Semiconductor Devices II

Part 2

Chapter 5: Beyond CMOS devices

Andras Kis

andras.kis@epfl.ch

EPFL – École Polytechnique Fédérale de Lausanne

Electrical Engineering Institute

[Outline] Chapter 5: Beyond CMOS devices

Fundamental limits to power dissipation

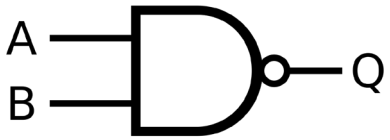
- Landauer limit
- Von Neumann architecture
- Joule heating

Emerging concepts based on 2D materials

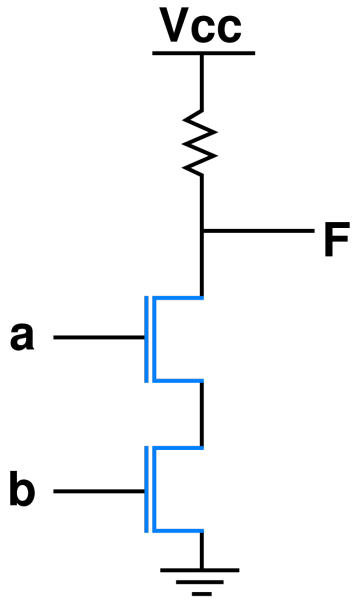
- Memory in logic
- Excitonic devices

Information processing, energy and reversibility

Consider the NAND logic gate and its truth table



A	B	A NAND B
0	0	1
0	1	1
1	0	1
1	1	0



NMOS
NAND gate

Logically irreversible operation: 3 different input configurations give the same output ("1")

If we "destroy" the inputs and only keep the result, we have increased the entropy

Limits on computing: The Landauer limit

Erasure of N bits of information from a physical system results in:

- An increase in entropy of

  $$\Delta S \geq k_B \ln(2) \cdot N$$

- Energy dissipation of

  $$\Delta E \geq k_B T \ln(2) \cdot N$$

- For 1 bit:

  $$\Delta E \geq k_B T \ln(2) = 2.86 \cdot 10^{-21} J$$

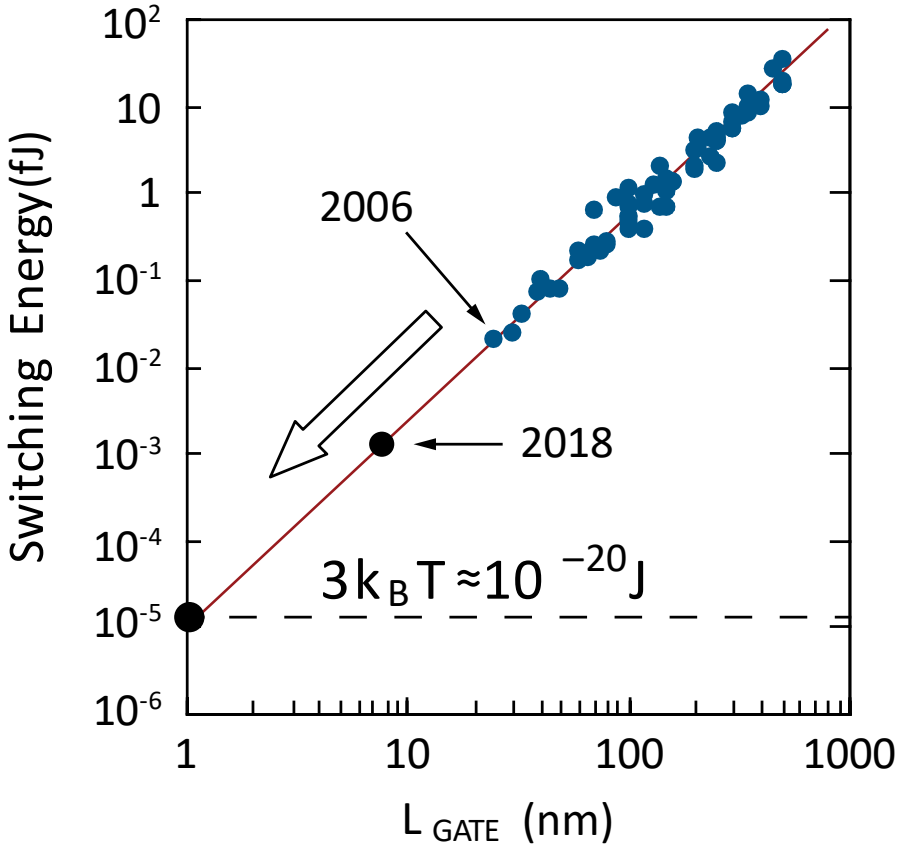
Rolf Landauer, IBM J. Res. & Dev. (1961)

k_B Boltzmann constant
T temperature

Limits on switching

For a charge-based switch

- Minimum switching energy is: $\Delta E_{sw_{min}} = 3k_B \mathrm{T} \cdot \ln(2) \approx 10^{-20}\, J$
- $\approx$ 3× worse than Landauer



Cavin et al. J Nanopart Res (2006)

k_B Boltzmann constant
T temperature

How about the entire system?

Intel Core i7 8700K processor (2017)

- Benchmark: 217 GOPS
- Dissipated power: 86.2 W

$$\frac{Dissipated\ power}{N\ operations} = \frac{86.2}{217 \cdot 10^9} = 3.97 \cdot 10^{-10}\,\frac{J}{flops}$$

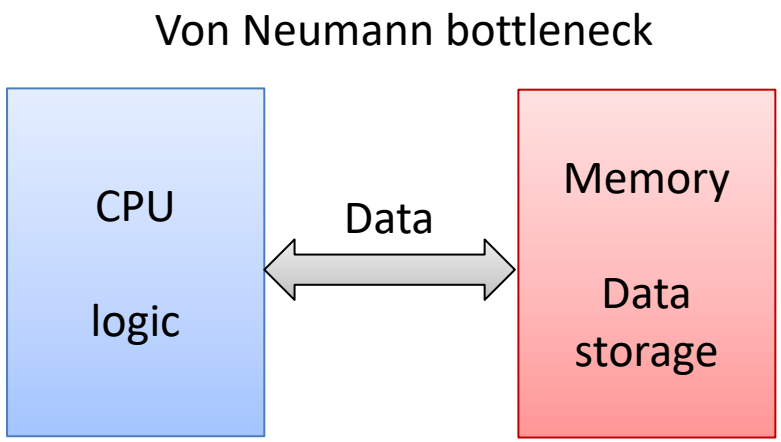
$$= 6.2 \cdot 10^{-12}\,\frac{J}{ops}\ (64\ \text{bits})$$

- Landauer limit:

$$2.86 \cdot 10^{-21}\ J/ops$$

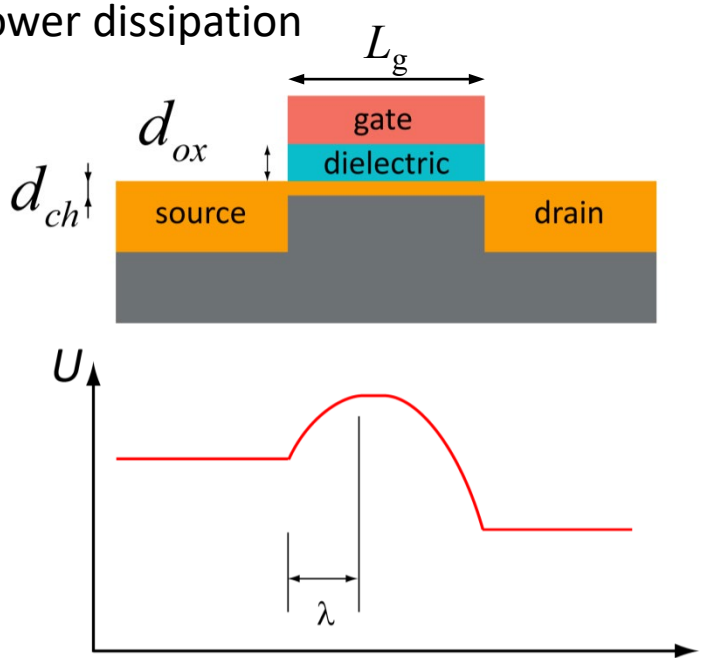
We are $10^9\times$ worse than the Landauer limit!

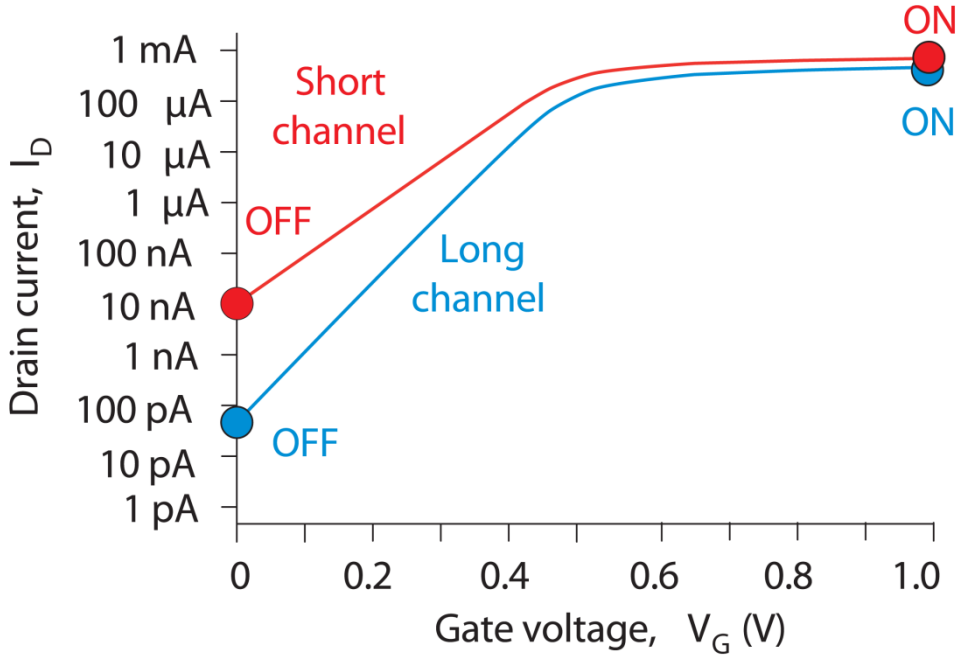
Von Neumann architecture



Ch 2.1: Heat Dissipation in Small Transistors

Planar MOSFET – short channel effects; standby power dissipation





$$\lambda = \sqrt{\frac{\varepsilon_{ch}}{\varepsilon_{ox}} d_{ox} d_{ch}}$$

ε(Si)=11.9

To deplete the channel: L_g at least 3–5× λ

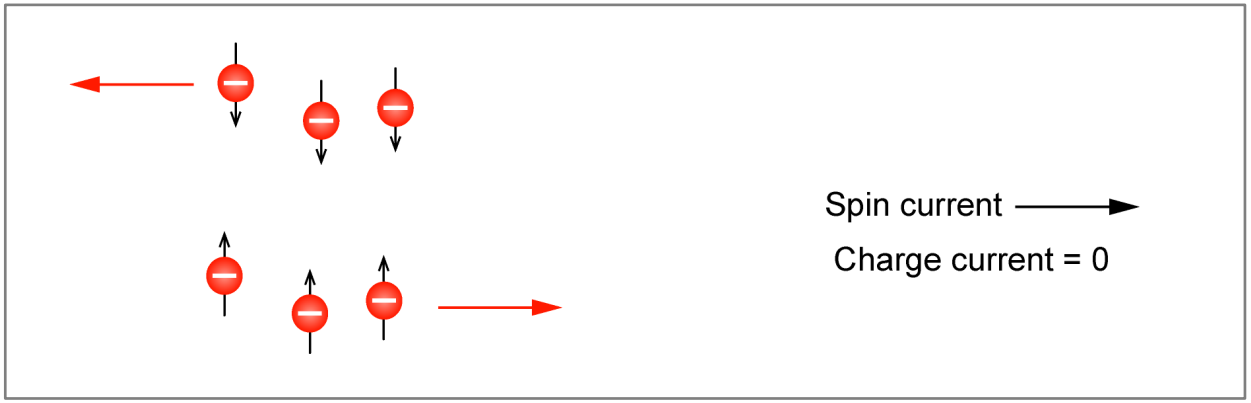
Example:

