Science Advances

MAAAS

Supplementary Materials for

Magnon Scattering Modulated by Omnidirectional Hopfion Motion in Antiferromagnets for Meta-Learning

Zhizhong Zhang *et al.*

Corresponding author: Yue Zhang, yz@buaa.edu.cn

The PDF file includes:

Sections S1 to S5
Figs. S1 to S10
Legends for movies S1 to S7
References

Other Supplementary Material for this manuscript includes the following:

Movies S1 to S7

I. STABILIZING A HOPFION IN AN ANTIFERROMAGNET

The energy minimization integrated in MuMax3[41] is employed to obtain hopfions in the lowest energy state. Material parameters and custom field settings are presented in Methods attached to the main text. The initial configuration of the Neel field is set to the following formulae in the form of cylindrical coordinates (ρ, φ, z)

$$\Theta = 2 \arcsin \rho^2 e^{1-\rho^2-z^2} \tag{S1a}$$

$$\Phi = P \operatorname{Im}\left(\ln \frac{\rho + iz + 1}{\rho + iz - 1}\right) + Q\varphi \tag{S1b}$$

where $i = \sqrt{-1}$, Θ and Φ represent respectively the polar angle and the azimuthal angle of spin orientation at each point, P and Q are the numbers of times the spin rotates after running around a meridian circle or a parallel circle. The H-index of the hopfion is calculated using the following procedure[58] and turns out to be $P \times Q$.

1. Calculate the fictitious magnetic field $\boldsymbol{F}$.

$$F_i = \epsilon_{ijk} \boldsymbol{n} \cdot \left(\partial_j \boldsymbol{n} \times \partial_k \boldsymbol{n}\right) \tag{S2}$$

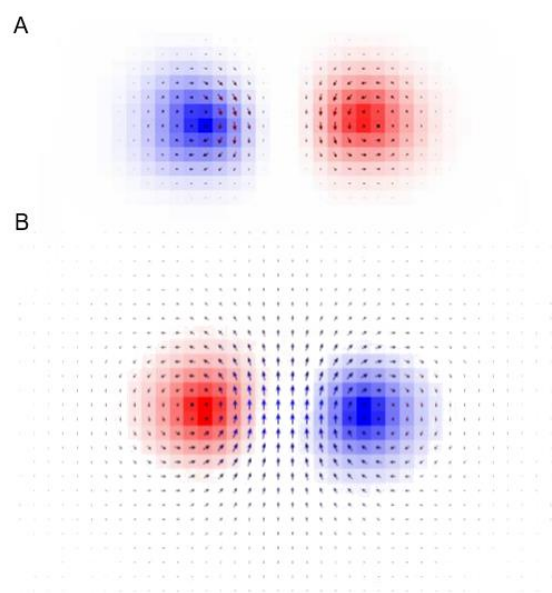2. Write out the fictitious vector potential $\boldsymbol{A}$ according to the spin configuration of the hopfion.

$$\boldsymbol{A} = \begin{bmatrix} -\dfrac{4}{\rho}\sin^2\dfrac{\Theta}{2}\sin\varphi\,\partial_\varphi\Phi - 4\cos^2\dfrac{\Theta}{2}\cos\varphi\,\partial_\rho\Phi \\[2mm] -\dfrac{4}{\rho}\sin^2\dfrac{\Theta}{2}\cos\varphi\,\partial_\varphi\Phi - 4\cos^2\dfrac{\Theta}{2}\sin\varphi\,\partial_\rho\Phi \\[2mm] -4\cos^2\dfrac{\Theta}{2}\partial_z\Phi \end{bmatrix} \tag{S3}$$

3. Calculate the Hopf index by integrating the inner product of $\boldsymbol{F}$ and $\boldsymbol{A}$ over the whole space.

$$H = -(8\pi)^{-2} \int \boldsymbol{F} \cdot \boldsymbol{A}\, dV \tag{S4}$$

where $\boldsymbol{n} = (\sin\Theta\cos\Phi, \sin\Theta\cos\Phi, \cos\Theta)$ is the Neel field, $\boldsymbol{F} = \sum_i F_i \hat{e}_i$ $(i = x, y, z)$ is the fictitious magnetic field, and $\boldsymbol{A}$ is the vector potential of $\boldsymbol{F}$ which is subject to $\nabla \times \boldsymbol{A} = \boldsymbol{F}$. The calculation of $\boldsymbol{A}$ is given in cylindrical coordinates (ρ, θ, φ). The formula of $\boldsymbol{A}$ is given in Ref. 58. The yz-profiles of the field $\boldsymbol{F}$ and $\boldsymbol{A}$ are shown in Fig. S1.

**Fig. S1** *yz*-profile of the fictitious magnetic field and its vector potential.

(**A**) Profile of the fictitious magnetic field $\boldsymbol{F}$. (**B**) Profile of the vector potential $\boldsymbol{A}$ of F. Arrows represent the *y*- and *z*-component of spin. Colors from blue, white, to red represents the *x*-component of spin from -1, 0 to 1.

The lowest energy states of H-1 (P=1, Q=1) and H-2 (P=1, Q=2) hopfions are thus obtained and shown in Fig. 1 in the main text, from which can be seen the characteristic entanglement between preimages of spins. Although the hopfions may be deformed or

reoriented by the high-order exchange interaction[33], which is anisotropic and non-quadratic in our system, these well-protected topologically non-trivial structures still demonstrate undoubtedly the existence and stability of hopfions in AFM systems.

Magnetic frustration and anisotropy both play important roles in stabilizing hopfions. As previously demonstrated, frustrated exchange interactions allow magnetization to twist spatially and form a knot-like structure[30-31]. It is noteworthy that the anisotropy energy is as necessary as frustration for hopfions' existence. In a system with chiral exchange instead of frustrated exchange, hopfions stretch themselves along the easy-axis direction until they meet the boundary. Without the protection from the anisotropy energy barrier at the boundary, a magnetic switching would occur at the boundary and cause the annihilation of hopfions. According to our simulation results, the problem of instability also arises in a frustrated system without anisotropy. The dynamical process is visualized by showing the evolution of the *x*-component of the Neel field (see Fig. S2). The exchange interaction breaks the symmetry of the initial spin configuration, exacerbates the asymmetrical distribution, and finally causes the annihilation of the hopfion. It should be noted that the symmetry breaking originates from the spin alignment at the boundary and is possible to be prevented as long as the boundary vanishes. Therefore, an anisotropy energy, no matter if it is cubic or uniaxial, should be introduced into the system so that the boundary is replaced with a *de facto* point at infinity. Besides, the introduction of anisotropy will lift a restriction on the size of the hopfion to the order of magnitude of the domain wall thickness, making the hopfion both rigid and compact.

Fig. S2 Annihilation of a hopfion in a frustrated magnet without anisotropy energies
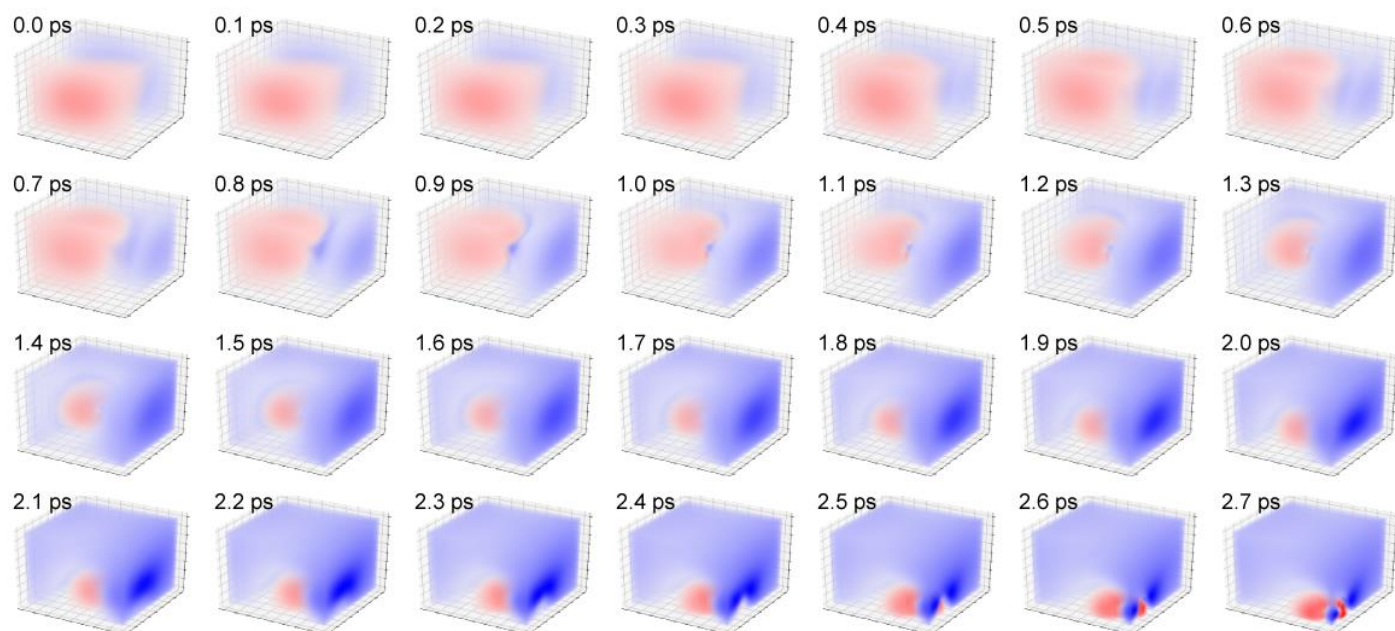
The color from blue, white, to red represents the x-component of spin from -1, 0, to 1.

