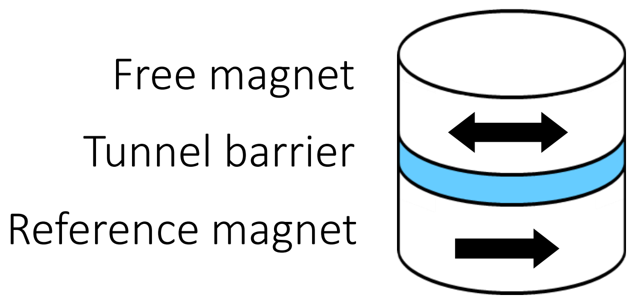
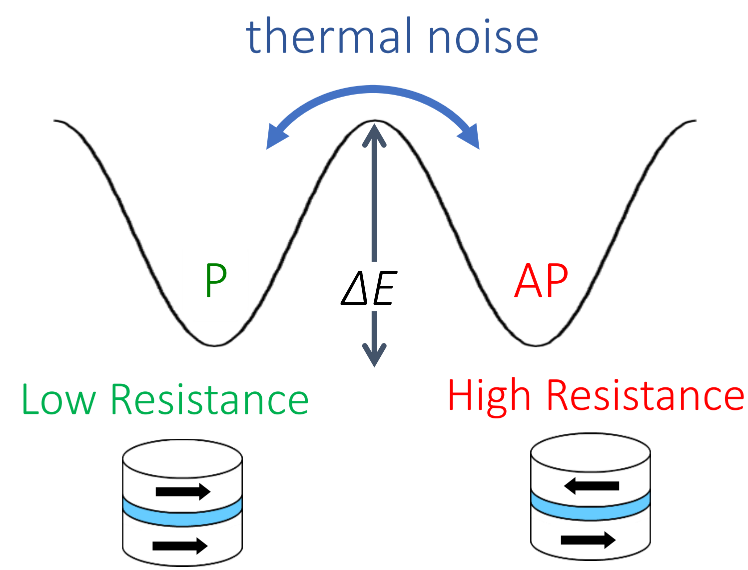
Unconventional computing with stochastic magnetic tunnel junctions

Alice Mizrahi, Tifenn Hirtzlin, Matthew Daniels, Nicolas Locatelli, Akio Fukushima, Hitoshi Kubota, Shinji Yuasa, Mark Stiles, Julie Grollier, Damien Querlioz

Lowering the energy consumption reduces the stability





Small size → Low ΔE

Low write energy → Low ΔE

Low ΔE → Unstable

Different energy barriers for different applications

Energy barrier

$60\ k_B T$ years — Memory for storage (MRAM)