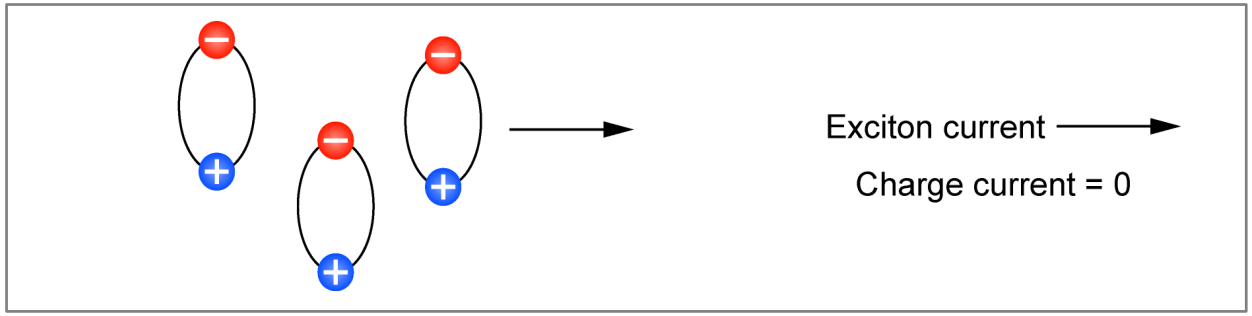
2nm thin Si, 1nm SiO_2: L_g>10nm

Ferain et al., Nature 479, 310 (2011)

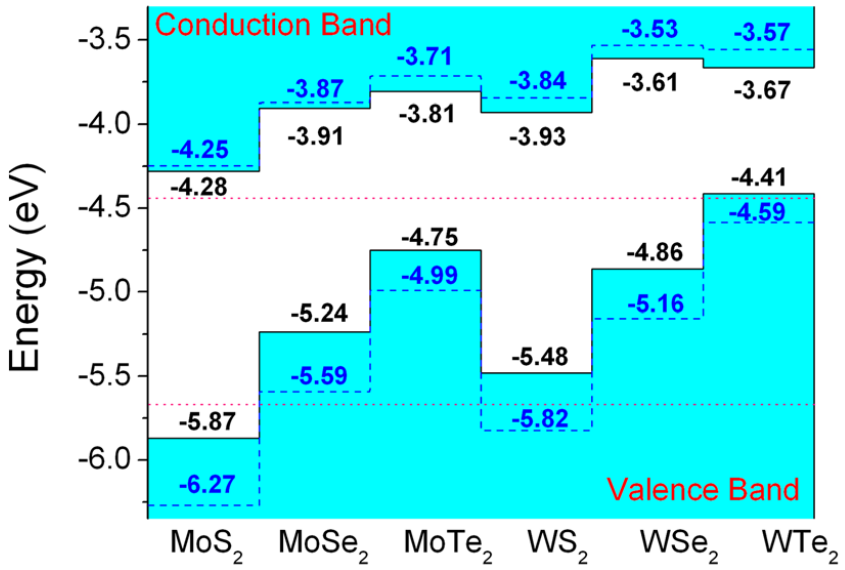
Colinge, Sol. State El. 48, 897 (2004)

Other Types of Current

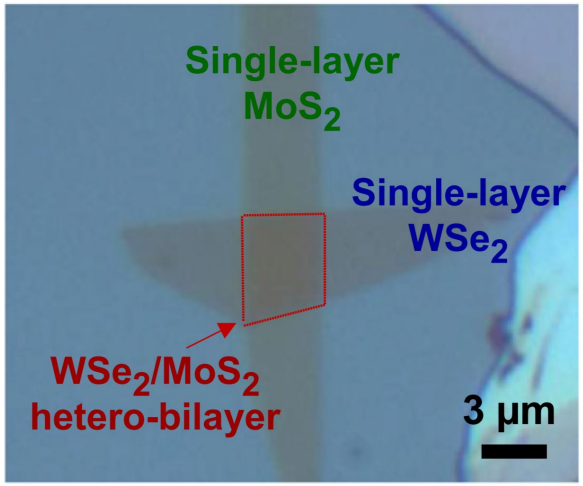




Interlayer Excitons in TMDCs

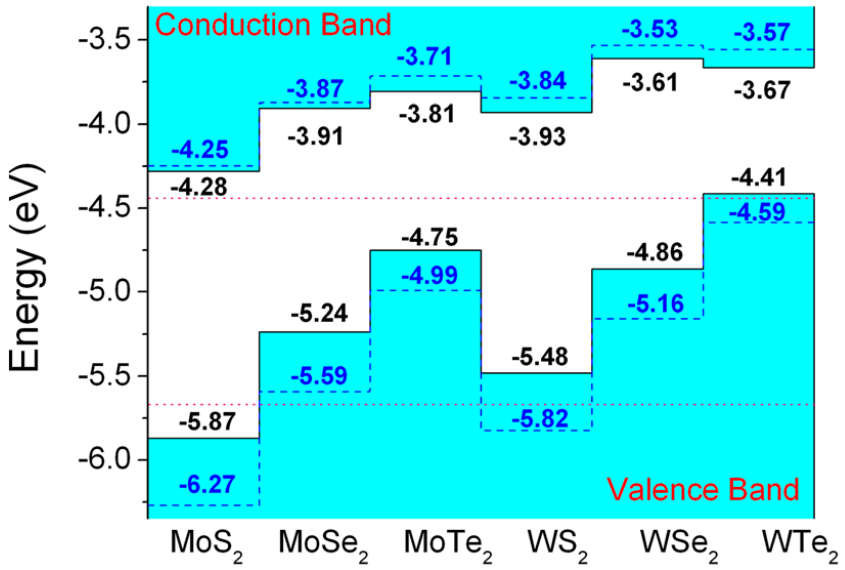


Kang, Jun, et al, APL (2013)



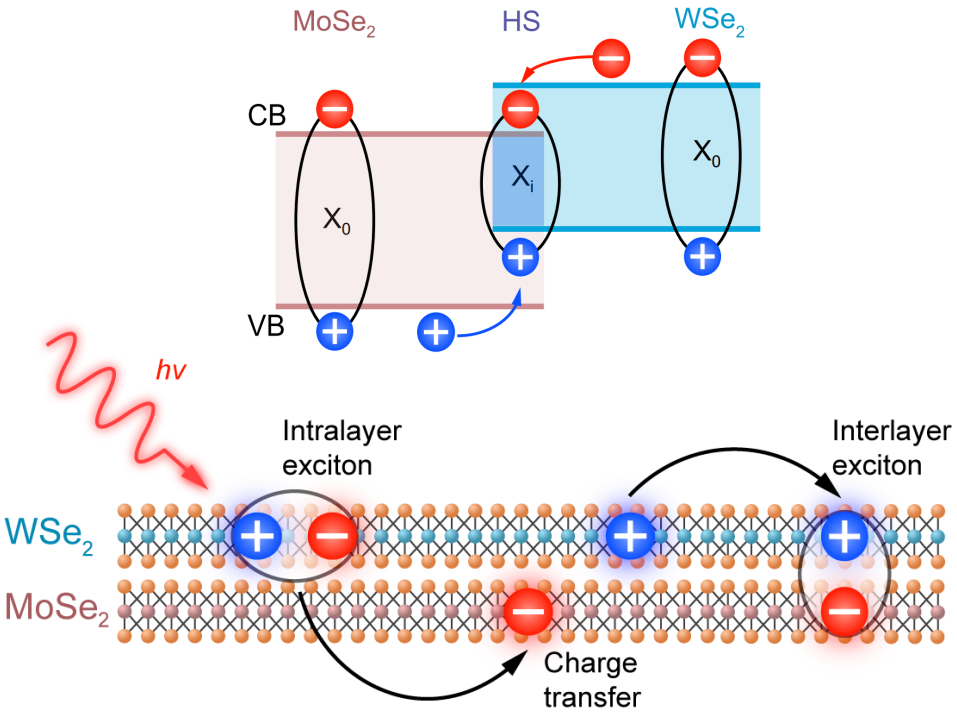
Fang…Javey; PNAS (2014)

Interlayer Excitons in TMDCs

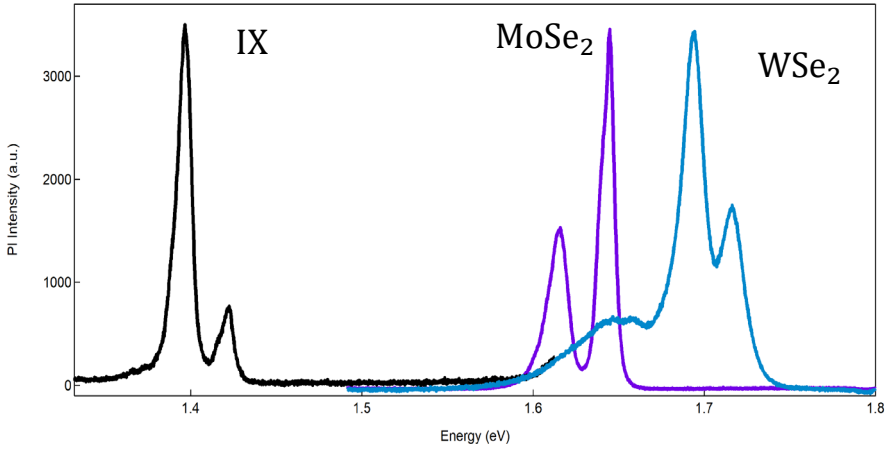


Kang, Jun, et al, APL (2013)

Band Diagram



Photoluminesce spectra



Spatial separation of charges:

- Built-in dipole moment -> manipulation
- Long lifetime ~ ns
- Large binding energy ~ 100s of meV

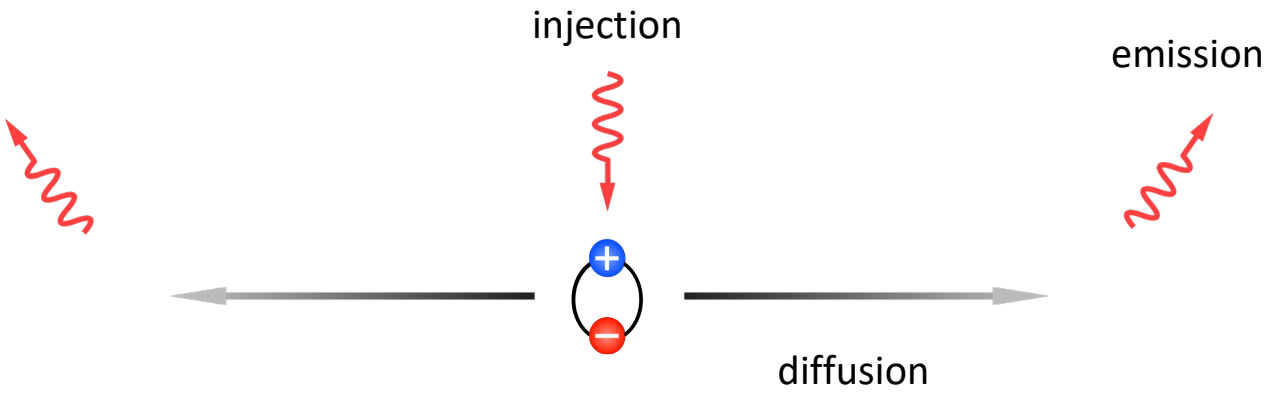
Related work:

Fang et al, PNAS (2014)
P. Rivera, et al. *N. Com* (2015)
P. Rivera, et al. *Science*(2016)
Hanbicki, et al. *ACS Nano* (2016)

