Department of Materials Science and Engineering

PhD Thesis

Reservoir Computing with Connected Magnetic Nanoring Ensembles

Ian T Vidamour

10/09/2023

Abstract

The concept of *in-materia* computing uses the natural complexity of material systems to perform computational operations naturally as part of the system's inherent response to input stimuli. However, to implement a material system for computation effectively, the physical response of the system must be understood and exploited under a suitable computational framework. This thesis explores the application of arrays of interconnected magnetic nanorings for computation under the framework of reservoir computing. By using a combination of experimental and simulation techniques, the work presented here aims to explore and understand the response of the nanoring arrays, exploit their interesting dynamic properties for computation, and expand upon the computational power achievable with the system.

Firstly, the implementation of a phenomenological model of the nanoring arrays is described, then validated against a range of experimental data covering the static, dynamic, and microstate response of the nanoring arrays with good agreement. This model then serves as a testbed for establishing a suitable paradigm for computing with the nanorings and exploring the computational properties of different regimes of response, ending with a proof-of-concept demonstration of reservoir computing with the nanorings on a benchmark spoken digit recognition task.

Next, the findings made in simulation are used to inform the development of an experimental demonstration of computation. This involved the creation of experimental apparatus to apply stimuli to the nanorings via rotating magnetic fields, and to measure the evolving anisotropic magnetoresistance of the device. Interesting dynamic properties of the system's resistance response are identified and paired with specific reservoir architectures that leverages them to provide different computational properties, evidenced by state-of-the art performances in a range of standard tasks. Finally, the changes in physical behaviour due to manipulations of the array's lattice structure are explored at the microstate level as well as their macroscale response. Computational properties of the different arrangements are evaluated, and lack of microstate resolution in the readout mechanism is attributed to the subtlety of the differences. However, the additional computational power available when different arrangements are combined shows promising scalability for devices of the nanoring arrays.

Declaration

This thesis is the result of my own work, all works done in collaboration with others are accredited accordingly either in the preamble to research chapters, or in the 'Author Contributions' section of the corresponding paper. No part of this thesis has already or is currently being submitted for any other qualification than the degree of Doctor of Philosophy at the University of Sheffield.

Acknowledgements

I'd like to open this thesis by thanking all of the people who made it possible. Firstly, I'd like to thank my supervisors, Dr Tom Hayward, Prof Eleni Vasilaki, and Prof Dan Allwood, for their incredible guidance and advice. Their enthusiasm, wisdom, and guidance have been instrumental in the completion of this thesis. Thank you for making it a pleasure to come into work every day, for inspiring me to further pursue a career in academia, and for making every meeting interesting, especially the ones that derailed into 20-minute conversations about video games.

I would also like to give a special thanks to Dr Guru Venkat and Dr Charles Swindells, whose help has been invaluable along the way. I am very grateful to be involved in a project with this level of collaboration, and I can confidently say the work presented here has benefitted greatly from their unwavering support. Now that the thesis is submitted, hopefully I can return the favour without the need for seven reminder emails. I would also like to thank Dr Matthew Ellis and Dr Luca Manneschi for helping me in coming to grips with the machine learning side of things, and I'm excited to work more with you guys in the future. I'd like to thank everyone present and past who has worked with me in J10, for all the good laughs along the way, and all the listening ears while I vented about having commented out the 'save' command for the third time this month.

I'd like to thank my friends and family for everything they have done not just during the PhD, but at every step of the way. I'd especially like to thank my Mum, Dad, Sister, and Nan, their faith in me has always been a motivation at every stage of study. I can't express how grateful I am for every one of you. Finally, I'd like to thank my amazing girlfriend, Leah. Your selfless support is inspirational, and I hope I can return the favour during the remainder of your own PhD. Over to you now!

Table of Contents

1-Introduction

The rapid growth of the fields of machine learning and artificial intelligence in the last decade has brought astounding developments in capabilities. However, there are also significant concerns over the energy usage for training these models. While many of the tasks that these models perform resemble those performed by the brain, such as speech and language processing[1], vision[2], and pattern recognition[3], there is a huge disparity in the efficiency of these models compared to the brain. For example, in a famous exhibition of man-versus-machine in 2016 where Google DeepMind's AlphaGo became the first machine to defeat a world champion player in the game of Go, there was a huge mismatch in the computational power available to each competitor; AlphaGo is capable of 30 trillion computations per second, requiring 170kW of power[4], while the brain of the machine's opponent, Mr. Lee Sedol, required just 20 W of power[5], and even beat the machine in one of the five matches.

In typical machine learning approaches, the neurons and synapses are represented numerically in silico, with large matrices storing the activities in neurons, as well as weights of the synaptic connections between them[6]. While vectorisation of mathematical operations provides rapid evaluation, with further acceleration through use of graphical processing units (GPUs) or tensor processing units (TPUs) designed to perform rapid tensor operations[7], the growth in computation required by modern models has outpaced the development of hardware, shown in figure 1. It is clear that a paradigm shift is also required in machine learning hardware to match the demands of novel machine learning techniques.

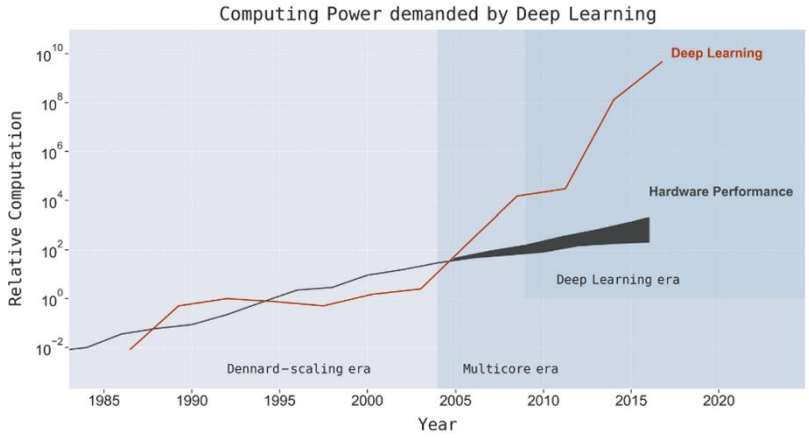


Figure 1- Graph comparing the relative capabilities of computing hardware (black) to the relative computational demands of machine learning (red), normalised against values from 1994. Taken from[8].

One critical factor in this efficiency mismatch owes to the difference in the types of computational operation performed by the brain compared to conventional computers. Away from the realm of binary abstraction, the brain operates in a distributed, highly parallelised manner; the high degree of interconnectivity and inherent structures of the brain providing specialised roles in transforming sensory data[5]. This is the opposite of how computation is performed in silico, with an exact, procedural approach towards processing information, with discrete components for computation and memory. Figure 2 shows the drastic difference between biological systems both in terms of power usage as well as distance between memory and processing units. This creates the so-called "von-Neumann bottleneck"; the computational overhead associated with the need for shuttling data between these discrete components[9].

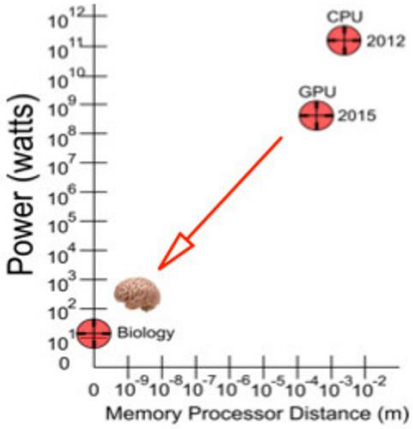


Figure 2- Comparison between the power usage/memory processor distance for state-of-the-art machine learning implementations on CPU/GPU hardware and biological systems. Adapted from[9]

By taking inspiration from the brain, the field of neuromorphic engineering attempts to replicate the functionality of brains by designing hardware that operates in functionally similar ways[10]. At the heart of this process was the same conventional computing components but operating in an analogue fashion, replicating the sensory processes that occur in biology to create machines that can respond to pressure[11], sound[12], and light[13] using with great success. However, these approaches were generally limited to sensing of data rather than processing it, only encapsulating part of the brain's functionality.

Unconventional or *In-Materia* computing aims to build on the field of neuromorphic engineering using unconventional material platforms, where the dynamic properties of the material substrate are chosen for their synergy with the dynamic nature of signal processing tasks[14]. In these approaches, the nonlinear and hysteretic properties of the material substrate are used to perform memory and computation operations in tandem, functionally akin to how the brain processes information. Additionally, by responding to continuous input stimuli directly, material systems can be exploited to perform computation without the need for analogue-to-digital conversion, allowing more direct interfacing between the computational platform and the environment in which it operates.

Many different types of material system have been proposed for computational purposes, including memristors, micro-electromechanical systems (MEMS)[15], analogue electronics[16], optical systems[17], and spintronic platforms[18]. Each of these systems come with their own inherent strengths and weaknesses: memristors are easy to miniaturise and interconnect, though suffer from degradation over many usage cycles, MEMS systems can be similarly integrated with standard electronics, but are prone to damage from mechanical stress and adverse environments[19]. Analogue electronics offer low power consumption and can be constructed from common off-the-shelf components[20]. Optical systems offer high data throughput but are often relatively large due to the need for optical-fibre delay loops to induce complex transient behaviours[21]. Spintronic platforms, such as the interconnected nanoring arrays studied here, are promising due to the possible non-volatility of their magnetic state[22], ability to provide low power input and readout by driving via spin-torque effects, and simple measurement via magnetoresistance effects[23].

Each of these families of device have seen recent proof-of-concept deployments in relatively simplistic computational settings, often under the paradigm of 'Reservoir Computing', RC. In RC's initial algorithmic implementation of echo state networks (ESNs)[24], the network consists of three distinct layers: an input layer, a reservoir layer, and a readout layer, shown in figure 3. The input layer passes information to an untrained, randomly connected recurrent neural network, termed a reservoir. This

reservoir provides temporal transformations of data via its recurrent connections and the leakage rate of nodes in the network, with nonlinearity provided by the nodes' activation function. This is used to provide a higher-dimensional representation of data which enables the linear readout layer of the reservoir states to make regression/classification decisions with increased accuracy than would be achievable on the unprocessed input data.

RC is especially harmonious with computation with physical systems because the requirements for the reservoir layer are easily met by a physical dynamical system. A recurrent neural network is mathematically a dynamical system, and the *specific* transformation is not critical, only that the transformation is suitably nonlinear and has timescales matching that of the task[25]. Contrary to standard neural networks, reservoirs are tuned not by training the connectivity, but by selecting initialisation parameters which scale the magnitudes of input weights and internal weights, as well as how long each node in the system retains information. Analogies to this can be easily achieved in physical systems by controlling the input scaling and duration of each input to the system. Secondly, the availability of time-multiplexing procedures to generate virtual networks of interconnected nodes from a single dynamical system means that proof-of-concept demonstrations of computation can be achieved with any system that exhibits nonlinearity and hysteresis, can be stimulated with a controllable input signal, and has some state variable that depends upon the input and can be measured reliably.

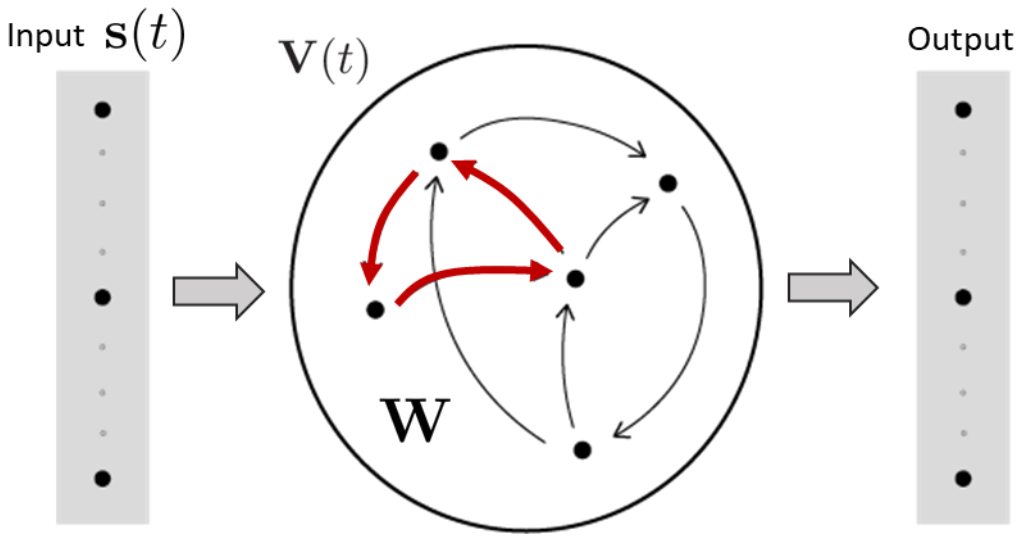


Figure 3- Schematic diagram of an Echo State Network, showing input, reservoir, and output layers. A recurrent connection loop is shown by the red arrows.

While initial demonstrations of computation with an arbitrary material may be relatively simple, for the proposed system to form a plausible candidate for future computing applications, additional criteria must be met. For example, initial exploratory demonstrations of unconventional computing showed the ability to recognise patterns from the interference patterns generated when water waves produced by controllable paddles interact on the surface a bucket of water[26]. Clearly, buckets of water will not form the next generation of machine learning architectures, suggesting that a more pertinent question than *can* a material system perform computation is *should* a material system be used for computation? The impact of a physical implementation of RC often hinges upon two key evaluations of a proposed system: Firstly, is there a route to achieving a system in which both input and output can be addressed in a device-tractable manner, allowing physical demonstrations to be achieved outside

of simulations. Secondly, does the substrate itself offer computationally interesting properties that can be exploited, such as non-volatility of state, controllability of timescales of response, or reconfigurability of regimes of operation that can be accessed post-manufacture.

The physical system explored in this thesis consists of an array of interconnected magnetic nanorings[27]. The arrays have sizes on the order of 100 micrometres, and proven routes to manufacture via electron beam lithography and the lift-off method. The continuous nature of the arrays mean that the system state can be easily evaluated via electrical transport measurements. The system has been shown to exhibit complex, emergent behaviours when driven by rotating magnetic fields due to interactions between the individual rings, with a non-volatile spatially distributed state-space determined by the domain structure of the nanorings. Hence, this system represents an exciting candidate for reservoir computing as it meets both of the criteria outlined above.

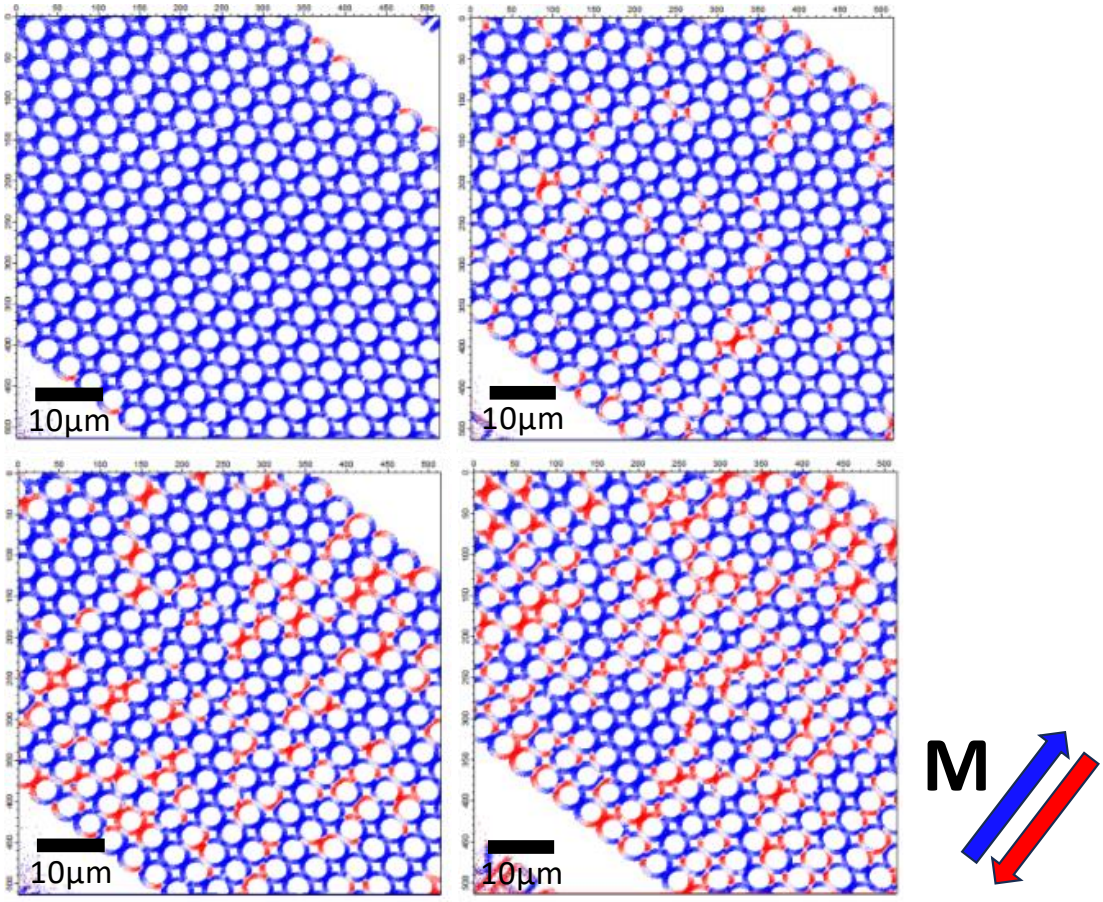


Figure 4- X-ray photoelectron emission microscopy images of nanoring networks after successive applications of a rotating magnetic field. Colour contrast represents magnetisation direction.

The objective of this thesis is to outline a pipeline from exploration of the physical processes of nanoring arrays to an experimental demonstration of how these systems can be used for computation, and how these computational capabilities can be evaluated and expanded. Each of the four main research chapters in this thesis addresses a specific research question, which each represent a key stage in the progression of the overall design process.

The first question that was explored was: How can the physical processes that occur in the nanoring network be represented so that the response of the system can be described effectively in a model? This is critical for developing an understanding of how the nanoring arrays respond to magnetic fields,

as well as providing a means for exploring the rings in simulation. The standard micromagnetic approaches used for simulating the response of magnetic materials are entirely unsuitable for the relatively large scale of the entire network, as well as the relatively long durations of response which require simulating. It is clear that a new methodology for modelling such a system was required. The agent-based model

The second research question was how can the dynamic response of the ring arrays be exploited in order to implement reservoir computing, and what computational properties can the system exhibit? To answer this, the model that was constructed to answer research question 1 was then used to explore potential paradigms for encoding information to the system, as well as which physical properties of the system would be useful as readout variables for RC. It was also critical to establish which regimes of device response were suited to which type of tasks, and the use of the model allowed extensive exploration into the different computational properties that are available to the nanoring arrays, serving as a proof-of-concept for computation with the nanoring arrays.

The third research question focussed on how can the proof-of-concept demonstration in simulation be actualised for experimental demonstrations of computation in the laboratory. This is critical for demonstrating the viability of magnetic nanoring arrays as a potential platform for unconventional computing, as it shows that the findings provided by the model can be extended to a real-world setting. It is also key to demonstrate that this is achievable with an input and measurement apparatus that has realistic routes to be engineered into a device that would be suitable for implementations outside of the laboratory.

The final research question was how can the magnetic system be engineered to expand upon the computational properties that were achieved in the initial demonstration. This is an important development as it is key that these systems must show scalability if they are to be implemented as a novel computation platform. To do this, an understanding of how the response of a nanoring array can be manipulated to provide different transformations of input, as well as a mechanism for exploiting these different transformations in tandem, must be developed.

1.1 – Thesis Outline

Chapter 2 contains an outline of the basic theory of magnetism, as well as more detailed information on phenomena used to drive the dynamic magnetic nanoring system and measure its response. Additionally, the fundamental machine learning principles used in the thesis are described, including the theory of constructing, training, and performing inference with artificial neural networks and reservoir computers.

Chapter 3 presents a literature review. This begins by outlining behaviours observed in magnetic nanowire systems, acting as a basis for understanding the nanorings studied here, before outlining existing literature on the single nanorings, small nanoring networks, and finally large arrays. Then advancements in reservoir computing paradigms are explored, focussing on those which would aid in the implementation of hardware RC. Finally, a survey of the current state of research into in-materia computing is presented, assessing the relative strengths and weaknesses of existing proposals, and highlighting gaps in current understanding that can be expanded upon.

Chapter 4 outlines the experimental techniques employed in this thesis. It will describe both the physical principles of different approaches, as well as the specific procedures that were performed here. The chapter will cover device manufacture, measurement of magnetic state via electrical transport measurements, as well as specialist techniques for characterisation performed at external facilities, such as X-ray photoelectron emission microscopy and polarised neutron reflectometry.

The main research results of the thesis will be presented as four publication-format chapters, each covering a different phase of the deployment of the nanoring system for computation.

Chapter 5 presents an article that describes the construction and validation of an agent-based model of the emergent behaviour in interconnected nanoring arrays. The software models the outcomes of stochastic pinning events, domain wall-domain wall interactions, and phenomenological nucleation and annihilation of domain walls. The model is then validated against a range of experimental data on the equilibrium response of the arrays, the dynamic timescales associated with a change in input, and the types of microstates formed under a range of applied fields.

Chapter 6 uses the model described above to demonstrate the feasibility of reservoir computing with the nanorings in simulation. These utilise a simple encoding/decoding paradigm, where input data is used to modulate the magnitude of a rotating magnetic field, and data is read out from the average properties of the arrays magnetic state (net magnetisation components, number of domain walls in the system). The simulations explore the range of computational properties that are available in the system, characterising these using task-agnostic metrics that evaluate the ability of the system to separate, generalise, and remember arbitrary inputs to the system. It is then shown how these metrics are correlated to performance in a spoken digit recognition task, thus demonstrating the how evaluating performance in metric space can expediate the process of hyper-parameter tuning.

Chapter 7 focuses on experimental demonstrations of computation in real devices. Firstly, anisotropic magnetoresistance response of the nanoring arrays is presented, and the basic physics underlying these explained. Next, the range of computational properties that can be extracted from the devices by exploiting three different reservoir architectures are explored. Each of these architectures are designed to exploit a specific dynamic property of the system's response and these allow performance matching or improving upon the state-of-the-art in for spintronic devices in three distinct tasks to be demonstrated.

Chapter 8 explores how the lattice arrangement (Kagome, trigonal, and square lattices) of the nanoring arrays influences their magnetic behaviour, and how these changes in behaviour affect their

computational capabilities. It is shown that behaviours are markedly different at the microstates level, that this difference manifest less strongly in some the global electrical measurements of the nanoring arrays. The computational properties of each of the lattice arrangements is then evaluated, with the results showing that although they show broadly similar metrics individually these can be expanded upon by combining the lattices together.

Chapter 9 presents the conclusions to the thesis, summarising the key findings presented in each of the research chapters, their impact, and the follow-on work that its conclusions led to. It also features a reflection upon potential studies for future work, the critical next steps in the development of the devices, and an evaluation of the credibility of the system as a potential computing platform.

References:

1. Floridi, L. & Chiriatti, M. GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds Mach.* **30**, 681–694 (2020).
2. Deep Learning for Computer Vision: A Brief Review. https://www.hindawi.com/journals/cin/2018/7068349/.
3. Wang, J., Chen, Y., Hao, S., Peng, X. & Hu, L. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognit. Lett.* **119**, 3–11 (2019).
4. Chen, J. X. The Evolution of Computing: AlphaGo. *Comput. Sci. Eng.* **18**, 4–7 (2016).
5. Hubel, D. The Brain. *Sci. Am.* **241**, 45–53 (1979).
6. Goos, G. *et al.* Neural Networks: Tricks of the Trade. doi:10.1007/978-3-642-35289-8.
7. Shi, S., Wang, Q., Xu, P. & Chu, X. Benchmarking State-of-the-Art Deep Learning Software Tools. in *2016 7th International Conference on Cloud Computing and Big Data (CCBD)* 99–104 (2016). doi:10.1109/CCBD.2016.029.
8. Thompson, N. C., Greenewald, K., Lee, K. & Manso, G. F. The Computational Limits of Deep Learning. Preprint at https://doi.org/10.48550/arXiv.2007.05558 (2022).
9. Zou, X., Xu, S., Chen, X., Liang, Y. & Han, Y. Breaking the von Neumann bottleneck: architecture-level processing-in-memory technology. *Sci. China Inf. Sci.* **64**, (2021).
10. Mead, C. How we created neuromorphic engineering. *Nat. Electron.* **3**, 434–435 (2020).
11. Markoff, J. Pad to Replace Computer Mouse Is Set for Debut. *The New York Times* (1994).
12. Lyon, R. F. & Mead, C. An analog electronic cochlea. *IEEE Trans. Acoust. Speech Signal Process.* **36**, 1119–1134 (1988).

13. Mead, C. A. & Mahowald, M. A. A silicon model of early visual processing. *Neural Netw.* **1**, 91–97 (1988).
14. Adamatzky, A. *Advances in Unconventional Computing: Volume 1: Theory*. (Springer, 2016).
15. Dion, G., Oudrhiri, A. I.-E., Barazani, B., Tessier-Poirier, A. & Sylvestre, J. Reservoir Computing in MEMS BT - Reservoir Computing: Theory, Physical Implementations, and Applications. in (eds. Nakajima, K. & Fischer, I.) 191–217 (Springer Singapore, 2021). doi:10.1007/978-981-13-1687-6_9.
16. Soriano, M. C. *et al.* Delay-Based Reservoir Computing: Noise Effects in a Combined Analog and Digital Implementation. *IEEE Trans. Neural Netw. Learn. Syst.* **26**, 388–393 (2015).
17. Paquot, Y. *et al.* Optoelectronic reservoir computing. *Sci. Rep.* **2**, (2012).
18. Allwood, D. A. *et al.* A perspective on physical reservoir computing with nanomagnetic devices. *Appl. Phys. Lett.* **(in review)**, (2022).
19. Dion, G., Mejaouri, S. & Sylvestre, J. Reservoir computing with a single delay-coupled non-linear mechanical oscillator. *J. Appl. Phys.* **124**, 152132 (2018).
20. Soriano, M. C., Brunner, D., Escalona-Morán, M., Mirasso, C. R. & Fischer, I. Minimal approach to neuro-inspired information processing. *Front. Comput. Neurosci.* **9**, (2015).
21. Jacobson, P. L., Shirao, M., Yu, K., Su, G. L. & Wu, M. C. Hybrid Convolutional Optoelectronic Reservoir Computing for Image Recognition. *J. Light. Technol.* (2021) doi:10.1109/JLT.2021.3124520.
22. Gartside, J. C. *et al.* Reconfigurable training and reservoir computing in an artificial spin-vortex ice via spin-wave fingerprinting. *Nat. Nanotechnol.* **17**, 460–469 (2022).
23. Torrejon, J. *et al.* Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017).
24. Jaeger, H. *The "echo state" approach to analysing and training recurrent neural networks- with an erratum note*. *GMD Technical Report* (2001).

25. Lukoševičius, M., Jaeger, H. & Schrauwen, B. Reservoir Computing Trends. *KI - Kunstliche Intell.* **26**, 365–371 (2012).

26. Fernando, C. & Sojakka, S. Pattern Recognition in a Bucket. in *Advances in Artificial Life* (eds. Banzhaf, W., Ziegler, J., Christaller, T., Dittrich, P. & Kim, J. T.) 588–597 (Springer, 2003). doi:10.1007/978-3-540-39432-7_63.

27. Dawidek, R. W. *et al.* Dynamically-Driven Emergence in a Nanomagnetic System. *Adv. Funct. Mater.* **31**, (2021).

2 - Theory

2.0 - Outline

Due to the interdisciplinary nature of the work, this thesis draws from concepts based both in materials science and computer science. The following chapter will focus on outlining the theoretical frameworks behind key concepts in both magnetism and machine learning and will consist of separate sections addressing each of these topics respectively in relatively phenomenological terms to provide accessible background information for readers of either discipline.

In the first section, we will consider the atomic origins of magnetism and the different types of magnetic material, before delving deeper into the energy contributions that dictate the magnetisation state of ferromagnetic materials, as well as the formation of domains and domain walls. Next, the dynamics of magnetic moment vectors in response to applied fields will be described using the framework of the Landau-Lifshitz-Gilbert equation. Finally, the mechanisms that lead to changes in electrical resistance associated with changing magnetic state will be defined for both anisotropic magnetoresistance (AMR) as well as giant magnetoresistance (GMR), demonstrating how magnetic states can be measured electrically.

The second section will take a didactic approach to introducing the topics of artificial neural networks, supervised learning, and reservoir computing. The process of error backpropagation will be introduced, describing how we can train feedforward networks under the supervised learning paradigm. Next, the concept of recurrence and time-dependent networks will be introduced, as well as the adaptations to learning rules that must be applied to train these networks with algorithms such as backpropagation through time. Finally, the machine learning paradigms of reservoir computing (RC) framework will be introduced, explaining how they differ in terms of training and structure from standard recurrent neural networks (RNNs), and the archetypal reservoir architecture of echo state networks (ESNS) will be outlined in detail.

2.1 – Magnetism

The magnetic properties of a material arise from the collective interaction of magnetic moments of the individual atoms that constitute the material, with the electronic configuration of the atoms determining the net moment of the atom. The magnetic behaviour a material system depends not only upon the presence of net magnetic moments from its atoms, but also their ordering. This section will introduce the origins of magnetic moments and how to determine the filling of electronic orbitals (and hence net moment) for atoms, before exploring the different types of magnetic material. Ferromagnetism will then be explored in detail, outlining the key interactions that underpin the behaviours observed in the magnetic nanorings studied in this thesis.

2.1.1 – Atomic Magnetic Moments

Fundamentally, the magnetic moments of individual atoms originate from the electrons within that atom. In classical mechanics, the generation of magnetic fields is associated with the movement of charged particles. Analogously, the quantum mechanical description of magnetic moments arises from the angular momentum terms associated with each charged electron. The angular momenta of electrons can be considered as two separate contributions: the ‘spin’ and ‘orbital’ angular momentum terms, labelled s_i and l_i for the spin and orbital momenta contributions of a single electron i respectively.

The spin angular momentum is an intrinsic, quantised property of the electron itself. Each electron has a spin angular momentum equivalent to $\pm\hbar/2$, where $\hbar$ denotes the reduced Planck's constant. These positive/negative momenta are commonly referred to as 'spin-up' and 'spin-down' respectively. Via the Pauli exclusion principle, a pair of electrons occupying any orbital in the electronic band structure must have opposite spin states, which cancel one another. Hence, only systems with unfilled orbitals may have a net contribution of spin angular momentum.[1]

The orbital angular momentum contribution is not an intrinsic property of the electron but depends upon the motion of an electron around the nucleus of the atom. Quantum mechanically, the movement of an electron is described by its wavefunction which describes relative probabilities of an electron occupying a particular state, with the orbital angular momentum given by the rotational motion associated with the wavefunction. The orbital angular momentum is determined by the azimuthal quantum number, l, which takes integer values, determined by the sub-shell the electron occupies. Only atoms with unfilled shells have zero net orbital angular momentum.[1]

In atoms/molecules with multiple electrons, the calculation of the total angular momentum depends upon the degree of spin-orbit coupling. For lighter atoms, electrons' spins interact amongst themselves, and the total spin angular momentum **S** is given by the sum of each electron's spin (2.1):

$$\boldsymbol{S} = \sum_i s_i \tag{2.1}$$

Similarly, for light atoms the overall orbital angular momentum vector **L** of an atom is the sum of individual contributions by each electron (2.2):

$$\boldsymbol{L} = \sum_i l_i \tag{2.2}$$

The total angular momentum, J, is then given by the sum of **S** and **L** (2.3):

$$\boldsymbol{J} = \boldsymbol{S} + \boldsymbol{L} \tag{2.3}$$

However, for larger atoms, there is significant spin-orbit coupling and hence interactions between the spin and orbital contributions must be considered. Here, the total angular momentum for each electron, j_i, is determined first, and then J is taken as the sum of all individual contributions (2.4):

$$\boldsymbol{J} = \sum_i j_i = \sum_i (s_i + l_i) \tag{2.4}$$

The magnetic moment, **μ**, of an atom is directly proportional to the total angular momentum vector, **J**, of the atom, and the gyromagnetic ratio of the system, γ, which has dependencies upon relativistic effects, but is largely governed by the charge and mass of the system: it is directly proportional to charge, and inversely proportional to the mass (2.5):

$$\boldsymbol{\mu} = -\gamma \boldsymbol{J} = \frac{g\mu_B}{\hbar}\boldsymbol{J} \tag{2.5}$$

where μ_B denotes the Bohr magneton, and g the Landé g-factor.

Since **J** determines the magnetic moment of a system, and **J** depends upon the electronic structure of the system, the rules which govern the filling of electronic orbitals are critical to understanding the magnetic properties of a system.[1]

2.1.2- Hund's Rules

The electronic structure of a system is ordered such that the overall energy of the system is minimised, whilst ensuring that the Pauli exclusion principle is preserved. Hund's rules provide an overview of principles which determine the electronic structure of a system's ground state, and are as follows:

1. The spin angular momentum of the system **S** is to be maximised, which in turn minimises the Coulombic repulsion by increasing average distance between electrons, reducing the shielding they provide to one another and hence yielding a lower energy state. To maximise spin angular momentum, orbitals within a shell are first occupied with parallel spins.[2]
2. The orbital angular momentum **L** is to be maximised, which again minimises Coulombic repulsion. This can be understood classically as electrons orbiting in the same direction will meet less frequently than electrons orbiting opposite to one another. To maximise orbital angular momentum, electrons will first occupy orbitals with the same sign of l.[2]
3. In the case of systems with less than half-full shells, the arrangement with the lowest allowed **J** value will be energetically favourable. If a shell is less than half full, the lowest value J can take is $|\boldsymbol{L} - \boldsymbol{S}|$. For more than half filled shell, the reverse is true and **J** is maximised, with maximum $\boldsymbol{J} = |\boldsymbol{L} + \boldsymbol{S}|$.

While Hund's rules provide a good basis for determining the electronic structure of light elements, they assume that spin-orbit coupling is weaker than the coupling between individual orbital angular momenta and individual spins. Since the relationship between the magnitude of spin-orbit coupling and the atomic number of the atom is quartic, spin-orbit contributions are negligible for lighter atoms, but become critical for heavier atoms, acting contrary to the underlying assumption of Hund's rules.

2.1.3- Types of Magnetism

All materials- even those with no net magnetic moments- interact with magnetic fields in one way or another. This section will outline the different types of magnetic materials and a phenomenological description of how their magnetic properties arise. The differences in magnetic behaviours between the different forms of magnetism can be easily described via the magnitude and sign of their susceptibility response, which describes how the material is magnetised in the presence of external fields, shown in figure 1.

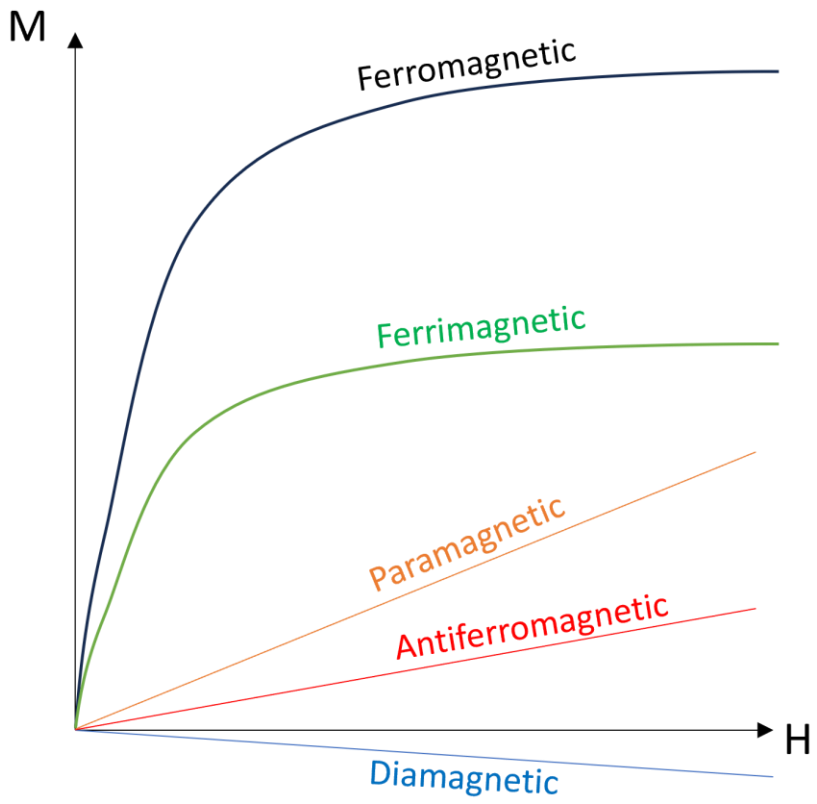


Figure 1- Characteristic magnetisation M vs applied field H responses for the different type of magnetic materials. The susceptibility, χ, is given by the gradient of each line.

2.1.3.1- Diamagnetism

While materials with full electron shells don't exhibit any net magnetic moments in isolation, the change in orbital motion that occurs under the influence of an external magnetic field can induce a moment in the material. While this effect occurs in all materials, it is such a weak interaction that only materials with no intrinsic net magnetic moments (those with full valence shells) are referred to as diamagnetic. The phenomenon can be interpreted classically via Lenz's law; the orbital motion of electrons in the atoms induces currents in the presence of an applied field, producing a force in opposition to the applied field. The magnitude of this force varies linearly with applied field, producing a tiny, negative magnetic susceptibility χ[3], shown in figure 1.

2.1.3.2- Paramagnetism

Paramagnetism occurs in materials with unpaired electrons, and hence net magnetic moments. Here, moments are so weakly coupled to one another that they become practically independent, and the presence of thermal energy causes fluctuations in the material which leads to random alignment of magnetic moments. In the presence of external fields, interactions between the magnetic moments and the field leads to a biasing of alignment along the direction of the applied field, as shown in figure 2, leading to a small positive susceptibility (see figure 1).[4]

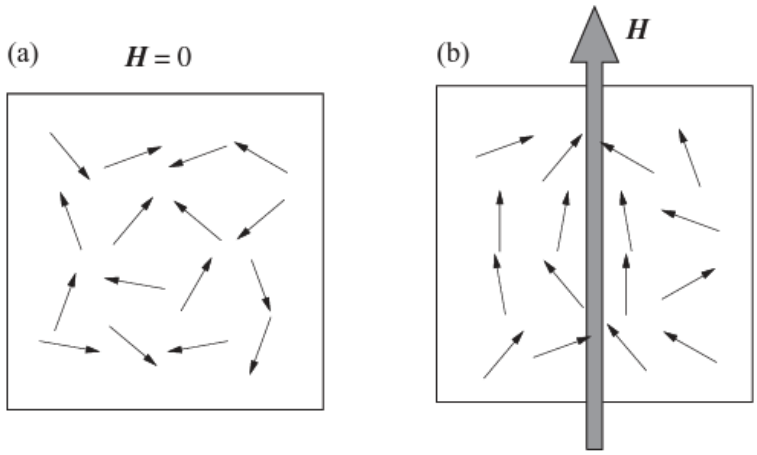


Figure 2- (a) Schematic diagram of magnetic moments in a paramagnetic material with no external field applied, with thermal fluctuations leading to random moment direction (black arrows). (b) The same material under an external magnetic field, H, in the direction of the grey arrow. Although moments do not completely align, a biasing in their direction can be observed. Taken from[4].

2.1.3.3- Ferromagnetism

Whilst also possessing positive susceptibility, ferromagnetic materials differ from paramagnetic materials in that they are able to retain net magnetic moments in the material even in absence of external fields. This arises from the presence of short-range coupling via the exchange interaction, which favours parallel alignment of local magnetic moments in ferromagnetic materials. This coupling is able to overcome the disordering presence of thermal energy, but to a finite extent. At temperatures beyond the Curie temperature, T_C, of a ferromagnetic material, the thermal fluctuations once again overpower the coupling between moments, and the material becomes paramagnetic.[5]

The magnetic behaviours of ferromagnetic materials are often characterised by what is known as a hysteresis loop, which explores how magnetisation changes with respect to externally applied fields and is shown in figure 3. Under strong enough external fields, the magnetic moments become aligned, with any further increase in field causing little change to the magnetisation of the material. This point is called saturation magnetisation, M_s. With the external field removed, exchange coupling means that some net magnetisation is maintained, called the remenant magnetisation, M_r. If the applied field is reversed, the magnetisation reduces once again, and the point at which net magnetisation becomes zero is termed the coercive field of the material, H_c. Increasing this negative field further begins to magnetise the material in the opposite direction, until the material once again becomes saturated. The resulting plot of magnetisation with respect to applied field when taken to saturation in both directions is called the major hysteresis loop, and loops that don't result in saturation are termed minor loops. The concept of ferromagnetic domains is crucial to describing remenant magnetisation and will be discussed in section 2.1.5.

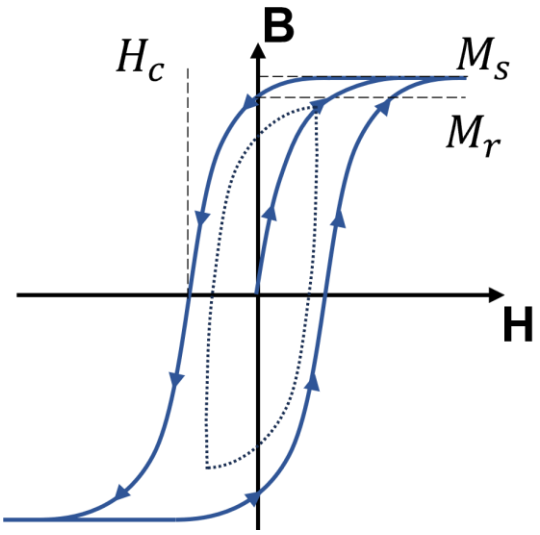


Figure 3- Diagram of typical magnetic hysteresis loop. Solid blue line reflects a major hysteresis loop, with arrows showing direction of magnetisation loop as field is swept. Dotted line reflects a typical minor hysteresis loop. Critical magnetic properties that can be determined from the loop, such as coercive field, saturation magnetisation, and remenant magnetisation, H_c, M_s, and M_r respectively, are labelled.

2.1.3.4- Antiferromagnetism

Antiferromagnetic materials exhibit exchange coupling like ferromagnetic materials, however antiparallel alignment is preferred instead, shown in figure 4. In these materials, external magnetic fields will deflect the antiparallel magnetic moments of the individual atoms away from one another, again producing positive magnetic susceptibility, but the material is unable to retain magnetisation when the field is removed. Analogously to the Curie temperature of ferromagnetic materials, the Néel temperature of antiferromagnetic materials is the point at which the material becomes paramagnetic, reducing the magnetic susceptibility of the material, though remaining positive.[6]

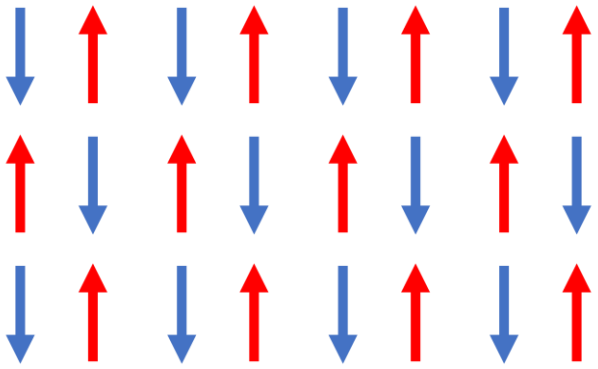


Figure 4- Typical spin structure of an antiferromagnetic material, with alternating spin up (red arrows) and spin down (blue arrows) of equal magnitude. The moments cancel one another out leading to no net magnetisation in absence of external fields.

2.1.3.5- Ferrimagnetism

Ferrimagnetism can be seen as somewhat of an intermediate between ferromagnetism and antiferromagnetism and arises from the combination of heterogenous magnetic elements within the lattice structure of a material. Like ferromagnetic materials, ferrimagnetic materials possess a net magnetisation in the absence of applied fields, whilst the local arrangement of magnetic moments is antiparallel as with antiferromagnets. The resulting magnetic moment arises from a disparity between

the magnitudes of the oppositely oriented moments within the lattice, shown in figure 5. This leads to magnetisation in the direction of the stronger moment within the lattice. [7]

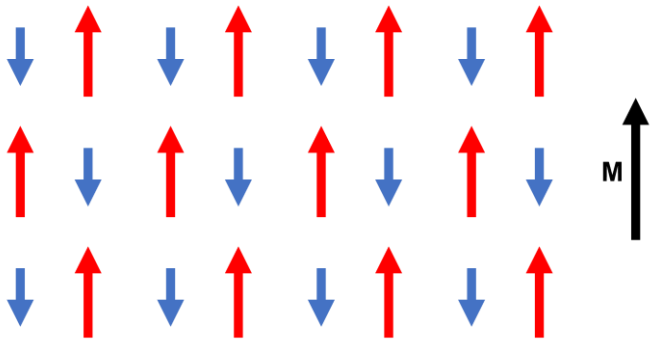


Figure 5- Typical lattice structure of spins in a ferrimagnetic material. Although similar ordering between spin up (red arrows) and spin down (blue arrows) elements occur, due to the increased magnitude of the spin up spins, there is a net magnetic moment in system.

2.1.4- Energy Contributions in Ferromagnetic Materials.

As alluded to earlier, there are many contributing energetic factors which determine the magnetisation state of a ferromagnetic material. These contributions depend upon the underlying crystal structure of the material, the magnitude of stray fields generated by locally ordered magnetisation, interactions with external fields, and quantum-mechanical interactions between neighbouring spins, amongst other factors[5]. These contributions will be described in turn in this section.

2.1.4.1- Magnetostatic Energy

Magnetostatic energy describes the potential energy that is stored in the stray field generated by a material with macroscopic net magnetisation. Here, the magnetic field forms a closed loop from the end of the locally ordered region to the beginning, as shown in figure 6. It can be seen in the figure that the stray field acts opposite to the magnetised region, acting to reverse the magnetisation, and hence is termed the demagnetisation field. The demagnetisation field, and hence the magnetostatic energy contributions, is proportional to the magnetisation of the material. The formation of domains (explained more extensively in section 2.1.5) aims to provide flux closure and minimise magnetostatic energy.[5]

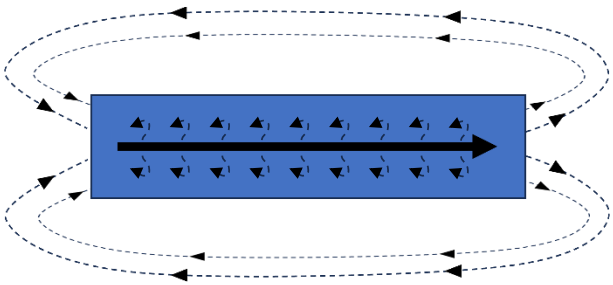


Figure 6- Demagnetisation field (black dotted lines) closing flux from the uniformly magnetised material (black arrow) and attempting to cause reversal.

The shape of the magnetic material also plays a key role in determining the magnetostatic energy of the material, termed shape anisotropy, which aims to minimise the formation demagnetising fields both internally and at the edge of the sample. This leads preferential alignment along spatial axes for samples that are not perfectly spherical, typically preferring axial directions with higher aspect ratios. For example, in the magnetic nanorings studied in this thesis, the large disparity between the thickness

of the nanostructures compared to the track widths ($\sim$10 nm compared to $\sim$400 nm) means that the magnetisation lies in plane, and the narrow circular structure means that the magnetisation in the plane tends to follow the circumference of the ring.

2.1.4.2- Magnetocrystalline Anisotropy Energy

Due to the underlying spatial arrangement of atoms within the lattice of a crystalline magnetic material, materials have a tendency for preferential magnetisation along certain crystallographic directions. The directions which face the least resistance to magnetisation are termed the 'easy' axes. The specific directions of the easy axes depend upon the lattice arrangement of the material itself, with different axes expected for body-centred cubic (BCC), face-centred cubic (FCC), and hexagonal close-packed (HCP) lattice arrangements. The consequence of this preferential alignment is that it requires more external field to be able to saturate the material along directions other than the easy axis, with the energy difference per unit volume of magnetised material aligned both along and away from the easy axis termed the magnetocrystalline anisotropy energy.[8]

For materials with cubic crystal structures, the crystalline anisotropy energy density of the crystal can be described via equation (2.6):

$$\varepsilon_C = K_1(cos^2\alpha\, cos^2\beta + cos^2\alpha\, cos^2\gamma + cos^2\beta\, cos^2\gamma) + K_2(cos^2\alpha\, cos^2\beta\, cos^2\gamma) \quad (2.6)$$

where K_1 and K_2 describe the 1st and 2nd order anisotropy constants, and α, β, and γ represent the angles between the magnetisation and the easy crystallographic axes running along the cubic edges of the crystal structures for BCC, or along the cube diagonals in FCC. Figure 7 shows the easy, intermediate, and hard magnetocrystalline axes for a BCC material.

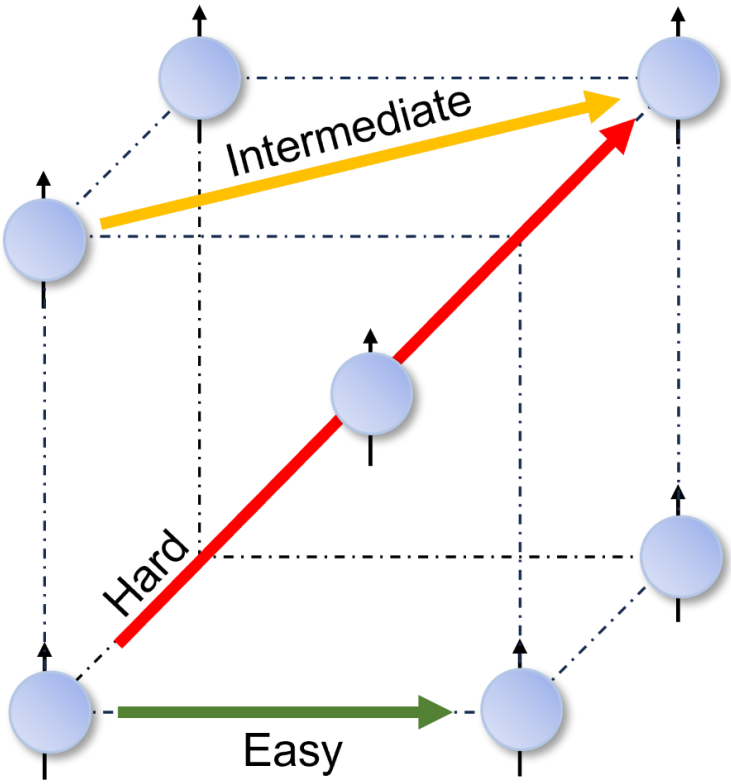


Figure 7- Magnetocrystalline anisotropy axes for a body centred cubic material. Atoms are shown by blue spheres, and their moments by black arrows. The easy axes (green arrow) lie along the edges of the cube structure, the intermediate axes (yellow arrow) between atoms on a face, and the hard axis (red arrow) between corners of the cube and the central atom.

For crystal structures with uniaxial symmetry (e.g., hexagonal close packed), there is only one easy axis and hence a simpler calculation of anisotropy energy density shown in equation 2.7:

$$\varepsilon_u = K_u(sin^2\alpha) \quad (2.7)$$

where K_u represents the uniaxial anisotropy constant.

2.1.4.3- Magnetostrictive Energy

When a material is magnetised, the magnetisation exerts a stress which strains the material via a process called magnetostriction. Whether this stress is compressive or tensile depends upon the material, and hence materials are deemed to have positive or negative magnetostriction depending upon whether the material elongates or constricts respectively. In multidomain magnetic materials, the change in magnetisation between the domains leads to stresses at different directions, shown in figure 8, and the elastic potential energy associated with these competing stresses is termed magnetostrictive energy.[9]

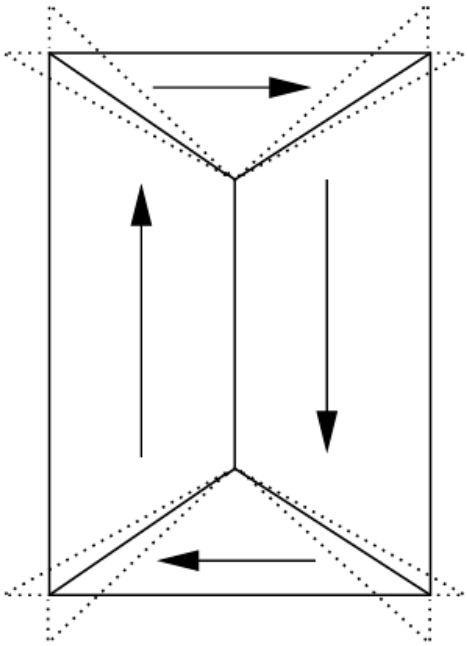


Figure 8- Domain structure of a ferromagnet with positive magnetostriction. The true domain wall position (solid black lines) is shown as a compromise between the attempts of adjacent domains attempting to elongate along their magnetisation direction, with the 'ideal' position of the elongated domains shown by the dotted lines. Taken from[9].

2.1.4.4- Zeeman Energy

The energetic state of a magnetic material depends not only upon the internal magnetisation of the material itself, but also the energy that is stored when a magnetic material interacts with an external field. This torque τ between magnetic moments in the material and an external field H is dependent upon the angle between the external field and magnetisation of the material θ, shown in equation (2.8), and the energy stored in the material, called the Zeeman energy, E_{Zeeman}, is calculated via equation (2.9):

$$\tau = \mu_0 Hm \sin\theta \quad (2.8)$$

$$E_{\text{Zeeman}} = -\mu_0 \int \boldsymbol{M} \cdot \boldsymbol{H}\, dV \quad (2.9)$$

where **M** is the magnetisation vector of the material, **H** is the applied field vector, V is the volume of the material, and μ_0 is the vacuum magnetic permeability.[5]

2.1.4.5- Exchange Energy

The exchange interaction is the quantum mechanical effect that is responsible for ferromagnetic or antiferromagnetic ordering of neighbouring spins in a material. It arises as a consequence of exchange symmetry between the wave functions electrons in the magnetic material, which forces either symmetric or antisymmetric spins when electrons are exchanged. Although other types of exchange exist, only Heisenberg exchange will be described here as it provides the simplest explanation for (anti)ferromagnetic alignment of spins in common magnetic materials. Here, direct exchange is considered between electrons which occupy overlapping orbitals, with the energy density contribution, $\varepsilon_{\mathrm{ex}}$ for a pair of spins, $\hat{S}_i$ and $\hat{S}_j$ given via equation (2.10):

$$\varepsilon_{\mathrm{ex}} = -2J_{\mathrm{ex}} \sum_{ij} \hat{S}_i \cdot \hat{S}_j \tag{2.10}$$

where J_{ex} denotes the exchange constant, which quantifies the strength of the exchange interaction. The sign of J_{ex} determines whether it is energetically favourable for either parallel or anti-parallel alignment of spins, with a positive J_{ex} favouring parallel (ferromagnetic) alignment, and a negative J_{ex} favouring antiparallel (antiferromagnetic) alignment of spins.[5]

In micromagnetics, the energetic contribution of the exchange interaction approximated over a continuum is given by evaluating the exchange energy from the divergence between the magnetisation direction vectors and an exchange constant A_{ex}, as per equation (2.11):

$$E_{ex} = \int_V A_{ex}[(\nabla \vec{m}_x)^2 + (\nabla \vec{m}_y)^2 + (\nabla \vec{m}_z)^2]dV \tag{2.11}$$

A method for predicting the sign of the exchange constant J_{ex}, was outlined by Slater in 1930[10] by calculating the ratio between interatomic distance of the material and the radius of the 3d orbital, shown in figure 9. This is useful for predicting Heisenberg exchange, which considers the Coulombic repulsion of a pair of electrons obeying the Pauli exclusion principle. As the ratio of interatomic distance to 3d radius decreases, the wavefunctions increasingly overlap, which favours antiparallel, antiferromagnetic alignment. As the degree of overlap is reduced, parallel alignment becomes favourable due to decreased repulsion between the electrons, and hence ferromagnetic ordering is observed. While this is useful for predicting change in ordering of 3d transition metals, it neglects more complex interactions that arise from different overlapping orbitals[11].

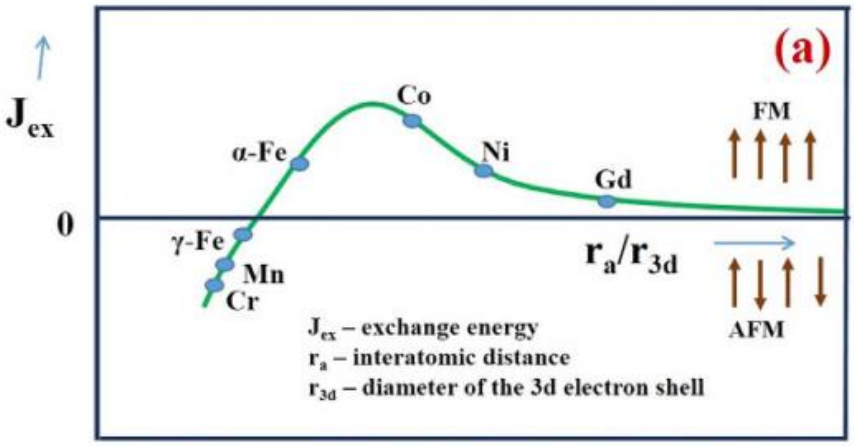


Figure 9- Predicted values of exchange constant J for different metals calculated from the ratio of atomic separation to 3d orbital diameter. Materials with values above zero on the curve suggest ferromagnetic alignment, while those below will exhibit antiferromagnetic alignment. Taken from[12].

2.1.5- Domains and Domain Walls

The magnetic microstates that magnetic materials tend to form are a consequence of the interplay between each of the competing energy terms listed in the previous section. One such example of the compromise between magnetostatic energy and exchange energy in a ferromagnetically ordered material arises in the formation of domains. The underlying domain structure of the nanomagnetic ring arrays studied in this thesis is critical to understanding both the system's response to external magnetic fields, as well as the anisotropic magnetoresistance response of the devices. This section will explore the formation of domains and domain walls in more detail.

2.1.5.1- Formation of Magnetic Domains

In order to reduce the potential energy stored in the demagnetisation field, magnetic materials tend to partition into sub-regions of locally uniform magnetisation which reduce magnetostatic energy via flux closure. These sub-regions are known as magnetic domains. At the interface between domains, moments locally rotate, forming domain walls. Here, the misalignment of local moments increases the exchange energy of the material, hence domain walls exist with finite width as a compromise between exchange, magnetostatic, and anisotropic energies.

While the competition between exchange and magnetostatic energies are the key driver behind domain formation, the size and shape of domains/domain walls are in part determined by the contributions of magnetocrystalline anisotropy energy and magnetostrictive energy. For example, in materials with a bcc lattice arrangement, the easy anisotropy axes lie perpendicular from one another, which leads to a tendency for domain walls to form along these perpendicular axes with domains separated by a 90° domain wall, with the domain walls providing flux closure known as closure domains[9]. The process of partitioning from a single domain state to a multidomain state to minimise total energy contributions is shown in figure 10.

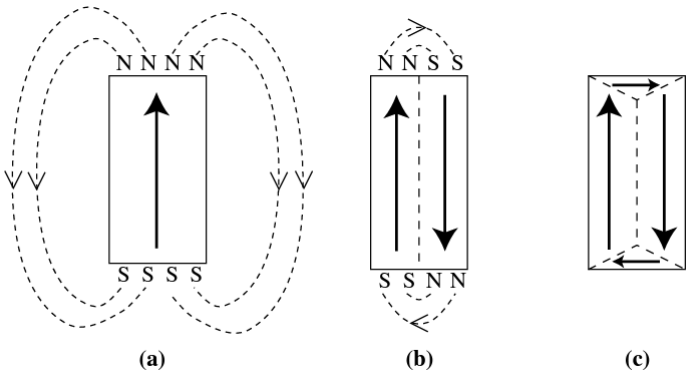


Figure 10- Process of reducing stray fields, and hence magnetostatic energy, of (a) a single ferromagnetic domain via flux closure as it is partitioned into (b) two domains, then (c) four domains. Taken from[9].

While this arrangement of domains provides a favourable compromise for each of magnetostatic, exchange, and anisotropic energy terms, the formation of the triangular domains creates a competition between the magnetostrictive contributions between each domain, with each domain either contracting or elongating along the magnetisation axis of the domain, depending upon whether the material has positive or negative magnetostriction. Since the magnetostrictive energy is proportional to the volume of closure domains, this creates a tendency for minimising the closure domains.[9]

2.1.5.2- Domain Walls

Similar to domains, the size, shape, and mode of rotation of domain walls are determined by competing energy contributions, namely between anisotropic and exchange energy terms in bulk materials.[9] Here, the magnetic moments in domain walls rotate along the plane of the domain wall, termed a Bloch wall, shown in figure 11. While exchange interactions favour small angular rotations between moments and large domain walls, the magnetocrystalline anisotropy favours alignment along the easy axes, and hence sharp transitions between the axes. This leads to larger domain walls in materials with higher exchange stiffness, and smaller domain walls in materials with strong magnetocrystalline anisotropy.

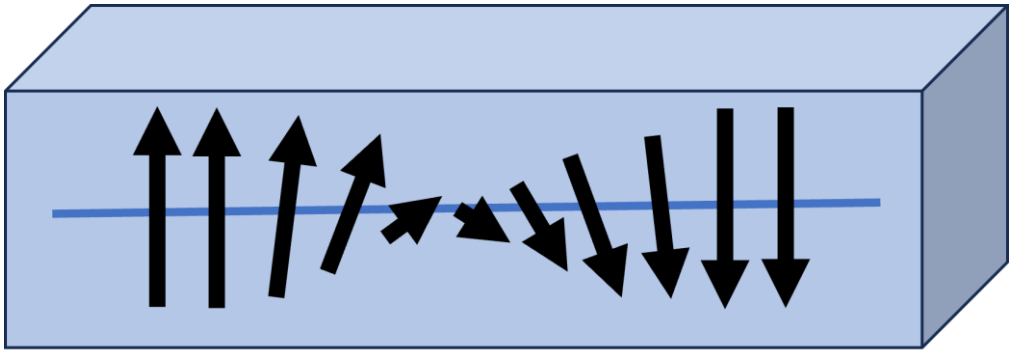


Figure 11- A Bloch wall in a bulk magnetic material, with the local magnetic moments (black arrows) rotating about the plane of the domain wall.

In thin films, the considerable shape anisotropy that confines the magnetisation in plane inhibits the rotation of moments along the domain wall axis, as this would lead to some moments in the domain wall pointing out of plane, which is magnetostatically unfavourable.[9] Instead, the moments rotate in the plane of the film magnetisation, shown in figure 12, and are called Néel walls.

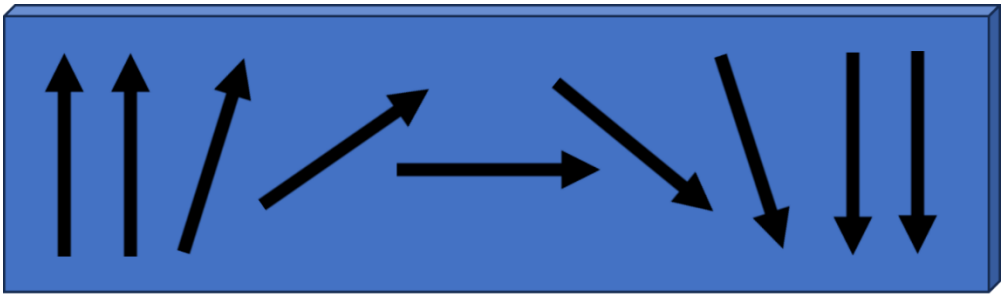


Figure 12- A Néel wall in a magnetic thin-film, showing rotation of magnetic moments (black arrows) in the plane of the magnetic film.

2.1.5.3- Interactions of Domains with External Fields

When a magnetic field is applied to a ferromagnetic material it is favourable for magnetic moments to rotate towards the field to reduce the system's Zeeman energy.[5] This creates a driving force for domains aligned with the field to grow, causing DWs to move in response to the field. This movement is initially reversible, with the original demagnetised domain structure reforming upon relaxation of the field. Figure 13 shows this process of domain wall growth, before hysteretic processes lead to DW annihilation as the material reaches saturation magnetisation.

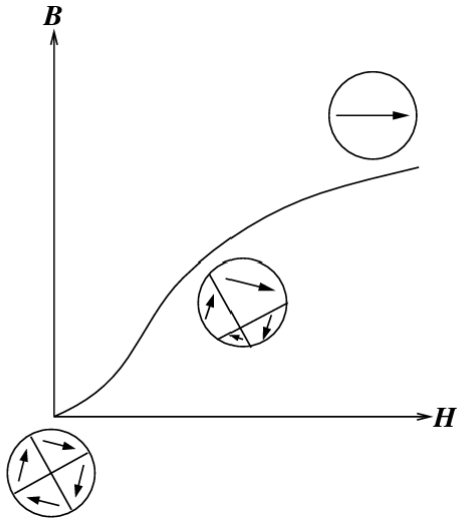


Figure 13- Schematic diagram of domain growth as a sample is magnetised towards saturation. Initially, domains along the direction of the external field grow whilst surrounding domains shrink, until saturation is reached, and the material is uniformly magnetised along the external field direction. Taken from[9].

As the domains continue to grow, DWs encounter energy barriers from material defects or geometrical changes that restrict the propagation of DWs, causing them to pin if the energy barriers are sufficiently large. This pinning behaviour is easily observed in bulk materials in the form of Barkhausen noise. The pinning and depinning causes sudden, small changes in magnetisation (Figure 14 inset) as the DWs hop between pinning sites. In nanowire devices such as the rings studied here, the reduction of the system to a very small number of domains (almost always either 1 or 2 domains per ring) means this pinning/depinning behaviour leads to DW movement that resembles ballistic jumps as the DW hops between pinning sites.

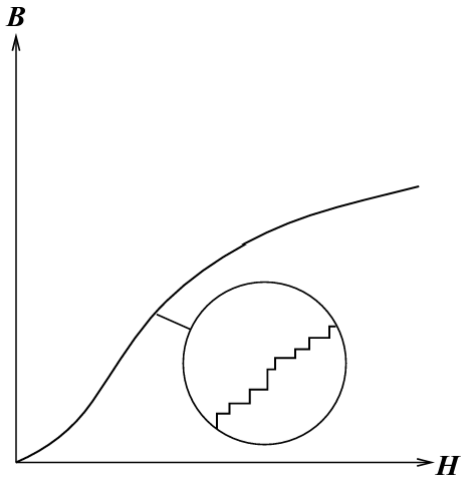


Figure 14- Magnetisation of a ferromagnetic material B versus applied field magnetic field H. Inset shows local magnification of magnetisation curve, showing erratic increase in magnetisation with respect to field due to the ballistic movement of DWs and subsequent domain growth as DWs overcome energy barriers with increasing applied field. Taken from[9].

2.1.6- Landau-Lifschitz-Gilbert Description of Magnetisation Dynamics

The most commonly used framework employed to describe the dynamic response of magnetic moments was proposed by Landau and Lifschitz[13], which describes the precessional motion of magnetisation about an external magnetic field. This description involves a semi-classical approach with phenomenological quantum-mechanical considerations and serves as the foundation for micromagnetic approaches for modelling magnetic materials. The Landau-Lifschitz equation models the decay of this precessional motion over time when damping is considered, describing how the moment dynamically tends towards an equilibrium position when aligned with the externally applied field, shown in equation (2.12):

$$\frac{d\boldsymbol{M}}{dt} = \gamma(\boldsymbol{M} \times \boldsymbol{H}_{\text{eff}}) + \frac{\lambda}{M_s^2}\boldsymbol{M} \times (\boldsymbol{M} \times \boldsymbol{H}_{\text{eff}}) \tag{2.12}$$

where **M** and **H** represent the magnetic moment and external field vectors respectively, γ represents the gyromagnetic ratio of the material, M_s the saturation magnetisation of the material, and λ represents the damping frequency. This formulation was later adapted by Gilbert by modifying the damping term to include a dimensionless damping parameter α, with the resulting equation commonly known as the Landau-Lifschitz-Gilbert equation[14] (2.13):

$$\frac{d\boldsymbol{M}}{dt} = \gamma(\boldsymbol{M} \times \boldsymbol{H}_{\text{eff}}) + \frac{\alpha}{M_s}(\boldsymbol{M} \times \frac{d\boldsymbol{M}}{dt}) \tag{2.13}$$

The damping terms in both of these equations govern how fast the system tends towards equilibrium in response to a change in magnetic field, with greater damping leading to shorter relaxation times. When considered in tandem with the effective field from the different energetic contributions, this equation is the basis of micromagnetic models, discussed in detail in chapter 4 section 6.

2.1.7- Ferromagnetic Resonance

Ferromagnetic resonance (FMR) is a phenomenon that arises from the precessional motion of magnetic moments about an external field when irradiated with microwave radiation. Here, the amount of energy dissipated by the precessing magnetic moments depends strongly upon the frequency of the applied field, with a sharp increase in dissipation when resonance occurs between the microwaves and the precessing magnetic moments. The resonant frequency is typically on the

order of 1 GHz – 40 GHz[15], with the specific frequency dependent upon the material's gyromagnetic ratio, anisotropy, as well as the size, shape, and orientation of the domains in the material. This allows FMR to be used as a tool for quantifying the magnetic state of a material, as specific magnetic structures can be tied to a given resonant frequency, with dissipation proportional to structure population, and has recently been employed for neuromorphic computing applications[16].

2.1.8- Magnetoresistance Effects

The influence of magnetisation on the electrical resistance of a material provides a means for simple measurements which depend upon the magnetic state of a system. These resistance changes arise from different physical effects, and the magnitude of resistance changes spans orders of magnitude from around 5% for anisotropic magnetoresistance (AMR) in permalloy[17], to upwards of 1000% for tunnel magnetoresistance (TMR) in CoFeB/MgO multilayers at cryogenic temperatures[18]. A phenomenological description of each of these effects will be described here.

2.1.8.1- Anisotropic Magnetoresistance

AMR is the weakest of the magnetoresistance effects in ferromagnetic materials and was first observed by Lord Kelvin in the 1850s[19], though the determination of its physical origin was described much later by Kondo in the 1960s[20]. It arises from spin-orbit coupling effects, where conduction electrons are scattered by the unquenched orbital angular momentum of 3d electrons in the ferromagnetic material, with maximum scattering (and hence highest resistance) occurring when the magnetisation of the material is parallel to the direction of electric current[21].

2.1.8.2- Giant Magnetoresistance

Giant magnetoresistance (GMR) occurs between layers of magnetic materials separated by a thin conductive spacer layer and can be understood via interlayer exchange coupling and spin-dependent transport[21]. In these multilayers, the thickness of the spacer layer determines the interlayer exchange coupling, with coupling alternating between ferromagnetic and antiferromagnetic alignment as spacer thickness is varied. Spin-dependent transport leads to the variation in resistance between antiparallel and parallel alignment of the ferromagnetic layers. Conduction electrons are spin-polarised by the magnetisation of the first layer. In the second layer, conduction electrons are more strongly scattered when their magnetic moments are anti-parallel to the magnetic moments of the magnetised material. This leads to low resistivity for parallel alignment, and high resistivity for anti-parallel alignment of the two layers, since the electrons are able to maintain their spin polarisation across finite length scales when transitioning between the layers, given by the spin diffusion length of the material.

In devices, the GMR effect can be utilised to measure the relative magnetisation of the material along a given axis via a spin valve.[22] Here, an antiferromagnet is used to pin one of the ferromagnetic layers in a given direction by increasing its coercivity, termed the 'pinned' layer. The angle between the magnetisation of the pinned layer and the measured 'free' (magnetically soft) layer leads to variation in resistivity across the layers via GMR.

2.1.8.3- Tunnel Magnetoresistance

TMR multilayers resemble those of GMR, but instead of a metallic spacer an insulator is used instead, forming a magnetic tunnel junction (MTJ). When a bias voltage is applied across the multilayer, electrons preferentially tunnel towards the positive electrode, with probability dependent upon the alignment between the ferromagnetic layers. This is again dependent upon the density of states at the Fermi level, with more electrons able to tunnel for parallel alignment as with GMR.[18]

2.2- Machine Learning and Artificial Neural Networks

The following section will take a didactic approach to introduce fundamental concepts in machine learning, starting with the biological inspiration behind artificial neural networks, the concept of linear separability of data, and methods for training networks under the supervised learning framework. The concept of recurrence will be introduced, and then the archetypal reservoir computing architecture of echo state networks will be defined in detail.

2.2.1 - What is an Artificial Neural Network?

An artificial neural network is an abstract model designed to mimic some of the form and function of the biological brain. It features an interconnected network of nodes, joined by weighted connections, resembling the neurons and synapses of brains respectively, and as such, the biological nomenclature is often used equivalently for describing artificial neural networks. Figure 15 displays schematic diagrams of biological and artificial neural networks. Nodes in the network provide nonlinearity between inputs and the activation of the neuron via their activation function. The weighted connections scale interactions between the nodes, and the weights are adapted according to the learning rule of the network.

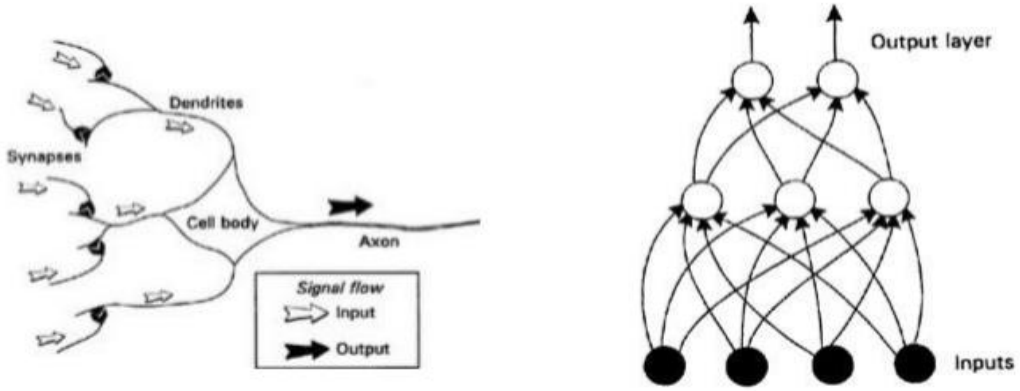


Figure 15- Left: Simplified diagram showing connections between neurons (black circles) in the brain. Arrows denote flow of electric current along synapses (lines). Right: Schematic of a typical artificial neural network. Black circles represent input nodes, white circles hidden nodes, and arrows show weighted connections, showing the direction that activation values are passed onwards. Taken from[20].

Neural networks typically find use in processing or modelling complex data and offer advantages over typical algorithmic approaches to such problems where the input tends to be inconsistent or noisy, such as for image or speech recognition tasks, or if the system to be modelled is highly nonlinear or seemingly chaotic, such as individual medical diagnostics or financial market prediction. Neural networks are trained to find patterns and can often be used to find valuable information hidden in large, generally poorly understood data[23].

2.2.2 - Supervised Learning in Neural Networks

The most common form of training a neural network, and the method employed in this thesis, comes in the form of supervised learning. Supervised learning utilises a teacher signal which provides a target for the output of the neural network for a given input. For classification tasks, this is commonly achieved via 'one-hot' encoding, where each of the N target classes are represented by N output nodes, where the target output is 1 for the node corresponding to the correct class, and 0 elsewhere. To perform classification, a 'winner-takes-all' approach is employed, where the predicted class is

determined by the output node with the highest activation. For regression tasks, the target signal is given by the desired output, with the network tasked to match output activation with the target signal.

The network is then trained to minimise the error between the target signal and the network output, with error defined via a loss function. For example, the means squared error (MSE) function, equation (2.14):

$$E = \frac{1}{2N}\sum_{i}^{N}(y_i - a_i)^2 \quad (2.14)$$

where y_i represents the target output activation on node i, and a_i represents the model's activation on node i, for each of N output nodes. There are many different loss functions available, with each having varying suitability and efficiency for training a network depending on the task. For example, the binary cross entropy loss function is especially effective for classification tasks, as it strongly penalises incorrect classification[24], or the Connectionist Temporal Classification function, which is designed for use in continuous time series analysis where outputs do not necessarily align with inputs, such as online speech processing[25].

The most common method for minimising an error is gradient descent[23]. Gradient descent uses an iterative process in order to find minima in complex functions by making small changes to the weighted connections between nodes. To minimise the MSE loss, E, we make a series of small changes to weights in the network in the direction that reduces E. To ensure this change is in the right direction for every weight, we need to find the gradient of the error with respect to each parameter.

2.2.3 - Backpropagation

Following on from the error calculated by a loss function, the backpropagation algorithm provides a method for minimising this error. A simple summary of key parts of the backpropagation algorithm for a typical network serves as an ideal way to understand both how and why the algorithm works. The original derivation from Rumelhart *et al.*'s 1986 paper can be found at[26]. Building upon the notation established earlier for the MSE function, we will train a network of nodes, x_i^n, where n and i represent layer and node index, respectively. Nodes are joined by weights, where weight w_{ij}^k joins node j in layer k-1 to node i in layer k. For simplicity, nodes in this network are linear, and the network is feedforward. A schematic of the network labelled with this notation is shown in figure 16.

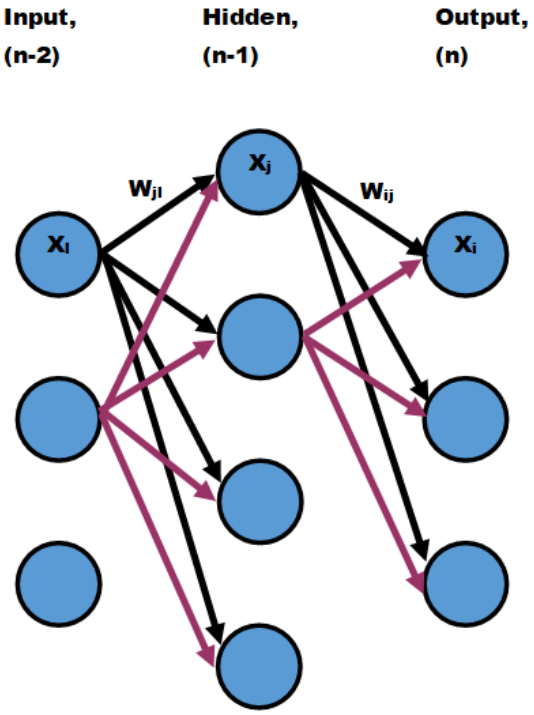


Figure 16- Simple network showing nodes (circles) and weights (arrows) in different network layers, illustrating the direction activation is passed through connections. Connectivity is only shown in the first 2 nodes of each layer for simplicity.

To find the direction with which to make small changes, we need to calculate how the error changes with respect to a change in a given weight, then make a small change, Δw, to the old weight (equation 2.15):

$$(2.15) \quad \Delta w_{ij}^{k} = -\alpha * \frac{\partial E}{\partial w_{ij}^{n}}$$

where α is the learning rate of the algorithm. However, the partial derivative, $\frac{\partial E}{\partial w_{ij}^{n}}$, is difficult to calculate a priori. Instead, we apply the chain rule to pass derivatives back through the network as a series of more simple calculations.

