

Descriptors for Machine Learning Model of Generalized Force Field in Condensed Matter Systems

Puhan Zhang, Sheng Zhang, and Gia-Wei Chern

Department of Physics, University of Virginia, Charlottesville, VA 22904, USA

(Dated: January 6, 2022)

We outline the general framework of machine learning (ML) methods for multi-scale dynamical modeling of condensed matter systems, and in particular of strongly correlated electron models. Complex spatial temporal behaviors in these systems often arise from the interplay between quasi-particles and the emergent dynamical classical degrees of freedom, such as local lattice distortions, spins, and order-parameters. Central to the proposed framework is the ML energy model that, by successfully emulating the time-consuming electronic structure calculation, can accurately predict a local energy based on the classical field in the intermediate neighborhood. In order to properly include the symmetry of the electron Hamiltonian, a crucial component of the ML energy model is the descriptor that transforms the neighborhood configuration into invariant feature variables, which are input to the learning model. A general theory of the descriptor for the classical fields is formulated, and two types of models are distinguished depending on the presence or absence of an internal symmetry for the classical field. Several specific approaches to the descriptor of the classical fields are presented. Our focus is on the group-theoretical method that offers a systematic and rigorous approach to compute invariants based on the bispectrum coefficients. We propose an efficient implementation of the bispectrum method based on the concept of reference irreducible representations. Finally, the implementations of the various descriptors are demonstrated on well-known electronic lattice models.

I. INTRODUCTION

Machine learning (ML) is emerging as a new paradigm for scientific research and engineering [1–14]. In particular, ML methods are increasingly employed in recent years to drastically speedup various computational tasks in quantum chemistry and materials science [15–23]. A ML model can be viewed as a complex high-dimensional function with numerous tunable parameters. Highly efficient methods have been developed to optimize these model parameters from large number of training dataset. Among the various ML models, deep neural networks (NN) [24, 25] represent the most powerful and versatile tools, which, in principle, can approximate any continuous function with arbitrary accuracy [26–28]. One of the most remarkable applications along this line is the development of ML models that can emulate the time-consuming first-principles electronic structure calculations based on, e.g. the density functional theory (DFT), thus significantly surpassing the size and time scales accessible to such accurate methods. Notably, the advent of interatomic potentials based on ML models has made it possible to perform large-scale molecular dynamics (MD) simulations with the accuracy of DFT and beyond [29–41].

The success of ML methods in quantum chemistry and quantum MD simulations has motivated similar applications in condensed-matter physics. For example, the utilization of ML model to emulate the complicated many-body calculation could potentially offer the tantalizing potential for accurate multi-scale dynamical modeling of interacting electron systems. A particularly important application is the large-scale simulations of the spatio-temporal dynamics of complex patterns that are

prevalent in correlated electron materials [42–52]. The intriguing nanoscale textures in such systems are believed to arise from the nontrivial interplay between quasi-equilibrium electrons and emergent classical fields or bosonic degrees of freedom whose dynamics is much slower than the relaxation of electrons. An example of such slow dynamical variables is the magnetic moments associated with localized *d* or *f* electrons immersed in the Fermi sea of fast-moving conducting electrons in the *s-d* or double-exchange model [53–57]. Another representative case is the order-parameter fields which couple to equilibrium quasi-particles in a symmetry-breaking phase [58–62]. The well separated time scales in such electron models is similar to the Born-Oppenheimer approximation underlying the *ab-initio* molecular dynamics methods [63, 64]. Instead of the atomic dynamics, the goal then is to model the adiabatic time evolution of slow dynamical variables such as local spins and order-parameter fields under the influence of the fast electron degrees of freedom.

Conventionally, an empirical or effective energy model as a function of the classical fields, such as an effective classical spin Hamiltonian or Ginzburg-Landau energy functional, is employed for large-scale dynamics simulations [65, 66]. Special care is taken to properly incorporate the symmetry of the original quantum Hamiltonian into the effective model. However, while the classical energy models coupled with phenomenological dynamics capture some universal features qualitatively, such empirical approach lacks the predictive power. Moreover, the effective classical energy, which can be viewed as integrating out the electrons beforehand, fails to describe the subtle interplay between the electron and the classical degrees of freedom during the dynamical evolution.

In order to more accurately simulate the dynamics of the classical fields, one needs to integrate out the fast electrons or quasi-particles on the fly. This means that the fermionic Hamiltonian, characterized by the instantaneous classical fields, needs to be solved at every time-step of the dynamical simulations. Compared with the classical energy model discussed above, this *quantum* dynamical approach is similar in spirit to the quantum MD methods in which the atomic forces are obtained by solving, e.g. the self-consistent Kohn-Sham equation at every time-step [63, 64]. Naturally, the huge computational overhead due to the repeated electronic structure calculations significantly limits the system size and simulation time accessible by such quantum approaches. For electron systems with e.g. Hubbard-type interactions, more sophisticated, hence more time-consuming, many-body methods such as the dynamical mean-field theory [67, 68], density-matrix renormalization group [69, 70], or quantum Monte Carlo [71], are required to properly include the strong electron correlation effects.

As mentioned above, the modern ML methods offer a promising solution to this computational difficulty in multi-scale quantum dynamical modeling of classical fields, as demonstrated by the ML-based interatomic potential for quantum MD simulations. Indeed, recent works [72, 73] have demonstrated the use of deep-learning NN models to enable large-scale quantum Landau-Lifshitz-Gilbert dynamics simulation of phase separation phenomena in a correlated electron system known as the double-exchange model [53–56]. Moreover, ML energy model was used to achieve large-scale quantum kinetic Monte Carlo simulations which reveals unexpected phase ordering dynamics in the Falicov-Kimball model [74], another canonical example of correlated electron systems [75, 76]. The classical degrees of freedom in the former case are local magnetic moments, while they are equivalent to a classical lattice gas mode used to describe the heavy f electrons in the latter case. In both applications, the electronic subsystem is described by quadratic fermionic Hamiltonians, which can be exactly diagonalized. These pilot studies, however, demonstrate the plausibility of applying ML methods to lattice models with strong electron-electron interactions, such as the Hubbard-type models. The general framework of the ML energy model for multi-scale dynamical simulations is discussed in Sec. II.

It is worth noting that the ML approach to the multi-scale modeling discussed above is essentially to develop a classical energy model based on the superb approximation power of modern learning models such as the NN. Large-scale dynamical simulations are possible mainly because of the efficiency of computing forces based on the classical effective energy. However, even with the general approximation capability of ML methods, it is not guaranteed that the symmetry of the original electron Hamiltonian can be properly included in the effective ML model. In order to ensure the symmetry properties of energy model, one needs to first construct a proper rep-

resentation of the classical field configuration to be used as input to the learning models. A good representation is *invariant* with respect to transformations of the symmetry group of the classical fields as well as those of the lattice point group. This crucial step of the ML model, namely the construction of the proper representation, is often referred to as feature engineering and the resultant feature variables, also called the generalized coordinates, are termed a descriptor [77–79].

Similar issues have also arisen in the context of ML interatomic potentials for quantum MD simulations. There, a proper descriptor of the atomic configuration should be invariant under rotational and permutational symmetries, while retaining the faithfulness of the Cartesian representation. Over the past decade, a number of descriptors have been proposed together with the learning models based on them [15, 29, 30, 35, 77–87]. The bond-order parameters, originally developed to characterize short-range structural order in liquid and glasses [80], are one example. A relatively simple approach is to use the ordered eigenvalues of the correlation matrix, such as the Coulomb or Edward sum matrices, as the descriptor [15]. Another example, which is physically intuitive, is the atom-centered symmetry functions (ACSFs) built from the two-body (relative distances) and three-body (relative angles) invariants of the atomic configurations. Because of its simplicity and flexibility, the ACSF descriptor is widely used in various learning models [29, 81]. The group-theoretical method, on the other hand, offers a more controlled approach to the construction of atomic representation based on the power-spectrum and bispectrum coefficients [30, 78]. In addition to these relatively well-established methods, other notable descriptors include moment tensor potential [82], atomic cluster expansion [83], and deep-potential representation [35]. It is worth noting that the research of atomic descriptor is an active ongoing field.

In this paper, we develop a general theory of the descriptors for the classical fields in condensed matter systems, with a special focus on the lattice models. Several specific approaches are also presented; some are motivated by and generalized from the atomic descriptors discussed above. Our main focus, however, is on the group-theoretical method, which can in principle provide a faithful representation of the local classical fields. The resultant descriptor in terms of the bispectrum coefficients of the irreducible representations (IRs) of the lattice point group is rigorous, but over-complete and cumbersome to implement. We next discuss the concept of the reference IRs which can significantly simplify the implementation of the bispectrum descriptor. The proposed descriptors are then applied to lattice models with varying symmetries and complexity of the classical fields.

The rest of the paper is organized as follows. In Sec. II, we outline the framework of utilizing ML energy model to achieve multi-scale dynamical modeling of condensed matter systems. We also discuss the similarities and differences between descriptor for ML-based quantum MD

methods and that of classical degrees of freedom in lattice electronic models. Sec. III presents a general formulation of the descriptors for the dynamical classical fields. While the majority of the discussion is on the group-theoretical method for the bispectrum coefficients, we also present other physically intuitive descriptors motivated by the studies of ML-based interatomic potentials for MD simulations. In Sec. IV, we demonstrate the bispectrum descriptor, together with the idea of reference IRs, to lattice models with a simple scalar classical field representing local breathing-type lattice distortions. Sec. V discusses the case of cooperative Jahn-Teller coupling which is an example of doublet classical field that transforms simultaneously with the point group. Sec. VI is devoted to the systems with a classical vector fields such as the local magnetic moments in double-exchange models. We demonstrate how to properly account for the global rotation symmetry on top of the discrete lattice symmetry. We conclude our work in Sec. VII.

