

NeuralMag: an open-source nodal finite-difference code for inverse micromagnetics

C. Abert,^{1, a)} F. Bruckner,¹ A. Voronov,¹ M. Lang,^{2,3} S. A. Pathak,^{2,3} S. Holt,^{2,3} R. Kraft,¹ R. Allayarov,¹ P. Flauger,¹ S. Koraltan,⁴ T. Schrefl,⁵ A. Chumak,¹ H. Fangohr,^{2,3,6} and D. Suess¹

¹⁾*Faculty of Physics, University of Vienna, 1090 Vienna, Austria*

²⁾*Max Planck Institute for the Structure and Dynamics of Matter, 22761 Hamburg, Germany*

³⁾*Center for Free-Electron Laser Science, Luruper Chaussee 149, 22761 Hamburg, Germany*

⁴⁾*Institute of Applied Physics, TU Wien, 1040 Vienna, Austria*

⁵⁾*Department for Integrated Sensor Systems, University for Continuing Education Krems, 3500 Krems, Austria*

⁶⁾*University of Southampton, SO17 1BJ Southampton, United Kingdom*

(Dated: 19 November 2024)

We present NeuralMag, a flexible and high-performance open-source Python library for micromagnetic simulations. NeuralMag leverages modern machine learning frameworks, such as PyTorch and JAX, to perform efficient tensor operations on various parallel hardware, including CPUs, GPUs, and TPUs. The library implements a novel nodal finite-difference discretization scheme that provides improved accuracy over traditional finite-difference methods without increasing computational complexity. NeuralMag is particularly well-suited for solving inverse problems, especially those with time-dependent objectives, thanks to its automatic differentiation capabilities. Performance benchmarks show that NeuralMag is competitive with state-of-the-art simulation codes while offering enhanced flexibility through its Python interface and integration with high-level computational backends.

I. INTRODUCTION

Micromagnetic simulations are a fundamental tool in the study of magnetization dynamics and play a crucial role in understanding and designing magnetic materials and devices. These simulations model the behavior of magnetic and magnonic systems at the nanoscale, providing insight into phenomena such as domain wall motion, magnetization reversal, and spin wave propagation. The field relies on various computational methods, with finite-difference and finite-element schemes being widely used. Notable examples of established finite-difference codes include OOMMF¹ and fidimag² for CPU-based simulations and mumax3³ and magnum.np⁴ for GPU-accelerated simulations. Finite-element-based methods, such as those implemented in NMag⁵, Tetramag⁶, FastMag⁷, FinMag⁸, and magnum.fe⁹, provide greater flexibility in handling complex geometries but can be computationally more expensive.

In addition to standard micromagnetic simulations, inverse problems have attracted considerable attention in recent years. These problems involve determining the optimal parameters — such as material properties, external fields, or device geometries — that lead to a desired magnetic configuration or device functionality. A significant body of work has focused on inverse modeling of the demagnetization field, a static inverse problem. This has been particularly useful in the context of magnetic 3D printing, where topology optimiza-

^{a)}Electronic mail: claas.abert@univie.ac.at

tion techniques are employed to design optimal material layouts, and the inverse modeling is used to infer the magnetization configuration of printed samples^{10–12}.

More recently, research in the emerging field of inverse magnonics has gained momentum, focusing on optimizing the functionality of magnonic devices. Magnonics uses spin waves (magnons) for information processing, and designing efficient magnonic devices poses complex nonlinear optimization challenges. Inverse-design approaches have been increasingly applied to magnonics, allowing researchers to automate the design of devices by specifying a desired functionality and using computational algorithms to find the optimal configuration^{13–16}.

In this paper, we present a novel discretization strategy for micromagnetic simulations, adjoint-state algorithms for efficiently solving time-dependent inverse problems, and the software design of NeuralMag, which integrates these advancements into a flexible and high-performance computational framework.

II. MICROMAGNETICS

The micromagnetic model provides a semi-classical continuum description of magnetization dynamics in ferromagnetic systems, as originally formulated by Brown¹⁷. The key governing equation is the Landau-Lifshitz-Gilbert (LLG) equation, which reads

$$\frac{\partial \mathbf{m}}{\partial t} = -\frac{\gamma}{1+\alpha^2} \mathbf{m} \times \mathbf{H}_{\text{eff}} - \frac{\alpha\gamma}{1+\alpha^2} \mathbf{m} \times (\mathbf{m} \times \mathbf{H}_{\text{eff}}) \quad (1)$$

with \mathbf{m} being the unit-vector field representation of the magnetization, γ being the reduced gyromagnetic ratio, and α being a dimensionless damping parameter. The effective field \mathbf{H}_{eff} accounts for all relevant interactions within the system and derives from the total energy as

$$\mathbf{H}_{\text{eff}} = -\frac{1}{\mu_0 M_s} \frac{\delta E}{\delta \mathbf{m}} \quad (2)$$

with M_s being the saturation magnetization and $\delta E/\delta \mathbf{m}$ denoting the variational derivative of the energy with respect to the magnetization. When the energy E depends on spatial derivatives of the magnetization field \mathbf{m} , additional boundary conditions must be imposed to solve Eqs. (1) and (2). One such example is the micromagnetic exchange energy, which is defined as

$$E_{\text{ex}} = \int_{\Omega} A (\nabla \mathbf{m})^2 dx, \quad (3)$$

where A is the exchange stiffness constant. The variation of the exchange energy with respect to \mathbf{m} yields

$$\delta E_{\text{ex}}(\mathbf{m}, \delta \mathbf{m}) = \int_{\Omega} \underbrace{-2[\nabla \cdot (A \nabla \mathbf{m})]}_{\equiv \delta E/\delta \mathbf{m}} \cdot \delta \mathbf{m} dx + \int_{\partial \Omega} \underbrace{2A \frac{\partial \mathbf{m}}{\partial \mathbf{n}}}_{\equiv \mathbf{B}(\mathbf{n})} \cdot \delta \mathbf{m} ds \quad (4)$$

leading to the exchange field definition

$$\mathbf{H}_{\text{ex}} = -\frac{1}{\mu_0 M_s} \frac{\delta E_{\text{ex}}}{\delta \mathbf{m}} = \frac{2}{\mu_0 M_s} \nabla \cdot (A \nabla \mathbf{m}). \quad (5)$$

The boundary term in Eq. (4) defines the appropriate exchange boundary condition. To satisfy equilibrium conditions in micromagnetics, the system must fulfill Brown's conditions, which require $\mathbf{m} \times \delta E/\delta \mathbf{m} = 0$ for $\mathbf{x} \in \Omega$, and $\mathbf{m} \times \mathbf{B} = 0$ for $\mathbf{x} \in \partial \Omega$.

