



This is a repository copy of *Voltage-controlled superparamagnetic ensembles for low-power reservoir computing*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/175436/>

Version: Accepted Version

Article:

Welbourne, A., Levy, A.L.R., Ellis, M.O.A. orcid.org/0000-0003-0338-8920 et al. (5 more authors) (2021) Voltage-controlled superparamagnetic ensembles for low-power reservoir computing. *Applied Physics Letters*, 118 (20). 202402. ISSN 0003-6951

<https://doi.org/10.1063/5.0048911>

This article may be downloaded for personal use only. Any other use requires prior permission of the author and AIP Publishing. This article appeared in 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, 202402 (2021) and may be found at <https://doi.org/10.1063/5.0048911>.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Voltage-controlled superparamagnetic ensembles for low-power reservoir computing

A. Welbourne,^{1, a)} A. L. R. Levy,^{2, 1} M. O. A. Ellis,³ H. Chen,¹ M. J. Thompson,¹ E. Vasilaki,³ D. A. Allwood,¹ and T. J. Hayward¹

¹⁾*Department of Material Science and Engineering, University of Sheffield, Sheffield, United Kingdom,*

S1 3JD

²⁾*Department of Physics, École Polytechnique, Paris, France, 91128 Palaiseau cedex*

³⁾*Department of Computer Science, University of Sheffield, Sheffield, United Kingdom,*

S1 4DP

(Dated: 24 April 2021)

We propose thermally-driven, voltage-controlled superparamagnetic ensembles as low-energy platforms for hardware-based reservoir computing. In the proposed devices, thermal noise is used to drive the ensembles' magnetization dynamics, while control of their net magnetization states is provided by strain-mediated voltage inputs. Using an ensemble of CoFeB nanodots as an example, we use analytical models and micromagnetic simulations to demonstrate how such a device can function as a reservoir and perform two benchmark machine learning tasks (spoken digit recognition and chaotic time series prediction) with competitive performance. Our results indicate robust performance on timescales from microseconds to milliseconds, potentially allowing such a reservoir to be tuned to perform a wide range of real-time tasks, from decision making in driverless cars (fast) to speech recognition (slow). The low energy consumption expected for such a device makes it an ideal candidate for use in edge computing applications that require low latency and power.

Artificial intelligence, and in particular machine learning (ML) techniques, are widely used across a variety of sectors, but the associated energy cost of training highly complex models has become a significant challenge¹. This has led to a desire to produce efficient, hardware-based neuromorphic platforms to replace the current state of the art i.e. software-based models evaluated on traditional CMOS computing architectures².

Recurrent neural networks (RNN) are inspired by the high interconnectivity of biological systems and can solve complex time-dependent tasks, such as speech recognition. However, their temporal interconnectivity requires complex training methods that are computationally expensive and difficult to implement on hardware³. The reservoir computing (RC) paradigm provides a solution to this by using a RNN with fixed synaptic weights (the reservoir) to transform inputs into a higher dimensional representation prior to them being passed to a simple, feed-forward output layer with trainable weights⁴. Time multiplexing allows RC to be performed even if the reservoir consists of only a single dynamical node⁵.

RC is also particularly well suited to hardware-based implementations as the reservoir can be replaced with any suitable physical system providing it has the correct properties: (a) the ability to produce nonlinear transformations of input data and (b) fading memory of past inputs. Such implementations have the potential for lower energy costs than reservoirs simulated on conventional computers as they can directly utilise the intrinsic functionalities of the physical systems. This has led to a wide range of systems being proposed as suitable

reservoirs⁶. In particular, magnetic systems exhibit non-volatility and nonlinearity, and thus are positioned as ideal candidates⁷. Previous investigations have demonstrated the suitability of a wide range of magnetic systems for use as reservoirs, including spin-torque nanoo oscillators, dipole-coupled nanomagnets, and arrays of interconnected nanowires⁸⁻¹⁴.

Here, we propose strain-mediated, voltage-controlled superparamagnetic ensembles as platforms for creating ultra-low-energy hardware-based reservoirs. Using a combination of micromagnetic simulations and analytical modelling, we will show that these ensembles possess the reproducible behaviour, non-linearity and fading memory required for RC and demonstrate (using a model of the system) competitive performance on two benchmark machine learning tasks: spoken digit recognition and chaotic time series prediction.

The modelled systems consist of ensembles of circular CoFeB nanodots with diameters in the range 40 nm to 90 nm and thicknesses of 4 nm (Fig. 1(a)). The dots' diameter is fixed within a single ensemble, but different values allow access to different timescales of the ensembles' responses. We assume a large ensemble (over 1000 dots, or a 6.3 μm square array) with nanodots spaced by 160 nm to minimise dipolar coupling. The dots are assumed to have saturation magnetisation $M_s = 1200 \text{ kA m}^{-1}$, exchange stiffness $A_{ex} = 10 \text{ pJ m}^{-1}$ and a growth or shape-induced^{15,16} uniaxial anisotropy $K = 8 \text{ kJ m}^{-3}$.

A schematic diagram illustrating the geometry of an individual dot is shown in Fig. 1(b). A fixed magnetic field (40% of the anisotropy field) is applied at 90° to the anisotropy axis. Magnetoelastic control of the arrays' magnetic states is provided by voltage-controlled uniaxial strain applied at 45° to the anisotropy axis (an