The source code with frustrated exchange interactions[42] to stabilize a hopfion is shown below.

```
// This code is published in [V. M. Kuchkin, K. Chichay, B. Barton-Singer, F. N. Rybakov, S. Blügel, B. J. Schroers,
N. S. Kiselev, Geometry and symmetry in skyrmion dynamics. Physical Review B 104, 165116 (2021)].


// Simulation Settings
CellSize      := 0.5e-9
SetGridSize(100, 100, 100)
SetCellSize(CellSize, CellSize, CellSize)
OpenBC        = true                // Open Boundary Condition
EnableDemag   = false               // No Demagnetization Field

// Material Parameters
Ms            := 1.51e5             // Saturation Magnetization
Msat          = Ms
A             := 2e-12              // Exchange Stiffness of Regular Exchange
Aex           = -A
Ku1           = 1e4                 // Uniaxial Anisotropy
anisU         = vector(0, 0, 1)
alpha         = 0.1                 // Damping Coefficient

// Arithmetic Expressions for Expressing High-order Exchange Interaction
mx2           := Shifted(m,2,0,0)
mx_2          := Shifted(m,-2,0,0)
my2           := Shifted(m,0,2,0)
```

```
my_2        := Shifted(m,0,-2,0)
mz2         := Shifted(m,0,0,2)
mz_2        := Shifted(m,0,0,-2)
laplacian2  := Add(Add(Add(Add(Add(mx2, mx_2), my2), my_2), mz2), mz_2)


// Effective Field and Energy Density of the High-order Exchange
Bex         := A*0.248                  // Exchange Stiffness of High-order Exchange
BField      := Mul( Const(-2.0/Ms * Bex/(CellSize*CellSize)), laplacian2)
BEdens      := Mul( Const(Bex/(CellSize*CellSize)), Dot(laplacian2, m))
AddFieldTerm(BField)
AddEdensTerm(BEdens)
```

II. SPIN WAVE CONFIGURATION FOR OMNIDIRECTIONAL MOTION IN 3D SPACE

In this section, we prove that the spin-wave configuration in Eqs. 8A and 8B in the main text results in a velocity with the azimuthal angle ϕ and the polar angle θ. The two hypotheses shown below are used in the proof.

1. CCW-C and CW-C waves will be scattered towards opposite transverse directions after crossing the hopfion, causing the hopfion to move along opposite transverse directions.

2. The driving effect of spin waves is a summation of the driving effects of two kinds of C-Waves, and the strength of each effect is proportional to the intensity of each C-wave component.

The first hypothesis embodies the magnon scattering origin of hopfions' motion. It is concluded empirically from Fig. 2B in the main text and is strictly demonstrated by Section III in this Supplementary Materials. The second hypothesis is verified by the omnidirectional motion, which is realized by the spin wave configuration derived here.

First, according to Hypothesis 1, the spin waves consisting of two L-waves with a phase offset should be decomposed into a linear combination of two C-waves which contribute a pair of velocities in different directions. Therefore, an oscillating spin in the incoming waves should be decomposed as

$$\psi = A\cos\omega t + iB\cos(\omega t + \varphi) = A_1 e^{i\omega t} + A_2 e^{-i\omega t} \tag{S5}$$

where

$$|A_1|^2 = \frac{1}{4}A^2 - \frac{1}{2}AB\sin\varphi + \frac{1}{4}B^2 \tag{S6a}$$

$$|A_2|^2 = \frac{1}{4}A^2 + \frac{1}{2}AB\sin\varphi + \frac{1}{4}B^2 \tag{S6b}$$

Second, according to Hypothesis 2, the intensities of C-waves should be regarded as the magnitudes of these two velocities whose vector sum is consequently the hopfion's velocity. This indicates that the longitudinal velocity of the hopfion is proportional to $|A_1|^2 + |A_2|^2$, while the transverse velocity is proportional to $|A_1|^2 - |A_2|^2$. As a result,

$$v_\parallel = \frac{1}{2}v_\parallel^0(A^2 + B^2) \tag{S7a}$$

$$v_\perp = v_\perp^0 AB\sin\varphi \tag{S7b}$$

where $v_{\parallel(\perp)}$ stands for the longitudinal (transverse) velocity, and $v_{\parallel(\perp)}^0$ is the scaling factor subject to $\tan\alpha_{\text{Hall}}^0 = v_\perp^0/v_\parallel^0$.

By the way, Eq. 4 in the main text is derived by letting $A = e^\theta$ and $B = e^{-\theta}$

$$v_\parallel = v_\parallel^0\cos 2\theta \tag{S8a}$$

$$v_\perp = v_\perp^0\sin\varphi \tag{S8b}$$

which leads to

$$\frac{\tan\alpha_{\text{Hall}}}{\tan\alpha_{\text{Hall}}^0} = \frac{\sin\varphi}{\cos 2\theta} \tag{S9}$$

The motion of the hopfion with a given velocity is treated as a compound motion driven by forces originating from two spin-wave sources. Here, the scope of Hypothesis 2 is extended from a pair of C-Waves to a pair of spin-wave sources.

Let $A = \cos(\theta/2)$, $B = \sin(\theta/2)\tan(\vartheta/2)$ and $\varphi = \varphi_1$. We obtain the velocity for the first source

$$v_{\parallel,1} = \frac{1}{2}v_\parallel^0\left(\cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2}\tan^2\frac{\vartheta}{2}\right) \tag{S10a}$$

$$v_{\perp,1} = \frac{1}{2}v_\perp^0\sin\theta\tan\frac{\vartheta}{2}\sin\varphi_1 \tag{S10b}$$

Let $A = \sin(\theta/2)$, $B = \cos(\theta/2)\tan(\vartheta/2)$ and $\varphi = \varphi_2$. We obtain the velocity for the second source

$$v_{\parallel,2} = \frac{1}{2}v_{\parallel}^0 \left(\sin^2\frac{\theta}{2} + \cos^2\frac{\theta}{2}\tan^2\frac{\vartheta}{2}\right) \tag{S11a}$$

$$v_{\perp,2} = \frac{1}{2}v_{\perp}^0 \sin\theta \tan\frac{\vartheta}{2}\sin\varphi_2 \tag{S11b}$$

Note that the two $v_{\perp}$'s are not in the same direction but form an angle of $135°(= \varphi_{\text{Hall}})$, and the two $v_{\perp}$'s are in the opposite direction of each other. As a result, the total force is

$$v_x = \frac{1}{2}v_{\parallel}^0 \cos\theta \left(1 - \tan^2\frac{\vartheta}{2}\right) \tag{S12a}$$

$$v_y = -\frac{1}{2}v_{\perp}^0 \sin\theta \tan\frac{\vartheta}{2}\sin\frac{\varphi}{2}(\sin\varphi_1 - \sin\varphi_2) \tag{S12b}$$

$$v_z = -\frac{1}{2}v_{\perp}^0 \sin\theta \tan\frac{\vartheta}{2}\cos\frac{\varphi}{2}(\sin\varphi_1 + \sin\varphi_2) \tag{S12c}$$

Let $\varphi_1 = \phi + \varphi_{\text{Hall}}/2$, $\varphi_2 = -\phi + \varphi_{\text{Hall}}/2$ and ϑ subject to $\cot\vartheta = (1/2)\sin\varphi_{\text{Hall}}\tan\alpha_{\text{Hall}}^0$. We prove that Eqs. 8A and 8B will lead to the velocity as follows

$$v_x = \frac{1}{2}v^0 \cos\theta \tag{S13a}$$

$$v_y = -\frac{1}{2}v^0 \sin\theta \sin\phi \tag{S13b}$$

$$v_z = -\frac{1}{2}v^0 \sin\theta \cos\phi \tag{S13c}$$

where

$$v^0 = v_{\perp}^0 \tan\frac{\vartheta}{2}\sin\phi = v_{\perp}^0 \left(1 - \tan^2\frac{\vartheta}{2}\right) \tag{S14}$$

Since ϑ is an intrinsic parameter of hopfions' motion and is solely determined by the geometric parameters of the Hall plane, it is referred to as θ_{Hall} in the main text.

III. LINEAR MODEL OF MAGNON SCATTERING IN FRUSTRATED MAGNET

A linear model is proposed to describe the magnon scattering by hopfions, and it is capable of predicting the magnon-driven velocity of any kind of magnetic solitons, especially hopfions with high H-indices. This model will reveal the particle-like aspect of spin waves, from which they are called "magnons". The formulation consists of two major steps. In the first step, spin waves in frustrated AFMs will be described as classical particles in the phase space. We will write the dynamic equation of Neel fields in frustrated AFMs, derive the dynamic equation of spin waves in the complex form, and finally develop them into the form of particle-like equations. In the second step, the driving force is calculated from the momentum exchange between these particles and the soliton.

The construction of our model starts from the Lagrangian density of AFMs written in terms of the Neel field[51-52]

$$\mathcal{L} = \frac{\rho}{2}\dot{\boldsymbol{n}}^2 - \frac{A}{2}\sum_{\mu=x,y,z}(\partial_\mu\boldsymbol{n})^2 + \frac{B}{4}\sum_{\mu=x,y,z}(\partial_\mu^2\boldsymbol{n})^2 + \frac{K}{2}n_z^2 \tag{S15}$$

where $\boldsymbol{n}$ represents the Neel field, $\rho = J^2\chi/M^2$ quantifies its inertia, and the last three terms comprise the potential energy density. Using the Rayleigh dissipation function $\mathcal{D} = \alpha\dot{\boldsymbol{n}}^2/2$, we have the magnetic dynamic equation if the contribution from the anisotropy is neglected

$$\alpha\boldsymbol{n}\times\dot{\boldsymbol{n}} + \rho\boldsymbol{n}\times\ddot{\boldsymbol{n}} = \boldsymbol{n}\times\sum_{\mu=x,y,z}(A\partial_\mu^2\boldsymbol{n} + B\partial_\mu^4\boldsymbol{n}) \tag{S16}$$

For a widely used AFM material, it is reasonable to suppose that the damping is so small that the first term in the LHS can be abandoned.