Starting at the output layer, we take the derivative of the error function with respect to the activation on nodes in the output layer for each input datapoint μ. Here we combine the difference between target and activation for a given node, denoted as e_i. In online training, the error is calculated for each input pattern μ, as is presented in (16). Alternatively, summing over either all of the patterns or subgroups of all patterns allows for batch and minibatch training respectively, shown in (17). On-line training is often better for avoiding local minima but offers slower convergence than batch methods[23].

$$(2.16) \quad \frac{\partial E}{\partial e_i(\mu)} = e_i(\mu) = y_i(\mu) - a_i(\mu)$$

$$(2.17) \quad \frac{\partial E}{\partial e_i} = \sum_{\mu} e_i(\mu) = \sum_{\mu} (y_i(\mu) - a_i(\mu))$$

Applying the chain rule, we can establish the dependence of output node errors on connections to those nodes in the previous layer, as shown in (2.18) and (2.19). Here, the updates for on-line training are used.

$$\frac{\partial E}{\partial x_i^n(\mu)} = \frac{\partial E}{\partial e_j(\mu)} * \frac{\partial e_j(\mu)}{\partial x_i^n(\mu)} \tag{2.18}$$

$$\frac{\partial E}{\partial w_{ij}^n} = \frac{\partial E}{\partial x_i^n(\mu)} * \frac{\partial x_i^n(\mu)}{\partial w_{ij}^n} \tag{2.19}$$

For the connections in the hidden layers, there exist pathways to all nodes in the output layer, and hence changing weights in the hidden layer effects all of the output nodes. To account for this, errors on hidden nodes are summed across nodes ahead of them in the network, shown in equations (2.20-2.22):

$$\frac{\partial E}{\partial w_{jl}^{n-1}} = \frac{\partial E}{\partial x_j^{n-1}} * \frac{\partial x_j^{n-1}}{\partial w_{jl}^{n-1}} \tag{2.20}$$

$$\frac{\partial E}{\partial x_j^{n-1}} = -\sum_i e_i * \frac{\partial x_i^n}{\partial x_j^{n-1}} \tag{2.21}$$

$$x_i^n = \sum_j x_j^{n-1} * w_{ij}^n \tag{2.22}$$

As each calculation of local gradient at a weight relies on the gradient of the connections ahead of it, we summarise these local gradients using the 'delta rule'. Equations (2.23) and (2.24) show how local error dependence, δ_j^{k-1}, can be calculated simply from the local gradient of the nodes ahead of it, and this tells how to make small changes to minimise error (for batch training, also sum over μ in both equations). The process propagates this error, δ, backwards through the layers, giving the name to the algorithm: Backpropagation.

$$\delta_j^{k-1} = \sum_i \delta_j^k x_l^{k-2} \tag{2.23}$$

$$\Delta w_{jl}^{k-1} = -\alpha * \delta_j^{k-1} x_l^{k-2} \tag{2.24}$$

This process is iterative, where training will continue to make small changes until a given number of iterations are completed or some convergence criterion for the error is met.

2.2.4 - Recurrence

In recurrent neural networks (RNNs), feedback is added by including weights which connect backwards through the network. Here, future activations depend upon previous states, contrary to the static nature of feedforward neural networks. The recurrent connections can be seen as loops in the network, shown schematically in figure 17a. RNNs find use in areas such as control of systems, speech recognition, or time series prediction, where the dynamic nature of the network synergises with the dynamic dependencies of the task[27]. Lukoševičius[28] notes that RNNs are able to sustain activity through recurrent pathways even in absence of input; mathematically this renders RNNs as dynamical systems, as opposed to feedforward networks which behave as functions.

One of the biggest problems with RNNs, and the main factor in their limited use, is the difficulty associated with their training[29]. The time dependence adds an extra level of complexity to finding the gradient of the error with respect to a change in weights, increasing the computational expense. The

most common training method for RNNs is backpropagation through time (BPTT), shown schematically in figure 17[27]. The BPTT algorithm is exact mathematically, but inefficient and lacks biological plausibility, hence the emergence of bioinspired alternatives available to improve upon the algorithm[30]. Conceptually, BPTT works by unravelling the dependencies through time, generating an equivalent feedforward network by accounting for recurrent connections with an additional hidden layer representing the activities in the previous time step.

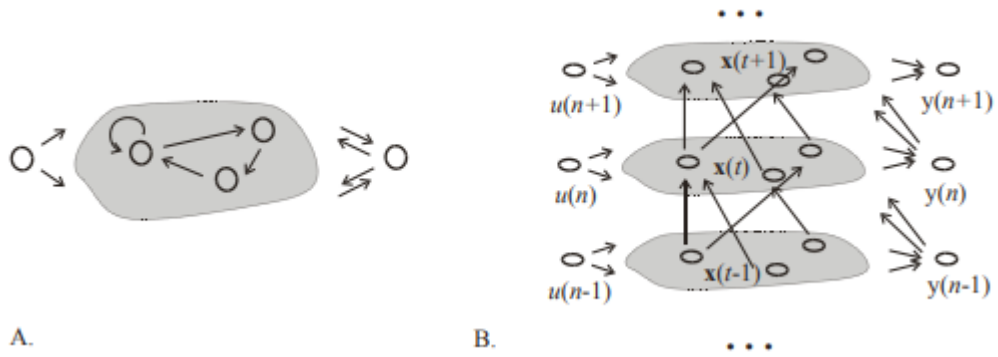


Figure 17- A. Schematic of a recurrent neural network. Arrows denote flow of activity between nodes. B. 'Unfolding' of a recurrent neural network in time via the BPTT process, showing reconnection of recurrent connections to nodes in future or past network states. U(n) represents input data point n, x(t) the reservoir state at time t, and y(n) represents the target for data point n[27].

From the schematic, it is clear to see the increased computational expense that adding recurrent connections adds to the training of a neural network, as network size is vastly increased. In order to relieve this computational expense, some alternative methods have been proposed in order to gain the computational advantage that adding a time dependence to the modelled system brings, whilst mitigating the computational cost of more difficult training methods that time dependence often brings. This is the motivation behind reservoir computing which will be discussed in depth in section 2.2.6.

2.2.5 - Linear Separability

Classification tasks rely upon the fact that a neural network is able to provide transformations on input data so that a linear classifier can effectively separate each class. For a given dataset with N features and a binary label, if the dataset is plotted in N dimensional space, the data is said to be linearly separable if a hyperplane can be drawn through the dataset, separating the two classes by their features. Figure 18 shows a dataset with 3 features, plotted in 3d space. The data is not linearly separable due to the misclassification of two points.

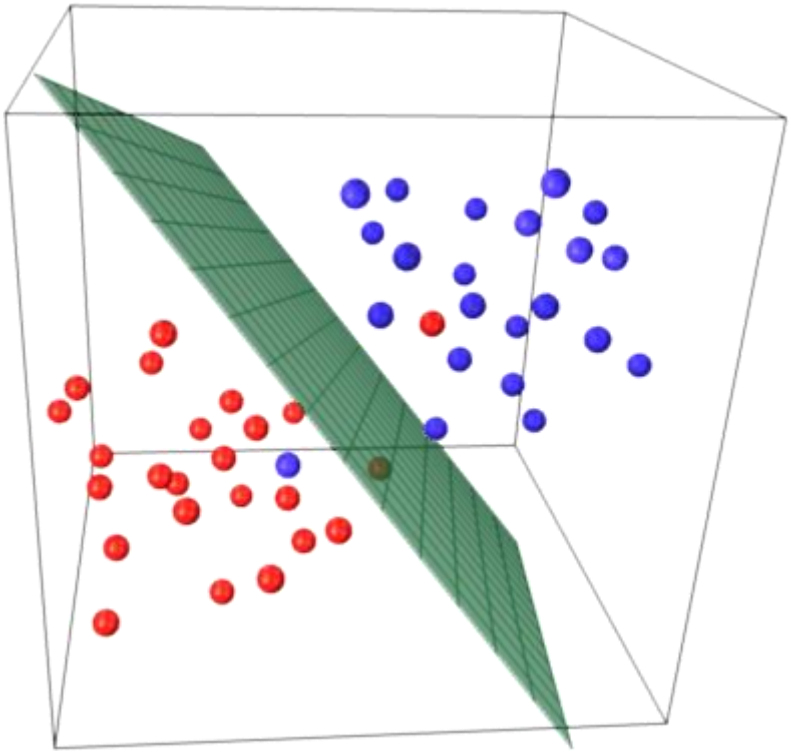


Figure 5- A 3-dimensional dataset showing linear separability between the two classes by a hyperplane. The class of the datapoints are represented by red and blue colour of circles [9].

While this this dataset can be classified rather effectively with a simple linear classifier, other datasets may prove more problematic to separate with a linear hyperplane. For example, consider the 2D data in figure 19. The classes shown in the toy dataset have definite class dependent structure, though cannot be separated linearly by their original features shown in (a). In order to remediate this, a nonlinear transformation of the features is required. This projects the data into different space where it may be more easily separated. Here, with a simple nonlinear function of $\hat{f}_i = f^2{}_i$ on each feature i, the dataset is transformed to the representation shown in (b), which can now be separated by a linear divider.

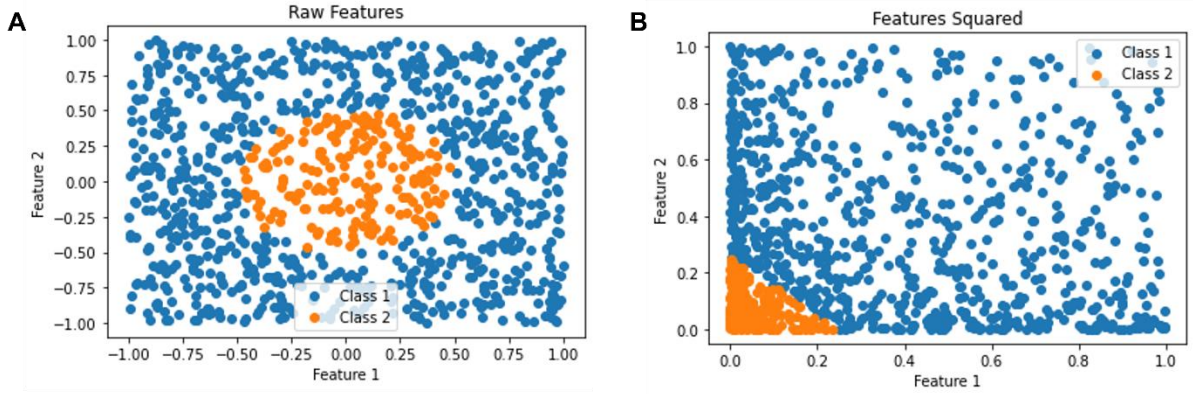


Figure 6- (A) Scatterplot of raw features for two classes, showing linear inseparability. (B) Nonlinear transform where each feature's transformed value is the square of the raw feature. This projection of the original data is now linearly separable.

Another geometric property of the transform provided by a neural network that improves classification accuracy is the increased dimensionality inherent to the network's representation. While the benefits of high-dimensional representations for data analysis remains an active area of research in mathematics[31–33], and mathematical analyses of the separation capacity of classifiers versus the

number of dimensions exists in the form of Vapnik-Chervonenkis (VC) dimension[34,35], the basic intuition behind the power of dimensionality can be described phenomenologically: A decision boundary (hyperplane) in N-dimensional space has itself dimensions (N-1), hence for larger N the decision boundary has more degrees of freedom with which to separate the data. However, with increasingly complex models, the tendency to overfit to noise within the training data becomes more likely.

2.2.6- Introduction to Reservoir Computing

Reservoir computing (RC) is a machine learning paradigm which exploits both the aforementioned properties of nonlinearity and high dimensionality in order to provide enhanced representations of input data to perform machine learning tasks. Unlike support vector machines, the complex transformation provided by the reservoir is time dependent, acting as a temporal kernel rather than a static one[29]. The simplest way for describing RC for both *in-silico* and *in-materia* applications is by considering the three distinct layers that constitute an RC platform: An input layer or mask, a reservoir layer, and an output or readout layer, shown schematically in figure 20.

The input layer provides weighted connections of input data to nodes within the reservoir. These weights are fixed and typically randomly sampled from various distributions, dependent upon the type of reservoir used and the nature of the data. In architectures that use multiplexing techniques to generate network complexity, this input layer is commonly referred to as a 'mask' (details on popular multiplexing techniques are outlined in the Literature Review, section 3.4).

The reservoir layer is where all of the time-dependent nonlinear processes in the reservoir occur and is the key to providing computationally useful representations of input data for the readout to exploit. In algorithmic implementations, the reservoir consists of a network of nodes with sparse, random connections. This maintains activity in the reservoir over time via the closed loops in connectivity between the nodes, as well as the time-dependent properties of the nodes themselves, with their activation functions (typically sigmoidal) providing nonlinearity. Since a recurrent network is itself a dynamical system, *in-materia* applications of RC substitute this reservoir layer with a physical dynamical system, where nonlinear physical processes occur, and the hysteresis of the system acts as a pathway to memory.

The readout layer is the only part of the network that is trained, and typically consists of a single linear layer. Since changes in the readout weights have no effect on the response of the reservoir layer, there is no need to unravel dependencies through time as is the case with recurrent neural networks. This becomes equivalent to the backpropagation decorrelation algorithm for recurrent neural networks, where only the final feedforward connections in the network are considered. Additionally, in the case where there is only a single linear output layer, a closed-form ordinary least squares solution exists between weighted outputs and target signals, described section 2.2.8.

The concept of reservoir computing was formulated to try and harness some of the computational power shown by recurrent neural networks for processing time dependent data, while mitigating the high complexity of training these recurrent networks. The method for achieving this reduction in training cost is the defining characteristic that connects different types of reservoir architectures; the decoupling of time-dependent dynamic processes which occur in the 'reservoir' from a time invariant 'readout' which is trained[28]

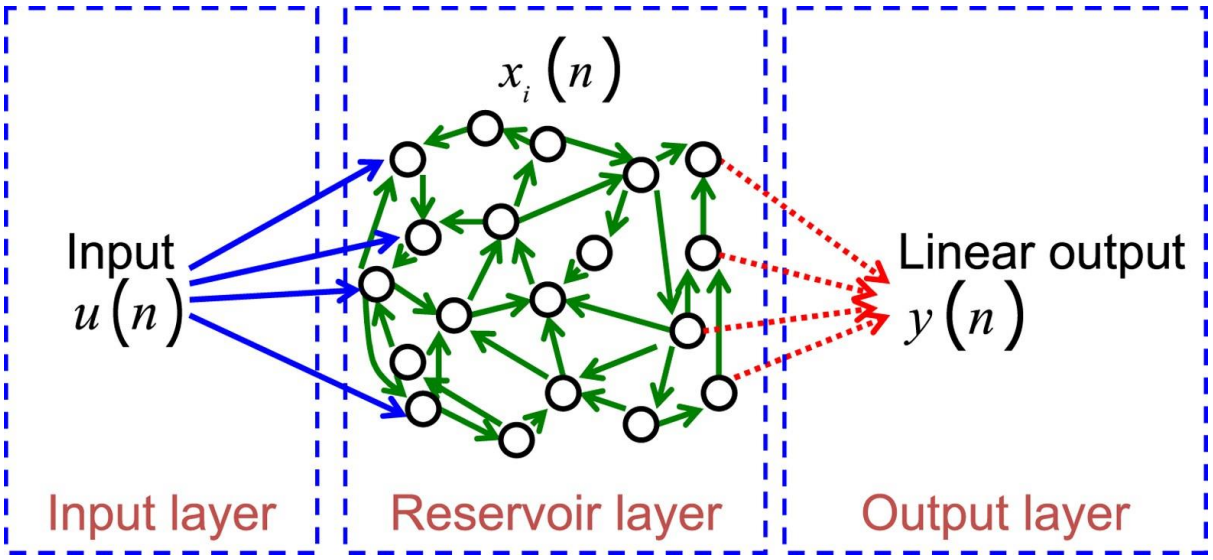


Figure 7- Schematic of reservoir architecture. Data is inputted via blue connections from input layer into nodes in the reservoir layer. Green connections show how these activations are passed between nodes within the reservoir. Dotted red lines show connections between output nodes in the reservoir system, on which activity will be periodically measured to produce reservoir state as the system evolves[36].

2.2.7- Echo State Networks

The first RC frameworks arose independently from one another in the early 2000s, consisting of echo state networks (ESNs), and liquid state machines (LSMs), with their inception coming from the fields of machine learning and computational neuroscience respectively, though drawing from the same underlying principles of fixed random connectivity and simple linear readout. The leaky-integrator neurons that are typical to ESNs bear far more dynamical similarity to nanomagnetic systems than the biologically inspired spiking neurons of LSMs, and as such LSMs will not be explored in depth here.

In ESNs, each of the three layers introduced in section 2.2.7 has its own associated weights, denoted here as $\mathbf{W}^{in}$, $\mathbf{W}^{res}$, and $\mathbf{W}^{out}$ for the input, reservoir, and output weights respectively. $\mathbf{W}^{in}$ is a fixed, dense random matrix that connects all input dimensions to all reservoir nodes. $\mathbf{W}^{res}$ is a fixed, sparse random matrix that defines the internal connectivity of the ESN, resembling an Erdos-Renyi graph structure where each node has a fixed probability for a non-zero weighted connection to another node in the network defined by some sparsity parameter. $\mathbf{W}^{out}$ is a dense matrix which connects each reservoir node to each output node, and the weights are trained in order to minimise the error between network output and some teacher signal.

Echo state networks exhibit what is called the 'echo state property', which can be phenomenologically described as a diminishing dependence of reservoir state upon any initial conditions. This means that any information provided to the system asymptotically fades from the system's state as new information is provided. Thus, ESNs are said to possess 'fading memory'. In order to ensure this property occurs, and that the reservoir state avoids chaotic regimes, the magnitude of weights in $\mathbf{W}^{res}$ must be normalised with respect to its spectral radius, defined by the largest eigenvalue of the weight matrix. Empirically speaking, the echo state property is maintained for any input providing the spectral radius of $\mathbf{W}^{res} < 1$[29].

The update rule for the activities on each node in a reservoir consisting of leaky integrator neurons with a tanh activation function are given by the following iterative equation (2.25):

$$\boldsymbol{x}(t+1) = (1-\alpha)\boldsymbol{x}(t) + \alpha\, tanh(\gamma \boldsymbol{W}^{\mathrm{in}}\boldsymbol{s}(t+1) + \rho \boldsymbol{W}^{\mathrm{res}}\boldsymbol{x}(t)) \quad (2.25)$$

where $\boldsymbol{x}(t)$ represents the reservoir state at time t, $\boldsymbol{s}(t)$ represents the input signal at time t, for a given set of network hyperparameters representing the leak rate of the leaky integrator neurons, input scaling, and spectral radius, represented by α, γ, and ρ respectively.

The update rule and the role each hyperparameter has on network dynamics can be understood by breaking the equation into three parts. Firstly, the $(1-\alpha)\boldsymbol{x}(t)$ term describes how much activity from the previous timestep is retained in each of the neurons and is controlled by the leak rate of the neurons. Higher leak rate means less activity is retained from timestep-to-timestep, and generally leads to shorter-term memory. The $\gamma \boldsymbol{W}^{\text{in}}\boldsymbol{s}(t+1)$ term describes how the strongly activity depends upon the input signal alone and is controlled by the input scaling parameter, with large γ leading to large dependence upon the current input. Finally, the $\rho \boldsymbol{W}^{\text{res}}\boldsymbol{x}(t)$ term dictates how much the activations depend upon the ESNs internal connectivity, with larger ρ leading to a greater dependence. The performance of ESNs in a given task is strongly dependent upon the ability to effectively match the dynamic properties of the network via its hyperparameters to the demands of the task.

2.2.8 – Training Reservoir Computers

For a linear readout, a simple relationship between target outputs and feature vectors extracted from reservoir state, x_n, and a target vector, y_n, can be achieved through regression, outlined in[28] and described here.

By stacking the reservoir state vectors and target vectors for each sequence of a training data set into matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ respectively, a closed-form ordinary-least-squares solution exists for finding the weights. Supposing that a weight matrix, $\boldsymbol{W}^{out}$, exists which connects reservoir state matrix $\boldsymbol{X}$ to target matrix $\boldsymbol{Y}$, this connection can be summarised by equation 2.26:

$$\boldsymbol{W}^{out}\boldsymbol{X} = \boldsymbol{Y} \tag{2.26}$$

The weight matrix can be calculated directly by post-multiplying either side of the equation by the inverse of $\boldsymbol{X}$. To avoid issues with singular matrices, pseudoinverses such as the Moore-Penrose pseudoinverse (MPP) are often used instead (equation 27).

$$\boldsymbol{W}^{out} = \boldsymbol{Y}\,\boldsymbol{X}^{+} \tag{2.27}$$

where $\boldsymbol{X}^{+}$ represents the MPP of $\boldsymbol{X}$.

To reduce the size of the matrix which requires inverting, and hence the computational demands, the normal equations solution using the transpose of the matrices is often taken instead. This also opens up the availability to include a regularisation parameter which helps reduce any problems with multicollinearity of solutions[28] (equation 2.28):

$$\boldsymbol{W}^{out} = (\boldsymbol{Y}\boldsymbol{X}^{T}) * (\boldsymbol{X}\boldsymbol{X}^{T} + \lambda \boldsymbol{I})^{+} \tag{2.28}$$

where λ represents the ridge or Tikhonov regularisation parameter, and $\boldsymbol{I}$ the identity matrix, which regularises along the diagonal 'ridge' of the weight matrix, serving as the namesake of the algorithm.

This batch training method forms the most popular method of training RC networks, thanks to its simplicity and ability to form a fully trained weight matrix in a single step. However, there are many other training methods available, depending on the requirements of the task, with any standard machine learning techniques for feedforward networks applicable to the readout. For tasks requiring online adaptation of parameters, such as adaptive channel equalisation, more familiar machine learning algorithms have been used, such as stochastic gradient descent or recursive least squares[37].

Additionally, taking reservoir transformations as an analogy to a temporal kernel, techniques suitable for support vector machines have been applied to reservoir training[38].

References:

1. Atomic origins of magnetism. in *Magnetic Materials: Fundamentals and Applications* (ed. Spaldin, N. A.) 22–37 (Cambridge University Press, 2010). doi:10.1017/CBO9780511781599.003.
2. Kutzelnigg, W. & Morgan, J. D. Hund's rules. *Z. Für Phys. At. Mol. Clust.* **36**, 197–214 (1996).
3. Diamagnetism. in *Magnetic Materials: Fundamentals and Applications* (ed. Spaldin, N. A.) 38–47 (Cambridge University Press, 2010). doi:10.1017/CBO9780511781599.004.
4. Paramagnetism. in *Magnetic Materials: Fundamentals and Applications* (ed. Spaldin, N. A.) 48–64 (Cambridge University Press, 2010). doi:10.1017/CBO9780511781599.005.
5. Interactions in ferromagnetic materials. in *Magnetic Materials: Fundamentals and Applications* (ed. Spaldin, N. A.) 65–78 (Cambridge University Press, 2010). doi:10.1017/CBO9780511781599.006.
6. Antiferromagnetism. in *Magnetic Materials: Fundamentals and Applications* (ed. Spaldin, N. A.) 96–112 (Cambridge University Press, 2010). doi:10.1017/CBO9780511781599.008.
7. Ferrimagnetism. in *Magnetic Materials: Fundamentals and Applications* (ed. Spaldin, N. A.) 113–129 (Cambridge University Press, 2010). doi:10.1017/CBO9780511781599.009.
8. Anisotropy. in *Magnetic Materials: Fundamentals and Applications* (ed. Spaldin, N. A.) 135–144 (Cambridge University Press, 2010). doi:10.1017/CBO9780511781599.011.
9. Ferromagnetic domains. in *Magnetic Materials: Fundamentals and Applications* (ed. Spaldin, N. A.) 79–95 (Cambridge University Press, 2010). doi:10.1017/CBO9780511781599.007.
10. Slater, J. C. Cohesion in Monovalent Metals. *Phys. Rev.* **35**, 509–529 (1930).
11. Cardias, R. *et al.* The Bethe-Slater curve revisited; new insights from electronic structure theory. *Sci. Rep.* **7**, 4058 (2017).
12. Chaudhary, V. & Ramanujan, R. V. Magnetocaloric Properties of Fe-Ni-Cr Nanoparticles for Active Cooling. *Sci. Rep.* **6**, 35156 (2016).

13. Landau, L. & Lifshitz, E. 3 - On the theory of the dispersion of magnetic permeability in ferromagnetic bodies Reprinted from Physikalische Zeitschrift der Sowjetunion 8, Part 2, 153, 1935. in *Perspectives in Theoretical Physics* (ed. Pitaevski, L. P.) 51–65 (Pergamon, 1992). doi:10.1016/B978-0-08-036364-6.50008-9.
14. Lakshmanan, M. The fascinating world of the Landau–Lifshitz–Gilbert equation: an overview. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **369**, 1280–1300 (2011).
15. Kittel, C. On the Theory of Ferromagnetic Resonance Absorption. *Phys. Rev.* **73**, 155–161 (1948).
16. Gartside, J. C. *et al.* Reconfigurable training and reservoir computing in an artificial spin-vortex ice via spin-wave fingerprinting. *Nat. Nanotechnol.* **17**, 460–469 (2022).
17. Goncharov, A. V. *et al.* Anisotropy of Magnetization Reversal and Magnetoresistance in Square Arrays of Permalloy Nano-Rings. *IEEE Trans. Magn.* **42**, 2948–2950 (2006).
18. Jiang, L., Naganuma, H., Oogane, M. & Ando, Y. Large Tunnel Magnetoresistance of 1056% at Room Temperature in MgO Based Double Barrier Magnetic Tunnel Junction. *Appl. Phys. Express* **2**, 083002 (2009).
19. Thomson, W. XIX. On the electro-dynamic qualities of metals:—Effects of magnetization on the electric conductivity of nickel and of iron. *Proc. R. Soc. Lond.* **8**, 546–550 (1997).
20. Kondo, J. Anomalous Hall Effect and Magnetoresistance of Ferromagnetic Metals. *Prog. Theor. Phys.* **27**, 772–792 (1962).
21. Magnetoresistance. in *Magnetic Materials: Fundamentals and Applications* (ed. Spaldin, N. A.) 156–168 (Cambridge University Press, 2010). doi:10.1017/CBO9780511781599.013.
22. Dieny, B. Giant magnetoresistance in spin-valve multilayers. *J. Magn. Magn. Mater.* **136**, 335–359 (1994).
23. Gurney, K. *An introduction to neural networks*. (CRC press, 2018).
24. Gordon-Rodriguez, E., Loaiza-Ganem, G., Pleiss, G. & Cunningham, J. P. Uses and Abuses of the Cross-Entropy Loss: Case Studies in Modern Deep Learning. in *Proceedings on 'I Can't Believe It's*

Not Better!' at NeurIPS Workshops (eds. Zosa Forde, J., Ruiz, F., Pradier, M. F. & Schein, A.) vol. 137 1–10 (PMLR, 2020).

25. Graves, A., Fernández, S., Gomez, F. & Schmidhuber, J. Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks. in *Proceedings of the 23rd International Conference on Machine Learning* 369–376 (Association for Computing Machinery, 2006). doi:10.1145/1143844.1143891.
26. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
27. Jaeger, H. Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the" echo state network" approach. (2002).
28. Lukoševičius, M., Jaeger, H. & Schrauwen, B. Reservoir Computing Trends. *KI - Kunstliche Intell.* **26**, 365–371 (2012).
29. Jaeger, H. *The "echo state" approach to analysing and training recurrent neural networks- with an erratum note*. *GMD Technical Report* (2001).
30. Manneschi, L. & Vasilaki, E. An alternative to backpropagation through time. *Nat. Mach. Intell.* **2**, 155–156 (2020).
31. Anderson, J., Belkin, M., Goyal, N., Rademacher, L. & Voss, J. The More, the Merrier: the Blessing of Dimensionality for Learning Large Gaussian Mixtures. in *Proceedings of The 27th Conference on Learning Theory* 1135–1164 (PMLR, 2014).
32. Gorban, A. N., Tyukin, I. Yu. & Romanenko, I. The Blessing of Dimensionality: Separation Theorems in the Thermodynamic Limit**The work is partially supported by Innovate UK, Technology Strategy Board, Knowledge Transfer Partnership grant KTP009890. *IFAC-Pap.* **49**, 64–69 (2016).
33. Blessing of dimensionality: mathematical foundations of the statistical physics of data | Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. https://royalsocietypublishing.org/doi/full/10.1098/rsta.2017.0237.

34. Vapnik, V. N. An overview of statistical learning theory. *IEEE Trans. Neural Netw.* **10**, 988–999 (1999).

35. Learnability and the Vapnik-Chervonenkis dimension | Journal of the ACM. https://dl.acm.org/doi/abs/10.1145/76359.76371.

36. Paquot, Y. *et al.* Optoelectronic reservoir computing. *Sci. Rep.* **2**, (2012).

37. Jaeger, H. & Haas, H. Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication. *Science* **304**, 78–80 (2004).

38. Lukoševičius, M. & Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **3**, 127–149 (2009).

3 – Literature Review

3.1- Introduction

The field of *in materio* computation has seen rapid development within the magnetics community, with a plethora of magnetic RC architectures proposed in both simulation and in experiments. To develop functional RC platforms on novel substrates, a cross-disciplinary approach must be taken to enable a harmonious interplay between understanding the physical response of the substrate, engineering of devices and measurement platforms, and application of appropriate machine learning paradigms. To address this, the following chapter will explore the literature for these key aspects in turn.

The first section is focussed upon the physics that governs the response of the nanoring arrays, exploring domain walls (DWs) at the nanoscale. DW dynamics in response to input will be introduced in more detail, focussing on the theory behind stochastic processes governing DW propagation. Next, techniques employed for manipulating DWs will be explored, as well as developments in methods used for reading the magnetic state of nanostructures. Finally, there will be a deeper exploration into the magnetic nanorings that have been explored in this thesis, both in terms of the behaviour and responses of individual nanorings, as well as the changes to response that occur when many rings are connected to form large arrays.

The next section will focus on RC, building upon the archetypal network configurations introduced in the previous chapter, as well as more recent iterations that have been proposed for both physical as well as algorithmic implementations of RC. The applications of RC will be explored, detailing the areas of machine learning where RC is able to represent state-of-the-art performance despite its relatively low computational cost. Methods for evaluating the computational properties of an RC platform in different regimes will be introduced, allowing the range of performance to be assessed in a more general manner than simply reporting accuracies in benchmark tasks. Finally, we will explore novel learning rules and methods for evolving RC platforms which have been designed to maximise the computational capability that can be extracted from a platform and make them more robust against well-documented shortcomings of RC such as catastrophic forgetting.

The final section will draw on the concepts outlined in the first two sections, covering the current state of *in materia* computing in both magnetic and non-magnetic platforms, forming a critical evaluation of the strengths and weaknesses of different proposed systems, as well as highlighting important developments on specific platforms that can be applied to *in materia* computing platforms more generally. This will provide important context for where the developments highlighted in later chapters sit within the field.

3.2 – Domain Walls at the Nanoscale

3.2.1- Domain Walls in Magnetic Nanowires

In planar nanowires made from soft magnetic materials such as Permalloy (Py), the low magneto-crystalline anisotropy means that the magnetic ground states of geometries are determined predominately by a balancing between magnetostatic and exchange interactions. The minimisation of magnetostatic energy confines the spins to the plane of the nanowire, and results in an in-plane rotation of magnetic moments between domains (Néel DWs). The high aspect ratio and resulting high shape anisotropy means magnetisation in domains follows the wires' longitudinal direction, leading to

a 180-degree rotation of magnetisation at DWs. Two types of DW arise depending on whether magnetisation is converging (head-to-head, H2H, converging magnetisation) or diverging (tail-to-tail, T2T, diverging magnetisation) at the DW, shown in figure 1.

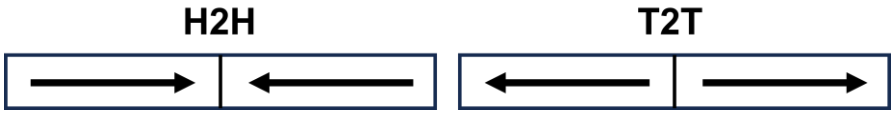


Figure 1- Schematic diagrams of Head-to-Head (left) and Tail-to-Tail (right) domain wall structure. Arrows show direction of magnetic moments within domains.

As well as direction of magnetisation, DWs also exist in two distinct classes depending upon the shape of the DW[1–3], forming either vortex or transverse DWs. Generally, vortex DWs result in reduced magnetostatic energy, as they are formed of closed loops of magnetic flux and are more commonly found in thicker, wider nanowires[2,4]. Conversely, transverse DWs reduce exchange energy, and are typical in thinner nanowires[3]. Figure 2 shows these DW structures in two different nanoring geometries.

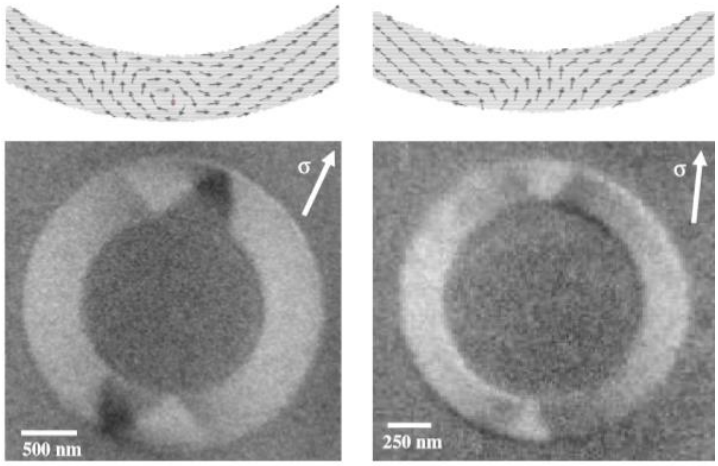


Figure 2- Upper: micromagnetic simulation of local magnetic moment configuration (shown by black arrows) for a vortex domain wall (left) and a transverse domain wall (right). Lower: X-Ray Photoemission electron microscopy images [X-PEEM] of both vortex (left) and transverse (right) domain wall structures. Axis of magnetic sensitivity is shown by arrow labelled σ, with moments pointing along this direction coloured white, and moments anti-parallel coloured black[5].

3.2.2- Domain Wall Motion in Nanowires

In nanowire devices, the isolated DWs within the constrained geometry can be probed with a high degree of precision, establishing a more detailed understanding of DW movement and how it changes with applied field. In figure 3, Beach *et al.* showed that DWs propagate in three distinct regimes in nanowire devices depending on the strength of the applied field. They expressed these regimes with respect to the changing DW mobility parameter, μ, which is equal to the gradient of the mean velocity plot shown below[6]. These three regimes are useful to reference when discussing the varying degrees of stochasticity of DW pinning events, as they each feature different structure and dynamics important to understanding their motion and associated stochasticity, which is addressed in the next section.

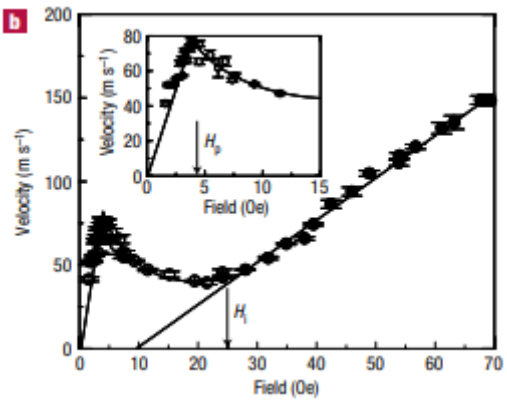


Figure 3- Average DW velocity versus driving field strength, showing initial linear regime of high mobility, subsequent decrease in velocity due to onset of Walker breakdown, before reduced linear mobility of turbulent DWs. Datapoints reflect experimental measurements, with associated error bars[6].

3.2.3 - Stochasticity in DW Pinning

Experimentally, DWs almost always experience some degree of stochasticity during pinning/depinning events in nanowires. This stochasticity has two main sources: the dependence of interactions between DWs and pinning sites upon DW structure, as well as thermal activation. The former will be discussed in detail first. Coupled with finite edge roughness in real nanowire devices, stochasticity is further enhanced, meaning that for finite temperature devices, the role of this stochasticity is often huge[7]. However, through engineering of the materials as well as careful application of driving stimuli, the effects of this stochasticity can be reduced.

3.2.3.1 - Stochasticity due to Dynamic DW Structure

At lower driving fields, the motion of DWs is generally well defined, with velocity scaling linearly with applied field. The mobility parameter μ, which describes the change in velocity with respect to applied field, in this regime is large and constant[6,8,9]. DW structure remains consistent as it propagates, meaning DWs arrive at pinning locations with similar structures across repeated events. Using the phenomenological Gilbert damping parameter, α, DW mobility can be defined in this regime via equation (3.1):

$$(3.1) \quad \mu = \frac{\gamma \Delta}{\alpha}$$

where γ is the gyromagnetic ratio, and Δ is the DW width. Through extensive control over DW injection and propagation, the associated stochasticity with DW pinning/depinning can be reduced to effectively zero in nanowires thanks to consistency of DW magnetisation state, demonstrated experimentally by Munoz *et al.*[10].

As driving fields increase, the configuration of DWs becomes oscillatory in nature beyond a characteristic field strength known as the Walker breakdown field[11]. Beyond this point, the mobility parameter, μ, becomes negative and varies with applied field. In this regime, the width, velocity, and structure of DWs all become transient, leading to highly stochastic pinning events as the DW interacts with pinning sites in a distribution of magnetisation states[11]. Hence, probabilistic pinning/depinning events occur according to this distribution, with individual outcomes for a single DW impossible to predict deterministically in this regime[7,8].

The onset of Walker breakdown can be offset via a variety of methods. Broomhall *et al.* demonstrated that it is possible to extend the viscous driving regime of permalloy nanowires by doping with heavy rare-earth metals such as terbium[12]. This was achieved by increasing the Gilbert damping constant of the material, since the field at which Walker breakdown occurs at is inversely proportional to the damping constant.[13] Figure 4 shows the result of this delayed onset of walker breakdown via increased damping.

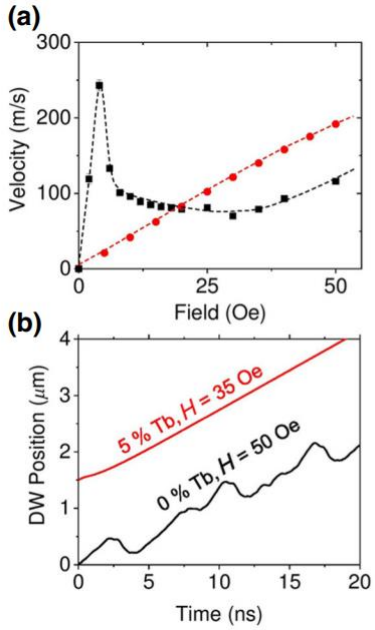


Figure 4- (a) comparison of domain wall velocity versus applied field strength for undoped permalloy (black line) versus permalloy doped with 5% terbium (red line). (b) domain wall position with respect to time for the doped sample driven at 35 Oe (red), and the undoped sample at 50 Oe (black). Taken from[12].

Additionally, the application of transverse applied field components has also been shown to reduce the onset of Walker breakdown by Glathe *et al.*[14]. This is due to the broadening of the domain wall width with applied field, which reduces the effects of edge roughness and acts to stabilise the domain wall structure. Due to the presence of considerable transverse fields in the rotating magnetic field schemes used to drive the nanoring arrays in this thesis, as well as the relatively low driving fields used, Walker breakdown will be largely suppressed.

Beyond this regime, the DW velocity profile once again becomes linear with applied field, but with a mobility an order of magnitude lower than the low field regime, due to turbulent DW structure[6,11,15]. According to Beach *et al.*, this large decrease comes as the periodic torques become balanced, leaving only the damping torque term to drive the DW forward[6]. The characteristic equation for mobility changes in this regime is given by equation (3.2), which highlights the dependence of the reduction in mobility on the damping parameter; materials with low damping characteristics have a greater reduction to mobility.

$$\mu = \frac{\gamma \Delta}{(\alpha + \alpha^{-1})} \quad (3.2)$$

3.2.3.2 - Stochasticity due to Thermal Activation

Another source of stochasticity in magnetic systems comes from thermal activation. Thermal fluctuations in any material with above absolute zero temperature cause individual magnetic moments

to flip, exciting spin waves and reducing magnetic order, aiding switching[16]. These random perturbations of spin structure act as a source of stochasticity in domain wall pinning[17], with the thermal energy of the atoms in the system being probabilistic in nature. Mathematically, the frequency that these perturbations lead to switching agrees with Arrhenius-Néel models[18], where the characteristic timescale of a thermal activation event occurring is exponentially dependent upon temperature, as shown by equation (3.3)

$$\tau = \tau_0 e^{\frac{\Delta E}{k_B T}} \quad (3.3)$$

where τ is the characteristic timescale of reversal, τ_0 is the reciprocal of the attempt frequency of the material, ΔE is the associated energy barrier, kB is the Boltzmann constant, and T is temperature[19]. Wuth *et al.* demonstrated this temperature dependence for depinning events in nanowires[20]. This characteristic timescale can be tied to the probability a pinning event is expected to occur, P, over a given duration of time, t, shown in equation (3.4):

$$P = 1 - e^{-\frac{t}{\tau}} \quad (3.4)$$

Building upon this theory of thermal activation, the Néel-Brown model develops a field dependence to the probability of switching[21]. This considers the effect that applied magnetic fields have on the energy barrier associated with DW pinning in Arrhenius-Néel models, ΔE, through equation (3.5):

$$E(H) = E_0 \left(1 - \frac{H}{H_{sw}^0}\right)^{\propto} \quad (3.5)$$

where E0 represents the energy barrier at zero field, H is the strength of the applied field, H_{sw}^0 is the switching field of the particle at zero kelvin, and α is an exponent describing the process of magnetic switching, usually around 1.5 for a single domain particle with uniaxial anisotropy[21]. Wernsdorfer *et al.* confirmed this relationship experimentally in[22]. Thermal activation will likely be the key source of stochasticity in the ring arrays studied here, as the transverse components of the applied cyclic driving fields increases the stability of propagating DWs and inhibits the onset of Walker breakdown, discussed in more detail in later sections.

3.2.4- Driving Domain Walls in Nanodevices

While electromagnets can provide a simple pathway to controllable driving of DWs in ferromagnetic materials, the relative size and power consumption of the electromagnets compared to proposed functional nanostructures means they are not the ideal candidate for DW-based devices. Alternatively, there are three commonly used methods for nanodevices that either provide magnetic field locally to the patterned nanostructures, or that use interactions between the magnetic moments of a material and electrical currents to produce spin-transfer torques. These torques can be generated via spin-polarising in pure charge currents, or via pure spin currents resulting from spin-hall effects. The mechanisms and applications of each of these methods will be briefly outlined in this section.

3.2.4.1- Oersted fields from local current striplines

While external electromagnets can easily generate magnetic fields required to provide stimulus to magnetic devices, only a tiny proportion of the energy stored in the generated magnetic fields goes on to cause a useful perturbation in the system. This poses a huge efficiency problem for proposed computational devices that use external electromagnets to provide input, since the magnets will consume energy at orders of magnitude higher than the computational platform itself. One possible

method for mitigating this using local current striplines to provide Oersted fields directly to the device. By minimising the distance between the electric current and the magnetic nanodevice, the amount of energy lost to generating fields away from where they are useful is minimised.

Previous work has shown that these local Oersted fields are capable of both generating and annihilating DWs in patterned nanowires[23], whilst requiring currents on the order of sub-milliamps. It has also been shown that these current striplines are able to reliably control magnetisation switching and domain wall propagation[24]. While these methods are promising for low-power modulation of the system, they are usually used to provide short-lived pulses of strong magnetic field rather than a continuously modulated driving field. At the current densities required to provide continuous fields on the order of 50 Oe, Joule heating would likely cause the current-carrying nanowires to become damaged.

3.2.4.2- Spin-transfer torques via spin-polarised charge currents.

The spin-transfer effect was theoretically predicted in 1996 by Berger[25] and Slonczewski[26], with huge promise for the application of spintronic devices. The effect arises from the transferring of momentum from a spin-polarised charge current to a non-orthogonal magnetic moment. Grollier *et al.* explain the process by which spin-transfer effects can be used to manipulate domain walls[27]: In magnetic nanowires with in-plane magnetisation, magnetic moments in the domains align pointing along the length of the nanowire. When charge currents are passed through the nanowires, s-d interactions lead to spin polarisation in the current. Since the size of DWs in nanowire geometries (on the order of wire width, generally 10s to 100s of nanometres) is significantly larger than the spin-diffusion length (around 4nm for Py at room temperature) or the coherence length (~1 or 2 nm), the change in direction of the magnetic moments presented by the DW is gradual enough for the conducting electrons to adiabatically follow the rotating moments in the DW. This imparts a transfer of momentum as the current traverses the DW, causing the rotation of moments in the DW. With sufficiently large currents, this rotation is enough to lead to DW propagation.

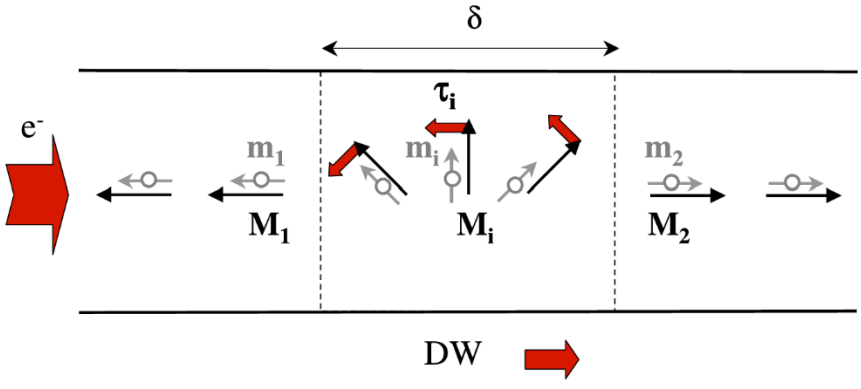


Figure 6 -Schematic diagram showing the resulting spin-transfer torques (red arrows) that arise across a domain wall (region between dotted lines) due to the spin polarisation that occurs when a charge current (denoted e^-) flows through a nanowire with aligned magnetic moments (shown by M_1). Taken from[27].

Whilst a viable route for manipulating DWs in nanowires, driving the ring array system with spin-transfer torques would still require large amplitude, short duration pulses of electric current to achieve the required current densities for DW propagation. Additionally, the continuously connected nature of

the nanoring system would pose difficulty when attempting to address subsections of an array individually, which could limit the use of spin-transfer torques to global input.

3.2.4.3- Spin-orbit torques via pure spin currents generated in a heavy-metal underlayer

An alternative method for utilising the spin-transfer effect to manipulate DWs in nanodevices involves the generation of pure spin currents generated via the spin-Hall effect when current is passed through a heavy metal[28]. Here, spin-orbit interactions lead to spin-dependent scattering of conducting electrons, meaning spin-up and spin-down electrons are scattered in opposing directions, generating a pure spin current orthogonal to the flow of conduction electrons, and is shown in figure 7.

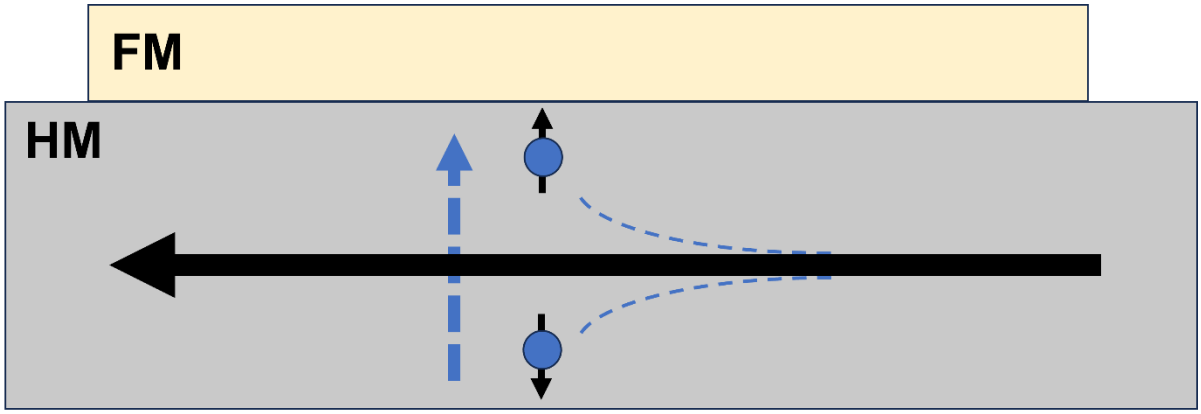


Figure 7- Spin dependent scattering of a charge current (black arrow) passed through a heavy metal underlayer, which causes accumulation of spin-up electrons at the interface of the heavy metal and the ferromagnet here. The preferential motion of the spin-up and spin-down electrons in opposite directions creates an effective spin current in the direction of the blue dotted arrow.

The relative efficiency of converting a charge-current to a pure spin-current for a given material, as well as the direction in which spin-up/spin-down electrons are deflected, depend upon the magnitude and sign of a dimensionless parameter known as the spin-Hall angle, where experimentally measured and theoretically determined values often have large discrepancies[29]. Experimentally determined spin-Hall angles often determine which heavy-metal underlayers are used, with platinum being the most frequent choice due to its large spin-Hall angle (ranging typically between 0.01 to 0.20[30]) and high conductivity.

In nanodevices, multilayers of ferromagnetic material deposited upon heavy-metal underlayers have been employed to utilise spin-orbit torques to cause DW propagation or magnetic reversal. This has seen implementations of DW-based logic[31], racetrack memory[32], and magnetic random-access memory[33], which use spin-orbit torques to electrically encode information into magnetic media at energies significantly lower than those of spin-transfer torques. However, since the direction of the spin current is oriented out-of-plane with respect to the magnetic material, the driving torques applied to domain walls are very weak for in-plane domain walls, and hence are primarily used to drive out-of-plane Néel walls[34]. In order to transfer sufficient torques, the spins of an in-plane material must be canted slightly out-of-plane via strong magnetic fields.

3.3 – Magnetic Nanorings

The system explored in this thesis consists of arrays of overlapping permalloy nanorings. In essence, an individual nanoring behaves like a nanowire looped back upon itself, providing a means of flux closure. The following section will explore the ground state and state space of individual magnetic nanorings, the transitions between these states, as well as their response to rotating magnetic fields. Next, the interactions that occur between connected rings will be examined, highlighting the changes in response that occur due to these interactions, before finally examining the emergent response of large arrays of interconnected nanorings arising from many interacting elements.

3.3.1 - Single Magnetic Nanorings

In thin-film nanostructures made of soft magnetic materials such as permalloy, magnetisation states are primarily controlled by the magnetostatic energy, and hence are strongly dependent on the shape of the nanostructure. The nanorings explored here resemble high aspect-ratio wires looped upon themselves, with the relatively narrow track widths leading to a significant shape anisotropy and a strong tendency for circumferential alignment of magnetic moments[5]. The ground state of the system is a closed loop around the ring circumference, forming a single magnetic domain (figure 8). This is an energy minimum as there are no poles from DWs (and hence no magnetostatic contributions) and is named the 'vortex' state. Vortex states have twofold chirality depending on the direction of magnetisation, leading to distinct clockwise and anticlockwise vortices.

Figure 8- Schematic diagram *of a "vortex" ring* configuration. *Arrow shows direction of magnetisation[35].*

As well as the vortex state, there also exist metastable, multidomain states (figure 9). The most frequent of these is the 'onion' state[5,35]. The onion state forms when a sufficiently strong magnetic field causes ferromagnetic alignment in the ring and is then relaxed, forming a pair of DWs along the direction of magnetic saturation. The 'twisted' state arises when a DW is driven around the system until it comes into proximity with a second DW in the ring and interacts at a fixed length with behaviour like a 360° DW[5,36].

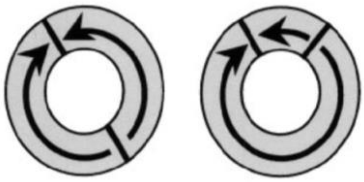


Figure 9- Schematics of "onion" ring (left) and "twisted" ring (right) configurations. Arrows show magnetisation direction, lines across width of ring represent domain walls[35].

Under the application of a linear field sweep on an isolated nanoring, the transition between onion and vortex states can be observed as they mediate the switching of magnetic alignment. These transitions are inferred from characteristic hysteresis loops, illustrated in figure 10. The way in which these changes occur is generally controlled by the geometry of the ring, strength of applied field, and

degree of edge roughness[36]. Micromagnetic simulations of field-driven motion in the absence of pinning suggest that for onion state rings there exists two mechanisms by which the DWs propagate around the ring: coherent rotation of the DWs around the ring, or the forming of an intermediate vortex state[36].

The mechanisms, shown in figure 10, depend on the movement the DW pair: if they move in the same direction, they collide and annihilate, forming the intermediate vortex state; if they move in opposite directions, there is uniform rotation of the onion state as magnetisation direction reverses. The interchangeability between onion and vortex state is key to the dynamics of the arrays studied in this report.

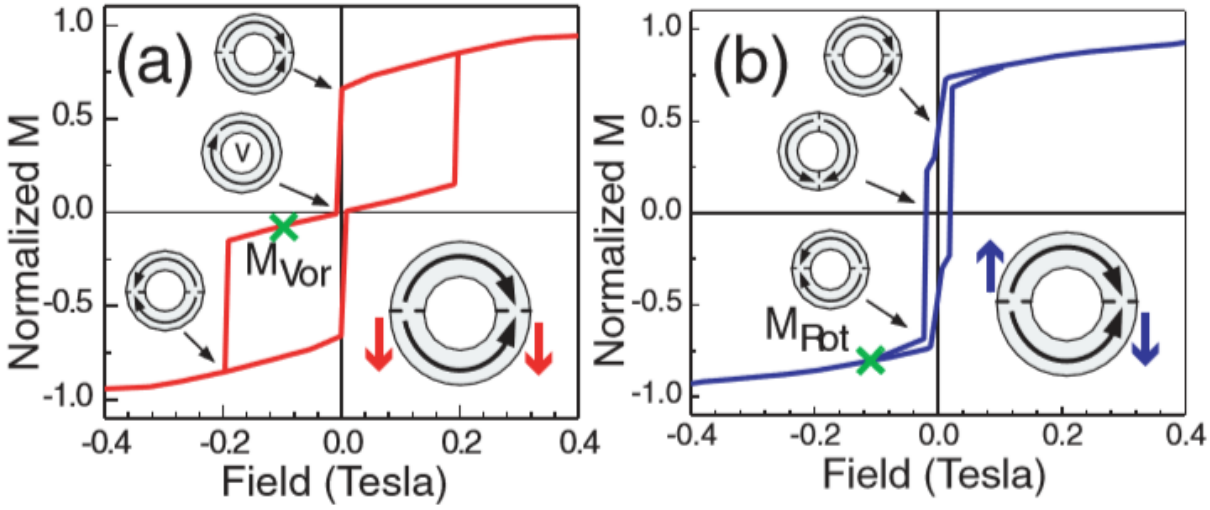


Figure 10- Micromagnetically obtained hysteresis loops for reversal mechanisms. Vortex formation, (a), has step-like behaviour in magnetization as domain configuration changes, whereas (b) has less abrupt changes with constant DW configuration. Inset figures show domain wall state at various stages of the hysteresis loops as well as direction of DW movement[36].

A study by Negoita *et al*. explored how DWs move when driven by an in-plane rotating magnetic field. Under a rotating driving field, DWs are expected to rotate along with the field to maintain minimum Zeeman energy. However, it was observed that DW pinning due to edge roughness causes the DWs to lag behind the driving field[37]. According to the schematic shown in figure 11, lag angle is defined as the polar angle from the ring's centre between DW location and the location of the energy minima. It was found that, as a result of increasing the transverse component of applied field, the DWs would depin with either sufficient applied field or large enough lag angle, with a sinusoidal relationship tying the parameters together, shown in (3.6)[37].

$$\theta_{lag} = sin\left(H_{app}\right) \quad (3.6)$$

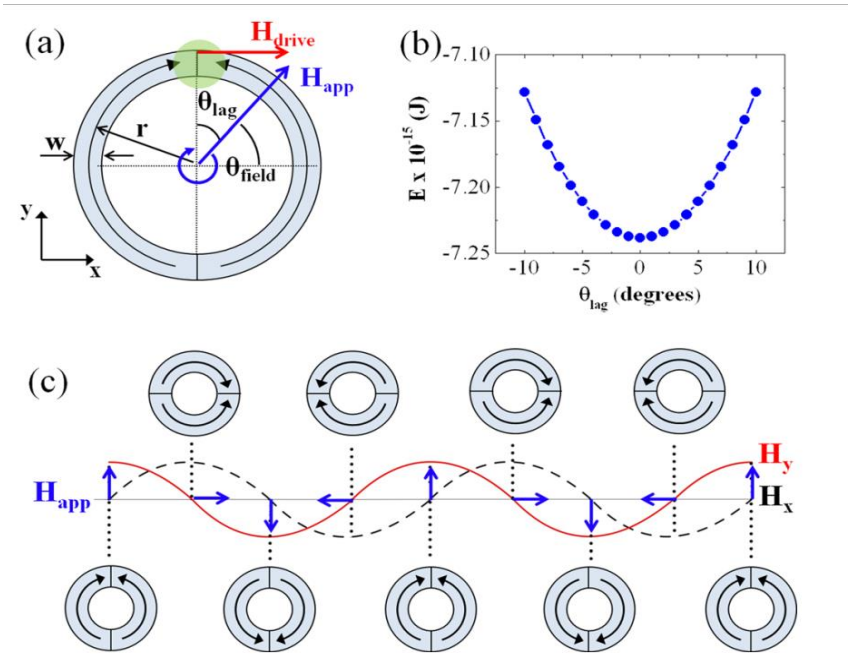


Figure 11- a) Geometrical definition of lag angle for ring structures b) Micromagnetically calculated Zeeman energy versus lag angle, and c) Components of applied field and DW structure an applied field rotates[37].

The cyclic field also influences the stochasticity of propagation in the ring systems due to the presence of transverse components of applied field. A study by Glathe *et al.* looked at the effect of these transverse fields on DW propagation in permalloy nanostrips and found that the applied transverse fields delayed onset of Walker breakdown, suppressing it completely with sufficiently strong transverse fields[38]. These findings suggest the presence of considerable transverse components in the rotating driving field applied to the ring system will likely mean Walker breakdown is less significant, as the onset of breakdown is delayed. Additionally, the reduction in driving component that occurs as DWs depin and lag angle reduces means Walker breakdown is also less likely to occur.

3.3.2 - Connected Nanorings

A key area for this thesis is defining how nanorings interact with adjacent rings in the connected arrays. Junctions where the rings overlap will have different magnetic characteristics to the rest of the ring: the local change in geometry creates a defect-like energy barrier for DW propagation. The increased area of magnetic material at the junction causes changes to DW structure upon entering a junction, with the additional material acting similarly to an 'anti-notch', disrupting local magnetic order[39,40]. In multi-ring systems, the junctions are critical to the mechanisms behind DW annihilation and nucleation, as they act as the primary pinning sites and cause differential movement of DWs about the rings[41].

The geometry of connections has significant influence on their behaviour[42]. For continuous connections with no separation of material, strong exchange coupling is present between the rings. For spatially separated rings, magnetostatic coupling acts across the separation. The strength of coupling is dependent upon distance of separation, with stronger coupling the closer the rings are located. Ren *et al.* measured MOKE hysteresis loops on four arrangements of interacting nanorings in response to linear field sweeps, shown in Figure 12, with the geometry of the arrangements shown in the micrographs above.

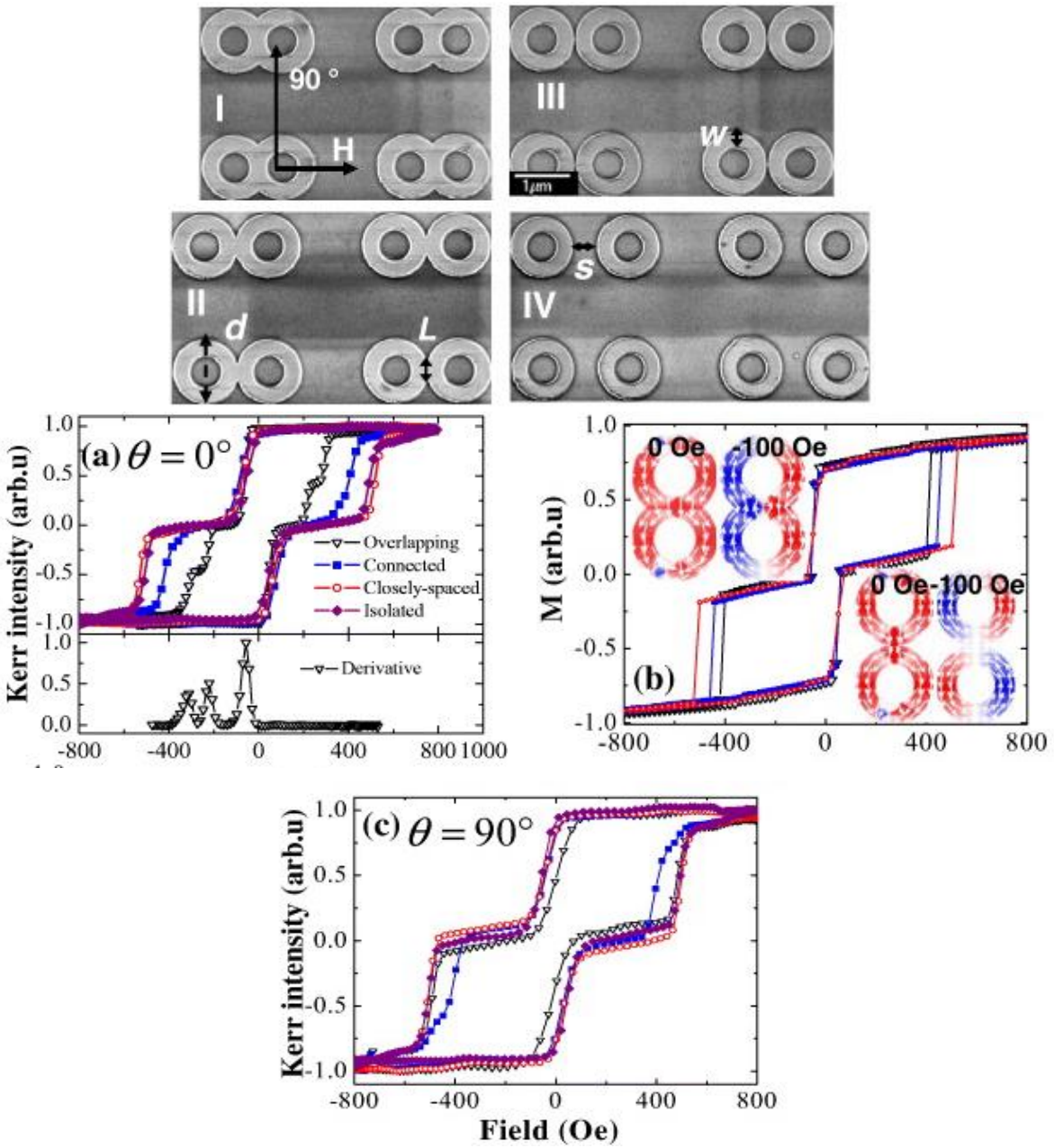


Figure 12- *Upper Image - Scanning electron microscopy (SEM) micrographs of pairs of interacting nanorings: Top Left - Overlapping, Bottom Left - Connected, Top Right - Separated, Bottom Right – Isolated*[42]*. Lower Image(a) MOKE hysteresis loops for the four connectivity arrangements for 0 angle. The derivative plot below highlights local minima representing switching for the 'overlapping' ring pairs b) Micromagnetically simulated hysteresis loops of the same arrangements, with images of simulated DW arrangements of overlapping (left) and connected (right) ring pairs in the inset, (c) MOKE hysteresis loops for the four bi-ring configurations for 90 angle*[42]*.*

Important features to note here include the limited difference between the 'closely spaced' and 'isolated' ring pairs, and the significant difference between the 'connected' and 'overlapping' ring pairs. This highlights the much greater influence of exchange coupling over magnetostatic coupling on the switching process, as well as the different reversal mechanics observed between the two magnetically contacted ring pairs. The derivative plot highlights a clear 3-step reversal process for overlapping rings, showing a difference in reversal mechanism that is dependent on the degree of overlap between rings[42], though the study does not allude to critical values where this change occurs.

The differences in the 0- and 90-degree plots would be expected to be of less importance to the response of square arrays of interconnected rings as there will be fourfold symmetry in a uniform square lattice, removing differences between 0 and 90 directions. Understanding of these reversal processes will be key to engineering the array dynamics, as it could allow the energy barriers to be manipulated between rings, changing the nature of device response, and the relative favourability of different states.

The number of connections between rings also affects expected magnetisation state within nanorings. Rose *et al.* showed that with increasing connections, the stability of the vortex state decreases[41], highlighting the influence the number of connections has on the reversal properties of individual rings, which will be important to consider for rings on the perimeter of the interconnected ring arrays. Other critical conclusions of the work of Rose *et al*. includes the finding that propagating DWs in one ring are able to cause movement of DWs in adjacent rings, and that there is a reduction in switching field and coercivity as connectivity between rings increases[41]. This is highlighted by the hysteresis loops shown in figure 13.

However, it is important to also understand how magnetic frustration may affect the rings' switching behaviour, where geometry is a key component in resolving locally frustrated rings. Frustration arises when changes in magnetisation cause regions of opposite polarity sit opposite to one another, forcing creation of a DW along the change in magnetic direction, shown in figure 14. In the triad structure studied here and the square structures in the arrays discussed later, there are different pathways available to form closed magnetic loops across rings, which in turn leads to greater degrees of freedom with which to alleviate frustration in square arrays, reducing the number of expected frustrated states.

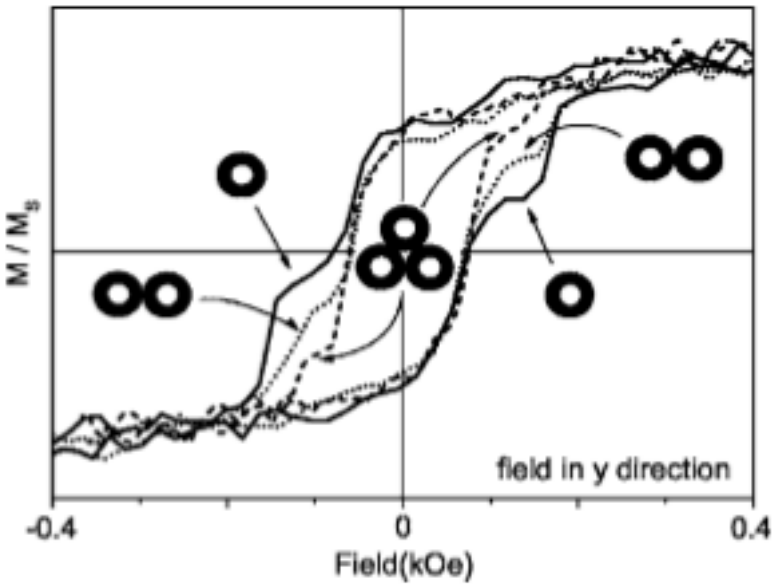


Figure 13- *MOKE hysteresis loops for single ring (solid line), bi-ring structure (dotted line), and tri-ring structure (dashed line)*[41].

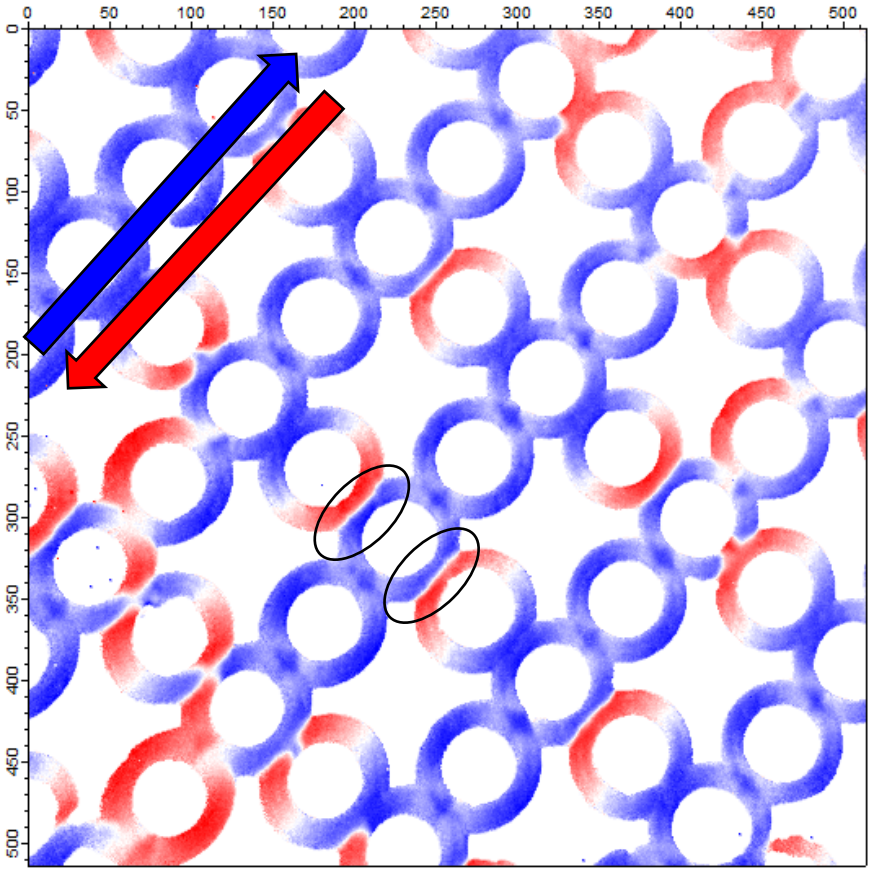


Figure 14- X-PEEM images of connected nanoring chains, black highlighted domain walls are formed due to local frustration. Red and blue colour denotes direction of magnetic alignment, shown by blue and red arrows.

To understand the nature of DW annihilation and nucleation mechanisms in connected bi-rings under rotating fields, Dawidek conducted micromagnetic simulations at zero Kelvin of connected ring pairs driven by rotating fields with different field strengths and initial conditions[43], shown in figure 15. The first case featured a fully populated pair of connected rings and a lower, 50 Oe driving field. The second simulation featured DWs in one ring only, and a higher, 80 Oe driving field. As shown by figure 15a, for rings driven by fields below a critical value, the magnetic field is insufficient to drive the DWs beyond the potential barrier of the junction without assistance from thermal effects, so remain pinned. The other DWs in the pair have no junctions to pass, so continue to propagate with the rotating field until they reach the pinned DWs and annihilate to form vortex states in both rings.

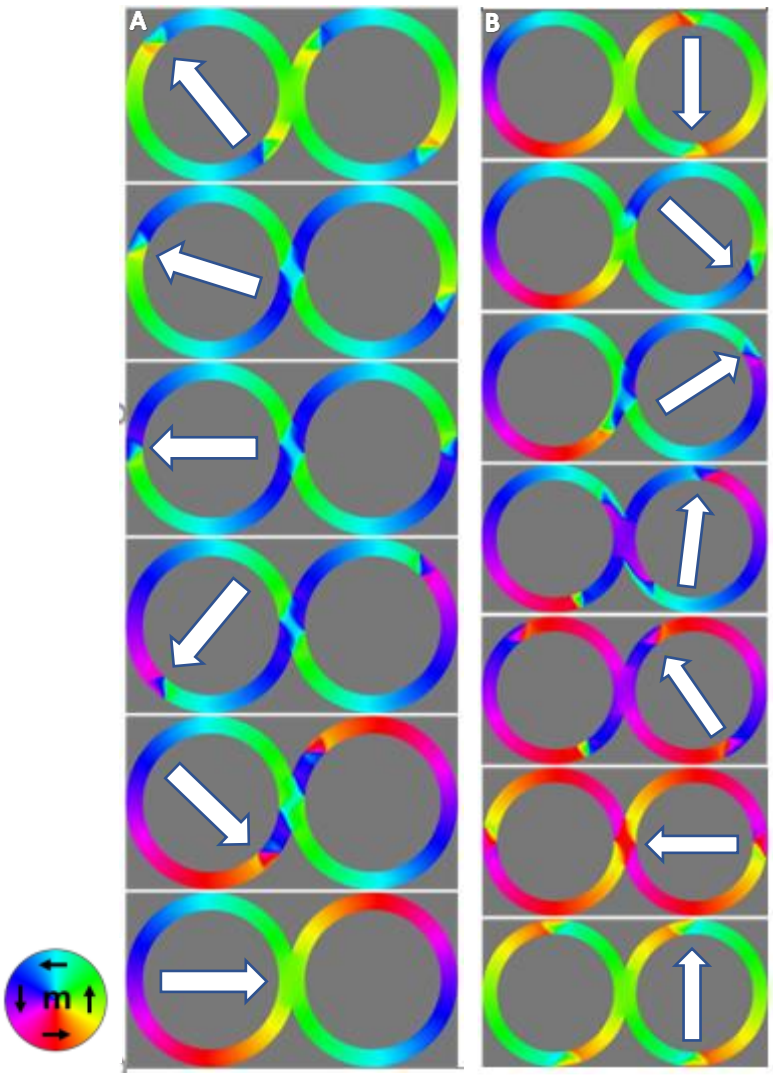


Figure 15- Micromagnetic simulations of connected bi-ring systems of 4□m diameter, 400nm track width, 20nm thickness and 50% overlap. A) Annihilation mechanism driven by a field of 50 Oe, B) Repopulation mechanism, driven by a field of 80 Oe. Colour on rings demonstrates magnetisation direction (see colour wheel), white arrows show direction of applied field[43].

The second simulation has DWs present in one ring only, and a driving field high enough to overcome the energy barrier of the junction (figure 15b). This means the DW passes the junction. As it passes, it causes magnetic reversal across the junction, causing frustration to occur between the populated ring and the empty ring. The frustration is immediately alleviated by the formation of a pair of DWs in the empty ring, which go on to follow the field as in the populated ring. This shows a mechanism for nucleating DWs in the system, relying on the movement of DWs in adjacent rings.

The micromagnetic simulations here neglect the effect of thermal noise, and as such are deterministic in nature. In systems at room temperature, the transition between pinning and propagating regimes is not absolute about a critical value, but rather a probabilistic process thanks to thermal activation, discussed earlier in sections 2.3 and 2.4.1. The consequences of the stochasticity associated with thermal activation will be addressed in the following section.

3.3.3 - Arrays of Connected Nanorings

As connectivity plays such a key part in the response of individual rings, it is important to consider some of the differences in behaviour expected from a large network of interconnected ring as opposed to isolated groups of 2 or 3 nanorings. Across large arrays, consistent global phenomena were observed[43], arising as a consequence of the individual stochastic interactions as well as the local influence a ring has over its neighbours.

These stochastic events meant that different percentages of the DWs are expected to depin with every rotation depending on the strength of the applied field. This leads to mechanisms for DW loss or gain depending on differential DW motion, with different domain states expected from ring to ring. The rates of DW loss and gain varies with applied field strength, leading to well-defined global states emerging as a statistical phenomenon across the entire array. This leads to expected levels for array magnetisation and normalised populations of each ring state as the ring arrays are driven to equilibrium, depending upon the magnitude of applied field.

Dawidek *et al.* explored the magnetisation characteristics of ultra-large samples of interconnected ring arrays over a $2cm^2$ wafer after the application of 50 successive rotations of magnetic fields, shown in figure 16[43], measured using polarised neutron reflectometry (PNR). At lower fields, the system struggled to deviate from the saturated state. At higher fields, the DWs always had sufficient driving force to overcome the energy barriers and follow the rotation of the magnetic field, returning to their initial conditions after the final rotation. The large drop in magnetisation at intermediate fields arises as individual stochastic events lead to a large distribution of magnetic states over the array. This resulted in a disordered magnetic texture across the array, leading to zero net magnetisation without forming exclusively vortex states (shown in later X-PEEM images). This complex nonlinear transformation is advantageous for machine learning applications, discussed later.

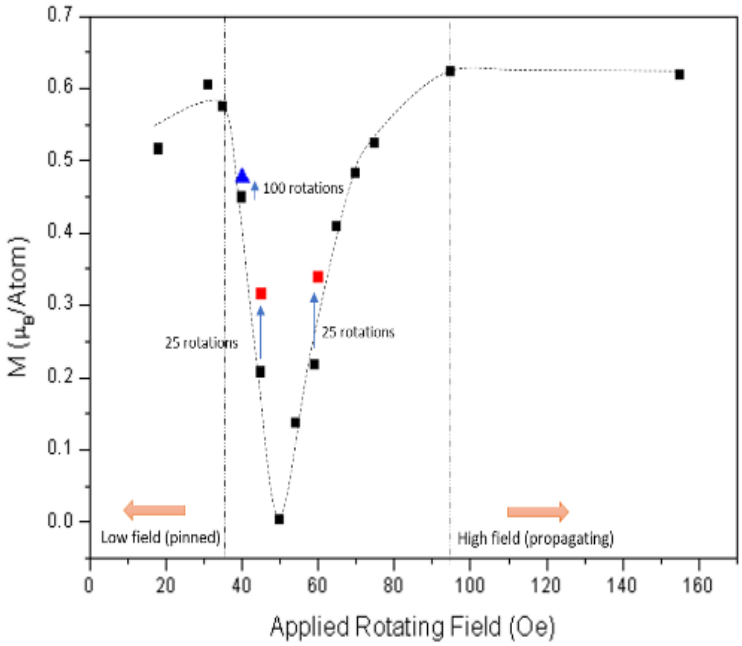


Figure 16- Array magnetisation state after 50 rotations of magnetic fields of varying strength. Blue triangles represent results from 100 rotations, red squares from 25 rotations[43].

Dawidek also conducted a similar experiment, where the array was driven over a series of varying applied field strengths without reinitialisation by a saturation pulse, shown below in figure 17. The differences in response between applications of the same field implies dynamic behaviour beyond a simple transition between well-established equilibrium positions, with the response at a given field being dependent upon the previous state off the array[43]. This hysteretic behaviour could prove useful as a source of memory in a computing system based on the arrays, with the state at a given time depending upon both the current input as well as past inputs to the system.

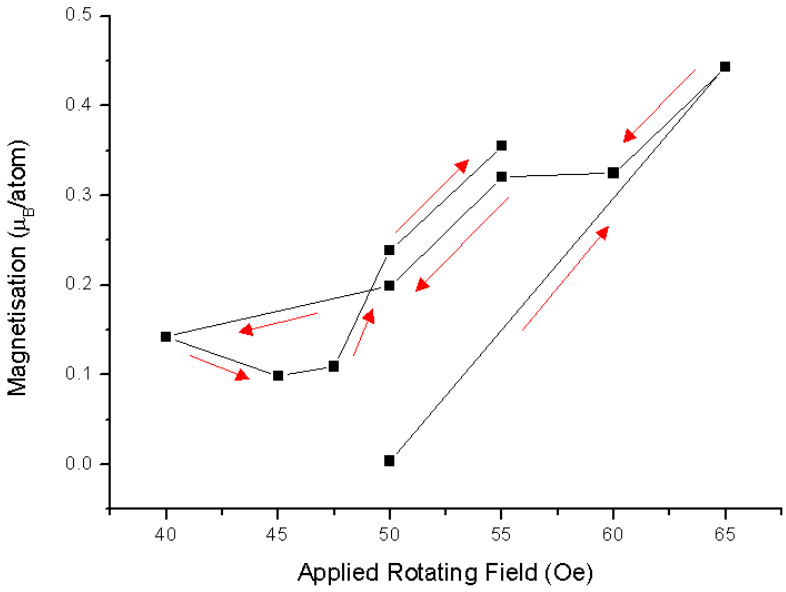


Figure 17- Magnetisation of ring arrays after 50 rotations of magnetic fields of varying strength with no saturation pulse between field sequences. Red arrows show the order in which measurements were taken.[43]

In the same work, X-PEEM imaging is used to provide greater insight to the magnetisation state after the application of cyclic fields (figure 18). This allows observation of the magnetic state of individual rings within the array. While the images are not identical, there exists regions of consistent behaviour, where rings of similar magnetisation tend to group together. This causes patterns to form within the images and demonstrates the influence of local magnetic structure on the ensemble response. This consistent complex structure on a global scale highlights an 'emergent' response of the arrays, with complexity arising because of many individual interactions[43].

