Magnetism for Neuromorphic Computing

Tom Hayward

Functional Magnetic Materials Group

Department of Materials Science and Engineering

The University of Sheffield

Outline

1. Introduction to Neuromoprhic Hardware.
2. Feed Forward Neural Networks.
3. Spintronic Neural Networks.
4. Reservoir Computing.
5. Spintronic Reservoir Computing.

AI is Amazing...

Write me a short poem about why AI is cool

The Wonder of AI

Born of code, a spark in the mind,
Infinite answers, endlessly kind.
From puzzles unsolved to dreams unfurled,
AI breathes life into a boundless world.

It learns, it grows, it shares, it speaks,
Reaching the heights that curiosity seeks.
A partner in thought, in art, in play,
It shapes tomorrow from seeds of today.

With logic and wonder, it breaks every rule,
That's why we say: AI is cool.

Healthcare

Transport

Customer Service

Finance

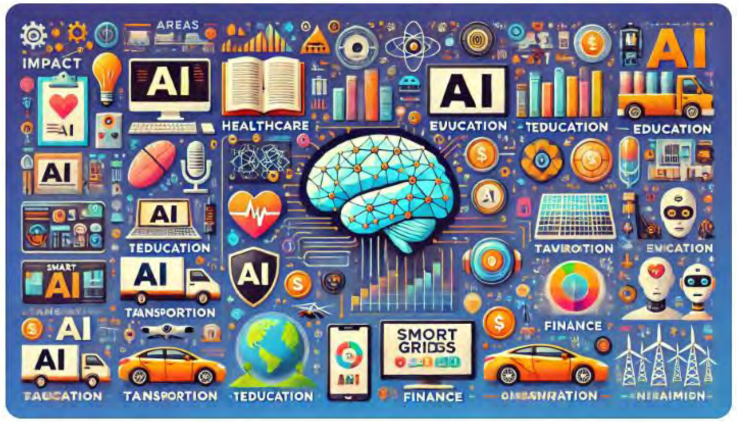
Entertainment

Energy

Manufacturing

Security

Make me an info-graphic showing the areas AI may have impact



Here is an infographic showcasing the areas where AI has significant impact. It visually represents sectors like healthcare, education, transportation, environment, finance, and entertainment with clear icons and vibrant design elements. Let me know if you'd like any edits or additional sections!

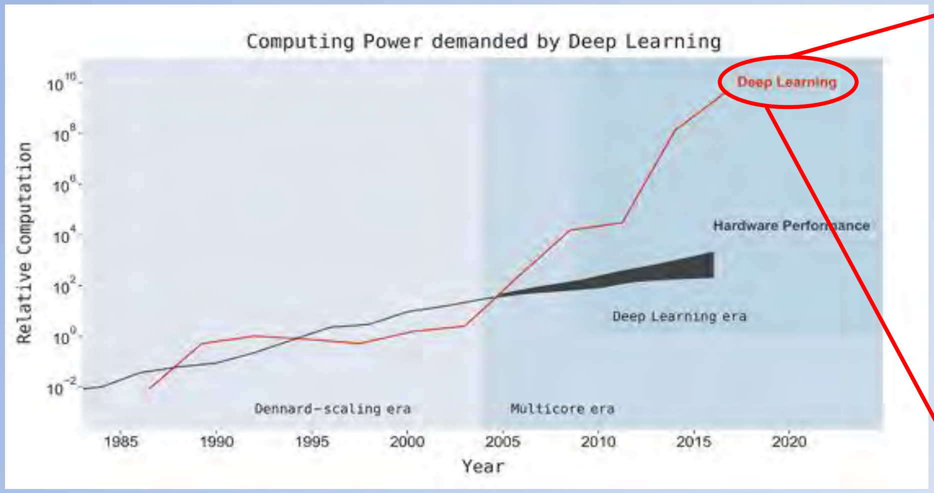
Can you write me code to generate a pytorch neural network with 2 inputs, 2 hidden layers with ten nodes and 1 output node?

Here's the code to define a PyTorch neural network with the specified architecture:

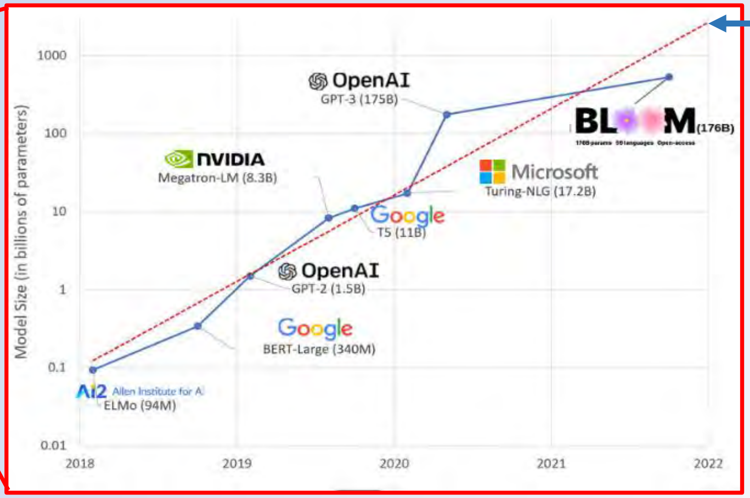
Pytorch Nn 2inputs

```python
import torch
import torch.nn as nn

class SimpleNeuralNetwork(nn.Module):
    def __init__(self):
        super(SimpleNeuralNetwork, self).__init__()
        # Define the network layers
        self.fc1 = nn.Linear(2, 10)  # Input layer to first hidden layer
        self.fc2 = nn.Linear(10, 10)  # First hidden layer to second hidden layer
        self.fc3 = nn.Linear(10, 1)  # Second hidden layer to output layer
        self.activation = nn.ReLU()  # Activation function

    def forward(self, x):
        # Define the forward pass
```

…but has problematic impacts…



‘The Computational Costs of Deep Learning’, N.C. Thompson *et al. (2022)*

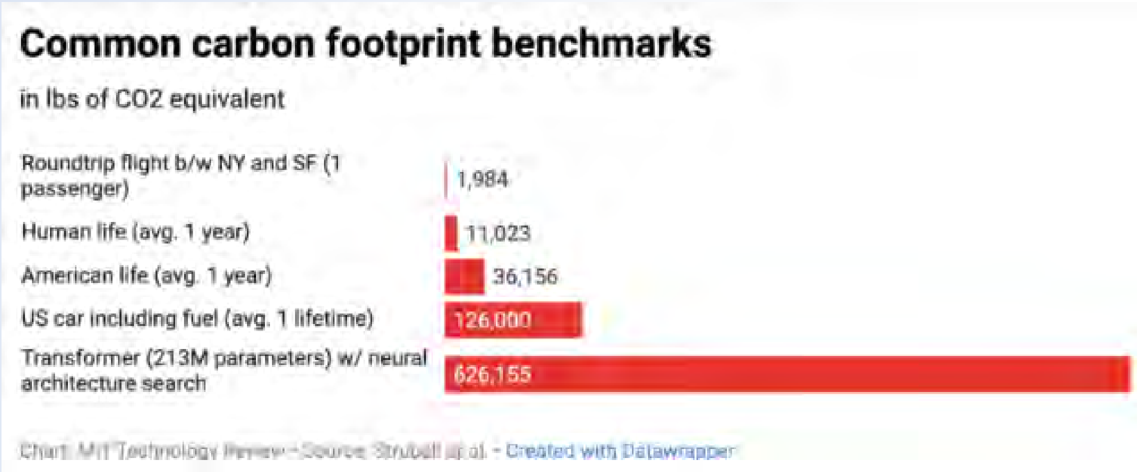


GTP-4o 1.8T params

“The mounting human and environmental costs of generative AI” Ars Technica (2023)

Scale of AI models is increasing faster than performance of hardware…

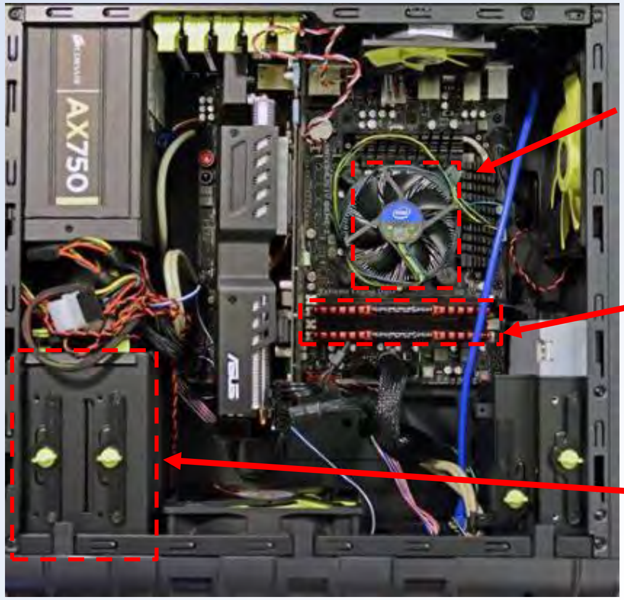
With terrifying consequences for power usage and carbon emissions…

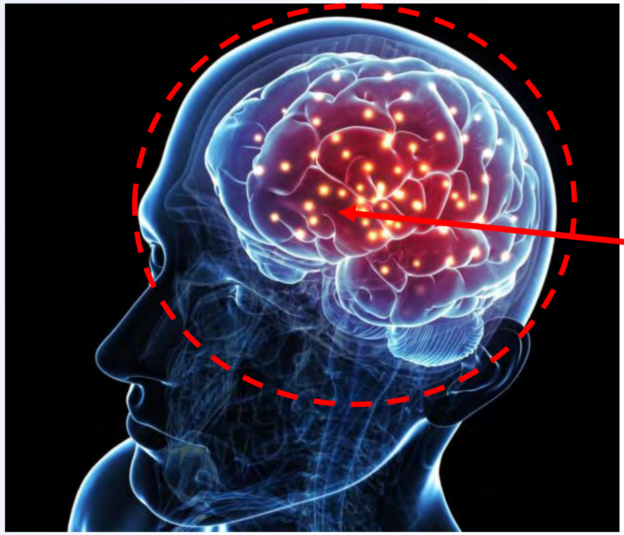


…because computers aren't like brains!

- While **modern CPUs/GPUs** can train and run large ANN this is only because CMOS technology is so good at maths!
  - Brute force **simulate** of neural architectures by doing vector maths.
- **Power inefficient**!
  - Training one deep neural network has carbon footprint of entire lifecycle of a car.
  - Human brain consumes ~20 W, equivalent HPC would consume ~10 MW.

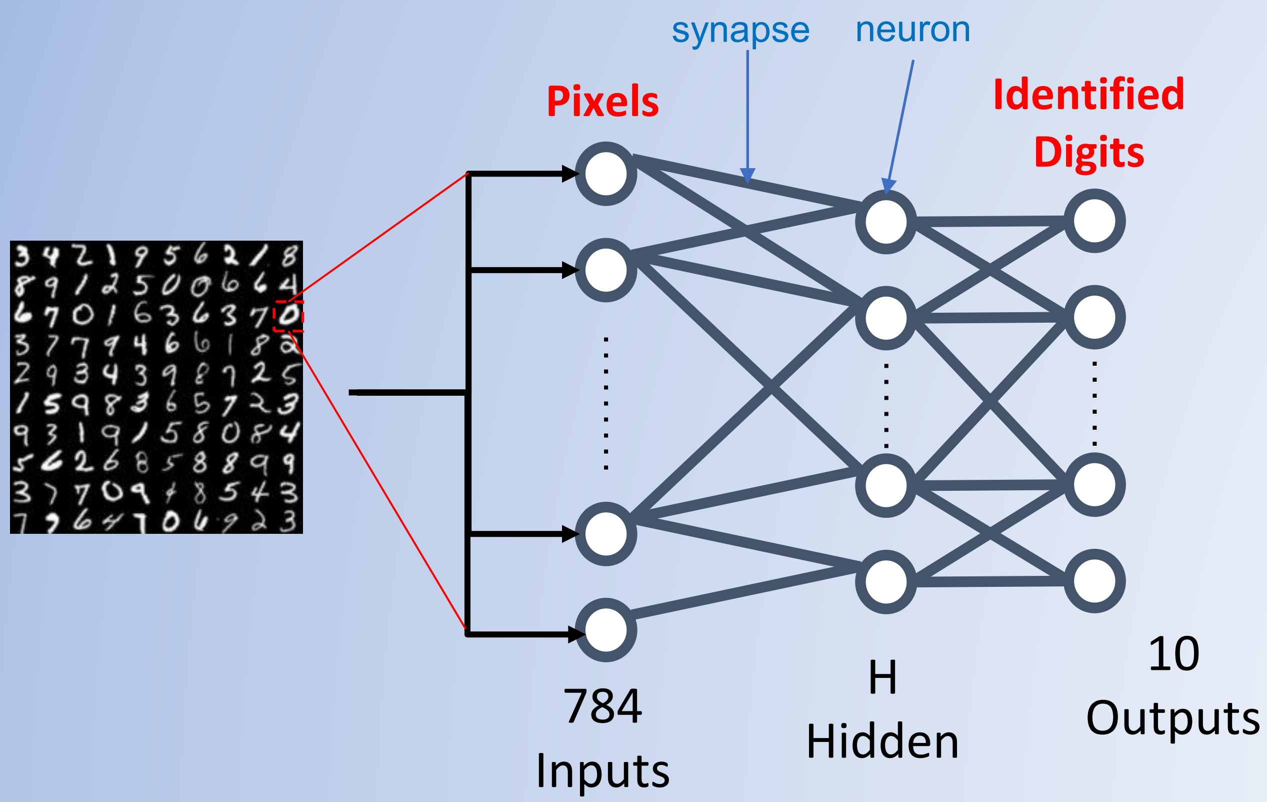
Neuromorphic technologies use the properties of functional materials to attempt to more directly **emulate neural architecture**.





Feed-forward Neural Networks

- Typical approach represented by **feed-forward neural network**.
  - Consists of layers of **neurons** connected by **unidirectional synapses** with analogue tuneable **weights w_i**.
  - Weights are "**trained**" to allow network to perform task (e.g. identification of hand-written digits).