Before considering the dynamics of spin waves, an orthonormal frame $\hat{\boldsymbol{e}}_{x,y,z}$ is defined at each point of the magnet, based on the spin configuration of the lowest energy state. $\hat{\boldsymbol{e}}_z$ is always pointing along the Neel field of the lowest energy state, denoted as $\boldsymbol{n}_0$, while $\hat{\boldsymbol{e}}_x$ and $\hat{\boldsymbol{e}}_y$ are arbitrarily chosen. Then, any vector fields on the magnet are represented in terms of this frame after a corresponding orthogonal transformation on their global representation. The frame is used to separate from the total magnetization the spin-wave component, which is treated as a perturbation disturbing $\boldsymbol{n}_0$ and is now represented by vectors on the tangent plane to the unit sphere at $\boldsymbol{n}_0$, or, in other words, vectors on the xOy-plane in the frame. Either of the planes can be regarded as a complex plane, on which vectors can be represented by complex numbers. The spin-wave chirality is defined by the handedness of the frame, or equivalently, the sign of the determinant of the orthogonal transformation mentioned previously. In this language, we rewrite the differential operators $\partial_\mu\boldsymbol{n}$, $\partial_\mu^2\boldsymbol{n}$, $\partial_\nu(\partial_\mu^2\boldsymbol{n})$ and $\partial_\nu^2(\partial_\mu^2\boldsymbol{n})$ in terms of the local coordinates[53]. The first two operators in the new form have been given in Ref. 53, and the rest of them are presented here.

$$\partial_\nu(\partial_\mu^2\boldsymbol{n}) = R(\partial_\nu + A_\nu)(\partial_\mu + A_\mu)^2\boldsymbol{n}' \tag{S17a}$$

$$\partial_\nu^2(\partial_\mu^2\boldsymbol{n}) = R(\partial_\nu + 2A_\nu)(\partial_\nu + A_\nu)(\partial_\mu + A_\mu)^2\boldsymbol{n}' \tag{S17b}$$

Suppose the spin-wave component $\delta\boldsymbol{n}$ (or $\delta\boldsymbol{n}' = R^{-1}\delta\boldsymbol{n}$ in the local form) varies much faster than $\boldsymbol{n}_0$ (or $\hat{\boldsymbol{e}}_z$ in the local form) both spatially and temporally. Substitute Eq. S13 into Eq. S12, leave only the fast components and rewrite Eq. S12 in terms of the complex field $\psi_q = n_x'$. We obtain the dynamic equation of spin waves in the complex form

$$\rho(i\partial_t - q\phi/\hbar)^2\psi_q = \sum_{\mu=x,y,z}\left[A(-i\partial_\mu - qa_\mu/\hbar)^2 + B(-i\partial_\mu - 2qa_\mu/\hbar)(-i\partial_\mu - qa_\mu/\hbar)^3\right]\psi_q \tag{S18}$$

where ψ_q is the spin waves in the complex form, $\phi = \hbar[R^{-1}\partial_t R]_{12}$ and $a_\mu = -\hbar[R^{-1}\partial_\mu R]_{12}$ constitute the emergent electromagnetic potential, and R is the matrix of the orthogonal transformation whose sign has been taken out and denoted as q. It should be noted that the spin-wave component for each kind of chirality is being treated separately and the motion of the hopfion is, in this sense, a combined effect of them. To simplify the solving procedure, $(-i\partial_\mu - 2a_\mu/\hbar)$ is replaced with an approximate term $(-i\partial_\mu - a_\mu/\hbar)$ so that Eq. S14 contains only two monomials of canonical momentum with the proper orders for completing the square.

Eq. S14 is solved by treating spin waves as light rays at first and then classical particles. To ease a further discussion, we take $q = 1$ and $\phi \equiv 0$ for simplicity. An ansatz is taken for the spin waves as $\psi(\boldsymbol{r},t) = A(\boldsymbol{r},t)\exp[i\varphi(\boldsymbol{r},t)]$. Supposing the amplitude A, the wave vector $\boldsymbol{k} = \nabla\varphi$, the angular frequency $\omega = \partial_t\varphi$, and the emergent vector potential $\boldsymbol{a}$ vary slowly in comparison with the

phase φ, substitute the ansatz into Eq. S14 and drop those terms higher than zeroth order. We obtain the eikonal equation

$$\omega^2 = \frac{1}{\rho} \sum_{\mu=x,y,z} \left[A(k_\mu - a_\mu)^2 + B(k_\mu - a_\mu)^4 \right] \tag{S19}$$

The higher order terms are neglected here, including the nonlinear effect of spin waves, the anisotropy field, an electromagnetic field caused by the motion of the hopfion, and the coupling between different frequencies. These phenomena will be considered and play an important role in constructing an all-hopfion neural network, but it is too far to go there simply to predict the velocity of the hopfion. The eikonal equation can be seen as the dispersion relation of spin waves after taking square roots on both sides, from which follows immediately Hamilton's equations describing classical particles moving in the traditional phase space

$$\frac{\mathrm{d}x_\nu}{\mathrm{d}t} = 2c \frac{K_\nu(K_\nu^2 + \eta)}{\sqrt{\sum_\mu (K_\mu^2 + \eta)^2 - 3\eta^2}} \tag{S20a}$$

$$\frac{\mathrm{d}K_\nu}{\mathrm{d}t} = \sum_\mu F_{\mu\nu} \frac{\mathrm{d}x_\mu}{\mathrm{d}t} \tag{S20b}$$

where μ and ν are the indices of coordinates, x_ν is the ray or the path of the classical particle, to which we use the term "magnon" to refer, $\dot{x}_\nu$ is the group velocity of the ray, $K_\nu = k_\nu - a_\nu$ is the canonical momentum of the ray, $F_{\mu\nu} = \mathbf{n} \cdot (\partial_\mu \mathbf{n} \times \partial_\nu \mathbf{m})$ is the antisymmetric matrix form of the emergent magnetic field $\mathbf{F} = \nabla \times \mathbf{a}$, $c = \sqrt{B/\rho}$, and $\eta = A/2B$. In our system, $c = 1.06 \sim 1.25$ nm$^2 \cdot$ ps^{-1}, $\eta = 0.161$ nm^{-2}, two orders of magnitude lower than that of $|\mathbf{K}|^2$, which is around $(2\pi)^2$ nm^{-2}. These two equations describe the dynamical behavior of magnons moving in the background of a hopfion. Eq. S16a reveals the inconsistence between the group velocity and the canonical momentum. Due to the magnetic frustration in our system, the group velocity is anisotropically dependent on the canonical momentum: the momentum along coordinate axes results in the largest group velocity, while those along other directions are weakened and guided to the nearest coordinate axes. Eq. S16b characterizes the force the magnon experiences as Lorenz-like, justifying the name of the emergent "magnetic" field. Our method is a variant of pilot wave theory without quantum potential. It emphasizes the particle aspect of magnons, an aspect more natural and engineer-friendly than the wave aspect once we decide to develop the physical process into a device with an "injector" to "emit" magnons and a "detector" to "capture" them.

Having described the behavior of magnons, we are ready to calculate the driving force acting on the hopfion. The incident wave is treated as an array of rays or a beam of magnons. Each magnon has an initial linear momentum equal to the wave vector of the incident wave and, after crossing the hopfion, transfers part of its momentum to the hopfion[54-55]. Adding up all the transferred momenta, we obtained the value of the driving force. Supposed that the hopfion experiences damping linear in its velocity, it will be reasonable to reach the conclusion that the stable velocity of the hopfion is proportional to the driving force.
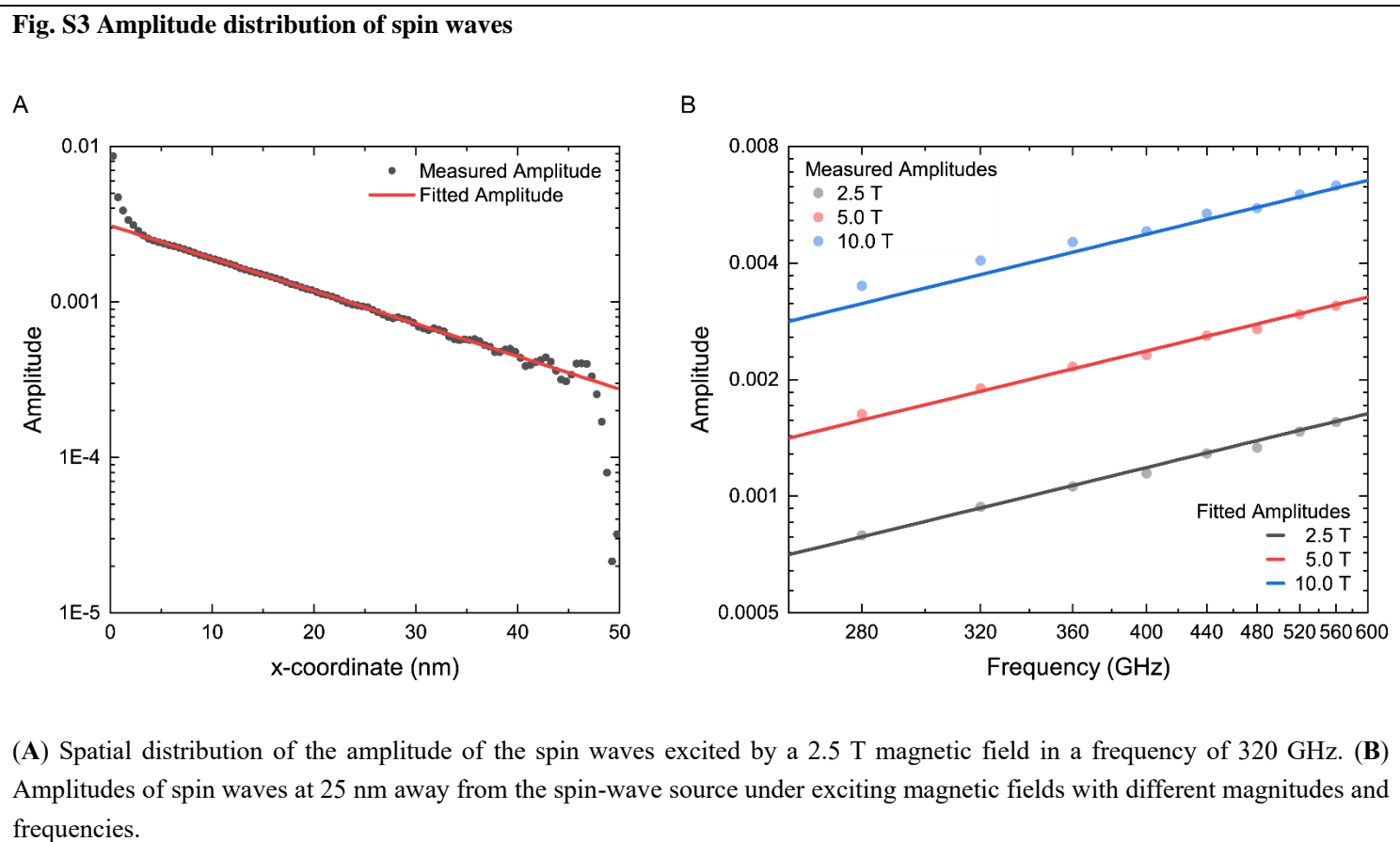
IV. MEASUREMENT OF THE VELOCITY OF A HOPFION

The velocity of the hopfion will be measured under different frequencies and then compared with the driving force calculated from our theoretical model. However, the instantaneous velocity is not available for the comparison since it varies with the amplitude of spin waves, which depends on the position of the hopfion as well as the frequency-determining decay length. An available definition of velocity should be the one normalized by the square of the amplitude of spin waves. In light of this, we will measure the amplitude of spin waves propagating in a uniformly oriented magnet and then determine the normalized velocity by fitting the displacement of the hopfion with the solution of the following differential equation

$$\dot{x} = v_x(f)|A(x)|^2 \tag{S21}$$

where f is the frequency of spin waves, x is the distance between the hopfion and the wave source, $\dot{x}$ is the instantaneous velocity, and v_x is the x-component of the normalized velocity, which is independent of the polarization. To determine the wave amplitude $A(x)$, we simulate the propagation of plane waves on a uniformly magnetized background and calculate the amplitude at each point. After an exponential fitting (see Fig. S3), the decay length for each frequency is found to be $\lambda = 1.55(f/\text{GHz})^{-1}$ nm in our system. In this way, the normalized velocity can be solved from Eq. S17

$$v_x = \frac{\lambda}{2A^2\tau}\left(e^{2x_0/\lambda} - e^{2x_1/\lambda}\right) \tag{S22}$$

where A is the amplitude of the wave source, and τ is the simulation time it takes for the hopfion to move from x_0 to x_1. The normalized velocity is independent of the wave amplitude or the position of the hopfion, and is thus suitable for comparison between different frequencies.