^{a)}a.welbourne@sheffield.ac.uk

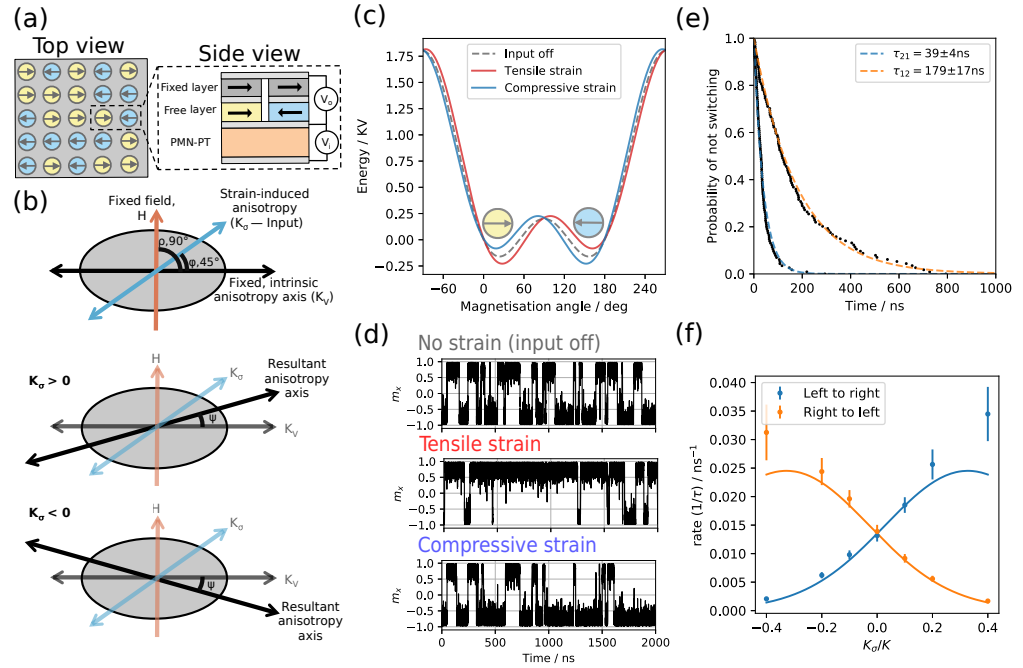


FIG. 1. (a) Schematic of a 25 dot ensemble. V_i and V_o show where voltage is input and output. The remainder of this figure considers single nano-dots from the larger ensemble. (b) Schematic of the control mechanism for the dots. Application of an additional strain-mediated anisotropy acts to rotate the anisotropy axis with respect to the fixed field. Behaviour in a single dot is shown, but the strain is applied to the whole ensemble. (c) Energy landscape for a nano-dot under strain as given by Eq. 1 for $k_\sigma = \pm 0.2$, $h = 0.4$, $\phi = 45^\circ$ and $\rho = 90^\circ$. (d) Micromagnetic simulations of thermally activated switching in a single 40 nm diameter, 4 nm thickness CoFeB cylinder with 8.2 kJ m^{-3} intrinsic uniaxial anisotropy and a field ($h = 0.4$) applied at 90° to the anisotropy axis. Top, no strain ($k_\sigma = 0$), unbiased. Middle, tensile applied strain ($k_\sigma = +0.2$), rightward pointing state preferred. Bottom, compressive applied strain ($k_\sigma = -0.2$), leftward pointing state preferred. (e) Multiple switches are sampled from the tensile strain micromagnetic simulation to plot the probability of not-switching as a function of time in the state. Average dwell time in the leftward state is shorter than in the rightward state. (f) Switching rate versus applied anisotropy for a single nano-dot. The points are for the micromagnetic simulations. The curves are rates for the energy barriers given by the Stoner-Wohlfarth model (Eq. 1).

ultra-lower-power consumption technique¹⁷). We base our parameters for this on the artificial multiferroic heterostructure PMN-PT (011)/CoFeB where strains of up to 0.175% can be applied, leading to a strain anisotropy of comparable magnitude to typical growth-induced intrinsic anisotropies^{18,19}. Input is via voltage applied perpendicularly across the PMN-PT layer. An output derived from, for example, giant magnetoresistance would be proportional to the magnetisation of the array (see Fig. 1(a)).

To model strain-mediated voltage control of the CoFeB nanodots, we express their energy using the Stoner-Wohlfarth model^{20,21}. The application of strain to isotropic, amorphous CoFeB results in the addition of a uniaxial magnetoelastic anisotropy term with mag-

nitude $K_\sigma = 3/2\lambda_s E \epsilon$, where λ_s is the magnetostriction coefficient (31–75 ppm^{19,22}), E the Young's modulus (216 GPa²³), and ϵ the voltage-induced strain²⁴. When combined with an intrinsic uniaxial anisotropy (K) the normalized energy per unit volume $e(\theta) = E(\theta)/KV$ is then:

$$e(\theta) = \sin^2(\theta) + k_\sigma \sin^2(\theta - \phi) - 2h \cos(\theta - \rho) \quad (1)$$

where θ is the angle of the magnetization to the intrinsic anisotropy axis, $k_\sigma = K_\sigma/K$, $h = \mu_0 M_s H / 2K = H/H_K$ is a fixed applied field, and ϕ , ρ are the angles of the strain anisotropy axis and field direction to the intrinsic anisotropy respectively. The two anisotropies can be combined into a single term²⁵ (see supplementary), resulting in a anisotropy axis that is rotated by chang-

ing the magnitude of K_σ through application of strain (Fig. 1(b)).

Fig. 1(c) shows how the energy barriers of the system evolve under tensile and compressive strains. Here, $\phi = 45^\circ$, $\rho = 90^\circ$, $h = 0.4$ and $k_\sigma = \pm 0.2$. The strain changes the energy profile such that the energy barriers are no longer equal for the two states. The superparamagnet's magnetization exhibits stochastic behavior, switching between the minimum energy states with a characteristic dwell time controlled by these energy barriers. Using the Néel-Arrhenius law, the switching rate from state i (magnetization right) to j (magnetization left) is:

$$\omega_{ij} = \frac{1}{\tau_{ij}} = f_0^{ij} \exp\left(-\frac{E_b^{ij}}{k_B T}\right) = f_0^{ij} \exp(-\beta' e_b^{ij}) \quad (2)$$

where f_0^{ij} is the attempt frequency (typically 10^9 – 10^{11} s $^{-1}$ ²⁶), k_B the Boltzmann constant, E_b^{ij} is the energy barrier from state i to j , e_b^{ij} the reduced energy, and $\beta' = KV/k_B T$ the scaled thermal energy. The thermal switching of superparamagnets has previously been used for random number generation^{27,28}, adiabatic computing algorithms²⁹, and encoding nonlinear functions³⁰.

This stochastic switching is illustrated in Fig. 1(d), which presents micromagnetic simulations ($T = 300$ K) of thermal switching between two ground states in a 40 nm diameter CoFeB cylinder performed using the MuMax3^{31,32} software package. Changing the strain applied manipulates the dwell time of the magnetisation in the two states: tensile(compressive) strain results in an increased dwell time in the right(left) magnetization state.