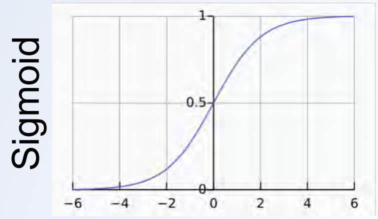
$$y_i = f(k_i) \qquad k_i = \sum_j w_{ij} x_j + b_i$$

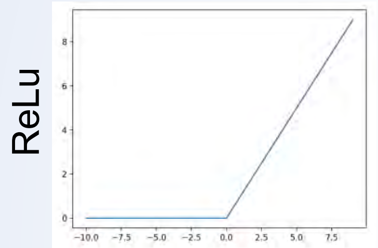
Neuron output

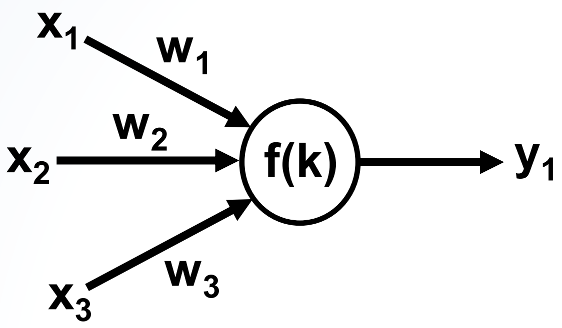
Weight

Synapse input

Bias







Neuron/Perceptron

Training FFNN

- Most commonly train FFNN/Multilayer Perception **weights** by **supervised learning**.
- Minimise error of network against a **training set (x_{target},y_{target})**:

$$C = (\mathbf{y}_{\mathbf{target}} - \mathbf{y}_{\mathbf{output}})^2$$

loss function

- Train weights by iterative **gradient descent**:

$$\mathbf{w}_{\mathbf{n+1}} = \mathbf{w}_{\mathbf{n}} - \gamma\, \nabla C(\mathbf{w}_{\mathbf{n}})$$

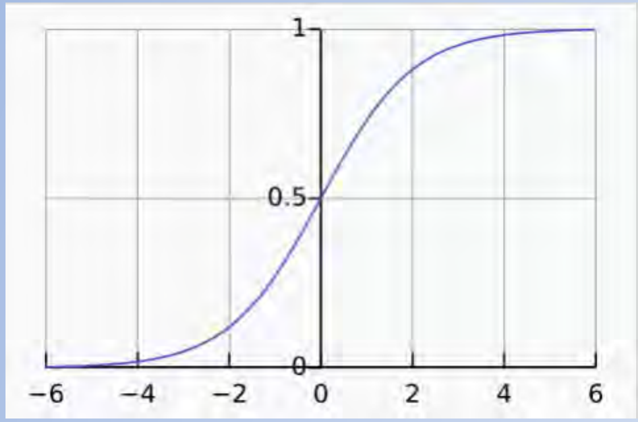
learning rate

$$\nabla C(\mathbf{w}_{\mathbf{n}}) = \frac{\partial C}{\partial \mathbf{w}_{\mathbf{n}}}$$

- Gradient function calculated using **backpropagation algorithm** (applies chain rule through network to understand how loss depends on weights).
- Iterate until error converges and test network on **unseen data (test set)**.

What do we need to make a magnetic neural network?

Neuron

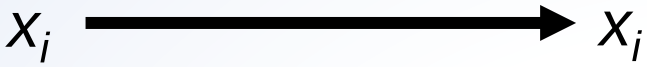


- Addition of **synaptic inputs**.
- Apply **non-linear activation**.

Synapse

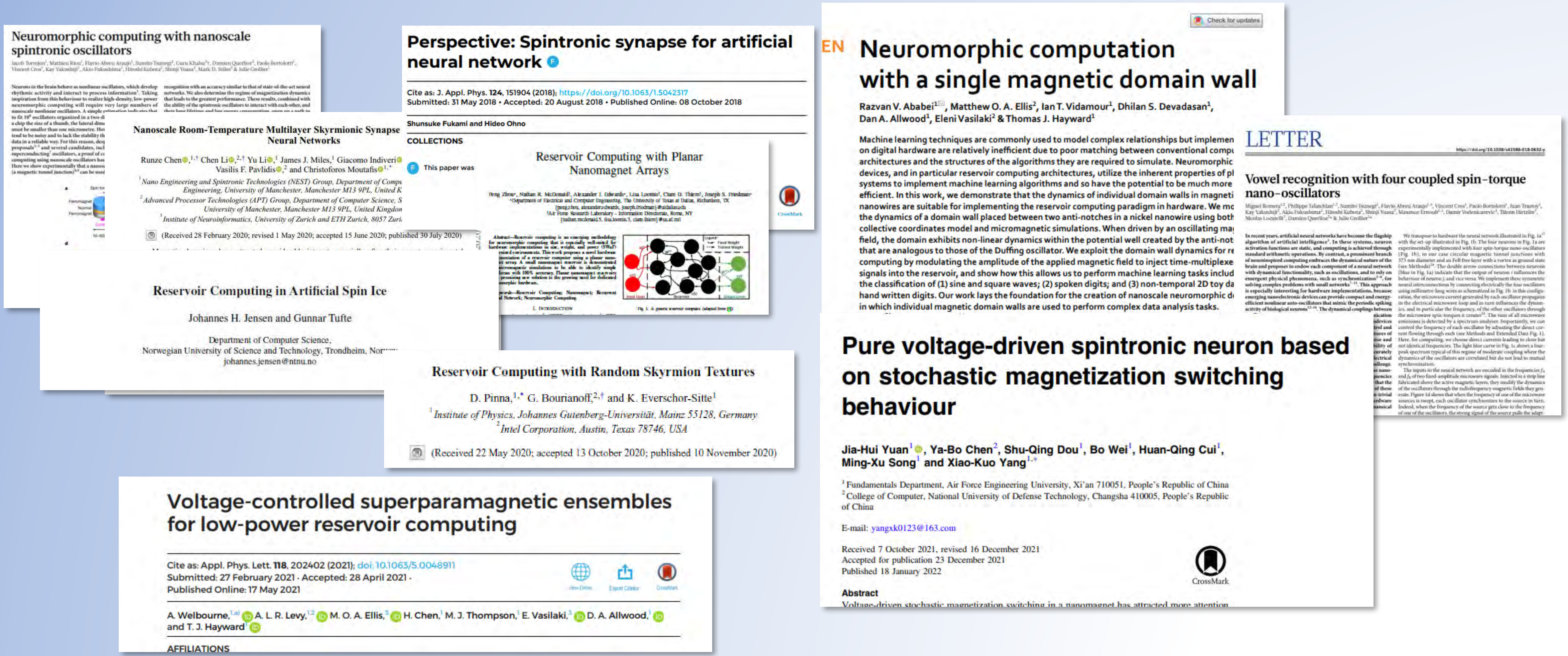
- **Multiplies** input by **weight**.
- Weight ideally local and non-volatile

Interconnect



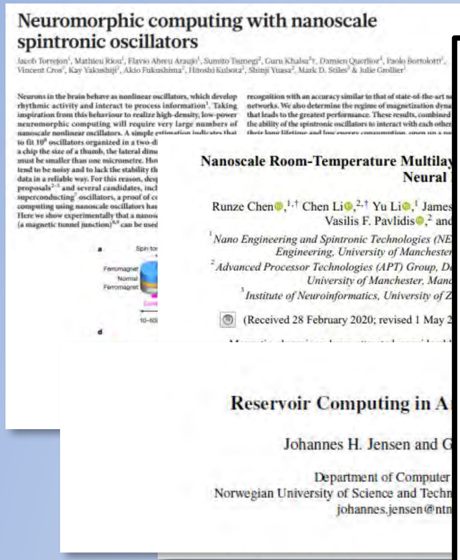
- **Transports** data through the network.

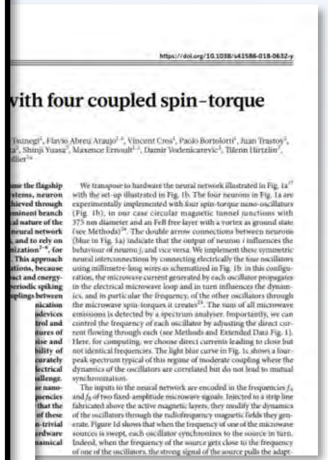
Lots of magnetism groups working on this!



- I'm just going to **mostly talk** about **our efforts** in the area…

Lots of magnetism groups working on this!



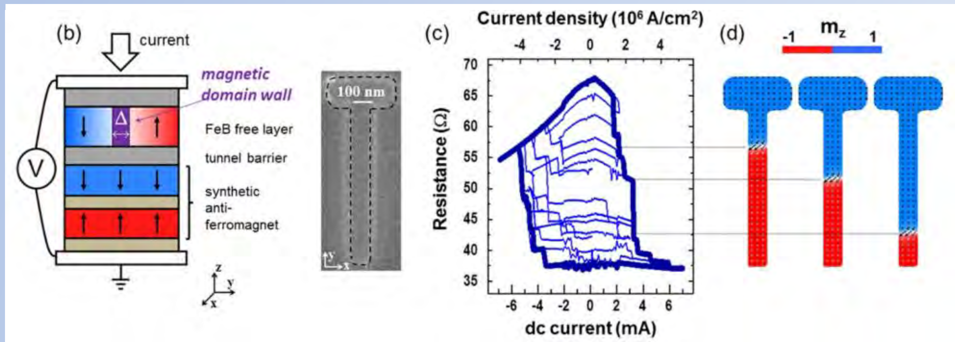


- **Why?**
  - Non-volatile memory!
  - Inherent non-linearity!
  - CMOS compatibility!
  - High speed (up to THz)!
  - Low power!
  - Physics based interconnects (exchange, magnetostatic coupling)!

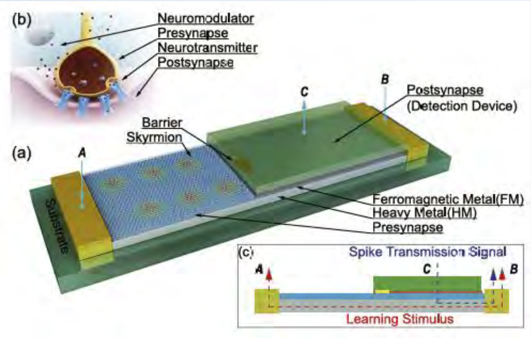
- I’m just going to **mostly talk** about **our efforts** in the area…

Spintronic Synapses

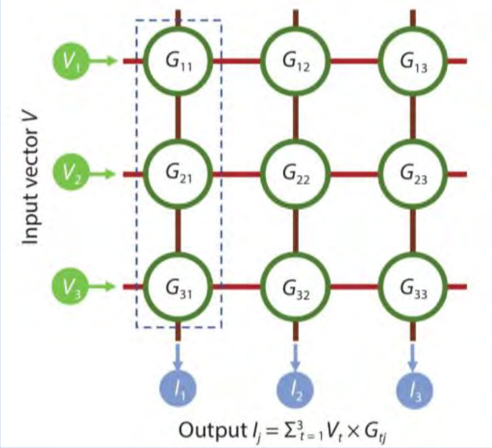
- Most proposals for **magnetic synapses** are based magnetoresistive memristors.
- Stores an analogue weight using e.g. **position of DWs/Skyrmions** in a multilayer nanotrack…



S. Lequeux *et. al.* Sci. Rep. **6,** 31510 (2016)



Y. Huang *et. al.* Nanotechnology **28**, 08LT02 (2017)

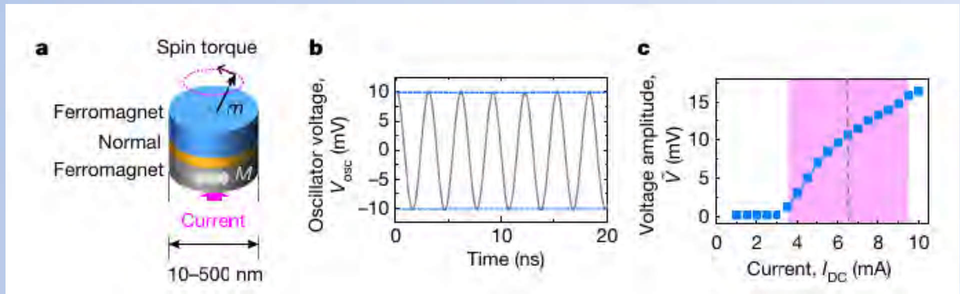


Integrating into **cross-bar arrays** allows **multiply and accumulate** via Kirchoff's law…

Spintronic Neurons

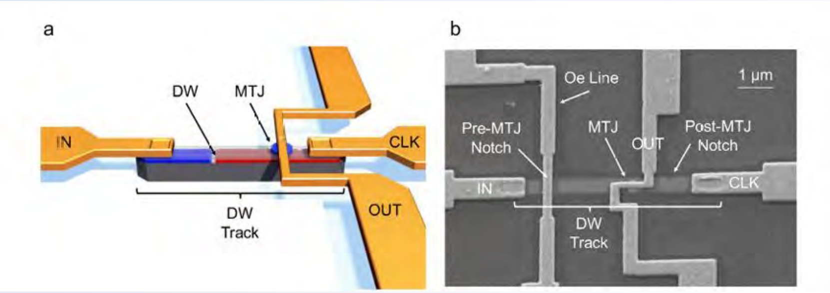
- **Spintronic neurons** produce a non-linear or step-like response to current input.

