Josephson Junction Based Neuromorphic Computing

Mike Schneider, Stephen Russek, Christine Donnelly, Burm Baek, Matt Pufall, Ian Haygood, Pete Hopkins, Paul Dresselhaus, Sam Benz, Bill Rippard
Artificial Intelligence

Machine Learning

Bio – Inspired

Neuromorphic Hardware

Deep Neural Networks

eyeriss.mit.edu
Software neural networks are mainstream

- Apple: A11 bionic SoC, Siri
- Google: Search results (RankBrain), translate, street view, ...
- Microsoft: Computational Network Toolkit (speech recognition)
- Amazon: from product recommendation to robotic picking routines
- Facebook Ranking, Language translation, ...

M. Nielsen, “Neural Networks and Deep Learning” (2015)
Neural networks dominate modern image recognition

ImageNet Competition

- Deep Convolutional Neural Network (Trained with GPUs)
- Human Accuracy ~ 5% error rate

Feature based algorithms
What interaction is being modeled in neural networks?

$w_0 \times x_0$

$w_1 \times x_1$

$f \sum_i w_i x_i + b$

Output
## Neuromorphic Single Flux Quantum

**Information transfer:** quantized pulse trains

**Long distance “lossless” pulse transmission**

**3D architecture**

**Memory/ plasticity**

<table>
<thead>
<tr>
<th>Neural</th>
<th>Single Flux Quantum</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Neuronal activity" /></td>
<td><img src="image2" alt="Josephson/JJ synapse" /></td>
</tr>
</tbody>
</table>

- **Neuronal axons:**
  - Voltage (mV) vs. Time (ms)

- **Josephson/JJ synapse:**
  - Voltage (mV) vs. Time (ps)

- **Magnetic JJ synapse**
Josephson junction (JJ)

\[ \psi_1 = \psi_{01} e^{-i\theta_1} \]
\[ \psi_2 = \psi_{02} e^{-i\theta_2} \]

\[ I_s = I_c \sin \theta \]

- Barrier = insulator, semiconductor, normal metal, ferromagnet, ...
Circuit Model of a Josephson Junction

\[ \frac{\varphi_0}{2\pi} C \ddot{\theta} + \frac{\varphi_0}{2\pi R_n} \dot{\theta} + I_c \sin(\theta) - I_b = 0 \]

Mass Damping

\[ U = \frac{I_c \varphi_0}{2\pi} (-\cos(\theta) - \frac{I_b}{I_c} \theta) \]

where \( \varphi_0 = \frac{\hbar}{2e} = 2.03 \times 10^{-15} \text{Vs} \)
Interaction of order parameters

\[ \uparrow \downarrow - \downarrow \uparrow \]

Singlet (S = 0)

\[ \uparrow \downarrow e^{iQ \cdot R} - \downarrow \uparrow e^{-iQ \cdot R} \]

\[ = (\uparrow \downarrow - \downarrow \uparrow) \cos(Q \cdot R) \]

\[ + i(\uparrow \downarrow + \downarrow \uparrow) \sin(Q \cdot R) \]

**Singlet** + Triplet

(S = 0) (S = 1)

Fulde-Ferrell-Larkin-Ovchinnikov (FFLO) state

Superconducting Spintronics Reviews

- Buzdin, Rev Mod Phys 2005
- Bergeret, Rev Mod Phys 2005
- Eschrig, Phys Today 2011
- Blamire, J of Phys Cond Matt 2014
Magnetic nanoclusters in a Josephson junction

Thin Films

Si

Mn

Nb

ordered

disordered

Moment (nAm²)

Temperature (K)

Moment (nAm²)

Temperature (K)

0 100 200 300
0
30
60
90
10 mT
5 mT
0.5 mT
0 mT

10 mT
5 mT
0.5 mT
0 mT
Changing magnetic order with fast electrical pulses in devices

Small applied magnetic field + 250 ps Electrical pulse

H=0

Only the small applied magnetic field
The basic neuro-inspired cell

\[ f \sum_i w_i x_i + b \]

Presynaptic SFQ Neuron

Postsynaptic SQUID Neuron
Magnetic Josephson junction neural network simulation

- 9 input pixel (JJs)
- 27 weights (MJJJs)
- 3 outputs (MJJ SQUIDs)
Real time (~ 3ns) recognition

Prezioso et al. Nature 2015
Real time (~3ns) recognition

Prezioso et al. Nature 2015
Real time (~3ns) recognition

Prezioso et al.
Nature 2015
Large scale image recognition

- ARGUS
- 1.8 billion pixels at 12 fps
- resolution of 6 inches over 10 square miles

- Currently pre-processed on board using kilowatts
- Full stream processed on a supercomputer
Scale to a large network

- Recent ImageNet winners use $\sim 10^{10}$ MA-operations / image

- $\sim 10^{17}$ MA-operations /sec to process ARGUS real time

- JJ neural network would take $\sim 3$ watt (including cooling) (1 spike / MA-operation)
Benchmarks

• **Energy**
  • Operational energy < $10^{-18}$
  • Training energy < $10^{-17}$
  • Cooling overhead $\sim 10^3$
  • (Human brain $\sim 10^{-15}$)

• **Speed**
  • Operational (device) $\sim 100$ GHz
  • Operational (circuit) $\sim 10$ GHz

• **Size**
  • Demonstrated 1.5 x 3 µm
  • Demonstrated for MJJ 100 nm

• **Scalability**
  • Demonstrated $\sim 10^1$
  • There is a lot of scaling to go...
Conclusions

• Josephson junctions are a very promising neuromorphic computing technology

• Magnetic Josephson junctions are a cryogenic non-volatile memory

• Magnetic Josephson junctions can be used for the synaptic function

• There is a lot of scaling needed for Josephson junction based neuromorphic computing