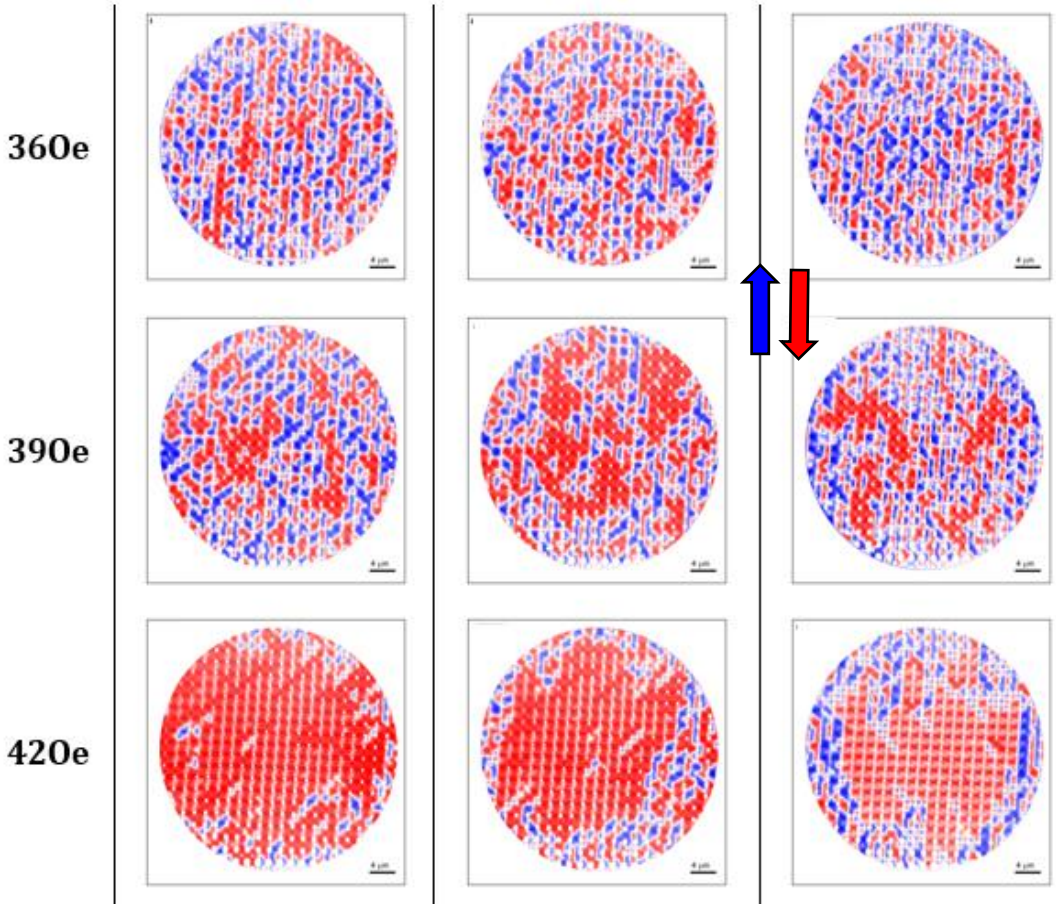


Figure 18- X-PEEM images of ring array samples following: Column 1; 30 rotations of specified field, Column 2; 30 rotations with +1.5Oe offset, Column 3; 30 rotations of original field. Red and blue colour denotes magnetic alignment with the provided arrows[43].

Agent based models (ABMs) are especially useful for exploring such a system where a set of well-defined agent rules can be established and defined phenomenologically. In ABMs, the complex behaviour 'emerges' because of these rules rather than being programmed in directly. The simplicity of agent-based models is critical, allowing a functional model of these large array systems to be built, as it can give rapid insight to the array dynamics where traditional micromagnetic modelling techniques would be unfeasible due to the scale of both time of evolution and array geometry. ABMs will be described in detail in the 'Methods' chapter.

3.4- Reservoir Computing

Building upon the theoretical frameworks for RC and their archetypal implementations of echo state networks introduced in the previous chapter, the following section will explore novel reservoir topologies, methods for characterising the performance of RC platforms, and novel learning rules designed to maximise reservoir performance and mitigate their shortcomings.

3.4.1 – Reservoirs with Ordered Connectivity

While ESNs and LSMs are generally randomly initialised, alternative reservoir paradigms with ordered connectivity have been proposed. These reservoir structures aim to maintain the computational properties of randomly connected ESNs, though with connectivity between the nodes which follow predetermined rules. This simplification of the connectivity is useful for physical proposals of reservoir

computing, since it limits device-to-device variability and provides synaptic connections that are much more feasible for hardware realisation.

One early example of ordered ESNs are the feed-forward ESNs (FF-ESN) proposed by Cernansky and Makula[44]. In these networks, nodes can be visualised in a linear chain, with connections only occurring in the 'down' direction of the chain (figure 19). Whilst this removes recurrent connection 'loops', there is still time-dependence in the reservoir nodes as information cascades downwards through the chain over time. To compare the performance of the FF-ESN with standard ESNs, the authors performed stability evaluation by training the FF-ESN to generate a periodic attractor[45] (in this case, a series of inputs correlated to the melody of 'The House of The Rising Sun'), with similar performance to standard ESNs. Comparisons were furthered using Mackey-Glass time series prediction, where the FF-ESNs outperformed standard ESNs.

The dynamics of the FF-ESNs are somewhat simpler due to the lack of looped recurrent connections, meaning that current activity can be described exactly as a function over a finite number of past inputs, dependent upon the connectivity. This differs from ESNs where there are infinitesimal contributions from all previous inputs. This opens interesting questions about the echo-state property, since the FF-ESNs exhibit strong computational performance without technically meeting the echo-state criteria.

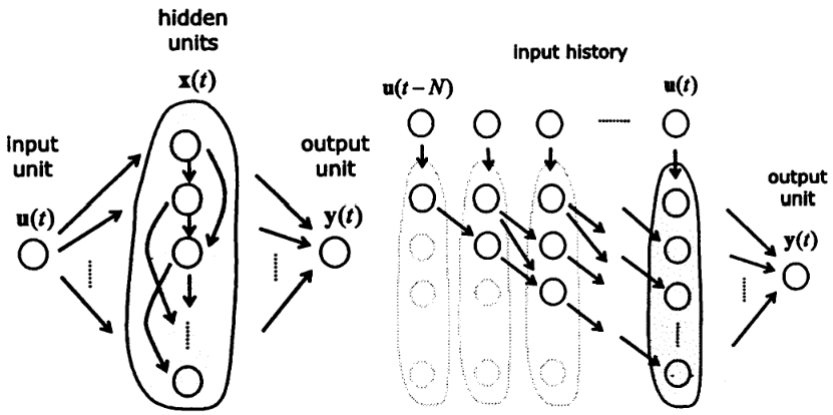


Figure 19- Schematic diagram showing the connectivity between nodes in the feed-forward ESN. Nodes are shown by circles and connections shown with arrows. Left diagram shows static connections, while the right diagram shows how inputs to the first node propagate through the array over several timesteps. Taken from[44].

Following the feedforward ESNs, other simple topologies such as the delay-line reservoir and its variants were introduced by Rodan and Tino[46]. Here, the connectivity of each node in the 'forward' direction is defined unidirectionally, forming a single chain of connected neurons with all weights of equal value. Figure 20 shows three different variations on this topology depending on the nature of feedback, consisting of the simple delay-line reservoir (20a), the delay-line with feedback (20b), and the simple cyclic reservoir (SCR, 20c). It can be observed that the delay line reservoir is a restricted case of the FF-ESN, where each node only connects to its adjacent forward neighbour.

In all cases, the input weights are binary and randomly sampled. For the standard case delay-line and SCR, the connections between neurons are in a single direction visualised orthogonally to the flow of input to output. This generates memory as each node receives information of the previous timestep from the previous node in the chain, increasing in delay along the chain with respect to the current input (the cyclic reservoir is identical to the simple delay-line though with an additional connection

linking the final node in the chain to the first). The delay-line with feedback operates similarly, though with an additional backwards connection accompanying the forward connections, with a smaller fixed weight.

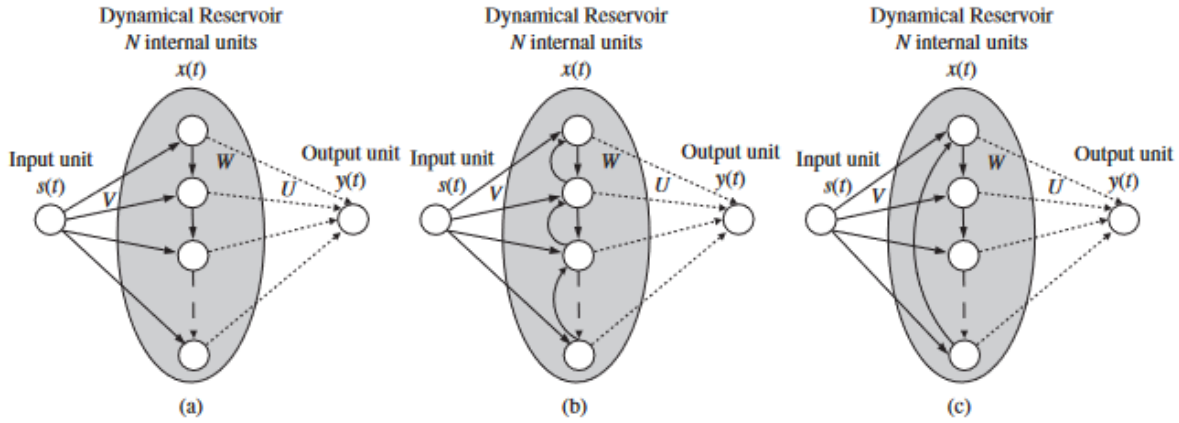


Figure 20- Network topology schematics showing connectivity between dynamical neurons for (a) delay-line reservoir, (b) delay-line reservoir with feedback, and (c) simple cyclic reservoir. Taken from[46].

To assess whether computational performance was maintained in these simple topologies compared to standard randomly generated ESNs, the authors performed a series of standard benchmark tests and evaluated the memory capacity (introduced in detail later in section 3.4.4) of the different configurations, as well as for standard ESNs, while controlling for network size. The simple topologies performed competitively, with the cyclic reservoir especially able to match standard ESN performance in almost all cases. This is a significant finding for the implementation of *in materio* reservoir platforms such as the ring system studied here, as ordered topologies of identical connectivity between discrete units is far simpler to implement in hardware than the random varied connectivity typical to standard ESNs.

3.4.2- Single Dynamical Node Reservoir

While the variants of the delay-line reservoir drastically reduced the network complexity of ESNs, they still require distinct dynamical nodes with synaptic connections. Another architecture, the single dynamical node reservoir (SDN) proposed by Appeltant *et al.*[47] (figure 21) reduces this complexity even further. This reservoir consists of 'virtual' nodes, generated from observing the state of a dynamic system as it evolves under input. Due to the temporal dependencies of the underlying dynamic system, the state of a virtual node at a given time depends not only on the current input, but also the state of the previous virtual node, hence this reservoir structure resembles the loop-like structure of the simple cyclic reservoir. To introduce longer-term dependencies to the state of the reservoir, delayed feedback is often included to connect each virtual node to itself in the past, adding a fraction of its previous state to the current input, which mimics leaky-integrator behaviour for each virtual node.

Inputs are provided to each of the virtual nodes through a time multiplexing procedure: a fixed mask modulates the inputs in a similar manner to the input weights in standard RC, with the masked inputs being fed to the dynamical node sequentially via a single input channel. This can also be extended to provide multiple input dimensions to a single multiplexed signal by altering the dimensions of the mask, providing a series of linear combinations of the multiple input dimensions over the duration of the sequential input.

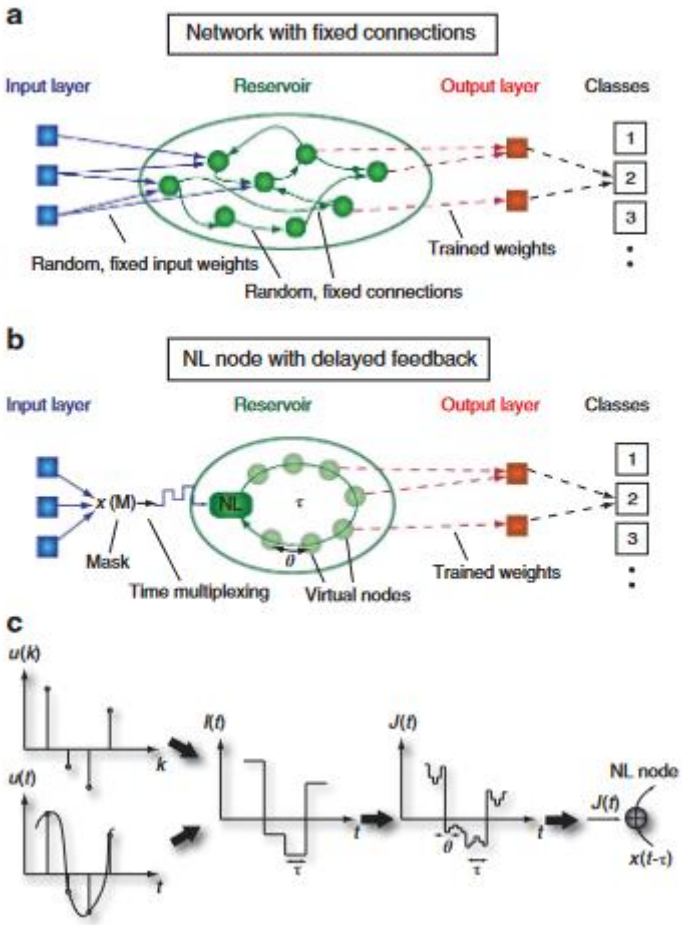


Figure 21- (a) Standard reservoir computing configuration. (b) Implementation of 'virtual node' cyclic delay line reservoir with a single dynamical node. (c) Schematic diagram showing combination of a time-discretised input signal u(k) to a continuous signal of steady state changes I(t) before combination with a fixed mask to produce a time-multiplexed signal J(t). Taken from[47].

Like with the delay-line reservoirs, the authors performed benchmark tasks and achieved competitive accuracies with SDN, highlighting that the simplifications made to the network topology did not result in a significant loss of performance. This development vastly improved the feasibility of physical dynamical systems for RC, as it removed the need for synaptic connections altogether; only methods for inputting data via a scalable stimulus and measuring an evolving physical property are required to realise a computationally useful RC platform. This led to implementations of single dynamical node reservoirs in both magnetic[48] and non-magnetic[49] oscillators, which will be explored in more detail in later sections.

3.4.3- Rotating Neurons Reservoir

Another recent implementation inspired by the initial cyclic reservoir aimed at providing simpler hardware implementation is the rotating neurons reservoir (RNR) introduced by Liang *et al.*[50] This configuration features distinct nodes like the original implementation of SCR, and similar binary input weights. The crucial difference is the implementation of the loop structure: instead of having explicit connections between the neurons that form a loop, the input and output connections synchronously 'rotate' (figure 22), meaning the readout of state for each input/output weight pair occurs on the adjacent node in each successive timestep. With the leaky-integrator neurons used in the RNR implementation, this becomes analogous to the SCR, since the current state readout will be dependent

on the input to the prior node in the 'loop' at the previous timestep via the leaking of past information, rather than directly receiving an input from time the previous timestep from a direct synaptic connection.

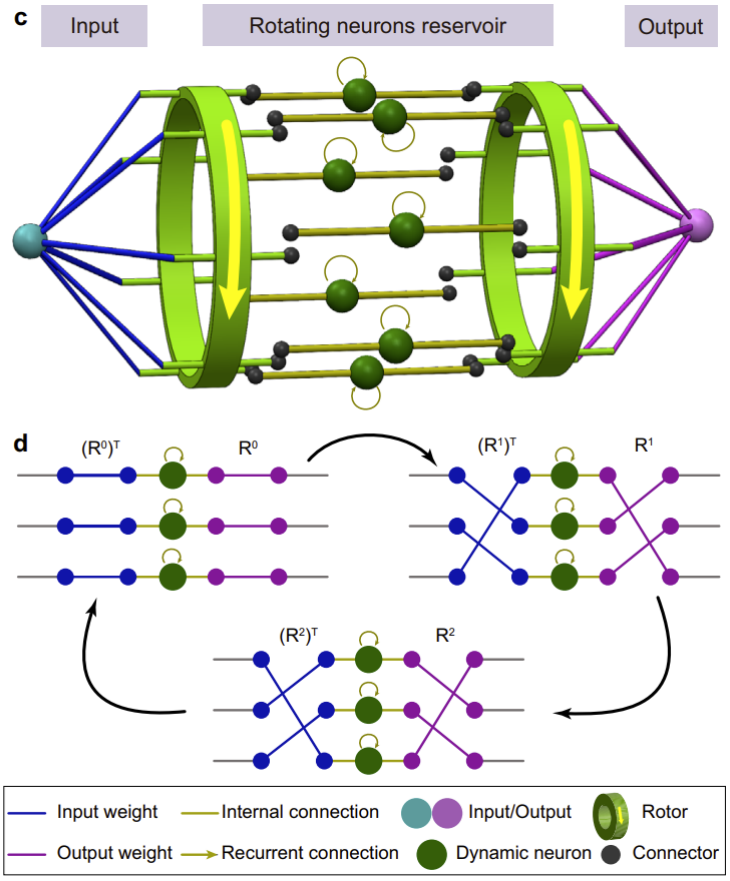


Figure 22- Schematic diagrams of the rotating neurons reservoir. Upper shows a visualisation of the rotating input and output connections (blue and purple respectively) connected to the dynamic neurons. Lower shows how the connections change for a three node network over three timesteps.

Again, performance was verified to the standard topologies using a combination of metrics and benchmark tests. Here, parameters on the dynamic nodes used in the RNR system were matched to the standard SCR case, and near identical performance obtained. While this implementation requires the use of multiplexer circuit components to create the 'rotating' effect, they are far simpler than the external delay lines used in the SDN case. Since data is no longer multiplexed in time, there are also gains in throughput compared to the SDN case (though still limited by the rate of the multiplexers). However, distinct nodes need to be manufactured and measured independently, leading to device-dependant variance in whether the RNR or SDN method has the simplest implementation when considering hardware realisations.

3.4.4- Characterising Reservoir Performance

While performance in task-based environments provide direct evidence of computation performed by a reservoir computing platform, they do little to inform on the range of computational properties a system can achieve, nor describe their computing power in general terms. As a result, various metrics have been proposed to statistically evaluate the computational power a reservoir transformation

provides based on arbitrary inputs, attempting to measure reservoir capability without the dependence on task-based data.

Kernel rank (KR) and generalisation rank (GR) are two common complimentary metrics which measure the ability of an RC platform to separate and generalise information respectively, introduced by Busing *et al.*[51] Both metrics are generated from measuring the end state of a reservoir evolving under many input streams consisting of uniformly distributed random data between 0 and 1. Singular value decomposition is used to estimate the rank of the resulting state matrices, giving an idea of the number of linearly uncorrelated sequences in the output. Where the two metrics differ is in the correlation between input sequences: In KR, the input sequences are entirely uncorrelated, whereas in GR, the final 3 input data are fixed for every sequence. Conceptually, KR measures the degree of nonlinearity between measured node states for uncorrelated inputs, while GR measures the reservoir's ability to converge to similar states from different past trajectories.

While these metrics represent an effective method for measuring phase transitions between regimes of reservoir operation, they are sensitive to a range of parameters that are rather arbitrarily defined across the field. For example, the output of both metrics is bounded by the number of nodes in a system, and some small threshold is chosen to determine which singular values constitute matrix rank, which can drastically skew results- especially in physical systems where experimental noise contributes to the high output matrix ranks.

Another commonly employed metric for reservoir computers is linear memory capacity (MC)[52]. This metric directly assesses the reservoir's ability to reconstruct delayed inputs to the system from the system's current state. The system is provided with uncorrelated inputs randomly sampled from a uniform distribution, and reservoir state is measured as the system evolves. A simple linear readout is then trained to reconstruct the inputs over a range of delays, and memory capacity evaluated via the following equation (3.7):

$$(3.7) \quad MC = \sum_{k=1}^{N} MC_k = \sum_{k=1}^{N} \frac{cov^2(u_{i-k}, y_k)}{\sigma^2(u_i)\sigma^2(y_k)}$$

where u_i represents the input at time i, u_{i-k} represents the true input at delay k, and y_k represents the reconstruction of input at delay k. While linear MC evaluates the reservoir's ability to reconstruct exact reconstructions of past inputs, often tasks require a nonlinear representation of past inputs instead. To evaluate this, the linear MC metric has been extended to include quadratic MC, cubic MC, et cetera, by simply raising the delayed input to the desired power.

Another approach used to evaluate the processing power of a dynamical system is the information processing capacity (IPC), introduced by Dambre *et al.* in 2012[53]. Functionally, this metric can be described as the total number of linearly independent output functions a system can generate from its input stimuli. It draws upon similar mathematical concepts to the memory capacity measurements described above but combines the different polynomial terms to give a more complete description of the memory of a system. This approach looks at many different nonlinear combinations of past inputs, and estimates each the capacity for each component C_m via equation (3.8)

$$(3.8) \qquad C_m = 1 - \frac{min_{w_i} MSE[y_{IPC,k}^m, y_{RC,k}]}{mean(y_{IPC}^m)^2}$$

where $min_{w_i} MSE[\gamma^m_{IPC,k}, \gamma_{RC,k}]$ represents the minimum mean-squared error when reproducing the information process $\gamma^m_{IPC,k}$ from the true inputs and the reservoir's prediction $\gamma_{RC,k}$. As the combination of past timesteps becomes increasingly nonlinear, the capacity of each component tends towards 0. Therefore, the generation of new components can be reasonably stopped once the calculated capacities tend towards zero. The total IPC is then calculated via summing over all components (equation 3.9):

$$C_{total} = \sum_{m=1}^{M-1} C_m \quad (3.9)$$

Recently, more novel metrics have been proposed by Love *et al.*[54] in 'nonlinearity' and 'complexity' measures. Like MC, these metrics are constructed by measuring the system state as it evolves under randomly generated inputs from a uniform distribution. The nonlinearity metric is measured by using a Volterra series[55] (can be roughly thought of as the dynamical system analogue for the relationship between a function and a Taylor series) to approximate the transformation the reservoir provides, and evaluated by calculating the R^2 correlation coefficient between the true signal and a truncated Volterra series consisting of only linear terms. The complexity measure is almost identical the KR metric proposed earlier in terms of evaluation, though instead of measuring the end state of all nodes after various time sequences, the evolving state of nodes over input of length equal to the number of nodes is used to produce matrix M with which rank is estimated.

While these novel metrics have yet to see widespread implementation in the community, both can be calculated using the same signal as used for MC, allowing rapid evaluation of different computational properties for physical systems where the rate of data throughput is a limitation. However, the set of metrics still lacks a method of evaluating the ability of the reservoir to generalise information. Nevertheless, they are useful in drawing a more comprehensive map of reservoir performance and the range of computationally distinct regimes of operation a reservoir platform can achieve.

One tool that has been designed to optimise the *cha*racterisation of *r*eservoir *c*omputers is CHARC[56]. This platform utilises genetic algorithms to explore the state space of a reservoir computer with respect to parameters that control the input scaling or the parameters of a reservoir's design/response. Where this platform gains its power is in its novelty search function- the objective that the genetic algorithm 'optimises' is separation between the space each initialisation occupies in metric space. This enables the algorithm to push the boundaries of the transforms a reservoir can achieve and can help to evaluate computational range of a reservoir computing platform. While the initial paper only explores over KR, GR, and MC, the software is suitable for expansion to cover other metrics. However, to employ CHARC as a useful search tool, models of an underlying dynamic system must be parallelisable and quick to evaluate.

3.4.5- 'SpaRCe' Algorithm for Sparse Reservoir Computing

While the simplicity of training offered by ridge regression is often an attractive property of RC, novel learning rules have been developed to increase the performance of RC platforms, and to mitigate shortcomings such as catastrophic forgetting[57,58]. One such example is the SpaRCe algorithm[59], a bio-inspired technique that uses learnable thresholds on output nodes to suppress certain outputs and induce sparse representations of information, shown in figure 23.

The algorithm can be described via the competition of two opposing update rules for the threshold: the first update rule promotes activity in neurons that are contributing to correct classification, while the second inhibits activity in correlated neurons. This approach rewards correct classification whilst reducing redundant activations and leads to fewer nodes being active in response to a given input at the output level. This specialisation of output nodes was shown to be beneficial for two standard tests prone to the effects of catastrophic forgetting: learning classes of information sequentially and learning distinct fixed permutations on data sequentially. The sparsity induced by the algorithm allows specialised nodes to learn the sequentially provided information, reducing the interference caused by learning this information sequentially, as correct classification is independent of many of the output nodes.

One of the key benefits of this approach is that it can mimic the effectiveness of the inclusion of an additional hidden layer in the readout with significantly fewer parameters: the number of additional learnable parameters is equal to the number of nodes for SpaRCe, meaning it scales linearly with reservoir size. On the other hand, an additional hidden layer's trainable parameters scales quadratically, requiring more computational cost for equal performance gains. Additionally, the increased stability of learned representations might prove useful in hardware, where the algorithm can suppress the contributions of noisy state readouts that have limited contribution towards correct classification.

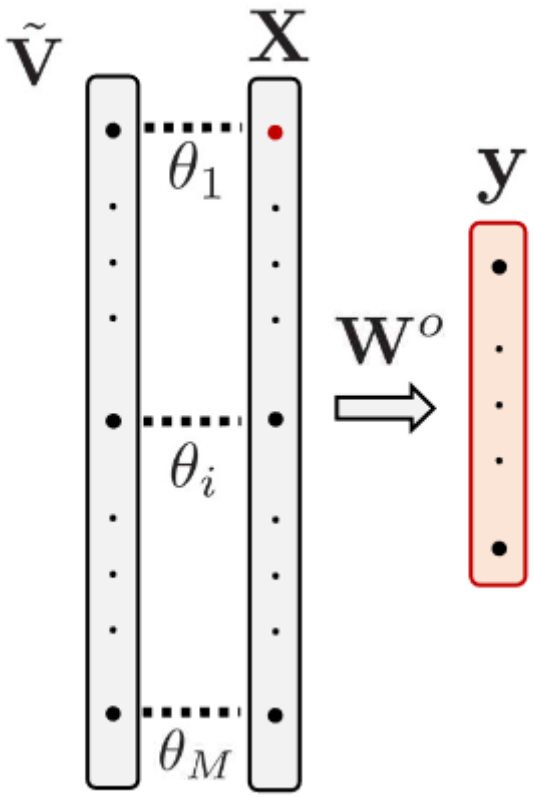


Figure 23- Schematic diagram of the thresholds, showing one-to-one connectivity between raw reservoir state vector $\tilde{V}$ and X, connected by learnable thresholds θ_i. Taken from[59]

3.4.6- Hierarchical Reservoir Computers

One common difficulty when designing/optimising reservoir computers is matching the timescale of dynamics of the reservoir to those of the task. This becomes increasingly difficult when the task being solved evolves with dynamics of multiple timescales. One possible method for building a model with multiple distinct reservoirs initialised with different hyperparameters to create subsystems that evolve with different dynamic characteristics. In the work of Manneschi *et al.*[60], the authors compared how hierarchical ESNs (where the output of one ESN feeds directly into the input of a second ESN) to parallel ESNs (where input is passed directly to two unconnected ESNs). It was found that the hierarchical reservoirs outperformed the parallel reservoirs, and that the ideal reservoir configuration was a quickly responding (short timescale) reservoir to receive the input initially and provide a higher dimensional representation, followed by a slowly responding (long timescale) reservoir which acts as a secondary nonlinear filter on the first representation and provided long-term memory.

Additionally, an online learning mechanism is proposed which allows the initialisation parameters to be automatically optimised for a given task instead of selected from a grid search. Since learning the initialisation parameters had a direct influence over the temporal parameters of the ESNs, training must also occur with respect to time, as with RNNs. Rather than use the hugely computationally demanding backpropagation through time (BPTT) algorithm, the authors opt for an approximation – e-prop[61] – which shows good performance in optimising the network while mitigating computational cost.

The ability to optimise the leakage rate for connected reservoirs to improve performance improves their suitability for applications by non-experts, as usually the parameter selection process largely determines the performance of an ESN, requiring knowledge of both the natural timescales of the task at hand, as well as how to achieve similar timescales within the ESNs. However, for physical reservoirs, the timescales of the underlying system may not be as easily controlled as in ESNs, especially in cases where the system evolves with fixed dynamics post-manufacture. An analogue suitable for online learning of timescales in physical systems may be achieved through controlling the input rate of input stimuli, though this has not been demonstrated in literature thus far.

3.4.7- Reservoir Computing with Heterogeneous Neurons

An alternative approach that has been proposed for achieving a more diverse range of timescales within a reservoir system is through the employment of heterogeneous neurons in the reservoir layer. Instead of having a globally defined leakage parameter for the network, Tanaka *et al.* proposed a reservoir architecture where each node had its own distinct leakage parameter[62]. In their work, they show that the range of timescales achieved at the reservoir's output are significantly expanded through this distribution of leak rate, and that this considerably improved performance in the case of chaotic time series prediction tasks which evolve with a mixture of short and long-term dynamic timescales.

This approach may be more viable for physical reservoir computers, as it would be possible to produce heterogeneous nodes via manipulations of, for example, device geometries or material parameters. This would enable the engineering of a more dynamically complex reservoir system, though would also require synaptic connections between the heterogeneous nodes that resemble the standard ESN model. One interesting study may be to combine this approach with the revolving neurons reservoir

approach by designing the distinct dynamical nodes to have different properties, mitigating the need for multiple synaptic connections but maintaining the diversity of network dynamics.

3.5- Physical Reservoir Computing

The field of reservoir computing with physical substrates has grown rapidly over recent years. Whilst this thesis will cover the applications of a specific magnetic system for RC, many of the foundational works were performed on non-magnetic substrates and contain many useful advances that are applicable to material computation more generally. The following section will introduce the foundational works in material computation, before reviewing the current state of the art for magnetic systems and critically evaluating the proposed systems' relative strengths and weaknesses.

3.5.1- The Liquid Brain

The first demonstration of RC with a physical substrate occurred even before the phrase 'Reservoir Computing' was coined to describe computation using a linear readout on an untrained dynamic system. In their 2003 paper, 'Pattern Recognition in a Bucket'[63], Fernando and Sojakka used the inherent interactive dynamics of propagating water waves as a basis for nonlinear computation, perhaps taking the name of Wolfgang Maass' 'liquid state machine' a little *too* literally. While buckets of water, shown in figure 24, may not be replacing CMOS circuits in devices any time soon, the conclusions of this paper are profound: the complexity of dynamical systems can offer higher-dimensional representations of data with similar effectiveness to algorithmic RC, with the physical relationships that govern the system's response providing continuous complex transformations of data inherently.

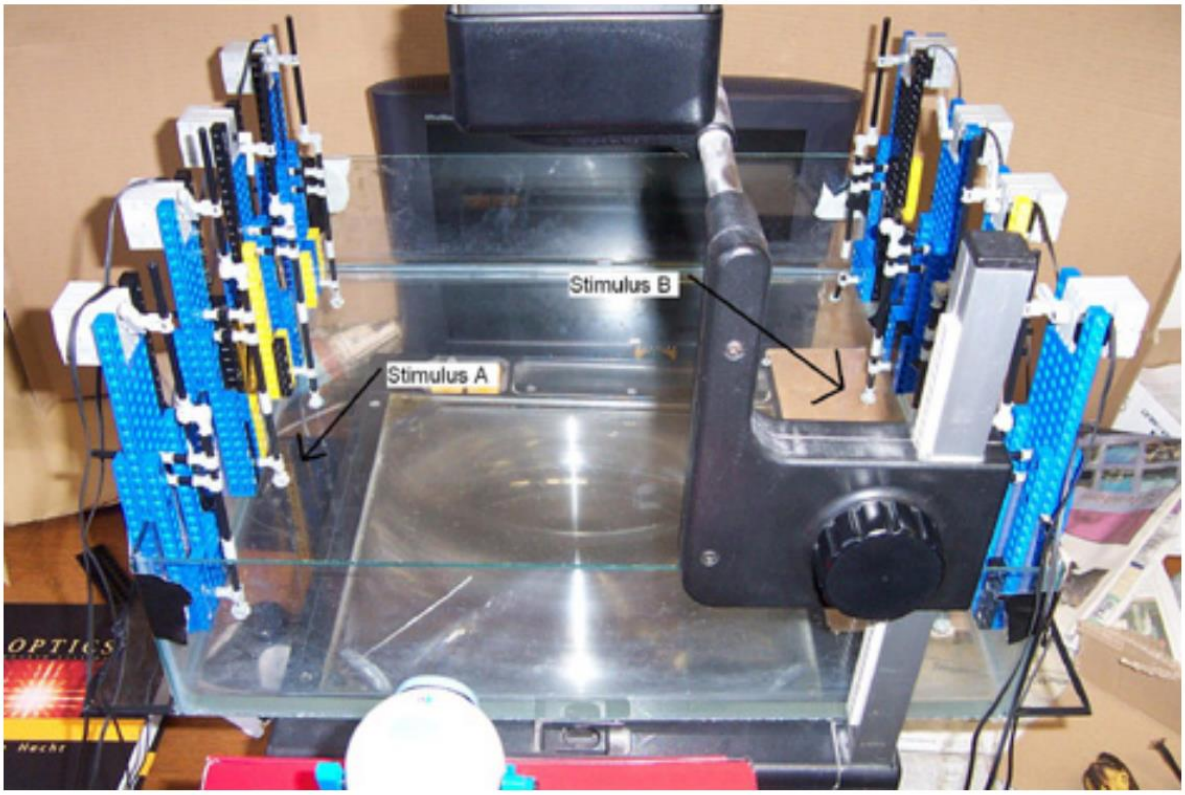


Figure 24- Experimental setup of 'the liquid brain', showing paddles used for input stimuli as well as the camera above used as 'readout'. Taken from[63].

The authors explain that the key underlying properties behind the water waves' ability to perform computational operations is their ability to provide a complex, higher-dimensional representation of input via the interactions that are intrinsic to the physical substrate. They assess these properties in their system by evaluating the neuronal complexity of the medium, showing that even with many closely spaced inputs, the system remains highly complex- though not chaotic.

3.5.2- Analogue Electronic RC

At the heart of all implementations of physical RC is analogue transformations of input signals provided by a complex physical system. One method for achieving these complex physical systems is to use well-understood electronic building blocks such as resistors, capacitors, and diodes to implement complex, transient circuit responses. Soriano *et al.* explored a range of different electronic hardware implementations in order to evaluate the role nonlinearity, noise, and connectivity played on the information processing capability of their RC platforms[64]. They found that a single oscillator with delay-based dynamics by itself lacked the properties required to handle complex time signals. However, through use of time-multiplexing approaches and delayed feedback of the system's state, the computational power of the system was drastically improved. This is important for the application of a whole range of physical RC systems, since relatively simplistic dynamical systems can be expanded to provide useful computational platforms with appropriate external RC architectures. An example of how a Mackey-Glass oscillator can be implemented in analogue electronic components is shown in figure 25.

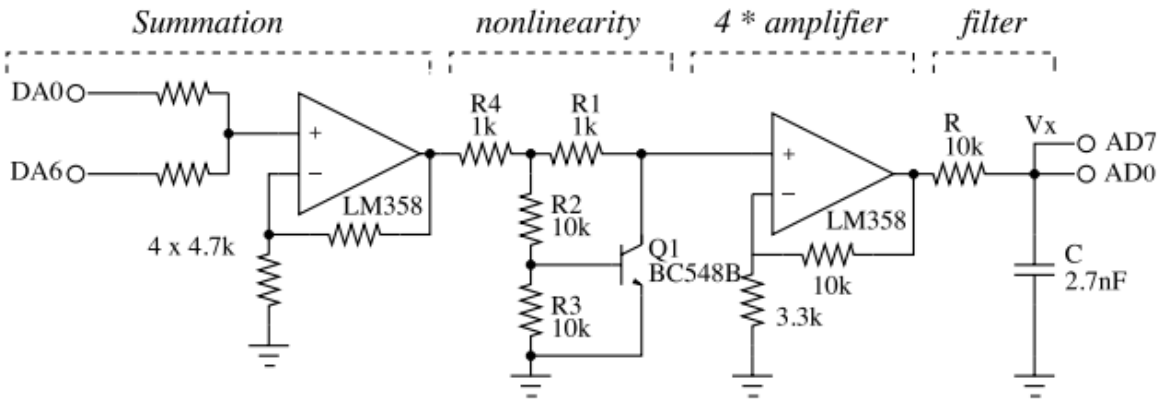


Figure 25- Circuit diagram of the analogue electronic components used to implement the Mackey-Glass delayed differential equations as the basis of a dynamical node. Taken from[65].

Other works in the area used multiple similar delay-coupled oscillators to produce structures resembling deep neural networks[66] showing improved computational performance on complex nonlinear problems, or chaotic circuit components[67] to provide additional dynamic complexity to the underlying system in order to solve complex tasks. These findings show that two contrasting approaches have typically been used to provide computational functionality with electronic devices: either using complex system dynamics or complex network structure to expand the functionality of the physical systems.

3.5.3- Memristive RC

The memristor is a long theoretically proposed[68] and relatively recently implemented[69] circuit component whose electrical resistance varies dependent upon the past inputs provided to it. Typically, memristors consist of either an oxide layer(s) or a polymer sandwiched between electrode terminals. Figure 26a-c shows a schematic diagram of a memristor, the manufactured device, and the relative percentage of constituent elements versus position for the manufactured device.

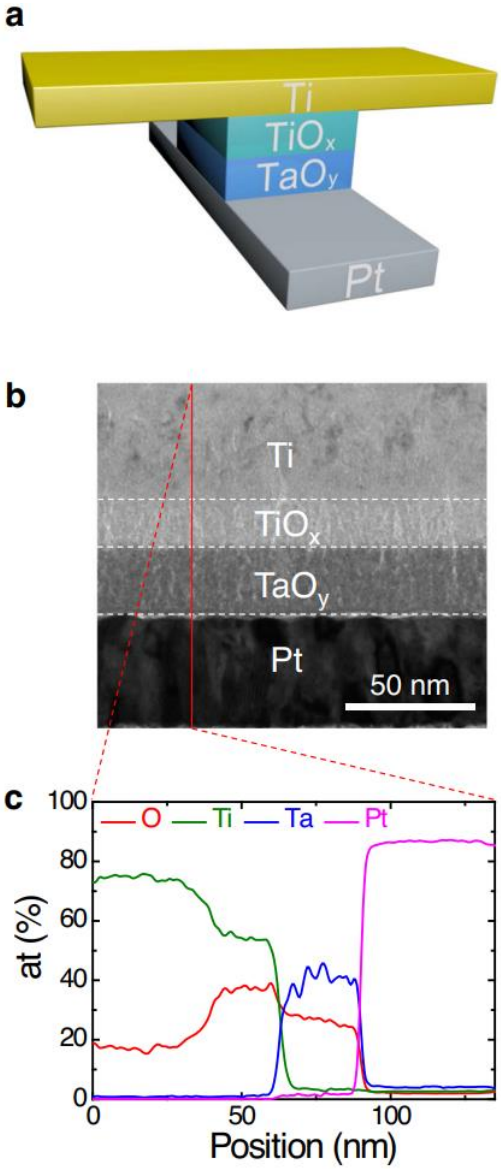


Figure 26-(a) Schematic overview of a memristor stack. (b) SEM micrograph of the manufactured device. (c) Elemental distribution profile taken from energy-dispersive spectroscope. Taken from.[70]

Under the influence of an electrical field, there is a physical change in the material due to the migration of ions within the material which causes a change in resistance, hence making the current resistance dependent upon the previous inputs to the system. This is a promising building block for reservoir computers due to the temporally complex transformations offered by each individual memristor[71]. Whilst memristors have often been used to create ordered networks resembling the functionality of ESNs[72–75], memristor crossbar arrays are useful for vector-matrix multiplication, allowing them to be implemented as all-analogue readouts for RC systems, with the variable resistance providing a

tuneable parameter for training[50]. While memristors form computationally useful, variation-tolerant[76] RC platforms, their potential for use in devices is currently held back by their reliability, scalability, and limited lifespan of operation[77].

3.5.4- Optoelectronic RC

Another family of devices that have promising properties for reservoir computing are optoelectronic components. These devices make use of the nonlinear properties of standard circuit components, the interference interactions that occur between light waves, and easily implementable delayed feedback methods made possible through use of fibre-optic spools and beam splitters. These properties mean optoelectronic platforms are especially harmonious with the single dynamical node paradigm[47] introduced earlier, providing a simple path to realisation of physical RC using optical systems. A schematic diagram for an optoelectronic system is shown in figure 27.

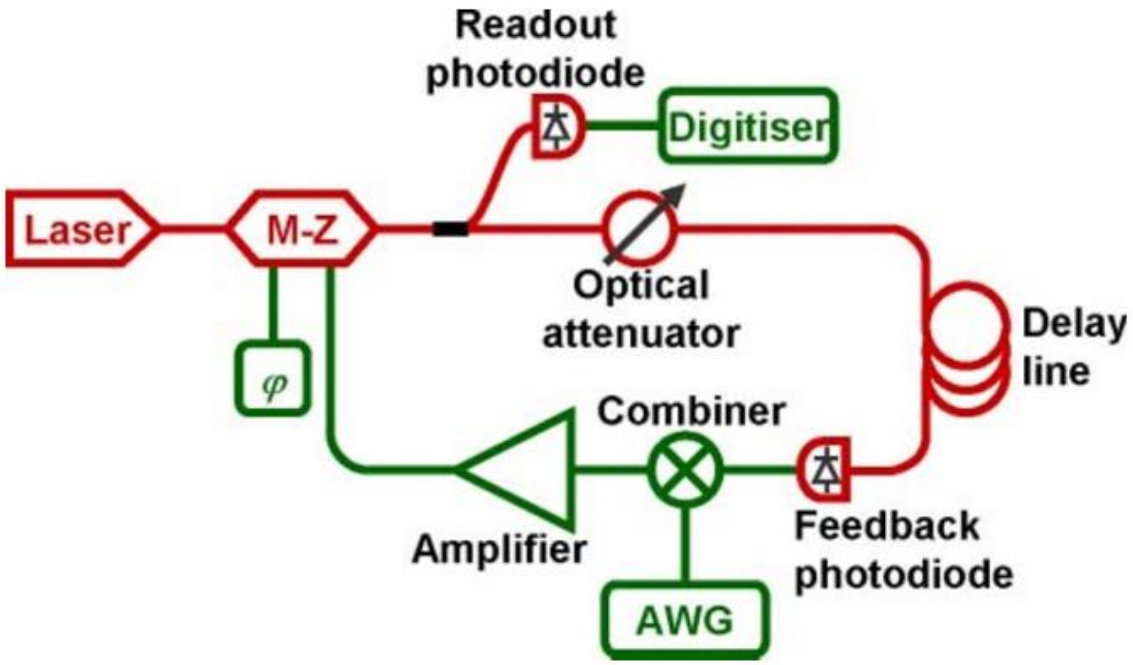


Figure 27- Schematic diagram of an implementation of optoelectronic reservoir computing using standard off-the-shelf electronic components. Taken from[49].

Reservoir computing using optoelectronics was outlined as part of numerical studies as early as 2008, with Vandoorne *et al.* showing the similarity between tanh activation functions and the power output of semiconductor optical amplifiers as input power becomes saturated[78]. The proposed networks consist of distinct dynamical nodes with nonlinear activation, connected in an ordered 'waterfall' structure reminiscent of the FF-ESNs described earlier[44].

However, the relatively high power required to saturate these amplifiers to produce nonlinearity meant this method was not preferred for the first implementations of optoelectronic RC[79]. Instead, a linear network was constructed using simple waveguides, splitters, and combiners to emulate the role synapses in accumulating input activations from prior nodes and propagating activity forwards as in standard neural networks. Whilst seemingly lacking the critical RC property of nonlinearity, this is mitigated by using a nonlinear photodetector, with readout proportional to the square of the activity.

The system performed competitively for spoken digit classification, though extensive study into the effects of removing nonlinearity from the internal reservoir processes were not conducted.

Another hardware implementation of optoelectronic RC came from Paquot *et al.*[49], featuring nonlinearity from a simple Mach-Zehnder interferometer[80] (which modulates amplitude according to the phase of light waves) coupled with an optical delay line to provide a fading memory of past states, realising the single dynamical node paradigm with delayed feedback as in[47]. Several tasks are performed, including distinguishing between sine and square waves, spoken digit recognition using the TI-46 database, nonlinear channel equalisation, and nonlinear auto-regressive moving average (NARMA) series reconstruction. This demonstrates a broad range of computational capability, but solely in a task-based environment. As one of the earliest high-profile implementations of physical RC with device-feasible hardware (i.e., not a bucket of water). Their work has since set a precedent for both the implementation methods of physical computing platforms under the SDN paradigm, as well as a precedent for the specific tasks used for demonstrating computation in material systems.

3.6- Magnetic RC

The interfacing of magnetic systems with conventional electronics for computing purposes has been long established due to computationally useful properties offered by magnetic devices: the inherent non-volatility of magnetic domain structures provides pathways for memory storage, and transfer-torques/magnetoresistance effects provide avenues for data writing and reading respectively. These properties meant that most existing implementations of magnetic devices for computation have focussed on data storage, with information encoded as a binary in the spin degree of freedom in technologies such as the magnetic hard disk drive (HDD), and magnetic random-access memory (MRAM).

While these technologies exploit the inherent memory effects present in magnetic materials, another ubiquitous property of magnetic materials, nonlinearity, is not leveraged in the data storage examples. However, the emerging field of magnetic reservoir computing utilises both properties in tandem. Nanomagnetic/spintronic platforms exploit a plethora of phenomena, ranging from transient dynamics and the interference behaviours present in interacting spin-wave systems, to complex metamaterial systems with spatially distributed responses such as skyrmion textures and artificial spin ices.

In their recent perspective on magnetic reservoir computing[81], Allwood *et al.* introduce a taxonomy for classifying magnetic RC platforms across two categories which are critical to their application as computing platforms: the input/output dimensionality (IOD) of their reservoir representations (1D- single scalar output for each input, or ND- an N-dimensional expansion for each input), as well as the timescales of their dynamical responses (DR) (LLG- dynamics on the order of 10s to 100s of nanoseconds resulting from relaxing precessional behaviours governed by the LLG equation, T- dynamics governed by thermal processes, ranging from seconds to milliseconds, and D- driven processes where the timescales of evolution match those of the driving stimulus). This provides a basis for evaluating proposed RC systems both in terms of their current and potential implementations, as well as for establishing what types of computing capabilities may be achievable for a given platform.

The following sections will utilise an adaptation of this taxonomy to facilitate the comparison of proposed nanomagnetic RC platforms. The same axes of IOD and DR will be used, though the subcategories for IOD are adapted slightly to make the differences between different implementations

more apparent. Instead of a single value distinguishing between single and multiple dimensionalities for both input and output together, here we instead consider this for both the input and the output of the reservoir, giving rise to three classes: SISO (single in, single out), SIMO (single in, multiple out), MIMO (multiple in, multiple out).

3.6.1- Spin-Torque Nano-Oscillators

Seminal work on the implementation of nanomagnetic platforms for RC revolved around spin-torque nano-oscillators (STNOs)[48]. The underlying system at the heart of these demonstrations is a magnetic tunnel junction (MTJ), consisting of two thin ferromagnetic layers, separated by a thin insulating barrier. The two ferromagnetic layers are designed to exhibit different properties: a 'free' layer in which magnetisation is free to change direction, and a 'fixed' layer, where magnetisation is locked in a given direction via interactions with an adjacent antiferromagnetic layer.

When DC electrical currents are provided to the multilayer, spin-torque effects cause the magnetisation of the free layer to precess. These precessions can be measured via the tunnel magnetoresistance (TMR) effect, causing changes in the electrical resistance across the MTJ which varies nonlinearly with the magnitude of the DC input current. This nonlinear variation coupled with short-term history-dependent behaviours arising due to damped precessional motion meets the broad RC criteria of nonlinearity and fading memory. This has allowed implementation of STNOs as reservoir computers using simple inputs of amplitude modulated DC currents, and readouts using power diodes to give a scalar output from the oscillating resistance. The single input/output dimension, as well as dynamics governed by LLG processes, means STNOs have a reservoir categorisation of IOD-SISO and DR-LLG. Figure 28 shows an overview of the STNO, with details on its response to input.

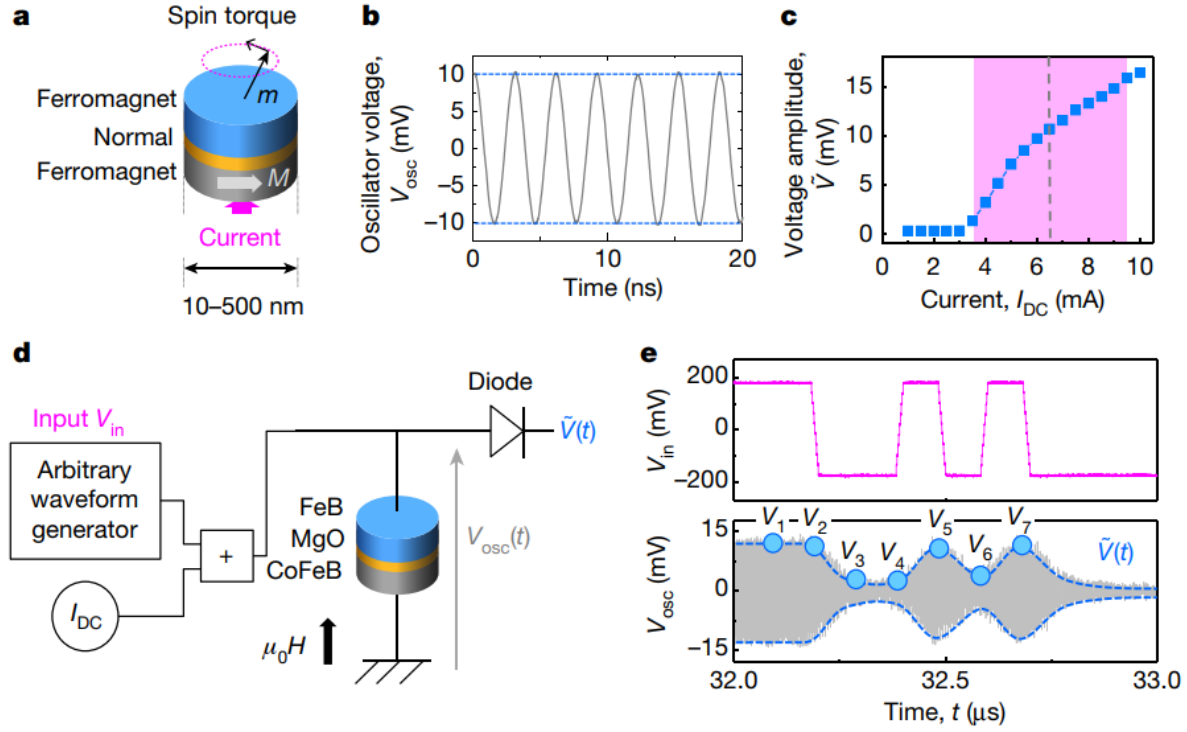


Figure 28- (a) Overview of a spin-torque nano-oscillator. (b) Oscillator voltage when driven continuously with a DC current. (c) amplitude of oscillator voltage at steady state when driven with DC current of varying amplitude. (d) Schematic for inputting data to the STNO and reading out via a power diode. (e) Dynamic response when oscillator is stepped between different continuous DC levels. Taken from[48].

The first implementation of this system for RC used the SDN reservoir paradigm without external delay lines to perform effective spoken digit recognition[48], and later studies by the same group showed the enhanced temporal properties gained via inclusion of delayed feedback[82], allowing performance of temporal pattern recognition. Other groups have also used numerical simulations to evaluate memory capacity and parity-check capability for binary inputs[83], and later with physical devices[84] without external feedback. Both studies show memory capacities of 2<MC<3 for systems without delayed feedback, highlighting the limited memory capacity offered by a single oscillator.

Other RC configurations of STNOs have been proposed, including reservoirs consisting of arrays of interconnected STNOs[85]. Here, the network is analogous to ESNs constructed from leaky-integrator neurons, with similar changes to memory compared to ESNs observed when varying both network size, as well as the input-rate with respect to an oscillator with a fixed timescale of decay- a property analogous to the leak-rate of standard ESNs.

At a device level, STNOs boast a range of attractive properties, including proven scalability down to 10s of nms per oscillator, low power consumption on the order of 1 uW per oscillator, and perhaps most significantly their ease of interfacing with existing electronics offered by all-electrical input/output paradigms and large signals created via TMR effects. Additionally, multiple STNOs can be connected to one another to form larger networks, creating IOD-MIMO reservoirs, expanding their computational capability. However, since the internal dynamic processes are determined by relaxation of precessing spins, only dynamic timescales faster than these processes are achievable. This does not cause problems for processing of data that has already been gathered, though renders direct processing of data from real-time sensory inputs impossible when intrinsic timescales are longer than those of the STNO.

3.6.2- Other Magnetic Nano-Oscillators

Following the success of STNOs, other magnetic oscillators have been proposed for RC applications. In the numerical studies of Ababei *et al.*[86], a single DW is locally confined by the energy barriers presented by a pair of anti-notches. The DW is then driven by a sinusoidally varying magnetic field at 500 MHz (DR-LLG), causing oscillations in the DW's position. Reservoir output is taken as the scalar DW position (IOD-SISO). Under low driving fields, the DW motion is harmonic within a single energy minimum of the double-well energy potential. As driving field increases, the DW gains enough energy to overcome the small energy barrier between the two wells and motion tends to bifurcate, inducing chaotic dynamics. As field increases further, motion once again becomes harmonic, though over the entire range of the potential landscape. The energy landscape, as well as a schematic of the oscillator, are shown in figure 29.

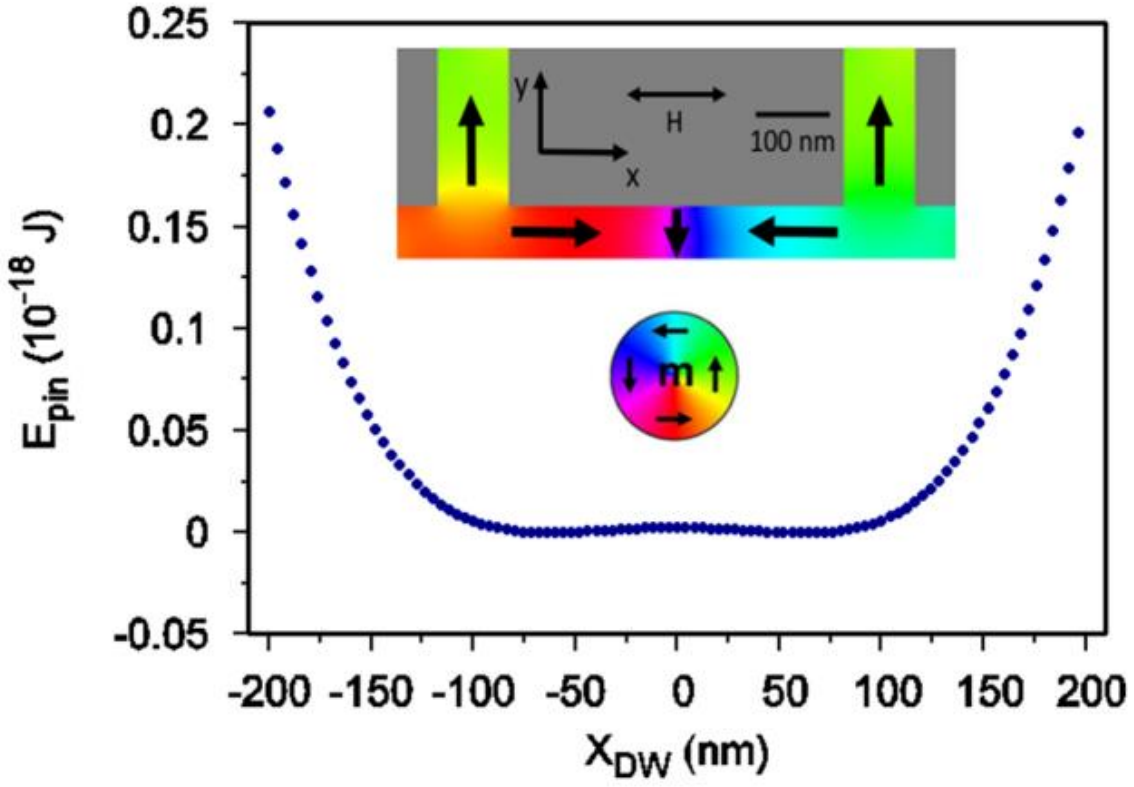


Figure 28 – Schematic diagram of the domain wall oscillator, with magnetisation shown by the colour wheel. Graph shows the pinning energy of the oscillator with respect to domain wall position. Taken from.[86]

The authors evaluated computational performance under a range of driving conditions spanning all three regimes for both temporal and non-temporal classification problems. Whilst performance centred in the chaotic regime produced generally poor results, it was shown that the optimal performing regions were located near the transitions between the regimes. This alludes that the highly nonlinear response generated by the change between regimes is promising for RC applications. The authors also demonstrated both handwritten and spoken digit recognition classification, with performance exceeding that of a linear control network, highlighting the role of the reservoir in improving classification accuracy.

Whilst the DW oscillator system simulated the effects of applied fields, similar behaviours could be driven using spin-torques or spin-orbit torques, improving device viability and reducing energy cost. However, the readout of position would remain challenging due to the high frequency of oscillation, though similar approaches to that of the STNOs may be taken via inclusion of spin-valve stacks to utilise TMR effects, giving resistance readouts proportional to the DW position used in this study. The geometric dependence of the energy landscape opens up a high degree of tunability through control of the size and separation of the anti-notches. This would allow different dynamics to be engineered into different oscillators, which can then be tailored for given tasks or used simply in parallel to produce a richer representation of input information.

Other similar demonstrations have been made using geometrically confined skyrmions (a topologically protected magnetic soliton), driven via spin-Hall interactions between the skyrmion and current pulses[87]. Image recognition was performed on a binarised version of the MNIST database. However, in

this study, the representation of the handwritten digits into 196 different binary time signals of length 4 is atypical within literature and a control network was not established to evaluate the role of the reservoir.

Another proposed nano-oscillator system for neuromorphic computing is the spin-Hall nano-oscillator (SHNO)[88]. This system features nano-constrictions in a thin-film multilayer consisting of a ferromagnetic layer on top of a heavy-metal underlayer, which uses spin-Hall effects to generate pure spin currents into the ferromagnetic layer. The constrictions result in large charge densities at the location of the constriction, producing subsequent spin currents which cause magnetic moments in the ferromagnetic layer to steadily precess. Arrays featuring multiple adjacent SHNOs have been shown to exhibit mutual synchronisation, increasing the complexity of the system response. While the authors demonstrate and expand upon the computationally useful properties exhibited by the arrays, thus far there have been no demonstrations of computation with SHNOs in a task-based environment.

3.6.3- Spin-Wave Based Magnetic RC

Like the early 'computation in a bucket' proposals featuring the interactions between propagating water waves, a spintronic analogue utilising spin waves have been proposed. Spin waves are phase-coherent, collective excitations of magnetic moments, which are generated through resonance-like effects when magnetic materials are driven at microwave frequencies (DR-LLG). When propagating spin waves collide, they interact similarly to water waves, with constructive/destructive interference patterns depending on the relative phase of the colliding waves.

In their 2018 study, Nakane *et al.*[89] propose devices consisting of Yttrium Iron Garnet (YIG) films acting as a medium for spin-wave propagation chosen for their low Gilbert damping ($\alpha = 0.001$ in[89]), and a magneto-electric coupling layer to allow generation and measurement of said spin waves, shown in figure 30. Both input and readout connections are distributed randomly in the film, with a spatially distributed response in the film due to the interference interactions (IOD-MIMO). The authors applied the simulated system to perform estimation of the length of input time signals, however the role of the reservoir is difficult to estimate due to lack of comparisons with alternative systems.

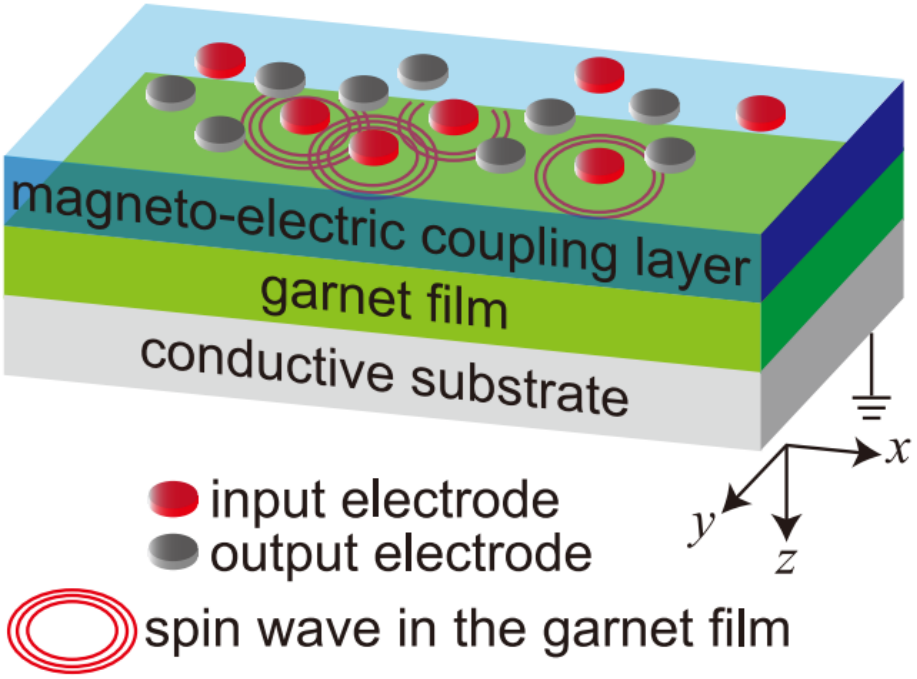


Figure 30- Schematic diagram of the spin-wave reservoir computing platform introduced in[89], showing input electrodes in red, with readout electrodes in grey, providing spatially dependent readout.

Similar proposals by Papp *et al.*[90] again use YIG films for RC, however with different mechanisms for creating inputs and interference. Here, input is instead provided as a uniform linear wavefront generated by an RF waveguide, and the resulting spin-waves are scattered using a randomly initialised 'checkerboard' pattern of magnetic nanodots with perpendicular magnetic anisotropy (PMA). Depending on whether the PMA dots have magnetisation pointing into or out of the plane of the checkerboard, the dispersion of the propagating spin waves is locally altered, leading to complex interference patterns on the other side of the scatterer where multiple readout antennas are located (IOD-SIMO). Depending upon the strength of input excitations, the degree of nonlinearity of the reservoir system can be controlled.

To evaluate the computational properties of the different system configurations, task-independent metrics of KR and GR are used, with notably different performance across driving conditions. This highlights the tunability of the system to produce a range of reservoir computers with different computational properties on a single sample. The authors then go further, using genetic algorithms to evolve the checkerboard configuration to maximise KR. Under the limitation of nanodots with a single magnetisation strength in a binary orientation, the evolutionary algorithm can only make minor improvements to KR compared to the random structure. However, when the input field scaling (a continuous parameter) is trained, a significant improvement is observed. This is attributed to the increased degrees of freedom offered by continuous parameter space and may extend to other magnetic platforms; high degrees of continuous tunability likely makes for more versatile computational properties than restricted binary implementations.

3.6.4- Artificial Spin Ices

Artificial spin ices (ASIs) are magnetic metamaterials consisting of large networks of interacting nanoislands arranged in regular lattices[91]. Each nanoisland is of sufficiently small size and high aspect ratio that shape anisotropy leads to the formation of single domain states. Dipolar interactions

between islands at each vertex lead to varying degrees of energetic favourability for given orderings of converging or diverging spins, which in turn leads to complex emergent behaviours in ASIs when the systems are driven via thermal annealing or with external magnetic fields.

Due to the spatially distributed nonlinear responses arising from these local dipole interactions, and history-dependent behaviours as the system passes through low or high energy states, ASIs have natural pathways for forming spatially distributed RC platforms with high dimensionality. Additionally, the driven nature of their dynamics (DR-D) allows for operation at a wide range of timescales that are unachievable for the previously introduced systems. The high degree of geometric tunability available to ASIs (size, pitch, angle of islands, alternative lattices etc.) is also very promising for RC applications.

Jensen *et al.*[92] have used simulations to model the behaviour of ASIs in regular square grids in response to applied magnetic fields (IOD-SIMO). In their work, they show that the ASI occupies a broad range of unique states depending upon the sequence of binary inputs applied to the system, which is also highly dependent upon the magnitude of the applied field strength and frequency.

They extend this work to perform reservoir computing with the ASI system with similar input encoding for 'pinwheel' geometries[93] (square arrangement with islands rotated 45 degrees about their centre). To evaluate computational properties of the ASIs, task independent metrics in KR and GR are calculated over a range of driving and coupling parameters, showing tuneable computational behaviour. The authors also explored the effects of output granularity by averaging over increasingly larger areas of the array. Given the emergent behaviours of ASI systems, this output granularity is key since local emergent ordering occurs over various length scales. As output granularity increases (i.e., more magnets in each average), the generalisation rank of the reservoir computer decreases (better generalisation), with an associated but lesser reduction in kernel rank. This represents tuneable control over both the separation and generalisation properties of the ASI reservoir. This tunability is likely transferrable to the magnetic ring array systems, where similar emergent behaviours will lead to more general aggregate responses over longer length scales.

More recently, physical realisations of RC with ASIs have been performed using artificial spin-vortex ices (ASVIs, spin ices where vortex core configurations, as well as bistable macrospin states, are available for the islands)[94], shown in figure 31. The arrays consist of alternating rows of low and high aspect ratio islands, which favour macrospin and vortex states respectively. The system is initialised into an all-macrospin state, and the array exhibits ratchet-like behaviour when successively driven with linear magnetic field pulses as the wider bars tend towards vortices. The relative rates of vortex-to-macrospin or macrospin-to-vortex transitions depend upon the applied field strength.

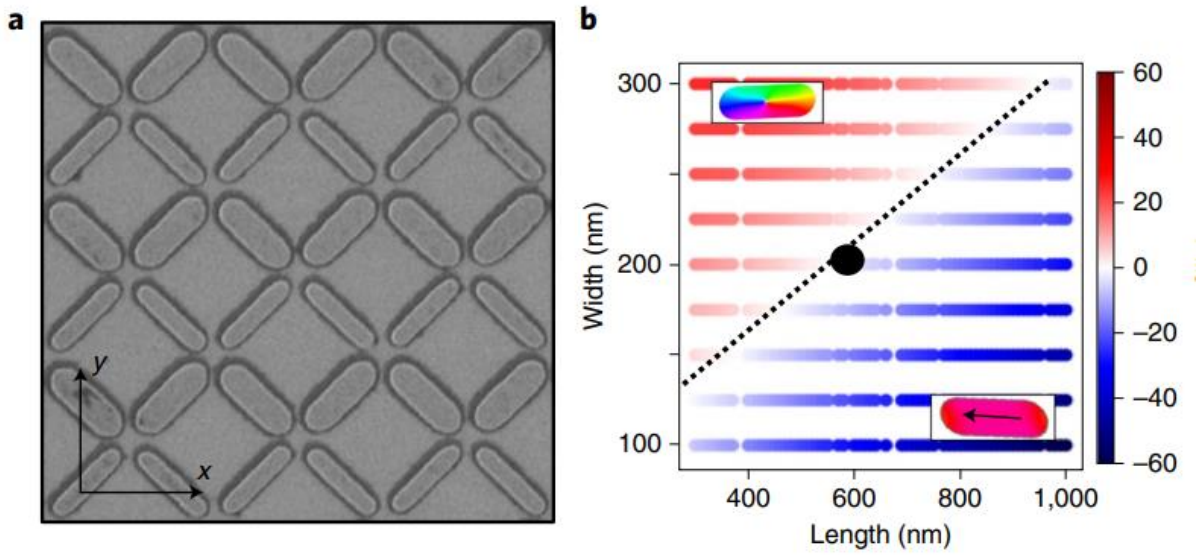


Figure 31- (a) SEM micrograph of artificial spin-vortex ice. (b) Energetic favourability of the macrospin and vortex states of a bar as width and length is varied. Black dot shows selection of thicker bar width in ASVI.

The RC paradigm applied to this system involves encoding data within the strength of the applied field pulses, and then evaluate the reservoir state via a process the authors refer to as 'spin-wave fingerprinting'. This readout technique uses the derivative of spin-wave dispersion with respect to perturbation field (dP/dH) across a frequency spectrum, with the spin structure of islands in the array giving rise to spin-wave modes with different resonant frequencies. By evaluating the magnitude of dP/dH across a range of frequency bins, an experimentally viable method for producing an IOD-SIMO reservoir can be constructed. This method addresses common limitation that electrical discontinuity of ASI presents when establishing an experimental readout and is extensible to other nanomagnetic platforms where system state results in varying spin-wave modes in the frequency domain. Similar approaches would likely be useful for the ring arrays, as the different domain configurations expected in each ring will each have distinct spin-wave modes, leading to a method for measuring relative populations of each configuration.

Using this paradigm, the authors demonstrate computational functionality by performing signal processing tasks, namely nonlinear signal transformation and Mackey-Glass series future prediction. The nonlinearity within the high-dimensional transform provided by the ASVI is showcased by the large decrease in mean-squared error provided by the reservoir compared to the control input. The performance of Mackey-Glass prediction up to 10 steps into the future initially seems to suggest long-term memory in the system. However, the plots provided in supplementary material, showing an oscillating error profile, alludes to this being a local minimum in accuracy generated by a nonlinear transform of the current input (the Mackey-Glass system is semi-periodic) rather than memory of past inputs and learning of the underlying dynamics. It is therefore difficult to assess the extent of the system's memory without more direct measures such as linear memory capacity.

References:

1. Hayward, T. J. *et al.* Direct imaging of domain-wall interactions in Ni80Fe 20 planar nanowires. *Phys. Rev. B - Condens. Matter Mater. Phys.* **81**, (2010).
2. Estévez, V. & Laurson, L. Head-to-head domain wall structures in wide permalloy strips. *Phys. Rev. B - Condens. Matter Mater. Phys.* **91**, 54407 (2015).
3. McMichael, R. D. & Donahue, M. J. Head to head domain wall structures in thin magnetic strips. *IEEE Trans. Magn.* **33**, 4167–4169 (1997).
4. Kläui, M. *et al.* Head-to-head domain-wall phase diagram in mesoscopic ring magnets. *Appl. Phys. Lett.* **85**, 5637–5639 (2004).
5. Vaz, C. A. F. *et al.* Ferromagnetic nanorings. *J. Phys. Condens. Matter* **19**, (2007).
6. Beach, G. S. D., Nistor, C., Knutson, C., Tsoi, M. & Erskine, J. L. Dynamics of field-driven domain-wall propagation in ferromagnetic nanowires. *Nat. Mater.* **4**, 741–744 (2005).
7. Hayward, T. J. Intrinsic Nature of Stochastic Domain Wall Pinning Phenomena in Magnetic Nanowire Devices. *Sci. Rep.* **5**, (2015).
8. Hayward, T. J. & Omari, K. A. Beyond the quasi-particle: Stochastic domain wall dynamics in soft ferromagnetic nanowires. *J. Phys. Appl. Phys.* **50**, 84006 (2017).
9. Hayashi, M. *et al.* Dependence of current and field driven depinning of domain walls on their structure and chirality in permalloy nanowires. *Phys. Rev. Lett.* **97**, (2006).
10. Muñoz, M. & Prieto, J. L. Suppression of the intrinsic stochastic pinning of domain walls in magnetic nanostripes. *Nat. Commun.* **2**, (2011).
11. Mougin, A., Cormier, M., Adam, J. P., Metaxas, P. J. & Ferré, J. Domain wall mobility, stability and Walker breakdown in magnetic nanowires. *Europhys. Lett. EPL* **78**, 57007 (2007).
12. Broomhall, T. J., Rushforth, A. W., Rosamond, M. C., Linfield, E. H. & Hayward, T. J. Suppression of Dynamically Induced Stochastic Magnetic Behavior Through Materials Engineering. *Phys. Rev. Appl.* **13**, 024039 (2020).

13. Broomhall, T. J. & Hayward, T. J. Suppression of Stochastic Domain Wall Pinning Through Control of Gilbert Damping. *Sci. Rep.* **7**, 17100 (2017).
14. Glathe, S., Mattheis, R. & Berkov, D. V. Direct observation and control of the Walker breakdown process during a field driven domain wall motion. *Appl. Phys. Lett.* **93**, 072508 (2008).
15. Lee, J. Y., Lee, K. S., Choi, S., Guslienko, K. Y. & Kim, S. K. Dynamic transformations of the internal structure of a moving domain wall in magnetic nanostripes. *Phys. Rev. B - Condens. Matter Mater. Phys.* **76**, (2007).
16. Martinez, E., Lopez-Diaz, L., Torres, L. & Garcia-Cervera, C. J. Thermal activation in Permalloy nanorectangles at room temperature. *Phys. B Condens. Matter* **372**, 286–289 (2006).
17. Laufenberg, M. *et al.* Observation of thermally activated domain wall transformations. *Appl. Phys. Lett.* **88**, 052507 (2006).
18. Jiang, X. *et al.* Enhanced stochasticity of domain wall motion in magnetic racetracks due to dynamic pinning. *Nat. Commun.* **1**, (2010).
19. Bance, S. *et al.* Thermal Activation in Permanent Magnets. *JOM* **67**, 1350–1356 (2015).
20. Wuth, C., Lendecke, P. & Meier, G. Temperature-dependent dynamics of stochastic domain-wall depinning in nanowires. *J. Phys. Condens. Matter* **24**, (2012).
21. Wegrowe, J.-E. *et al.* Magnetic relaxation of nanowires: beyond the Néel-Brown activation process. *Europhys. Lett.* **38**, 329 (1997).
22. Wernsdorfer, W. *et al.* Experimental evidence of the Néel-Brown model of magnetization reversal. *Phys. Rev. Lett.* **78**, 1791–1794 (1997).
23. Stein, F. U., Bocklage, L., Matsuyama, T. & Meier, G. Generation and annihilation of domain walls in nanowires by localized fields. *Appl. Phys. Lett.* **100**, (2012).
24. Burn, D. M., Arac, E. & Atkinson, D. Magnetization switching and domain-wall propagation behavior in edge-modulated ferromagnetic nanowire structures. *Phys. Rev. B* **88**, 104422 (2013).
25. Berger, L. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B* **54**, 9353–9358 (1996).

26. Slonczewski, J. C. Current-driven excitation of magnetic multilayers. *J. Magn. Magn. Mater.* **159**, L1–L7 (1996).
27. Grollier, J. *et al.* Magnetic domain wall motion by spin transfer. *Comptes Rendus Phys.* **12**, 309–317 (2011).
28. Sinova, J., Valenzuela, S. O., Wunderlich, J., Back, C. H. & Jungwirth, T. Spin Hall effects. *Rev. Mod. Phys.* **87**, 1213–1260 (2015).
29. Mosendz, O. *et al.* Quantifying Spin Hall Angles from Spin Pumping: Experiments and Theory. *Phys. Rev. Lett.* **104**, 046601 (2010).
30. Self-consistent determination of spin Hall angle and spin diffusion length in Pt and Pd: The role of the interface spin loss. https://www.science.org/doi/10.1126/sciadv.aat1670 doi:10.1126/sciadv.aat1670.
31. Luo, Z. *et al.* Current-driven magnetic domain-wall logic. *Nat. 2020 5797798* **579**, 214–218 (2020).
32. Zhang, Y. *et al.* Ring-shaped Racetrack memory based on spin orbit torque driven chiral domain wall motions. *Sci. Rep.* **6**, 35062 (2016).
33. Cubukcu, M. *et al.* Ultra-Fast Perpendicular Spin–Orbit Torque MRAM. *IEEE Trans. Magn.* **54**, 1–4 (2018).
34. Zhang, Y., Luo, S., Yang, X. & Yang, C. Spin-orbit-torque-induced magnetic domain wall motion in Ta/CoFe nanowires with sloped perpendicular magnetic anisotropy. *Sci. Rep.* **7**, 2047 (2017).
35. Castaño, F. J. *et al.* Metastable states in magnetic nanorings. *Phys. Rev. B - Condens. Matter Mater. Phys.* **67**, (2003).
36. Zhu, F. Q. *et al.* Magnetic bistability and controllable reversal of asymmetric ferromagnetic nanorings. *Phys. Rev. Lett.* **96**, (2006).
37. Negoita, M., Hayward, T. J., Miller, J. A. & Allwood, D. A. Domain walls in ring-shaped nanowires under rotating applied fields. *J. Appl. Phys.* **114**, (2013).

38. Glathe, S., Berkov, I., Mikolajick, T. & Mattheis, R. Experimental study of domain wall motion in long nanostrips under the influence of a transverse field. *Appl. Phys. Lett.* **93**, 162505 (2008).
39. Kunz, A. & Priem, J. D. Dynamic notch pinning fields for domain walls in ferromagnetic nanowires. in *IEEE Transactions on Magnetics* vol. 46 1559–1561 (2010).
40. Chang, L. J., Yao, Y. D., Lin, P. & Lee, S. F. Magnetic Interaction in Domain Wall Depinning at Square Notch and Antinotch Traps. *IEEE Trans. Magn.* **47**, 2519–2521 (2011).
41. Rose, V. *et al.* Frustrated magnetic vortices in a triad of permalloy rings: Magneto-optical Kerr effect, magnetic force microscopy, and micromagnetic simulations. *Phys. Rev. B - Condens. Matter Mater. Phys.* **73**, (2006).
42. Ren, Y., Jain, S., Adeyeye, A. O. & Ross, C. A. Magnetization states in coupled Ni80Fe20 bi-ring nanostructures. *New J. Phys.* **12**, (2010).
43. Dawidek, R. W. *et al.* Dynamically-Driven Emergence in a Nanomagnetic System. *Adv. Funct. Mater.* **31**, (2021).
44. Cernansky, M. & Makula, M. Feed-forward echo state networks. in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.* vol. 3 1479–1482 vol. 3 (2005).
45. Jaeger, H. *The "echo state" approach to analysing and training recurrent neural networks- with an erratum note*. *GMD Technical Report* (2001).
46. Rodan, A. & Tino, P. Minimum Complexity Echo State Network. *IEEE Trans. Neural Netw.* **22**, 131–144 (2011).
47. Appeltant, L. *et al.* Information processing using a single dynamical node as complex system. *Nat. Commun.* **2**, (2011).
48. Torrejon, J. *et al.* Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017).
49. Paquot, Y. *et al.* Optoelectronic reservoir computing. *Sci. Rep.* **2**, (2012).
50. Liang, X. *et al.* Rotating neurons for all-analog implementation of cyclic reservoir computing. *Nat. Commun.* **13**, 1549 (2022).