STO-based leaky integrator neuron



Torrejon et. al. Nature **547**, 428 (2017)

DW-based Spiking Neuron

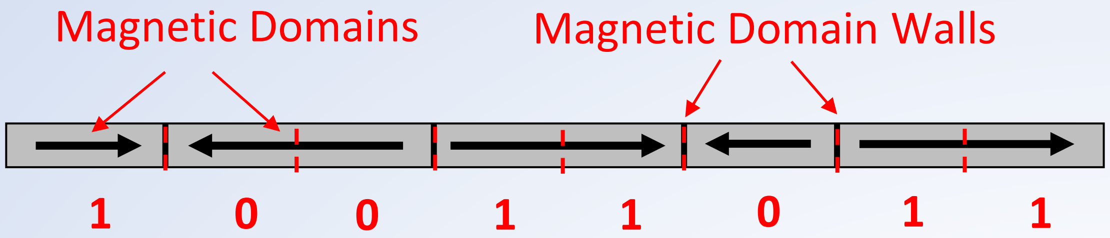


Leonard et. al. Appl. Phys. Lett. **122**, 262406 (2023)

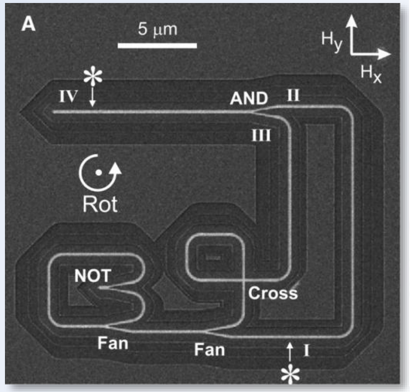
- Spintronic synapses and neurons would typically be integrated into a **CMOS architecture** to create neural circuits (i.e. **interconnects are electrical**)

Domain Wall Devices

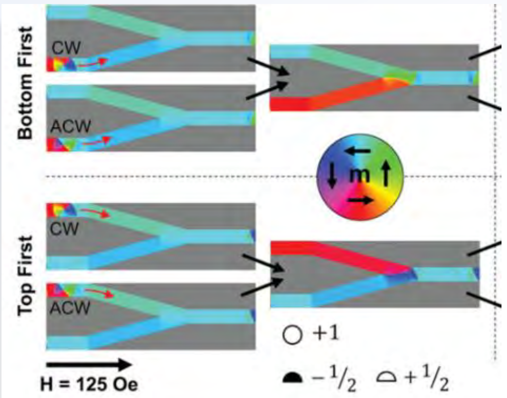
- The devices we've looked at so far could **integrate** with **conventional microelectronics**.
- But could we create **neural networks** that are **all magnetic**?



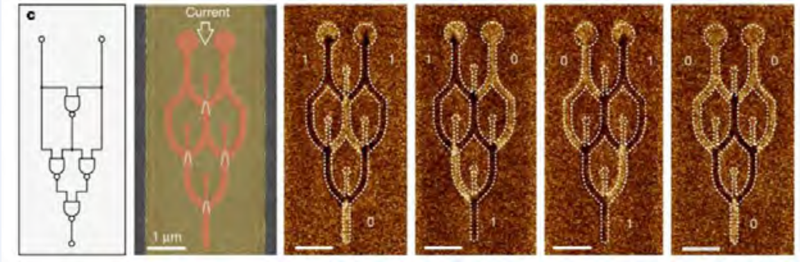
Domain Wall Logic



Allwood et al. Science 309, 1688 (2005)



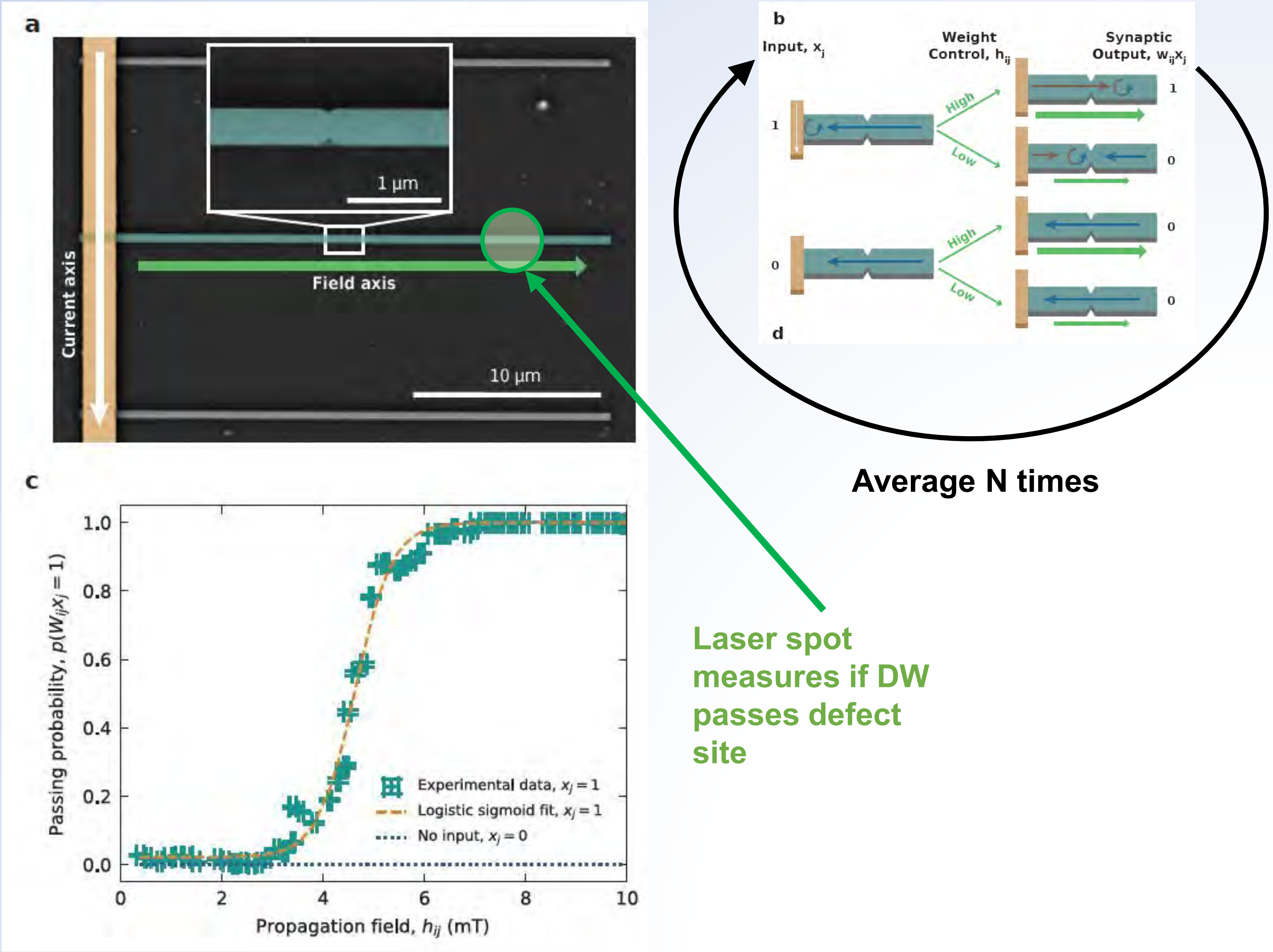
Omari et al. Adv. Func. Mat. 29, 1807282 (2019)



Luo et al. Nature 579, 214 (2020)

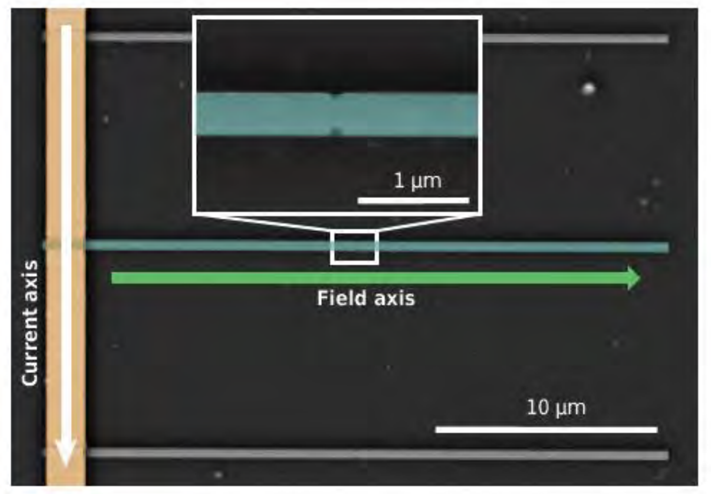
Stochastic Magnetic Synapses

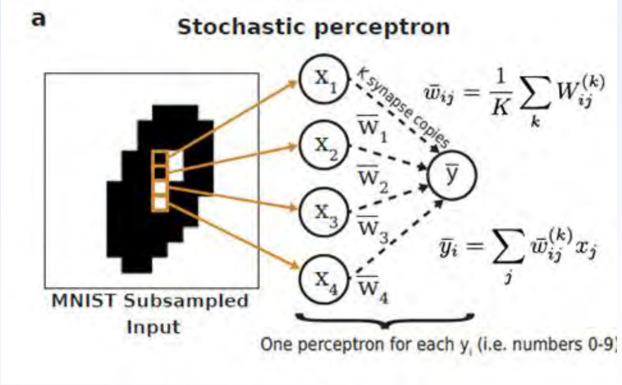
- **How can we create an analogue weight from a digital system?**
  - Domain walls either present or absent!
- DW pinning at defects is stochastic!.
  - Probability of DW passing defect **P_{pass}**.
  - **Problem for digital devices!!**
- Experimental measurements show that **P_{pass}** can be **tuned sigmoidally** between 0 and 1 using externally applied bias field H_{bias}.
- Over repeat measurements create an **analogue response** from the synapse to represent its weight!

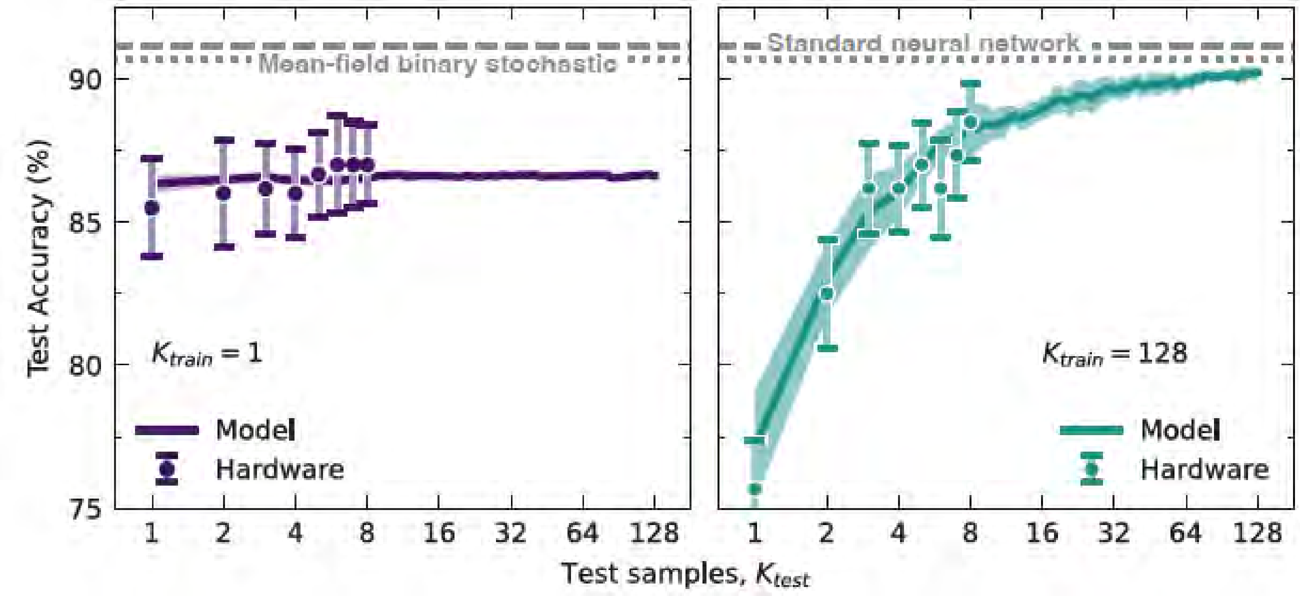


Experimental Demonstration

- Experimentally demonstrate **feasibility** using serial measurements of a **single DW synapse** to mimic network.
  - Identify **written digits** from the MNIST database
- **Successful** experimental implementation of stochastic synapses.
  - **87 % accuracy in MNIST task.**
- Can adapt training to **accuracy/latency** required.
  - Tuneable power/precision!



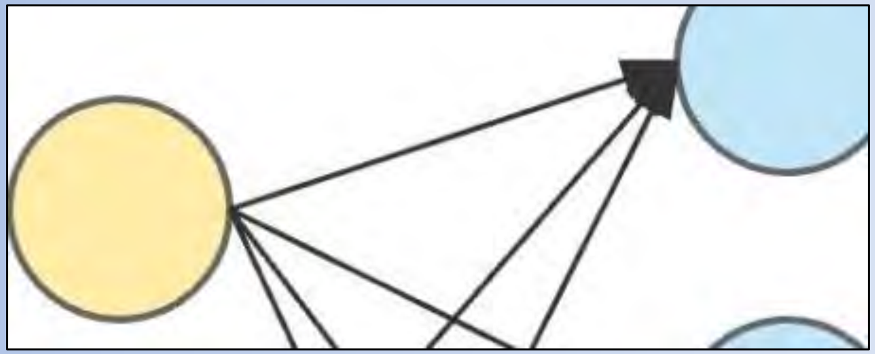


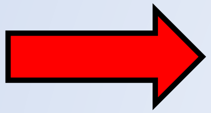


All Magnetic FFNN?

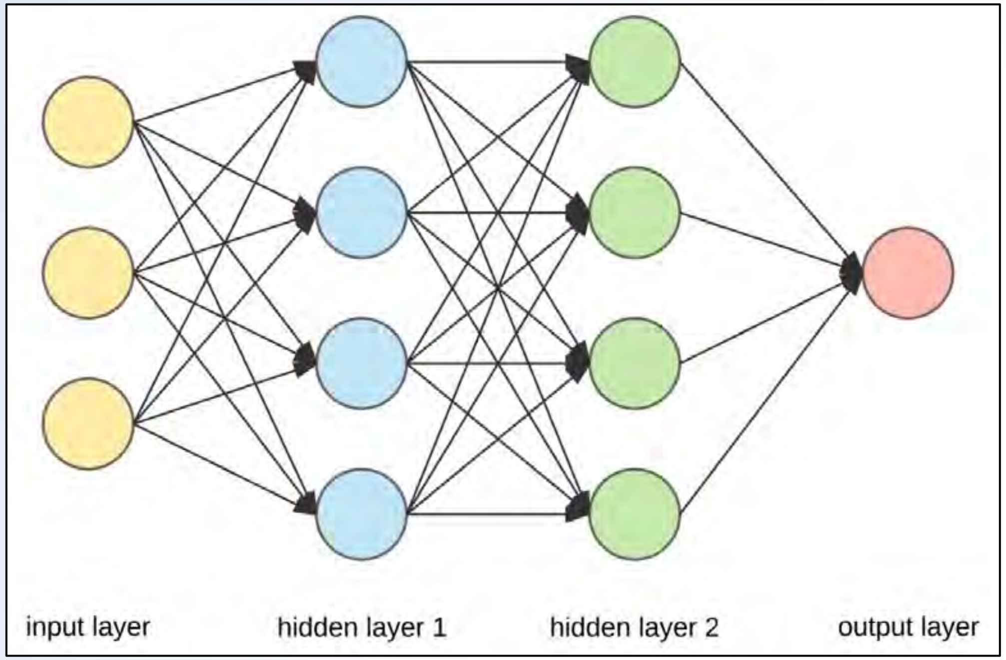
- We’ve made a **synapse**, but how do we create a whole neural network using **magnetic materials alone**?
  - Remove interconnects/Von Neuman bottleneck!