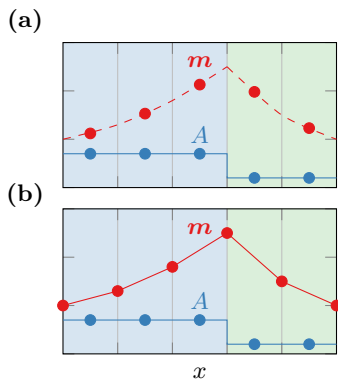


FIG. 1. Illustration of the discretization of the magnetization \mathbf{m} and the material parameter A for a one-dimensional representation of a two-phase magnetic system, using different numerical schemes: (a) Standard finite differences: Both the material parameter and the magnetization are discretized with a single value per simulation cell. The magnetization degrees of freedom are treated as sample points of a continuous function. (b) Finite elements: The material parameters are discretized using piecewise constant functions, while the magnetization is represented as piecewise affine, with degrees of freedom located at the vertices.

A similar variational treatment at internal interfaces, where material parameters vary discontinuously, introduces additional interface conditions¹⁸. Assuming a continuous magnetization across such interfaces and dividing the domain into regions of continuous material parameters, the corresponding interface condition can be written as $\mathbf{m}_1 \times \mathbf{B}_1(\mathbf{n}) = \mathbf{m}_2 \times \mathbf{B}_2(\mathbf{n})$, where $\mathbf{B}_1(\mathbf{n})$ and $\mathbf{B}_2(\mathbf{n})$ represent the boundary terms on either side of the interface.

In case of the exchange energy being the only energy contribution introducing spatial derivatives and furthermore considering $\mathbf{m} \perp \partial\mathbf{m}/\partial n$, this leads to the well known exchange jump condition¹⁷

$$A_1 \frac{\partial \mathbf{m}_1}{\partial \mathbf{n}} = A_2 \frac{\partial \mathbf{m}_2}{\partial \mathbf{n}}. \quad (6)$$

In addition to satisfying equilibrium conditions, the boundary and interface conditions must be consistently fulfilled at all times when solving the Landau-Lifshitz-Gilbert (LLG) equation¹⁸.

III. NODAL FINITE-DIFFERENCE SCHEME

Existing micromagnetic simulation software usually employs either a finite-difference discretization on regular grids^{18,19} or a finite-element discretization on irregular grids^{18,20}. The use of regular cuboid grids in the case of finite-difference micromagnetics allows for a very efficient computation of the demagnetization field by means of an FFT accelerated convolution. On the other hand the finite-element method allows for the accurate modeling of complex structures due to the use of irregular meshes.

Moreover, finite-element micromagnetics provides a more subtle but sometimes highly relevant advantage over finite-difference micromagnetics: In finite-element micromagnetics, the magnetization is usually explicitly defined on each mesh-vertex whereas standard finite-difference tools store one magnetization vector per simulation cell, which is typically taken to be the magnetization in the center of this cell. While this difference appears to be insignificant for the micromagnetic modeling in the bulk, it plays a crucial role when considering material interfaces where the magnetization is subject to boundary and jump conditions. Consider e.g. the exchange jump condition (6), which prescribes a discontinu-

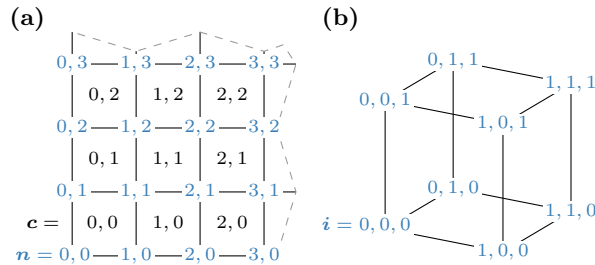


FIG. 2. Cell and vertex numbering using multiindices for nodal finite-differences. (a) Two-dimensional representation of global cell and node indices denoted by \mathbf{c} (black) and \mathbf{n} (blue). (b) Three-dimensional representation of local vertex numbering denoted by index \mathbf{i} (blue).

ity in the first spatial derivative of the magnetization across material interfaces. Choosing the degrees of freedom of the magnetization in the cell centers as illustrated in Fig. 1(a) requires a careful treatment of the boundary conditions that are defined on the vertices²¹. Similar considerations apply to interfacial energy contributions such as the RKKY coupling between two ferromagnetic layers²². Inaccurate modeling of the boundary conditions can lead to a loss of convergence order and consequently introduce significant numerical errors. Introducing more energy contributions depending on surface integrals or spatial derivatives of \mathbf{m} result in more complex boundary conditions¹⁸ that become unfeasible to handle in standard finite-difference micromagnetics.

In contrast, the finite-element method allows for the choice of tailored function spaces for the magnetization and material parameters, as shown in Fig. 1(b). Moreover, the inherently variational nature of the finite-element method allows to solve for the effective-field contributions by directly considering the variation of the energy¹⁸ resulting in the weak form

$$-\int_{\Omega} \mu_0 M_s \mathbf{H}(\mathbf{m}) \cdot \mathbf{v} \, dx = \delta E(\mathbf{m}, \mathbf{v}) \quad \forall \quad \mathbf{v} \in V \quad (7)$$

with V being a sufficiently smooth function space referred to as test space. By a proper choice of function spaces for the material parameters and fields, this procedure does not require to explicitly account for the boundary conditions at all.

A. Local Field Terms

The nodal finite-difference scheme proposed in this work applies the finite-element method for local field contributions on a regular cuboid grid. This enables the use of an FFT accelerated demagnetization-field computation as in standard finite-difference micromagnetics, see Sec. III B, while providing the rigorous and accurate handling of material interfaces for all local field contributions due to finite-element modeling. In order to address the cells and nodes of the regular grid, we introduce multi-indices \mathbf{c} , \mathbf{n} and \mathbf{i} as

$$\mathbf{c} = (c_1, c_2, c_3) \in \{0, \dots, N_1 - 1\} \times \{0, \dots, N_2 - 1\} \times \{0, \dots, N_3 - 1\} \quad (8)$$

$$\mathbf{n} = (n_1, n_2, n_3) \in \{0, \dots, N_1\} \times \{0, \dots, N_2\} \times \{0, \dots, N_3\} \quad (9)$$

$$\mathbf{i} = (i_1, i_2, i_3) \in \{0, 1\} \times \{0, 1\} \times \{0, 1\} \quad (10)$$

with N_1, N_2, N_3 being the number of simulation cells in the respective mesh dimension. The indices \mathbf{c} and \mathbf{n} are used to address simulation cells and nodes respectively according to the numbering introduced in Fig. 2(a). The index \mathbf{i} either acts as local vertex number in a simulation cell according to Fig. 2(b) or more general as a relative index to address neighborhoods.

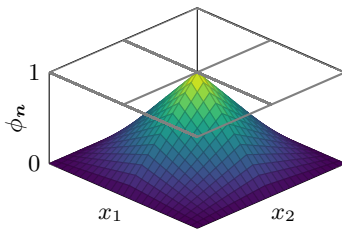


FIG. 3. Two-dimensional representation of a basis function $\phi_{\mathbf{n}}$ in nodal finite-differences with a support spanning 4 simulation cells.

