CMMRC

Frontiers of Memristive Materials for Neuromorphic Processing

TEXAS A&M UNIVERSITY
Engineering

Overview of Memristors & Neuromorphic Processing

R. Stanley Williams
Professor of Electrical and Computer Engineering
Hewlett Packard Enterprise Chair

February 28, 2020

How to get from memristors to neuromorphic computing?

What qualifies as a neuromorphic process?

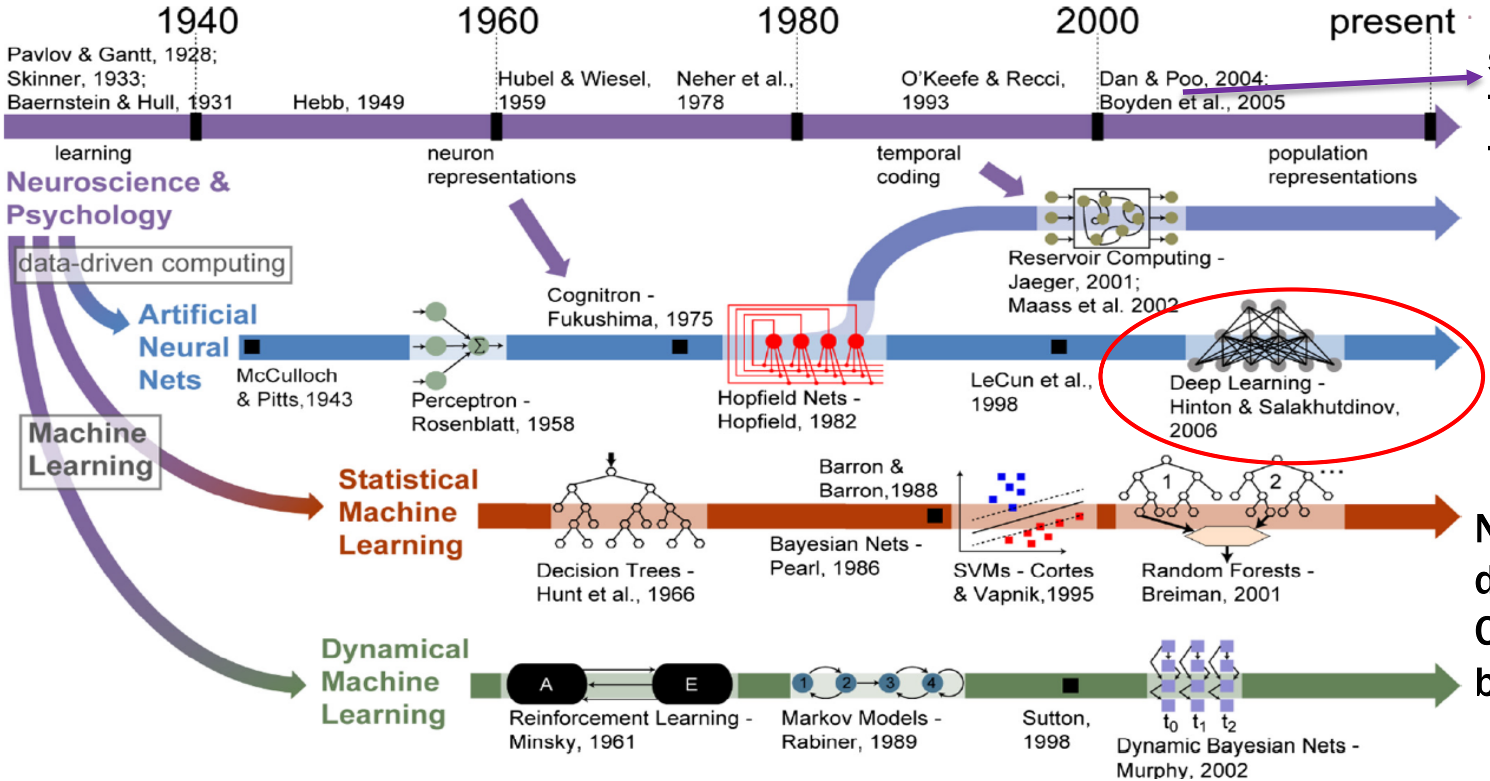
A memristor is a nonlinear dynamical system

Two types of memristors: nonvolatile and ‘locally active’

Examples of mem-elements and computation using them

neuroscience ↔ ‘neuromorphic computing’

TEXAS A&M UNIVERSITY Engineering



spiking, SyNAPSE, True North, Loihi, Tianjic

Enabled by Linear Algebra Accelerators: GPU and TPU

Nonlinear dynamical systems, Chaotic attractor based computing

From “A historical survey of algorithms and hardware architectures for neural-inspired and neuromorphic computing applications.” C. D. James, et al. *Biologically Inspired Cognitive Architectures* (2017)

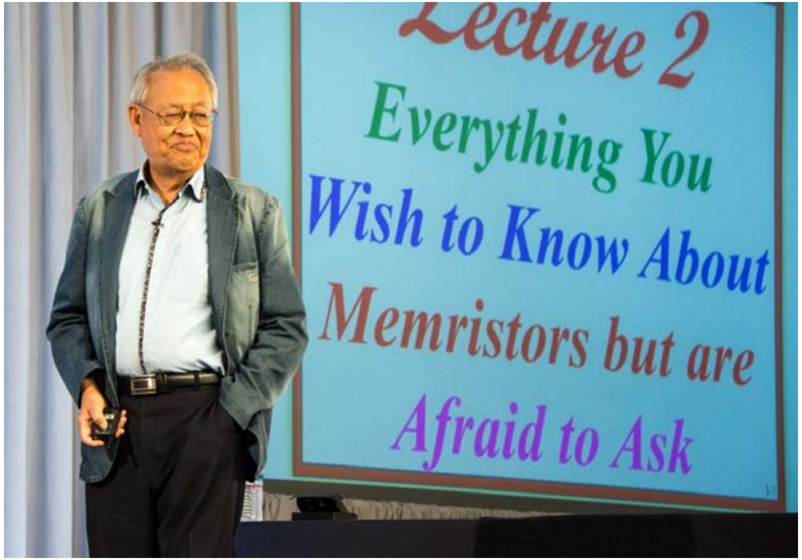
The Chua Lectures: A 12-Part Series with HP Labs

TEXAS A&M UNIVERSITY Engineering

From Memristors and Cellular Nonlinear Networks to the Edge of Chaos
https://www.youtube.com/playlist?list=PLtS6YX0YOX4eAQ6IrOZSta3xjRXzpcXyi
or enter "The Chua Lectures" into your favorite browser

'Linearize then analyze' is not valid for understanding nanodevices or neurons – a nonlinear dynamical theory of electronic circuits is needed, and was developed 50 years ago by Leon Chua.

The memristor is one of many nonlinear dynamical circuit elements with memory!



Mathematical (Axiomatic) <u>Definition</u> of a Memristor:

$$v = R(w, i)i$$ Quasi-static conduction eq. – Ohm’s Law

$$\frac{dw}{dt} = f(w, i)$$ Dynamical eq. – evolution of state under stimulus

L. O. Chua, “Memristor - the missing circuit element,” IEEE Trans. Circuit Theory 18, 507–519 (1971).
L. O. Chua and S. M. Kang, "Memristive devices and systems," *Proc. IEEE,* 64 (2), 209-23 (1976). –

w is the <u>state</u> <u>variable</u> (or variables)

Instead of a disembodied mathematical entity, *w* should describe real physical properties of the circuit element

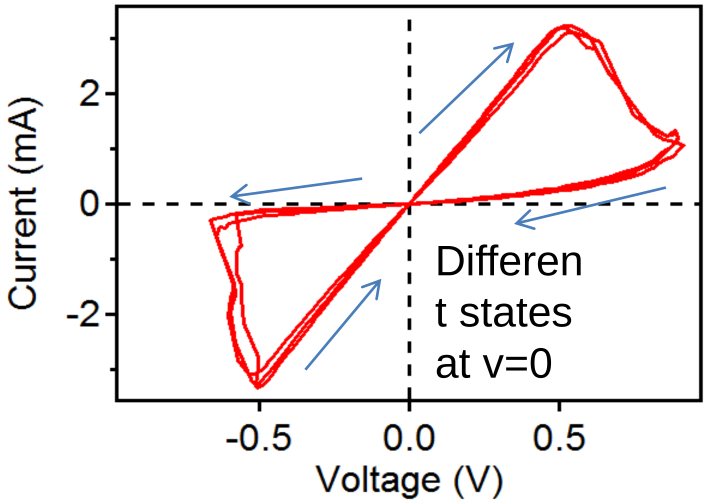
Need to use correct physics so that it applies in this universe!

Memristors have ‘pinched’ hysteresis loops

When driven by a cyclic voltage or current

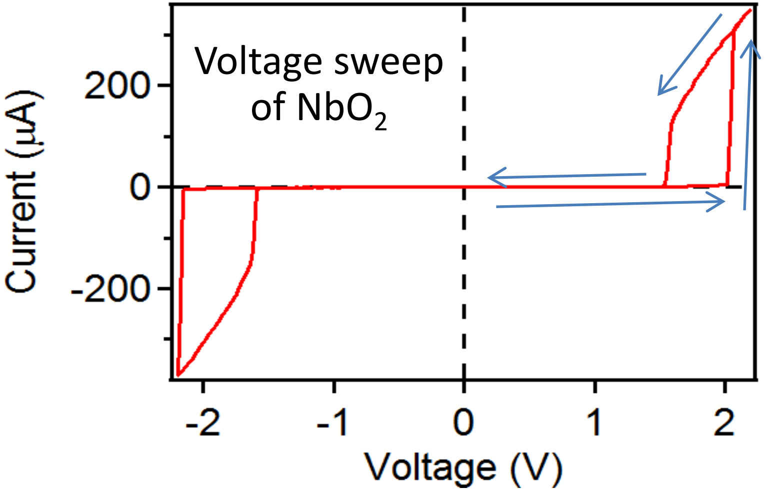
Nonvolatile Memristor

- Digital memory/storage device
- Synapse in neuromorphic circuit



Locally Active Memristor

- “Selector” in crossbar memories
- Emerging neuronal computing devices



Find or Invent Materials that Posses Memristance:

TEXAS A&M UNIVERSITY Engineering

Nonvolatile:

‘Synaptic’

State stored as resistance

Continuously variable

- Real numbers

Memory and storage
ReRAM, PC RAM, STT RAM

TaO_x, $Ge_2Sb_2Te_5$, magnetics

Locally Active:

‘Neuronic’

State transmitted as spike

Threshold switching, NDR

- Looks digital

Gain, logic, chaos

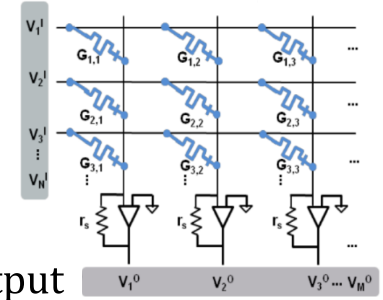
Mott transitions, mobile ions

VO_2 & NbO_2, molecular redox, CDW, ionic diffusion

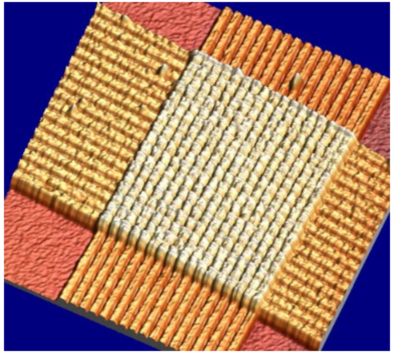
Memristor Crossbars for Computational Acceleration

Dot Product Engine: memristor arrays accelerate vector-matrix multiplication

Input Voltage vector



Output current $I_j^O = \sum_i G_{ij} \cdot V_i^I$



- Parallel multiply & add through Kirchoff's and Ohm's laws

  1961, K. Steinbuch "*Die Lernmatrix*"– suggests using "ferromagnetic toroids"

- Memristors as highly scalable, tunable analog resistors

  High ON/OFF ratio (~10^5), supporting multiple levels

- Well suited for streaming workloads like neural nets
- Many ways to scale up

  Memristor levels, array size, wire pitch, 3D layer, DAC/ADC speed & width etc.