Fig. S3 Amplitude distribution of spin waves

(**A**) Spatial distribution of the amplitude of the spin waves excited by a 2.5 T magnetic field in a frequency of 320 GHz. (**B**) Amplitudes of spin waves at 25 nm away from the spin-wave source under exciting magnetic fields with different magnitudes and frequencies.

V. REALIZATION OF ALL-HOPFION META-LEARNING DEVICE

1. Implementation of a Neural Network

We construct a neural network to learn a vectorial time series $s(t)$. The training of the neural network requires a training dataset D_{train} and a test dataset D_{test}. D_{train} and D_{test} consist of a pair of input and output vectors. An input vector is a time sampling of $s(t)$ denoted as $\boldsymbol{s}_t = [s(t - \Delta t), s(t - 2\Delta t), \ldots, s(t - n\Delta t)]^{\text{T}}$, where Δt is the sampling interval. An output vector o_t (temporarily written as a scalar) is the value of $s(t)$ at time t. Our network will learn the features of the signal from $\boldsymbol{s}_t$ and give its own prediction $\tilde{o}_t$.

The nonlinear part of the neural network is realized using magnon scattering by the hopfion. Before scattering, the input vector $\boldsymbol{s}_{m\Delta t}$ should be encoded into the frequency spectrum of the incident magnons. Here, we denote the frequency spectrum between $(m - 1)\Delta t$ and $m\Delta t$ as $[a_{m,1}, a_{m,2}, \ldots, a_{m,n}]$, etc. As a result, the incident spin waves have the form $\psi_m(t) = \sum_k a_{m,k} \exp i(\omega_k t + \varphi_{m,k})$, and they will last for Δt. After these waves or magnons are scattered by the hopfion, they are detected at the other side of the device, hence the outgoing spin waves $\psi'_m(t) = \sum_k a'_{m,k} \exp i(\omega_k t + \varphi'_{m,k})$ with the same time duration. The outgoing waves are then decoded using the inverse Fourier transformation and expressed in a vector's form $\boldsymbol{p}_{m\Delta t} = [a'_{m,1}, a'_{m,2}, \ldots, a'_{m,n}]^{\text{T}}$. To summarize, an input series $\{\boldsymbol{s}_{\Delta t}, \boldsymbol{s}_{2\Delta t}, \ldots, \boldsymbol{s}_{m\Delta t}, \ldots\}$ is encoded into a sequence of incoming waves $\{\psi'_{\Delta t}(t), \psi'_{2\Delta t}(t), \ldots, \psi'_{m\Delta t}(t), \ldots\}$ with the same duration Δt (here $\Delta t = 20$ ps), and an output series $\{\boldsymbol{p}_{\Delta t}, \boldsymbol{p}_{2\Delta t}, \ldots, \boldsymbol{p}_{m\Delta t}, \ldots\}$ is decoded from the outgoing waves $\{\psi'_{\Delta t}(t), \psi'_{2\Delta t}(t), \ldots, \psi'_{m\Delta t}(t), \ldots\}$ generated by the scattering process. Since the neural network is built into the scattering process, the meta-learning procedure is actually the procedure of modulating the parameters of the scattering process. The spectra and waveforms after encoding and before decoding are shown in Fig. S4 to clarify the abovementioned procedure.

**Fig. S4 Frequency spectra and waveforms after encoding and before decoding**
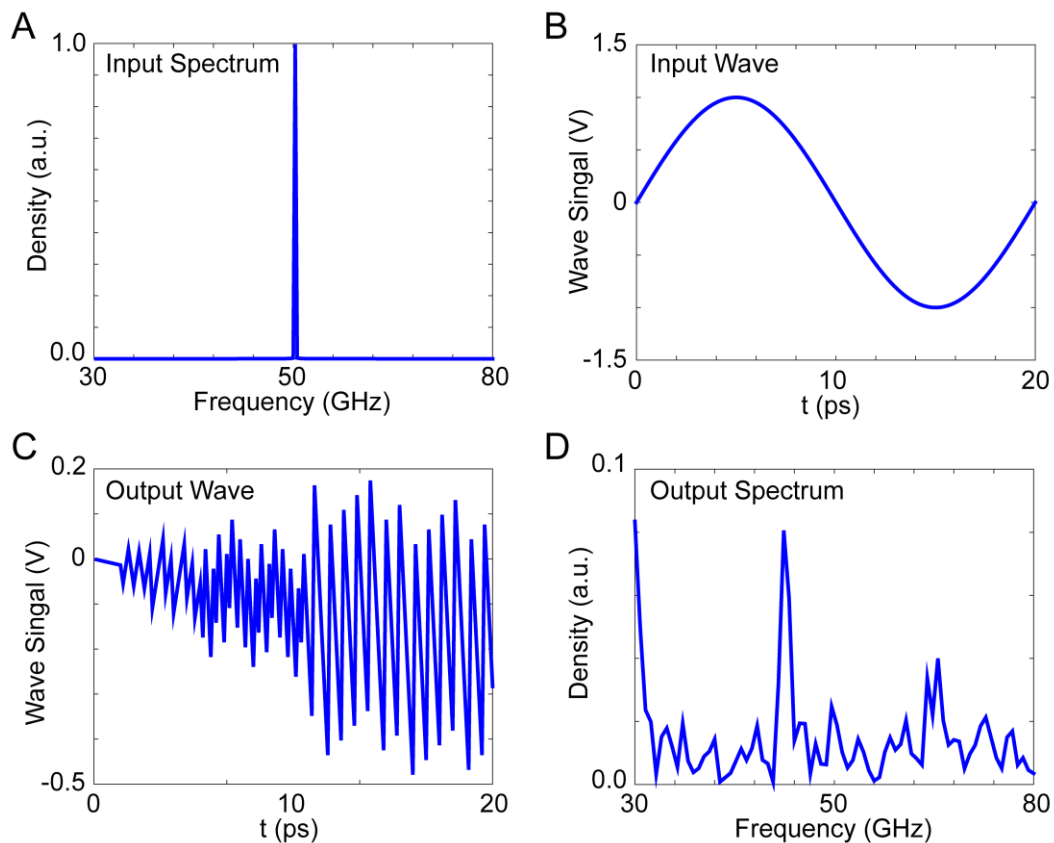
(**A** and **B**) Encoding the training dataset into the incident spin waves of our device. The incident spin waves are shown in (A) frequency spectrum and (B) time domain. (**C** and **D**) Decoding the prediction signal from the outgoing spin waves of our device. The outgoing spin waves are shown in (C) frequency spectrum and (D) time domain.

The linear part of the network is to calculate the final prediction $\{\tilde{o}_{\Delta t}, \tilde{o}_{2\Delta t}, \ldots, \tilde{o}_{m\Delta t}, \ldots\}$ from $\{\boldsymbol{p}_{\Delta t}, \boldsymbol{p}_{2\Delta t}, \ldots, \boldsymbol{p}_{m\Delta t}, \ldots\}$. The calculation happens in the output layer of the physical neural network in the form of $\tilde{o}_{m\Delta t} = \boldsymbol{w}_{\text{out}}^{\text{T}} \boldsymbol{p}_{m\Delta t}$ for each m, where $\boldsymbol{w}_{\text{out}}$ is the weights of the output layer and is independent of time. The optimization of this part is done by minimizing the residual error $\sum_{m \in \text{Batch}} \|\tilde{o}_{m\Delta t} - o_{m\Delta t}\|^2$ for every batch of vectors in the training dataset. The values of the elements in $\boldsymbol{w}_{\text{out}}$ are determined using the stochastic gradient descent (SGD) algorithm[39]. Several periodic signals are given in Fig. S5 to demonstrate the function of linear regression. The meta-learning step is skipped, leaving only the linear regression step responsible for predicting the target signals. The neuron number of the neural network is set to 10, so that the frequency spectrum has only ten components. The input

signals are sampled every 50 ps. The output layer of neural network is then trained to predict the given signals. It turns out that, with a fixed nonlinear mapping, a linear network can still give good predictions of periodic signals[56]. However, the modulation of the nonlinear mapping becomes necessary if we want to improve the learning capability of the network, especially the learning capability for complex prediction tasks, through multi-task training.