Fig. 1(e), plots the probability of not switching as a function of time in a state for the tensile strain case. The decay rate of these exponential distributions gives the dwell time and switching rate for the state. Fig. 1(f) compares rates extracted from the micromagnetic simulations (points) with those derived from energy barriers defined by the macrospin model (Eq. 1). A value of $f_0 = 0.5 \times 10^9$ Hz is selected such that the values of w_{ij} at zero applied strain match between the two models. The macrospin model agrees reasonably with the micromagnetic simulations, however, the data match less well at higher magnitudes of strain. The deviation occurs for values of k_σ (high) and β' (low) where the assumption of Néel-Arrhenius like switching between two states approaches the limit of its validity ($E_b \sim 3k_b T$)³³. We do not calculate micromagnetic rates for all values of β' as the run time required for the simulations scales exponentially. However, we expect the agreement is as good or better for larger β' where the assumption $E_b \gg 3k_b T$ is always met. The macrospin model is, therefore, used to model the CoFeB dots throughout the remainder of this paper.

We now consider how the stochastic behaviour of single nanodots manifests in an extended ensemble as predictable collective behavior. The average magnetization

obeys simple rules that, when subject to a temporally varying input in the form of a globally applied strain, produce a reproducible, complex, nonlinear response with fading memory, thus fulfilling the basic requirements of a hardware-based reservoir.

We model the extended ensemble as a two-state system governed by the master equation³³:

$$\frac{dp_1}{dt} = p_2 \omega_{21} - p_1 \omega_{12} \quad (3)$$

where p_i is the normalized population in state i such that $p_1 + p_2 = 1$, and ω_{ij} is the transition rate as defined in Eq. 2. This model does not take into account the error introduced by a finite system size. However, it is appropriate for systems with a large number of nanodots (as we assume here) since the error scales as $1/\sqrt{n}$. Integrating Eq. 3 gives the time-dependent probability for fixed transition rates:

$$p_1(t) = \underbrace{\frac{w_{21}}{w} [1 - \exp(-wt)]}_{\text{Excitation}} + \underbrace{p_1(0) \exp(-wt)}_{\text{Decay}} \quad (4)$$

where $\omega = \omega_{21} + \omega_{12}$. Given our system with a fixed applied field, the reduced magnetization (unit length) along the x direction (magnetization right) can be written:

$$m_x(t) = p_1(t) \cos(\delta_1(t)) - p_2(t) \cos(\delta_2(t)) \quad (5)$$

where δ_i refers to the slight rotation of the energy minimum away from the x axis due to the fixed field.

Fig. 2(a) plots the equilibrium ($t \rightarrow \infty$) response of reduced magnetization to an applied strain anisotropy (k_σ). The response is nonlinear and while the exact form depends on the scaled thermal energy (β'), it maintains sigmoidal like features that sharpen with increasing $KV/k_B T$. Importantly, the curve remains nonlinear as required for reservoir computing. Fig. 2(b) illustrates the complex response of the reduced magnetisation to a temporal input sequence. Section A, where a single input is turned on and held, demonstrates excitation of the magnetisation away from zero net magnetisation. Section B, where no input is present, demonstrates decay behavior back towards zero net magnetisation. In all other sections, applied strain inputs are held for a duration less than the zero-input base time (T_0) of the system, and thus the system response exhibits both excitation and decay terms. The zero-input timescale scales as $T_0 \propto \exp(0.36\beta')$, a variation from nanoseconds to milliseconds for $\beta' = 10$ to 50, which allows a reservoir to be designed with a timescale to match a particular real-time task (see supplementary for further details). The sharp transitions upon changes of input are due to the slight rotations of the minimum energy states within single dots (change in δ_i) and are expected to occur on a timescale much faster than the thermally driven response of the ensemble. Since $p_i(t)$ depends on $p_i(t - \Delta t)$, the system exhibits a fading memory: the response of the reservoir depends on the current and previous inputs.

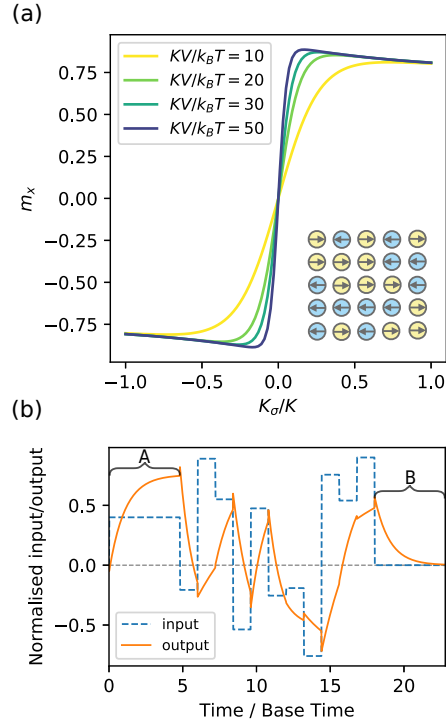


FIG. 2. Collective response from an ensemble of nanodots. (a) The equilibrium magnetization response of the reservoir to input strain anisotropy for $\beta' = 10$ to 50. (b) Reservoir magnetization over time in response to an input sequence. Sections A and B show the excitation and decay responses of magnetisation as described in Eqs. 4 & 5.

Together, the reproducible response, nonlinearity, and fading memory of the superparamagnetic ensembles suggest they have the correct properties to be used as hardware reservoirs. Furthermore, as discussed by Hu *et al.*¹⁷ the estimated power consumption for magnetoelastically driven MRAM is 160 aJ per bit written (a rotation of 90°). Our system uses ~ 1000 such bits, but less than 15% of the strain required for a 90° rotation; this gives an approximate estimate of 24 fJ per input. This compares favorably to the write energy per bit of STT-RAM of 100 fJ¹⁷. While this figure is not precise, it indicates the potential for ultra-low-energy consumption. Furthermore, Safranski *et al.*³⁴ have verified that uncorrelated thermally-driven switching can be observed in CoFeB MTJs for dwell times of 5 ns and above, which we can take as a floor for the operational speed of these nanodots. Longer dwell times can easily be induced by scaling up the size of the magnetic elements, thus increas-

ing the energy barrier between states. Therefore, the timescale of the ensembles' responses could be easily engineered across many orders of magnitude to allow them to perform a wide range of real-time tasks.

