Nanodevices for bio-inspired computing

Julie Grollier

CNRS/Thales lab,
Palaiseau, France

bioSPINspired

Collaborations

CNRS/Thales spintronics team:
Alice Mizrahi, Steven Lequeux, Jacob Torrejon, Mathieu Riou, Philippe Talatchian, Miguel Romera, Flavio Abreu Araujo, Daniele Pinna, Romain Lebrun, Paolo Bortolotti, Vincent Cros

CNRS/Thales oxide team:
Vincent Garcia, Manuel Bibes, Sören Boyn, André Chanthbouala, Karin Garcia, Cécile Carretero, Stéphane Fusil, Stéphanie Girod, Eric Jacquet, Hiroyuki Yamada, Agnès Barthélémy

AIST Tsukuba Japan:
Sumito Tsunegi, Rie Matsumoto, Akio Fukushima, Kay Yakushiji, Hitoshi Kubota, Shinji Yuasa

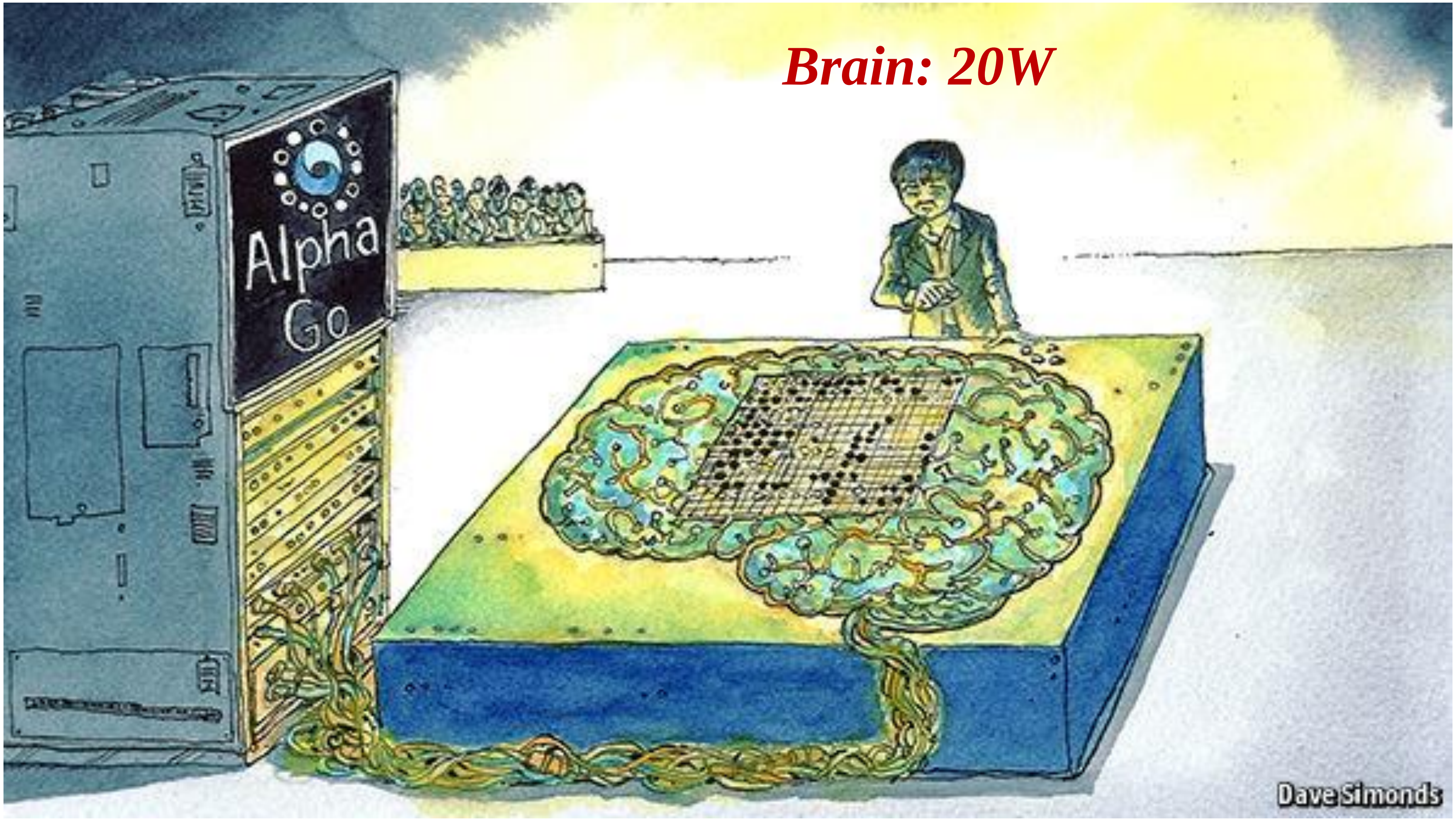
IEF Orsay France:
Damir Vodenicarevic, Joo-Von Kim, Nicolas Locatelli, Damien Querlioz

IMS Bordeaux France:
Gwendal Lecerf, Jean Tomas, Sylvain Saïghi

NIST Gaithersburg:
Guru Khalsa, Mark Stiles

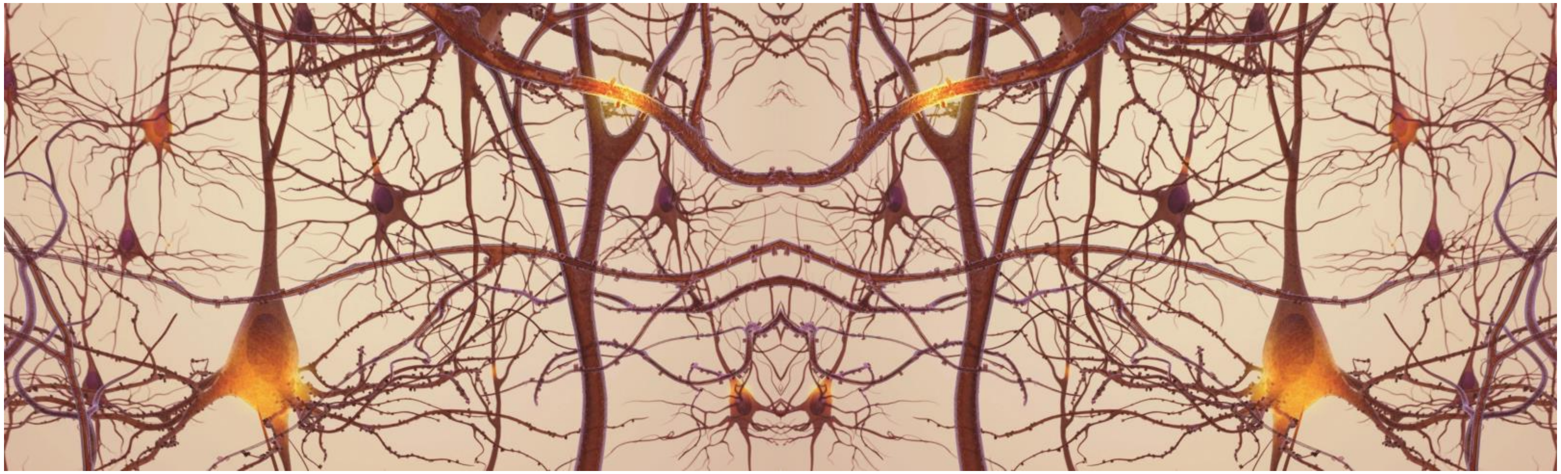
Impressive progress in artificial intelligence but in terms of power efficiency, the brain is the winner

AlphaGo: 150 kW



Computing « like the brain » requires increasing parallelism

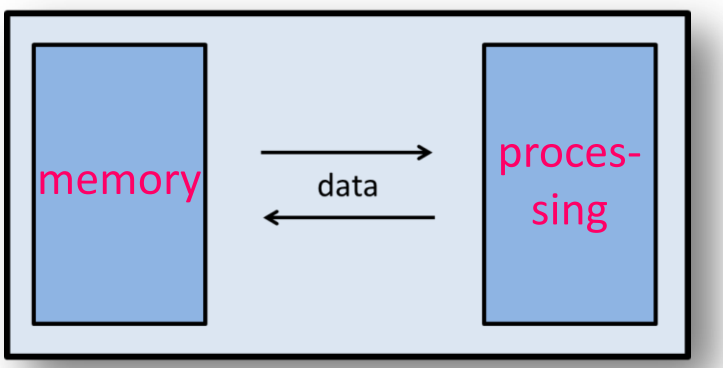
- The brain is massively parallel
- Artificial Neural networks as well
- Recent progress in AI: new models, but also increasing parallelism of information processing (GPUs)
- Current trend : FPGAs



Computing with low energy requires entangling memory with processing

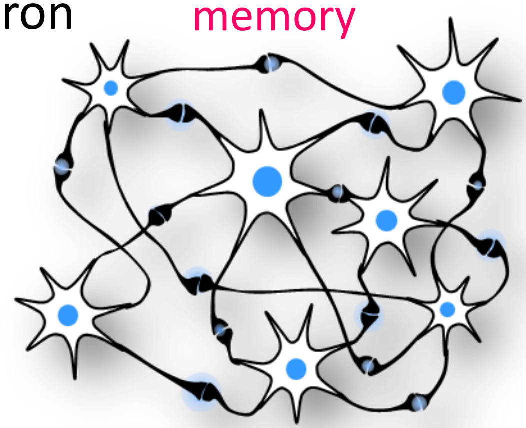
Power consumption in current computers is high because of Von Neumann’s bottleneck: separation between memory and processing

Digital computer



100 W/cm^2

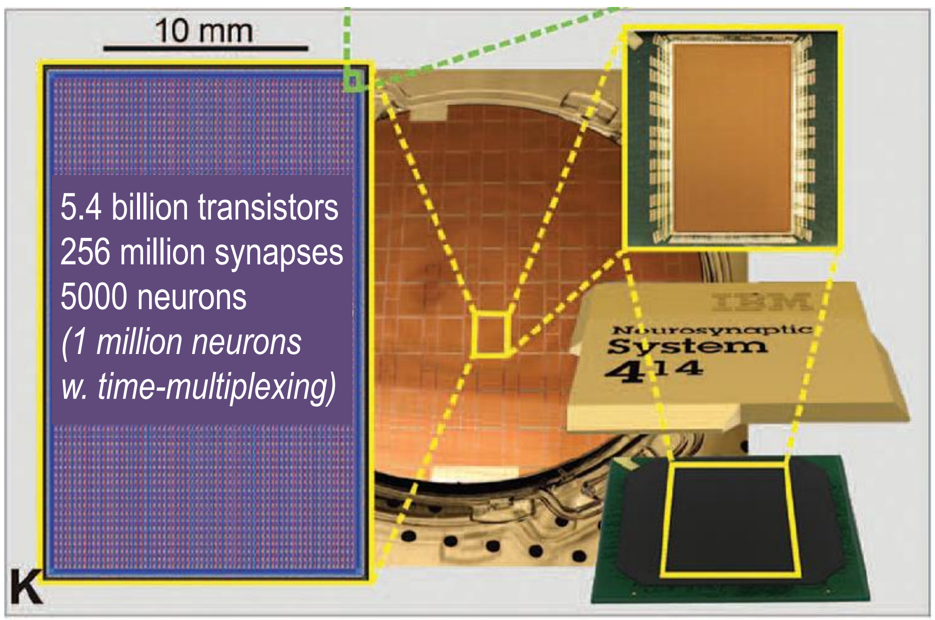
Brain

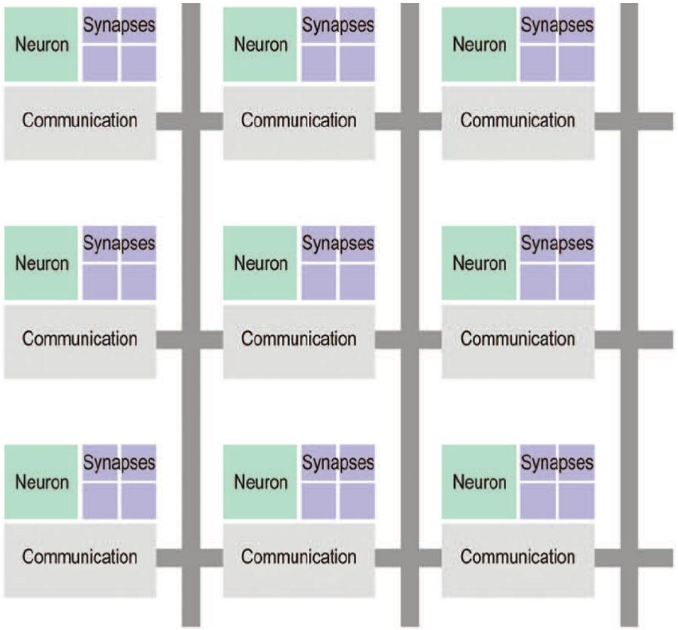


20 W in total !

Recent progresses with CMOS technology : example of IBM's TrueNorth chip

- Highly parallel, colocalized memory and processing
- Low power consumption 20 mW/cm^2 (processor 100 W/cm^2)
- Cannot learn



