

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.

AI is Amazing…

 \circledS

…but has problematic impacts…

emissions…

With terrifying consequences for

power usage and carbon

'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 Ibs of CO2 equivalent

Charl Mill Technology Heyeev - Source: Strubell at al. - Created with Datawrapper

…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 wⁱ** .
	- 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**:

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

$$
\nabla C(\mathbf{w}_n) = \frac{\partial C}{\partial \mathbf{w}_n}
$$
 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 local and nonvolatile

• **Transports** data through the network.

Lots of magnetism groups working on this!

Neuromorphic computing with nanoscale spintronic oscillators

.
acob Torrejonⁱ, Mathieu Riou¹, Flavio Abreu Araujo¹, Sumito Tsunegi², Guru Khaisa²r, Damien Quetlioz⁴, Paolo Bortolotti¹.
Tucent Cros", Kay Yakoshiji?, Akio Fakushima", Huodu Kubora?, Shinji Yuasa?, Mark D. Neurons in the brain behere as nonlinear oscillators, which develop recognition with an accuracy similar in that of state of the precise information . Taking a meteories for experimental in the precise interaction of the to fit 10⁸ oscillators organized in a two-
a chip the size of a thumb, the lateral dim
must be smaller than one micrometre. He Nanoscale Room-Temperature Multilaver Skyrmionic Synapse tend to be noisy and to lack the stability th
data in a reliable way. For this reason, des **Neural Networks** proposals²⁻⁵ and several candidates, is sperconducting⁷ oscillators, a proof of co
onsputing using nanoscale oscillators has
kre we show experimentally that a nanos Runze Chen D, ^{1,†} Chen Liⁿ,^{2,†} Yu Liⁿ,¹ James J. Miles,¹ Giacomo Indiveriⁿ ² Advanced Processor Technologies (APT) Group, Department of Computer Science, S. and a complete the control of the state of the control of Johannes H. Jensen and Gunnar Tufte

neural network ^o

F This paper was Vasilis F. Pavlidis[®],² and Christoforos Moutafis^{®1,*} Nano Engineering and Spintronic Technologies (NEST) Group, Department of Compu Engineering, University of Manchester, Manchester M13 9PL, United K.

University of Manchester Manchester M13 9PL United Kinodon 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

Norwegian University of Science and Technology, Trondheim, Norjohannes.jensen@ntnu.no

Perspective: Spintronic synapse for artificial

Voltage-controlled superparamagnetic ensembles for low-power reservoir computing

Check for updates

Neuromorphic computation with a single magnetic domain wall

Razvan V. Ababei¹⁶⁶, Matthew O. A. Ellis², Ian T. Vidamour¹, Dhilan S. Devadasan¹, Dan A. Allwood¹, Eleni Vasilaki² & Thomas J. Hayward¹

Machine learning techniques are commonly used to model complex relationships but implemen on digital hardware are relatively inefficient due to poor matching between conventional comp 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 mo 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.

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 ² College 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

LETTER

Vowel recognition with four coupled spin-torque nano-oscillators

mera^{1,3}, Philippe Talatchlan^{1,2}, Sumito Tsanegi², Flavio Abreu Araujo^{1,2}, Vit sent Cros¹, Paolo Bortolotti⁵, Juan Trastov² y Yakınbiji", Akio Fukushima", Ritoshi Kubota", Shinji Yuasa", Maxenor Ernoult¹²³, Damir Vodenicarevke", Tifen
1988 i Octatelli", Damien Operlioz⁷⁴ & Julie Grollier³⁴

In recost years, artificial insural active
changes and the state of the state of the systems, neuron. Hith the
state of the state for the state and computing is achieved hought. We imagine the
interaction the state of the Hence the contribution of the contributio with dynamical functionality, such as o coefflutions, and to edy on
mergent physical phenomenos, such as synchromization
 $^{2.4}$, for subting complex problems with small networks
 $^{1.11}$. This approach is especially inte

proping to success the series of the ser sintrol the frequency of ands acultists he oftmores. The light blue curve in Fig. 1c sh

tps://doi.org/10.1038/s41586-018-0632-

of the oscillators through the nationway
erate, Figure 1d slares that when the freq week each oscillator synchron smical Indeed, when the frequency of the source yets close to the free of one of the oscillators, the strong signal of the source pulls the ada

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

Lots of magnetism groups working on this!

• 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.

• 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 i**ntegrate** with **conventional microelectronics**.
- But could we create **neural networks** that are **all magnetic**?

Domain Wall Logic $5 \mu m$ \mathbf{w} AND II \mathbf{m} \mathbf{m} \mathcal{C}
Rot **NOT** Cross $O+1$ Fan Fan $+$ \bf{I} $-1/2 \ \Omega + 1/2$ $H = 125$ Oe

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 **Ppass**.
	- **Problem for digital devices!!**
- Experimental measurements show that **Ppass** can be **tuned sigmoidally** between 0 and 1 using externally applied bias field H_{bias} .
- Over repeat measurements create an **analogue response** from the

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

Stochastic DW injection

All Magnetic FFNN

 1.0

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 Criffin,³ D Thomas J. Hayward,^{1,a)} D Luca Manneschi,² D Mohammad F. KH. Musameh, ⁴ Simon O'Keefe, ³ Susan Stepney,³ Charles Swindells, **19** Martin A. Trefzer.⁴ in Eleni Vasilaki.^{2,a)} in Guru Venkat.¹ in Ian Vidamour.¹² in and Chester Wringe³ in

AFFILIATIONS

¹Department of Materials Science and Engineering, University of Sheffield, Sheffield S1 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 Ferromagne Norma Ferromagne Curren

10-500 nm

STOs

 (a)

Ton view

Paramagnets

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**!

Micromagnetics (mumax³)

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

 10^5

- 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?**