51. Büsing, L., Schrauwen, B. & Legenstein, R. Connectivity, Dynamics, and Memory in Reservoir Computing with Binary and Analog Neurons. *Neural Comput.* **22**, 1272–1311 (2010).
52. Jaeger, H. *Short term memory in echo state networks*. *GMD Report 152* 60 (2002).
53. Dambre, J., Verstraeten, D., Schrauwen, B. & Massar, S. Information Processing Capacity of Dynamical Systems. *Sci. Rep.* **2**, 514 (2012).
54. Love, J., Mulkers, J., Bourianoff, G., Leliaert, J. & Everschor-Sitte, K. Task Agnostic Metrics for Reservoir Computing. Preprint at http://arxiv.org/abs/2108.01512 (2021).
55. Boyd, S. & Chua, L. Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Trans. Circuits Syst.* **32**, 1150–1161 (1985).
56. Dale, M., Miller, J. F., Stepney, S. & Trefzer, M. A. A substrate-independent framework to characterize reservoir computers. *Proc. R. Soc. Math. Phys. Eng. Sci.* **475**, 20180723 (2019).
57. McCloskey, M. & Cohen, N. J. Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem. in *Psychology of Learning and Motivation* (ed. Bower, G. H.) vol. 24 109–165 (Academic Press, 1989).
58. Ratcliff, R. Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychol. Rev.* **97**, 285–308 (1990).
59. Manneschi, L., Lin, A. C. & Vasilaki, E. SpaRCe: Improved Learning of Reservoir Computing Systems Through Sparse Representations. *IEEE Trans. Neural Netw. Learn. Syst.* (2021) doi:10.1109/TNNLS.2021.3102378.
60. Manneschi, L. *et al.* Exploiting Multiple Timescales in Hierarchical Echo State Networks. *Front. Appl. Math. Stat.* **6**, (2021).
61. Bellec, G. *et al.* A solution to the learning dilemma for recurrent networks of spiking neurons. *Nat. Commun.* **11**, 3625 (2020).
62. Tanaka, G., Matsumori, T., Yoshida, H. & Aihara, K. Reservoir computing with diverse timescales for prediction of multiscale dynamics. *Phys. Rev. Res.* **4**, L032014 (2022).

63. Fernando, C. & Sojakka, S. Pattern Recognition in a Bucket. in *Advances in Artificial Life* (eds. Banzhaf, W., Ziegler, J., Christaller, T., Dittrich, P. & Kim, J. T.) 588–597 (Springer, 2003). doi:10.1007/978-3-540-39432-7_63.

64. Soriano, M. C., Brunner, D., Escalona-Morán, M., Mirasso, C. R. & Fischer, I. Minimal approach to neuro-inspired information processing. *Front. Comput. Neurosci.* **9**, (2015).

65. Soriano, M. C. *et al.* Delay-Based Reservoir Computing: Noise Effects in a Combined Analog and Digital Implementation. *IEEE Trans. Neural Netw. Learn. Syst.* **26**, 388–393 (2015).

66. Li, J., Bai, K., Liu, L. & Yi, Y. A deep learning based approach for analog hardware implementation of delayed feedback reservoir computing system. in *2018 19th International Symposium on Quality Electronic Design (ISQED)* 308–313 (2018). doi:10.1109/ISQED.2018.8357305.

67. Jensen, J. H. & Tufte, G. Reservoir computing with a chaotic circuit. in 222–229 (MIT Press, 2017). doi:10.1162/isal_a_039.

68. Chua, L. Memristor-The missing circuit element. *IEEE Trans. Circuit Theory* **18**, 507–519 (1971).

69. Strukov, D. B., Snider, G. S., Stewart, D. R. & Williams, R. S. The missing memristor found. *Nature* **453**, 80–83 (2008).

70. Zhong, Y. *et al.* Dynamic memristor-based reservoir computing for high-efficiency temporal signal processing. *Nat. Commun.* **12**, 408 (2021).

71. Indiveri, G., Linares-Barranco, B., Legenstein, R., Deligeorgis, G. & Prodromakis, T. Integration of nanoscale memristor synapses in neuromorphic computing architectures. *Nanotechnology* **24**, 384010 (2013).

72. Hassan, A. M., Li, H. H. & Chen, Y. Hardware implementation of echo state networks using memristor double crossbar arrays. in *2017 International Joint Conference on Neural Networks (IJCNN)* 2171–2177 (2017). doi:10.1109/IJCNN.2017.7966118.

73. Yang, X., Chen, W. & Wang, F. Z. Investigations of the staircase memristor model and applications of memristor-based local connections. *Analog Integr. Circuits Signal Process.* **87**, 263–273 (2016).

74. Kulkarni, M. S. & Teuscher, C. Memristor-based reservoir computing. *Proc. 2012 IEEEACM Int. Symp. Nanoscale Archit. NANOARCH 2012* 226–232 (2012) doi:10.1145/2765491.2765531.

75. Moon, J. *et al.* Temporal data classification and forecasting using a memristor-based reservoir computing system. *Nat. Electron. 2019 210* **2**, 480–487 (2019).

76. Bürger, J. & Teuscher, C. Variation-tolerant Computing with Memristive Reservoirs. in *2013 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH)* 1–6 (2013). doi:10.1109/NanoArch.2013.6623028.

77. Mehonic, A. *et al.* Memristors—From In-Memory Computing, Deep Learning Acceleration, and Spiking Neural Networks to the Future of Neuromorphic and Bio-Inspired Computing. *Adv. Intell. Syst.* **2**, 2000085 (2020).

78. Vandoorne, K. *et al.* Toward optical signal processing using Photonic Reservoir Computing. *Opt. Express* **16**, 11182–11192 (2008).

79. Vandoorne, K. *et al.* Experimental demonstration of reservoir computing on a silicon photonics chip. *Nat. Commun.* **5**, 3541 (2014).

80. Zetie, K. P., Adams, S. F. & Tocknell, R. M. How does a Mach-Zehnder interferometer work? *Phys. Educ.* **35**, 46 (2000).

81. Allwood, D. A. *et al.* A perspective on physical reservoir computing with nanomagnetic devices. *Appl. Phys. Lett.* **(in review)**, (2022).

82. Riou, M. *et al.* Temporal pattern recognition with delayed-feedback spin-torque nano-oscillators. *Phys. Rev. Appl.* **12**, (2019).

83. Tsunegi, S. *et al.* Evaluation of memory capacity of spin torque oscillator for recurrent neural networks. *Jpn. J. Appl. Phys.* **57**, 120307 (2018).

84. Tsunegi, S. *et al.* Physical reservoir computing based on spin torque oscillator with forced synchronization. *Appl. Phys. Lett.* **114**, 164101 (2019).

85. Kanao, T. *et al.* Reservoir Computing on Spin-Torque Oscillator Array. *Phys. Rev. Appl.* **12**, 024052 (2019).

86. Ababei, R. V. *et al.* Neuromorphic computation with a single magnetic domain wall. *Sci. Rep. 2021 111* **11**, 1–13 (2021).

87. Jiang, W. *et al.* Physical reservoir computing using magnetic skyrmion memristor and spin torque nano-oscillator. *Appl. Phys. Lett.* **115**, 192403 (2019).

88. Zahedinejad, M. *et al.* Two-dimensional mutually synchronized spin Hall nano-oscillator arrays for neuromorphic computing. *Nat. Nanotechnol. 2019 151* **15**, 47–52 (2019).

89. Nakane, R., Tanaka, G. & Hirose, A. Reservoir Computing with Spin Waves Excited in a Garnet Film. *IEEE Access* **6**, 4462–4469 (2018).

90. Papp, A., Csaba, G. & Porod, W. Characterization of nonlinear spin-wave interference by reservoir-computing metrics. *Appl. Phys. Lett.* **119**, 112403 (2021).

91. Advances in artificial spin ice | Nature Reviews Physics. https://www.nature.com/articles/s42254-019-0118-3.

92. Jensen, J. H., Folven, E. & Tufte, G. Computation in artificial spin ice. in *ALIFE 2018 - 2018 Conference on Artificial Life: Beyond AI* 15–22 (MIT Press - Journals, 2020). doi:10.1162/isal_a_00011.

93. Jensen, J. H. & Tufte, G. Reservoir Computing in Artificial Spin Ice. in 376–383 (MIT Press, 2020). doi:10.1162/isal_a_00268.

94. Gartside, J. C. *et al.* Reconfigurable training and reservoir computing in an artificial spin-vortex ice via spin-wave fingerprinting. *Nat. Nanotechnol.* **17**, 460–469 (2022).

4- Experimental Techniques

This chapter will cover the range of experimental techniques that were applied in the realisation of this thesis. It will cover both the specific processes employed for each technique, as well as a brief overview of the physical principles that enable each technique. It will begin by covering each of the stages of device fabrication. Next, the process for performing electrical transport measurements of the ring arrays will be described.

4.1- Device Fabrication

The devices discussed in this thesis were fabricated using electron-beam lithography and lift-off processing, with metallisation via thermal evaporation. Figure 1 shows a schematic diagram of the basic fabrication procedure for a single lithography and metallisation step. Initially an atomically flat silicon substrate is coated with a polymer resist via spin coating. Patterning is then performed by locally exposing a polymer to radiation in the form of light or on electron beams. This exposure modifies the resist such that its solubility in a mild, selective developer solvent is modified, thus allowing only the regions of the substrate that are to be metallised to be uncovered. The substrate is metallised across the entire surface e.g., via physical vapour deposition, before a more aggressive solvent removes the remaining polymer, leaving behind only the patterned structures. The following section of this chapter will outline each phase of this process in more detail.

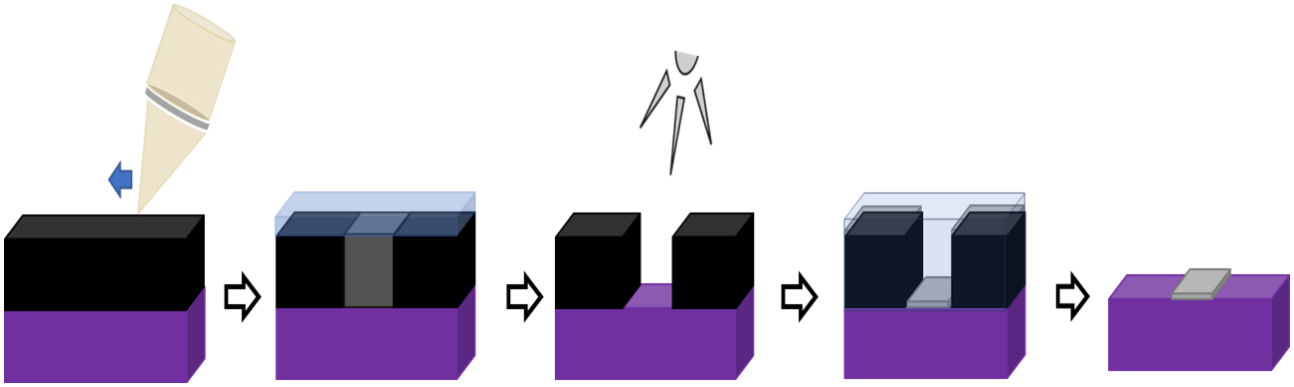


Figure 1- The manufacturing process for the nanoring devices. Left to right: first, an electron beam (yellow) is scanned over the polymer resist (black) above the silicon substrate (purple), a weak solvent (pale blue) then etches away the irradiated area (dark grey). Permalloy (light grey) is then metallised onto the surface of the sample, before the remaining resist is removed with a stronger solvent, leaving the patterned nanostructures behind.

4.1.1- Patterning Nanostructures.

The nanostructures considered in this thesis were grown on silicon wafers with a (100) crystallographic orientation and a thermally oxidised surface to facilitate electric transport measurements. The wafers were spin-coated with CSAR-62 resist to generate a thin (on the order of 100s of nanometres), homogenous layer of polymer. Prior to patterning, the resist undergoes a soft baking for three minutes at 180 ℃ to solidify the polymer and evaporate any solvent.

Patterning was achieved via electron beam lithography using a RAITH Voyager system, with typical beam dosages of $100\ \mu C/cm^2$ and beam voltages of $50\ kV$, which is scanned across the polymer according to the pattern geometry. This technique allows resolution of nanostructure features on the order of 10s of nanometres, though typical track widths for the rings exceeded ~200 nm in this work. For the positive resist used here, the patterning process causes chemical changes to the irradiated

material, breaking apart polymer chains and improving solubility. The patterned structure was then developed by heating to 23 ℃, and submerged in a xylene bath for 30 seconds, with light manual agitation of the solvent.

4.1.2- Metallisation via Thermal Evaporation

The next phase of device fabrication was the metallisation. This involved the controlled deposition of a homogenous metal surface upon the silicon substrate in a ultra-high vacuum environment. Here, the chosen metallisation technique was physical vapour deposition via thermal evaporation, which involved resistive heating of a tungsten wire basket crucible coated in aluminium oxide, which contained the deposition material in powder form. The combination of high crucible temperatures and low base pressures (here, below 10^{-6} mBar) caused the metal to evaporate, and the anaerobic environment prevented oxidation of the deposited film. The high vacuum also provided a long mean free path for the vapour, ensuring atom arrives at the substrate with uniform direction and energy, allowing for good compatibility with the lift-off process by minimising side-wall covering of the resist. To lift off the unwanted metal surface deposited on top of the remaining resist, organic solvent Microposit Remover 1165 was used. Typically, immersion in the solvent was sufficient for lift-off, though an ultrasound sonicator was used in cases where excess resist remained.

The nanostructures in this thesis were deposited in a Wordentec thermal evaporator with a base pressure ~10^{-7} mbar. Films were typically grown at pressures <5 x 10^{-5} mbar, and rates of 0.4 – 1.4 Å/s. A quartz micro-balance was used to measure deposition rates, allowing accurate measurement of the created film thicknesses, here between 10-20 nm of Permalloy. A secondary metallisation of ~3 nm of Aluminium was applied to provide a thin capping layer to the nanostructures, reducing the effects of oxidisation in altering the magnetic properties of the resulting devices.

4.1.3- Electrically Contacted Samples

For samples where electrical transport measurements were intended, the lithography and subsequent metallisation procedures were repeated, allowing patterning and deposition of electrical contacts on top of the ring structures. Here, reference markers were made in first lithography stage so that the second stage of lithography could be aligned.

Typically, a layer configuration of Ti (20 nm)/Au (200nm) was used, with the Ti layer acting as an adhesion layer for the Au. The resulting structure is shown in figure 2a. Electrical connections were then made between the contact pads (sample in a pseudo-4-point contact arrangement) joined a standard 20-pin leaded chip carrier (LCC) via wire bonding (Figure 2(b)). The 20 pins meant up to nine samples per wafer could be measured via connection to a pair of mutual ground pads.

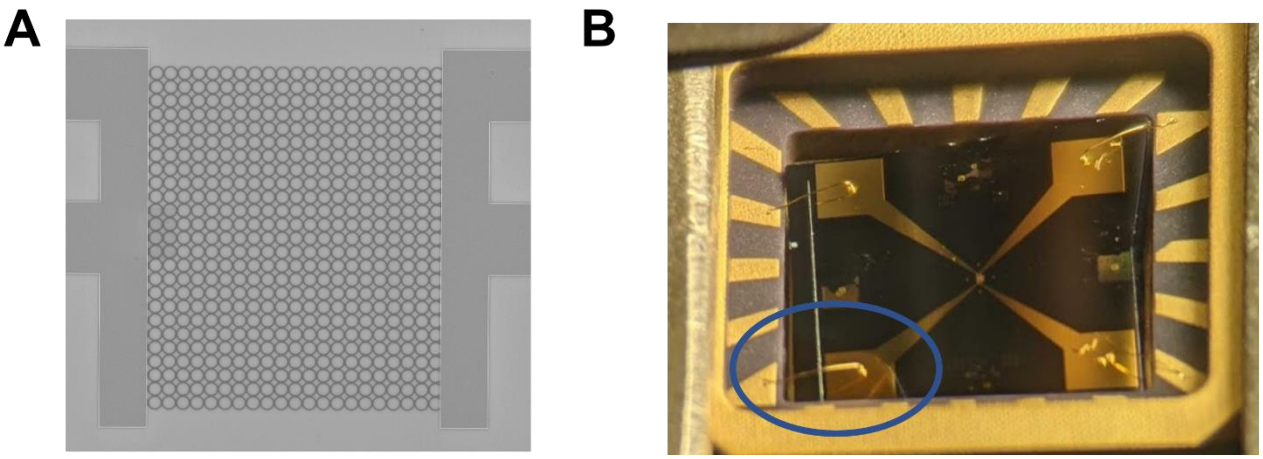


Figure 2- (A) Scanning electron micrograph of electrically contacted ring arrays, showing electrical gold contacts patterned upon the nanoring array. (B) Photograph showing wire-bonded connections (circled in blue) between the electrical contact pads of the device and the terminals on the chip-carrier.

4.2- Anisotropic Magnetoresistance Measurements of Ring Devices

Magnetoresistance measurements of the electrically contacted ring devices were performed using a custom-built electrical transport measurement system. A quadrupole electromagnet allowed the application of arbitrary, rotating magnetic fields to drive the nanoring arrays. The changes in resistance via anisotropic magnetoresistance were then measured using a lock-in amplifier. Figure 3 shows a schematic diagram of the complete rig, with signal paths for (blue) the driving field, (purple) measurement of the driving field, and (red) measurement of device AMR shown.

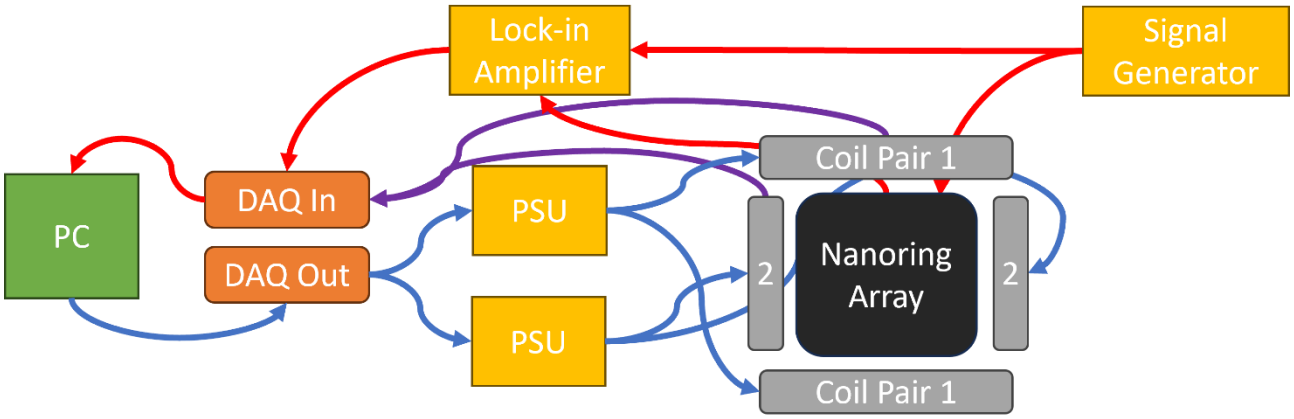


Figure 3- Schematic overview of electrical transport measurement rig. Follows signals that control electromagnet coils (blue), measurement of applied field (purple), and measurement of AMR signal (red).

4.2.1- Generating Rotating Magnetic Fields

Magnetisation dynamics in the nanorings were driven by rotating magnetic fields generated by two pairs of custom-built air-core electromagnets, shown in figure 3. The electromagnets driven by Kepco BOP 36-6D power supply units, which were run in constant current mode so that the magnetic fields applied were not affect by resistance changes in coils due to joule heating. The coils were mounted in a custom-built holder in a pseudo-Helmholtz arrangement to minimise the non-uniformity of the fields generated over the sample (Figure 3).

The current delivered to the coils were controlled by voltage outputs from a National Instruments DAQ card and were calibrated against the control voltages using a Hall probe situated above the mounted sample. To induce a rotating magnetic field, the two electromagnet pairs were driven with sinusoidal

inputs of equal magnitude with a 90° phase offset between the pairs. Due to the inductive load of the coils and the need to step quickly between arbitrary input magnitudes for machine learning tasks, the electromagnets were limited to operating speeds of around 100 Hz.

Figure 4- Photograph of the four electromagnet coils (outer) surrounding the nanoring device (centre).

4.2.2- Electrical Transport Measurements

In these measurements, sinusoidal current waves on the order of 40-50 kHz with amplitudes of ~1 mA were generated using a Keithley 6221 current source. Current was injected into the nanoring arrays, and potential difference over the device was measured via Au contact pads in a pseudo four-point arrangement, shown in Figure 2A/B. Due to the relatively small magnitude of the AMR effect (typically on the order of 5% for Permalloy[1]), coupled with the fact that that diversity of states presented by the rings only explored a faction of this maximum signal range, lock-in amplification was used to extract AMR signal signals from the rings.

This technique relies on the principle of orthogonality between sinusoidal functions, allowing signals that occur at frequencies or phases different to the reference signal to be attenuated to zero as signals are integrated over time. In practice, this detection is achieved by multiplying the reference signal and the input signal together, which means any signals that oscillate at the same frequency as the reference signal (the AMR-modulated reference signal) are maintained, while oscillations occurring at different frequencies (noise) are attenuated. A low-pass filter is used to remove the high-frequency modulation imparted by the carrier signal and 'integrate' the signal over time, which returns a DC signal proportional to the resistance of the device. The attenuation of noise is proportional to the time-constant of the low-pass filter, or mathematically speaking, the duration at which the sinusoids are integrated. In experiments, this leads to a trade-off between noise attenuation and responsivity of the lock-in, with longer integration windows (longer time-constants) leading to less noise but slower response of the DC output level (transforming the underlying signal which is amplified when the DC level changes).

For the rig used in this thesis, the reference signal is generated by the lock-in amplifier, which uses a step-edge clock generated from the trigger of the current source to match the reference signal frequency to that of the current source. The time constant of the low pass filter was set at either 1 or 3 µs, providing a rapid response time while enabling integration of many oscillations of the kHz signal. The lock-in used a sensitivity of 1 mV, and an internal operational amplifier was used to amplify the resulting signal output by the lock-in by a factor of 100. The resulting voltage signal was logged by the analogue input capability of the NI DAQ card.

4.3- X-ray Photoelectron Emission Microscopy

X-ray photoelectron emission microscopy (X-PEEM) is a powerful technique that allows contrast images of magnetisation states of to be acquired magnetic materials by monitoring the relative intensities of photoelectrons emitted from the surface of a magnetic material when illuminated with x-rays of opposite polarisations. Figure 5 shows a schematic diagram of a typical X-PEEM measurement. These measurements rely on several physical effects and experimental techniques, such as the photoelectric effect, x-ray absorption spectra (XAS) and x-ray circular magnetic dichroism (XMCD), which will be outlined in more detail here. Typically, synchrotron radiation is required to generate x-rays of sufficient brilliance and monochromaticity to perform magnetic contrast imaging, restricting X-PEEM experiments to major facilities. The measurements presented in this thesis were performed at beamline I06 at the Diamond Light Source and the CIRCE beamline at the ALBA synchrotron. The first sections will outline the physics behind X-PEEM, with a detailed experimental overview provided at the end.

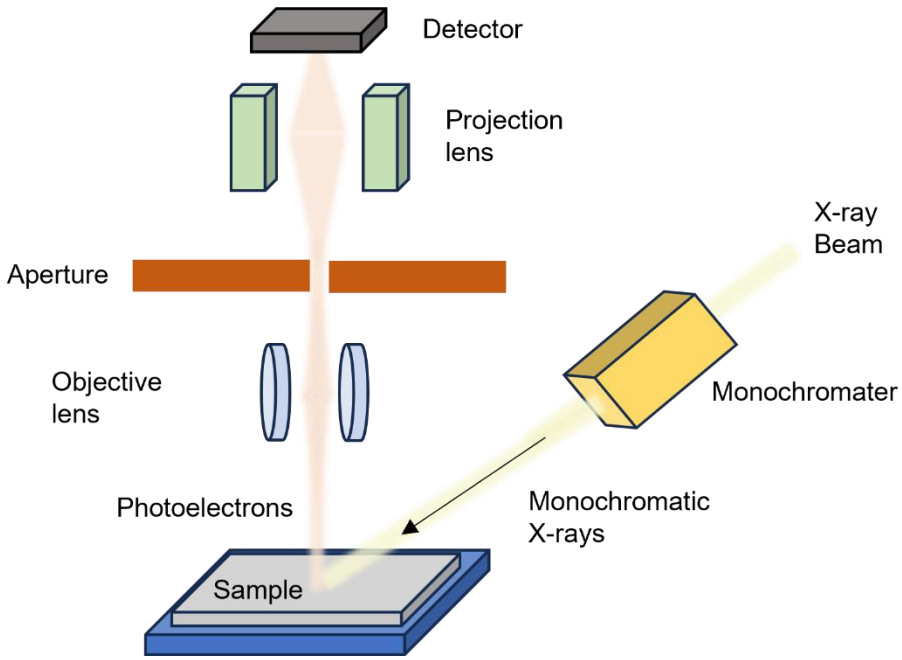


Figure 5- Schematic diagram of an X-ray photoelectron emission microscope, showing beam path and focussing lenses.

4.3.1- Photoelectric Effect

The photoelectric effect is a quantum mechanical process by which electrons of discrete energy levels become delocalised from matter after absorption of photons with energies exceeding the work function of the electrons. A schematic diagram of the emission of photoelectrons is shown in figure 6. Since this is a quantum effect, specific energy levels can be targeted by controlling the wavelength of incident light, which opens a range of possibilities for experiments, such as element-specific composition mapping, or the targeting of specific electrons which exhibit differential absorption depending upon magnetisation and polarisation of the light, enabling the X-PEEM measurements described later.

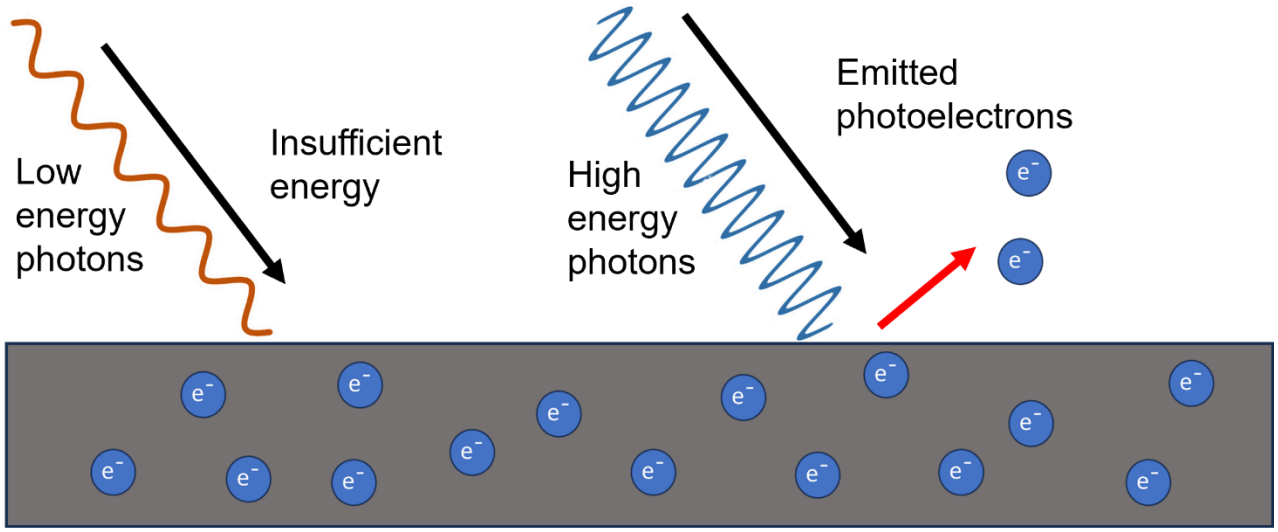


Figure 6- Schematic diagram showing long wavelength photons (orange), with photon energies below the work function of electrons occupying the metal surface, hence no photoelectrons emitted. Short wavelength, high energy photons (blue) with sufficient energy to delocalise electrons, resulting in photoemission.

4.3.2- X-ray Absorption Spectroscopy

XAS is a common technique used for determining the electron structure of a material by measuring the relative absorption of photons when the material is illuminated with monochromatic light of tuneable wavelengths. Here, the photon energy is gradually scanned, and the relative absorption of incident light is measured. This allows determination of the electron structure of the material by identifying a series of 'edges' associated with increased absorption of x-rays with energies corresponding to the discrete energy levels of electrons. In the context of X-PEEM imaging, XAS is used to identify the specific photon energies required to emit photoelectrons from energy levels which give rise to the underlying magnetic properties of the material.

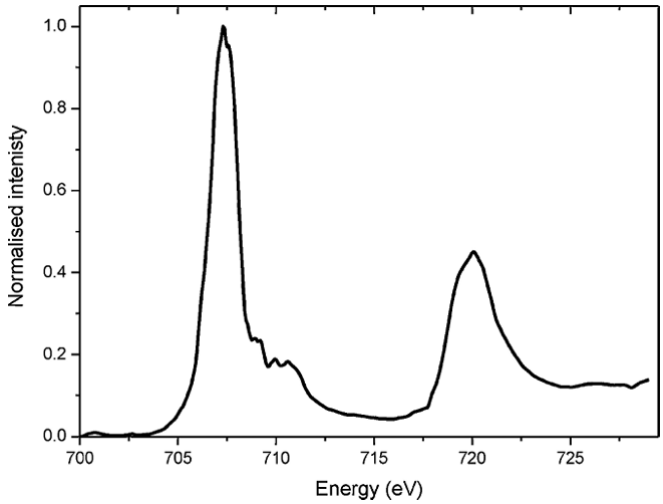


Figure 7- A typical x-ray absorption spectra for the L-3 edge of iron, showing a resonant peak at around 706 eV. Taken from[2].

4.3.3- X-ray Magnetic Circular Dichroism

In magneto-optics, XMCD refers to the difference in relative absorption of incident circularly polarised x-rays depending upon the magnetisation of the illuminated material. This phenomenon is critical to X-PEEM imaging as it allows contrast images to be drawn from the differential absorption of X-rays. In the case of the Permalloy structures studied here, X-PEEM experiments tend to use the L-3 edge of iron, identified via XAS. Despite Py being compositionally mostly nickel, the iron edge is chosen as it

provides a bigger contrast in photoemission via XMCD than nickel, and subsequently leads to bigger signals from which create contrast images.

4.3.4- Experimental Procedure

The process for creating magnetic contrast images via X-PEEM was similar at both Diamond and ALBA and both rely upon all of the phenomena detailed above. Firstly, an XAS spectrum of the material was generated by varying x-ray photon energy around the targeted edge using a monochromator, which allows calibration between the quoted energy level of the monochromator and the targeted edge at the point of maximum absorption. Next, a normalisation image was taken on the substrate away from the magnetic material as a control to improve image quality via flat-field correction (compensating for spatial variations in sensitivity of the detector), and illumination correction (correcting for non-uniformity of the illuminating beam).

To generate the contrast images, a series of four photoemission electron microscopy measurements were taken, which measured the relative intensities of emitted electrons both on and off resonance at the energy level of the chosen edge for x-rays with both clockwise and anticlockwise polarisation. The measurements at resonance allows determination of XMCD across the image via the difference in intensity, and the measurements off resonance give a background level with which to normalise the signal, generating a larger relative contrast. In the resulting image, magnetic contrast was generated along the axis of the incident electron beam, with magnetisation along the direction of the x-rays leading to the largest difference in intensity, and the sign of the difference when subtracting anticlockwise intensity from clockwise intensity determining whether magnetisation is aligned parallel or anti-parallel. Example images generated via X-PEEM at both Diamond and Alba are shown in Figure 8.

For the application of in-situ rotating magnetic fields, a specialist sample holder designed at ALBA[3] was used, which allowed generation of in plane fields via two pairs of orthogonal pole pieces. The applied fields were calibrated against the current provided to the electromagnets via a hall probe and a micropositioner, ensuring field was measured directly above the sample due to the large variation expected with the pole pieces in close proximity to one another. This also meant that any samples used for microscopy must occupy an area on the wafer under 1 mm to minimise the inhomogeneity of the field.

Due to the presence of remenant magnetisation in the iron pole pieces, an additional calibration was added to the experimental procedure which aimed to minimise the contribution of remanence to the applied field sequences. First, the pole pieces were demagnetised via rotating fields of decaying magnitude. The position of an identifiable feature of the ring arrays (usually the intersection between a corner ring and an adjacent edge ring) was established, and pixel locations recorded. Next, the field required to saturate a given array was found via gradual ramping of the applied fields, imaging between applications, and determining saturation via observation of a uniformly magnetised array. Finally, a compensatory field was determined via trial-and-error searching to find the required field to restore the calibration feature to its original location, hence removing the additional deflection of photoelectrons caused by remanent fields in the pole pieces.

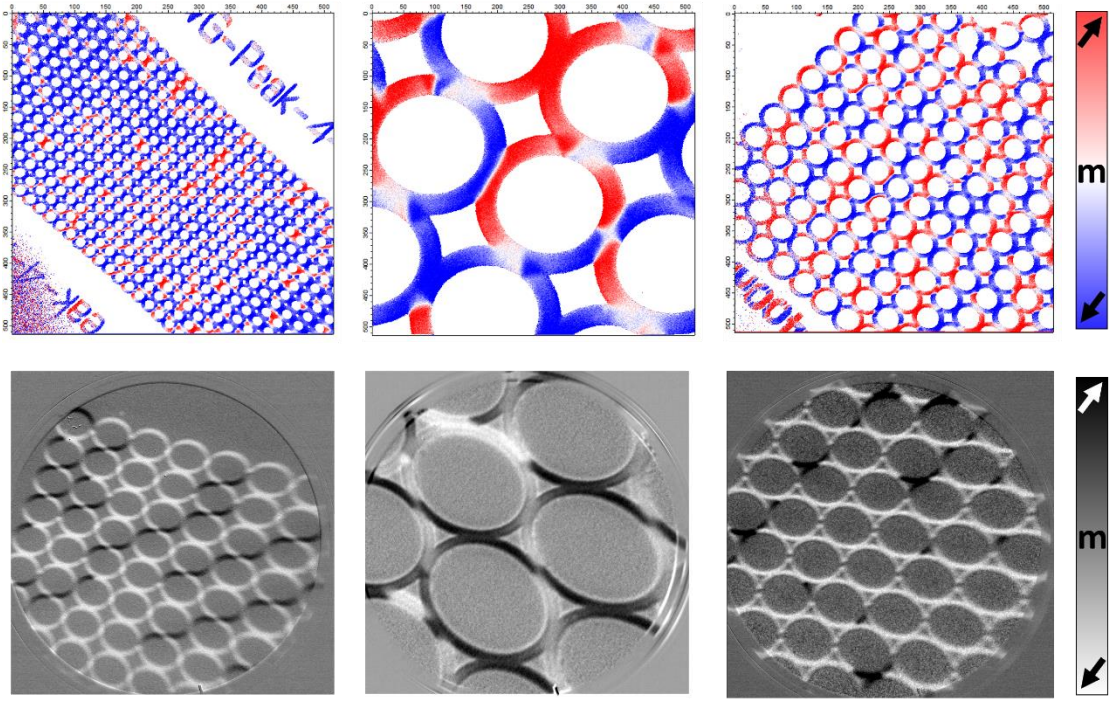


Figure 8- X-PEEM micrographs taken from Diamond (upper) and ALBA (lower) synchrotrons, showing a range of fields-of-view to demonstrate range of magnetisation resolution available. Colour bars show magnetisation direction.

4.4- Polarised Neutron Reflectometry

Polarised neutron reflectometry (PNR) is a characterisation technique that allows the absolute magnetisation per atom of magnetic materials to be determined in magnetic thin films. To generate the neutron beam, typically a spallation source at a specialist nuclear facility is used. The results used in this thesis were generated at ISIS muon and neutron source in the UK. The spallation source collides high energy protons with a heavy target material (usually mercury or tungsten), leading to a cascade of nuclear reactions that generate neutrons with a broad energy distribution. These neutrons are then moderated with water to lower their energies and are collimated and polarised to produce a beam. A schematic diagram of the beam path is shown in figure 9.

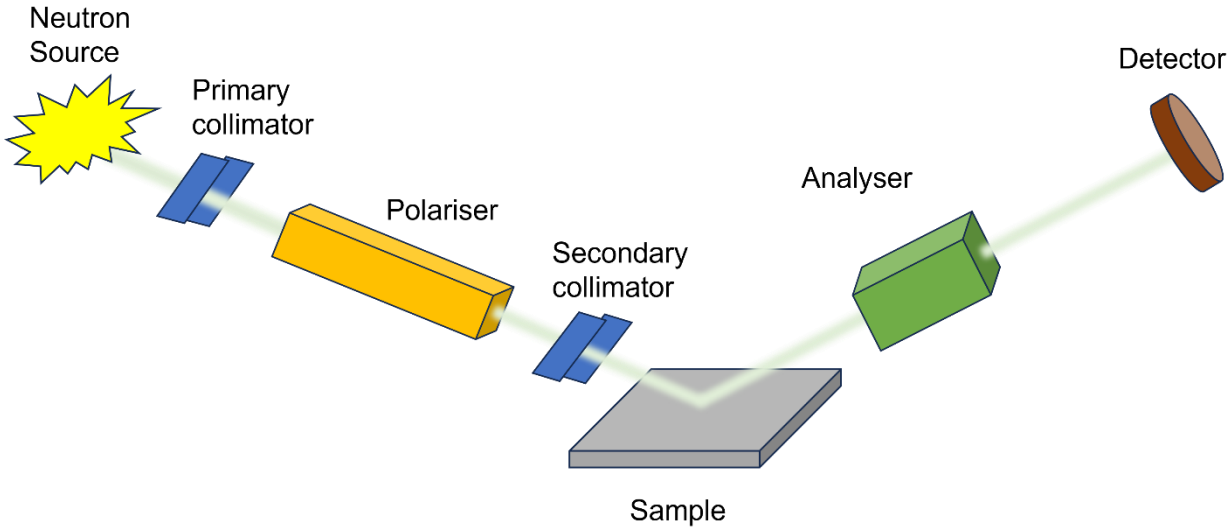


Figure 9- Schematic diagram showing the beam path of neutrons in a polarised neutron reflectometry experiment. Neutrons are generated by a spallation source, which are then collimated into a beam. The beam is

then polarised before focussing again prior to irradiating the sample. An analyser measures the polarisation of the reflected beam, and a detector measures the deflection angle.

The sample is held in the presence of an external magnetic field then irradiated with the neutrons polarised either parallel or antiparallel to the field, and the reflectivity of the neutron beam is measured as a function of the angle of reflection. Due to interactions between the neutrons and the magnetic moments of the material, the intensity and polarisation of the reflected neutrons vary with the magnetisation of the material. To determine magnetisation, measurements are taken for both parallel and anti-parallel alignment of the neutron polarisation and the magnetic field, and the spin-flipped and non-spin-flip components for a given spin state are summed, providing sensitivity to magnetisation along the polarisation direction.

To extract the magnetisation-dependent scattering from the measurement, models from the GenX[4] software package were employed, with the resulting fit from this process shown in figure 10, which plots neutron beam intensity as a function of the scattering vector of the beam. To fit the model, scattering contributions from the silicon wafer away from the magnetic material, as well as the exposed silicon in between the ring structures must be considered. Due to the large in-plane coherence length of silicon (~100µm) relative to the individual nanorings (2/4 µm), contributions from the rings themselves and the silicon between the rings were summed coherently in the model, and the contributions from the wafer away from the sample were summed incoherently. While the nanorings themselves were individually too small to resolve, PNR enabled the measurement of entire arrays. However, as the rings are sensitive to small magnetic fields, the external field was limited to around 18 Oe, which resulted in very small signals. To counteract this, the ring structures were patterned over an area on the order of a few centimetres squared, allowing effective measurement without perturbing the magnetic state of the sample. To initialise the ring arrays, a large in-plane field pulse of 1.9 kOe was applied, saturating the rings into the 'onion' state. The magnitude of this field was then reduced to the desired strength H_{app} between 20-150 Oe, and the sample itself was rotated for N revolutions, generating rotating fields in the reference frame of the sample. Sample magnetisation was then determined from the spin-asymmetry between the reflectivity profiles of opposite neutron polarisations.

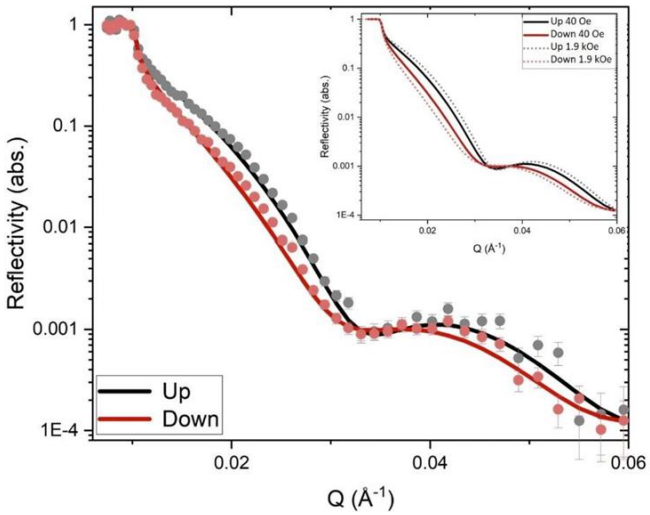


Figure 10- Plot of spin-up (black) and spin-down (red) neutron beam intensities as a function of neutron wavelength. The spin asymmetry, calculated from the separation between the two components, is used to determine the magnetisation of the sample. Taken from[5].

4.5- Micromagnetic Modelling

Micromagnetic modelling provides huge utility in simulating the magnetic response of nanoscale magnetic systems, and often serves as a basis for providing a detailed prediction of magnetisation dynamics observed experimentally, or as an explorative tool for designing and optimising magnetic devices. The models provide less detail than atomistic models that simulate the response of single spins but offer a vast speedup in simulation times (especially with recent GPU acceleration) and volume of material that can be simulated, whilst providing adequate detail to simulate nanoscale magnetic structures such as domain walls or skyrmions. Micromagnetic models are a general-purpose simulation platform due to the ability to arbitrarily define the geometry of magnetic material as well as parameters such as exchange stiffness, saturation magnetisation, and damping constant to reflect a broad range of material properties. The work in this thesis uses is the MuMax3[6] software package, which employs a finite-difference approach to solving interactions between cells. The specific considerations of MuMax3 will be described in detail.

The fundamental principle behind micromagnetic methods is a classical continuum approximation of magnetisation energies, which is then minimised. Here, energy contributions from external field, anisotropy, exchange, and magnetostatics are calculated via the following equations (4.1-4.4), respectively:

$$(4.1)\quad E_H = -\int_V \vec{H}_{ext}\vec{M}\,dV$$

$$(4.2)\quad E_{ani} = -\int_V K_{ani}(\vec{m}\vec{e})^2\,dV$$

$$(4.3)\quad E_{ex} = \int_V A_{ex}[(\nabla\vec{m}_x)^2 + (\nabla\vec{m}_y)^2 + (\nabla\vec{m}_z)^2]\,dV$$

$$(4.4)\quad E_{demag} = -\int_V \vec{H}_{demag}(\vec{r})\vec{M}(\vec{r})\,dV$$

where $\vec{H}_{ext}$ represents the external field vector, $\vec{M}$ is the magnetisation vector, K_{ani} is the anisotropy constant, $\vec{m}$ is a unit direction vector of magnetisation, $\vec{e}$ a unit vector describing anisotropy direction, A_{ex} is the exchange constant, and $\vec{H}_{demag}$ the demagnetisation field.

Numerical micromagnetic simulations involve approximating the spin structure of the underlying material into cells of uniform magnetisation, solving the above equations, and modelling the interactions between each cell and all other cells in the simulation. When defining the size of each of the cells within the mesh, a trade-off occurs between simulation times and level of detail encapsulated by the model, with smaller cell sizes providing greater resolution but increased computational overheads. One useful parameter for determining appropriate sizes of cells in the mesh is the exchange length, l_{ex}, which describes a characteristic at which exchange interactions dominate over magnetostatic interactions, and hence the mesh size should be selected to be smaller than l_{ex}. For 3d ferromagnetic materials, this is usually on the order of a few nanometres, and can be calculated via equation 4.5:

$$(4.5)\quad l_{ex} = \sqrt{\frac{A}{M_s^2}}$$

To determine the dynamic magnetisation response of the cells, MuMax3 calculates the rate of change of reduced magnetisation with respect to time, giving a torque which has three contributions consisting of Landau-Lifshitz torque, Zhang-Li spin-transfer torque, and Slonczewski spin-transfer torque. While the spin-transfer torque terms allow determination of the magnetisation dynamics induced by interaction with electrical currents, the simulations performed in this thesis involved only fields, and as such only the Landau-Lifshitz torque will be elaborated upon.

The explicit form of the Landau-Lifshitz torque, τ_{LL}, employed within MuMax3 is defined in equation (4.6):

$$\overrightarrow{\tau_{LL}} = \gamma_{LL} \frac{1}{1+\alpha^2}\left(\vec{\mathbf{m}} \times \vec{\mathbf{B}}_{\text{eff}} + \alpha\left(\vec{\mathbf{m}} \times \left(\vec{\mathbf{m}} \times \vec{\mathbf{B}}_{\text{eff}}\right)\right)\right) \quad (4.6)$$

where γ_{LL} describes the gyromagnetic ratio of the material, α the Gilbert damping parameter, $\vec{\mathbf{m}}$ the reduced magnetisation vector, and $\vec{\mathbf{B}}_{\text{eff}}$ the effective magnetic field acting upon the cell. The effective magnetic field considers a series of contributions, including the externally applied field, the magnetostatic field, the Heisenberg exchange field, the Dzyaloshinskii-Moriya exchange field, the magneto-crystalline anisotropy field, and the thermal field. Since the modelling performed in this thesis focussed solely on determination of ground states and zero-kelvin switching fields (chapters 5 and 7), thermal contributions were omitted. The equation is calculated iteratively for all cells. Here, an artificially large damping parameter of 1 was employed to provide a simplified model of the dynamic evolution of magnetisation, effectively removing Gilbert damping contributions and allowing the energy minimisation processes to occur over the order of 10ns. The effects of temperature were also ignored by performing simulations at zero Kelvin, removing thermal field contributions, and giving deterministic results. An example output showing the magnetisation of a single ring in the ‘onion’ state is shown in figure 11.

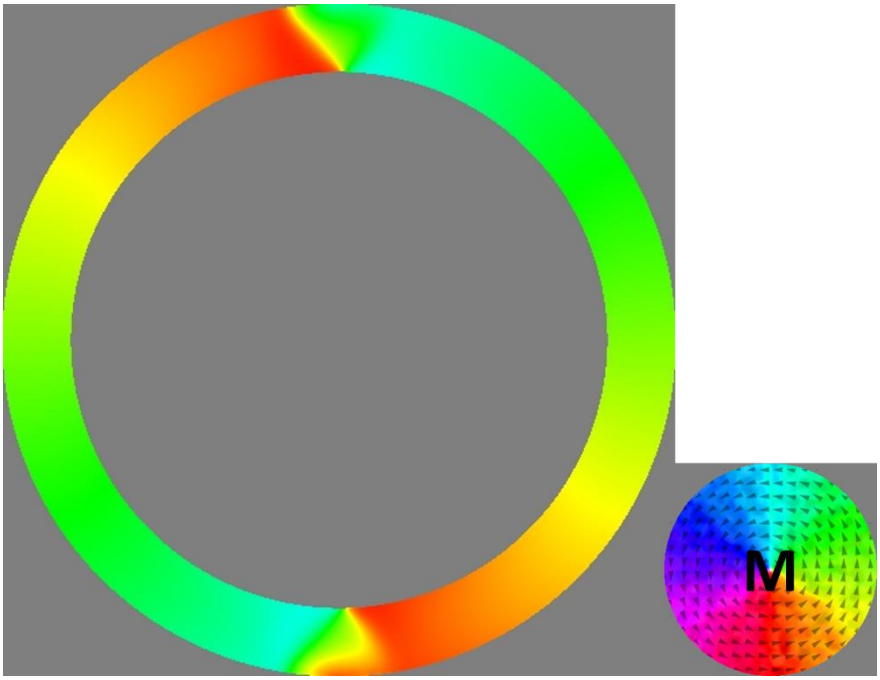


Figure 11- MuMax3 magnetisation output of a single nanoring in the 'onion' state. Colour wheel shows encoding of magnetisation direction to colour.

4.6- Agent-Based Modelling

While micromagnetic models form the standard practice for modelling magnetic systems, the size of the networks of ring arrays studied here, plus the many-second timescale of simulation makes micromagnetic approaches practically impossible. Instead, an agent-based approach was taken.

Agent-based modelling is a simulation technique that focusses upon the interactions between sets of individual 'agents' which often have distributed properties that describe their interactions with both other agents as well as the simulation environment stochastically. This approach differs from more standard equation-based approaches[7] as they focus upon the emergent behaviours that arise from stochastic interactions between individuals and enable complex system-wide phenomena that cannot be described by the actions of individuals alone. Agent-based models (ABMs) have seen applications in a wide range of different fields, including biology[8,9], ecology[10], epidemiology[11], and social science[12].

The ABM framework synergises well with the nanoring arrays as the response of the system can be described by the action of and interaction between domain walls in the system (the agents), which have inherently stochastic outcomes due to the presence of thermal effects in the system. The framework enables the distributed parameters that occur in real magnetic devices via manufacturing imperfections to be represented by instantiating the environmental parameters of the junctions by randomly sampling the properties that describe their magnetic interactions from a distribution.

By representing the nanoring array as a discretised system of agents interacting with a fixed environment, the processes that determine the magnetic state of the array can be vastly simplified, enabling simulation of much larger volumes of magnetic material evolving over much longer time durations than would be accessible via micromagnetic techniques. Additionally, this approach encapsulates how the outcome of each stochastic interaction influences the future evolution of the system, encapsulating the response as part of a Markov process, resembling the nature of evolution of the domain structure observed experimentally. The production of the model will be presented in detail in chapter 5.

References:

1. Nagura, H., Saito, K., Takanashi, K. & Fujimori, H. Influence of third elements on the anisotropic magnetoresistance in permalloy films. *J. Magn. Magn. Mater.* **212**, 53–58 (2000).
2. Miedema, P. S. & de Groot, F. M. F. The iron L edges: Fe 2p X-ray absorption and electron energy loss spectroscopy. *J. Electron Spectrosc. Relat. Phenom.* **187**, 32–48 (2013).
3. Foerster, M. *et al.* Custom sample environments at the ALBA XPEEM. *Ultramicroscopy* **171**, 63–69 (2016).
4. Björck, M. & Andersson, G. GenX: an extensible X-ray reflectivity refinement program utilizing differential evolution. *J. Appl. Crystallogr.* **40**, 1174–1178 (2007).
5. Dawidek, R. W. *et al.* Dynamically-Driven Emergence in a Nanomagnetic System. *Adv. Funct. Mater.* **31**, (2021).
6. Vansteenkiste, A. *et al.* The design and verification of MuMax3. *AIP Adv.* **4**, 107133 (2014).
7. Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide | SpringerLink. https://link.springer.com/chapter/10.1007/10692956_2.
8. Walker, D. C. *et al.* The epitheliome: agent-based modelling of the social behaviour of cells. *Biosystems* **76**, 89–100 (2004).
9. Agent-based modelling in synthetic biology | Essays in Biochemistry | Portland Press. https://portlandpress.com/essaysbiochem/article/60/4/325/78417/Agent-based-modelling-in-synthetic-biology.
10. McLane, A. J., Semeniuk, C., McDermid, G. J. & Marceau, D. J. The role of agent-based models in wildlife ecology and management. *Ecol. Model.* **222**, 1544–1556 (2011).

11. Cuevas, E. An agent-based model to evaluate the COVID-19 transmission risks in facilities. *Comput. Biol. Med.* **121**, 103827 (2020).
12. Agent-Based Models in Empirical Social Research - Elizabeth Bruch, Jon Atwell, 2015. https://journals.sagepub.com/doi/full/10.1177/0049124113506405.

5- “RingSim- An Agent-based Approach for Modelling Mesoscopic Magnetic Nanowire Networks”

5.0- Preamble

This chapter presents the implementation and validation of RingSim- a bespoke simulation platform designed to model the response of magnetic nanoring networks in response to rotating magnetic fields. This work follows on from the contributions of previous PhD/Masters students of the group who had extensively studied the physics of both single nanorings and nanoring arrays.

Attempts to model the system previously were limited simple numerical models that describe array-wide properties in terms of depinning probability and a balancing of rate equations that describe steady state populations. However, these numerical models neglect the complex, spatially distributed behaviours that had been observed in the past. Furthermore, by only considering equilibrium conditions and steady state populations, the dynamic properties that make the system interesting for reservoir computing are neglected.

While micromagnetic models had been used previously to resolve the fine-grain details of interactions between a single ring and its neighbours, it is a practical impossibility to extend the model to cover larger arrays where emergent behaviours arising from these interactions occur. This is compounded by the fact that rotational magnetic field inputs that had been provided to the system occur at 10s of Hz, meaning signals of many seconds would be required to encode any meaningful information to the arrays for reservoir computing purposes.

It was clear that a different approach to modelling the system was required. Agent-based models are especially suited to describing systems with emergent behaviours, as they focus on the actions and interactions between many individuals, and thus were selected as a promising candidate for simulating the nanoring arrays. Since the interactions of domain walls with both external fields and each other dictate the response of the system, it was proposed that the system could be reduced to simply modelling the domain walls and their interactions, then calculating the magnetisation state of the array from the position of the domain walls.

In the model, the stochastic depinning events observed previously were modelled via an Arrhenius-Néel law. The associated energy barriers, as well as how they are modulated by external fields, were calculated via a phenomenological approximation of Sharrock’s law. This allowed the model to tie realistic inputs to the expected probabilities of reversal for a given domain wall. The interactions between domain walls in the model were programmed to reproduce behaviours that had been observed in previous experiments. This allowed the emergent behaviours observed experimentally to arise from a few simple rules within RingSim.

The model was tuned to replicate steady-state experimental results, and then validated against both measurements of the systems dynamics as well as with observations of the microstates the ring arrays tend to form. Realising a faithful model of the ring arrays’ complex response was a vital first step that unlocked the ability to explore how the system could be used as a basis for reservoir computing. Author contributions are listed at the end of the article.

RingSim- An Agent-based Approach for Modelling Mesoscopic Magnetic Nanowire Networks

Ian T Vidamour[1,2], Guru Venkat[1], Charles Swindells[1], David Griffin[3], Paul W Fry[4], Richard M Rowan-Robinson[1], Alex Welbourne[1], Francesco Maccherozzi[5], Sarnjeet S Dhesi[5], Susan Stepney[3], Eleni Vasilaki[2], Dan A Allwood[1], and Thomas J Hayward[1]

Department of Materials Science and Engineering, University of Sheffield[1]

Department of Computer Science, University of Sheffield[2]

Department of Computer Science, University of York[3]

Centre for Nanoscience and Technology, University of Sheffield[4]

Diamond Light Source Ltd., Science and Technology Facilities Council UK[5]

Abstract

We describe 'RingSim', a phenomenological agent-based model that allows numerical simulation of magnetic nanowire networks with dimensions as large as hundreds of micrometres over time scales as long as seconds, which are a practical impossibility for general-purpose micromagnetic simulation tools. In RingSim, domain walls are instanced as mobile agents which respond to magnetic field and uses simple phenomenological rules to describe their stochastic interactions with pinning sites and other domain walls. We first present a detailed description of the model and its algorithmic implementation when simulating the behaviours of arrays of interconnected ring-shaped nanowires, which have previously been proposed as hardware platforms for unconventional computing applications. The model is then validated against a series of experimental measurements of an array's static and dynamic responses to rotating magnetic fields, and its field-dependent magnetic state populations. The results from RingSim agree well with those of experimental studies, demonstrating that agent-based modelling is a a powerful tool for the exploring mesoscale magnetic devices over time and length scales that are inaccessible to more conventional simulation techniques.

Introduction

The creation of models of system behaviour is critical to the development of emerging technologies, since they allow for rapid evaluation of device behaviour without the need for manufacturing samples or performing experimental measurements. For devices based around magnetic materials, the processes that underpin the device's response to external inputs often have well-defined physical descriptions. Examples such as spin-torque nano-oscillators [1, 2], domain wall oscillators [3, 4], and super-paramagnetic ensembles [5] have one-dimensional numerical descriptions that allow the state of the magnetic substrate to be approximated to predict device performance quickly with low computation cost.

For devices without well-defined dynamical equations, or when more detail on the underlying magnetic state of the system is required, the typical approach is to use general-purpose micromagnetic simulation packages such as MuMax3 [6] or OOMMF [7]. These platforms

approximate the spin structure of the magnetic materials into cells on the order of a few nanometres in size and model the evolution of spins in the presence of external driving fields or spin-torques via the Landau-Lifschitz-Gilbert equation [8, 9]. While these simulation packages provide a high level of detail on the magnetic response of a device, they are associated with a high computational overhead. For example, in simulations of a Skyrmion confined in a nanodisk of 80nm diameter, 1nm thickness, using a mesh size of $1nm^3$, simulating 50ms of dynamic response takes on the order of 40 minutes [10], a simulation duration 48,000 times longer than the physical processes being simulated, even despite the vast acceleration of these simulations possible with modern hardware [11].

While signal durations on the order of seconds are not appropriate for magnetic systems with fast dynamics such as spin-torque nano-oscillators, other systems with dynamics governed by thermal processes, such as the nanowire network presented here, have much slower timescales of response. In order to excite interesting dynamics in these systems, the systems must be driven with input stimuli that evolve over similar timescales. When coupled with the relatively larger size of these systems, simulation with micromagnetic approaches becomes practically impossible.

The complexity of responses exhibited in devices of large networks of interacting magnetic elements such as artificial spin-ice systems [12–15], or arrays of interconnected magnetic nanorings [16], present increased difficulty in describing the system's evolution, with each containing hundreds, thousands, or even millions of nanometre/micrometre-scale magnetic elements. These systems exhibit emergent behaviours, where interactions between elements in the system lead to global behaviours that cannot be described by the action of individual elements alone. While these complex dynamics provide technical difficulty for simulation, they exhibit interesting computational properties, with the dynamics exploited for neuromorphic computing purposes [17–19]. With the computational constraints of general-purpose simulation packages mentioned earlier, it is clear that alternative approaches for modelling such a system must be taken.

Agent-based models describe the evolution of complex, multi-element systems in terms of the interactions between individual agents, as well as external environmental parameters [20]. The agents are often instanced into the model with distributed parameters, and commonly feature stochastic behaviours. Interactions are programmed in phenomenologically, with the outcomes of interactions determined by a set of predefined rules that aim to encapsulate the behaviours of the system being simulated. These types of models are especially harmonious with systems that exhibit emergent behaviours and have been used extensively in modelling complex dynamic systems such as flocking birds [21], or even structures within the brain [22, 23].

In this paper, we introduce an agent-based model, RingSim, which allows rapid simulation of networks of up to thousands of interconnected magnetic nanorings for signal durations on the order of seconds. The response of the model is governed by the outcome of stochastic pinning events and subsequent propagation of domain walls. The local pinning sites that occur at junctions where the rings intersect are programmed as interaction sites, with domain walls instanced as agents that interact with both neighbouring domain walls as well as external driving fields. By fitting model parameters describing ring junction properties, as well as the degree of influence of domain wall-domain wall interactions, we show how the model provides excellent agreement between simulated responses and experimental data. Not only is the model capable of encapsulating the system's global response to external inputs both statically and dynamically, but also provides information on the microstate of each ring, visually showing similar local

agglomeration of domains and relative populations of individual ring states as observed experimentally.

Magnetic Nanoring Array Dynamics

RingSim describes the response of interconnected magnetic nanoring arrays (Figure 1a), experimentally detailed in previous works [16, 18, 24]. Conceptually, the response of the system can be described via the transitions between metastable domain configuration for each of the rings, mediated by stochastic pinning events. Figure 1b shows the three basic ring configurations: (i) a ’vortex’ ring containing a single domain and no domain walls, (ii) an ’onion’ ring containing two equally sized domains and a pair of domain walls at opposite ends of the ring, and (iii) a ’three-quarter’ ring, featuring two differently sized domains, with domain walls situated at angles of 90 degrees from one another.

The interconnected nature of the ring arrays leads to local pinning sites at the junctions between neighbouring rings. The change in geometry presented by the junction creates an energy barrier, similar to an ’anti-notch’ [25, 26], which tends to locally pin itinerant domain walls. When driven with sufficiently high amplitude rotating fields, the domain walls are able to overcome this energy barrier, and coherently propagate with the rotating field, shown in Figure 1c(i). Under smaller driving fields, domain walls can become locally pinned at these junctions, with a finite probability of depinning via thermal activation. This stochastic process leads to decoupled propagation of domain walls within the rings depending upon the outcomes of pinning events, which may lead to domain wall collision and subsequent annihilation, changing the ring from an onion/three-quarter state to a vortex state, shown in Figure 1c(ii).

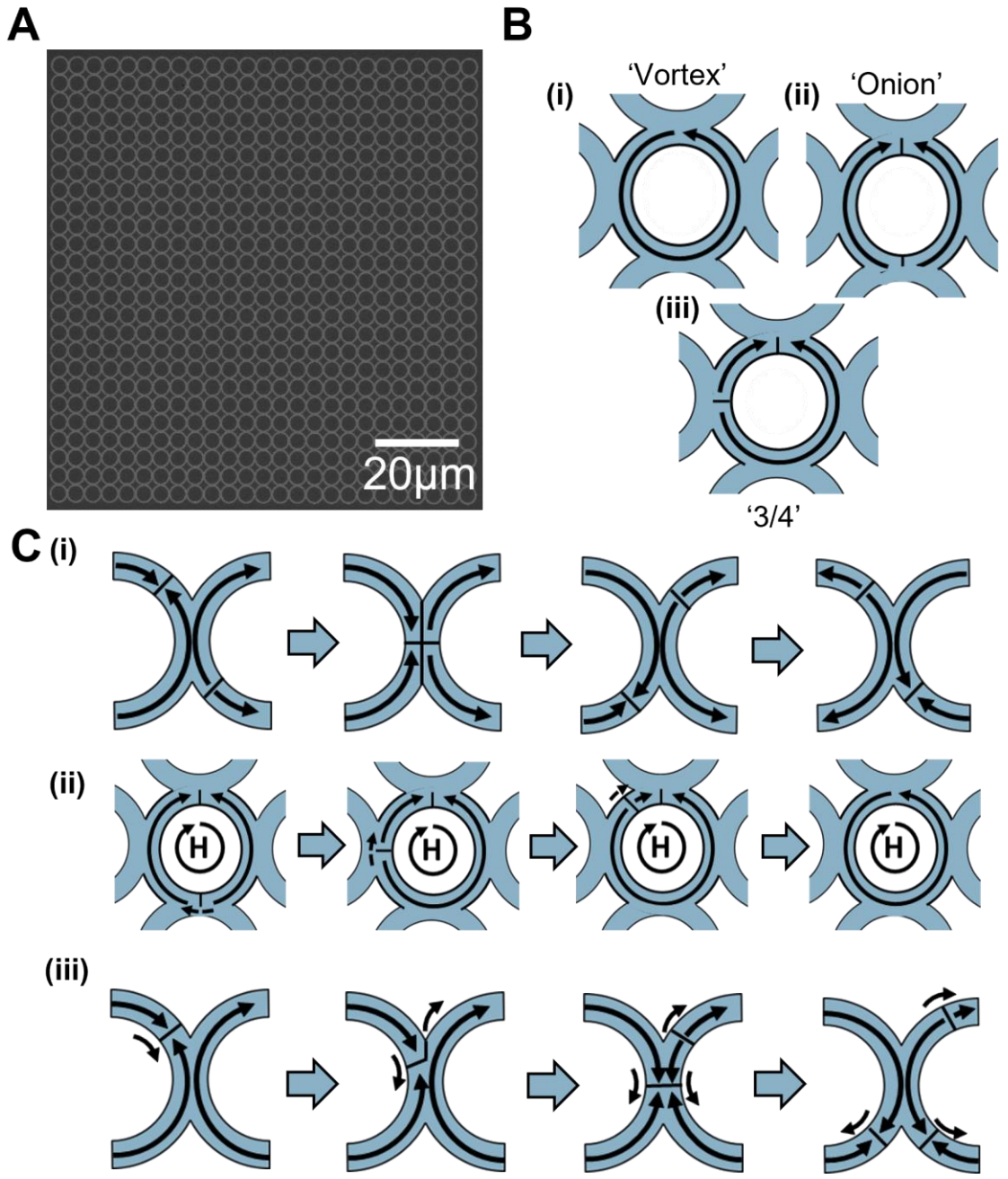


Figure 1: A- Scanning electron micrograph showing a 25x25 magnetic nanoring array. B- Schematic diagrams showing the three metastable domain states in an individual ring: (i)- Vortex configuration, (ii)- Onion configuration, and (iii)- Three-Quarter configuration. Arrows show direction of magnetisation, and lines normal to the circumference of the rings reflect the position of domain walls. C- Schematics showing the outcomes of different stochastic propagation events: (i)-Coherent propagation of domain walls with sufficiently strong magnetic fields, (ii)-Stochastic propagation of domain walls within a ring, with the upper domain wall becoming locally pinned before being annihilated by the propagating lower domain wall, and (iii)-The renucleation process when an itinerant domain wall causes local magnetic reversal, injecting a pair of domain walls into an adjacent ring. Arrows show magnetisation direction, with lines reflecting domain walls. Dotted arrows included to show propagation direction of domain walls, and large blue arrows represent the progression of time.

Domain walls may be reintroduced into the system when propagating domain walls lead to magnetic reversal across a junction, resulting in magnetic frustration between the two neighbouring rings. To alleviate this frustration, a pair of domain walls are nucleated in the ring adjacent to the propagating domain wall, shown in Figure 1c(iii). Between these mechanisms for annihilation and renucleation, a dynamic equilibrium is created between the rate of domain wall loss and gain, depending upon the relative probabilities for domain walls to propagate beyond a junction for a given applied field.

For very low or very high applied fields, the deterministic pinning/propagation respectively leads to few collisions, and hence the array exists as mainly onion or three quarter rings. For intermediate applied fields, the stochastic movement of domain walls leaves the array in a mixture of states from all three configurations depending upon the relative rates of collision and renucleation. Since the domain state of each ring is determined by the outcomes of the pinning events, the array behaviour can be approximated via simulation of these events. To achieve this, we use empirically verified relationships to calculate expected probabilities of domain walls in the system propagating beyond pinning sites, described in detail in the next section.

Modelling Stochastic Pinning Events

In magnetic materials, thermal energy introduces stochastic domain wall motion via the random fluctuation of individual magnetic moments which assist reversal processes. This results in a finite expected timescale for a reversal event to occur, depending upon the size of the associated energy barrier and the temperature of the system. Empirically, the Arrhenius-Néel relationship calculates the characteristic timescale of reversal via equation (5.1):

$$\tau_r = \tau_0^{\frac{\Delta E}{k_B T}} \quad (5.1)$$

where τ_r represents the expected reversal timescale, τ_0 the attempt frequency associated with the material (here taken to be ≈ 1GHz for $Ni_{80}Fe_{20}$ [27]), ΔE the magnitude of the effective energy barrier, k_B the Boltzmann constant, and T the temperature of the system. This relationship has been experimentally verified for the reversal of single magnetic domains, with excellent agreement [28].

As well as the temperature of the system, external magnetic fields also influence the outcome of pinning events by modulating the magnitude of the effective energy barrier. Previous work has shown that this modulation is dependent upon the component of applied field acting transverse to the domain wall axis, with the domain walls having lowest Zeeman energy when aligned with the field vector [29, 30]. The rotating magnetic fields used to drive the ring arrays means that this transverse component, $H_{transverse}$ is dependent upon the magnitude of the applied field, $H_{applied}$, and the sine of the angle between the applied field and the domain wall, ϑ_{lag}, shown schematically in Figure 2a and described mathematically via (5.2):

$$Htransverse = Happlied * sin(\theta lag) \quad (5.2)$$

The relationship between the transverse field, the magnitude of the modulated energy barrier, ΔE, and the initial energy barrier, E_0, is given via the phenomenological Sharrock equation [31]:

$$(5.3) \quad E = E_0 \left(1 - \frac{H_{\text{transverse}}}{H_{\text{sw}^0}}\right)^{\alpha}$$

where $H_{sw}{}^0$ represents the zero-Kelvin switching field of the magnetic element, and α is a geometrical constant, (here taken to be 3/2 [32]).

It has been observed in previous works that the presence of either one or two domain walls across a junction leads to a difference in domain wall structure [16], each with different energetic properties, and hence different depinning behaviours. Here, these interactions across rings have been modelled as a linear scaling of the energy barrier in the single domain wall case, shown in equations (5.4) & (5.5), and described in more detail in section 5.

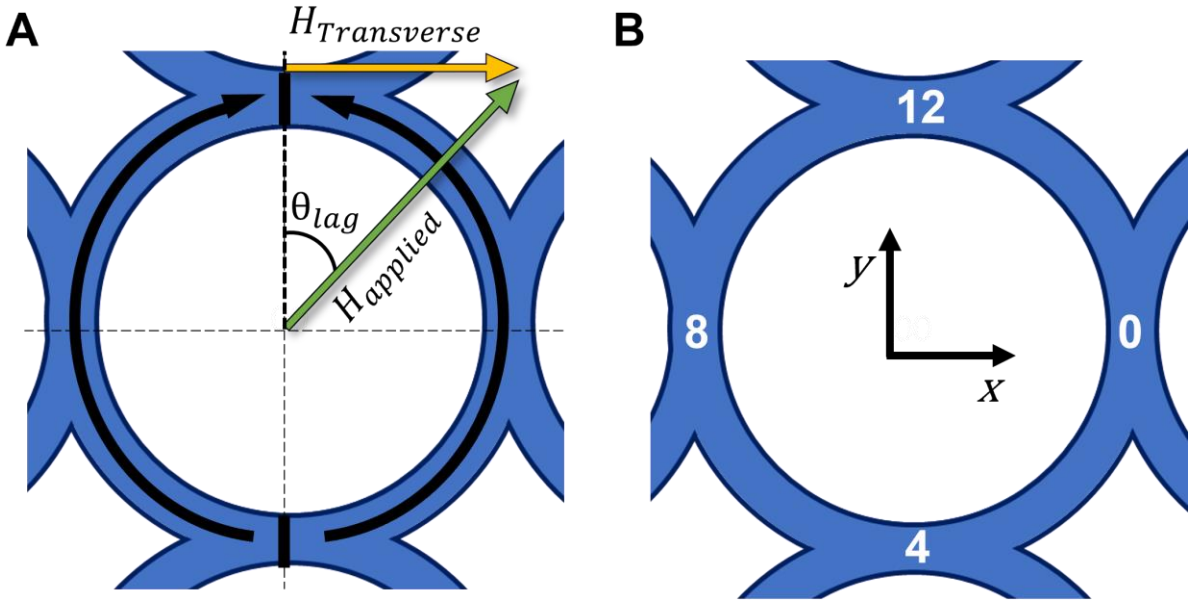


Figure 2: A- Schematic diagram showing the calculation of $H_{transverse}$ from an applied field $H_{applied}$ and the angular lag between the domain wall location and the direction of applied field, ϑ_{lag}. B- Schematic diagram representing the cardinal X and Y axes directions, as well as the corresponding junction indices at each of the intersections between the rings for $N_{segment}$ = 16.

$$(5.4) \quad H^0_{2\text{DW}} = k * H^0_{1\text{DW}}$$

$$(5.5) \quad E^0_{2\text{DW}} = k * E^0_{1\text{DW}}$$

where $H_1{}^0{}_{DW}$ and $H_2{}^0{}_{DW}$ represent the switching fields for one and two domain wall cases respectively, $E_1{}^0{}_{DW}$ and $E_2{}^0{}_{DW}$ the equivalents for initial energy barriers, and k a fixed scaling parameter, the fitting of which is discussed in section 5.

In order to approximate the varying $H_{transverse}$ (and consequently E_0) as the applied field rotates, the field is simply discretised into a series of angular steps, each held for a duration of $t_{\text{step}} = \frac{N_{\text{steps}}}{2\pi f}$ seconds, where N_{steps} represents the number of discrete steps the full rotation is split up into, and

f represents the rotational frequency of the applied field. This allows the approximated reversal probability P_{depin} for a given step to be calculated from equations (5.1 – 5.5) via equation (5.6):

$$(5.6)\quad P_{\text{depin}} = 1 - e^{-\frac{t_{\text{step}}}{\tau_r}}$$

Phenomenological Modelling of Magnetic Nanoring Arrays

Phenomenological descriptions of domain walls and their interactions allowed a simple method for programming the complex experimentally observed behaviours of collision, nucleation, and interaction. Parameters of the model are grouped via the properties they represent; array, ring, junction, and initialisation parameters have labels A, R, J, and N respectively.

The rings within the model are represented by vectors of length N_{segment}, where each entry to the vector represents a ring segment of arc length $2\pi/N_{\text{segment}}$ radians. The index of each of the entries to this vectors represents position within the ring, rotating clockwise from the positive x direction, shown in Figure 2b. Here, a value of 16 for N_{segment} was selected to provide a good trade-off between approximating a smooth rotation of field and matching the fourfold symmetry of the array, whilst keeping the number of simulation steps low.

The domain state of the nanoring array is expressed as an array of vectors A_{DW} of dimensions $[N_{\text{segment}} \times (N_r)^2]$, where N_r represents the number of rings in each row of the square array. Domain walls are instanced into the array by labelling an index in each ring vector with either a +1 or a -1, reflecting head-to-head and tail-to-tail domain walls respectively. Since many of the key behaviours of the ring array are determined by the points of interconnectivity between the rings in the array and the interactions that arise at them, three separate vectors of length $2N_r(N_r - 1)$ are created which record the of every junction in the network: J_{DW}, which tracks the number of domain walls occupying each junction, J_E, which reflects the magnitude of the energy barrier E_0 presented by each junction, and J_H, which reflects the zero-kelvin switching field H_{sw}^0 for each junction, as described in the previous section.

The state of the simulated nanoring array is initialised by instancing a head-to-head domain wall in every ring in A_{DW} at index $Nsegment/4$, and a tail-to-tail domain wall in every index $3 \times Nsegment/4$, corresponding to the positive/negative y direction respectively, and emulating the saturated state in the positive y direction. The magnetisation state of the array is then generated from the position and variety of all domain walls in the system. Firstly, an additional array, A_{dir}, of identical shape to A_{DW} is generated, which is used to keep track of the direction of magnetisation in each segment of the rings. Each index of A_{dir} records whether the magnetisation runs clockwise (+1) or anticlockwise (-1) over each segment, and is marked zero in the locations of domain walls. From this direction matrix, the magnetisation is calculated in terms of components in the x and y axes (M_x and M_y respectively) via:

$$(5.7)\quad M_x = \sum_{i=1}^{N_r^2} \sum_{s=1}^{N_{\text{segment}}} sin\left(\frac{2\pi s}{N_{\text{segment}}}\right) A_{dir}^{i,s}$$

$$(5.8)\quad M_y = \sum_{i=1}^{N_r^2} \sum_{s=1}^{N_{\text{segment}}} cos\left(\frac{2\pi s}{N_{\text{segment}}}\right) A_{dir}^{i,s}$$

This gives the magnetisation of the array in arbitrary units, which is then normalised against the magnetisation in the saturated state, determined as the value of M_y in the initialised array.

Figure 3a shows a high-level overview of the how domain state of the model evolves: Firstly, the external applied field is moved by a fixed angular step. The relative probabilities of all domain walls in the system depinning are then calculated via equations (5.1) to (5.6), before comparison with random variable drawn from a uniform distribution between 0 and 1 to determine which domain walls are free to propagate. The domain walls then propagate towards their respective energy minima parallel/anti-parallel to the field vector, with interactions programmed phenomenologically before finalising their positions and restarting the process. These processes will be discussed in detail next.

Figure 3b shows a flowchart for the process of calculating depinning probabilities. In RingSim, two sources of domain wall pinning are considered. Firstly, the effects of edge roughness are included by imposing a threshold field, below which domain wall propagation does not occur [33]. It has been experimentally observed that under low applied fields, very few domain walls depin even after many cycles[16]. Additionally, the domain walls that do depin seem to consistently do so after repeated cycles, suggesting that the threshold field for propagation is a distributed parameter across different rings due to manufacturing inconsistencies. To recreate this behaviour, a vector representing the edge roughness threshold field for each ring, R_{ER}, is created by sampling from a normal distribution with a fixed mean and standard deviation to resemble the expected variance in properties via manufacturing imperfections in experimental samples. The exact mean and variance were determined via correlation with experiments and discussed in more detail in section 5. This acts as a hard threshold by setting depinning probability to zero if $H_{transverse}$ (via equation (5.2)) is below the value of R_{ER}^{i} for a given ring i.

The second source of pinning comes from the energy barrier associated with the junctions between rings. Domain walls occupying the same junction are considered as coupled in RingSim, with the calculation of reversal probabilities occurring only once with coupled outcomes for both domain walls. The number of domain walls at each junction is given by the entry in J_{DW} for a particular junction. Each junction has its own energy barrier and switching field, stored in the vectors J_E and J_H, which are scaled for domain wall-domain wall interactions via equations (5.4) &(5.5) if two domain walls occupy the junction. Next, the effects of the external field are accounted for by scaling the energy barrier via equation (5.3), giving the magnitude of the effective energy barrier ΔE. As the depinning process is thermally activated, the expected transition time for reversal can be determined via equation (5.1) for a given temperature of the system. The rotational frequency of the applied field is then used to give an expected probability of reversal for a given angular step, described via equation (5.6).

Figure 3c shows a flow chart for stochastic depinning of the domain walls in the model. A random number sampled from a uniform distribution between zero and one, x, is generated for each junction, and is compared to P_{depin}, with the domain walls occupying any junction where $P_{depin} > x$ deemed free to propagate. The process for domain wall propagation is outlined in figure 3d. Here, all depinned domain walls propagate along their shortest route to their respective energy minima. Firstly, the magnetisation states of any propagating domain walls that occupied a junction are compared with the adjacent rings. If the subsequent propagation leads to magnetic frustration across the junction, then the nucleation process occurs. This begins by initialising a pair of domain walls, one head-to-head and one tail-to-tail, at the junction in the adjacent ring. One of these domain walls moves to the applicable energy minimum that occupies one of the quarter sections adjacent to the junction. The other domain wall is then flagged for an additional depinning check at the junction between where it was initialised and its energy minimum. In the

case where both minima lie at the junctions 90 degrees to the initialisation junctions, the domain walls propagate to their respective minima where they remain until the next field step.

