Nonlinear oxide devices and brain-like analog computing

Suhas Kumar
Thanks for the funding and user facilities
The voids of computing

Present digital computers

Floating point arithmetic
Arithmetic heavy
Boolean logic

Image processing
Real-time regex matching
Data heavy

Gene sequencing (+ other NP-class)
Nonlinear dynamics
Weather prediction

Major limitations of digital computers:

• End of Moore’s law
• Von Neumann architecture
• Boltzmann tyranny
• Boolean logic
• Turing limit
The post-Moore’s law physics-driven computer

Materials discovery
(new behaviors)

Interacting devices
(new algorithms)

Functional devices
(new functions)
Outline of the talk

1. Neuronic devices – information processing or communication

2. Synaptic devices – information storage

3. The fundamental connection between synaptic and neuronic devices

4. Synaptic + neuronic NP-hard accelerators
Neuronic devices
Mott Memristors

Volatile resistive switching can emulate neuron-like spiking!

NbO2 $\rightarrow$ Mott insulator

![Graph showing the relationship between voltage and current for different memristors and model.](image)

*Nature Materials* 12, 114 (2013)
Devices and static behavior

Kumar et al., *Nature Communications*, 8, 658 (2017)
The cause of NDR in NbO$_2$

Kumar, Nature Comms. 8, 658 (2017)
Thermal anomaly in NbO$_2$

\[ \sigma_{th}^{Metal} < \sigma_{th}^{Insulator} \]

Counter Wiedemann Franz postulate

Kumar et al., Nature Communications 8, 658 (2017)
Lee et al., Science 355, 371 (2017)
Model for the electrical behavior

Modified 3D Poole-Frenkel Equation

\[ i_m = \left[ \alpha e^{-\frac{E_a}{2k_B T}} \left( \frac{k_B T}{\beta \sqrt{v_m}} \right)^2 \left\{ 1 + \left( \frac{\beta \sqrt{v_m}}{k_B T} - 1 \right) e^{\frac{\beta \sqrt{v_m}}{k_B T}} \right\} + \frac{\alpha e^{-\frac{E_a}{k_B T}}}{2t} \right] v_m \quad \text{... Equation (1)} \]

Temperature dynamics

\[ \frac{dT}{dt} = \frac{i_m v_m}{C_{th}} - \frac{T - T_{amb}}{C_{th} R_{th}(T)} \quad \text{... Equation (2)} \]

Mott transition in \( R_{th} \)

\[ R_{th}(T) = \begin{cases} 1.4 \times 10^6 & \text{for } T \leq T_C \\ 2 \times 10^6 & \text{for } T > T_C \end{cases} \quad \text{... Equation (3)} \]

Kumar et al., Nature Communications, 8, 658 (2017)
Temperature-controlled instability (Chua Corsage)

Theorem: Chua corsage leads to chaos in dynamics

Kumar et al., Nature Communications, (2018)
Kumar et al., Nature Communications, 8, 658 (2017)
Modeling the dynamical behavior

\[
\frac{dT}{dt} = \frac{v_m^2}{R(T)C_{th}} - \frac{T - T_{amb}}{R_{th}C_{th}} + \eta(T)
\]

\[
\frac{dv_m}{dt} = \frac{v_{in}}{R_sC} - v_m \left( \frac{1}{R(T)C} + \frac{1}{R_sC} \right)
\]

What does it take to produce chaos?

1. Local activity
2. A minimum of three state variables; OR two state variables + a coupled dynamic signal

For chaos, where is the third state variable or oscillating force?

\[
\eta(T) = T \left( \frac{k_B}{C_{th}} \right)^{\frac{1}{2}} \frac{4\pi}{R_{th}C_{th}} \cos \left[ \frac{2\pi t}{R_{th}C_{th}} \right]
\]

Fluctuation-dissipation theorem of local activity

Local Activity:

The biology-based theory that now allows us to predict dynamics of electronic devices using their static behavior.

Local activity leads to chaos

Kumar et al., Nature Communications, (2018)
What is chaos?

– Extreme sensitivity to variations in initial conditions

– In between perfectly ordered and completely random behavior

– Very difficult to predict (but not impossible)

– Why is edge of chaos behavior favored in computing systems?
  – Units communicate/cooperate with to compute
  – But no global synchronization
  – Applies to neurons in a brain
Thermal fluctuations

\[ \eta(T) = T \left( \frac{k_B}{C_{th}} \right)^{\frac{1}{2}} \frac{4\pi}{R_{th}C_{th}} \cos \left[ \frac{2\pi t}{R_{th}C_{th}} \right] \]

Double pendulum with a large driving force!

Simulations

Chaotic oscillations

Are temperature fluctuations the only route to chaos?

