

Does the Future of Computing Lie in Software or in Hardware?

Waleed El-Geresy and Adil Malik

Imperial College
London

Outline

- Perspective: two philosophies (big data vs design)
- Is hardware or software the solution?
- Examples from our research
 - Memristor Spike Timing Dependent Plasticity
 - Joint Source Channel Coding on Memristors
- Quantum computing

Assumption: The Future of Computing is AI

- Artificial Intelligence
- Key features
 - Noise-tolerant computing
 - Complex and non-interpretable
 - Data driven
 - Customised
 - Requires data as a resource
 - Increasingly independent of human input

Imperial College
London

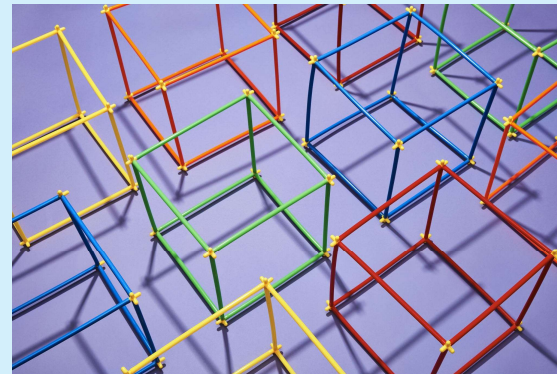
Two Philosophies

Power



*Better energy efficiency
and scaling*

Design



*New designs and
software approaches*

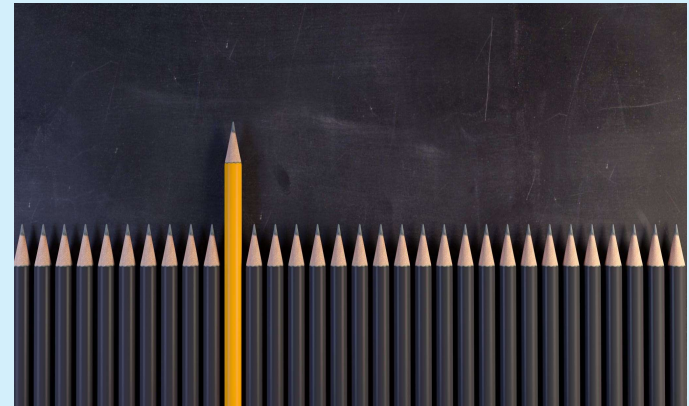
Energy Efficiency and Scale

- Transistors reaching smallest size possible
 - Energy density is a limitation
 - Quantum effects
- Hardware mismatch with novel algorithms leads to inefficiency
- What we need is more data
 - Fundamental statistical limits on generalisation and bias
 - Deep learning – needs more data
 - Large-scale language models (billions of parameters)
- Way to acquire more data - online data acquisition and training
- Pruning and model compression (PET's 223M vs GPT-3's 175B)



Design

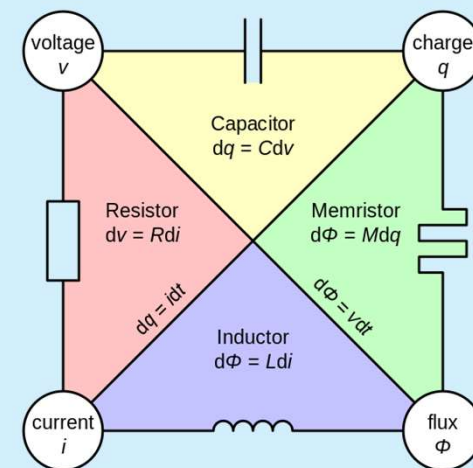
- Making better use of what we have
- Learning from nature and the brain
 - Millions of years of existing evolution
- Fundamental argument: need the right biases
 - Convolutional neural networks
 - Reinforcement learning
 - Attention mechanisms
- Algorithms that are robust to noise
 - Spatiotemporal data
 - Temporal coding (implicit temporal correlations)
- Hardware cannot be assumed



Quality over quantity

Memristor Devices

- Devices that were touted as the “fourth fundamental circuit element” by Leon Chua
- Applied voltage changes their resistance
- Storage and processing occur in the same location
- “Memory resistors”

