

## ARTICLE OPEN



# Machine learning nonequilibrium electron forces for spin dynamics of itinerant magnets

Puhan Zhang<sup>1</sup> and Gia-Wei Chern<sup>1</sup>✉

We present a generalized potential theory for conservative as well as nonconservative forces for the Landau-Lifshitz magnetization dynamics. Importantly, this formulation makes possible an elegant generalization of the Behler-Parrinello machine learning (ML) approach, which is a cornerstone of ML-based quantum molecular dynamics methods, to the modeling of force fields in adiabatic spin dynamics of out-of-equilibrium itinerant magnetic systems. We demonstrate our approach by developing a deep-learning neural network that successfully learns the electron-mediated exchange fields in a driven s-d model computed from the nonequilibrium Green's function method. We show that dynamical simulations with forces predicted from the neural network accurately reproduce the voltage-driven domain-wall propagation. Our work also lays the foundation for ML modeling of spin transfer torques and opens a avenue for ML-based multi-scale modeling of nonequilibrium dynamical phenomena in itinerant magnets and spintronics.

*npj Computational Materials* (2023)9:32; <https://doi.org/10.1038/s41524-023-00990-0>

## INTRODUCTION

In the past decade, machine learning (ML) techniques have greatly impacted many areas of industry and scientific research. The introduction of ML methods to the physical sciences has produced many fruitful results as well as opened several promising directions<sup>1–10</sup>. In particular, the utilization of ML models as universal approximations for high-dimensional functions has significantly improved the efficiency of complex numerical simulations<sup>11–22</sup>. Perhaps the most prominent and successful application in this direction is the ML prediction of energy and forces in quantum molecular dynamics (QMD) simulations<sup>23–36</sup>. Contrary to classical MD methods that are based on empirical force fields, the atomic forces in QMD are computed by integrating out electrons on-the-fly as the atomic trajectories are generated<sup>37</sup>. Various many-body methods, notably the density functional theory, have been used for the force calculation of QMD. However, the fact that most of these electronic structure methods are computationally very expensive significantly restricts the accessible scales of atomic simulations. The ML model offers a promising solution to this computational difficulty by accurately emulating the time-consuming many-body calculations, thus offering the possibility of large-scale QMD simulations with the desired quantum accuracy.

The central idea behind the remarkable scalability of ML force-field models is the principle of locality, or the nearsightedness, of electronic matters<sup>38,39</sup>, which, in the context of QMD simulations, assumes that the force acting on a given atom only depends on its immediate surroundings. A practical implementation of the ML model based on this principle was demonstrated in the pioneer work of Behler and Parrinello<sup>23</sup>, and Bartók et al.<sup>24</sup>. In this approach, the total energy of the system is partitioned as  $E = \sum_i \epsilon_i$ , where  $\epsilon_i$  is called the atomic energy and only depends on the local environment of the  $i$ -th atom<sup>23,24</sup>. The atomic forces are then obtained from derivatives of the predicted energy:  $\mathbf{F}_i = -\partial E / \partial \mathbf{R}_i$ , where  $\mathbf{R}_i$  is the atomic position vector. Crucially, the complicated dependence of atomic energy  $\epsilon_i$  on its neighborhood is approximated by the ML model, which is trained on the condition

that both the predicted individual forces  $\mathbf{F}_i$  as well as the total energy  $E$  agree with the quantum calculations.

Also importantly, by focusing on the local energy  $\epsilon_i$ , which, as a scalar, is invariant under symmetry transformations such as rotations, the symmetry properties of the system can be easily incorporated into the ML model in such Behler-Parrinello (BP) type schemes<sup>23,24</sup>. This approach also ensures that the predicted forces are conservative, a property that is important for Born-Oppenheimer molecular dynamics simulations. The BP scheme has been generalized to improve Monte Carlo simulations of lattice models in condensed matter physics<sup>40–43</sup>. Notably, ML force-field models based on the BP scheme have also been developed to enable large-scale Landau-Lifshitz dynamics simulations of quasi-equilibrium correlated electron magnets<sup>44,45</sup>.

The fact that the atomic forces are conservative in the BP-type approach, however, also significantly limits its capability to represent forces due to highly nonequilibrium electrons, such as in systems under external drive. This is because the energy  $E$  is not a well defined concept in such open systems. The resultant nonequilibrium electronic forces often cannot be written as a derivative of an effective potential energy. A case in point is the current-induced force<sup>46–49</sup> in, e.g. the molecular junctions, which has been shown to be nonconservative. Another important example is the spin-transfer torque<sup>50–54</sup> due to polarized electron current that plays a central role in nanomagnetism and spintronics. Consequently, it is unclear how all the well-developed machinery of ML techniques for quasi-equilibrium QMD can be applied to model the dynamics of electronic systems far from equilibrium.

In this paper, we propose a solution to this important problem in the context of quantum Landau-Lifshitz-Gilbert dynamics for itinerant magnets. We first show that general nonconservative forces in the Landau-Lifshitz equation can be expressed in terms of two scalar potentials. This formulation thus allows one to translate the prediction of exchange fields to that of two potential energies. Applying the locality principle, a generalized BP neural network is developed to predict two associated local energies, from which the forces acting on spins can be obtained through automatic differentiation. As discussed above, the scalar outputs also allow

<sup>1</sup>Department of Physics, University of Virginia, Charlottesville, VA 22904, USA. ✉email: gchern@virginia.edu

for an easier incorporation of symmetry into the ML models. Moreover, similar to the original BP-type schemes, our proposed ML approach enjoys the advantage of further physical constraints on the force prediction. As a demonstration, we apply our ML framework to model the exchange fields computed from the nonequilibrium Green's function method on the s-d system, a well-studied model for itinerant magnets. We further show that voltage-driven propagation of magnetic domain-walls can be accurately reproduced based on forces predicted by the trained neural-network model.

## RESULTS

### Generalized potential theory

The dynamics of a magnetic system is described by the Landau-Lifshitz-Gilbert (LLG) equation<sup>55,56</sup>:

$$\frac{d\mathbf{S}_i}{dt} = -\gamma\mathbf{S}_i \times \mathbf{H}_i + \alpha\mathbf{S}_i \times \frac{d\mathbf{S}_i}{dt}, \quad (1)$$

where  $\gamma$  is the gyromagnetic ratio,  $\alpha$  is an effective damping parameter, and  $\mathbf{H}_i$  is a local magnetic field. In analogy with the molecular dynamics, this local electron-mediated exchange field can be viewed as a force acting on spin  $\mathbf{S}_i$ . For a conservative exchange field, this local field is given by  $\mathbf{H}_i = -\partial E/\partial\mathbf{S}_i$ , where  $E = E(\{\mathbf{S}_i\})$  is the energy of the system which is either conserved or, in the presence of dissipation, decreases with time. Explicitly, the energy dissipation rate is  $dE/dt = -\frac{\alpha}{\gamma} \sum_i (d\mathbf{S}_i/dt)^2$ . Consequently, magnetization dynamics in an open system where energies can be pumped into spins from external sources is beyond the LLG equation governed by a conservative force.

As noted above, the nonequilibrium electronic forces are often nonconservative and *cannot* be expressed as derivatives of a single potential energy  $E$ . As a result, the BP method cannot be directly applied to model the nonequilibrium forces. An alternative approach is ML models that directly predict the non-conservative vector force  $\mathbf{H}_i$ <sup>57,58</sup>. However, besides the difficulty of incorporating the spin-rotation symmetry with a vector output, ML force-field model without additional energy constraints is prone to overfitting and hence less accurate. Indeed, in the so-called gradient-domain ML force-field models, additional constraint is introduced to ensure a curl-free conservative force field<sup>57,58</sup> for quasi-equilibrium electron systems.

For nonconservative forces originating from out-of-equilibrium electrons, there is no constraint on the force-field or the total energy. In order to impose similar physical conditions based on the potential theory, here we derive a general expression for the exchange fields acting on spins in terms of multiple scalar potentials. We first note that one of the most crucial features of the LLG dynamics is the preservation of the spin length, i.e.  $|\mathbf{S}_i(t)|$  is a constant. The most general dynamical equation that satisfies this constraint has the form

$$\frac{d\mathbf{S}_i}{dt} = \mathbf{T}_i = -\gamma\mathbf{S}_i \times \mathbf{V}(\mathbf{S}_i), \quad (2)$$

where  $\mathbf{T}_i$  is the torque and  $\mathbf{V}(\mathbf{S})$  defines a vector field on a unit sphere  $S^2$ . Applying the Helmholtz-Hodge theorem for the case of the  $S^2$  domain<sup>59–61</sup>, the vector field can be decomposed into the radial, gradient, and solenoidal components as:

$$\mathbf{V}(\mathbf{S}) = \mathbf{S}\mathcal{R}(\mathbf{S}) + \nabla_s \mathcal{E}(\mathbf{S}) + \nabla_s \times \mathcal{G}(\mathbf{S}), \quad (3)$$

where  $\mathcal{R}, \mathcal{E}$  and  $\mathcal{G}$  are three scalar functions of the spin  $\mathbf{S} = (S^x, S^y, S^z)$ . The surface gradient operator on a scalar function  $f(\mathbf{S})$  is

$$\nabla_s f = \frac{\partial f}{\partial \mathbf{S}} - \mathbf{S} \left( \mathbf{S} \cdot \frac{\partial f}{\partial \mathbf{S}} \right), \quad (4)$$

while the curl operator on the  $S^2$  sphere is given by

$$\nabla_s \times f = \mathbf{S} \times \frac{\partial f}{\partial \mathbf{S}}. \quad (5)$$

Here  $\frac{\partial f}{\partial \mathbf{S}} = \sum_{a=x,y,z} \frac{\partial f}{\partial S^a}$  is the normal gradient in three dimensions, without the restriction  $|\mathbf{S}| = \text{constant}$ .