$15\ k_B T$ ms — Synaptic weight for continuously learning neural net

$6\ k_B T$ μs — Stochastic regime

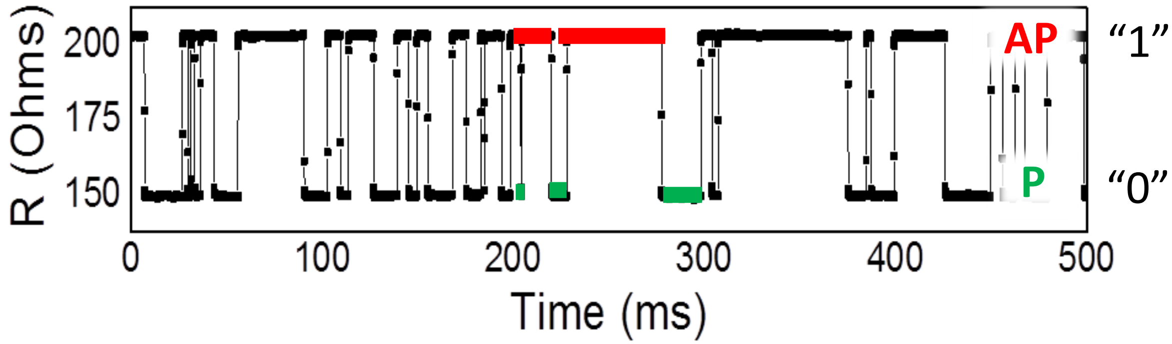
Stochastic bitstream generator

Stochastic spiking neuron

Stochastic oscillator

NIST UNIVERSITY OF MARYLAND

A stochastic oscillator powered by thermal noise

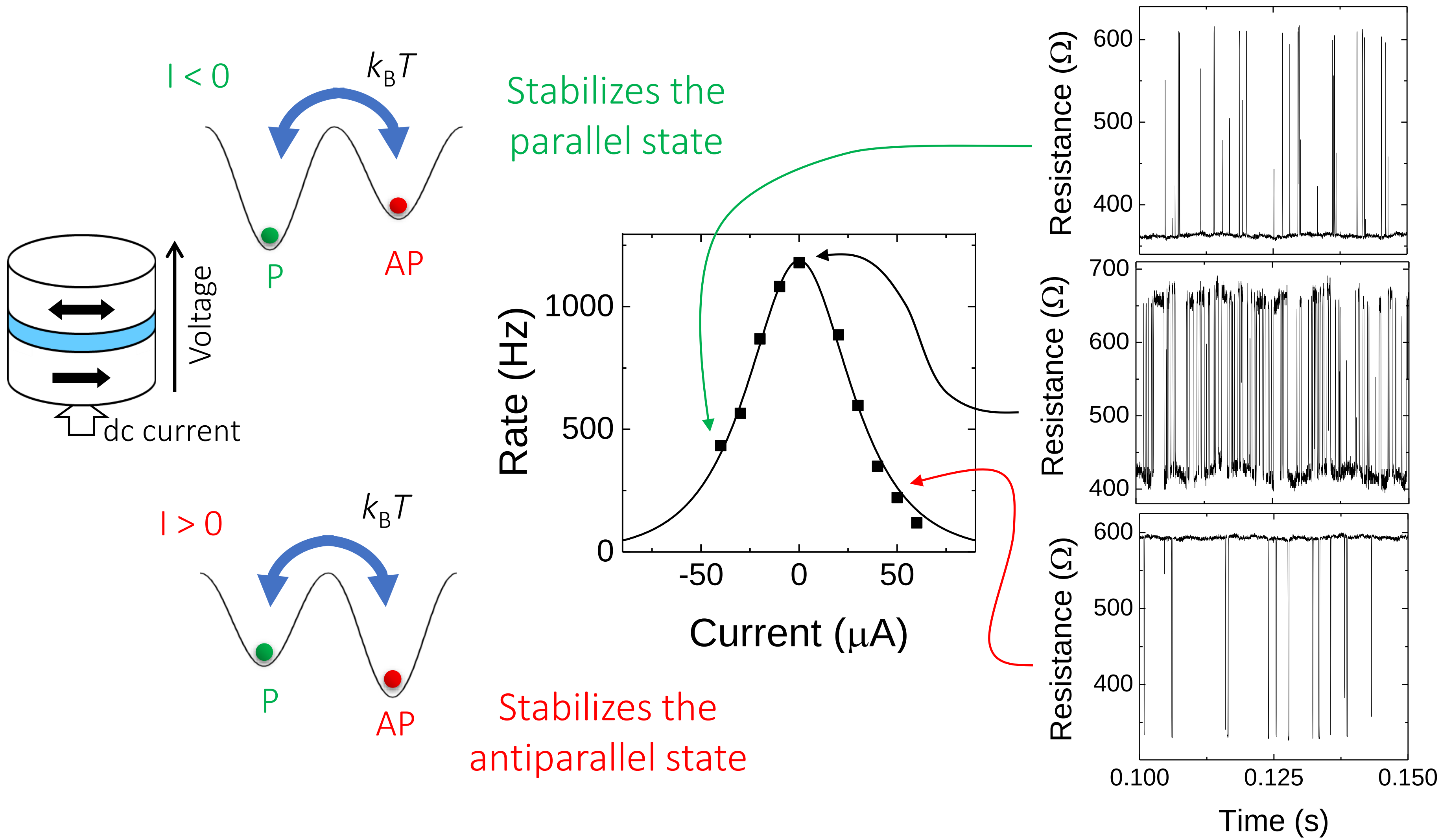


No external energy source

Two-state signal easy to convert to digital signal

Stochasticity well understood and controlled

Spin torque controls the stochasticity

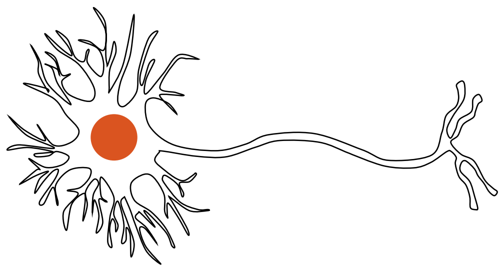


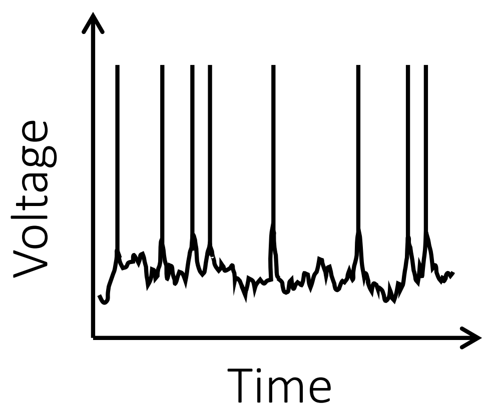
Stochastic analog to digital conversion

1. Population coding with stochastic spiking neurons

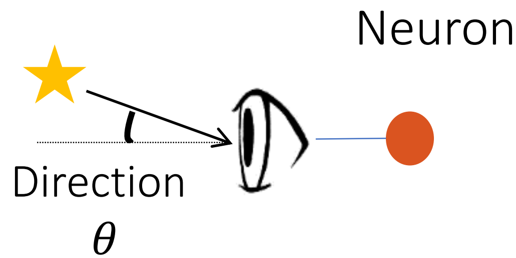
2. Noise-induced synchronization of a stochastic oscillator

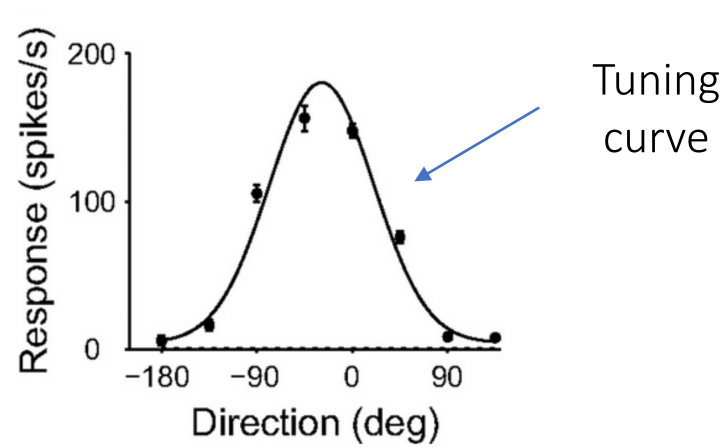
Encoding information in the spike rate of a stochastic neuron





Rate coding
- ✓ Robust to errors
- ✓ Fast approximate results

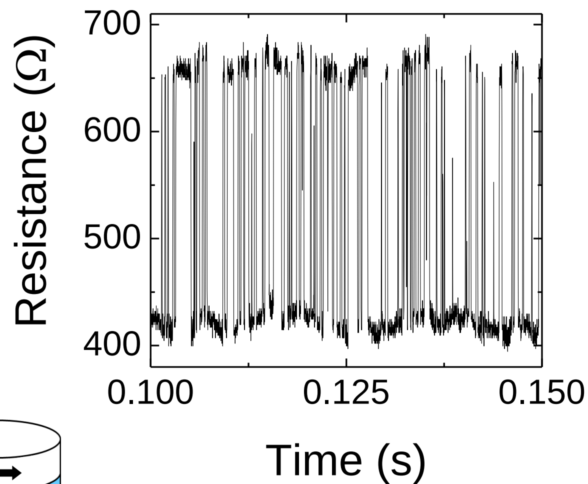


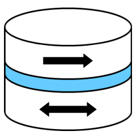


Kumano et al., J. Neurophysiology 2010

Rate varies with value of stimulus

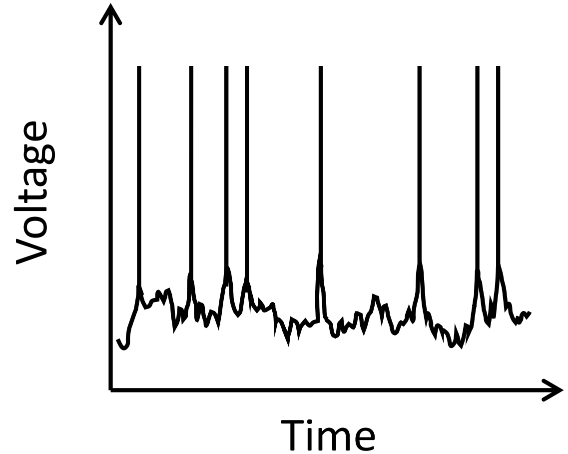
The stochastic magnetic tunnel junction emulates a stochastic spiking neuron