II. GENERAL FRAMEWORK

The rich and complex behaviors of several correlated electron systems arise from the emergence of slow classical degrees of freedom which couple to the electron liquid with a relatively short relaxation time. These classical variables could arise from the local lattice distortions or displacements which couples to electrons through deformation potential in, e.g. adiabatic Holstein or Jahn-Teller models. They could also correspond to local magnetic moments associated with localized core electrons in the double-exchange system [53–56]. The classical field can also represent the collective degrees of freedom such as order parameters of symmetry breaking phases [58–62], or amplitudes of slave bosons in Gutzwiller theory of Mott metal-insulator transitions [88–90].

In the following, we denote these emergent classical fields by an array of classical variables associated at every sites of the lattice:

$$\Phi(\mathbf{r}_i) = \Phi_i = (\Phi_{i,1}, \Phi_{i,2}, \dots, \Phi_{i,M}) \quad (1)$$

We then consider the following general fermionic Hamiltonian characterized by the classical fields:

$$\begin{aligned} \hat{\mathcal{H}} = & \sum_{mn} \sum_{\alpha\beta} t_{m\alpha,n\beta}(\{\Phi_i\}) \hat{c}_{m\alpha}^\dagger \hat{c}_{n\beta} \\ & + \sum_{mnkl} \sum_{\alpha\beta\gamma\delta} v_{m\alpha,n\beta,k\gamma,l\delta}(\{\Phi_i\}) \hat{c}_{m\alpha}^\dagger \hat{c}_{n\beta}^\dagger \hat{c}_{l\gamma} \hat{c}_{k\delta}, \end{aligned} \quad (2)$$

where $\hat{c}_{i,\alpha}^\dagger$ is the creation operator of electron with quantum number α at site- i . In the following we use the latin letters i, j, k, \dots to denote the lattice sites, and the Greek letters α, β, \dots for internal degrees of freedom, such as spins and orbitals, of the electrons. t and v are electron hopping and interaction coefficients that depend on Φ_i . Finally, there is in general also a “classical” potential energy $\mathcal{V}(\{\Phi_i\})$ for the classical fields, which is independent of the electrons.

Here we are interested in the adiabatic dynamics of such lattice fermion systems with mixed quantum (electron) and classical degrees of freedom. The adiabatic approximation, which is similar to the Born-Oppenheimer approximation in quantum molecular dynamics (MD) [63, 64], is based on a well separation of time scales for the electrons and the classical variables. It assumes that the relaxation of electrons is much faster than the dynamical evolution of the classical variables. As a result, the time evolution of the classical Φ_i field is determined by the quasi-equilibrium electronic state of the instantaneous Hamiltonian. Specifically, this electronic state is represented by the many-body density matrix

$$\hat{\rho}_e(\{\Phi_i\}) = \exp[-\beta \hat{\mathcal{H}}(\{\Phi_i\})] / \mathcal{Z}, \quad (3)$$

where $\beta = 1/k_B T$ is the inverse temperature, and $\mathcal{Z} = \text{Tr} e^{-\beta \hat{\mathcal{H}}}$ is the instantaneous partition function of the electrons. The evolution of the classical fields is governed by dynamical equations ranging from phenomenological relaxational and Metropolis/Glauber-type dynamics to Newton/Langevin equation of motion or Landau-Lifshitz-Gilbert (LLG) equation. For example, one of the simple pure relaxation equation is the time-dependent Ginzburg-Landau (TDGL) equation [65, 66], also known as the model-A dynamics [91]:

$$\frac{\partial \Phi_i}{\partial t} = -\lambda \frac{\partial E}{\partial \Phi_i}, \quad (4)$$

where λ is a dissipation constant, and the effective energy is obtained from the expectation value of the instantaneous Hamiltonian,

$$E = \langle \hat{\mathcal{H}}(\{\Phi_i\}) \rangle = \text{Tr}(\hat{\rho}_e \hat{\mathcal{H}}). \quad (5)$$

In general, the dynamics of the classical fields is determined by the “forces”, which is the derivative of the effective energy with respect to the classical variables. An example is given by the right-hand side of the TDGL equation (4). Given this electron density matrix, one can then compute the generalized electronic forces acting on the classical variables:

$$\mathcal{F}_i = -\frac{\partial \langle \hat{\mathcal{H}} \rangle}{\partial \Phi_i} = -\frac{\partial \text{Tr}(\hat{\rho}_e \hat{\mathcal{H}})}{\partial \Phi_i}. \quad (6)$$

It is worth noting that these forces have to be computed at every time-step of the dynamical simulations. Finally, the “classical” potential energy $\mathcal{V}(\{\Phi_i\})$ also contributes to the force $-\partial \mathcal{V} / \partial \Phi_i$, which can be easily computed and included in the equation of motion.

Compared with dynamical simulations based on an empirical energy model, the quantum dynamical approach here is to obtain the effective energy as well as forces by integrating out the electrons on the fly. However, the calculation of these effective forces is highly time-consuming, and could be prohibitively expensive for large systems. For example, for lattice models without electron-electron interactions, i.e. $v = 0$, the calculation of the effective forces only requires diagonalizing a

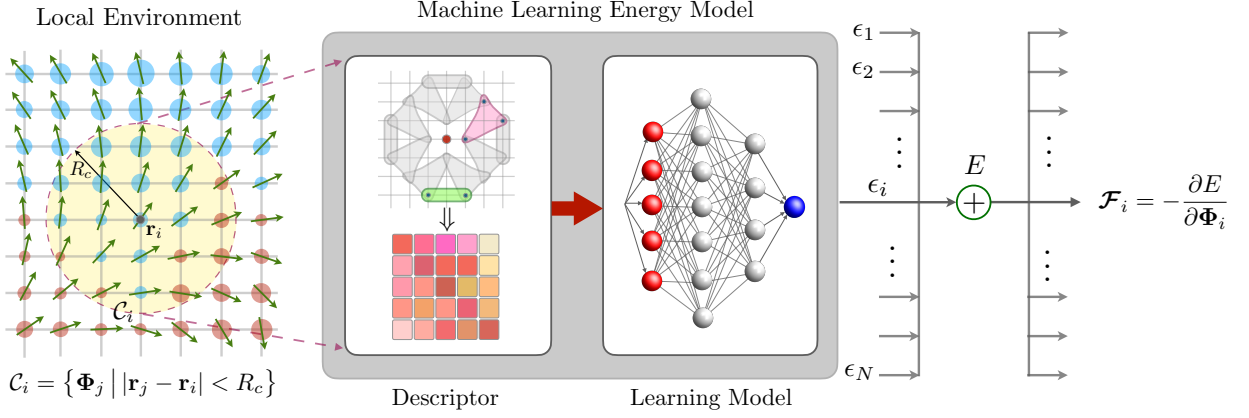


FIG. 1. Numerical framework of the machine-learning force-field model for dynamical simulations of condensed matter systems. At the center of this approach is the ML energy model which takes the local neighborhood \mathcal{C}_i of a given site as the input, and predicts a local energy ϵ_i . The generalized force \mathcal{F}_i is given by the derivative of the total energy $E = \sum_i \epsilon_i$. The ML energy model consists of two major and roughly independent components: the descriptor and the learning model.

quadratic fermionic Hamiltonian. The time complexity of direct diagonalization for a system of N sites scales as $\mathcal{O}(N^3)$. Even though linear-scaling techniques, such as the kernel polynomial method (KPM) [92], have been developed for computing the density matrices of such quadratic Hamiltonians, since the electronic problem has to be solved at every time-step of the dynamical simulation, sophisticated implementations, including for example GPU programming, are often required in order to meet the required efficiency.

For models with electron interactions $v \neq 0$, such as the Hubbard-Kanamori-type interactions, more sophisticated many-body methods are needed to solve the lattice electron model. One popular and widely used approach is the self-consistent methods which include the well-known Hartree-Fock mean-field for symmetry-breaking phases and the Gutzwiller/slave-boson methods for Mott transitions. The central idea of this approach is to reduce the many-body problem into an effective single-particle or quadratic Hamiltonian, which can then be solved by either exact diagonalization or KPM. However, the requirement of self-consistency means that solution of the quadratic Hamiltonian has to be computed multiple times through iteration until a convergence is reached. And this iteration has to be performed again at every time-step, which introduces an extra time complexity even with efficient techniques such as the KPM. The computational overhead is even more demanding for more advanced methods such as the dynamical mean-field theory and quantum Monte Carlo simulations.

As discussed in Section I, ML methods offer a promising solution to this computationally difficult problem. The central idea is the principle of locality, also called the nearsightedness of electronic matter [93, 94], which assumes that local properties such as the on-site forces \mathcal{F}_i only depends on classical fields in the neighborhood of the i -th site. This approach is similar to the Behler-Parrinello (BP) formulation which is fundamental to

the ML potential for *ab-initio* MD simulations [29, 30]. Specifically, the total energy Eq. (5) is first partitioned into local energies associated with individual sites:

$$E = \sum_i \epsilon_i. \quad (7)$$

Importantly, the local site-energy ϵ_i is assumed to depend only on the classical fields in the local environment through a universal function $\epsilon_i = \varepsilon(\mathcal{C}_i)$, where \mathcal{C}_i denotes the local configuration of the classical variables. In practical implementations, this is often defined as the Φ_i variables with a given cutoff radius R_c :

$$\mathcal{C}_i = \{\Phi_j \mid R_{ij} = |\mathbf{r}_j - \mathbf{r}_i| \leq R_c\}. \quad (8)$$

With this partitioning, the calculation of the total electron energy can now be significantly simplified by properly grouping the classical variables Φ_i and substituting into the universal function $\varepsilon(\mathcal{C}_i)$. Crucially, this complex function can now be accurately approximated by ML models, especially the deep-learning NN, thanks to their unprecedented expressive power. Practically, the ML model is derived through a training process based on solutions of the particular many-body method on small systems. Once this universal function is determined, the ML potential thus provides an effective energy model in terms of the classical fields

$$E(\{\Phi_i\}) = \sum_i \varepsilon(\mathcal{C}_i) \quad (9)$$

The effective forces Eq. (6) acting on the classical fields can now be efficiently computed from the derivatives of this classical energy. The general framework of the ML-based force field model for dynamical simulation is summarized in Fig. 1. For ML models based on the neural network, the forces can be readily obtained through the automatic differentiation. Importantly, the ML model offers the efficiency of classical energy model, yet with the accuracy of the many-body techniques employed for generating the training dataset.