- Performance (execution time) improvements >1000x and energy efficiency >100x over GPUs for particular applications
- Commercial products in development

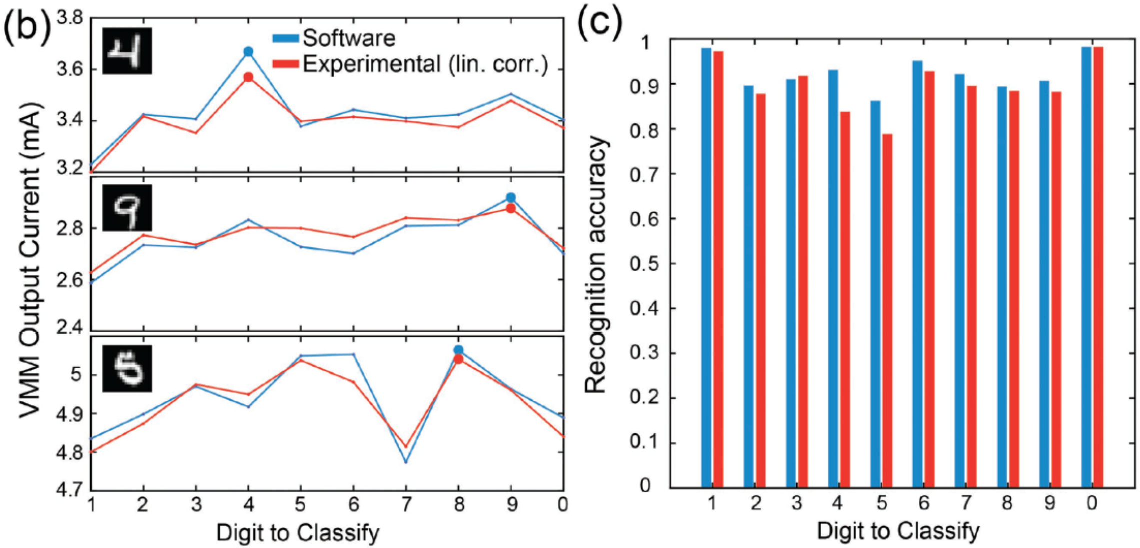
Memristor-Based Analog Computation and Neural Network Classification with a Dot Product Engine

TEXAS A&M UNIVERSITY
Engineering

Miao Hu, Catherine E. Graves, Can Li, Yunning Li, Ning Ge, Eric Montgomery, Noraica Davila, Hao Jiang, R. Stanley Williams, J. Joshua Yang, Qiangfei Xia,* and John Paul Strachan**

Adv. Mater. **2018**, 1705914

MNIST digit classification, single layer network



Crude system with ~2% bad devices

Vector-matrix multiplication:
~8x 4 bit digital ASIC

Full inference output (estimated):
15x inference rate
5.5x power efficiency
7.5x area efficiency

Significant improvements over time

IEEE JOURNAL OF SELECTED TOPICS IN QUANTUM ELECTRONICS, VOL. 26, NO. 1, JANUARY/FEBRUARY 2020

TEXAS A&M UNIVERSITY Engineering

Photonic Multiply-Accumulate Operations for Neural Networks

Mitchell A. Nahmias, Thomas Ferreira de Lima, Alexander N. Tait, Hsuan-Tung Peng, Bhavin J. Shastri, *Member, IEEE*, and Paul R. Prucnal, *Fellow, IEEE*

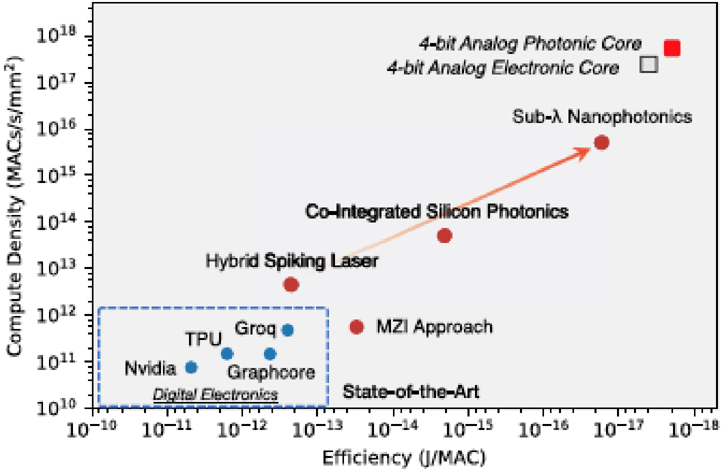


Fig. 7. Comparison of deep learning hardware accelerators with photonic platforms discussed in Section VI, modified from Ref. [7]. Photonic systems can support high bandwidth densities on-chip while consuming minimal energy both transporting data and performing computations. Metrics for digital electronic architectures taken from various sources [12], [124]–[127]. Also included are the analog limits for photonic and electronic matrix cores with $N = 1024$ and 4 bits of precision, from Table I.

TABLE I
COMPUTE DENSITY PERFORMANCE FOR IDEALIZED ELECTRONIC AND PHOTONIC MATRIX CORES WITH $N = 1024$, SUBJECT TO POWER DENSITY $<1\ \mathrm{W/mm^2}$

Technology	Noise Precision	Energy (aJ/MAC)	Compute Density (PMACs/s/mm²)
Electronic Crossbar	4 bit	4.0	250
	8 bit	5.0	198
Photonic Core	4 bit	2.0	513
	8 bit	81.9	12.2

At 8 bit precision a memristor crossbar has 16x better energy efficiency and compute density than an optimized photonic system.

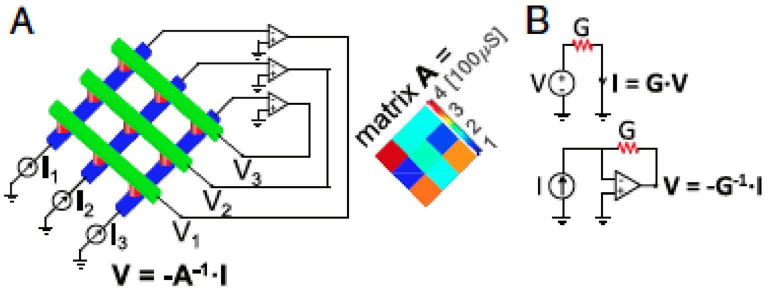
Solving matrix equations in one step with cross-point resistive arrays

Zhong Sun[a], Giacomo Pedretti[a], Elia Ambrosi[a], Alessandro Bricalli[a], Wei Wang[a], and Daniele Ielmini[a,1]

[a]Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy

Edited by Eli Yablonovitch, University of California, Berkeley, CA, and approved January 17, 2019 (received for review September 11, 2018)

Proceedings of the National Academy of Sciences **116**, 4123-4128 (2019)

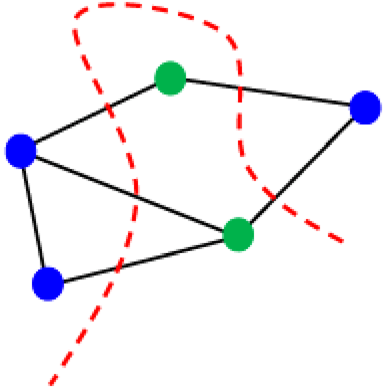


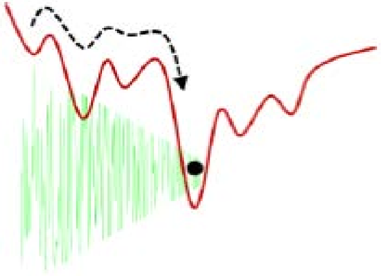
Can solve Ax=b for known A and b in O(1)
Can find maximal eigenvalue in O(1)
Can solve 1st order ODEs in O(1)
Can compute AB and A^{-1} in O(n)

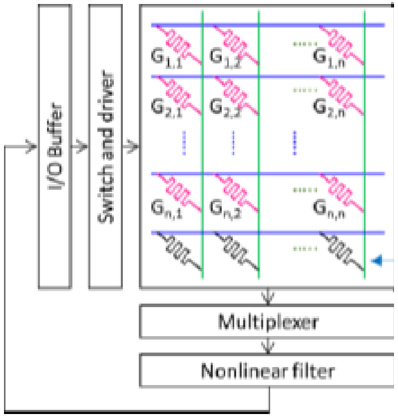
General idea: run a dot product engine in reverse!