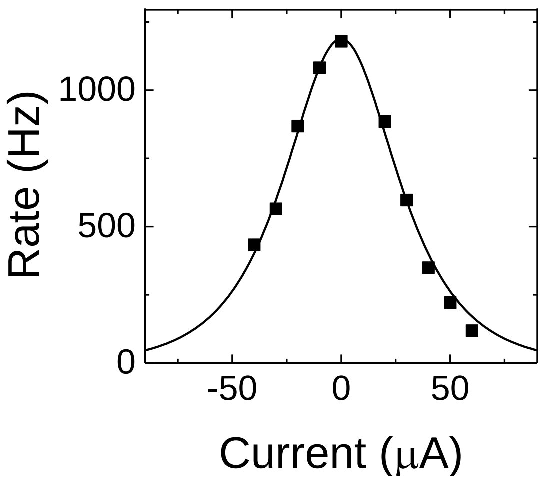


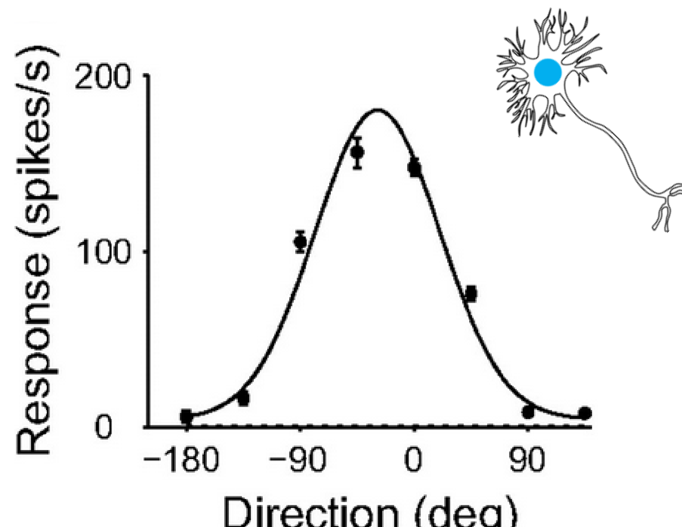


Switches ↔ spikes

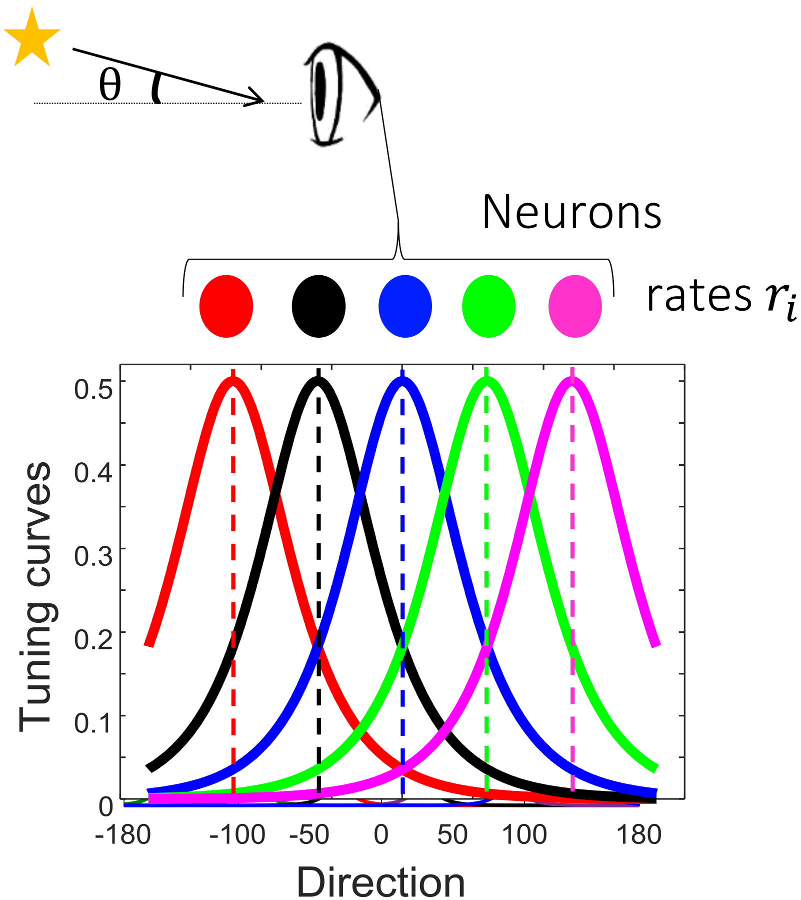
Stimulus = current



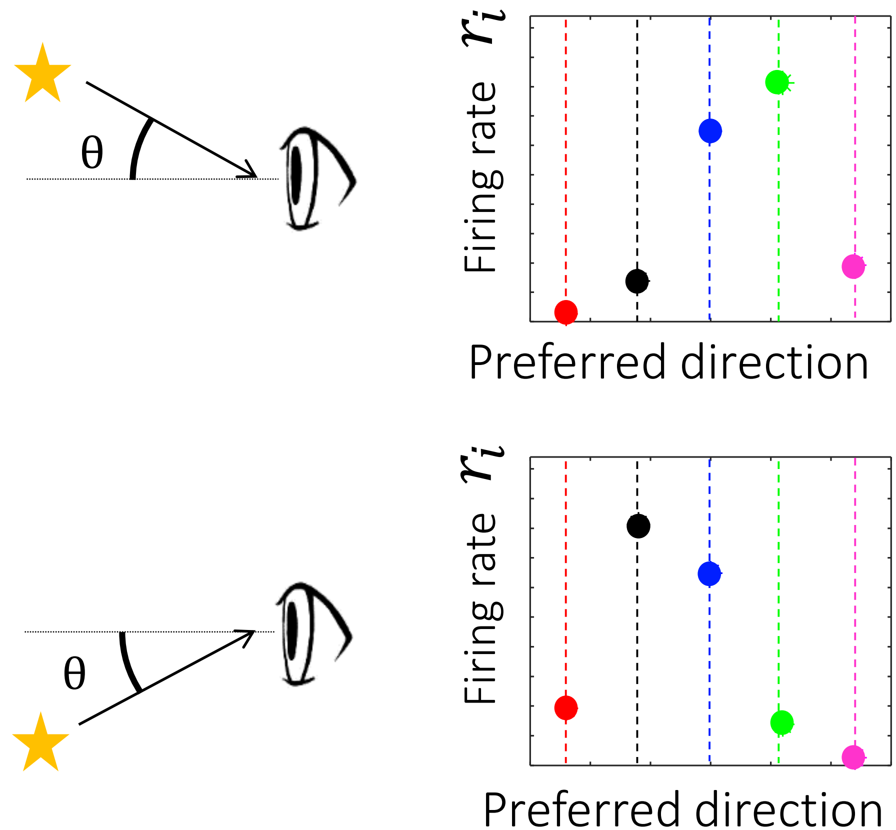




Robust coding of information by a population of neurons



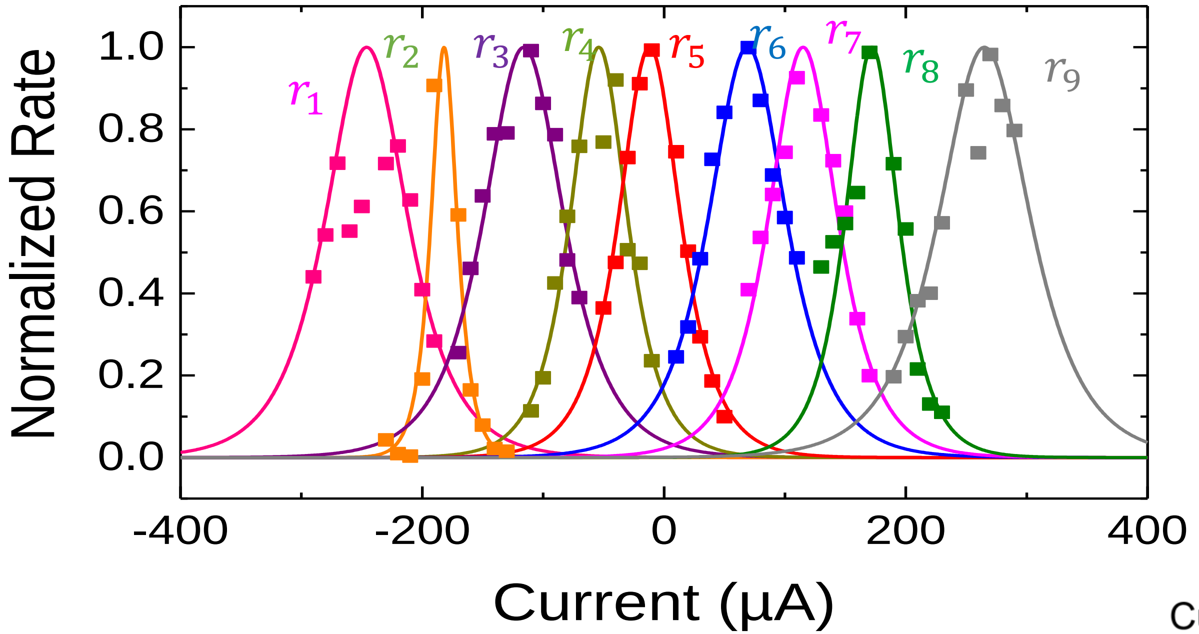
Each neuron processes a range of inputs (directions)



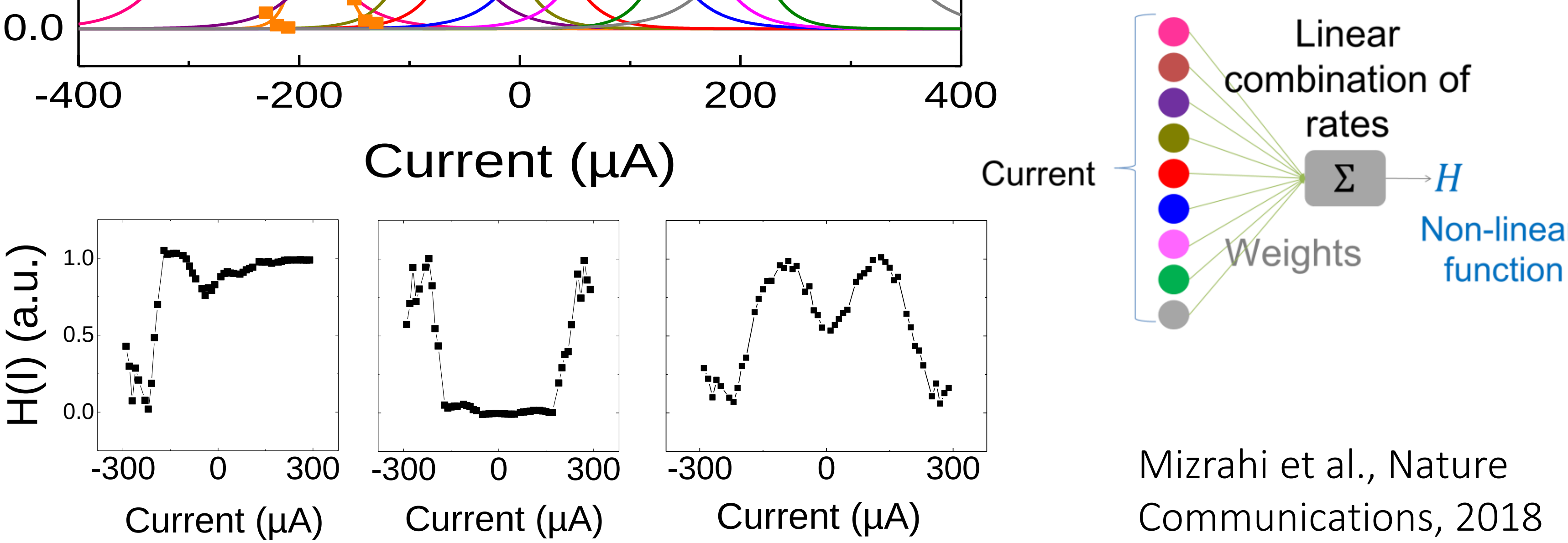
The direction can be inferred from the rates of the neurons

