?

5. ENERGY



Can we design a Perfect Thermal Glue?



Topological Insulators…
Topological Superconductors…
Weyl Semi-Metals…

Increased coherence, higher mobility
decreased scattering

Topology helps with materials classification and discovery

Can it help device performance more significantly?

High Density Memory, Low power Logic

SKYRMIONS and DIRAC FERMIONS

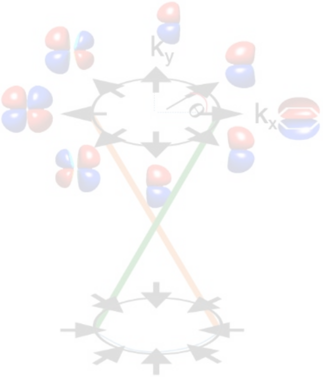
Talk Outline: Topology in classical computing

- Challenge of Moore's law: $E_{diss} = Q \,.\, \Delta V$

- Reducing charge Q → Ultrasmall Magnetic bits

Challenge is thermal fluctuations.
Real-space topological winding can circumvent them
→ **Magnetic Skyrmions for temporal memory**

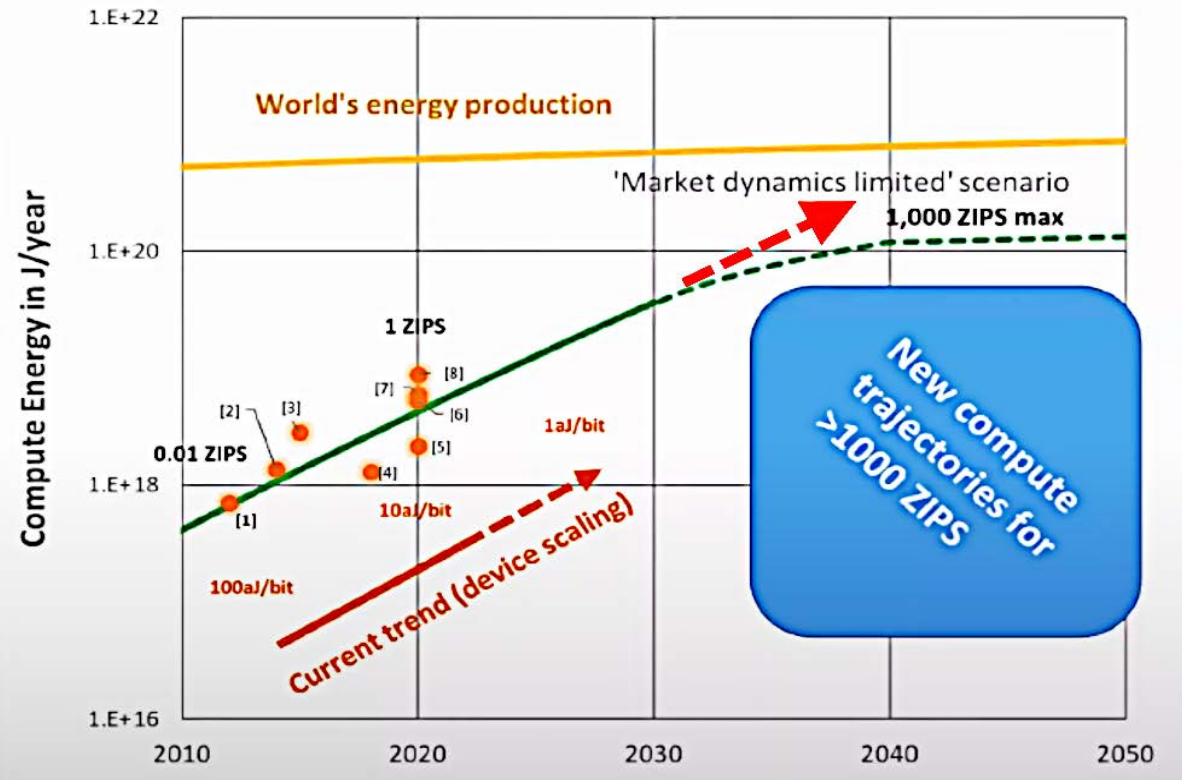
- Reducing Voltage ΔV → Operating below Boltzmann limit



Challenge is spontaneous transmission and error rate
Topological winding in k-space limits transmission
→ **2D Dirac Fermions for Klein Tunnel Switches**

Le roi est Moore, vive le roi

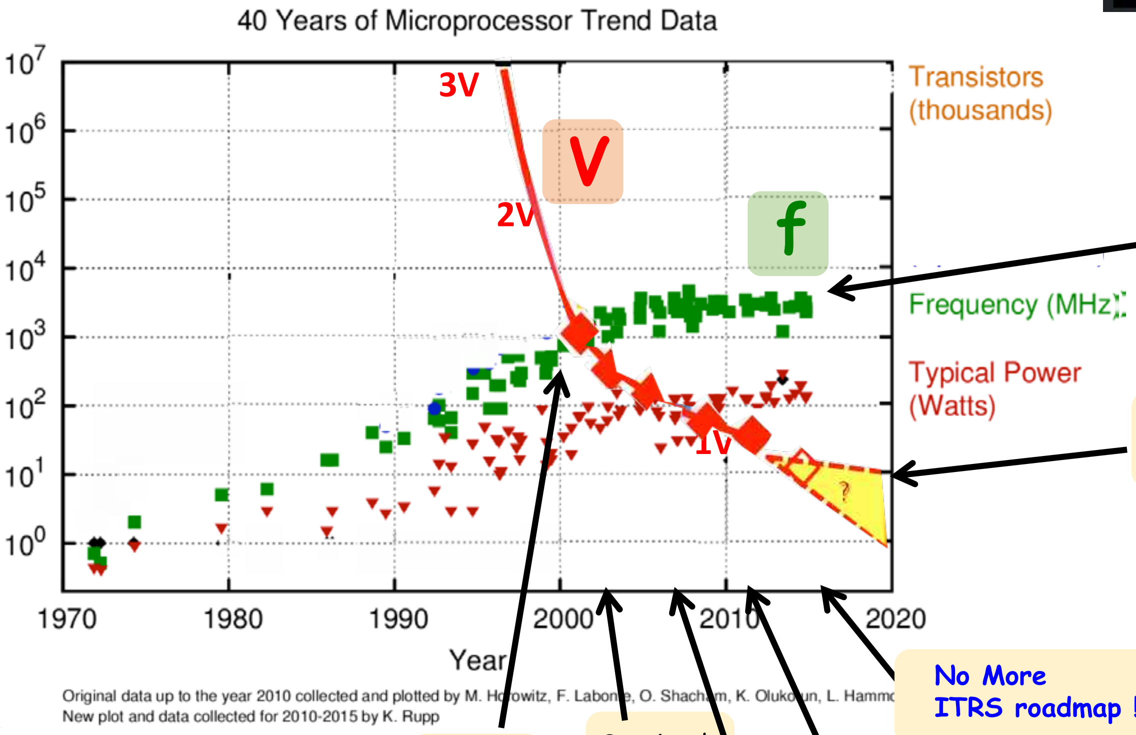
- Computing requires shuttling charges around
- Currents create ‘friction’ and heat
- This heat has now become prohibitive



www.diamandis.com/blog/proof-exponential-tech-growth
SRC Decadal Plan, 2020

Alive? Dead? Dead-Alive?

$$P \propto CV^2f$$



Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammo
New plot and data collected for 2010-2015 by K. Rupp

Limited by
Heat

Limited by
Error tolerance

Freq & V
Scaling –
Death of
Dennard
Scaling

Strained
Si

High-K
Metal Gate

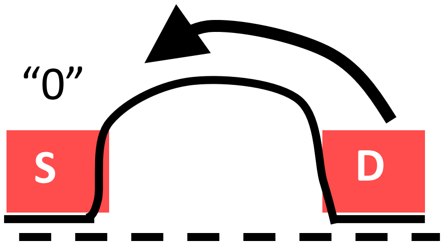
NonPlanar
Gate

**No More
ITRS roadmap !**

New world of
Software driven Hardware
(ASICs, accelerators
for IoT edge devices)

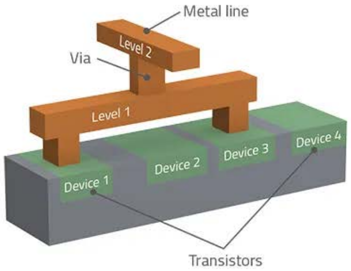
$$E_{diss} \propto Q\Delta V$$

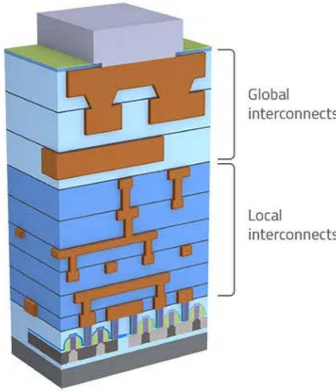
Need large ΔV to avoid bit flip errors (RELIABILITY)



ΔV >> thermal voltage ~ $k_B T$ ~25 mV to avoid error within 10^{-12}

Need large Q to charge up all the interconnects to ΔV (DRIVABILITY)





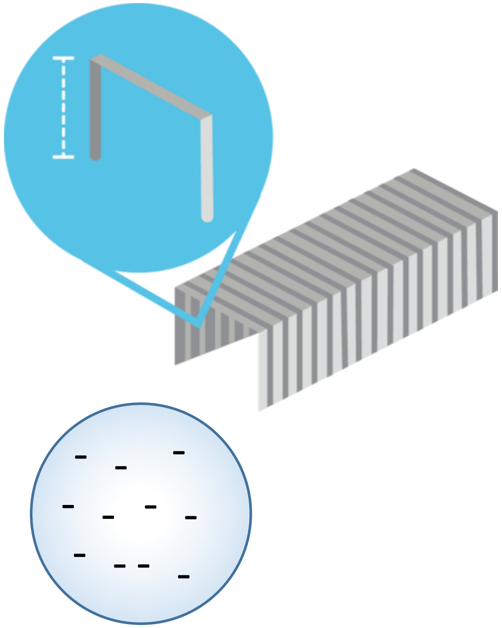
C ~ 1 fF for metal Interconnects

Q ~ 10,000 e !

