

Magnetism for Neuromorphic Computing

Tom Hayward

Functional Magnetic Materials Group Department of Materials Science and Engineering

The University of Sheffield



Engineering and Physical Sciences Research Council



Horizon 2020 European Union funding for Research & Innovation

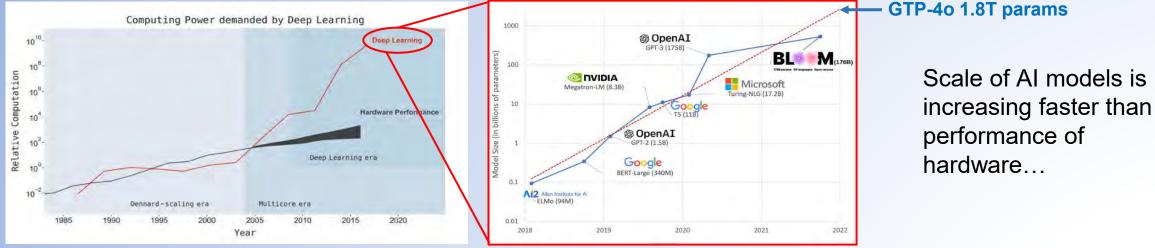
Outline

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

Al is Amazing...

Write me a short poem	about why Al is cool Healthcare	e Transport
 The Wonder of AI Born of code, a spark in the mind, Infinite answers, endlessly kind. From puzzles unsolved to dreams unfurled, 		Customer Service
Al 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.	Make me an info-graphic showing the areas AI may have impact	Finance
With logic and wonder, it breaks every rule, That's why we say: AI is cool.	HEALTHCARE EUUCATION TEDUCATION -EDUCATION EDUCATION EDUCATION TEDUCATION TEDUCATION TAXING TION TAXING TION TAXING TION TAXING TION TAXING TION	 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:
Energy	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!	<pre>(/> Pytorch Nn 2inputs 1 import torch 2 import torch.nn as nn 3 4 ~ class SimpleNeuralNetwork(nn.Module): 5 ~ definit(self): 6 super(SimpleNeuralNetwork, self)init() 7 # Define the network layers</pre>
Manufacturing	Security	<pre>8 self.fc1 = nn.Linear(2, 10) # Input layer to first hidden layer 9 self.fc2 = nn.Linear(10, 10) # First hidden layer to second hidden layer 10 self.fc3 = nn.Linear(10, 1) # Second hidden layer to output layer 11 self.activation = nn.ReLU() # Activation function 12 13 v def forward(self, x): 14 # Define the forward pass</pre>

...but has problematic impacts...



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

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

Common carbon footprint benchmarks

in lbs of CO2 equivalent



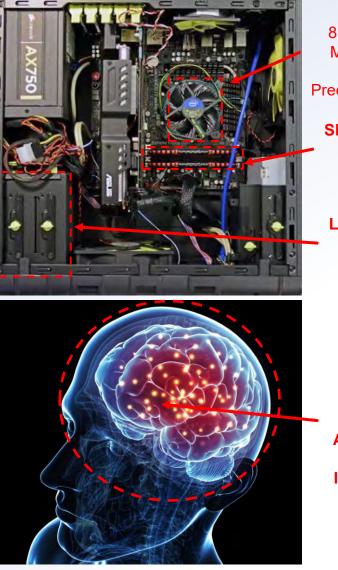
Chart Mit Technology Preserve Source Strucell up at - Created with Datawrapper

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**.



CPU 8 – 16 parallel threads Minimal local memory Strictly Digital Precise but noise intolerant

Short Term Memory ~4 GB - 32 GB

Long Term Memory ~256 GB to 2 TB

Massively Parallel!

Memory embedded with compute!

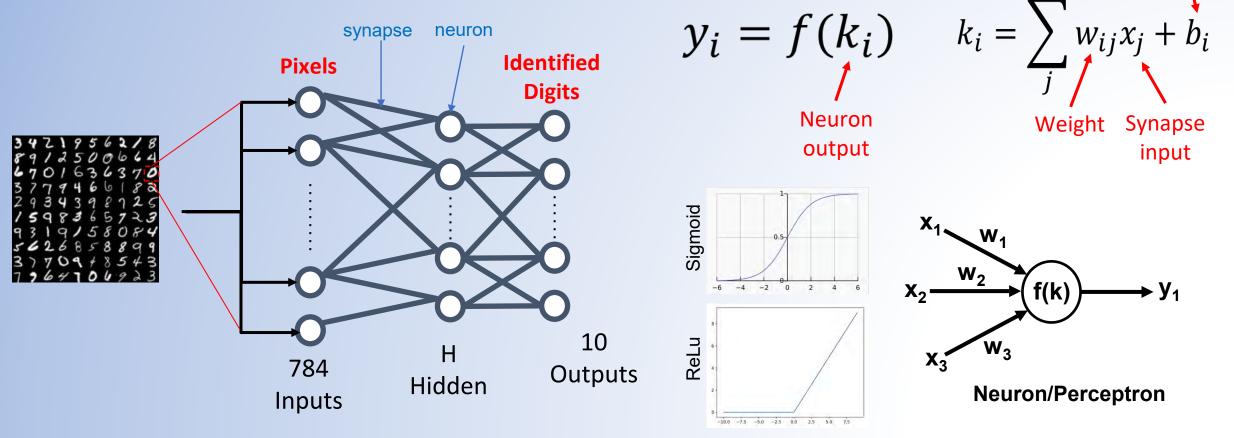
Analogue not Digital

Imprecise, but noise tolerant.

Feed-forward Neural Networks

Bias

- 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 handwritten digits).