III. DESCRIPTOR

The ML energy model in Eq. (9) naturally needs to preserve the symmetry of the original Hamiltonian, which includes both the symmetry of the dynamical variables and that of the underlying lattice. Nonetheless, despite the universal approximation capability of ML models, the symmetry of the original electron Hamiltonian is not automatically captured. Since the training of ML model is essentially an optimization process with randomly chosen datasets, the symmetry of the model can only be statistically approximated even with a large amount of training data. As discussed in Sec. I, a proper representation of the classical fields is required to ensure that the symmetry of the electron Hamiltonian is built into the ML model. A good representation, or descriptor, of the local environment must be invariant with respect to symmetry transformations of the system.

For condensed matter systems defined on a lattice, the ML energy model $\varepsilon(\mathcal{C}_i)$ must be invariant under the discrete transformations of the point group, denoted as G_L , associated with the center site- i . Moreover, for classical fields with a complex structure, one also needs to take into account the symmetry associated with transformations among the multiple components $\Phi_{i,1}, \Phi_{i,2}, \dots$ at the same site. We classify the classical fields into two types depending on whether these two symmetries are entangled to each other or not. Examples of these two types are illustrated in Fig. 2. For models of the first type, the internal symmetry of the classical variables is coupled to the lattice symmetry. Examples of type-I classical variables include on-site displacement vector fields $\mathbf{u}_i = (u_i^x, u_i^y, u_i^z)$ [95–97], where the transformation of the 3 components of the displacement vector is coupled to the discrete rotations of the point group. Another example is the Jahn-Teller doublet $\mathbf{Q}_i = (Q_i^{x^2-y^2}, Q_i^{3z^2-r^2})$ characterizing local structural distortion [98, 99]. The only relevant symmetry group for such type-I models is the on-site point group G_L . Under the symmetry operation $\hat{g} \in G_L$, the rearrangement of the classical fields at different lattice sites coincides with the transformation of the various components. Noting that $\Phi_{j,\alpha} = \Phi_\alpha(\mathbf{r}_j)$ with $\alpha = 1, 2, \dots, M$ being the index of the various components, the transformation of the classical fields is described by

$$\tilde{\Phi}_\alpha(O(\hat{g}) \cdot \mathbf{R}_{ij}) = \mathcal{M}_{\alpha\beta}(\hat{g})\Phi_\beta(\mathbf{R}_{ij}), \quad (10)$$

where $\mathbf{R}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ is the relative position vector of site- j , $\mathcal{M}_{\alpha\beta}(\hat{g})$ is the M -dimensional matrix representation of the symmetry operation \hat{g} , and $O(\hat{g})$ is the 3-dimensional orthogonal matrix transforming site- i to site- k , i.e. $\mathbf{R}_{ik} = O(\hat{g}) \cdot \mathbf{R}_{ij}$.

For type-II models, the classical fields are characterized by an independent internal symmetry group, which will be denoted as G_Φ . The most representative example, perhaps, is the models with local classical spins \mathbf{S}_i as illustrated in Fig. 2(b). For spins with n -component, the symmetry group of the system is a direct product of

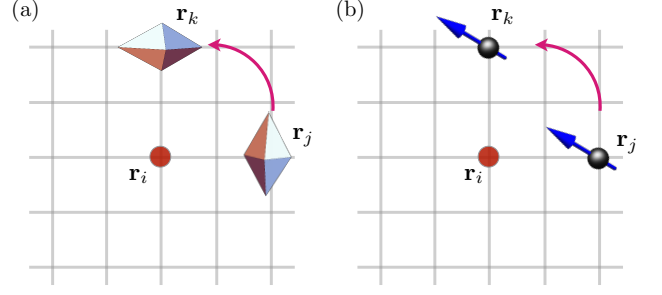


FIG. 2. Classical fields with different symmetry properties. An example of type-I case is the local Jahn-Teller distortion \mathbf{Q}_j as shown in panel (a). The lattice rotation/reflection is accompanied by a simultaneous transformation of the Jahn-Teller phonons. Panel (b) shows the type-II case exemplified by local spins \mathbf{S}_j . The global rotation symmetry of spins is independent of the discrete point-group symmetry of the lattice.

the lattice group G_L and the internal symmetry group $G_\Phi = O(n)$ describing the global rotation symmetry of the spins. The most general symmetry operation consists of the lattice rotation/reflection $\hat{g} \in G_L$, and the transformation $\hat{h} \in G_\Phi$:

$$\tilde{\Phi}_\alpha(O(\hat{g}) \cdot \mathbf{R}_{ij}) = \mathcal{M}_{\alpha\beta}(\hat{h})\Phi_\beta(\mathbf{R}_{ij}), \quad (11)$$

Note that $\mathcal{M}_{\alpha\beta}(\hat{h})$ is the matrix representation of the group G_Φ , which is independent of the lattice point group. It is important to note that, for type-II models, the ML potential energy $\varepsilon(\mathcal{C}_i)$ must be invariant under the general combined symmetry transformation $\hat{h} \otimes \hat{g}$.

It is worth noting that while our focus is on the lattice models which are prevalent in condensed matter physics, most of the analysis presented in this work can be generalized to the off-lattice models or disordered systems if the point group G_L is replaced by the continuous 3-dimensional rotation group $O(3)$, of course, assuming the system possesses such a global rotation symmetry. This means that our analysis can also be applied to electron models defined on an amorphous system or an atomic liquid state. In fact, the latter case can be viewed as a molecular dynamics system with an array of classical variables Φ_i associated with every atom. A particular interesting application would be the Gutzwiller MD method where the classical fields Φ_i corresponds to the slave-boson amplitudes [100].

Having discussed the general symmetry transformations for the two types of classical fields, we next describe a concise vector representation of the neighborhood \mathcal{C}_i with its center at site- i , as defined in Eq. (8). We first consider the type-I models; the case of the type-II model will be discussed in Sec. III C. For convenience, the site-indices of lattice points within \mathcal{C}_i are labeled as j_r , where $r = 1, 2, \dots, L = |\mathcal{C}_i|$. Essentially, the integer r offers an ordered list of lattice sites in the neighborhood. Under the symmetry operation \hat{g} of the point group, the lattice point j_r is mapped to j_s if and only if

$(\mathbf{r}_{j_s} - \mathbf{r}_i) = O(\hat{g}) \cdot (\mathbf{r}_{j_r} - \mathbf{r}_i)$. Consequently, \hat{g} can be represented by a $L \times L$ permutation matrix \mathcal{P} , which means the nonzero matrix elements are $\mathcal{P}_{sr}(\hat{g}) = 1$ if the two sites j_r and j_s are related by \hat{g} . Next we introduce a vector $\vec{\mathcal{U}}$ whose components are given by the classical fields in the neighborhood:

$$\mathcal{U}_{r,\alpha} = \Phi_\alpha(\mathbf{r}_{j_r}), \quad (12)$$

It is easy to see that $\vec{\mathcal{U}}$ offers a vector representation of dimension $L \times M$ for the point group G_L . And the corresponding matrix representation \mathcal{T} of the symmetry operation $\hat{g} \in G_L$ is given by

$$\vec{\mathcal{U}}_{r,\alpha} = \mathcal{T}_{\alpha,\beta}(\hat{g}) \mathcal{U}_{s,\beta} = \mathcal{P}_{rs}(\hat{g}) \mathcal{M}_{\alpha\beta}(\hat{g}) \mathcal{U}_{s,\beta}. \quad (13)$$

Also importantly, the matrix \mathcal{T} provides an orthogonal matrix representation of the point group, i.e. $\mathcal{T}^\dagger \mathcal{T} = \mathcal{T} \mathcal{T}^\dagger = \mathbb{I}$, where \mathbb{I} is the $L \times M$ -dimensional identity matrix. Next we present two descriptors based on this vector representation of the neighborhood.

A. Correlation matrix

The idea of correlation matrix is similar to the so-called Weyl matrix [101] for characterizing the local environment of a single-species molecular system. Specifically, for L atoms within a cutoff radius in the neighborhood of atom- i , the Weyl matrix is defined as $\Sigma_{jk} = (\mathbf{r}_j - \mathbf{r}_i) \cdot (\mathbf{r}_k - \mathbf{r}_i)$. Since the matrix elements are given by scalar products of relative position vectors, the Weyl matrix remains the same under rotation, reflection, and translation operations. However, Σ_{jk} is not a suitable descriptor because permutations of atoms change the order of rows and columns. On the other hand, such permutation operations correspond to a unitary or orthogonal transformation of the Weyl matrices. Consequently the eigenvalues $\{\lambda_m\}$ of the Σ -matrix are invariant under permutation and can be used as a descriptor. A generalization of the Weyl matrix, which can also treat multiple atom-species, is the Coulomb matrix [15]: $M_{ij} = Z_i Z_j / |\mathbf{r}_i - \mathbf{r}_j|$ for $i \neq j$, where Z_i is the nuclear charge of atom- i , and $M_{ii} = \text{const} \times Z_i^{2.4}$. The ‘‘Coulomb’’ interaction form of the off-diagonal matrix elements partly accounts for the Coulomb repulsion between the nuclei, which also highlights the importance of taking into account the atomic pair distances in the descriptor. Other forms of the pair correlation, such as Edward sum or sine-matrices, have also been proposed [79].