Constructing non-linear transformation with a population of artificial neurons

Shift the tuning curves $I_{\text{eff}}{}^{i} = I - I_{\text{bias}}{}^{i}$



$$H(I) = \sum_{i=1}^{9} w_i r_i(I_{\text{eff}}{}^{i})$$



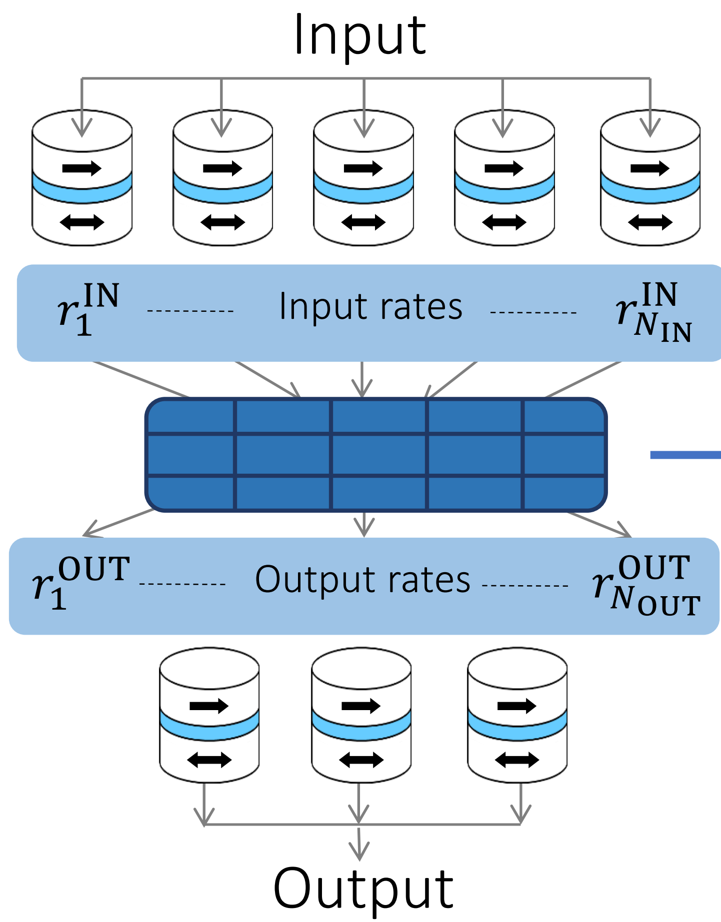
Mizrahi et al., Nature Communications, 2018

An artificial neural net with magnetic tunnel junctions

Stochastic junctions as neurons (N_{IN})

Stable junctions as binary encoded synaptic weights (N_{IN} x N_{OUT})

Stochastic junctions as neurons (N_{OUT})



Input value *I*

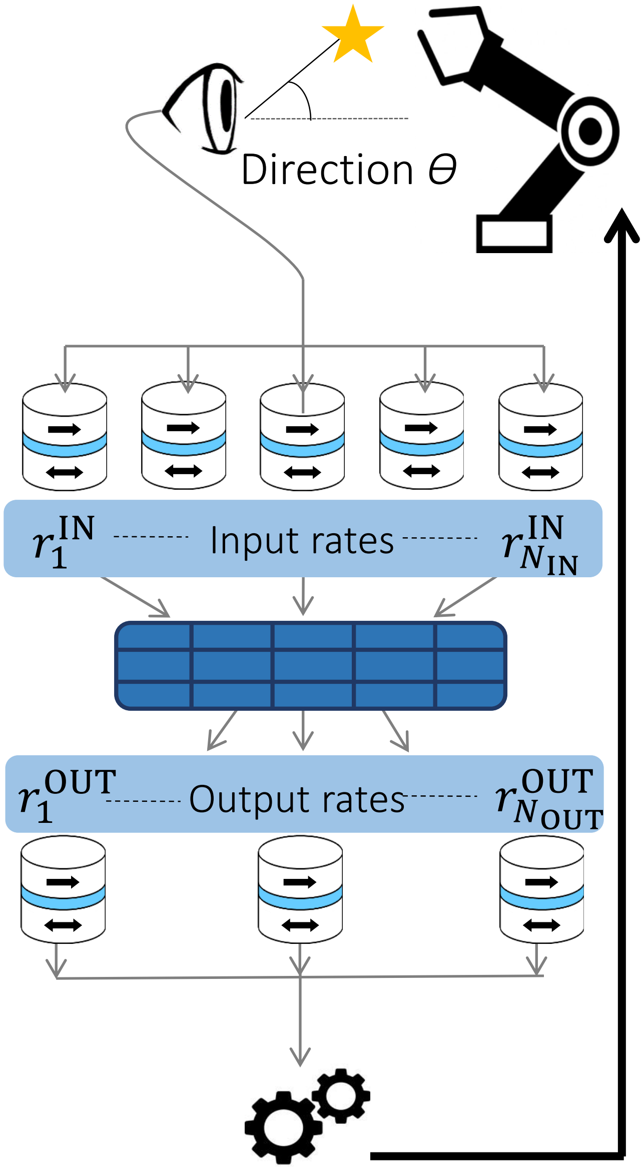
Transformation *F*

Output value *O = F(I)*

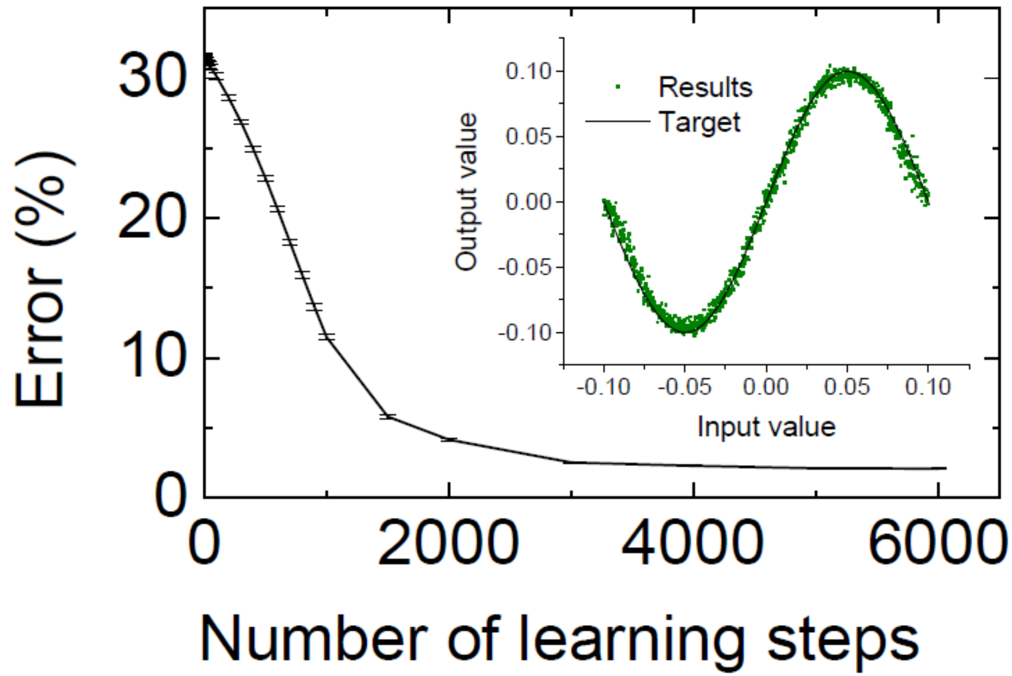
$$r_j^{\mathrm{OUT}} = \sum_{i=1}^{N} w_{ij} r_i^{\mathrm{IN}}$$