Magnets can have low "Q"

$E_{diss} \propto 10{,}000\ k_BT$

$= 250\ eV$



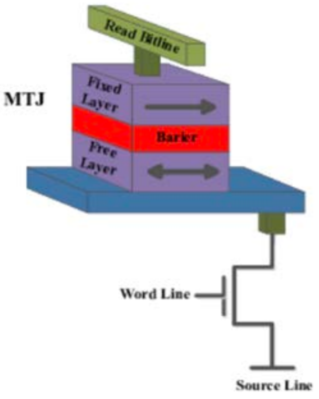
"Staple" information carriers together

Magnets ~ 10,000 spins
But they act like one giant spin
(Salahuddin, Behin-Aein, Datta)

Ion channels ~ 3 ions
Act like one single ion
IEEE TED 63, 1681 (2016)

What sets "Q" ?

1D Fokker Planck in perpendicular magnet STTRAM

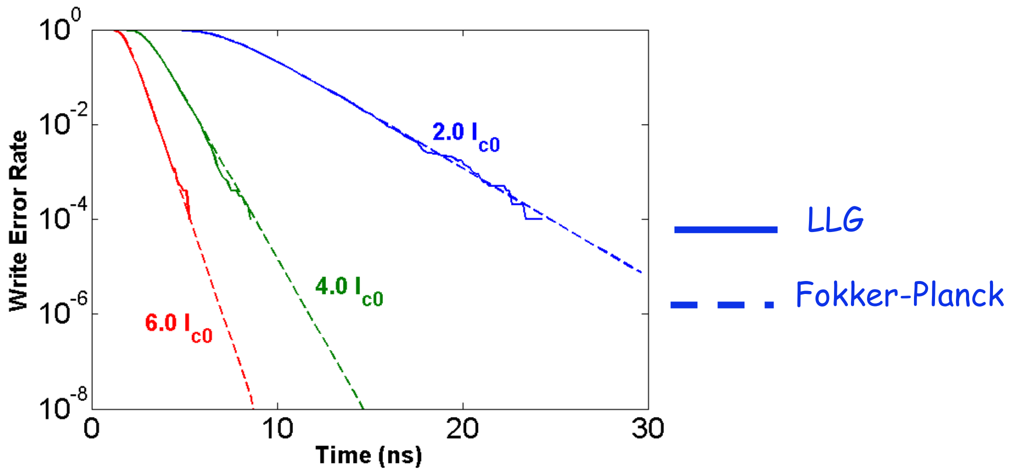


Angular momentum conservation (Drivability) & Boltzmann (Reliability)

$(Q_c/q) \times (P\hbar\gamma/2) = M_s\Omega(1+\alpha^2)$

WER = $(E_b/k_BT) \times e^{-2Q/Q_c}$

IEEE TED 59,2221 ('12)
64, 319 ('17)
arxiv:2101.09947



100 x 20 x 1 nm Fe magnet ~ 10^4spins, need 10^6 els @ WER~10^{-9}
1 ns switching, J ~ 50MA/cm^2 = 0.5 TA/m^2
High cost because of shared read-write line (several electrons x oxide barrier)

SOTRAM: Separate write/read lines, all metallic, I higher by ~10X, V lower by ~10X

Penalty for reducing Q → Reliability

$(Q_c/q) \times (P\hbar\gamma/2) = M_s\Omega(1+\alpha^2)$

$\mathrm{WER} = (E_b/k_BT) \times e^{-2Q/Q_c}$

Smaller Q → smaller volume Ω
But thermal stability barrier goes down $E_b \sim M_sH_k\Omega$

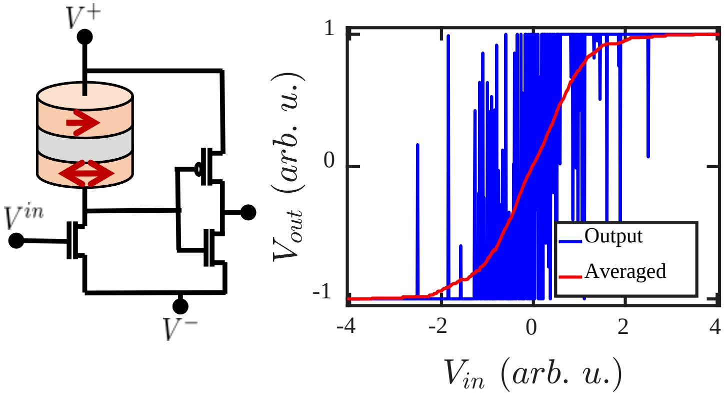
This reduces dynamic error rate WER
But it increases static error rate

'Superparamagnetic limit' $\tau \sim \tau_0 e^{E_b/k_BT}$

To get to small volume, 1-2Tb/in^2 – read/write/stability

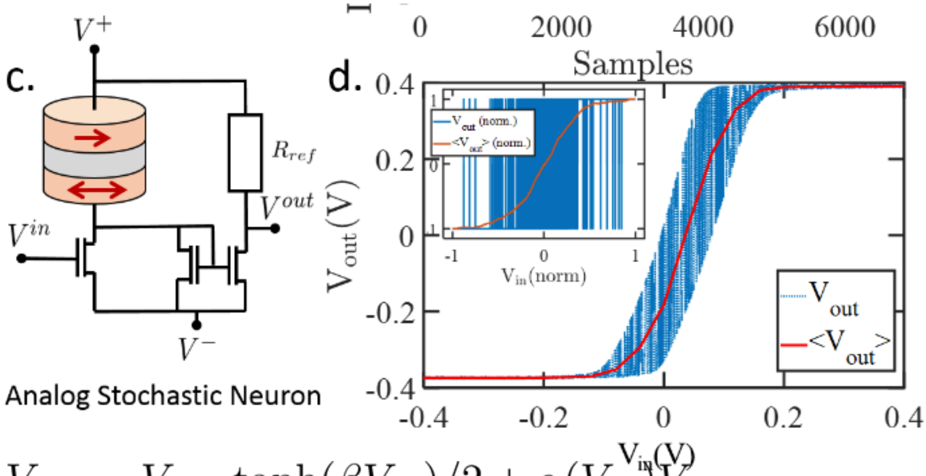
- Heat Assisted Magnetic Recording (HAMR)
- Bit Patterned Recording (BPR)
- Topological Distortions (Skyrmions)

Small barrier not necessarily bad !!



$$V_{out} = \mathrm{sgn}\Big(\tanh(\beta V_{in}) + \alpha(V_{in})V_{rnd}\Big)V_{DD}/2$$

Binary Stochastic Neurons



Analog Stochastic Neuron

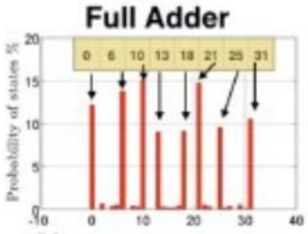
$$V_{out} = V_{DD}\tanh(\beta V_{in})/2 + \alpha(V_{in})V_{rnd}$$

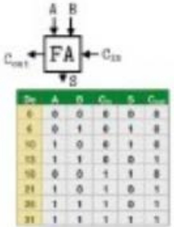
Analog Stochastic Neurons

Stochasticity allows you to sample the phase space quickly
Solve Quadratic Binary Optimization Problems (QUBO)

NP problems (TSP)

Complex Logic Operations with invertibility

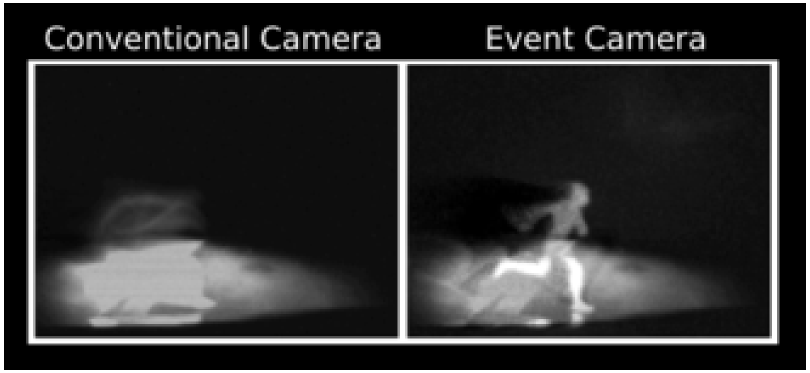




Dec	A	B	C_{in}	S	C_{out}
0	0	0	0	0	0
6	0	1	0	1	0
10	1	0	0	1	0
13	1	1	0	0	1
18	0	0	1	1	0
21	1	0	1	0	1
25	1	1	1	0	1
31	1	1	1	1	1

Kamsari, Ganguly, Sutton, Datta (Purdue)

Reservoir Computing → Time-domain learning and inferencing
(Possible Application – Fast Event Based Imaging)

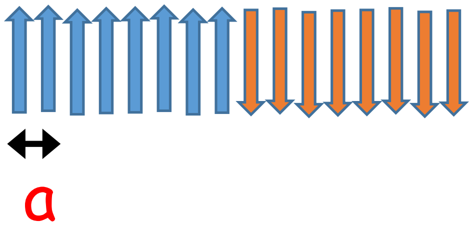


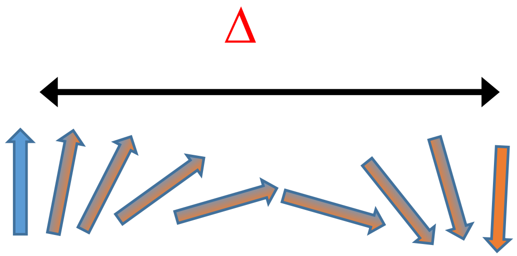
ICONS 2020, SPIE 2019, IECON 2018, SPIE 2018 …

IEEE Access 2021, IEEE JxCDC 2020, VLSI Soc 2020, ICONS 2020, SPIE 2019…

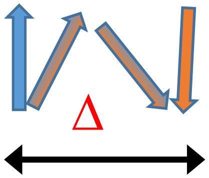
To create a large barrier in a scaled magnet → defect