We discretize all continuous fields appearing in weak forms with standard piecewise polynomial and globally continuous basis functions $\phi_{\mathbf{n}}$ that form a nodal basis on the cuboid mesh. Each basis function $\phi_{\mathbf{n}}$ is defined per simulation cell in terms of reference basis functions $\hat{\phi}_{\mathbf{i}}$ as

$$\phi_{\mathbf{n}}(\mathbf{x}) = \sum_{\mathbf{i}} \hat{\phi}_{\mathbf{i}} \left[\begin{array}{l} (x_1/\Delta x_1 - n_1 + i_1) \\ (x_2/\Delta x_2 - n_2 + i_2) \\ (x_3/\Delta x_3 - n_3 + i_3) \end{array} \right] \quad (11)$$

with Δx_k being the simulation-cell size in the k -th dimension. The reference basis functions $\hat{\phi}_{\mathbf{i}}$ are defined on the reference unit cell $\Omega_{\text{ref}} = [0, 1] \times [0, 1] \times [0, 1]$ as

$$\hat{\phi}_{\mathbf{i}}(\mathbf{x}) = \mathbb{1}_{\Omega_{\text{ref}}}(\mathbf{x}) \begin{bmatrix} 1 - i_1 + (2i_1x_1 - x_1) \\ 1 - i_2 + (2i_2x_2 - x_2) \\ 1 - i_3 + (2i_3x_3 - x_3) \end{bmatrix} \quad (12)$$

where $\mathbb{1}_{\Omega_{\text{ref}}}$ denotes the characteristic function of Ω_{ref} which evaluates to 1 if $\mathbf{x} \in \Omega_{\text{ref}}$ and to 0 else. This restricts the support of the reference basis functions $\hat{\phi}_{\mathbf{i}}$ to the reference cell Ω_{ref} . A 2D representation of a basis function is visualized in Fig. 3. Furthermore, we introduce vector basis functions as

$$\boldsymbol{\phi}_{\mathbf{n},j} = \phi_{\mathbf{n}} \mathbf{e}_j \quad (13)$$

with \mathbf{e}_j being the unit vector in direction $j \in \{1, 2, 3\}$. Continuous vector fields such as the magnetization \mathbf{m} and the effective field \mathbf{H}_{eff} are then discretized as

$$\mathbf{m} \rightarrow \mathbf{m}^{\text{h}} = \sum_{\mathbf{n},j} m_{\mathbf{n},j} \boldsymbol{\phi}_{\mathbf{n},j} \quad (14)$$

with the superscript h denoting the discretized version of a field and coefficients $m_{\mathbf{n},j}$ being the nodal values of this field.

For material parameters, such as the saturation magnetization M_s , we choose a piecewise constant function space in order to allow for the accurate modeling of rapid material interfaces. Namely, we define these parameters per simulation cell resulting in the following discretization

$$M_s \rightarrow M_s^{\text{h}} = \sum_{\mathbf{c}} M_{s,\mathbf{c}} \vartheta_{\mathbf{c}} \quad (15)$$

with basis functions

$$\vartheta_{\mathbf{c}} = \mathbb{1}_{\Omega_{\mathbf{c}}}. \quad (16)$$

Replacing all fields with their discretized counterparts in the weak form (7) and testing with individual basis functions instead of arbitrary test functions yields the discretized weak form

$$- \int_{\Omega} \mu_0 M_s \mathbf{H}^{\text{h}}(\mathbf{m}^{\text{h}}) \cdot \boldsymbol{\phi}_{\mathbf{n},j} \, d\mathbf{x} = \delta E(\mathbf{m}^{\text{h}}, \boldsymbol{\phi}_{\mathbf{n},j}) \quad \forall \quad \mathbf{n}, j. \quad (17)$$

For a given node \mathbf{n} , we split the variation $\delta E(\mathbf{m}^h, \phi_{\mathbf{n},j})$ into its contributions from the 8 simulation cells that share node \mathbf{n} and we address these cells by the local index $\mathbf{i} \in \{0,1\}^3$. In general, the variation over a single simulation cell depends on the magnetization values of all nodes of this cell. Considering the three components of the magnetization, the contribution of the cell \mathbf{i} to the variation can be written as

$$\delta E_j^{*\mathbf{i}} = F_{\mathbf{i},j}(m_{\mathbf{i}',j'}) \quad \text{for } \mathbf{i}' \in \{0,1\}^3 \text{ and } j' \in \{1,2,3\} \quad (18)$$

where $m_{\mathbf{i}',j'}$ denotes all nodal values of the magnetization in cell \mathbf{i} . If the energy E is quadratic in \mathbf{m} , the function F is linear in $m_{\mathbf{i}',j'}$ and can be described by a 24×24 matrix considering the $2^3 \cdot 3$ degrees of freedom defined by the index pairs \mathbf{i}, j and \mathbf{i}', j' . In the finite-element context this matrix is usually referred to as element matrix of the weak form.

In order to compute the variation at all nodes, we introduce the vector $\delta \mathbf{E}$ with components $\delta E_{\mathbf{n},j} = \delta E(\mathbf{m}^h, \phi_{\mathbf{n},j})$ and the auxiliary vectors $\delta \mathbf{E}^{*\mathbf{i}}$ containing the cell-wise variations according to (18) for all nodes. Considering the node and cell numbering introduced in Fig. 2, the global node index is given by the global cell index and the relative node index as $\mathbf{n}(\mathbf{c}, \mathbf{i}) = \mathbf{c} + \mathbf{i}$ resulting in

$$\delta E_{\mathbf{c}+\mathbf{i},j}^{*\mathbf{i}} = F_{\mathbf{i},j}(m_{\mathbf{c}+\mathbf{i}',j'}) \quad \text{for } \mathbf{i}' \in \{0,1\}^3 \text{ and } j' \in \{1,2,3\} \quad (19)$$

$$\delta \mathbf{E} = \sum_{\mathbf{i}} \delta \mathbf{E}^{*\mathbf{i}}. \quad (20)$$

Note that F only depends on the relative index \mathbf{i} and the component j . Eqs. (19) and (20) deliver a straight-forward strategy for a parallel evaluation over the cell index \mathbf{c} , see Sec. V A.

If the energy E depends on further fields, such as an external field or material parameters, the mapping function can be easily extended by adding additional arguments

$$F_{\mathbf{i},j}(m_{\mathbf{c}+\mathbf{i}',j'}, a_{\mathbf{c}+\mathbf{i}'}^1, a_{\mathbf{c}+\mathbf{i}'}^2, \dots, b_{\mathbf{c}}^1, b_{\mathbf{c}}^2, \dots) \quad (21)$$

where the variables $a_{\mathbf{i}+j}^l$ are the coefficients of arbitrary scalar fields discretized with nodal basis functions (11) and the variables $b_{\mathbf{i}}^l$ are the coefficients of arbitrary scalar fields discretized with cell basis functions (16). Since $F_{j,k}$ does not explicitly depend on the cell index \mathbf{i} , it is fully determined by the integrand of the weak form (17) and the dimensions of a single simulation cell $\Omega_{\mathbf{i}}$.