Training FFNN

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

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

loss function

• Train weights by iterative gradient descent:

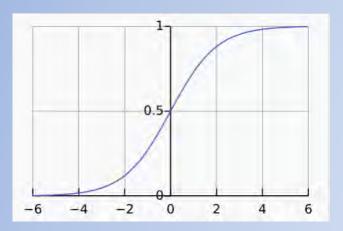
$$\mathbf{w_{n+1}} = \mathbf{w_n} - \gamma \nabla C(\mathbf{w_n})$$

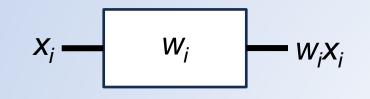
learning rate

- 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 Synapse





Interconnect



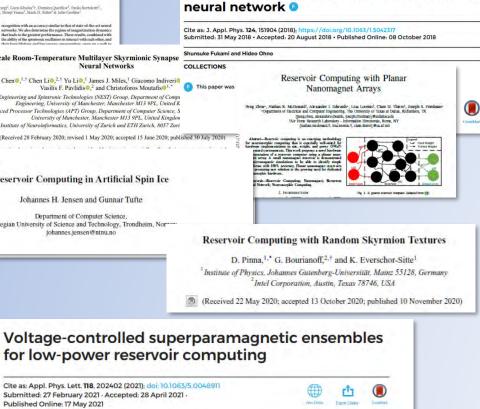
- Addition of synaptic inputs.
- Apply non-linear activation.
- Multiplies input by weight.
 Weight ideally least and per
- Weight ideally local and nonvolatile

• **Transports** data through the network.

Lots of magnetism groups working on this!

Neuromorphic computing with nanoscale spintronic oscillators

lacob Torrejon¹, Mathieu Riou¹, Flavos Abreu Araujo¹, Sumito Tiarnega², Guru Khalsa²r, Damier Querlior⁴, Paolo Bortolotti¹ Uncent Crwc², Kav Yakoshili¹¹, Akio Pakushima², Hurshi Kubura², Shimii Yuawa², Mark D. Selles³ & Julie Großler⁴ Surgers in the brain behave a multivers worlflature, which develops styrbmic activity and interacts to process information. Taking any structure to mitch head on the structure of the structure of the structure of the structure structure to mitch head on the structure of the structure of the structure of the structure of the structure structure to structure on the structure of the structure to fit 10⁴ oscillators organized in a two-di a chip the size of a thumb, the lateral dime Nanoscale Room-Temperature Multilayer Skyrmionic Synapse nust be smaller than one micr tend to be noisy and to lack the stability f Neural Networks lata in a reliable way. For this reason, dev oposals2-3 and several candidates, in perconducting⁷ oscillators, a proof of computing using nanoscale oscillators has are we show experimentally that a nanoscale magnetic tunnel junction)^{5,0} can be used Runze Chen[®],^{1,†} Chen Li[®],^{2,†} Yu Li[®],¹ James J. Miles,¹ Giacomo Indiveri[®] Vasilis F. Pavlidis⁰,² and Christoforos Moutafis^{01,*} ¹Nano Engineering and Spintronic Technologies (NEST) Group, Department of Compu Envineering, University of Manchester, Manchester M13 9PL, United K Advanced Processor Technologies (APT) Group, Department of Computer Science, S University of Manchester Manchester M13 9PL United Kinodon ³ Institute of Neuroinformatics, University of Zurich and ETH Zurich, 8057 Zuri (Received 28 February 2020; revised 1 May 2020; accepted 15 June 2020; published 30 July 2020) **Reservoir Computing in Artificial Spin Ice** Johannes H. Jensen and Gunnar Tufte Department of Computer Science Norwegian University of Science and Technology, Trondheim, Norjohannes.jensen@ntnu.no



A. Welbourne, 🔤 🝈 A. L. R. Levy, 🏆 🔞 M. O. A. Ellis, 🖏 🔀 H. Chen, M. J. Thompson, E. Vasilaki, 🕉 🔞 D. A. Allwood, 🔞 and T. J. Hayward

AFFILIATIONS

Published Online: 17 May 2021

Perspective: Spintronic synapse for artificial EN Neuromorphic computation

architectures and the structures of the algorithms they are required to simulate. Neuromorphic devices, and in particular reservoir computing architectures, utilize the inherent properties of pl systems to implement machine learning algorithms and so have the potential to be much more efficient. In this work, we demonstrate that the dynamics of individual domain walls in magneti nanowires are suitable for implementing the reservoir computing paradigm in hardware. We me

Dan A. Allwood¹, Eleni Vasilaki² & Thomas J. Hayward¹

the dynamics of a domain wall placed between two anti-notches in a nickel nanowire using both collective coordinates model and micromagnetic simulations. When driven by an oscillating ma field, the domain exhibits non-linear dynamics within the potential well created by the anti-not that are analogous to those of the Duffing oscillator. We exploit the domain wall dynamics for re computing by modulating the amplitude of the applied magnetic field to inject time-multiplexe signals into the reservoir, and show how this allows us to perform machine learning tasks includ the classification of (1) sine and square waves; (2) spoken digits; and (3) non-temporal 2D toy da hand written digits. Our work lays the foundation for the creation of nanoscale neuromorphic d in which individual magnetic domain walls are used to perform complex data analysis tasks.

with a single magnetic domain wall

Razvan V. Ababei¹⁵⁰, Matthew O. A. Ellis², Ian T. Vidamour¹, Dhilan S. Devadasan¹,

on digital hardware are relatively inefficient due to poor matching between conventional comp

Pure voltage-driven spintronic neuron based on stochastic magnetization switching behaviour

Jia-Hui Yuan¹, Ya-Bo Chen², Shu-Qing Dou¹, Bo Wei¹, Huan-Qing Cui¹, Ming-Xu Song and Xiao-Kuo Yang

¹ Fundamentals Department, Air Force Engineering University, Xi'an 710051, People's Republic of China ²College of Computer, National University of Defense Technology, Changsha 410005, People's Republic of China

E-mail: yangxk0123@163.com

Abstract

Received 7 October 2021, revised 16 December 2021 Accepted for publication 23 December 2021 Published 18 January 2022

Voltage-driven stochastic magnetization switching in a nanomagnet has attracted more attention

Machine learning techniques are commonly used to model complex relationships but implemen $\ \ LETTER$