Having established that strain-mediated, voltage-controlled superparamagnetic ensembles possess the qualities required for RC, we now test our system's performance against two common ML tasks. We treat the nanodot array as a single dynamical node with complex internal behavior and thus adopt the approach described by Appeltant *et al.*⁵ (Fig. 3(a)–(d)).

The RC process is as follows. A raw input signal with N_{in} dimensions is discretized such that $\mathbf{S}(n\delta t) = \mathbf{S}_n$ giving a 2d matrix of size $N_{samples}$ by N_{in} . For the training part, each input sample has a corresponding desired output \mathbf{y}_n of dimension N_{out} , forming a matrix of $N_{samples}$ by N_{out} . A mask matrix is used to project a random combination of the input dimensions on to each of the N_v virtual neurons and is constant for all samples of the input. For a 1d input signal, such as the NARMA10 task (Fig. 3(b)), the random nature of the mask is designed to stimulate a range of dynamics. For a sample index n and virtual neuron i the total input voltage is:

$$v_{ni} = \Delta v \sum_{j=1}^{N_{in}} S_{nj} M_{ij} + \gamma x_{(n-1)i}, \quad (6)$$

where Δv represents a scaling factor to give an input voltage into the reservoir, γ is the feedback strength and $x_{(n-1)i}$ the output value of the reservoir for virtual neuron i but for sample $n-1$.

This matrix of input voltages is serialized as 1d time series with each virtual node presented in order before the applying the next input sample. This is demonstrated in Fig. 3(c), which shows the masked input, masked input with added feedback term, and finally the raw output of the reservoir at the transition point between samples 56 and 57. The masked input voltage for each virtual neuron is held constant for the neuron duration θ , thus the reservoir time between each input sample is $\zeta = N_v \theta$. The raw output of the reservoir is measured at the end of each virtual neuron duration to give a transformed value, which for the n th input sample and virtual neuron i is

$$x_{ni} = m_x(v_{ni}) \quad (7)$$

The aim of RC is that transformed values from e.g. different input classes are now linearly separable by a hyperplane due to the higher dimensional transformation. Thus, a set of linear output weights are used to classify the output. The predicted output of sample n and output dimension k is

$$\tilde{y}_{nk} = \sum_{j=0}^{N_v} x_{nj} W_{kj}, \quad (8)$$

where W_{kj} is the output weight matrix (Fig. 3(d)). A constant bias term is included as the $j=0$ element of

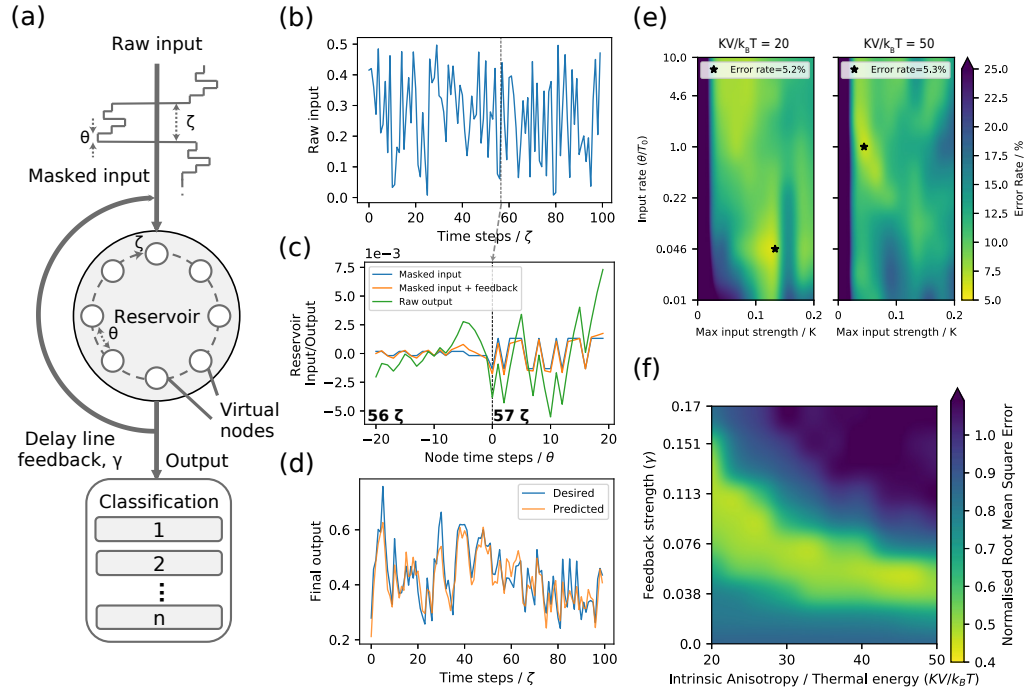


FIG. 3. Reservoir computing using the superparamagnetic ensemble. (a) Schematic of reservoir computing. (b) Raw input to the system; here, random numbers from the interval $[0, 0.5]$ for the NARMA10 task. (c) A single input value from (b) is transformed by the mask into masked inputs. A snapshot is shown from just before to just after the raw data time step 57ζ . Feedback is then added: the output from each virtual node one time step (c) previously is multiplied by a feedback strength and added to the masked inputs. The raw output is the response of the reservoir to the masked input with feedback. (d) Output after trained linear layer compared to desired output. Here, the desired NARMA10 output from Eq. 10 is compared to the output provided by the trained reservoir. (e) Machine learning performance on speech recognition task. (f) Performance on NARMA10 task.

the weights with an extra column in x as $x_{n0} = 1$ for all n . The output weights are found by using ridge (or Tikhonov) regression.³⁵ This gives a direct solution for the weights as

$$\mathbf{W} = \mathbf{y}^T \mathbf{x} (\mathbf{x}^T \mathbf{x} + \lambda \mathbf{I})^{-1}, \quad (9)$$

where λ is the L2-norm regularisation hyper-parameter and \mathbf{I} the identity matrix. λ is found through a grid search to find the lowest mean-square error which is sampled using 5-fold cross validation of the training data set for each task to prevent overfitting.