Let’s look at Domain walls between up and down spins





Ferromagnetic exchange J (Hund’s Rule/Pauli Exclusion)
wants to keep spins aligned, so spin flip costs energy.
Compromise → flip spins gradually

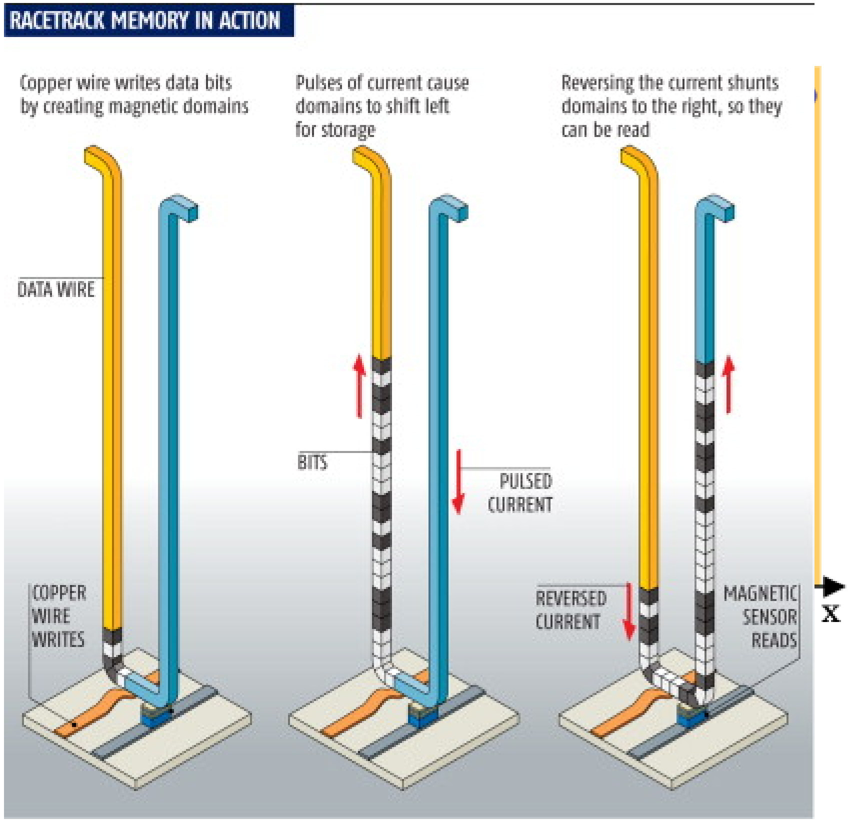


Uniaxial Anisotropy K (Magnetocrystalline, Shape)
wants to keep spins along a fixed axis

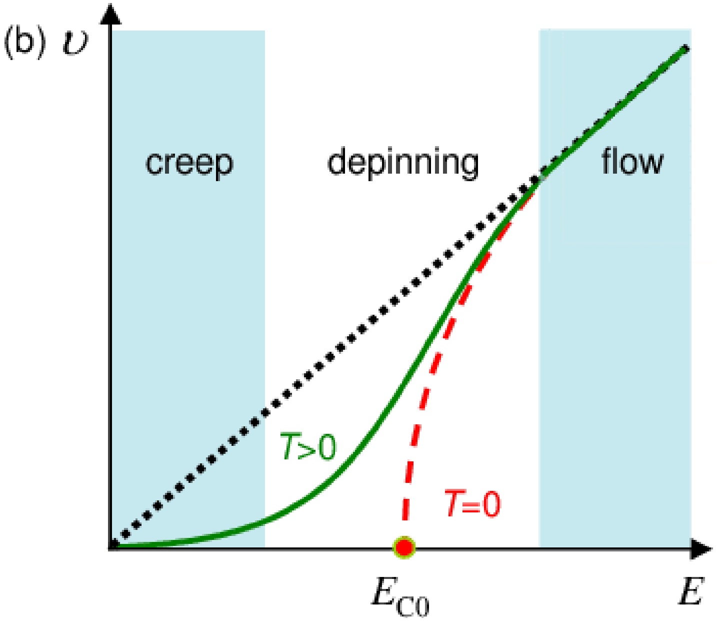
Domain Wall Width

$\Delta \sim \sqrt{J/Ka}$

Smaller Q in all metallic Domain Walls



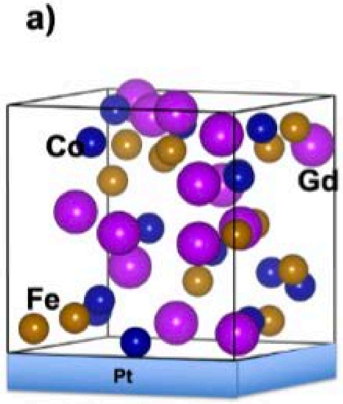
IBM Stuart Parkins

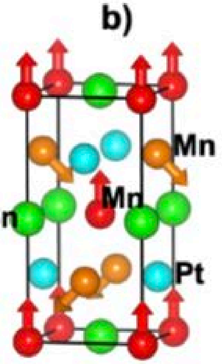


DW Pinning leads to creep – Nonlinear motion

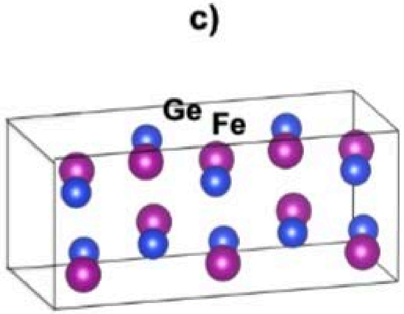
What happens if we add spatial inversion symmetry breaking?

a-FiM
on
Heavy
Metal





D_{2d} Mn_2PtSn



B20 solids (FeGe,MnSi)

Introduces
Dzyaloshinskii-Moriya
(DMI) field

$D = z \times r_{ij}$

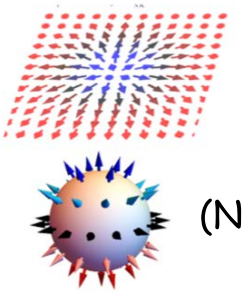
$D = \sigma_z . r_{ij}$

$D = r_{ij}$

Creates a small vortex like excitation (**SKYRMION**)
with winding texture in 2-D

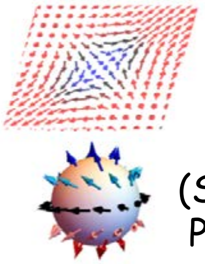
Spins must be
perpendicular to this field

$$d/dt \begin{pmatrix} x \\ y \end{pmatrix} = D_{\perp}$$



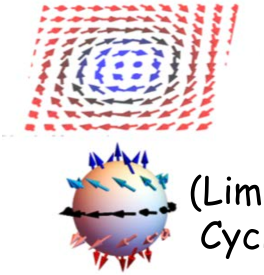
(a) Neel Skyrmions

(Nodes)



(d) AntiSkyrmions
(D2d tetragonal inverse

(Saddle
Pts)



(b) Bloch Skyrmions

(Limit
Cycles)

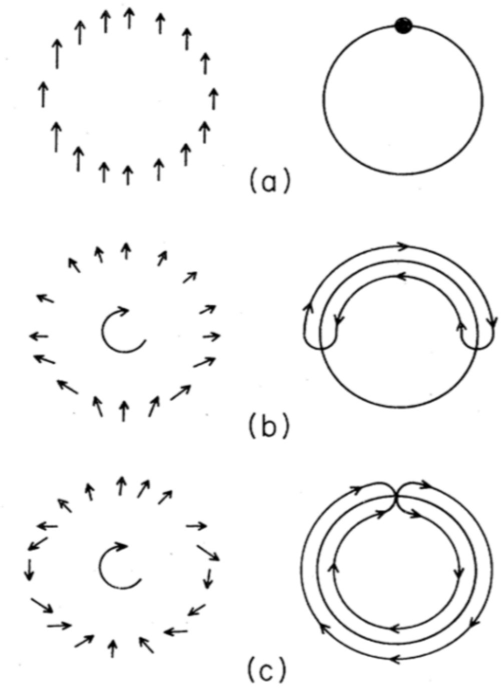
How much winding?

TOPOLOGY

$\langle \Psi_k | \Psi_{k+dk} \rangle \sim \exp[iA.dk]$

Winding # $N_{sk} = \oint A.dk/2\pi \sim M.(dM/dx \times dM/dy)$

Line integral of Berry curvature → topological invariant (like Gauss-Bonet Curvature)



Classify
by winding#

$N_{sk} = 0$ → Vortex, Bubble

$N_{sk} = 1/2$ → Meron

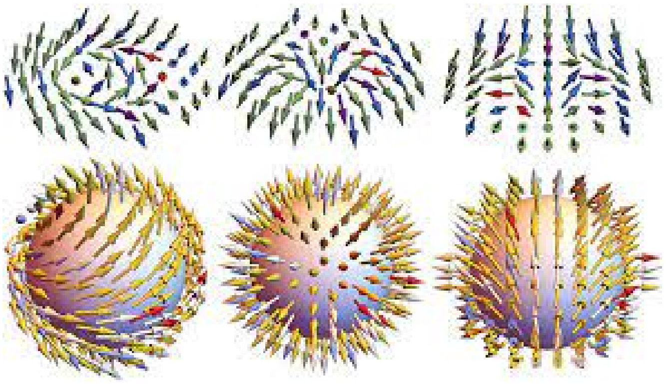
$N_{sk} = 1$ → Skyrmion

$N_{sk} = -1$ → Antiskyrmion

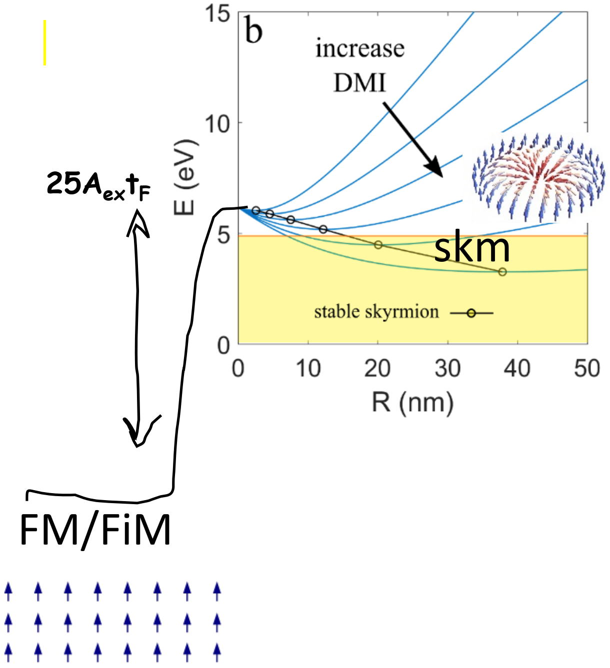
Sub-Classify
by domain angle
or helicity

Neel skyrmion
$\psi = 0$

Bloch skyrmion
$\psi = \pi/2$

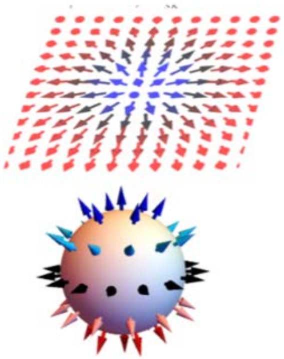


Winding of spins create a metastable state



- DMI, winding, exchange stiffness and curvature stabilize an isolated skyrmion by introducing rigidity
- Smaller skyrmions (lower DMI) are less stable.