Check for updates

mac-//doi.org/10.1038/s41586-018-0632-

Vowel recognition with four coupled spin-torque nano-oscillators

el Romera^{1,5}, Philippe Talatchian^{1,5}, Sumito Tsunegi², Flavio Abreu Araujo^{1,4}, V Yakunhiji", Akio Tukushima", Itinoshi Kabota", Shinji Yuasa", Maxenor Ernouht¹², Damir Vodenicarevic¹, Tilen das Locatelli¹, Damien Querlioz⁵* & Julie Groffier⁵*

in recent years, or utilization end network have become the flightly. We transport how darking the second of neuroinspired computing embraces the dynamical nature of the brain and proposes to endow each component of a neural network with dynamical functionality, such as oscillations, and to rely on wirr opinaties indechanges, seen is oscinatines, and no inter or emergent physical phenomenes, such as synchronization.¹⁴ A for and ing complex problems with small activation.¹¹ Thiss, presente emerging tamobies trends devices any power compact and energy-efficient nonlinear auto-socillators that minis the periods quality activity of hisbacking any such as the formation of the periods quality activity of hisbacking any such as the formation of the periods quality.

between nication adevices trol and tures of tise and sility of curately lectrical allenge. re nano-

(Fig. 1b); in our case circular magnetic tunnel junctions with 375 nm diameter and an FeB free layer with a vortex as ground sta-(see Methods)²⁶. The double arrow connections between neuroi

(blue in Fig. 3a) indicate that the output of neuron / infl

erate, Figure 1d slavys that when the fr west, each oscillator synchros amical Indeed, when the frequency of the source gets close to the free of one of the oscillators, the strong signal of the source palls the ada

I'm just going to mostly talk about our efforts in the area...

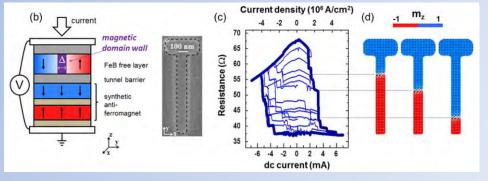
Lots of magnetism groups working on this!

Check for updates Neuromorphic computing with nanoscale Perspective: Spintronic synapse for artificial spintronic oscillators Why? interact to process information'. Taking networks. We also determine the regime of m chariour to realize high-density, low-power uting will require very large numbers of the ability of the spintronic oscillators to inter-rifilators. A junction extension of the spintronic oscillators to interhin the size of a thumh, the lateral din Nanoscale Room-Temperature Multil ist be smaller than one mice nd to be noisy and to lack the stability f Neura ta in a reliable way. For this reason, de Non-volatile memory! mac // doi org/10.1038/s41586-018-0632 • inducting oscillators, a proof of c ing using nanoscale oscillators ha Runze Chen[®],^{1,†} Chen Li[®],^{2,†} Yu Li[®],¹ Jar Vasilis F. Pavlidis .2 a vith four coupled spin-torque Nano Engineering and Spintronic Technologies (N Engineering, University of Manchest Advanced Processor Technologies (APT) Group, Inherent non-linearity! • University of Manchester, Ma Institute of Neuroinformatics, University of (Received 28 February 2020; revised 1 May with the set-up illustrated in Fig. 1b. The four neurons in Fig. Is are CMOS compatibility! Fig. 1h), in our case cit 5 nm diameter and an Fell free layer with a warter as ground at • blue to Fig. 3a) indicate that the output of neuron i Reservoir Computing in A High speed (up to THz)! Johannes H. Jensen and • Department of Compu Norwegian University of Science and Tech iohannes iensen@r Low power! Physics based interconnects (exchange, • magnetostatic coupling)! Voltage-cont for low-powe Accented for publication 23 December 2021 Cite as: Appl. Phys. Lett. 118, 202402 (2021); doi:10.1063/5.0048911 T Published 18 January 2022 Submitted: 27 February 2021 · Accepted: 28 April 2021 · Espert Casico Published Online: 17 May 2021 Abstract Voltage-driven stochastic magnetization switching in a nanomagnet has attracted more attention A. Welbourne, 🔤 🔞 A. L. R. Levy, 🍹 👩 M. O. A. Ellis, ³ 🔞 H. Chen, ¹ M. J. Thompson, ¹ E. Vasilaki, ³ 🔞 D. A. Allwood, ¹ and T. J. Hayward AFFILIATIONS

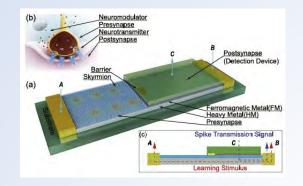
I'm just going to mostly talk about our efforts in the area...

Spintronic Synapses

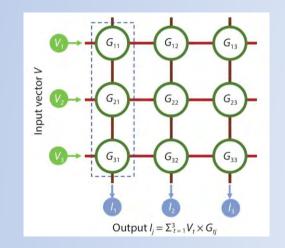
- Most proposals for magnetic synapses are based magnetoresistive memristors.
- Stores an analogue weight using e.g. position of DWs/Skyrmions in a multilayer nanotrack...



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



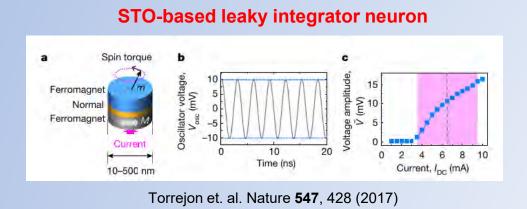
Y. Huang et. al. Nanotechnology 28, 08LT02 (2017)



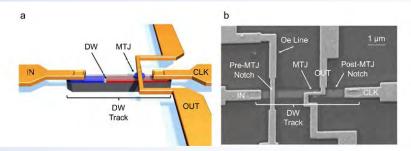
Integrating into cross-bar arrays allows multiply and accumulate via Kirchoff's law...