We use two common machine learning tasks as proof of principle of the reservoir's operation: spoken digit recognition (TI-46 dataset) and chaotic time series prediction (NARMA10). These require a reservoir with good non-linearity and memory respectively. Performance heat maps for the spoken digit task are shown in Fig. 3(e).

The task requires the recognition of the spoken digits 0–9 by five female speakers from the NIST TI-46 dataset³⁶. Preprocessing separates the data into 13 frequency bands using a MFCC filter³⁷ before it is masked. Training and validation was performed on 100 utterances per speaker, testing on 160. The input rate (θ/T_0) and the scaling of the input (maximum k_σ) are varied and the error rate for the test dataset is plotted. 50 virtual nodes was found to give good performance. The left hand heat map is for an ensemble with a base timescale $T_0 = 1.4 \mu\text{s}$ ($KV/k_B T = 20$, nanodot diameter 57 nm), the right hand heat map for $T_0 = 66 \text{ ms}$ ($KV/k_B T = 50$, nanodot diameter 90 nm). In both cases, an error rate of 5% can be achieved, showing that this reservoir can perform well at this task across a range of timescales. To provide a baseline, the error rate when the reservoir is removed and the masked input is treated as the output is 23%. The relatively high base level is due to the inherent nonlin-

earity in the preprocessing discussed by Abreu Araujo *et al.*³⁸. However, the important demonstration here is the increase in performance with the reservoir, which is comparable to other hardware based reservoirs^{5,8,38}. The complexity of the maps, while not discussed in further detail here, may be due in part to coupling between input strength and internal timescale (see supplementary).

The goal of the chaotic time series task (NARMA10) is to predict the response to white noise (random numbers drawn from $[0, 0.5]$) of a discrete-time tenth order nonlinear auto-regressive moving average^{39,40}:

$$y_n = y_{n-1} \left(0.3 + 0.05 \sum_{k=1}^{10} y_{n-k} \right) + 1.5S_{n-1}S_{n-10} + 0.1. \quad (10)$$

The task for the network is to predict the NARMA10 output y_n given the input S_n . We can characterize the success of the network via the metric, Normalized Root Mean Squared Error (NRMSE):

$$\text{NRMSE} = \sqrt{\frac{1}{N_s} \frac{\sum_n (\hat{y}_n - y_n)^2}{\text{Var}(y)}}, \quad (11)$$

where y_n is the true output (Eq. 10) and \hat{y}_n the predicted output from the network (Eq. 8). For this task, we make use of the delay line feedback (Fig. 3(a)) similar to the Mackey-Glass reservoir discussed by Appeltant *et al.*⁵. This delay line (an addition of the output at time $t - \zeta$ to the input at time t) serves to boost the memory capacity of the system; an important requirement for the NARMA10 task. Fig. 3(f) presents a heat map where NRMSE is plotted for varying feedback strength (γ) and intrinsic timescale (change in $KV/k_B T$). Previous optimisation lead to the use of 400 virtual nodes, input rate $\theta/T_0 = 0.3$, and input scaling of 0.003. 2000 samples were used for training and validation (10-fold), 1000 for testing. It can be seen that good performance (around $\text{NRMSE} = 0.42$, c.f Ref. 39) can be maintained from the μs to ms regime by varying the feedback strength. Fig. 3(d) demonstrates a typical predicted and true signal for this level of NRMSE. As with the speech recognition task, we can compare to the value achieved when the reservoir is removed and the masked input taken as the output to be trained: a $\text{NRMSE} \approx 1$. Using a Mackey-Glass reservoir (with optimisation parameters based on those in Appeltant *et al.*⁵) a value of 0.50 was achieved. It should be noted that this value differs from that given by Appeltant; this is believed to be due to a differing definition of NRMSE.

While we have demonstrated the feasibility of this system for machine learning, we now address two of the key potential challenges in realising it as an experimental device. Firstly, reproducibility of the magnetic behaviour. Whether due to deviations from the Stoner-Wohlfarth model, small perturbations from long-range magnetostatic coupling, or variation in PMN-PT response from nanodot to nanodot, the magnetic response curve may differ from

that presented in Fig. 2. However, one of the chief advantages of the system envisaged is that the behaviour does not have to be realised exactly in order to produce the behaviour required for machine learning, as long as the response of the system is nonlinear, reproducible and has fading memory. As the output is derived from an average across an ensemble, slight changes in behaviour from nanodot to nanodot are averaged out across the whole array. This should give the device high tolerance to small deviations. This is in contrast to typical magnetic logic and memory devices where slight device to device variations can cause significant problems with reproducibility. We have also demonstrated, in Fig. 3(e), that high performance can be achieved for systems with differing magnetization response curves. This leads to the second consideration: how reproducible the device will be run-to-run. The reproducibility of the signal will largely depend on the size of the array, with the error scaling as $1/\sqrt{n}$. By increasing the size of the array, the error can be made arbitrarily small. It will be important to determine how large an error can be tolerated, and this will be an avenue of future research. We anticipate that the threshold of the noise tolerance will be problem specific. The tolerance of the system to changes in magnetic response and the increase in reproducibility due to the aggregate nature of the device should help to mitigate the key challenges of realizing a physical device.

In conclusion, we have proposed strain-mediated voltage-controlled superparamagnetic ensembles as platforms for hardware-based reservoir computing. We have demonstrated competitive outcomes on two machine learning benchmark tasks (spoken digit recognition and chaotic time series prediction) as well as explaining the properties of the ensemble which give rise to their ability to act as a reservoir. By tuning internal parameters (input rate, input scaling, and feedback strength), this performance can be delivered across a range of timescales, from μs to ms . This would allow a physical realization of such a reservoir to be tuned to provide computation in real time for a wide range of possible physical inputs: from decision-making in driverless cars (fast) to speech recognition (slow). The simplicity of the system, coupled with the low energy consumption for such a voltage-controlled, thermally-driven device, makes it an ideal candidate for use in edge computing applications where high performance is needed at low latency and power.

SUPPLEMENTARY MATERIAL

See the supplementary material for a discussion of the full form of the anisotropy, and the tunable behaviour of the timescales.

ACKNOWLEDGMENTS

The authors thank the Engineering and Physical Sciences Research Council (grant numbers: EP/S009647/1, EP/V006339/1) for financial support. The project leading to this application has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 861618 (SpinENGINE).