Motivated by Weyl and Coulomb matrices, we propose a descriptor given by the ordered eigenvalues of the following correlation matrix of the classical fields

$$C_{\alpha,\beta} = \begin{cases} g(\mathcal{U}_{r,\alpha}) & (r\alpha) = (s\beta) \\ f(\mathbf{r}_{j_r}, \mathbf{r}_{j_s}) \mathcal{U}_{r,\alpha} \mathcal{U}_{s,\beta} & \text{otherwise} \end{cases} \quad (14)$$

where $g(\cdot)$ and $f(\cdot)$ are two functions depending on the model under consideration. For example, one can choose a Coulomb interaction $f(\mathbf{r}_1, \mathbf{r}_2) = 1/|\mathbf{r}_1 - \mathbf{r}_2|$. It is worth

noting that this correlation matrix is *not* invariant under symmetry operations of the point group G_L . Instead, from Eq. (13), the C matrix transforms as

$$\tilde{C} = \mathcal{T}(\hat{g}) C \mathcal{T}^\dagger(\hat{g}). \quad (15)$$

Nonetheless, since \mathcal{T} is an orthogonal matrix, the eigenvalues of the correlation matrix remains invariant with respect to symmetry operations \hat{g} of the point group. As a result, the ordered list of eigenvalues λ_m of the correlation matrix C can be used as a descriptor that preserves the symmetry of the system. In practical implementations, only a finite number of the largest eigenvalues are used as feature variables. Unlike the Coulomb matrix used for fitting the atomization energies of molecular systems, there is no physical basis for the choice of the pair function $f(R)$. An example of descriptor based on correlation matrix is given in Sec. IV below.

B. Bispectrum coefficients

In this Section, we present a more systematic method for constructing a descriptor based on the group-theoretical method. Specifically, the feature variables are given by the so-called bispectrum coefficients computed from the expansion coefficients of irreducible representations of the point group [102–104]. The bispectrum coefficients, which are invariant under the symmetry operations of the point group, are in a sense similar to the scalar triple product of three vectors which is invariant under arbitrary rotations. It is also worth noting that similar group-theoretical methods, with important modifications to simplify the implementation, have been proposed as descriptor for ML interatomic potentials in quantum MD simulations [30, 78].

In order to compute the bispectrum coefficients, the first step is to obtain the irreducible representations of the neighborhood. As discussed above, the vector $\vec{\mathcal{U}}$, defined in Eq. (12) provides a $L \times M$ -dimensional representation of the local environment \mathcal{C}_i . This high-dimensional representation can then be decomposed into irreducible representations (IRs) of the point group G_L following the standard procedures [105, 106]. Specifically, we use Γ to label the different IRs in the decomposition, and denote the corresponding basis vector of IR- Γ as

$$\vec{\Upsilon}^\Gamma = (\vec{\Upsilon}_1^\Gamma, \vec{\Upsilon}_2^\Gamma, \dots, \vec{\Upsilon}_{n_\Gamma}^\Gamma), \quad (16)$$

where n_Γ is the dimension of corresponding IR. Note that each ‘‘component’’ $\vec{\Upsilon}_\mu^\Gamma = \{\Upsilon_{\mu;r,\alpha}^\Gamma\}$ is itself a $(L \times M)$ -dimensional vector. The neighborhood vector is then decomposed as

$$\mathcal{U}_{r,\alpha} = \sum_{\Gamma} \sum_{\mu=1}^{n_\Gamma} f_\mu^\Gamma \Upsilon_{\mu;r,\alpha}^\Gamma. \quad (17)$$

The expansion coefficients f_μ^Γ of the IR, play a role similar to the Fourier coefficients for the translation group.

Using the orthogonality of the basis vectors of different IRs, the expansion coefficients are given by

$$f_\mu^\Gamma = \tilde{\Upsilon}_\mu^{\Gamma\dagger} \cdot \vec{\mathcal{U}} = \sum_{r=1}^L \sum_{\alpha=1}^M \Upsilon_{\mu,r\alpha}^{\Gamma*} \mathcal{U}_{r\alpha}. \quad (18)$$

For convenience, we can group the expansion coefficients of a given IR into a vector:

$$\mathbf{f}^\Gamma = (f_1^\Gamma, f_2^\Gamma, \dots, f_{n_\Gamma}^\Gamma). \quad (19)$$

In terms of the classical fields, see e.g. Eq. (12), the expansion coefficients are

$$f_\mu^\Gamma = \sum_{r=1}^L \sum_{\alpha=1}^M \Upsilon_{\mu,r\alpha}^{\Gamma*} \Phi_\alpha(\mathbf{r}_{j_r}). \quad (20)$$

Under symmetry operations \hat{g} of the point group G_L , different IRs transform independently of each other. Consequently, the transformation of the vector \mathbf{f}^Γ of a given IR is described by an $n_\Gamma \times n_\Gamma$ unitary matrix \mathbf{D}^Γ as

$$\tilde{\mathbf{f}}_\mu^\Gamma = \sum_{\mu'} D_{\mu\mu'}^\Gamma(\hat{g}) f_{\mu'}^\Gamma, \quad (21)$$

or the more concise vector equation: $\tilde{\mathbf{f}}^\Gamma = \mathbf{D}^\Gamma \cdot \mathbf{f}^\Gamma$. From the transformation relation Eq. (13) for the vector $\vec{\mathcal{U}}$, the transformation matrix \mathbf{D}^Γ can be explicitly computed.

$$D_{\mu\mu'}^\Gamma(\hat{g}) = \tilde{\Upsilon}_\mu^{\Gamma\dagger} \cdot \mathcal{T}(\hat{g}) \cdot \tilde{\Upsilon}_{\mu'}^\Gamma. \quad (22)$$

It is worth noting that the transformation matrices of a given IR have been tabulated for most point and double groups. Similar to the ordinary Fourier analysis, we define the power spectrum for a given IR as

$$p^\Gamma \equiv \mathbf{f}^{\Gamma\dagger} \cdot \mathbf{f}^\Gamma = \sum_{\mu=1}^{n_\Gamma} |f_\mu^\Gamma|^2 \quad (23)$$

Since the transformation matrices are unitary $\mathbf{D}^\dagger \mathbf{D} = 1$, it is easy to see that the power spectrum is invariant under symmetry operations:

$$\tilde{p}^\Gamma = \tilde{\mathbf{f}}^{\Gamma\dagger} \cdot \tilde{\mathbf{f}}^\Gamma = \mathbf{f}^{\Gamma\dagger} \mathbf{D}^{\Gamma\dagger} \mathbf{D}^\Gamma \mathbf{f}^\Gamma = \mathbf{f}^{\Gamma\dagger} \mathbf{f}^\Gamma = p^\Gamma, \quad (24)$$

This indicates that the amplitude of each IR can be used as the descriptor for the local environment \mathcal{C}_i . However, the power spectrum p^Γ is not a complete descriptor of the neighborhood function, since it neglects the weight distribution within each IR. Neither does it account for the relative phases between different IRs.

A more complete description, which consists of a larger set of invariants, is given by the bispectrum of the IRs. To this end, we first consider the tensor product of coefficient vectors $\mathbf{f}^{\Gamma_1} \otimes \mathbf{f}^{\Gamma_2}$, which can be viewed as the expansion coefficients of the tensor-product $\vec{\mathcal{U}} \otimes \vec{\mathcal{U}}$ of the vector representation with a tensor-product basis $\tilde{\Upsilon}_\mu^{\Gamma_1} \otimes \tilde{\Upsilon}_\nu^{\Gamma_2}$. Under a symmetry operation, according to Eq. (21), the tensor-product transforms as

$$\begin{aligned} \mathbf{f}^{\Gamma_1} \otimes \mathbf{f}^{\Gamma_2} &\rightarrow (\mathbf{D}^{\Gamma_1} \cdot \mathbf{f}^{\Gamma_1}) \otimes (\mathbf{D}^{\Gamma_2} \cdot \mathbf{f}^{\Gamma_2}) \\ &= (\mathbf{D}^{\Gamma_1} \otimes \mathbf{D}^{\Gamma_2}) \cdot (\mathbf{f}^{\Gamma_1} \otimes \mathbf{f}^{\Gamma_2}). \end{aligned} \quad (25)$$

As is well established in the representation theory of finite groups, the direct product of two IRs can be decomposed into a direct sum of IRs. This indicates the following decomposition of the direct-product matrices:

$$\mathbf{D}^{\Gamma_1} \otimes \mathbf{D}^{\Gamma_2} = (\mathbf{C}^{\Gamma_1, \Gamma_2})^\dagger \left[\bigoplus_\Gamma \mathbf{D}^\Gamma \right] \mathbf{C}^{\Gamma_1, \Gamma_2}, \quad (26)$$

where \oplus means direct sum over the IRs of the direct product. We note that IR of the same dimension and symmetry could appear more than once in the direct sum. The $\mathbf{C}^{\Gamma_1, \Gamma_2}$ is a unitary matrix of dimension $n_{\Gamma_1} \times n_{\Gamma_2}$; its matrix elements are known as the Clebsch-Gordan coefficients of the symmetry group under consideration. Explicitly, we have

$$\begin{aligned} D_{\mu\mu'}^{\Gamma_1}(\hat{g}) D_{\nu\nu'}^{\Gamma_2}(\hat{g}) \\ = \sum_\Gamma \sum_{\kappa, \kappa'} (C_{\kappa, \mu\nu}^{\Gamma; \Gamma_1, \Gamma_2})^* D_{\kappa\kappa'}^\Gamma(\hat{g}) C_{\kappa', \mu'\nu'}^{\Gamma; \Gamma_1, \Gamma_2} \end{aligned} \quad (27)$$