In order to determine the discretized effective field \mathbf{H}^h , the weak form requires the solution of a linear mass system defined by the left-hand side of (17). To avoid this costly procedure, we employ mass lumping to the left-hand side of (7) as described in Abert¹⁸ resulting in

$$H_{\mathbf{n},j} = - \left[\int_{\Omega} \mu_0 M_s^h \phi_{\mathbf{n}} \, d\mathbf{x} \right]^{-1} \delta E_{\mathbf{n},j} \quad (22)$$

where the saturation magnetization M_s is discretized cell-wise according to (15).

The proposed method is applicable to any energy contribution whose density depends solely on the magnetization and its first-order spatial derivatives, such as Zeeman energy, crystalline anisotropies, and both symmetric and antisymmetric exchange interactions. Due to the regularity of the cuboidal grid, a matrix-free implementation of the presented scheme is straight-forward. The local support of the basis functions results in a computational complexity of $\mathcal{O}(N)$ for the evaluation of any local field term with N being the number of simulation cells.

B. Demagnetization Field

To compute the demagnetization field, we employ the well-established FFT-accelerated method commonly used in standard finite-difference micromagnetic simulations¹⁹. This

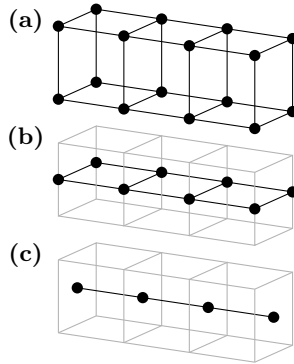


FIG. 4. Representation of the degrees of freedom for a square-shaped rod using the following methods: (a) full three-dimensional discretization, (b) two-dimensional discretization with basis functions that are constant along the third dimension, and (c) one-dimensional discretization with basis functions that are constant along both the second and third dimensions.

algorithm calculates the demagnetization field generated by homogeneously magnetized cuboidal simulation cells arranged on a regular grid through fast convolution. Since this method requires both the magnetization and the resulting field to be specified for each simulation cell, we introduce a straightforward pre- and post-processing step. This procedure averages the values to transition between nodal and cell-centered discretizations efficiently. FFT-accelerated methods that operate directly on node-wise discretized magnetizations have been proposed in previous studies^{23,24}. However, we opt for the standard method based on homogeneously magnetized cuboids due to its advantages in memory efficiency and computational performance specifically for 2D computations where the FFT also reduces to two dimensions.

C. Low-Dimensional Geometries

Discretizing a mesh with $N_1 \times N_2 \times N_3$ cells results in $(N_1 + 1) \times (N_2 + 1) \times (N_3 + 1)$ degrees of freedom when using a nodal basis for the function discretization. In bulk system simulations, this introduces only a negligible overhead in comparison to standard finite-difference schemes, where the degrees of freedom are equal to the number of simulation cells. However, a significant application area for micromagnetic simulations involves magnetic thin films, which are often discretized with just a single layer of simulation cells. In such cases, the 3D nodal discretization introduces a notable overhead – roughly doubling the computational cost – because it requires separate descriptions for the top and bottom surfaces of the thin film. This contrasts with standard finite differences, where the problem effectively reduces to a 2D formulation. By transitioning to 2D basis functions while maintaining full 3D integration in the weak form (17), the nodal finite-difference scheme can accurately describe magnetic thin films. This approach reduces the degrees of freedom to $N_1 \times N_2 \times 1$, making it more efficient for thin film simulations. Namely, the 2D basis function on the reference cell are chosen as

$$\hat{\phi}_{\mathbf{i}}(\mathbf{x}) = \mathbb{1}_{\Omega_{\text{ref}}}(\mathbf{x}) [1 - i_1 + (2i_1x_1 - x_1)] \cdot [1 - i_2 + (2i_2x_2 - x_2)] \quad (23)$$

with a 2D multiindex $\mathbf{i} = (i_1, i_2) \in \{0, 1\}^2$. As illustrated in Fig. 4 this approach can be also generalized to 1D problems leading to basis functions

$$\hat{\phi}_i(\mathbf{x}) = \mathbb{1}_{\Omega_{\text{ref}}}(\mathbf{x}) [1 - i + (2ix_1 - x_1)] \quad (24)$$

with a scalar index $i \in \{0, 1\}$.

IV. INVERSE MICROMAGNETICS

In addition to employing a nodal finite-difference scheme, NeuralMag is specifically designed to address inverse problems in both space and time domains. In this context, the computation of individual field terms or the solution of the Landau-Lifshitz-Gilbert (LLG) equation (1) is classified as a forward problem F . Given a vector of design variables $\boldsymbol{\theta}$, which may include material properties or the initial magnetization configuration, these forward problems yield well-defined outputs \mathbf{y} , such as effective field contributions or the resulting magnetization trajectory

$$F(\boldsymbol{\theta}) = \mathbf{y}. \quad (25)$$

An inverse problem is formulated to determine the design variables $\boldsymbol{\theta}$ that yield a specified result \mathbf{y} from the forward problem. This task is often challenging, as inverse problems are typically ill-posed, and their solution vectors may encompass a large number of degrees of freedom. The most common strategy to solve such problem is the reformulation in terms of a minimization problem

$$\min_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}) \quad \text{with} \quad \mathcal{L}(\boldsymbol{\theta}) = \|F(\boldsymbol{\theta}) - \mathbf{y}\| \quad (26)$$

that might be complemented by additional terms for regularization or smoothing purposes. In the case of a high dimensional input $\boldsymbol{\theta}$ and a nonlinear function F this problem is nontrivial. In such cases, iterative methods, typically based on the gradient of the functional $\nabla_{\boldsymbol{\theta}} \mathcal{L}$, are commonly employed to find a solution. NeuralMag uses automatic differentiation²⁵ for static problems such as inverse strayfield calculations. In contrast to the adjoint method that has been used in previous works^{10,11}, this approach performs the differentiation on the discrete level (discretize first). As for the adjoint method, the gradient computation requires a forward solve and a subsequent backward solve with the complexity of the backward solve being equivalent to that of the forward solve.

For time-dependent problems, NeuralMag implements the adjoint-state method²⁶. The adjoint-state method is a powerful tool for the solution of PDE-constrained optimization problems also referred to as optimal-control problems. Given a forward problem

$$\frac{\partial \mathbf{m}}{\partial t} = \mathbf{f}_{\boldsymbol{\theta}}(t, \mathbf{m}), \quad (27)$$

with design variables $\boldsymbol{\theta}$, we define an objective functional

$$\mathcal{L}(\boldsymbol{\theta}) = L(\mathbf{m}(T; \boldsymbol{\theta}), \mathbf{y}_{\text{target}}) \quad (28)$$

with $\mathbf{m}(T; \boldsymbol{\theta})$ being the solution of (27) for a final time T and $\mathbf{y}_{\text{target}}$ being the desired output of the forward problem. In order to compute the gradient of the objective functional with respect to the design variables $\nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta})$, the adjoint-state method requires two steps. In the first step, the forward problem (27) is solved for the given design variables $\boldsymbol{\theta}$, which results in the output $\mathbf{m}_{\text{output}} = \mathbf{m}(T)$. In the second step, the so-called adjoint problem is solved, which is given by the following system of ODEs