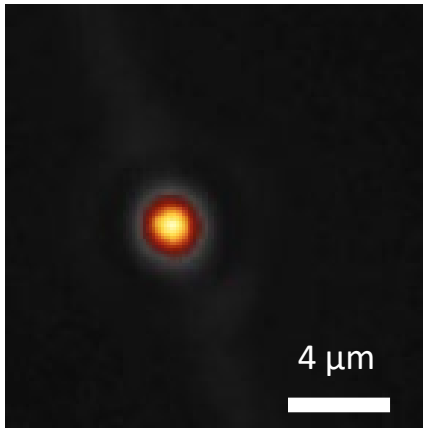
Exciton Manipulation in a Device

Exciton diffusion

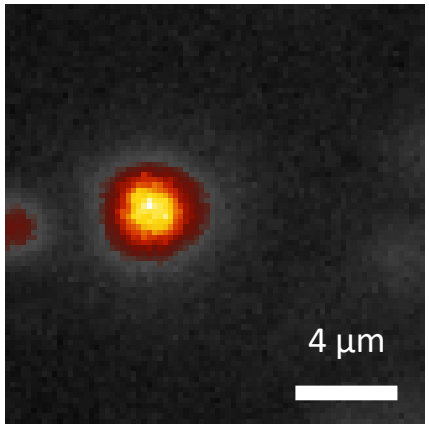
Driving forces: $\nabla T, \nabla \mu$



Excitation spot



IX emission

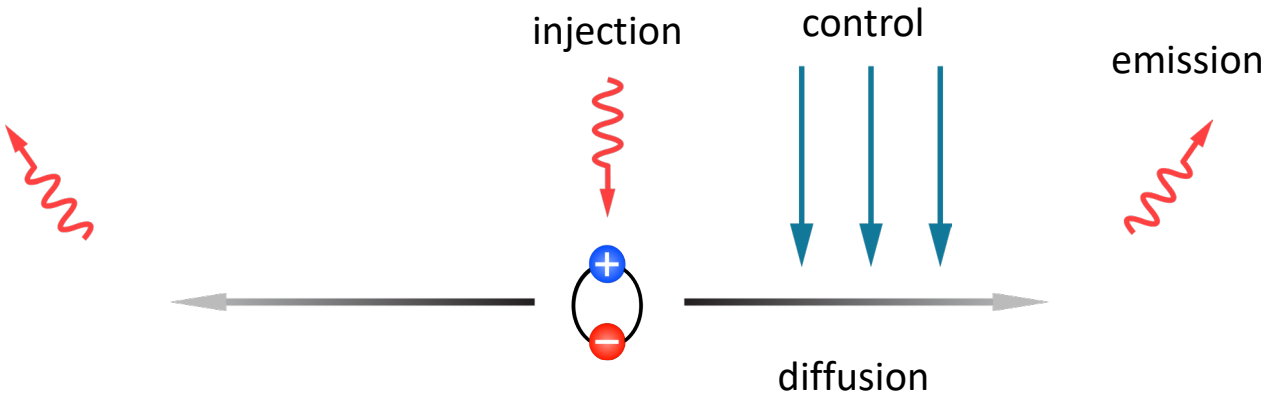


7 – 740 µW

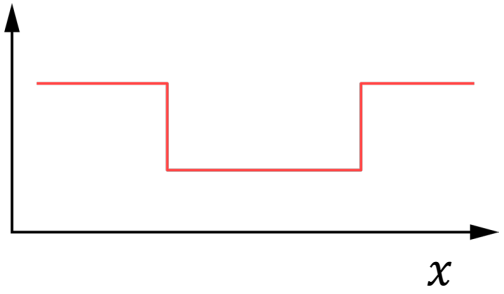
Exciton Manipulation in a Device

Exciton diffusion

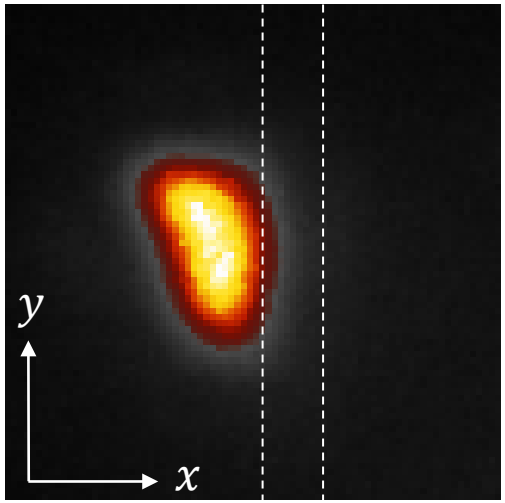
Driving forces: $\nabla T, \nabla \mu$

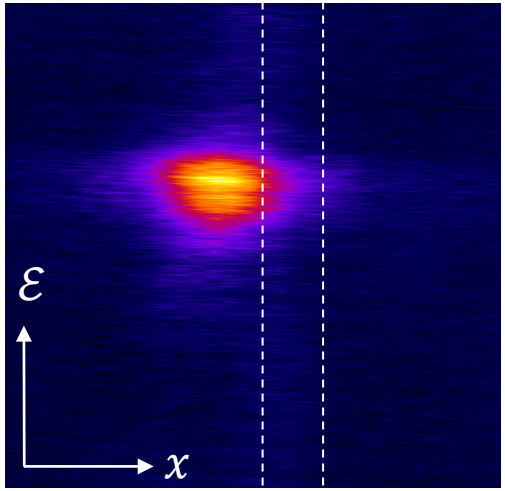


$$\varphi_{el} = \delta E = -p_z E_z(x, y)$$



IX emission





Simplest device:

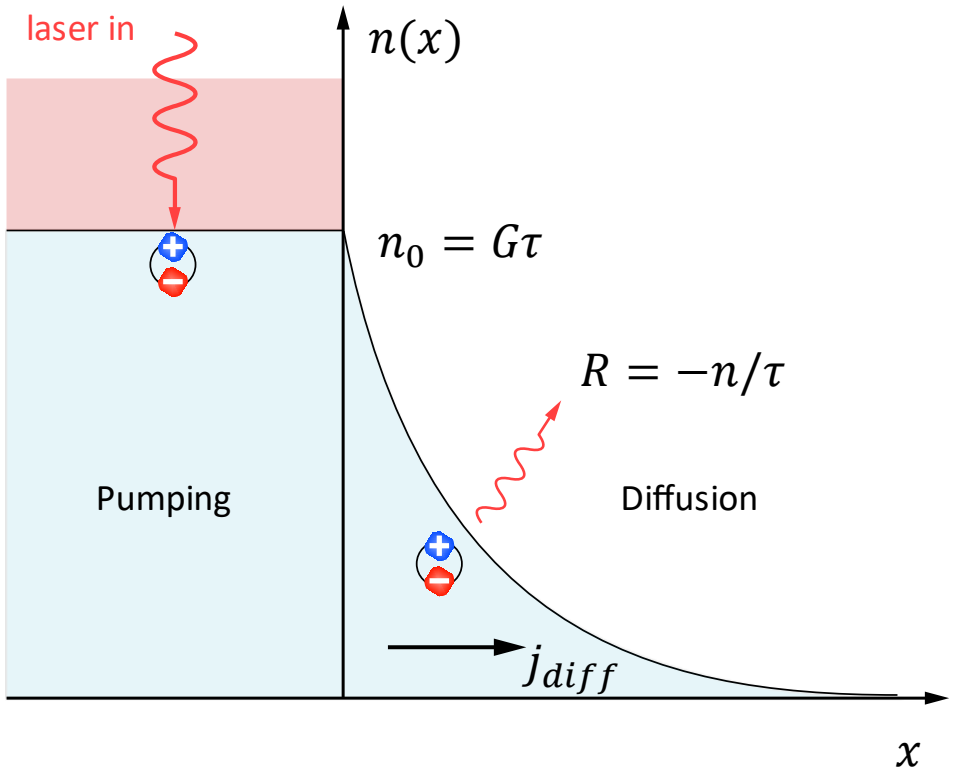
Excitonic transistor/switch

Exciton manipulation in a device

Exciton diffusion governed by diffusion equation with an external potential:

$$D\frac{\partial^2 n}{\partial x^2} - \frac{D}{k_B T}\frac{\partial}{\partial x}\left(\frac{\partial \varphi}{\partial x} n\right) + G - \frac{n}{\tau} = \frac{\partial n}{\partial t}$$

$$\varphi_{el} = \delta E = -p_z E_z(x, y)$$

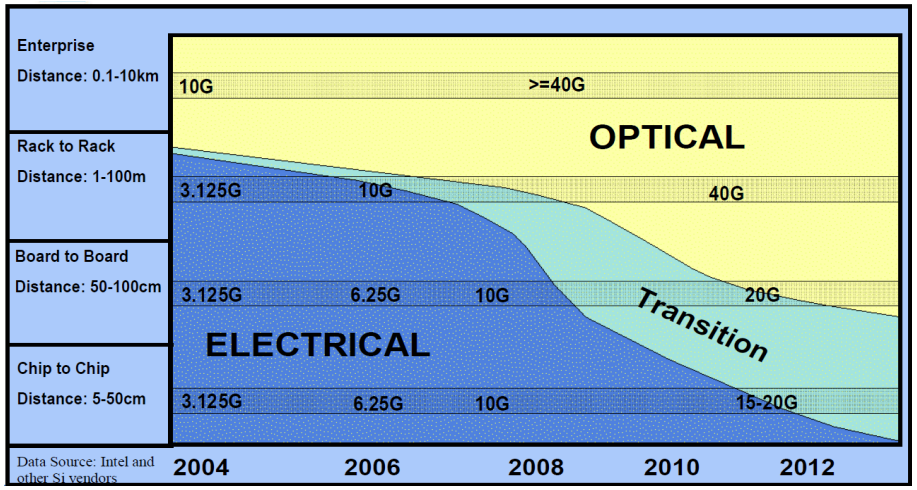


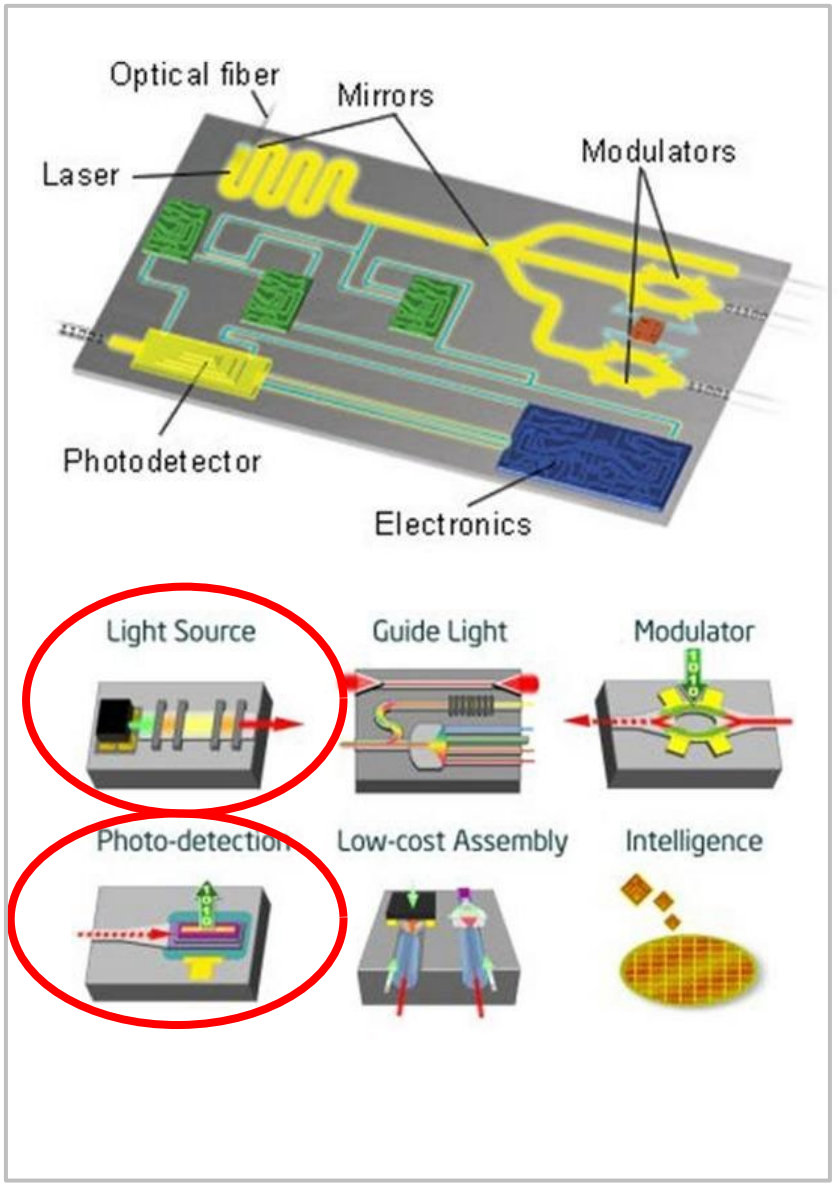
j_{diff}	current
n	exciton concentration
D	diffusion coefficient
τ	exciton lifetime
G	generation rate
R	recombination rate
p	dipole moment
φ	exciton potential
k_B	Boltzmann constant

Unuchek, Ciarrocchi et al. Nature (2018)