Since the radial component, which is parallel to the spin direction, does not contribute to the torque  $\mathbf{T}_i$ , the radial function  $\mathcal{R}$  behaves as a gauge transformation, which has no physical effects on the spin dynamics. This implies that one can define a physical exchange field  $\mathbf{H}$  consisting of only the gradient and solenoidal components in the expansion Eq. (3), i.e.  $\mathbf{H} = \nabla_s \mathcal{E}(\mathbf{S}) + \nabla_s \times \mathcal{G}(\mathbf{S})$ . On the other hand, compared with the surface gradient  $\nabla_s$ , the normal gradient  $\partial/\partial\mathbf{S}$  produces an additional radial component, which can then be gauged away, i.e. the difference between  $\nabla_s \mathcal{E}$  and  $\partial\mathcal{E}/\partial\mathbf{S}$ , according to Eq. (4), is a radial vector field, which again does not contribute to the spin dynamics. Consequently, the most general exchange field in the LLG equation can be expressed in terms of the two scalar fields as

$$\mathbf{H}_i = -\frac{\partial \mathcal{E}}{\partial \mathbf{S}_i} - \mathbf{S}_i \times \frac{\partial \mathcal{G}}{\partial \mathbf{S}_i} = \mathbf{h}_i^{\text{eq}} + \mathbf{h}_i^{\text{neq}}. \quad (6)$$

By analogy with the conservative force, the first term is called the quasi-equilibrium exchange field. The second term which comes from the curl-field is denoted as the nonequilibrium exchange field; see Fig. 1. The generalized LLG equation then reads

$$\frac{d\mathbf{S}_i}{dt} = \gamma\mathbf{S}_i \times \frac{\partial \mathcal{E}}{\partial \mathbf{S}_i} + \gamma\mathbf{S}_i \times \left( \mathbf{S}_i \times \frac{\partial \mathcal{G}}{\partial \mathbf{S}_i} \right) + \alpha\mathbf{S}_i \times \frac{d\mathbf{S}_i}{dt}, \quad (7)$$

The first term describes the conventional precessional dynamics in Eq. (1) with the scalar potential  $\mathcal{E}$  now playing the role of an effective conservative potential. Importantly, while the third Gilbert term accounts for universal dissipation of the energy  $\mathcal{E}$ , the second toroidal term can represent dynamical processes of both energy loss and gain. For example, by setting the potential  $\mathcal{G} = -\lambda\mathcal{E}$ , where  $\lambda$  is a positive parameter, the second term corresponds to a dissipation term introduced in LL's original work<sup>55</sup>. On the other hand, the nonequilibrium Slonczewski-Berger spin-torque<sup>50,51</sup> can also be expressed by the second term in Eq. (7) by identifying the vector  $\mathbf{m}_i = -\partial\mathcal{G}/\partial\mathbf{S}_i$  as the magnetization of the fixed layer in a magnetic tunnel junction.

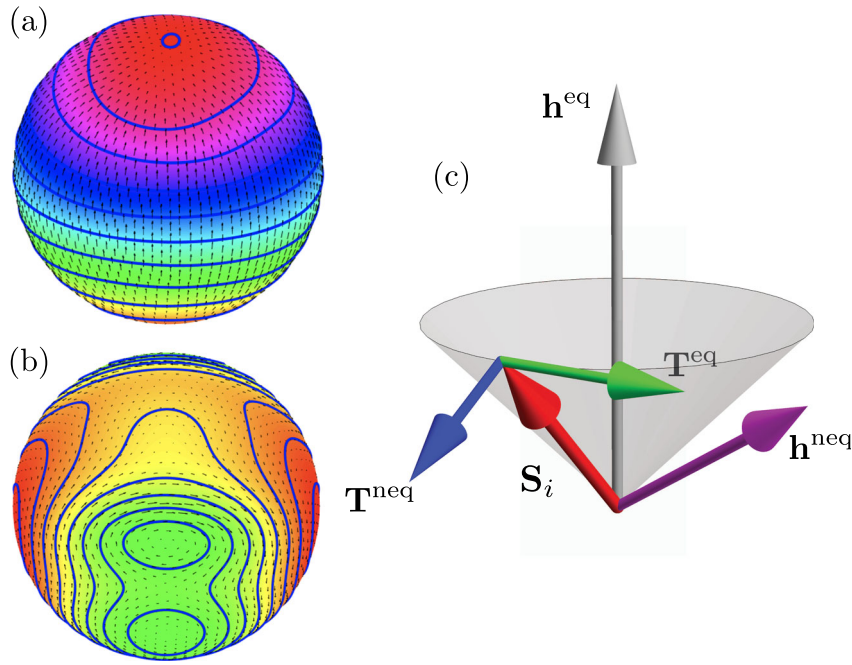
The fact that the generalized potential theory allows for dissipative mechanisms through the  $\mathcal{G}$  term suggests a potential alternative formulation of thermostat. However, further investigation is required in order to consistently include the stochastic thermal fields in this formulation. On the other hand, by focusing the generalized potentials  $\mathcal{E}$  and  $\mathcal{G}$  on the modeling of electron-mediated exchange fields, a stochastic thermal field can be straightforwardly incorporated into the formulation based on conventional Gilbert damping with a consistent Langevin-type thermostat<sup>62,63</sup>. Details of the stochastic LLG equation are discussed in the Methods section.

### Machine-learning exchange-field model for LL dynamics

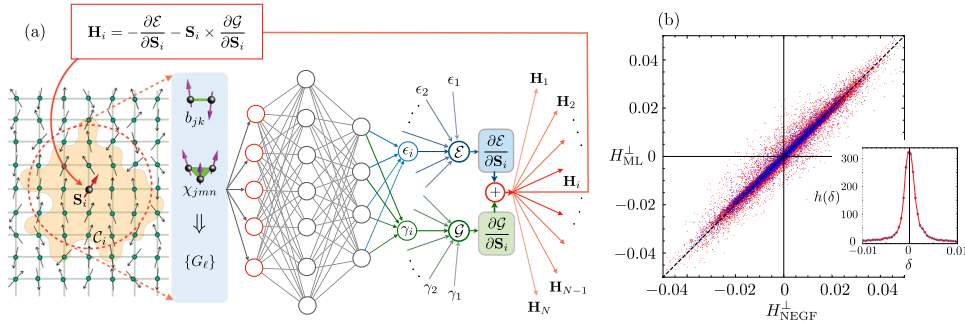
By expressing the general exchange fields in terms of the scalar potentials  $\mathcal{E}$  and  $\mathcal{G}$ , which correspond to the quasi-equilibrium and nonequilibrium components, respectively, one can now generalize the BP-type NN scheme for the forces arising from out-of-equilibrium electrons. To this end, we first partition the two potential energies into local contributions, namely  $\mathcal{E} = \sum_i \epsilon_i$  and  $\mathcal{G} = \sum_i \gamma_i$ . Based on the principle of locality<sup>38,39</sup>, these two local energies  $\epsilon_i$  and  $\gamma_i$  are assumed to depend only on the local magnetic environment  $C_i$  through two universal functions, i.e.  $\epsilon_i = \epsilon(C_i)$  and  $\gamma_i = \chi(C_i)$  for a given electronic model. The overall dependence of the two potential energies on the spin configuration  $\{\mathbf{S}_j\}$  of the system can be expressed as

$$\mathcal{E}(\{\mathbf{S}_j\}) = \sum_i \epsilon(C_i), \quad \mathcal{G}(\{\mathbf{S}_j\}) = \sum_i \chi(C_i). \quad (8)$$

In practice, the magnetic environment  $C_i$  can be defined as the spin configuration within some cutoff radius  $R_c$  from the  $i$ -th spin, i.e.  $C_i = \{\mathbf{S}_j | |\mathbf{r}_j - \mathbf{r}_i| \leq R_c\}$ . As discussed above, the complex



**Fig. 1 The Helmholtz-Hodge of vector fields on a sphere.** A tangential vector field on a sphere can be decomposed into (a) the curl-free component  $\nabla_s \mathcal{E}$  and (b) the divergence-free component  $\nabla_s \times \mathcal{G}$ . (c) shows the gradient field  $\mathbf{h}_i^{\text{eq}} = -\partial \mathcal{E} / \partial \mathbf{S}_i$ , which can be viewed as a quasi-equilibrium exchange field, and the curl-field  $\mathbf{h}_i^{\text{neq}} = -\mathbf{S}_i \times \partial \mathcal{G} / \partial \mathbf{S}_i$ , which corresponds to the nonequilibrium force, and their respective torques  $\mathbf{T}_i^{\text{eq}} = \mathbf{h}_i^{\text{eq}} \times \mathbf{S}_i$  and  $\mathbf{T}_i^{\text{neq}} = \mathbf{h}_i^{\text{neq}} \times \mathbf{S}_i$ .



**Fig. 2 A scalable neural-network force-field model for out-of-equilibrium itinerant spin system and benchmark of force prediction.** a Schematic diagram of the neural-network model. A descriptor transforms the neighborhood spin configuration  $C_i$  to effective coordinates  $\{G_e\}$  which are then fed into a NN. The two output nodes of the NN correspond to the local energy  $e_i = \epsilon(C_i)$  and  $\gamma_i = \chi(C_i)$  associated with site- $i$ . The corresponding total potential energies  $\mathcal{E}$  and  $\mathcal{G}$  are obtained from summation of these local energies. Automatic differentiation<sup>65,66</sup> is employed to compute the derivatives  $\partial \mathcal{E} / \partial \mathbf{S}_i$  and  $\partial \mathcal{G} / \partial \mathbf{S}_i$ , from which the local exchange fields  $\mathbf{H}_i$  are obtained according to the generalized force expression Eq. (6). The NN model is trained by datasets obtained from nonequilibrium Green's function (NEGF) calculation for a driven s-d model. Panel b shows the ML predicted forces versus those from the NEGF calculation for the s-d model with exchange coupling  $J = 3.8t_{\text{nn}}$ ; the blue and red data points correspond to the training and validation/test datasets, respectively. The inset shows the normalized distribution of the prediction error of the perpendicular components of the forces from the test dataset.