"Transfer of coded information from sensory to motor networks"
Salinas et Abbott, J. Neuroscience, 1995

The system is capable of learning a transformation



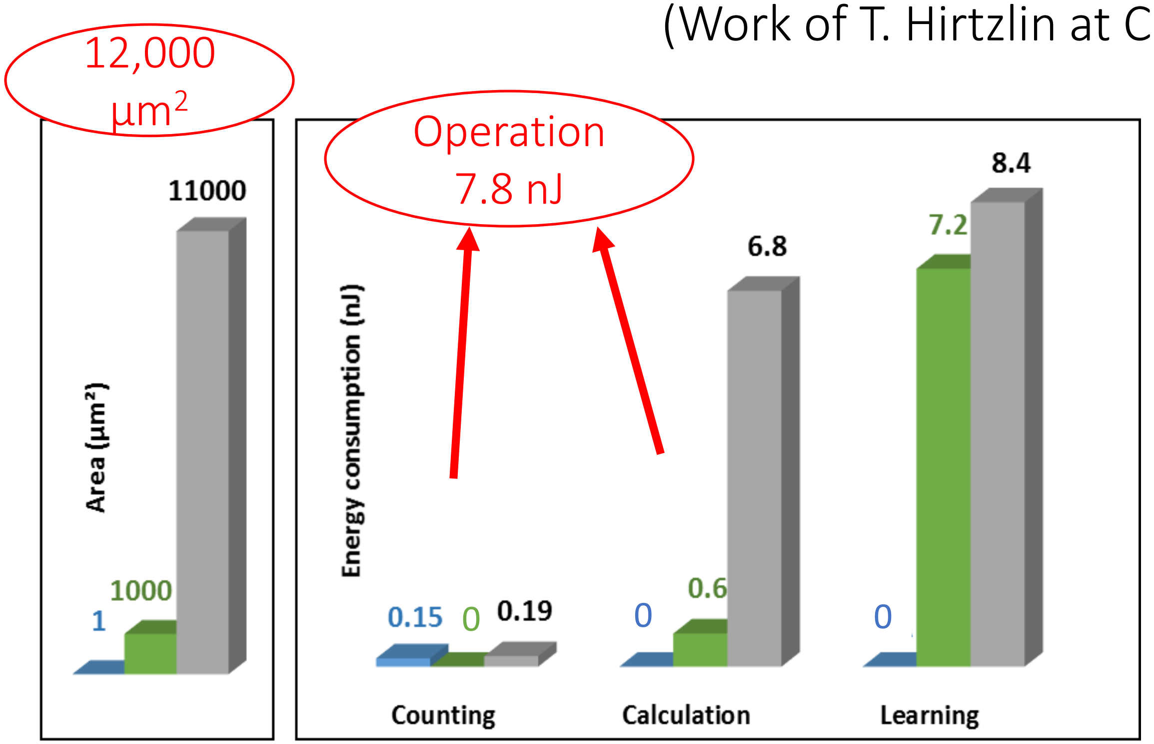
Learning rule:
Catch → do nothing
Miss → modify the weights



Simulations with experimentally verified model
$\Delta E_{\mathrm{neuron}} = 6\ k_B T$

The system consumes less energy and area than CMOS-only implementations

Area and energy consumption of our system
(Work of T. Hirtzlin at C2N)



Neurons

Synaptic Weights

CMOS overhead:

- Switch detection & count
- Calculation (rate and weights)
- Learning rule

Key asset = stochastic analog to digital conversion

CMOS-only neurons:
Area > 20,000 µm²
Energy of operation > 20 nJ

Using unreliable synapses lowers the power consumption

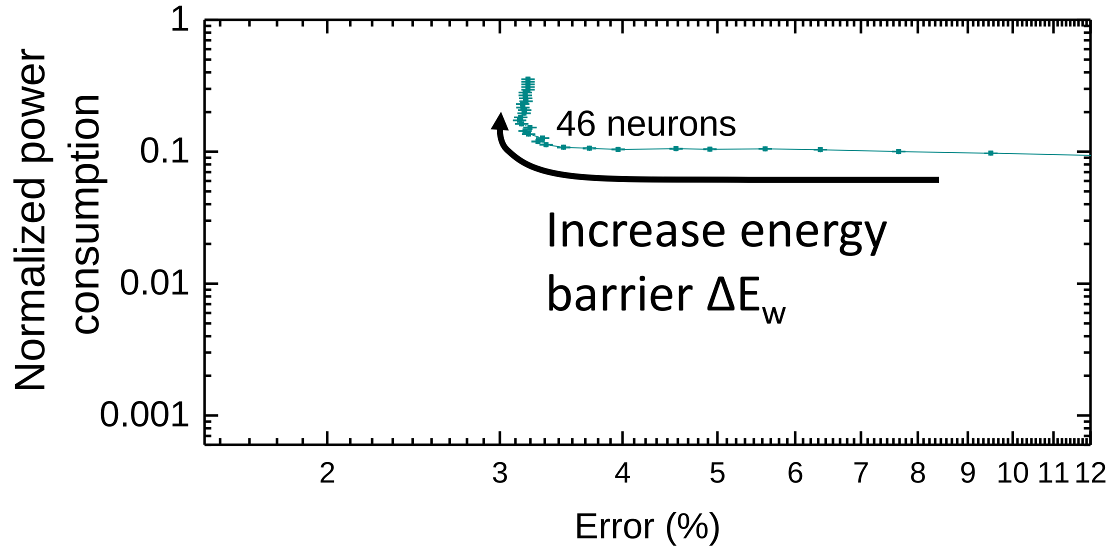
Write current

$$I \propto \Delta E_w$$

(Sato et al., APL 2014)

Write power

$$Power \propto {\Delta E_w}^2$$



Using unreliable synapses lowers the power consumption

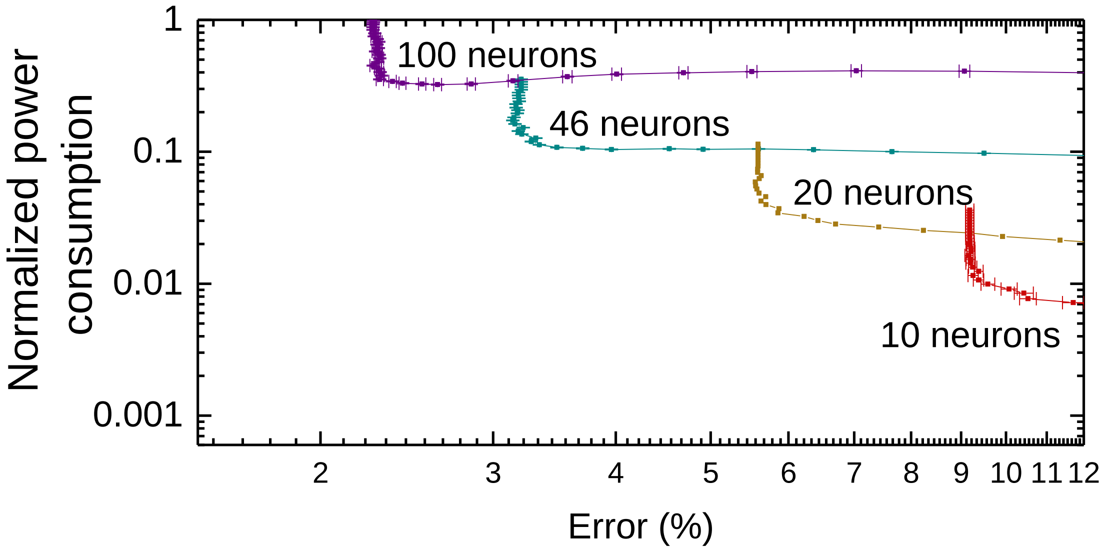
Write current

$$I \propto \Delta E_w$$

(Sato et al., APL 2014)