Fig. S5 Predicting periodic signals by means of linear regression
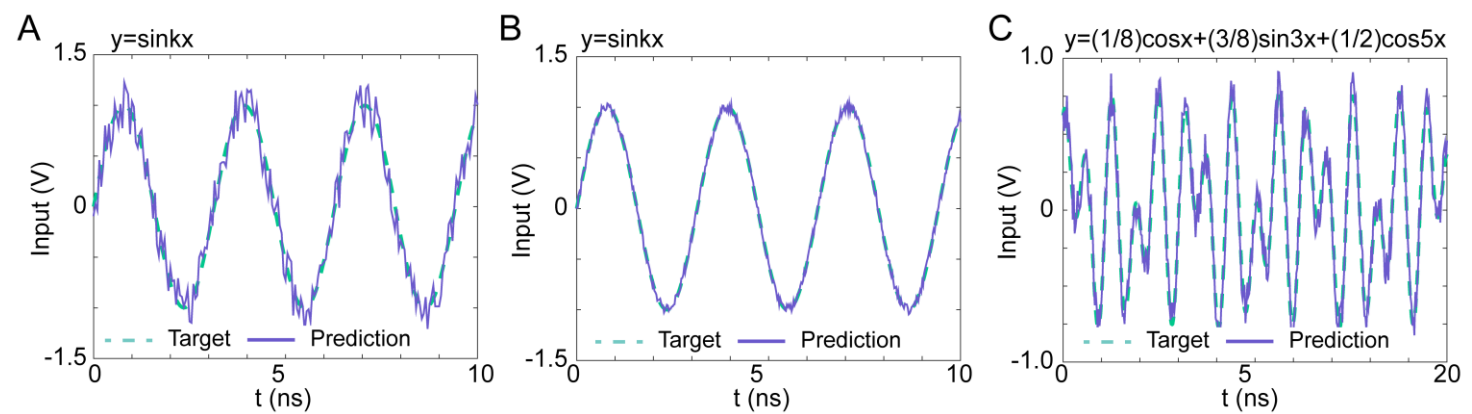
(**A** and **B**) Sinusoidal signal as the target signal and the predictions given by two encoding schemes. (**C**) Periodic signals as the target signal and the prediction given in the linear regression step. The periodic signal consists of three independent sinusoidal components.

---

2. Implementation of Meta-Learning

To understand the procedure of meta-learning, it is necessary to clarify how the neural network is built into the scattering process. Each frequency in the spectrum can be treated as a neuron in the neural network. Hence, the spectra of the incident waves and of the scattered waves can be seen respectively as the input and output of the hidden layers. These two layers are connected by the hidden layers, i.e., the nonlinear scattering process which maps the incident-wave spectrum to the scattering spectrum. The components in the scattering spectrum are then linearly recombined in the output layer to produce an output. By training the output layer, this architecture can perform as well as an ordinary deep neural network (DNN)[48].

A densely interconnected network can therefore be easily obtained if the number of neurons becomes large enough. The total number of neurons in the neural network can be modulated by the precision of detection and is possible to become very large. Since the magnons of the same frequency are basically reflected by the same isosurface of $|\boldsymbol{F}|$ during the scattering process, the connections between a node and the others are roughly shaped by the corresponding isosurface. On the other hand, the interference between magnons with different frequencies gives the fine structure of the network. This interference is described by nonlinear energy terms which leads to an interaction between magnons. As a result, there is a nonlinear mapping from $\{\boldsymbol{s}_{\Delta t}, \boldsymbol{s}_{2\Delta t}, \dots, \boldsymbol{x}_{m\Delta t}, \dots\}$ to $\{\boldsymbol{p}_{\Delta t}, \boldsymbol{p}_{2\Delta t}, \dots, \boldsymbol{p}_{m\Delta t}, \dots\}$, on which a neural network is based.

According to our model, the magnon scattering spectrum at the detector's position depends on the position of the hopfion. Denote all the state parameters of the hopfion as the vector $\boldsymbol{\omega}$. As the input series $\{\boldsymbol{s}_{\Delta t}, \boldsymbol{s}_{2\Delta t}, \dots, \boldsymbol{s}_{m\Delta t}, \dots\}$ is put into the device, the state of the hopfion is changed accordingly and forms a state series $\{\boldsymbol{\omega}_{\Delta t}, \boldsymbol{\omega}_{2\Delta t}, \dots, \boldsymbol{\omega}_{m\Delta t}, \dots\}$. Denote the nonlinear mapping from the frequency spectrum before scattering to that after scattering as $\boldsymbol{p}_{m\Delta t} = \mathcal{F}(\boldsymbol{s}_{m\Delta t}; \boldsymbol{\omega}_{m\Delta t})$. $\mathcal{F}$ is a nonlinear mapping from $\boldsymbol{s}_{m\Delta t}$ to $\boldsymbol{p}_{m\Delta t}$, depending on the state parameters of the hopfion $\boldsymbol{\omega}_{m\Delta t}$, which will be shown later, the hyperparameters of the neural network. The training of the meta-learning network is performed by optimizing the weights of the output layer $\boldsymbol{w}_{\text{out}}$ and the hyperparameters $\boldsymbol{\omega}$.

$$\{\boldsymbol{w}_{\text{out}}, \boldsymbol{\omega}\}^* = \underset{\boldsymbol{w}_{\text{out}}, \boldsymbol{\omega}}{\text{argmin}} \, \text{Loss}[\boldsymbol{w}_{\text{out}}, \boldsymbol{\omega}, D_{\text{train}}] \tag{S23}$$

This is realized by the so-called bilevel optimization[19, 57]. $\boldsymbol{w}_{\text{out}}$ is optimized using linear regression under the constraints that the nonlinear mapping is fixed by $\boldsymbol{\omega}$. In this sense, $\boldsymbol{w}_{\text{out}}$ can be seen as a function of $\boldsymbol{\omega}$ and is expressed in a function's form as $\boldsymbol{w}_{\text{out}}(\boldsymbol{\omega})$ to emphasize this fact. Since $\boldsymbol{w}_{\text{out}}$ itself plays the role of a neural network, the controller of the nonlinear mapping, $\boldsymbol{\omega}$, should be regarded as a hyperparameter of the neural network. The mathematical form of the optimization is written separately in two levels. For multi-task training, $\boldsymbol{\omega}$ is shared between different kinds of tasks, while $\boldsymbol{w}_{\text{out}}$ is divided into many versions and each kind of tasks possesses its own versions.

$$\boldsymbol{\omega} \leftarrow \underset{\boldsymbol{\omega}}{\text{argmin}} \, \text{Loss}[\boldsymbol{w}_{\text{out}}(\boldsymbol{\omega}), \boldsymbol{\omega}, D_{\text{train}}] \, \text{st.} \, \boldsymbol{w}_{\text{out}}(\boldsymbol{\omega}) = \underset{\boldsymbol{w}_{\text{out}}}{\text{argmin}} \, \text{Loss}[\boldsymbol{w}_{\text{out}}, \boldsymbol{\omega}, D_{\text{train}}] \tag{S24}$$

The minimization in the second level is conducted using linear regression. The minimization in the first part is much more complex since the loss function's dependence on $\boldsymbol{\omega}$ is extremely nonlinear. It is carried out using state space searching based on numerical simulation. Given an original state $\boldsymbol{\omega}_0$, the best increment $\delta\bar{\boldsymbol{\omega}}_{m\Delta t} = \boldsymbol{\omega}_{m\Delta t} - \boldsymbol{\omega}_{(m-1)\Delta t}$ is found by "trial-and-error". Given a trial increment $\delta\boldsymbol{\omega}$, the resulting scattering spectrum is found by micromagnetic simulation. The loss function is then minimized w.r.t $\boldsymbol{w}_{\text{out}}$ and compared with the original one

$$\delta\text{Loss} = \text{Loss}[\boldsymbol{w}_{\text{out}}(\boldsymbol{\omega}), \boldsymbol{\omega} + \delta\boldsymbol{\omega}, D_{\text{train}}] - \text{Loss}[\boldsymbol{w}_{\text{out}}(\boldsymbol{\omega}), \boldsymbol{\omega}, D_{\text{train}}] \tag{S25}$$

By searching $\delta\boldsymbol{\omega}$ in the state space, the best $\delta\boldsymbol{\omega}$ that leads to the fastest decrease of the loss function is found.

3. Working Principle and Training Method of Meta-Learning Network

The principle of our proposed neural network implementation is based on a physical neural network. As shown in Fig. S6, we illustrate the schematic and information flow of the neural network in the main text. In our neural network, the nonlinear mapping function implemented by the scattering action acts as the hidden layer of an ordinary DNN[48].