DATA AVAILABILITY

The data that support the findings of this study are openly available in ORDA at <http://doi.org/10.15131/shef.data.14130452>.

- ¹E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in nlp," (2019), [arXiv:1906.02243 \[cs.CL\]](https://arxiv.org/abs/1906.02243).
- ²"Big data needs a hardware revolution," *Nature* **554**, 145–146 (2018).
- ³G. Bellec, F. Scherr, A. Subramoney, E. Hajek, D. Salaj, R. Legenstein, and W. Maass, "A solution to the learning dilemma for recurrent networks of spiking neurons," *Nature Communications* **11**, 3625 (2020).
- ⁴H. Jaeger and H. Haas, "Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication," *Science* **304**, 78–80 (2004).
- ⁵L. Appeltant, M. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. Mirasso, and I. Fischer, "Information processing using a single dynamical node as complex system," *Nature Communications* **2**, 468 (2011).
- ⁶G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose, "Recent advances in physical reservoir computing: A review," *Neural Networks* **115**, 100–123 (2019).
- ⁷J. Grollier, D. Querlioz, K. Y. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, "Neuromorphic spintronics," *Nature Electronics* **3**, 360–370 (2020).
- ⁸J. Torrejon, M. Riu, 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**, 428–431 (2017).
- ⁹T. Furuta, K. Fujii, K. Nakajima, S. Tsunegi, H. Kubota, Y. Suzuki, and S. Miwa, "Macromagnetic Simulation for Reservoir Computing Utilizing Spin Dynamics in Magnetic Tunnel Junctions," *Physical Review Applied* **10**, 034063 (2018).
- ¹⁰D. Prychynenko, M. Sitte, K. Litzius, B. Krüger, G. Bourianoff, M. Kläui, J. Sinova, and K. Everschor-Sitte, "Magnetic Skyrmion as a Nonlinear Resistive Element: A Potential Building Block for Reservoir Computing," *Physical Review Applied* **9**, 014034 (2018).
- ¹¹H. Nomura, T. Furuta, K. Tsujimoto, Y. Kuwabiraki, F. Peper, E. Tamura, S. Miwa, M. Goto, R. Nakatani, and Y. Suzuki, "Reservoir computing with dipole-coupled nanomagnets," *Japanese Journal of Applied Physics* **58**, 070901 (2019).
- ¹²J. H. Jensen and G. Tufte, "Reservoir Computing in Artificial Spin Ice," in *The 2020 Conference on Artificial Life* (MIT Press, Online, 2020) pp. 376–383.
- ¹³M. Dale, R. F. L. Evans, S. Jenkins, S. O'Keefe, A. Sebald, S. Stepney, F. Torre, and M. Trefzer, "Reservoir Computing with Thin-film Ferromagnetic Devices," [arXiv:2101.12700 \[cond-mat\]](https://arxiv.org/abs/2101.12700) (2021).
- ¹⁴R. W. Dawidek, T. J. Hayward, I. T. Vidamour, T. J. Broomhall, G. Venkat, M. A. Mamoori, A. Mullen, S. J. Kyle, P. W. Fry, N.-J. Steinke, J. F. K. Cooper, F. Maccherozzi, S. S. Dhessi, 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*, 2008389.
- ¹⁵H. Raanaei, H. Nguyen, G. Andersson, H. Lidbaum, P. Korelis, K. Leifer, and B. Hjörvarsson, "Imprinting layer specific magnetic anisotropies in amorphous multilayers," *Journal of Applied Physics* **106**, 023918 (2009).
- ¹⁶A. T. Hindmarch, A. W. Rushforth, R. P. Campion, C. H. Marrows, and B. L. Gallagher, "Origin of in-plane uniaxial magnetic anisotropy in CoFeB amorphous ferromagnetic thin films," *Physical Review B* **83**, 212404 (2011).
- ¹⁷J.-M. Hu, Z. Li, L.-Q. Chen, and C.-W. Nan, "High-density magnetoresistive random access memory operating at ultralow voltage at room temperature," *Nature Communications* **2**, 553 (2011).
- ¹⁸T. Wu, P. Zhao, M. Bao, A. Bur, J. L. Hockel, K. Wong, K. P. Mohanchandra, C. S. Lynch, and G. P. Carman, "Domain engineered switchable strain states in ferroelectric (011) [Pb(Mg_{1/3}Nb_{2/3})O₃](1-x)-[PbTiO₃]_x (PMN-PT, x=0.32) single crystals," *Journal of Applied Physics* **109**, 124101 (2011), ISBN: 0021-8979.
- ¹⁹J. Wang, D. Pesquera, R. Mansell, S. van Dijken, R. P. Cowburn, M. Ghidini, and N. D. Mathur, "Giant non-volatile magnetoelectric effects via growth anisotropy in Co₄₀Fe₄₀B₂₀ films on PMN-PT substrates," *Applied Physics Letters* **114**, 092401 (2019).
- ²⁰E. C. Stoner and E. P. Wohlfarth, "A mechanism of magnetic hysteresis in heterogeneous alloys," *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **240**, 599–642 (1948).
- ²¹R. P. Cowburn, D. K. Koltsov, A. O. Adeyeye, M. E. Welland, and D. M. Tricker, "Single-Domain Circular Nanomagnets," *Physical Review Letters* **83**, 1042–1045 (1999).
- ²²D. Wang, C. Nordman, Z. Qian, J. M. Daughton, and J. Myers, "Magnetostriction effect of amorphous CoFeB thin films and application in spin-dependent tunnel junctions," *Journal of Applied Physics* **97**, 10C906 (2005).
- ²³Y.-T. Chen and C. C. Chang, "Effect of grain size on magnetic and nanomechanical properties of Co₆₀Fe₂₀B₂₀ thin films," *Journal of Alloys and Compounds* **498**, 113–117 (2010).
- ²⁴B. Cullity and C. D. Graham, *Introduction To Magnetic Materials* (John Wiley & Sons, Inc., 2009).
- ²⁵Z. Guo, X. Yang, X. Liu, J. Ou-Yang, B. Zhu, S. Chen, and Y. Zhang, "Electric field induced non-90 rotation of the easy axis of a ferromagnetic film," *Applied Physics Letters* **112**, 052904 (2018).
- ²⁶A. Aharoni, *Introduction to the Theory of Ferromagnetism* (Clarendon Press, 2000).
- ²⁷D. Vodenicarevic, N. Locatelli, A. Mizrahi, J. Friedman, A. Vincent, M. Romera, A. Fukushima, K. Yakushiji, H. Kubota, S. Yuasa, S. Tiwari, J. Grollier, and D. Querlioz, "Low-Energy Truly Random Number Generation with Superparamagnetic Tunnel Junctions for Unconventional Computing," *Physical Review Applied* **8**, 054045 (2017).
- ²⁸B. Parks, M. Bapna, J. Igbokwe, H. Almasi, W. Wang, and S. A. Majetich, "Superparamagnetic perpendicular magnetic tunnel junctions for true random number generators," *AIP Advances* **8**, 055903 (2017).
- ²⁹W. A. Borders, A. Z. Pervaiz, S. Fukami, K. Y. Camsari, H. Ohno, and S. Datta, "Integer factorization using stochastic magnetic tunnel junctions," *Nature* **573**, 390–393 (2019).
- ³⁰A. Mizrahi, T. Hirtzlin, A. Fukushima, H. Kubota, S. Yuasa, J. Grollier, and D. Querlioz, "Neural-like computing with populations of superparamagnetic basis functions," *Nature Communications* **9**, 1533 (2018).
- ³¹A. Vansteenkiste, J. Leliaert, M. Dvornik, M. Helsen, F. Garcia-Sanchez, and B. Van Waeyenberge, "The design and verification of MuMax3," *AIP Advances* **10**, 107133 (2014), 10.1063/1.4899186.