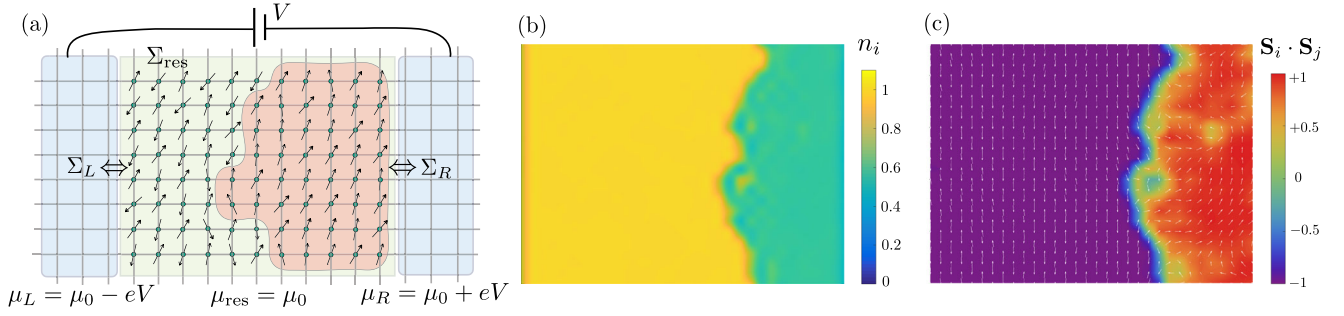
dependences of local energies on the local magnetic environment  $C_i$  are then approximated by a deep-learning NN as shown in Fig. 2.

To ensure that symmetries of the original electron Hamiltonian are preserved in the two energy functions, a magnetic descriptor developed in our previous work<sup>45</sup> is employed to translate the local magnetic environment  $C_i$  into a set of feature variables  $\{G_e\}$  that are invariant under symmetry operations of the system. In particular, for itinerant spin systems such as the well-studied s-d model, the global spin-rotation symmetry needs to be preserved in the ML force-field models. This SO(3) rotation symmetry can be manifestly maintained by using bond variables  $b_{jk}$  and scalar chirality  $\chi_{jkl}$  as building blocks for the construction of the feature variables; they are defined as

$$b_{jk} = \mathbf{S}_j \cdot \mathbf{S}_k, \quad \chi_{jkl} = \mathbf{S}_j \cdot \mathbf{S}_k \times \mathbf{S}_l. \quad (9)$$

Effectively, this means that the two local energies are functions only of these bond/chirality variables in the neighborhood, e.g.  $e_i = \epsilon(b_{jk}, \chi_{jkl})$ , where sites- $j$ ,  $k$ , and  $l$  are within the cutoff radius of the neighborhood.

The ML model also needs to respect the discrete lattice symmetries, such as described by the point group  $D_4$  for the case of square lattice. To obtain the relevant invariant variables, we first note that the collection of bond/chirality variables  $\{b_{jk}, \chi_{jkl}\}$  around the  $i$ -th spin form the basis of a high-dimensional representation of the  $D_4$  group. This reducible representation of the magnetic environment is then decomposed into the fundamental irreducible representations (IR)<sup>64</sup>. The basis of each IR  $f_r^{A_1}, f_r^{A_2}, \dots, f_r^E$ , where  $r$  enumerates the multiplicity in the IR in the decomposition, are proper linear combinations of the bond and scalar chirality variables. Finally, generalized coordinates  $\{G_e\}$  that are



**Fig. 3** Nonequilibrium Green's function (NEGF) calculation for the s-d model driven by an external voltage. **a** Schematic diagram of the driven s-d model in a metal-insulator-metal capacitor structure. The central region is described by the square-lattice s-d Hamiltonian in Eq. (10), while the two electrodes at left and right ends are shown here as modeled by simple square-lattice tight-binding model. An external voltage is applied to the two electrodes, giving rise to a difference in chemical potentials  $\mu_{L/R} = \mu_0 \mp eV/2$ , where  $\mu_0$  is the chemical potential of the background reservoir. Panels **b** and **c** show the on-site electron number  $n_i$  and nearest-neighbor spin-spin correlation  $\mathbf{S}_i \cdot \mathbf{S}_j$ , respectively, of a snapshot during the voltage-induced insulator-to-metal transition simulated by the NEGF-LLG method. The size of the square lattice is  $30 \times 24$ .

invariant under lattice symmetry operations are obtained from the amplitudes and relative phases of these IR basis<sup>41</sup>. More details of the lattice descriptor can be found in the Methods Section.

The resultant feature variables  $\{G_i\}$  are then fed into a fully connected NN, which in turn produces the two local energies  $\epsilon_i$  and  $\gamma_i$  associated with the  $i$ -th spin; see Fig. 2. Applying the NN model to compute all the local energies, the two potential energies  $\mathcal{E}$  and  $\mathcal{G}$  are then obtained through Eq. (8). The local exchange fields  $\mathbf{H}_i$  are computed from the derivatives of the two potentials via Eq. (6), where the two derivatives  $\partial\mathcal{E}/\partial\mathbf{S}_i$  and  $\partial\mathcal{G}/\partial\mathbf{S}_i$  can be efficiently and accurately computed using automatic differentiation techniques<sup>65,66</sup>.

We emphasize that, as in ML-based interatomic potentials for quantum MD simulations, the ML energy model of Eq. (8) essentially provides a classical spin model for an underlying driven electronic systems. However, the energy and force calculations based on the highly nonlinear neural-network model is computationally more demanding compared with classical simulations of short-ranged empirical spin models<sup>67,68</sup>. While computational efficiency of the neural net can be improved with GPU implementations, there are issues of limited memory storage, especially for models with a large cutoff radius. Nonetheless, the ML model is still significantly more efficient than the quantum calculations. Also importantly, the BP-type structure of the presented ML model allows for a linear-scaling implementation of the dynamical simulations.

It is worth noting that magnetic descriptors based on the above bond/chirality variables<sup>45</sup>, strictly speaking, cannot be applied to electron-spin Hamiltonians with magnetic anisotropy, such as spin-orbit coupling. In such systems, the spin-rotation symmetry is coupled to the discrete lattice symmetry. Feature variables that are invariant under the combined symmetry group can still be obtained based on the group-theoretical method described above<sup>69</sup>. However, for most s-d type models where the SU(2) spin-rotation symmetry is only slightly broken, the above descriptor is still a good approximation and a useful starting point for building more general feature variables. We note in passing that different approaches to magnetic descriptors have also been proposed in recent years<sup>70,71</sup>, often in conjunction with MD simulations. While similar bond-variables are also proposed as descriptors<sup>70</sup>, the inclusion of the scalar chirality  $\chi_{ijk}$  in our model plays a crucial role in the stabilization of complex non-coplanar magnetic structures<sup>45</sup>. Finally, off-lattice magnetic descriptors based on bond/chirality variables, which can then be used for combined LLG and MD simulations, are discussed in Ref. <sup>69</sup>.

### Machine-learning for nonequilibrium Green's function method

The above ML framework is general and can be used to represent exchange field in any nonequilibrium electron systems. As a

demonstration of our approach, here we apply it to model the forces computed from the nonequilibrium Green's functions (NEGF) method<sup>72–74</sup> for a driven s-d system<sup>50–52</sup>. The s-d model is widely used in the study of spintronics and spin transfer torques for itinerant magnets. The large  $J$  limit of the s-d model, also known as the double-exchange model, plays an important role in the physics of colossal magnetoresistance observed in several manganites<sup>75</sup>. Here we consider a square-lattice s-d system sandwiched by two electrodes in a capacitor structure shown in Fig. 3. The total Hamiltonian has two parts  $\mathcal{H}_{\text{tot}} = \mathcal{H}_{s-d} + \mathcal{H}_{\text{res}}$ , where the first part is the s-d Hamiltonian,

$$\mathcal{H}_{s-d} = -t_{\text{nn}} \sum_{\langle ij \rangle} (c_{ia}^\dagger c_{ja} + \text{h.c.}) - J \sum_i \mathbf{S}_i \cdot c_{ia}^\dagger \boldsymbol{\sigma}_{\alpha\beta} c_{i\beta}, \quad (10)$$

and  $\mathcal{H}_{\text{res}}$  describes the electrodes and reservoir degrees of freedom, as well as their coupling to the s-d model in the center. The effects of the reservoir fermions can be subsumed into a self-energy  $\boldsymbol{\Sigma}^r(\epsilon)$  in the retarded Green's function:

$$\mathbf{G}^r(\epsilon) = [\epsilon \mathbf{I} - \mathbf{H}_{s-d} - \boldsymbol{\Sigma}^r(\epsilon)]^{-1}, \quad (11)$$

where  $\mathbf{H}_{s-d}$  is matrix representation of the s-d Hamiltonian in the site-spin  $(i, a)$  space; more details can be found in the Method Section. Next, the lesser Green's function  $\mathbf{G}^<$ , which is important for computing physical observables, is obtained using the Keldysh formula for quasi-steady electron states:  $\mathbf{G}^<(\epsilon) = \mathbf{G}^r(\epsilon) \boldsymbol{\Sigma}^<(\epsilon) \mathbf{G}^a(\epsilon)$ , where the lesser self-energy  $\boldsymbol{\Sigma}^<$  is related to the  $\boldsymbol{\Sigma}^r$  through the dissipation-fluctuation theorem. For example, the on-site electron number is given by  $n_i = \sum_a \langle \hat{c}_{ia}^\dagger \hat{c}_{ia} \rangle = \sum_a \int \frac{d\epsilon}{2\pi i} G_{ia,ia}^<(\epsilon)$ . The exchange fields acting on spins in Eq. (1) are obtained using the generalized Hellmann-Feynman theorem, and are explicitly computed from the lesser Green's function<sup>76–79</sup>

$$\mathbf{H}_i = - \left\langle \frac{\partial \hat{\mathcal{H}}_{s-d}}{\partial \mathbf{S}_i} \right\rangle = J \sum_{\alpha\beta} \boldsymbol{\sigma}_{\beta\alpha} \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} G_{ia,i\beta}^<(\epsilon). \quad (12)$$

The above NEGF calculation is combined with the stochastic LLG dynamics to simulate the insulator-to-metal transition (IMT) of the s-d model driven by an external voltage<sup>79</sup>. A small yet finite Langevin-type stochastic field is added to the local exchange field at every site to account for the thermal effects. A second-order algorithm is then used to integrate the LLG equation<sup>80,81</sup>.