We’ve made this….





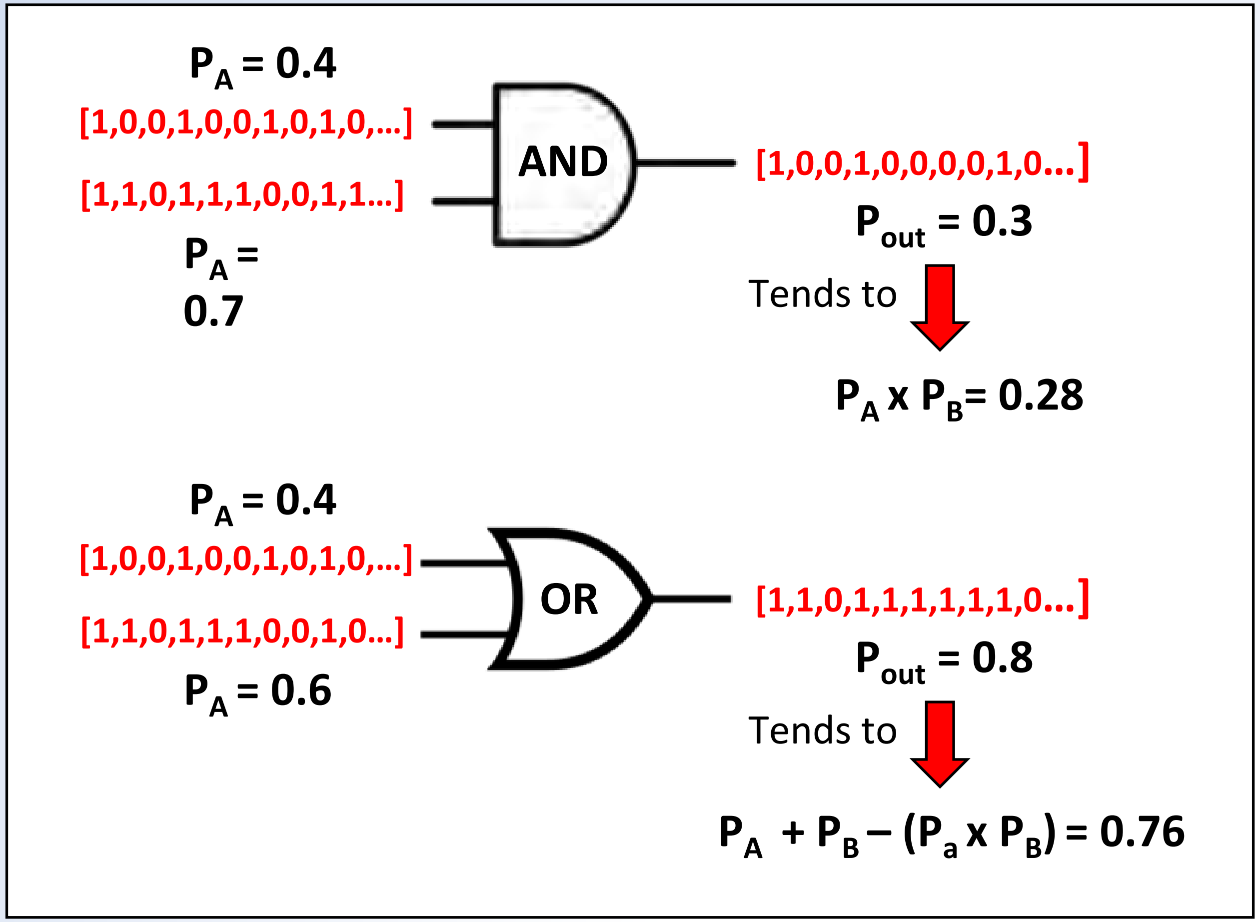
…but we want to make this.



Stochastic Computing

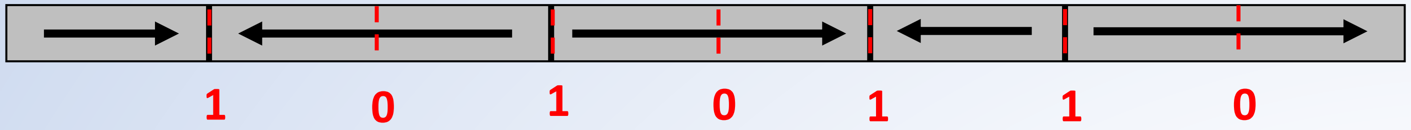
- In **stochastic computing** floating point numbers are represented by streams of binary bits.
  - Fraction of "1"s in a bitstream defines the number.
- Passing these through conventional **logic gates** naturally performs **numerical calculations**:
  - AND = P_A x P_B
  - OR = $P_A + P_B - (P_a \times P_B)$
  - NOT = $1 - P_A$

Provides **rapid estimates** that become more accurate over time.

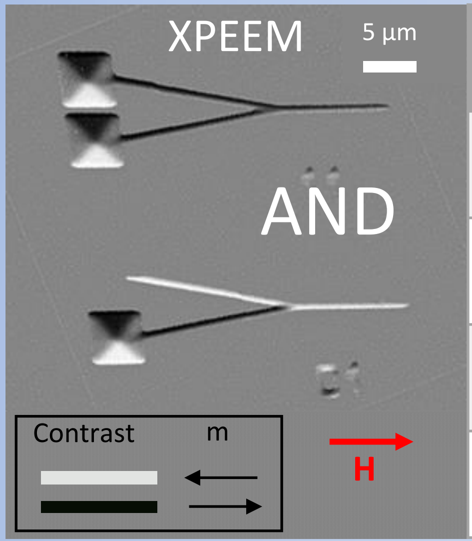


Stochastic Computing with DWs

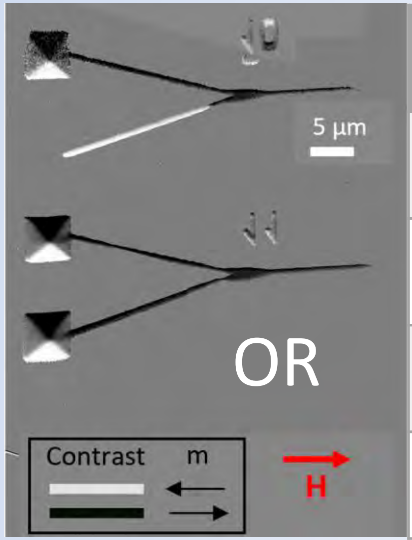
- Represent data by **presence (1) or absence (0) of a DW**.



Deterministic DW Logic Gates



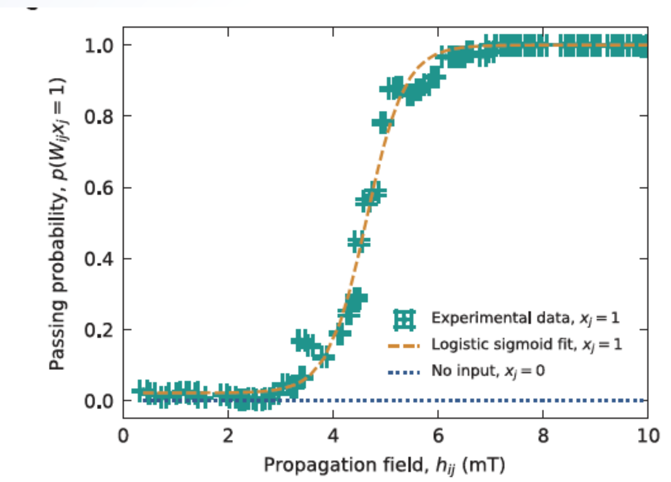
A	B	Out
0	0	0
1	0	0
0	1	0
1	1	1



A	B	Out
0	0	0
1	0	1
0	1	1
1	1	1

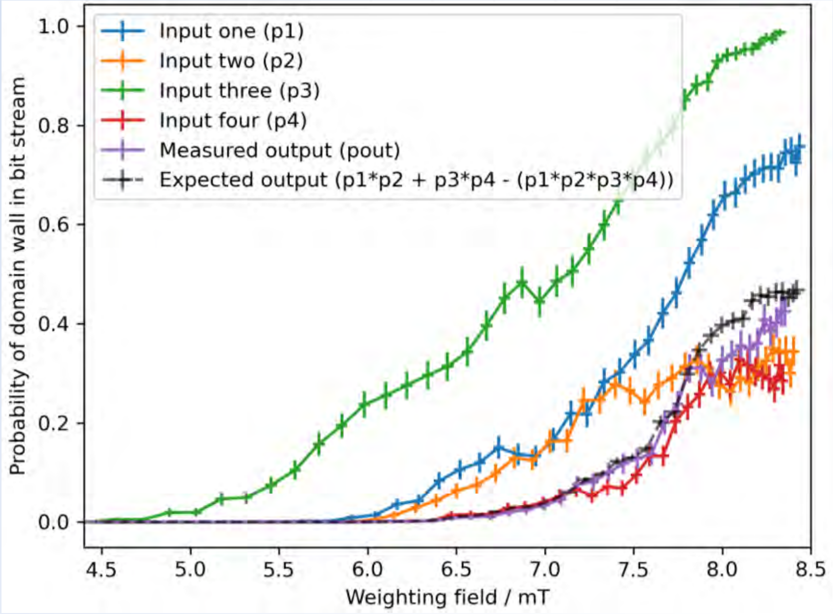
+

Stochastic DW injection



All Magnetic FFNN

Use magnetic nanowires to create a simple **2 synapse, 1 neuron, neural network**.

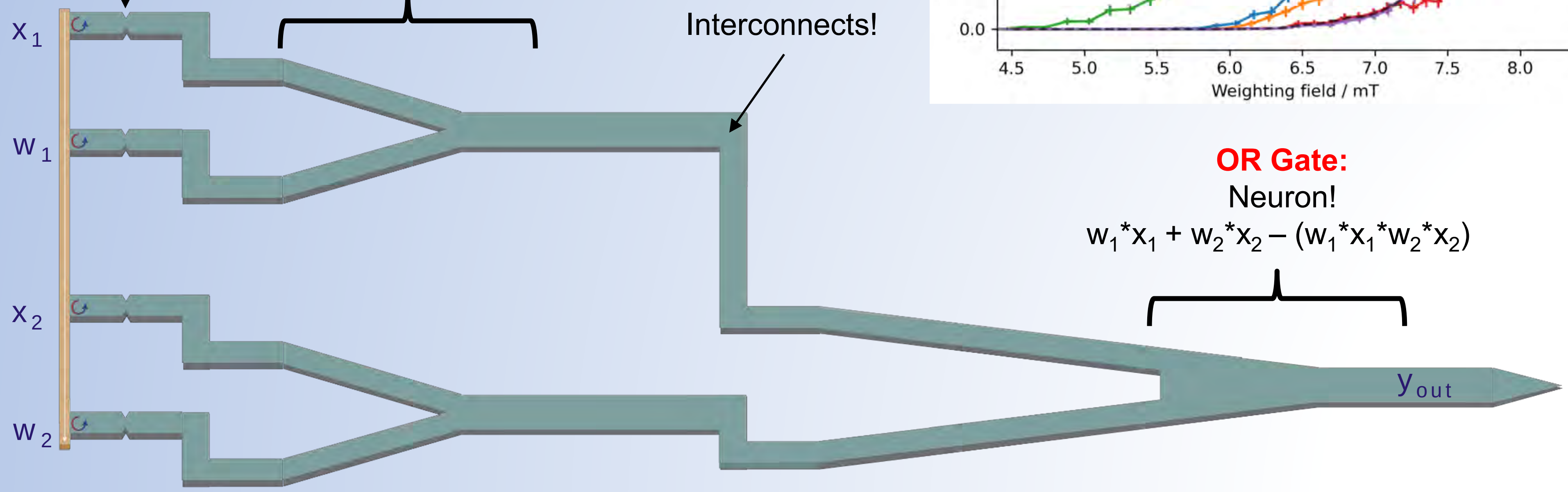


Defects:
Inputs!
Create stochastic bitstreams

AND Gate:
Synapse!
$w_i{}^*x_i$

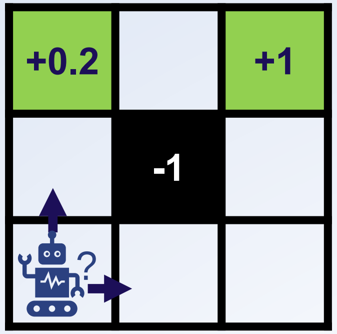
Nanowires
Interconnects!