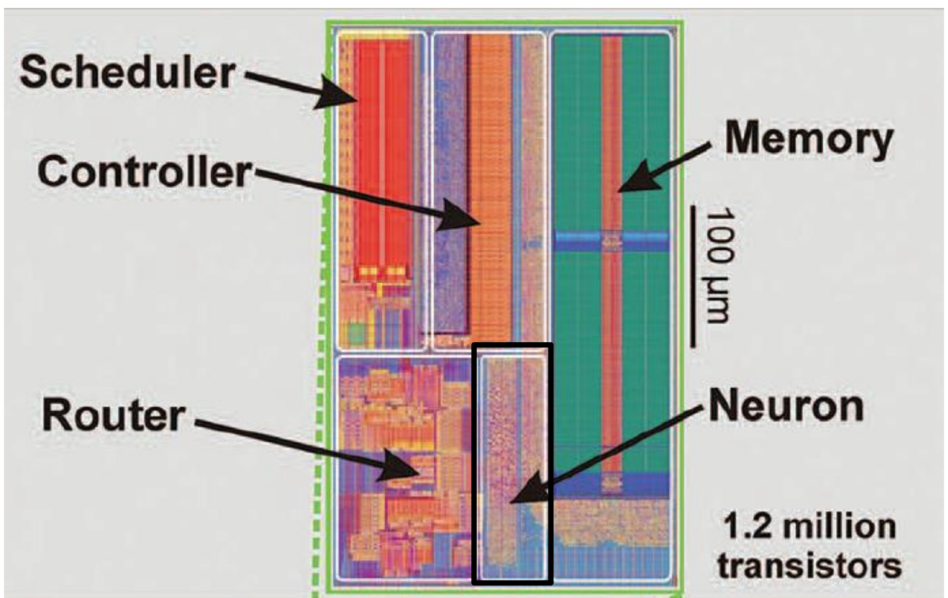


Merolla et al, *Science* **345**, 668 (2014)

Relying on current technology (CMOS) alone is not a long-term solution

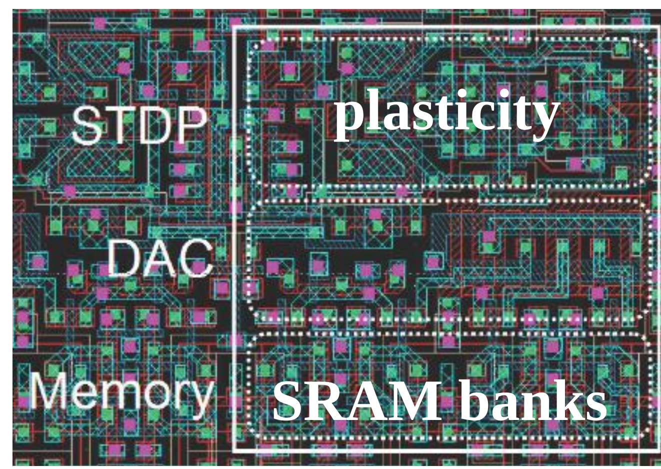
- A transistor is nanoscale but it is just a switch
- CMOS does not provide memory (volatile)

CMOS neuron **10-100 μm**



Merolla et al, *Science* **345**, 668 (2014)

CMOS synapse **10 μm**



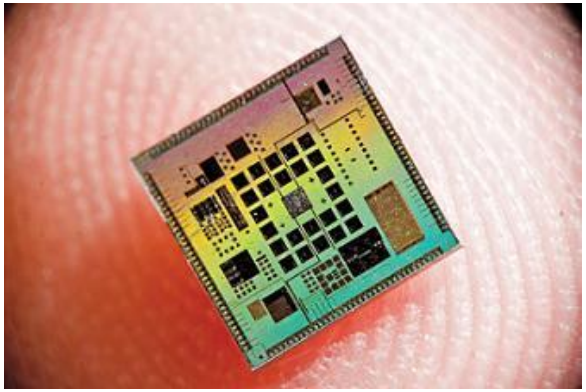
Schemmel et al., *IJCNN* (2006)

To build smart chips, novel nanoscale devices are needed to emulate neurons and synapses

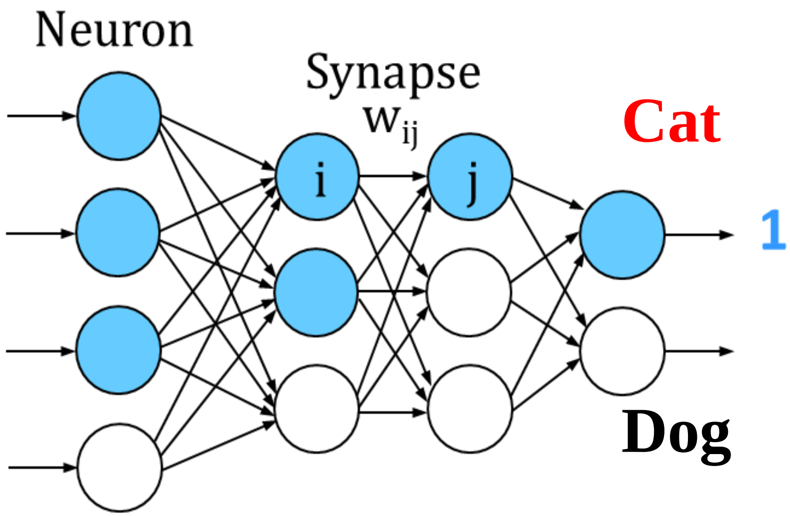
- Brain : 10^{11} neurons, 10^{15} synapses
- AlphaGo: millions of neurons and synapses
- Visual system: 500 millions of neurons

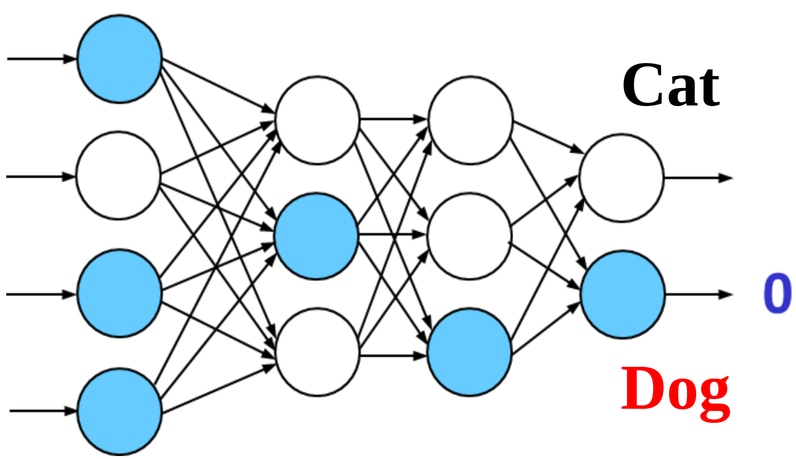
Hundreds of millions of neuron-like and synapse-like devices in a 1 cm^2 chip

→ Each device << 1 μm^2

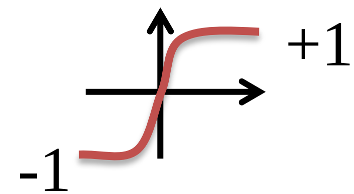


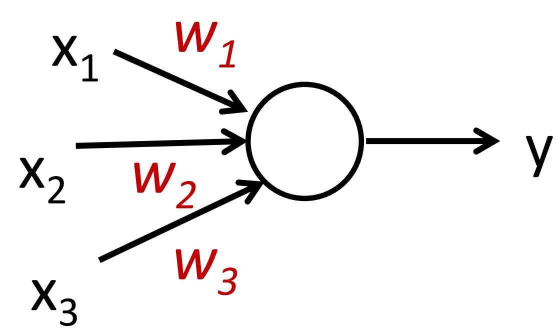
Ingredients needed for neural networks: non-linearity, memory and plasticity





- **Synapses**: analog valves (weights w)
- **Neurons**: non-linear





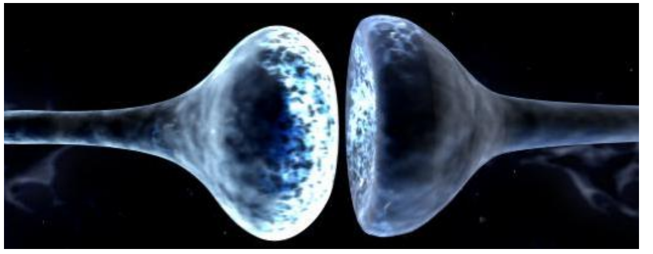
$$y = f\,(\,w_1\,x_1 + w_2\,x_2 + w_3\,x_3)$$

Many options ! → I will focus on electronic devices

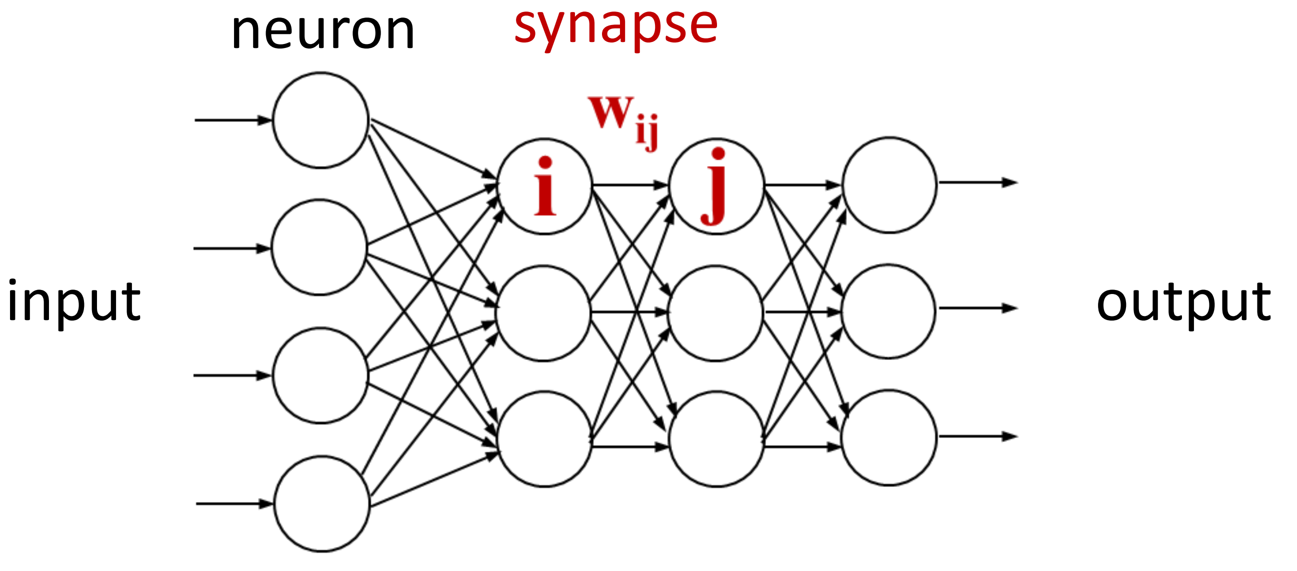
- Nano-Synapses
- Nano-Neurons
- Why neural networks and nanodevices are a great match

- Nano-Synapses
- Nano-Neurons
- Why neural networks and nanodevices are a great match

The fundamental ingredients of synapses are memory and plasticity



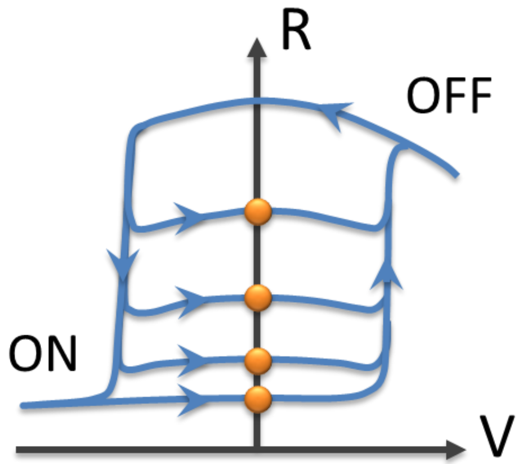
- Synapses: analog valves
- Learning: tuning synapses

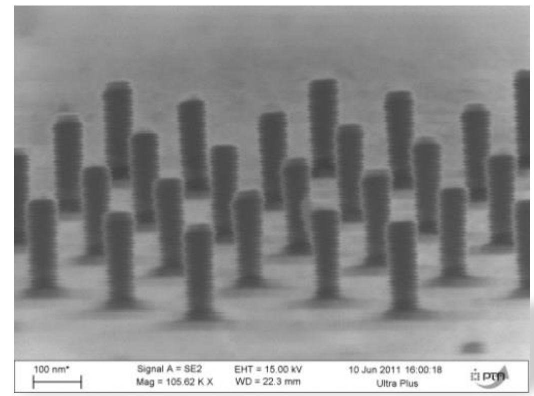


Memristors are tunable nano-resistors with memory

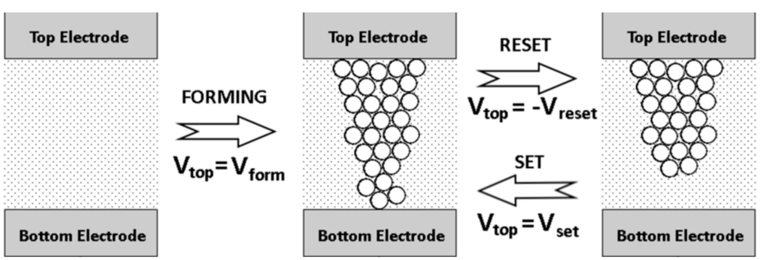
tunable
nano-resistor

Chua, *IEEE Trans. Circuit Theory* (1971)