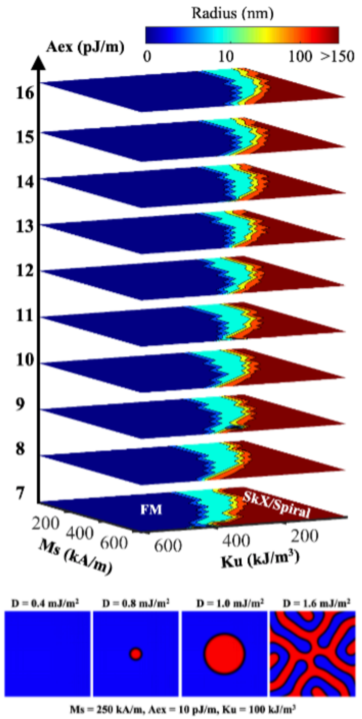
- Skm stereographically maps onto a sphere
- Nsk is a topological invariant
  Cannot deform spins continuously to melt into the background
- Will need large currents, edges, defects, or atomistic fluctuations to break it apart



(a) Neel Skyrmions
(e.g. CoGd on Pt)
$\psi = 0$, $N_{sk} = 1$

arxiv:1909.09446, 2101.09947
IEEE Trans Mag 56, 1500108 ('20)
Sci Rep 9, 9964 ('19)

How to get small and stable skyrmions?



Ms = 250 kA/m, Aex = 10 pJ/m, Ku = 100 kJ/m³

$R_{sk} \propto \Delta$ (DW width)

$\Delta \sim D/K$

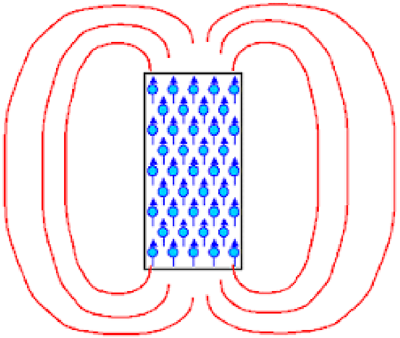
$K = K_u - \mu_0 M_s^2/2$

Large Magnetization creates stray fields that oppose anisotropy

Andy Kent (NYU)

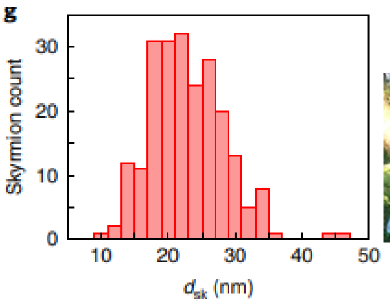
Joe Poon (UVA)

Geoff Beach (MIT)

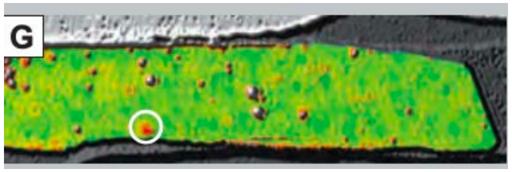


Reduce M_s → Ferrimagnets near compensation

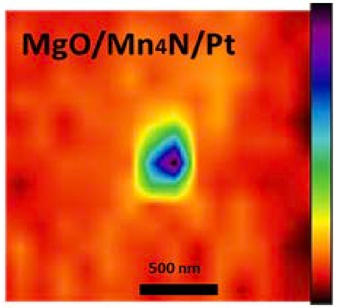
10-20nm STXM data on CoGd (Beach, MIT)
Vary Temp to hit compensation

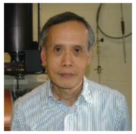


Romming et al
~6-7 nm in PdFe bilayers
at 4.2K



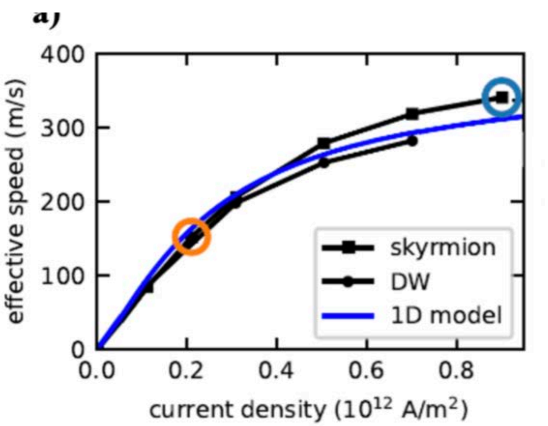
MFM data (Poon, UVA/Andy Kent NYU)
Mn4N + **Pt-W underlayer**





Fast, Non-Volatile and static Skyrmions

$$v \sim \theta_{sh} J / M_s \sqrt{[N_{sk}^2 + \alpha^2]}$$



Nice and Linear

Low Ms

~ 700m/s skms in CoGd
~1000m/s DWs in Garnet AFMs

Low α

MnGa Ferrimagnetic Heuslers

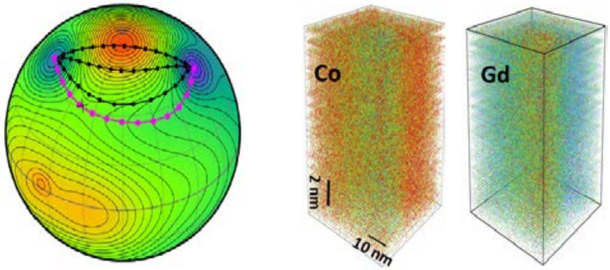
Y. Xie et al, arxiv:1909.09446

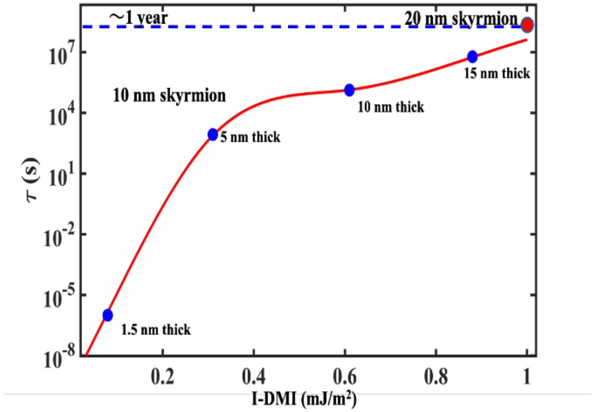
arxiv:2101.09947

activation energy

$$\Gamma(T) = \Gamma_0(T) e^{-\frac{\Delta E}{k_B T}}$$

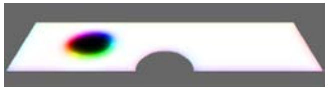
Attempt frequency (Entropy term)



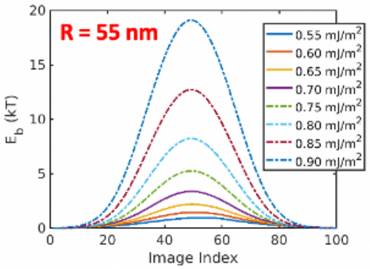


Decade-Long Skyrmion Lifetimes in Ferrimagnetic GdCo at Room Temperature, C. T. Ma, S. J. Poon.

Preliminary data from MIT show large ~100nm skyrmions with low diffusion constant < 10^{-20} m^2/s, << literature

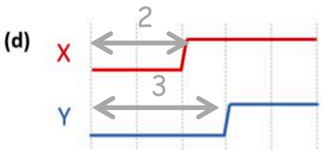


800 x 200 x 3 nm GdCo racetrack
+ 100nm circular "notch"



~40k_BT barrier (10 yr retention), allows ~50 nm skyrmions to unpin at 4 ns with 1.4×10^{11} A/m^2 < j < 10^{12} A/m^2

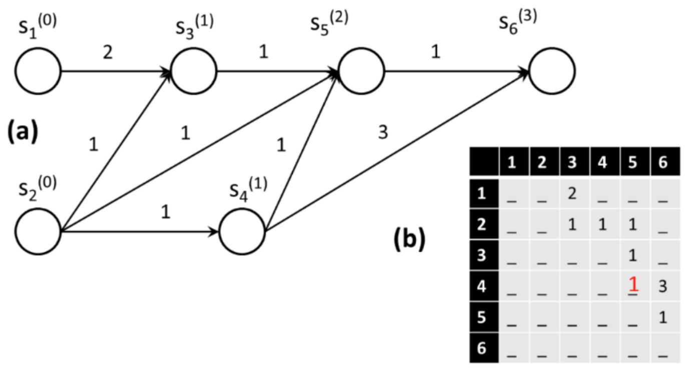
What can we use it for? HD memory and temporal memory