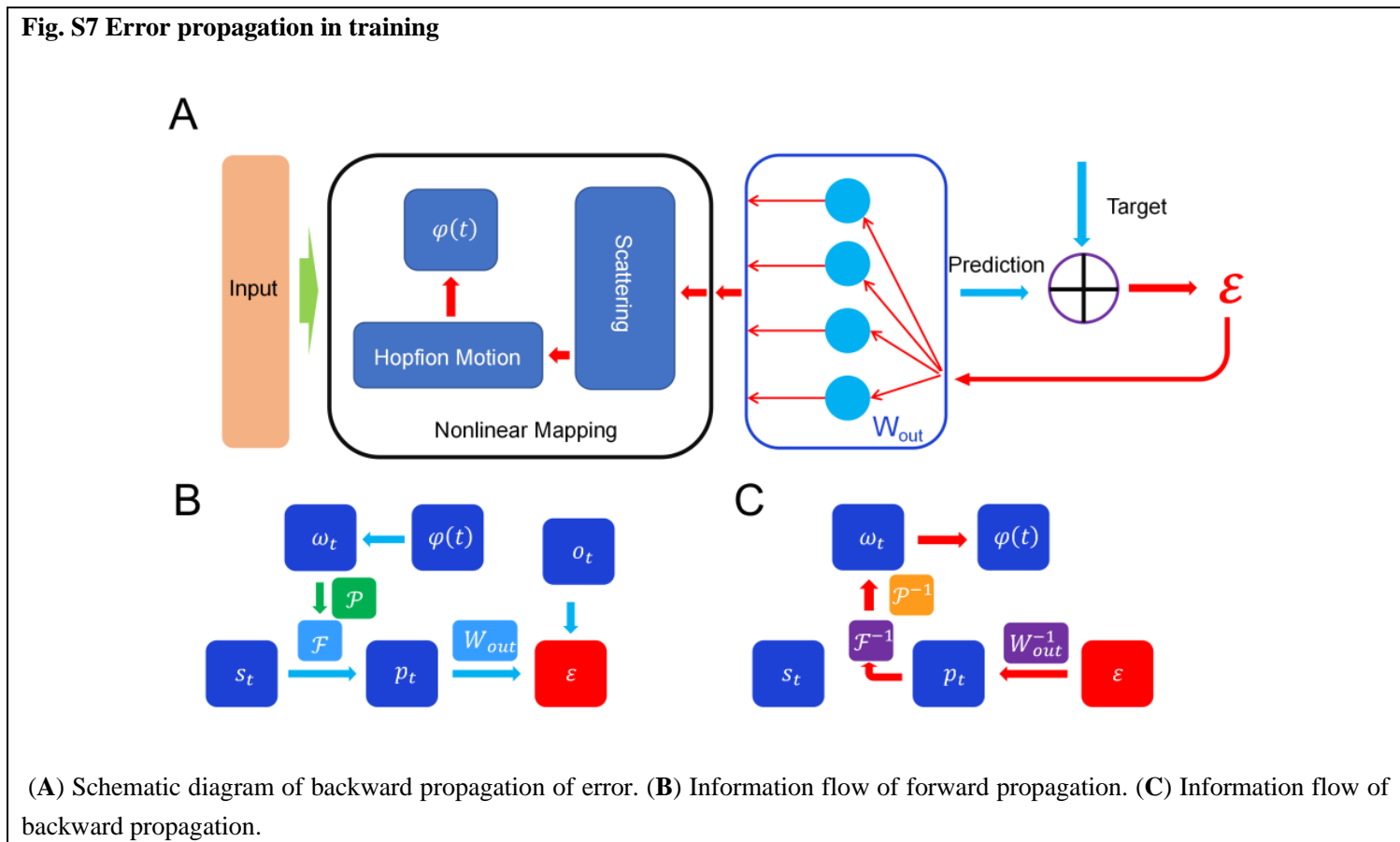**Fig. S6 Comparison between a DNN and a physical neural network**
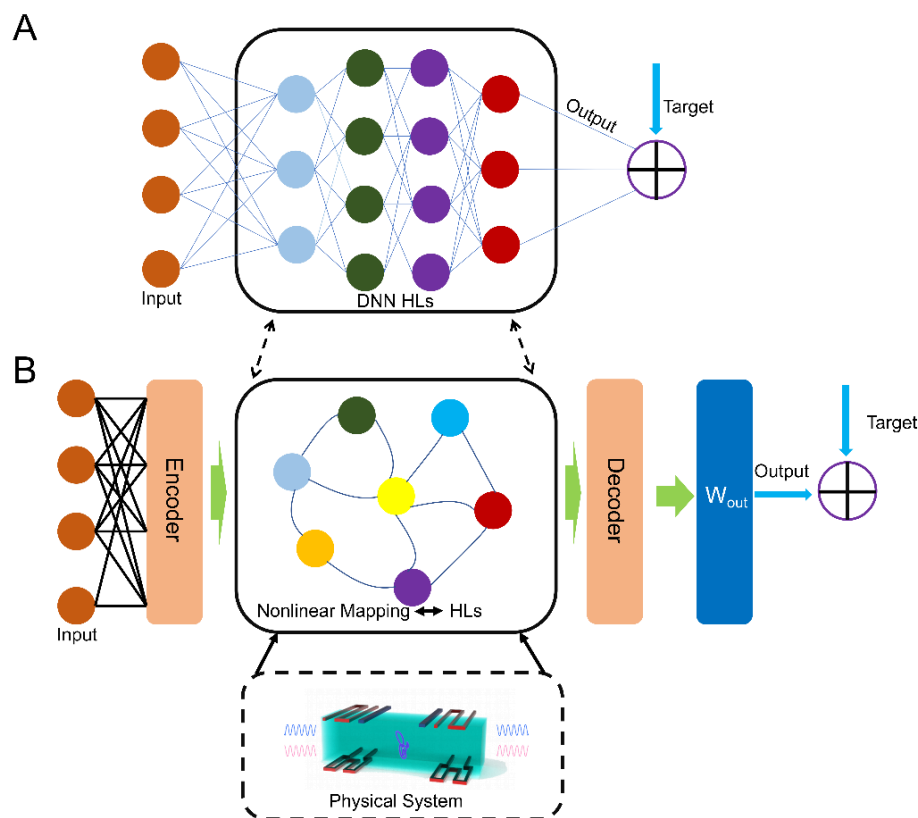
(**A**) Structure of a DNN. For a DNN, the parameters of all the neurons in the hidden layer need to be trained. (**B**) Structure of a physical neural network. For a physical neural network, the hidden layer is replaced by a nonlinear physical process of the physical system, which maps the encoded signals from the encoder to certain physical quantities observed by the decoder. Only the weights of the output layer need to be trained. In our work, the magnon scattering process is the nonlinear physical process that realizes nonlinear mapping in the physical neural network. It is trained using modulating the state parameters of the hopfion.
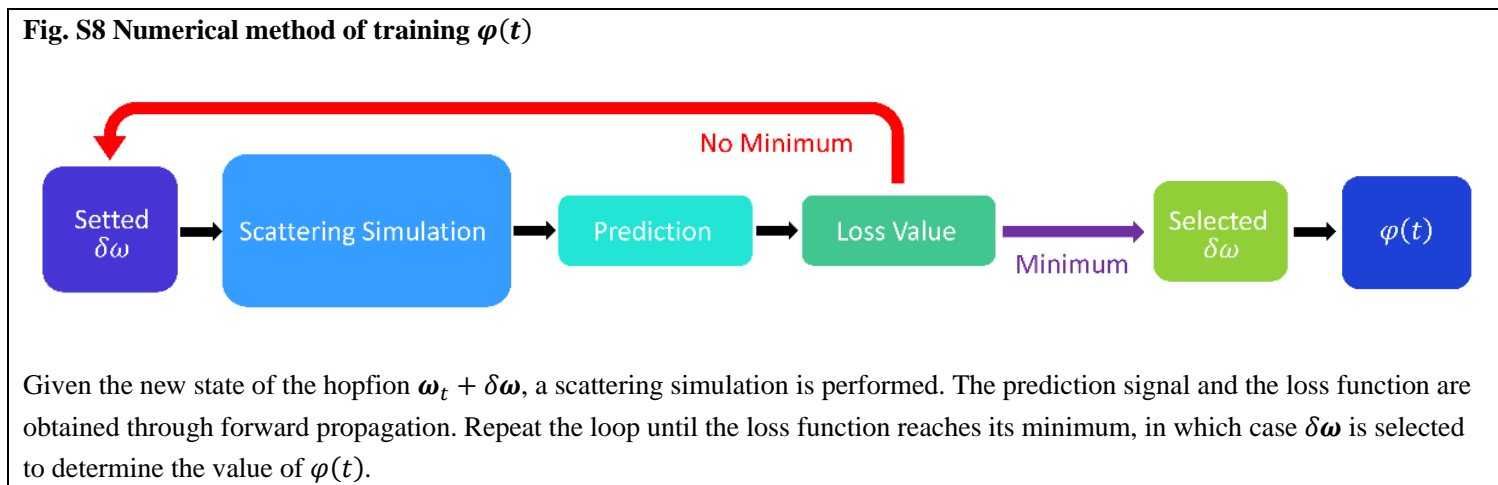
As shown in Fig. S7, we demonstrate the information flow and the back propagation process of prediction error in the neural network. From Fig. S7A, we can see that the training of basic neural network parameters can be completed by training the output layer. Furthermore, the training for magnon scattering is equivalent to hyperparameter training, and the meta-learning function can be realized based on this. However, it is a challenge to find the analytical form of derivatives of the nonlinear mapping required by the back-propagation. To solve this problem, we propose a training algorithm of back-propagation error combined with numerical simulation.

As illustrated in Fig. S8, we draw the flowchart of this algorithm. When training the hyperparameter $\boldsymbol{\omega}$ according to the backpropagation error ε, we first need to give the system a trial parameter $\delta\boldsymbol{\omega}$ and calculate the backpropagation error $\varepsilon + \delta\varepsilon$, at this moment by numerical simulation and compare it with the original error. By searching all the possible trial parameters $\delta\boldsymbol{\omega}$, the trial parameter $\delta\bar{\boldsymbol{\omega}}$ that makes the error ε decrease fastest will be found and it is used as to update hyperparameters. Finally, the variation of $\boldsymbol{\omega}$, $\delta\bar{\boldsymbol{\omega}}$, can be realized by modulating $\varphi(t)$ according to our research on omnidirectional motion (specifically, Eq. 8 in the main text). With this algorithm, it is possible to train any hyperparameter without solving the inverse function.

Fig. S7 Error propagation in training

(**A**) Schematic diagram of backward propagation of error. (**B**) Information flow of forward propagation. (**C**) Information flow of backward propagation.

Fig. S8 Numerical method of training $\varphi(t)$

Given the new state of the hopfion $\boldsymbol{\omega}_t + \delta\boldsymbol{\omega}$, a scattering simulation is performed. The prediction signal and the loss function are obtained through forward propagation. Repeat the loop until the loss function reaches its minimum, in which case $\delta\boldsymbol{\omega}$ is selected to determine the value of $\varphi(t)$.

4. Verifying the Properties of Meta-Learning in a Multi-task Scenario

The neural network we proposed is a typical meta-learning framework. With the nonlinear mapping, the output layer can exhaust a large proportion of configurations of a traditional DNN simply by adjusting its weights[48]. Therefore, we can treat the output layer with a fixed nonlinear mapping as a traditional DNN. In this manner, a parameter that can control the nonlinear mapping can be seen as a hyperparameter of this neural network. In our framework, this hyperparameter is $\boldsymbol{\omega}$, the hopfion's state. Therefore, our neural network has a necessary structure element of meta-learning.

Using multi-task scenarios as a criterion, we will demonstrate that our device realizes the function of meta-learning[19]. The multi-task learning ability refers to such a quality that network will gain a strengthened learning capability in a new task after its hyperparameters are trained on multiple tasks. Generally, the learning capability refers to learning speed or learning accuracy.

To verify that our network has such an ability, a set of trigonometric functions $\left\{\sin t, \cos t, \frac{1}{2}\sin 2t + \frac{1}{2}\cos t, \sin 3t\right\}$ is selected as the task family to train the hyperparameters in our neural network. The dataset is a set of temporal samples of these functions and is divided into the training dataset and the test dataset. The training is carried out by the bilevel optimization. $\boldsymbol{w}_{\text{out}}$ is trained using the training dataset in the task family, while $\boldsymbol{\omega}$ is trained using the test dataset. The loss function we used to train $\boldsymbol{w}_{\text{out}}$ for each task is

$$L_t^{\text{task}} = |\varepsilon_t|^2 = |o_t - \tilde{o}_t|^2 = \left|o_t - \boldsymbol{w}_{\text{out}}^{\text{T}}\boldsymbol{p}_t\right|^2 = \left|o_t - \boldsymbol{w}_{\text{out}}^{\text{T}}\mathcal{P}_{\boldsymbol{\omega}[\boldsymbol{\varphi}(t)]}(\boldsymbol{s}_t)\right|^2 \tag{S26}$$

while the loss function we used to train $\boldsymbol{\omega}$ is the sum of all tasks[19]

$$L_t^{\mathrm{meta}} = \sum_{i=1}^{N} \left| \varepsilon_t^{(i)} \right|^2 = \sum_{i=1}^{N} \left| o_t^{(i)} - \left(\boldsymbol{w}_{\mathrm{out}}^{(i)} \right)^{\mathrm{T}} \mathcal{P}_{\boldsymbol{\omega}[\boldsymbol{\varphi}(t)]} \left(\boldsymbol{s}_t^{(i)} \right) \right|^2 \tag{S27}$$

where i run over all the N tasks and the superscript (i) denotes a variable varying between different tasks. After that, a new test function, $\frac{2}{3}\sin 3t + \frac{1}{3}\cos t + \frac{1}{3}\cos 5t - \frac{1}{3}\sin 4t$, is used to estimate the learning capabilities.