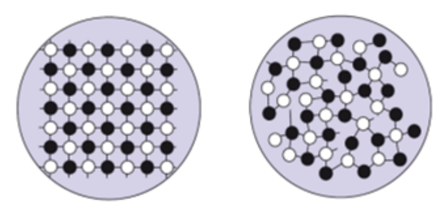




Red-Ox



Phase change

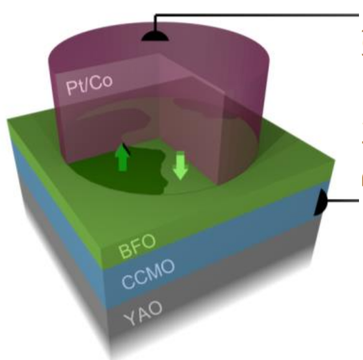


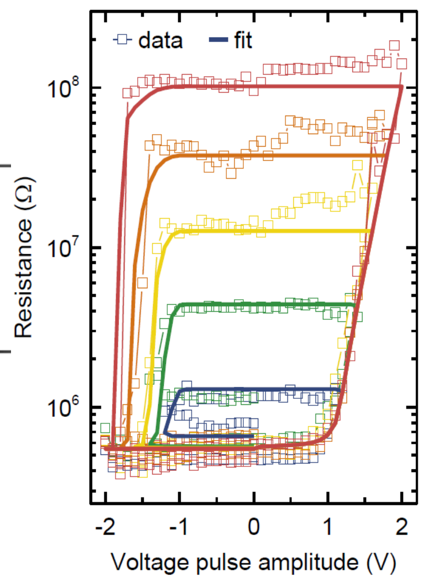
Yang et al., *Nat. Nano*. (2013)

Kuzum et al, *Nanotechnology* (2013)

Memristors that do not involve large ionic/atomic displacements are interesting for improved endurance and speed

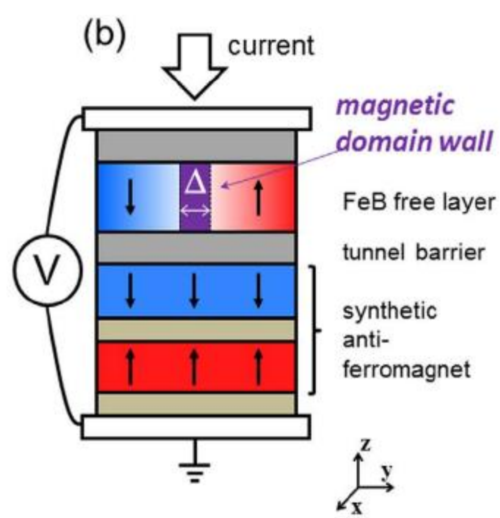
Ferroelectric

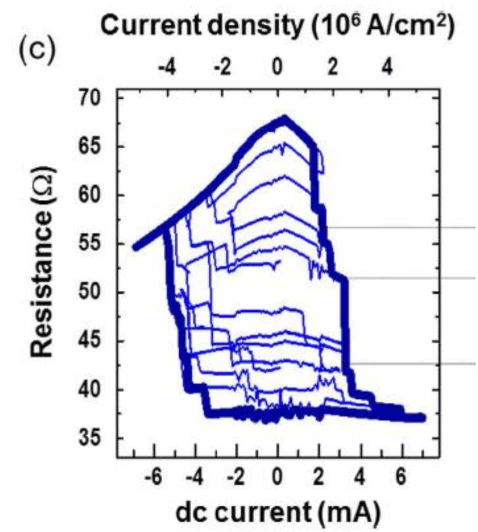




André Chanthbouala, **JG** et al, *Nat. Mat.* 11, 860 (2012)

Spintronic

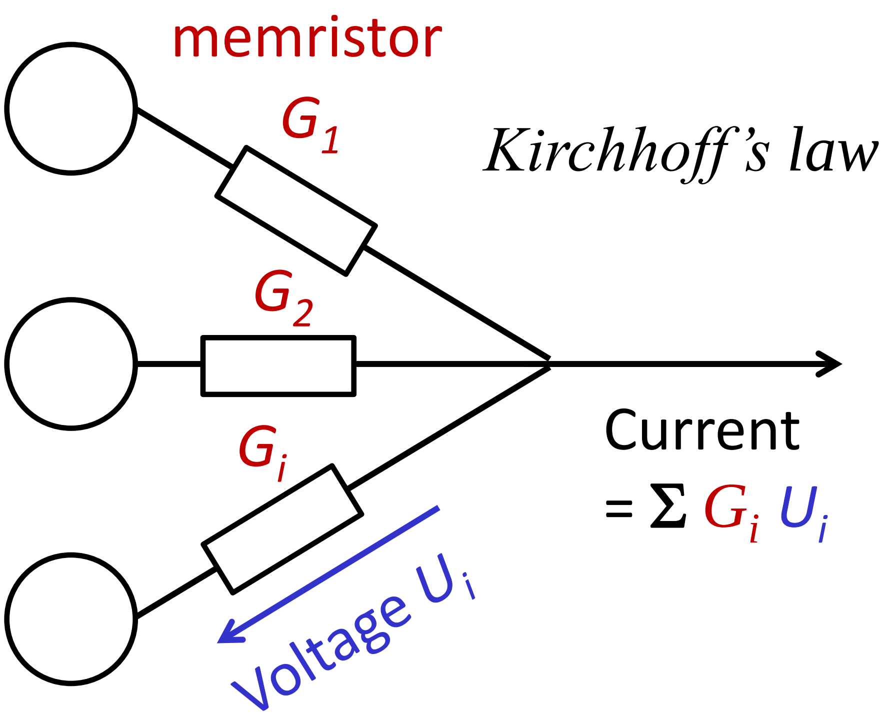


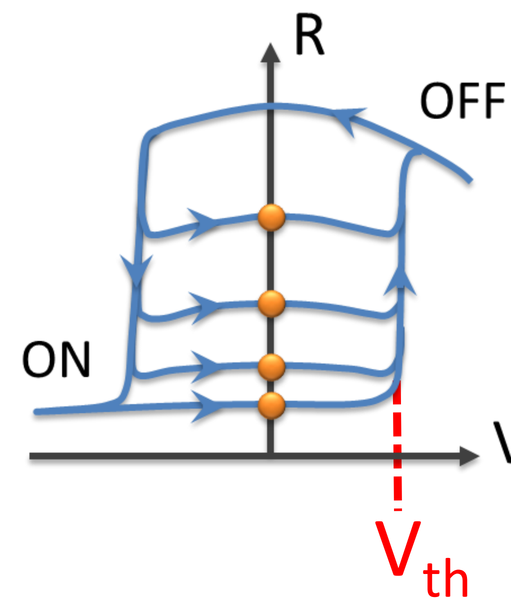


Steven Lequeux, **JG** et al, *Sci. Rep.* 6:31510 (2016)

Fukami et al, *Nat. Mater.* 15, 535 (2016)

Memristors emulate electronic synapses: the weight is their tunable conductance G





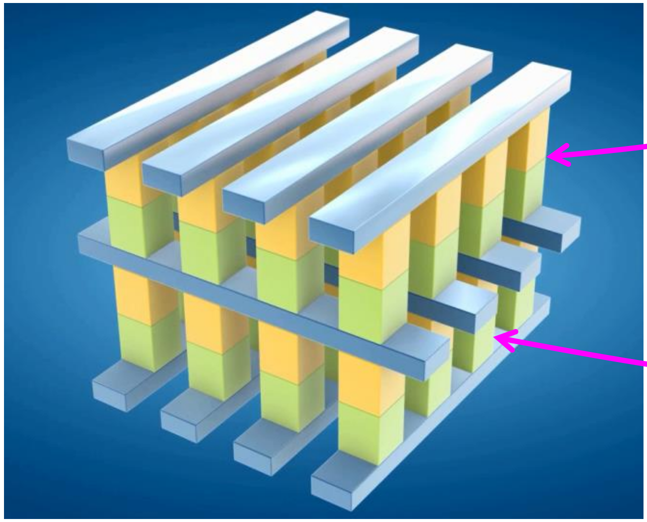
$U_i < V_{th}$: calculating mode

$U_i > V_{th}$: learning mode

One challenge being currently tackled is building large arrays of memristors

Recent progress has been achieved towards the commer-cialization of binary memories made of memristor arrays

- 3D Xpoint, Intel/Micron Optane Lenovo 32 Gbits

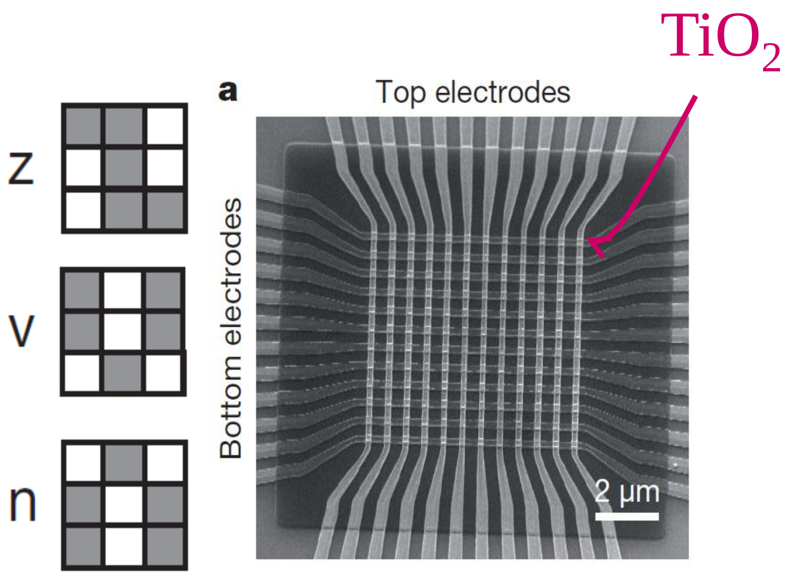


http://www.theregister.co.uk/2017/01/04/optane_arrives_at_ces/

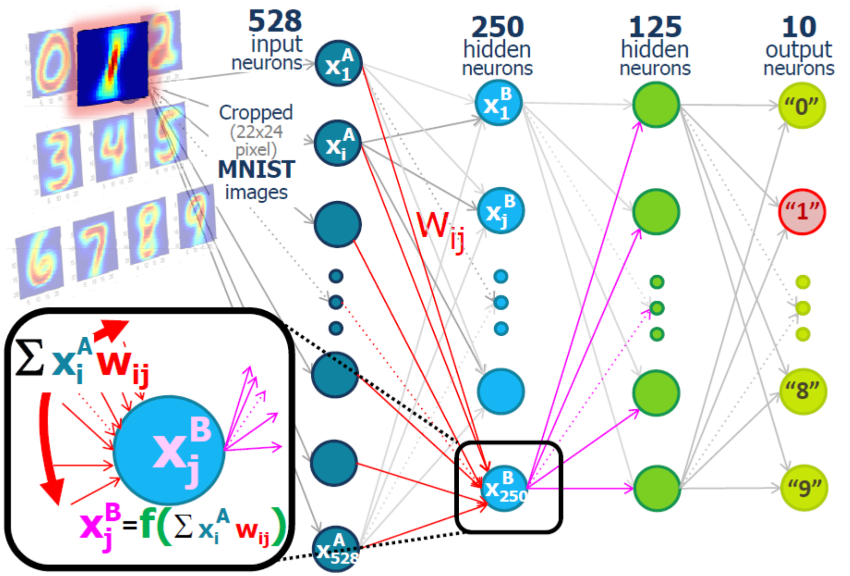
As for neuromorphic computing with memristors, the field is at its beginning

Supervised learning, back-propagation

Prezioso et al, *Nature* 521, 61 (2015)

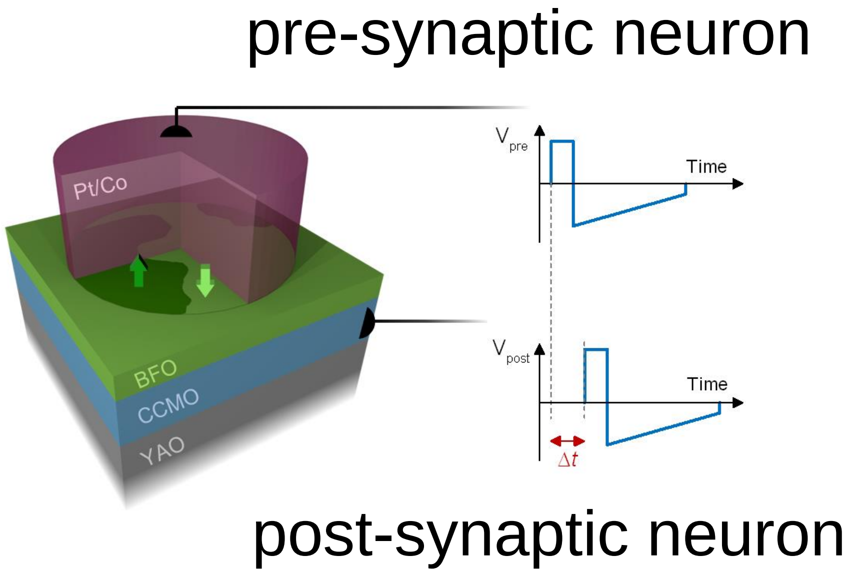


Burr et al, *IEEE IEDM* (2014)

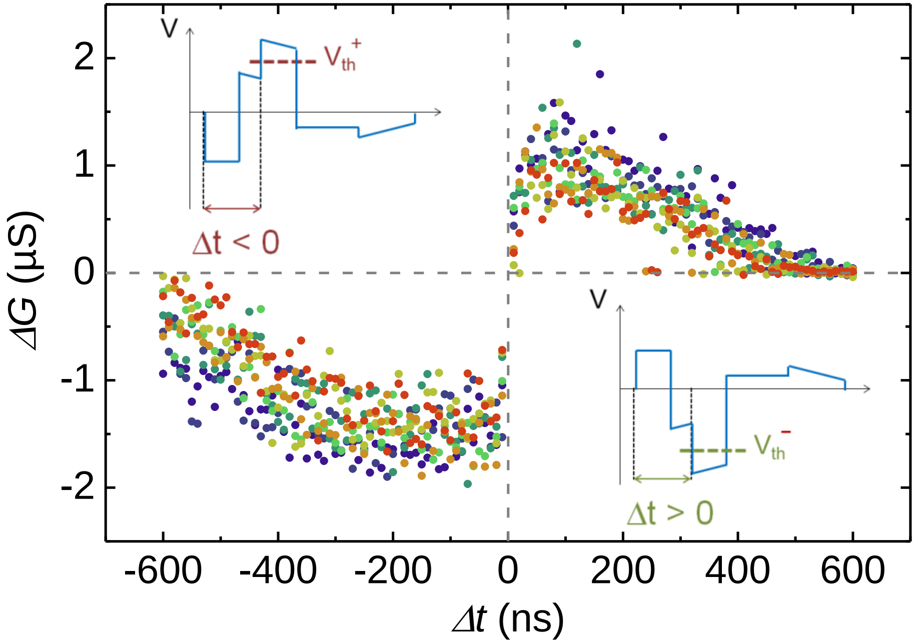


IBM experiments: Handwritten digit recognition with ~ 165 000 synapses (phase change with selector)