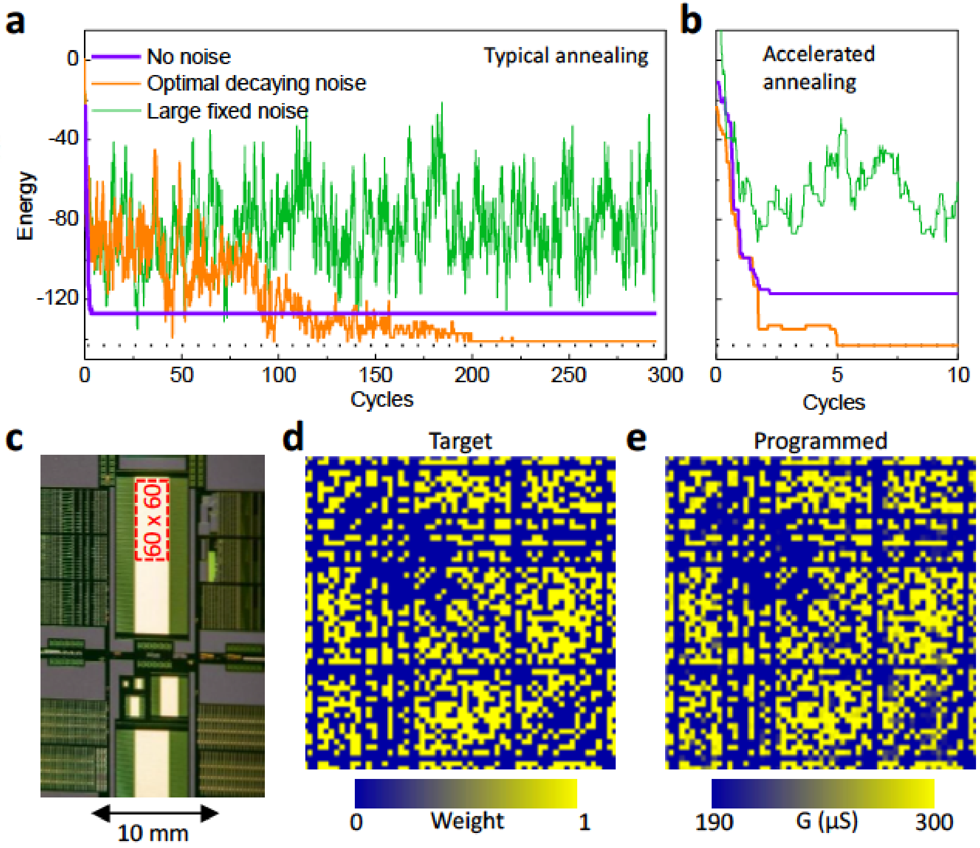
Store matrix as conductance values in crossbar
Inject currents corresponding to b values into columns
After transients dissipate, measure voltages on rows (x)
Can speed up any linear algebra operation by n^2
Highly parallel and reversible - almost no energy
Can sweep inputs and look for nulls dynamically
Precision can be improved by iteration with small cost

Hopfield Network solving 60-node max-cut problems (NP) using noise









Harnessing Intrinsic Noise in Memristor Hopfield Neural Networks for Combinatorial Optimization

TEXAS A&M UNIVERSITY
Engineering

Fuxi Cai[1,2,*], Suhas Kumar[1,*], Thomas Van Vaerenbergh[1,*], Rui Liu[1,3], Can Li[1], Shimeng Yu[4], Qiangfei Xia[5], J. Joshua Yang[5], Raymond Beausoleil[1], Wei Lu[2], and John Paul Strachan[1]

https://arxiv.org/abs/1903.11194

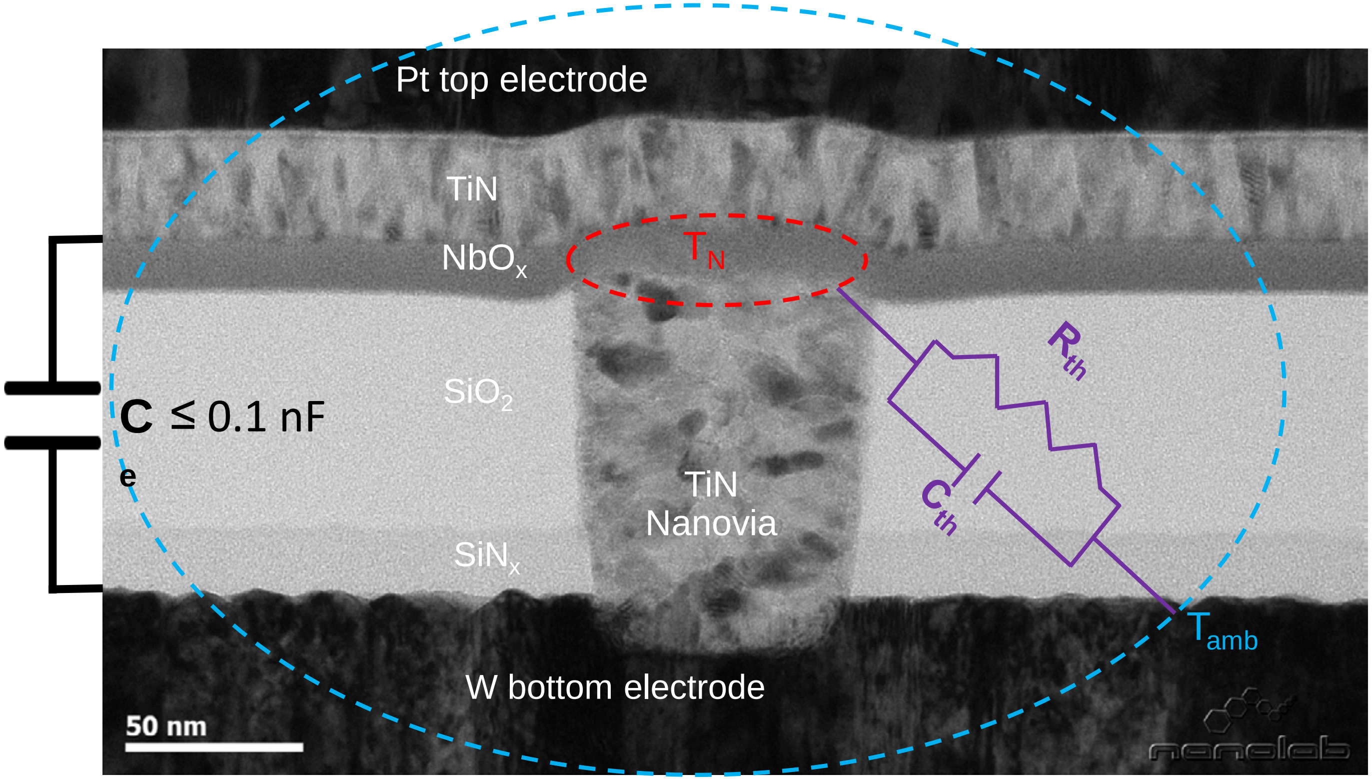
Table 1. Comparison of the mem-HNN and current state-of-the-art annealing accelerators, such as a GPU implementation of the noisy mean-field algorithm[10], our own simulations using the previously suggested[6] parallel tempering implementation on a CPU (cfr. SM 1.8.3) and experimental results for the D-wave annealer and the measurement-feedback CIM discussed in Ref[6]. A hybrid update mechanism updates some, but not all the nodes at a given iteration.

	mem-HNN (seq.) memristor	mem-HNN (par.) memristor	NMFA GPU	PT@UFO single-core CPU	D-wave 2000Q supercond. qubits	CIM fiber-optics
Clock frequency	1 GHz	1 GHz	1.582 GHz	2.6 GHz		1 GHz
Annealing time T_{ann}	300 ns	300 ns	12.3 μs ($N = 100$)	223.6 μs	1 ms (N=55)	150 μs
Time-to-solution	3.3 μs	0.3 μs	10 μs	223.6 μs	10^4 s (N=55)	600 μs
Power	66 mW	792 mW	<250 W	20 W	25 kW	
Energy-to-solution	0.22 μJ	0.22 μJ	<2.5 mJ	4 mJ	250 MJ	
Solutions/s/Watts	4.6×10^6	4.6×10^6	>400	250	4×10^{-9}	
Update mechanism	hybrid	hybrid	asynchronous	asynchronous	synchronous	asynchronous
Connectivity	all-to-all	all-to-all	all-to-all	all-to-all	Chimera	all-to-all
Scaling p_{TS}	ae^{-bN}	ae^{-bN}	ae^{-bN}	ae^{-bN}	ae^{-bN^c}	ae^{-bN}
Cryogenic cooling	no	no	no	no	yes	no

Dynamical measurements:

oscillations, neuristors, chaotic attractors

Integrated NbO_2 (Mott) Memristor and Capacitor



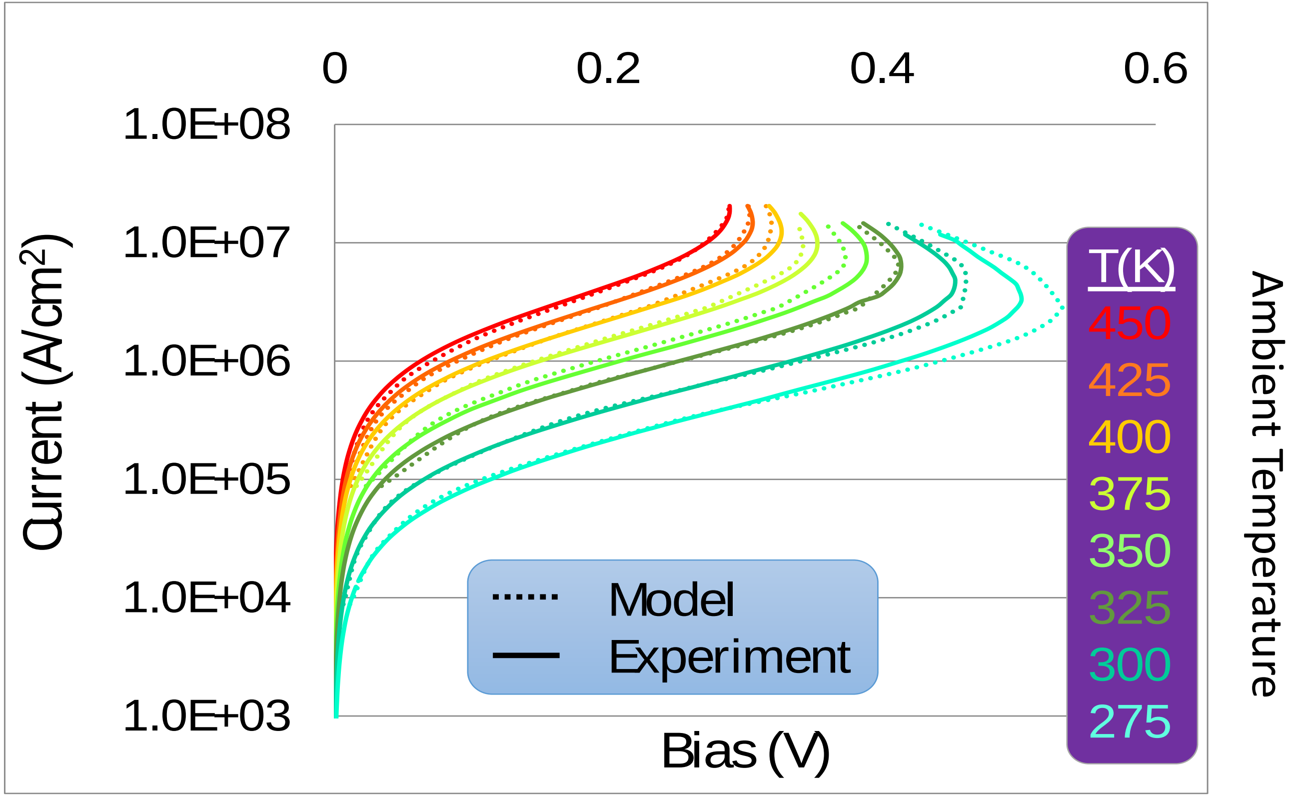
Thermal Design

$R_{th}C_{th} \leq 0.1$ ns

$R_{th} \geq 10^6$ K/W

$C_{th} \leq 10^{-16}$ J/K

Measured (solid) and calculated (dashed) V-I curves for CC-NDR



G Gibson et. al., *Appl. Phys. Letts.* **108** (2016)