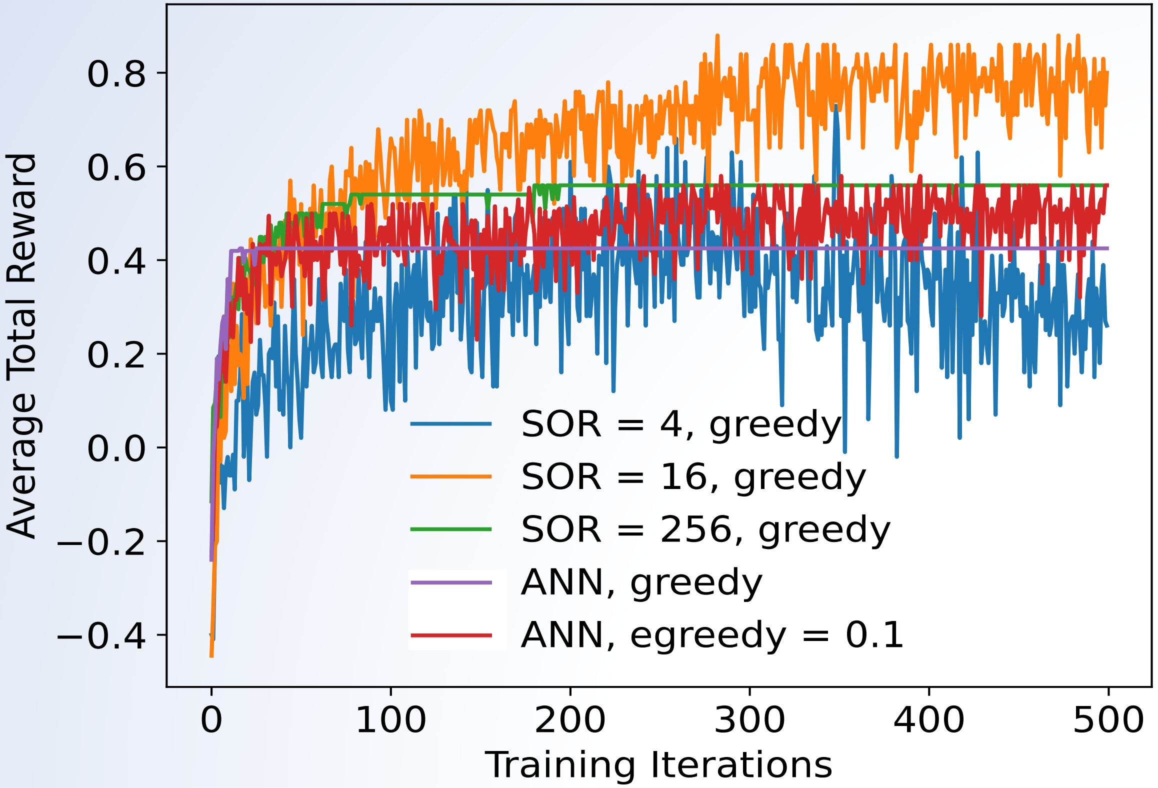
OR Gate:
Neuron!
$w_1{}^*x_1 + w_2{}^*x_2 – (w_1{}^*x_1{}^*w_2{}^*x_2)$



Exploiting Stochasticity: Reinforcement Learning

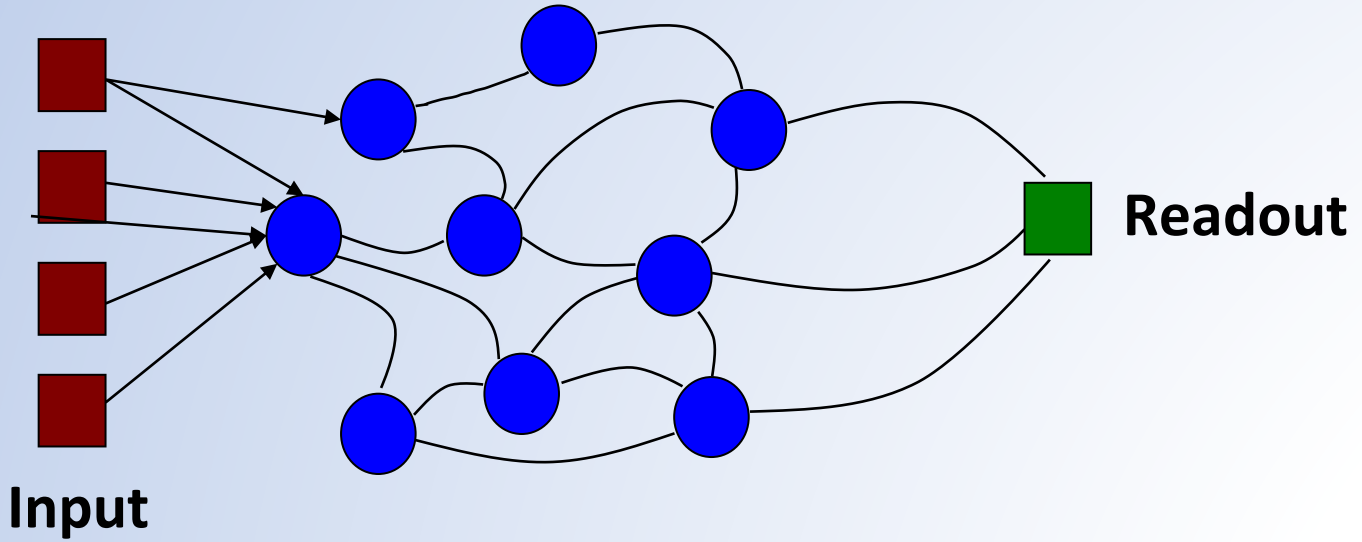
- Can the **randomness** of our devices be **useful**?
  - Yes – in reinforcement learning!
- Model for a simple maze **navigation** task.
- Our device can **explore** without any additional random number generation!





Recurrent Neural Networks

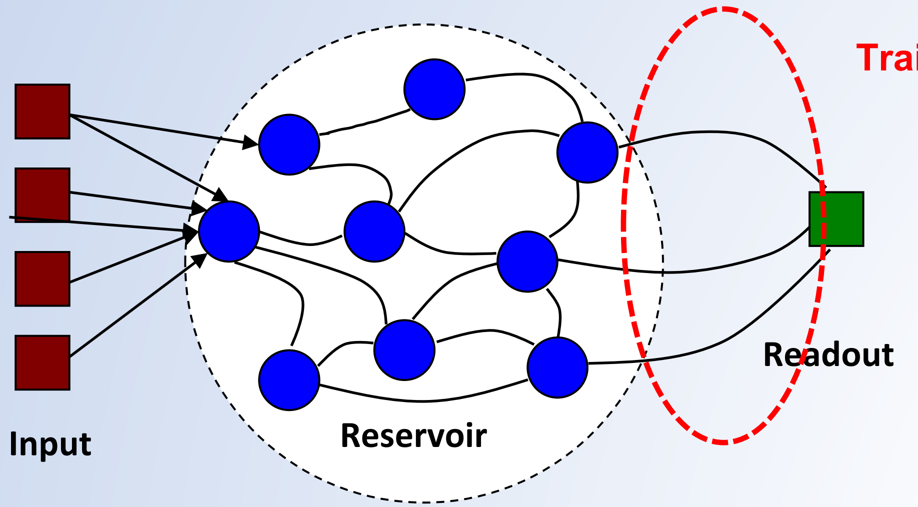
- Feed forward networks create a **static transform** of input data, but many interesting problems are time dependent.
- Recurrent networks create **time-dependent transforms** of data…



- Very computationally powerful, and well-suited to time series analysis.
- Much harder to train as need to “unravel” network in time to train (**backpropagation through time**).

Reservoir Computing

- **Reservoir computing** developed is a development of recurrent neural networks.
- Recurrent network with fixed synaptic weights (**the reservoir**) connected to a trainable **readout layer**.

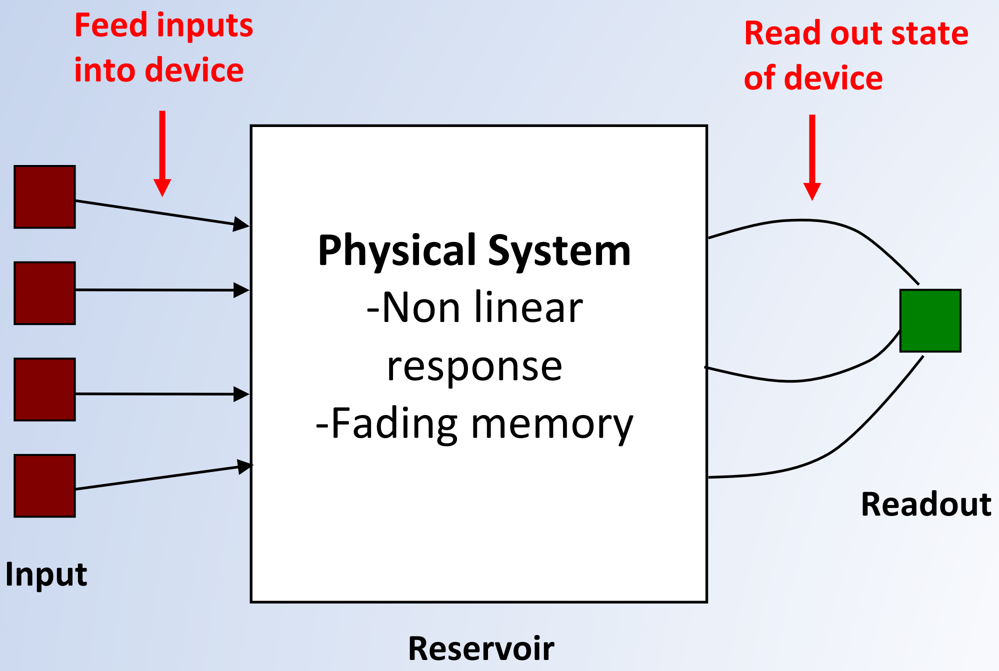


$$y(t) = \sum_{N} w_i u_i(t) + b$$

- Role of the reservoir is to **transform** input data into a form that is **more easily classifiable**.

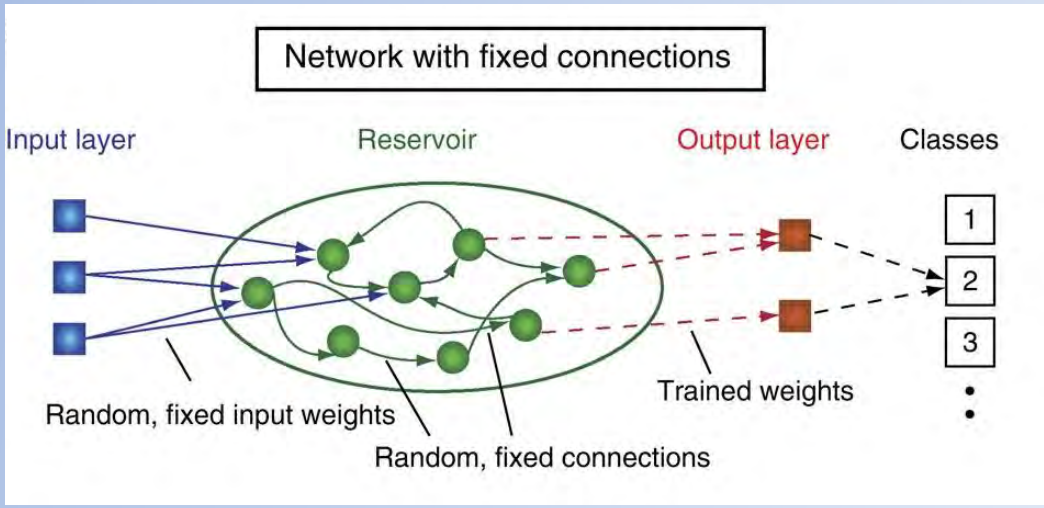
Physical Reservoir Computing

- For **hardware realisations** as the reservoir can be replaced any **physical system** that has the following properties:
  - **Non-linear response** to input.
  - **Fading memory** of past inputs.

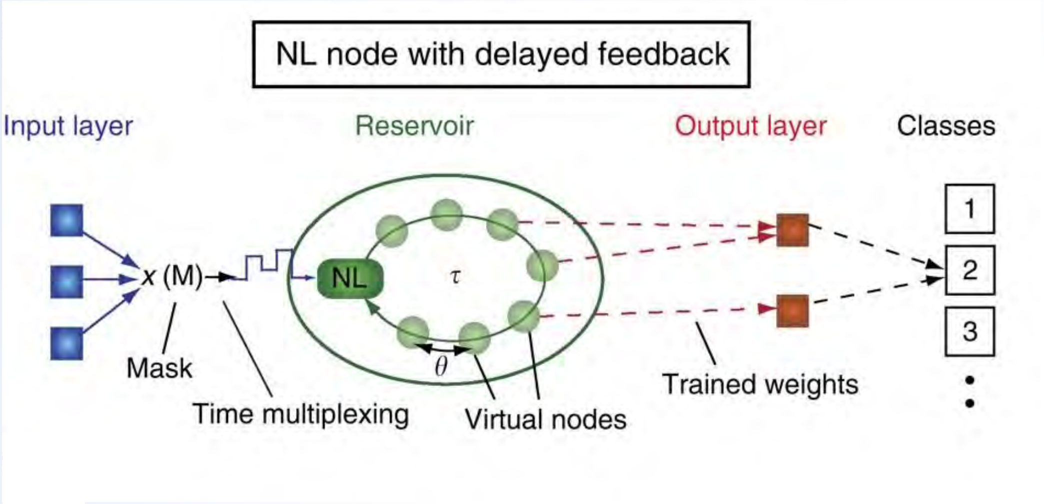


Time multiplexed RC

- Some devices only have **one input** and **one output**.
- How can we create a **multi-input** and **multi-output** device?
  - **Time multiplexing!**



Nodes of network spatially distributed



Nodes of network temporally distributed

Appeltant *et. al.* Nature Comms **2** 468 (2011)

Nanomagnetic RC Platforms

- There's lots of ideas out there! We published this last year, but it's probably already missing a lot!

A perspective on physical reservoir computing with nanomagnetic devices

Cite as: Appl. Phys. Lett. **122**, 040501 (2023); doi: 10.1063/5.0119040
Submitted: 7 August 2022 · Accepted: 31 December 2022 ·
Published Online: 23 January 2023

View Online Export Citation CrossMark

Dan A. Allwood,[1] Matthew O. A. Ellis,[2] David Griffin,[3] Thomas J. Hayward,[1,a)] Luca Manneschi,[2] Mohammad F. KH. Musameh,[4] Simon O'Keefe,[3] Susan Stepney,[3] Charles Swindells,[1] Martin A. Trefzer,[4] Eleni Vasilaki,[2,a)] Guru Venkat,[1] Ian Vidamour,[1,2] and Chester Wringe[3]

AFFILIATIONS

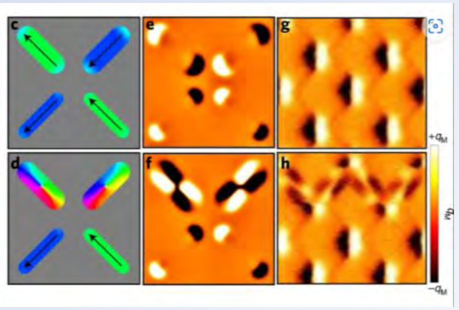
[1]Department of Materials Science and Engineering, University of Sheffield, Sheffield S1 3JD, United Kingdom
[2]Department of Computer Science, University of Sheffield, Sheffield S1 4DP, United Kingdom
[3]Department of Computer Science, University of York, York YO10 5GH, United Kingdom
[4]Department of Electronic Engineering, University of York, York YO10 5DD, United Kingdom

[a)]Authors to whom correspondence should be addressed: t.hayward@sheffield.ac.uk and e.vasilaki@sheffield.ac.uk

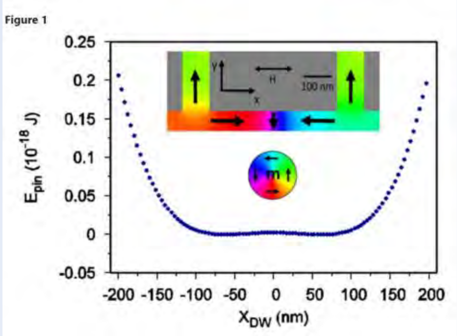
ABSTRACT

Neural networks have revolutionized the area of artificial intelligence and introduced transformative applications to almost every scientific field and industry. However, this success comes at a great price; the energy requirements for training advanced models are unsustainable. One promising way to address this pressing issue is by developing low-energy neuromorphic hardware that directly supports the algorithm's requirements. The intrinsic non-volatility, non-linearity, and memory of spintronic devices make them appealing candidates for neuromorphic devices. Here, we focus on the reservoir computing paradigm, a recurrent network with a simple training algorithm suitable for computation with spintronic devices since they can provide the properties of non-linearity and memory. We review technologies and methods for developing neuromorphic spintronic devices and conclude with critical open issues to address before such devices become widely used.