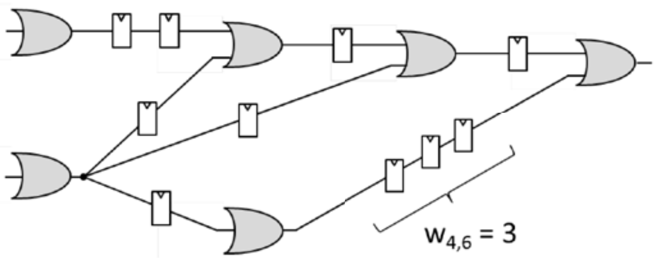
- Information stored as timing delays (energy at edges)

Strukov group, UCSB

Applications - Solving a MIN/MAX problem



	1	2	3	4	5	6
1	_	_	2	_	_	_
2	_	_	1	1	1	_
3	_	_	_	_	1	_
4	_	_	_	_	1	3
5	_	_	_	_	_	1
6	_	_	_	_	_	_



OR gates at nodes

Delay flip-flops for Edge weights

Input to output MIN path with weighted edges

Shortest path since one node activated activates OR gate (2 cycles). AND gate would give longest path

In-sensor analog data processing (continuous amplitude, discrete time)

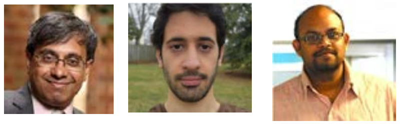
- Fast De novo alignment of millions of short DNA base pairs (NIST, UCSB)
- Text alignment (Pentti Kenarva)
- Dynamic time Warping (time series similarity – robotics, speech, ECG)
- Depth info from binocular image matching

Scalable because of Dynamic Programming (solutions of sub-computations are partial solutions of total computation)
Each node calculates shortest path from node to itself. Adjacent nodes use optimal solutions from diagonal predecessors to calculate their own scores.

Computation only at the wavefronts → effective gating to save energy (new way of representing data, not by energy-latency trade-off)

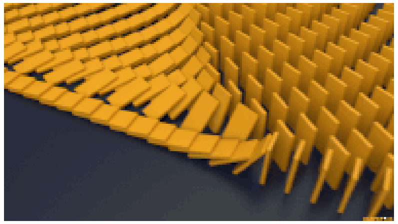
STORAGE: Flip-Flips (energy hungry), or memristors (nonlinear)

How to store Temporal Information (wavefront arrival times?)



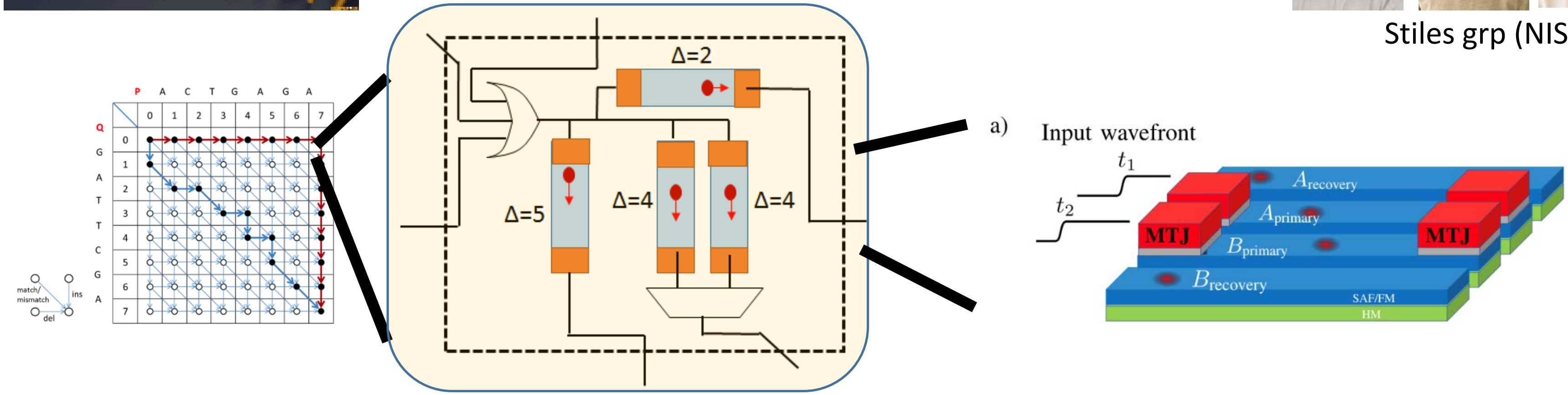
Ghosh grp

Stan grp



IEEE JxCDC 6, 107 ('20)

Stiles grp (NIST)

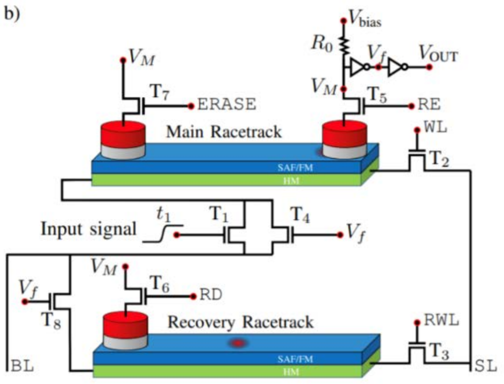


- For pattern matching we store delays (thresholds for decision trees)
- Each racetrack only needs ONE skyrmion (+ recovery racetrack with 1 skyrmion for copy, and 8 switching transistors)
- Four bit information in 16 possible positions of each skyrmion along racetrack
- Position of skyrmion == Duty cycle of input == arrival time of wavefront
- Utilize – Tunability and Linearity to convert time into space

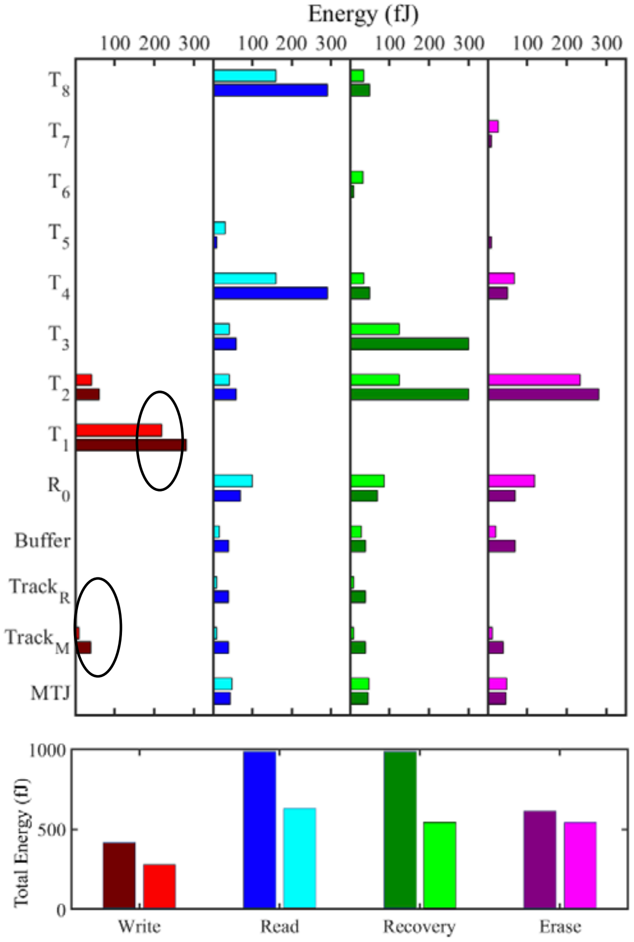
Encode t=1

Encode t=2

Encode t=16



Vakili, Sakib, Ganguly, Stan,
Daniels, Madhavan, Stiles, Ghosh, JxCDC 2020



- Analog storage -- Compact, potentially low energy How costly is this?

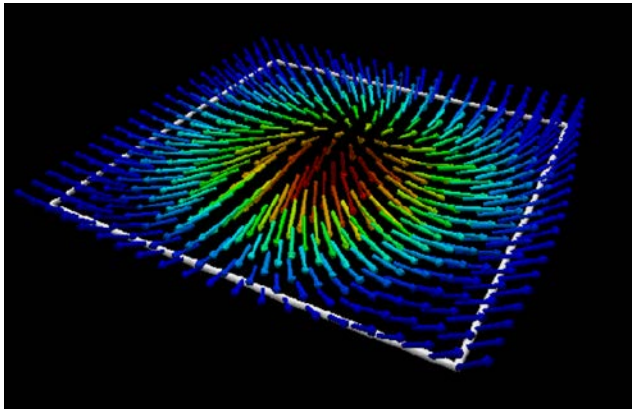
(< 10 fJ in racetrack metal ~2 pJ in Overhead for 8 transistors) footprint)

Doing this with equivalent 4-bit CMOS memory:

- Clock digitized with up-counter – 94 transistors
- S-R latches for storage – 32 transistors
- Readout with 4x16 decoder – 172 transistors

(~300 transistors)

Playback (low damping): EDP ~32X lower than 1GHz CMOS @ 1 cycle
10-100 reads per write cycle → EDP ~2400X-2700X lower
(Higher damping) trade-off footprint/resistance
Non-volatility: Turn off for long times (seconds-hrs for cache)