As mentioned above, the sum over Γ could include multiple IRs of the same transformation properties. To construct the bispectrum coefficients, we first consider the following vector:

$$\mathbf{v}^{\Gamma_1, \Gamma_2} = \mathbf{C}^{\Gamma_1, \Gamma_2} \cdot (\mathbf{f}^{\Gamma_1} \otimes \mathbf{f}^{\Gamma_2}) \quad (28)$$

Since the Clebsch-Gordan matrix is essentially a transformation of basis, vector \mathbf{v} is thus the expansion coefficients of the irreducible basis for the tensor-product $\vec{\mathcal{U}} \otimes \vec{\mathcal{U}}$. This can also be seen from the transformation of the \mathbf{v} vector. Substitute Eq. (26) into (25), and multiply the resultant expression by the $\mathbf{C}^{\Gamma_1, \Gamma_2}$ matrix from the left, we see that under symmetry operation \hat{g} , the vector $\mathbf{v}^{\Gamma_1, \Gamma_2}$ transforms according to

$$\tilde{\mathbf{v}}^{\Gamma_1, \Gamma_2} = \left[\bigoplus_\Gamma \mathbf{D}^\Gamma(\hat{g}) \right] \cdot \mathbf{v}^{\Gamma_1, \Gamma_2} \quad (29)$$

This result thus also indicates we can decompose \mathbf{v} into a direct sum of vectors each of which corresponds to an irreducible representation:

$$\mathbf{v}^{\Gamma_1, \Gamma_2} = \bigoplus_\Gamma \mathbf{u}^{\Gamma; \Gamma_1, \Gamma_2} \quad (30)$$

Each vector transforms under symmetry operation as

$$\tilde{\mathbf{u}}^{\Gamma; \Gamma_1, \Gamma_2} = \mathbf{D}^\Gamma(\hat{g}) \cdot \mathbf{u}^{\Gamma; \Gamma_1, \Gamma_2}. \quad (31)$$

From this equation and Eq. (21) for the transformation of the vector \mathbf{f} belong to the same IR- Γ , it is straightforward to see that the following “inner product” is a scalar invariant under any symmetry operation:

$$b^{\Gamma, \Gamma_1, \Gamma_2} = \mathbf{f}^{\Gamma\dagger} \cdot \mathbf{u}^{\Gamma; \Gamma_1, \Gamma_2}, \quad (32)$$

These coefficients are called the bispectrum of the expansion coefficients of the IRs. Using Eq. (28) to express the \mathbf{u} vectors, we obtain the following explicit formula for the bispectrum coefficients

$$b^{\Gamma, \Gamma_1, \Gamma_2} = \sum_{\kappa, \mu, \nu} C_{\kappa, \mu\nu}^{\Gamma; \Gamma_1, \Gamma_2} f_\kappa^{\Gamma*} f_\mu^{\Gamma_1} f_\nu^{\Gamma_2}. \quad (33)$$

The above expression shows the similarity of the b coefficients with the scalar triple of three $O(3)$ vectors. It should also be noted that the power spectrum p^Γ is part of the bispectrum coefficients. In fact, while formally the bispectrum coefficients are built from product of three IR-amplitudes, they can also be used to describe invariants consisting of two IR-coefficients. This corresponds to the case when the decomposition of the direct product representation $\Gamma_1 \otimes \Gamma_2$ includes the trivial one-dimensional representation, denoted as Γ_0 for convenience. By setting the corresponding coefficient to be a constant, e.g. $f_{\Gamma_0} = 1$, we see that the resultant bispectrum coefficient $b^{\Gamma_0, \Gamma_1, \Gamma_2}$ is nonzero only if the two IRs Γ_1 and Γ_2 transform in exactly the same way under symmetry operations, hence have the same dimension. Consequently, we can define the following generalization of power spectrum

$$p^{\Gamma_1, \Gamma_2} = \mathbf{f}^{\Gamma_1 \dagger} \cdot \mathbf{f}^{\Gamma_2} = \sum_{\mu} f_{\mu}^{\Gamma_1 *} f_{\mu}^{\Gamma_2}. \quad (34)$$

The standard power spectrum Eq. (23) of a given IR- Γ corresponds to the case $\Gamma_1 = \Gamma_2 = \Gamma$.

Importantly, since the bispectrum coefficients are invariant under symmetry operations of the point group, they serve as proper descriptor to be combined with the ML models. Moreover, it can be shown that the bispectrum provides a faithful representation of the original configuration in the sense that the vector \mathcal{U} can be rigorously reconstructed from all bispectrum coefficients [102–104]. For practical applications, however, there are a large number of the bispectrum coefficients for most models and point groups. For example, let N be the number of IRs from the decomposition of $\vec{\mathcal{U}}$, which is roughly of the order of $N \sim (L \times M)$, the number of bispectrum is of the order of N^3 , which in general is a rather large number. Moreover, as will be demonstrated in explicit examples in Sec. IV, the bispectrum is an over-complete representation with redundant information. Consequently, further simplification is often required for practical implementations.

As an application of the bispectrum method, here we briefly review its application to represent the atomic environment for ML interatomic potentials. The bispectrum method is often combined with the Gaussian kernel potential learning model and the so-called smooth overlap of atomic positions (SOAP) technique, which approximates atoms in the neighborhood by Gaussian functions of a finite width [30, 78]. For MD simulations, the local atomic configuration is described by the charge density $\rho(\mathbf{r})$ with the origin $\mathbf{r} = 0$ corresponding to the center atom. The symmetry group of three-dimensional free space is $G_L = SO(3)$, and the corresponding irreducible representations are labeled by an integer $\ell = 0, 1, 2, \dots$, which is essentially the angular momentum quantum numbers [107]. Indeed, the basis function Υ_{μ}^{Γ} for the $SO(3)$ group is simply the spherical harmonics $Y_{\ell, m}$. Choosing a proper radial basis $g_n(r)$,

the atomic neighborhood density is expanded as

$$\rho(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{n\ell m} g_n(r) Y_{\ell m}(\theta, \phi), \quad (35)$$

Note that there is an additional integer index n for the expansion coefficients due to the radial dependence. The bispectrum coefficients are then labeled by six integers [78]:

$$b_{n; n_1, n_2}^{\ell; \ell_1, \ell_2} = \sum_{m, m_1, m_2} f_{n\ell m}^* C_{m; m_1, m_2}^{\ell; \ell_1, \ell_2} f_{n_1 \ell_1 m_1} f_{n_2 \ell_2 m_2}, \quad (36)$$

where $C_{m; m_1, m_2}^{\ell; \ell_1, \ell_2}$ are Clebsch-Gordan coefficients of the $SO(3)$ group [107]. For a given set of radial indices (n, n_1, n_2) , the bispectrum coefficients are nonzero only when $\ell = \ell_1 + \ell_2$ due to conservation of angular momentum. However, there are still an infinite number of the b coefficients, and some cutoff ℓ_{\max} has to be introduced for practical implementation. To further simplify the calculation, one can consider only coefficients with $n_1 = n_2 = n$. This, however, implies that rotations of different radial basis are decoupled, thus introducing a spurious symmetry. Nonetheless, some simplifications can be achieved through special designs of the radial basis functions [78].

Instead of dealing with the natural $SO(3)$ group for the three-dimensional space, an alternative approach is to project the atomic environment within a cutoff R_c onto the surface of the four-dimensional sphere S^3 [30, 78]. Specifically, this means that the center-atom is at the north pole, while the cutoff radius, i.e. the 3-sphere specified by $|\mathbf{r}| = R_c$, is mapped to the south pole of the S^3 . Next assuming an approximate $SO(4)$ symmetry for the projected atomic density, one can then use the resultant bispectrum coefficients as the descriptor. As the IR of the $SO(4)$ group is again labeled by an integer j , the bispectrum coefficients are indexed by three integers b^{j, j_1, j_2} . It should be noted that although the projection to S^3 implicitly assumes a spurious $SO(4)$ symmetry, a most crucial advantage of this approach is the absence of the need for radial basis.

C. Internal symmetry

As discussed above, the type-II models are characterized by an internal symmetry group G_{Φ} , independent of the lattice point group, that governs the transformation of the classical fields Φ_i . The feature variables for the ML models need to be invariant with respect to transformations of both symmetry groups. As the multiple components of the local classical vector Φ_i do not transform simultaneously with the lattice symmetry operations, the method described in Sec. III B cannot be directly applied to the type-II models.