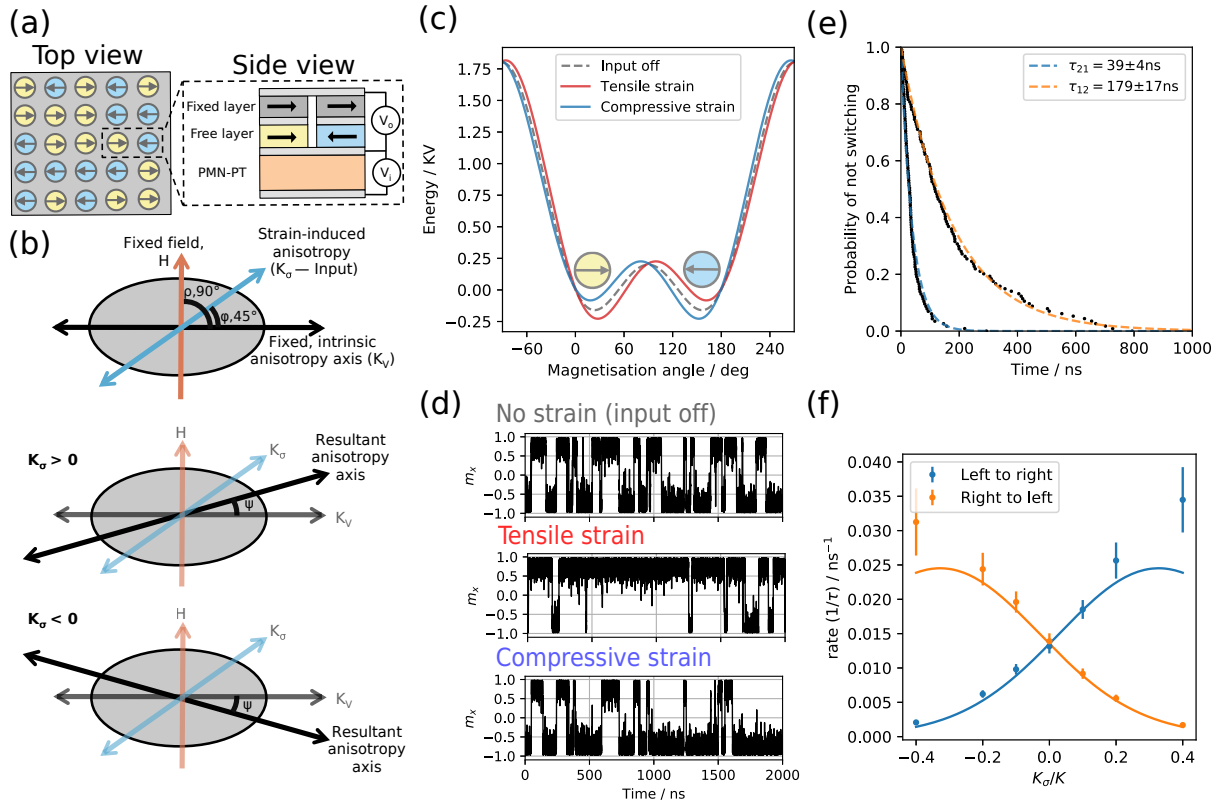
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0048911

- ³²J. Leliaert, J. Mulkers, J. De Clercq, A. Coene, M. Dvornik, and B. Van Waeyenberge, "Adaptively time stepping the stochastic Landau-Lifshitz-Gilbert equation at nonzero temperature: Implementation and validation in MuMax3," *AIP Advances* **7**, 125010 (2017).
- ³³R. W. Chantrell, N. Walmsley, J. Gore, and M. Maylin, "Calculations of the susceptibility of interacting superparamagnetic particles," *Physical Review B* **63**, 024410 (2000).
- ³⁴C. Safranski, J. Kaiser, P. Trouilloud, P. Hashemi, G. Hu, and J. Z. Sun, "Demonstration of nanosecond operation in stochastic magnetic tunnel junctions," [arXiv:2010.14393 \[cond-mat\]](https://arxiv.org/abs/2010.14393) (2020).
- ³⁵M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review* **3**, 127 – 149 (2009).
- ³⁶M. Liberman, R. Amsler, K. Church, E. Fox, C. Hafner, J. Klavans, M. Marcus, B. Mercer, J. Pedersen, P. Roossin, D. Walker, S. Warwick, and A. Zampolli, "TI 46-Word," [10.35111/ZX7A-FW03](https://arxiv.org/abs/10.35111/ZX7A-FW03) (1993).
- ³⁷J. Lyons, D. Y.-B. Wang, Gianluca, H. Shteingart, E. Mavrinac, Y. Gaurkar, W. Watcharawisetkul, S. Birch, L. Zhihe, J. Hölzl, J. Lesinski, H. Almér, C. Lord, and A. Stark, "james-lyons/python-speech-features: release v0.6.1," (2020).
- ³⁸F. Abreu Araujo, M. Riou, J. Torrejon, S. Tsunegi, D. Querlioz, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, and J. Grollier, "Role of non-linear data processing on speech recognition task in the framework of reservoir computing," *Scientific Reports* **10**, 328 (2020).
- ³⁹L. Manneschi, M. O. A. Ellis, G. Gigante, A. C. Lin, P. Del Giudice, and E. Vasilaki, "Exploiting multiple timescales in hierarchical echo state networks," *Frontiers in Applied Mathematics and Statistics* **6**, 76 (2021).
- ⁴⁰A. Atiya and A. Parlos, "New results on recurrent network training: unifying the algorithms and accelerating convergence," *IEEE Transactions on Neural Networks* **11**, 697–709 (2000).