In the simulations of the voltage-induced IMT, the system is initially in an insulating antiferromagnetic (AFM) state with an energy gap  $\Delta E_g = 2J$ . An external voltage  $V$  is applied to the two electrodes, which couple to the system at the left and right edges. When the chemical potential of the right electrode is lowered to the eigen-energies of the in-gap edge modes, an instability towards the ferromagnetic (FM) ordering is triggered as electrons

are drained from the edge of the system into the electrode<sup>79</sup>. This instability leads to the nucleation of the FM domains at the edge of the sample. The voltage-driven expansion of the FM domains transforms the system into the low-resistant metallic state. Panels (b) and (c) of Fig. 3 show the on-site electron number  $n_i$  and the nearest-neighbor spin-spin correlation  $b_{\langle ij \rangle} = \mathbf{S}_i \cdot \mathbf{S}_j$ , respectively, of an intermediate state during the IMT. A rather sharp interface separating two domains of distinct electron densities is developed. The insulating AFM region is half-filled with exactly one electron per site, while the nucleated FM domains are characterized by low electron density and tend to be metallic.

The real-space NEGF calculation for a medium size lattice, e.g. less than 1000 spins, is already time-consuming by itself. This is mainly because the calculation of the retarded Green's function  $\mathbf{G}^r(\epsilon)$  requires the inversion of a large matrix that has to be carried out for thousands of different energies  $\epsilon$ ; see Eq. (11). In the NEGF-LLG simulation of driven itinerant magnets, the above NEGF calculation has to be repeated at every time-step of the dynamical simulation. The resultant computational overhead is thus rather substantial. Even with 200 parallel cores, it often takes up to two weeks to perform a complete IMT simulation. As a result, only relatively small scale simulations with less than 1000 spins can be achieved even with highly parallelized programming<sup>79</sup>. As discussed in the Introduction, by accurately emulating the expensive NEGF calculations, the ML approach to nonequilibrium electron forces offers a promising solution to overcome this difficulty of multi-scaling modeling.

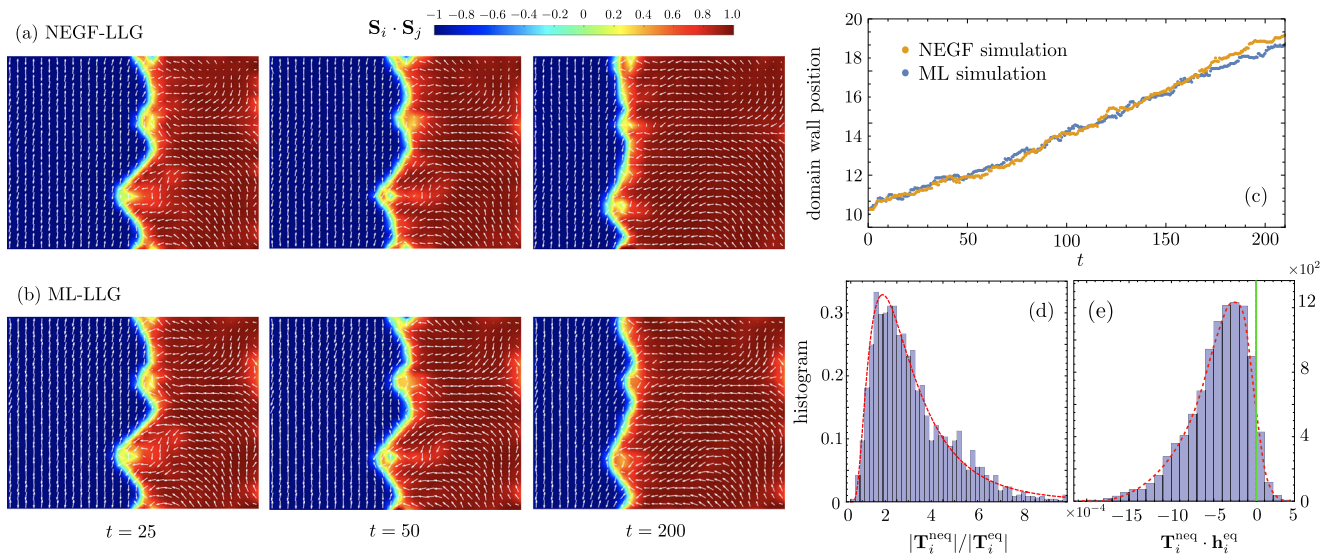
Here we build a six-layer NN to implement the learning model shown in Fig. 2(a). The electronic exchange fields computed from the NEGF method are used to train the NN model based on our generalized force formula Eq. (6). A total of 3200 snapshots, each of which provides roughly 600 force data, are used for the training. Contrary to the standard BP method where both forces and total energy are included in the training of the NN model, the loss function in our case is entirely given by the mean squared error of the forces since the concept of total energy is not well defined for

such open systems. Figure 2(b) shows the componentwise torques  $\mathbf{S}_i \times \mathbf{h}_i$  predicted from our trained NN model versus the exact results. An excellent MSE of  $8.97 \times 10^{-6}$  is obtained from the trained NN model. The normalized distribution of the prediction error obtained from the validation dataset, shown in the inset of Fig. 2(b), is characterized by a rather small standard deviation of  $\sigma = 0.0014$ . More details of the ML training is discussed in the Method Section.

We note that the NN model is trained by dataset from NEGF-LLG simulations with a fixed external voltage  $eV = 3.2$ . As a result, it is designed to specifically learn the out-of-equilibrium electron states with this particular driving voltage, and cannot be used as an effective model for simulations of different  $V$ . Nonetheless, similar to ML force-field models for quantum MD simulations, our trained NN model is scalable, which means it can be used to simulate much larger systems with the same applied voltage. Moreover, the NN model is also transferrable in the sense that it can be used in ML-LLG simulations with different thermal fluctuations or classical magnetic disorder, such as random on-site anisotropy. In the latter case, more diverse and general datasets (with different temperatures or disorder realizations) have to be used for training the NN model. It is also possible to incorporate the driving voltage  $V$  as one of the inputs to the NN, assuming a smooth and continuous  $V$ -dependence of the exchange-fields. We will leave the development of such ML model for future studies.

#### Machine-learning spin dynamics simulations: Quasi-equilibrium vs Nonequilibrium torques

We next incorporate the NN exchange field model into the LLG dynamics for the simulation of the voltage-driven domain-wall propagation in the square-lattice s-d model. As discussed above, the kinetics of the nonequilibrium insulator-to-metal transition is essentially governed by the propagation of the FM-AFM domain walls. We focus our ML model on the force prediction of the interface region where the two distinct magnetic phases coexist. Figure 4(a) and (b) shows the propagation of domain walls obtained from the NEGF-LLG as well as the ML-LLG simulations on



**Fig. 4 Benchmark and analysis of ML-LLG simulations of a driven itinerant magnet.** **a** and **b** Domain wall propagation in a s-d system driven by an external voltage  $eV = 3.2t_{\text{nn}}$ . Comparison between **(a)** NEGF-LLG simulations and **(b)** ML-LLG simulations. The lattice size is  $30 \times 24$ . The color bar indicates the local nearest-neighbor spin correlation  $b_{ij} = \mathbf{S}_i \cdot \mathbf{S}_j$ . Blue (red) area corresponds to AFM (FM) domains. The NEGF-LLG simulation was carried out at a low temperature of  $k_B T = 0.01 t_{\text{nn}}$ , while the ML simulation was performed without Langevin noise. **c** Average position of FM-AFM domain, obtained from NEGF-LLG and ML-LLG simulations, as a function of time during the voltage-driven insulator to metal transition of the s-d model. **d** histogram of the ratio  $|\mathbf{T}_i^{\text{neq}}|/|\mathbf{T}_i^{\text{eq}}|$  predicted by the NN model for the simulation of domain-wall propagation. The dashed lines are guide for eye. Here  $\mathbf{T}_i^{\text{eq}} = \mathbf{S}_i \times \partial \mathcal{E} / \partial \mathbf{S}_i$  is the quasi-equilibrium torque, and  $\mathbf{T}_i^{\text{neq}} = \mathbf{S}_i \times (\mathbf{S}_i \times \partial \mathcal{G} / \partial \mathbf{S}_i)$  is the non-equilibrium torque in the generalized LLG equation. **e** histogram of the scalar product  $\mathbf{T}_i^{\text{neq}} \cdot \mathbf{h}_i^{\text{eq}}$  obtained from the NN model for domain-wall propagation.

a  $30 \times 24$  square lattice. The same initial state with a well-developed FM-AFM domain wall was used for both simulations. In the NEGF-LLG simulations, a small thermal noise is introduced which serves as a small perturbation to the unstable Néel order of the driven system. A Langevin-type thermostat, corresponding to a low temperature of  $k_B T = 0.01 t_{nn}$  is employed in the stochastic LLG dynamics<sup>62,63</sup>. On the other hand, the statistic error associated with the force prediction of the NN model, as shown in Fig. 2(b), can be treated as an effective thermal noise<sup>45</sup>. Indeed, as the prediction error seems to be well approximated by a Gaussian distribution, its effect resembles the addition of normal-distributed random noise added to every site at each time-step in a Langevin thermostat<sup>63</sup>.

The domain-wall positions averaged over the transverse  $y$ -direction, obtained from LLG simulations using NEGF forces and ML-predicted forces, are plotted in Fig. 4(c) as functions of time. The two trajectories agree well with each other with a small discrepancy that can be attributed to the random Langevin noise in the LLG simulation and the force prediction error of the ML model. This overall agreement might indicate that the prediction error happens to mimic the small temperature used in the NEGF-LLG simulations. But more likely, this is an indication that thermal effect of this magnitude is not a dominant factor, but mainly serves as a seed to induce the instability of the Néel state.

A useful by-product of our NN model is the partitioning of the electron-mediated exchange fields into the quasi-equilibrium  $\mathbf{h}^{\text{eq}}$  and nonequilibrium  $\mathbf{h}^{\text{neq}}$  components according to the decomposition in Eq. (6). It is worth noting that such partitioning is often impossible in the microscopic approaches such as the NEGF calculation for the exchange field in Eq. (12). The introduction of these two potentials  $\mathcal{E}$  and  $\mathcal{G}$  is in fact similar in spirit to the partitioning of the total electronic energy into atomic or site energies in the original Behler-Parrinello ML model. These atomic energies also cannot be directly obtained from the DFT calculations. Yet the trained BP-type ML model could predict such local energies associated with individual atoms, thus providing useful information on the energy distribution of the atomic system.