Memristors' resistance can evolve autonomously through spikes of neighbouring neurons: unsupervised learning possible



Spike timing plasticity



Jo et al, *Nanoletters* 10, 1297 (2010)
Zamarrenos-Ramos et al, *Frontiers in Neuroscience* 5, 26 (2011)
Sören Boyn, **JG** et al, Nature Com. 8, 14736 (2017)

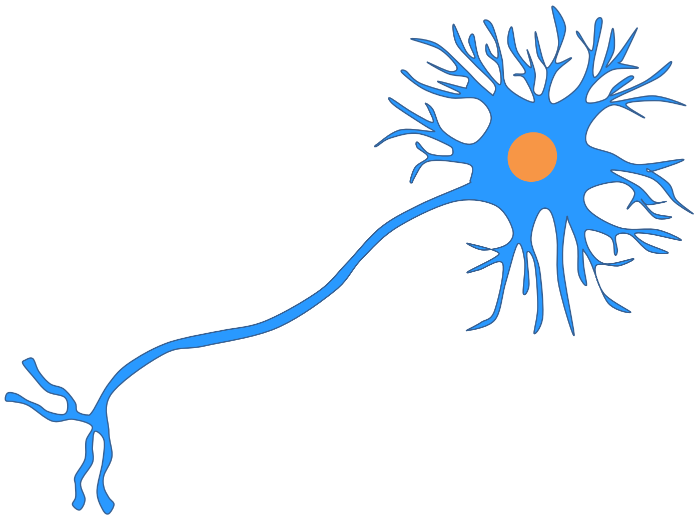
Perspectives

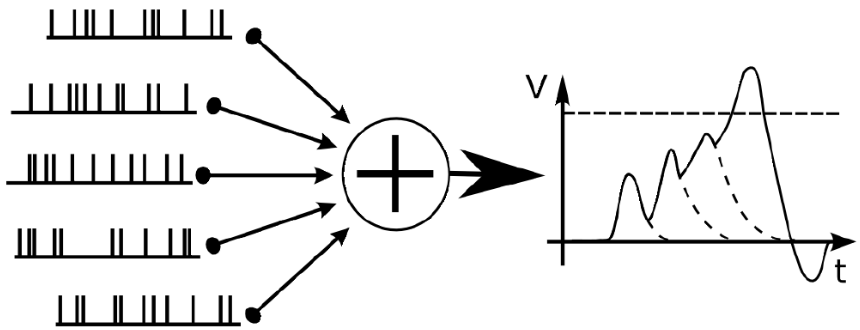
- Lots of technical work needed for improving memristors
- Experimental demonstration of unsupervised learning (coming soon)
- More physics for more synapse-like functionalities
- Novel types of devices

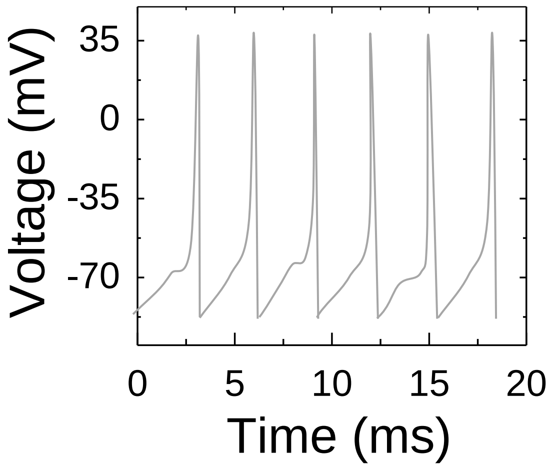
- Nano-Synapses
- Nano-Neurons
- Why neural networks and nanodevices are a great match

Biological neurons are oscillators

- Integration
- Spikes
- Non-linear oscillator
- Rate coding

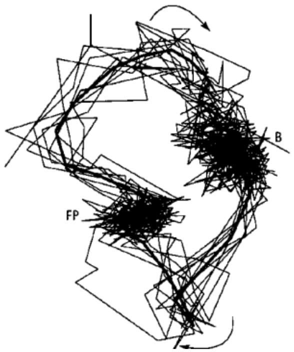




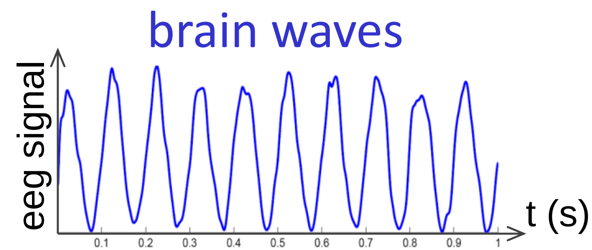


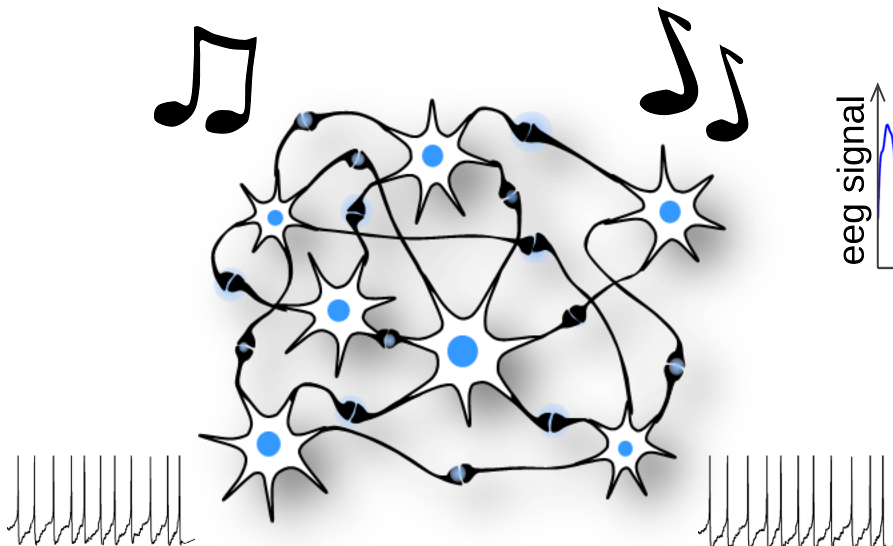
Non-linear dynamics in the brain has inspired many computing models

Complex transients



Synchronization





neuron: oscillator

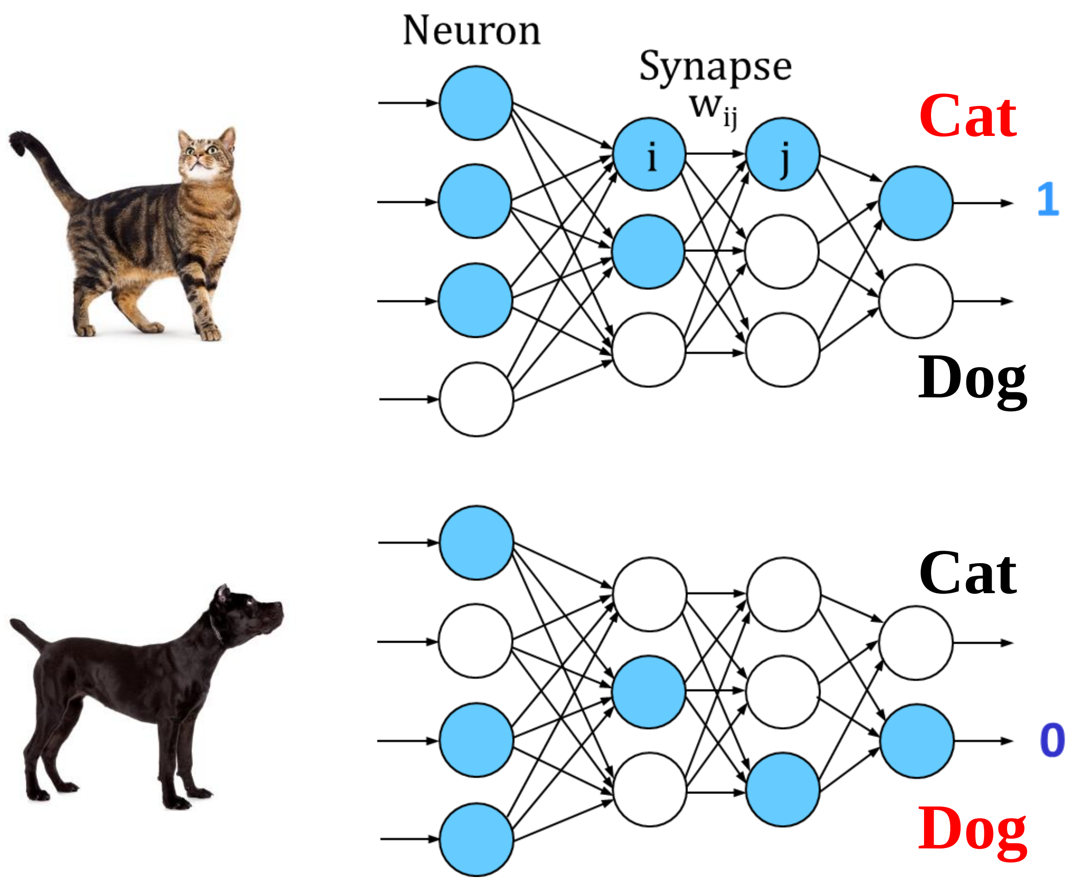
synapse: coupling

neuron: oscillator

Many attemps to realize neuromorphic computing with nanoscale oscillators: memristive oscillators, magnetic oscillators, MEMS…

Hoppensteadt et al, *PRL* (1999), Aonishi et al, *PRL* (1999) Jaeger et al, *Science* (2004)

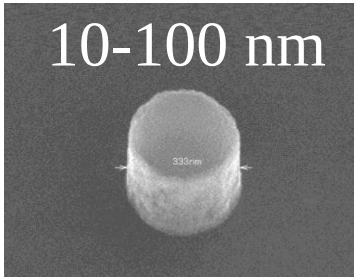
Nanoscale oscillators are noisy/unreliable
Neural networks are tolerant of input noise, but not of component unreliability



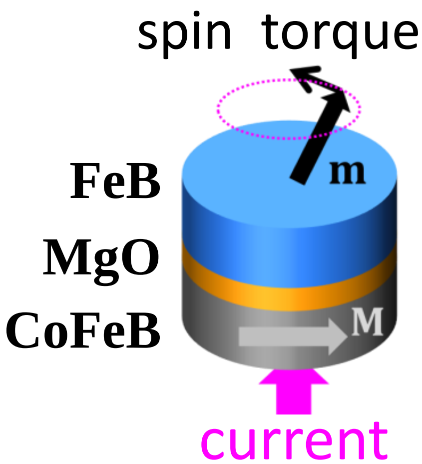
Magnetic oscillators are nanodevices with well controlled dynamics

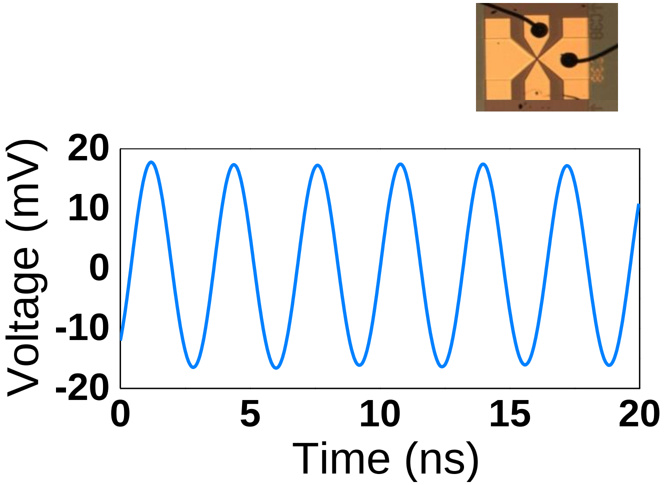
Nanoscale, fast (GHz) and easily measurable

magnetic tunnel junction



compatible with CMOS

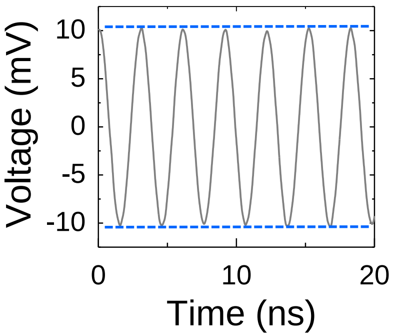


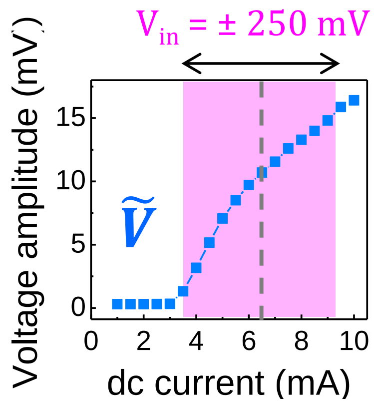


Same structure as magnetic memories

Nicolas Locatelli, V. Cros and J. Grollier, Spin-torque building blocks, Nat. Mat. 13, 11 (2014)