Write power

$$Power \propto \Delta E_w{}^2$$



Using unreliable synapses lowers the power consumption

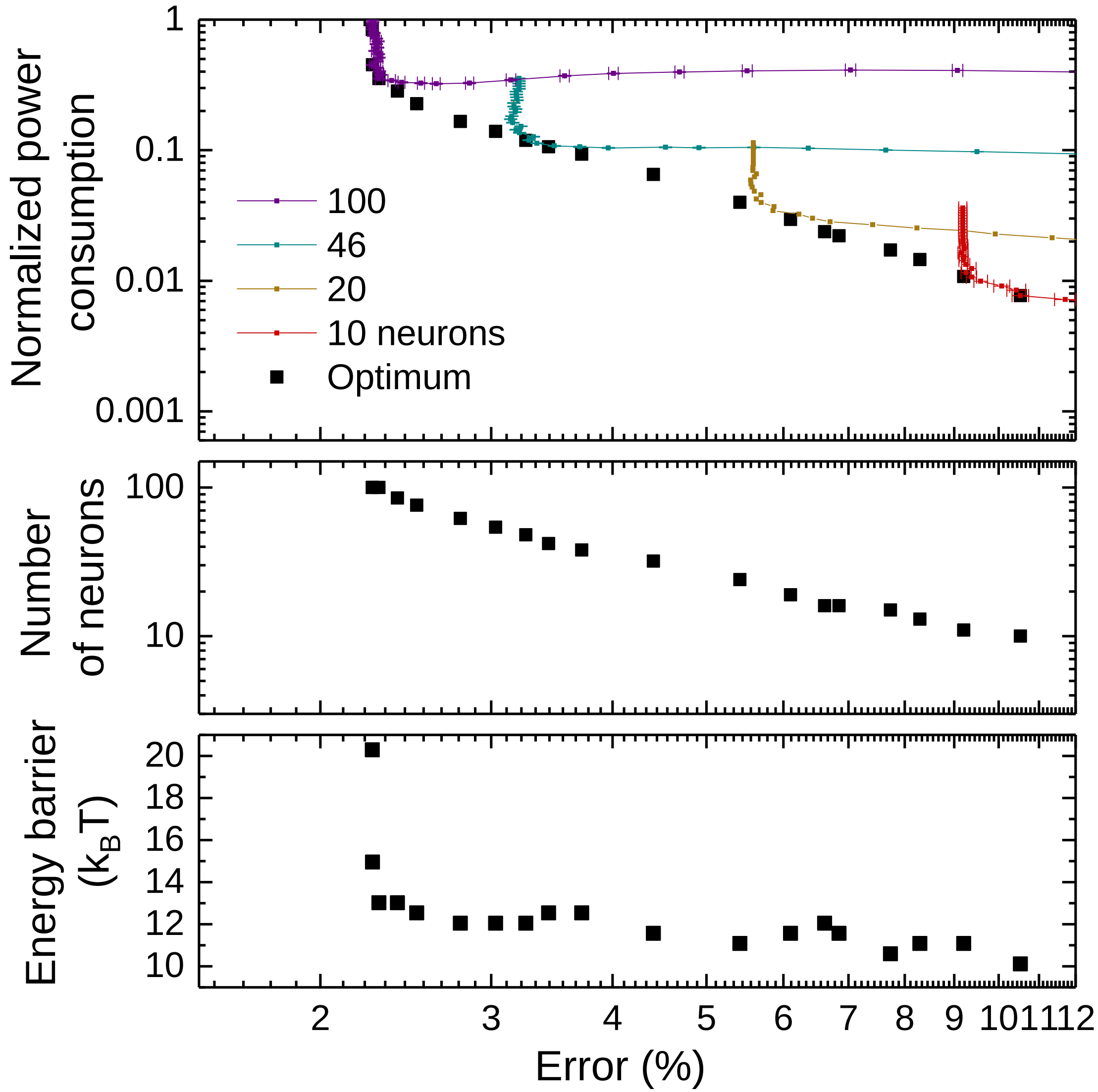
Write current

$$I \propto \Delta E_w$$

(Sato et al., APL 2014)

Write power

$$Power \propto \Delta E_w{}^2$$



Using unreliable synapses lowers the power consumption

Write current

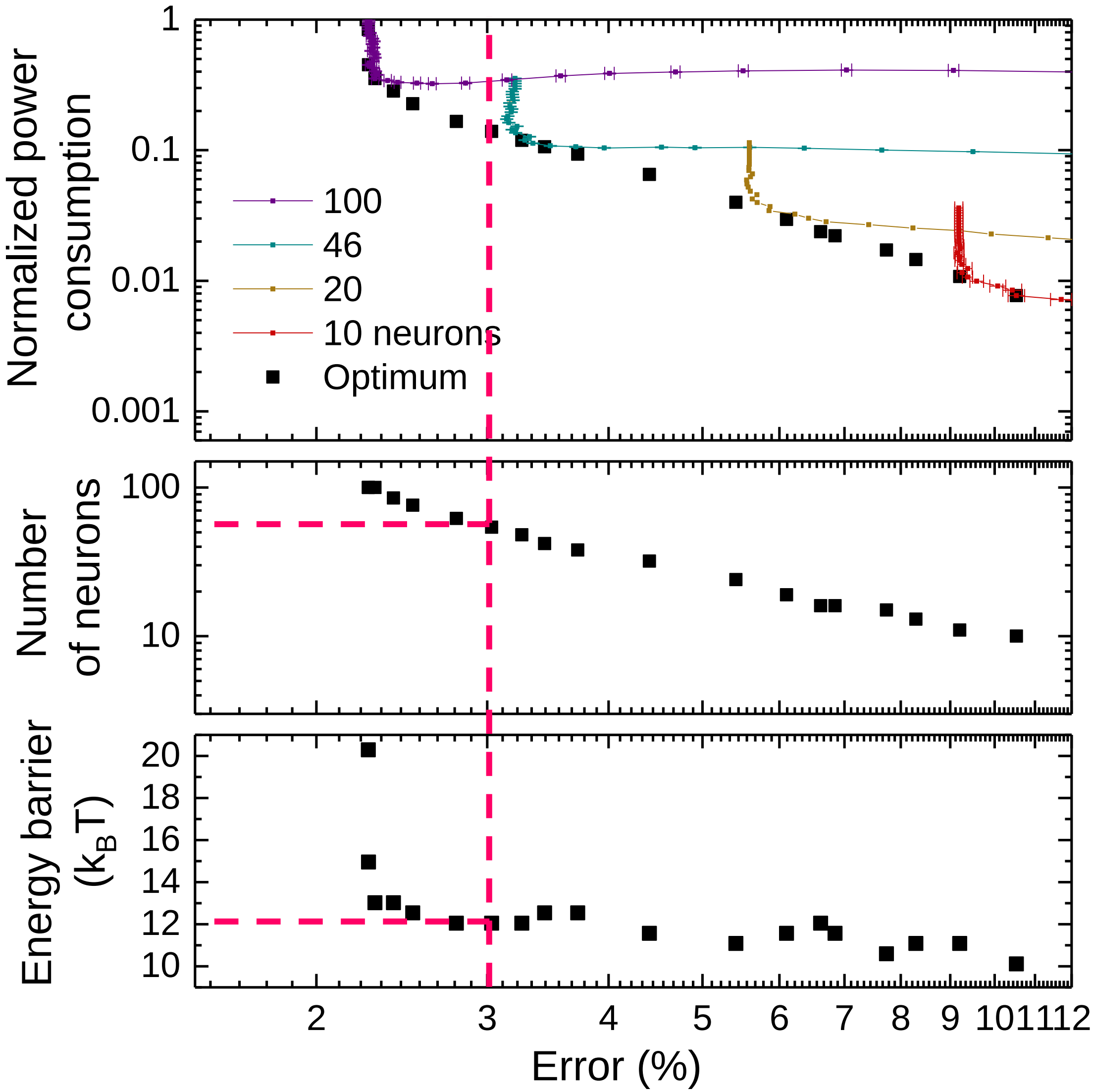
$$I \propto \Delta E_w$$

(Sato et al., APL 2014)