We first compare the change in learning speed with and without multi-task training. The learning speed is characterized by the descending speed of the loss ratio (the normalized loss function). The change of the loss ratio without and with multi-task training are plotted respectively in Fig. S9 (A and C). Without multi-task training, it takes 90 iterations for the loss ratio to converge to 0.1 and the loss ratio keeps bouncing between 0.05 and 0.15. In comparison, after multi-task training, it only takes 30 iterations for the convergence and the loss ratio is then kept at an extremely low level, smaller than 0.05. Since the rate of decline of the loss ratio can be seen as an indicator of the learning speed, we can conclude that the learning speed of our neural network is improved by the multi-task training.

We then estimate the learning accuracy for the test task by the normalized root mean squared error (NRMSE). NRMSE is a useful metric to quantify the similarity between two temporal signals. As shown in Fig. S9 (B and D), NRMSE is reduced from 0.046 to 0.031 after the multi-task training. The results demonstrate that multi-task training does increase the prediction accuracy of the neural network.
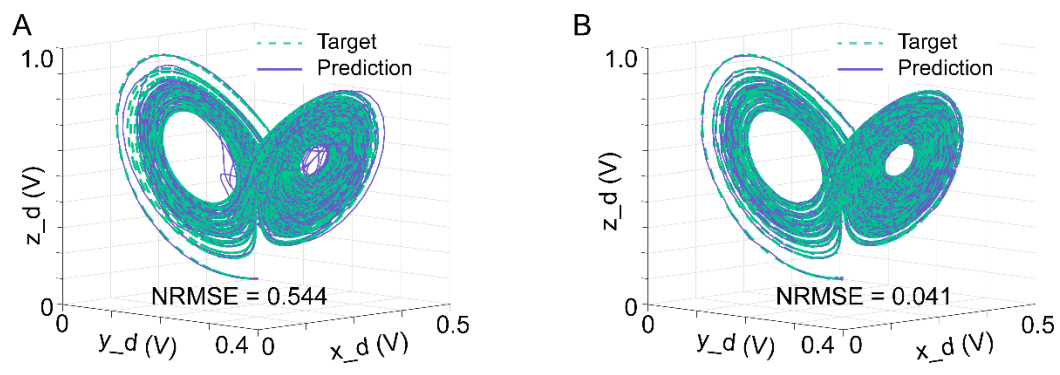
Fig. S9 Learning capability with and without multi-task training
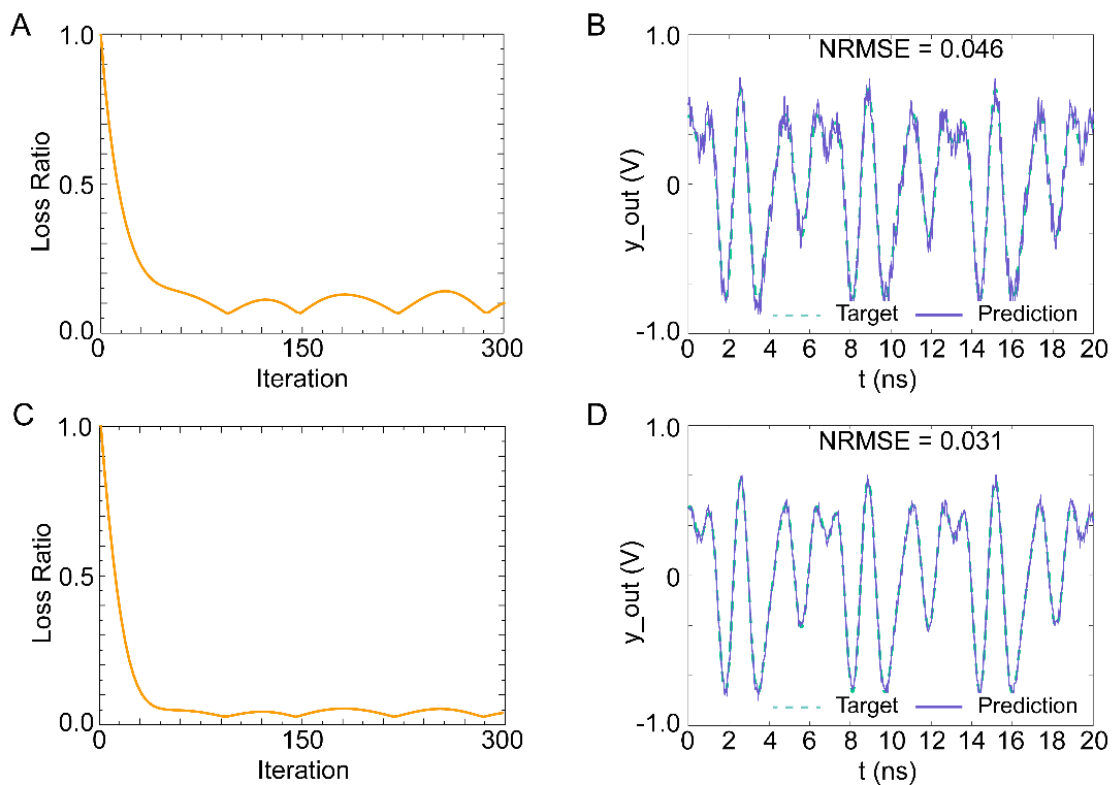
(**A** to **B**) Learning capability without multi-task training. The learning capability is quantified by (A) the convergence of the loss ratio during the test training and (B) the prediction accuracy. (**C** to **D**) Learning capability after multi-task training. The learning capability is quantified by (C) the convergence of the loss ratio during the test training and (D) the prediction accuracy.

According to meta-learning theory, the hyperparameters training using the bilevel optimization can also be applied for single-task scenarios to improve the capability for complex learning tasks[49-50]. Our neural network also has this advantage brought by meta-learning. To verify that, a numerical experiment is performed comparing the learning capability of the Lorenz Butterfly. Since Iteration 2 is the optimizing step responsible for training hyperparameters, we will compare the prediction precision with and without Iteration 2. As shown in Fig. S10, a purely physical neural network with only Iteration 1 gives much less accurate predictions (NRMSE = 0.544) than the predictions of our neural network with Iteration 2 (NRMSE = 0.041). Our results undoubtedly demonstrates that the significant improvement in prediction accuracy in single-task scenarios is attributed to meta-learning.

Fig. S10 Prediction accuracy for the Lorenz Butterfly

A

z_d (V)
1.0
0

Target
Prediction

NRMSE = 0.544

y_d (V)   0.4   0

x_d (V)   0.5

B

z_d (V)
1.0
0

Target
Prediction

NRMSE = 0.041

y_d (V)   0.4   0

x_d (V)   0.5

(**A**) Prediction without Iteration 2. (**B**) Prediction with Iteration 2.

Movies:

Movie S1. Circular trajectory of a hopfion driven by magnons.

Movie S2. Helical trajectory of a hopfion driven by magnons.

Movie S3. Chinese knot trajectory of a hopfion driven by magnons.

Movie S4. Trefoil knot trajectory of a hopfion driven by magnons.

Movie S5. Schematic animation of magnon scattering by a hopfion at 220 GHz.

Movie S6. Schematic animation of magnon scattering by a hopfion at 340 GHz.

Movie S7. Schematic animation of magnon scattering by a hopfion at 500 GHz.

REFERENCES AND NOTES

1. R. Friedberg, T. D. Lee, A. Sirlin, Gauge-field non-topological solitons in three space-dimensions (II). *Nucl. Phys. B.* **115**, 32–47 (1976).

2. M. Rho, A. S. Goldhaber, G. E. Brown, Topological soliton bag model for baryons. *Phys. Rev. Lett.* **51**, 747–750 (1983).

3. G. Finocchio, F. Büttner, R. Tomasello, M. Carpentieri, M. Kläui, Magnetic skyrmions: From fundamental to applications. *J. Phys. D Appl. Phys.* **49**, 423001 (2016).

4. A. Fert, V. Cros, J. Sampaio, Skyrmions on the track. *Nat. Nanotechnol.* **8**, 152–156 (2013).

5. S. M. Mohseni, S. R. Sani, J. Persson, T. N. A. Nguyen, S. Chung, Y. E. Pogoryelov, P. K. Muduli, E. Iacocca, A. Eklund, R. K. Dumas, S. Bonetti, A. Deac, M. A. Hoefer, J. Åkerman, Spin torque–generated magnetic droplet solitons. *Science.* **339**, 1295–1298 (2013).

6. J. Zázvorka, F. Jakobs, D. Heinze, N. Keil, S. Kromin, S. Jaiswal, K. Litzius, G. Jakob, P. Virnau, D. Pinna, K. Everschor-Sitte, L. Rózsa, A. Donges, U. Nowak, M. Kläui, Thermal skyrmion diffusion used in a reshuffler device. *Nat. Nanotechnol.* **14**, 658–661 (2019).

7. Y. Zhou, M. Ezawa, A reversible conversion between a skyrmion and a domain-wall pair in a junction geometry. *Nat. Commun.* **5**, 4652 (2014).

8. X. Zhang, M. Ezawa, Y. Zhou, Magnetic skyrmion logic gates: Conversion, duplication and merging of skyrmions. *Sci. Rep.* **5**, 9400 (2015).

9. M. A. P. Gonçalves, C. Escorihuela-Sayalero, P. Garca-Fernández, J. Junquera, J. Íñiguez, Theoretical guidelines to create and tune electric skyrmion bubbles. *Sci. Adv.* **5**, eaau7023 (2019).

10. N. Romming, A. Kubetzka, C. Hanneken, K. von Bergmann, R. Wiesendanger, Field-dependent size and shape of single magnetic skyrmions. *Phys. Rev. Lett.* **114**, 177203 (2015).

11. J. Iwasaki, M. Mochizuki, N. Nagaosa, Current-induced skyrmion dynamics in constricted geometries. *Nat. Nanotechnol.* **8**, 742–747 (2013).