$$\begin{cases} \frac{\partial \mathbf{m}}{\partial t} = \mathbf{f}_{\boldsymbol{\theta}}(t, \mathbf{m}) & \text{with} \quad \mathbf{m}(T) = \mathbf{m}_{\text{output}} \\ \frac{\partial \mathbf{a}}{\partial t} = -\frac{\partial \mathbf{f}_{\boldsymbol{\theta}}(t, \mathbf{m})}{\partial \mathbf{m}} \mathbf{a} & \text{with} \quad \mathbf{a}(T) = \nabla_{\mathbf{y}} L(\mathbf{y}, \mathbf{y}_{\text{target}}), \mathbf{y} = \mathbf{m}_{\text{output}} \\ \frac{\partial \mathbf{u}}{\partial t} = -\frac{\partial \mathbf{f}_{\boldsymbol{\theta}}(t, \mathbf{m})}{\partial \boldsymbol{\theta}} \mathbf{a} & \text{with} \quad \mathbf{u}(T) = \mathbf{0}. \end{cases} \quad (29)$$

with \mathbf{a} being the so-called adjoint variable. This system is solved backwards in time starting from the final time T used in the forward pass. Successful integration of the system yields the output $\mathbf{u}(0)$ which can be identified as the desired gradient of the objective

$$\mathbf{u}(0) = \nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}). \quad (30)$$

While the objective (28) depends solely on the magnetization at the final time T , extending this method to objectives depending on multiple time points T_i can be done in a straightforward fashion by adding appropriate terms depending on $\mathbf{m}(T_i; \boldsymbol{\theta})$ to (28). The computational and storage complexity of the adjoint system is comparable to that of the forward problem, yielding an exceptionally efficient strategy for the gradient computation of PDE-constrained optimization. This method is superior to the backpropagation method^{13,14} with regard to the storage requirements that are similar to a regular forward pass. However, this advantage comes at the cost of a reduced accuracy which is caused by the backwards pass that reconstructs the magnetization trajectory by inverse integration instead of using the exact values from the forward pass.

V. IMPLEMENTATION

NeuralMag is a Python library designed specifically for micromagnetic simulations, with a focus on high-performance tensor computations. A key feature of NeuralMag is its ability to operate with either PyTorch²⁷ or JAX^{28,29} as computational backend, allowing users to select the framework that best suits their needs. By leveraging these modern machine learning frameworks, NeuralMag achieves efficient computations on a variety of parallel hardware, including CPUs, GPUs, and TPUs. This versatility is complemented by the advantages these frameworks offer, such as optimized performance for large-scale simulations and built-in support for automatic differentiation, which simplifies solving inverse problems. Through the modular design, the software is prepared to simplify the use of other computational backends in the future.

The use of either PyTorch or JAX as backends allows NeuralMag to fully exploit the unique strengths of each framework. PyTorch’s `torch.compile()` feature enables just-in-time (JIT) compilation, optimizing the computational workflow by reducing operation overhead and enabling kernel fusion for faster execution on compatible hardware. However, PyTorch currently has limitations when compiling complex functions, such as those involving the demagnetization field, which means `torch.compile()` can only be applied to certain field terms.

In contrast, JAX’s `jit()` function can be applied to the entire right-hand side of the LLG equation. This capability allows JAX to significantly reduce Python overhead and leads to notable performance gains, particularly for smaller systems where the overhead would otherwise be a bottleneck.

Both backends support single- and double-precision computations, enabling NeuralMag to offer users flexibility in balancing computational speed with numerical accuracy according to the requirements of each simulation. The dual-backend approach ensures that NeuralMag can adapt to the user’s preferred ecosystem while maintaining high computational efficiency and flexibility.

A. Form Compilation

```
def e_expr(m, dim):
    A = Variable("material_A", "c" * dim)
    return (A * (
        m.diff(N.x).dot(m.diff(N.x)) +
        m.diff(N.y).dot(m.diff(N.y)) +
        m.diff(N.z).dot(m.diff(N.z))
    ) * dV(dim))
```

Listing 1. Symbolic definition the exchange energy (3) in NeuralMag.

```
def h(dx, m, material_A, material_Ms, rho):
    h = torch.zeros_like(m)
```

```

h[: -1, : -1, : -1, 0] += material__A[...] * rho[...] * (
    m[: -1, : -1, : -1, 0] * (
        0.2222222222222222 * dx[0] * dx[1] / dx[2] +
        0.2222222222222222 * dx[0] * dx[2] / dx[1] +
        0.2222222222222222 * dx[1] * dx[2] / dx[0]) +
    m[: -1, : -1, 1 :, 0] * (
        ...
    )
    ...
)
h[: -1, : -1, : -1, 1] += ...
h[: -1, : -1, : -1, 2] += ...
h[: -1, : -1, 1 :, 0] += ...
...
return h / mass

```

Listing 2. Automatically generated code for the computation of the exchange field.

At the heart of NeuralMag is a form compiler that translates a symbolic representation of a finite-element weak form into efficient tensor operations tailored to the chosen backend. For symbolic computation, NeuralMag leverages the Python library SymPy³⁰. SymPy provides a powerful framework for representing the mathematical structures involved in micromagnetic simulations. Specifically, NeuralMag introduces custom SymPy symbols to represent functions that are discretized either node-wise or cell-wise, as described in Sec. III of this paper. Users can define the weak form of the micromagnetic problem using SymPy’s symbolic language, allowing them to work in an intuitive mathematical formulation.

In addition to defining weak forms symbolically, NeuralMag leverages SymPy to automatically perform the variation of a symbolic energy expression, allowing it to derive the corresponding weak form. This capability streamlines the process of converting complex energy functionals into their weak form representations. For instance, in Lst. 1, the exchange energy is defined symbolically using SymPy, demonstrating how users can express physical energy terms within the framework.

NeuralMag’s form compiler processes the symbolic weak form and transforms it into the discrete mapping function $F_{i,j}$, as defined in (19). This transformation is achieved by applying Gauss quadrature to integrate over the finite elements, converting the weak form into a set of tensor operations—primarily multiplications and summations—that can be efficiently executed by the selected backend. The role of the relative cell index, as discussed in equations (19) and (20), is handled by tensor slicing. This involves slicing along specific tensor dimensions by removing either the first `[1:]` or the last `[:-1]` value of the tensor in that dimension, which is necessary for handling the spatial relationships between adjacent cells in the discretized domain. This systematic conversion of symbolic expressions into backend-specific tensor operations is key to NeuralMag’s high-performance computational capabilities. An example code snippet for the PyTorch backend, generated from the exchange energy defined in Lst. 1, is shown in Lst. 2. The generated function is highly optimized, as it operates solely on raw tensor objects without introducing any loops or conditional statements. This structure ensures that the function is ideally suited for optimization by the just-in-time (JIT) compilers of both PyTorch and JAX. By avoiding control flow statements, the generated code can be compiled into efficient low-level machine instructions, maximizing performance on parallel hardware architectures.