Spintronic Neurons

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



DW-based Spiking Neuron

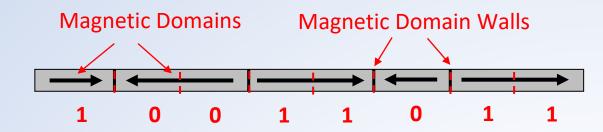


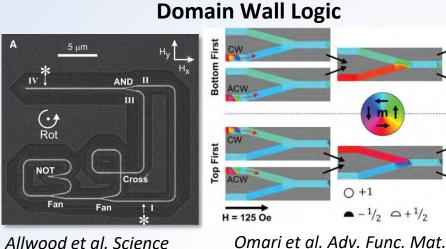
Leonard et. al. Appl. Phys. Lett. 122, 262406 (2023)

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

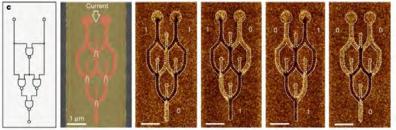
Domain Wall Devices

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





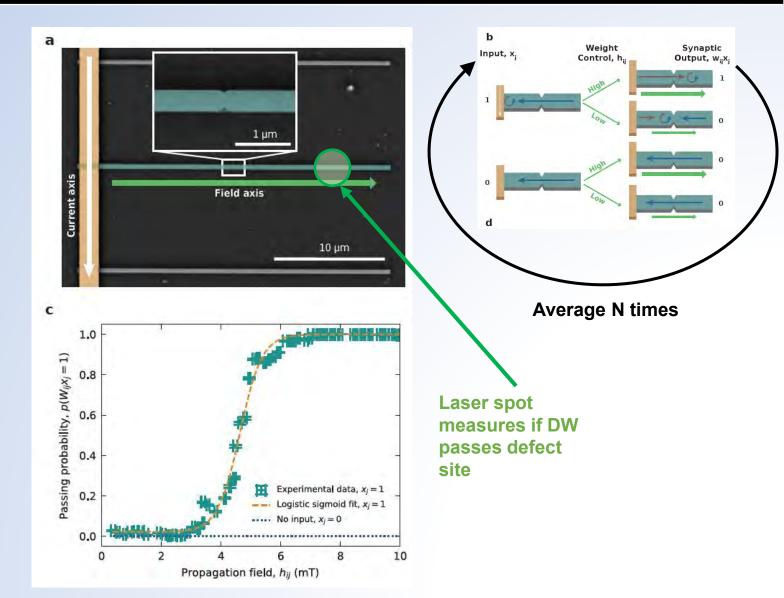
Allwood et al. Science 309, 1688 (2005) Omari et al. Adv. Func. Mat. 29, 1807282 (2019)



Luo et al. Nature 579, 214 (2020)

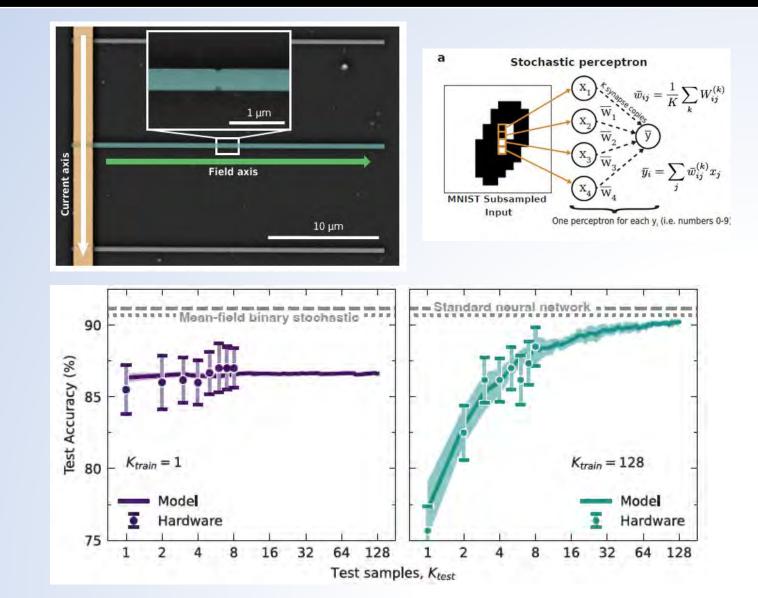
Stochastic Magnetic Synapses

- 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!



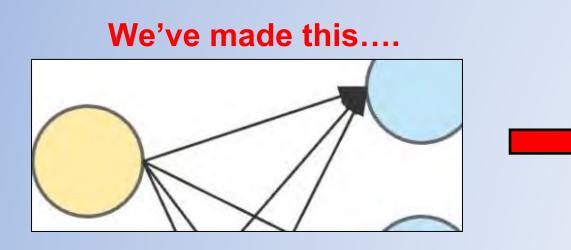
Experimental Demonstration

- 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!

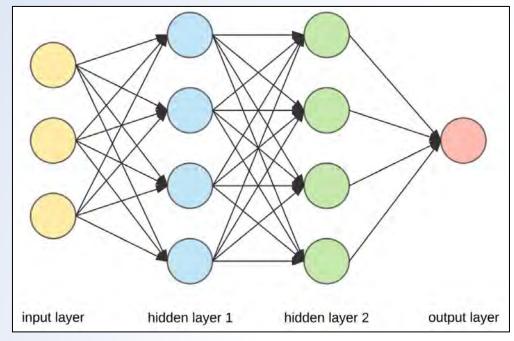


All Magnetic FFNN?

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



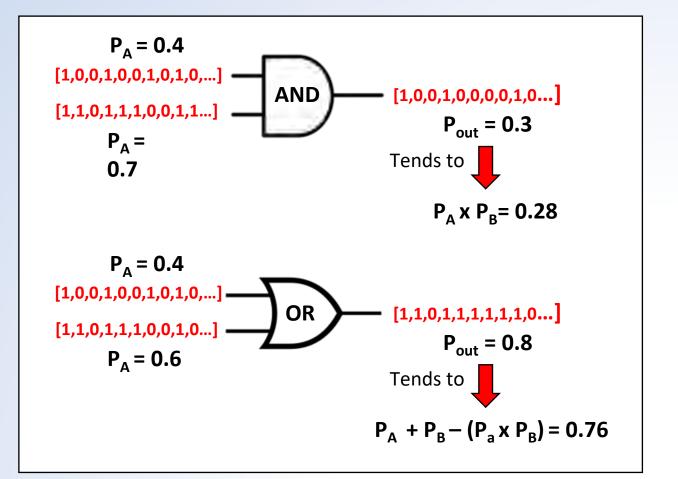
...but we want to make this.