From this decomposition, one can compute the quasi-equilibrium torques  $\mathbf{T}_i^{\text{eq}} = \mathbf{h}_i^{\text{eq}} \times \mathbf{S}_i$ , as well as the nonequilibrium ones  $\mathbf{T}_i^{\text{neq}} = \mathbf{h}_i^{\text{neq}} \times \mathbf{S}_i$ . Figure 4(d) shows the histogram of the ratio  $|\mathbf{T}_i^{\text{neq}}|/|\mathbf{T}_i^{\text{eq}}|$  of these two torque components for spins in the vicinity of the AFM-FM domain walls. As expected, the driving force of the domain-wall propagation is dominated by the nonequilibrium exchange fields. As demonstrated in Fig. 1(c), the quasi-equilibrium torque  $\mathbf{T}_i^{\text{eq}} = \mathbf{S}_i \times \partial \mathcal{E} / \partial \mathbf{S}_i$  is responsible for the precession motion of spins along contours of constant energy  $\mathcal{E}$ . The nonequilibrium torque  $\mathbf{T}_i^{\text{neq}} = \mathbf{S}_i \times (\mathbf{S}_i \times \partial \mathcal{G} / \partial \mathbf{S}_i)$ , on the other hand, often points to a direction opposite to that of the Landau-Lifshitz damping torque  $\mathbf{T}_i^{\text{damping}} = \lambda \mathbf{S}_i \times \mathbf{T}_i^{\text{eq}}$ , where  $\lambda = \gamma \alpha / (1 + \alpha^2)$  is the effective damping coefficient, computed from the (quasi) equilibrium exchange field. This is confirmed by the histogram of the scalar product  $(\mathbf{T}_i^{\text{neq}} \cdot \mathbf{h}_i^{\text{eq}})$  obtained from the NN model for spins in the vicinity of the domain walls, which is shown in Fig. 4(e). The predominantly negative values of this scalar product indicate the nonequilibrium torques are mostly pulling the spins away from the local field direction  $\mathbf{h}_i^{\text{eq}} = -\partial \mathcal{E} / \partial \mathbf{S}_i$ , due to the quasi-equilibrium potential, thus acting in a way similar to the so-called anti-damping torques<sup>52–54</sup>.

## DISCUSSION

The machine-learning force-field models have revolutionized atomistic simulation methods which are crucial to several fields of biological and physical sciences. In particular, taking advantage of the nearsightedness property of electronic matter, the widely-used Behler-Parrinello scheme and other similar approaches allow one to implement transferrable and scalable ML force field models, thus enabling large-scale molecular dynamics simulations with the accuracy of the state-of-the-art quantum calculations. Yet,

despite significant progress in recent years, the majority of research focus on conservative forces due to quasi-equilibrium electrons. This is partly because, by focusing on the prediction of local atomic energies, the BP-type approaches are restricted to forces which can be expressed as derivatives of an effective energy. An important challenge in this field is the generalization of the BP-type schemes to represent non-conservative forces originating from out-of-equilibrium electrons, such as those in a driven system.

Our work marks a crucial step toward ML modeling of nonequilibrium nonconservative force fields of functional electronic materials. Thanks to the special property that the magnitude of magnetization is conserved by the Landau-Lifshitz dynamics, a generalized potential theory is developed for both conservative and non-conservative forces for spin dynamics. More importantly, this formulation allows one to generalize the BP-type schemes to the ML modeling of electronic forces in highly nonequilibrium itinerant magnets. We demonstrate our approach by developing a neural network model that successfully predicts the electronic forces computed from the nonequilibrium Green's function method for a driven s-d model. LLG simulations using the NN-predicted forces also accurately reproduce the voltage-driven domain-wall propagation.

The ML framework developed in this work can also be used to implement accurate and efficient modeling of spin-transfer torques (STT)<sup>50–53</sup>, which plays a central role in the emerging field of spintronics. It is worth noting that most LLG simulations of magnetic systems involving STT are based on empirical formulas<sup>50–53,82–84</sup>, which are similar to the empirical force-field models used in classical MD simulations. While LLG simulations with empirical STT formulas can be achieved on rather large systems, such classical simulations could not describe the subtle interplay between spins and electrons. On the other hand, although STT can be more accurately computed using the NEGF method, its combination with LLG dynamics simulations is computationally very demanding and has so far only been achieved with a hybrid classical-quantum implementation or applied to relatively small systems<sup>76–78,85–89</sup>. We envision ML-based STT models will open an avenue to achieve large-scale dynamical simulations of magnetic textures and spintronic devices with the accuracy of nonequilibrium quantum methods.

While our work provides an elegant implementation of ML force models for general spin dynamics, it remains unclear whether and how similar approaches can be applied to the molecular force fields. The fact that a generalized BP method for spin dynamics is possible is because the exchange field is defined on the two-dimensional surface  $S^2$  of a sphere. This suggests that a similar approach can be applied to the force fields of 2D molecular systems. Indeed, a general 2D force field can be decomposed as  $\mathbf{F}(x, y) = -\nabla_{2D} \phi + \nabla_{2D} \times (A_2 \hat{\mathbf{z}})$ , where  $\nabla_{2D} = (\partial_x, \partial_y)$ , and  $\phi(x, y)$  and  $A_2(x, y)$  are two scalar functions. The ML framework developed here can be straightforwardly adopted to represent nonconservative and nonequilibrium forces for MD simulations of such driven 2D molecular systems. However, this approach cannot be directly applied to 3D systems as the representation of a general force field requires both a scalar and a vector potential:  $\mathbf{F} = -\nabla \phi + \nabla \times \mathbf{A}$ . One possible solution is to employ a NN model with a vector output. However, the preservation of spin rotation symmetry requires more sophisticated descriptors. Further work is required to develop a general ML force field model for QMD simulations of out-of-equilibrium electronic systems.

## METHODS

### Stochastic Landau-Lifshitz-Gilbert dynamics

The NEGF-LLG simulations of the resistance transition are carried out at finite temperatures. To incorporate the stochastic thermal

field into the LLG equation, an additional time-dependent term is added to the local exchange field  $\mathbf{H}_i$  in Eq. (1), giving rise to the following stochastic LLG equation<sup>62,63</sup>

$$\frac{d\mathbf{S}_i}{dt} = -\gamma \mathbf{S}_i \times (\mathbf{H}_i + \boldsymbol{\zeta}_i) + \alpha \mathbf{S}_i \times \frac{d\mathbf{S}_i}{dt}. \quad (13)$$

There are two contributions to the local exchange fields: the deterministic exchange field  $\mathbf{H}_i$  and a random thermal field  $\boldsymbol{\zeta}_i$ . The deterministic term is given by  $\mathbf{H}_i = -\partial E/\partial \mathbf{S}_i$  for conservative field, or Eq. (12) for the non-conservative case of out-of-equilibrium system. For ML-LLG simulations, the deterministic exchange field  $\mathbf{H}_i$  is obtained from the generalized potentials as shown in Eq. (6). The thermal fields at different sites are independent of each other and their time-dependence is modeled by a white noise with zero mean. Specifically they satisfy the following statistic properties

$$\begin{aligned} \langle \zeta_i^m(t) \rangle &= 0, \\ \langle \zeta_i^m(t) \zeta_j^n(t') \rangle &= \delta_{ij} \delta_{mn} \delta(t-t') \frac{\alpha k_B T}{V} \end{aligned} \quad (14)$$

where  $m, n = x, y$ , and  $z$  denote the Cartesian components of the thermal fields. A second-order semi-implicit finite-difference method with special care taken to conserve the spin length is employed to integrate the stochastic LLG equation<sup>80,81</sup>.

### NEGF calculation of the exchange fields

In this section we outline spin dynamics with forces computed from the nonequilibrium Green's function (NEGF) method. We consider a two-dimensional capacitor structure, shown in Fig. 2, described by a total Hamiltonian  $\mathcal{H} = \mathcal{H}_{s-d} + \mathcal{H}_{res}$ , where the two terms correspond to the s-d model in the center and the reservoir including the two electrodes at the two ends of the capacitor structure. The Hamiltonian of the s-d model is described in Eq. (10), and that of the reservoir is given by

$$\mathcal{H}_{res} = \sum_{k,a,i} \epsilon_k d_{i,k,a}^\dagger d_{i,k,a} - \sum_{i,k,a} V_{k,i} \left( d_{i,k,a}^\dagger c_{i,a} + \text{h.c.} \right). \quad (15)$$

Here  $d_{i,\alpha,k}$  represents non-interacting fermions from the bath ( $i$  inside the bulk) or the leads (for  $i$  on the two open boundaries),  $\alpha$  is the spin index, and  $k$  is a continuous quantum number. For example,  $k$  encodes the band-structure of the two leads.

As the s-d Hamiltonian is quadratic in the electron operators, which means there is no direct electron-electron interactions, it can be written as

$$\mathcal{H}_{s-d} = \hat{\mathbf{c}}^\dagger \mathbf{H}_{s-d} \hat{\mathbf{c}}. \quad (16)$$

where  $\hat{\mathbf{c}} = (\hat{c}_{1,\uparrow}, \hat{c}_{1,\downarrow}, \dots, \hat{c}_{N,\uparrow}, \hat{c}_{N,\downarrow})$  is a vector of the electron annihilation operators, and we have introduced a "first-quantized" Hamiltonian  $\mathbf{H}_{s-d}$ , which is a  $2N \times 2N$  matrix in the lattice site-spin space with the following matrix elements:

$$(\mathbf{H}_{s-d})_{ia,j\beta} = t_{ij} \delta_{\alpha\beta} - J_H \delta_{ij} \mathbf{S}_i \cdot \boldsymbol{\sigma}_{\alpha\beta}, \quad (17)$$

To simulate the time-evolution of the s-d model, we first note that the relatively slow dynamics of spins allows us to employ the adiabatic approximation, which is analogous to the Born-Oppenheimer approximation in quantum molecular dynamics. In this approximation, the electrons are assumed to quickly reach a quasi-steady state, which could be in quasi-equilibrium thermodynamically or out of equilibrium as in a driven system, with respect to the instantaneous spin configuration. The semiclassical or adiabatic dynamics of local spins in the s-d model is described by the stochastic LLG equation in Eq. (13). Computationally, the most crucial step is the calculation of the exchange field  $\mathbf{H}_i$ . For a conservative force, e.g. due to electrons in quasi-equilibrium, the exchange field is given by the partial derivative of a potential energy:  $\mathbf{H}_i = -\partial E/\partial \mathbf{S}_i$ , where  $E = \langle \mathcal{H}_{s-d} \rangle = \text{Tr}(\rho_{eq} \mathcal{H}_{s-d})$  is the energy of the quasi-equilibrium electron liquid<sup>90,91</sup>. Often this is obtained using exact diagonalization or more efficient linear-scaling techniques such as the kernel polynomial method.