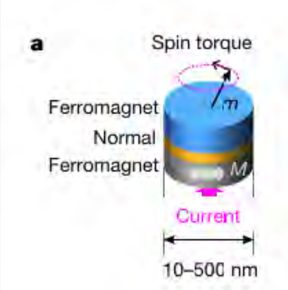
D.A. Allwood, T.J. Hayward *et. al.*
Appl. Phys. Lett. **122**, 040501 (2023)



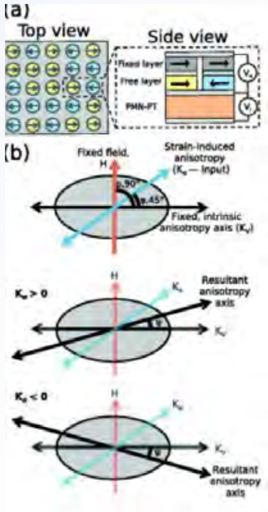
Spin Ices



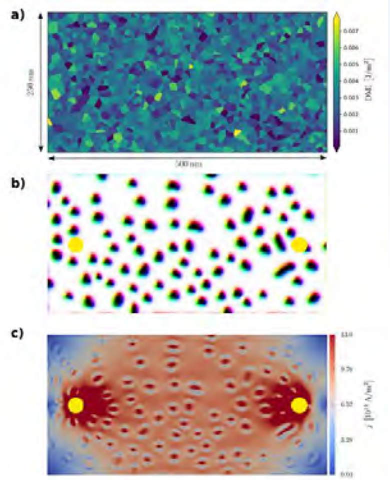
Domain Walls



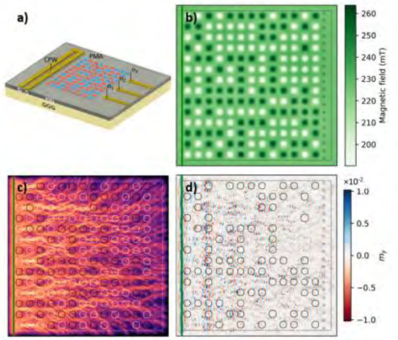
STOs



Super Paramagnets



Skyrmions

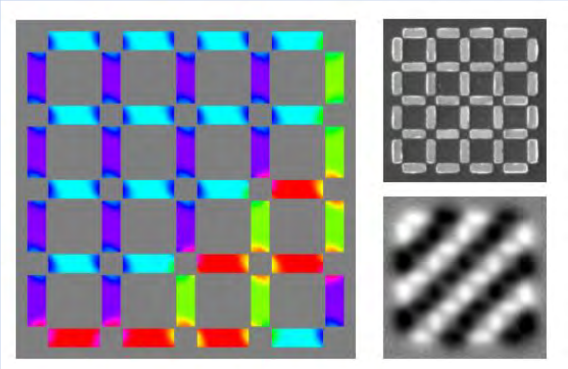


Magnons

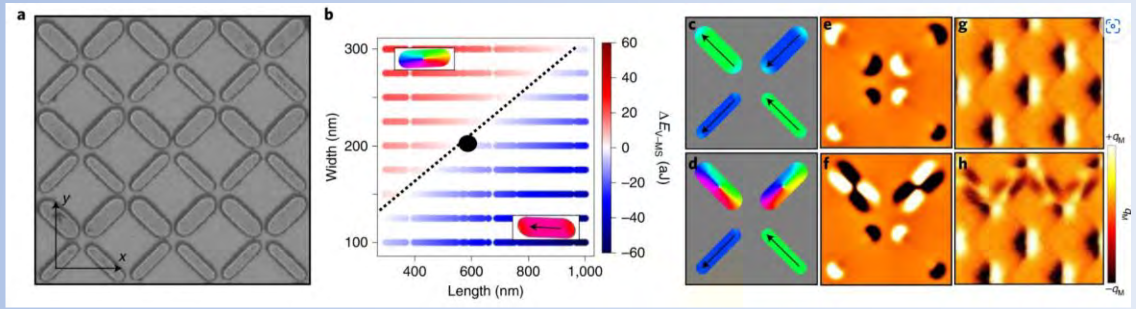
RC with Magnetic Metamaterials

- Magnetic Meta Materials like **Artificial Spin Ices** and **Nanoring Ensembles** consist of large numbers of interconnected elements.
- Huge **state space**, **emergent behaviour** → potential ideal for reservoir computing!

ASI

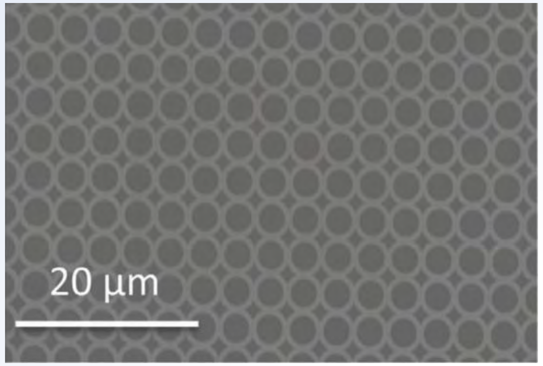


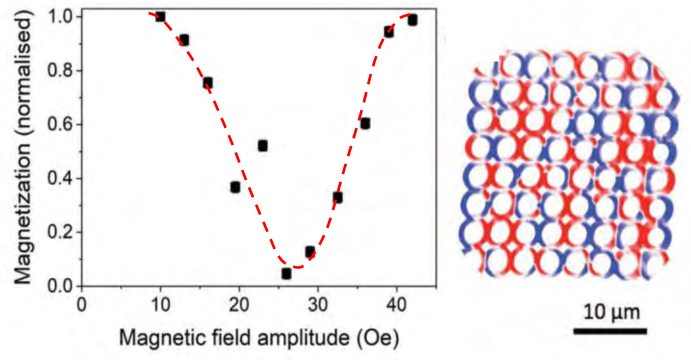
J.H. Jensen *et. al.* Artificial Life Conference Proceedings, 15-22 (2018)



J.C. Gartside *et. al.* Nature Nanotechnology 17 (5), 460-469 (2022)

NRE

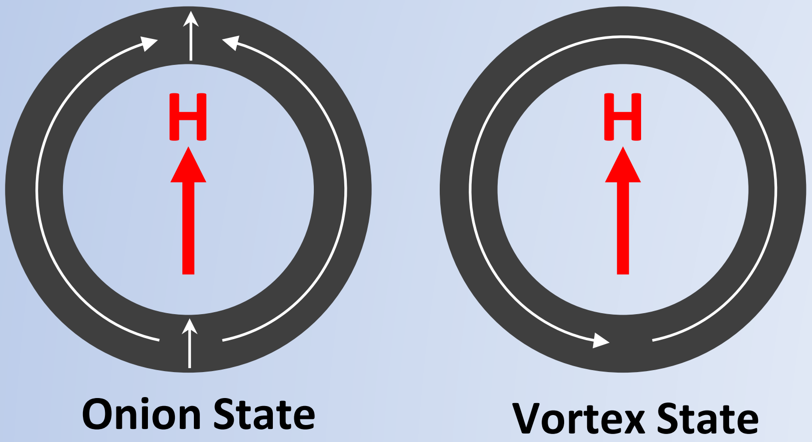


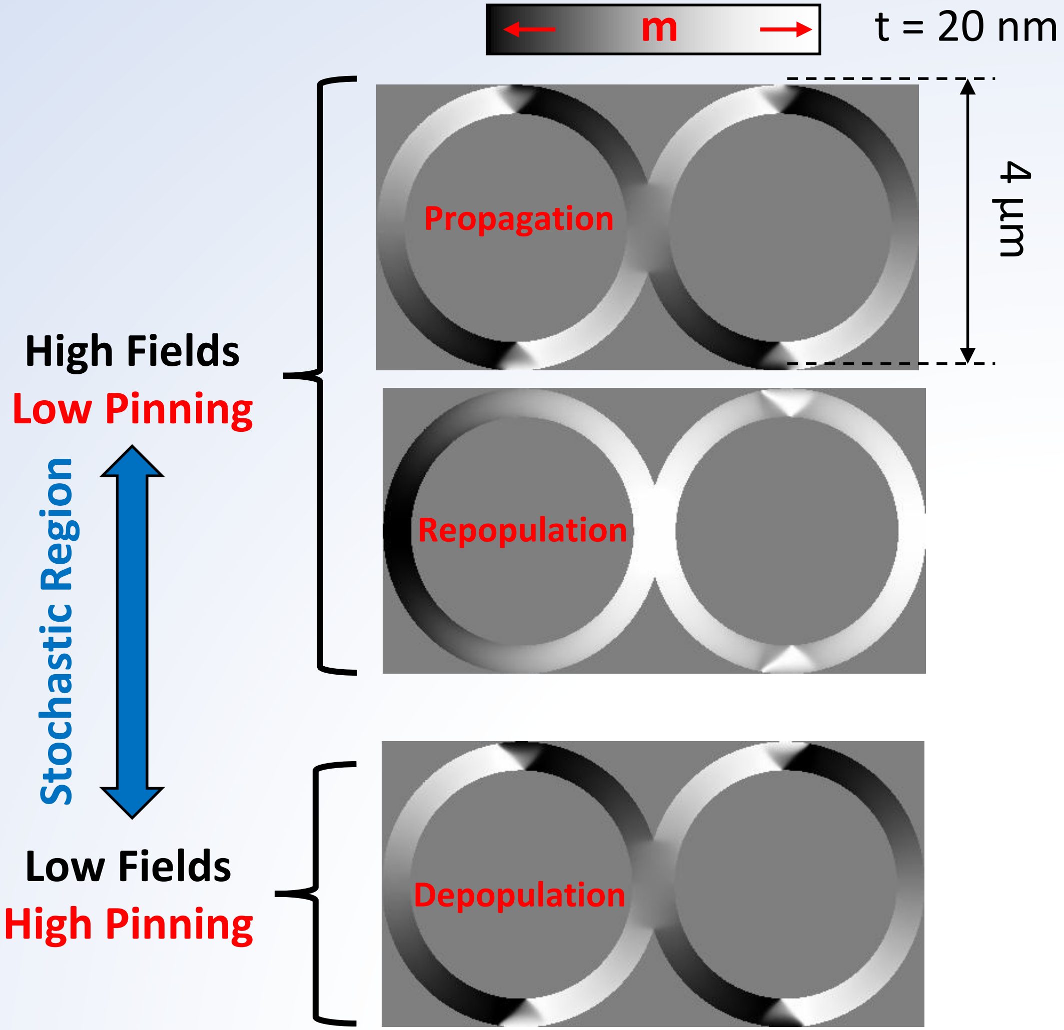


R.W. Dawidek, T.J. Hayward et. al. Adv. Func. Mater. 31, 2008389 (2021)

Domain Walls in Ring-shaped Nanowires

- **Ring-shaped nanowires** form two magnetic states.
  - **Onion state** – DWs rotate with rotating applied field.
  - **Vortex state** – circulating magnetisation.
- When multiple rings are connected the **junctions act as pinning sites**.
  - Produce **domain wall interactions** that may cause **population** or **depopulation** of DWs in the rings.
- At intermediate applied fields these processes will be **stochastic**!