B. Dynamic Attributes

```

>>> state = State(...)
>>> state.a = 1.0
>>> state.b = lambda a: 2.0 * a
>>> state.c = 5.0
>>> state.d = lambda b, c: b + c
>>> print(state.d)

```

7.0

Listing 3. Example usage of dynamic attributes in NeuralMag.

```
def lmda(a, c):
    b = __b(a)
    return __lmda(b, c)
```

Listing 4. Automatically generated function for the evaluation of the dynamic attribute d.

NeuralMag introduces the concept of dynamic attributes through its state object, which allows attributes to be either tensors or functions that depend on tensors and return tensors. This flexible design enables dynamic relationships between attributes, where some can be defined as functions of others, with NeuralMag automatically managing these dependencies. For example, consider the code in Lst. 3: attributes `a`, `b`, `c`, and `d` are defined, where `b` depends on `a`, and `d` depends on both `b` and `c`. When `state.d` is accessed, NeuralMag resolves these dependencies, and the output is `7.0` because `d` is computed as the sum of `b` (which is $2 * a = 2.0$) and `c` (which is 5.0). Importantly, instead of scalar values, any tensor can be used as an attribute, allowing for more complex operations on multidimensional data.

When defining such dynamic attributes, NeuralMag analyzes the function signatures to identify all dependencies in a recursive manner. It then generates a new Python function at runtime that only relies on pure tensors and eliminates any control structures, such as loops or conditionals, ensuring the function remains optimal for high-performance tensor computation. In the case of the example from Lst. 3, the dynamically created function looks like Lst. 4, where `d` depends on `b`. Although `b` is not explicitly listed in the function arguments, its dependency on `a` is automatically resolved within the body of the function. This approach simplifies the handling of complex dependencies, while maintaining the computational efficiency needed for the PyTorch and JAX backends.

C. Automatic Differentiation and Time Integration

In the context of inverse problems, NeuralMag leverages automatic differentiation and efficient time integration to solve complex optimization tasks. Both PyTorch and JAX offer powerful automatic differentiation capabilities, which are crucial for computing gradients with respect to parameters in inverse problems. For time integration, NeuralMag integrates with `torchdiffeq`³¹ (for PyTorch) and `diffrax`³² (for JAX), both of which provide support for solving ordinary differential equations (ODEs).

Time integration is essential in dynamic micromagnetic problems, where the system’s evolution must be accurately tracked. Both libraries support a variety of numerical schemes for time stepping, including Euler methods, Runge-Kutta methods (such as RK4), and adaptive solvers like the Dormand-Prince method. These methods ensure that NeuralMag can flexibly adapt to different accuracy and performance requirements in dynamic simulations.

For gradient-based optimization in inverse problems, NeuralMag supports both the adjoint method²⁶ and traditional backpropagation. The adjoint method is particularly well-suited for problems with long time horizons or large state spaces, as it computes gradients more efficiently by solving an adjoint ODE backward in time. Both `torchdiffeq` and `diffrax` support the adjoint method for time integration, offering an efficient way to compute gradients when optimizing over dynamic systems. At the same time, they also allow for direct backpropagation through the time integration process, which can be more straightforward for shorter time intervals or simpler problems.

By combining automatic differentiation with advanced time integration techniques, NeuralMag can effectively tackle inverse problems in micromagnetic simulations, allowing users to optimize parameters while ensuring accurate numerical solutions over time.

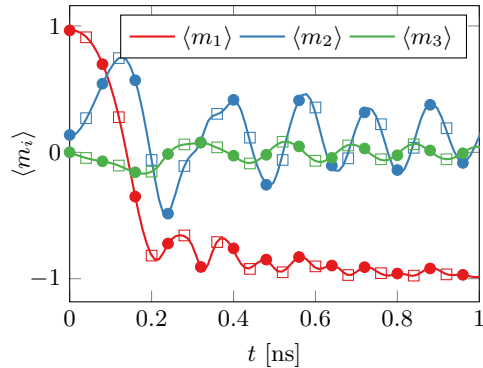


FIG. 5. Results for MuMag Standard Problem #4 are presented using both 2D and 3D discretizations as computed by NeuralMag. The reference solution, computed with OOMMF¹, is depicted by solid lines for comparison. The NeuralMag solutions are illustrated using circles for the 3D discretization and squares for the 2D discretization.

| discontinuous parameters | analytical [T] | magnum.af [T] | NeuralMag [T] |
|--------------------------|----------------|---------------|---------------|
| $A/K/M_s$ | 1.568 | 1.585 | 1.580 |
| A/K | 1.089 | 1.116 | 1.112 |
| A/M_s | 1.206 | 1.256 | 1.205 |
| A | 0.838 | 0.868 | 0.867 |
| K/M_s | 1.005 | 1.020 | 1.012 |
| K | 0.565 | 0.582 | 0.571 |

TABLE I. Depinning fields for a domain wall in a two-phase magnet, as defined in²¹, computed with NeuralMag and compared to the analytical and numerical reference solutions computed with magnum.af.

VI. VALIDATION AND BENCHMARKS

To validate the accuracy of NeuralMag, we solve two significant micromagnetic problems. These tests showcase NeuralMag’s ability to handle both standard and advanced cases, verifying its precision and computational efficiency.

The first validation case is MuMag Standard Problem #4³³, which simulates the dynamic behavior of a thin ferromagnetic film under an applied magnetic field. The focus is on the time evolution of the averaged magnetization components. We solve this problem using a full 3D spatial discretization and compare the results to a 2D simulation as described in Sec. III C of the paper. The results, displayed in Fig. 5, show excellent agreement with the reference solutions from the MuMag community, demonstrating the precision of NeuralMag in simulating the time dynamics of micromagnetic systems both with the 3D as well as 2D thin-film approximation.

The second validation case involves solving the domain wall pinning problem proposed by Heistracher et al.²¹. This problem focuses on calculating the coercive field required to unpin a domain wall at the interface between two magnetic phases with varying material properties, such as exchange interaction, uniaxial anisotropy, and spontaneous magnetization. This problem is sensitive to discontinuities in these parameters, making it an ideal test for NeuralMag’s handling of complex material boundaries.

In this validation, we compare the switching fields calculated by NeuralMag with the analytical results provided in Tab. 1 of the original paper. We varied the material parameters (exchange constant A , anisotropy constant K , and saturation magnetization M_s) in different combinations across the two magnetic phases. Tab. I compares the switching fields obtained

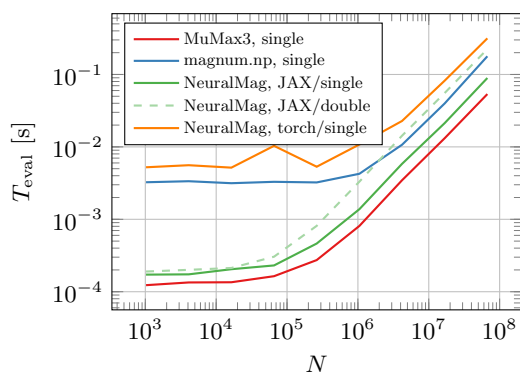


FIG. 6. Comparison of the computation time for evaluating the right-hand side of the Landau-Lifshitz-Gilbert (LLG) equation, including both the exchange and demagnetization fields, across various system sizes N , in comparison with other finite-difference codes. The legend indicates the code as well as the floating-point precision used for the computation.

using NeuralMag with those presented in the reference paper. Our results closely match the analytical solutions, with minor deviations likely due to the time integration method and field rate used during the simulation. These successful validations confirm that NeuralMag correctly handles discontinuities at material interfaces and provides accurate predictions for complex micromagnetic systems.