Ch 2.4: Optoelectronics and IT devices

Data load

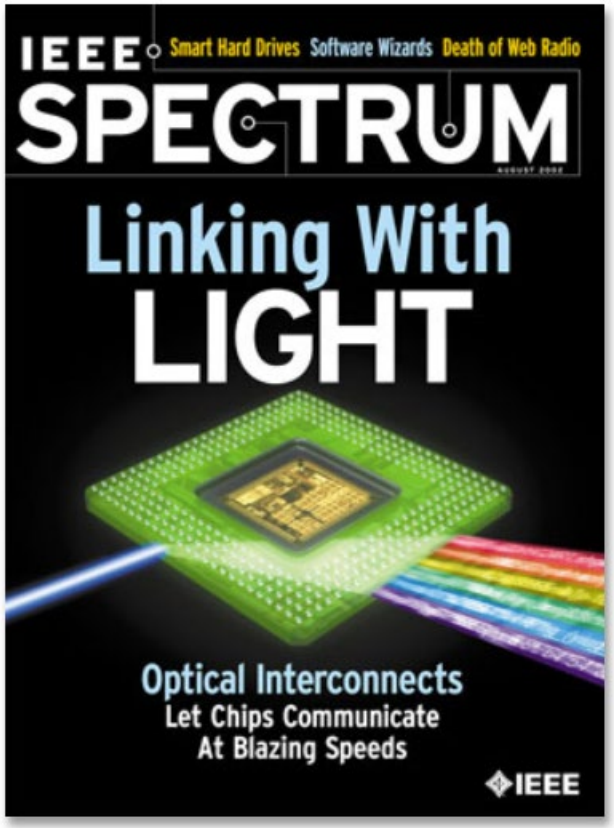




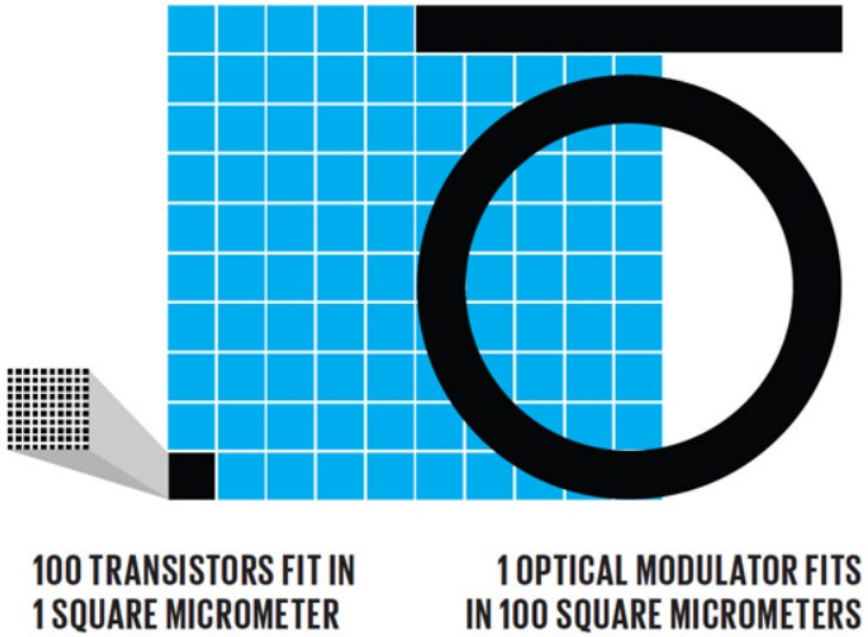
Flexible devices

From the TV show Westworld

Excitonic Devices



Savage; IEEE Spectrum (2002)

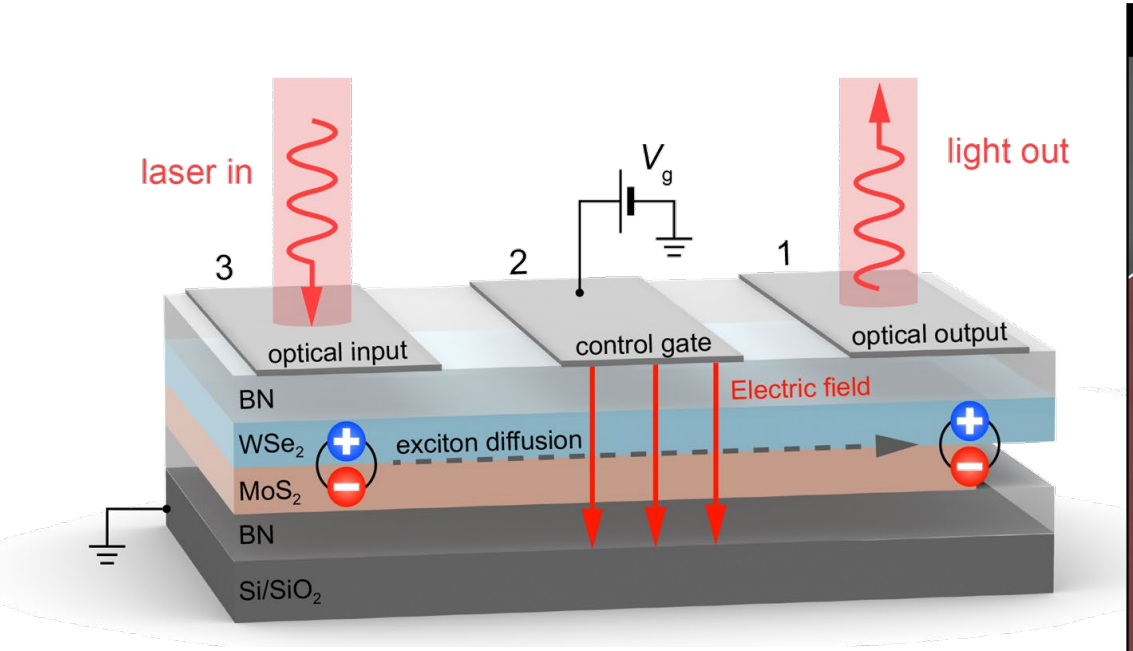


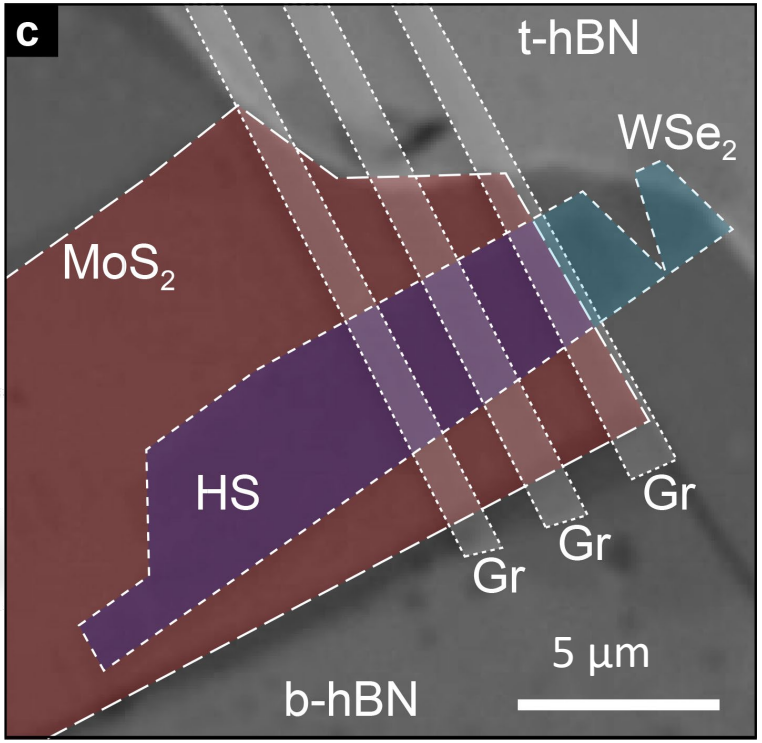
Levi; IEEE Spectrum (2018)

$$device\ size \gg particle\ size$$

$$photons: \lambda$$
$$excitons: Bohr\ radius$$

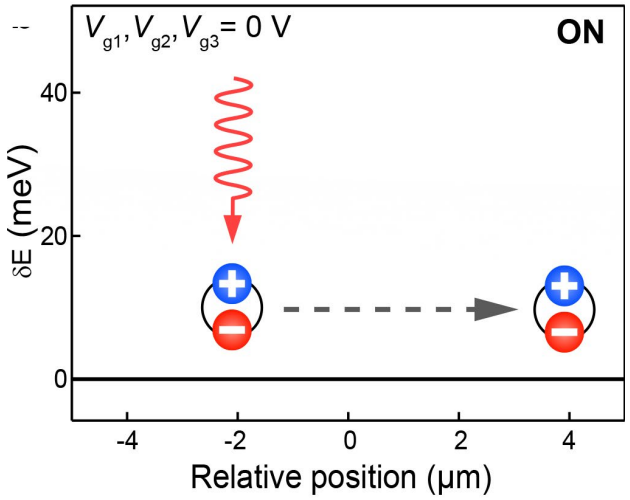
Room-temperature Excitonic Devices

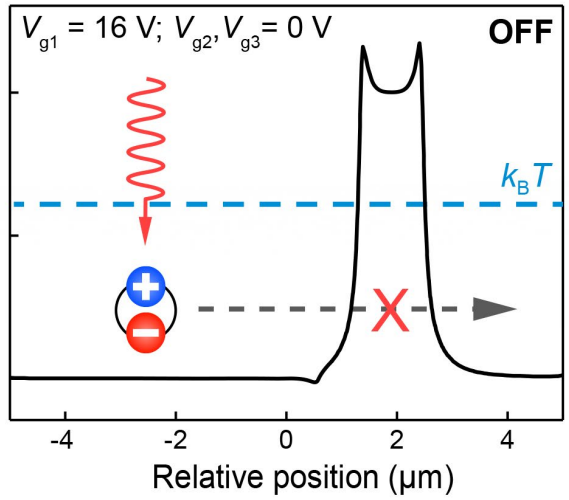


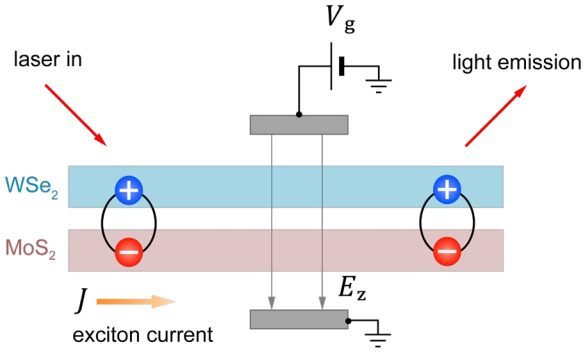


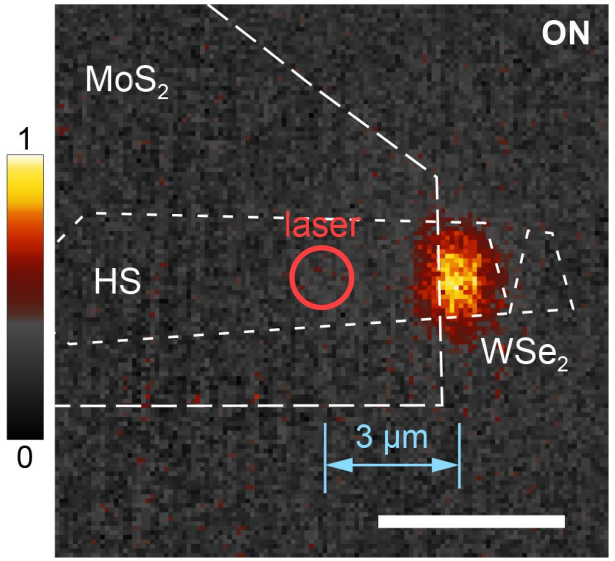
Unuchek, Ciarrocchi,…,Kis; Nature (2018)