On the other hand, for an out-of-equilibrium quantum state  $|\Psi\rangle$  such as the one driven by two electrodes in our case, the energy  $E$  of the system is not well defined. However, the exchange field can still be computed using the generalized Hellmann-Feynman theorem<sup>76,77</sup>,

$$\mathbf{H}_i = - \left\langle \Psi \left| \frac{\partial \mathcal{H}_{s-d}}{\partial \mathbf{S}_i} \right| \Psi \right\rangle = J_H \rho_{ia,j\beta}(\{\mathbf{S}_i\}) \boldsymbol{\sigma}_{\beta\alpha}. \quad (18)$$

Here we have introduced the single-particle density matrix  $\rho_{ia,j\beta}(t) = \langle \Psi(t) | c_{j\beta}^\dagger c_{ia} | \Psi(t) \rangle$ . It is worth noting that this electron-induced nonequilibrium exchange field is related to the spin-transfer torques and current-induced phenomena such as tunneling magnetoresistance<sup>50-53</sup>.

The general nonequilibrium density matrix  $\rho_{ia,j\beta}(t)$  can be expressed in terms of the equal-time lesser Green's function  $\rho_{ia,j\beta}(t) = G_{ia,j\beta}^<(t, t)$ , where the general two-time Green's function, defined as  $G_{ia,j\beta}^<(t_1, t_2) = i \langle \Psi(t) | c_{ia}^\dagger(t_1) c_{j\beta}(t_2) | \Psi(t) \rangle$ , is computed using the NEGF method. First, assuming the various reservoir parts are in thermal equilibrium with their respective local chemical potentials, we integrate out these reservoir degrees of freedom and obtain the Fourier-transformed retarded Green's function matrix for the central region.

$$\mathbf{G}^r(\epsilon) = [\epsilon \mathbf{I} - \mathbf{H}_{s-d} - \boldsymbol{\Sigma}^r(\epsilon)]^{-1}, \quad (19)$$

where  $\mathbf{H}_{s-d}$  is the first-quantized Hamiltonian matrix introduced in Eq. (17) and  $\boldsymbol{\Sigma}^r$  is the matrix representation of the dissipation-induced self-energy; its explicit matrix elements are

$$\Sigma_{ia,j\beta}^r(\epsilon) = \delta_{ij} \delta_{\alpha\beta} \sum_k \frac{|V_{i,k}|^2}{\epsilon - \epsilon_k + i0^+}. \quad (20)$$

The resultant level-broadening matrix  $\boldsymbol{\Gamma} = i(\boldsymbol{\Sigma}^r - \boldsymbol{\Sigma}^a)$  is diagonal with  $\Gamma_{ia,ia} = \pi \sum_k |V_{i,k}|^2 \delta(\epsilon - \epsilon_k)$ . For simplicity, we assume flat wide-band spectrum for the reservoirs, which leads to a frequency-independent broadening factor with two different values  $\Gamma_{lead}$  and  $\Gamma_{bath}$ . Next, using the Keldysh formula for quasi-steady state, the lesser Green's function is obtained from the retarded/advanced Green's functions:

$$\mathbf{G}^<(\epsilon) = \mathbf{G}^r(\epsilon) \boldsymbol{\Sigma}^<(\epsilon) \mathbf{G}^a(\epsilon), \quad (21)$$

and the lesser self-energy is related to the  $\boldsymbol{\Sigma}^{r/a}$  through dissipation-fluctuation theorem:

$$\Sigma_{ia,j\beta}^<(\epsilon) = 2i \delta_{ij} \delta_{\alpha\beta} \Gamma_i f_{FD}(\epsilon - \mu_i). \quad (22)$$

Here  $\Gamma_i = \Gamma_{lead}$  or  $\Gamma_{bath}$  depending on whether site- $i$  is at the boundaries or in the bulk, and  $f_{L,R}(\epsilon) = f_{FD}(\epsilon - \mu_{L,R})$  are the Fermi-Dirac distribution functions. The local chemical potential  $\mu_i = \mu_0$  for the bath, and  $\mu_i = \mu_{L/R} = \mu_0 \mp eV/2$  for the two electrodes, where  $V$  is the applied voltage.

Given the retarded Green's function  $\mathbf{G}^r(\epsilon)$  in frequency domain, the density matrix  $\rho_{ia,j\beta}$ , which is the equal-time retarded Green's function, is then given by the integral

$$\rho_{ia,j\beta}(\{\mathbf{S}_i\}) = \int \frac{d\epsilon}{2\pi i} G_{ia,j\beta}^<(\epsilon; \{\mathbf{S}_i\}), \quad (23)$$

for quasi-steady electron state. Here we have explicitly shown the dependence of both the Green's function and the density matrix on the instantaneous spin configuration  $\{\mathbf{S}_i\}$ . The density matrix is used in the computation of the exchange field Eq. (18) acting on spins.

### Group theoretical method for lattice descriptor

The s-d model is characterized by two independent symmetry groups: the global SO(3) rotation symmetry and the point group symmetry of the lattice. Consequently, the feature variables or effective coordinates characterizing the magnetic environment  $C_i$  of the neighborhood need to be invariant under transformations

of both symmetry groups. Here we outline the implementation of such a magnetic descriptor<sup>45</sup>; more details can be found in Supplemental information. As discussed above, instead of directly using the spin vectors  $\mathbf{S}_j$  as input, the spin-rotation symmetry can be preserved by using the scalar variables as building blocks for the magnetic descriptor. Two types of fundamental scalars that can be obtained from vector spins include the inner products, or bond variables,  $b_{jk} = \mathbf{S}_j \cdot \mathbf{S}_k$  of a spin-pair, and the triple-product, also known as the scalar chirality,  $\chi_{jkl} = \mathbf{S}_j \cdot \mathbf{S}_k \times \mathbf{S}_l$  of a spin-triplet.

Next we construct feature variables that are invariant under the discrete point group symmetry, which is  $D_4$  in the case of square lattice. The group-theoretical method provides a rigorous and systematic approach to obtain general invariants of a given symmetry group. The first step is to obtain the basis of irreducible representations (IRs) of the point group. In our case, the symmetry-related bond and scalar chirality variables constructed from the magnetic environment  $C_i$  form a finite-dimensional representation of the point group. They can be decomposed into IRs through proper combinations. For example, consider the four bonds  $b_m \equiv b_{im}$  between the center spin  $\mathbf{S}_i$  and the four nearest neighbors  $\mathbf{S}_m$  with  $m = 1, \dots, 4$ . The 1-dimensional IR  $A_1$  is given by  $f^{A_1} = b_1 + b_2 + b_3 + b_4$ , while the 2-dimensional irreducible IR is  $\mathbf{f} = (b_1 - b_3, b_2 - b_4)$ . More examples are given in the supplemental information. For convenience, we arrange the basis functions of a given IR in the decomposition into a vector  $\mathbf{f}_r^\Gamma = (f_{r,1}^\Gamma, f_{r,2}^\Gamma, \dots, f_{r,D_r}^\Gamma)$  where  $\Gamma$  labels the IR,  $r$  enumerates the multiple occurrences of IR  $\Gamma$  in the decomposition, and  $D_r$  is the dimension of the IR. Given these basis functions, one can immediately obtain a set of invariants called power spectrum  $\{p_r^\Gamma\}$ , which are the amplitudes of each individual IR coefficients, i.e.  $p_r^\Gamma = |\mathbf{f}_r^\Gamma|^2$ . However, feature variables based only on power spectrum are incomplete in the sense that the relative phases between different IRs are ignored. For example, the relative "angle" between two IRs of the same type:  $\cos \theta = (\mathbf{f}_{r_1}^\Gamma \cdot \mathbf{f}_{r_2}^\Gamma) / (|\mathbf{f}_{r_1}^\Gamma| |\mathbf{f}_{r_2}^\Gamma|)$  is also an invariant of the symmetry group. Without such phase information, the NN model might suffer from additional error due to the spurious symmetry, namely two IRs can freely rotate independent of each other.

A more general set of invariants of a symmetry group is called the bispectrum coefficients<sup>92</sup>, which are triple products of the IR coefficients; the difference in the transformation properties of the three IRs is accounted for by the Clebsch-Gordon coefficients of the symmetry group. The power spectrum  $p_r^\Gamma$  is a special subset of the bispectrum coefficients. It is also worth noting that the bispectrum coefficients are complete in the sense that they can be used to faithfully reconstruct the neighborhood configuration up to the symmetry operations. Indeed, it has been demonstrated that atomic descriptors for ML-based molecular dynamics can be obtained by applying the bispectrum method to the three-dimensional rotation group which is an intrinsic symmetry of interatomic interactions<sup>93</sup>.