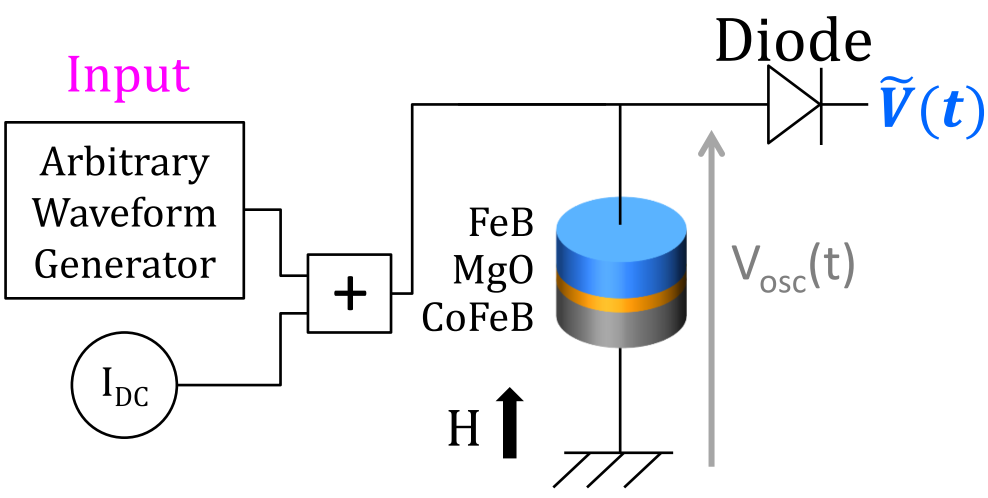
Spin-torque nano-oscillators can emulate neruons because their amplitude dynamics is non-linear and well-controlled

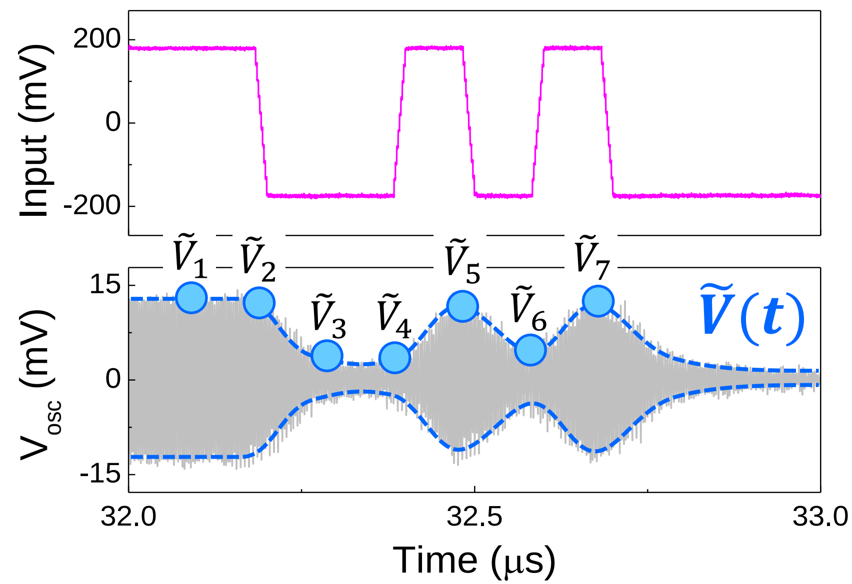




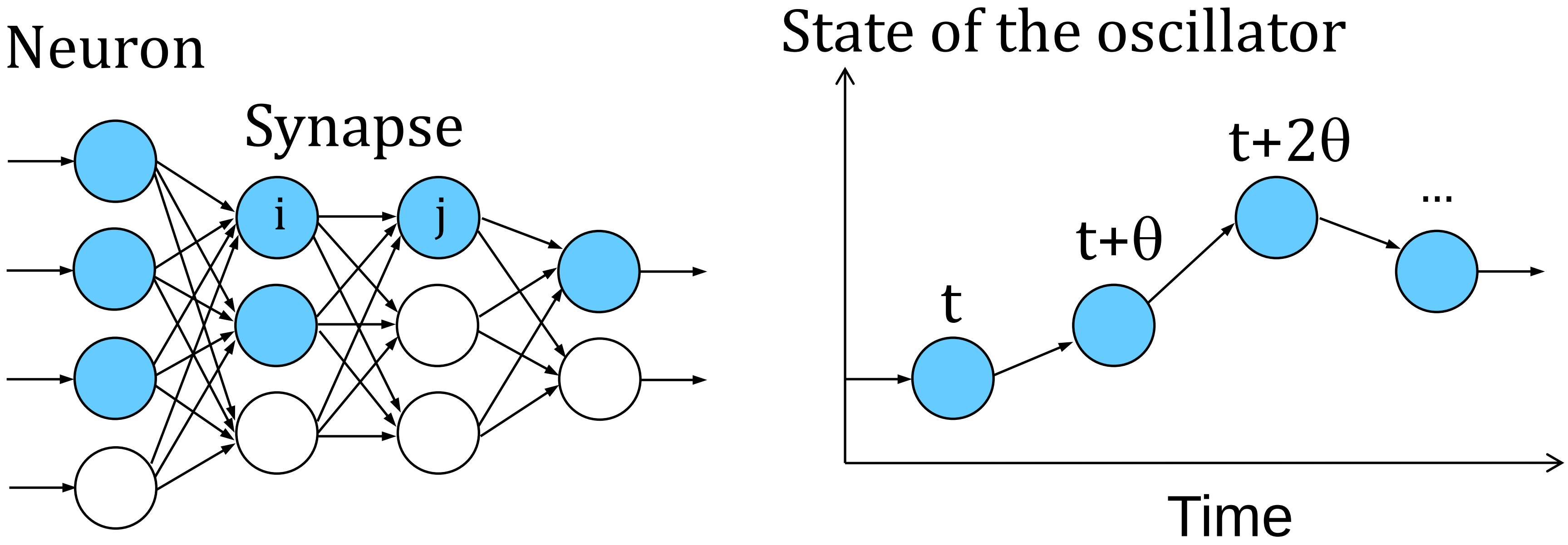
Input: current

Output: amplitude of the oscillator’s voltage





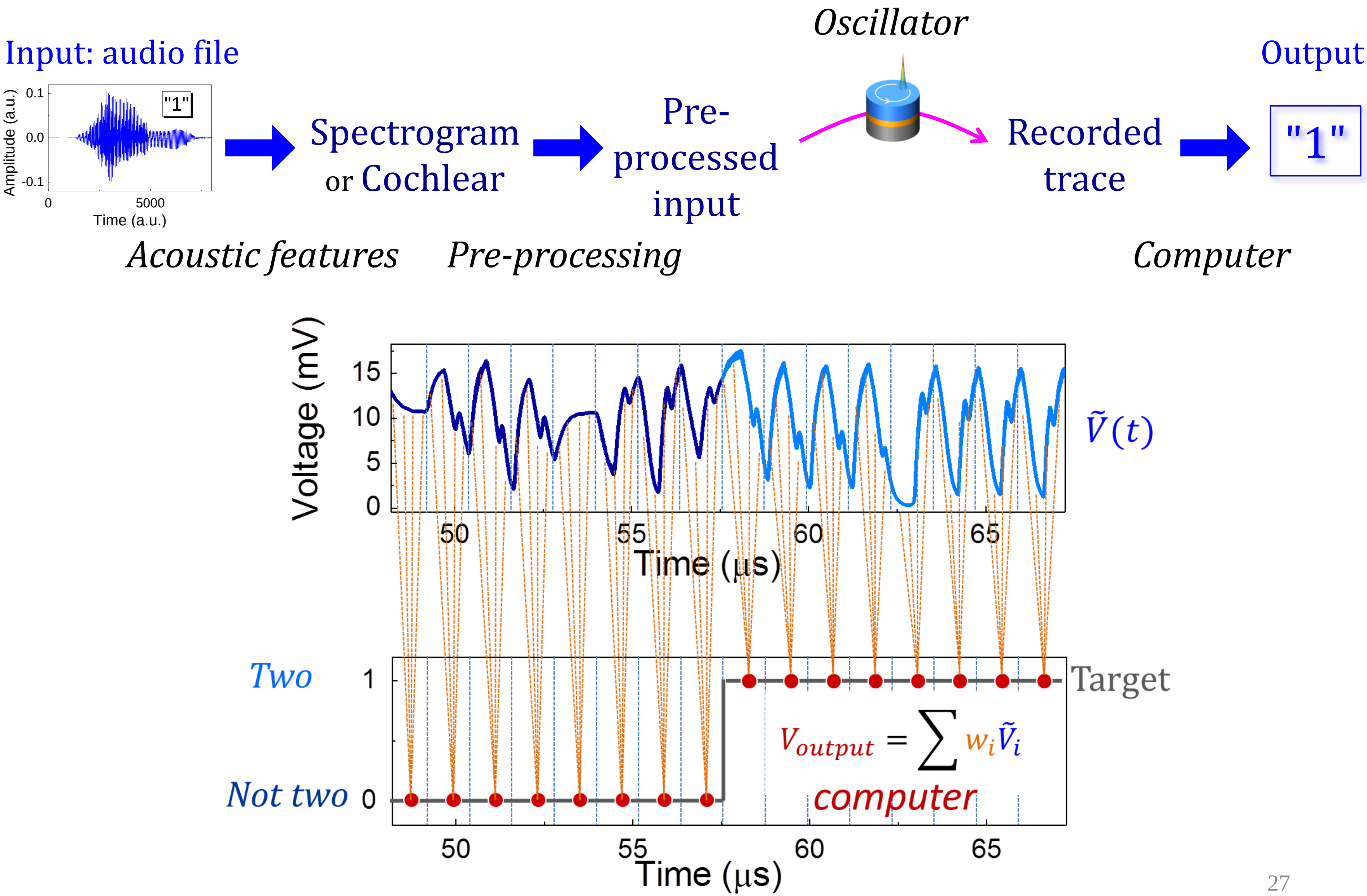
Proof of neuromorphic computing with spin-torque nano-oscillators: computing with a single oscillator through time-multiplexing



Reservoir computing approach

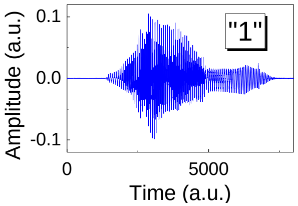
Appeltant et al, *Nat. Com* 2:468 (2011); Martinenghi et al, *PRL* 108, 244101 (2012)

Task: spoken digit recognition (NIST TI-46 corpus)



Experimental results of spoken digit recognition

Input: audio file



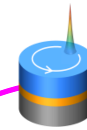
Spectrogram or Cochlear

Acoustic features

Pre-processed input

Pre-processing

Oscillator



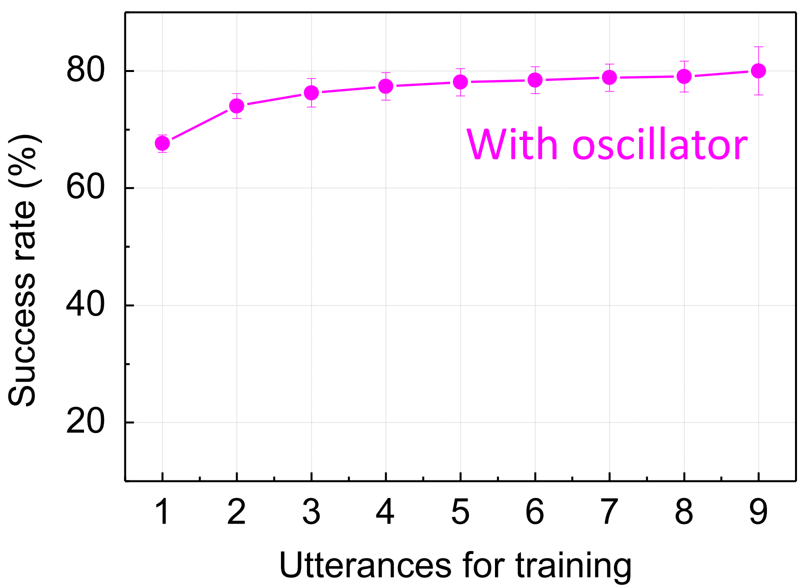
Recorded trace

Output

"1"