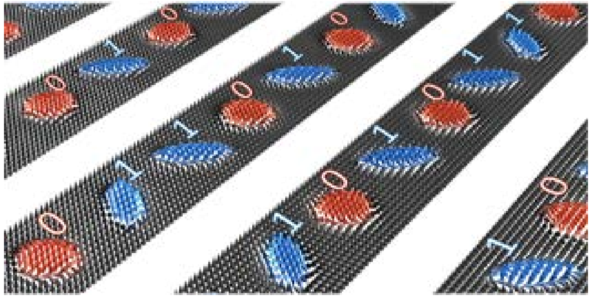


Topology of spin winding in-real space
stabilizes tiny mobile magnets

Small Q, with topology enabled stability barrier

Talk Outline: Topology in classical computing

- Death of Moore's law: $E_{diss} = Q \,.\, \Delta V$

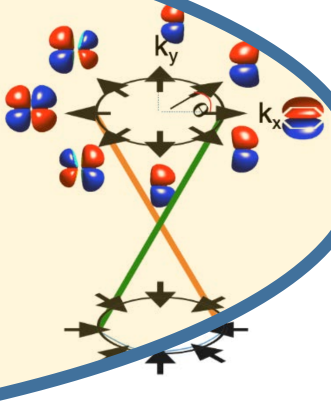


- Reducing charge Q →

Magnetic Excitations for temporal memory

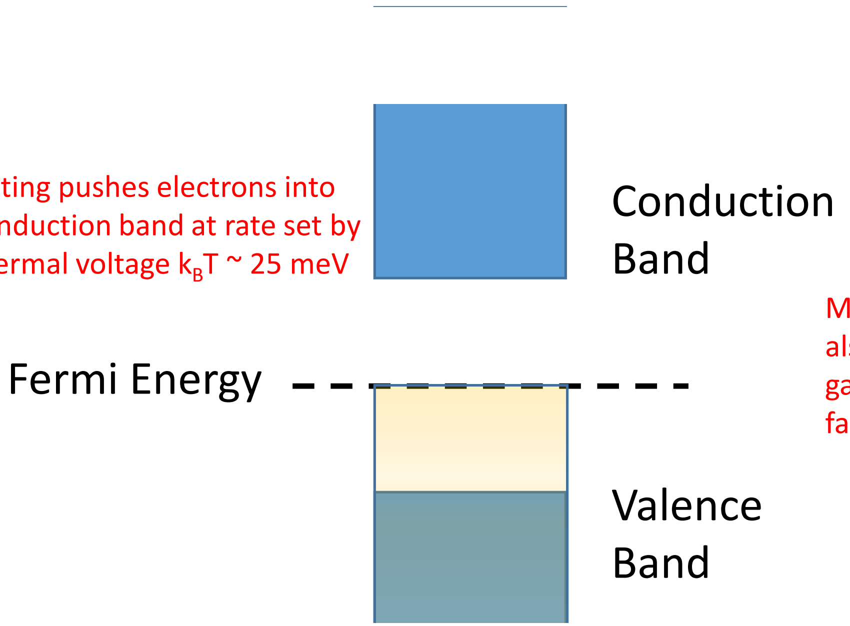
- Reducing Voltage ΔV → Operating below Boltzmann limit

2D Materials for Klein Tunnel Switches



How to Reduce ΔV ?

$E_{diss} \propto Q\Delta V$



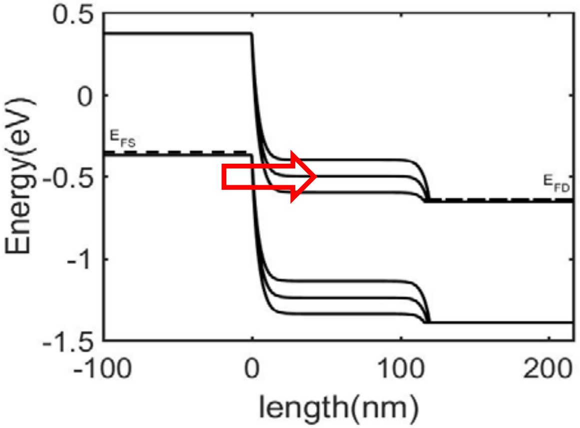
Gating pushes electrons into conduction band at rate set by thermal voltage $k_BT \sim 25$ meV

Conduction Band

Fermi Energy

Making gap/barrier also change with gate bias will give faster turn-on

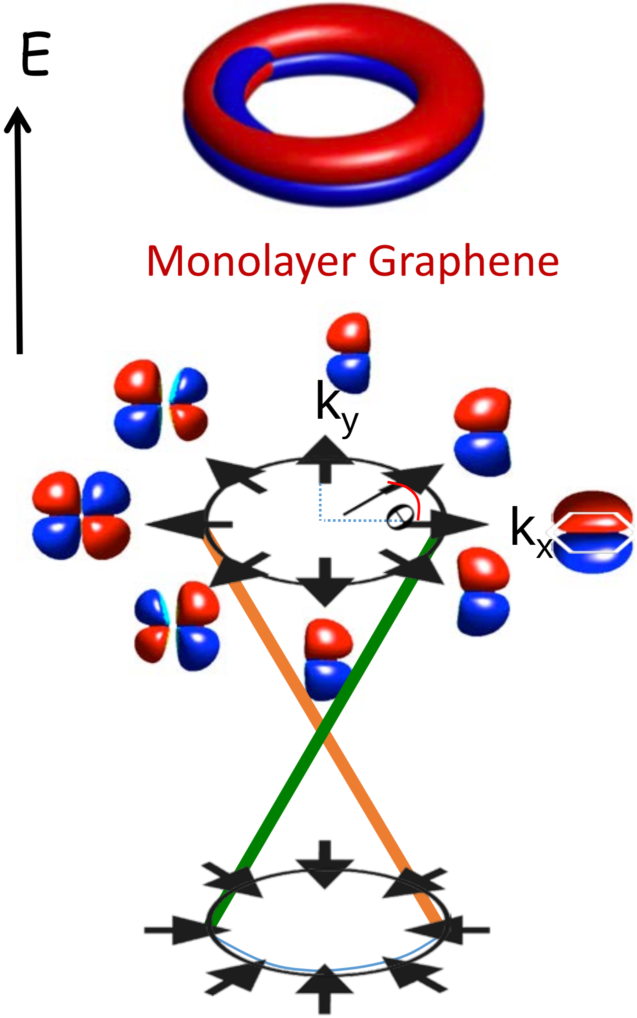
Valence Band



Tunnel-FETs (Premature Tunneling due to traps, Auger)

Topology of emerging materials

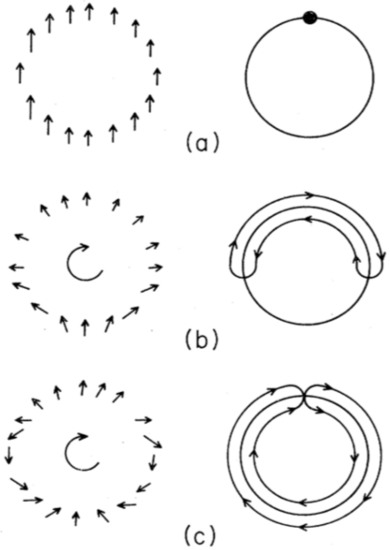
Q: Line Integral of Berry Curvature



Monolayer Graphene

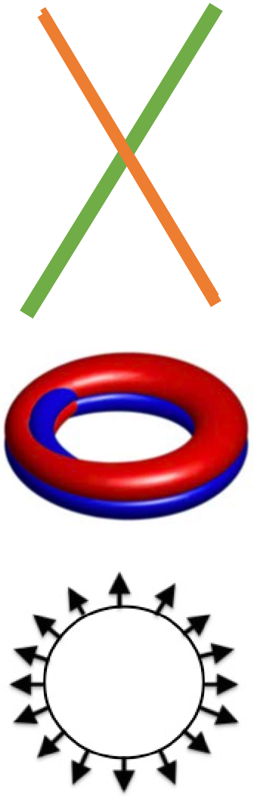
Electrons wind around Fermi surface and become orthogonal (this is responsible for the gaplessness of graphene)

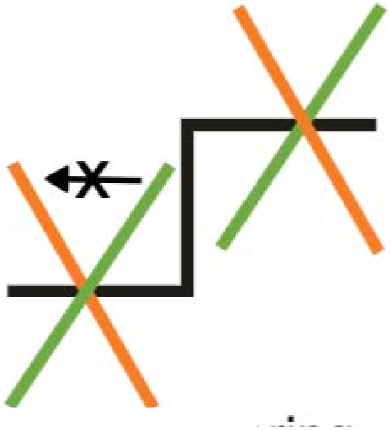
Q = 1/2



Transmitting electrons at a graphene PN junction

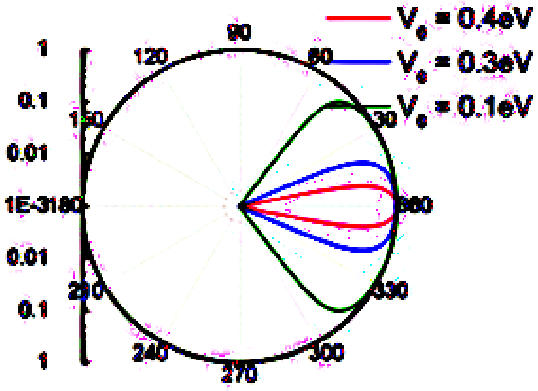
Pseudospin conservation and Klein tunneling sets transmission across GPNJ





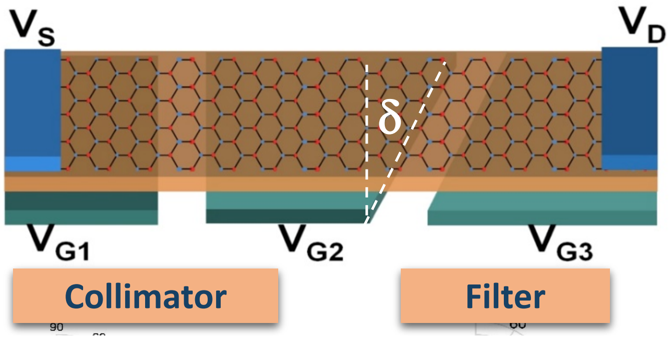
No backscattering at normal incidence (Klein Tunneling)

Can use to collimate electrons – fairly robust with roughness, gate split

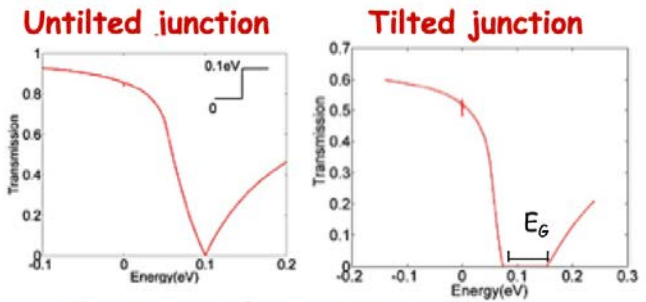


Cheianov/Falko/Altshuler (Graphene)

Switching with geometry: Klein Tunnel Transistor



nnn → regular graphene, large ON current
npn → collimation lobes misaligned → opens a gate dependent transmission gap