Locally-active memristor model for S-type NDR

TEXAS A&M UNIVERSITY Engineering

Temperature (State) Dependent Ohm's Law: 3D Frenkel-Poole Conduction Model

$$j(F,T) = \sigma F = \sigma_0(T)\left(\frac{k_B T}{\beta}\right)^2 \left\{1 + \left(\frac{\beta\sqrt{F}}{a k_B T} - 1\right) e^{\frac{\beta\sqrt{F}}{a k_B T}}\right\} + \frac{\sigma_0(T) F}{2}$$

$$\sigma_0(T) = e\mu N_c \left(\frac{N_d}{N_t}\right)^2 e^{-\frac{E_d + E_t}{2 k_B T}}$$

Dynamical Equation: Newton's Law of Cooling (Sasha Alexandrov Model)

$$C_{th}\frac{dT_N}{dt} = \frac{T_{amb} - T_N}{R_{th}} + IV$$

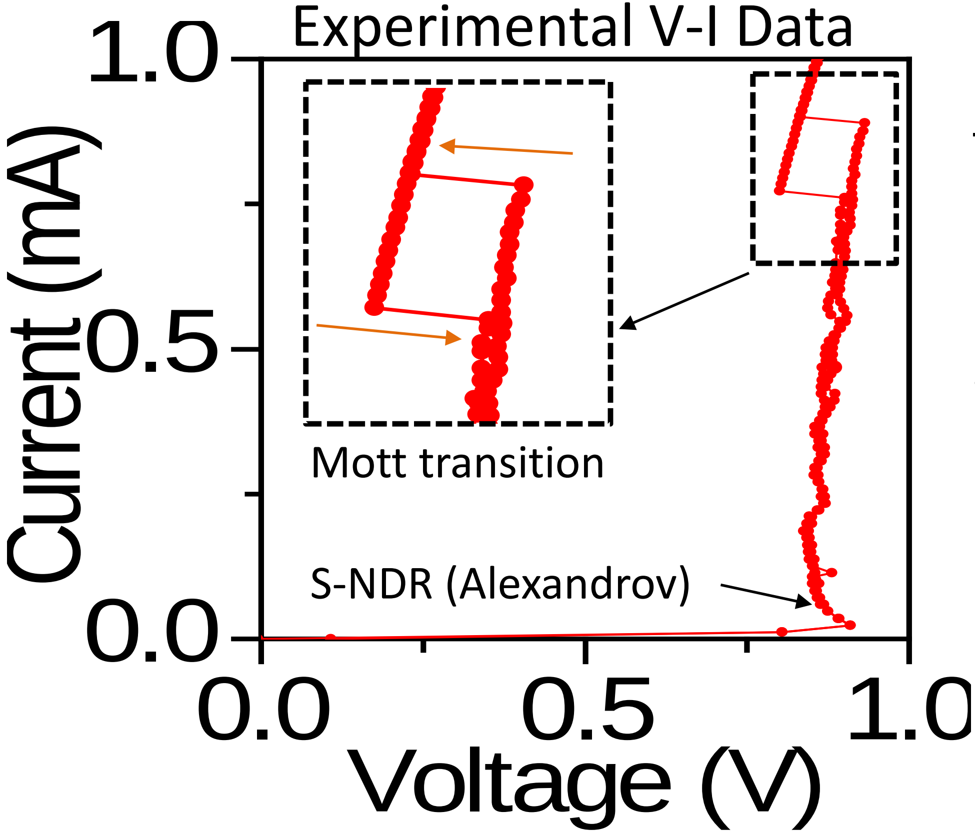
Negative Differential Resistance (Local Activity) is a result of feedback among thermally activated transport, Joule heating and heat transport!

What happened to the Mott transition?

G Gibson et. al., *Appl. Phys. Letts.* **108** (2016)

'DC' Electrical Characterization of a NbO_2 'Mott' Memristor

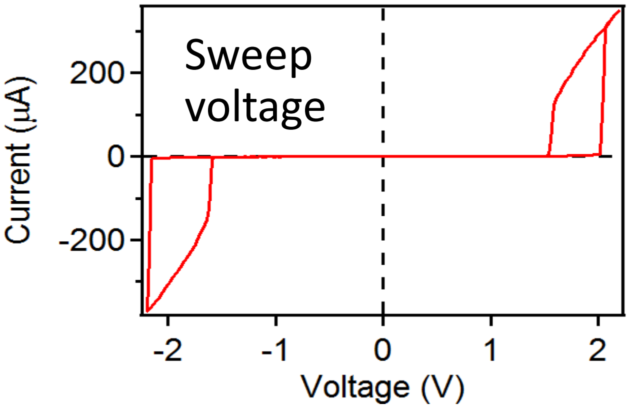
Experimental V-I Data



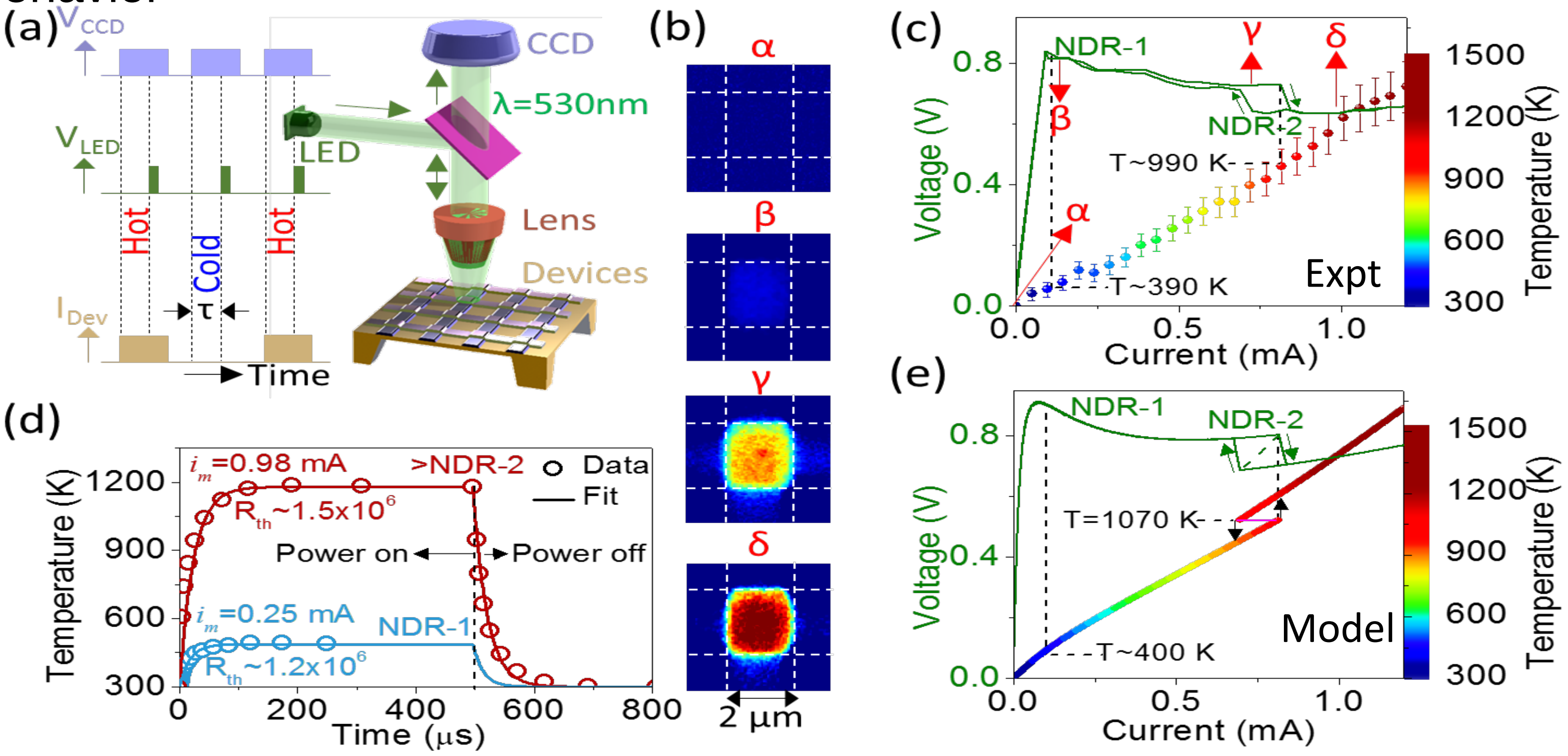
Data collected by slowly sweeping current and measuring voltage.

Pushed to much higher currents and thus internal temperatures.

See two 'NDR' regions!



New tools for experiments: Thermoreflectance Imaging of NbO_2 current source behavior

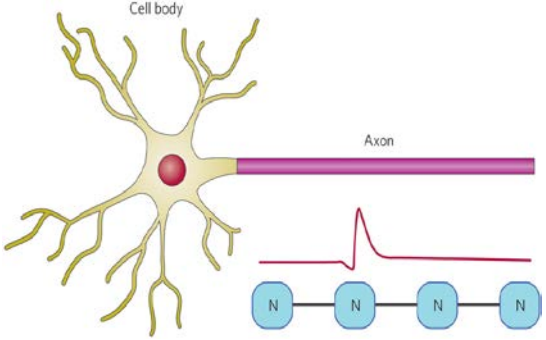


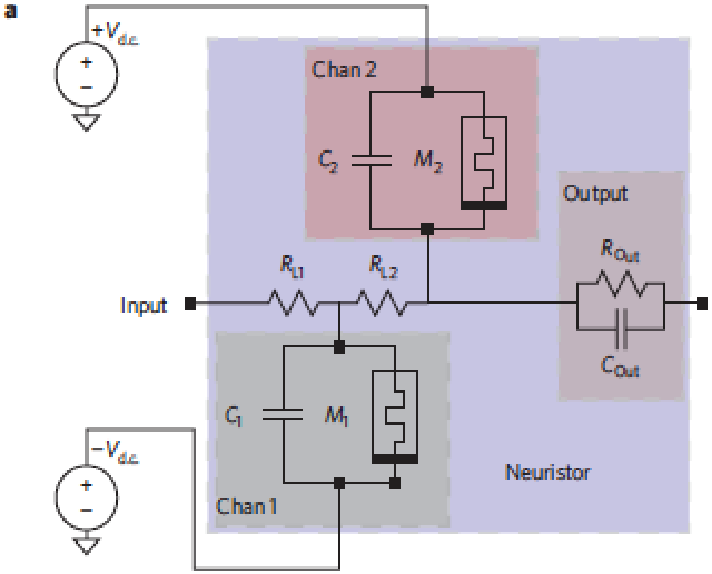
nature materials
LETTERS
PUBLISHED ONLINE: 16 DECEMBER 2012 | DOI: 10.1038/NMAT3510