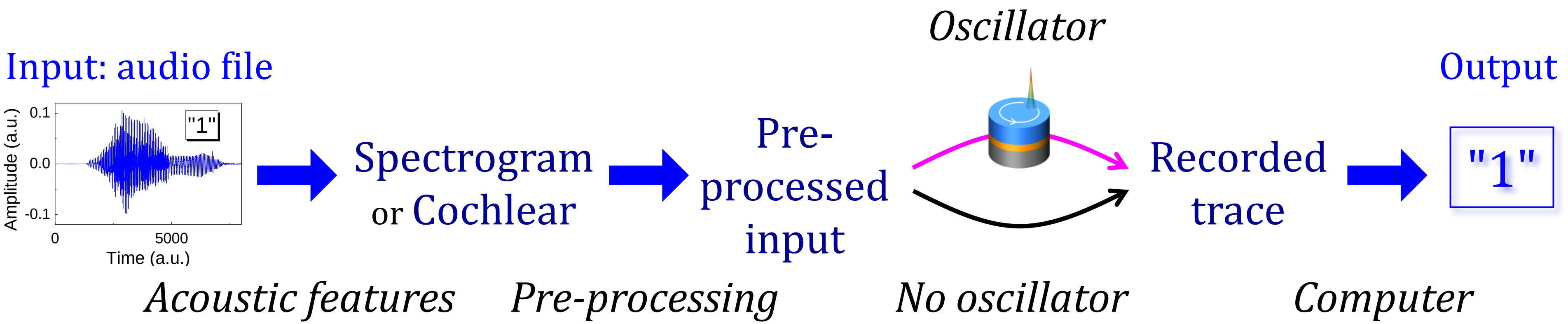
Computer

Spectrogram

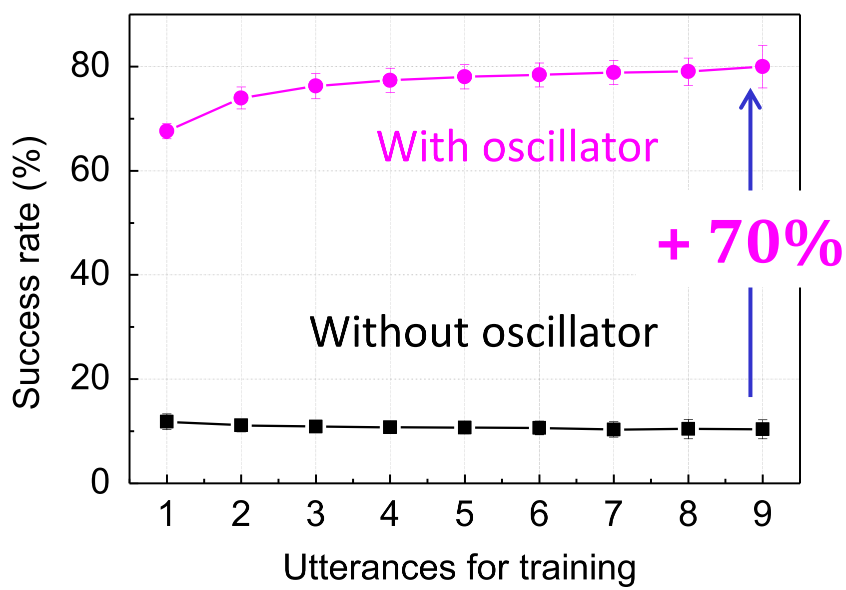


80%

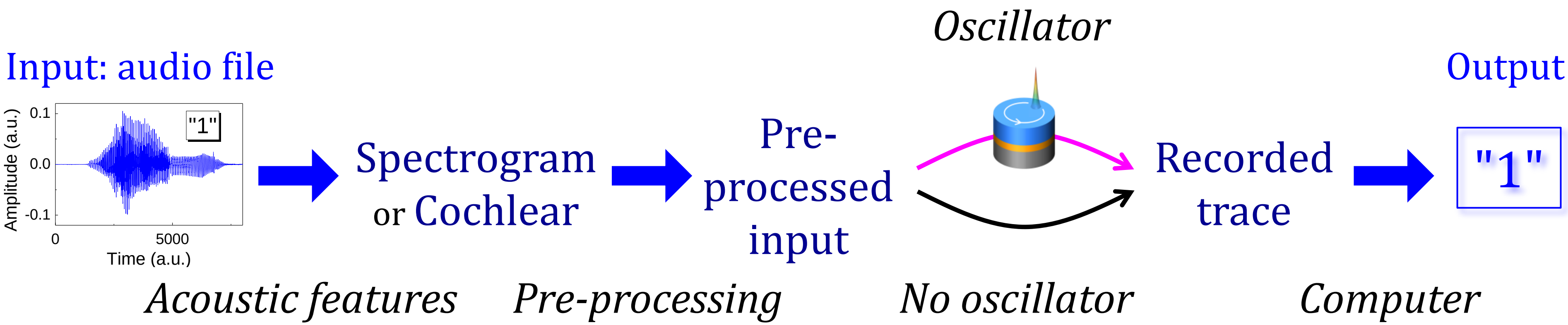
Experimental results of spoken digit recognition



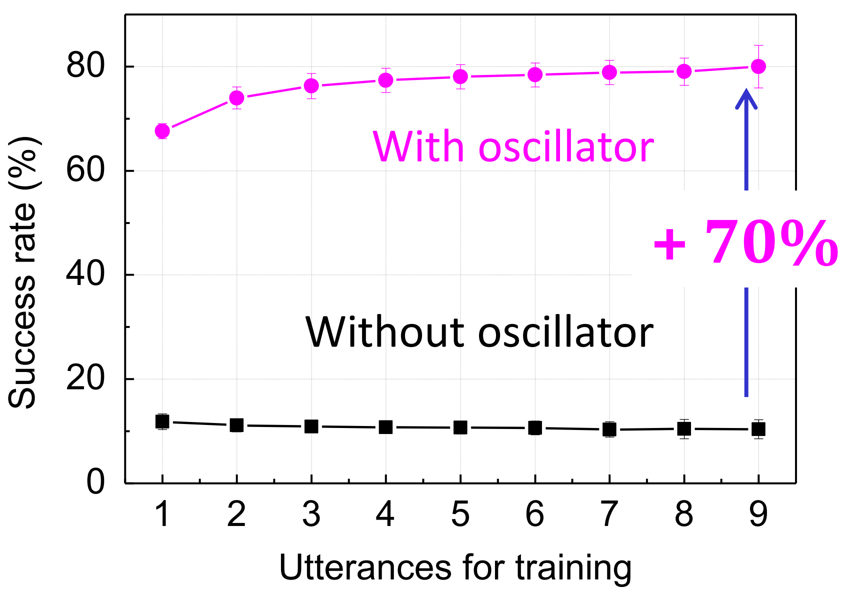
Spectrogram



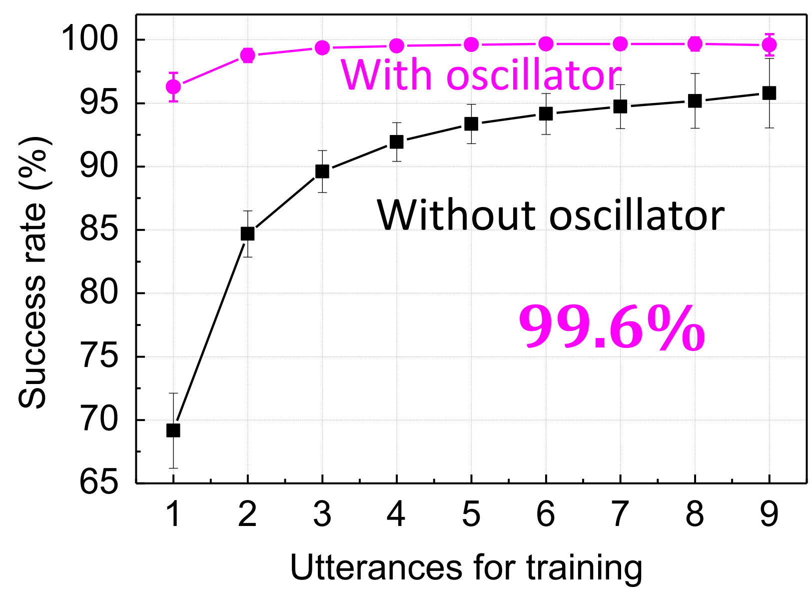
Experimental results of spoken digit recognition



Spectrogram



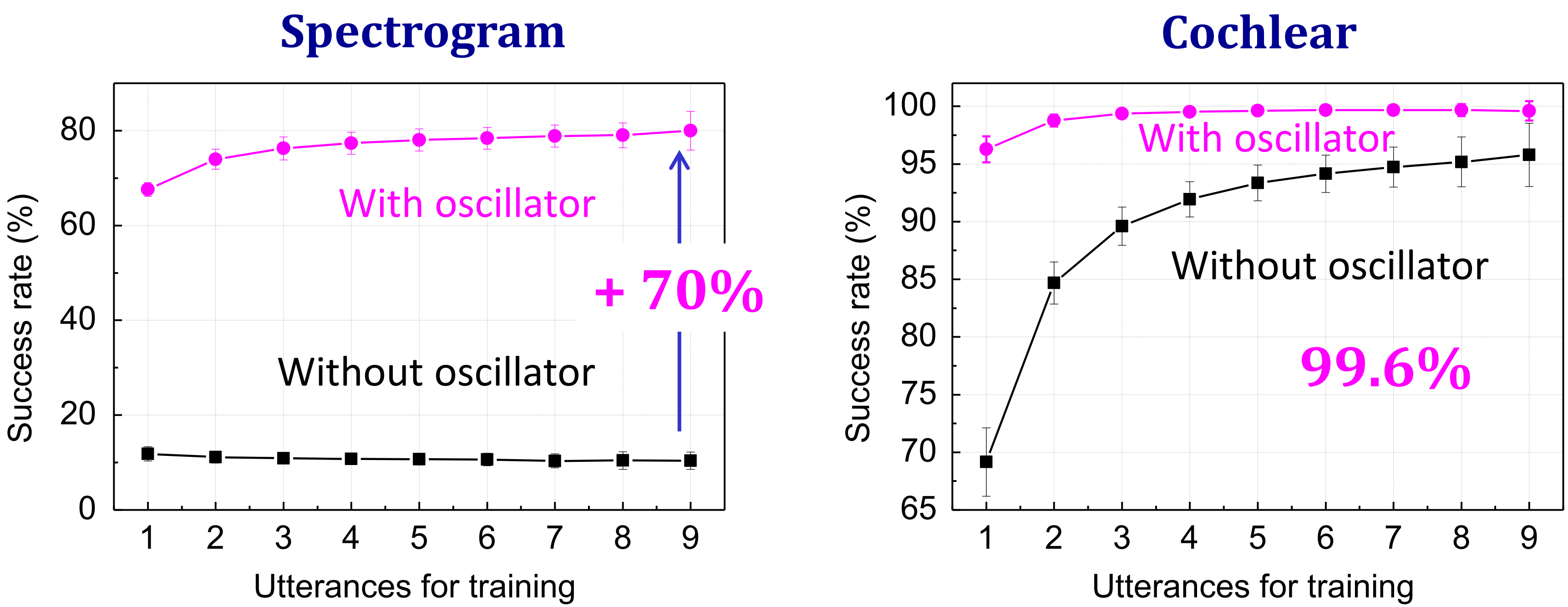
Cochlear



First demonstration of neuromorphic computing with a nanoscale oscillator

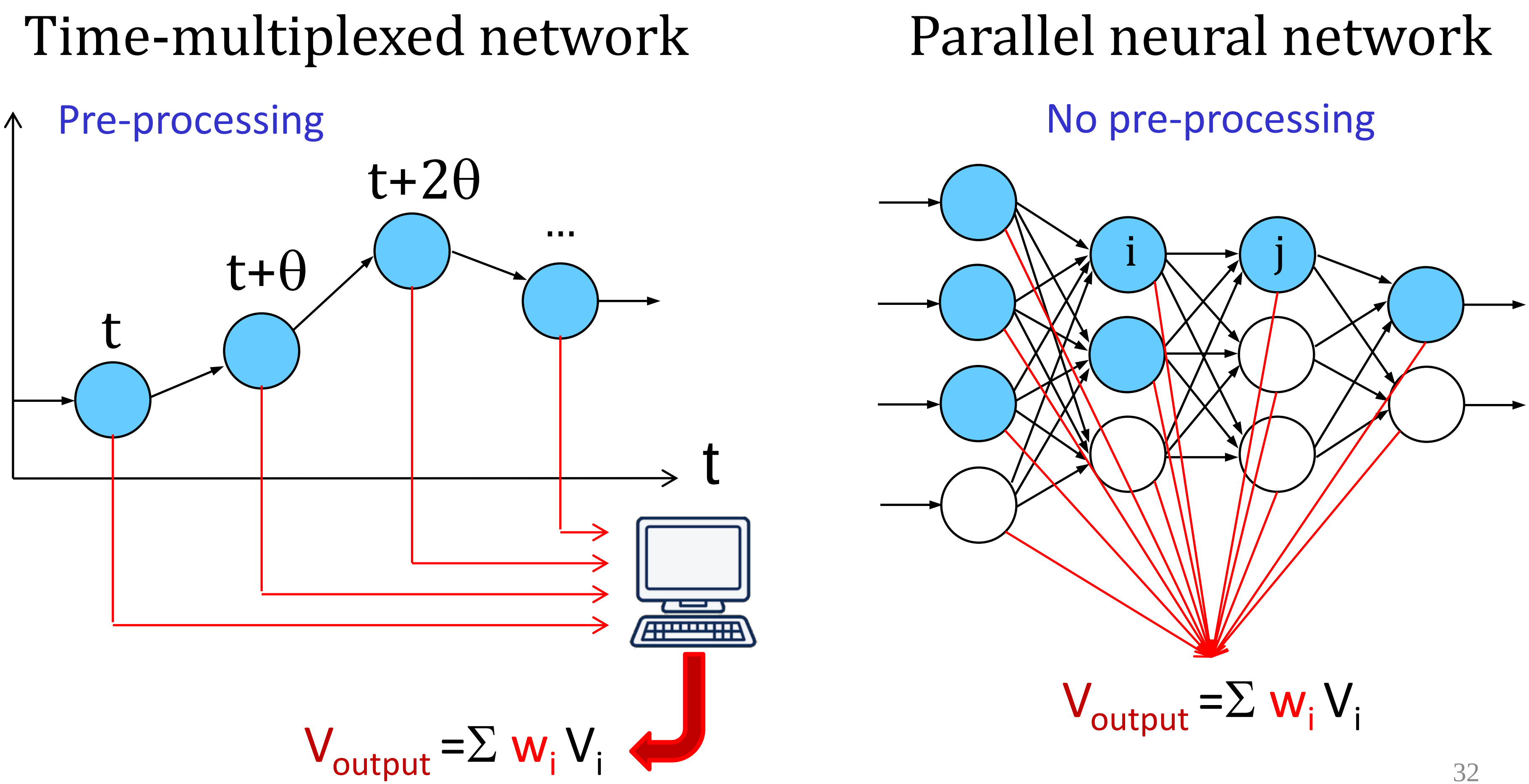
Jacob Torrejon-Diaz, Mathieu Riou, Flavio Abreu-Araujo, **JG** et al, arXiv:1701.07715