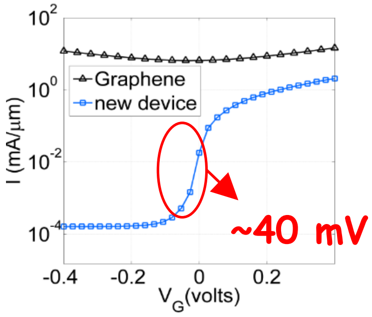


Gate Tunable Transport Gap

$E_G = V_0 \sin\delta / \cos^2\delta$

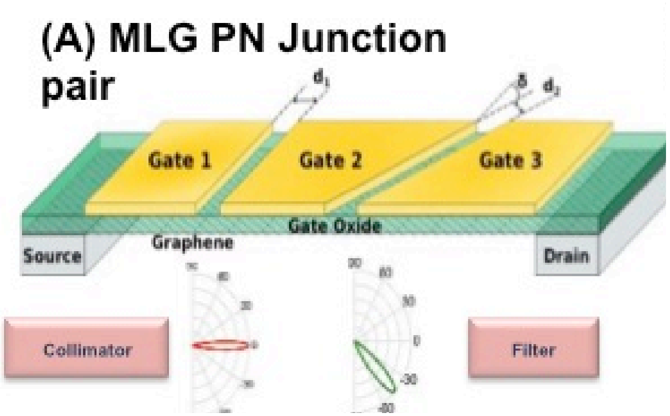
Sajjad, Ghosh APL '11;
SG, Cond-mat'13

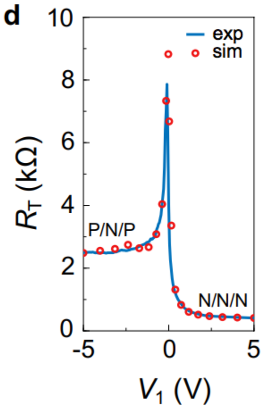
APL '11 PRB '12
ACS Nano '13 DRC '14
Patent # US9570559B2
PRL '15
Science '16
Sci Rep '17
PNAS '19 APL '19
ACS Nano Article ASAP '19
Physics '20



~40 mV ~ $2k_BT/3$

Experimental switching

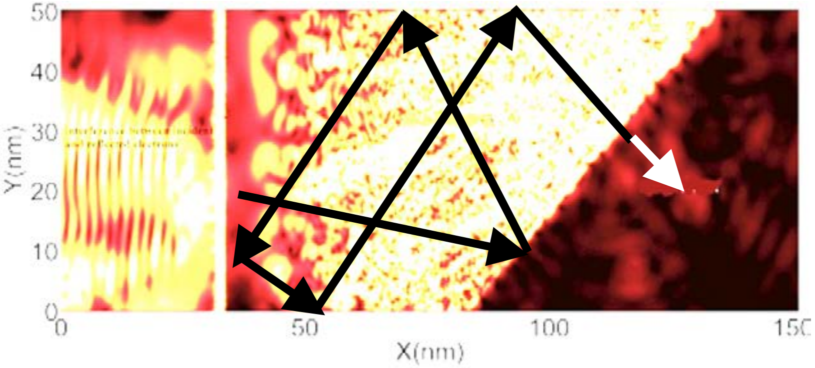




ON-OFF
~ 6-13

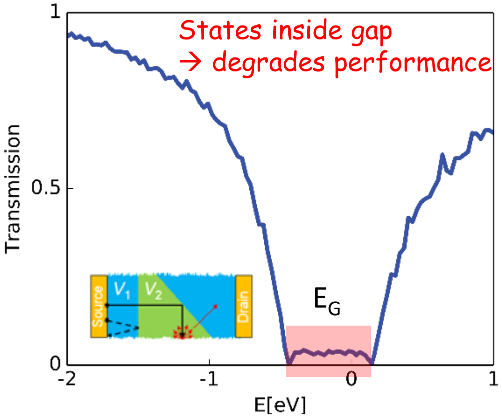
$\lambda_{sc} > 1\ \mu m$

PNAS '19



Edge reflections
bleed states
into gap

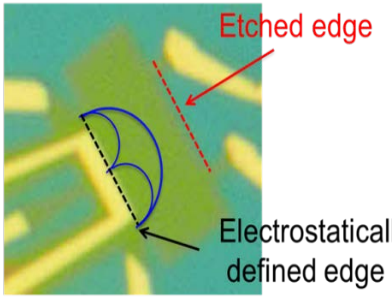
Need aggressive
collimation

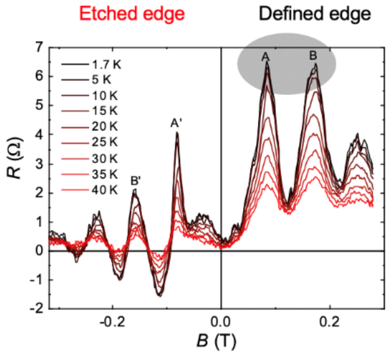


Causes saturation in the output characteristics
(Good for RF applications) Sci Rep 7, 9714 (2017)

A path towards higher performance KTFETs

Electrostatically defined edge for reduced edge scattering (2X)



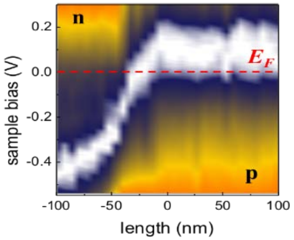


$S_{etched} = A'/B' \sim 0.5$

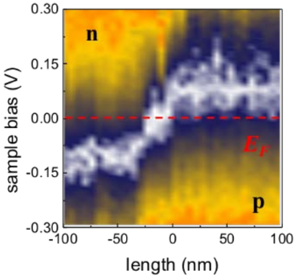
$S_{electrostatic} = A/B \sim 1.0$

Graphite gates for better Line-edge roughness (2X)

Buried Gate

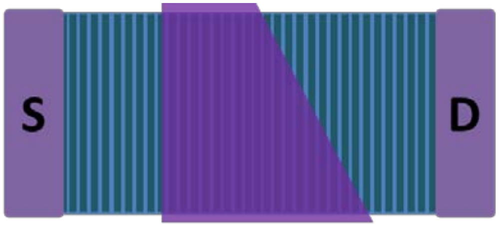


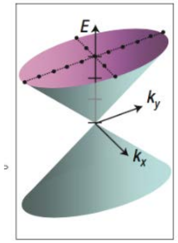
Graphite Gate



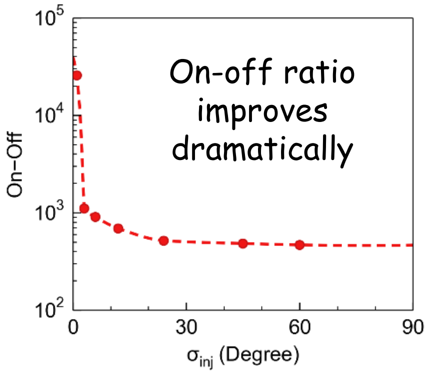
Kerelsky et al, ACS Nano 13, 2558 (2019)

Superlattice for better Collimation (Park et al Nat Physics '08) (1000X)