Stochastic Computing

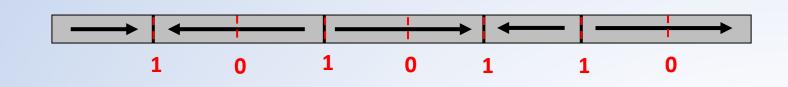
- In stochastic computing floating point numbers are represented by streams of binary bits.
 - Fraction of "1"s in a bitstream defines the number.
- Passing these through conventional logic gates naturally performs numerical calculations:
 - AND = $P_A \times P_B$
 - OR = $P_A + P_B (P_a \times P_B)$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

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



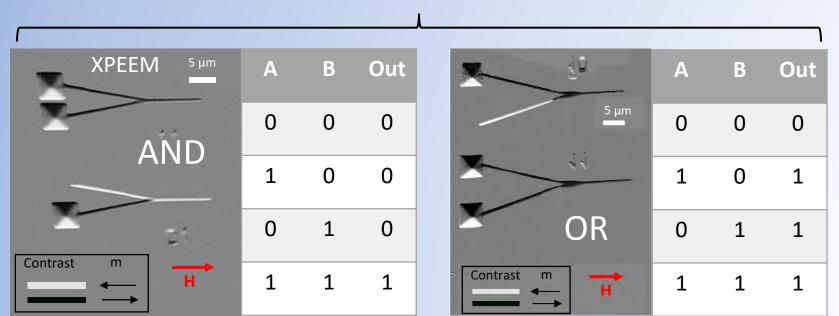
Stochastic Computing with DWs

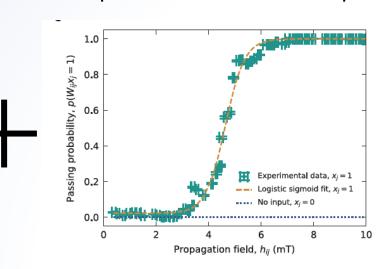
 Represent data by presence (1) or absence
 (0) of a DW.



Deterministic DW Logic Gates







All Magnetic FFNN

1.0

0.8

stream

Input one (p1)

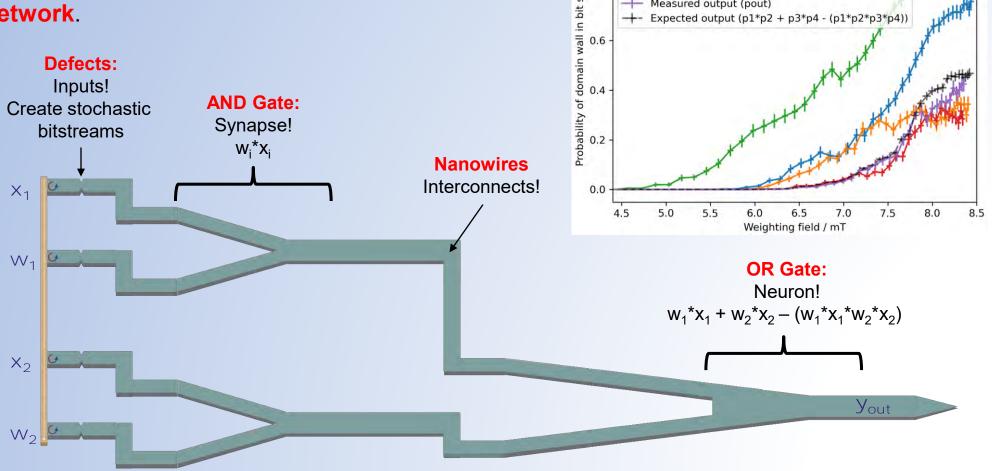
Input two (p2)

Input three (p3)

Input four (p4)

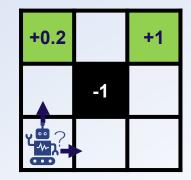
Measured output (pout)

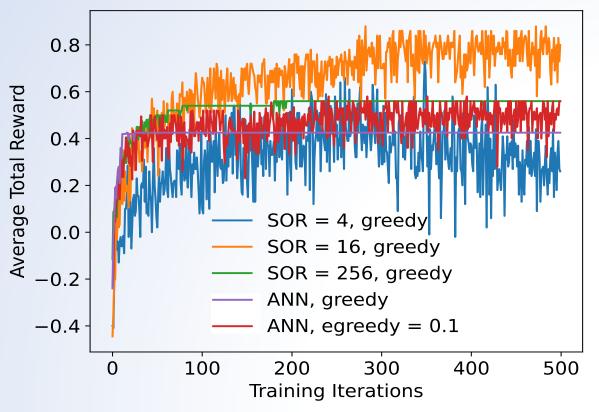
Use magnetic nanowires to create a simple 2 synapse, 1 neuron, neural network.



Exploiting Stochasticity: Reinforcement Learning

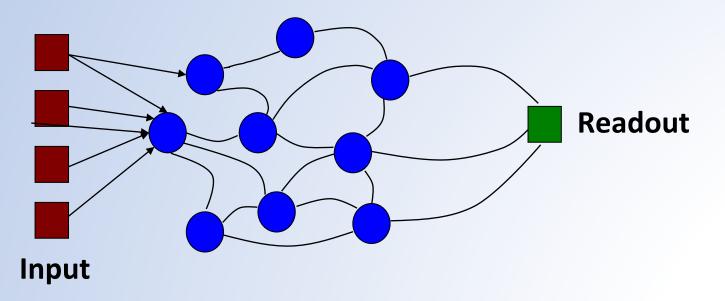
- Can the randomness of our devices be useful?
 - Yes in reinforcement learning!
- Model for a simple maze navigation task.
- Our device can explore without any additional random number generation!