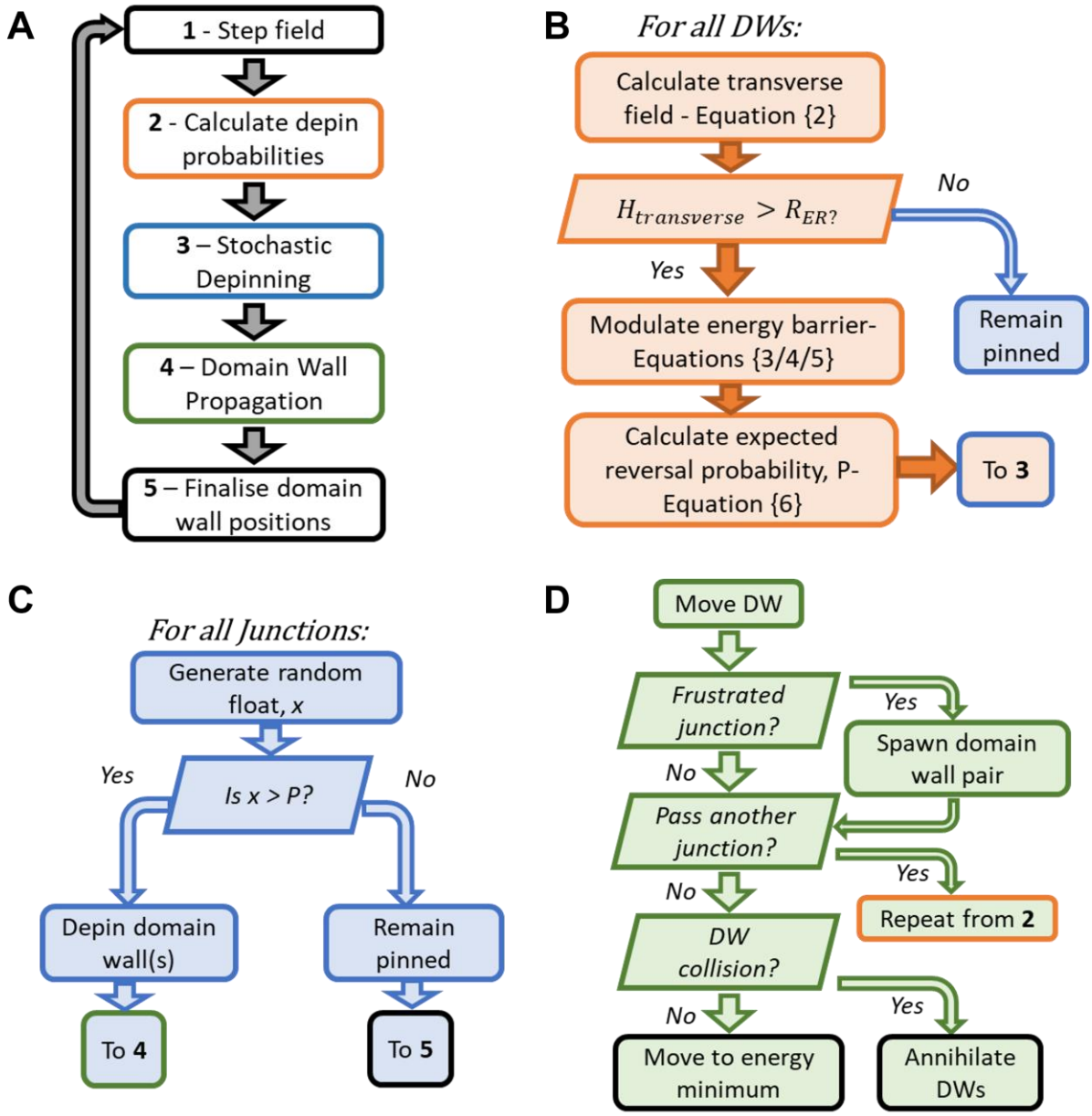


Figure 3: A- Overview of the key steps the taken in the modelling procedure. B-Process for determining the probabilities P for each of the domain walls to propagate within the model. C-Process for deciding the stochastic outcomes of pinning events based on previously calculated probability P. D- Process for determining the next state of the array from the outcome of depinning events, and any collisions or additional junctions that may be passed on the domain wall's path to the energy minimum.

For domain wall propagation in a given ring, the domain walls propagate towards their energy minimum until they meet one of three outcomes: if they collide with another domain wall before reaching their respective minimum, the domain walls annihilate, and leave the ring in a vortex state, with the chirality depending upon the itinerant domain wall. If the domain wall reaches

another junction before its energy minimum, then the domain wall is flagged for an additional depinning check at that junction since it may become locally pinned again. Finally, if the domain wall reaches its energy minimum, it remains there until the next field step.

The process of pinning checks and subsequent propagation then reoccurs, before finalising domain wall positions for the field step. Junctions are then checked for the number of domain walls occupying them, and the J_{DW} vector is updated accordingly. The magnetisation state of the array is then calculated from the final positions and returned in terms of normalised M_x and M_y components, and all state matrices are updated. The field then moves another angular increment, and the process starts over.

Fitting Model Parameters

In order to tie the predictions of RingSim to real-world devices and their responses to external fields and temperatures, a series of experimental and modelling techniques were used in order to assert many of the model parameters, with the remaining free parameters iteratively tuned by fitting the model's magnetisation output to experimentally gathered data. Firstly, the values of τ_0 and α were taken from literature, as an attempt frequency of 1 GHz for Permalloy [27], and an alpha value of 3⁄2 [32].

Micromagnetic simulations using the MuMax3 [6] software package were performed on a pair of overlapping half-rings, representing a single junction, but extending to ring properties via symmetry. The dimensions of the simulated rings reflected the geometry of the manufactured devices which were later used for further characterisation, with nominal diameters of 4μm, track widths of 400 nm, thicknesses of 10 nm, and with each ring overlapping 50% of its track width with its neighbour. Material parameters of the system were set to reflect $Ni_{80}Fe_{20}$ in line with the manufactured devices (M_s = 860 kA/m, A_{ex} = 13 pJ/m, α_G = 1, an artificially large damping parameter to speed up simulation). The simulations were used to establish the effects of domain wall-domain wall interactions on the switching field of the junction in absence of thermal effects. The simulations, shown in Figures 4A/B were initialised with a single domain/two domains occupying the junction respectively, and field is ramped in 1 Oe increments every 8 nanoseconds. The domain walls were deemed to depin at the field when they became fully delocalised from the junction, and the strength of the applied field recorded, with depinning fields of 79 Oe and 64 Oe for the one and two domain wall cases respectively. While these results reflect the zero-kelvin switching field for an idealised material, imperfections from the manufacturing process (lower true saturation magnetisation, imperfect geometry, presence of grains etc.) mean these values will not be numerically identical to those of a manufactured device. However, these values were used to be indicative of the ratio between the two processes, and hence used to determine the value of k in equation (5.4). The effects on the energy barrier are assumed to be equivalent, giving a k value of 0.81 for equations (5.4) and (5.5).

The remaining free parameters of R_{ER}, E_0, and H_{sw}^0 were fit to magnetoresistance (AMR) measurements of the nanorings. As introduced in previous works [18], the AMR response of the nanoring array has two distinct responses to rotating fields, one at the same frequency of the rotating field (1f response) and another at twice the field frequency (2f response). The 1f response occurs due to susceptibility effects, with elastic stretching and contraction of the domain walls in the system in response to the rotating field, with MuMax3 simulations of this

effect shown in Figure 4c. The 2f signal depends upon the propagation of domain walls as they move between junction sites where pinned domain walls sit either orthogonal or parallel to the current density, shown in Figure 4d.

Figure 5a shows the relative magnitudes of the Fourier components of the AMR response of the 1f and 2f frequency components across a range of applied fields over 30 rotations. Two key features of this response were used to fit model parameters. Firstly, the end of the linear regime of the 1f response reflects the onset of domain wall motion, as the change from linear increase is due to the addition of incomplete propagation of domain walls around the rings. This allows determination of the R_{ER} at 20.5 Oe. Secondly, the magnitude of the 2f signal reflects the number of domain walls propagating in the

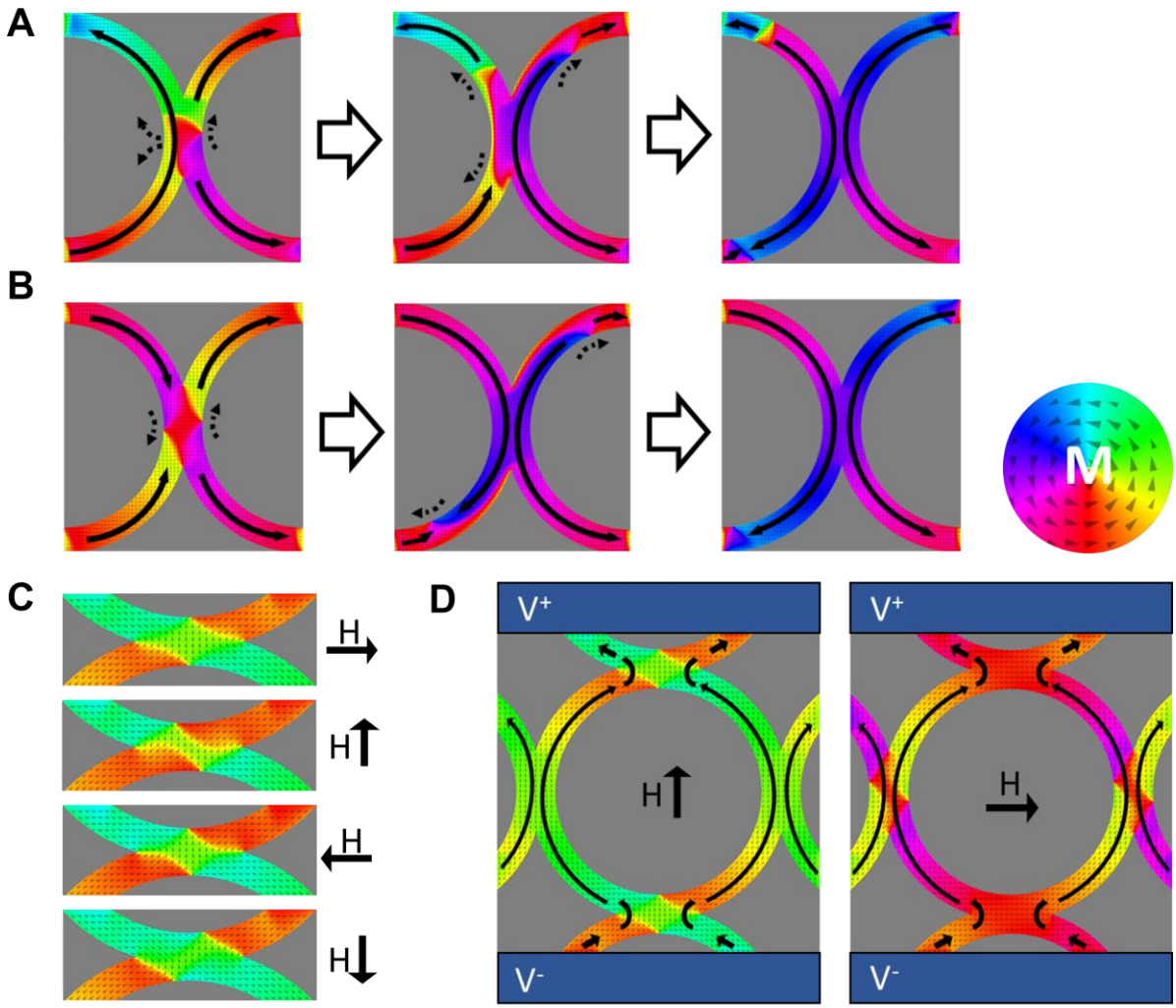


Figure 4: A/B- Domain wall reversal processes for one and two domain wall cases respectively, produced using MuMax3. Black arrows show direction of domains, white arrows reflect increasing of the applied field. CMumax simulations of domain wall structure across a junction during applications of 20 Oe rotating field in $\pi/2$ radian steps, showing the expansion and contraction of the domain wall via magnetic susceptibility. DMuMax simulations of domain state with domains pinned in orientations parallel to current density (left) and orthogonal to current density (right). Black arrows reflect flow of current density. Colour wheel reflects direction of magnetisation in all plots, with local direction shown by grey triangles on the colour wheel.

system. This is proportional to the amplitude of the magnetisation response of the array in the absence of susceptibility effects, which is given by M_x and M_y in RingSim. This equivalence allows tuning of the remaining free parameters of E_0 and H_0 by comparing the amplitude of the magnetisation output in RingSim to the experimental 2f data for the same applied fields, and selecting E_0 and H_0 values which provide the same response.

The experimental procedure used to generate the 2f data (30 rotations of applied fields between 25-30 Oe) were repeated in RingSim, for simulations initialised with a range of E_0 and H_0 values. Figures 5b and 5c show the magnitude of the magnetisation response over a few example E_0 and H_0 pairs compared with the device's 2f response, and the mean-squared error between the simulated magnetisation and the experimental response across all E_0 and H_0 pairs respectively. It can be observed that a region of E_0 and H_0 pairs are able to fit the experimental data well, reflected by the region of low mean-squared error in Figure 5c.

Since temperature also modulates the relative depinning probabilities which determine the number of propagating domain walls, the AMR response of the experimental system over a range of applied temperatures must also be determined to find the specific E_0 and H_0 pair that describes the system best. The temperature of the system was controlled by mounting the device on a Peltier cell, with temperature measured via a pyrometer positioned above the device. From these measurements, a linear shift in the point of maximum gradient of the 2f response was observed and shown in Figure 5d. Similar to the previous fitting process, these experiments were repeated within RingSim, and the gradient of this linear shift calculated across a range of E_0 and H_0 pairs and compared to the experimentally gathered data. Crucially, the change in thermal gradient shown in figure 5e is different when compared to the 2f fit shown in figure 5c. This allowed determination of which of E_0 and H_0 pairs satisfied the response of the experimental system best across all temperatures via generation of a fitting metric from the combined errors between the field dependent and the temperature dependent measurements, shown in figure 5f. An optimal E_0 and H_0 pair was chosen which reconciles both experiments, here determined to be $H_0 = 55$ Oe and $E_0 = 2.625 \times 10^{-19}$ J.

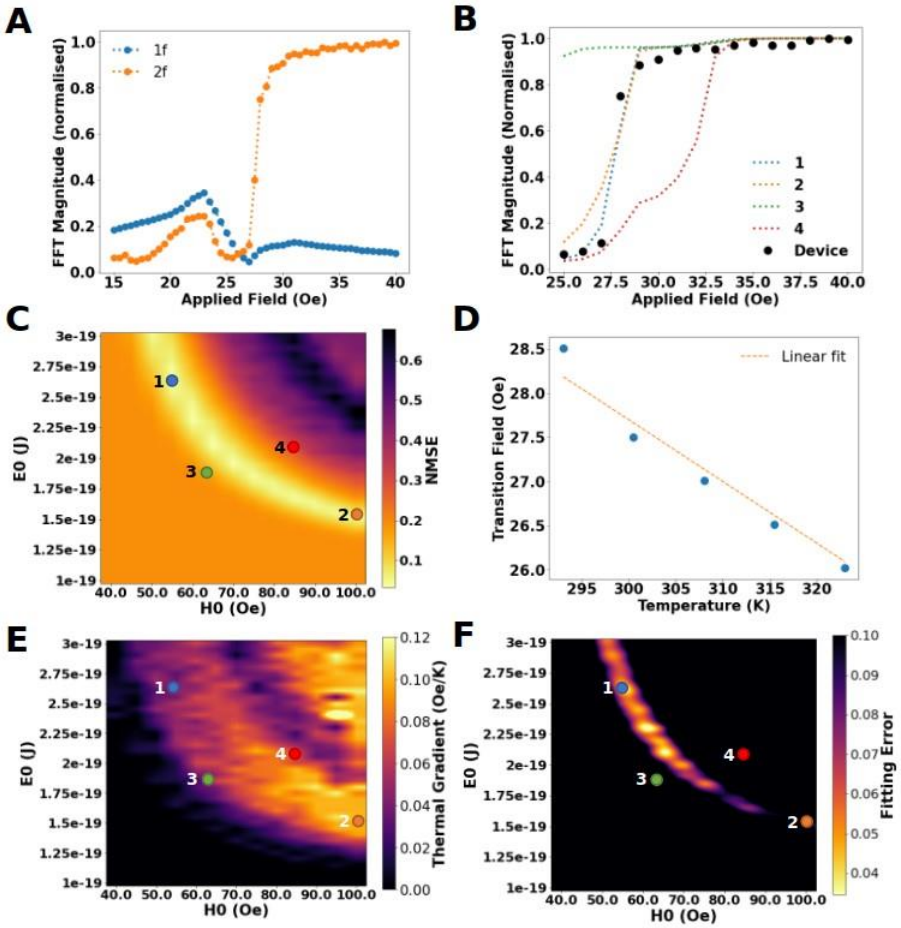


Figure 5: A- Normalised magnitudes of Fourier components corresponding to the clock frequency (1f) and twice the clock frequency (2f) over 30 rotations of varying magnetic field. B- Example fits of RingSim magnetisation output to the 2f transition in experimental device. The four coloured lines reflect different E_0/H_0 pairs, shown in the map in C. C- Colourmap of mean-squared error between model's reproduction of magnetisation and experimentally gathered response via AMR for all explored E_0/H_0 values. D-Plot of extracted transition fields by varying the temperature of the magnetic ring array. Transition field determined from the point of maximum gradient in the 2f response, with linear fit used to extract the change in transition field with temperature shown. E-Colourmap showing gradient of transition field with respect to temperature, shown for all E_0/H_0 pairs. F- Combined fitting metric made by combining difference in gradient of transition field between simulated system and experimental system and the mean-squared error between the simulated 2f transition and experimental data.

Validating the Model

The previous section outlined a procedure for fitting model parameters to the equilibrium response of the dynamic system of magnetic nanorings. However, the ring arrays are known to exhibit dynamic responses to input, as well as different populations of the three domain states (Figure 1B) with respect to driving field [16]. To confirm the model encapsulates these behaviours, the model's predictions need to be validated against additional experimental data that measure the dynamic timescales of the system's response, as well as the populations of domain states observed experimentally.

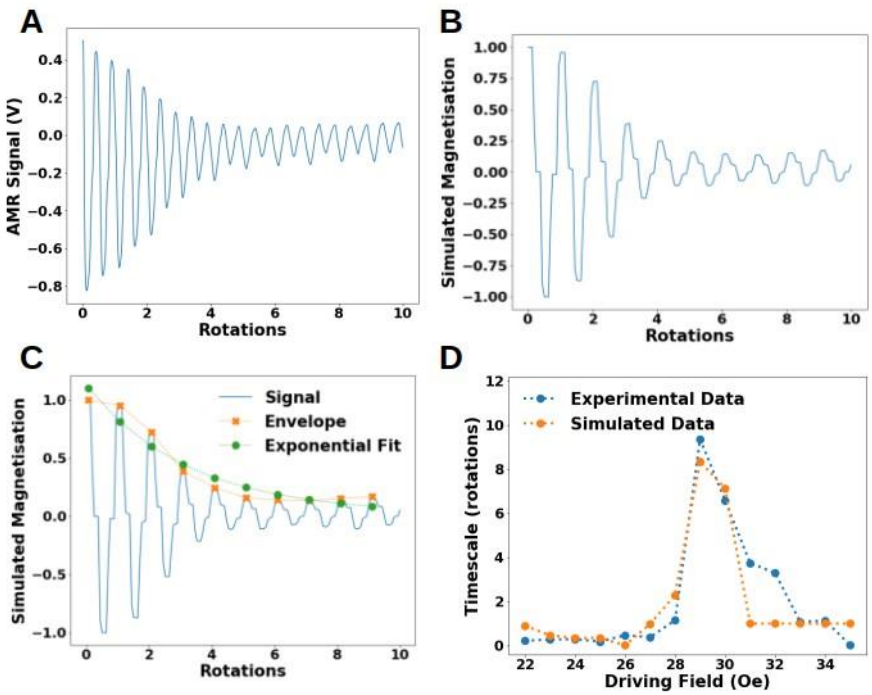


Figure 6: A- Measured voltage signal from experimental measurement of the ring devices under 10 rotations of 28 Oe applied field. B- Simulated magnetisation response via RingSim for 10 rotations of 28 Oe applied field. C- Outline of the procedure for measuring dynamic timescales. First, the envelope of the underlying signal is generated by marking the maximum magnetisation over a rotation of applied field, shown by the orange markers. Then, an exponential fit is generated to replicate the envelope of the magnetisation/AMR signal, shown in green. D- Comparison of the resulting decay timescales for experimentally gathered data (blue), and the simulated magnetisation output (orange). Timescale is presented with respect to number of rotations rather than in time.

To establish the dynamic behaviours of the physical system, further AMR measurements were performed to determine the rate at which the ring array reaches equilibrium response from a saturated state over a range of applied fields. In order to measure these dynamic timescales τ_d, the envelope of the AMR was calculated over successive rotations, and an exponential function of the form $X(t) = X_0 - ae^{-}\frac{t}{\tau_d} + c$ was fitted to the resulting decay curve. This was compared to a similar exponential fit to the envelope of the magnetisation signal generated by RingSim.

Figure 6a/b shows comparisons between the AMR response of the physical device at 29 Oe (the longest decay timescale) and the equivalent magnetisation response generated by RingSim. Although the RingSim magnetisation response lacks the period-doubling effect seen in the AMR response, functionally the signals are very similar in terms of decay time and steady-state level. Figure 6c shows the generated exponential fits for an example wave, while Figure 6d shows the comparison between the fitted τ_d parameters across these fields for both simulated and experimental data. There is excellent agreement between the two at lower and intermediate fields, showing that RingSim effectively simulates the regions of highly stochastic propagation well. However, there are longer timescales observed 30-32 Oe in experiments than in simulation, with the simulation predicting the equilibrium amplitude is reached instantaneously. This suggests that RingSim underestimates the field at which uniform propagation occurs.

In order to determine the microstates formed by the ring array in response to various applied fields, magnetic contrast micrographs were generated using X-ray photo-emission electron

microscopy (X-PEEM) on subsections of the nanoring array. To validate these behaviours in RingSim, a visualisation tool was developed which emulates the magnetic contrast observed in X-PEEM according to the values of M_y in each segment of all of the rings in the simulated system. Figure 7a shows a comparison between data generated via X-PEEM and RingSim's visualisation tool after 30 rotations of a range of applied fields. Similar grouping of larger domains locally reflects a similar tendency for the domain wall-domain wall interactions in both the physical device and within RingSim to lead to local regions of magnetic order. However, there seems to be increased disorder at lower fields for the outputs generated with RingSim.

From the position and location of domain walls shown in images such as in figure 7a (and at other fields not presented in the figure), it is possible to determine the relative proportions of vortex, onion, and three-quarter rings over the subsection of the array after 30 rotations of field over a range of magnitudes. The application of field was similarly repeated within RingSim, and the relative populations of each ring state counted. Figures 7b and 7c show the variation of these populations for simulated 25x25 ring arrays and experimentally gathered data on a sub-sample of the 25x25 array respectively. While there is some variance between the relative populations observed in experiments and simulated by RingSim, the general trends are similar given significantly fewer rings sampled in the experimental data.

One possible source of the differences between the X-PEEM measurements and RingSim model is that the sample used for X-PEEM is different to the sample used for fitting the model. Although the nominal dimensions of the arrays' designs were the same, they were manufactured in different lithography and deposition runs, which could lead to some slight variation in ring width/thickness, accounting for the slight shifts in field. Additionally, remanence in the iron cores of the electromagnets used to generate the applied fields in the experimental data may have led to slightly asymmetrical field rotations, which could explain a biasing in the formation of vortex states in the experimental data which is not seen in the model. In combination with the previous results, we believe this shows good agreement between the microstates formed in experiments and simulated via RingSim, validating that the simulation of depinning processes that RingSim is based upon provide a good description of the overall processes that dictate the response of the magnetic nanoring arrays.

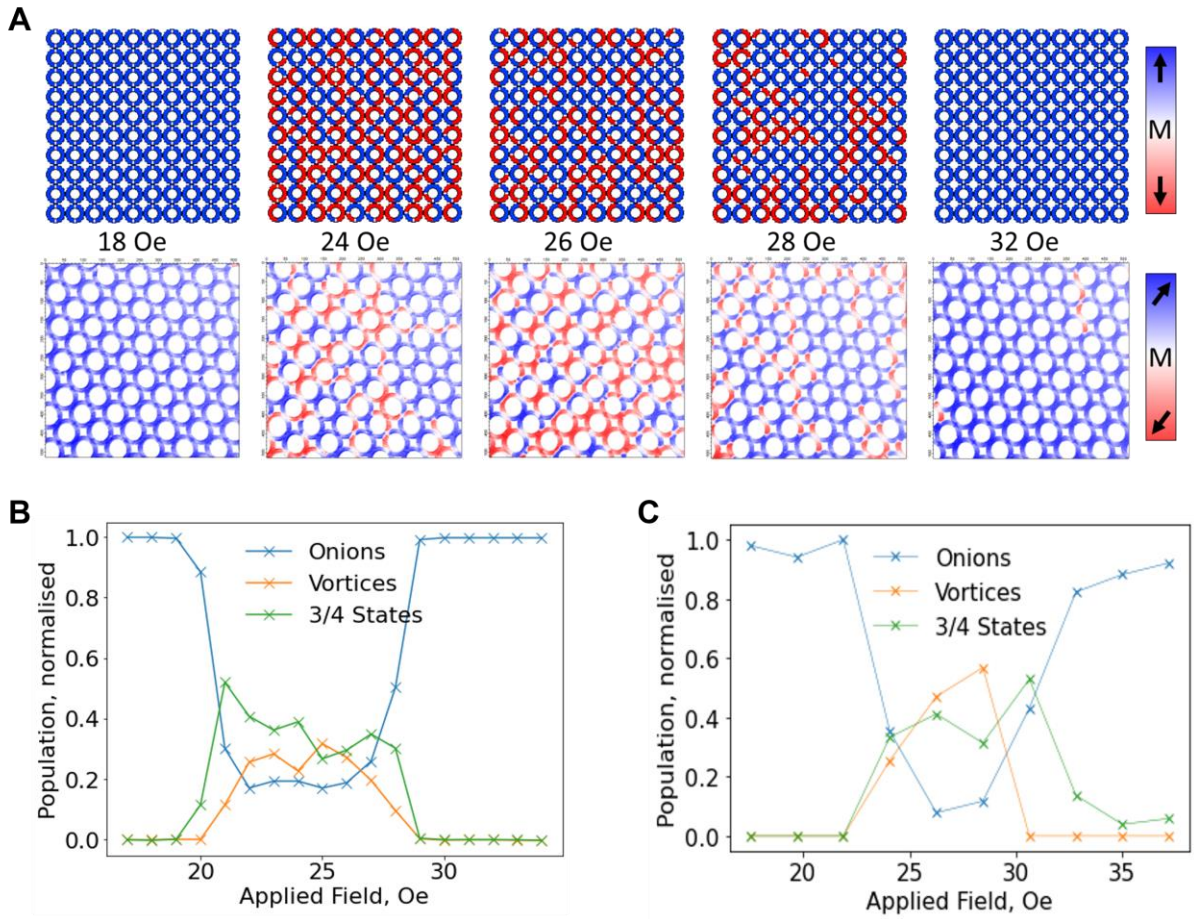


Figure 7: A- Comparison of magnetisation state between ring structures generated via RingSim (upper) and experimentally gathered data via X-PEEM imaging (lower) for applied fields between 18 and 32 Oe. In both cases, colour reflects magnetisation direction along the vertical axis, reflected in the colour bar on the right. State count of the three different ring domain configurations (onion, vortex, and three-quarter) with respect to the magnitude of the simulated rotating field. Generated over a 25x25 array, normalised against ring number. D-State count of ring domain configurations with respect to the magnitude of current provided to driving electromagnet [34]. Generated over subsection of 25x25 ring array containing 40-50 nanorings, normalised against ring number for a given image.

Conclusions

In this paper, we have outlined a methodology for modelling extended nanowire networks. With a combination of numerical models for calculating reversal probabilities and directly programmed phenomenological behaviour, RingSim was able to emulate many of the behaviours observed in real nanoring devices.

The resulting model provides excellent agreement with experimentally gathered data, not only in the equilibrium response of the system, but also with the dynamic timescales associated with reaching equilibrium as well as the typical domain microstates that are formed across different driving fields. RingSim is able to model relatively large areas of magnetic materials with modest computational overheads, allowing predictions to be made for device-level responses that would be practically impossible to achieve with general-purpose micromagnetic simulators. To draw a rough comparison, RingSim is able to simulate the response of a 25-by-25 array of nanorings at

speeds of 1.4 rotations per second on an Intel i7 processor, allowing rapid evaluation of device response to arbitrary field inputs.

While the exact formulation and phenomenology featured within RingSim is specific to the system of interconnected magnetic nanorings, we believe that the general methodology of reducing the simulated magnetic processes to the modelling key agents within the system, and programming interactions phenomenologically, can be applied to many other similar systems, such as connected artificial spin-ice networks, domain wall logic networks, etc. Simulation tools such as RingSim can provide the opportunity to perform rapid exploration of device properties for applications such as neuromorphic computing, especially for systems with large parameter spaces to explore that would be infeasible to measure experimentally.

Author Contributions

ITV designed and programmed RingSim and performed all simulations. DG helped optimise the speed of the model. ITV, GV, and CS designed the setup for performing AMR measurements. ITV performed all AMR measurements. GV, CS, and PWF designed and manufactured all samples. ITV, RMRR, AW, TJH, and DAA performed X-PEEM measurements, which were overseen by DB, FM, and SSD. DAA and TJH conceptualised the work.

Acknowledgements

The authors thank STFC for beam time on the Offspec beamline at the ISIS Neutron and Muon Source (https://doi.org/10.5286/ISIS.E.RB1810656) and on beamline I06 at the Diamond Light Source, and thank Jordi Prat, Michael Foerster and Lucia Aballe from ALBA for providing quadrupole sample holders [34]. R.W.D., T.J.B., and I.T.V. acknowledge DTA-funded PhD studentships from EPSRC. The authors gratefully acknowledge the support of EPSRC through grant EP/S009647/1, EP/V006339/1 and EP/V006029/1. This project has received funding from the European Union's Horizon 2020 FET-Open program under grant agreement No 861618 (SpinEngine).

References

[1]J. Torrejon, M. Riou, F. A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, and J. Grollier, "Neuromorphic computing with nanoscale spintronic oscillators", Nature **547**, Publisher: Nature Publishing Group, 428–431 (2017).

[2]T. Kanao, H. Suto, K. Mizushima, H. Goto, T. Tanamoto, and T. Nagasawa, "Reservoir computing on spin-torque oscillator array", Physical Review Applied **12**, Publisher: American Physical Society, 024052 (2019).

[3]A. Pivano and V. O. Dolocan, "Chaotic dynamics of magnetic domain walls in nanowires", Phys. Rev. B **93**, 144410 (2016).

[4]R. V. Ababei, M. O. A. Ellis, I. T. Vidamour, D. S. Devadasan, D. A. Allwood, E. Vasilaki, and T. J. Hayward, "Neuromorphic computation with a single magnetic domain wall", Scientific Reports 2021 11:1 **11**, Publisher: Nature Publishing Group, 1–13 (2021).

[5]A. Welbourne, A. L. R. Levy, M. O. A. Ellis, H. Chen, M. J. Thompson, E. Vasilaki, D. A. Allwood, and T. J. Hayward, "Voltage-controlled superparamagnetic ensembles for low-power reservoir computing", Applied Physics Letters **118**, Publisher: AIP Publishing, 202402 (2021).

[6]A. Vansteenkiste, J. Leliaert, M. Dvornik, M. Helsen, F. Garcia-Sanchez, and B. Van Waeyenberge, "The design and verification of MuMax3", AIP Advances **4**, Publisher: American Institute of Physics, 107133 (2014).

[7]M. J. Donahue and D. G. Porter, *Oommf user's guide, version 1.0*, 1999.

[8]L. Landau and E. Lifshitz, "On the theory of the dispersion of magnetic permeability in ferromagnetic bodies", in *Perspectives in theoretical physics* (Elsevier, 1992), pp. 51–65.

[9]T. L. Gilbert, "A lagrangian formulation of the gyromagnetic equation of the magnetization field", Phys. Rev. **100**, 1243 (1955).

[10]X. Chen, F. A. Araujo, M. Riou, J. Torrejon, D. Ravelosona, W. Kang, W. Zhao, J. Grollier, and D. Querlioz, "Forecasting the outcome of spintronic experiments with neural ordinary differential equations", Nature Communications **13**, Number: 1 Publisher: Nature Publishing Group, 1016 (2022).

[11]J. Leliaert and J. Mulkers, "Tomorrow's micromagnetic simulations", Journal of Applied Physics **125**, 180901 (2019).

[12]J. H. Jensen and G. Tufte, "Reservoir computing with a chaotic circuit", in (Sept. 1, 2017), pp. 222–229.

[13]J. H. Jensen and G. Tufte, "Reservoir computing in artificial spin ice", in (July 1, 2020), pp. 376–383.

[14]J. C. Gartside, K. D. Stenning, A. Vanstone, H. H. Holder, D. M. Arroo, T. Dion, F. Caravelli, H. Kurebayashi, and W. R. Branford, "Reconfigurable training and reservoir computing in an artificial spin-vortex ice via spin-wave fingerprinting", Nature Nanotechnology **17**, 460–469 (2022).

[15]K. Hon, Y. Kuwabiraki, M. Goto, R. Nakatani, Y. Suzuki, and H. Nomura, "Numerical simulation of artificial spin ice for reservoir computing", Applied Physics Express **14**, 033001 (2021).

[16]R. W. Dawidek, T. J. Hayward, I. T. Vidamour, T. Broomhall, G. Venkat, M. Al Mamoori, A. Mullen, S. J. Kyle, P. Fry, N.-J. Steinke, J. Cooper, F. Maccherozzi, S. S. Dhesi, L. Aballe, M. Foerster, J. Prat, E. Vasilaki, M. O. A. Ellis, and D. A. Allwood, "Dynamically-driven emergence in a nanomagnetic system", Advanced Functional Materials **31** (2021).

[17]K. D. Stenning, J. C. Gartside, L. Manneschi, C. T. S. Cheung, T. Chen, A. Vanstone, J. Love, H. H. Holder, F. Caravelli, K. Everschor-Sitte, E. Vasilaki, and W. R. Branford, *Adaptive programmable networks for in materia neuromorphic computing*, Nov. 11, 2022.

[18]I. Vidamour, C. Swindells, G. Venkat, L. Manneschi, P. Fry, A. Welbourne, R. Rowan-Robinson, D. Backes, F. Maccherozzi, S. Dhesi, E. Vasilaki, et al., "Reconfigurable reservoir computing in a magnetic metamaterial", Communications Physics **6**, 230 (2023).

[19]D. A. Allwood, M. O. A. Ellis, D. Griffin, T. J. Hayward, L. Manneschi, M. F. K. Musameh, S. O'Keefe, S. Stepney, C. Swindells, M. A. Trefzer, E. Vasilaki, G. Venkat, I. T. Vidamour, and C. Wringe, "A perspective on physical reservoir computing with nanomagnetic devices", Applied Physics Letters **(in review)** (2022).

[20]C. M. Macal and M. J. North, "Agent-based modeling and simulation", in Proceedings of the 2009 winter simulation conference (wsc) (2009), pp. 86–98.

[21]F. Stonedahl and U. Wilensky, "Finding forms of flocking: evolutionary search in abm parameterspaces", in Multi-agent-based simulation xi: international workshop, mabs 2010, toronto, canada, may 11, 2010, revised selected papers 11 (Springer, 2011), pp. 61–75.

[22]K. E. Joyce, P. J. Laurienti, and S. Hayasaka, “Complexity in a brain-inspired agent-based model”, Neural Networks **33**, 275–290 (2012).

[23]L. Zhang, C. G. Strouthos, Z. Wang, and T. S. Deisboeck, “Simulating brain tumor heterogeneity with a multiscale agent-based model: linking molecular signatures, phenotypes and expansion rate”, Mathematical and computer modelling **49**, 307–319 (2009).

[24]I. Vidamour, M. O. Ellis, D. Griffin, G. Venkat, C. Swindells, R. W. S. Dawidek, T. J. Broomhall, N.-J. Steinke, J. Cooper, F. Maccherozzi, S. Dhesi, S. Stepney, E. Vasilaki, D. A. Allwood, and T. J. Hayward, “Quantifying the computational capability of a nanomagnetic reservoir computing platform with emergent magnetisation dynamics”, Nanotechnology, Publisher: IOP Publishing, 10.1088/13616528/ac87b5 (2022).

[25]L. J. Chang, Y. D. Yao, P. Lin, and S. F. Lee, “Magnetic interaction in domain wall depinning at square notch and antinotch traps”, IEEE Transactions on Magnetics **47**, Conference Name: IEEE Transactions on Magnetics, 2519–2521 (2011).

[26]C. W. Sandweg, N. Wiese, D. McGrouther, S. J. Hermsdoerfer, H. Schultheiss, B. Leven, S. McVitie, B. Hillebrands, and J. N. Chapman, “Direct observation of domain wall structures in curved permalloy wires containing an antinotch”, Journal of Applied Physics **103**, Publisher: American Institute of PhysicsAIP, 93906 (2008).

[27]J. W. Lau, M. Beleggia, and Y. Zhu, “Common reversal mechanisms and correlation between transient domain states and field sweep rate in patterned permalloy structures”, Journal of Applied Physics **102**, 043906 (2007).

[28]W. Wernsdorfer, E. B. Orozco, K. Hasselbach, A. Benoit, B. Barbara, N. Demoncy, A. Loiseau, H. Pascard, and D. Mailly, “Experimental evidence of the n´eel-brown model of magnetization reversal”, Physical Review Letters **78**, 1791–1794 (1997).

[29]M. Negoita, T. J. Hayward, and D. A. Allwood, “Controlling domain walls velocities in ferromagnetic ring-shaped nanowires”, Applied Physics Letters **100**, 10.1063/1.3685467 (2012).

[30]M. Negoita, T. J. Hayward, J. A. Miller, and D. A. Allwood, “Domain walls in ring-shaped nanowires under rotating applied fields”, Journal of Applied Physics **114**, 10.1063/1.4812388 (2013).

[31]M. P. Sharrock, “Time dependence of switching fields in magnetic recording media (invited)”, Journal of Applied Physics **76**, 6413–6418 (1994).

[32]J.-E. Wegrowe, J. P. Meier, B. Doudin, J.-P. Ansermet, W. Wernsdorfer, B. Barbara, W. T. Coffey, Y. P. Kalmykov, and J.-L. D´ejardin, “Magnetic relaxation of nanowires: beyond the n´eel-brown activation process”, Europhysics Letters **38**, Publisher: IOP Publishing, 329 (1997).

[33]S. Dutta, S. A. Siddiqui, J. A. Currivan-Incorvia, C. A. Ross, and M. A. Baldo, “The spatial resolution limit for an individual domain wall in magnetic nanowires”, Nano Letters **17**, PMID: 28813156, 5869– 5874 (2017).

[34]M. Foerster, J. Prat, V. Massana, N. Gonzalez, A. Fontsere, B. Molas, O. Matilla, E. Pellegrin, and L. Aballe, “Custom sample environments at the ALBA XPEEM”, Ultramicroscopy **171**, Publisher: North-Holland, 63–69 (2016).

6- “Quantifying the computational capability of a nanomagnetic reservoir computing platform with emergent magnetisation dynamics”.

6.0- Preamble

Following on from the previous chapter which outlined the development of the phenomenological model of the nanoring arrays, this chapter uses the model as a basis for exploring the application of nanoring arrays for reservoir computing. The objective of this study was to explore the range of dynamic responses exhibited by the nanoring arrays and determine the range of computational properties that were available.

In order to implement the nanoring arrays as a reservoir computing platform, a paradigm was required for encoding information into the array, as well as extracting meaningful state information from the array. Since the arrays respond strongly to rotating magnetic fields, input was directly encoded into the magnitude of the applied rotating field. To read out information that was dependent upon the magnetic state of the system, the RingSim variables of domain wall population and the components of magnetisation of the array along and normal to the saturation direction were chosen. To cover the full range of transformations available to the ring arrays, and to ensure that the applied fields were operating in a range that excited dynamic behaviours in the ring arrays, inputs were scaled simple linear transform using a pair of input scaling parameters, H_C and H_R, which defined the central field for zero input, and the total range the input spanned in Oe respectively.

In order to leverage the system as a reservoir computer, a standard approach introduced by Appeltant et al. in their 2011 paper ‘Information processing with a single dynamical node’ was employed. This approach combines the input data at each timestep with a fixed input mask to multiplex inputs over time, with output of each ‘virtual’ node taken measured after each masked input. Due to the intrinsic dynamics of the ring array, the response of each virtual node depends not only on the current input, but also the state of the system at the end of the last input. This emulates synaptic connections between each virtual node, resembling a network with fixed connectivity determined by the dynamic properties of the ring arrays.

To establish the computational properties of each transformation for each of the three state readout variables, task independent metrics of kernel rank, generalisation rank, and memory capacity were employed. These metrics provide a good evaluation of the two most important properties of a reservoir computer: nonlinearity and memory. Since they require the processing of significantly less data than most machine learning tasks, they are a useful tool for rapidly evaluating useful regimes of operation and can be used to expedite the parameter selection procedure.

To evidence that the rings are useful in task-based environments, and that the task-independent metrics were indeed good predictors of task performance, a standard benchmark task of spoken digit recognition was performed. When taking each readout variable alone, task performance was modest, demonstrating a useful but perhaps somewhat simplistic transformation was achieved via each of the state properties. When combining these outputs however, the reservoir performed significantly better, showcasing the increased computational capacity of the reservoir when a more sophisticated evaluation of magnetic state was made. Finally, it was shown that by taking heuristic measurements of ‘good’ reservoir performance from the measurements in metric space, the regions of best

performance could be isolated from the metrics, which is a useful tool for expediting the input scaling parameter search for tasks when data throughput is a limiting factor.

In this work, my personal contributions were the development of the model, design of the reservoir framework, the implementation and simulation of all metric and task-based analysis, and the drafting of the article. This paper was published in IOP Nanotechnology on 07/09/2022 and can be found online at doi.org/10.1088/1361-6528/ac87b5.

Quantifying the Computational Capability of a Nanomagnetic Reservoir Computing Platform with Emergent Magnetisation Dynamics

I T Vidamour[1], M O A Ellis[2], D Griffin[3], G Venkat[1], C Swindells[1], R W S Dawidek[1], T J Broomhall[1], NJ Steinke[4], J F K Cooper[4], F Maccherozzi[5], S S Dhesi[5], S Stepney[3], E Vasilaki[2,6], D A Allwood[1], T J Hayward[1]

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- ISIS Neutron and Muon Source, Rutherford Appleton Lab, Didcot, OX11 0QX, United Kingdom
5- Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxfordshire OX11 0DE, United Kingdom
6- Institute of Neuroinformatics, University of Zurich and ETH Zurich, 8057 Zürich, Switzerland

Abstract

Devices based on arrays of interconnected magnetic nano-rings with emergent magnetization dynamics have recently been proposed for use in reservoir computing applications, but for them to be computationally useful it must be possible to optimise their dynamical responses. Here, we use a phenomenological model to demonstrate that such reservoirs can be optimised for classification tasks by tuning hyperparameters that control the scaling and input-rate of data into the system using rotating magnetic fields. We use task-independent metrics to assess the rings' computational capabilities at each set of these hyperparameters and show how these metrics correlate directly to performance in spoken and written digit recognition tasks. We then show that these metrics, and performance in tasks, can be further improved by expanding the reservoir's output to include multiple, concurrent measures of the ring arrays' magnetic states.

Introduction

Neuromorphic devices use inherent material properties to perform brain-like computational operations *in materio*. This allows for improvements in efficiency over standard artificial neural networks as neural architectures are directly *emulated* in hardware, rather than *simulated* using conventional computers [1].

Reservoir computing (RC) is a machine learning paradigm that is well-suited to *in materio* implementations. In RC, a fixed dynamical system (the reservoir) transforms input signals into higher dimensional representations, facilitating classification in cases where input data is linearly inseparable. In the archetypal Echo State Networks (ESNs) [2], the reservoir takes the form of a recurrent neural network (RNN) initialised with a sparse, random connectivity matrix. A linear readout layer provides output from the weighted sum of activity across nodes within the reservoir [3]. ESNs address the well-known difficulties of training RNNs, and recent models have improved both their applicability to classification tasks and their robustness against catastrophic forgetting [4].

While ESNs are typically simulated on conventional computers, recent studies have shown that computational ability is preserved if the RNNs are replaced by physical systems [5]–[12] with the correct properties: nonlinearity between input and output, and 'fading' memory of past inputs. With typical hardware implementations of reservoir computing there is no separation between the components used for computation and those used for memory, mitigating the von-Neumann bottleneck associated with discrete memory and computation units. This offers potential benefits in

terms of reduced latency, increased computational power per unit area, and improved energy efficiency of the system.

Different classes of physical systems have been proposed for RC, each with their own advantages and technological challenges [13]. For example, optoelectronic systems [6], [14] utilise the nonlinear properties of off-the-shelf intensity modulation components coupled with fibre optic spools as delay lines to create reservoirs that can perform computational tasks with performance that rivals the state-of-the-art [14]. However, these suffer from substantial challenges to minimisation due to the size of optical delay lines. Molecular platforms use molecules such as proteins [15] and enzymes[16] as the computational building blocks for RC. These offer advantages in terms of the complexity of the reservoir and the feasibility of miniaturisation, though throughput is often slow and interfacing with standard electrical components is challenging. Memristors have also been proposed for RC [17], [18], as well as other forms of computation [19]. Memristors use short-term memory effects created from their variable resistance over time and are a particularly promising implementation of RC due to their suitability for interfacing with standard electronics as well as the ease of miniaturisation.

Nanomagnetic platforms are also well-suited to creating hardware-based reservoirs, offering many desirable properties including non-volatility, which provides a natural path to memory, and inherent non-linearity in their dynamics. Furthermore, methods for electrical reading [20], [21] and writing [22], [23] data are well-established from both the development of commercial magnetic random access memory (MRAM) [24], as well as research into more novel nanomagnetic logic [25]–[27] and memory [28], [29] devices. A wide range of nanomagnetic systems have been proposed for use as reservoirs including spin-torque oscillators (STOs) [12], super-paramagnetic arrays [10], skyrmion textures [30], single domain walls [11], artificial spin-ices [8], and garnet films [9]. While other nano-scale magnetic systems, such as spin-hall nano-oscillators (SHNOs) [31], and DW based spin-memristors [32] have been used for other neuromorphic computing implementations, they also show the nonlinear, time-dependent responses necessary for potential RC implementation.

Recently, we have proposed arrays of interconnected magnetic nano-rings as candidate platforms for nanomagnetic RC (Figure 1(a)) [33]. The arrays, which are lithographically patterned from thin films, consist of planar, ring-shaped nanowires of the soft magnetic material $Ni_{80}Fe_{20}$ with typical ring diameters <5 μm, linewidths <500 nm and film thicknesses <40 nm. For these dimensions the magnetic ground state of the rings in the array are “vortex” states (Figure 1(a)), where the local magnetization vector rotates in a closed loop around the rings’ circumferences. However, they can also support meta-stable, bi-domain “onion” states where magnetic domains with anti-parallel circulation are separated by a pairs of magnetic domain walls (DWs) (Figure 1(b)).

The soft magnetic properties of the nanowires means that the DWs are highly mobile and propagate through the nanowires like rigid quasi-particles when subjected to applied magnetic fields [17]. Specifically, in-plane rotating fields can drive the DWs pairs to coherently and continuously rotate around the ring circumferences [34] (Figure 1(c)). While in isolated rings DW motion is relatively unimpeded, in interconnected ring arrays junctions between the rings act as pinning sites that present localised energy barriers against DW propagation. The interaction of the DWs with such pinning sites are highly stochastic [28], [29] such that when DWs encounter junctions during their rotation around the ring they have a finite probability of becoming pinned temporarily in place, with pinning becoming less likely as the rotating field amplitude is increased. These pinning events lead to field-dependent stochastic interactions between pinned and propagating DWs at the rings’ junctions which can cause both loss of DW pairs from the array (i.e., increasing in the number of vortex states) or gain of DW pairs (i.e., decreasing in the number of vortex states).

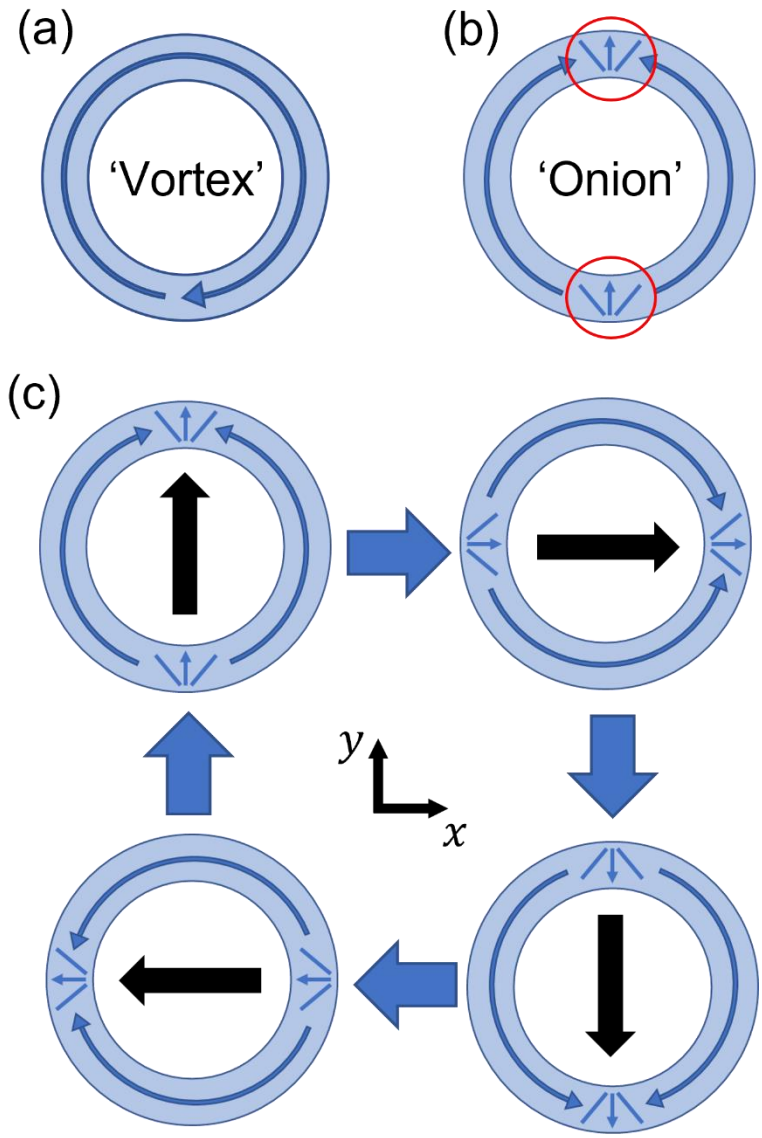


Figure 1- (a) Schematic diagram of a 'vortex' state. The blue arrows represent the local magnetisation direction (b) An 'onion' state, featuring two domains (separated by a head-to-head domain wall (upper red circle), and a tail-to-tail domain wall (lower red circle). (c) Coherent rotation of domain walls as the applied magnetic field direction (black arrows) rotates through 360°.

Collectively, these effects result in the ring arrays' magnetisation states exhibiting emergent responses to rotating magnetic fields, where the simple interactions of DWs at the junctions between rings results in complex collective behaviour of the arrays as a whole [33]. In our previous work we have shown that these emergent dynamics result in both a highly non-linear response of an arrays' magnetisation states to the rotating field amplitude, and fading memory of previous magnetisation states, thus meeting the two primary criteria for a dynamical system to be used for RC [33]. The ring arrays have a proven manufacturing route using electron-beam lithography and the lift-off method, as demonstrated in the previous experimental studies on the system[33]. The transformations provided by the rings' response can be varied by controlling how input data scales the applied field and its input rate. This offers the possibility of tuning their responses for different computational tasks. Furthermore, it is well established that the magnetisation states of magnetic nanorings can be characterised electrically using either anisotropic magnetoresistance measurements [21], [37], or by giant magnetoresistance if the rings are patterned from multilayer films with spin valve properties [38], [39], providing a potential readout mechanism for the ring arrays and making them highly

suitable for device integration. Together, these properties suggest interconnected rings have great potential for in-materio reservoir computing.

In a previous work we used a phenomenological model of a ring array's dynamics to demonstrate that these systems could be used as reservoirs [33]. This was achieved by treating an array as a single dynamical node into which time-multiplexed data [5] was input via the amplitude of a rotating applied magnetic field. Our simulations demonstrated that the ring arrays could successfully perform benchmark classification tasks such as spoken digit recognition. However, quantification of performance in any given task does not represent a comprehensive evaluation of the computational capabilities of a reservoir.

Assessing the computational capabilities is challenging. In general, different devices will provide different reservoir transformations, with different dynamical regimes of a given device offering further flexibility. To overcome this, task-independent metrics of Kernel Rank (KR), Generalisation Rank (GR) [40], [41], and Linear Memory Capacity (MC) [42] can be employed. These allow empirical measurement of the reservoir's ability to separate, generalise, and remember input respectively. These metrics provide an insight to the properties of a given reservoir configuration along three different computational axes, calculated directly from the transformations the dynamical system provides. The findings of these metrics can be utilised to provide a more informed starting point when optimising these systems to perform machine learning tasks, based on the assumed demands of a given task (e.g., high KR where data is linearly inseparable, high MC for regression tasks with long-term temporal dependencies).

In this paper, we use task-independent metrics to assess the computational capabilities of a modelled interconnected magnetic nanoring array. We show how controlling the scaling and input-rate of data allows these metrics to be tuned, and how their variation correlates to performance in a pair of benchmark classification tasks (spoken and handwritten digit recognition). We then demonstrate how expanding the reservoir's output to include multiple, concurrent measures of the array's magnetic state further improves upon these reservoir metrics and performance in classification tasks.

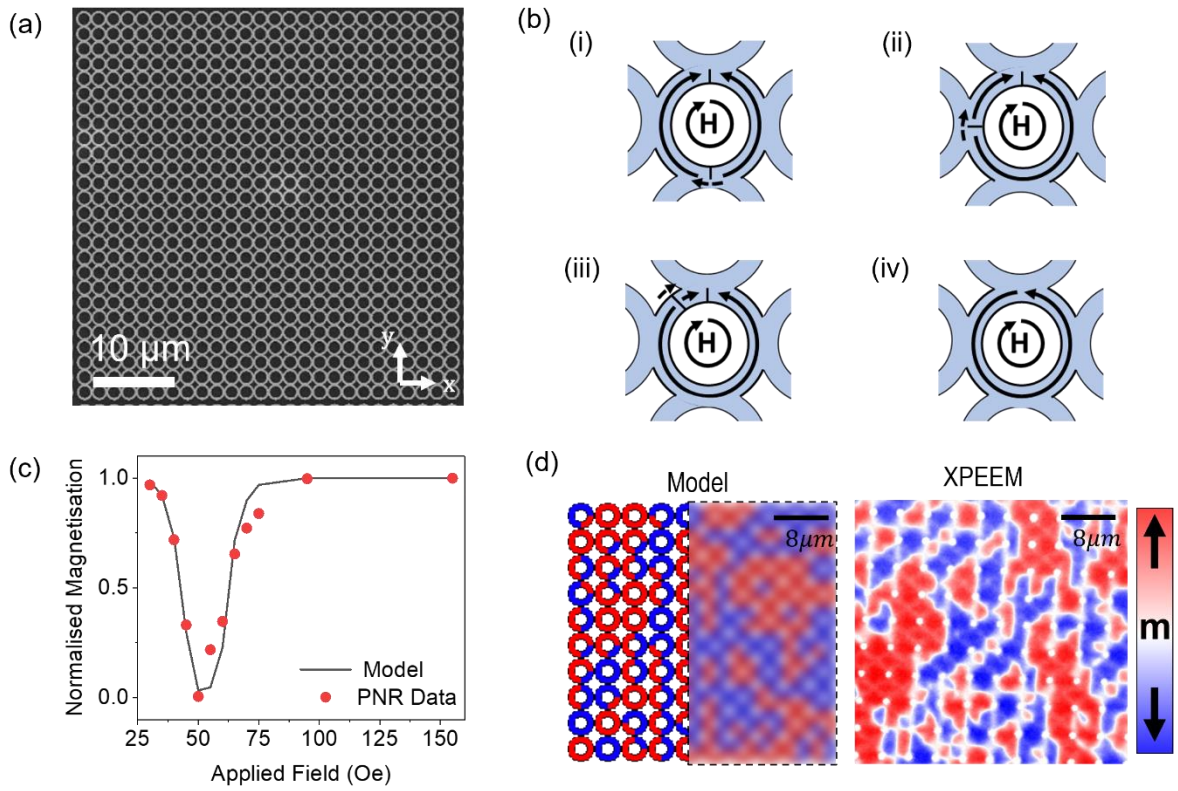


Figure 2: (a) Scanning Electron Micrograph of a typical nanoring array which has been modelled. (b) Schematics illustrating the DW annihilation process. The upper domain wall stays pinned whilst the lower domain wall moves, leaving the ring in the following states: (i) 'onion' state, (ii) '3/4'state (iii) DW collision, and (iv) formation of a 'vortex' state. (c) Equilibrium magnetisation of the array as a function of applied field as measured by polarised neutron reflectometry (PNR, red symbols) [33] and simulated equilibrium magnetisation given by the fitted model (black line). *(d) Magnetisation images generated by the model and by X-ray photoelectron emission microscopy (X-PEEM) after 10 rotations of 50 Oe applied field. Red/Blue colour denotes magnetisation along the axes given by the colour bar. The modelled image includes a blurred region to aid visual comparisons.*

Methodology

Simulating Magnetic Nanoring Arrays

The system modelled consisted of a 25 x 25 square array (Figure 2(a)) of 4 μm diameter, 400 nm line width $Ni_{80}Fe_{20}$ rings, with thickness t = 20 nm. This system was experimentally characterised in [33], where we also created and validated a phenomenological model of its behaviour. Here, we used this model, RingSim, to simulate the response of a ring system to streams of data encoded using the amplitude of a rotating magnetic field, with the simulated magnetic states of the array acting as output.

In RingSim, rings existed as either 'onion' states, containing two DWs or 'vortex' states containing no DWs. DWs were instanced into RingSim as agents which attempted to follow the rotating field to minimise Zeeman energy [43]. DWs within RingSim existed as pairs, with one DW instanced as a 'Head-to-Head' DW (H2H, converging magnetisation), and the other as a 'Tail-to-Tail' DW (T2T, diverging magnetisation). When a uniform magnetic field was applied to a ring, the minimum Zeeman energy positions for H2H DWs and T2T DWs were located at opposite sides of the ring, and rotated with the direction of the applied field. For example, for the coordinate system shown in Figure 1(c), magnetic field applied (black arrows) in the positive y direction would have energy minima at the top of each ring for H2H DWs, and the bottom for T2T DWs. Differential movement of DWs as a result of stochastic

pinning events (Figure 2(b.i)) led to the formation of '3/4' states (i.e. onion states with one DW displaced by 90°, Figure 2(b.ii)), or collapse of the ring into a vortex state upon DW collision and annihilation (Figure 2(b.iii/iv)).

Junctions between rings created anti-notch-like energy barriers against DW propagation [44], [45]. The size of these energy barriers was modulated by the tangential component of applied field in accordance with Sharrock's equation [35], [46]:

(6.1) $$\Delta E = E_0 \left(1 - \frac{H_{drive}}{H_0}\right)^{\alpha}$$

where ΔE represented the field-modulated energy barrier, E_0 is the magnitude of the unmodulated energy barrier, H_{drive} is the component of applied field acting tangentially to the ring at the DW's position, H_0 is the zero-temperature depinning field, and α is a geometrical exponent that controls the variation of the energy barrier with applied field. E_0 also depended on whether a DW was present in the neighbouring ring on the other side of the junction; where this was the case reducing the energy barrier was reduced by a factor 0.75.

The expected timescale of thermally activated reversal, t_r, was calculated from ΔE barrier using the Arrhenius-Néel law, [47]:

(6.2) $$t_r = t_0 e^{\frac{\Delta E}{k_B T}}$$

where t_0 represented the inverse of the attempt frequency (~ 1 GHz for $Ni_{80}Fe_{20}$ [35]), k_B represented the Boltzmann constant, and T was the temperature (taken to be fixed at 293 K).

The rotating magnetic field was modelled as series of discrete steps of π/8 radians. At each step, the field was held for a duration of $t_H = 1/(16 * f)$ seconds, where f was the frequency of rotation, taken here to be 5 Hz, resembling the order of the rotational frequencies used in our characterisation experiments. For each of these field steps, the probability, P of a DW depinning from by the energy barrier was calculated using:

(6.3) $$P = 1 - e^{-t_H/t_r}$$

The stochastic nature of DW pinning was modelled by comparing random floating points to pinning probability P. If the generated random floating point exceeded P, then the DW was considered free to propagate via the shortest path to either the appropriate Zeeman energy minima, or an intermediate junction. Interactions between DWs were introduced phenomenologically: If a H2H/T2T DW collided with T2T/H2H DW in the same ring, both DWs were annihilated, leaving the ring in an onion state; If a DW passed a junction where a no DW was present in the neighbouring ring it nucleates a H2H and T2T DW pair into that ring.

RingSim was defined by 4 free parameters: $H_0, E_0, \alpha,$ and ΔH_0 . The first three parameters were defined in the previous equations, and the final parameter, ΔH_0 represented the standard deviation of a gaussian distribution of H_0 across the junctions of the array and was included to approximate the variation in junction properties arising from material or lithographic defects. In this study, these parameters were fitted to the results of polarised neutron reflectivity (PNR) measurements of the array shown in Figure 2(a), where the array's net magnetisation along the y-axis was measured as a function of rotating field amplitude. The PNR measurements were taken following 50 rotations at each applied field amplitude, with saturation and relaxation of the array occurring before each measurement [33]. The data thus represented a dynamic equilibrium magnetisation state of the array at each applied field amplitude.

Figure 2(c) presents PNR data, along with the fit produced by RingSim ($H_0 = 14.25\ mT, E_0 = 1.05eV, \alpha = 1.1,\ \Delta H_0 = 1.25mT$). The model fitted the data well, with agreement being particularly strong in the region H_{drive} = 35 - 70 Oe, where the system exhibited an emergent response. Furthermore,

images generated from RingSim in this regime of behaviour showed good qualitative agreement with X-ray photo-electron emission microscopy images (X-PEEM), (Figure 2(d)), with both showing extended magnetic domains forming over similar length scales in the array. Further details of the validation of RingSim against experimental results can be found in [33].

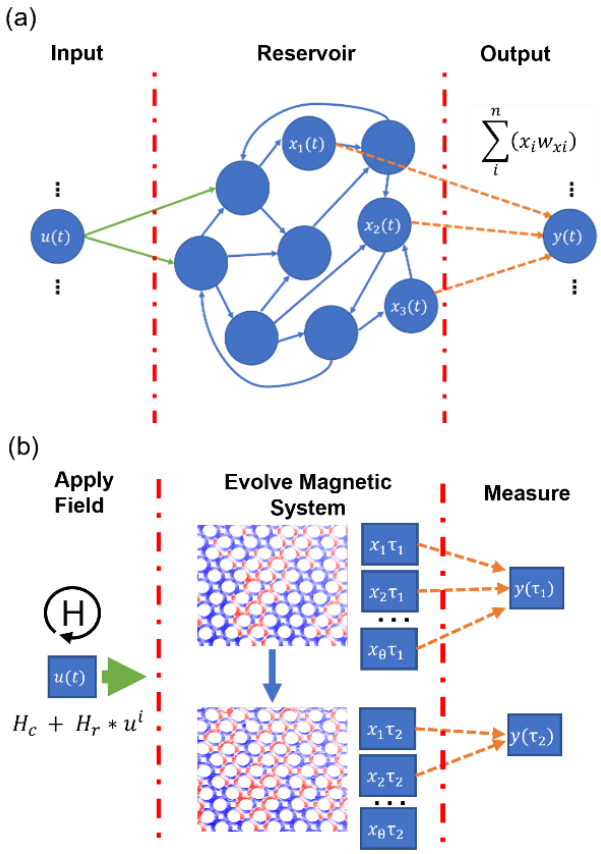


Figure 3: (a) The ESN approach to RC, showing the layered structure of the model: a fixed reservoir layer is provided with time-varying input via weighted connections. A linear output layer then provides a weighted sum of activities from nodes within the reservoir layer. (b) Our approach, where the ring array acts as a single dynamical node, into which time-multiplexed data is input using a rotating magnetic field, and output is extracted by measuring magnetic properties of the array at the end of each time-multiplexed input.

Simulating Reservoir Computing with Ring Arrays

RC involves the transformation of discrete-time input signals, $u(t)$, to reservoir states, $x(t)$ (Figure 3(a)). The reservoir configuration employed here follows the paradigm of a single dynamical node, as introduced by Appeltant et al. [5] (Figure 3(b)). Here, the network was constructed of 'virtual' nodes, created by observing a physical property of a dynamic system as it responds to time-multiplexed input. This approach has been used in a wide range of physical reservoirs due to its ease of implementation [5], [6], [10]–[12].

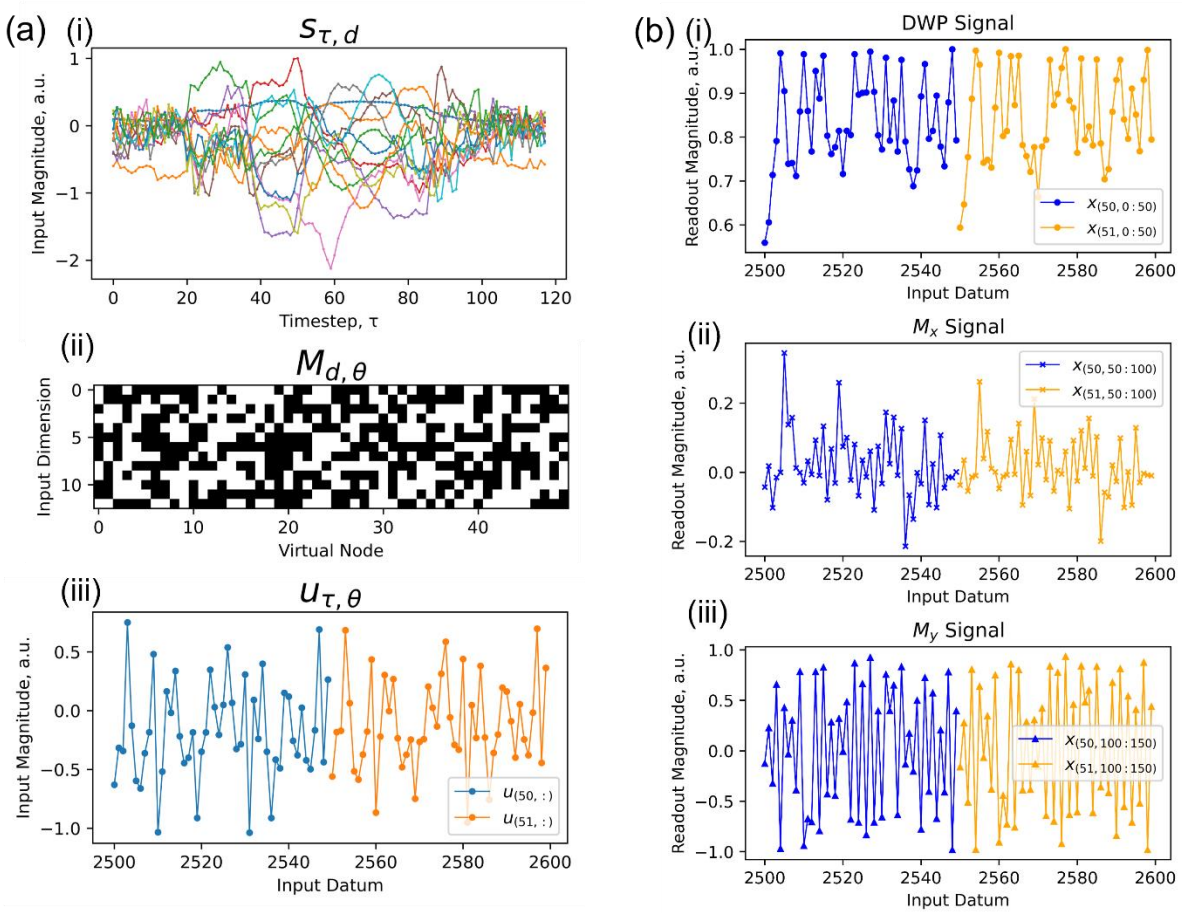


*Figure 4- (a) Outline of the masking procedure for data representing an utterance of the digit 'zero', showing filtered input signal **s** (top, different colours represent different input dimensions d), which is combined with binary input mask **M** (middle), to form masked input signal **u** (bottom, masked signals shown for τ=50/51). (b) Example signals for the three state readout variables, DWP (top), M_x (middle), and M_y (bottom) when driven with $\boldsymbol{u}_{50}$ (blue) and $\boldsymbol{u}_{51}$ (orange) of a spoken digit 'zero'.*

Figure 4 outlines the reservoir computing procedure. The raw input signal $\boldsymbol{s}_{\tau,d}$ (Figure 4(a.i)) is first combined with a randomly generated mask matrix $\boldsymbol{M}_{d,\theta}$ (Figure 4(a.ii) to create a time-multiplexed input matrix, $\boldsymbol{u}_{\tau,\theta}$, (Figure 4(a.iii)) via equation (6.4),

(6.4) $$\boldsymbol{u}_{\tau,\theta} = \boldsymbol{s}_{\tau,d} * \boldsymbol{M}_{d,\theta}$$

where d = number of dimensions of the input signal, τ = the number of time-steps of the input signal, and θ = the number of virtual nodes. Fifty virtual nodes were chosen here to provide adequate dimensionality expansion while keeping simulation durations short. Matrix $\boldsymbol{u}_{\tau,\theta}$ was then flattened by concatenating row by row, producing a 1D input signal u_n of length $\tau * \theta$. For $d > 1$, $\boldsymbol{M}_{d,\theta}$ was filled with random binary digits (figure 4(a.ii)). Here, the virtual nodes provide different linear combinations of input dimensions over time, allowing the reservoir to provide a non-linear representation of these inputs via the system's intrinsic nonlinearity. For single-dimensional inputs, a mask of random floating points was used instead to excite different responses in the reservoir over time.

For an input datum i from u_n, the applied field amplitude $H^i_{applied}$ was given by:

(6.5) $$H^i_{applied} = H_{centre} + H_{range} * u_i$$

where H_{centre} and H_{range} represented the offset and scaling of the rotating field sequence. Each input was applied for a given number of quarter-rotations of field, where N_q denoted the number

of quarter-rotations (with quarter-rotations chosen to reflect the fourfold rotational symmetry of the array). Here, values of N_q were chosen such that N_q was smaller than the number of quarter rotations required to reach an equilibrium state. This connected the states of the virtual nodes to one another by maintaining the reservoir in transience [5], [12].

Three variables correlated to the magnetic state of the array were logged at the end of each input, the number of DWs currently in the system, normalised to the number found at saturation (Domain Wall Population, DWP, Figure 4(b.i)); and the array's net magnetisation components in the x and y directions, M_x and M_y, Figure 4(b.ii/b.iii)). These were concatenated, producing a reservoir state vector x_{3n} which was three times the length of u_n. We note that while these output variables were chosen as they were easily available in RingSim, they are all potentially accessible in electronic measurements of real devices. For example, values representative of M_x and M_y could be measured in spin valve stacks exhibiting GMR multilayers with appropriately aligned pinned layers [20], [38], [39], while a proxy for DWP could be obtained through anisotropic magnetoresistance (AMR) measurements [21], [48], [49].

The output from the reservoir, $\boldsymbol{Y}_{\omega,n}$, was constructed by combining the reservoir state matrix, $\boldsymbol{X}_{\theta,n}$, with output weights, $\boldsymbol{W^{out}}_{\omega,\theta}$, where ω reflects the number of output nodes, and n reflects the number of input patterns used to construct reservoir state matrix ***X.*** Output weights were calculated using an ordinary-least-squares method with Tikhonov regularisation, commonly referred to as 'Ridge Regression' and described by the following equation:

$$\boldsymbol{W}^{out} = \boldsymbol{Y}\boldsymbol{X}^T * (\boldsymbol{X}\boldsymbol{X}^T + \gamma^2\boldsymbol{I})^{\dagger} \tag{6.6}$$

where γ^2 represented the regularisation parameter, $\boldsymbol{I}$ the identity matrix, and † the Moore-Penrose pseudo-inverse operation. Regularisation was performed by selecting the γ^2 with highest average classification accuracy on the training set, evaluated across multiple shuffles of the training data.

Task-Independent Metrics

We estimated the computational properties of the reservoirs using task independent metrics: KR, GR [40], [41], and MC [42]. Both KR and GR rely on estimations of the number of non-linearly related output patterns are that generated when a system is provided with input sequences with distinct characteristics. The differences between input characteristics were designed to assess how the reservoir responds to distinct input patterns compared to nominally similar patterns. Hence, KR estimated a reservoir's ability to map distinct inputs to different reservoir states, while GR estimated the ability to generalise noisy versions of the same input to similar reservoir states. Generally, higher KR scores mean a better ability to separate data, while lower GR scores reflect a better ability to generalise data. The ability of the reservoir to separate input data effectively (High KR) whilst not being overly sensitive to small changes in initial conditions (Low GR) is what makes them effective as classifiers, though the exact ratio of KR to GR for optimal performance is dependent upon the given task [41].

To measure both KR and GR, N x 1-dimensional input signals of length M were generated from independent and identically distributed (i.i.d.) floating points uniformly distributed between ±1 and applied to the reservoirs. Here, N = 200 and M = 10. For KR, the sequences were unique and uncorrelated from one another; in GR, the sequences were uncorrelated except for the final three inputs, which were identical for each sequence. Hence, the output states from KR show the reservoir's ability to map each unique input pattern to non-linearly related output states, and the output states from GR demonstrate both the system's ability to converge to similar output states from random starting points, as well as the system's overall robustness to noise.

Before the sequences were applied to the reservoir, they were combined with a fixed mask of random floating points as described earlier for the case where input dimensionality, d = 1, thus producing N input signals of length M x θ. These were then scaled into rotating field sequences using Equation 5. The reservoir was reinitialised with a strong pulse of magnetic field, saturating the system, and forming onion states uniformly aligned along +y prior each sequence being inputted. The final reservoir states following each sequence were generated by taking measurements of DWP, M_x and M_y across all θ virtual nodes, thus generating an output matrix $\boldsymbol{O}_{N,3\theta}$. KR and GR were defined as the ranks of these matrices, and thus represented the number of linearly independent responses produced at the reservoir's output when driven with input signals with the characteristics described above. In addition to generating KR and GR for matrices $\boldsymbol{O}_{N,3\theta}$ containing all three state variables, ranks were also generated for $\boldsymbol{O}_{N,\theta}$ matrices for each of DWP, M_x and M_y alone, in order to highlight the effect on computation of evaluating additional reservoir properties for each input datum.

The ranks of $\boldsymbol{O}$ were estimated using singular value decomposition, calculated as the number of singular values above an arbitrary small noise threshold, here 0.1. To alleviate biasing higher metric scores to regions of operation where the readout state variables are numerically higher, all output matrices $\boldsymbol{O}$ were normalised against the maximum value in $\boldsymbol{O}$ prior to singular value decomposition to provide fair rank estimation between field profiles.

To evaluate the ring array's memory capacity, it was driven with an i.i.d. input u_i, where i denoted each input datum before masking. Here, i=550, but the outputs from the first 50 inputs were discarded to wash out any initial conditions from the reservoir's response. We then trained ${\boldsymbol{W}^{out}}_{\omega,\theta}$ to recover past inputs u_{i-k} for each delay k as output y_k, with a 250:250 train/ test split. MC was evaluated from the covariance between the delayed inputs u_{i-k} and the trained reconstruction of input y_k for summed across all nodes θ for each delay via the following formula:

$$MC = \sum_{k=1}^{\theta} \frac{cov^2(u_{i-k},y_k)}{\sigma^2(u_i)\sigma^2(y_k)} \tag{6.7}$$

MC provided a basic insight into the memory properties of the reservoir, with MC approximating the number of time-steps in the past over which the network could reliably recall previous inputs.

Heatmaps were generated showing metric values (KR, GR, MC) for a range of driving field parameters H_{centre}, H_{range}, and N_q. In these heatmaps, and the subsequent performance heatmaps for the digit recognition tasks, the scaling parameters were instead expressed as the mean field applied, as well as the standard deviation of the field sequence for a given set of input parameters. This is done to account for the reduction in effective range caused by the masking process, as well as the non-uniform distribution of the task data, allowing more effective correlation between the metrics and the task performance.

Digit Recognition Tasks

A pair of benchmark classification tasks were chosen to assess reservoir performance: spoken digit recognition (NIST TI-46 database, [50]) and handwritten digit recognition (MNIST, [51]). Both consisted of a total of 500 utterances/images of the digits 0 – 9. For TI-46, each of five female speakers provided ten utterances of each digit. Inputs were created using a Mel-Frequency Cepstral filter [52] to produce responses in 13 frequency bands across 50 ms windows, generating a raw input signal s with d=13 and τ equal to the number of windows generated by a given utterance. For MNIST, 50 images of each digit were taken randomly from the 'training' set of the database. The [28x28] pixel images were considered as signals where d=τ=28, and multiplexing was performed column-by-column. For both tasks, the array was initialised to contain uniform 'onion' states aligned along +y prior to each input sequence.

Performance was assessed across multiple 80:20 splits of training and testing data for each digit. The 500 datapoints were split into train/test groups randomly, and performance was averaged over 100 different shuffles of the data. For training/testing, the output for every τ in each signal was labelled according to the digit they represent using one-hot encoding, generating state matrix $\boldsymbol{X}_{\theta,m}$ and target matrix $\boldsymbol{Y}_{\omega,m}$ where m was the total number of time-steps across all utterances.

To classify unseen data, the activation for each of the ω=10 outputs was calculated as the cumulative sum across all time-steps τ, meaning that the classification algorithm could handle inputs of different number of timesteps. Classification was performed using a winner-takes-all approach based on the output with the highest activation. We assessed performance for output vectors consisting of each of the three state variables independently, as well as for all three simultaneously.

Results and Discussion

Task Independent Metrics

Figure 5 presents heatmaps for KR, GR and MC created by of varying H_{centre} and H_{range} values for each of the three state variable measurements. Data are presented for a fixed input rate of N_q = 2, which exhibited best overall performance in the tasks, while heatmaps for input rates N_q = 1 and N_q = 4 can be found in the supplementary material (Figures S1 and S2). Each metric was bounded by number of nodes, here 50, as the maximum rank of output matrices is limited to the smallest dimension of $\boldsymbol{O}_{N,\theta}$, and MC is similarly bounded as the maximum covariance for a given delay is 1, summed across all nodes θ. The three input rates show broadly the same behaviour, with some shifting peak values towards higher mean fields and ranges for the fastest input rate of N_q = 1 since higher fields perturbed the system more significantly when each input was applied for more rotations.