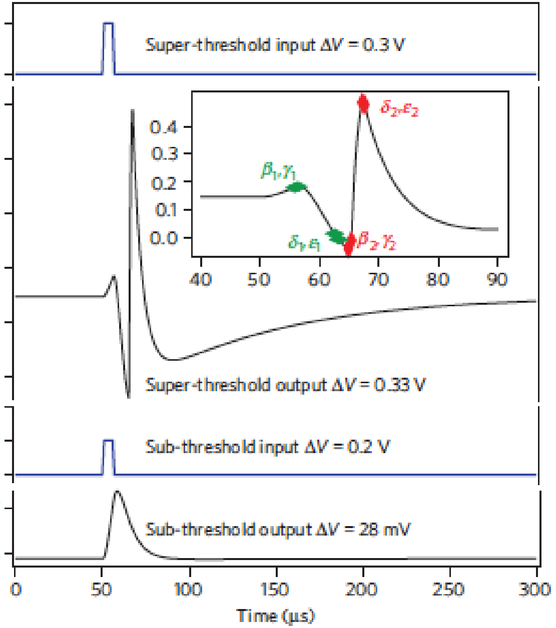
A scalable neuristor built with Mott memristors (NbO_2)

Matthew D. Pickett*, Gilberto Medeiros-Ribeiro and R. Stanley Williams

'integrate and fire' pulse amplifier for threshold logic and communication





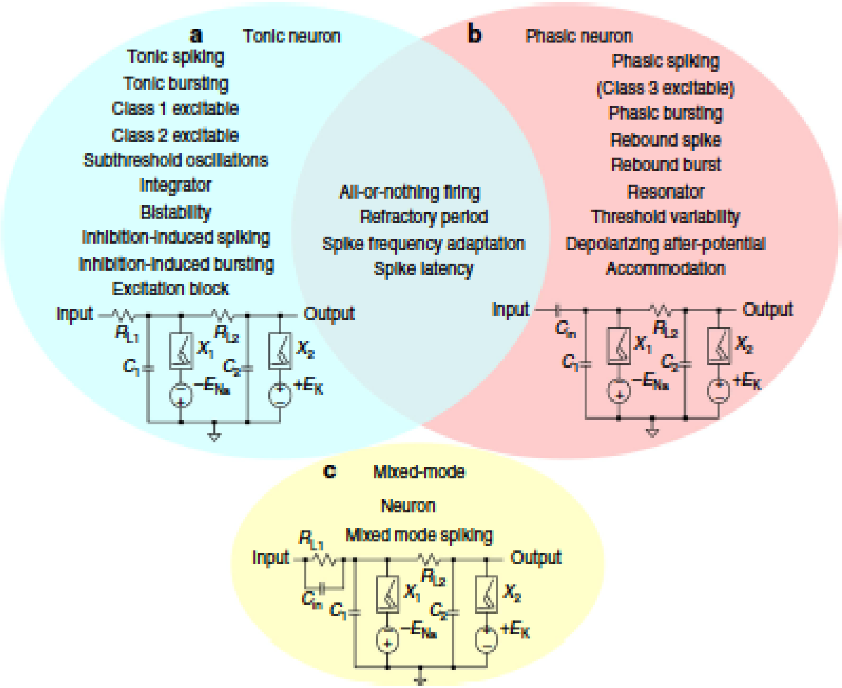


NATURE COMMUNICATIONS | (2018)9:4661 | DOI: 10.1038/s41467-018-07052-w

ARTICLE
DOI: 10.1038/s41467-018-07052-w OPEN

Biological plausibility and stochasticity in scalable VO_2 active memristor neurons

Wei Yi[1], Kenneth K. Tsang[1], Stephen K. Lam[1], Xiwei Bai[1], Jack A. Crowell[1] & Elias A. Flores[1]



"Here we report that neurons built with nano-scale vanadium dioxide active memristors possess all three classes of excitability and most (23) of the known biological neuronal dynamics, and are intrinsically stochastic."

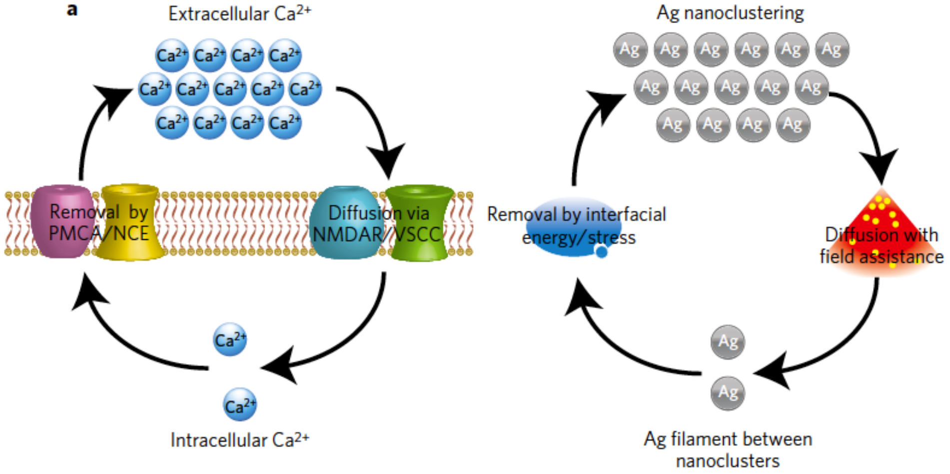
Wei Yi et al., HRL!

Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing

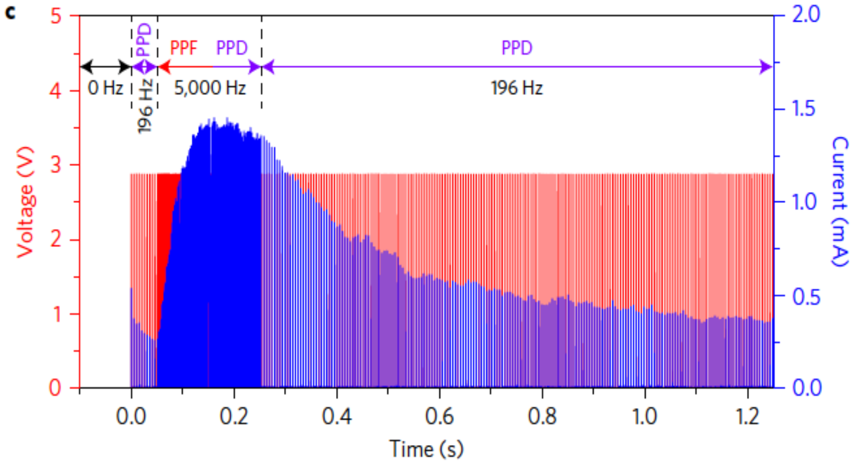
Zhongrui Wang[1†], Saumil Joshi[1†], Sergey E. Savel'ev[2], Hao Jiang[1], Rivu Midya[1], Peng Lin[1], Miao Hu[3], Ning Ge[3], John Paul Strachan[3], Zhiyong Li[3], Qing Wu[4], Mark Barnell[4], Geng-Lin Li[5], Huolin L. Xin[6], R. Stanley Williams[3], Qiangfei Xia[1] and J. Joshua Yang[1*]

NATURE MATERIALS 2017

Emulating Ca ions in biosynapses with Ag in Oxides



Ion dynamics produce synaptic plasticity
Quantitative models mirror experiments



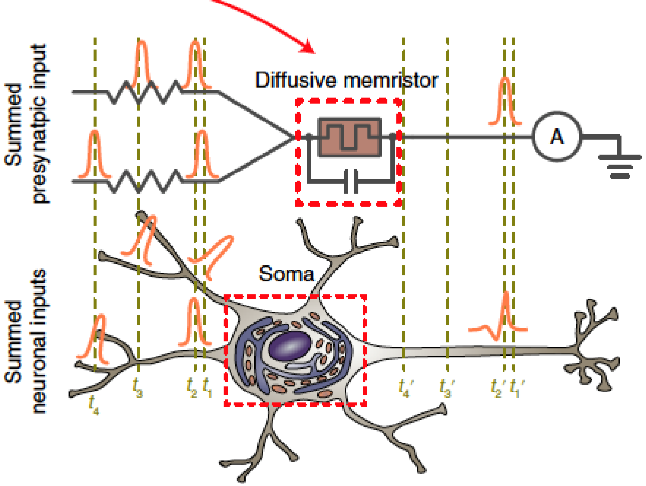
Fully memristive neural networks for pattern classification with unsupervised learning

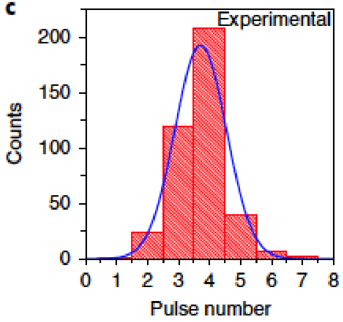
TEXAS A&M UNIVERSITY
Engineering

Zhongrui Wang[1,6], Saumil Joshi[1,6], Sergey Savel'ev[2], Wenhao Song[1], Rivu Midya[1], Yunning Li[1], Mingyi Rao[1], Peng Yan[1], Shiva Asapu[1], Ye Zhuo[1], Hao Jiang[1], Peng Lin[1], Can Li[1], Jung Ho Yoon[1], Navnidhi K. Upadhyay[1], Jiaming Zhang[3], Miao Hu[3], John Paul Strachan[3], Mark Barnell[4], Qing Wu[4], Huaqiang Wu[5], R. Stanley Williams[3]*, Qiangfei Xia[1]* and J. Joshua Yang[1]*

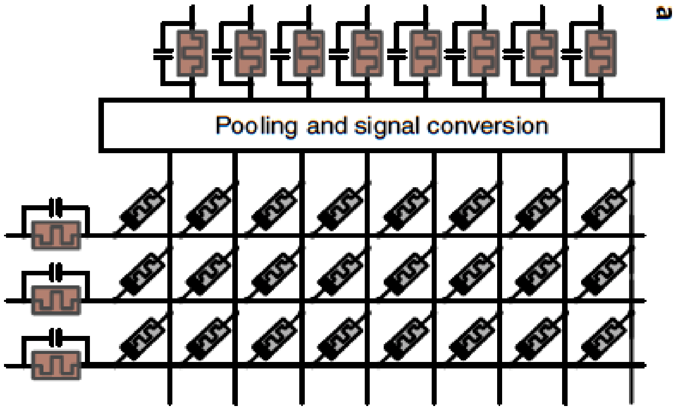
NATURE ELECTRONICS | VOL 1 | FEBRUARY 2018 | 137–145 |

Emulate a leaky integrate and fire neuron





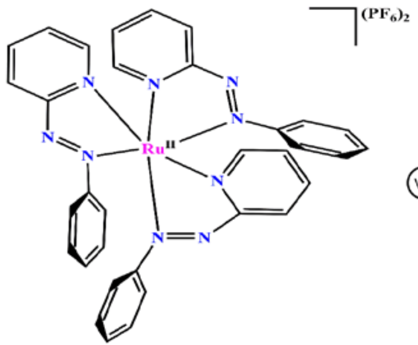
Unsupervised learning in all-memristor network



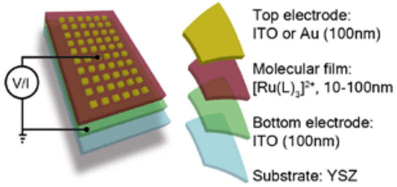
Robust resistive memory devices using solution-processable metal-coordinated azo aromatics, S. Goswami et al.