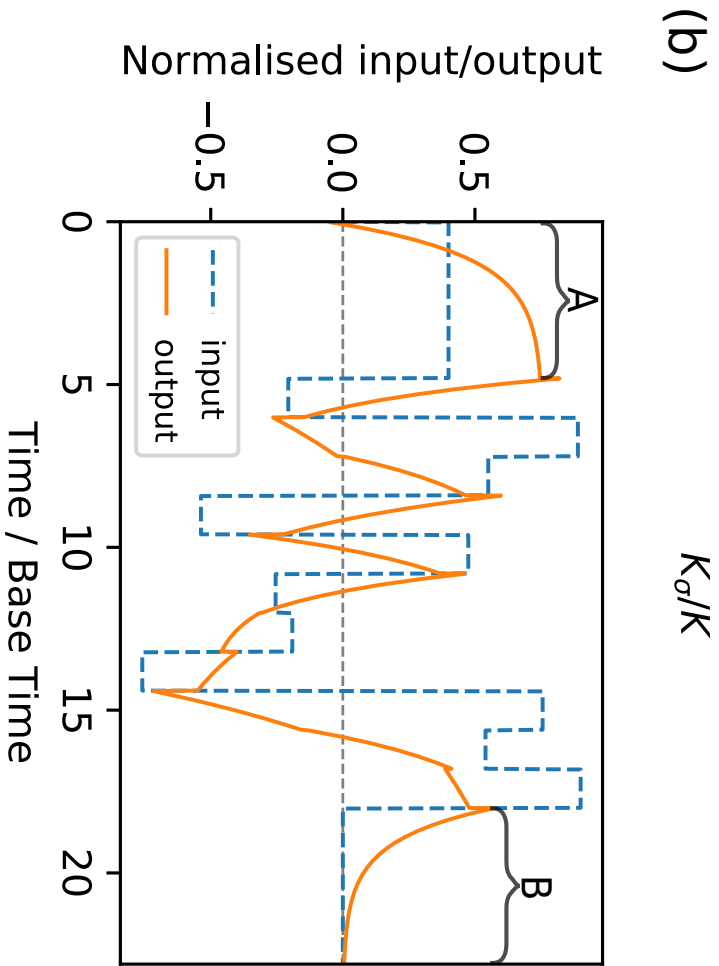
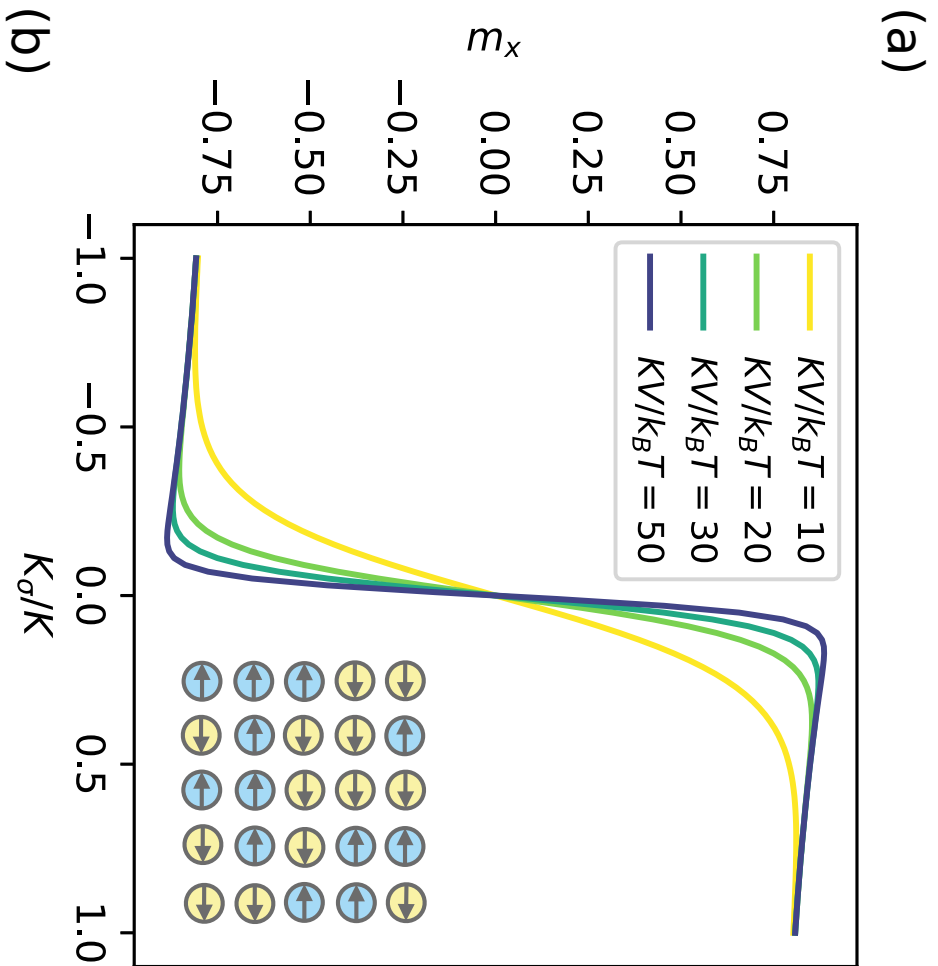
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0048911



This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0048911



This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0048911