To evaluate the performance of NeuralMag, we conducted a throughput benchmark, shown in Fig. 6, where we compare the time required for evaluating the right-hand side (RHS) of the Landau-Lifshitz-Gilbert (LLG) equation across different system sizes. Specifically, we measure the time for the integration of the full LLG including the exchange and demagnetization field and then divide by the number of field evaluations. This procedure can be easily applied to any micromagnetic code without the need to modify it and provides a robust measure of the over-all performance at the same time. In this benchmark, NeuralMag is compared to two widely-used micromagnetic simulation tools: mumax3³ and magnum.np⁴. mumax3 shows the best performance due to its highly optimized GPU implementation. However, NeuralMag, when using JAX as the backend, almost matches the performance of mumax3, being less than a factor of 2 slower.

Remarkably, NeuralMag maintains this competitive performance even for small system sizes, despite the computational overhead typically associated with a Python implementation. This performance can be attributed to the just-in-time (JIT) compilation feature of JAX, which optimizes the entire RHS of the LLG equation at runtime. Thanks to NeuralMag’s architecture, JAX is able to analyze and compile the full computation into highly optimized machine code, reducing overhead and achieving near-optimal execution times. This demonstrates the strength of NeuralMag’s design in leveraging modern machine learning frameworks to achieve high-performance computations while maintaining flexibility.

The remaining performance gap of approximately a factor of two compared to MuMax3 likely arises from the demagnetization field computation, as the current FFT interface in JAX is limited, preventing certain optimizations. However, as JAX’s FFT capabilities expand, this gap could narrow significantly — or even vanish completely — in the future.

VII. INVERSE PROBLEMS

```
state = State(...)
state.angles = [jnp.pi / 2, jnp.pi / 2]
h_ext = lambda angles: jnp.stack(
    [
        Hc / 2 * jnp.sin(angles[0]) * jnp.cos(angles[1]),
```

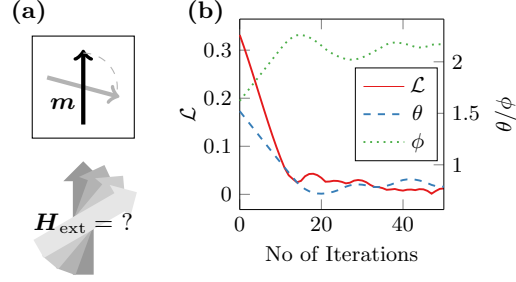


FIG. 7. Simple inverse micromagnetic problem for the optimization of the external field direction in order to align the magnetization of a single-domain particle in a given direction. (a) Sketch of the problem setup. (b) Convergence of the objective function \mathcal{L} and the optimized field angles θ and ϕ .

```

    Hc / 2 * jnp.sin(angles[0]) * jnp.sin(angles[1]),
    Hc / 2 * jnp.cos(angles[0]),
    ]
)
...
llg = nm.LLGSolver(state, parameters=["angles"])
m_target = nm.VectorFunction(state).fill((0.5**0.5, 0, 0.5**0.5)).tensor

def loss(angles, args):
    m_pred = llg.solve(state.tensor([0.0, 0.05e-9]), angles).ys[-1]
    return jnp.mean((m_target - m_pred) ** 2)

solver = optx.BFGS(1e-3, 1e-3, optx.max_norm)
result = optx.minimise(loss, solver, state.angles)

```

Listing 5. Simulation script for inverse problem.

To demonstrate the solution of time-dependent inverse problems using NeuralMag, we solve a straightforward optimization problem. Specifically, we aim to optimize the direction of an external magnetic field to align the magnetization of a single-domain particle with a target configuration $\mathbf{m}_{\text{target}}$ after a given time T , see Fig. 7(a). The optimization minimizes the objective function \mathcal{L} with respect to the field angles θ and ϕ , as defined by the system

$$\mathbf{H}_{\text{ext}}(\theta, \phi) = H_c \begin{pmatrix} \sin(\theta) \cos(\phi) \\ \sin(\theta) \sin(\phi) \\ \cos(\theta) \end{pmatrix}, \quad (31)$$

$$\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{aniso}} + \mathbf{H}_{\text{exchange}} + \mathbf{H}_{\text{ext}}, \quad (32)$$

$$\mathcal{L}(\theta, \phi) = \int_{\Omega} \|\mathbf{m}(T) - \mathbf{m}_{\text{target}}\| d\mathbf{x} \quad (33)$$

with $\mathbf{m}(t)$ being constrained by the LLG (1). A shortened code listing demonstrating the setup for this inverse problem is provided in Lst. 5. In this simple optimization, convergence is achieved after 30-50 gradient-descent steps, see Fig. 7(b). NeuralMag computes the gradient of the objective function by performing one forward and one backward simulation of the dynamic problem.

VIII. CODE DEVELOPMENT AND AVAILABILITY

The source code of NeuralMag is publicly available under the GNU LGPL License on GitLab at <https://gitlab.com/neuralmag/neuralmag>³⁴. Comprehensive documentation, an API reference, and tutorials are available at <https://neuralmag.gitlab.io/>.

NeuralMag can be installed via standard package managers such as pip or conda. Users and community members are encouraged to contribute to the codebase, tutorials, and documentation. Continuous integration workflows are set up using GitLab CI/CD to automatically run tests after every code change. These tests are run with both the PyTorch and the JAX backend and include unit, integration, and system tests, covering both the standard problems and benchmarks. The repository includes all numerical problems discussed in this paper, as well as the code to reproduce the benchmarks.

IX. CONCLUSION

In this paper, we have introduced NeuralMag, an open-source Python library for micromagnetic simulations that leverages modern machine learning frameworks such as PyTorch and JAX to achieve high performance. NeuralMag implements a novel nodal finite-difference discretization scheme, which provides a rigorous numerical description of continuous fields such as the magnetization as well as discontinuous material parameters. This approach is particularly useful for the accurate modeling of material interfaces, while maintaining the same computational complexity as standard finite-difference schemes. Its performance is competitive with state-of-the-art micromagnetic simulation codes, yet it offers unparalleled flexibility due to its Python-based interface and support for optimized tensor operations on a variety of hardware platforms.

NeuralMag is especially well-suited for solving inverse problems, particularly those with time-dependent objectives, thanks to its ability to seamlessly compute gradients using automatic differentiation. This makes it a powerful tool for a wide range of optimization and simulation tasks in micromagnetics. NeuralMag is freely available³⁴, making it accessible to the broader research community for further development and application.

X. ACKNOWLEDGEMENTS

This research was funded in whole or in part by the Austrian Science Fund (FWF) 10.55776/P34671, 10.55776/I6068, and 10.55776/PAT3864023. For open access purposes, the author has applied a CC BY public copyright license to any author-accepted manuscript version arising from this submission. We gratefully acknowledge the wedding of the Koraltans for fruitful discussions and great time which led to this publication.