State of the art: 96 to 99.8 %



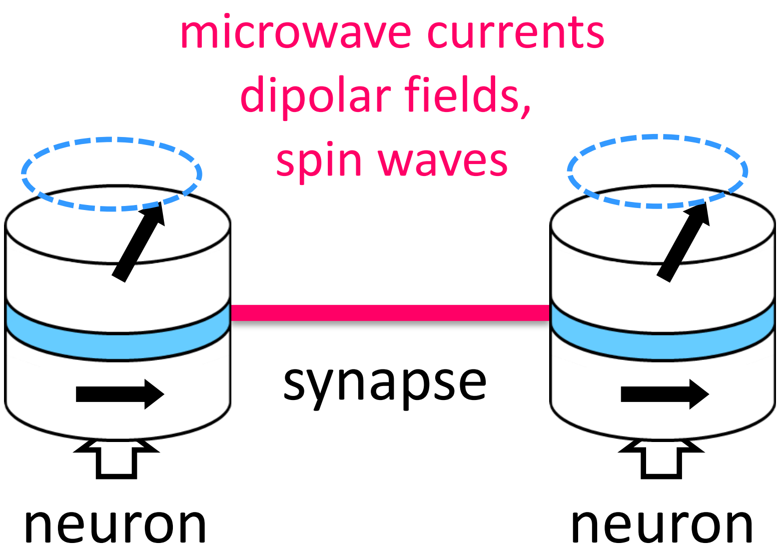
Fully parallel networks are necessary to speed up computing

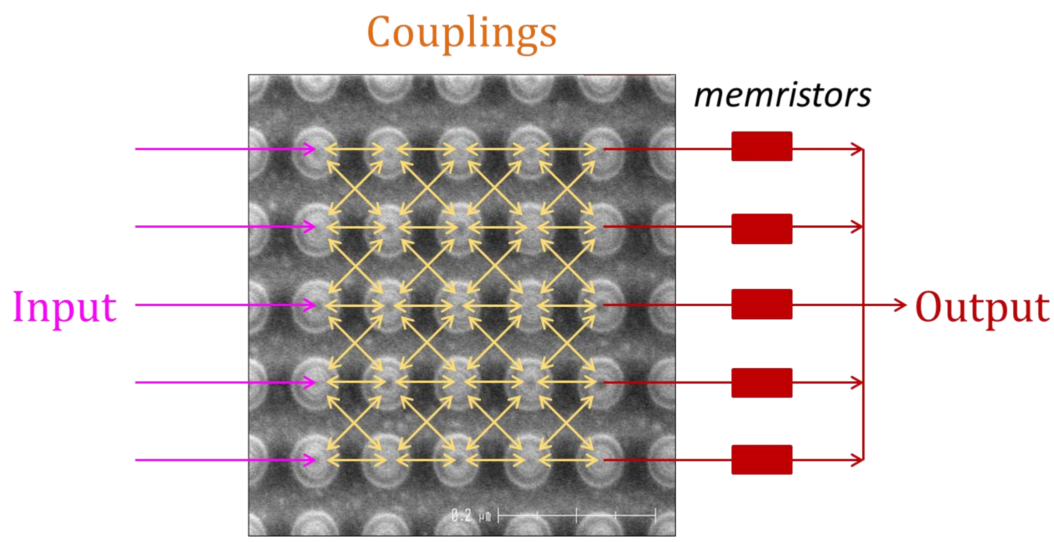
1 s ← **One million oscillators** → 1 μs



Coupling spin-torque nano-oscillators towards parallel neural networks

Enhanced ability to interact and synchronize





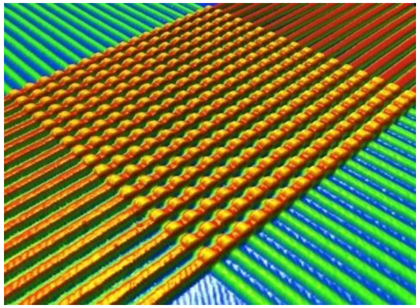
Slavin et al, *IEEE Trans Mag* 45, 1875 (2009)
Awad et al, *Nat. Phys*, 10.1038/NPHYS3927

Hynix/Toshiba IEDM 2016

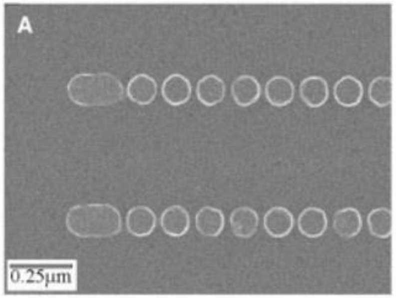
- Nano-Synapses
- Nano-Neurons
- Why neural networks and nanodevices are a great match

Novel nanodevices: interesting features but….

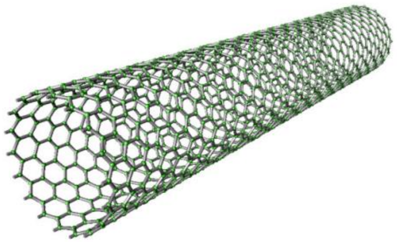
→ High device variability
→ High sensitivity to noise
→ Stochastic behavior



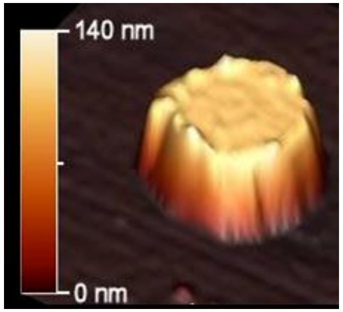
(Memristors, Hewlett-Packard)



(Nanomagnetic logic, Cowburn et al.)



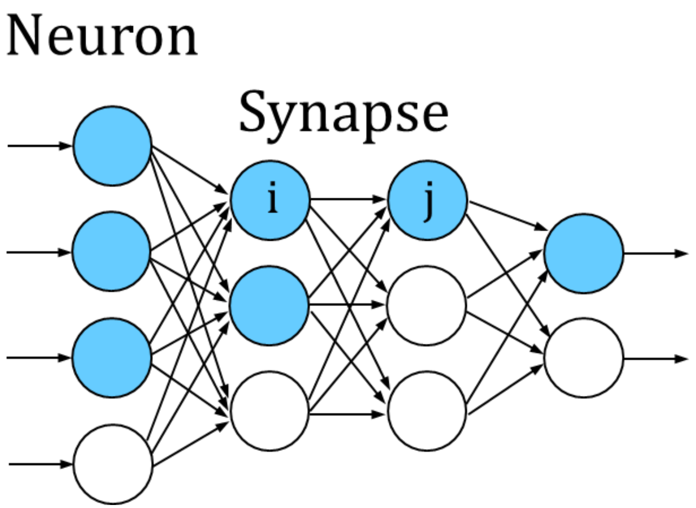
(Carbon nanotube)

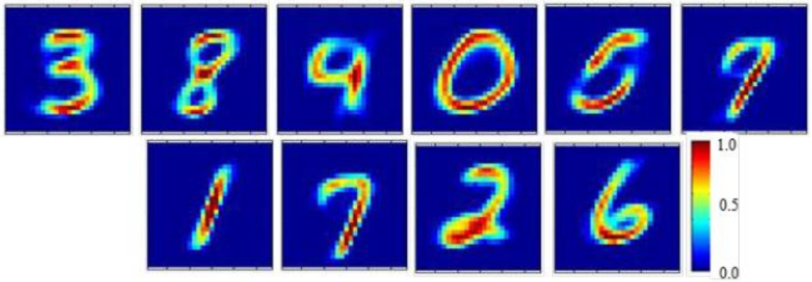


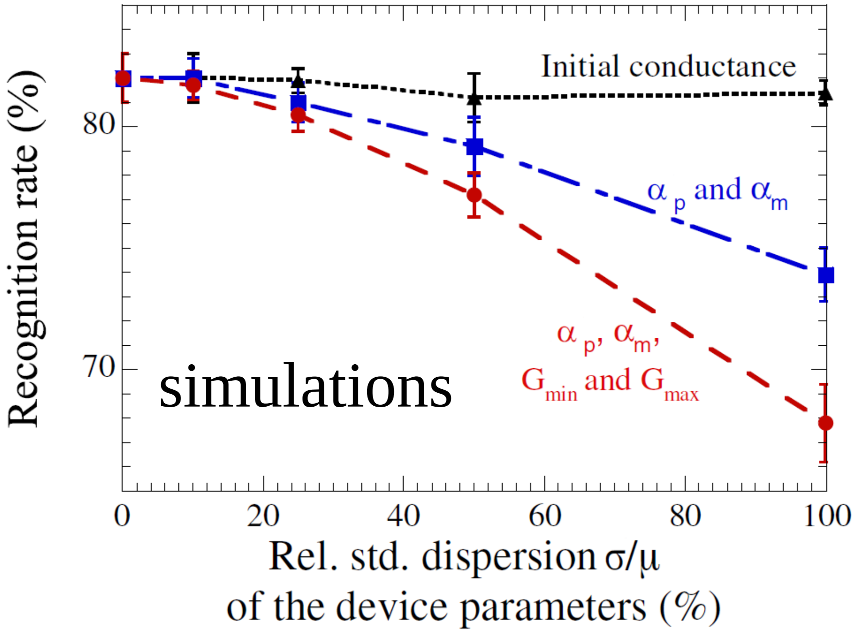
(Spin transfer oscillator, CNRS/Thales)

Not adapted to boolean computing but what about neuromorphic computing ?

Thanks to their plasticity, neural networks are highly resilient against nanodevice variability !



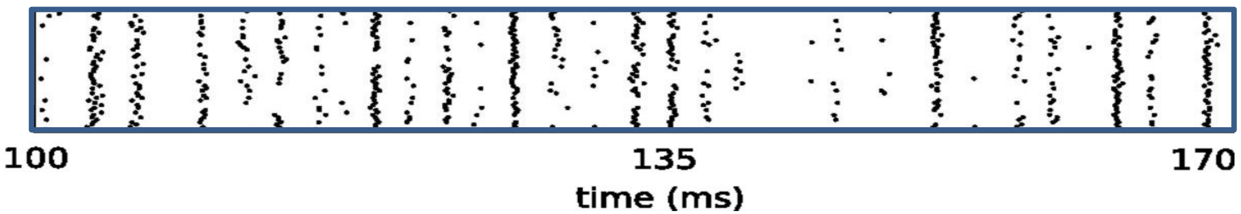


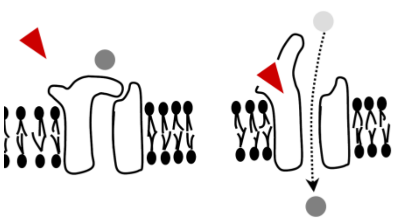


Damien Querlioz et al, *IEEE Trans. Nano.*, vol. 12, num. 3, p. 288 (2013)

Biological synapses and neurons are noisy: the brain seems to operate at the thermal limit to minimize its power consumption

Neural spikes in response to the same input recorded 50 times

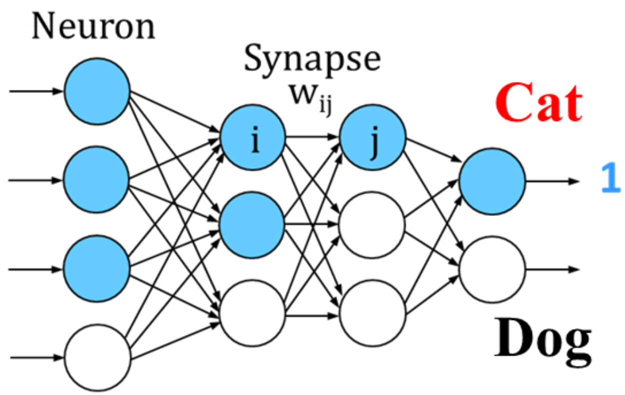


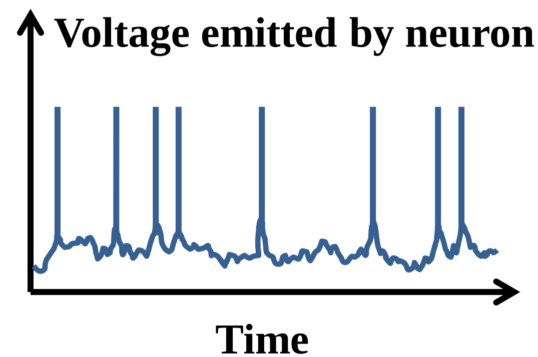


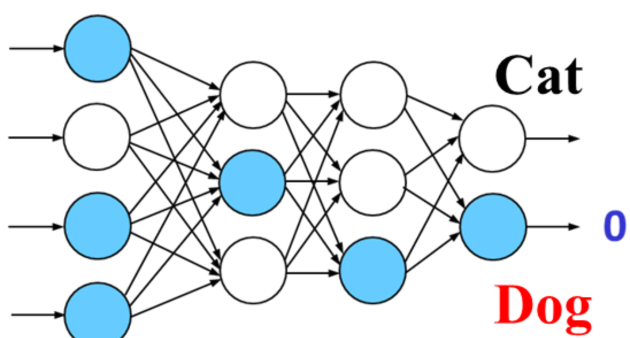
Computing at low power with stochastic components is possible

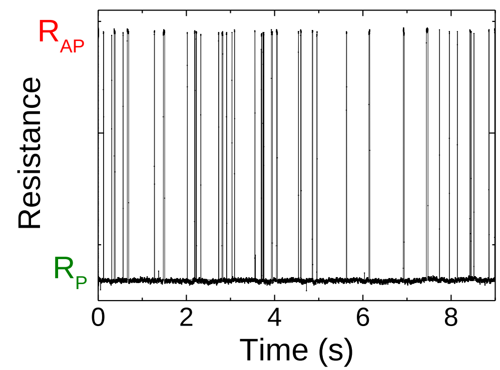
How can we realize reliable computations with stochastic devices ?