One solution is to treat each of the M components of the classical fields $\Phi_i = \{\Phi_{i, \alpha}\}$ ($\alpha = 1, 2, \dots, M$) as

independent. We then view the neighborhood configuration $\mathcal{U}_\alpha = (\mathcal{U}_{1,\alpha}, \mathcal{U}_{2,\alpha}, \dots, \mathcal{U}_{L,\alpha})$ as M independent L -dimensional representations of the neighborhood \mathcal{C}_i . Each component is then decomposed into the IRs of the lattice group (c.f. Eq. (17) for the type-I case)

$$\mathcal{U}_{r,\alpha} = \sum_{\Gamma} \sum_{\mu=1}^{n_{\Gamma}} f_{\mu,\alpha}^{\Gamma} \Upsilon_{\mu;r}^{\Gamma}, \quad (37)$$

Note the basis function Υ of the IR now only depends on the site-index r . The coefficients of the IRs are similarly obtained based on the orthogonality of the basis functions

$$f_{\mu,\alpha}^{\Gamma} = \sum_{r=1}^L \Upsilon_{\mu;r}^{\Gamma*} \mathcal{U}_{r,\alpha} = \sum_{r=1}^L \Upsilon_{\mu;r}^{\Gamma*} \Phi_{\alpha}(\mathbf{r}_{j_r}). \quad (38)$$

For each of the IR Γ in the decomposition (with respect to point group), there are M components indexed by α . As each can be viewed as a M -dimensional representation of the internal symmetry group, it can be decomposed into the IR of G_{Φ} labeled by K :

$$f_{\mu,\alpha}^{\Gamma} = \sum_K \sum_{m=1}^{n_K} F_{\mu,m}^{\Gamma,K} \mathcal{Y}_{m,\alpha}^K. \quad (39)$$

Here \mathcal{Y}_m^K is the basis function of the K -th IR whose dimension is n_K . Using the orthogonality of the basis functions, the expansion coefficients are given by

$$F_{\mu,m}^{\Gamma,K} = \sum_{\alpha=1}^M \mathcal{Y}_{m,\alpha}^{K*} f_{\mu,\alpha}^{\Gamma} = \sum_{\alpha=1}^M \sum_{r=1}^L \mathcal{Y}_{m,\alpha}^{K*} \Upsilon_{\mu;r}^{\Gamma*} \Phi_{\alpha}(\mathbf{r}_{j_r}). \quad (40)$$

Here we have used Eq. (38) in the second equality to express $F_{\mu,m}^{\Gamma,K}$ in terms of the classical fields. It is worth noting that this mixed expansion coefficients, expressed as a special combination of the classical fields, have well defined transformation properties, indicated by the IR indices Γ and K , under both the point group of the site-symmetry and the internal symmetry group. However, since the two set of symmetry transformations are independent of each other, one cannot obtain simultaneous bispectrum coefficients with respect to both symmetry groups. To proceed, we can first “trace out” the point group indices μ by forming the bispectrum coefficients of the point group first

$$B_{K_1,l;K_2,m;K_3,n}^{\Gamma,\Gamma_1,\Gamma_2} = \sum_{\kappa,\mu,\nu} C_{\kappa,\mu\nu}^{\Gamma;\Gamma_1,\Gamma_2} F_{\kappa,l}^{\Gamma,K_1*} F_{\mu,m}^{\Gamma_1,K_2} F_{\nu,n}^{\Gamma_2,K_3}. \quad (41)$$

These coefficients with three indices l, m, n can be viewed as a tensor-product representation $K \otimes K_1 \otimes K_2$ of the internal symmetry group G_{Φ} . Next we decompose this tensor-product representation into a direct sum of IRs of the group G_{Φ} . For convenience of the discussion, we denote the coefficients of the IR in the direct sum as F_q^K . Then invariants with respect to the internal symmetry are given by the bispectrum coefficients from the “triple” product of these F_q^K coefficients. Importantly, these bispectrum coefficients are now invariant with respect to

both the lattice and internal symmetry groups. Since the F_q^K coefficients themselves are already triple product of the field variables, the final invariants in general are composed of 9 classical variables; although some of them can be reduced. Since the number of the coefficients increases even more dramatically with the cutoff radius R_c for type-II models, further approximations are necessary to simplify the implementation of the descriptor.

A second approach, which is physically more intuitive and transparent, is to start from the symmetry of the classical fields and first construct building blocks that are already invariant under the transformations of the internal symmetry group. The group-theoretical method discussed in Sec. III B is then applied to these building blocks for the lattice symmetry. To this end, we again note that the classical fields $\Phi_j = \{\Phi_{j,\alpha}\}$ at every sites in the neighborhood \mathcal{C}_i is obviously an M -dimensional representation of the internal symmetry group, and can be decomposed into IRs of the G_{Φ} group:

$$\Phi_{j,\alpha} = \sum_K \sum_{m=1}^{n_K} f_{j,m}^K \mathcal{Y}_{m,\alpha}^K, \quad (42)$$

It is worth noting that the expansion coefficients $f_{j,m}^K$ acquires a site index j . Again, using the orthogonality of \mathcal{Y} , we have

$$f_{j,m}^K = \sum_{\alpha=1}^M \mathcal{Y}_{m,\alpha}^{K*} \Phi_{j,\alpha}. \quad (43)$$

If the decomposition in Eq. (42) includes the trivial representation K_0 which is by definition a one-dimensional IR, then the coefficients $f_j^{K_0}$ are automatically invariant with respect to the internal symmetry group and are part of the building blocks for the lattice group.

Other invariants of the internal symmetry group are provided by the generalized power spectrum Eq. (34) and the bispectrum coefficients. The crucial difference here is that these invariants are to be built from different sites, thus also serving as many-body correlation functions. First, we consider the generalized power spectrum obtained from a pair of sites (jk)

$$p_{jk}^{K_1,K_2} = \sum_m f_{j,m}^{K_1*} f_{k,m}^{K_2}, \quad (44)$$

Again, the generalized power spectrum coefficient is nonzero only if the two IRs K_1 and K_2 have the same transformation properties. Similarly, one can build invariants of internal symmetry from a triplet (jkl) of lattice sites based on the bispectrum coefficients

$$b_{jkl}^{K,K_1,K_2} = \sum_{l,m,n} c_{l,m,n}^{K;K_1,K_2} f_{j,l}^{K*} f_{k,m}^{K_1} f_{l,n}^{K_2}. \quad (45)$$

where $c_{l,m,n}^{K;K_1,K_2}$ are the Clebsch-Gordan coefficients of the internal symmetry group G_{Φ} . Fig. 3 shows examples of the atomic pairs (jk) and triplets (jkl) related by the lattice rotation and reflection in the neighborhood of the

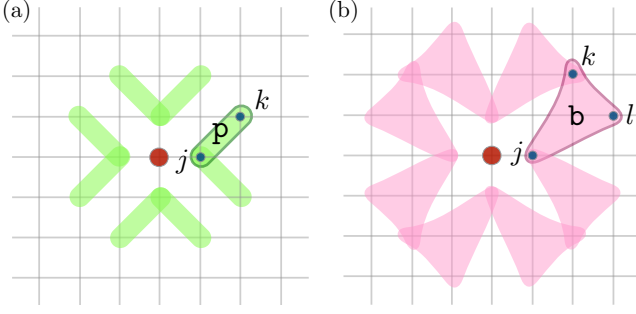


FIG. 3. Examples showing the atomic pair (jk) and triplet (jkl), which are related by the lattice rotation and reflection symmetries, in the neighborhood of the center site on a square lattice.

center site. As mentioned above, these quantities \mathbf{p} and \mathbf{b} also encode the two-body and three-body correlations, respectively, of the neighborhood. Also importantly, they remain unchanged under operations of the internal symmetry group and can be used as building blocks for constructing the invariants of the lattice point group. To this end, we arrange them, including the single-site trivial IR, into a vector of dimension \mathcal{N} :

$$\vec{\mathcal{U}} = (\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_{\mathcal{N}}) = (\mathbf{f}_j^{K_0}, \mathbf{p}_{jk}^{K_1, K_2}, \mathbf{b}_{jkl}^{K, K_1, K_2}). \quad (46)$$

Here we use \mathcal{U}_J to denote the components of this vector, where the index J is used to label either a site j , a pair (jk), or a triplet (jkl). The dimension \mathcal{N} is dominated by the number of atomic pairs and triplets in the neighborhood. For a neighborhood consisting of L sites, these two number scale as L^2 and L^3 , respectively. Moreover, one also needs to take into account the number of different IRs. As the total classical degrees of freedom is $L \times M$, the set of all \mathbf{f} , \mathbf{p} , and \mathbf{b} invariants obviously is an over-complete representation of the neighborhood. Practically, one needs to introduce further constraints in order to reduce this number, for example, by restricting distances between the pairs or triplets to be smaller than another cutoff, or to avoid too many overlaps of the pairs and triples.

Irrespective of the approximations, by keeping all symmetry related pairs and triples, as shown in Fig. 3, in Eq. (46), the vector $\vec{\mathcal{U}}$ forms an \mathcal{N} -dimensional representation of the lattice point group G_L . We next apply the same group-theoretical method discussed in Sec. III B to obtain the bispectrum coefficients of the point group. We again decompose $\vec{\mathcal{U}}$ into the IRs

$$\mathcal{U}_J = \sum_{\Gamma} \sum_{\mu=1}^{n_{\Gamma}} f_{\mu}^{\Gamma} \Upsilon_{\mu;J}^{\Gamma}. \quad (47)$$

where $\Upsilon_{\mu;J}^{\Gamma}$ are the appropriate basis functions. It is worth noting that the IRs of the single sites, pairs, and triplets are decoupled from each other. The expansion

coefficients are then obtained separately as

$$f_{\mu}^{\Gamma} = \begin{cases} \sum_j \Upsilon_{\mu;j}^{\Gamma*} \mathbf{f}_j^{K_0} \\ \sum_{(jk)} \Upsilon_{\mu;jk}^{\Gamma*} \mathbf{p}_{jk}^{K_1, K_2} \\ \sum_{(jkl)} \Upsilon_{\mu;jkl}^{\Gamma*} \mathbf{b}_{jkl}^{K, K_1, K_2} \end{cases} \quad (48)$$

Given these IR coefficients, Eqs. (33) and (34) can then be used to compute the generalized power spectrum and bispectrum coefficients, respectively, which are invariant with respect to both the internal and the lattice symmetry groups of the type-II systems.