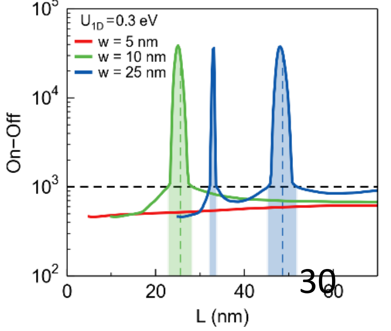


Collimator filter the injections with large angle



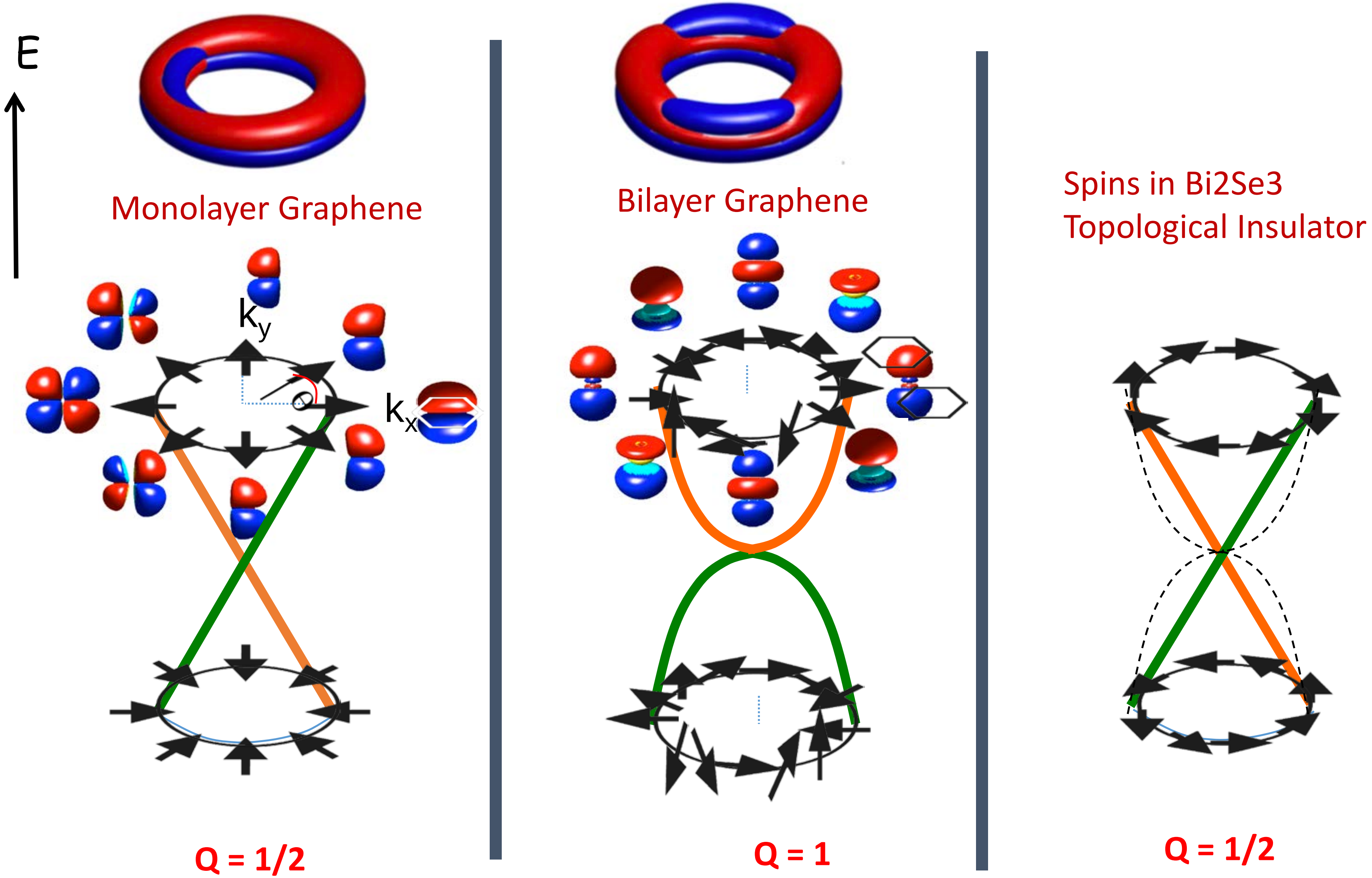
On-off ratio improves dramatically

Min^m 10^3 On-Off ratio



Switching is symmetry driven

Similar Dirac Fermions physics in other Q-Materials



Monolayer Graphene

Bilayer Graphene

Spins in Bi2Se3
Topological Insulator

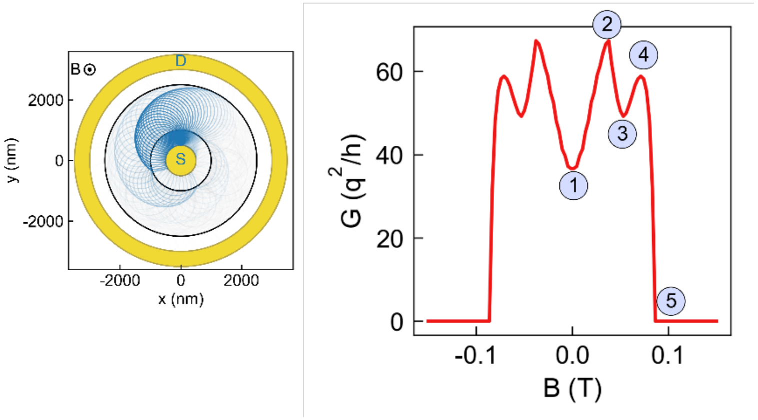
Q = 1/2

Q = 1

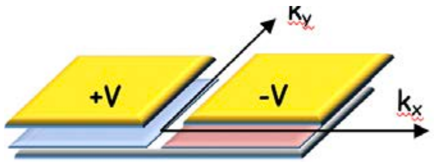
Q = 1/2

Q: Line Integral of
Berry Curvature ~ $\langle n(k) \mid \nabla_k \mid n(k) \rangle$

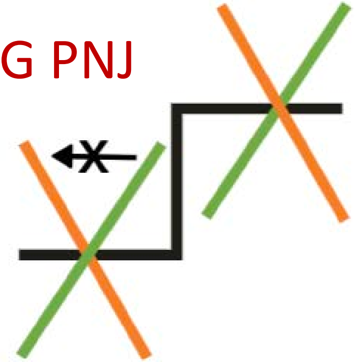
Topology → Universal Small Reflectivity

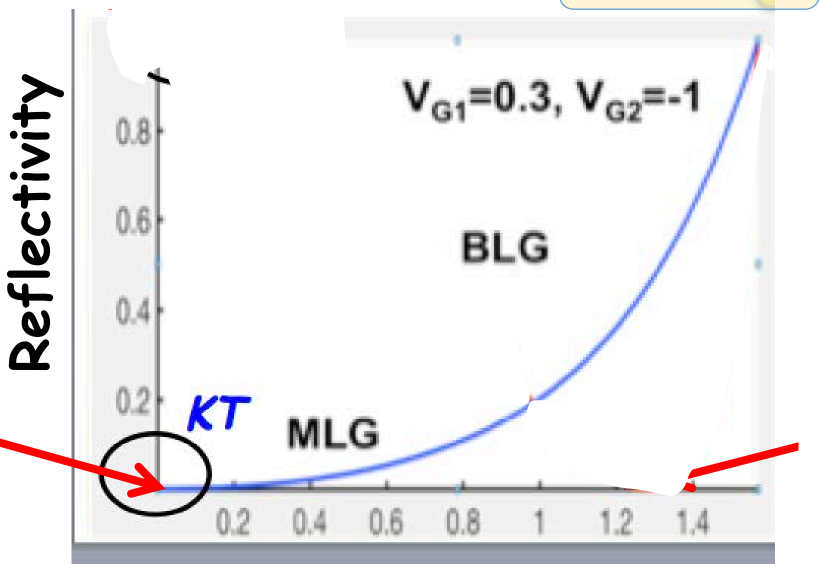


Proposed Expt: Corbino dips in MR data



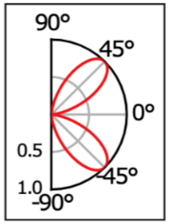
MLG PNJ



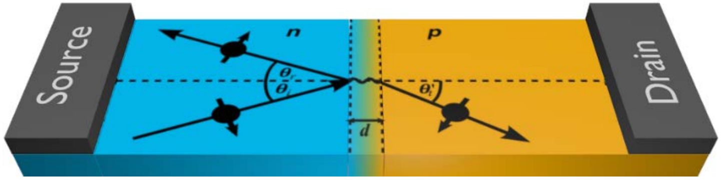


GRAPHENE

Electronic Brewster Angle, $\tan\theta = n_2/n_1$

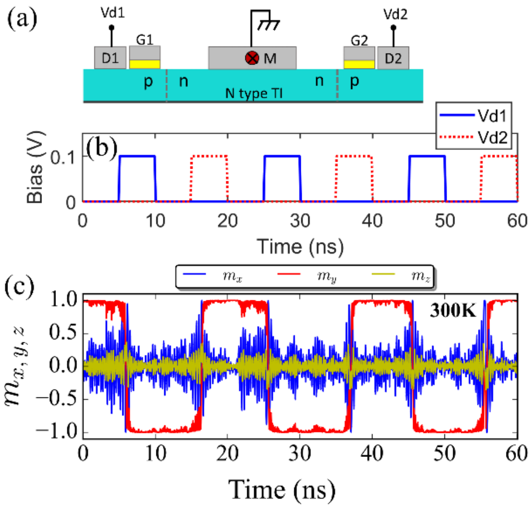


See this with spins in TIs/WSMs as well

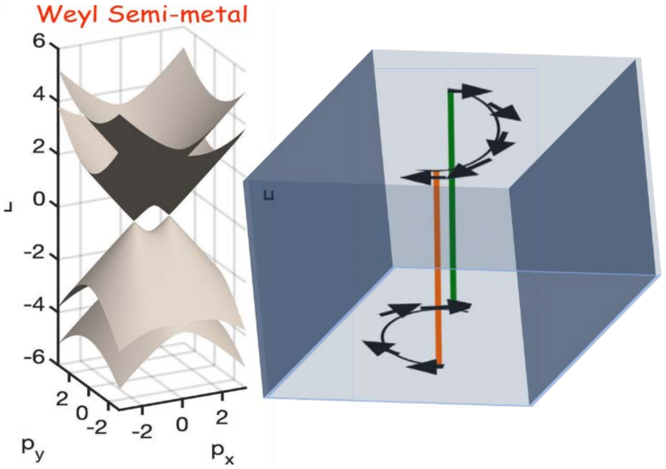


- Klein tunneling/collimation of spins creates a gate tunable charge reflection
- This flips spin because of spin-momentum locking
- This flips soft magnets

PRB 96, 205151 (2017)
PRL 114, 176801 (2015)



Proposed Expt: Bidirectional switching of in-plane FM with current reversal + gated PNJ



DSM/WSMs → spin-momentum locked surface states

Expts demo gating at ~10V (Lifschitz transition)

Split gated WSM structures → low bias gating

Physics ABOUT BROWSE PRESS COLLECTIONS

VIEWPOINT

Spin Control with a Topological Semimetal

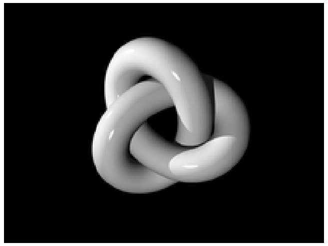
Avik Ghosh

Department of Electrical and Computer Engineering and Department of Physics, University of Virginia, Charlottesville, VA, USA

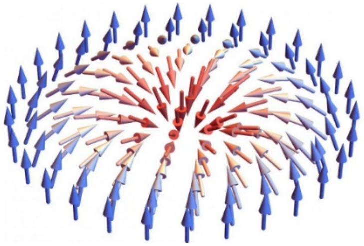
March 16, 2020 • *Physics* 13, 38

Physics 13, 38 (2020)

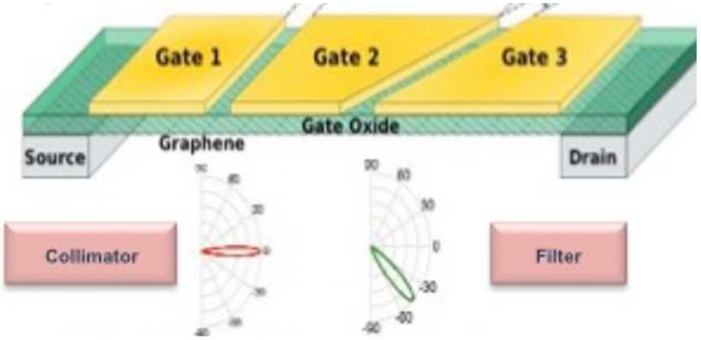
Topology has traditionally helped with materials classification and discovery



Topology can be useful & practical in a commercial sense



High density Racetrack Memory with Skyrmions



Klein Tunnel Transistors with Dirac Fermions

SPECIAL THANKS TO MY MENTORS !