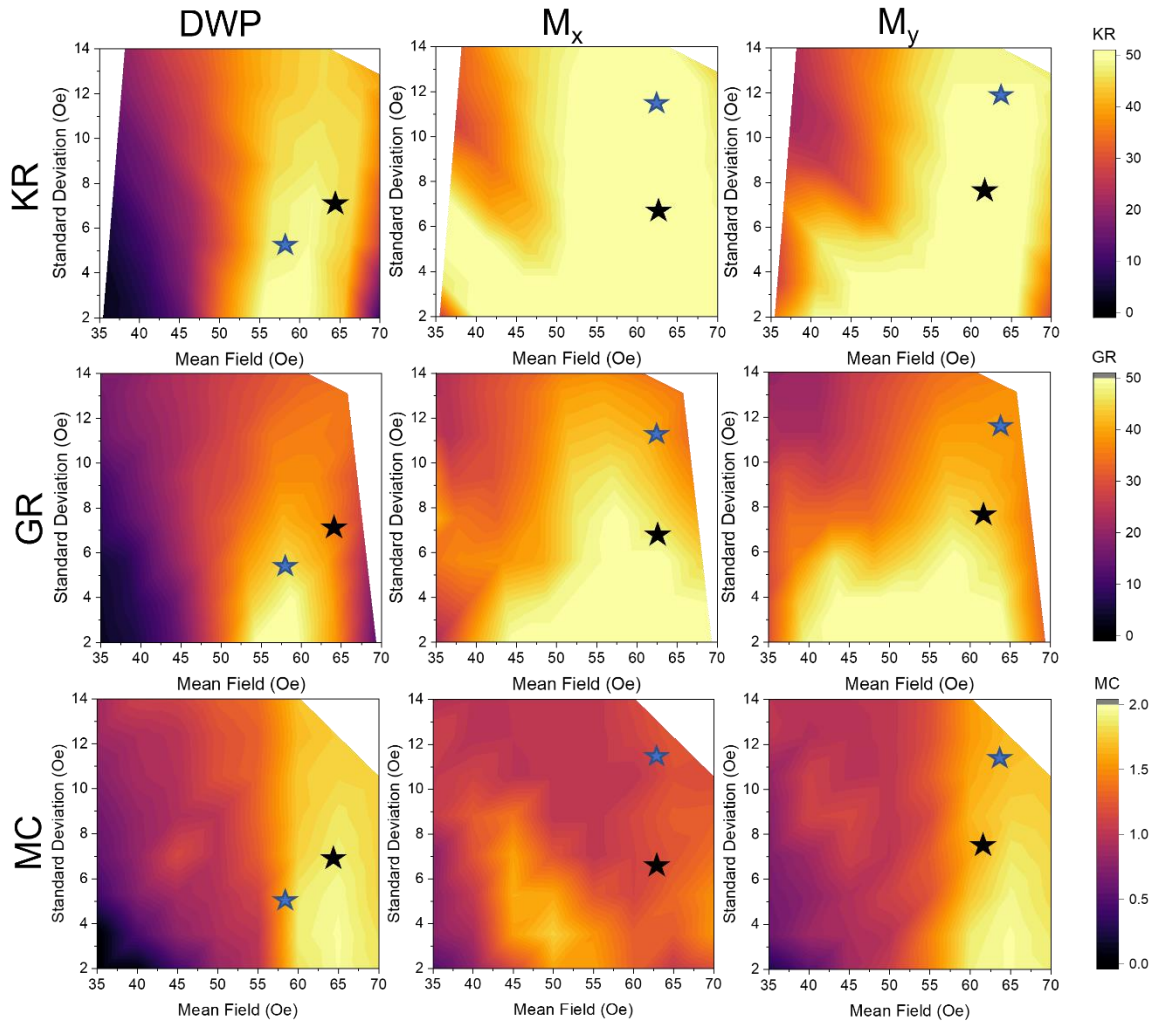


Figure 5- Empirically-measured reservoir metrics Kernel Rank (KR), Generalisation Rank (GR), and Memory Capacity (MC) for an input rate of $N_q = 2$, or a half rotation of field per input, for each of the output parameters of Domain Wall Population, X Magnetisation, and Y Magnetisation taken independently. Black/blue stars represent peak performance in spoken/handwritten digit recognition for each reservoir state variable, while metric scores are reflected according to the colour bars on the right.

A wide distribution of different reservoir properties were observed in the heatmaps. The configurations most suitable for reservoir computing were clustered around mean field = 50-70 Oe, with KR reaching its maximum value and MC > 1, indicating the presence of both nonlinearity and memory in the reservoirs' responses.

Useful reservoirs were not expected in the 'hotspots' of relatively high GR, as in those regions the reservoirs could not map similar input sequences to similar output states. These regions likely corresponded to dynamical regimes where the inherent stochastic noise of the system was large and obscured the underlying signal, preventing effective generalisation. For all three state variables, there was a wide region of high KR. This represented regimes where the reservoirs' dynamics were suitably non-linear due to a combination of the ring array responding non-linearly to rotating field amplitude (Figure 2(c)) and the connections between virtual nodes produced by the system being maintained transience. Effective reservoir computers were likely to be found in regions where KR > GR as these reflected the system operating in a complex yet ordered state. When optimising task performance, this criterion allows a reduction of the search space without discarding useful reservoir configurations.

All reservoir configurations exhibited low MC, with a maximum MC = 2. This feature can be explained by the time-multiplexed approach to RC used: consecutive timesteps of the input data were separated by the entire sequence of virtual nodes and so the memory of the system was able to connect virtual nodes together, though struggled to connect inputs from more than one timestep in the past. In other time-multiplexed networks with a single dynamical node, a delay line is often included to provide feedback of output from previous timesteps [5], [10], [53], and would likely augment the memory characteristics of the ring array reservoirs similarly. However, this is beyond the scope of this paper, as the goal was to characterise the computational properties of the dynamical system itself, without aid of peripheral feedback methods. As memory is critical to reservoir computation, a decision criterion for restricting the hyperparameter space search as part of task optimisation can be drawn where MC > 1.25 in order to eliminate reservoirs without effective memory.

There were noticeable differences between the metric maps for DWP and M_x /M_y, with the directional magnetisation components having higher KR/GR values in general. One possible reason for this discrepancy related to the relative complexity of each state variable, since the DWP measure was indifferent to the direction and size of domains in rings, while the magnetisation components were sensitive to these additional factors, and hence exhibited a richer dependence on the system's state. Additionally peak values of MC and KR were maintained at higher input standard deviations for M_x /M_y. This arose from differences between the expected equilibrium values of DWP and magnetisation for a given field; DWP saturated to maximum values at lower applied fields than magnetisation as the rings form '3/4' configurations (thus maximising DWP) at lower fields than they formed uniformly aligned onion states (thus maximising magnetisation).

There were also differences between the metric maps for M_x and M_y which can be seen most clearly for N_q = 2 (Figure 5) and N_q = 4 (Figure S2) where the regions with high KR and GR extended over larger proportions of the maps for M_x than M_y. We suspect these differences occurred because readouts were always performed when the field was aligned along the y-axis for even values of N_q, meaning the system was less significantly perturbed in x. Smaller perturbations meant a smaller signal-to-noise ratio for the state variable compared to the inherent noise of the system, and hence higher values of KR and GR, especially when the mean field and standard deviation were lower. M_x also produced a lower peak value in MC than M_y. Similarly, this is likely due to M_x having small magnitude, and hence hindering reconstruction with worse noise properties, when operating in the regime where the system seems to show the biggest dependence on past input (Mean Field > 60 Oe, evidenced by peak MC scores for other two variables).

The differing metric distributions for each of the state variables suggested that they provided different data transformations, and thus could be combined to create an enhanced reservoir. This was demonstrated through production of reservoir metric heatmaps for output vectors containing all three state variables together, as shown in Figure 6.

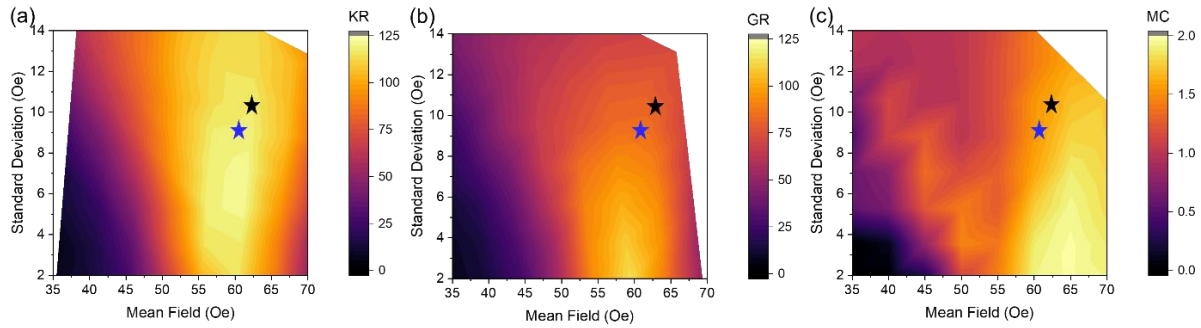


Figure 6- Three-output-variable metric maps for (a) KR, (b) GR and (c) MC for an input rate of N_q = 2. Black/Blue stars represent the highest performing configurations in the spoken/handwritten digit recognition tasks respectively.

The three-output-variable metric maps showed substantial improvements over those for the individual variables. Both KR and GR increased, highlighting the additional nonlinear mappings provided by using the three state measurements concurrently. However, not all the additional nodes contributed additional nonlinearity. As noted previously, the upper bound for KR and GR was equal to the number of nodes/weights in the virtual network, i.e. 150 when using all three output variables. Both KR and GR peaked below this maximum value, with maximum ranks of 121 and 117 respectively, illustrating the diminishing returns of adding additional virtual nodes. The maximum MC of the system did not increase. This was likely due to the three state variables having similar rates of change with respect to changing inputs.

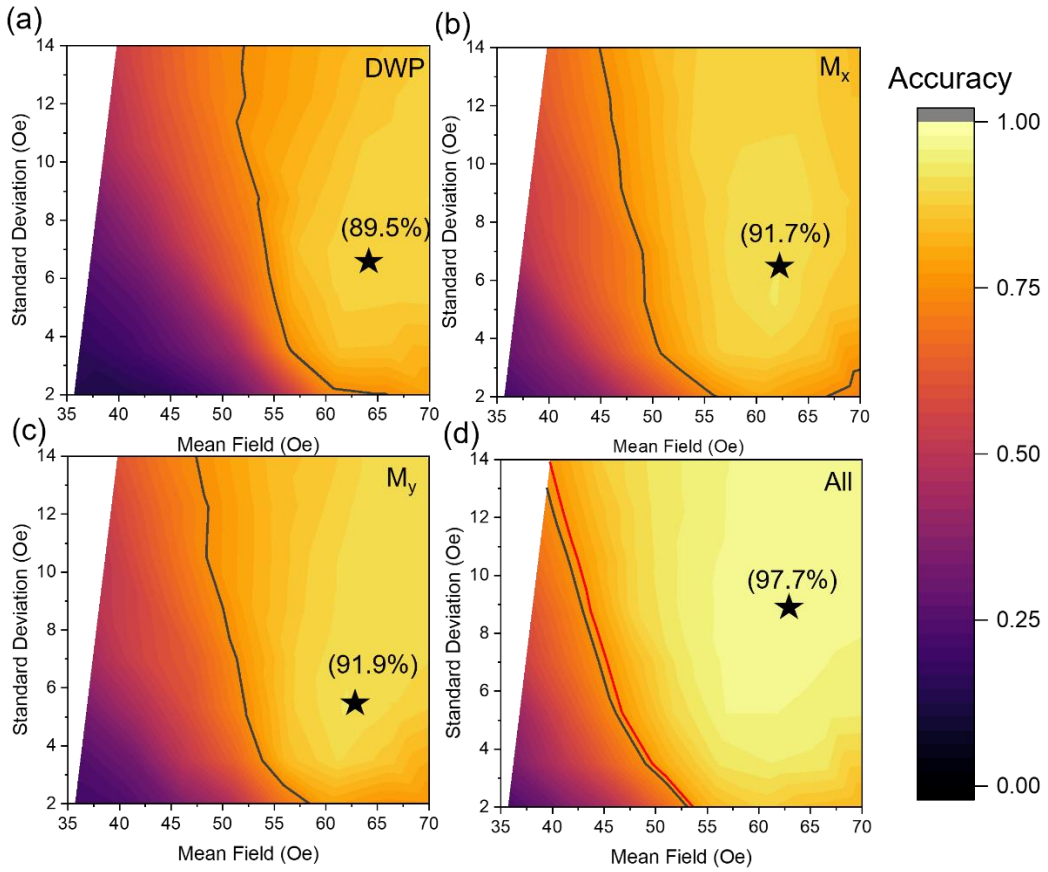


Figure 7- Test accuracy for TI-46, 100 utterances by five different speakers, 80:20 training: testing split. Quoted accuracies are for 100 different shuffles of training/test data. Four maps represent which outputs constituted the features used for classification: (a) DWP output (b) M_x (c) M_y (d) combined outputs. Blacks line = performance without reservoir transformation, where θ = 50. Red line = baseline performance for input data generated from a mask with θ = 150.

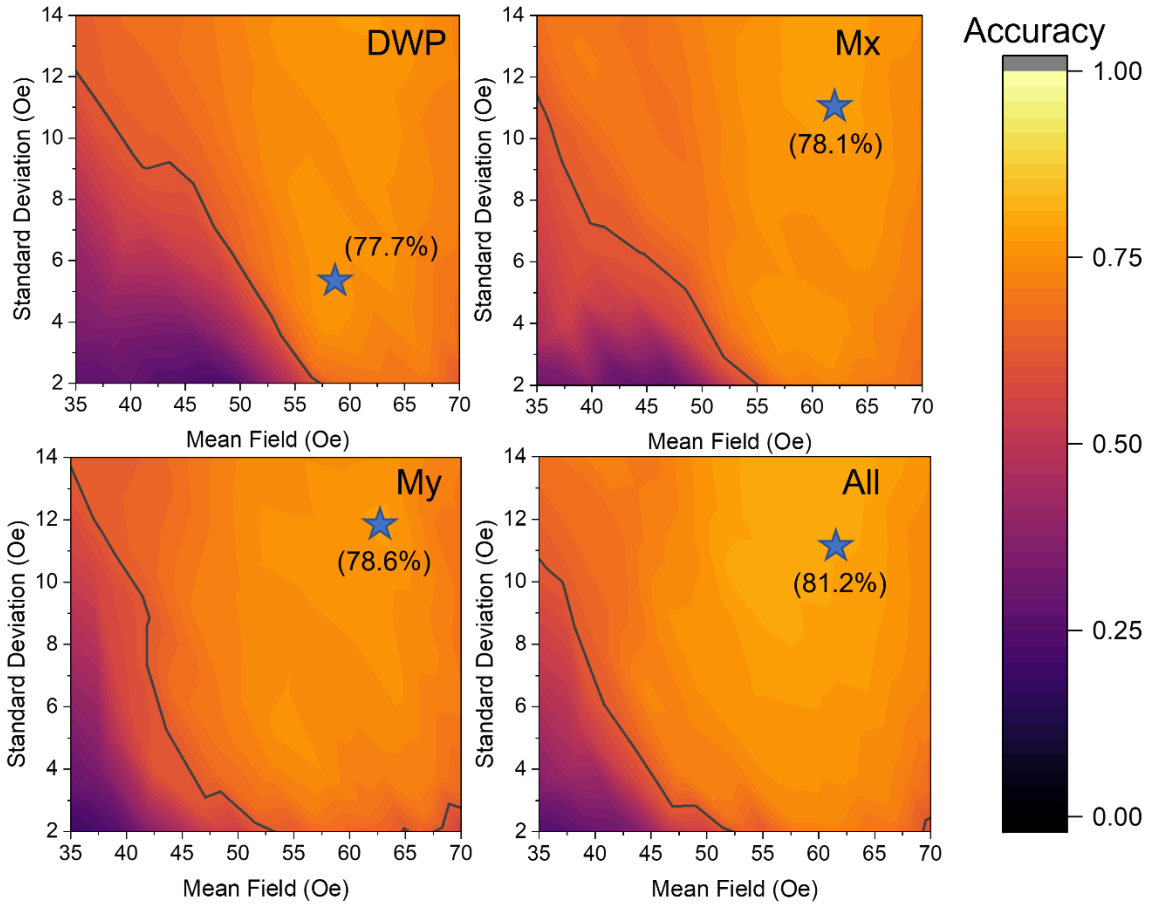


Figure 8- Task performance for 100 different shuffles of training/testing data of the MNIST task. Blue stars and percentages represent peak classification accuracies for each configuration. Black line represents control performance generated by skipping reservoir transformation, where θ = 50.

Digit Recognition Tasks

Figure 7 presents maps of TI-46 performance for each of the output variables, as well for all three combined. Peak accuracies were 89.5%, 91.7%, and 91.9% for the DWP, M_x and M_y outputs respectively and increased to 97.7% for the three properties combined. As a control, performance tests were performed where the output weights were trained directly on the masked input data, meaning all the pre-processing steps were included, but the reservoir's transformation was skipped. Control measurements were created for the three-output-variable case by generating masked inputs with θ = 150 to ensure parity in trainable parameters. The 50 and 150 virtual node control setups achieved average accuracies of 75.8% and 77.2% respectively. Thus, all the output configurations substantially outperformed the control. Reservoir configurations which outperformed the control data are bounded on the heatmaps by black and red lines, showing 50 and 150 node controls respectively.

Peak accuracies were competitive with proposed architectures with similar pre-processing (94.8% with 500 training samples and 50 virtual nodes on a superparamagnetic array [10], 99.8% with 900 training samples and 400 virtual nodes on a spin-torque nano-oscillator [54]). While the STNOs and superparamagnetic systems can respond to stimuli on nano-second timescales, these speeds are a requirement for these systems. The response of the ring system is more closely tied to the number of rotations rather than the frequency of input, with faster frequencies increasing the required field for a given amplitude of oscillation increasing with input rate. This is shown in figure S5 where the field

requirement increases for similar perturbations in the system's net magnetisation for 5 compared to 500 Hz. The tuneability of driving frequency of the ring system, coupled with the non-volatility of the system's response, means the ring system is suitable for a wide range of input, broadening its range of potential applications where real-world sensory data is involved.

The sequential MNIST handwritten digit recognition task was performed similarly, with both individual state variables, as well as the three outputs combined (Figure 8). Again, each state variable provided considerable improvement over the control configuration, with scores of 77.7%, 78.1%, and 78.6% for the DWP, M_x and M_y outputs respectively, compared to a control accuracy of 61.2%. Performance in this task rivalled an ESN with slightly fewer nodes (100 nodes, 79.43% accuracy [55]). In both tasks, M_y slightly outperformed the other two output measures, owing to its greater expressivity than DWP, and improved signal-to-noise properties compared to M_x, properties which are reflected in the metric maps where M_y exhibited a higher KR than DWP, and a lower GR than M_x.

While both tasks showed an improvement in performance when the concurrent state variables were combined, the MNIST task showed a smaller increase in performance compared to the TI-46 task, rising to a peak accuracy of 81.2%. The task independent metrics can explain the discrepancy between the improvement provided to the two tasks; metric evaluation showed a gain in the KR and the GR of the system with combined outputs, but no change to the system's MC. The sequential MNIST task requires effective correlations to be drawn over long separations for successful classification (e.g., the left-most, 'earlier' columns for the digit '3' are crucial to avoid confusion with an '8'), which require longer memory capacities than were provided by the ring array. Adding the additional readouts did not improve memory capacity, hence there was a limited improvement to performance.

Strong correlations were observed between the reservoir metric maps and task performance maps. The regions of highest performance in both digit recognition task (indicated by symbols on the metric plots in Figures 5 and 6) were found at points that had high KR scores but avoided areas with high GR. MC was also strongly correlated to performance, with the highest performing reservoirs all having a memory capacity above 1.5. This indicated the importance of having both memory and nonlinearity in the system to provide useful transformations for classification. Practically, a hyperparameter search to optimise these reservoirs for a task could be confined within the decision boundaries outlined in the task-independent metrics section: KR – GR > 10, and MC > 1.25. This would reduce the parameter space of the original search considerably whilst still capturing the peak performance in both tasks. Decision boundaries overlayed on task performance heatmaps are shown in Figure 9.

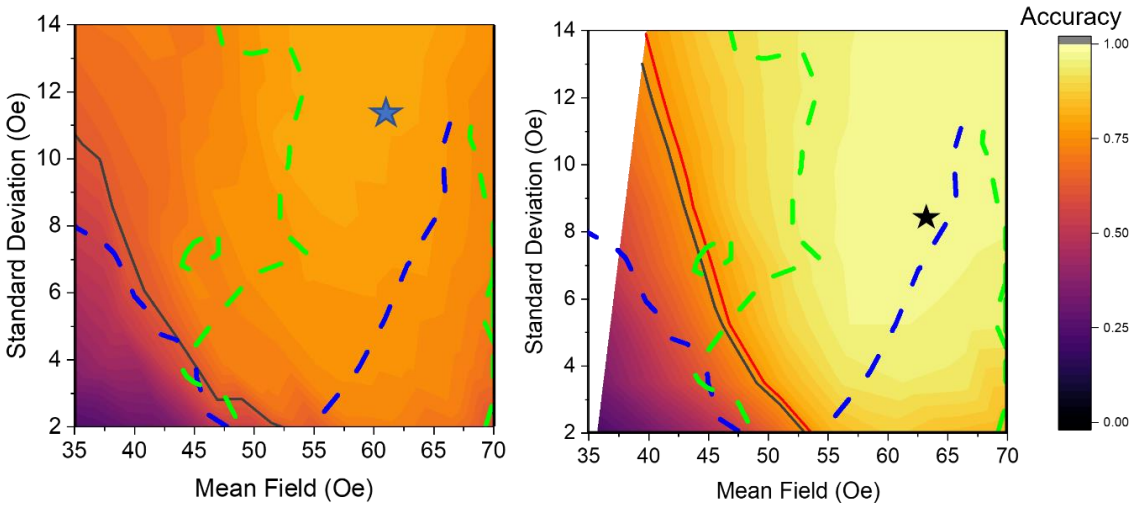


Figure 9- Decision boundaries of KR – GR > 10 (blue) and MC > 1.25 (green) overlayed on performance heatmaps for the MNIST task (left) and TI-46 task (right), showing peak performance is captured for both tasks by these decision boundaries.

In conclusion, we have shown that reservoirs based on interconnected magnetic nano-rings arrays can achieve a broad range of signal transformations, and explored the suitability of these for reservoir computing by calculating task independent metrics KR, GR, and MC. We then showed how the range of available metrics could be expanded by taking multiple concurrent measurements of the system's magnetic state. Finally, we demonstrated that these metrics correlated to performance in classification tasks and highlighted the substantial increase in performance that using the additional measurements of system state brought. These ring arrays are particularly interesting as RC candidates due to the non-volatility of the system's state when stimulus is removed, as well as the large scale of geometrical manipulation that is possible in order to tune the system's dynamic response. The continuous nature of the ring structure is useful as it allows for easy measurement of the system's state via electrical transport measurements. By employing two-stage lithographic techniques typical for magnetic nanostructure devices, non-magnetic electrical contacts can be superimposed upon the ring structures, allowing injection of current to the ring arrays for simple transport measurements via anisotropic magnetoresistance without altering the magnetic response of the array. Additionally, patterning the ring arrays into a multilayer with pseudo-spin-valve properties will enable a readout via giant magnetoresistance effects, which will significantly improve the magnitude of electrical signals resulting from state changes compared to anisotropic magnetoresistance. While the current state readout variables of domain wall population and magnetisation aren't directly accessible electrically, the resistance changes in the array due to magnetoresistance effects will be correlated to the magnetic state of the array and should provide sufficient richness for experimental emulation of the input/output paradigm demonstrated here in simulation. However, the current input mechanism limits inputs to a single dimension and creates large temporal separation between inputs, hindering the memory of the system. Since the magnetisation dynamics of the ring ensembles are spatially distributed, they are naturally well-suited to spatially multiplexed approaches where local data inputs are used to address discrete regions of the array as physical, rather than virtual, nodes. This should allow substantial enhancement of the arrays' memory characteristics. Our work represents an important step towards realising RC in magnetic ensembles with emergent magnetisation dynamics.

Acknowledgements

The authors thank STFC for beam time on the Offspec beamline at the ISIS Neutron and Muon Source (https://doi.org/10.5286/ISIS.E.RB1810656) and on beamline I06 at the Diamond Light Source, and

thank Jordi Prat, Michael Foerster and Lucia Aballe from ALBA for providing quadrupole sample holders[56]. I.T.V., R.W.D., and T.J.B., acknowledge DTA-funded PhD studentships from EPSRC. The authors gratefully acknowledge the support of EPSRC through grants EP/S009647/1, EP/V006339/1 and EP/V006029/1. This project has received funding from the European Union's Horizon 2020 FET-Open program under grant agreement No 861618 (SpinEngine).

Data Availability

The data that support the findings of this study are openly available in ORDA, at https://figshare.shef.ac.uk, reference number 10.15131/shef.data.16904806.

References:

[1] C. Mead, "How we created neuromorphic engineering," *Nature Electronics*, vol. 3, no. 7. Nature Research, pp. 434–435, Jul. 01, 2020. doi: 10.1038/s41928-020-0448-2.

[2] H. Jaeger, "The 'echo state' approach to analysing and training recurrent neural networks- with an erratum note," Bonn, Germany, 2001.

[3] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review*, vol. 3, no. 3, pp. 127–149, Aug. 2009, doi: 10.1016/j.cosrev.2009.03.005.

[4] L. Manneschi, A. C. Lin, and E. Vasilaki, "SpaRCe: Improved Learning of Reservoir Computing Systems Through Sparse Representations," *IEEE Transactions on Neural Networks and Learning Systems*, 2021, doi: 10.1109/TNNLS.2021.3102378.

[5] L. Appeltant *et al.*, "Information processing using a single dynamical node as complex system," *Nature Communications*, vol. 2, no. 1, 2011, doi: 10.1038/ncomms1476.

[6] Y. Paquot *et al.*, "Optoelectronic reservoir computing," *Scientific Reports*, vol. 2, 2012, doi: 10.1038/srep00287.

[7] M. Dale, J. F. Miller, S. Stepney, and M. A. Trefzer, "Evolving carbon nanotube reservoir computers," 2016. doi: 10.1007/978-3-319-41312-9_5.

[8] J. H. Jensen, E. Folven, and G. Tufte, "Computation in artificial spin ice," in *ALIFE 2018 - 2018 Conference on Artificial Life: Beyond AI*, Jul. 2020, pp. 15–22. doi: 10.1162/isal_a_00011.

[9] R. Nakane, G. Tanaka, and A. Hirose, "Reservoir Computing with Spin Waves Excited in a Garnet Film," *IEEE Access*, vol. 6, pp. 4462–4469, Jan. 2018, doi: 10.1109/ACCESS.2018.2794584.

[10] A. Welbourne *et al.*, "Voltage-controlled superparamagnetic ensembles for low-power reservoir computing," *Applied Physics Letters*, vol. 118, no. 20, p. 202402, May 2021, doi: 10.1063/5.0048911.

[11] R. v. Ababei *et al.*, "Neuromorphic computation with a single magnetic domain wall," *Scientific Reports 2021 11:1*, vol. 11, no. 1, pp. 1–13, Aug. 2021, doi: 10.1038/s41598-021-94975-y.

[12] J. Torrejon *et al.*, "Neuromorphic computing with nanoscale spintronic oscillators," *Nature*, vol. 547, no. 7664, pp. 428–431, Jul. 2017, doi: 10.1038/nature23011.

[13] M. Dale, J. F. Miller, S. Stepney, and M. A. Trefzer, "Reservoir Computing in Material Substrates," *Natural Computing Series*, pp. 141–166, 2021, doi: 10.1007/978-981-13-1687-6_7.

[14] M. Rafayelyan, J. Dong, Y. Tan, F. Krzakala, and S. Gigan, “Large-Scale Optical Reservoir Computing for Spatiotemporal Chaotic Systems Prediction,” *Physical Review X*, vol. 10, no. 4, p. 041037, Nov. 2020, doi: 10.1103/PHYSREVX.10.041037/FIGURES/7/MEDIUM.

[15] K. A. Tsakalos, G. C. Sirakoulis, A. Adamatzky, and J. Smith, “Protein Structured Reservoir computing for Spike-based Pattern Recognition,” *IEEE Transactions on Parallel and Distributed Systems*, 2021, doi: 10.1109/TPDS.2021.3068826.

[16] W. Yahiro, N. Aubert-Kato, and M. Hagiya, “A reservoir computing approach for molecular computing,” *Artificial Life*, 2018, Accessed: Jan. 26, 2022. [Online]. Available: http://direct.mit.edu/isal/proceedings-pdf/alife2018/30/31/1904938/isal_a_00013.pdf

[17] M. S. Kulkarni and C. Teuscher, “Memristor-based reservoir computing,” *Proceedings of the 2012 IEEE/ACM International Symposium on Nanoscale Architectures, NANOARCH 2012*, pp. 226–232, 2012, doi: 10.1145/2765491.2765531.

[18] J. Moon *et al.*, “Temporal data classification and forecasting using a memristor-based reservoir computing system,” *Nature Electronics 2019 2:10*, vol. 2, no. 10, pp. 480–487, Oct. 2019, doi: 10.1038/s41928-019-0313-3.

[19] A. Mehonic, A. Sebastian, B. Rajendran, O. Simeone, E. Vasilaki, and A. J. Kenyon, “Memristors—From In-Memory Computing, Deep Learning Acceleration, and Spiking Neural Networks to the Future of Neuromorphic and Bio-Inspired Computing,” *Advanced Intelligent Systems*, vol. 2, no. 11, p. 2000085, Nov. 2020, doi: 10.1002/AISY.202000085.

[20] L. Piraux *et al.*, “Giant magnetoresistance in magnetic multilayered nanowires,” *Applied Physics Letters*, vol. 65, no. 19, pp. 2484–2486, Jun. 1994, doi: 10.1063/1.112672.

[21] G. Bordignon *et al.*, “Analysis of magnetoresistance in arrays of connected nano-rings,” in *IEEE Transactions on Magnetics*, Jun. 2007, vol. 43, no. 6, pp. 2881–2883. doi: 10.1109/TMAG.2007.892597.

[22] M. Ono *et al.*, “Ultrafast and energy-efficient all-optical switching with graphene-loaded deep-subwavelength plasmonic waveguides,” *Nature Photonics 2019 14:1*, vol. 14, no. 1, pp. 37–43, Nov. 2019, doi: 10.1038/s41566-019-0547-7.

[23] N. Locatelli, V. Cros, and J. Grollier, “Spin-torque building blocks,” *Nature Materials 2014 13:1*, vol. 13, no. 1, pp. 11–20, Dec. 2013, doi: 10.1038/nmat3823.

[24] A. D. Kent and D. C. Worledge, “A new spin on magnetic memories,” *Nature Nanotechnology 2015 10:3*, vol. 10, no. 3, pp. 187–191, Mar. 2015, doi: 10.1038/nnano.2015.24.

[25] D. A. Allwood, G. Xiong, C. C. Faulkner, D. Atkinson, D. Petit, and R. P. Cowburn, “Magnetic domain-wall logic,” *Science (1979)*, vol. 309, no. 5741, pp. 1688–1692, Sep. 2005, doi: 10.1126/SCIENCE.1108813/ASSET/4E1829EF-7E73-4727-995D-718FEF5B4EDD/ASSETS/GRAPHIC/309_1688_F3.JPEG.

[26] Z. Luo *et al.*, “Current-driven magnetic domain-wall logic,” *Nature 2020 579:7798*, vol. 579, no. 7798, pp. 214–218, Mar. 2020, doi: 10.1038/s41586-020-2061-y.

[27] D. Bhowmik, L. You, and S. Salahuddin, “Spin Hall effect clocking of nanomagnetic logic without a magnetic field,” *Nature Nanotechnology 2013 9:1*, vol. 9, no. 1, pp. 59–63, Nov. 2013, doi: 10.1038/nnano.2013.241.

[28] S. S. P. Parkin, M. Hayashi, and L. Thomas, “Magnetic domain-wall racetrack memory,” *Science (1979)*, vol. 320, no. 5873, pp. 190–194, Apr. 2008, doi: 10.1126/SCIENCE.1145799/SUPPL_FILE/PARKIN.SOM.PDF.

[29] A. v. Kimel and M. Li, “Writing magnetic memory with ultrashort light pulses,” *Nature Reviews Materials 2019 4:3*, vol. 4, no. 3, pp. 189–200, Feb. 2019, doi: 10.1038/s41578-019-0086-3.

[30] D. Pinna, G. Bourianoff, and K. Everschor-Sitte, “Reservoir Computing with Random Skyrmion Textures,” *Physical Review Applied*, vol. 14, no. 5, p. 054020, Nov. 2020, doi: 10.1103/PhysRevApplied.14.054020.

[31] M. Zahedinejad *et al.*, “Two-dimensional mutually synchronized spin Hall nano-oscillator arrays for neuromorphic computing,” *Nature Nanotechnology 2019 15:1*, vol. 15, no. 1, pp. 47–52, Dec. 2019, doi: 10.1038/s41565-019-0593-9.

[32] T. Shibata *et al.*, “Linear and symmetric conductance response of magnetic domain wall type spin-memristor for analog neuromorphic computing,” *Applied Physics Express*, vol. 13, no. 4, p. 043004, Mar. 2020, doi: 10.35848/1882-0786/AB7E07.

[33] R. W. Dawidek *et al.*, “Dynamically-Driven Emergence in a Nanomagnetic System,” *Advanced Functional Materials*, vol. 31, no. 8, 2021.

[34] M. Negoita, T. J. Hayward, and D. A. Allwood, “Controlling domain walls velocities in ferromagnetic ring-shaped nanowires,” *Applied Physics Letters*, vol. 100, no. 7, Feb. 2012, doi: 10.1063/1.3685467.

[35] C. Wuth, P. Lendecke, and G. Meier, “Temperature-dependent dynamics of stochastic domain-wall depinning in nanowires,” *Journal of Physics Condensed Matter*, vol. 24, no. 2, 2012, doi: 10.1088/0953-8984/24/2/024207.

[36] T. J. Hayward and K. A. Omari, “Beyond the quasi-particle: Stochastic domain wall dynamics in soft ferromagnetic nanowires,” *Journal of Physics D: Applied Physics*, vol. 50, no. 8, p. 84006, 2017, doi: 10.1088/1361-6463/aa553f.

[37] C. A. F. Vaz *et al.*, “Ferromagnetic nanorings,” *Journal of Physics Condensed Matter*, vol. 19, no. 25, Jun. 2007, doi: 10.1088/0953-8984/19/25/255207.

[38] T. J. Hayward *et al.*, “Reading and writing of vortex circulation in pseudo-spin-valve ring devices,” *Applied Physics Letters*, vol. 89, no. 11, p. 112510, Sep. 2006, doi: 10.1063/1.2349305.

[39] T. J. Hayward *et al.*, “Switching behavior of individual pseudo-spin-valve ring structures,” *Physical Review B*, vol. 74, no. 13, p. 134405, Oct. 2006, doi: 10.1103/PhysRevB.74.134405.

[40] L. Büsing, B. Schrauwen, and R. Legenstein, “Connectivity, Dynamics, and Memory in Reservoir Computing with Binary and Analog Neurons,” *Neural Computation*, vol. 22, no. 5, pp. 1272–1311, May 2010, doi: 10.1162/NECO.2009.01-09-947.

[41] M. Dale, J. F. Miller, S. Stepney, and M. A. Trefzer, “A substrate-independent framework to characterize reservoir computers,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 475, no. 2226, Jun. 2019, doi: 10.1098/rspa.2018.0723.

[42] H. Jaeger, “Short term memory in echo state networks,” 2002. [Online]. Available: papers://78a99879-71e7-4c85-9127-d29c2b4b416b/Paper/p14153%5Cnhttp://neuron-ai.tuke.sk/~bundzel/diploma_theses_students/2006/Martin Sramko- Echo State NN in Prediction/STMEchoStatesTechRep.pdf

[43] M. Negoita, T. J. Hayward, J. A. Miller, and D. A. Allwood, “Domain walls in ring-shaped nanowires under rotating applied fields,” *Journal of Applied Physics*, vol. 114, no. 1, Jul. 2013, doi: 10.1063/1.4812388.

[44] A. Kunz and J. D. Priem, “Dynamic notch pinning fields for domain walls in ferromagnetic nanowires,” in *IEEE Transactions on Magnetics*, Jun. 2010, vol. 46, no. 6, pp. 1559–1561. doi: 10.1109/TMAG.2010.2041044.

[45] C. W. Sandweg *et al.*, “Direct observation of domain wall structures in curved permalloy wires containing an antinotch,” *Journal of Applied Physics*, vol. 103, no. 9, p. 93906, 2008, doi: 10.1063/1.2913318.

[46] M. P. Sharrock, “Time dependence of switching fields in magnetic recording media (invited),” *Journal of Applied Physics*, vol. 76, no. 10, pp. 6413–6418, 1994, doi: 10.1063/1.358282.

[47] W. Wernsdorfer *et al.*, “Experimental evidence of the Néel-Brown model of magnetization reversal,” *Physical Review Letters*, vol. 78, no. 9, pp. 1791–1794, Mar. 1997, doi: 10.1103/PhysRevLett.78.1791.

[48] S. Jain, C. C. Wang, and A. O. Adeyeye, “Magnetoresistance behavior of ferromagnetic nanorings in a ring-wire hybrid configuration,” *Nanotechnology*, vol. 19, p. 085302, 2008, doi: 10.1088/0957-4484/19/8/085302.

[49] A. v. Goncharov *et al.*, “Anisotropy of Magnetization Reversal and Magnetoresistance in Square Arrays of Permalloy Nano-Rings,” *IEEE Transactions on Magnetics*, vol. 42, no. 10, pp. 2948–2950, 2006, doi: 10.1109/TMAG.2006.878429.

[50] M. Liberman *et al.*, “TI 46-Word Corpus.” Linguistic Data Consortium, Philadelphia, 1993.

[51] L. Deng, “The MNIST database of handwritten digit images for machine learning research,” *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012, doi: 10.1109/MSP.2012.2211477.

[52] S. Molau, M. Pitz, R. Schlüter, and H. Ney, “Computing mel-frequency cepstral coefficients on the power spectrum,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2001, vol. 1, pp. 73–76. doi: 10.1109/icassp.2001.940770.

[53] M. Riou *et al.*, “Temporal pattern recognition with delayed-feedback spin-torque nano-oscillators,” *Physical Review Applied*, vol. 12, no. 2, Aug. 2019, doi: 10.1103/PhysRevApplied.12.024049.

[54] F. Abreu Araujo *et al.*, “Role of non-linear data processing on speech recognition task in the framework of reservoir computing,” *Scientific Reports*, vol. 10, no. 1, pp. 1–11, Dec. 2020, doi: 10.1038/s41598-019-56991-x.

[55] N. Schaetti, M. Salomon, and R. Couturier, “Echo State Networks-Based Reservoir Computing for MNIST Handwritten Digits Recognition,” in *2016 IEEE Intl Conference on Computational Science and Engineering (CSE) and IEEE Intl Conference on Embedded and Ubiquitous Computing (EUC) and 15th Intl Symposium on Distributed Computing and Applications for Business Engineering (DCABES)*, Aug. 2016, pp. 484–491. doi: 10.1109/CSE-EUC-DCABES.2016.229.

[56] M. Foerster *et al.*, “Custom sample environments at the ALBA XPEEM,” *Ultramicroscopy*, vol. 171, pp. 63–69, Dec. 2016, doi: 10.1016/J.ULTRAMIC.2016.08.016.

Supplementary Material

Further Metric Plots

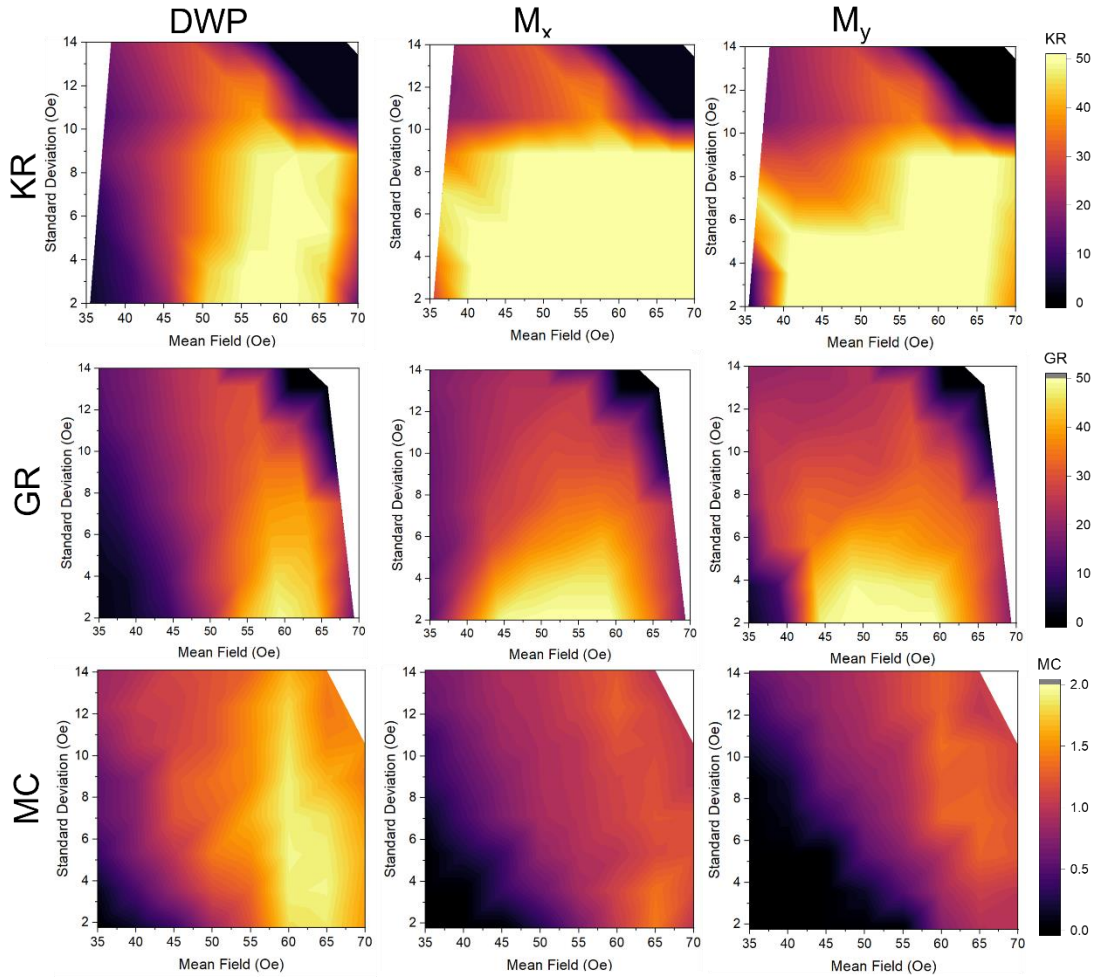


Figure S1 - Empirically measured reservoir metrics for an input rate of Nqrot = 1. Rows reflect kernel rank, generalisation rank, and memory capacity respectively. Columns reflect results produced for domain wall population, x magnetisation, and y magnetisation taken as output respectively.

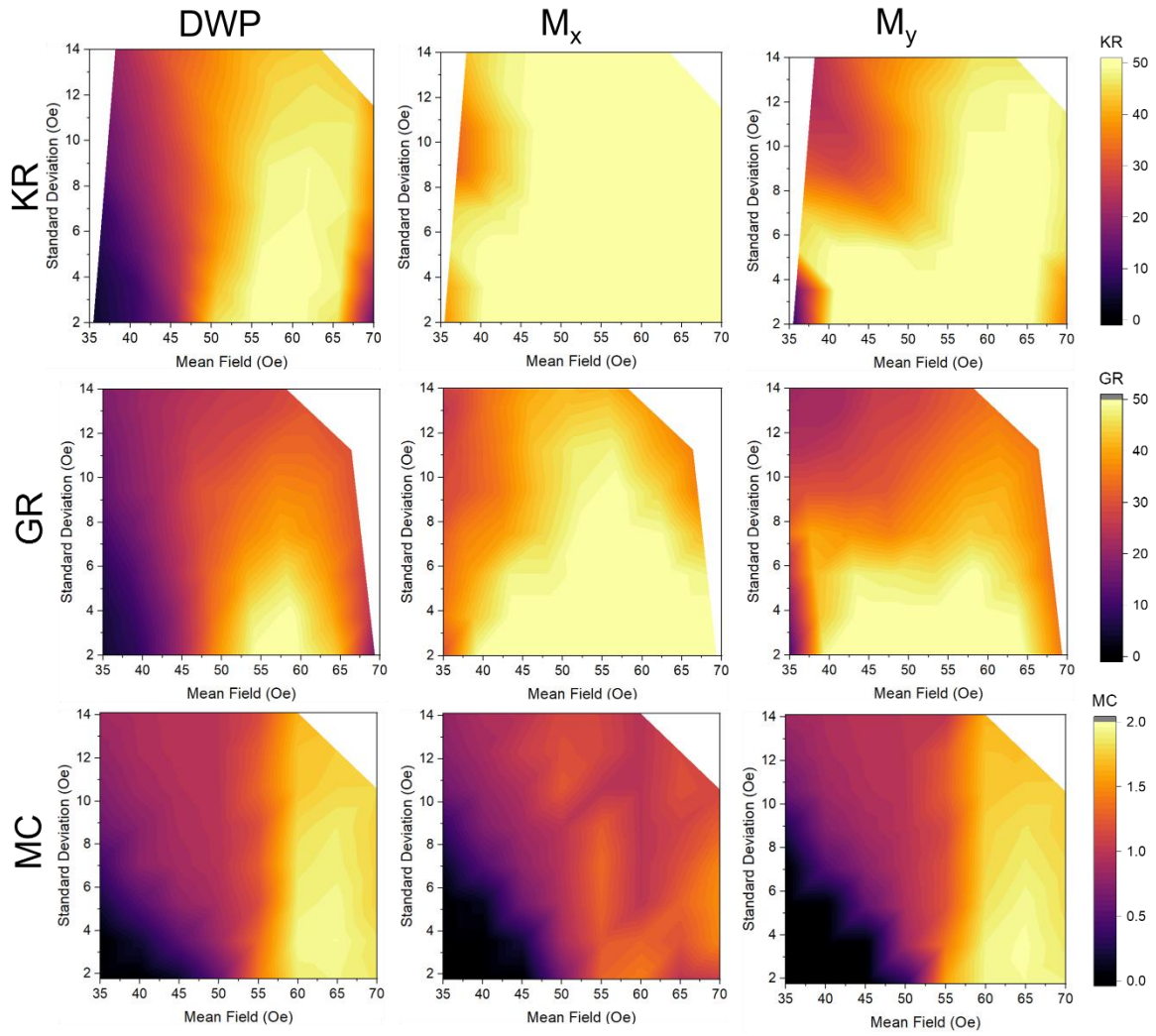


Figure S2 - Empirically measured reservoir metrics for an input rate of Nqrot = 4. Rows reflect kernel rank, generalisation rank, and memory capacity respectively. Columns reflect results produced for domain wall population, x magnetisation, and y magnetisation taken as output respectively.

Further 3-Variable TI-46 Heatmaps

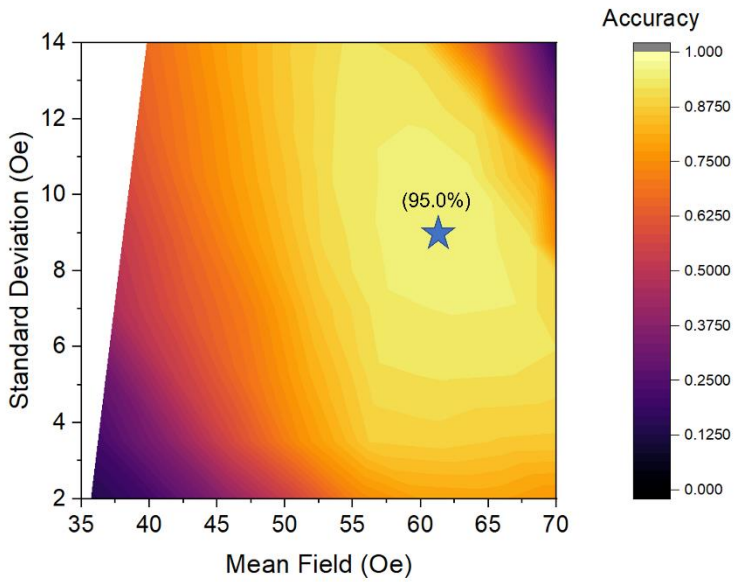


Figure S3- 3-State-variable output accuracy heatmap for TI-46 task. $N_q = 1$

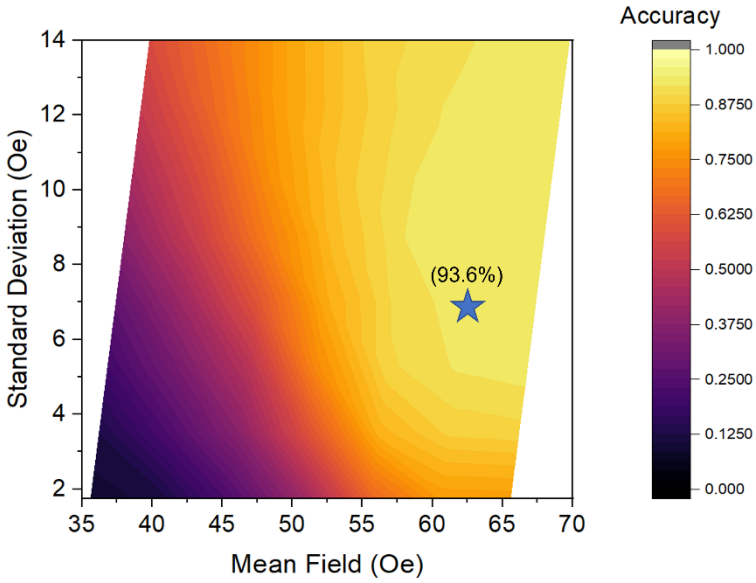


Figure S4- 3-State-variable output accuracy heatmap for TI-46 task. $N_q = 4$

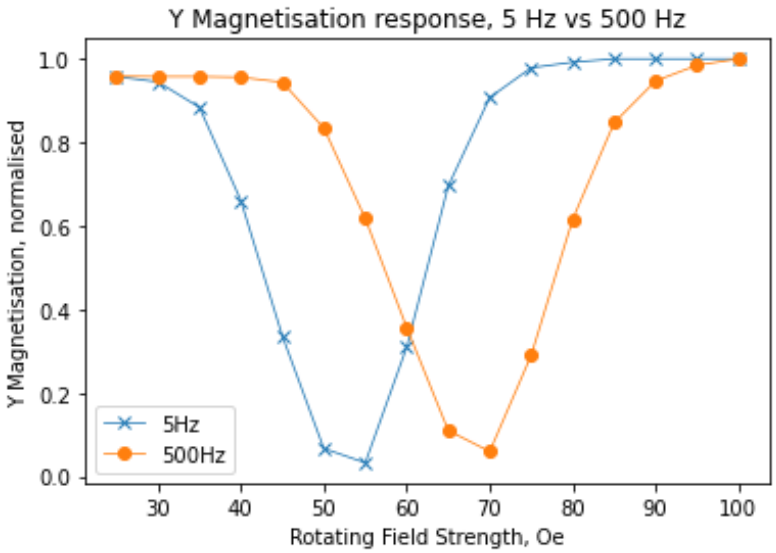


Figure S5- Plots showing simulated Y Magnetisation after 50 rotations of varying applied field for systems driven with drive frequencies of 5Hz (blue), and 500 Hz (orange).

7- ‘Reconfigurable Reservoir Computing with a Magnetic Metamaterial’

7.0- Preamble

The objective of this chapter was to take what we had learned from the simulations of the ring arrays and develop a comprehensive computational platform that could be reconfigured to solve a wide range of tasks on a single device. This required the development of an experimental rig that could input arbitrary field sequences to the ring arrays and provide a device tractable way of reading out information of the magnetic state of the array (i.e., no beamlines or large optical setups!).

To apply the rotating magnetic fields to stimulate the array, I designed a set of custom-built air coil electromagnets, allowing applications of arbitrary rotating fields to the nanoring arrays. In order to have full controllability over applying long field sequences and measuring device response, a codebase for interfacing with experimental equipment was designed. While I had created a rudimentary method for both controlling the electromagnets and logging measurement data within LabView, a postdoc within the group, Charles Swindells, provided a complete overhaul of this framework into an easily accessible and general set of python scripts which were then used for all future experiments.

Charles and I then worked extensively to optimise the measurement of anisotropic magnetoresistance of the ring arrays as fields were applied. Once a reliable method for measuring the devices was established, I undertook an exploration into the physical origins of characteristic features that were observed in the magnetoresistance response of the devices, using both experimental measurements and micromagnetic simulations, which were described in the article. From this exploration, a number of computationally interesting dynamical properties were identified.

Firstly, that there were two key frequency components of the AMR signal, each with a distinct nonlinear relationship with respect to applied field. The components arose from different mechanisms of magnetic response: the deformation of pinned domain walls in the system via susceptibility effects, and the propagation of domain walls around the rings. This was interesting for computation as it led to drastically different shapes in the resistance oscillations observed with respect to applied field. This meant that by taking fixed samples of the resistance level across a rotation, each sample had a distinct nonlinear response to the applied field, providing a means of providing nonlinearity and dimensionality expansion with the ring arrays. This led to the generation of a simplistic but novel method for reservoir computing with continuous-time oscillating responses; by recording the oscillation over time and having each time step represent a reservoir state readout.

The second interesting property was the observation of drastically different settling times in response to a step change in input depending on the magnitude of the input. This meant that the ring arrays had a range of different timescales of response that could be exploited under the common reservoir paradigm of networks constructed from a single dynamical node, as introduced in the previous chapter.

Finally, the ring arrays exhibited non-volatility of magnetic state when fields below a threshold value were applied. This was useful in generating memory behaviours, since it allowed storage of past inputs in the magnetic state of the array until another input that exceeds this threshold was provided. However, to exploit this, a network architecture consisting of distinct real nodes was required, as under the previous two architectures, new information would always wash out the state of the single node. To achieve this, the framework of the rotating neurons reservoir was employed. This allowed the

construction of a network of multiple nodes without having to directly provide connections between each of the nodes; instead a rotating pair of input/output connections emulate the interconnectivity. This provided a solution to the problem of low memory with the previously employed single-node arrangements.

In order to demonstrate the broad range of computational properties that could be achieved by changing the reservoir architecture around the dynamic system, metric space assessment was again performed and corroborated via performance of benchmark tasks with drastically different computational requirements. This showcased the increased flexibility offered when the choice of reservoir architecture synergised well with both the dynamic properties of the reservoir, as well as the demands of the task. This study marked a significant development for the thesis: the demonstration of a working device that provided flexible computational properties.

Specific details on author contributions can be found at the end of the article. This paper published in Communications Physics on 26/08/2023 and can be found online at https://doi.org/10.1038/s42005-023-01352-4.

Reconfigurable Reservoir Computing in a Magnetic Metamaterial

Vidamour, I. T.[1,2*]; Swindells, C.[1]; Venkat, G.[1]; Manneschi, L.[2]; Fry, P. W.[3]; Welbourne, A.[1]; Rowan-Robinson, R. M.[1]; Backes, D[4.]; Maccherozzi, F[4.]; Dhesi, S. S.[4]; Vasilaki, E.[2]; Allwood, D. A.[1]; and Hayward, T. J.[1]

1- Department of Materials Science and Engineering, University of Sheffield, Sheffield, UK
2- Department of Computer Science, University of Sheffield, Sheffield, UK
3- Nanoscience and Technology Centre, University of Sheffield, Sheffield, UK
4- Diamond Light Source, Harwell Science and Innovation Campus, Didcot, UK
*- Corresponding author, contactable at i.vidamour@sheffield.ac.uk

Abstract

In-materia reservoir computing (RC) leverages the intrinsic physical responses of functional materials to perform complex computational tasks. Magnetic metamaterials are exciting candidates for RC due to their huge state space, nonlinear emergent dynamics, and non-volatile memory. However, to be suitable for a broad range of tasks, the material system is required to exhibit a broad range of properties, and isolating these behaviours experimentally can often prove difficult. By using an electrically accessible device consisting of an array of interconnected magnetic nanorings- a system shown to exhibit complex emergent dynamics- here we show how reconfiguring the reservoir architecture allows exploitation of different aspects the system's dynamical behaviours. This is evidenced through state-of-the-art performance in diverse benchmark tasks with very different computational requirements, highlighting the additional computational configurability that can be obtained by altering the input/output architecture around the material system.

Introduction

In-materia computation, where the responses of material systems are exploited to perform computational operations, offers a potential alternative to conventional CMOS computing. Here, like in biological neurons, data processing operations are performed intrinsically via the physics governing the system's response to inputs. This offers potential improvements in both latency and power efficiency, as dynamical complexity and memory are inherent properties of the substrate. This removes the need to shuttle data between discrete memory and computational units, which can cost up to 100 times the energy of the computation itself when discrete memory units are located off-chip[1].

Reservoir Computing (RC)[2,3] is a bio-inspired computational paradigm which is especially harmonious with *in-materia* computation. In RC, a time-dependent 'reservoir' layer (typically a recurrent neural network, RNN) provides complex nonlinear representations of input data, and a time-invariant readout layer provides a weighted output of the evolving state of the reservoir. Only the readout layer is trained, alleviating the training difficulties associated with standard RNNs since temporal dependencies of the reservoir layer are decoupled from the simple linear output[4].

As the response of the RNN is mathematically analogous to that of a dynamic system, it can be substituted with a real-world dynamic system with appropriate properties, namely nonlinearity between input and output, and a dependence on previous state that asymptotically diminishes over time, termed a 'fading memory'. This has led to a plethora of proposed implementations, with platforms including optoelectronic[5–7], molecular[8], mechanical[9–11], biological[12,13], memristive[14–16], and magnetic[17–23] systems.

Nanomagnetic platforms are of particular interest for RC due to their inherent hysteretic behaviours and nonlinearity of system dynamics, satisfying the two broad criteria necessary for RC. Many magnetic systems have been proposed as reservoirs and come with their own strengths and weaknesses. Spin-torque nano-oscillators[17,24,25] offer high data-throughput and passive

synchronization, and can be characterised using simple electrical measurements. The all-electric nature of the input/output to these oscillators has allowed for small artificial neural networks (<10 nodes) to be demonstrated experimentally[26,27], and larger networks have been simulated for RC with binary inputs[28,29]. However, the intrinsic dynamics of single oscillators are relatively simple (though they can be augmented via external delayed feedback[30,31]) and have durations on the order of nanoseconds, limiting their suitability to processing applications where sensory data arrives with characteristic timescales on the order of seconds- far beyond the intrinsic decay times of these systems. Magnetic metamaterials (materials which are engineered to exhibit complex physical responses beyond their underlying material properties) such as artificial spin-ice systems[19,20,32] and skyrmion textures[33], represent an exciting subcategory for magnetic RC, boasting complex, spatially distributed responses. However, interfacing with these materials is challenging, since spin-ices are electrically discontinuous and skyrmion textures require sub-100K temperatures, inhibiting device-tractable measurement approaches.

While there have been many recent, important developments showcasing device-specific RC performance in a range of physical systems, many more general questions remain, such as how different RC architectures can be used to extract different computational properties, and how these architectures can best synergise with the underlying system dynamics. Frequently, the 'single dynamical node' paradigm[34] is employed with little attention to its role in the computation or to the alternative computational properties that could be extracted with different reservoir architectures. This leaves some of the broader potential of nanomagnetic RC as reconfigurable computational platforms untapped.

In this paper, we experimentally demonstrate a pipeline from characterization of device physics, to reservoir design, to state-of-the-art performance in several, diverse computational tasks with a single magnetic device consisting of an array of interconnected magnetic nanorings[35]. The nanoring system boasts a combination of highly complex system response and simple electrical readout: strong coupling between individual ring elements produces complex 'emergent' dynamics (where large-scale responses arise from the collective effects of simple interactions between elements, rather than the properties of the elements themselves), while the continuous nature of the patterned nanostructure facilitates electrical transport measurements. Additionally, non-volatile domain configurations formed in response to input provides a natural means of generating system memory at driving fields an order of magnitude smaller than spin-ice systems[36]. To harness these emergent behaviours, we employ the device in three distinct reservoir architectures that each leverage different aspects of its dynamical properties. We then demonstrate how this provides flexible computational functionality by performing benchmark tasks with contrasting computational requirements on a single device, achieving state-of-the-art accuracies. This highlights the reconfigurability achievable in *in materio* platforms via careful choice of the accompanying RC architecture.

Results

I- Response of Nanoring Arrays

The devices studied here consist of arrays of 10nm thick $Ni_{80}Fe_{20}$ (Permalloy, Py) nanorings, patterned into a square lattice with each ring having nominal diameters of 4μm and track widths of 400nm, each overlapping with its nearest neighbours across 50% of their track widths[35,36]. The arrays were fabricated by electron beam lithography with lift-off processing and metallised via thermal evaporation. Ti/Au electrical contacts were then added via additional lithography and deposition steps, allowing measurements of the device's anisotropic magnetoresistance (AMR). Despite typical AMR ratios of 3-4% for Py[37], shape anisotropy in the rings means that magnetisation typically runs parallel to the applied currents. This meant only the domain walls which present local changes in magnetisation direction can be detected via AMR, leading to an effective AMR ratio of 0.2% for the

device, with the signal quality improved via lock-in amplification techniques (see Methods- Electrical transport measurements). The samples have saturation magnetisation $\mu_0 M_s$ of 0.969 ± 0.006 T, determined via broadband ferromagnetic resonance measurements (see supplementary note 4, and supplementary figure S6). Figure 1a shows a scanning electron microscope image of the device.

In previous studies[35,36], we have shown that interconnected nanoring arrays exhibit emergent magnetization dynamics under rotating in-plane magnetic fields. At the microstate, each ring exists in one of three metastable configurations, defined by the number and position of domain walls (DWs) it possesses, with configurations for 'vortex' (zero DWs), 'onion' (two DWs, 180º separation), and 'three-quarter' (two DWs, 90º separation) shown in figure 1b. To initialise the ring arrays, a strong pulse of magnetic field and subsequent relaxation leads to a uniform state of aligned onion rings, with DWs pointing along the direction of the saturation pulse. Under high driving fields, the DWs can coherently propagate with the applied field, maintaining onion configuration. However, under lower driving fields, stochastic pinning events cause differential movement of DWs within a ring (onion to three-quarter transition), potentially leading to DW annihilation (three-quarter to vortex transition) when itinerant DWs in the same ring collide. DWs can be restored in rings via the propagation of a DW in neighbouring ring, with the magnetic reversal across the junction between the two rings leading to injection of a pair of DWs in the empty ring (vortex to onion/three-quarter transition). Schematics for these processes are shown supplementary figure 2g, 2h. Whilst these behaviours are stochastic at the local scale, interactions between many rings lead to a well-defined global emergent response, providing a complex yet repeatable dynamic state evolution (Figure 1c-1e).

To evaluate the evolving magnetic states of the arrays for computation, AMR measurements performed via the electrical contacts shown in Figure 1a. This gives a single global readout for each array, which varies over a given input rotation. Initially, the device's response as a function of rotating field amplitude was surveyed to determine the characteristics of the responses and identify computationally useful features (Figure 1f). Fourier analysis of the AMR response led to observation of two distinct signals with frequencies that match (1f signal) as well as double (2f signal) the frequency of the rotating magnetic field, with the relative magnitude of the two signals with respect to driving field amplitude shown in Figure 1f(i) (see Supplementary note 1 for further Fourier analysis). Physically, these processes can be separated into elastic deformation of the rings' domain structures due to susceptibility effects (1f, dominant at lower fields), and irreversible DW propagation between pinning sites in the rings (2f, dominant at higher fields). Further details of these mechanisms can be found in the supplementary note 2. The dynamic nature of the system's response was evaluated by measuring the number of rotations that were required for the AMR signal to reach dynamic equilibrium (<2% amplitude variance between cycles) from saturation, with the measured timescales and the underlying signals shown in figures 1f(ii), and 1f(iii) – 1f(vi) respectively. The onset of DW motion can also be observed at a ~22 Oe, marked by the nonlinear increase of 1f signal in figure 1F(i), as well as the start of varying time-signals between cycles in figure 1f(iii) – 1f(vi).

Figure 1: Overview of static and dynamic responses of nanoring arrays

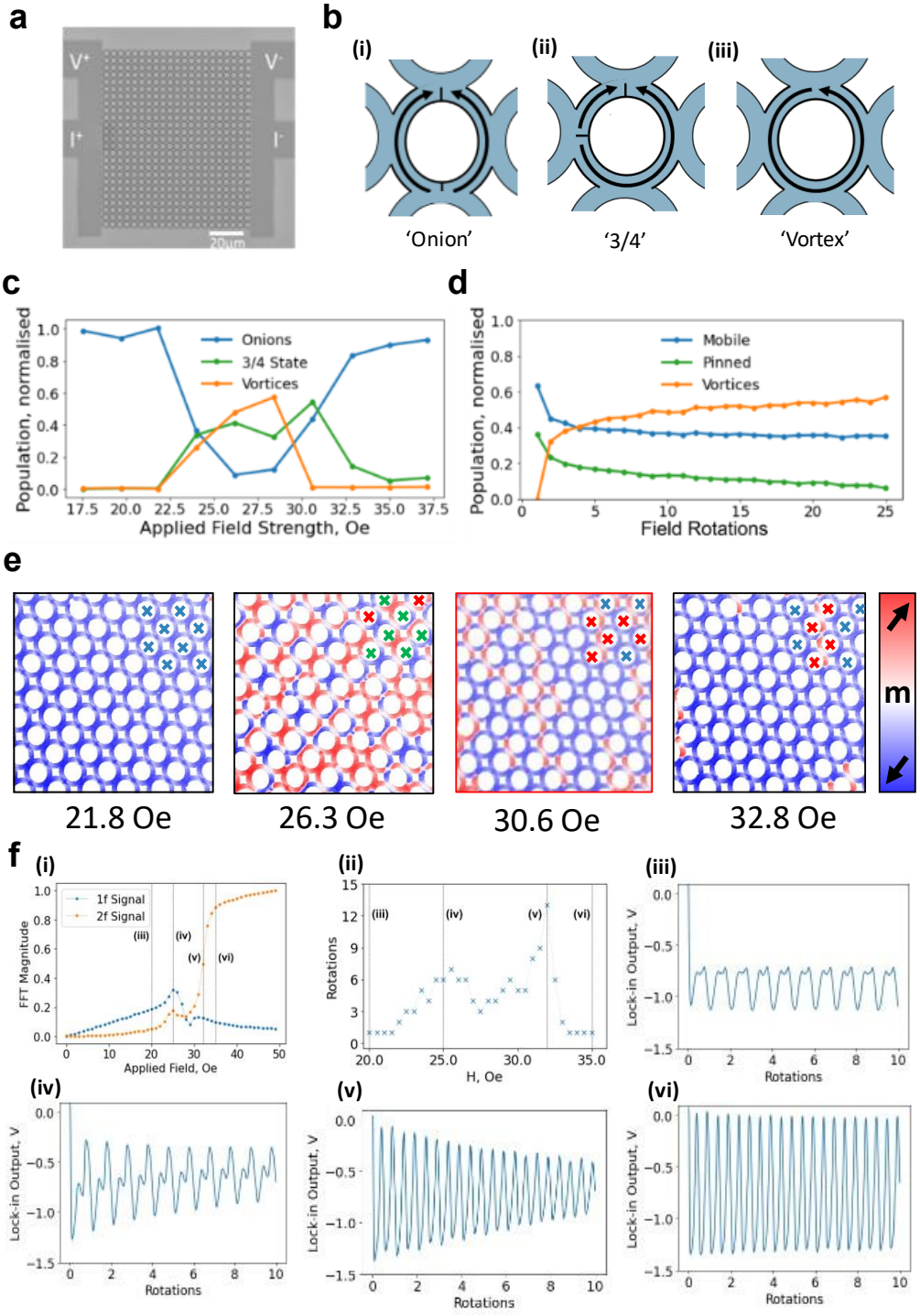


Figure 1a- Scanning electron microscope image showing a nanomagnetic ring array and electrical contacts. 1b- Schematics of available ring domain states, showing (i) Onion, (ii) Three-quarter, and (iii) Vortex. 1c- Varying state population of an array as driving field is increased, taken by counting populations of X-Ray photoemission electron microscopy images (X-PEEM) images after 30 rotations of applied field. 1d- Population of mobile, pinned, and vortex state rings over 25 successive cycles of 27 Oe rotating applied field, inferred from time-varying AMR signals. 1e- X-PEEM images of ring arrays when driven with 30 field rotations of amplitude 21.8, 26.3, 30.6, and 32.8 Oe of applied field. Magnetic contrast is given by the direction arrows on the colour bar, crosses in the top right corner rings denote (blue) onion, (green) vortex, and (red) three-quarter ring configurations. 1f(i)- Fourier components of AMR signal of arrays driven with 10 rotations of magnetic field at various applied fields. Blue datapoints show Fourier component at the driving field frequency (1f), while orange datapoints show

component at double the driving field frequency (2f). (ii)- Number of rotations of field required for the system to reach an equilibrium response (<2% peak-to-peak deviation between cycles) from saturation for a range of applied fields. (iii)-(vi) - Measured lock-in voltage of the array when driven with 10 field rotations of amplitude (iii) 20 Oe, (iv) 25 Oe, (v) 32 Oe, and (vi) 35 Oe from saturation.

From these measurements, three computationally promising properties can be identified. Firstly, the distinct variation of the AMR frequency components with respect to field provides crucial nonlinearity. Secondly, the dependence of the device's response on its past states, as evidenced by the range of timescales observed in the AMR signals, allows information to be connected across time in manner reminiscent of the echo-state property of echo state networks (ESNs). Finally, the presence of a threshold field below which no irreversible DW motion occurs shows a non-volatility of system state, providing pathways to longer-term storage of information.

The key demonstration of this paper is how these physical behaviours of the nanoring devices can be harnessed in different ways to create RCs with different computational properties, and thus tackle problems with different computational requirements. We achieve this by incorporating the device into three distinct reservoir architectures: an approach which takes advantage of the time-continuous oscillations of the nanoring array (signal sub-sample reservoir), the 'single dynamical node' architecture introduced by Appeltant et al.[34], and the recently proposed 'rotating neurons reservoir' of Liang et al.[38], yet to be deployed outside of analogue electronic RC. These architectures are presented schematically in Figure 2 and described in their respective Methods sections. In the following, we will explain how each of these architectures allows different computational properties to be emphasised and then exploited to perform challenging computational tasks. For further details on the methods employed for the machine learning tasks, see Supplementary note 5.

II- Signal Sub-sample Reservoir

One foundational task for RC platforms is nonlinear signal transformation[20,32,39–41]. In this problem, the system is provided with input of a given periodic response and is tasked with transforming the input signal into a different target signal. To perform this task, the reservoir should provide a higher-dimensional, nonlinear representation of the input signal so that the transformation between the input and the target can be computed via a simple linear readout.

To meet these computational demands, we designed a simple reservoir input/output architecture that directly exploited the non-linear variation of the 1f/2f frequency signals (Figure 3a). Here, each input datum scaled the field amplitude for a single rotation, and the resulting AMR response was sampled at 32 times per input, expanding input dimensionality 32-fold. The two frequency components have different nonlinear variations with respect to input magnitude, meaning that the relative magnitude of the continuous signal at fixed sample points will have nonlinear variation with respect to each other, providing dimensionality expansion of the input data. This offers a very simple method for providing increased nonlinearity in physical systems with continuous signals, obtained by leveraging a phase transition in system response.

Figure 2- Schematic diagrams of each reservoir architecture

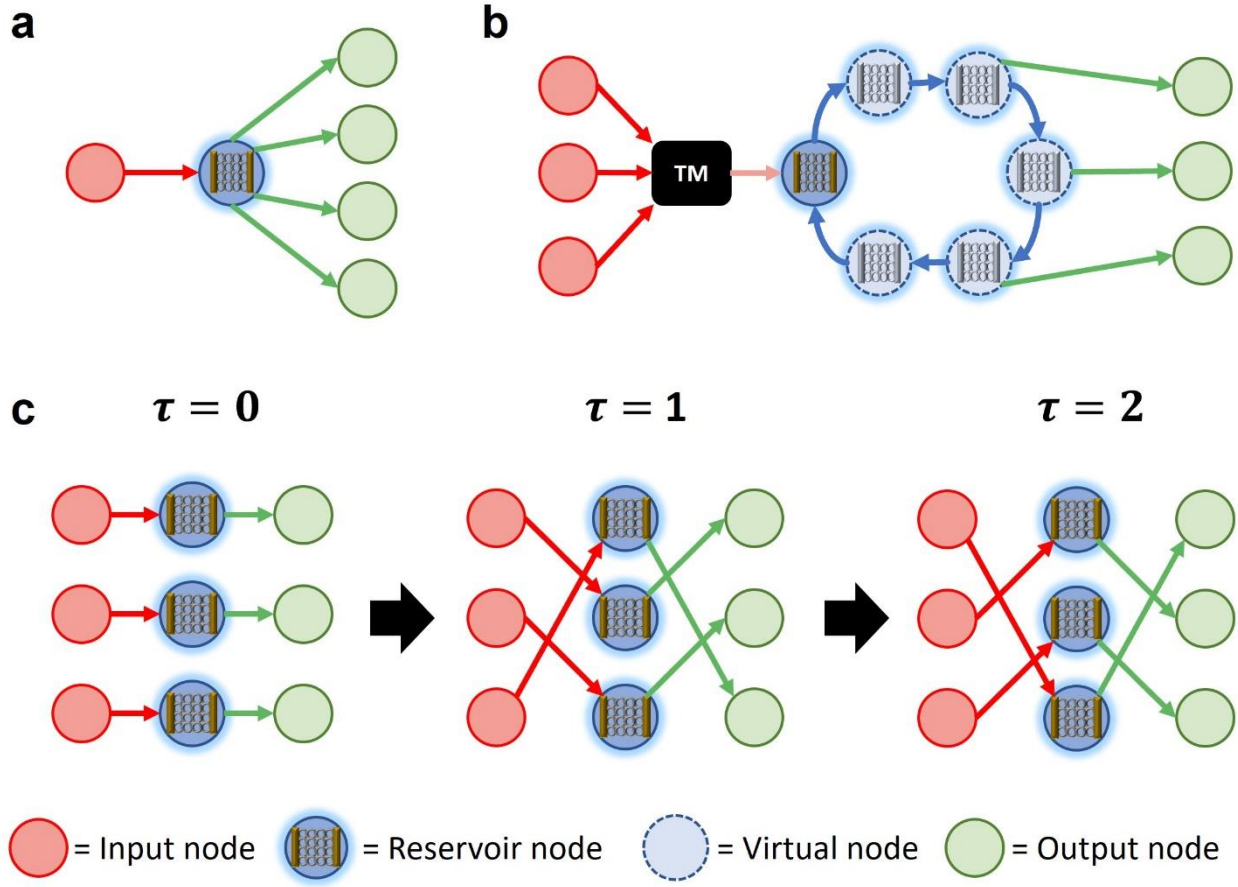
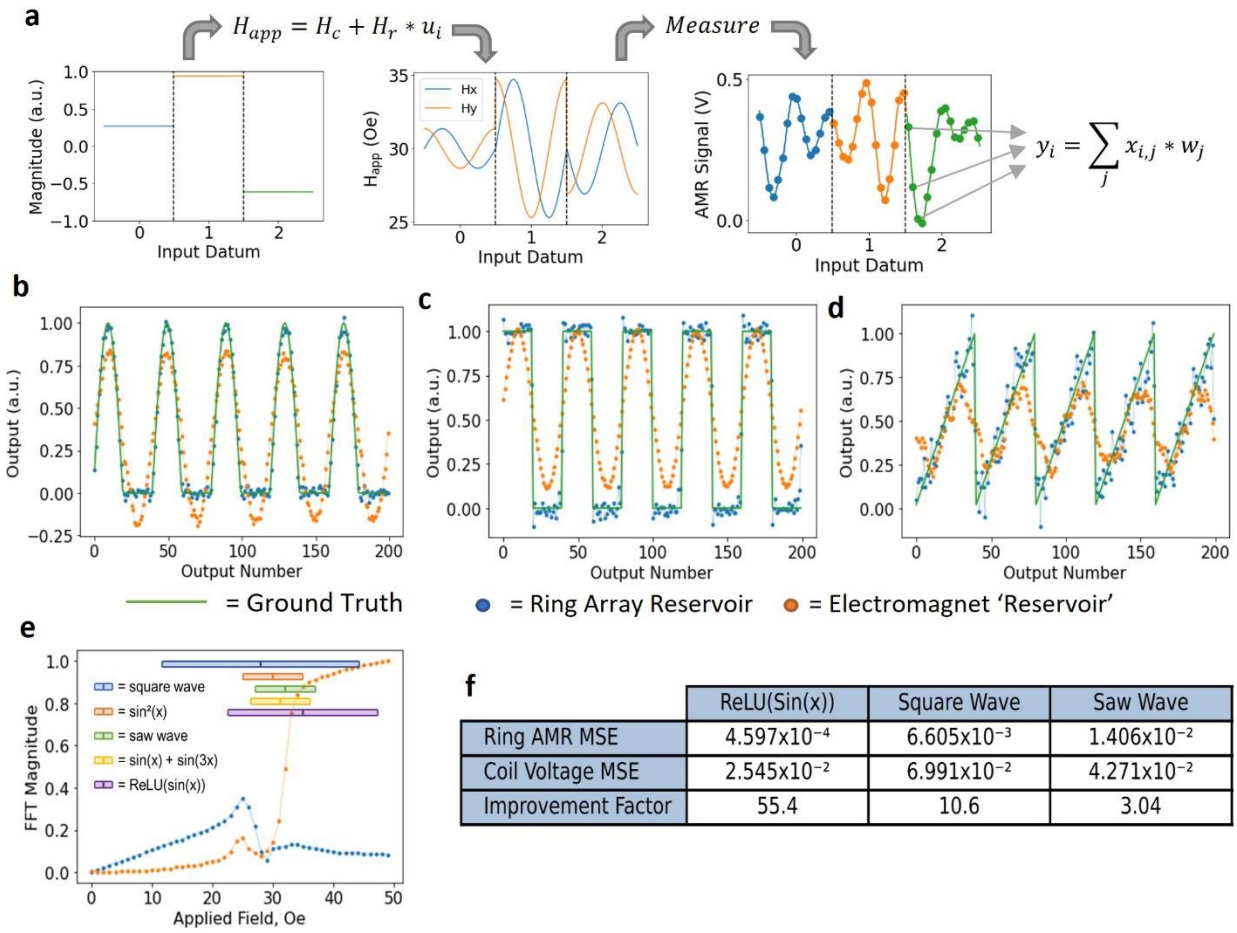


Figure 2- Schematic diagram showing three different reservoir architectures (a-c), with differing methods for providing input data (red circles) into reservoir nodes (blue circles) and reading reservoir state as output (green circles). a- Signal sub-sample architecture, showing a singular input datum fed into the ring arrays, with multiple state readouts taken from the single node. b- Single dynamical node architecture, where multiple input dimensions are time-multiplexed (black rectangle), before being fed into a single node. 'Virtual nodes' (pale blue circles), are generated from the dynamical node as input varies over time, generating outputs for each virtual node. c- Revolving neurons architecture, where the weighted connections between input-to-reservoir and reservoir-to-output change consecutively with each input timestep τ.