However, the number of bispectrum coefficients is often too large for practical applications, and some of them are redundant. Here we have implemented a descriptor that is modified from the bispectrum method<sup>45</sup>. We introduce the reference basis functions  $\mathbf{f}_{\text{ref}}^\Gamma$  for each distinct IR of the point group. These reference basis are computed by averaging large blocks of bond and chirality variables, such that they are less sensitive to small changes in the neighborhood spin configurations. We then define the relative "phase" of an IR as the projection of its basis functions onto the reference basis:  $\eta_r^\Gamma \equiv \mathbf{f}_r^\Gamma \cdot \mathbf{f}_{\text{ref}}^\Gamma / |\mathbf{f}_r^\Gamma| |\mathbf{f}_{\text{ref}}^\Gamma|$ . The effective coordinates are then the collection of power spectrum coefficients and the relative phases:  $\{G_\rho\} = \{p_r^\Gamma, \eta_r^\Gamma\}$ . The various steps of the descriptor are summarized in the following

$$C_i \rightarrow \{b_{jk}, \chi_{jmn}\} \rightarrow \{\mathbf{f}_r^\Gamma\} \rightarrow \{p_r^\Gamma, \eta_r^\Gamma\} \quad (24)$$

The generalized coordinates  $\{G_\rho\}$ , or feature variables characterizing the neighborhood spins, are then forwarded to the neural network which produces the local energies at its output node. For the cutoff radius  $R_c = 5a$  used in this work, there is a total of 539 bond/chirality variables in each neighborhood.

### Neural network model and training

A six-layer NN model with four hidden layers composed of  $1024 \times 512 \times 256 \times 128$  neurons is constructed and trained on PyTorch<sup>94</sup>. A schematic diagram of the NN is shown in Fig. 2(a). The size of the input layer size is given by the number of feature variables  $G_\rho$ , which is 539 in this work. The NN performs a series of linear transformations on the input neurons where the ReLU function<sup>95</sup> is used as the activation function between layers. The output layer consists of two neurons whose values correspond to the two local energies  $e_i$  and  $y_i$ . Since only the perpendicular component of the exchange field  $\mathbf{H}_{i,\perp}^+$  enters the torque  $\mathbf{T}_i = \mathbf{S}_i \times \mathbf{H}_i$ , that drives LL dynamics, the loss function is given by

$$L = \sum_{i=1}^N \left| \mathbf{H}_{i,\perp}^{\text{NEGF}} - \mathbf{H}_{i,\perp}^{\text{ML}} \right|^2. \quad (25)$$

The parameter of the NN is optimized by the Adam stochastic optimizer<sup>96</sup> at a learning rate of 0.0001. For the training of the NN, 3200 snapshots from the NEGF-LLG simulations are used as the training dataset. A 5-fold cross-validation and early stopping regularization are performed to prevent overfitting. More details can be found in supplemental information.

### DATA AVAILABILITY

Sample trained models and dataset can be found at [https://github.com/chemgroupUVA/ML\\_non\\_equilibrium\\_de](https://github.com/chemgroupUVA/ML_non_equilibrium_de).

### CODE AVAILABILITY

The source codes used in this work can be downloaded from the GitHub repository [https://github.com/chemgroupUVA/ML\\_non\\_equilibrium\\_de](https://github.com/chemgroupUVA/ML_non_equilibrium_de). These include C codes for spin dynamics and PyTorch codes for the machine learning models and the LLG equation.

Received: 23 July 2022; Accepted: 21 February 2023;

Published online: 03 March 2023

### REFERENCES

- Carleo, G. et al. Machine learning and the physical sciences. *Rev. Mod. Phys.* **91**, 045002 (2019).
- Sarma, S. D., Deng, D.-L. & Duan, L.-M. Machine learning meets quantum physics. *Phys. Today* **72**, 48–54 (2019).
- Bedolla, E., Padierna, L. C. & Castaneda-Priego, R. Machine learning for condensed matter physics. *J. Phys.: Condens. Matter* **33**, 053001 (2021).
- Carrasquilla, J. & Melko, R. G. Machine learning phases of matter. *Nat. Phys.* **13**, 431–434 (2017).
- Van Nieuwenburg, E. P., Liu, Y.-H. & Huber, S. D. Learning phase transitions by confusion. *Nat. Phys.* **13**, 435–439 (2017).
- Zhang, Y. & Kim, E.-A. Quantum Loop Topography for Machine Learning. *Phys. Rev. Lett.* **118**, 216401 (2017).
- Schindler, F., Regnault, N. & Neupert, T. Probing many-body localization with neural networks. *Phys. Rev. B* **95**, 245134 (2017).
- Venderley, J., Khemani, V. & Kim, E.-A. Machine Learning Out-of-Equilibrium Phases of Matter. *Phys. Rev. Lett.* **120**, 257204 (2018).
- Carleo, G. & Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017).
- Nomura, Y., Darmawan, A. S., Yamaji, Y. & Imada, M. Restricted Boltzmann machine learning for solving strongly correlated quantum systems. *Phys. Rev. B* **96**, 205152 (2017).
- Rupp, M., Tkatchenko, A., Müller, K.-R. & von Lilienfeld, O. A. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys. Rev. Lett.* **108**, 058301 (2012).



12. Snyder, J. C., Rupp, M., Hansen, K., Müller, K.-R. & Burke, K. Finding Density Functionals with Machine Learning. *Phys. Rev. Lett.* **108**, 253002 (2012).
13. Brockherde, F. et al. Bypassing the Kohn-Sham equations with machine learning. *Nat. Commun.* **8**, 872 (2017).
14. Schütt, K. T., Gastegger, M., Tkatchenko, A., Müller, K.-R. & Maurer, R. J. Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions. *Nat. Commun.* **10**, 5024 (2019).
15. Wang, S. et al. Massive computational acceleration by using neural networks to emulate mechanism-based biological models. *Nat. Commun.* **10**, 4354 (2019).
16. Tsubaki, M. & Mizoguchi, T. Quantum Deep Field: Data-Driven Wave Function, Electron Density Generation, and Atomization Energy Prediction and Extrapolation with Machine Learning. *Phys. Rev. Lett.* **125**, 206401 (2020).
17. Bürkle, M. et al. Deep-Learning Approach to First-Principles Transport Simulations. *Phys. Rev. Lett.* **126**, 177701 (2021).
18. Huang, L. & Wang, L. Accelerated Monte Carlo simulations with restricted Boltzmann machines. *Phys. Rev. B* **95**, 035105 (2017).
19. Liu, J., Qi, Y., Meng, Z. Y. & Fu, L. Self-learning Monte Carlo method. *Phys. Rev. B* **95**, 041101(R) (2017).
20. Liu, J., Shen, H., Qi, Y., Meng, Z. Y. & Fu, L. Self-learning Monte Carlo method and cumulative update in fermion systems. *Phys. Rev. B* **95**, 241104(R) (2017).
21. Nagai, Y., Shen, H., Qi, Y., Liu, J. & Fu, L. Self-learning Monte Carlo method: Continuous-time algorithm. *Phys. Rev. B* **96**, 161102(R) (2017).
22. Chen, C. et al. Symmetry-enforced self-learning Monte Carlo method applied to the Holstein model. *Phys. Rev. B* **98**, 041102 (2018).
23. Behler, J. & Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
24. Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
25. Li, Z., Kermode, J. R. & De Vita, A. Molecular Dynamics with On-the-Fly Machine Learning of Quantum-Mechanical Forces. *Phys. Rev. Lett.* **114**, 096405 (2015).
26. Botu, V., Batra, R., Chapman, J. & Ramprasad, R. Machine Learning Force Fields: Construction, Validation, and Outlook. *J. Phys. Chem. C* **121**, 511–522 (2017).
27. Li, Y. et al. Machine Learning Force Field Parameters from Ab Initio Data. *J. Chem. Theory Comput.* **13**, 4492–4503 (2017).
28. Smith, J. S., Isayev, O. & Roitberg, A. E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **8**, 3192–3203 (2017).
29. Zhang, L., Han, J., Wang, H., Car, R. & Weinan, E. J. P. R. L. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Phys. Rev. Lett.* **120**, 143001 (2018).
30. Behler, J. Perspective: Machine learning potentials for atomistic simulations. *J. Chem. Phys.* **145**, 170901 (2016).
31. Deringer, V. L., Caro, M. A. & Csányi, G. Machine learning interatomic potentials as emerging tools for materials science. *Adv. Mater.* **31**, 1902765 (2019).
32. McGibbon, R. T. et al. Improving the accuracy of Moller-Plesset perturbation theory with neural networks. *J. Chem. Phys.* **147**, 161725 (2017).
33. Suwa, H. et al. Machine learning for molecular dynamics with strongly correlated electrons. *Phys. Rev. B* **99**, 161107 (2019).
34. Mueller, T., Hernandez, A. & Wang, C. Machine learning for interatomic potential models. *J. Chem. Phys.* **152**, 050902 (2020).
35. Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **99**, 014104 (2019).
36. Thompson, A. P., Swiler, L. P., Trott, C. R., Foiles, S. M. & Tucker, G. J. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *J. Comput. Phys.* **285**, 316–330 (2015).
37. Marx, D. & Hutter, J. *Ab initio molecular dynamics: basic theory and advanced methods* (Cambridge University Press, Cambridge, 2009).
38. Walter, K. Density functional and density matrix method scaling linearly with the number of atoms. *Phys. Rev. Lett.* **76**, 3168 (1996).
39. Prodan, E. & Walter, K. Nearsightedness of electronic matter. *Proc. Natl. Acad. Sci.* **102**, 11635–11638 (2005).
40. Nagai, Y., Okumura, M. & Tanaka, A. Self-learning Monte Carlo method with Behler-Parrinello neural networks. *Phys. Rev. B* **101**, 115111 (2020).
41. Ma, J., Zhang, P., Tan, Y., Ghosh, A. W. & Chern, G.-W. Machine learning electron correlation in a disordered medium. *Phys. Rev. B* **99**, 085118 (2019).
42. Liu, Y.-H., Zhang, S., Zhang, P., Lee, T.-K. & Chern, G.-W. Machine learning predictions for local electronic properties of disordered correlated electron systems. *Phys. Rev. B* **106**, 035131 (2022).
43. Zhang, S., Zhang, P. & Chern, G.-W. Anomalous phase separation in a correlated electron system: Machine-learning enabled large-scale kinetic Monte Carlo simulations. *Proc. Natl. Acad. Sci.* **119**, e2119957119 (2022).
44. Zhang, P., Saha, P., Chern, G.-W. Machine learning dynamics of phase separation in correlated electron magnets. Preprint at <https://doi.org/10.48550/arXiv.2006.04205> (2020).
45. Zhang, P. & Chern, G.-W. Arrested Phase Separation in Double-Exchange Models: Large-Scale Simulation Enabled by Machine Learning. *Phys. Rev. Lett.* **127**, 146401 (2021).
46. Lü, J.-T., Brandbyge, M., Hedegard, P., Todorov, T. N. & Dundas, D. Current-induced atomic dynamics, instabilities, and Raman signals: Quasiclassical Langevin equation approach. *Phys. Rev. B* **85**, 245444 (2012).
47. Todorov, T. N., Dundas, D. & McEniry, E. J. Nonconservative generalized current-induced forces. *Phys. Rev. B* **81**, 075416 (2010).
48. Dundas, D., McEniry, E. J. & Todorov, T. N. Current-driven atomic waterwheels. *Nat. Nanotech.* **4**, 99–102 (2009).
49. Di Ventra, M. & Pantelides, S. T. Hellmann-Feynman theorem and the definition of forces in quantum time-dependent and transport problems. *Phys. Rev. B* **61**, 16207 (2000).
50. Slonczewski, J. C. Current-driven excitation of magnetic multilayers. *J. Magn. Magn. Mater.* **159**, L1–L7 (1996).
51. Berger, L. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B* **54**, 9353–9358 (1996).
52. Brataas, A., Kent, A. D. & Ohno, H. Current-induced torques in magnetic materials. *Nat. Mater.* **11**, 372–381 (2012).
53. Ralph, D. C. & Stiles, M. D. Spin transfer torques. *J. Magn. Magn. Mater.* **320**, 1190–1216 (2008).
54. Salahuddin, S., Datta, D. & Datta, S. Spin Transfer Torque as a Non-Conservative Pseudo-Field. Preprint at <https://doi.org/10.48550/arXiv.0811.3472> (2008).
55. Landau, L. D. & Lifshitz, E. M. Theory of the dispersion of magnetic permeability in ferromagnetic bodies. *Phys. Z. Sowjetunion.* **8**, 153–169 (1935).
56. Gilbert, T. L. A Lagrangian formulation of the gyromagnetic equation of the magnetic field. *Phys. Rev.* **100**, 1243 (1955). A phenomenological theory of damping in ferromagnetic materials. *IEEE Trans. Mag.* **40**, 3443–3449 (2004).
57. Chmiela, S. et al. Machine learning of accurate energy-conserving molecular force fields. *Sci. Adv.* **3**, e1603015 (2017).
58. Chmiela, S., Sauceda, H. E., Müller, K.-R. & Tkatchenko, A. Towards exact molecular dynamics simulations with machine-learned force fields. *Nat. Commun.* **9**, 3887 (2018).
59. Adams, J. F. Vector fields on spheres. *Ann. Math.* **75**, 603–632 (1962).
60. Swartztrauber, P. N. The approximation of vector functions and their derivatives on the sphere. *SIAM J. Numer. Anal.* **18**, 191–210 (1981).
61. Fan, M., Paul, D., Lee, T. C. M. & Matsuo, T. Modeling tangential vector fields on a sphere. *J. Am. Stat. Assoc.* **113**, 1625–1636 (2018).
62. Brown Jr, W. F. Thermal fluctuations of a single-domain particle. *Phys. Rev.* **130**, 1677–1686 (1963).
63. Garcia-Palacios, J. L. & Larazo, F. J. Langevin-dynamics study of the dynamical properties of small magnetic particles. *Phys. Rev. B* **58**, 14937–14958 (1998).
64. Mameresh, M. *Group Theory and Its Application to Physical Problems* (Dover, New York, 1962).
65. Griewank, A. & Walther, A. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation* (SIAM, Philadelphia, 2008).
66. Paszke, A. et al. Automatic differentiation in PyTorch. *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA. (2017).
67. Evans, R. F. L. et al. Atomistic spin model simulations of magnetic nanomaterials. *J. Phys.: Condens. Matter.* **26**, 103202 (2014).
68. Vansteenkiste, A. et al. The design and verification of MuMax3. *AIP Adv.* **4**, 107133 (2014).
69. Zhang, P., Zhang, S. & Chern, G.-W. Descriptors for Machine Learning Model of Generalized Force Field in Condensed Matter Systems. Preprint at <https://doi.org/10.48550/arXiv.2201.00798> (2022).
70. Brännvall, M. A., Gambino, D., Armiento, R. & Alling, B. Machine learning approach for longitudinal spin fluctuation effects in bcc Fe at  $T_c$  and under Earth-core conditions. *Phys. Rev. B* **105**, 144417 (2022).
71. Novikov, I., Grabowski, B., Körmann, F. & Shapcev, A. Magnetic Moment Tensor Potentials for collinear spin-polarized materials reproduce different magnetic states of bcc Fe. *npj Comput. Mater.* **8**, 13 (2022).
72. Datta, S. *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, 1995).
73. Haug, H. & Jauho, A.-P. *Quantum Kinetics in Transport and Optics of Semiconductors*, Springer Series in Solid-State Sciences **123** (Springer-Verlag, Berlin, 2008).
74. Di Ventra, M. *Electrical Transport in Nanoscale Systems* (Cambridge University Press, Cambridge, 2008).
75. Dagotto, E. *Nanoscale phase separation and colossal magnetoresistance* (Berlin, Springer, 2002).
76. Stamenova, M., Sanvito, S. & Todorov, T. N. Current-driven magnetic rearrangements in spin-polarized point contacts. *Phys. Rev. B* **72**, 134407 (2005).
77. Salahuddin, S. & Datta, S. Self-consistent simulation of quantum transport and magnetization dynamics in spin-torque based devices. *Appl. Phys. Lett.* **89**, 153504 (2006).