Recurrent Neural Networks

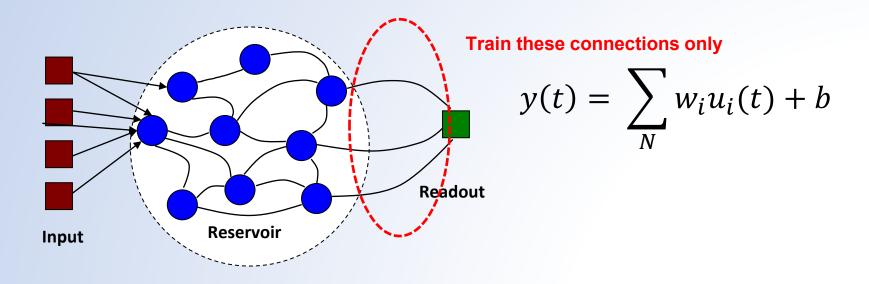
- Feed forward networks create a static transform of input data, but many interesting problems are time dependent.
- Recurrent networks create time-dependent transforms of data...



- 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

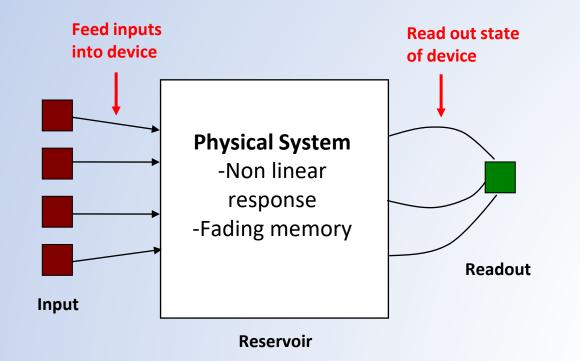
- **Reservoir computing** developed is a development of recurrent neural networks.
- Recurrent network with fixed synaptic weights (the reservoir) connected to a trainable readout layer.



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

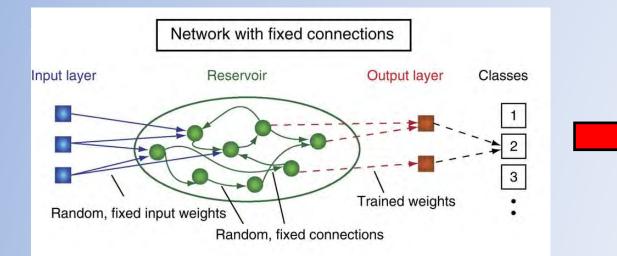
Physical Reservoir Computing

- 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.

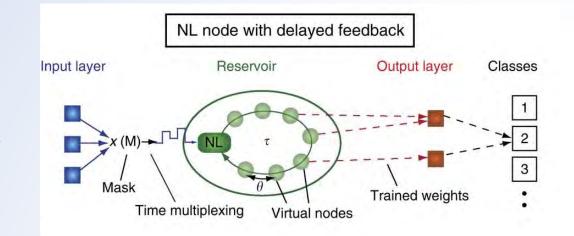


Time multiplexed RC

- 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

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

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 (a Eleni Vasilaki,^{2,a)} (a Guru Venkat,¹ (b) Ian Vidamour,¹² (b) and Chester Wringe³ (b)

AFFILIATIONS

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

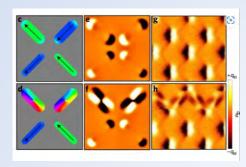
^{a)}Authors to whom correspondence should be addressed: thayward@sheffield.ac.uk and e.vasilaki@sheffield.ac.uk

ABSTRACT

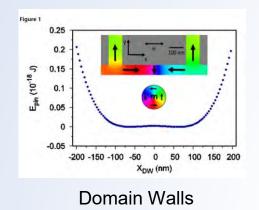
Neural networks have revolutionized the area of artificial intelligence and introduced transformative applications to almost every scientific field and industry. However, this success comes at a great price; the energy requirements for training advanced models are unsustainable. One promising way to 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.

© 2023 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http:// creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/5.0119040

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



Spin Ices

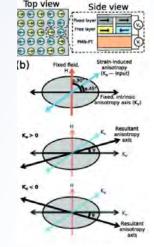


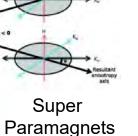
Spin torque Ferromagnet Norma Ferromagne Curren

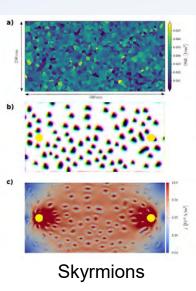
STOs

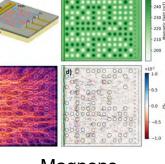
(a)

10-500 nm





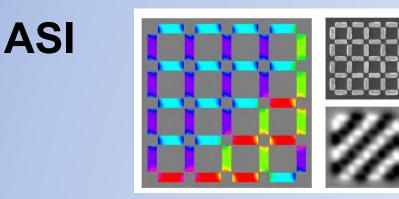




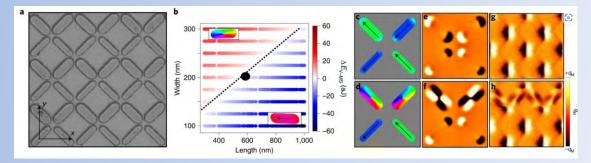
Magnons

RC with Magnetic Metamaterials

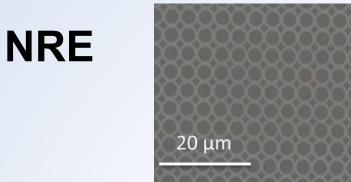
- Magnetic Meta Materials like Artificial Spin Ices and Nanoring Ensembles consist of large numbers of interconnected elements.
- Huge state space, emergent behaviour → potential ideal for reservoir computing!