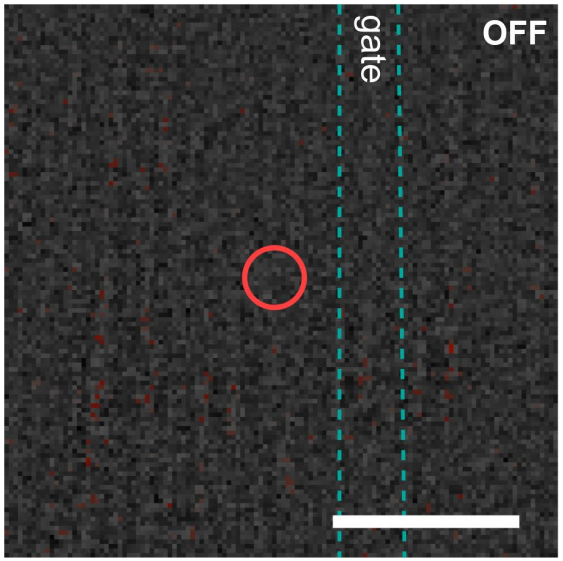
Gate Control: Switching





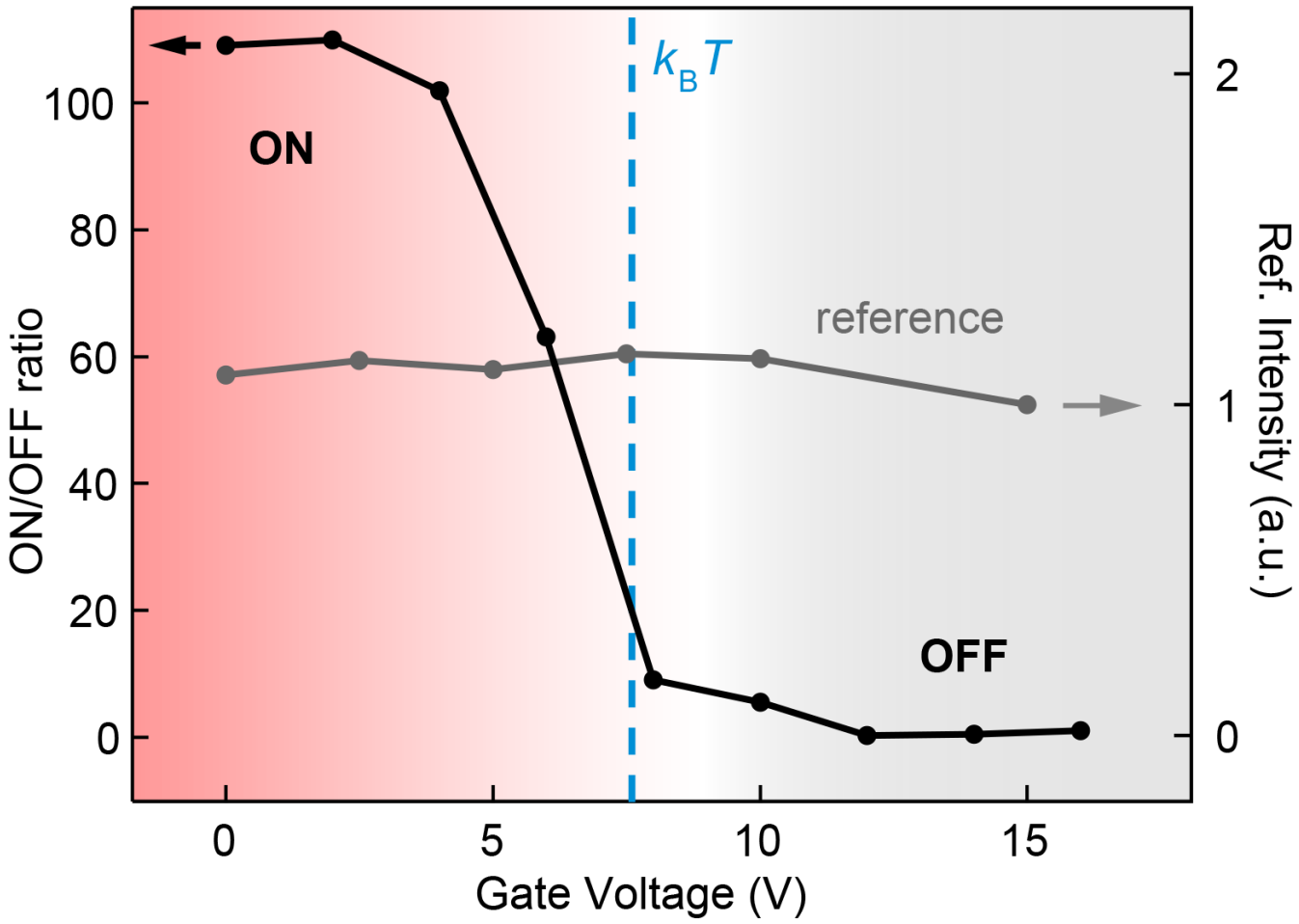




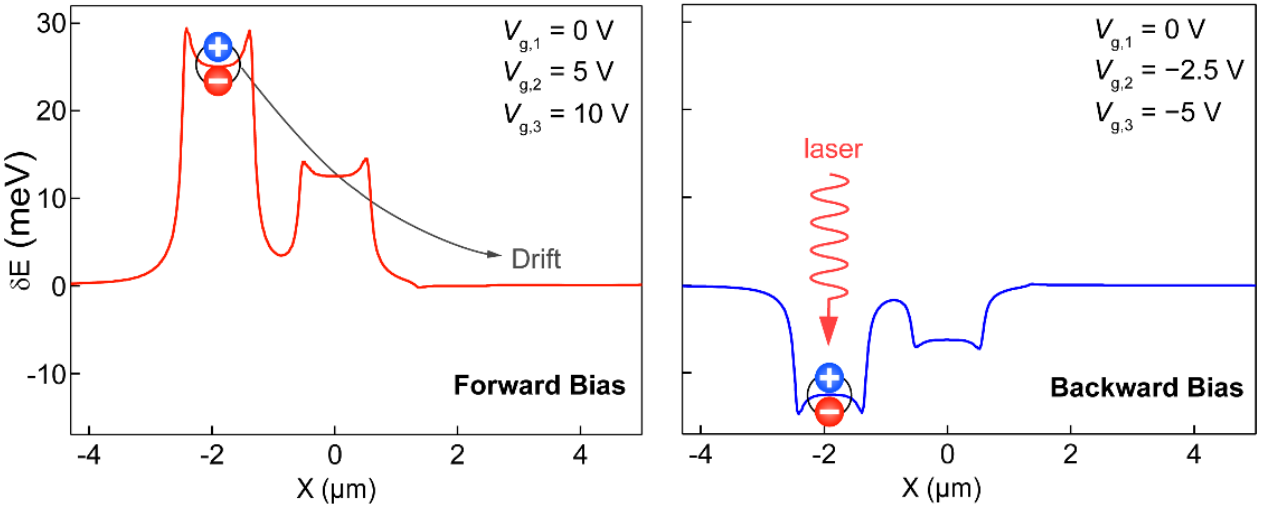


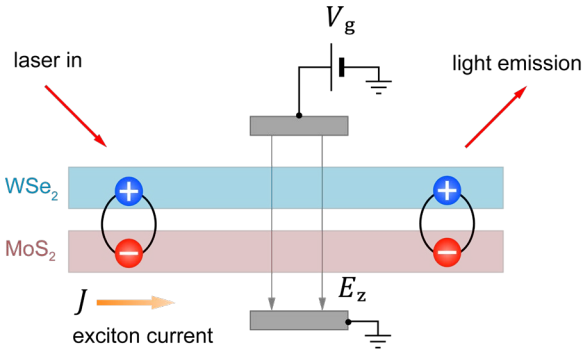
Unuchek, Ciarrocchi,…,Kis; Nature (2018)