Write power

$$Power \propto {\Delta E_w}^2$$

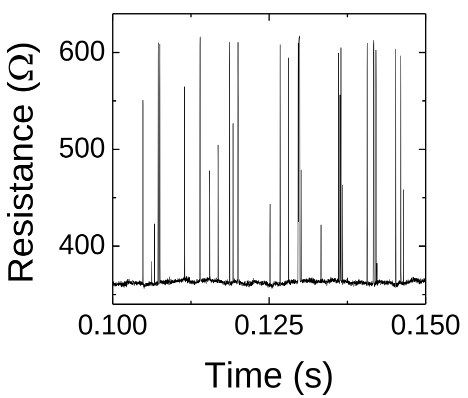
Example:
For 3% precision
→ 54 neurons in each population (2916 weights)
→ $\Delta E_w = 12\ k_B T$



A low energy continuously learning neural net with magnetic tunnel junctions

Fully stochastic magnetic tunnel junction as neuron + Continuous learning + Slightly stochastic magnetic tunnel junction as synapse

=



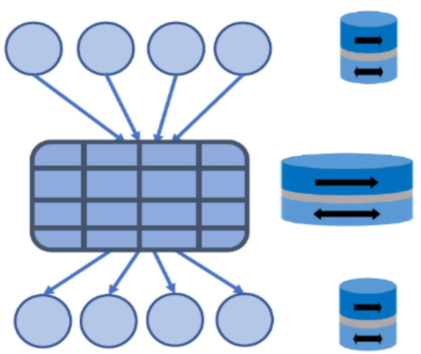
Artificial neural network

Low energy

Capable of learning

Resilient to variability & unreliability

Adaptable to changes



Key application: smart sensors

Mizrahi et al., Nature Communications, 2018

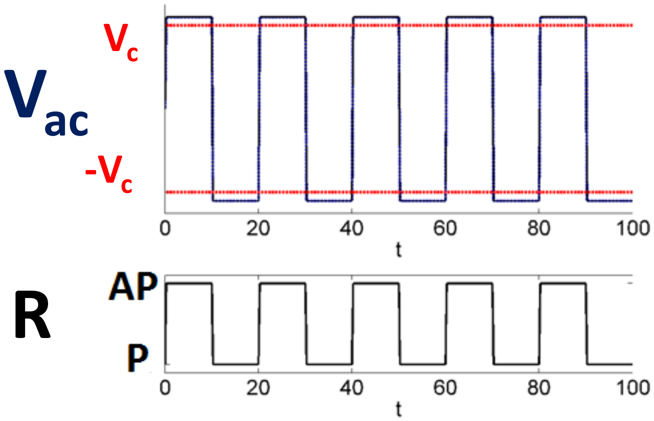
Mizrahi et al., J. Applied Physics, 2018

1. Population coding with stochastic spiking neurons

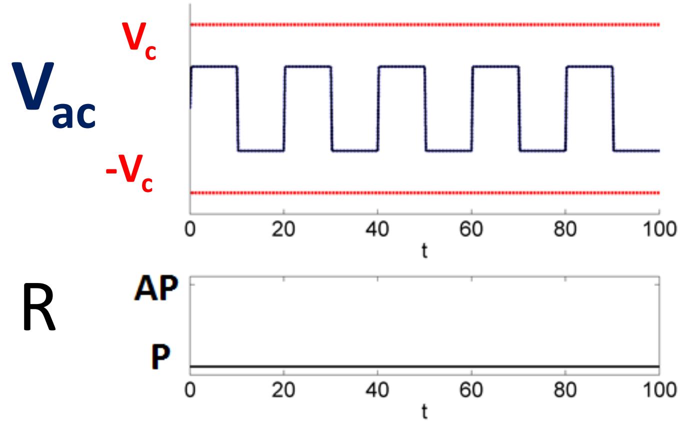
2. Noise-induced synchronization of a stochastic oscillator

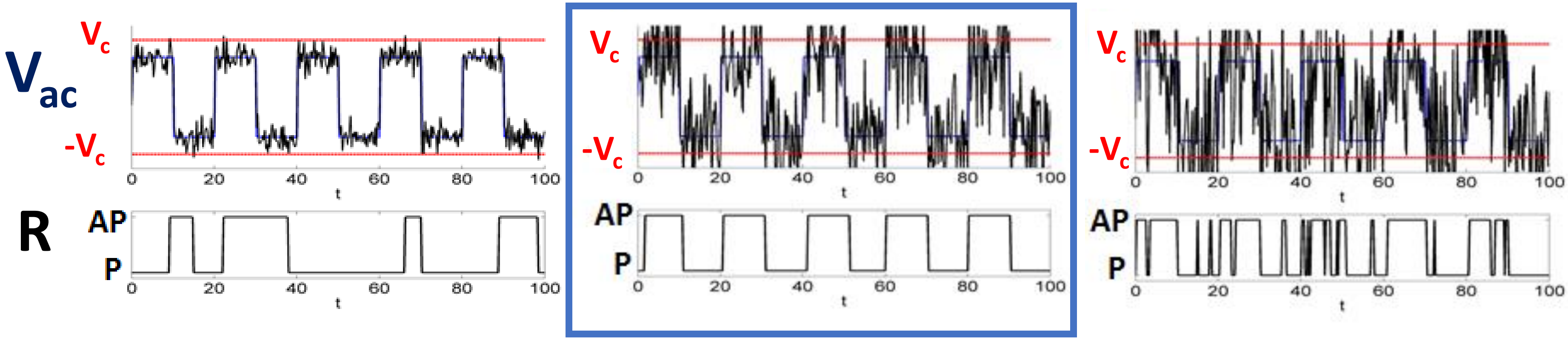
Noise can induce low-power synchronization