D. Atom-centered symmetry functions

The building blocks introduced in Eqs. (44) and (45) above also offer the basis for a descriptor which can be viewed as the generalization of the atom-centered symmetry function (ACSF) originally proposed to describe the atomic configurations [29, 36]. Unlike the group-theoretic methods, the ACSF approach is physically more intuitive and relatively simple to implement. On the other hand, it is more difficult to control the errors due to the *ad hoc* parameterizations of the symmetry functions. Nonetheless, ACSF has been successfully applied to the ML interatomic potential for a wide range of materials. We first briefly review the basic features of ACSF using the example of mono-atomic systems. For a given atomic configuration $\{\mathbf{r}_j\}$ in the vicinity of a center atom- i , the fundamental invariants that are invariant under rotations and reflections of the $O(3)$ group are the distances $R_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$ from the center atom, and the angles $\theta_{ijk} = \arccos[(\mathbf{r}_j - \mathbf{r}_i) \cdot (\mathbf{r}_k - \mathbf{r}_i) / R_{ij} R_{ik}]$. Based on these quantities, two kinds of symmetry functions are introduced. The first type is the two-body (between atoms j and the center atom- i) symmetry function

$$G_2(\{\xi_m\}) = \sum_{j \neq i} F_2(R_{ij}; \{\xi_m\}), \quad (49)$$

where $F_2(R; \xi_m)$ is a user-defined function, parameterized by $\{\xi_m\}$ to extract atomic structures at certain distances from the center atom. One popular choice, proposed in the original work [29], is a Gaussian with a soft cutoff at radius R_c

$$F_2(R; \{\xi_m\}) = e^{-(R - \xi_1)^2 / \xi_2^2} f_c(R). \quad (50)$$

Here $f_c(r) = \frac{1}{2} [\cos(\frac{\pi r}{R_c}) + 1]$ for $R \leq R_c$ and zero otherwise. The two parameters ξ_1 and ξ_2 specify the center and width, respectively, of the Gaussian function. The 3-body symmetry functions are defined as

$$G_3(\{\xi_m\}) = \sum_{j, k \neq i} F_3(R_{ij}, R_{ik}, R_{jk}, \theta_{ijk}; \{\xi_m\}), \quad (51)$$

An example of the three-body envelop function characterized by three parameters is [29, 36]

$$F_3(R_1, R_2, R_3, \theta; \{\xi_m\}) = 2^{1-\xi_1} (1 + \xi_2 \cos \theta)^{\xi_1} \quad (52)$$

$$\times \exp[-(R_1^2 + R_2^2 + R_3^2)/\xi_3^2] f_c(R_1) f_c(R_2) f_c(R_3).$$

We note that generalizations to take into account the different atom species have also been made [79]. Moreover, depending on the problems at hand, it might be more convenient to use different F_2 and F_3 functions, and several variants of these functions have been proposed [79].

Next we present a generalization of the ACSF for condensed-matter systems, where each atom is now associated with a dynamical classical field Φ_j . We emphasize that the formulation presented here can also be used for disordered systems, where the “lattice” point group is replaced by the 3D rotation group $SO(3)$. Moreover, for applications to MD simulation of liquid systems with a dynamical classical fields, the generalized ACSF provides a convenient descriptor for ML energy models for both the atomic dynamics and the classical fields. In order to incorporate the internal symmetry, our approach is to define a set of symmetry functions based on the building blocks in Eq. (46). We start with the two-body symmetry functions that include the coefficients of the trivial IR at every sites:

$$G_{2a}(\{\xi_m\}) = \sum_{j \neq i} \mathbf{f}_j^{K_0} F_2(R_{ij}; \{\xi_m\}), \quad (53)$$

This is the direct generalization of the original two-body symmetry functions that incorporates the on-site classical fields. Another way to build the 2-body symmetry functions is to use the invariants $\mathbf{p}_{ij}^{K_1, K_2}$ between the center site- i and a neighboring site- j :

$$G_{2b}^{K_1, K_2}(\{\xi_m\}) = \sum_{j \neq i} \mathbf{p}_{jk}^{K_1, K_2} F_2'(R_{ij}; \{\xi_m\}), \quad (54)$$

The envelope function $F_2'(R)$ is not necessarily the same as the one for G_{2a} . A three-body symmetry function based on single-site invariants is

$$G_{3a}^{K_1, K_2}(\{\xi_m\}) = \sum_{jk \neq i} \mathbf{f}_j^{K_1} \mathbf{f}_k^{K_2} \quad (55)$$

$$\times F_3(R_{ij}, R_{ik}, R_{jk}, \theta_{ijk}; \{\xi_m\}),$$

The pair-wise invariants can also be combined with the center atom to define a three-body symmetry function:

$$G_{3b}^{K_1, K_2}(\{\xi_m\}) = \sum_{jk \neq i} \mathbf{p}_{jk}^{K_1, K_2} \quad (56)$$

$$\times F_3'(R_{ij}, R_{ik}, R_{jk}, \theta_{ijk}; \{\xi_m\}),$$

A second type of 3-body symmetry functions is obtained from the invariants $\mathbf{b}_{ijk}^{K, K_1, K_2}$ that involves the center atom

$$G_{3c}^{K, K_1, K_2}(\{\xi_m\}) = \sum_{jk \neq i} \mathbf{b}_{ijk}^{K, K_1, K_2} \quad (57)$$

$$\times F_3''(R_{ij}, R_{ik}, R_{jk}, \theta_{ijk}; \{\xi_m\}),$$

Finally, several four-body symmetry functions can be defined based on the fundamental invariants of the internal symmetry group. For example, combining the triplet (jkl) with the center site, we have

$$G_4^{K, K_1, K_2}(\{\xi_m\}) = \sum_{jkl \neq i} \mathbf{b}_{jkl}^{K, K_1, K_2} \quad (58)$$

$$\times F_4(R_{ij}, R_{ik}, R_{il}, \dots; \theta_{ijk}, \theta_{ikl}, \dots).$$

It is worth noting that most of the symmetry functions also depend on the IR indices \mathbf{K} of the internal symmetry group. We also note that since the relative angles θ_{ijk} are pre-defined constants for models on a regular lattice, the dependence of the F functions on these angles is trivial. More importantly, these F functions are used to select the more relevant pairs or triplets to be included in the symmetry functions.

In particular, the symmetry functions can be simplified to a sum over the symmetric-IR for lattice models. Take G_{3b} as an example, we first divide all atomic pairs (jk) in the neighborhood into inequivalent classes such that pairs within the same class are related by the point group symmetry. Moreover, since pairs belong to the same class are related by rotations or reflections that preserve the distance from the center site, they share the same value of the F_3 function; see Fig. 3(a) for an example of the symmetry-related pairs on a square lattice. Using π to denote the inequivalent classes of pairs, we then have

$$G_{3b}^{K_1, K_2}(\{\xi_m\}) = \sum_{\pi} F_3'(\pi; \{\xi_m\}) \sum_{\hat{g}} \mathbf{p}_{\pi(\hat{g})}^{K_1, K_2}. \quad (59)$$

Here $\pi(\hat{g})$ denotes atomic pairs (jk) related to a reference pair in the class π by the symmetry operation \hat{g} . The sum over \hat{g} , which is the symmetric sum of the pair-wise invariants \mathbf{p} , corresponds to the 1D trivial IR of the lattice point group. Consequently, the symmetry function G_{3b} is manifestly an invariant of both the internal and lattice symmetry groups.

To briefly conclude this Section, we have formulated a general theory of descriptors for characterizing dynamical classical fields in condensed matter systems, and presented various different, yet related, approaches for computing the invariant feature variables. By generalizing the concept of the Weyl matrix for atomic environment, we show that the ordered eigenvalues of a correlation matrix can be used to characterize the classical fields in a local neighborhood. The group-theoretical method offers a rigorous and systematic approach to derive a descriptor based on the bispectrum coefficients. Finally, we discuss a descriptor that incorporates the symmetry of the classical fields into the atom-centered symmetry functions. Explicit implementations of these descriptors are demonstrated for well-studied correlated electron systems in the following sections.

IV. EXAMPLE: ADIABATIC DYNAMICS OF CLASSICAL SCALAR FIELD

We first discuss descriptors for the simplest classical field: a dynamical scalar variable $Q_i = Q(\mathbf{r}_i)$ associated with every lattice sites in a square lattice. Physically, such dynamical scalar field can be viewed as describing the local isotropic structure distortion, for example, the breathing mode of the MO_6 octahedron in transition metal oxide. Specific example is given by the Holstein model [109] with spinless electrons:

$$\hat{\mathcal{H}} = -t \sum_{\langle ij \rangle} (\hat{c}_i^\dagger \hat{c}_j + \text{h.c.}) - g \sum_i Q_i \hat{n}_i \quad (60)$$

where $\hat{c}_i/\hat{c}_i^\dagger$ is the annihilation/creation operators of spinless electron at site- i , and $\hat{n}_i = \hat{c}_i^\dagger \hat{c}_i$ is the corresponding number operator. The first-term describes electron hopping between nearest-neighbor sites $\langle ij \rangle$, t is the nearest-neighbor hopping coefficient. The second term denotes phonon-electron interaction with a coupling constant g . The Hamiltonian is supplemented by the classical potential energy that describes the elastic energies of the local structural distortion,

$$\mathcal{V}(\{Q_i\}) = \frac{K_0}{2} \sum_i Q_i^2 + K_1 \sum_{\langle ij \rangle} Q_i Q_j. \quad (61)$$

where K_0 and K_1 are the effective spring constants. Inclusion of electron-electron interaction leads to the Holstein-Hubbard model [110] with spinful electrons

$$\begin{aligned} \hat{\mathcal{H}} = & -t \sum_{\langle ij \rangle} \sum_{\sigma=\uparrow, \downarrow} (\hat{c}_{i,\sigma}^\dagger \hat{c}_{j,\sigma} + \text{h.c.}) \\ & + U \sum_i \hat{n}_{i,\uparrow} \hat{n}_{i,\downarrow} + V \sum_{\langle ij \rangle} \hat{n}_i \hat{n}_j - g \sum_i Q_i \hat{n}_i, \end{aligned} \quad (62)$$

Here $\hat{n}_{i,\sigma} = \hat{c}_{i,\sigma}^\dagger \hat{c}_{i,\sigma}$ is the number operator of electron with spin- σ , and $\hat{n}_i = \hat{n}_{i,\uparrow} + \hat{n}_{i,\downarrow}$, the U term describes the well known on-site Hubbard repulsion, and V represents short-range Coulomb interactions.