¹M. J. Donahue and D. G. Porter, “OOMMF user’s guide, version 1.0,” (1999).

²M.-A. Bisotti, D. Cortés-Ortuño, R. A. Pepper, W. Wang, M. Beg, T. Kluyver, and H. Fangohr, “Fidimag—a finite difference atomistic and micromagnetic simulation package,” arXiv preprint arXiv:2002.04318 (2020).

³A. Vansteenkiste, J. Leliaert, M. Dvornik, M. Helsen, F. Garcia-Sanchez, and B. Van Waeyenberge, “The design and verification of MuMax3,” AIP Adv. **4** (2014), 10.1063/1.4899186.

⁴F. Bruckner, S. Koraltan, C. Abert, and D. Suess, “magnum. np: a pytorch based gpu enhanced finite difference micromagnetic simulation framework for high level development and inverse design,” Scientific Reports **13**, 12054 (2023).

⁵T. Fischbacher, M. Franchin, G. Bordignon, and H. Fangohr, “A systematic approach to multiphysics extensions of finite-element-based micromagnetic simulations: Nmag,” IEEE Transactions on Magnetics **43**, 2896–2898 (2007).

⁶A. Kakay, E. Westphal, and R. Hertel, “Speedup of fem micromagnetic simulations with graphical processing units,” IEEE transactions on magnetics **46**, 2303–2306 (2010).

⁷R. Chang, S. Li, M. Lubarda, B. Livshitz, and V. Lomakin, “Fastmag: Fast micromagnetic simulator for complex magnetic structures,” Journal of Applied Physics **109** (2011).

⁸M.-A. Bisotti, M. Beg, W. Wang, M. Albert, D. Chernyshenko, D. Cortés-Ortuño, R. A. Pepper, M. Vouden, R. Carey, H. Fuchs, *et al.*, “Finmag: finite-element micromagnetic simulation tool (version 1.0),” Zenodo (2018).

⁹C. Abert, L. Exl, F. Bruckner, A. Drews, and D. Suess, “magnum. fe: A micromagnetic finite-element simulation code based on fenics,” Journal of Magnetism and Magnetic Materials **345**, 29–35 (2013).

¹⁰F. Bruckner, C. Abert, G. Wautischer, C. Huber, C. Vogler, M. Hinze, and D. Suess, “Solving large-scale inverse magnetostatic problems using the adjoint method,” Scientific reports **7**, 40816 (2017).

- ¹¹C. Abert, C. Huber, F. Bruckner, C. Vogler, G. Wautischer, and D. Suess, “A fast finite-difference algorithm for topology optimization of permanent magnets,” *Journal of Applied Physics* **122** (2017).
- ¹²C. Huber, C. Abert, F. Bruckner, C. Pfaff, J. Kriwet, M. Groenefeld, I. Teliban, C. Vogler, and D. Suess, “Topology optimized and 3d printed polymer-bonded permanent magnets for a predefined external field,” *Journal of Applied Physics* **122** (2017).
- ¹³Á. Papp, W. Porod, and G. Csaba, “Nanoscale neural network using non-linear spin-wave interference,” *Nat Commun* **12**, 6422 (2021).
- ¹⁴Q. Wang, A. V. Chumak, and P. Pirro, “Inverse-design magnonic devices,” *Nat Commun* **12**, 2636 (2021).
- ¹⁵Z. Yan, Y. Xing, and X. Han, “Inverse design of magnonic filter,” *Journal of Magnetism and Magnetic Materials* **563**, 169976 (2022).
- ¹⁶N. Zenbaa, C. Abert, F. Majcen, M. Kerber, R. Serha, S. Knauer, Q. Wang, T. Schrefl, D. Suess, and A. Chumak, “Magnonic inverse-design processor,” arXiv preprint arXiv:2403.17724 (2024).
- ¹⁷W. F. Brown Jr, “Micromagnetics, domains, and resonance,” *Journal of Applied Physics* **30**, S62–S69 (1959).
- ¹⁸C. Abert, “Micromagnetics and spintronics: Models and numerical methods,” *Eur. Phys. J. B* **92**, 1–45 (2019).
- ¹⁹J. E. Miltat and M. J. Donahue, “Numerical micromagnetics: Finite difference methods,” *Handbook of magnetism and advanced magnetic materials* **2**, 742–764 (2007).
- ²⁰T. Schrefl, G. Hrkac, S. Bance, D. Suess, O. Ertl, and J. Fidler, “Numerical micromagnetics (finite element method),” *Handbook of magnetism and advanced magnetic materials* **2**, 765 – 795 (2007).
- ²¹P. Heistracher, C. Abert, F. Bruckner, T. Schrefl, and D. Suess, “Proposal for a micromagnetic standard problem: domain wall pinning at phase boundaries,” *Journal of Magnetism and Magnetic Materials* **548**, 168875 (2022).
- ²²D. Suess, S. Koraltan, F. Slanovc, F. Bruckner, and C. Abert, “Accurate finite-difference micromagnetics of magnets including rky interaction: Analytical solution and comparison to standard micromagnetic codes,” *Physical Review B* **107**, 104424 (2023).
- ²³D. Berkov, K. Ramstöck, and A. Hubert, “Solving micromagnetic problems. towards an optimal numerical method,” *physica status solidi (a)* **137**, 207–225 (1993).
- ²⁴K. Ramstöck, T. Leibl, and A. Hubert, “Optimizing stray field computations in finite-element micromagnetics,” *Journal of magnetism and magnetic materials* **135**, 97–110 (1994).
- ²⁵A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, “Automatic differentiation in pytorch,” (2017).
- ²⁶R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, “Neural ordinary differential equations,” *Advances in neural information processing systems* **31** (2018).
- ²⁷A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, *et al.*, “Pytorch: An imperative style, high-performance deep learning library,” *Advances in neural information processing systems* **32** (2019).
- ²⁸R. Frostig, M. J. Johnson, and C. Leary, “Compiling machine learning programs via high-level tracing,” *Systems for Machine Learning* **4** (2018).
- ²⁹J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, *et al.*, “Jax: Autograd and xla,” *Astrophysics Source Code Library*, ascl-2111 (2021).
- ³⁰A. Meurer, C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, *et al.*, “SymPy: symbolic computing in python,” *PeerJ Computer Science* **3**, e103 (2017).
- ³¹P. Kidger, R. T. Chen, and T. J. Lyons, ““ hey, that’s not an ode”: Faster ode adjoints via seminorms.” in *ICML* (2021) pp. 5443–5452.
- ³²P. Kidger, *On Neural Differential Equations*, Ph.D. thesis, University of Oxford (2021).
- ³³“ μ MAG standard problem #4,” <https://www.ctcms.nist.gov/~rdm/std4/spec4.html> (2024).
- ³⁴“Neuralmag: A machine learning framework for micromagnetic simulations,” <https://gitlab.com/neuralmag/neuralmag> (2024).