Deterministic $V_{ac} > V_c$



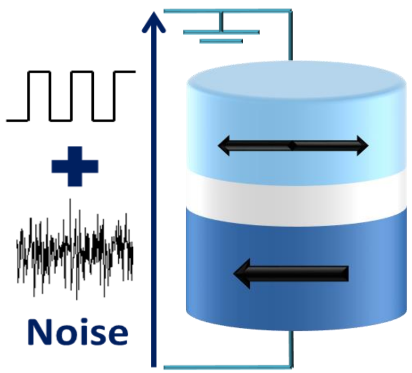
Subthreshold $V_{ac} < V_c$





Optimal noise range

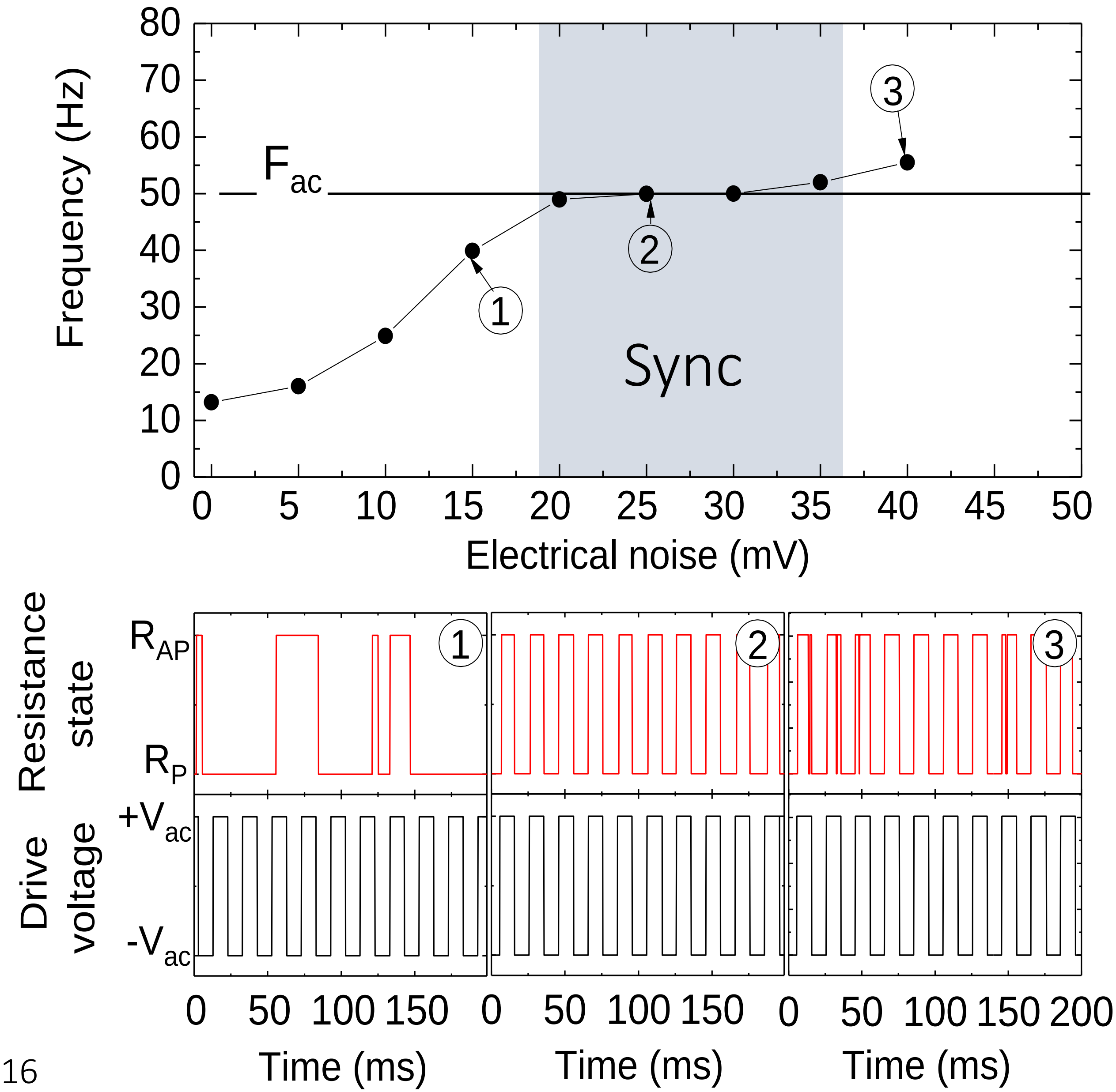
Noise controls frequency and phase locking



V_{ac} = 63 mV while
V_c = 235 mV @0K

$\Delta E = 22.5\ k_BT$
Natural frequency ≈ 0.1 Hz

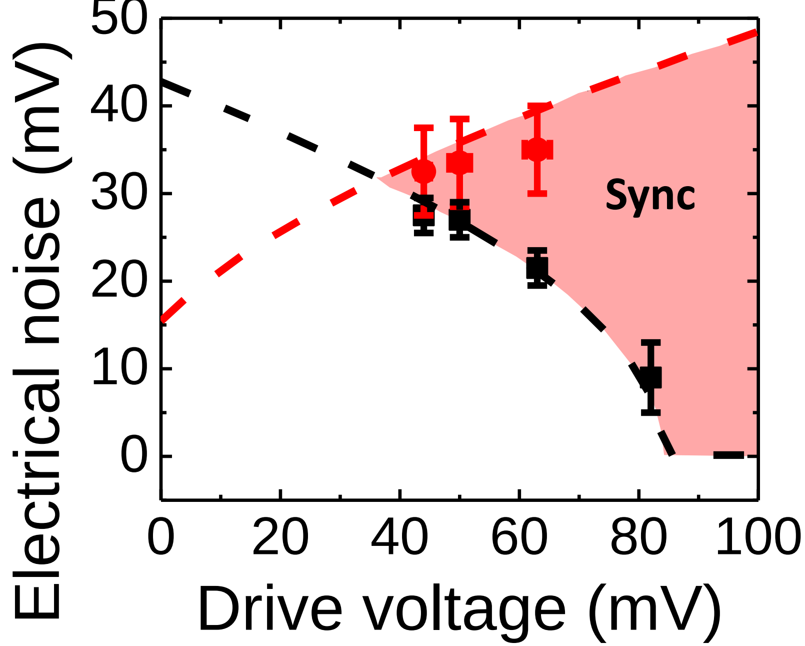
Thermal noise (room temperature)
+
Electrical noise (white Gaussian)



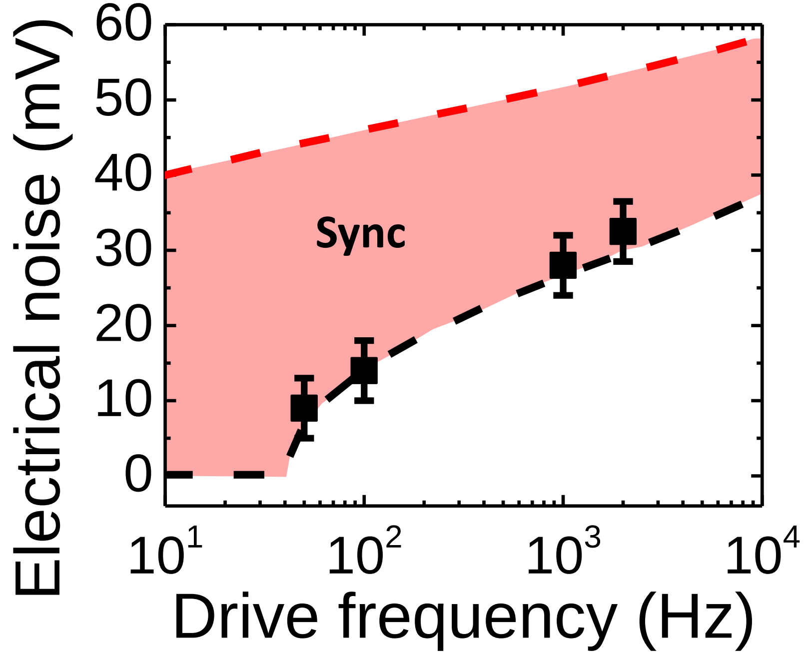
Mizrahi et al., Scientific Reports, 2016

Synchronization is possible at broad ranges of amplitudes and frequencies

F_{ac} = 50 Hz



V_{ac} = 82 mV



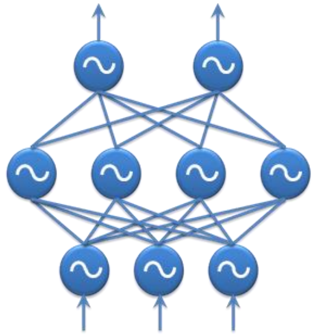
Boundaries of synchronization

Model Experiments

Natural frequency ≈ 0.1Hz

Low-energy synchronization for computing

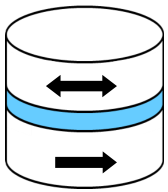
How to use it for
associative memory,
pattern classification etc.?



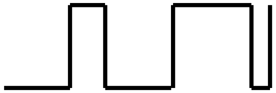
Need to reinvent computing schemes for
several coupled stochastic oscillators!

Mizrahi et al., Scientific Reports, 2016 ; Mizrahi et al., IEEE Transactions on Magnetism, 2015

The stochastic magnetic tunnel junction is a promising building block for low-power computing



Endurance
Reliability
CMOS compatibility
New handles (spin-orbit etc.)



Powered by thermal noise
Driven by small signals
Analog to digital conversion
Stochasticity understood and controlled

Stochastic spiking neuron

Stochastic oscillator for synchronization

NIST UNIVERSITY OF MARYLAND