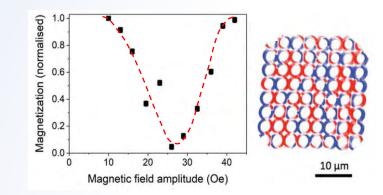


J.H. Jensen et. al. Artificial Life Conference Proceedings, 15-22 (2018)



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

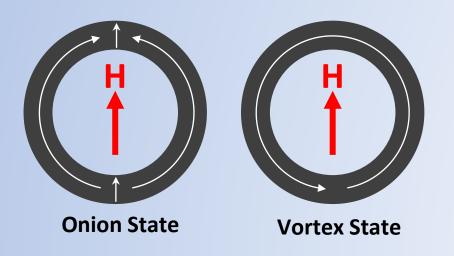


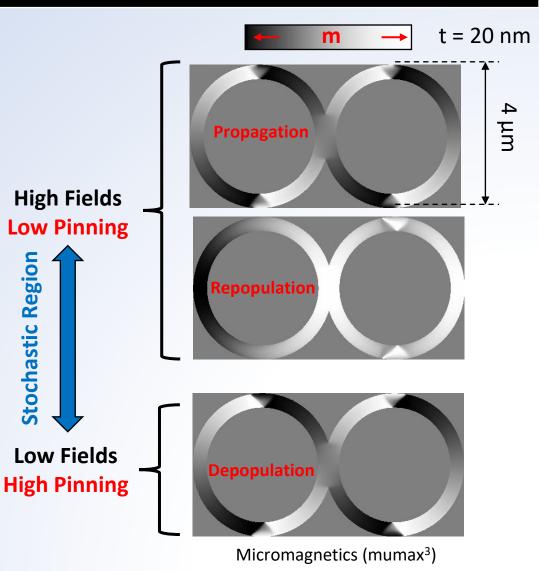


R.W. Dawidek, T.J. Hayward et. al. Adv. Func. Mater. 31, 2008389 (2021)

Domain Walls in Ring-shaped Nanowires

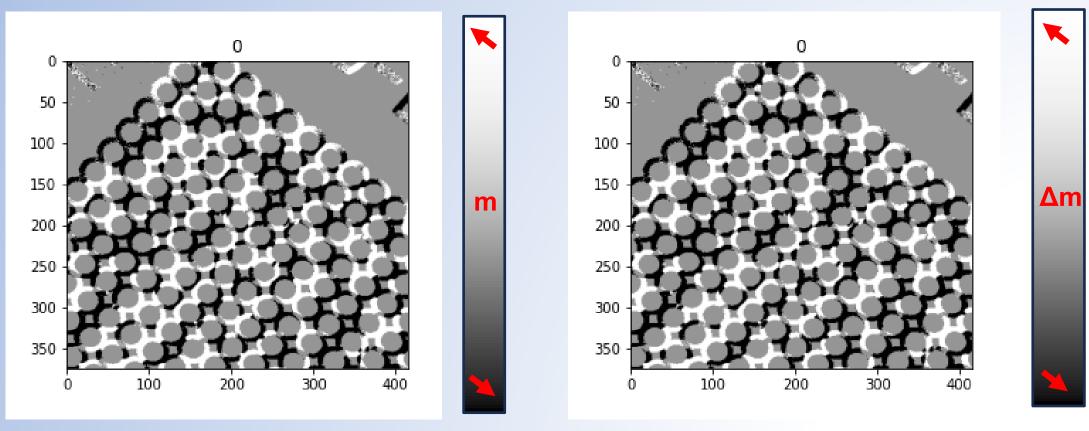
- **Ring-shaped nanowires** form two magnetic states.
 - **Onion state** DWs rotate with rotating applied field.
 - Vortex state circulating magnetisation.
- When multiple rings are connected the junctions act as pinning sites.
 - Produce domain wall interactions that may cause population or depopulation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





Emergent Behaviour in Ring Ensembles

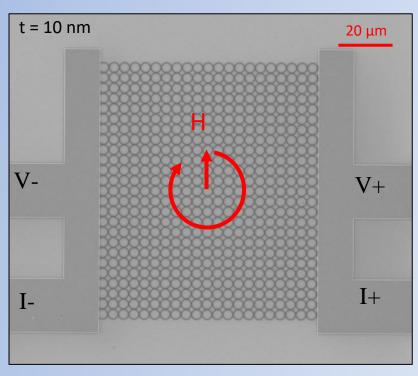
- Complex, emergent interactions in large, interconnected arrays!
- PEEM measurements allow us to **directly visualise** the behaviours!



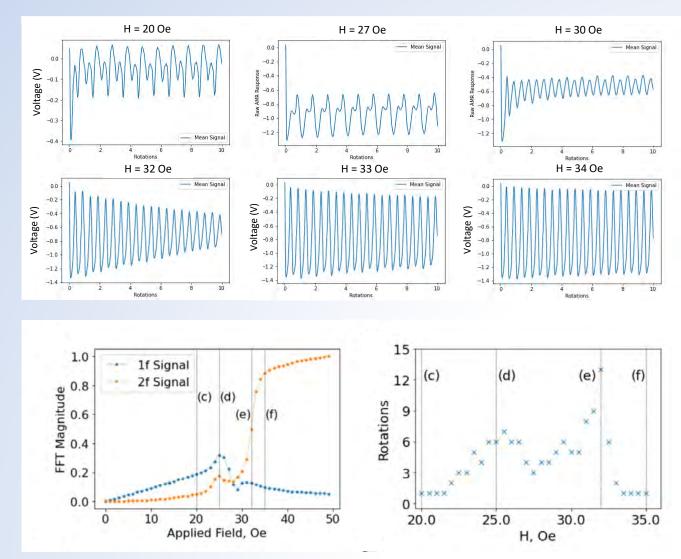


Creating Electrical Devices

- 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)