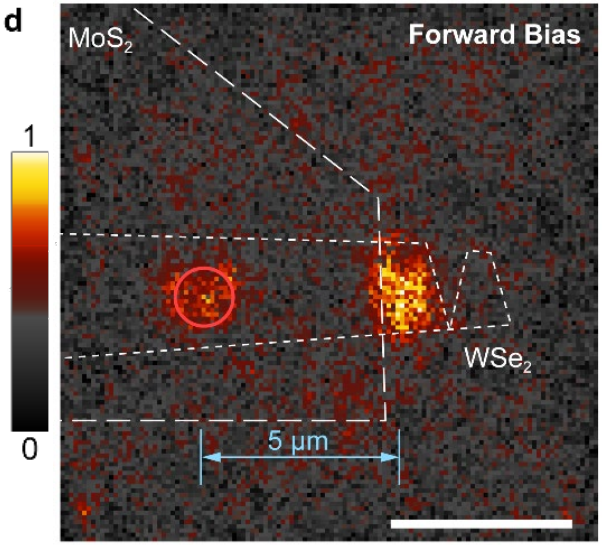
Gate Control: Switching

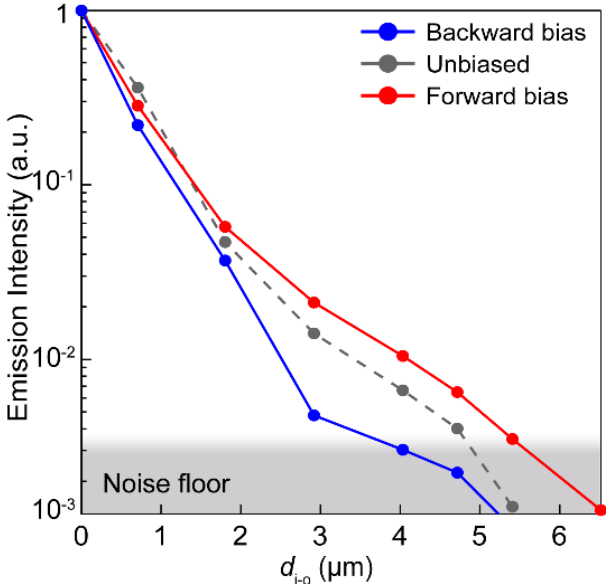


Gate Control: Enhanced Diffusion

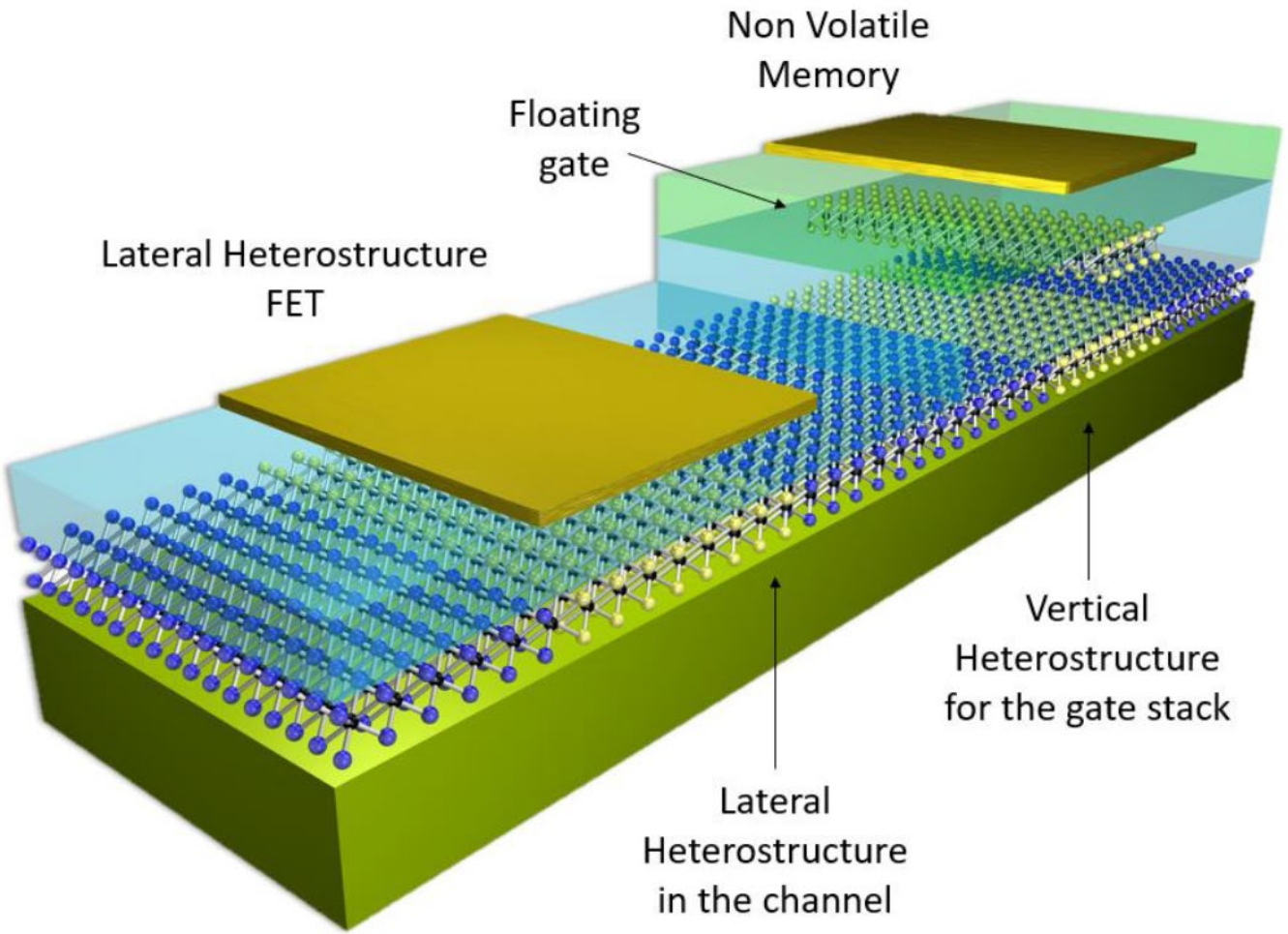








Logic in Memory Concept

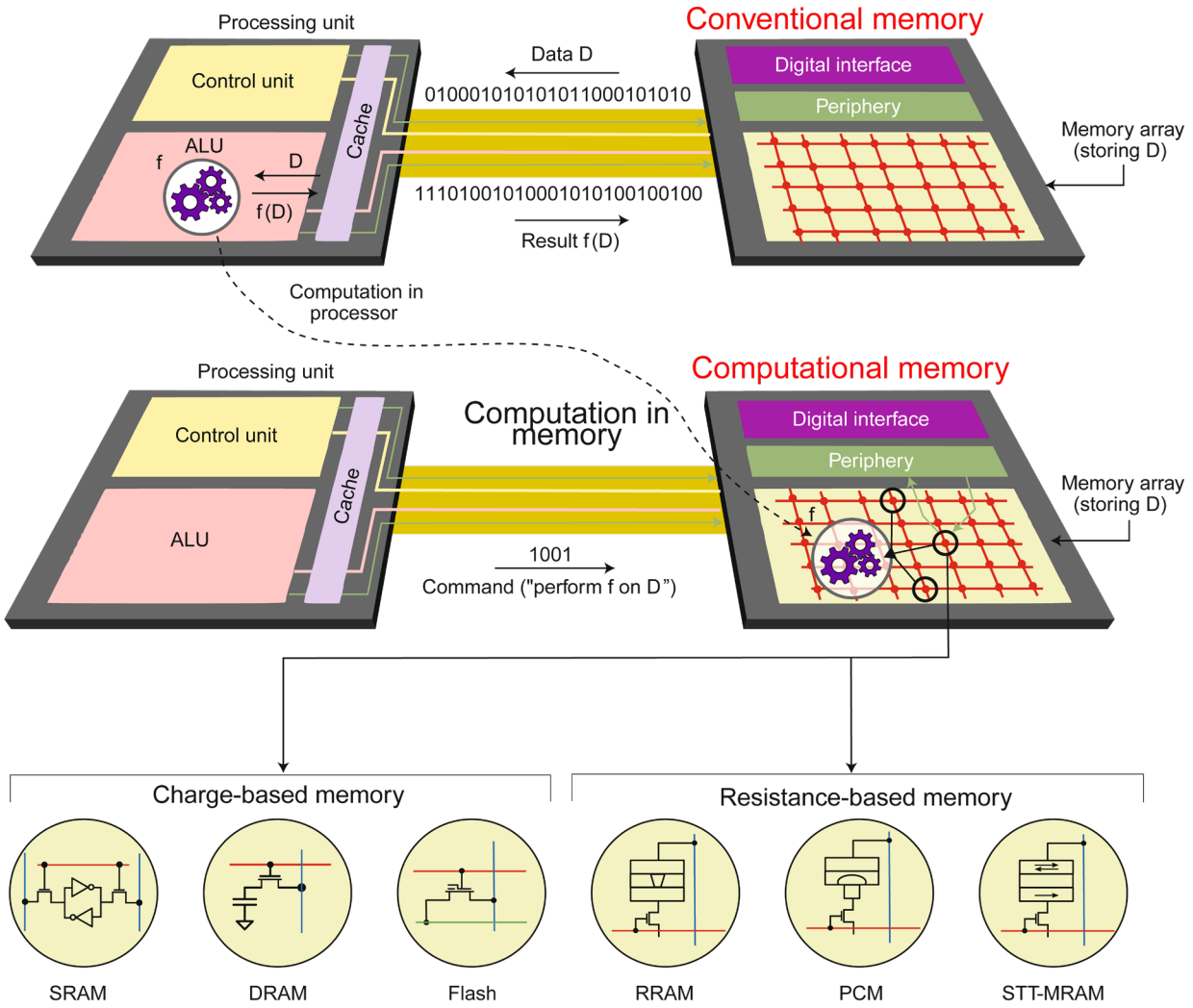


Credit: Giuseppe Iannaccone, University of Pisa

Logic in Memory Concept

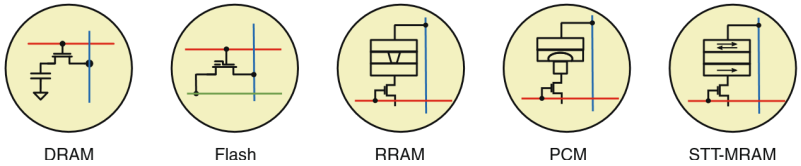
Von Neumann

Logic in memory



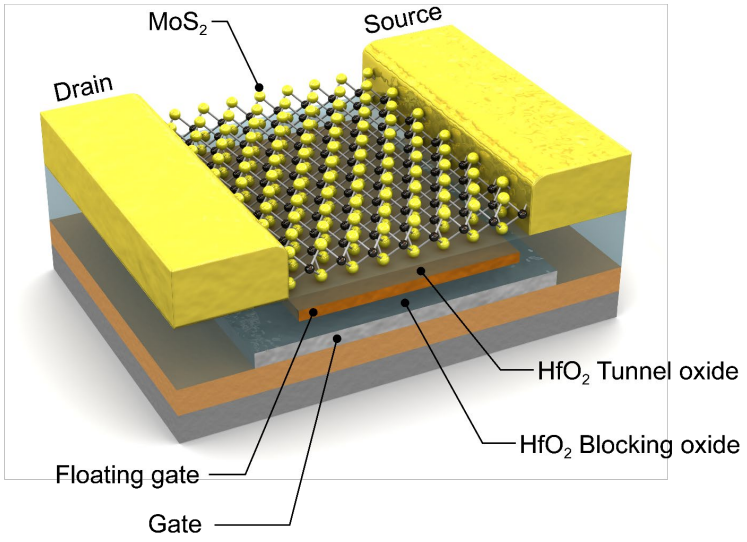
Sebastian…Eleftheriou, Nat. Nanotech. (2020)

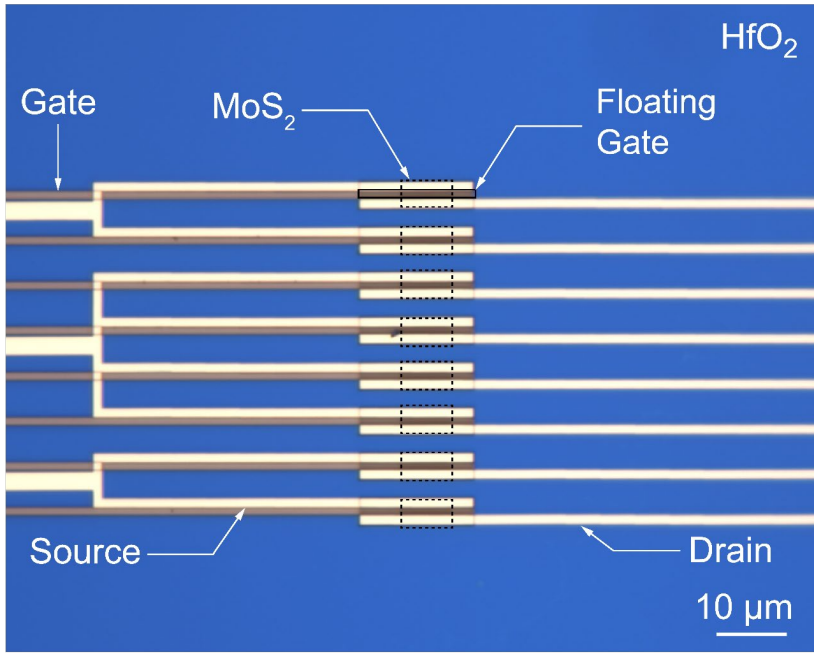
Logic in Memory

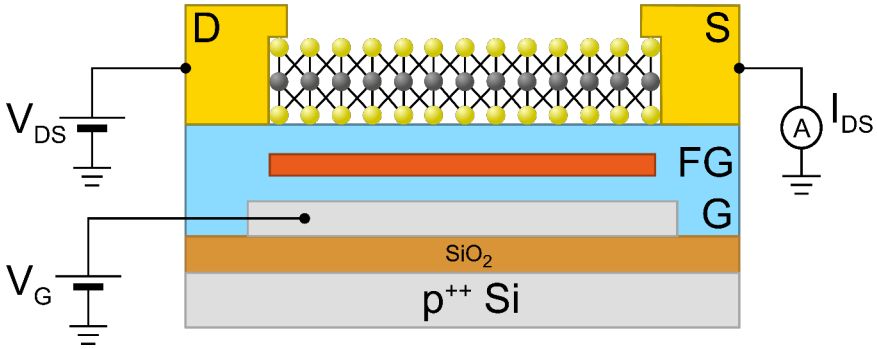


Performance metrics	DRAM	Flash	PCM	STT-MRAM	RRAM	HDD
Feature size (nm)	36	22	45	95	9	NA
Cell Area	$6F^2$	$4F^2$	$4F^2$	$4F^2$	$4F^2$	~256*
Write/Erase Time	< 10ns	1/0.1ms	100ns	<10ms	<1ns	5ms
Retention	64ms	>10y	>10y	>10y	>10y	>10y
Endurance	>1E16	1E4	1E9	>1E12	1E12	>1E16
Nonvolatility	N	Y	Y	Y	Y	Y
Multi-level capability	N	Y	Y	N	Y	-
Write Energy (J/bit)	4E-15	> 2E-16	6E-16	2.5E-12	1E-13	-
Standby Power (W/Gb)	1E-1	1E-3	1E-3	1E-3	1E-3	110

2D Logic with Memory



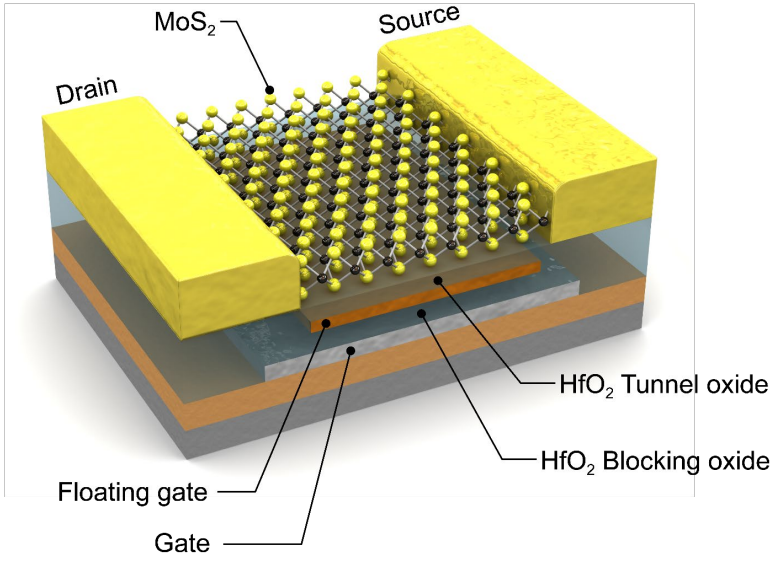


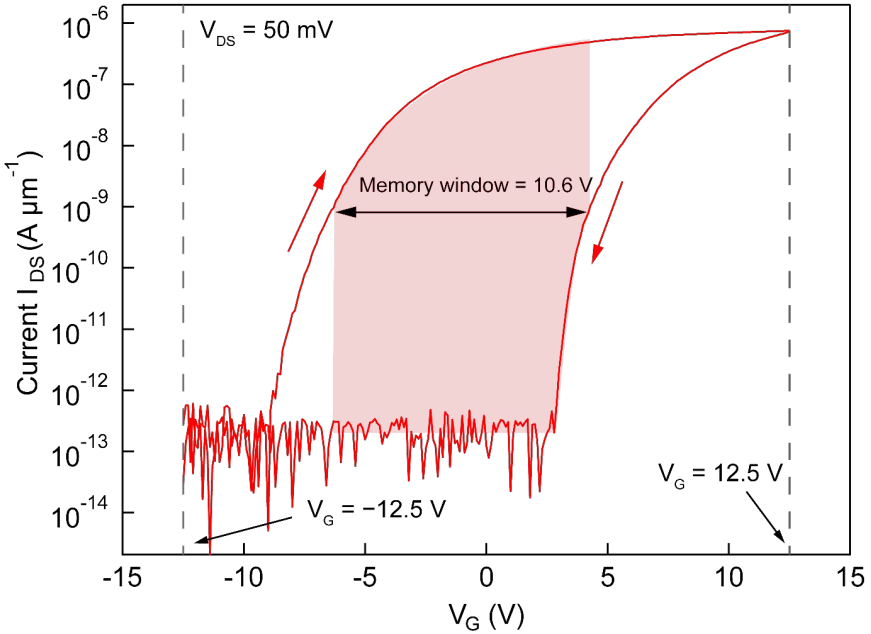


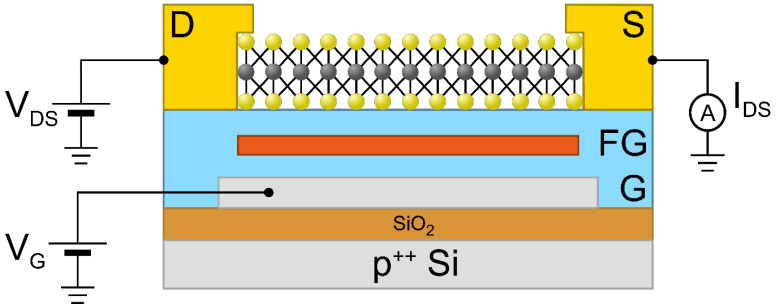
Advantages of 2D for logic in memory and flash memory:

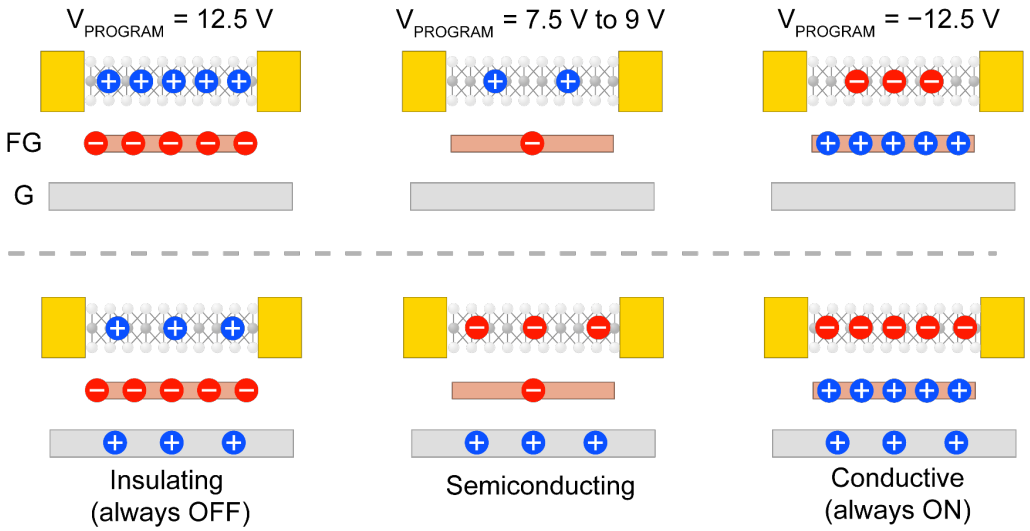
- Reduced cell-to-cell interference
- Multilevel storage
- Possible scaling beyond 12 nm

2D Logic with Memory

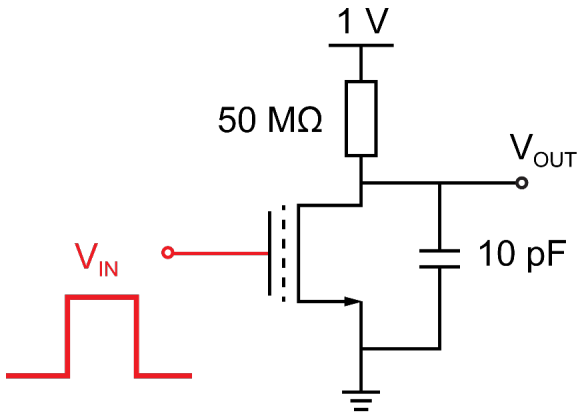








Programmable Inverter

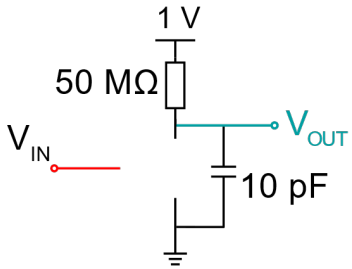


Q	IN		$X^{(Q)}$	
1	0	1	0	0
2	0	1	0	1
3	0	1	1	1

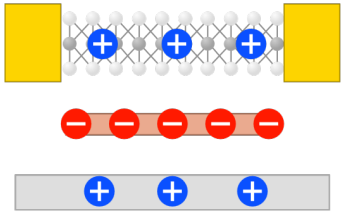
Logic input and memory states

$X^{(Q)}$	OUT
0	1
1	0

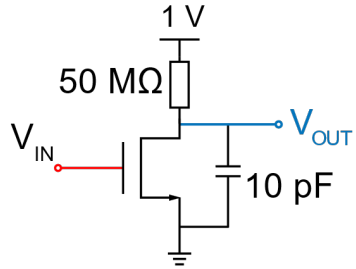
Memory states and logic output



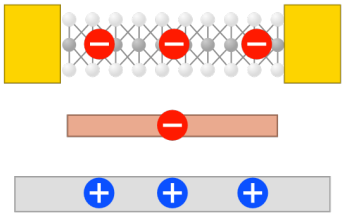
Q = 1 : memory state 0



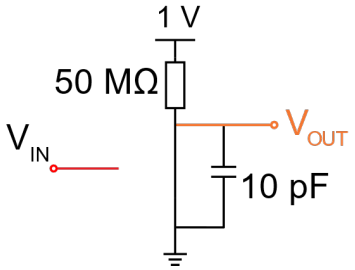
Insulating
(always OFF)



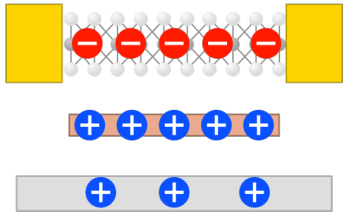
Q = 2 : memory state 0 OR 1



Semiconducting

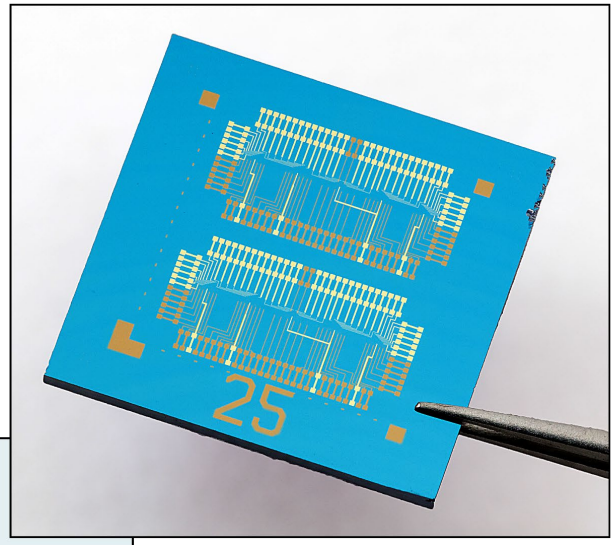


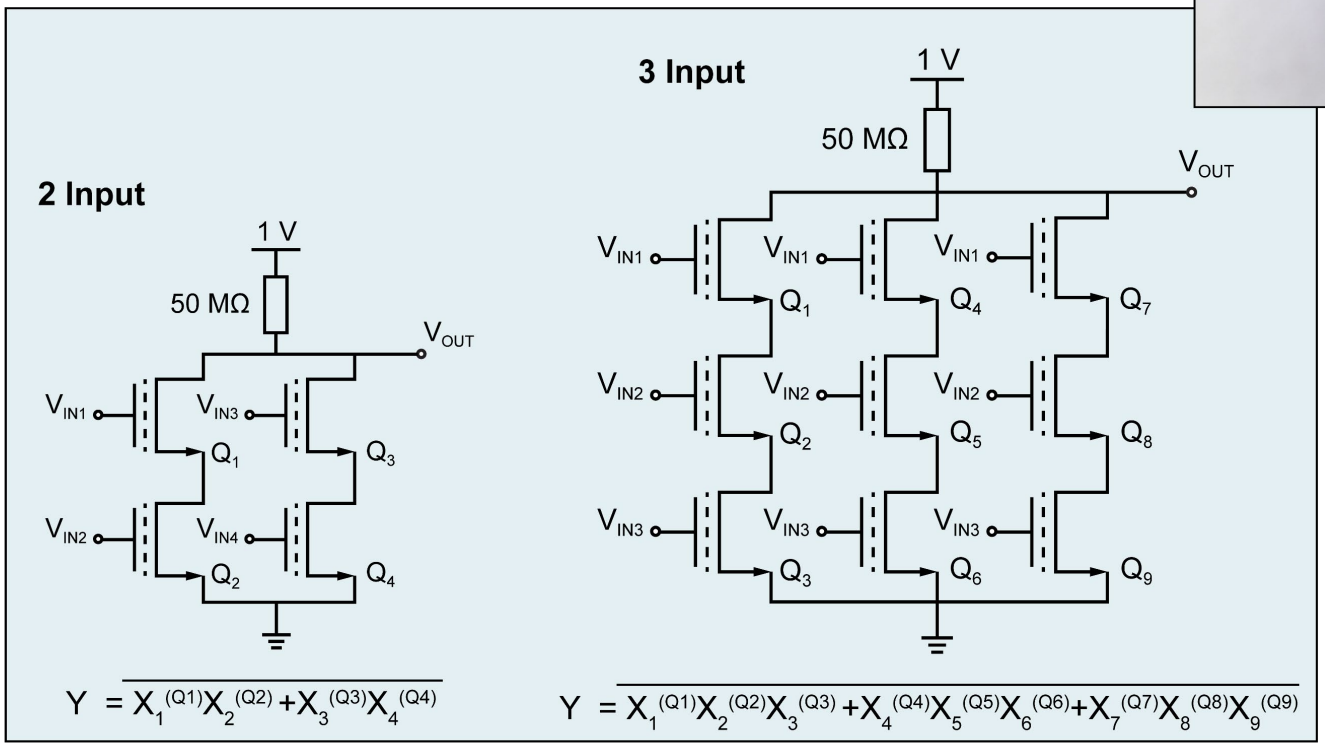
Q = 3 : memory state 1



Conductive
(always ON)

Universal Logic Gates

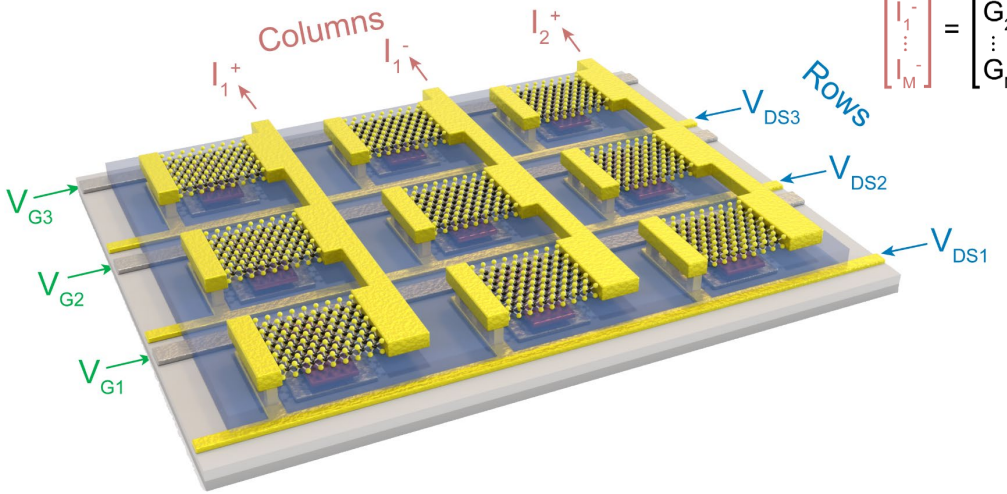




Marega…Kis; Nature (2020)