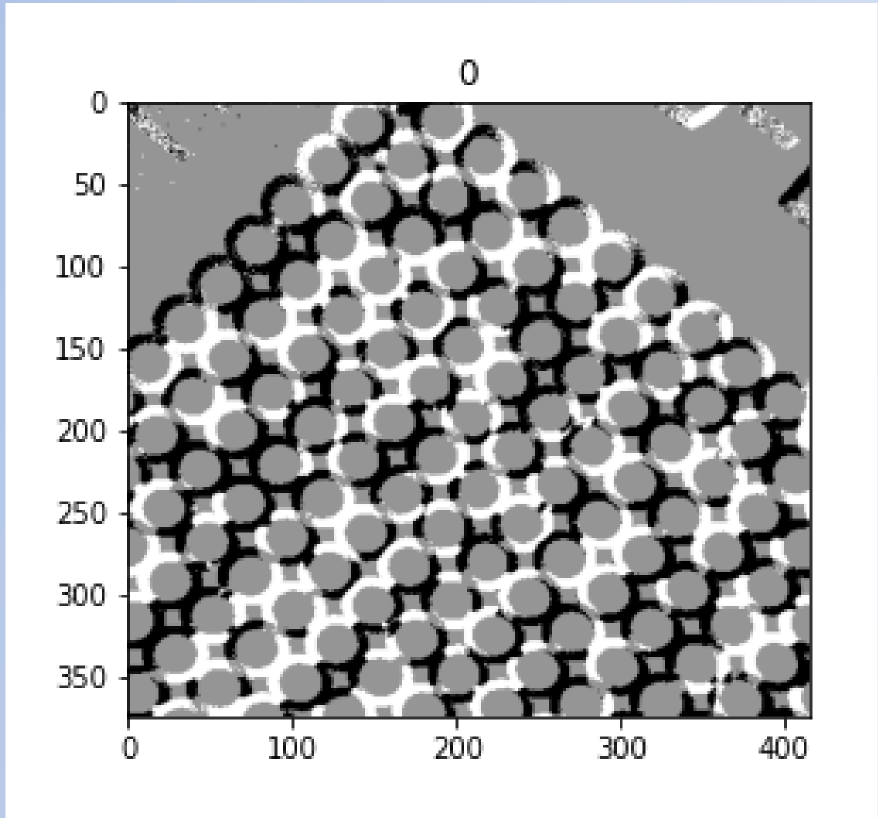


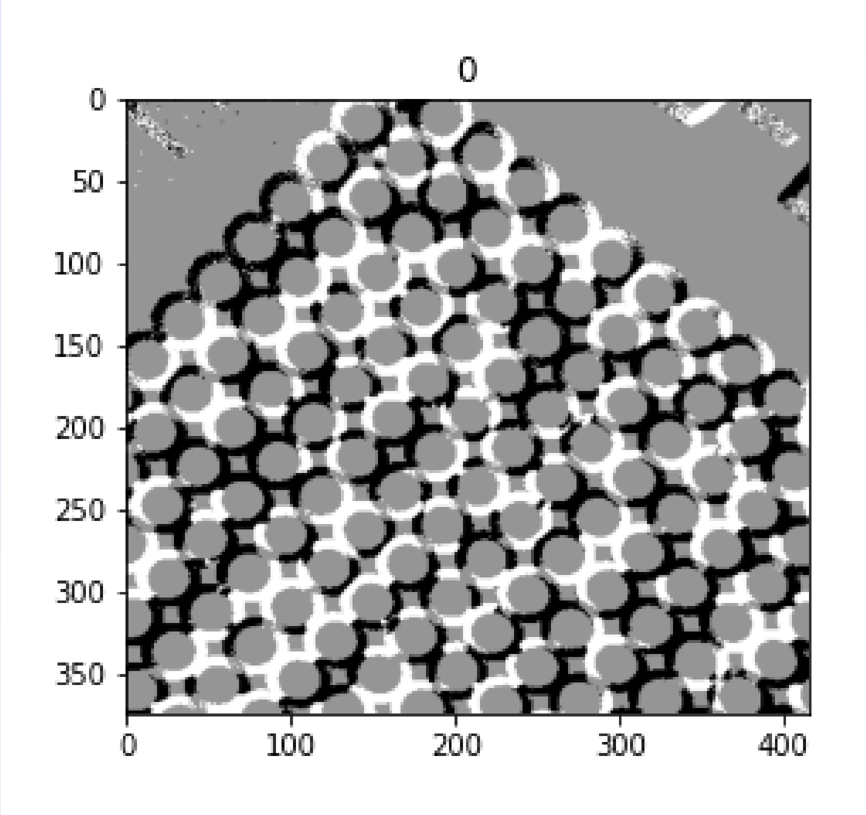


Micromagnetics (mumax^3)

Emergent Behaviour in Ring Ensembles

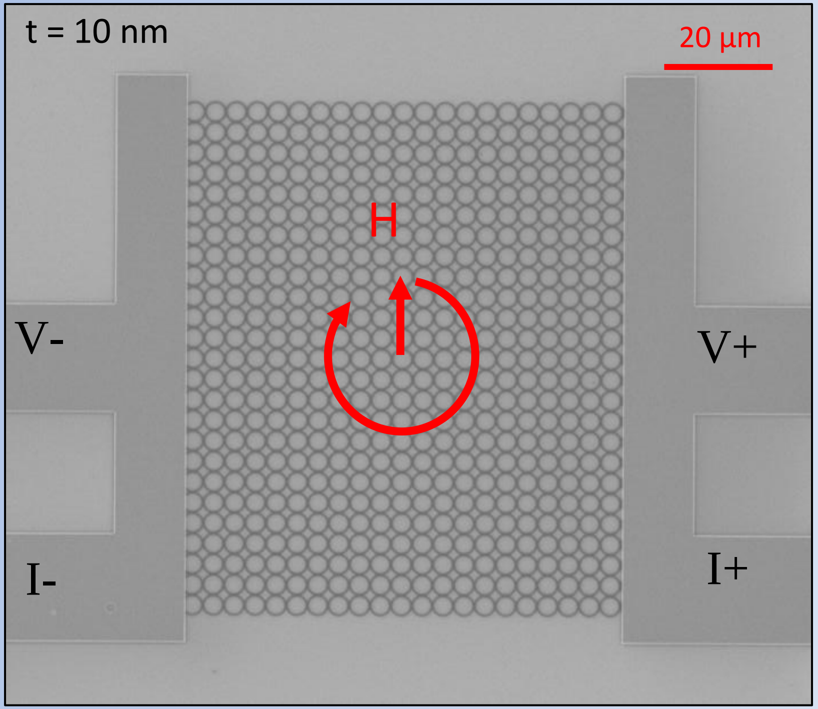
- Complex, emergent interactions in large, **interconnected arrays**!
- PEEM measurements allow us to **directly visualise** the behaviours!



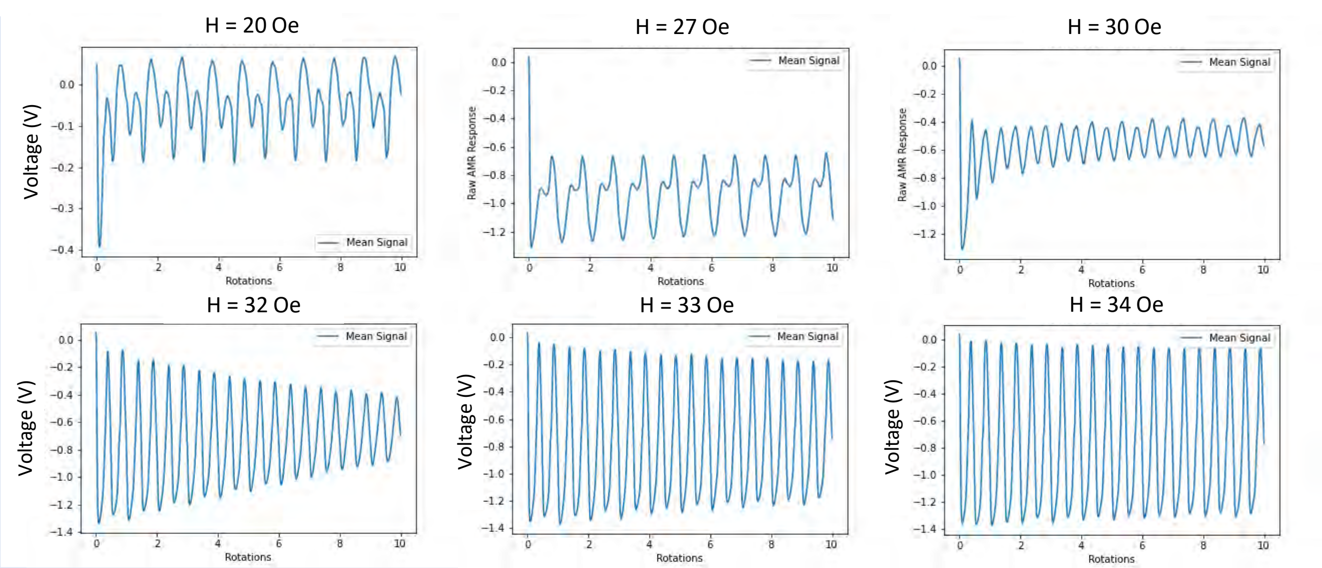


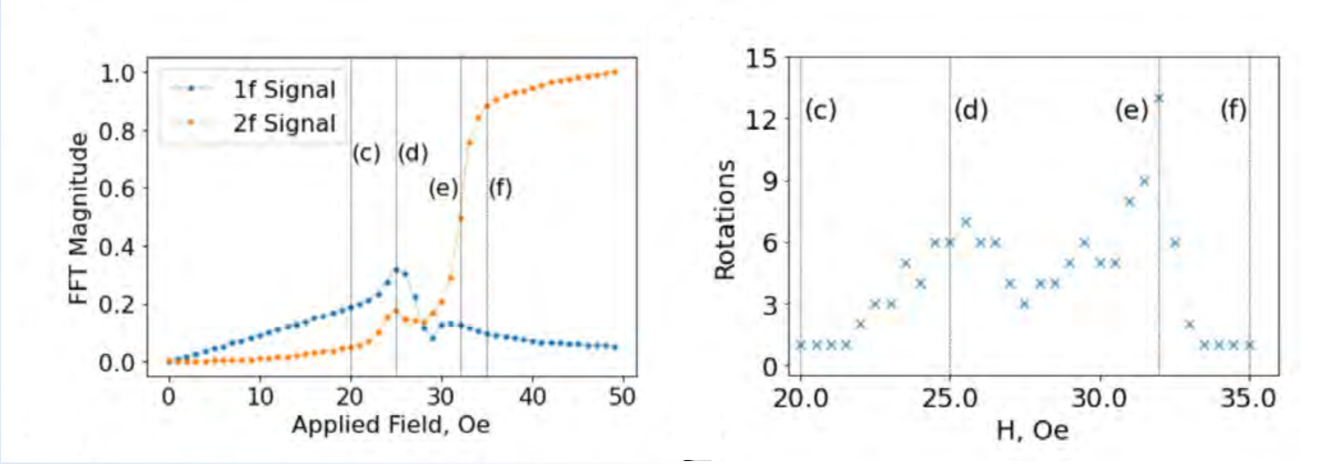
Creating Electrical Devices

- Create **electrically contacted** arrays.
- Measure **anisotropic magnetoresistance** (AMR).
  - Essentially measures number, size and position of DWs in the system.
- Rich signals produced when rotating magnetic fields applied.
  - Frequency components at **1x** and **2x** field frequency.
  - Non-linear variation, fading memory! ✓✓



I.T. Vidamour et. al Comms. Phys. 6, 230 (2023)





RC: Signal Reconstruction

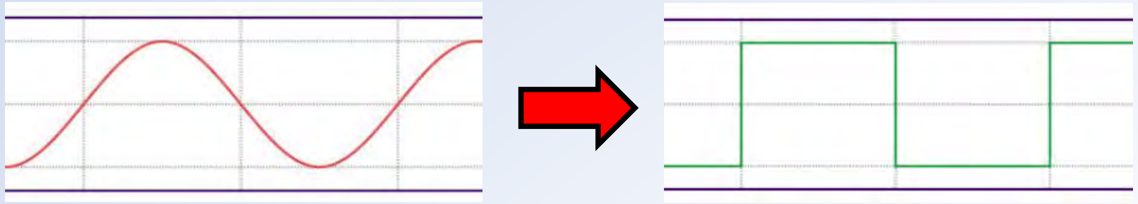
- **Transform** sine wave into other periodic waveforms.
- Simple approach:
  - Encode data in **amplitude of rotating field**.

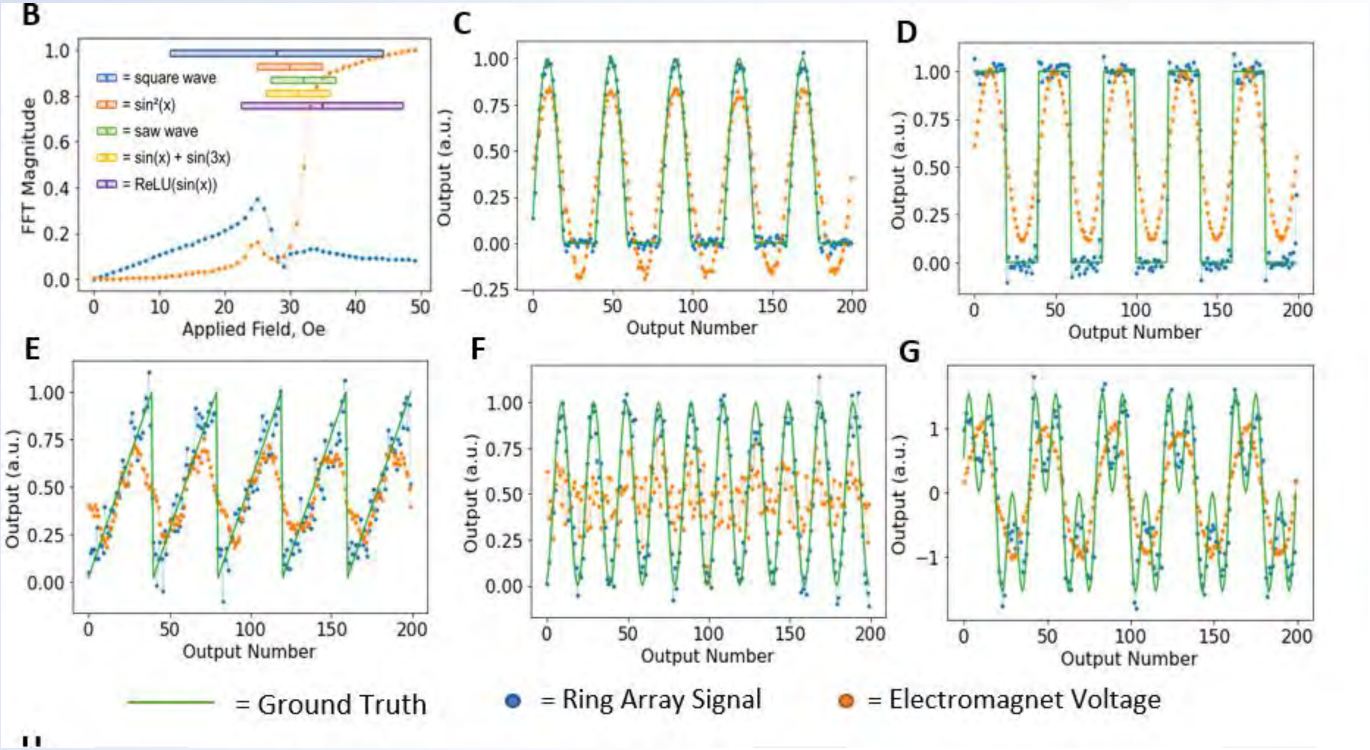
$$H = H_0 + \Delta H * x(t)$$

  - One input datum per cycle.
  - Sample **AMR response 32x per input** to create output vector
- Low mean square error (**MSE**) for all waveforms tested.
  - Referenced to equivalent operation on electromagnet voltage.