Nature Materials **16**, 1216–1224 (2017)

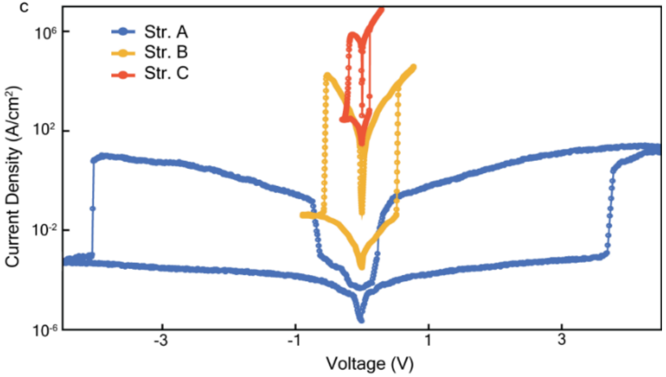
TEXAS A&M UNIVERSITY Engineering

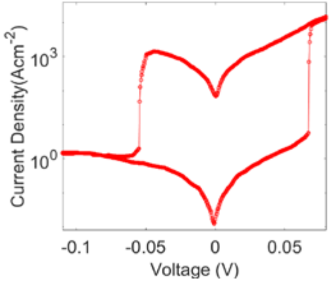


device structure



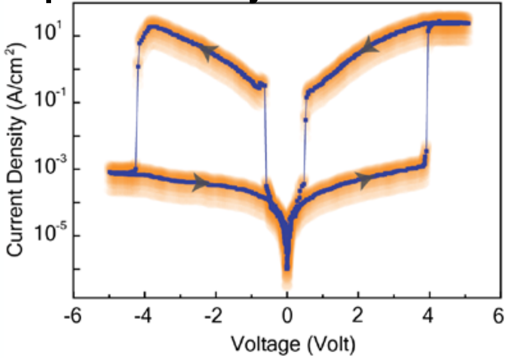
Switching is field driven



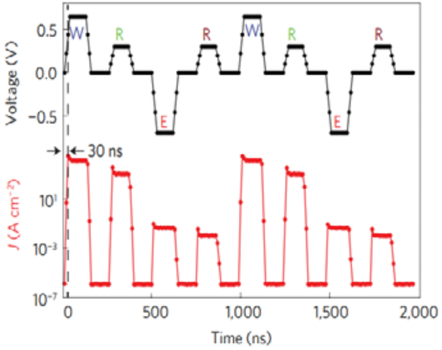


Switching energy ~6aJ

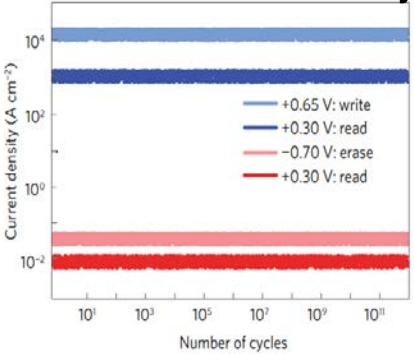
reproducibility – 321 devices



# Devices	Top Electrode Size (um)
122	100
47	80
38	60
29	40
31	20
26	10
28	1
Total- 321	



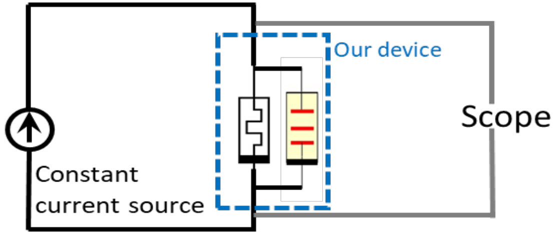
endurance & stability



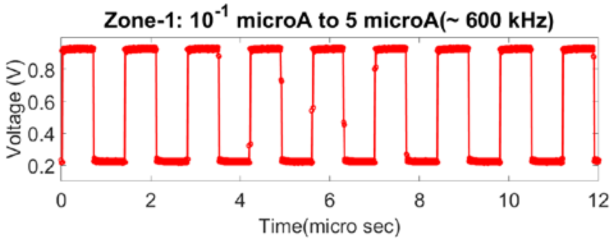
A self oscillating molecular-film device (Memrisys 2019)

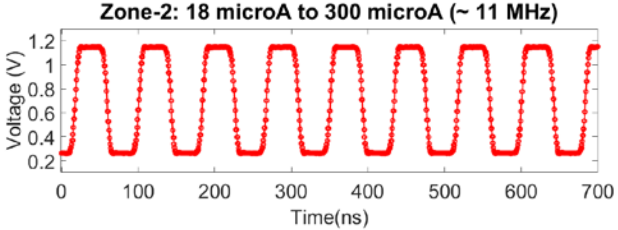
Sreetosh Goswamy and T. Venkatesan, NUS

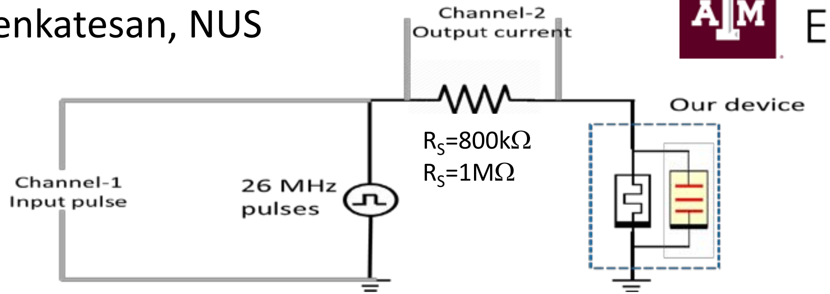
TEXAS A&M UNIVERSITY
Engineering



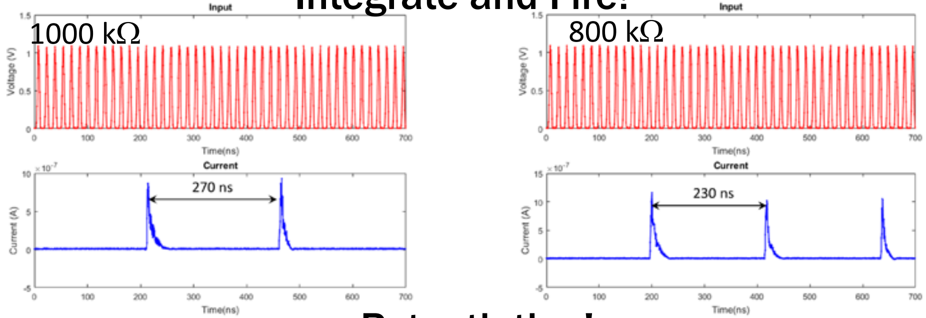
Single device bistable oscillator



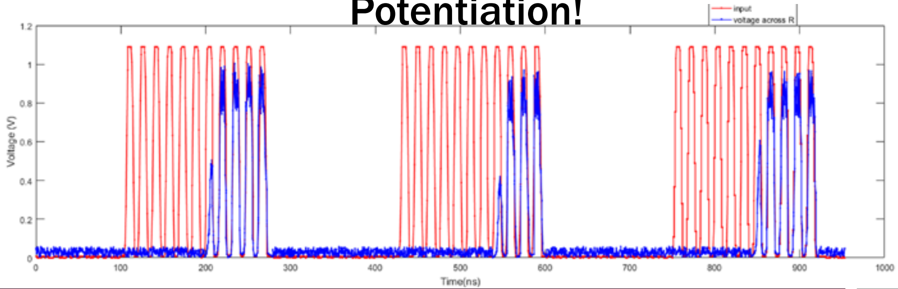




Integrate and Fire!



Potentiation!



Dynamical nonlinear memory capacitance in biomimetic membranes

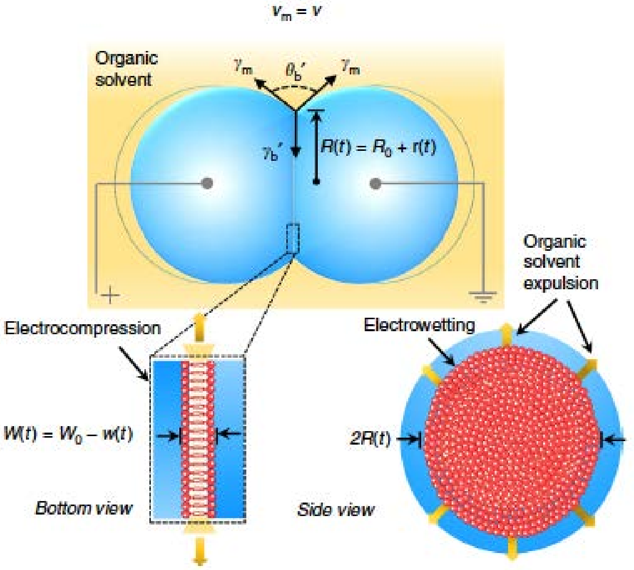
Joseph S. Najem[1,2], Md Sakib Hasan[3], R. Stanley Williams[4], Ryan J. Weiss[3], Garrett S. Rose[3], Graham J. Taylor[1,5], Stephen A. Sarles[1] & C. Patrick Collier[6]

NATURE COMMUNICATIONS | (2019)10:3239 | https://doi.org/10.1038/s41467-019-11223-8

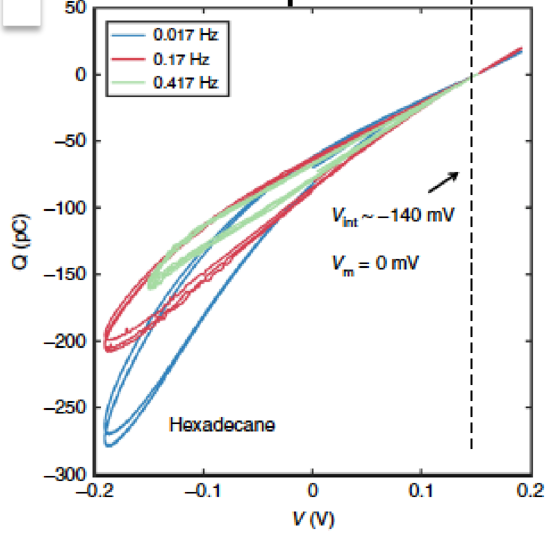
TEXAS A&M UNIVERSITY Engineering

Need to revise the Hodgkin-Huxley Model?