Figure 3b-3d shows the resulting signal reconstruction when the ring array system was tasked with transforming sinusoidal input to ReLU(sin(x)) (rectified linear unit), square wave, sawtooth waveforms. To evidence the impact of the metamaterial on computation, a control experiment was performed by recording the voltage of one of the driving electromagnets as the measured reservoir state instead of the resistance of the nanoring array. This provided equal dimensionality expansion as the nanoring array transformation, but without the nonlinearities contributed by the nanoring system. However, these measurements do contain any hardware-based nonlinearities in the electromagnets such as slew-rate between inputs and inductive effects, accounting for any nonlinearities provided by the experimental equipment. The ring array network outperformed the control network in all cases, offering up to a 55-fold reduction in MSE ($4.6\text{x}10^{-4}$ compared to $2.5\text{x}10^{-2}$) when replicating the ReLU function. The rings also perform favourably compared to proposed spin-ice platforms, with lower errors for Sawtooth ($1.406\ \text{x}10^{-2}$ vs $1.919\text{x}10^{-2}$) and Square ($6.605\ \text{x}10^{-3}$ vs $2.429\ \text{x}10^{-2}$) waves[20]. The different reconstruction tasks are performed optimally at very different ranges of applied field (Figure 3e), highlighting how the ring array's dynamics can be further tuned to for better performance in a range of similar problems even when held within a consistent reservoir architecture. The accuracies for all transformations for both the ring array and control network, as well as the ratio between them, are shown in the figure 3f.

Figure 3- Performance of signal transformation task



	ReLU(Sin(x))	Square Wave	Saw Wave
Ring AMR MSE	4.597×10^{-4}	6.605×10^{-3}	1.406×10^{-2}
Coil Voltage MSE	2.545×10^{-2}	6.991×10^{-2}	4.271×10^{-2}
Improvement Factor	55.4	10.6	3.04

Figure 3a- Schematic diagram showing scaling of input data u_i to applied field H_{app}, application of field rotations as components of field in x and y, H_x and H_y respectively, followed by sampling of resulting anisotropic magnetoresistance (AMR) signals to produce features, combined via a weighted sum to produce output. 3b-f: optimal reconstructions obtained from the Ring Array (blue) as well as the control measurements of electromagnet voltage (orange) compared to the desired target signal (green), for (b) rectified linear unit(sin(x)), (c) square wave, (d) saw wave. 3e- Input scaling parameters corresponding to reservoir configurations with minimum error for the signal reconstruction task, overlayed on relative 1f (blue dotted line) and 2f (orange dotted line) signal magnitudes over a range of applied fields. Bar width demonstrates applied field range (H_r), with central field (H_c) marked by the solid line. 3f- Comparison of mean-squared error between target signal and reconstructions drawn from the measured Ring Array voltages, as well as a control measurement taken from voltage measurements of the driving electromagnets.

III- Single Dynamical Node Reservoir

Another key application for RC is the classification of time varying signals such as spoken digits, a task which has been previously used to benchmark a variety of RC platforms[17,21,25,34]. While input data for the previous task was 1-dimensional, input data for speech recognition tasks are typically multi-dimensional. Furthermore, non-linear interactions between these input dimensions in the reservoir are essential to successful classification. Here, we consider classification of the spoken digits 0-9 from the TI-46 database (see Supplementary Note 5- Spoken Digit Recognition Task for details). The input data was 13-dimensional, consisting of the results of applying Mel-frequency cepstral filters[42] to each utterance. The data is linearly inseparable, with classification accuracy being limited to around 75%[25] if input data is passed directly to a linear readout layer. The role of the reservoir is to provide a non-

linear mapping of input data into higher dimensional reservoir space, thus allowing the linear readout layer to establish hyperplanes which can classify the data accurately.

Tackling this problem requires a reservoir architecture that expresses the non-linearity of the device's AMR response, can accommodate multiple input dimensions, and allows nonlinear combinations of these input dimensions, properties that cannot be provided by the signal sub-sample architecture. To satisfy these requirements we adopted the single dynamical node approach (Figure 4a) initially proposed by Appeltant *et. al.*[34] and detailed in the Methods section. Multidimensional input data was fed sequentially into the device, creating a reservoir constructed of 'virtual' nodes that convolves inputs temporally via the ring array's transient dynamics. Thus, this approach leveraged both the non-linear response of the device's AMR signal to input (via the activation of the virtual nodes), and its transient nature (which allowed interaction between virtual nodes).

As shown in the previous task application, our device exhibited a broad range of responses that were potentially useful for computation. Searches over parameter space can be performed for simple tasks such as signal transformation, however for more data-intensive tasks, this process is inefficient. Previous studies have shown that task-agnostic metrics, which can be found via statistical analysis of small random datasets[36,43], can speed up parameter selection by identifying promising regions of parameters space. Using metrics of kernel rank (KR, the ability of the reservoir to separate different input classes) and generalization rank (GR, and the ability of the reservoir to generalise inputs of the same class), we evaluated the computational properties of the device's transformations for a range of scalar parameters controlling the scaling (H_r) and offset (H_c) of inputs (see supplementary note 3). As the spoken digit recognition task required improving the linear separability of input data, KR was chosen to be the key identifier of promising performance, with a comparatively lower GR also needed to generalise between the different speakers.

To highlight the single dynamical node approach's better suitability to the spoken digit recognition task, metric maps were also drawn similarly for the other reservoir architectures (Supplementary figure S4). While the revolving neurons reservoir showed good separation properties (high KR), and the signal sub-sample reservoir good generalisation properties (low GR), only the single dynamical node architecture exhibited a balance of the two, showing better suitability for classification tasks. This is likely due to the rotating neuron reservoir's increased dependency upon past states reducing its ability to generalise, and the relatively smaller dimensionality of the signal sub-sampling reservoir leading to poorer ability to separate information. Conversely, the single dynamical node approach both provides good dimensionality expansion, as well as having decreased dependence on past states via the increased separation of inputs over time provided by the time multiplexing procedure.

Figure 4b shows the error rates versus training samples for spoken digit recognition, obtained using both a 'promising' (H_c = 29 Oe, H_r = 10 Oe, KR = 76) and arbitrarily chosen reservoir configurations (e.g., H_c = 21 Oe, Scaling, H_r = 7.5 Oe, KR = 52). 100-fold cross-validation was performed to evaluate general performance and find a suitable regularization parameter, selected for best performance on the training set to prevent overfitting. Again, a reservoir constructed from the voltage signals across the driving coils was used as a control, effectively skipping the reservoir transformation whilst including the same pre-processing steps. A significant reduction of word-error-rate is observed moving from 'control' to 'arbitrary' to the 'promising' case, with error-rates of 24.8%, 10.4%, and 4.6% respectively. This demonstrates not only the effectiveness of the reservoir's transformations in improving the linear separability of the data, but also the utility of evaluating metric scores to expedite system parameter selection.

One method for further improving performance commonly employed in conventional RC settings is the use of bespoke learning rules instead of standard regression-based training methods. Here, the

SpaRCe[44] algorithm was used, which was developed for use on ESNs though thus far has not been applied to physical systems, and its online nature synergises well with life-long learning paradigms especially useful for system-level device applications[45]. The algorithm aims to suppress confounding information and induce sparse output representations. Here, these properties help to mitigate the effects of experimental noise and remove redundant virtual node outputs. With SpaRCe, the accuracy was improved to 99.8%, as shown in Figure 4c. The ring arrays matched state-of-the-art performance compared to other magnetic architectures, even with fewer (50) virtual nodes used in the time-multiplexing procedure (STNOs with 400 virtual nodes, 99.8%[17], simulations of superparamagnetic arrays with 50 virtual nodes, 95.7%[21]), and improved upon the performance achieved in simulations of the ring system (97.7%[36]).

Figure 4- Performance of spoken digit recognition task

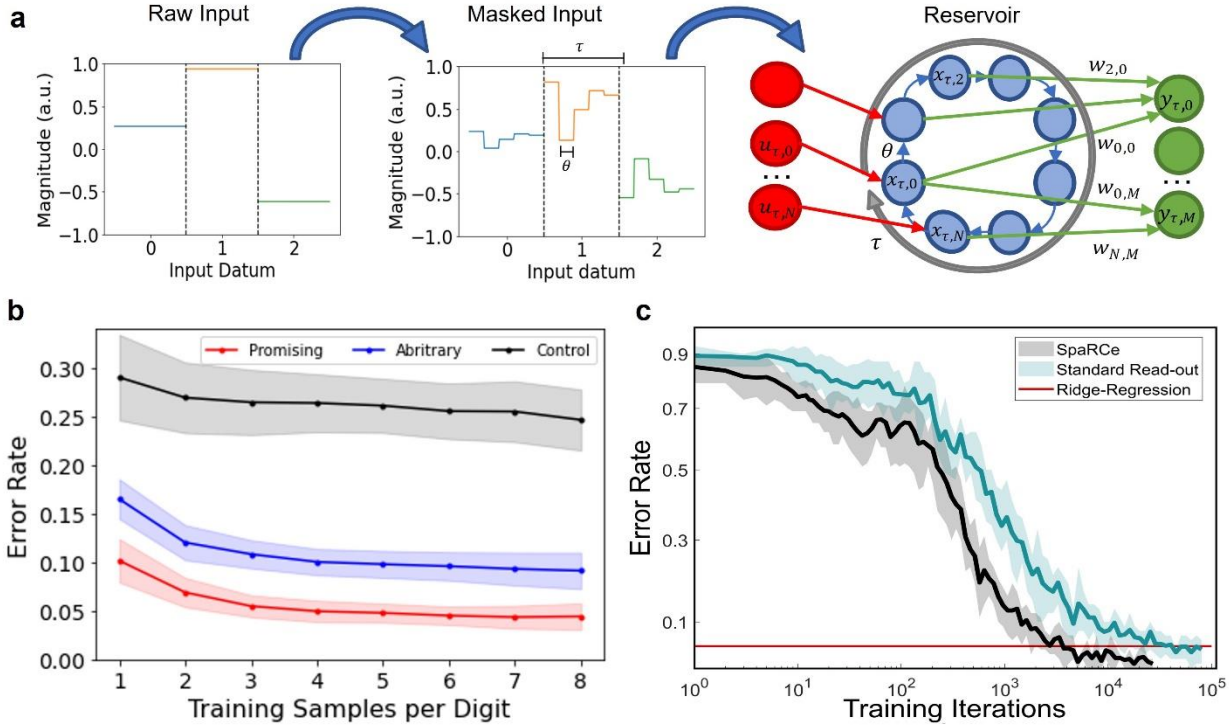


Figure 4a- Process showing time multiplexing procedure, taking raw inputs, combining them with a fixed mask to produce masked input (each of N virtual nodes has an input of duration θ, totalling to a duration of $\tau = N\theta$ per unmasked input), then inputting those inputs to the reservoir and measuring evolving reservoir state. 4b- Error rate versus number of sequences used for training for 'promising' (H_c / H_r = 29±10 Oe, (red)) and 'arbitrary' (H_c / H_r = 21±7.5 Oe, (blue)) reservoir parameters, and control measurements taken from voltage readings of the input electromagnets for the 'promising' case (black). The shaded regions show the standard deviation of performance over the 100-fold cross-validation. 4c- Error rate vs training iteration comparison between online learning methods using the SpaRCe algorithm (black) and standard online learning (blue) for a system driven with H_c and H_r values of 29±10 Oe. The shaded region shows minimum and maximum accuracies over 10-fold cross-validation. Red line shows accuracy achieved with ridge regression.

IV- Revolving Neurons Reservoir

In addition to data classification tasks, RC is also highly applicable to time series prediction problems. To be successful in these tasks, RC platforms often require fading memory of past inputs to correctly predict future trajectories, in addition to the non-linear properties that were exploited in the previous two tasks. The memory of a reservoir can characterised by evaluating the linear memory capacity[46] (MC), which measures the ability to reconstruct past inputs from the current reservoir state over increasing delays. Typically, nanomagnetic RC platforms exhibit low MC without the inclusion of

delayed feedback due to the short timescale of intrinsic dynamic behaviors[23]. Additionally, reservoirs constructed under the single dynamical node paradigm struggle to recall previous input datapoints due to the long temporal separations between each input created by the time-multiplexing procedure. For example, the prior architectures presented here exhibited peak MC < 3, meaning they could only reliably recall the previous two inputs (see supplementary figure S5).

To utilise the system's non-volatile properties and create a architecture better suited to time series prediction tasks, the recently proposed 'rotating neurons reservoir'[38] (RNR, see Methods) configuration was employed. Here, the system was constructed from 50 distinct dynamical nodes, with inputs to each node modulated by a fixed, rotating input (output) mask which multiplied input (reservoir state) values by ±1 (weight value), shifting the input/output connections to each node by one position every timestep (Figures 2c, 5a).

The memory effects exhibited in this configuration emerge from the ratchet-like nature of the device's non-volatile response: small inputs cause reversable perturbations while larger inputs cause non-volatile changes to underlying domain structure. In the RNR configuration, this means the system's evolution is dependent upon the sign of the input at a given time, determined by the mask. For negative mask values, the low applied field strengths leave the rings' domain structures unchanged through multiple timesteps until a positive input is applied to the system, where the higher applied fields cause DW propagation which is then measured as a change in the system's resistance. This allows inference of the previous inputs applied to the system from the current states of the dynamical nodes, increasing MC. This architecture hence synergises well systems where activity decays slowly in the absence of large inputs.

Figure 5b shows the MCs calculated from the ring array system using the RNR approach. A peak MC of around 11.5 was found at H_c = 21 Oe and H_r = 10 Oe, showing that the device's non-volatile properties were being harnessed to provide much greater memory of past inputs than the other approaches, and thus extending applicability to problems with longer-term temporal dependencies. The region of maximum MC here is correlated to the central field at which DW motion starts to occur (Figure 1f(ii)). This corroborates the reasoning that the movement of DWs into different non-volatile configurations at fields above this value is where the system is 'storing' its memory of past states.

While MC can quantify the extent of linear memory (direct reconstruction of past inputs) in the system, real-world regression problems often require nonlinear memory (nonlinear representations of past inputs) for accurate prediction. To demonstrate the extent of nonlinear memory available to the system, we trained the system to reproduce a nonlinear auto-regressive moving average (NARMA-N) of input signals with varying degrees of autocorrelation (NARMA-5 and NARMA-10). For this problem, a system with perfect linear memory of equal degree to the autocorrelation (i.e., a shift-register of length N) can only achieve normalised means squared errors (NMSE) of around 0.4[34]. To improve upon this, a system needs to store nonlinear representations of past inputs. Figure 5c, 5d presents heatmaps of NMSE achieved over a range of field scaling parameters for NARMA-5 (5c), and NARMA-10 (5d), as well as examples of the reconstructed signals (5e, 5f). Regions where the ring array system outperforms the shift register in the NARMA-5 and NARMA-10 tasks are shown by the grey lines in Figure 5c, 5d, achieving peak NMSEs of 0.265 and 0.359 respectively. The combination of MC and performance of NARMA-N demonstrated that the system had been effectively reconfigured into a configuration with both linear and nonlinear memory without the aid of external delayed feedback lines that have typically been used in other demonstrations.

Figure 5- Performance of linear and nonlinear memory tasks.

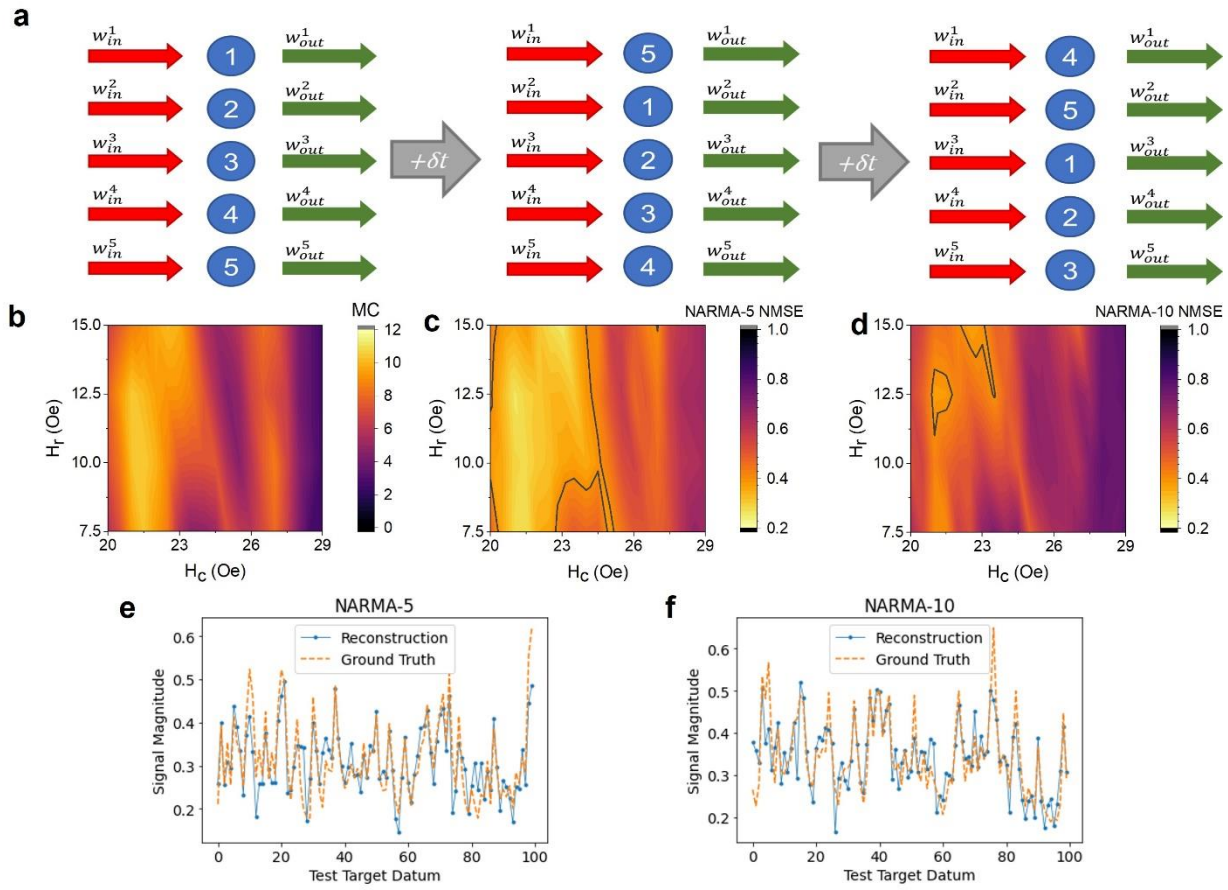


Figure 5a- Schematic diagram for simplified revolver setup consisting of three nodes, showing how input (red arrows) and output (green arrows) change with each timestep τ with respect to fixed dynamical nodes (blue circles). 5b- Memory Capacity (MC) over a range of field scaling parameters under the rotating neurons reservoir construction. 5c/d - Performance heatmaps for the (c) nonlinear autoregressive moving average (NARMA) -5 and (d) NARMA-10 system approximation task. Regions inside the grey line show configurations outperforming the score of a shift register with equal degree to the NARMA problem. 5e/f – NARMA signal reconstruction for optimally performing ring array reservoirs (blue, Normalised mean squared error (NMSE) = 0.265) compared to ground truth (orange, NMSE = 0.359) for (e) NARMA-5 and (f) NARMA-10.

Conclusion

In this paper we have demonstrated how a range of different RC architectures allow exploitation of different underlying dynamic properties in a complex magnetic system. This reconfigurability allowed the platform to achieve state-of-the art performance in three diverse tasks with differing computational requirements. To summarise the key correlations between underlying dynamics and suitable reservoir architectures, we found that the signal subsampling architecture synergises with phase transitions in the system's response to provide nonlinear mappings of input, the single dynamical node paradigm synergises with transient responses to connect different input dimensions across time, and the rotating neurons reservoir scheme synergises well with regimes where reservoir state changes slowly with small/zero inputs, allowing information from past inputs to be sustained over time via the rotating input mask.

The synergy between these dynamic properties is also directly correlated to the type of task that the resulting reservoir is suitable for solving: the dimensionality expansion and nonlinear dynamics provided by the signal subsampling architecture allows for effective 1D signal processing, the temporal mixing of input dimensions in the single dynamical node architecture enables classification on multivariate data, and the slow dynamics modulated by the rotating input mask in the rotating neurons reservoir architecture allows for effective performance in memory based tasks. Aside from the architecture choice, the selection of suitable scaling parameters for the input data is also critical to performance. To address this, we used task-independent metrics to provide a more holistic mapping of the computational properties of the reservoir across a range of scaling parameters and demonstrated the additional performance attainable via selecting promising parameters from the resulting metric maps for both classification-based tasks (KR/GR) and memory-based tasks (MC), with additional comparisons between each of the architectures' scores in these metrics.

We believe that the range of dynamical regimes offered by the system, combined with the ability to address each of these properties separately and extract distinct computational properties via controlling the external reservoir architecture, makes the ring system a candidate for reservoir computing with complex dynamic substrates. Additionally, the effectiveness of synergising the reservoir architecture with the dynamic properties of the underlying system makes for an effective methodology for extracting a broad range of computational capability for other similar devices. The ring devices are not without their limitations however, with the current device being driven external rotating magnetic fields, which provides both a limitation on the throughput on data input to the system (on the order of 100s of Hz), and power wastage in generating the magnetic fields over areas orders of magnitude larger than the nanoring array itself. Additionally, the current electrical readout provides a single scalar readout on the entire system state at a given point in time, which is sub-optimal for extracting complex state information on a system which exhibits spatially distributed responses like the ring system here. The feasibility of the ring system as a complex RC device that would be applicable to real-world settings hinges upon the ability to respond to electrical inputs such as spin-orbit torque driven DW motion, as well as expanding upon the readout mechanism to provide spatially resolved measurement of magnetic state.

To expand the computational capabilities of the ring arrays, the complex behaviours outlined here should operate concurrently as part of a larger system. The changes in magnetic responses offered via geometric changes to the system could enable multiple devices to operate in different regimes of dynamics and emphasise different computational properties under a single input field. Other magnetic metamaterial platforms have been shown to be useful in 'deep' reservoir networks with distributed reservoir properties[32], which the ring system would also likely benefit from. We believe that this work marks a significant step forward towards the realisation of metamaterial systems as computational platforms that are device-compatible, and that the rich playground of computationally useful dynamics they offer makes the ring system a promising candidate for physics-based neuromorphic computation platforms.

Methods

Device Fabrication

The ring array devices were fabricated using two-stage electron-beam lithography, with layouts patterned using a RAITH Voyager system. Wafers of Si (001) with a thermally oxidized surface were spin-coated with a positive resist. The ring structures were metallized to thicknesses of 10nm via thermal evaporation of $Ni_{80}Fe_{20}$ powder using a custom-built (Wordentec Ltd) evaporator with typical base pressures of below 10^{-7} mBar. The initial resist went through lift-off, leaving the ring structures before re-application of the resist and further electron-beam lithography. Electrical contacts were metallized in two stages of thermal evaporation, first with 20nm titanium to form a seed layer, before

growth of 200nm of gold. Electrical connections were provided between the device and a chip carrier through bonding of gold wire between contact pads on the device and the chip carrier.

Electrical Transport Measurements of Ring Arrays

Currents of 1.4mA were provided to the arrays as a 43117 Hz sine wave into the patterned contacts (Figure 1a) on the device using a Keithley 6221 current source. Resistance changes via AMR effects were measured using a Stanford Research SR830 lock-in amplifier. A National Instruments NI DAQ card was used to log the output voltage of the lock-in amplifier 64 times per rotation of applied field, and the data were then saved on a personal computer. The rotating magnetic fields were generated using two pairs of custom-built air-coil electromagnets in a pseudo-Helmholtz arrangement. The electromagnets were driven by a pair of Kepco BOP 36-6D power supplies and were controlled via voltage signals generated using a computer and the analogue output functionality of the NI card. A rotating field frequency of 37 Hz was chosen as a compromise between data throughput and signal fidelity.

Reservoir Computing

In RC, the fixed reservoir layer provides a transformation of discrete-time input signals $u(t)$, to reservoir states, $x(t)$, according to the internal dynamics of the reservoir layer. The readout layer (here, a single-layer linear perceptron) provides a weighted sum of the reservoir states as output, $y(t)$. The transformation provided by the reservoir layer results in a higher-dimensional mapping of the input signals. This aids the output layer in classifying the input signals by allowing selection of hyperplanes in higher-dimensional space to correctly classify data that was previously linearly inseparable.

In this work, the RNN that constitutes the reservoir layer of the typical echo state network (ESN) was replaced with the magnetic nanoring device. The reservoir transformation was provided by the physical processes that govern the array's magnetic response to field, as well as the changes to electrical resistance that consequently occur. Methods for inputting and extracting data are outlined for each reservoir configuration:

Signal Subsample Reservoir

Input sequences u_τ are transformed to give an applied field sequence via a pair of scalar parameters H_c and H_r, shown in the following equation, which represent the zero-input field offset and the field scaling factor respectively:

$$H_{input} = H_c + H_r * u_\tau \quad (7.1)$$

Each input was applied for a single rotation of magnetic field. The reservoir states were then extracted by sampling the lock-in voltage signal 32 times per rotation, producing a 32-node output.

Single Dynamical Node Reservoir

This approach uses 'virtual' nodes[34], where the reservoir states are generated from observing the state of the nanoring array as it evolves under time-multiplexed input. The generation of the time-multiplexed sequence of applied field magnitudes, H_{input}, (a vector of length $\theta * \tau$, where θ represents the desired number of virtual nodes, and τ the number of discrete-time windows the initial input sequence contains) was accomplished by combining the d-dimensional input vector for each timestep in $u_{\tau d}$ with a fixed input mask matrix, $M_{d,\theta}$, and flattened into a 1D sequence by concatenating timestep-by-timestep via:

$$H_{input} = H_c + H_r \sum_{k=1}^{\tau} u_{k,d} * M_{d,\theta} \quad (7.2)$$

where $M_{d,\theta}$ consisted of randomly generated 0's and 1's. The field sequence was then input to the system by rotating the field at magnitudes specified by H_{input} for a given number of quarter-rotations per input datum.

The resulting voltage signals provided by the lock-in amplifier underwent some simple processing steps: Firstly, a high-pass filter with a low cut-off frequency of 3Hz was used to centre the signals about zero and remove any low-frequency noise in the system. Band-pass filters were used to extract the 1f and 2f components separately. The pass-windows for each of these filters were centred about the input frequency and twice the input frequency, with band widths of 25% of the centre frequency to capture the damped dynamics of the oscillatory system. The outputs of the high-pass, and each of the band-pass filters, were sampled twice per input, forming a complete reservoir state vector six times the length of H_{input}.

Rotating Neurons Reservoir

This technique employs a shifting input/output mask[38], functionally analogous to rotating the input and output weights synchronously while keeping the dynamical neurons fixed. The procedure for this 'rotation' can be described as follows: Consider a system of θ dynamical nodes η^i, where i denotes the index of each node. An input signal $u_{\tau,d}$ is combined with mask $M_{d,\theta}$, to produce input dimensions $s_{\tau\theta}$. The input to node η^i at timestep t, $\tilde{s}_{t,i}$, is given via

$$\tilde{s}_{t,i} = s_{t,(i+t)\%\theta} \quad (7.3)$$

where '%' represents the modulo operation. The resulting output matrix, $\tilde{X}_{\tau\theta}$, is generated by vertically concatenating the output of vectors all θ nodes as they evolve, and is 'unraveled' similarly to form reservoir state matrix X via:

$$X_{t,i} = \tilde{X}_{t,(i-t)\%\theta} \quad (7.4)$$

Additional information on each of the machine learning tasks, details of training methods employed, and any data processing steps taken can be found in Supplementary Methods - Machine Learning Tasks.

Data Availability

A repository containing all data used in this study can be found on the University of Sheffield's online database, ORDA, DOI: 10.15131/shef.data.23815434

Code Availability

A repository containing all code used in this study can be found on the University of Sheffield's online database, ORDA, DOI: 10.15131/shef.data.23815434

Author Contributions

ITV performed all experiments, analysis, and modelling, and drafted the article. ITV and CS designed and optimised the AMR measurement apparatus. ITV implemented and performed all machine learning frameworks excluding the SpaRCe algorithm, implemented by LM. GV, CS, and PWF designed and manufactured all devices. ITV, AW, RMRR, DAA, and TJH performed X-PEEM measurements, which were overseen by DB, FM, and SSD. EV, DAA, and TJH conceptualised and supervised the work. All authors reviewed the manuscript.

Acknowledgements:

The authors would like to thank Jack C. Gartside for helpful discussions. The authors thank STFC for beam time on beamline I06 at the Diamond Light Source, and thank Jordi Prat, Michael Foerster and

Lucia Aballe from ALBA for providing quadrupole sample holders[47]. I.T.V. acknowledges a DTA-funded PhD studentship from EPSRC. The authors gratefully acknowledge the support of EPSRC through grants EP/S009647/1, EP/V006339/1, and EP/V006029/1. This project has received funding from the European Union's Horizon 2020 FETOpen program under grant agreement No 861618 (SpinEngine). This work was supported by the Leverhulme Trust (RPG-2018-324, RPG-2019-097).

Competing Interests
The authors have no competing interests to declare.

References:

1. Zou, X., Xu, S., Chen, X., Liang, Y. & Han, Y. Breaking the von Neumann bottleneck: architecture-level processing-in-memory technology. *Sci. China Inf. Sci.* **64**, (2021).
2. Jaeger, H. *The "echo state" approach to analysing and training recurrent neural networks- with an erratum note*. *GMD Technical Report* (2001).
3. Lukoševičius, M., Jaeger, H. & Schrauwen, B. Reservoir Computing Trends. *KI - Kunstliche Intell.* **26**, 365–371 (2012).
4. Lukoševičius, M. & Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **3**, 127–149 (2009).
5. Paquot, Y. *et al.* Optoelectronic reservoir computing. *Sci. Rep.* **2**, (2012).
6. Jacobson, P. L., Shirao, M., Yu, K., Su, G. L. & Wu, M. C. Hybrid Convolutional Optoelectronic Reservoir Computing for Image Recognition. *J. Light. Technol.* (2021) doi:10.1109/JLT.2021.3124520.
7. Sande, G. V. der, Brunner, D. & Soriano, M. C. Advances in photonic reservoir computing. *Nanophotonics* **6**, 561–576 (2017).
8. Yahiro, W., Aubert-Kato, N. & Hagiya, M. A reservoir computing approach for molecular computing. *Artif. Life* (2018).
9. Dion, G., Mejaouri, S. & Sylvestre, J. Reservoir computing with a single delay-coupled non-linear mechanical oscillator. *J. Appl. Phys.* **124**, 152132 (2018).

10. Coulombe, J. C., York, M. C. A. & Sylvestre, J. Computing with networks of nonlinear mechanical oscillators. *PLOS ONE* **12**, e0178663 (2017).
11. Dion, G., Oudrhiri, A. I.-E., Barazani, B., Tessier-Poirier, A. & Sylvestre, J. Reservoir Computing in MEMS BT - Reservoir Computing: Theory, Physical Implementations, and Applications. in (eds. Nakajima, K. & Fischer, I.) 191–217 (Springer Singapore, 2021). doi:10.1007/978-981-13-1687-6_9.
12. Tsakalos, K. A., Sirakoulis, G. C., Adamatzky, A. & Smith, J. Protein Structured Reservoir computing for Spike-based Pattern Recognition. *IEEE Trans. Parallel Distrib. Syst.* (2021) doi:10.1109/TPDS.2021.3068826.
13. Liu, X. & Parhi, K. K. Reservoir Computing Using DNA Oscillators. *ACS Synth. Biol.* **11**, 780–787 (2022).
14. Kulkarni, M. S. & Teuscher, C. Memristor-based reservoir computing. *Proc. 2012 IEEEACM Int. Symp. Nanoscale Archit. NANOARCH 2012* 226–232 (2012) doi:10.1145/2765491.2765531.
15. Hassan, A. M., Li, H. H. & Chen, Y. Hardware implementation of echo state networks using memristor double crossbar arrays. in *2017 International Joint Conference on Neural Networks (IJCNN)* 2171–2177 (2017). doi:10.1109/IJCNN.2017.7966118.
16. Mehonic, A. *et al.* Memristors—From In-Memory Computing, Deep Learning Acceleration, and Spiking Neural Networks to the Future of Neuromorphic and Bio-Inspired Computing. *Adv. Intell. Syst.* **2**, 2000085 (2020).
17. Torrejon, J. *et al.* Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017).
18. Nakane, R., Tanaka, G. & Hirose, A. Reservoir Computing with Spin Waves Excited in a Garnet Film. *IEEE Access* **6**, 4462–4469 (2018).

19. Jensen, J. H., Folven, E. & Tufte, G. Computation in artificial spin ice. in *ALIFE 2018 - 2018 Conference on Artificial Life: Beyond AI* 15–22 (MIT Press - Journals, 2020). doi:10.1162/isal_a_00011.
20. Gartside, J. C. *et al.* Reconfigurable training and reservoir computing in an artificial spin-vortex ice via spin-wave fingerprinting. *Nat. Nanotechnol.* **17**, 460–469 (2022).
21. Welbourne, A. *et al.* Voltage-controlled superparamagnetic ensembles for low-power reservoir computing. *Appl. Phys. Lett.* **118**, 202402 (2021).
22. Ababei, R. V. *et al.* Neuromorphic computation with a single magnetic domain wall. *Sci. Rep. 2021 111* **11**, 1–13 (2021).
23. Allwood, D. A. *et al.* A perspective on physical reservoir computing with nanomagnetic devices. *Appl. Phys. Lett.* **122**, 040501 (2023).
24. Riou, M. *et al.* Temporal pattern recognition with delayed-feedback spin-torque nano-oscillators. *Phys. Rev. Appl.* **12**, (2019).
25. Abreu Araujo, F. *et al.* Role of non-linear data processing on speech recognition task in the framework of reservoir computing. *Sci. Rep.* **10**, 1–11 (2020).
26. Leroux, N. *et al.* Hardware realization of the multiply and accumulate operation on radio-frequency signals with magnetic tunnel junctions. *Neuromorphic Comput. Eng.* **1**, 011001 (2021).
27. Ross, A. *et al.* Multilayer spintronic neural networks with radio-frequency connections. *Nature Nanotechnology* 1-8 (2023).
28. Kanao, T. *et al.* Reservoir Computing on Spin-Torque Oscillator Array. *Phys. Rev. Appl.* **12**, 024052 (2019).
29. Nomura, H. *et al.* Reservoir computing with two-bit input task using dipole-coupled nanomagnet array. *Jpn. J. Appl. Phys.* **59**, SEEG02 (2019).

30. Williame, J., Difini Accioly, A., Rontani, D., Sciamanna, M. & Kim, J.-V. Chaotic dynamics in a macrospin spin-torque nano-oscillator with delayed feedback. *Appl. Phys. Lett.* **114**, 232405 (2019).
31. Taniguchi, T. *et al.* Chaos in nanomagnet via feedback current. *Phys. Rev. B* **100**, 174425 (2019).
32. Stenning, K. D. *et al.* Adaptive Programmable Networks for In Materia Neuromorphic Computing. Preprint at https://doi.org/10.48550/arXiv.2211.06373 (2022).
33. Pinna, D., Bourianoff, G. & Everschor-Sitte, K. Reservoir Computing with Random Skyrmion Textures. *Phys. Rev. Appl.* **14**, 054020 (2020).
34. Appeltant, L. *et al.* Information processing using a single dynamical node as complex system. *Nat. Commun.* **2**, (2011).
35. Dawidek, R. W. *et al.* Dynamically-Driven Emergence in a Nanomagnetic System. *Adv. Funct. Mater.* **31**, (2021).
36. Vidamour, I. *et al.* Quantifying the Computational Capability of a Nanomagnetic Reservoir Computing Platform with Emergent Magnetisation Dynamics. *Nanotechnology* (2022) doi:10.1088/1361-6528/ac87b5.
37. Nagura, H., Saito, K., Takanashi, K. & Fujimori, H. Influence of third elements on the anisotropic magnetoresistance in permalloy films. *J. Magn. Magn. Mater.* **212**, 53–58 (2000).
38. Liang, X. *et al.* Rotating neurons for all-analog implementation of cyclic reservoir computing. *Nat. Commun.* **13**, 1549 (2022).
39. Daniels, R. K. *et al.* Reservoir computing with 3D nanowire networks. *Neural Netw.* **154**, 122–130 (2022).

40. Fu, K. *et al.* Reservoir Computing with Neuromemristive Nanowire Networks. in *2020 International Joint Conference on Neural Networks (IJCNN)* 1–8 (2020). doi:10.1109/IJCNN48605.2020.9207727.
41. Dale, M., Miller, J. F., Stepney, S. & Trefzer, M. A. *Evolving carbon nanotube reservoir computers*. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* vol. 9726 49–61 (2016).
42. Molau, S., Pitz, M., Schlüter, R. & Ney, H. Computing mel-frequency cepstral coefficients on the power spectrum. in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* vol. 1 73–76 (2001).
43. Dale, M., Miller, J. F., Stepney, S. & Trefzer, M. A. A substrate-independent framework to characterize reservoir computers. *Proc. R. Soc. Math. Phys. Eng. Sci.* **475**, (2019).
44. Manneschi, L., Lin, A. C. & Vasilaki, E. SpaRCe: Improved Learning of Reservoir Computing Systems Through Sparse Representations. *IEEE Trans. Neural Netw. Learn. Syst.* 1–15 (2021) doi:10.1109/TNNLS.2021.3102378.
45. Liu, B. Lifelong machine learning: a paradigm for continuous learning. *Front. Comput. Sci.* **11**, 359–361 (2017).
46. Jaeger, H. *Short term memory in echo state networks*. *GMD Report 152* 60 (2002).
47. Foerster, M. *et al.* Custom sample environments at the ALBA XPEEM. *Ultramicroscopy* **171**, 63–69 (2016).

Supplementary Material – Reconfigurable Reservoir computing with a magnetic metamaterial

Vidamour, I. T.[1,2*]; Swindells, C.[1]; Venkat, G.[1]; Manneschi, L.[2]; Fry, P. W.[3]; Welbourne, A.[1]; Rowan-Robinson, R. M.[1]; Backes, D[4,]; Maccherozzi, F[4,]; Dhesi, S. S.[4]; Vasilaki, E.[2]; Allwood, D. A.[1]; and Hayward, T. J.[1]

1- Department of Materials Science and Engineering, University of Sheffield, Sheffield, UK
2- Department of Computer Science, University of Sheffield, Sheffield, UK
3- Nanoscience and Technology Centre, University of Sheffield, UK
4- Diamond Light Source, Harwell Science and Innovation Campus, Didcot, UK
*- Corresponding author, contactable at i.vidamour@sheffield.ac.uk

Supplementary Note 1- Fourier Analysis of AMR Signals

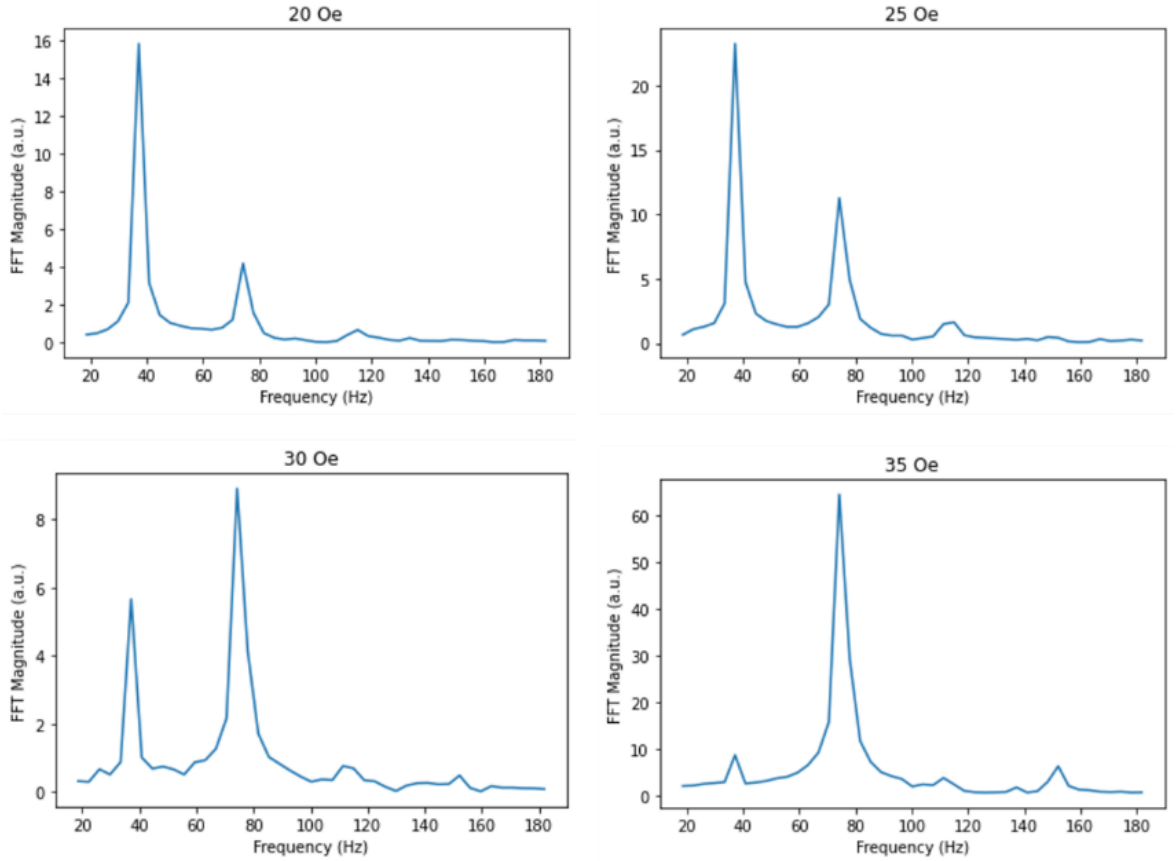


Figure S1- Fast-Fourier transformation of AMR signals over 20 rotations of applied rotating fields of magnitude (top left to bottom right) 20, 25, 30, and 35 Oe at 37Hz, with peaks at 37 Hz (1f) and 74 Hz (2f).

Supplementary Note 2- Anisotropic magnetoresistance response of ring arrays.

The anisotropic magnetoresistance (AMR) effects observed in Py provide a simple measure of state space via resistance changes over time. In the ring system, shape anisotropy forces magnetic moments to follow the track of the nanorings, resulting in a small angle between current density and magnetic moments inside of domains. However, DWs in the system

provide a discontinuity of local magnetization direction, where moments are aligned perpendicular to the flow of current, meaning reduced resistance is observed due to the presence of DWs. Hence, the measured resistance across the array depends upon the number, size, and positions of DWs in the system.

Figure S1 shows the Fourier components of the AMR response for various applied fields. Major peaks can be observed at the clock frequency (1f) and twice the clock frequency (2f). Figure S2A shows the 1f and 2f Fourier components of the AMR response when saturated and then driven by successive rotations of different magnetic field strengths, while Figure S2B shows the number of rotations required for the system to reach dynamic equilibrium from saturation over a range of applied fields. Using the correlation between DW pinning/propagation and the 1f and 2f Fourier components (Figure S2A), we can infer DW dynamics in different regimes of response (Figure S2C-F). At low driving fields (20 Oe, Figure S2C), there is a 1f signal due to elastic stretching and contracting of pinned DWs, with no irreversible change to the magnetic state. At 25 Oe (Figure S2D), DWs propagation occurs stochastically, leading to a superposition of the frequency components, dominated by the 1f signal. The onset of vortex formation causes oscillation magnitude to diminish as the number of DWs tends towards equilibrium levels. At fields of 30 Oe (Figure S2E), DWs propagation increases, leading to a 2f dominated signal, again decreasing over time due to vortex formation. At higher driving fields (35 Oe, Figure S2F), all DWs in the system coherently follow the rotating field, with no transient behaviours as no vortices are formed. The transitions between these regimes can be corroborated by the state population vs applied field data shown in Figure 1D. The transient behaviours rely on vortex formation, with no transience in the purely pinned/propagating regimes (Figure S2B).

Micromagnetic simulations of permalloy nanorings were performed using MuMax3[1] with material parameters: saturation magnetization = 860 kA/m and exchange stiffness = 13 pJ/m. An artificially large damping parameter of 1 was used to reduce simulation times. provide insight into how the magnetization (and hence resistance) state of a simplified ring geometry evolves under the propagating DW (Figure 2G), and the pinned DW (Figure S2H) regimes. Figure S2G shows different angles between current flow (white arrows) and magnetic moments (black arrows) depending on the junction location, leading to high and low resistance states. When DWs coherently follow the rotating field, they propagate through the four junctions, leading to a pair of resistance minima/maxima, generating a resistance signal of twice the clock frequency. When the DWs remain pinned (Figure 2H), the applied field causes deformation of the DW and the surrounding domains, with expansion/contraction when the applied field lies parallel/antiparallel to the direction of the magnetization of the DW respectively. This leads to oscillations matching the frequency of the rotating field for pinned DWs.

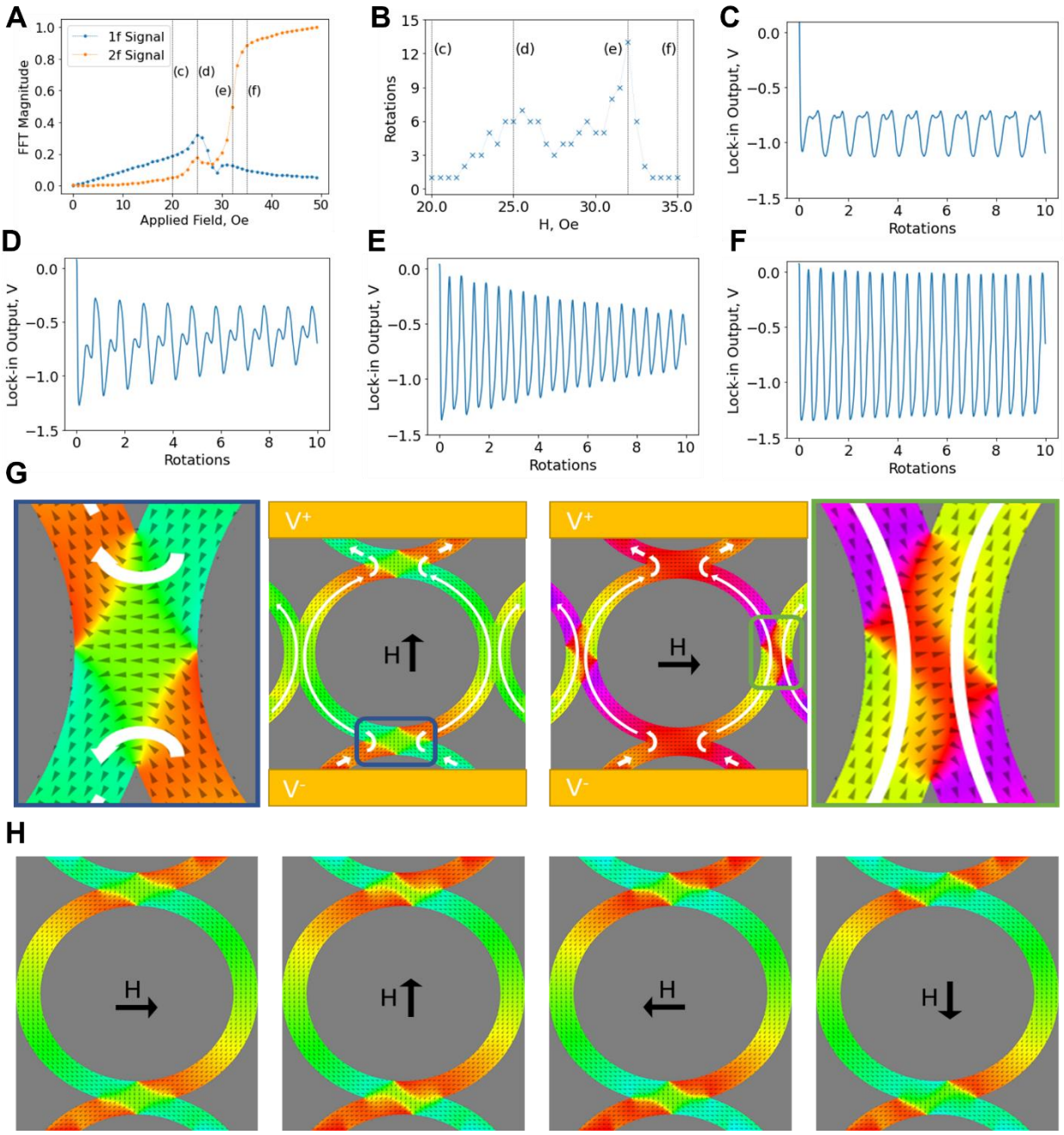


Figure S2A- Fourier components of systems driven with 10 rotations of magnetic field at various applied fields. Blue datapoints show Fourier component at the driving field frequency (1f), while orange datapoints show component at double the driving field frequency (2f). S2B- Number of rotations of field required for the system to reach an equilibrium response (<2% peak-to-peak deviation between cycles) from saturation for a range of applied fields. S2C-F - Measured lock-in voltage of the array when driven with 10 field rotations of amplitude (C) 20 Oe, (D) 25 Oe, (E) 32 Oe, and (F) 35 Oe from saturation. S2G- Mumax3 simulations[1] showing angle between magnetization direction (black triangles) compared to direction of current flow (white arrows) for systems initialized with a pulse of magnetization along the flow of current direction (left) and perpendicular to the current direction (right). S2H- Mumax3 simulations showing the deformation of DWs under the influence of external fields of 25 Oe in the direction shown by the black arrows.

Supplementary Note 3- Metric Space Assessment

The specific transformation provided by the nanoring system is dependent upon the scaling and input rate of data. Consequently, this affects computational capability as each transformation will have varying degrees of nonlinearity and different timescales of dependence on past input. Previous work has shown that these changes in capability can be quantified in task-independent metric space, and metric scores were a good predictor of task performance[2]. These metrics provide a rough evaluation of reservoir performance along contrasting computational axes: KR measures the reservoir's ability to represent distinct input patterns in different reservoir states, while GR measures ability to generalize similar input patterns to similar reservoir states. A simple heuristic of computational quality, CQ, can be constructed by subtracting GR from KR, with 'ideal' configurations having a larger CQ due to high KR (good separation properties) and low GR (good generalization properties)[9]. By changing the duration each input is applied for, the degree of 'mixing' across virtual nodes can be controlled, with shorter input durations leading to greater mixing and reduced sensitivity to a single input. Figure S3 shows calculated metrics for a range of field scaling parameters H_c and H_r.

Since the spoken digit recognition task requires good separation properties from the reservoir, the 'promising' parameters (red cross) were chosen for their high KR. The 'arbitrary' parameters (blue star) were chosen for their similar CQ properties, though low KR compared to the 'promising' parameters. As well as the single dynamical node reservoir, the metric assessment with respect to input parameters was also performed on the signal sub-sample and rotating neurons reservoir, which showed good separation and generalisation capability respectively, but poor separation between the two resulting in a low CQ for these reservoir configurations.

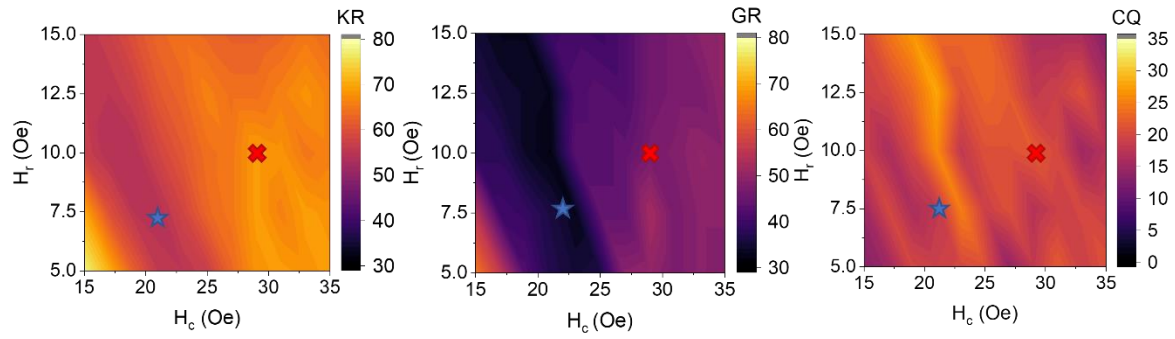


Figure S3- Evaluated KR (a), GR (b), and CQ (c) metric scores for a range of field scaling parameters. 'Promising' and 'Arbitrary' field scaling parameters are indicated by red crosses and blue stars respectively.

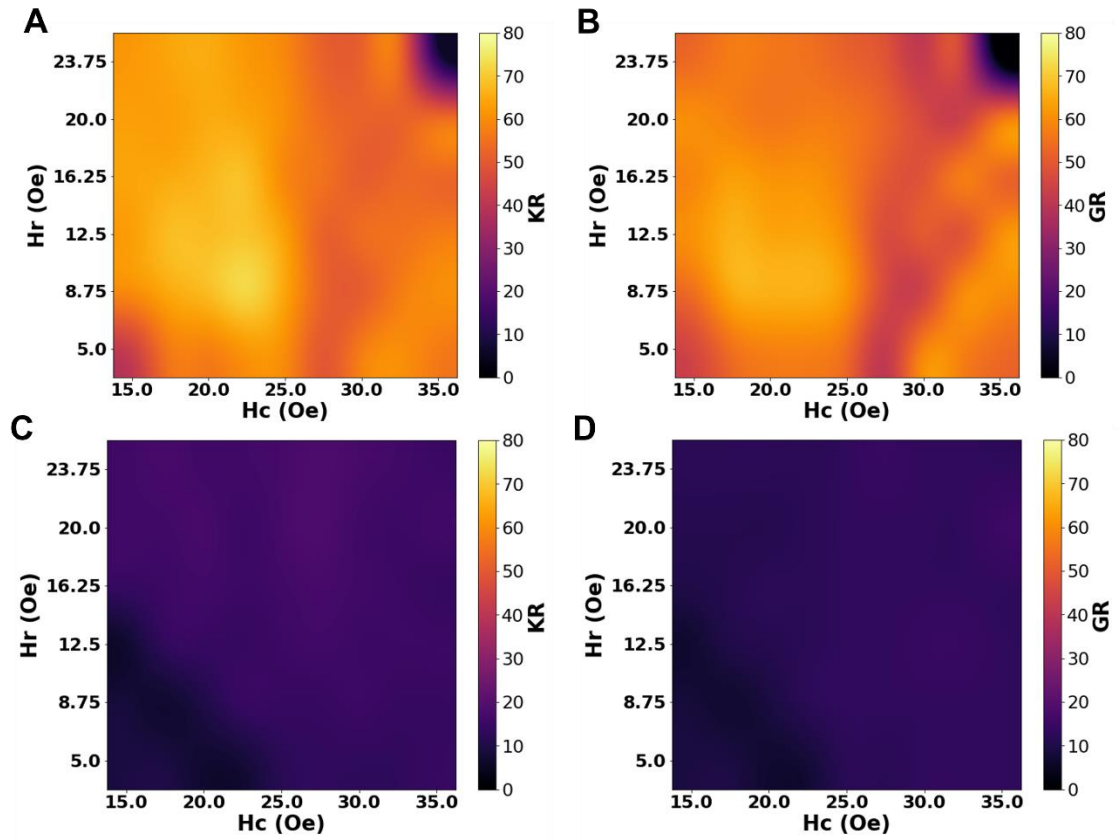


Figure S4- Metric heatmaps over a range of field scalings, for the revolving neurons reservoir (KR- A, GR- B) and the signal sub-sample reservoir (KR- C, GR- B) architectures.

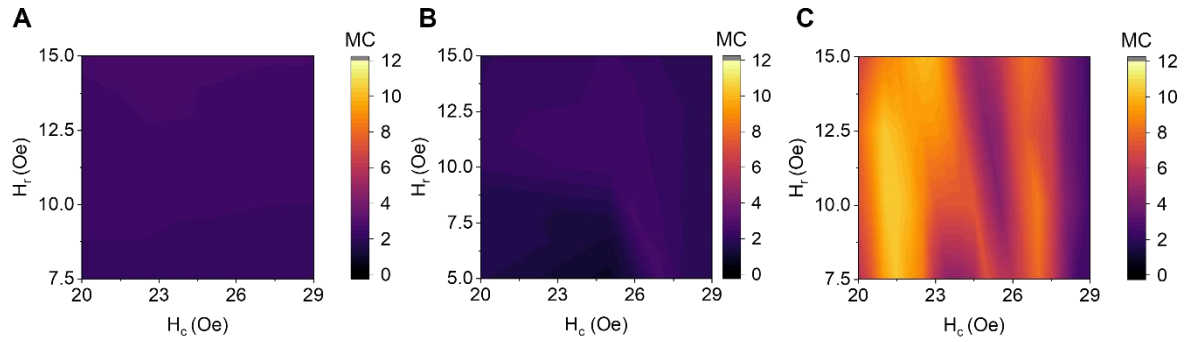


Figure S5- Comparison of linear memory capacity between the three reservoir configurations, (A) signal sub-sample reservoir, (B) Single dynamical node reservoir, and (C) rotating neurons reservoir.

Supplementary Note 4- Calculation of Saturation Magnetisation

Saturation magnetisation was calculated from broadband ferromagnetic resonance data of the fundamental mode. From best fits of the peak absorption as a function of external magnetic field and applied microwave frequency to the Kittel curve, $\mu_0 M_s$ was determined to be 0.969 ± 0.006 T, representing slightly Nickel-rich Py.

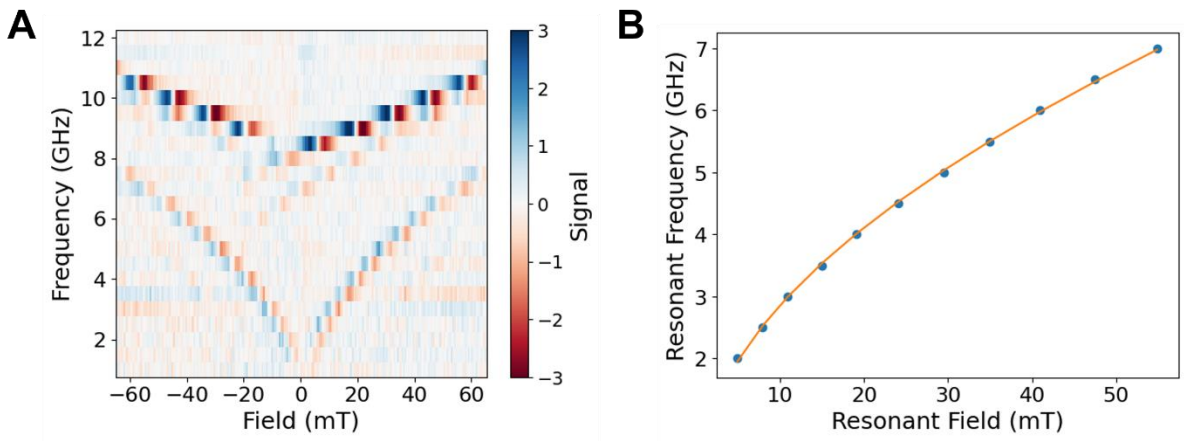


Figure S6- (A)- Peak microwave absorption of nanoring devices across various applied frequencies and fields. (B)- Fit of fundamental mode, giving a $\mu_0 M_s$ of 0.969 ± 0.006 T.

Supplementary Note 5- Supplementary Methods - Machine Learning Tasks

Offline Training

The weights associated with the readout layer, W^{out}, for the reservoir configurations described here were trained offline using an ordinary-least-squares method with Tikhonov regularization, commonly referred to as 'ridge regression'[3,4]. The process is described by equation:

$$\text{S.E. (7.1.1)} \quad W^{out} = YX^T * (XX^T + \lambda I)^{\dagger}$$

where X represents a matrix created from the reservoir states produced from the training data, Y represents the target labels associated with the training data, λ represents the regularization parameter, I the identity matrix, and † the Moore-Penrose pseudo-inverse operation.

Online Training

The output weights W^{out} were optimised via gradient descent to minimize the sigmoid cross entropy loss function for both the standard readout and as well as the SpaRCe algorithm, given by

$$\text{S.E. (7.1.2)} \quad E = -\left\{ \sum_j \tilde{y}_j \, log \, \left(\sigma(y_j)\right) + (1 - \tilde{y}_j) log \, \left(1 - \sigma(y_j)\right) \right\}$$

where $\tilde{y}_j$ is the desired output for class j.

SpaRCe

While simple linear regression techniques like ridge regression can be used to analytically find the weights that provide the least-squares difference between training data and target labels, more advanced learning rules are able to converge to different sets of output weights which provide a more robust classification platform. One such learning rule that has been demonstrated to be effective for RC is SpaRCe[5] - a bio-inspired algorithm that utilizes learnable thresholds on output neurons to introduce sparse representations of data and promote more specialized features. SpaRCe has been shown to improve classification

accuracy in common benchmark tasks, as well as alleviating the problem of catastrophic forgetting associated with online learning of sequential tasks/classes[6,7]. It is an on-line algorithm, which is likely more suited to real-world implementations of RC where the system learns continuously as new data is presented, rather than through storing many input patterns before changing parameters.

The output weights used for the SpaRCe algorithm were also trained in an online manner using the sigmoid cross-entropy loss function outlined earlier. Additionally, a learnable threshold, $\hat{\theta}_k$, was applied to each output neuron k. The output for each thresholded node, $\hat{x}_k$, is given from its activity, x_k, via

$$\text{S.E. (7.1.3)} \quad \hat{x}_k = sign(x_k)\, ReLU\{|x_k| - \theta_k\}$$

$$\text{S.E. (7.1.4)} \quad \theta_k = P_n(|x_k|) + \hat{\theta}_k$$

where ReLU represents the rectified linear unit, and $P_n(|x_k|)$ is defined as the percentile n of the activity of neuron k across the inputs, computed across the training data. The choice of n provides an initial sparsity level for the thresholds. $P_n(|x_k|)$ remains constant throughout training since node output x_k with respect to input is constant for RC.

It is possible to show how adaptation of the thresholds θ_k through gradient descent is characterized by two terms: the first is decreasing the thresholds of neurons that are correctly contributing to the classification process; the second is increasing the threshold of nodes that are correlated and thus carry redundant information. We refer to the original publication[4] for more details. Hyperparameters of starting percentile and learning rate were selected by choosing the best performance on the training data set, with reported accuracies given by 10-fold cross validation.

Signal Transformation Task

The raw input signals for the signal transformation task were generated from equation $s(t) = sin\, sin\left(\frac{\pi}{20}t\right)$ where t represents the discrete timestep associated with each input. Target signals were generated using similar time discretization. For each discrete point, a single rotation of magnetic field was applied, and 32 samples were taken as reservoir states from the lock-in voltage signal. A control measurement was taken by similarly measuring the voltage across one of the electromagnets to exclude the nonlinear transformation provided by the nanoring arrays. Both the nanoring array reservoir and the control 'reservoir' readouts were trained using ridge regression. The system was saturated and then driven for 51 full sine waves, with a range of field scaling parameters H_c and H_r values explored, with the first sine wave discarded to wash-out the dependence of the reservoir on the initial saturated state. Training states consisted of the measurements produced from the first 45 sine waves, and the associated training labels were defined by the target signal being reproduced. Reported test accuracies were generated by evaluating the mean-squared error between the target signal and a signal produced by multiplying the trained readout weights and the measured states over the remaining five sine waves, with the accuracy for the highest performing H_c and H_r pair recorded.

Assessing Computational Capability of Arrays

KR and GR are calculated by assessing the number of linearly independent output sequences present at the output when driven with input signals of given characteristics. The raw input matrix for each metric consisted of 100 sequences of 10 randomly sampled floating points from a uniform distribution between ±1. For KR, all the sequences and points within them are completely uncorrelated, whilst for GR the final 3 floating points in every sequence were identical to those of the first sequence. These input sequences were then combined with a fixed mask consisting of 50 random floating points, again sampled from a uniform distribution between ±1. Hence, each point in the raw input sequence was linearly correlated to each other via the fixed mask.

The signals were input to the reservoirs for a range of H_c and H_r values, for input rates of half a rotation and one full rotation per input, saturating between sequences. The AMR signals for the final 50 masked inputs (corresponding to the final unmasked input) of each sequence were logged. The AMR signals were passed through a high-pass filter, before being normalized against the maximum value of the signal. Features were extracted from the signals by sampling four times per complete rotation, meaning 200 features for a full rotation per input, 100 features for half a rotation. Output matrices were then constructed from the extracted features for each of the 100 sequences. Singular value decomposition was performed on the output matrices, and the rank of the matrices estimated by counting the number of singular values above an arbitrary small threshold value of 0.1.

Spoken Digit Recognition Task

Input signals were created by combining filtered audio signals with a fixed mask according to the single dynamical node paradigm. Details of how filtered signals were generated from raw audio data can be found in the supplementary material. Before each input sequence was applied, the system was saturated with a strong (~150 Oe) pulse of field. The system was then driven with applied fields scaled about various H_c and H_r pairs. Each input was applied for either a half-rotation or full rotation of applied field. Features were extracted according to the method described in the single dynamical node paradigm. To avoid biasing towards certain speakers or digits in the training procedure, k 'training' samples were chosen for each digit for each speaker at random, with performance assessed for various values of k. From the selected training samples, reservoir state matrix X and target matrix Y were constructed by concatenating vertically, producing matrices of dimensions [300, L] and [10, L] respectively, where L represents the total number of timesteps τ across all the training samples. These matrices were then used to calculate output weights W^{out} via ridge regression as described earlier. To classify the remaining 'test' digits, the trained output weights were multiplied with the reservoir states for each timestep, producing an output activation for each node for every timestep. Classification was performed using a winner-takes-all approach by selecting the node with the highest average activation across all timesteps, with correct classification occurring if the node with the highest activation matched the digit being spoken. Quoted performances were generated using 100-fold cross validation by repeating the random training/testing sampling process 100 times, with the regularization hyperparameter λ selected according to highest average performance over the repetitions to avoid overfitting.

Linear Memory Capacity

To calculate the system's MC, an input consisting of 2500 random floating points uniformly distributed between 0 and 1 was used to drive the system for a range of H_c and H_r values, with inputs masked according to the RNR paradigm. The AMR signal over the duration of input was logged, and samples corresponding to the first 200 inputs were discarded as a 'wash out' of initial system conditions. Features were sampled from the AMR signal twice per node per input. The resulting 100x2300 feature matrix was split into training and testing data by separating at the data corresponding to the 1800th input. A target signal was generated for each input, consisting of the current input, plus the 50 inputs that came previously. The system was then trained to reproduce the delayed input signal from the 100 reservoir features. MC was estimated using the following equation:

$$\text{S.E. (7.1.5)} \quad MC = \sum_{k=1}^{50} MC_k = \sum_{k=1}^{50} \frac{cov^2(u_{i-k}, y_k)}{\sigma^2(u_i)\sigma^2(y_k)}$$

where u_i is the raw input sequence, u_{i-k} is the raw input sequence at delay k, and y_k is the reservoir's reconstruction of the input sequence at delay k. Small uncorrelated MC_k values below 0.2 were excluded from the sum to negate any random correlations between the reconstructed signal and the actual target.

NARMA Prediction

The objective of this task is to predict the output of a nonlinear autoregressive moving average y_i of randomly generated inputs u_i, drawn from a uniform distribution between 0 and 0.5, with correlations of different orders N. The target sequence is generated from the following iterative equation:

$$\text{S. E. (7.1.6)} \quad y_i = y_{i-1} * (0.3 + 0.05 * \sum_{k=1}^{N} y_{i-k}) + 1.5 * u_{i-1} * u_{i-N} + 0.1$$

These tasks were performed under the rotating neurons reservoir paradigm, with four measurements taken of the evolving AMR signal for each of the 50 dynamic neurons, producing a reservoir state vector of length 200 for each timestep. The system was driven with u_i of length 1500, with the first 100 timesteps discarded as a 'wash-out' of initial conditions. The system's output is calculated for each timestep by combining the reservoir state vector with weight vector W^{out}, trained with ridge regression over 1200 data points. Quoted accuracies were given by calculating the normalised mean squared error (NMSE) of the reconstructed signal $\tilde{y}_i$ over the remaining 200 unseen data points, where:

$$\text{S. E. (7.1.7)} \quad NMSE = \frac{\sum_i^{200} (|y_i - \tilde{y}_i|)^2}{200 * \sigma^2(y)}$$

Pre-Processing of Spoken Digit Recognition Data

The speech data was taken from a subsample of the TI-46 database[8], where each of five female speakers provided ten utterances of each digit. Audio features were extracted from the raw-audio waves using a Mel-Frequency Cepstral filter[9] which produced a discrete-time power response in 13 frequency bands across 50 ms windows. This produced an unmasked signal of length [13, τ], where τ is equal to the number of windows generated by a given utterance. These signals were time-multiplexed using the masking procedure described previously and a mask of size [13, 50], producing a 1D input vector of length $50 * \tau$. The number of virtual nodes was kept at 50 as a compromise between providing a thorough mix of input dimensions, while keeping the number of datapoints low. The process was repeated for each of the 500 raw input signals, which were then normalized against the maximum value from the resulting pre-processed dataset.

Supplementary References

1. Vansteenkiste, A. *et al.* The design and verification of MuMax3. *AIP Adv.* **4**, 107133 (2014).
2. Dale, M., Miller, J. F., Stepney, S. & Trefzer, M. A. A substrate-independent framework to characterize reservoir computers. *Proc. R. Soc. Math. Phys. Eng. Sci.* **475**, (2019).
3. Lukoševičius, M., Jaeger, H. & Schrauwen, B. Reservoir Computing Trends. *KI - Kunstliche Intell.* **26**, 365–371 (2012).
4. Hoerl, A. E. & Kennard, R. W. Ridge Regression: Applications to Nonorthogonal Problems. *Technometrics* **12**, 69–82 (1970).
5. Manneschi, L., Lin, A. C. & Vasilaki, E. SpaRCe: Improved Learning of Reservoir Computing Systems Through Sparse Representations. *IEEE Trans. Neural Netw. Learn. Syst.* 1–15 (2021) doi:10.1109/TNNLS.2021.3102378.
6. McCloskey, M. & Cohen, N. J. Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem. in *Psychology of Learning and Motivation* (ed. Bower, G. H.) vol. 24 109–165 (Academic Press, 1989).
7. Kirkpatrick, J. *et al.* Overcoming catastrophic forgetting in neural networks. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 3521–3526 (2017).
8. Liberman, M. *et al.* TI 46-Word Corpus. (1993).

9. Molau, S., Pitz, M., Schlüter, R. & Ney, H. Computing mel-frequency cepstral coefficients on the power spectrum. in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* vol. 1 73–76 (2001).

8- “Tuning reservoir computing performance - modifying physical device response and reservoir architecture.”

8.0- Preamble

With a working experimental demonstrator of the nanoring devices as a reservoir computing platform, we then focussed upon how we could expand the computational capabilities of the nanoring arrays by engineering the nanoring arrays to provide different responses. Some obvious choices for manipulating the behaviour of the arrays were to change the track widths of rings, the degree of overlap between neighbouring rings, and the lattice arrangement of the array. While work is ongoing both at the University of Sheffield and with external collaborators with explorations of the former two array manipulations, the final article of this thesis focusses on changes to the lattice arrangement.

This work was primarily a joint effort between a postdoc in the group, Guru Venkat, and I. The work covers the exploration of three different lattice arrangements; the standard square lattice that had been used thus far in the thesis, a trigonal arrangement where each ring attaches to six other rings, and a Kagome arrangement where each ring attaches to three other rings. The first part of the paper, led by Guru, focussed on exploring the physical response of the different lattice arrangements.

Firstly, the types of different microstates that were formed under different applied fields were explored via X-PEEM imaging of the different lattice arrangements. While it was obvious that the different arrangements would form different specific microstates depending upon the availability of more/fewer pinning locations than the square arrays, they each showed markedly different responses with respect to applied field. This seemed promising for extracting different behaviours from the arrays.

However, when the microscopic array response was measured via AMR, the different lattice arrangements appeared broadly similar. While the specific fields at which each array became dynamically active differed slightly, the three arrangements showed both similar shaped equilibrium responses as well as similar ranges of timescales of settling. It was clear that the global measurement technique as blurring together most of the differences that were observed at the microstate.

The second part of the study, led by me, focussed upon evaluating the computational differences between the three arrays, again using task independent metrics to measure the nonlinearity and memory available to the three arrangements. As the previous paper highlighted, the choice of reservoir architecture is critical when assessing the complete range of computational properties a system can exhibit, and hence each of the three lattices were explored under each of the three computational architectures outlined previously.

Whilst the choice of architecture appeared to be the primary factor leading to different computational properties, there were some differences observed between the different lattices. These differences were mainly in the specific fields at which the different lattice arrangements had optimal properties, due to the differences in when each array became dynamically active as observed earlier. The single dynamical node architecture also shown some significant differences, with the computational quality observed inversely correlated to each lattice’s tendency to form vortices. This is likely due to the fact domain walls play such a key role in the AMR signal.

Finally, the paper shows how despite the lattices having broadly similar peak metric scores, the difference in the regime of operation for each of the lattices at a given field is a functional property

that can be exploited to improve computational capability. This was achieved by combining the outputs of each lattice for the same input sequence, with the different nonlinearities provided by each lattice leading to improved computational quality.

This was an important finding as it shows promise for the scalability of the ring array devices- by expanding the computational capabilities by exploiting the different nonlinearities generated via geometrical manipulation of the arrays. Additionally, it also importantly highlighted that global readout methods like AMR tend to obfuscate the differences between each system, suggesting that additional computational power, or a more diverse range of computational properties, may be attained by having a microstate dependent readout. This conclusion is similar to the expansion in performance that was observed in chapter 6, where concatenating the different state readouts from RingSim led to largely improved performance.

In terms of personal contributions to the paper, I took part in the X-PEEM measurements of the different devices, performed some of the AMR measurements (the data was then analysed by Guru), and performed all implementations of the computational metrics. The paper was co-drafted by Guru and I.

Tuning reservoir computing performance modifying physical device response and reservoir architecture

G. Venkat[1], I. T. Vidamour[1,2], C. Swindells[1], P. W. Fry[3], M. C. Rosamond[4], M. Foerster[5], M. A. Nin˜o[5], A. Bischoff[6], R. Allenspach[6], D. Griffin[7], S. Stepney[7], D. A. Allwood[1], E. Vasilaki[2], T. J. Hayward[1]

1 Department of Materials Science and Engineering, University of Sheffield, S1 3JD,UK
2 Department of Computer Science, University of Sheffield, Sheffield S1 4DP, UK
3 Nanoscience and Technology Centre, University of Sheffield, Sheffield, S3 7HQ, UK
4 School of Electronic and Electrical Engineering, University of Leeds, Leeds, LS2 9JT, UK
5 ALBA Synchrotron Light Facility, 08290, Cerdanyola del Valles, Spain
6 IBM Research-Zurich, 8803 Rüschlikon, Switzerland
7 Department of Computer Science, University of York, YO10 5GH, UK

Abstract.

Physical reservoir computing (RC) is an attractive machine learning technique that is ideal for processing of time dependent data series. It is also uniquely well-aligned to in-materio computing realisations that allow the inherent memory and non-linear responses of functional materials to be directly exploited for computation. We have recently shown that square arrays of interconnected magnetic nanorings are attractive candidates for in-materio reservoir computing, and experimentally demonstrated their strong performance in a range of benchmark tasks. Here, we extend these studies to other lattice arrangements of rings, including trigonal and Kagome grids, to explore how these affect both the magnetic behaviours of the arrays, and their computational properties. We show that while lattice geometry substantially affects the microstate behaviour of the arrays, these differences manifest less profoundly when averaging magnetic behaviour across the arrays. Consequently the computational properties (as measured using task agnostic metrics) of devices with a single electrical readout are found to be only subtly different, with the approach used to time-multiplex data into and out of the arrays having a stronger effect on properties than the lattice geometry. However, we also find that hybrid reservoirs that combine the outputs from arrays with different lattice geometries show enhanced computational properties compared to any single array.

Introduction.

Neuromorphic computing is being actively pursued for artificial intelligence and machine learning applications and investment in it is projected to grow significantly in the coming decade [1]. However, the cost associated with training large neural networks for such applications can be significant and this has led to investigation into unconventional computing approaches with lower energy footprints.

Reservoir computing [2–4] (RC) is a computing paradigm which uses the dynamics of a recurrent neural network (RNN) or another dynamical system (algorithmic or physical), often referred to as the reservoir, to transform input data to a higher dimensional space where it may be classified more easily. It has attracted interest in recent years mainly because the internal weights of the RNN are fixed and hence do not need to be trained, creating substantial energy savings when compared to a conventional RNN [5]. Significantly, the ‘black-box’ nature of the reservoir means that the RNN can be replaced any dynamical system that has (a) a non-linear response to stimuli, (b) a state space rich enough to allow input data to be expanded into higher dimensional space where classification becomes easier and (c) an asymptotic washing out of system states with stimuli (typically referred to as ‘fading memory’). This has led to a wide variety of different implementations of *in materio* RC using e.g., photonic [6], mechanical [7] and memristive [8,9] systems.

In materio RC offers potential advantages of increased computational efficiency in temporal tasks compared to other static *in-materio* computing paradigms which resemble standard neural networks, with the material acting simply as a nonlinear activation function and can even compete in performance with *in-silica* computing approaches [10]. Benchmark tasks demonstrated using RC approaches include signal transformation, speech and image recognition and time series prediction [11]. Reservoirs need to have different computational properties to be effective in different tasks. For example, tasks such as signal transformation and spoken digit recognition primarily utilise the non-linear transform provided by the reservoir, while the well-known NARMA-10 task requires the reservoir to provide both non-linearity and memory of past inputs [11].

There are mainly two ways to tune an *in materio* reservoir's properties: Firstly, intrinsic dynamics of the reservoir can be changed e.g., by making changes to the physical system used, or by replacing it with another system altogether. For example, it has been reported that magnetic nanodots [12] and artificial spin ice [13,14] systems exhibit different computational properties when their lattice arrangements are varied. Alternatively, the reservoir architecture (i.e. the way data is interfaced with the reservoir) can be changed, for example by using different time-multiplexing approaches [11]. However, the relative effectiveness of these two tuning approaches have yet to be explicitly compared for a given type of physical system.

We have recently shown that magnetic domain wall (MDW) dynamics in interconnected arrays of $Ni_{80}Fe_{20}$ (Permalloy) magnetic nanorings (NRAs) of can be exploited for reservoir computing [11,15,16]. In these studies, information was encoded in the amplitude of rotating applied magnetic fields, which then drove emergent DW interactions within the arrays. We have shown that the magnetic response of the arrays exhibited both the non-linearity and fading memory required for a useful reservoir, and demonstrated state-of-the-art performance for several benchmark tasks, including signal transformation, spoken digit recognition and time series prediction.

In this paper we explore how variations in the lattice arrangement of NRAs change both their physical behaviours, and their resulting computational properties when used as reservoirs. We study these alongside three different time-multiplexed RC architecture [11] to understand how the effects of changing the physical form of the reservoirs compare to those produced by change the way we interface data with them.