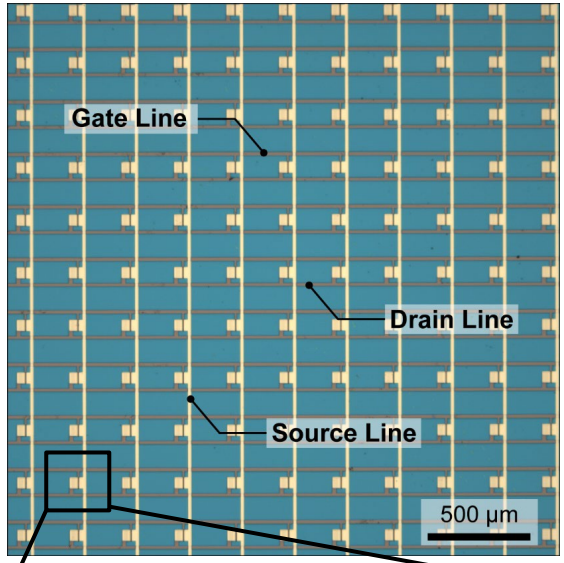
Large-scale Integrated MoS_2 circuits

FGFET Matrix

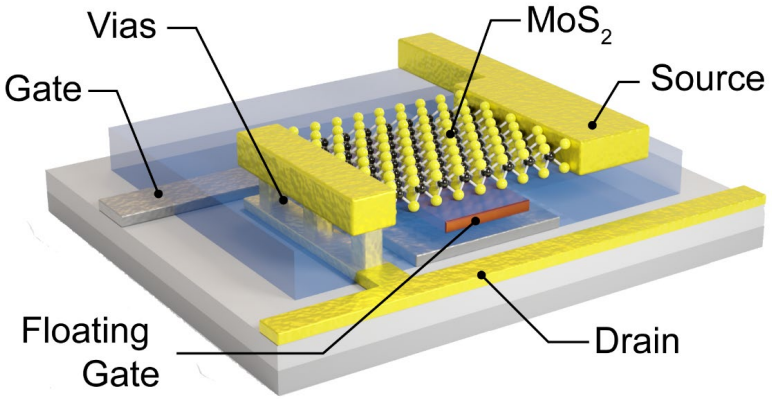


Vector-Matrix Multiplication

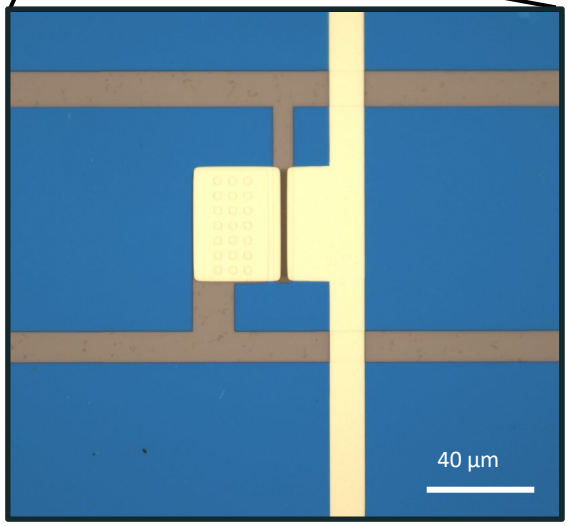
$$\begin{bmatrix} I_1^+ \\ I_1^- \\ \vdots \\ I_M^- \end{bmatrix} = \begin{bmatrix} G_{11}^+ & G_{12}^+ & \cdots & G_{1N}^+ \\ G_{21}^- & G_{22}^- & \cdots & G_{2N}^- \\ \vdots & \vdots & \ddots & \vdots \\ G_{M1}^- & G_{M2}^- & \cdots & G_{MN}^- \end{bmatrix} \cdot \begin{bmatrix} V_{DS1} \\ V_{DS2} \\ \vdots \\ V_{DS(N)} \end{bmatrix}$$



32×32 FGFET array
1024 devices



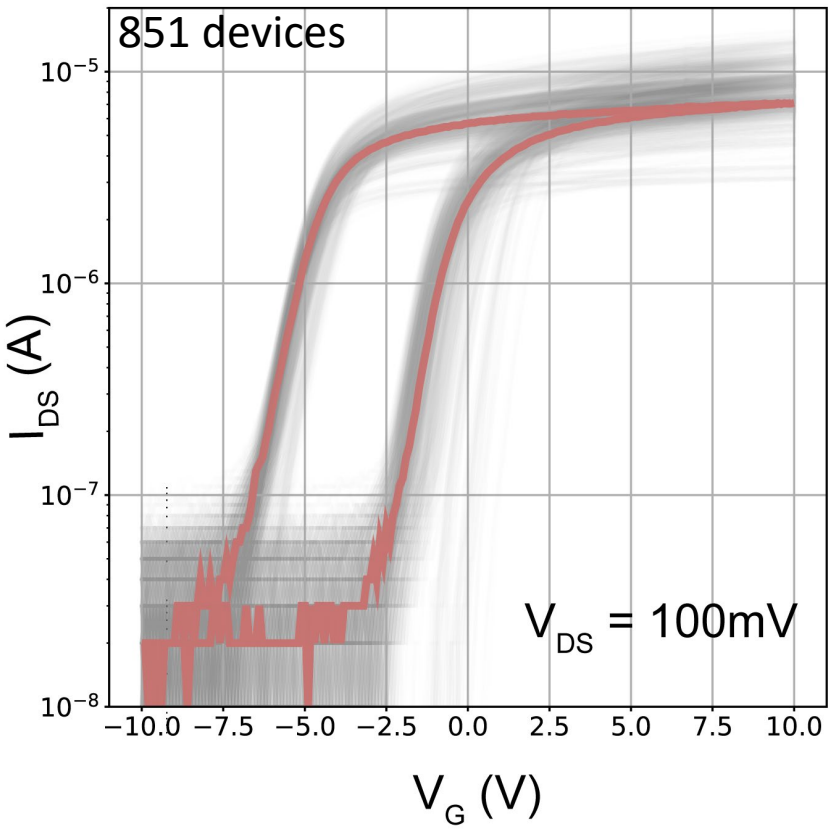
Floating-gate Field-Effect Transistor

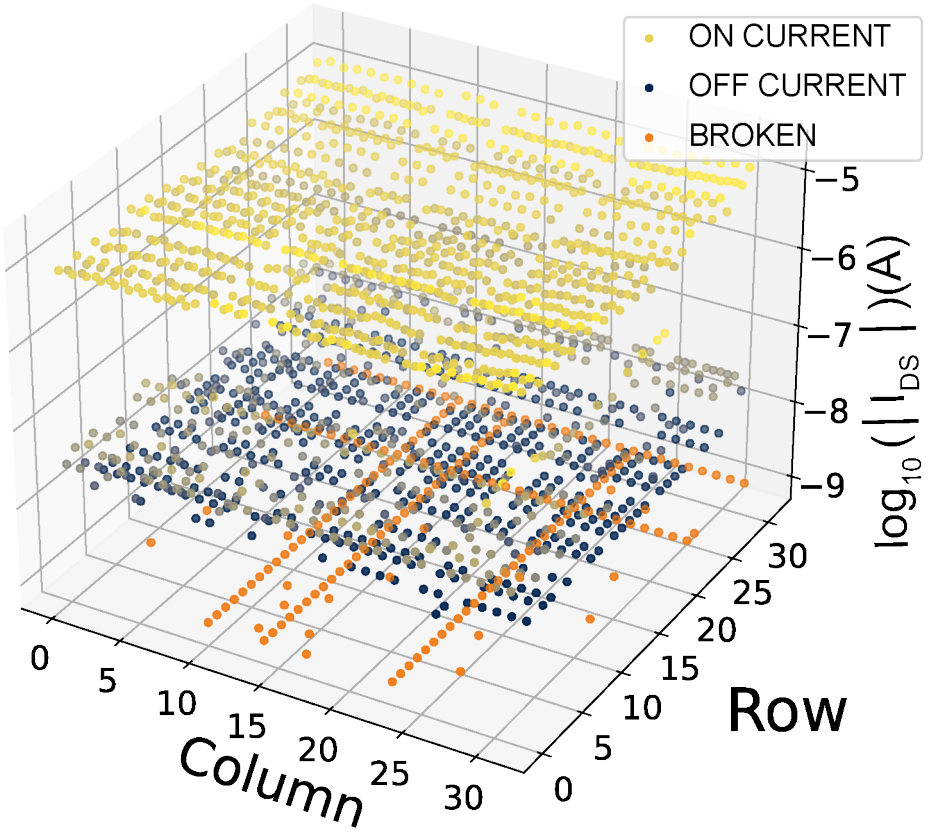


Marega…Kis; Nature Electronics (2023)

Large-scale Integrated MoS_2 circuits

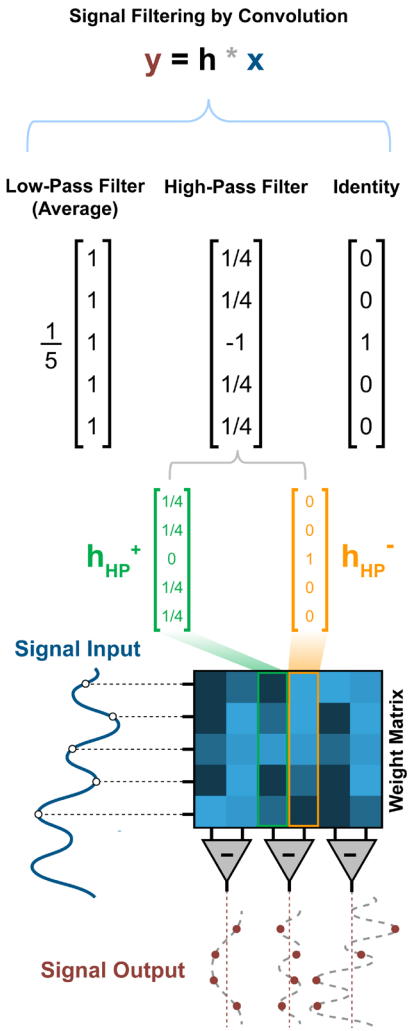
32×32 FGFET array
1024 devices
83% yield

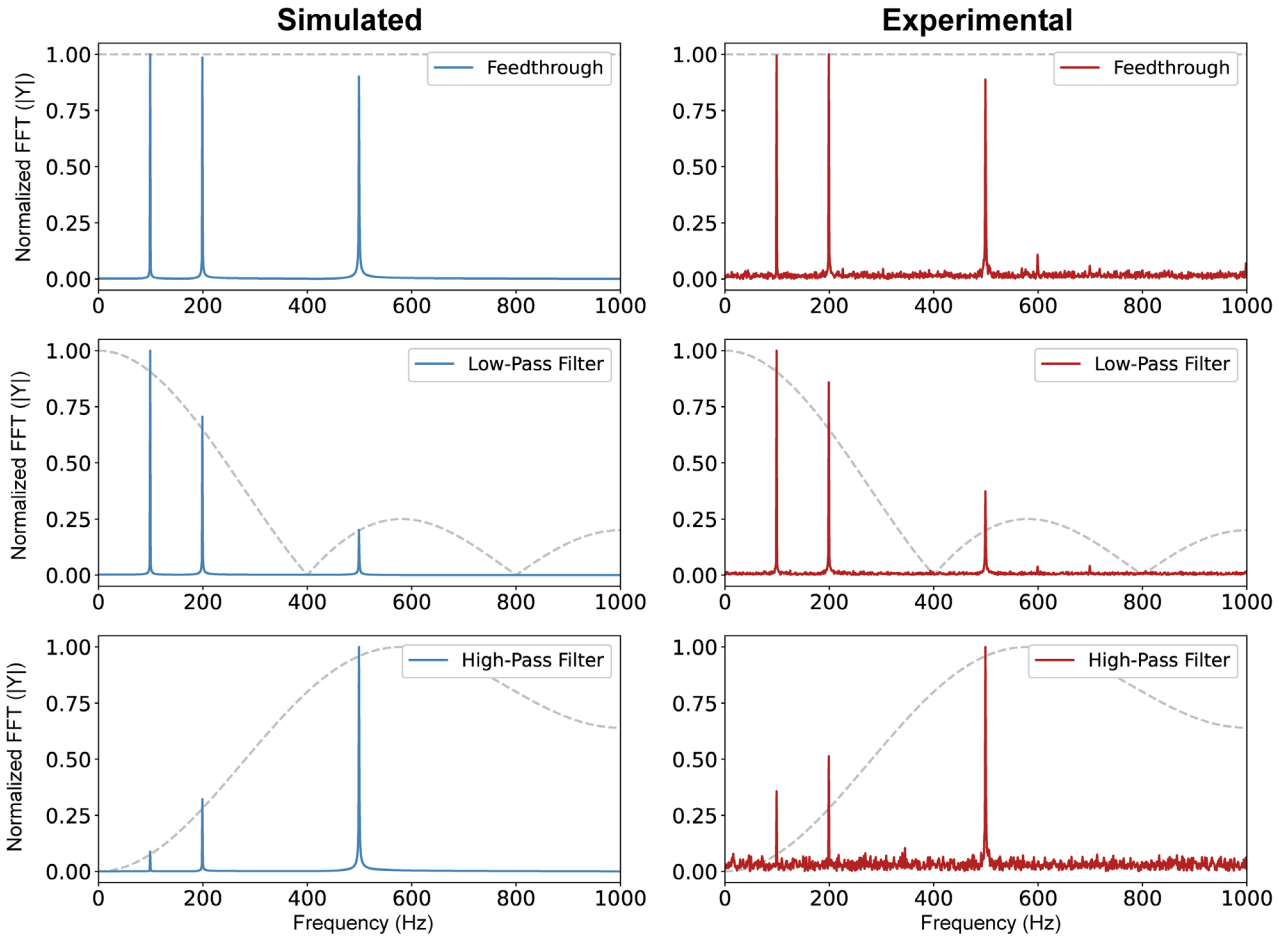




Marega…Kis; Nature Electronics (2023)

In-Memory Signal Processing





Marega…Kis; Nature Electronics (2023)

Recapitulation

Fundamental limits to power dissipation

- Landauer limit
- Von Neumann architecture
- Joule heating

Emerging concepts based on 2D materials

- Valley/spintronics
- Excitonic devices: currents, polarisation, valley polarisation
- Memory in logic