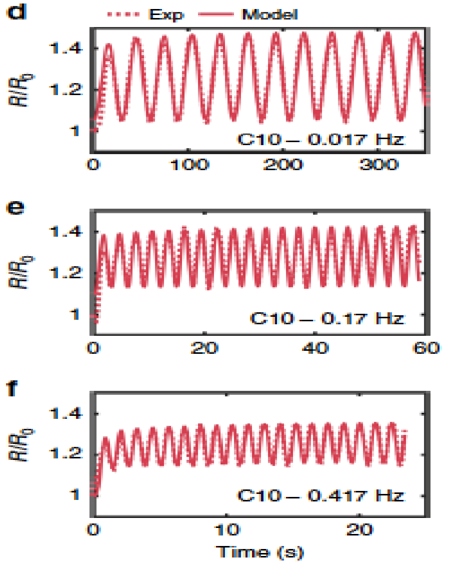
Reductionist biology: simplest cell model



Signal processing via memcapacitance

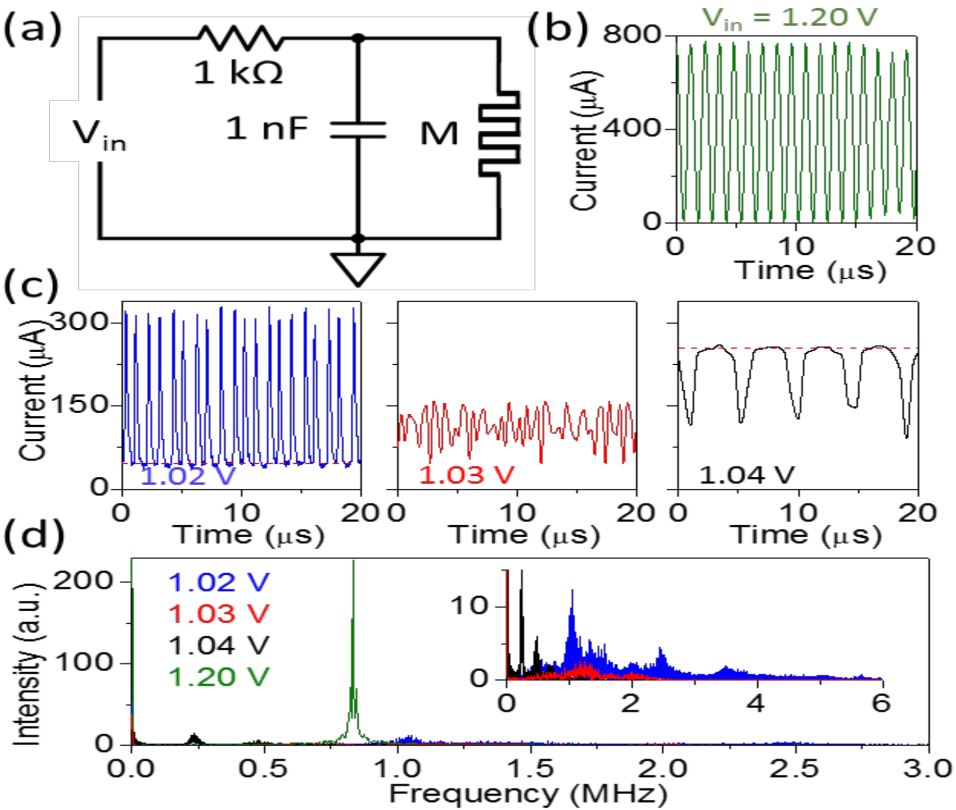


Quantitative modeling

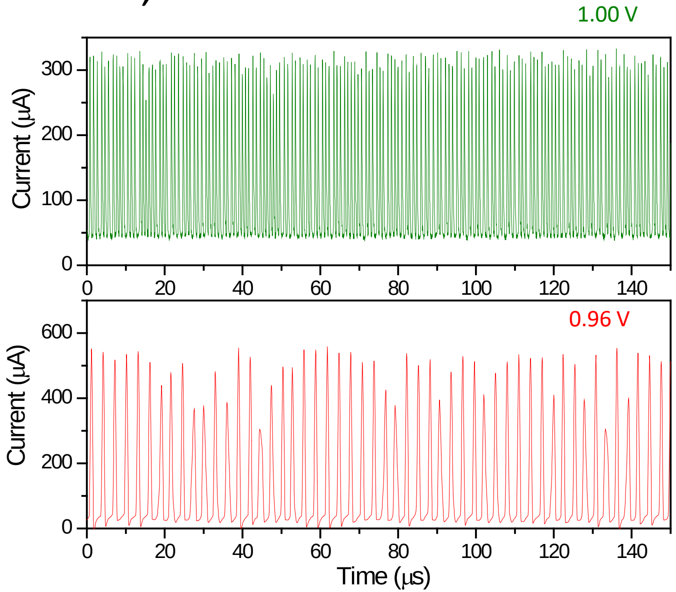


Chaotic oscillations in the NbO_2 S-NDR

TEXAS A&M UNIVERSITY Engineering



Extreme voltage sensitivity – oscillation amplitudes, frequency and regularity – sinusoidal, chaotic and random



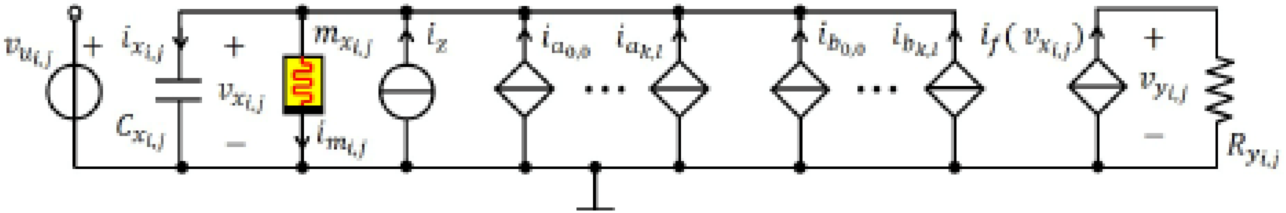
S. Kumar et al., Nature (2017)

Locally active memristors at edge of chaos for computation

TEXAS A&M UNIVERSITY Engineering

Theoretical Foundations of Memristor Cellular Nonlinear Networks: Memcomputing With Bistable-Like Memristors

Ronald Tetzlaff, *Senior Member, IEEE*, Alon Ascoli, Ioannis Messaris, and Leon O. Chua, *Fellow, IEEE*

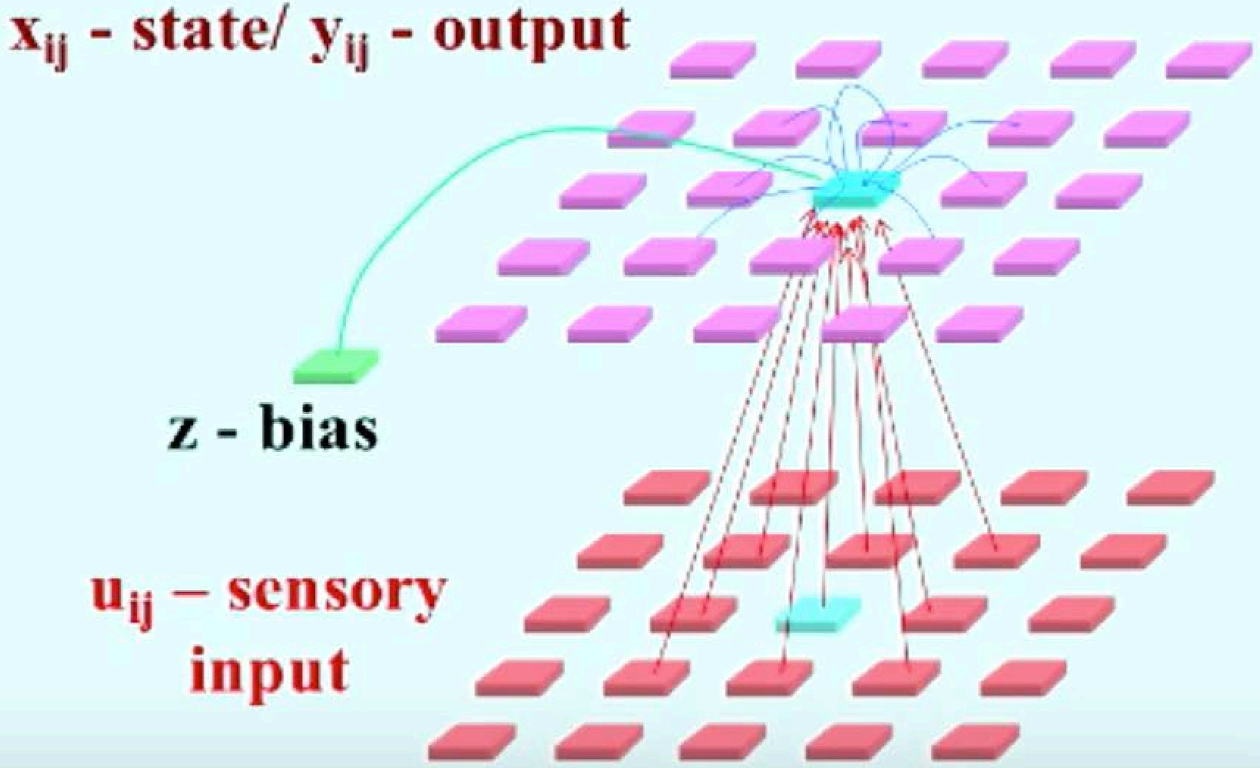


A single cell of a cellular nonlinear neural network
Power, capacitor, memristor(s), inputs and outputs

By itself, this cell will settle into a stable steady state – it will 'go to sleep'

If the memristor is on the 'edge of chaos', connecting two or more cells with resistors will 'wake them up'