We consider three different NRA lattices with different numbers of nearest neighbours (NN) for each ring: (a) square (with NN = 4), (b) trigonal (with NN = 6) and (c) Kagome (with NN = 3). X-ray photo-emission electron microscopy (X-PEEM) imaging of the arrays' microstates show that the different lattices both exhibit different characteristic configurations of MDWs and differing evolution of these with applied stimulus strength, thus showing the strong influence of the number and position of NN rings on NRA behaviour. We then use anisotropic magnetoresistance measurements (AMR) to probe the global responses of each NRA lattice arrangement and use these with the three reservoir architectures to evaluate task agnostic metrics that express their computational properties. We observe that the time-multiplexing approach has a more profound influence on computational properties than the type of NRA lattice used, suggesting that our global readout mechanism does not allow differences in the rich microstate behaviours of the arrays to be fully captured. However, reservoirs constructed using the combined outputs of NRAs with different lattice arrangements do show superior performance in metric space to any reservoir constructed from a single NRA, indicating the utility of combining the dynamics of multiple material reservoirs for improved computation.

Methodology.

All measurements were performed on 25 × 25 NRAs. Each NRA had rings of diameter 4μm and ring widths of 300 − 400nm. The overlap between rings was 150 − 200nm. Fig. 1 (a)-(c) shows scanning electron microscopy (SEM) images of the arrays square, trigonal and Kagome lattices respectively.

NRAs were patterned using electron beam lithography with lift-off processing. After electron beam exposure and development 10nm of $Ni_{80}Fe_{20}$ was thermally (AMR measurements) or electron beam (XPEEM measurements) evaporated (at a baked out base pressure of $< 10^{-7}$ mbar) before lift-off. The samples used for AMR measurements were patterned on thermally oxidised silicon substrates, and underwent a second stage of lithography and metallisation to pattern Ti(20nm)/Au(100nm) electrical contacts at the edges of the NRAs (Fig. 5 (b)). The samples for X-PEEM measurements were patterned on silicon substrates with native oxide layers and had an additional ∼ 2nm Aluminium capping layer to avoid charging of the samples during X-PEEM imaging.

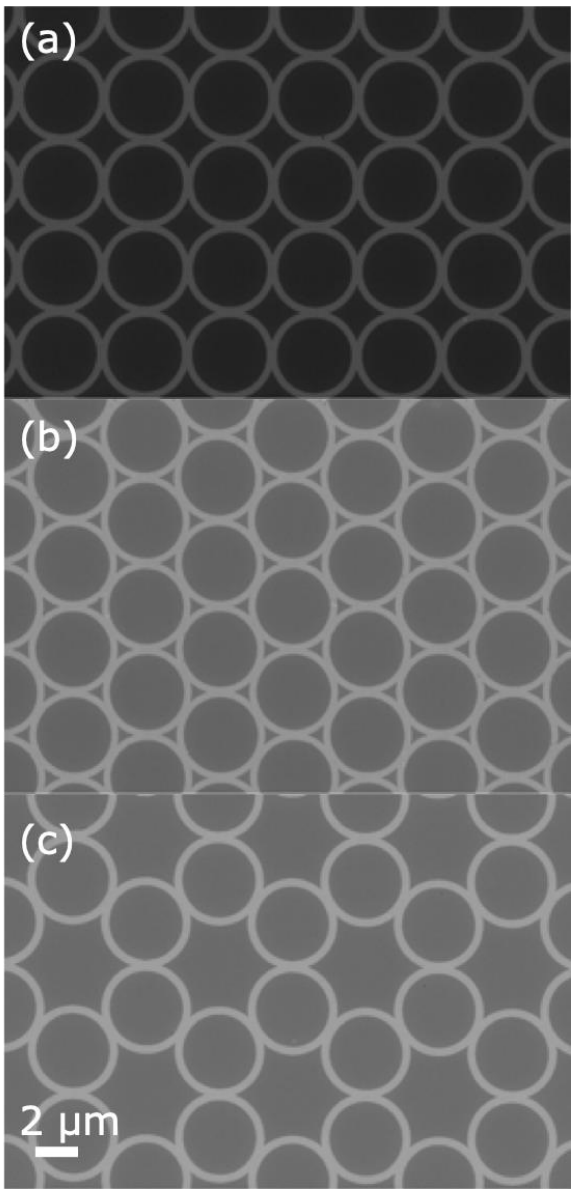


Figure 1. SEM images of ring arrays with (a) square, (b) trigonal and (c) Kagome lattice arrangements.

AMR measurements were performed using a custom-built electric transport measurement rig. A sinusoidal probe current of 1 mA was provided to the NRA contacts at a frequency of 43117 Hz using a Keithley 6221 current source. Resistance changes due to AMR effects [17,18] were measured using a Stanford Research Model SR830 lock-in amplifier.

Rotating magnetic fields were generated using two pairs of air-coil electromagnets in Helmholtz-like configurations. The electromagnets were driven by a pair of Kepco BOP 36-6D power supplies and were controlled via voltage signals supplied via a National Instruments acquisition card.

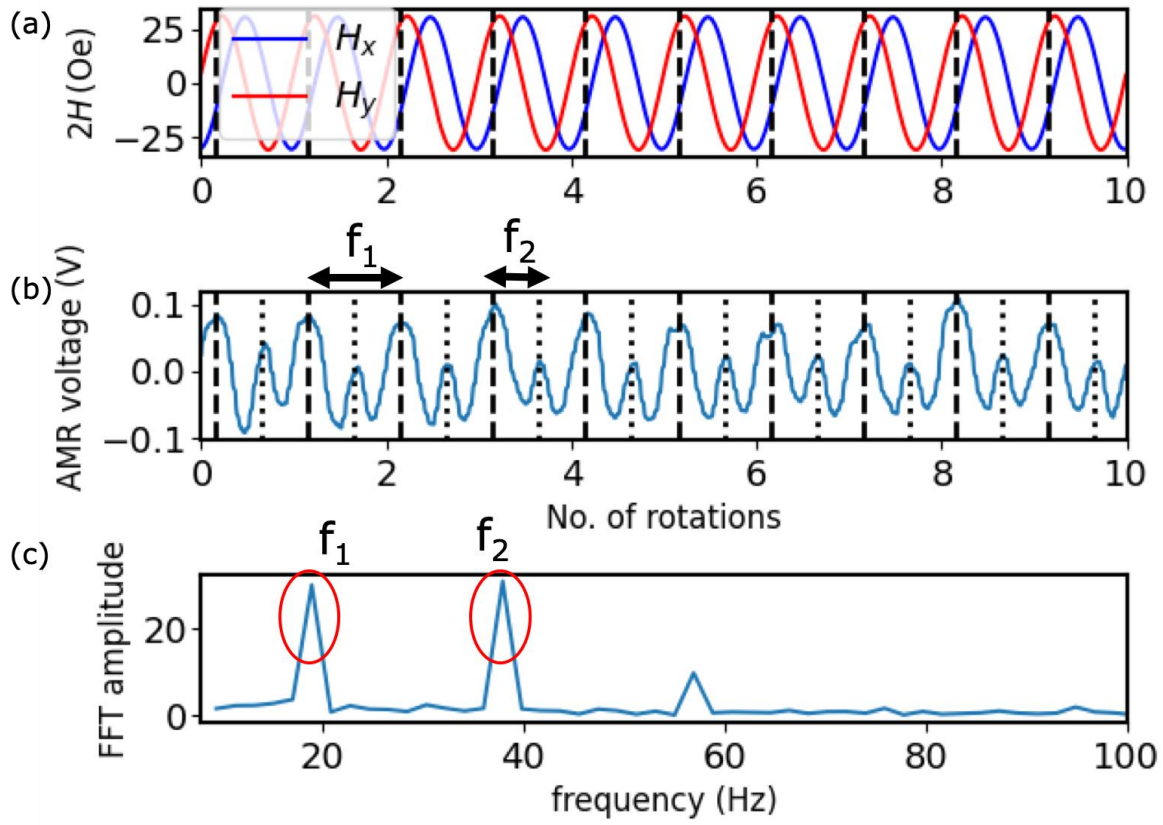


Figure 2. (a) The H_x and H_y components of the field profile that was applied to the sample and (b) the measured AMR signal for the first 10 rotations. (c) The Fourier transform of the signal with the f_1 and f_2 components marked.

AMR characterisation of the NRAs was performed as follows: The magnetisation of the NRAs were initially saturated (not shown in Fig. 2) and brought to remanance which created MDWs in the ring. Subsequently 30 cycles of a rotating magnetic field at a frequency of 19 Hz, which were created by applying 90° phase shifted waveforms to the *x* and *y* axis electromagnet coils (Fig. 2 (a)), were applied and the AMR signal was recorded (Fig. 2 (b)). The FFT of the last 10 cycles of the AMR signal was then obtained and the amplitudes of the f_1 = 19 Hz (input field frequency) and f_2 = 38Hz (2× input field frequency) components (marked in Fig. 2 (c)) were extracted. This was repeated for different field amplitudes. The f_1 and f_2 frequencies correspond to MDW contortion due to susceptibility effects and irreversible DW propagation transitions respectively and are prominent in the measured AMR response [11].

X-ray photo-emission electron microscopy (X-PEEM) was performed at the CIRCE beamline at the ALBA synchrotron. Magnetic domains images were obtained by averaging a series of X-ray absorption (XAS) images on and off the Fe-L_3 resonance and with left and right X-ray circular polarization in order to generate contrast by X-ray magnetic circular dichroism (XMCD). Samples were mounted on cartridges with a quadrupole magnet [19], which was used to generate in-plane rotating magnetic field at a frequency of 1 Hz. In the X-PEEM measurements, the rings were first all driven into saturation and then 30 cycles of rotating fields of different amplitudes were applied with images being taken after rotation. Microstate populations were extracted from these images by counting the magnetic states using Python custom image processing libraries [20]. Confidence intervals for the state populations were calculated by measuring standard deviations of the counts of each state over multiple runs of the measurements. The net magnetisation of the array was calculated by finding the weighted sum of the signed magnetisation components along the PEEM sensitivity direction for the relative population of different magnetic states.

In order to give a broader description of the range of computational behaviours available to a given reservoir, rather than simply performance in a given task, here we employ task independent metrics to characterise the different lattice arrangement's computational properties. These metrics used here, kernel rank (KR), generalisation rank (GR) [21,22] , and linear memory capacity (MC) [23] evaluate the computational properties of a reservoir along three different axes using randomly generated data, and have been shown to be good predictors of task performance [22]. For full details on the implementation of each of the metrics, see Supplementary Note 1.

The metrics of KR and GR are evaluated similarly but measure opposite properties. Both evaluate the number of linearly independent output states when driven with uncorrelated (KR)/ correlated inputs (GR). This measures the reservoir's ability to separate distinct input sequences (KR), as well as generalise similar inputs (GR). Both metrics are bounded by the number of output nodes the reservoir has. A system with good separation properties will have a high KR, while a system with good generalisation properties will have a low GR. While the exact balance between ability to separate and generalise will vary from task to task, a basic heuristic, computing quality (CQ) [24] was constructed by calculating the difference between KR and GR.

Another important property of reservoir computers is their dependence upon past information. To measure this, MC evaluates how well the reservoir is able to reconstruct past inputs from its current reservoir state with a linear output layer. To do this, the readout layer of the reservoir is trained to reproduce delayed states over an input signal, then evaluated on an unseen test set. Again, MC is bounded by the number of nodes in the network.

To explore the full range of dynamic regimes of the NRA, and hence measure the different computational properties of each regime, the input data was scaled to cover different ranges of the NRA's response. Input data u_τ was encoded into the amplitude of the global rotating field at a frequency of 37Hz. The encoding was linear and of the form $H_{rot} = H_c + H_r \times u_\tau$, where H_c is the centre field and H_r is field range of the transformation.

Previous work has shown that the full range of computational properties available to a given material cannot be accessed under a single reservoir architecture, since different architectures are able to better able to exploit given dynamic properties of a system's response for computational advantage [11]. Hence, it is important to test a range of different architectures in order to explore the full range of computational behaviours a given system can exhibit. Here we explore three reservoir architectures: the signal sub-sample reservoir (SSR), the single dynamical node reservoir (SDN), and the rotating neurons reservoir (RNR). The following section will provide a phenomenological overview of each of

the reservoir architectures. For details on implementation, and a schematic diagram of each of the architectures, see supplementary note 2.

The SSR architecture harnesses the oscillatory dynamics of the NRAs. Due to the different nonlinear relationships between field and the 1f/2f frequency components in the AMR signal, the shape of the AMR trace for a given rotation changes drastically according to field strength. When taking the amplitude of signal at fixed points within the rotation as output, both nonlinearity and dimensionality expansion is provided.

The SDN architecture, introduced by Appeltant et al. [25] utilises the transient behaviours of the dynamic system to transform input data. The network consists of a single node multiplexed in time, generating 'virtual' nodes. A fixed random mask provides different linear combinations of input dimensions to each virtual node, which are are connected to one another sequentially via the dependence of current system state on its past states. Again, both nonlinearity and dimensionality expansion are provided, plus the ability for the multiple dimensions of input data to interact with each other across time to form a richer representation.

Introduced by Liang et al. [26], the RNR architecture employs multiple dynamical nodes unconnected from one another. Instead, the input and output connections to each node synchronously rotate, which changes the input/output weights associated with each node over time. Considering the time series of a given output dimension, the output will have contributions from different nodes over time, emulating connections between nodes. By distributing information across the many real nodes, coupled with each node's inherent non-volatile response described in [11], information can stay in the system for long periods of time, generating effective memory.

Results and discussion.

Microstate Characterisation

We begin by studying the differences between the microstates formed in NRAs with the different lattice arrangements (Fig. 3). As MDWs in the NRAs tend to pin in the junctions between rings, and these junction points differ in number and position between the three lattice-types, we defined a sub-set of states for each lattice-type which are described in the table in (Fig. 3). Here, we refer to three different types of domain microstate: the 'onion' microstate which reflects a pair of domain walls at opposite ends of the ring, and hence maximum net magnetisation in the direction of the domain walls, a 'fractional' microstate which represents a shifted variant of the onion state (with the fraction denoting the relative size of the larger domain in the ring), and the 'vortex' microstate with no domain walls and flux closure within the ring.

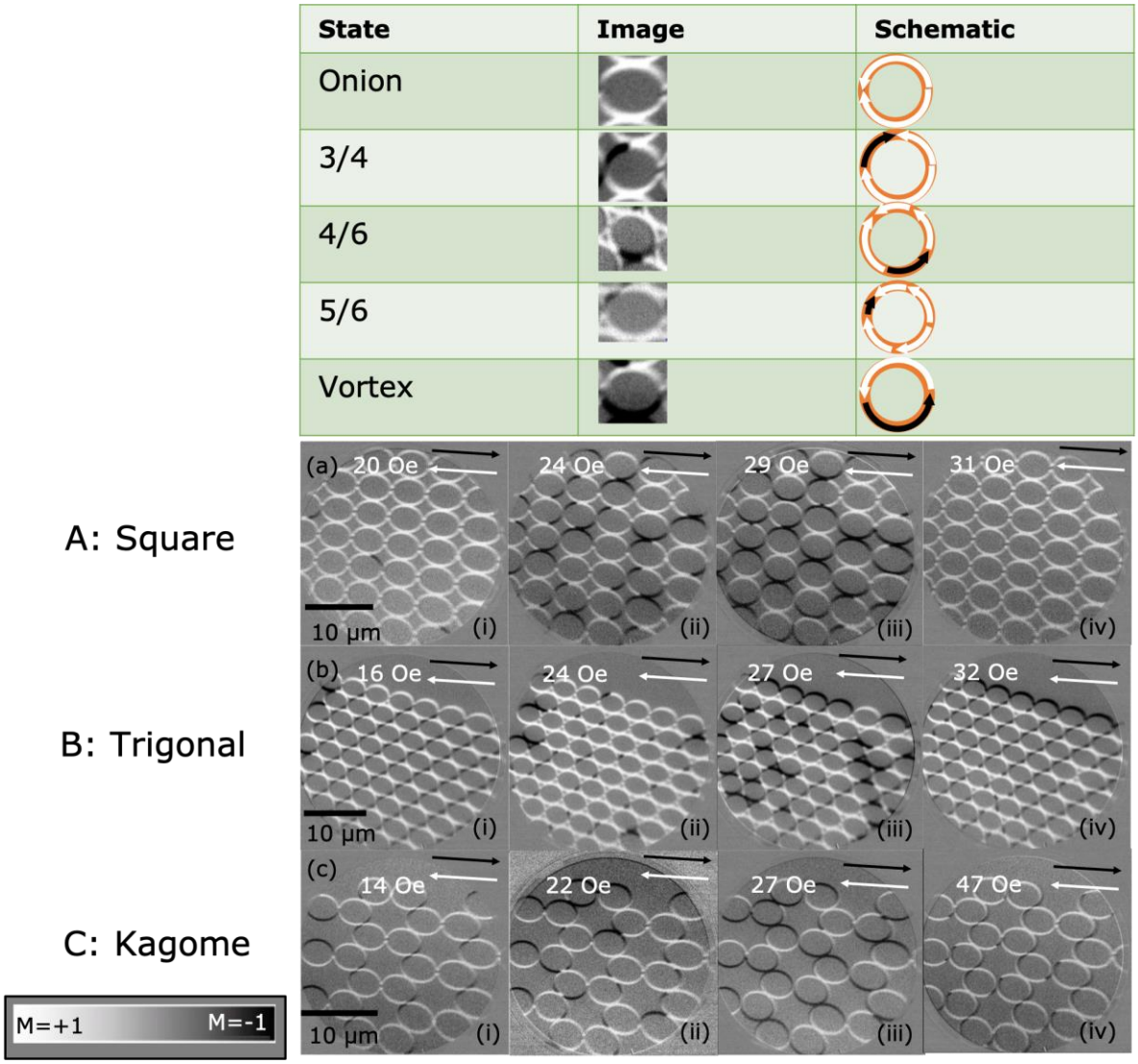


Figure 3. The X-PEEM responses for the (A) square, (B) trigonal and (C) Kagome lattice arrangements for as the rotating field amplitude was increased ((i)-(iv)). The table specified types of magnetic states seen in the different arrays along with schematics as well as representative images.

Fig. 3 (a) presents example microstates observed in the square lattice as a function of rotating field amplitude. At H_{rot} = 20Oe the field was too weak to depin MDWs from the junctions within which they were initialised, and hence all rings were in bi-domain states ('onion' states [27]), with net magnetisation aligned along the direction of initial saturation (Fig. 3(a)(i)). At H_{rot} = 24Oe the field became strong enough to cause occasional MDW movement in some rings, thus forming '3/4' states with individual MDWs rotated by 90° from their initial position (Fig. 3(a)(ii)). As the field was increased to H_{rot} = 29Oe stochastic MDW depinning events increased in frequency, leading to MDW pairs colliding and annihilating to form flux-closed 'vortex' states (Fig. 3(a)(iii)). The ground state of a square-lattice NRA is a checker-board pattern of 'vortex' states with alternating circulation direction, but this is not typically reached due to the re-nucleation of MDWs pairs into 'vortex' states by activity in neighbouring rings. Full details of these emergent effects can be found in our previous work [15]. At higher fields still (H_{rot} = 29−31Oe) large numbers of MDWs were active, meaning that MDW re-nucleation events dominated over annihilation events, and the array progressively repopulated with MDWs. By H_{rot} = 31

Oe, this re-population was complete and the NRA was saturated with ‘onion’ states, which rotated coherently with the applied field (Fig.3(a)(iv)).

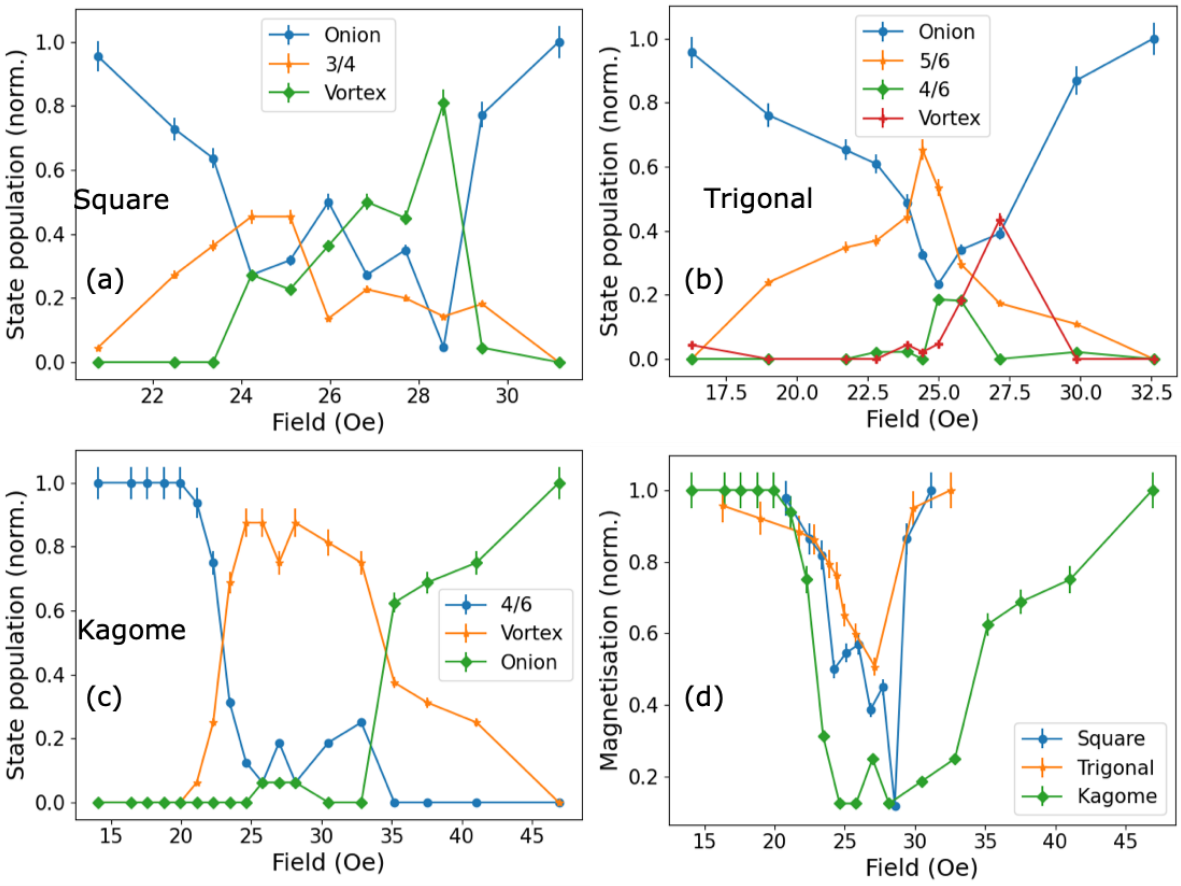


Figure 4. The magnetic state populations for the (a) square, (b) trigonal and (c) Kagome lattice NRAs as H_{rot} is varied. The magnetic states are schematically shown in Fig. 3. (d) The variation of the net magnetisation along the saturation direction with H_{rot} for the three lattice arrangements.

Field-dependent state populations for the square lattice are shown in Fig. 4 (a). A strongly non-linear variation of the these was observed in the region of emergent dynamics, resulting in a plurality of MDW states, being observed at intermediate applied fields (H_{rot} = 24 − 31Oe). The decrease in ‘onion’ states (between H_{rot} = 20 − 24Oe) was accompanied by an increase in ‘3/4’ states. The ‘3/4’ states then gave way to a majority of ‘vortex’ states, the population of which continued to increase until they peaked at H_{rot} = 29Oe. Further increases in field resulted in a relatively sharp repopulation with ’onion’ states. These behaviours were reflected in the variation of the NRAs net magnetisation with field (Fig. 4 (d)) which was significantly reduced in the emergent regime due to the large population of ‘3/4’ and ‘vortex’ states.

The trigonal NRA exhibited significant differences in behaviour to the square array. This was to be expected as the increased number of nearest neighbour rings in these arrays (NN = 6) meant that the MDW motion was inhibited by a greater number of pinning sites, and the formation of interlocking vortex states in triads of adjacent rings was inherently be frustrated [28]. Following saturation the rings

adopted configurations, which were broadly identical to the bidomain 'onion' states observed in the square array (Figs 3(b)(i)). As the applied field increased beyond the onset of MDW depinning (H_{rot} = 24−27Oe), new microstates were formed by MDWs moving either one ('5/6') or two ('4/6') junctions around the rings. Notably, many less 'vortex' states were formed than in the square lattice, reflecting the inherent geometric frustration of the lattice (Figs. 3(b)(ii)&(iii)). At higher fields still (H_{rot} = 32Oe) the array was repopulated with 'onion' states in a similar manner to square lattice (Fig. 3(a)(iv)). Field-dependent state populations again showed strongly non-linear trends, with a progressive evolution from majority 'onion' to '5/6', then '4/6' to 'vortex' as MDW depinning became more likely before an eventual re-saturation with 'onion' microstates when the field was strong enough to reliably overcome pinning at all junctions (Fig. 4 (b)). The variation in net magnetisation was broadly similar to that observed in the square lattice, with a dip occurring at intermediate fields, but with dramatically reduced magnitude due to the comparatively low 'vortex' population in the trigonal array.

The Kagome NRA exhibited the simplest microstate behaviour of the three lattices studied. Following the saturation, rings adopted '4/6' states (Fig. 3(c)(i)). These differed from the 'onion' state configurations observed post-saturation in the other lattices, as the lower symmetry of the Kagome lattice meant no two junction sites were directly opposite each other. As the applied field increased (H_{rot} = 22 − 27Oe) almost all rings progressively fell into interlocking 'vortex' states, thus reaching the magnetic ground state of the array (Fig. 3(c)(ii) & (iii)). This was enabled by the lower number of pinning sites in the Kagome lattice, which made it easier for MDW pairs to meet, and annihilate with each other. Notably, a much higher applied field (H_{rot} = 47Oe) was required to fully populate the array with 'onion' states than in the other two geometries (Fig. 3(c)(iv)). This was because the square and trigonal lattices always contained a residual population of MDWs that could assist with MDW repopulation, while these had to be nucleated from the ground state in the Kagome lattice. Plots of field-dependent microstate populations confirmed that the Kagome array showed the largest population of 'vortex' states of all three lattice geometries (Fig. 4(c)). This was also reflected in the array's net magnetisation, which reached lower values at intermediate fields than either of the other lattices (Fig. 4(d)).

Collectively, the analysis of microstates showed that the different lattices differed substantially in both the microstates they formed, and the way the populations of these varied with applied fields. Such differences would be expected to result in differences in computational behaviour were they able to be accessed by a tractable readout mechanism. However, the net magnetisation data presented in Fig. 4(d) showed that when the average properties of the arrays were considered these differences became less profound. There, all three lattices showed a broadly similar non-linear trends, differing only in the magnitude and field scaling of their responses.

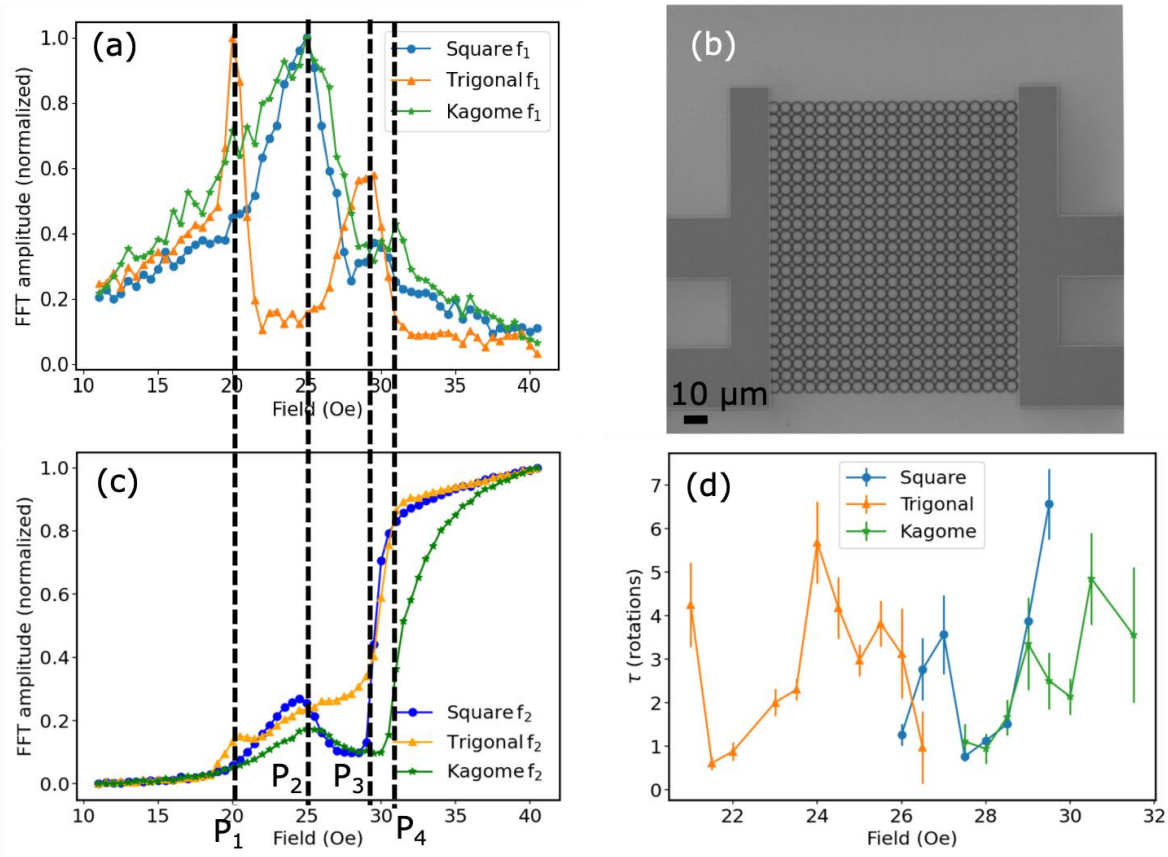


Figure 5. The (a) f_1 and (c) f_2 AMR responses as a function of H_{rot} show the different transitions in MDW dynamics in the NRAs. (b) An SEM micrograph of the trigonal NRA showing the Au contacts used for the AMR measurements. (d) The time scales of settling for the different NRAs extracted from the AMR measurements.

Having explored the detailed microstate behaviour of the arrays we now explore the responses that are accessible via magnetoresistance measurements and are therefore device tractable.

Macroscopic Characterisation

The macroscopic electric responses of the NRAs were obtained using Au contacts on either side of the NRA (shown in the SEM micrograph in Fig. 5 (b)) are shown in Figs. 5 (a) and (c). Using the procedure described in Sec. 2, the f_1 = 19Hz and f_2 = 38Hz components were monitored as a function of H_{rot}. As mentioned before, the f_1 component contains features of the distortion the MDWs undergo as a rotating magnetic field is applied and the f_2 response denotes the irreversible transitions when MDWs overcome the energy barrier at ring junctions and propagate [11]. All the different lattice arrangements show peaks in the f_1 response. Similarly, the f_2 responses for all the lattice arrangements start with a slow increase in amplitude and shows a large transition at higher field amplitudes which corresponds to when all the MDWs in the array overcome the energy barriers at junctions and start propagating with the rotating field.

The square and Kagome lattices show a prominent peak in the f_1 response (marked by P_2) at $H_{rot} \approx$ 25Oe and a harmonic of this peak is seen in the f_2 response. This marks the field at which some of the MDWs depin from junctions stochastically (after maximum distortion) and start propagating [11]. For the trigonal lattice, this peak happens at $H_{rot} \approx$ 20Oe (marked by P_1) suggesting a lower energy barrier for the trigonal lattice that the MDWs have to overcome to start propagation. The sharp transition in the

f_2 response occurs at the $H_{rot} \approx 29$Oe (marked by P_3) for the square and trigonal lattices and the f_1 responses at P_3 show a large peak for the trigonal lattice and a small peak for the square. This transition for the Kagome lattice happens at a higher field of $H_{rot} \approx 32$Oe (marked by P_4 with a similar small peak in the f_1 response) and is a consequence of the higher field required to repopulate the array with MDWs from the interlocking 'vortex' states as described above. The field regime between the 1st and 2nd peaks in the f_1 response is when a variety of the different microstates shown in the PEEM images are stochastically formed. The square and Kagome arrays show a decrease in the f_2 response in this field regime due to the formation of 'vortex' states and this is not seen in the f_2 response of the trigonal lattice due to the reduced number of vortex states formed. We thus see that these results broadly correlate with the magnetisation variation of the arrays of different lattice arrangements (shown in Fig. 4 (d)). However, the features of the microstates seen in the magnetisation variation of Fig. 4 (d) do not manifest in the macroscopic ensemble AMR response of these arrays.

In addition to the non-linearity of its transfer function, the transient nature of a reservoir's dynamics is key to its computational properties [4]. To understand this, we also studied the time scales over which the NRAs' AMR signals settled at select field points in their emergent regimes. In these measurements, the magnetisation of the array was again initially saturated and relaxed, creating MDWs at remanence. Then 50 cycles of a rotating field were applied on the array and the AMR signal simultaneously recorded. The time constants (τ) were obtained by fitting an exponential dependence of the form $y = A\left(1 - e^{-\frac{x}{\tau}+B}\right)$ to the envelope of the AMR signal. Note that above and below these field values, the time constants could not be extracted. This was because at lower fields, the change in AMR signal was not appreciable and at higher fields, the signal settled into high amplitude oscillations (corresponding to propagating MDWs) by one field rotation. It was observed from Fig. 5 (c) that all the lattice arrangements exhibit a similar range of maximum and minimum settling times.

Collectively the analysis of the AMR responses of the arrays shows that, while the different lattices show differences in their microstate behaviours, the global AMR measurements are insufficiently rich to capture these differences, and all three lattices show broadly similar behaviours.

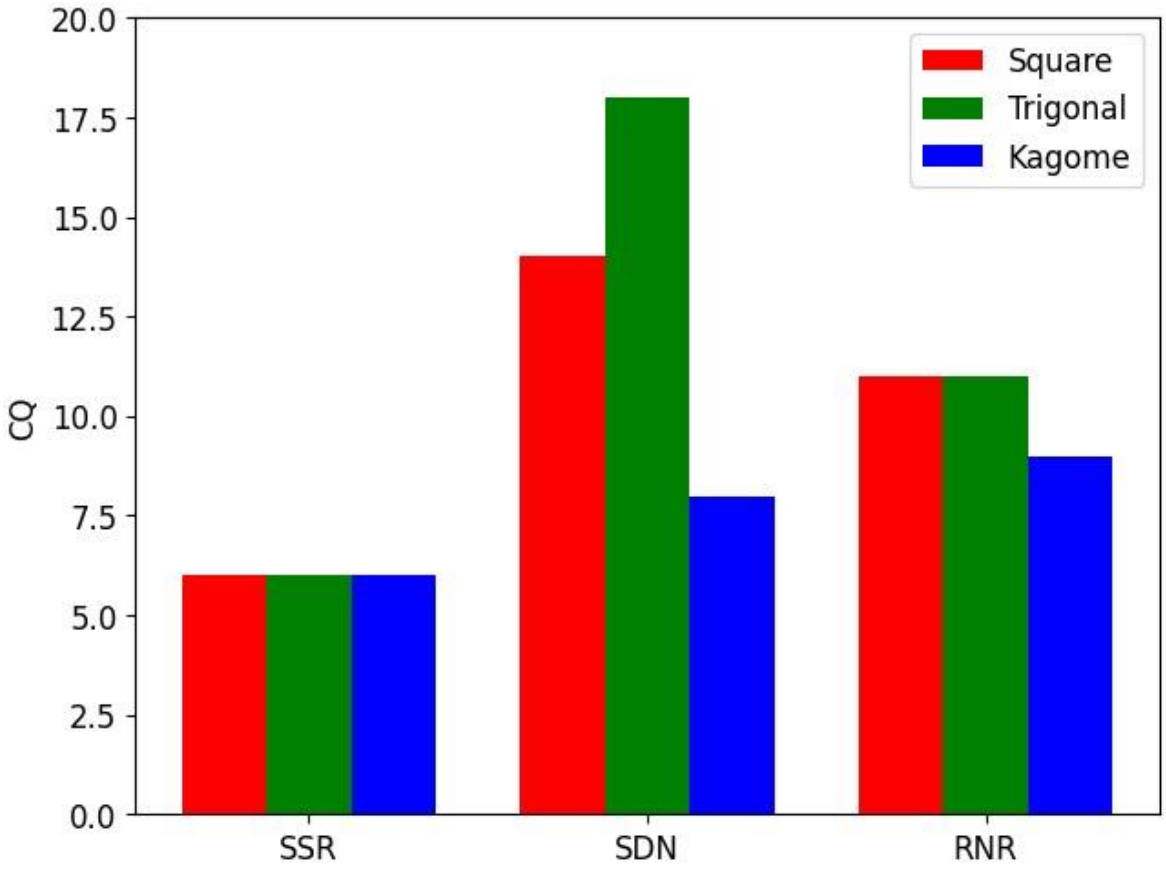


Figure 6. Comparison of peak computational quality (CQ) scores between each lattice arrangement for each of the the single subsample reservoir (SSR), single dynamical node (SDN), and rotating neurons reservoir (RNR) architectures.

Computational Evaluation

Here, we quantify the computational properties of the NRAs in each of the three lattice arrangements according to the metrics and architectures defined in 2. Fig. 6 plots the peak CQ metric score for each of the three lattice arrangements for each of the three reservoir architectures. It can be observed that architectural choice is the dominant factor in general, with smaller variance between the lattice arrangements. CQ here is correlated with the number of states available for each ring within the three lattices: trigonal has the highest number of available states, and the highest metric score, while Kagome lattices have the fewest available states and the lowest metric scores, tying the greater computational complexity offered by the square and trigonal arrays to the greater microstate complexity available to these arrays. However, there remains significant differences between the lattices under the SDN architecture, especially with the relatively poor performance of the Kagome lattice. This is likely due to increased tendency of Kagome lattices to form vortex states compared to the other arrays, since the lack of mobile domain walls with increased vortex states leads to a suppression the dynamic behaviours which the SDN architecture relies upon [11].

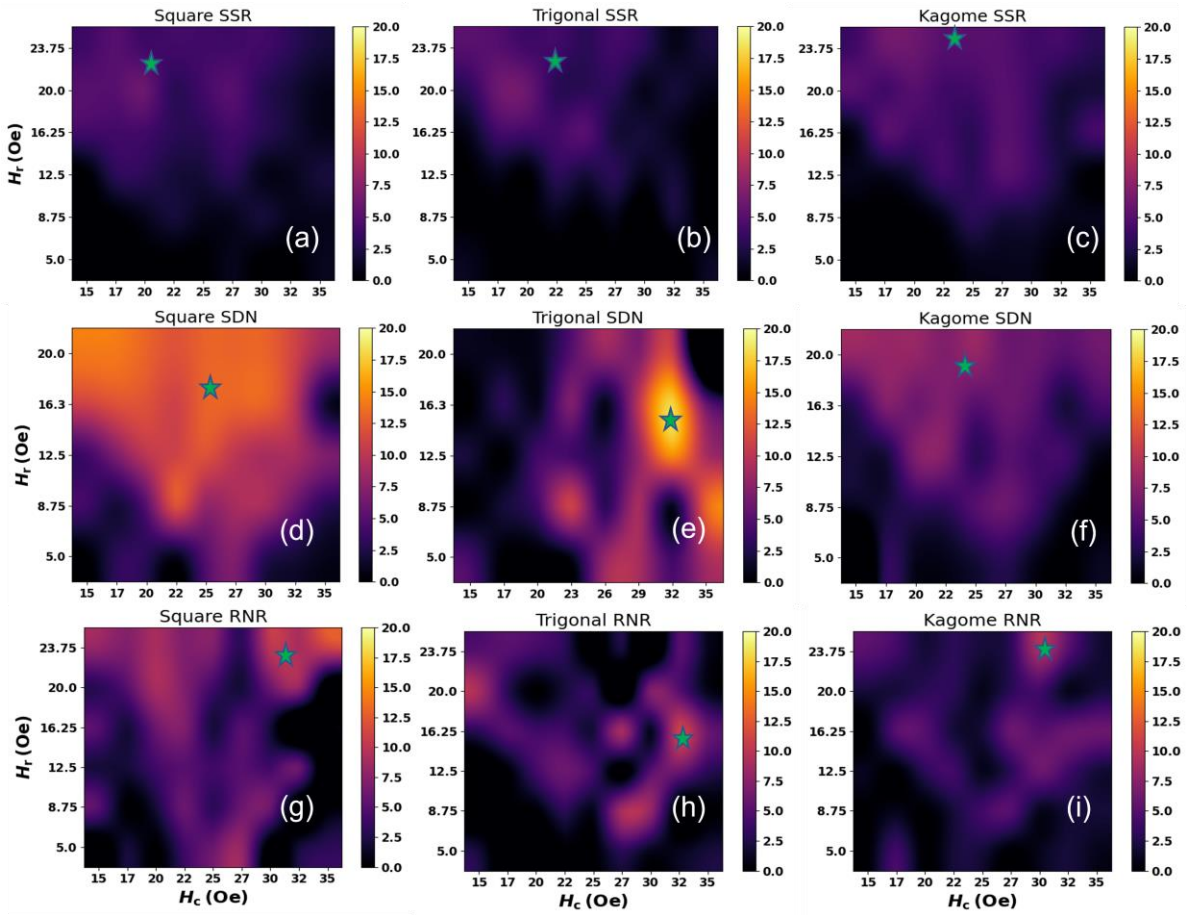


Figure 7. The computational quality metric maps as a function of the centre field (H_c) and field range (H_r) for the SSR (a-c), SDN (d-f) and RNR (g-i) RC architectures for the different lattice arrangements. The stars mark the positions of maximum CQ.

Fig. 7 shows the calculated CQ for each of the nine lattice and architecture combinations with respect to input scaling parameters. The colour maps for the trigonal array show a very different shape compared to the square and Kagome arrays, which can be correlated to the difference in the f_1 and f_2 responses in Fig. 5 (a/c): The region of poor CQ for trigonal arrays aligns with a region of the two responses where the behaviour is broadly linear ($20 < H_c < 26$ Oe), offering poor computational properties compared to the highly nonlinear responses of both Kagome and square arrays.

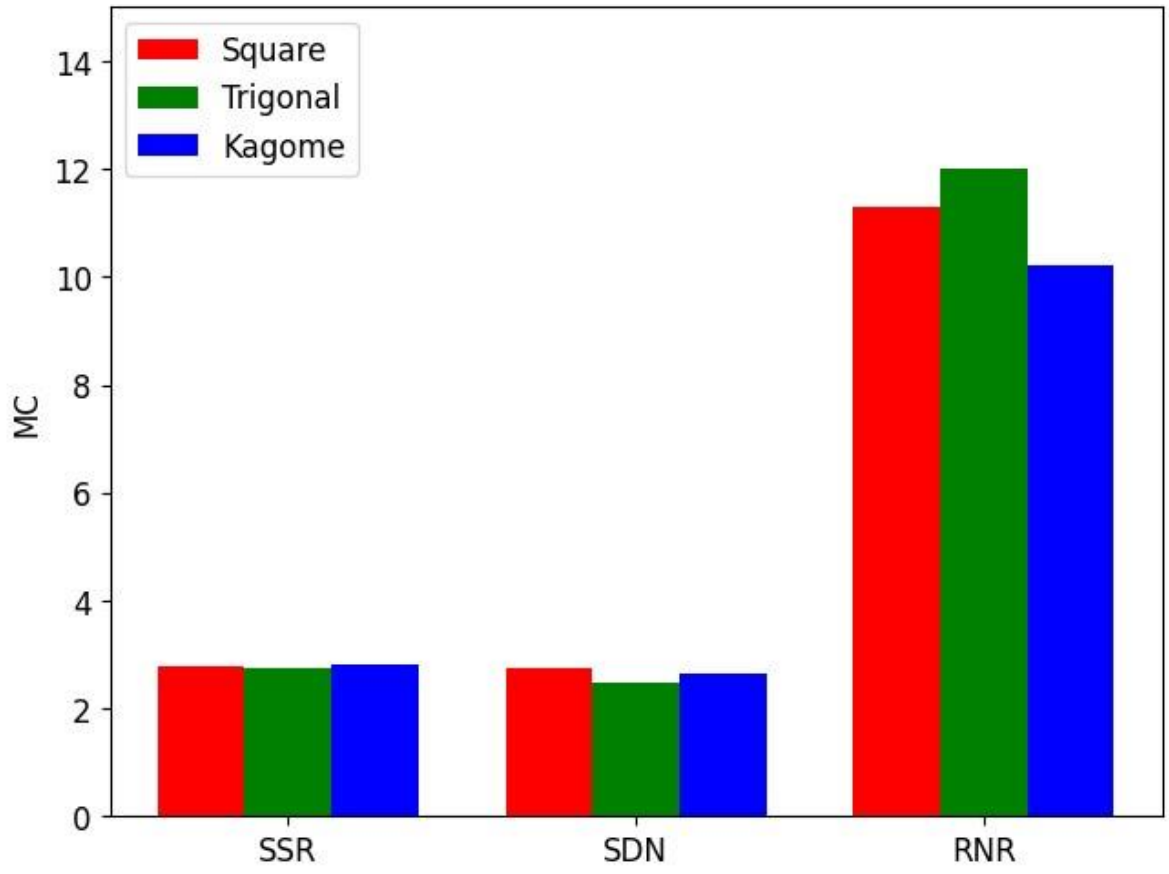


Figure 8. Peak memory capacity (MC) metric maps for the nine architecture/lattice arrangement pairs.

Memory is another computational property of reservoirs, as it determines the timescale at which the reservoir responds to input, and for how long information stays in the reservoir. This is useful for performing time-series analysis tasks more effectively, where it is critical to match the timescales of the reservoir to the timescales of the task [29] . Fig. 8 shows the highest linear MC obtained lattice and architecture combinations. Each of the lattice arrangements achieved broadly similar MC values within the same architecture, with the RNR providing the highest degree of memory and the square and trigonal arrays performing slightly better than the Kagome lattice in this architecture. This is due to the ability of the RNR architecture to store information across the distinct real nodes, whereas the single/virtual nodes in the SSR and SDN will have information washed out of the node as more inputs are provided.

Fig. 9 shows a heatmap of MC with respect to field scaling parameters for the nine lattice/reservoir pairs. For the RNR architecture, all three lattice arrangements show best performance for H_c values centred about the point at which the array starts to become dynamically active. This correlated with observations from our previous study where we showed that peak memory was obtained when input fields periodically traversed the field at which DWs became dynamically active, thus allowing information to be retained over multiple field cycles [11]. This further evidences that MC is strongly coupled to the ability to store information within the network, as well as the ability to exploit the non-volatile properties exhibited by all lattice arrangements.

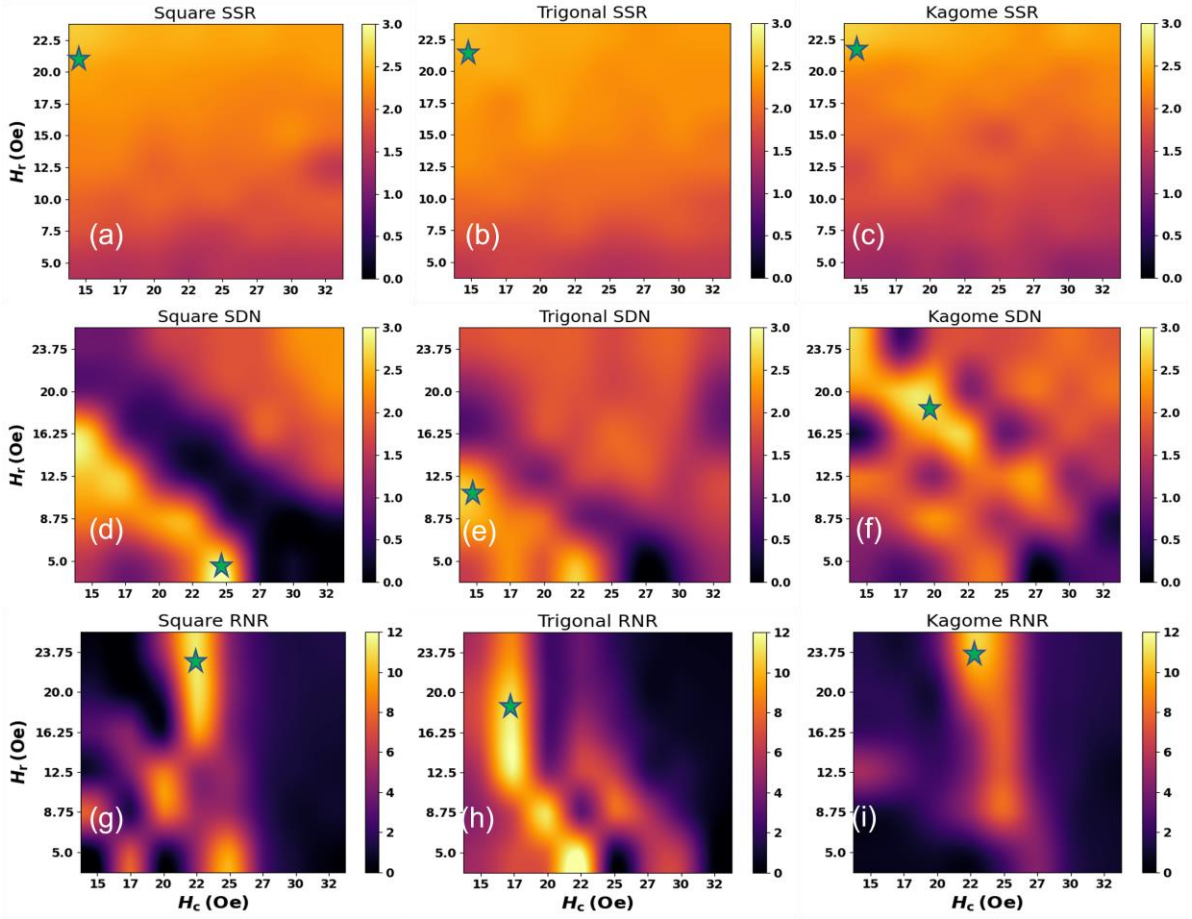


Figure 9. The memory capacity metric maps as a function of the centre field (H_c) and field range (H_r) for the SSR (a-c), SDN (d-f) and RNR (g-i) RC architectures for the different lattice arrangements.

Combining Lattice Outputs

Whilst the lattice arrangements showed slight differences in terms of maximum scores achieved in metrics, they often showed considerable differences in the input scaling parameters at which these areas of peak performance were reached, implying different behaviours in different lattices at a given applied field. This enabled multiple behaviours to be captured at a given field by combining the different lattices' responses. To show the difference in computational properties resulting from the combination of the lattice arrangements, the metric calculations were repeated with the output for each lattice concatenated together to form a single reservoir state matrix with three times as many outputs per input.

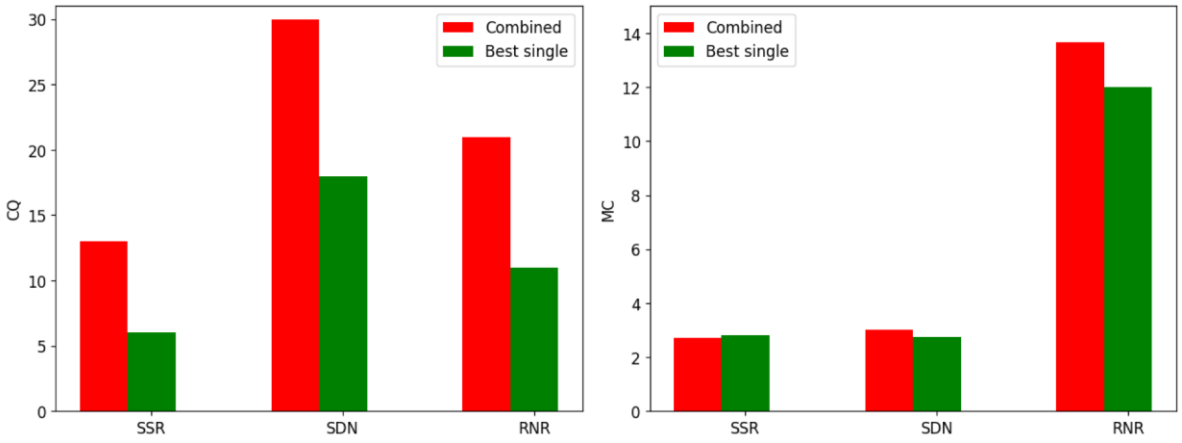


Figure 10. Comparison between peak performances of the combined networks (red) and the highest performing single network (green) for both computational quality (CQ) and memory capacity (MC).

Fig. 10 shows the resulting peak CQ and MC values compared to the best performing single lattice/architecture pair for the three reservoir architectures. It can be observed that there is a considerable increase in the CQ for all three reservoir architectures compared to the best performing single case for a given architecture. This is due to the different nonlinear representations provided by each lattice, providing better ability to separate/generalise data via these different nonlinearities. When considering MC, the combined reservoirs performed similarly in terms of peak score. For the SSR and SDN cases, this is likely tied to the inability to store information beyond 2 inputs in any of the lattices in a single node/single time multiplexed node configuration. There is slight improvement for the RNR, likely due to the presence of varied timescales for the three lattices, shown in figure 5(d).

Fig. 11 shows the metric heatmaps of the combined reservoir. There is a broadening of the suitable regions of operation compared to the single lattices due to the different ranges of activity of each lattice arrangement, ensuring at least one is operating in a dynamically interesting regime for a broader range of input scaling parameters. This highlights the potential scalability of the NRAs in terms of expanding computational capability, as improved CQ is obtained when combining the different nonlinear relationships between input and output provided by each of the different lattice arrangements. Combining the outputs also provides a means of exploiting the different timescales of response from the three lattice arrangements highlighted in Fig. 5 (d). This behaviour is functionally similar to the unconnected hierarchical ESNs presented in [30], which showed improvement in solving tasks on data with multiple timescales of autocorrelation.

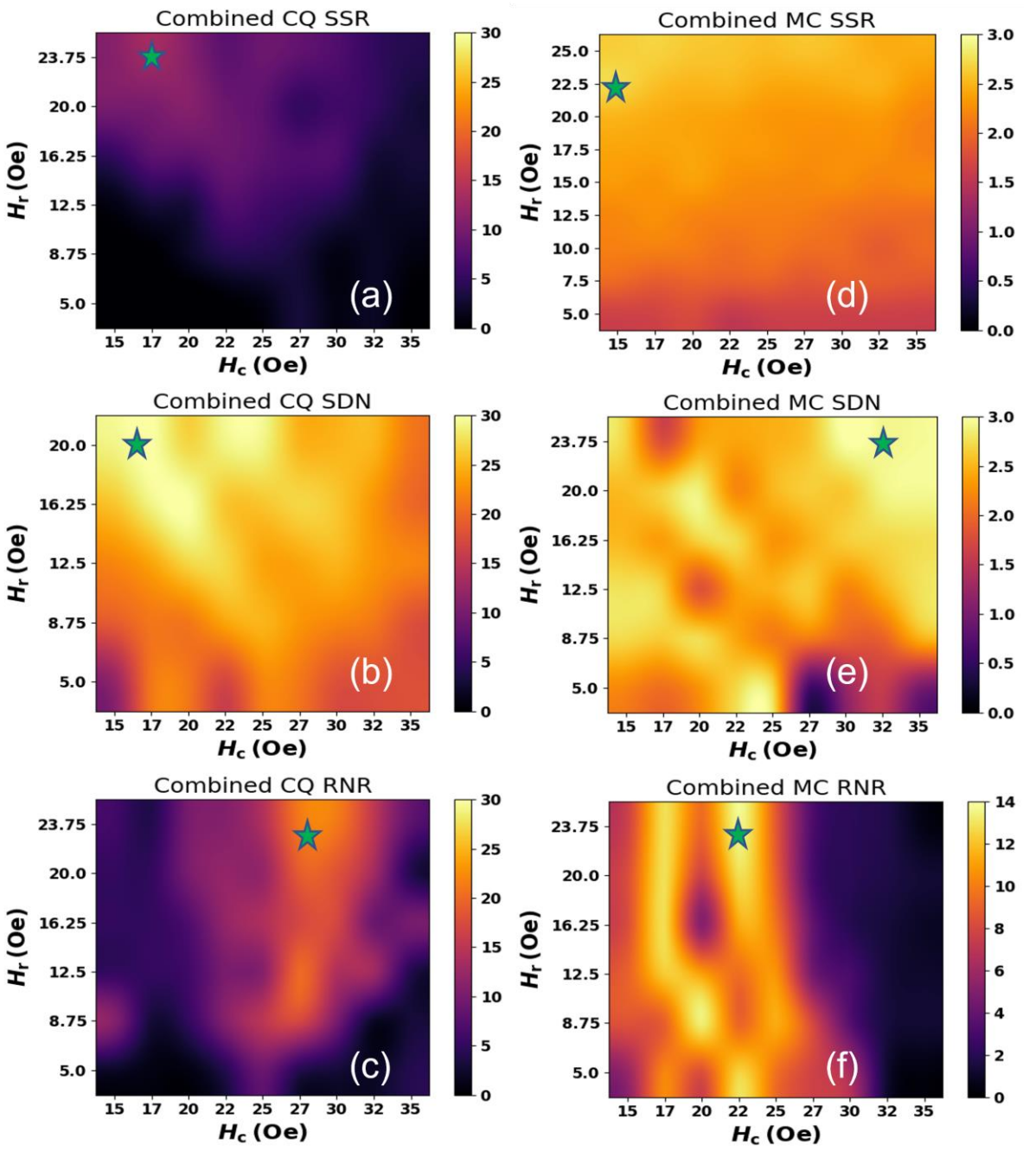


Figure 11. The memory capacity metric maps as a function of the centre field (H_c) and field range (H_r) for the SSR (a-c), SDN (d-f) and RNR (g-i) RC architectures for the different lattice arrangements.

Conclusions.

In conclusion, we have explored the magnetic responses and computational properties of ring arrays with different lattice arrangements. XPEEM measurements showed a wide variety of microscopic magnetic states for the different lattices. We extracted the state populations for the different lattices, and they showed significantly different variation in the evolution of states. We also measured the macroscopic AMR electrical response of the different lattice NRAs and observed there were differences

in the fields at which given responses occur. However, these differences were manifested less strongly in the array-averaged AMR responses.

When the computational capabilities of the arrays were measured via their AMR responses, similar computational metric scores were observed within a given reservoir architecture, highlighting the role of array-wide measurements in obscuring the different microstate behaviours. In spite of this, the ability to exploit different dynamical regimes between the lattices at a given field allowed improved computational performance when the different lattices were combined.

Collectively, our measurements show that although the different lattice arrangements have considerably different microstate responses, global 1D measurements of array state such as AMR provide limited differences in computational behaviour. In order to exploit the microstate differences observed between the lattice arrangements, a measurement technique which is microstate-sensitive (such as ferromagnetic resonance, [31]) should be employed. Additionally, we anticipate that if such arrays are driven by local stimuli (using current carrying microstrip lines or spin-orbit torque), we could make use of the high spatial inhomogeneity of NRA states in these lattice arrangements for improved computational capability.

Acknowledgments.

We acknowledge funding from the Horizon 2020 FET-Open SpinEngine (Agreement no 861618) and the EPSRC MARCH project EP/V006339/1.

References.

[1] Insights S 2023 Global Neuromorphic Computing market Size https://www.sphericalinsights.com/reports/neuromorphic-computing-market accessed: 202306-13

[2] Jaeger H 2001 *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report* **148** 13

[3] Lukoˇseviˇcius M, Jaeger H and Schrauwen B 2012 *KI-Ku¨nstliche Intelligenz* **26** 365–371

[4] Allwood D A *et al.* 2023 *Applied Physics Letters* **122** 040501

[5] Tanaka G, Yamane T, H´eroux J B, Nakane R, Kanazawa N, Takeda S, Numata H, Nakano D and Hirose A 2019 *Neural Networks* **115** 100–123 ISSN 0893-6080 URL https://www.sciencedirect.com/science/article/pii/S0893608019300784

[6] Paquot Y *et al.* 2012 *Scientific reports* **2** 287

[7] Dion G *et al.* 2021 *Reservoir Computing: Theory, Physical Implementations, and Applications* 191–217

[8] Mehonic A *et al.* 2020 *Advanced Intelligent Systems* **2** 2000085

[9] Kulkarni M S and Teuscher C 2012 Memristor-based reservoir computing *Proceedings of the 2012 IEEE/ACM International Symposium on Nanoscale Architectures* pp 226–232

[10] Dale M, Stepney S, Miller J F and Trefzer M 2017 Reservoir computing in materio: An evaluation of configuration through evolution *2016 IEEE Symposium Series on Computational Intelligence, SSCI 2016* (Institute of Electrical and Electronics Engineers Inc.) ISBN 978-1-5090-4240-1

[11] Vidamour I, Swindells C, Venkat G, Manneschi L, Fry P, Welbourne A, Rowan-Robinson R, Backes D, Maccherozzi F, Dhesi S *et al.* 2023 *Communications Physics* **6** 230

[12] Wang C, Adeyeye A and Singh N 2006 *Nanotechnology* **17** 1629

[13] Skjærvø S H, Marrows C H, Stamps R L and Heyderman L J 2020 *Nature Reviews Physics* **2** 13–28

[14] Le´on A 2013 *Current Applied Physics* **13** 2014–2018

[15] Dawidek R W *et al.* 2021 *Advanced Functional Materials* **31** 2008389

[16] Vidamour I T *et al.* 2022 *Nanotechnology* **33** 485203

[17] Bordignon G *et al.* 2007 *IEEE Transactions on Magnetics* **43** 2881–2883

[18] Ross C *et al.* 2006 *Journal of Applied physics* **99** 08S501

[19] Foerster M *et al.* 2016 *Ultramicroscopy* **171** 63–69

[20] Venkat G *et al.* 2021 Peem-data-analysis https://gitlab.com/spintronic-computing-group/ data-analysis/peem-data-analysis

[21] Bu¨sing L, Schrauwen B and Legenstein R 2010 *Neural Computation* **22** 1272–1311 ISSN 0899-7667 publisher: MIT Press

[22] Dale M *et al.* 2019 *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **475** ISSN 14712946

[23] Jaeger H 2002 Short term memory in echo state networks Tech. rep. publication Title: GMD Report 152

[24] Jensen J H and Tufte G 2020 Reservoir Computing in Artificial Spin Ice (MIT Press) pp 376–383 URL https://direct.mit.edu/isal/proceedings-abstract/isal2020/32/376/98434

[25] Appeltant L, Soriano M C, Van Der Sande G, Danckaert J, Massar S, Dambre J, Schrauwen B, Mirasso C R and Fischer I 2011 *Nature Communications* **2** ISSN 20411723 URL www.nature.com/ naturecommunications

[26] Liang X, Zhong Y, Tang J, Liu Z, Yao P, Sun K, Zhang Q, Gao B, Heidari H, Qian H and Wu H 2022 *Nature Communications* **13** 1549 ISSN 2041-1723 number: 1 Publisher: Nature Publishing Group URL https://www.nature.com/articles/s41467-022-29260-1

[27] Negoita M, Hayward T J and Allwood D A 2012 *Applied Physics Letters* **100** ISSN 00036951

[28] Rose V *et al.* 2006 *Physical Review B* **73** 094442

[29] Luko˘sevi˘cius M and Jaeger H 2009 *Computer Science Review* **3** 127–149 ISSN 15740137

[30] Manneschi L *et al.* 2021 *Frontiers in Applied Mathematics and Statistics* **6** ISSN 2297-4687

[31] Gartside J C, Stenning K D, Vanstone A, Holder H H, Arroo D M, Dion T, Caravelli F, Kurebayashi H and Branford W R 2022 *Nature Nanotechnology* **17** 460–469 ISSN 1748-3395 URL https://doi.org/10.1038/s41565-022-01091-7

Tuning reservoir computing performance modifying physical device response and reservoir architecture

Supplementary Material

Calculating Metrics

The computational metrics of KR and GR were evaluated similarly though they describe opposite properties. For KR, the system was driven by *M* sequences of length *N* independent and identically distributed (i.i.d.) random inputs sampled from a uniform distribution between zero and one, where M = the number of output nodes N_{nodes} in the network, and N > the number of inputs required to wash out initial conditions of the reservoir. The input sequences for GR were generated similarly, though the final three entries to each input sequence were identical to the first sequence. Each sequence was transformed sequentially, with the reservoir being reinitialised to the saturated state between sequences.

The nonlinear dynamics of the system provided a transformation of these input signals and an output matrix X of dimensions [$M \times N_{nodes}$] was taken from the node activities after the final input. Singular value decomposition was then performed on the output matrix X, and number of the resulting singular values above a threshold value (typically ≈ 0.1) was used to estimate the rank of matrix X, giving the degree of nonlinearity between each node's representation of different inputs. For KR, a high ability to provide nonlinear representations was reflected by a high KR value, while for GR, a high rank reflects a high sensitivity to past inputs and an inability to converge to linearly related states. A classification heuristic, computing quality (CQ) [1], was constructed by calculating the difference between KR and GR. All three metrics were bounded by the smallest dimension of *M* or N_{nodes}.

For MC, a similar i.i.d. input vector s_i of length *L* was used (here *L* = 2000), again sampled from a uniform distribution between zero and one. The input vector was converted into appropriate driving fields (using the linear encoding described above) and the resulting reservoir states are gathered after every input, forming an output matrix X of dimensions [$L \times N_{nodes}$]. The memory capacity of the system was calculated by training a linear readout layer to reconstruct a target signal Y, consisting of the previous N_{nodes} inputs to the reservoir at each time-step. After discarding an initial washout period, matrices X and Y were split into a training and testing dataset, with the former being used to optimise the weights of the linear readout layer, and the latter being used for evaluation of performance. The value of MC is given via:

(8.1)

$$MC = \sum_{k=0}^{N_{nodes}} MC_k = \sum_{k=0}^{N_{nodes}} \frac{\mathrm{cov}^2(\tilde{\mathrm{Y}}_{\mathrm{t-k}}, \mathrm{Y}_{\mathrm{t-k}})}{\sigma^2(\tilde{Y}) \times \sigma^2(Y)}$$

where cov() denotes the covariance of two sequences and $\tilde{Y}_{t-k}$ is the prediction of the linear readout for delay k. MC is hence bounded by the number of nodes in the system.

Implementation of reservoir architectures

Signal sub-sample reservoir.

In this paradigm, each datum *i* in a 1D discrete input signal s_i is used to scale the magnitude of a single rotation of magnetic field, and the AMR signal is logged over the rotation. Features are extracting by taking a fixed number *N* (*N* = 32 here) samples per input, creating a feature vector of length *N*. These features are then used to generate reservoir states matrix $X_{i,N}$ by taking concatenating these feature vectors for each timestep *i*.

Single dynamical node reservoir.
This approach was introduced by Appeltant et al. [2] and involves the time-multiplexing of input signals $s_{i,d}$ via a fixed input mask $M_{d,N}$ consisting of randomly sampled numbers from a uniform distribution between -1 and 1, where *d* = number of input dimensions, and *N* = the number of desired virtual nodes in the resulting reservoir. At each timestep, the reservoir input $\tilde{s}_{i,N}$ is given by $\tilde{s}_{i,N} = s_{i,d} \times M_{d,N}$. Each entry of $\tilde{s}_{i,N}$ is then used to scale the magnitude of rotating field for a single rotation per input. Reservoir state matrix $X_{i,N}$ is generated by taking the peak-to-peak amplitude of the resulting AMR signal for each input in $\tilde{s}_{i,N}$.

Rotating neurons reservoir
This approach was introduced by Liang et al. [3], where the reservoir consists of multiple dynamical nodes, and inputs $s_{i,d}$ are scaled via a rotating input mask $M_{d,N}$ similarly to the SDN reservoir (again, *N* = 50), except mask values are instead randomly sampled from binary values of -1 or 1. At each timestep, the input mask 'rotates' by shifting each row upwards by one position, changing connection between each mask value and node by one position every timestep. The output weights undergo the opposite transformation, ensuring that correlation between input mask and output weight is maintained. This process is analogous to keeping the input/output values fixed, and changing which node receives the input is measured for output at each timestep.

References:

[1] Jensen J H and Tufte G 2020 Reservoir Computing in Artificial Spin Ice (MIT Press) pp 376–383 URL https://direct.mit.edu/isal/proceedings-abstract/isal2020/32/376/98434

[2] Appeltant L, Soriano M C, Van Der Sande G, Danckaert J, Massar S, Dambre J, Schrauwen B, Mirasso C R and Fischer I 2011 *Nature Communications* **2** ISSN 20411723

[3] Liang X, Zhong Y, Tang J, Liu Z, Yao P, Sun K, Zhang Q, Gao B, Heidari H, Qian H and Wu H 2022 *Nature Communications* **13** 1549 ISSN 2041-1723 number: 1 Publisher: Nature Publishing Group URL https://www.nature.com/articles/s41467-022-29260-1

9- Conclusions and Future Work.

This thesis has described the investigation and application of arrays of interconnected permalloy nanorings as a reservoir computing platform. Through both simulations and experiments, a deep understanding has been developed in how the nanoring systems respond to input stimuli of rotating magnetic fields, and how the observed dynamics can be best exploited under the framework of reservoir computing. Key developments include a simulation platform for probing the response of the nanoring arrays, methodologies that can expedite the process of determining optimal dynamical regimes for computation, an experimental setup for measuring the magnetoresistance response of the nanorings, and demonstrations of computation in both simulation and hardware. The findings of this work will serve as the foundation for future exploration into the development of the nanoring arrays as a potential computing platform. The key findings of each of the research chapters presented here, as well as both the currently occurring and the potential future studies that follow the work presented in this thesis.

Chapter 5 presented the design and validation of RingSim, a bespoke simulator for modelling the state of nanoring networks when driven by rotating magnetic fields. It was shown that the complex response of the arrays can be represented by modelling the outcome of stochastic pinning events, with phenomenological consideration of the interactions between domain walls in the system. RingSim showed good agreement with a range of experimental data, representing the equilibrium responses, the dynamic responses, as well as the microstate responses of the ring arrays.

In order to refine the model further, a more complete analysis into the distributed properties of the nanoring arrays that arises from the manufacturing procedure would be useful, as the inclusion of distributed properties in RingSim were simply fit to best approximate experimental data. Alternatively, a thorough investigation into the consistency of the fabrication procedure, as well as the consequences it has on the energetic landscape of the junctions within the array, would provide a more concrete description of the variance observed experimentally and implemented in RingSim.

At present, the formulation of RingSim only covers the square lattice arrangement. Since the physical processes that dictate the responses of the arrays will translate to both the trigonal and Kagome lattice arrangements (and even more arbitrary, aperiodic tessellations), it is likely that an extension of RingSim will serve as a good approximator of these arrangements also. Indeed, the generalisation of RingSim is being actively pursued, and is in the stage of validation, with collaborators in the University of York.

Finally, it would also be useful to extend the outputs of RingSim to resemble the experimental measurements used for computation, at present AMR but potentially with other readouts in the future. This would mean RingSim would be a more direct predictor of device behaviour, as it would be able to produce similar responses as those measured experimentally, rather than as simply proxies of properties like magnetisation, as is currently available. With collaborators from the University of Ghent, RingSim is currently being coupled with additional software that would enable the AMR state of the array to be calculated from the position of domain walls within RingSim. This work has used MuMax3 simulations of the AMR response of each of the possible configurations a single ring can exhibit, then translating this into a resistor network of the entire array that matches the output of RingSim to produce realistic AMR signals. This work is again in the validation stage.

Chapter 6 focussed upon the initial investigations into quantifying the computational capabilities of the nanoring arrays in simulations, with some demonstrations of task performance in a standard benchmark task of spoken digit recognition. The key outcomes of this chapter were the development of a paradigm for computing with the nanoring arrays, the identification of ideal regimes for computation from the array's response, and a methodology for speeding up the parameter selection process for a given task via the rapid evaluation of task-independent metrics. It was also observed here that with increased information on the magnetic state of the array, the computational capability of the array in a task-based environment was expanded significantly. This suggests that a potential method for expanding the computational power of the nanorings is to extract more information on the complex magnetic state of the system.

A potential extension of this work would be to investigate different readouts of the system state, or different initialisations of the array itself. Since RingSim is a good approximator of system behaviour, and the energetic properties of specific junctions can be arbitrarily defined, the potential computational benefits that could be obtained from, for example, arrays with distributed properties could be explored without the need for manufacture. Additionally, the granularity of readout could be explored, since arbitrary subsections of the arrays can be defined and measured, which would guide the implementation of locally addressable readouts of system state.

While systematic explorations of these questions would provide useful insight into how the arrays change with respect to varying parameters, a more 'black box' approach could be taken. In collaboration with the University of York, there is intention to couple the RingSim platform to CHARC- an evolutionary algorithm which is able to manipulate initialisation properties of a given system in order to maximise the range of computational properties available. This could lead to generation of creative structures in simulation, which could then be reproduced in experiments. Any potential differences between the simulation prediction and the experimental realisation could then also be used to refine the models in order to form a closed-loop optimisation process.

Chapter 7 presented an experimental demonstration of the nanoring arrays as a computational platform. To achieve this, an experimental rig for driving and measuring the arrays was created and optimised. This allowed for easy generation of driving fields according to the arbitrary input signals of a given task, the readout of system state via AMR measurements and conversion of measurements into features for machine learning, and the training of resulting networks to be performed. It was found that through synergising the demands of a task with both the input/output paradigm to the ring arrays and the dynamic properties of the array's response, the system could be reconfigured into different computational platforms capable of solving a wide range of machine learning tasks. This was crucial as it demonstrated that the different aspects of the ring array's complex emergent response could be leveraged in a controllable manner, demonstrating the nanoring array's capabilities as a flexible computing platform.

The elephant in the room when considering the nanoring arrays as a supposed low-power computing platform is the power demand of the electromagnets used to drive the system. Part of this inefficiency comes from the large area of the generated fields compared to the size of the arrays themselves. The field requirements to stimulate the arrays (10s of Oe) is relatively modest, and so it is not inconceivable that the electromagnets could be engineered to be much smaller in size and in close proximity to the ring arrays, which would significantly reduce their energy demands.

However, a more elegant solution could be to used spin-currents to drive the system. The complete removal for the need of external magnetic fields is definitely desirable for future all-electrical implementations of the system. Work is ongoing into the possibility of using the spin-currents

generated from a platinum underlayer with nanoring arrays patterned on top of the platinum. Since the nanorings have in-plane magnetisation, these implementations would require the inclusion of strong magnetic fields acting perpendicular to the plane of the rings in order to cant magnetic moments out of plane, and greatly increase the efficiency of spin-transfer. However, the contact between the underlayer and the nanorings would cause most of the measurement currents applied for AMR to be shunted through the low resistance platinum, reducing the already low percentage difference of resistance states in the AMR measurements.

Chapter 8 explored one of the possible methods for manipulating the array geometry in order to get different responses out of the system: changing the degree of interconnectivity between the nanorings. First, the study focussed on exploring the microstate differences observed between the three different lattice arrangements: square, trigonal and Kagome. The three arrangements had drastic different relationships between their respective domain states and applied field strength, suggesting that the lattice arrangements had functionally different responses to field. However, when they were measured via AMR, the three lattice arrangements showed a lot of similarities. This translated to similarities in computational performance when AMR was used as readout. This suggested that a more sophisticated, microstate sensitive readout would be required to exploit the differences between the lattices for computation. In spite of this, the shifts in field strengths at which the different arrangements became dynamically active meant that for a given applied field, each arrangement would have a different nonlinear transformation of the input. This led to improved computational properties when the different outputs were combined.

As mentioned in the preamble to chapter 8, there are currently two separate ongoing studies into other geometrical manipulations to the ring arrays: changes to the track width, which is being explored at the University of Sheffield, and also an investigation into the change in overlap of the junctions between the rings, being conducted at IBM Zurich. These studies will help inform on the extent to which the response of the ring arrays can be controlled and modified and will help engineer more sophisticated devices with distributed responses across larger arrays. This will likely boost the computational capabilities of said devices compared to the regular arrays studied thus far, evidencing the potential scalability of the ring arrays as more powerful computing platforms.

One of the key takeaway points of this chapter, which corroborates the findings presented in chapter 6, is that single readouts lead to reduced computational capabilities. Ideally, as many meaningful properties of the system state (provided they are not linearly related to one another) should be taken. This increases the dimensionality expansion provided by the reservoir as well as the degree of nonlinearity of the transformation, useful properties for both regression and classification-based tasks. One method for having a higher-resolution readout of system microstates that has been exploited for artificial spin-ices is spectral fingerprinting via ferromagnetic resonance measurements. These measurements provide a huge dimensionality of output in frequency space. It is expected that the ring arrays will benefit greatly from a similar readout, since there is huge diversity of states in the array which would be detected by spectral fingerprinting. This is actively being pursued at Sheffield currently.

Other potential avenues for further work using the nanoring arrays include moving away from the reservoir computing paradigm. Although for certain tasks reservoir computers represent the state of the art, these are often limited to temporal signal processing tasks with short term dependencies. Outside of these tasks, reservoirs are somewhat limited in terms of performance, and reservoir computers suffer diminishing returns in performance when simply increasing network size due to the random connectivity. Instead of having random networks constructed via architectural tricks to exploit the dynamic properties of a given material system, a potential paradigm shift could come from

constructing standard neural networks from dynamical nodes. Here, the underlying time-dependent properties and complex transformations provided by the material systems would still be at the heart of computation, but the networks could be tuned to the specific dynamic requirements of the task via the weighted connections between the nodes, rather than via informed selection of initialisation parameters as is the case with reservoir computing.

The big obstacle to the implementation of this approach is that almost all standard machine learning techniques used to train neural networks require explicit knowledge of the gradient of the error with respect to a change in weights and to current/past inputs. For systems like the nanorings where there are no analytical equations that describe their evolution, there is no method for calculating these gradients, hence alternative approaches must be taken. In the field of machine learning, there are two obvious data-driven candidates for approximating the gradient instead: using equation discovery approaches to find a set of differential equations that describe the system or model the system using neural networks that emulate the device's response. Alternatively, approaches such as cascade learning could be deployed, in which a multilayer network is built by finding solutions layer by layer, training at every step in a similar manner to reservoir computing. This would circumvent the need for gradients and offer some computational advantages compared to standard RC, but will eventually face similar diminishing returns with increasing network size.

As a whole, the rings exhibit many desirable properties that make them interesting as a computational platform: their driven dynamics can be effectively matched to a wide range of timescales of input data, their non-volatility is very useful for more passive forms of computation in cases where data arrives sporadically, the continuous nature of the lattice structure facilitates electrical measurement of the system, and the broad range of geometrical manipulations available to the system means that their response can be tuned and expanded to produce a range of transformations. However, at present there are engineering constraints that will limit their possible future applicability. The main concerns at present are the generation of driving stimuli for the arrays without magnetic fields, the relatively poor signal to noise ratio at present with AMR measurements despite the use of lock-in amplification, and lack of ability to locally excite and locally measure the ring arrays to take advantage of their spatially distributed response.

While it may not be likely that we have nanoring-based computing chips in our smartphones in the near future, I believe that the exploration into physics-based computation, both with spintronic platforms as well as with other families of device, will prove useful in solving certain computational problems in a passive low-power manner. If the properties of a given device are especially suited for a task, or if their response to input allows both direct sensing and computation, physical substrates can conceivably find a role in, for example, signal processing, edge computation or smart sensing.