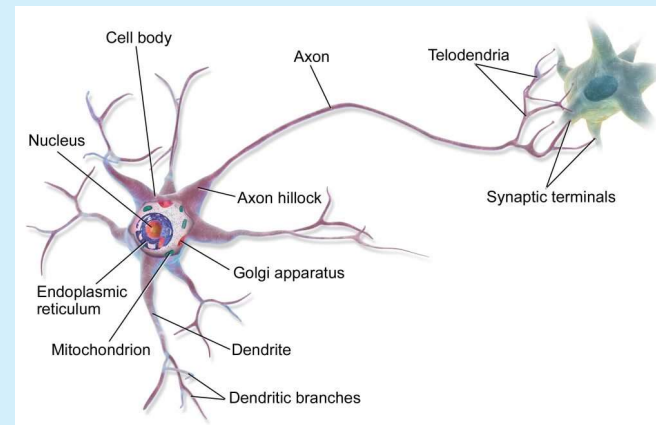
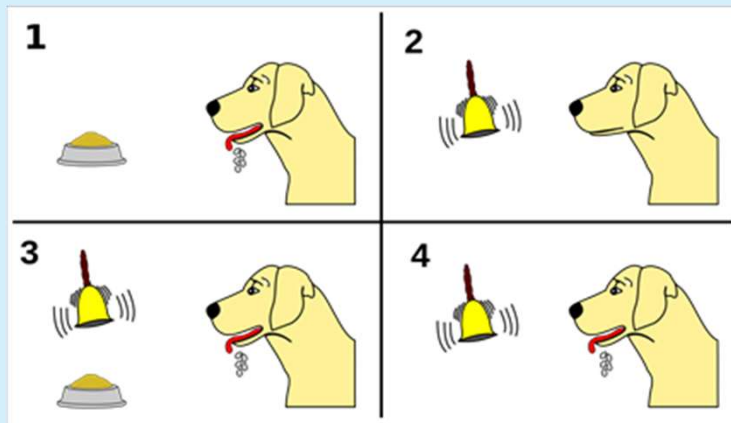


Neuromorphic Computing with Memristors

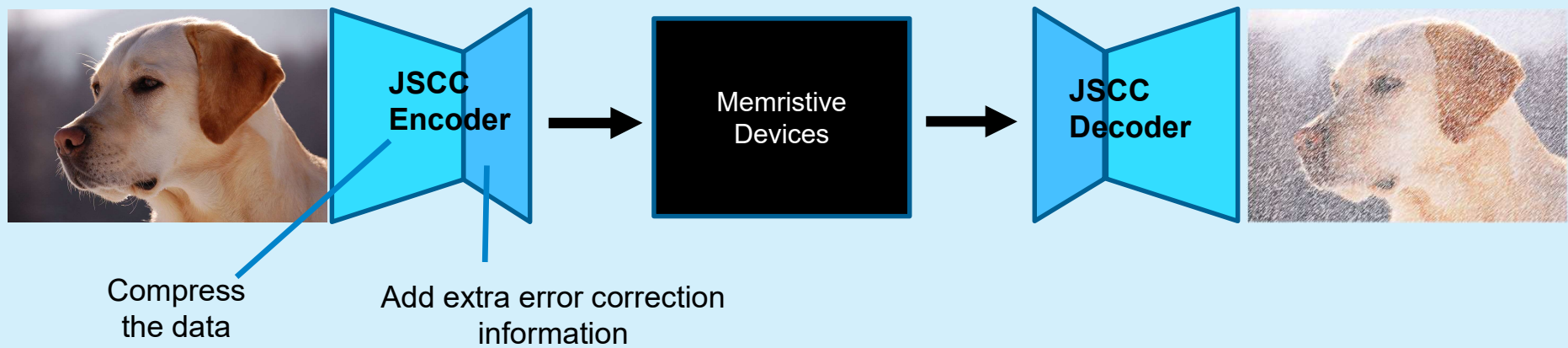
- In the words of Carver Mead: “But we pay... for taking all the beautiful physics that is built into those transistors, mashing it down into a 1 or a 0, and then painfully building it back up.” (Mead, 1990)
- **Elementary operations** are dictated by individual device physics, cutting out the costly digital conversion and deconversion. We use the operations in their native, analogue form, since we represent information in an analogue form.
- Increased power efficiency for the types of software that AI is seen to be dependent upon.

STDP on Memristors

- Neuromorphic computing: fundamental computational primitives change
- A method of online learning which naturally interacts with stimulus



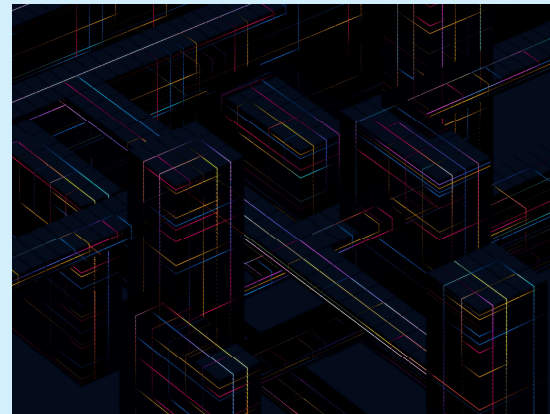
Joint Source-Channel Coding on Memristors



Imperial College
London

Quantum Inspired Approaches

- Massive parallelism
- Non-determinism
- Quantum consciousness and free will
- Quantum computing
- Security and cryptography (RSA)



Imperial College
London

Questions and Discussion