Stochastic computing: sampling multiple times, averaging









Can we even leverage noise for computing ?

Randomness can be useful
Example: stochastic resonance

Low noise **Optimal noise** High noise

Gammaitoni et al, Reviews Of Modern Physics, 70(1), 223–287 (1998)

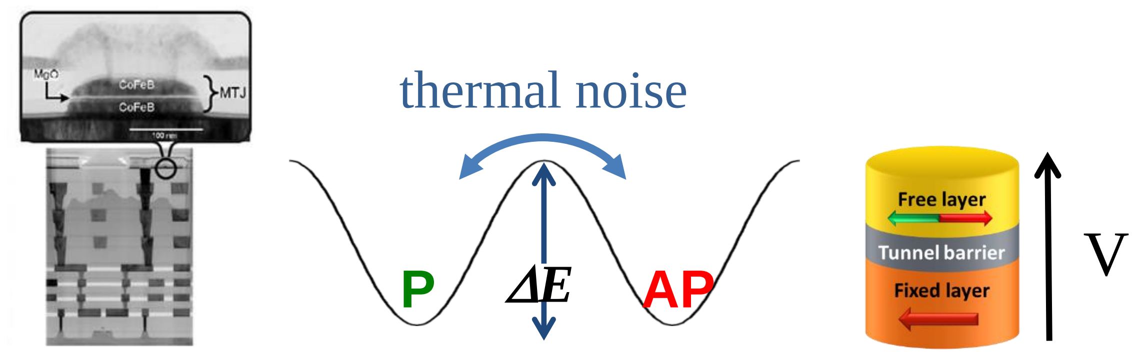
Superparamagnetic tunnel junctions:

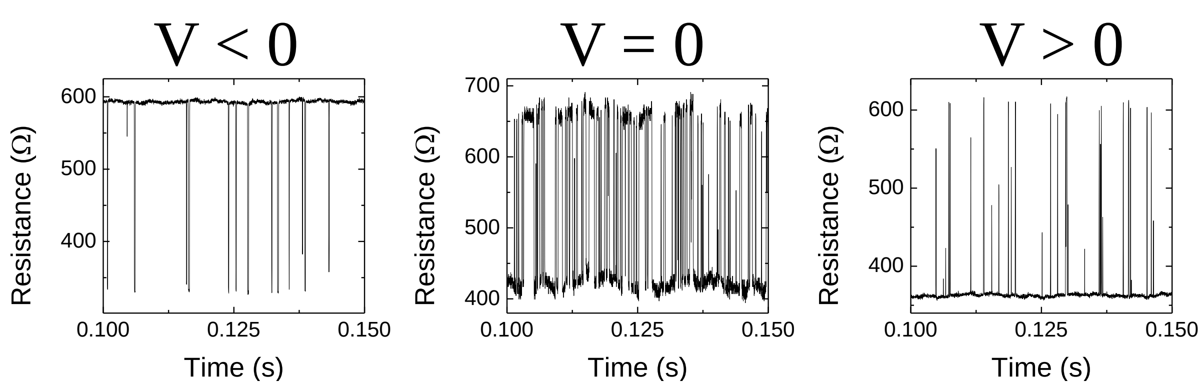
✓ **Can synchronize to small periodic signals**

Alice Mizrahi, Damien Querlioz, **JG** et al, Sci. Rep. 6:30535 (2016)

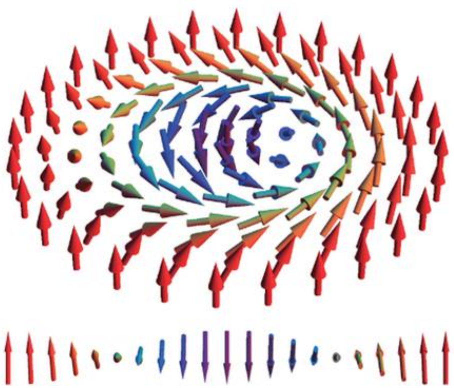
✓ **Are promising building blocks for Boltzmann machines**

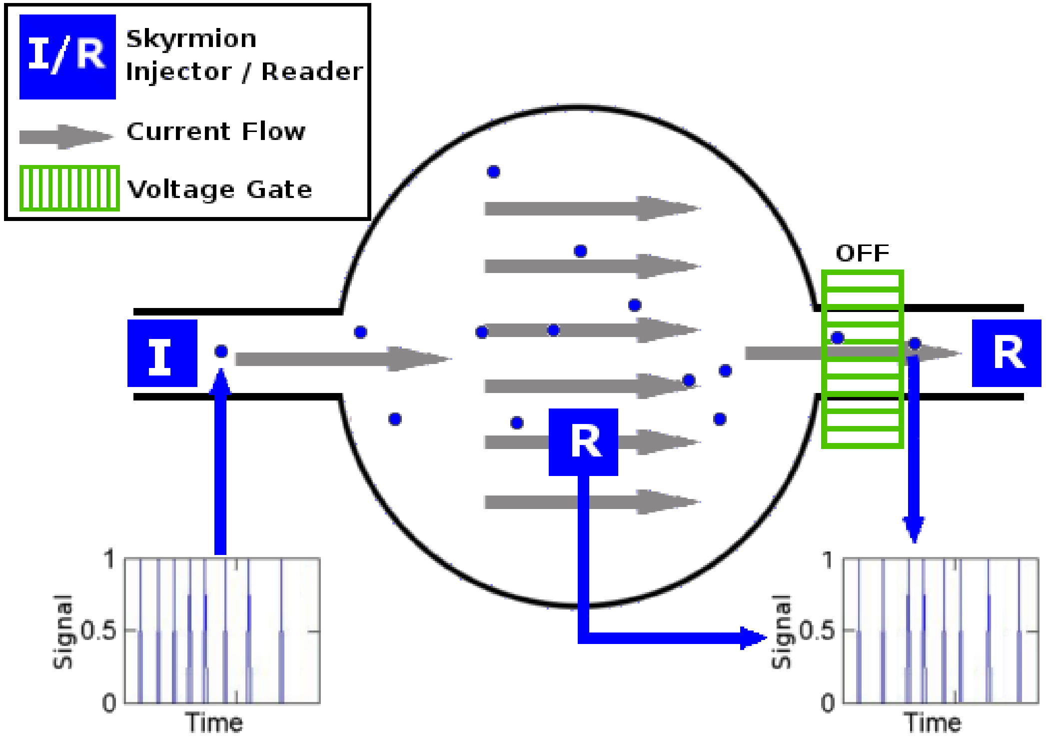
Camsari et al, arXiv:1610.00377





Thermal diffusion of Skyrmions can be used for reshuffling signals or as neurons for stochastic computing



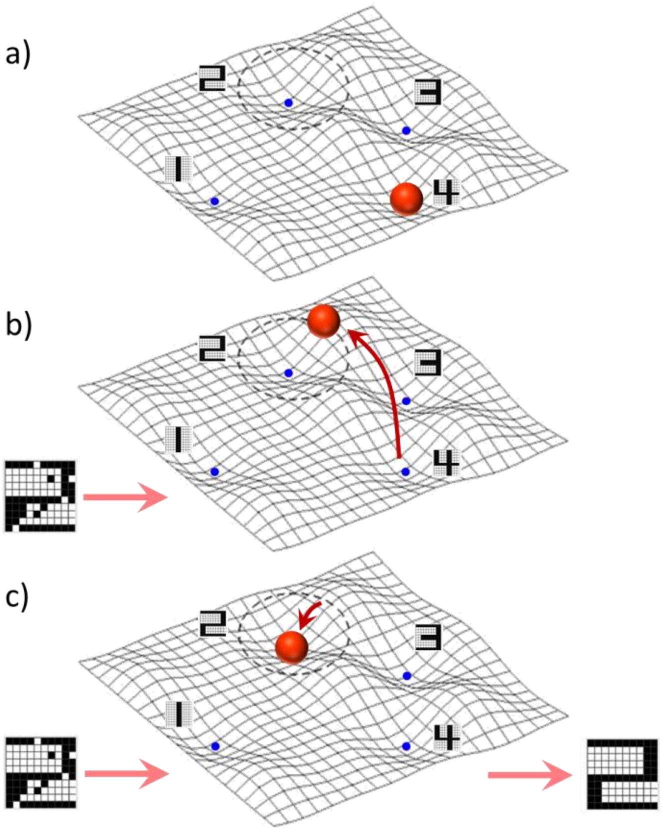


Daniele Pinna, **JG** et al, arXiv:1701.07750

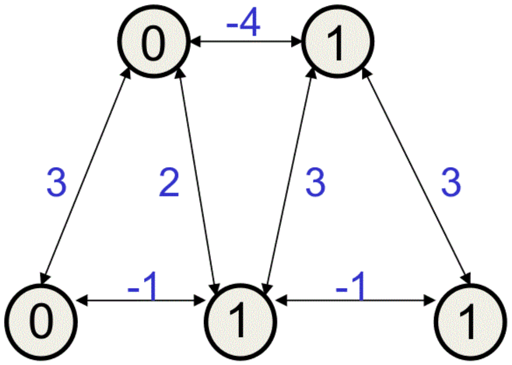
Nanodevices allow using physics for computing

Hopfield nets physics

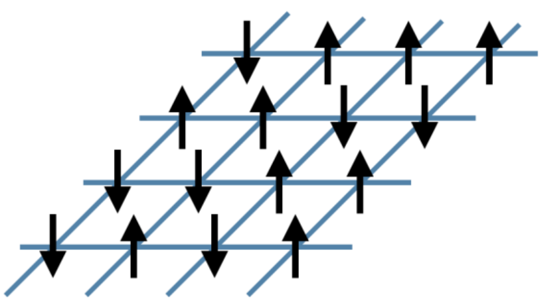
attractors = energy minima



Hopfield neural nets



Ising spin system inspiration



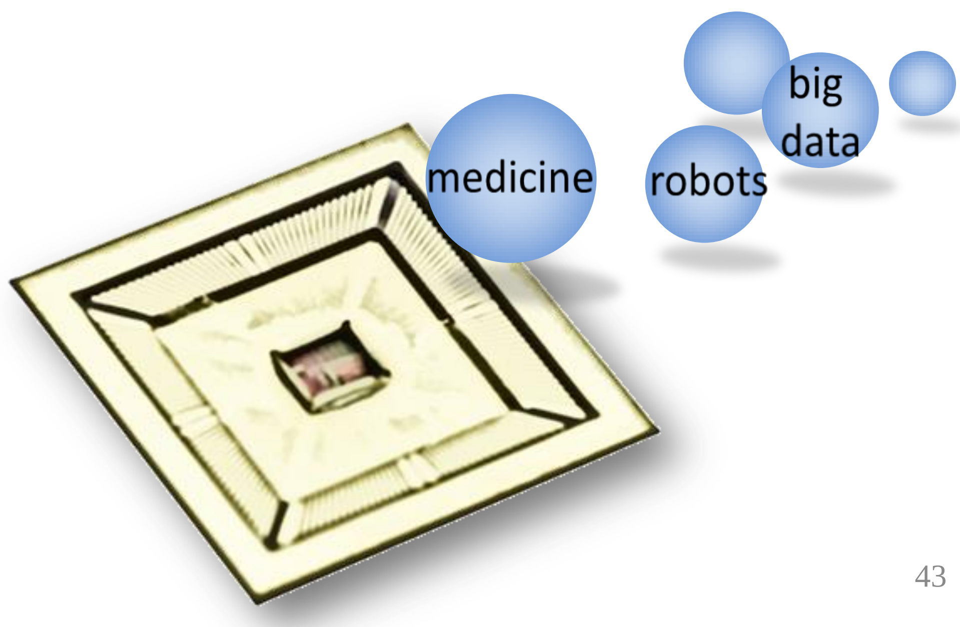
Back to the spin system !

J. Grollier, D. Querlioz, M. D. Stiles, *PIEEE* vol.104, n°10, 2024 (2016)

Conclusion

- Nanodevices open new perspectives for neuromorphic computing
- Working at the thermal limit → low power consumption
- Memristors can emulate synapses
- Nano-oscillators can emulate neurons
- The challenge is to assemble them densely together for computing

- If we succed: smart chips that can learn and adapt autonomously



Comparison between oscillators

	Lateral dimension	Energy / oscillation	Freq.	Power consumption	Operating time	Ability to sync.
CMOS neuron	> 30 µm	265 pJ	10 Hz	265 nW	Infinite	Yes (AER)
Accelerated CMOS neuron	≈ 10 µm	8.5 pJ	1-10 MHz	8-40 µW	Infinite	Yes (AER)
CMOS ring oscillator	6 µm	6 nJ	200 KHz	1.2 nW	Infinite	?
CMOS ring oscillator	6 µm	30 fJ	1.5 GHz	50 µW	Infinite	?
CMOS ring oscillator	≈ 300 µm	≈ 1 pJ	8-16 GHz	20 mW	Infinite	Yes
Electro-optical oscillator	> 1 cm	1 nJ	≈ GHz	≈ 1 W (laser)	Infinite	Yes
Memristive oscillator	> 100 nm	> 33 pJ	300 MHz	> 150 µW	Hours	Yes
Our spin-torque oscillator	300 nm	3 pJ	300 MHz	1 mW	> Months	Yes
Next gen spin-torque oscillator	10 nm	100 aJ	10 GHz	1 µW	> Months	Yes