Constructing a Cellular Nonlinear Network

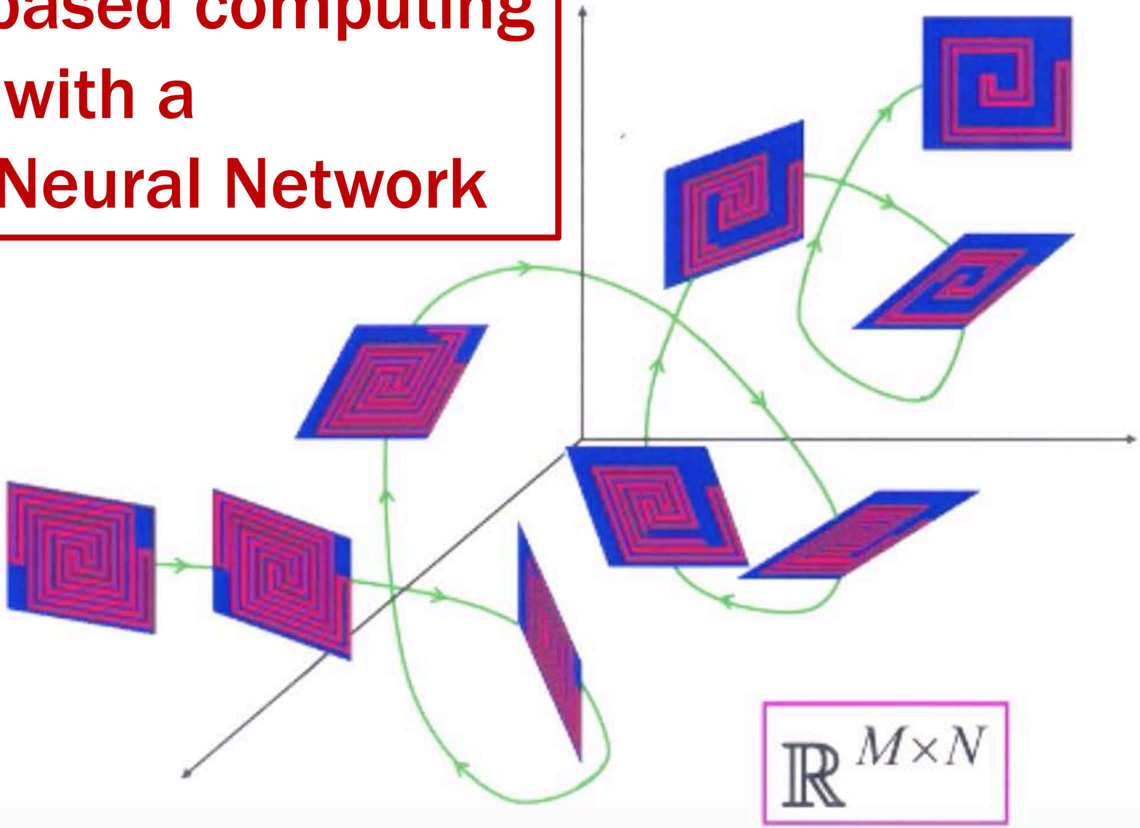


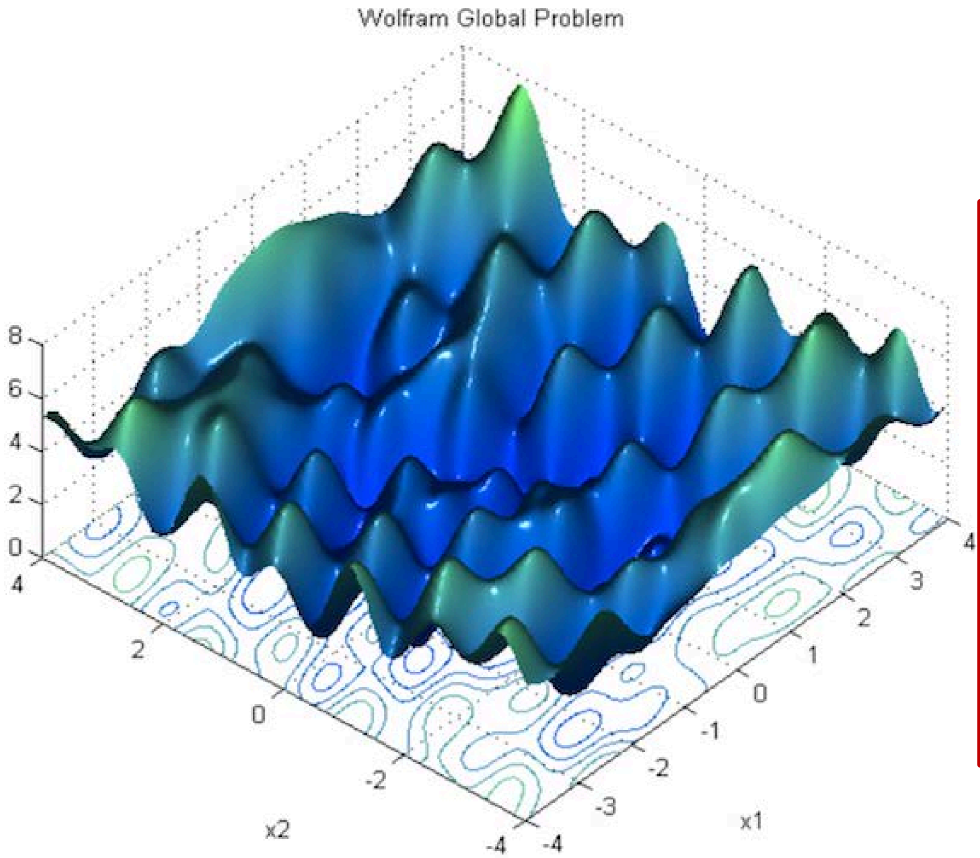
$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

$$z = -0.5$$

Attractor based computing with a Cellular Neural Network

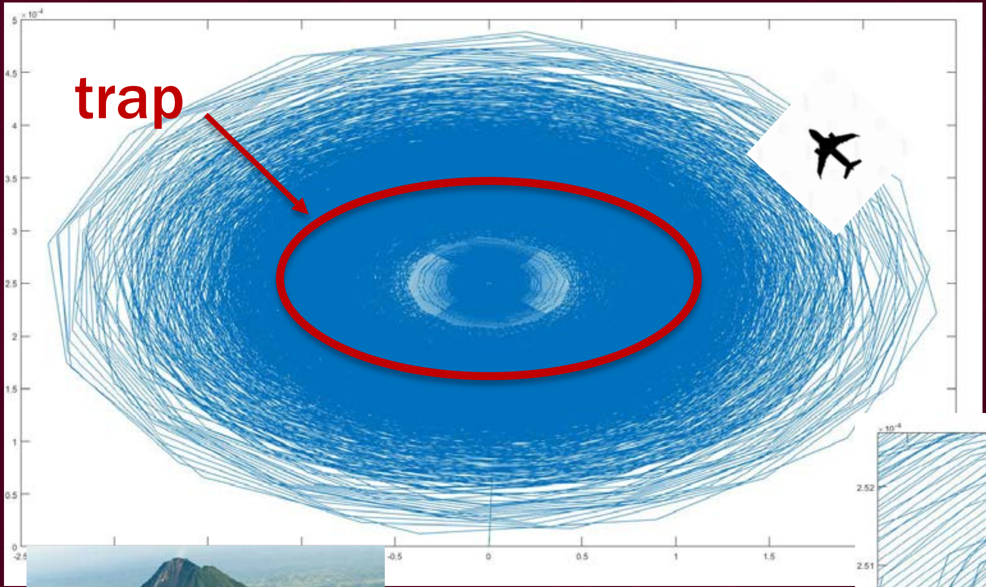


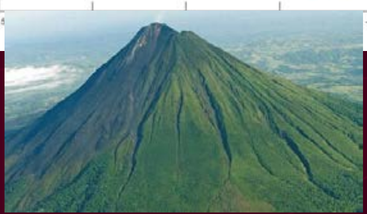


Most neural networks rely on multidimensional minimization techniques. That is like driving around on a mountainous terrain with no roads and no maps, always trying to figure out which way is 'down'.

Yoshio Bengio, Deep Learning Summer School, Aug. 15, 2015

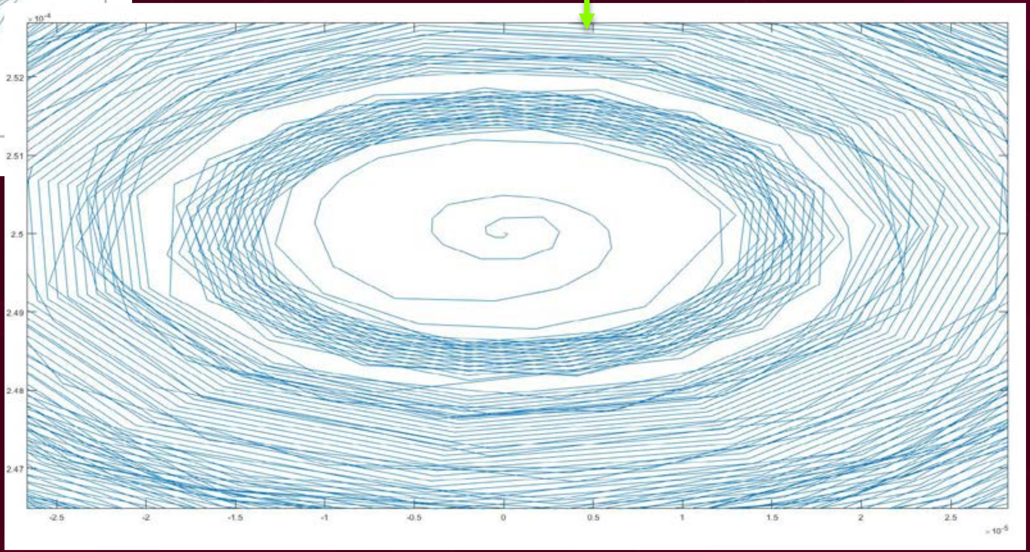
‘volcano in a basin’





Strange Attractor-based computing is more like flying

← The transient chaotic trajectory of a nonlinear dynamical system



~10^6 faster than random search

Memristors, Nonlinear Dynamics, Neuromorphic Computing

TEXAS A&M UNIVERSITY Engineering

Computation (Brain Inspired and Other)

Turing O-machines
Multinary Logic (ternary and higher)
Neural Networks (of all kinds)
Hebbian Learning, STDP
Boltzmann/Ising Machines
Hopfield Networks (NP problems)
Bayesian Inference, Markov Chains

Electro-Iono-Thermo-Device Physics

Drift-Diffusion-Thermophoresis
Thermally activated transport
Phase Transitions (e.g. Mott)
Negative Differential Resistance
Negative Differential Capacitance
Non-equilibrium Thermodynamics
Spontaneous Symmetry Breaking

Nonlinear Dynamical Circuit Theory

Principle of Local Activity
Oscillators and Amplifiers (e.g. neuristors)
Chaos and Edge of Chaos
Complexity and Emergent Phenomena
Connection to Non-equilibrium Thermo

Materials Science and Discovery and Design

Materials by Design and Discovery
Correlated Electronic and Spin States
Topological Properties of Materials
Tailored Bulk and Interface Properties
Nanoscale Structural Phenomena

TEXAS A&M UNIVERSITY

Engineering