Input

Output





RC: More Complex Tasks

Spoken Digit Recognition

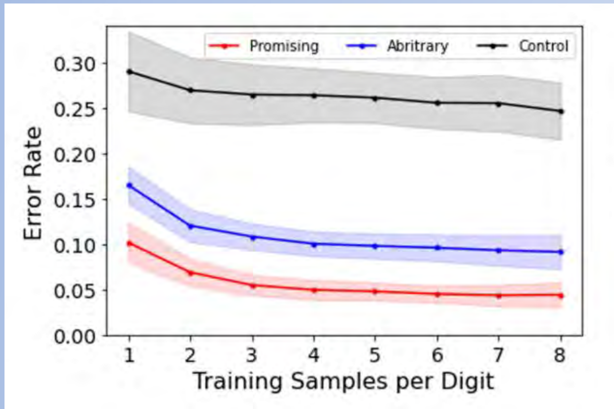
Input

Output

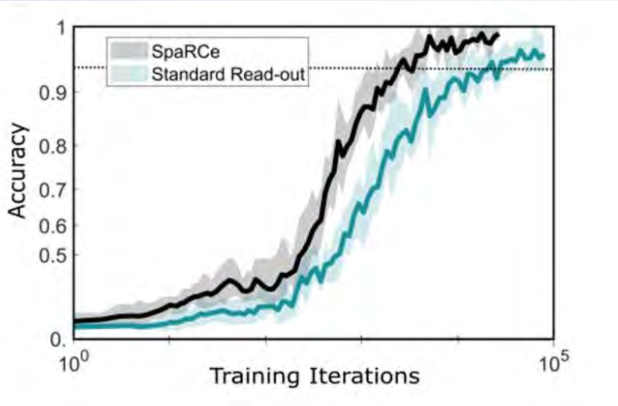
3!

- Classify digits 0-9 for 5x female speakers from TI-46 database
- Data preprocessed via a Mel-Frequency Cepstral filter.

Offline Training



Online Training



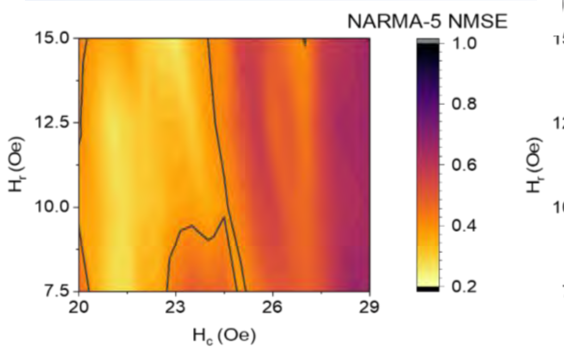
I.T. Vidamour et. al Comms. Phys. 6, 230 (2023)

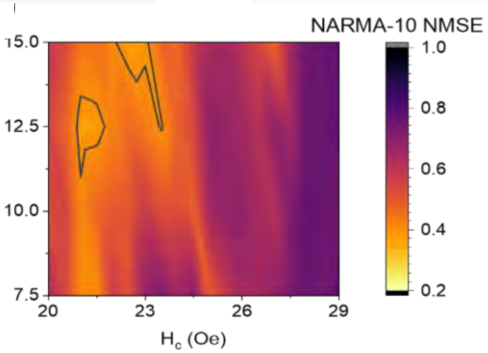
NARMA-N Time Series

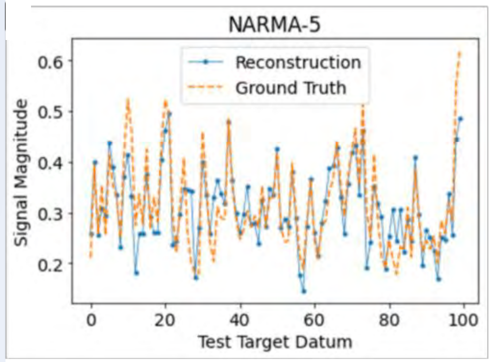
Input | **Output**

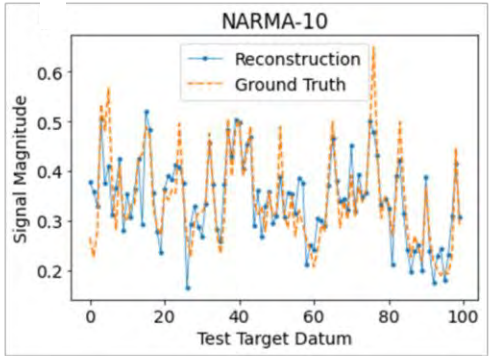
x_t → $y_{t+1} = F(y_t, y_{t-1}, .., y_{t-N}, x_t, x_{t-1}, .., x_{t-N})$

(White noise) | (Non-linear transform of last N inputs and outputs)



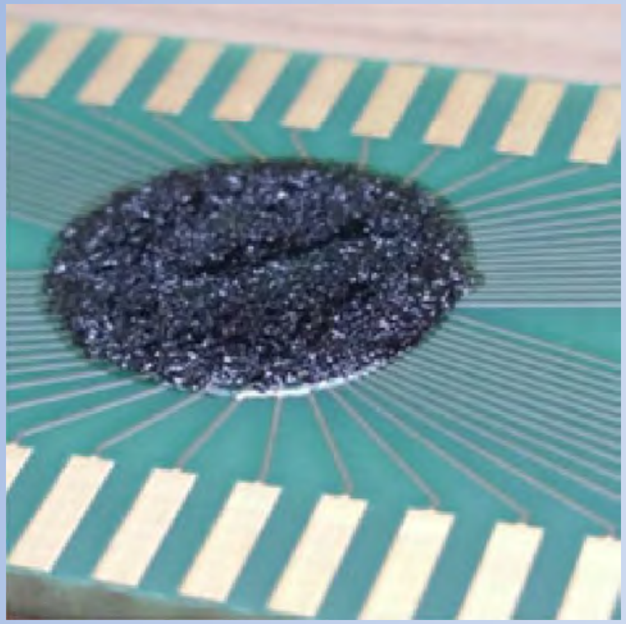


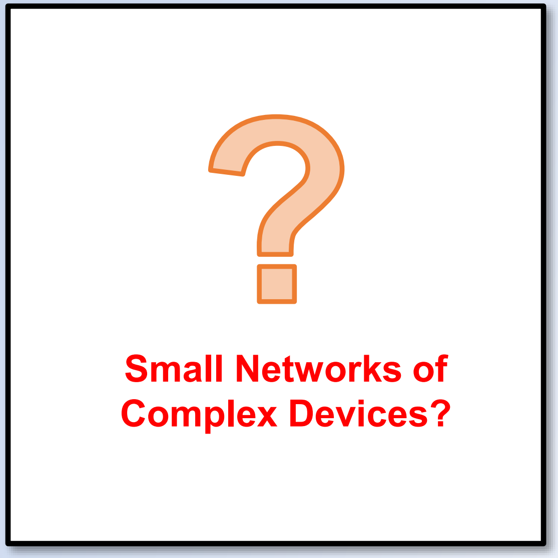


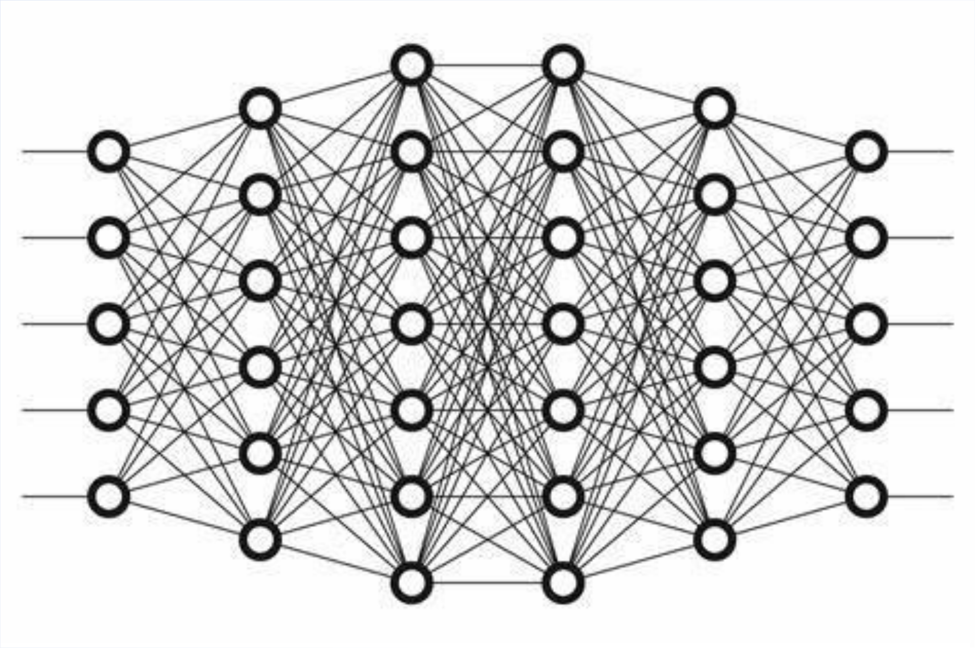


Dynamic System Networks

- Tendency towards a **false dichotomy** in how we think about physical neuromorphic computing…
- But what if we didn't think **so digitally….**





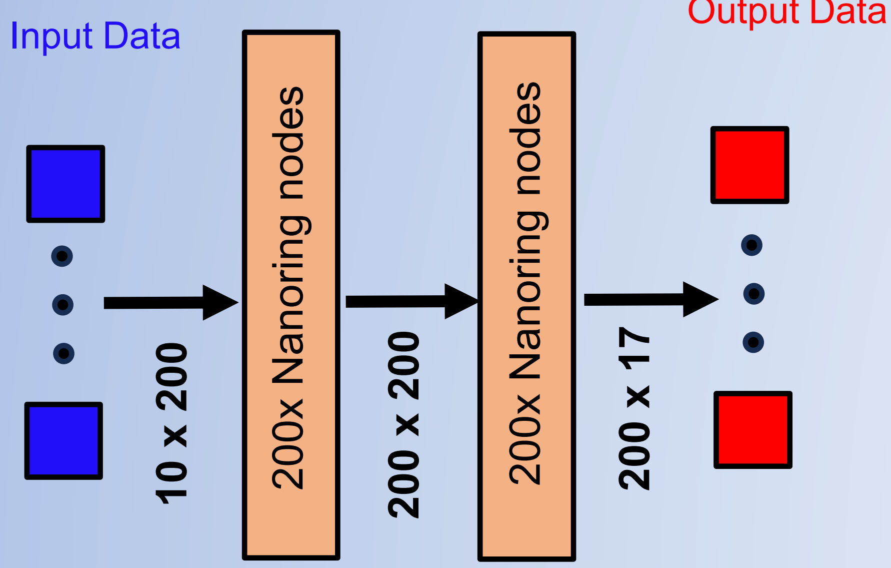


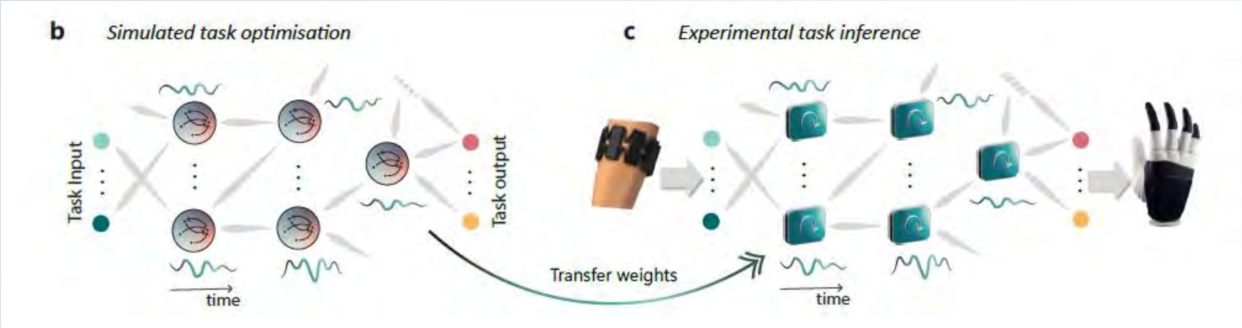
Complex Device, Monolithic

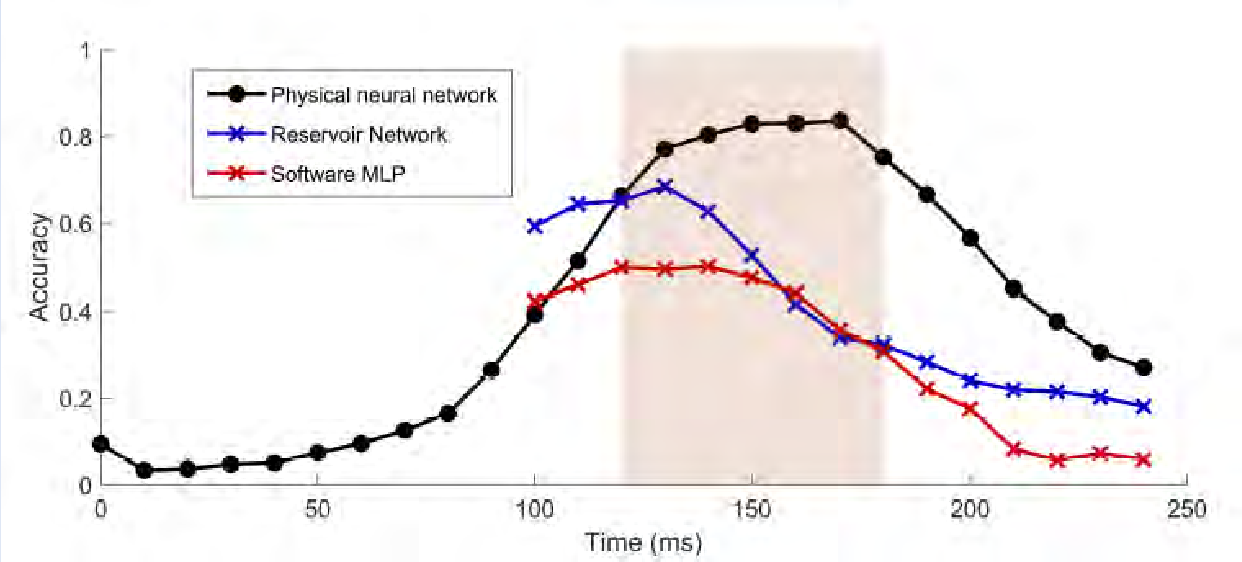
Simple Devices, Networked

Real World Task: Neuroprosthetics

- Train nanoring array **network** to recognise **hand gestures**:
  - Ninapro database.
  - Recognise 17 hand gestures.
  - 27 subjects.
  - 10 channels of electrode input data.







I.T. Vidamour et. al. *under review*

Conclusions

- Magnetic materials offer exciting possibilities for create new types of **neuromorphic hardware**!
- **Do the materials/devices/phenomena you are studying have any exciting functionality that could be explored for these applications?**