Controlled chaos in a single electronic device!

\[ \lambda = \lim_{t \to \infty} \lim_{\delta Z_0 \to 0} \frac{1}{t} \ln \frac{|\delta Z(t)|}{|\delta Z_0|} \]

Nonlinearity and bifurcations

**The missing pieces of device models:**

1. Minimization of internal energy or $\Delta H$
2. Spontaneous symmetry breaking during nonlinearity
3. Amplification and coupling of ambient thermal fluctuations in nanoscale devices

\[ j_U \cdot A = j_L \cdot (A - x \cdot A) + j_H \cdot (x \cdot A) \]

A different route to chaos in nonlinear devices.

Kumar *et al.*, *Nature Communications*, (2018)
Kumar *et al.*, *Advanced Materials*, 28, 2772 (2016)
Synaptic devices
Ion migration in Hf and Ta oxides

HfOx+Hf device
Material stack: Pt/HfOx/Hf/Pt

Operated
Pristine/Virgin

THERMAL RESET
200 C

X-ray intensity (a.u.)
2 µm

I_{570eV}
300 nm

O K-edge
Non-volatile switching and neuronic devices

The origin of nonvolatile storage in ReRAM

(a) Graph showing current-voltage characteristics with i-sweep and v-sweep.

(b) Image with color scale indicating temperature.

(c) Image with 2 μm scale showing High and Low states.
The origin of non-volatile storage
Fundamental requirement for S-NDR $\rightarrow$ non-volatility

\[ i = Gv \]
\[ G = \psi T^\xi \]

Steady state: $\frac{dT}{dt} = 0$

\[ T = T_{amb} + R_{th} v^2 \psi T^\xi \]

\[ \frac{dT}{di} = R_{th} v^2 \psi \xi T^{\xi-1} \frac{dT}{di} + 2R_{th} \psi T^\xi \frac{dv}{di} \]

\[ R_{th} v^2 \psi \xi T^{\xi-1} = 1 \]

\[ T = T_{NDR} = \frac{\xi T_{amb}}{\xi - 1} \quad \rightarrow \xi > 1 \]

Estimate only this material parameter and one can predict most of the nonlinear and information storage behaviors.
A brain-like computer using synaptic and neuronic oxide memristors
The traveling salesman problem

Objective:

Find the shortest path

Constraints:

1. Visit every city once
2. Visit every city no more than once
3. Do not visit more than one city in a given stop

A traveling salesman wants to visit every city in his territory.

Finding the shortest route is easy for a few cities. But the problem grows complex rapidly.
“Hard” problems

It is non-deterministic polynomial (NP) complete.

Other NP-complete/hard problems:

Gene sequencing/traveling salesman

Sudoku

Vehicle routing

Open shop scheduling

Bandwidth allocation

NFL scheduling

256 games

20k variables

50k constraints

Takes ~3 months on a 1000-core system to solve!
Hopfield network

Recursive feedback neural network

Energy function:

\[
E = -\frac{1}{2} \sum_i \sum_j s_{i,j} \sum_k \sum_l s_{k,l} w_{(i,j),(k,l)} + \sum_i \sum_j s_{i,j} \theta
\]

An incredibly compact transistorless all-analogue implementation of a Hopfield network

US Patent App. 15/141,410
Example of one solution with chaos

The traveling salesman problem

Statistics of many solutions with and without chaos

We only want *better* solutions quickly.

High precision $\rightarrow$ prohibitive slow downs

Literally annealing the system into its solution!

Room temperature analogue of quantum adiabatic annealing

Mem-HNN outperforms digital hardware and quantum annealing

Benchmarking on a suite of 60 node max-cut problems

<table>
<thead>
<tr>
<th></th>
<th>Analog</th>
<th>Digital</th>
<th>Quantum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mem-HNN</td>
<td>GPU</td>
<td>D-wave 2000Q</td>
</tr>
<tr>
<td>Clock frequency</td>
<td>1 GHz</td>
<td>1.5 GHz</td>
<td></td>
</tr>
<tr>
<td>Time-to-solution</td>
<td>0.3 µs</td>
<td>10 µs</td>
<td>$10^4$ s</td>
</tr>
<tr>
<td>Power</td>
<td>792 mW</td>
<td>&lt;250 W</td>
<td>25,000 W</td>
</tr>
<tr>
<td>Solutions/s/Watts</td>
<td>$4.6 \times 10^6$</td>
<td>&gt;400</td>
<td>$4 \times 10^{-9}$</td>
</tr>
</tbody>
</table>

10,000x better Solutions/sec/Watt !!

arXiv:1903.11194: Kumar et. al. Harnessing Intrinsic Noise in Memristor Hopfield Neural Networks for Combinatorial Optimization

Major Caveat: Circuit level analysis only – no scalable architecture analyzed yet

Reason for optimism: NP-hard problems are compute-intensive, not data-intensive
Summary

– Analog, brain-like computing can outperform any digital alternative in solving certain classes of intractable problems

– But we need new device physics to do that
  – Controlled chaos
  – Nano-size coupling with ambient thermal fluctuations
  – Enthalpy minimization and symmetry breaking
  – Super-linearity and non-volatility (the connection between neurons and synapses)