78. Xie, Y., Ma, J., Ganguly, S. & Ghosh, A. W. From materials to systems: a multiscale analysis of nanomagnetic switching. *J. Comput. Electron.* **16**, 1201–1226 (2017).
79. Chern, G.-W. Spatio-temporal dynamics of voltage-induced resistance transition in the double-exchange model. *Phys. Rev. B* **106**, 245146 (2022).
80. Serpico, C., Mayergoyz, I. D. & Bertotti, G. Numerical technique for integration of the Landau-Lifshitz equation. *J. Appl. Phys.* **89**, 6991–6993 (2001).
81. Mentink, J. H., Tretyakov, M. V., Fasolino, A., Katsnelson, M. I. & Rasing, T. Stable and fast semi-implicit integration of the stochastic Landau-Lifshitz equation. *J. Phys.: Condens. Matter.* **22**, 176001 (2010).
82. Bazaliy, Ya. B., Jones, B. A. & Zhang, S.-C. Modification of the Landau-Lifshitz equation in the presence of a spin-polarized current in colossal- and giant-magnetoresistive materials. *Phys. Rev. B* **57**, R3213–R3216 (1998).
83. Zhang, S. & Li, Z. Roles of Nonequilibrium Conduction Electrons on the Magnetization Dynamics of Ferromagnets. *Phys. Rev. Lett.* **93**, 127204 (2004).
84. Tataru, G., Kohno, H. & Shibata, J. Microscopic approach to current-driven domain wall dynamics. *Phys. Rep.* **468**, 213–301 (2008).
85. Chen, S.-H., Chang, C.-R., Xiao, J. Q. & Nikolić, B. K. Spin and charge pumping in magnetic tunnel junctions with precessing magnetization: A nonequilibrium Green function approach. *Phys. Rev. B* **79**, 054424 (2009).
86. Ellis, M. O. A., Stamenova, M. & Sanvito, S. Multiscale modeling of current-induced switching in magnetic tunnel junctions using ab initio spin-transfer torques. *Phys. Rev. B* **96**, 224410 (2017).
87. Petrović, M. D., Popescu, B. S., Bajpai, U., Plechac, P. & Nikolić, B. K. Spin and Charge Pumping by a Steady or Pulse-Current-Driven Magnetic Domain Wall: A Self-Consistent Multiscale Time-Dependent Quantum-Classical Hybrid Approach. *Phys. Rev. Appl.* **10**, 054038 (2018).
88. Dolui, K. et al. Proximity Spin-Orbit Torque on a Two-Dimensional Magnet within van der Waals Heterostructure: Current-Driven Antiferromagnet-to-Ferromagnet Reversible Nonequilibrium Phase Transition in B1ayer CrI<sub>3</sub>. *Nano Lett.* **20**, 2288–2295 (2020).
89. Nikolić, B. K. et al. First-Principles Quantum Transport Modeling of Spin-Transfer and Spin-Orbit Torques in Magnetic Multilayers. In: Andreoni, W., Yip, S. (eds) *Handbook of Materials Modeling* (Springer Verlag, 2020).
90. Antropov, V. P., Tretyakov, S. V. & Harmon, B. N. Spin dynamics in magnets: Quantum effects and numerical simulations. *J. Appl. Phys.* **81**, 3961–3965 (1997).
91. Ma, P.-W. & Dudarev, S. L. Langevin spin dynamics. *Phys. Rev. B* **83**, 134418 (2011).
92. Kondor, R. A novel set of rotationally and translationally invariant features for images based on the non-commutative bispectrum. Preprint at <https://doi.org/10.48550/arXiv.cs/0701127> (2007).
93. Bartók, A., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).
94. Paszke, A. et al. PyTorch: An imperative style, high-performance deep learning library. *Adv. Neural Inform. Process. Sys.* **32**, 8024–8035 (2019).
95. Barron, J. Continuously differentiable exponential linear units. Preprint at <https://doi.org/10.48550/arXiv.1704.07483> (2017).
96. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. Preprint at <https://doi.org/10.48550/arXiv.1412.6980> (2014).

## ACKNOWLEDGEMENTS

The authors thank Sheng Zhang and Avik Ghosh for useful discussions. This work was supported by the US Department of Energy Basic Energy Sciences under Award No. DE-SC0020330. The authors also acknowledge the support of Research Computing at the University of Virginia.

## AUTHOR CONTRIBUTIONS

G.W.C designed the research. G.W.C developed and integrated the nonequilibrium Green's function (NEGF) codes into the Landau-Lifshitz-Gilbert (LLG) simulations. P.Z. worked on the training of neural network model, performed the NEGF-LLG simulations using the ML-potentials, and analyzed the results. P.Z. and G.W.C. wrote the paper.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41524-023-00990-0>.

**Correspondence** and requests for materials should be addressed to Gia-Wei Chern.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023