12. M. Finazzi, M. Savoini, A. R. Khorsand, A. Tsukamoto, A. Itoh, L. Duo, A. Kirilyuk, T. Rasing, M. Ezawa, Laser-induced magnetic nanostructures with tunable topological properties. *Phys. Rev. Lett.* **110**, 177205 (2013).

13. O. Boulle, J. Vogel, H. Yang, S. Pizzini, D. de Souza Chaves, A. Locatelli, T. O. Menteş, A. Sala, L. D. Buda-Prejbeanu, O. Klein, M. Belmeguenai, Y. Roussigné, A. Stashkevich, S. M. Chérif, L. Aballe, M. Foerster, M. Chshiev, S. Auffret, I. M. Miron, G. Gaudin, Room-temperature chiral magnetic skyrmions in ultrathin magnetic nanostructures. *Nat. Nanotechnol.* **11**, 449–454 (2016).

14. S. Woo, K. Litzius, B. Krüger, M.-Y. Im, L. Caretta, K. Richter, M. Mann, A. Krone, R. M. Reeve, M. Weigand, P. Agrawal, I. Lemesh, M.-A. Mawass, P. Fischer, M. Kläui, G. S. D. Beach, Observation of room-temperature magnetic skyrmions and their current-driven dynamics in ultrathin metallic ferromagnets. *Nat. Mater.* **15**, 501–506 (2016).

15. J. Grollier, D. Querlioz, K. Y. Camsari, K. Everschor-Sitte, S. Fukami, M. D. Stiles, Neuromorphic spintronics. *Nat. Electron.* **3**, 360–370 (2020).

16. K. Yue, Y. Liu, R. K. Lake, A. C. Parker, A brain-plausible neuromorphic on-the-fly learning system implemented with magnetic domain wall analog memristors. *Sci. Adv.* **5**, eaau8170 (2019).

17. S. Luo, N. Xu, Z. Guo, Y. Zhang, J. Hong, L. You, Voltage-controlled skyrmion memristor for energy-efficient synapse applications. *IEEE Electron Device Lett.* **40**, 635–638 (2019).

18. K. M. Song, J. S. Jeong, B. Pan, X. Zhang, J. Xia, S. Cha, T. E. Park, K. Kim, S. Finizio, J. Raabe, J. Chang, Y. Zhou, W. Zhao, W. Kang, H. Ju, S. Woo, Skyrmion-based artificial synapses for neuromorphic computing. *Nat. Electron.* **3***, 148–155 (2020).

19. T. Hospedales, A. Antoniou, P. Micaelli, A. Storkey, Meta-learning in neural networks: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 5149–5169 (2022).

20. R. Vilalta, Y. Drissi, A perspective view and survey of meta-learning. *Artif. Intell. Rev.* **18**, 77–95 (2002).

21. K. Hsu, S. Levine, C. Finn, Unsupervised Learning via Meta-Learning; https://arxiv.org/abs/1810.02334 (2018).

22. A. Luedtke, M. Carone, N. Simon, O. Sofrygin, Learning to learn from data: Using deep adversarial learning to construct optimal statistical procedures. *Sci. Adv.* **6**, eaaw2140 (2020).

23. A. Rajeswaran, C. Finn, S. M. Kakade, S. Levine, Meta-learning with implicit gradients. *Adv. Neural Inf. Process Syst. Title.* **32**, 97817 (2019).

24. V. E. E. Korepin, L. D. Faddeev, Quantization of solitons. *Theor. Math. Phys.* **25**, 1039–1049 (1975).

25. L. D. Faddeev, Some comments on the many-dimensional solitons. *Lett. Math. Phys.* **1**, 289–293 (1976).

26. J. S. Tai, I. I. Smalyukh, Static Hopf solitons and knotted emergent fields in solid-state noncentrosymmetric magnetic nanostructures. *Phys. Rev. Lett.* **121**, 187201 (2018).

27. J. S. B. Tai, P. J. Ackerman, I. I. Smalyukh, Topological transformations of Hopf solitons in chiral ferromagnets and liquid crystals. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 921–926 (2018).

28. P. Sutcliffe, Hopfions in chiral magnets. *J. Phys. A: Math. Theor.* **51**, 375401 (2018).

29. R. Voinescu, J. S. B. Tai, I. I. Smalyukh, Hopf solitons in helical and conical backgrounds of chiral magnetic solids. *Phys. Rev. Lett.* **125**, 057201 (2020).

30. I. L. Bogolubsky, Three-dimensional topological solitons in the lattice model of a magnet with competing interactions. *Phys. Lett. A* **126**, 511–514 (1988).

31. P. Sutcliffe, Skyrmion knots in frustrated magnets. *Phys. Rev. Lett.* **118**, 247203 (2017).

32. N. Kent, N. Reynolds, D. Raftrey, I. T. G. Campbell, S. Virasawmy, S. Dhuey, R. V. Chopdekar, A. Hierro-Rodriguez, A. Sorrentino, E. Pereiro, S. Ferrer, F. Hellman, P. Sutcliffe, P. Fischer, Creation and observation of Hopfions in magnetic multilayer systems. *Nat. Commun.* **12**, 1562 (2021).

33. F. N. Rybakov, N. S. Kiselev, A. B. Borisov, L. Döring, C. Melcher, S. Blügel, Magnetic hopfions in solids. *APL Mater.* **10**, 111113 (2022).

34. X. S. Wang, A. Qaiumzadeh, A. Brataas, Current-driven dynamics of magnetic hopfions. *Phys. Rev. Lett.* **123**, 147203 (2019).

35. Y. Liu, W. Hou, X. Han, J. Zang, Three-dimensional dynamics of a magnetic hopfion driven by spin transfer torque. *Phys. Rev. Lett.* **124**, 127204 (2020).

36. L. Bo, L. Ji, C. Hu, R. Zhao, Y. Li, J. Zhang, X. Zhang, Spin excitation spectrum of a magnetic hopfion. *Appl. Phys. Lett.* **119**, 212408 (2021)

37. D. Raftrey, P. Fischer, Field-driven dynamics of magnetic Hopfions. *Phys. Rev. Lett.* **127**, 257201 (2021).

38. S. Grytsiuk, J.-P. Hanke, M. Hoffmann, J. Bouaziz, O. Gomonay, G. Bihlmayer, S. Lounis, Y. Mokrousov, S. Blügel, Topological-chiral magnetic interactions driven by emergent orbital magnetism. *Nat. Commun.* **11**, 511 (2020).

39. M. Lukoševičius, J. Herber, Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **3**, 127–149 (2009).

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

41. A. Vansteenkiste, J. Leliaert, M. Dvornik, M. Helsen, F. Garcia-Sanchez, B. V. Waeyenberge, The design and verification of MuMax3. *AIP Adv.* **4**, 107133 (2014).

42. V. M. Kuchkin, K. Chichay, B. Barton-Singer, F. N. Rybakov, S. Blügel, B. J. Schroers, N. S. Kiselev, Geometry and symmetry in skyrmion dynamics. *Phys. Rev. B* **104**, 165116 (2021).

43. T. Dombre, N. Read. Nonlinear σ models for triangular quantum antiferromagnets. *Phys. Rev. B* **39**, 6797–6801 (1989).

44. C. Naya, D. Schubring, M. Shifman, Z. Wang. Skyrmions and hopfions in three-dimensional frustrated magnets. *Phys. Rev. B* **106**, 094434 (2022).

45. S. Wienholdt, D. Hinzke, U. Nowak, THz switching of antiferromagnets and ferrimagnets. *Phys. Rev. Lett.* **108**, 247207 (2012).

46. R. Cheng, D. Xiao, A. Brataas, Terahertz antiferromagnetic spin hall nano-oscillator. *Phys. Rev. Lett.* **116**, 207603 (2016).

47. J. Li, C. B. Wilson, R. Cheng, M. Lohmann, M. Kavand, W. Yuan, M. Aldosary, N. Agladze, P. Wei, M. S. Sherwin, J. Shi, Spin current from sub-terahertz-generated antiferromagnetic magnons. *Nature* **578**, 70–74 (2020).

48. G. Marcucci, D. Pierangeli, C. Conti, Theory of neuromorphic computing by waves: Machine learning by rogue waves, dispersive shocks, and solitons. *Phys. Rev. Lett.* **125**, 093901 (2020).

49. M. Andrychowicz, M. Denil, S. G. Colmenarejo, M. W. Hoffman, D. Pfau, T. Schaul, B. Shillingford, N. de Freitas, Learning to learn by gradient descent, in *Proceedings of the 30th International Conference on Neural Information Processing Systems* (NIPS'16), Barcelona, Spain, 5 to 10 December 2016, pp. 3988–3996.

50. C. Finn, P. Abbeel, S. Levine, Model-agnostic meta-learning for fast adaptation of deep networks, in *Proceedings of the 34th International Conference on Machine Learning* (ICML'17), Sydney, NSW, Australia, 6 to 11 August 2017, pp. 1126–1135.

51. S. H. Oh, S. K. Kim, K. J. Xiao, K. J. Lee, Bidirectional spin-wave-driven domain wall motion in ferrimagnets. *Phys. Rev. B* **100**, 174403 (2019).

52. P. Shen, Y. Tserkovnyak, S. K. Kim, Driving a magnetized domain wall in an antiferromagnet by magnons. *J. Appl. Phys.* **127**, 223905 (2020).

53. S. K. Kim, K. Nakata, D. Loss, Y. Tserkovnyak, Tunable magnonic thermal Hall effect in skyrmion crystal phases of ferrimagnets. *Phys. Rev. Lett.* **122**, 057204 (2019).

54. P. Yan, A. Kamra, Y. Cao, G. E. Bauer, Angular and linear momentum of excited ferromagnets. *Phys. Rev. B* **88**, 144413 (2013).

55. S. K. Kim, Y. Tserkovnyak, O. Tchernyshyov, Propulsion of a domain wall in an antiferromagnet by magnons. *Phys. Rev. B* **90**, 104406 (2014).

56. G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, A. Hirose, Recent advances in physical reservoir computing: A review. *Neural Netw.* **115**, 100–123 (2019).

57. F. Zhou, W. Bin, Z. Li, Deep Meta-Learning: Learning to Learn in the Concept Space; https://arxiv.org/abs/1802.03596 (2018).

58. J. Gladikowski, M. Hellmund, Static solitons with non-zero Hopf number. *Phys. Rev. D* **56**, 5194–5199 (1997).