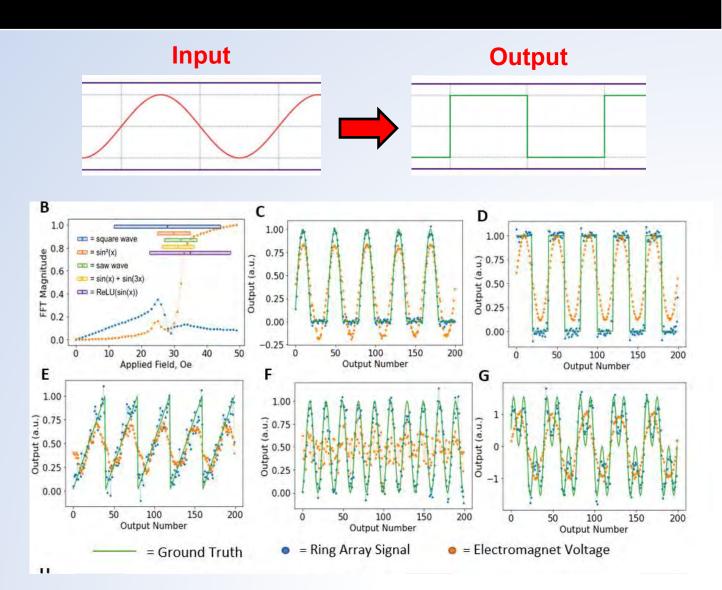


RC: Signal Reconstruction

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

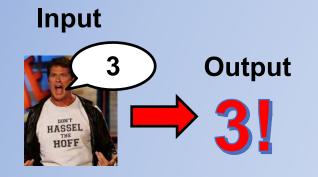
 $H = H_0 + \Delta H * x(t)$

- 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.



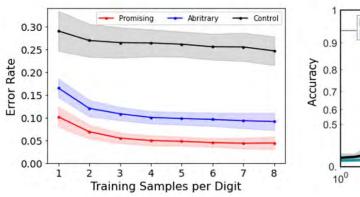
RC: More Complex Tasks

Spoken Digit Recognition

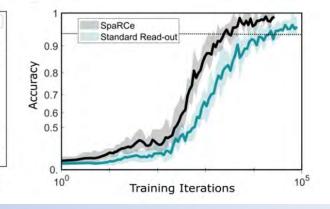


- Classify digits 0-9 for 5x female speakers from TI-46 database
- Data preprocessed via a Mel-Frequency Cepstral filter.

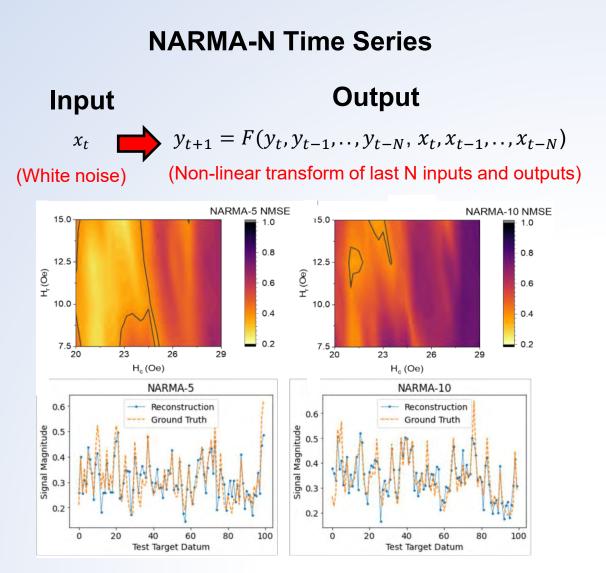
Offline Training



Online Training

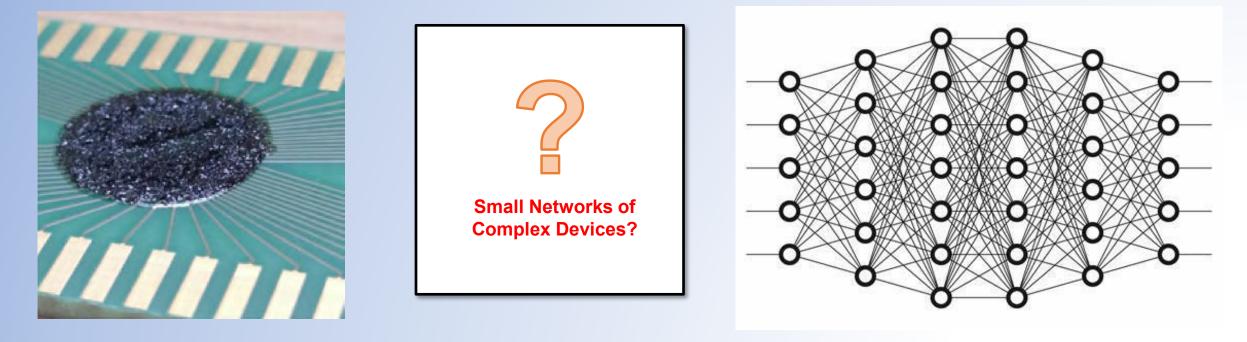


I.T. Vidamour et. al Comms. Phys. 6, 230 (2023)



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**....

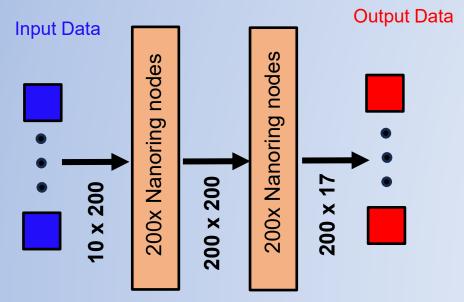


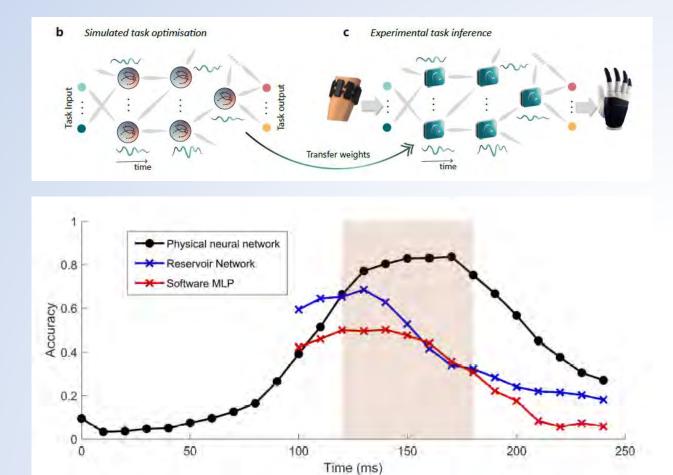
Complex Device, Monolithic

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?