The Holstein models in which the lattice degrees of freedom are treated quantum mechanically are used to study phenomena related to electron-phonon coupling, such as polaron physics and superconductivity. On the other hand, Holstein models with classical phonons also serve as simple model systems to investigate the effects of structural distortions on the electronic properties. Indeed, the Jahn-Teller model, which is the multi-orbital generalization of the Holstein model, plays an important role in the physics of colossal magnetoresistance effect. In particular, as discussed in Sec. I, complex inhomogeneous states can arise from the interplay between the fast electron and slow classical lattice dynamics.

Here we are interested in the adiabatic dynamics of these models and treat the lattice distortions as classical

dynamical variables. Their time evolution is then governed by the Langevin equation

$$\mu \frac{d^2 Q_i}{dt^2} + \lambda \frac{d Q_i}{dt} = -\frac{\partial \mathcal{V}}{\partial Q_i} - \frac{\partial \langle \hat{\mathcal{H}} \rangle}{\partial Q_i} + \eta_i(t). \quad (63)$$

where μ is the effective mass and λ is the dissipation constant, and $\eta_i(t)$ represents the stochastic thermal forces. The first term on the right hand side describes the classical elastic restoring force, while the second term is due to the electron-lattice coupling. Explicit calculation gives

$$F_i^{\text{elec}} = -\frac{\partial \langle \hat{\mathcal{H}} \rangle}{\partial Q_i} = g \langle \hat{n}_i \rangle, \quad (64)$$

The electron force is proportional to the on-site electron density. Similar to the Born-Oppenheimer approximation in *ab initio* or quantum MD simulations, the electrons are assumed to quickly reach quasi-equilibrium of the instantaneous Hamiltonian. For a given classical field, the Holstein model describes a quadratic fermionic Hamiltonian, which can be solved by, e.g. exact diagonalization. In the presence of Hubbard interaction U , many-body methods such as the real-space Gutzwiller/slave boson [87], or DMFT are required to solve the electron Hamiltonian and compute the electron force. As discussed in Sec. II, these are time-consuming computations for large system sizes, and ML methods can be employed to achieve large-scale dynamical simulations. In the following, we implement the descriptors discussed in Sec. III to the Holstein-type models.

A. Correlation matrix

We first discuss the descriptor based on the correlation matrix. Since here we are dealing with a scalar field, there is no internal index. Using the notations introduced to label the neighborhood sites in Sec. III, we arrange the lattice distortions into a vector \mathcal{U} with elements $\mathcal{U}_r = Q_{j_r} = Q(\mathbf{r}_{j_r})$. The matrix index $r = 1, 2, 3, \dots, L$, where L is the total number of sites in the neighborhood. We reserve $r = 1$ for the center site, i.e. $j_1 = i$. For convenience, we also define $R_{rs} = |\mathbf{r}_{j_r} - \mathbf{r}_{j_s}|$. The explicit definition of the correlation matrix is (c.f. Eq. (14))

$$\begin{aligned} C_{11} &= \mathcal{U}_1^2 \\ C_{rr} &= \mathcal{U}_r^2 / R_{1r}^2 & (r \neq 1) \\ C_{r1} &= C_{1r} = \mathcal{U}_1 \mathcal{U}_r / R_{1r}^2 & (r \neq 1) \\ C_{rs} &= C_{sr} = \mathcal{U}_r \mathcal{U}_s / R_{1r} R_{1s} R_{rs} & (r, s \neq 1, r \neq s) \end{aligned} \quad (65)$$

The dimension of the correlation matrix is given by the number of lattice sites L in the local neighborhood, which in our implementation contains a total of $L = 89$ sites (up to the 14th neighbors). Since this is a relatively small number, all eigenvalues of the C matrix are used for the descriptor.

We integrate the correlation-matrix descriptor with a neural network (NN) learning model to predict the electron force. As a proof of principle, we consider the Hol-

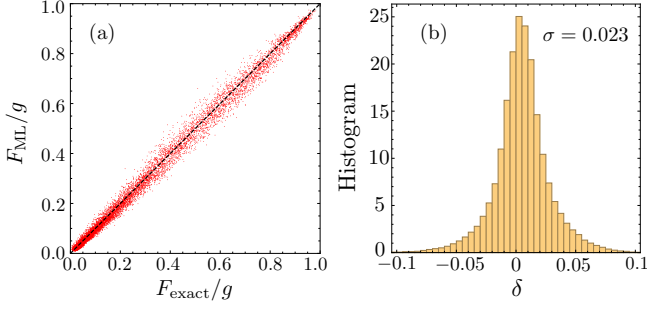


FIG. 4. (a) Electron forces predicted from ML model with the correlation-matrix descriptor versus exact solutions for test dataset of the Holstein model with $g = 1.5t$. Here the forces are normalized by the coupling constant g , hence are the same as the on-site electron density n_i . (b) Histogram of the force error $\delta = (F_{\text{ML}} - F_{\text{exact}})/g$.

stein model Eq. (60) without electron-electron interaction. By exactly diagonalizing the quadratic Hamiltonian, the electron force, which is proportional to the on-site electron density, is obtained from the eigenvectors. A six-layer NN model is trained from 2000 snapshots of a 30×30 system. Fig. 4 shows the ML predictions versus the exact forces. Here we plot the dimensionless forces normalized by the coupling constant g , which is the same as the on-site electron density. The histogram of the prediction error exhibits a small standard deviation $\sigma = 0.023$, indicating very good accuracy of the ML predicted forces.

A remark about the ML model. The results shown in Fig. 4 were obtained from an ML energy model based on the BP scheme shown in Fig. 1. The output of the NN is the local energy ϵ_i , and the force is obtained via automatic differentiation of the total energy. Since the scalar force is simply proportional to the electron density, one can apply the supervised learning to build a NN which directly predicts the on-site density $\langle \hat{n}_i \rangle$ from the lattice distortions $\{Q_j\}$ in the neighborhood. Interestingly, we found that the accuracy of this direct approach is worse than that based on the BP method. As already noted in previous works, the BP method ensures that the predicted forces are conservative as they are given by the derivative of an effective energy. The more constrained supervised learning in the BP scheme also helps with the prediction accuracy.

B. Bispectrum

Next we discuss the bispectrum descriptor of the Holstein-type models based on the group-theoretical method. The site-symmetry of the square lattice is described by the D_4 point group. As discussed above, the collection of on-site lattice distortions $\{Q_j\}$ in the neighborhood forms a high-dimensional representation of the D_4 group, which can be decomposed into the five irreducible representations: A_1 , A_2 , B_1 , B_2 , and E ; see Ta-

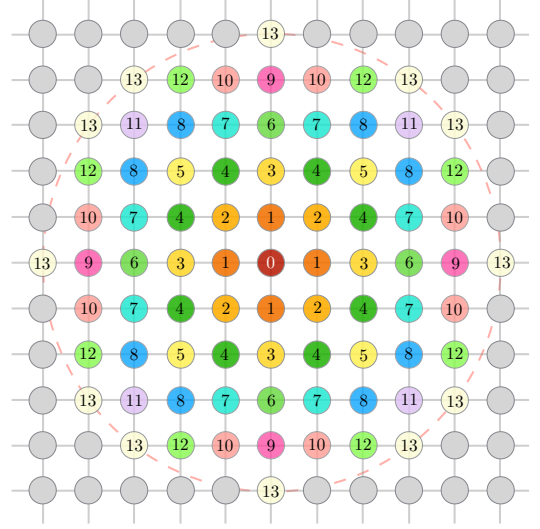


FIG. 5. Partition of lattice sites in the neighborhood C_i into groups (nearest neighbors, 2nd nearest neighbors, and so on) depending on their distance to the center site. Lattice sites belong to the same neighboring group are also related by symmetry operations of the point group.

ble I for the character table of the point group D_4 . The first four are singlet IRs, while E is a doublet representation. The task of the decomposition is made easier by noting that the $\{Q_j\}$ of same radius from the center form invariant blocks under the symmetry operations, i.e. the matrix representation of symmetry operations of D_4 are block-diagonalized with each block corresponding to a given radius; see Fig. 5.

In fact, direction examination shows that there are only two kinds of invariant blocks, one of size 4 and the other 8, illustrated in Fig. 6(a) and (b), respectively. Consequently, one only needs to decompose the resultant 4- and 8-dimensional reducible representations. The decomposition of the 4-site blocks in Fig. 6(a) is $4 = A_1 \oplus B_1 \oplus E$, with the following coefficients:

$$\begin{aligned} f^{A_1} &= Q_a + Q_b + Q_c + Q_d, \\ f^{B_1} &= Q_a - Q_b + Q_c - Q_d, \\ f_1^E &= Q_a - Q_c, \quad f_2^E = Q_b - Q_d. \end{aligned} \quad (66)$$

The decomposition of the 8-site block shown in Fig. 6(b)

D_4	E	$2C_4(z)$	$C_2(z)$	$2C_2'$	$2C_2''$	Linear func	Quadratic func
A_1	+1	+1	+1	+1	+1	—	$(x^2 + y^2), z^2$
A_2	+1	+1	+1	-1	-1	z	—
B_1	+1	-1	+1	+1	-1	—	$x^2 - y^2$
B_2	+1	-1	+1	-1	+1	—	xy
E	+2	0	-2	0	0	(x, y)	(xz, yz)

TABLE I. Character table for point group D_4 . Also shown are the linear and quadratic function representations of the various IRs.

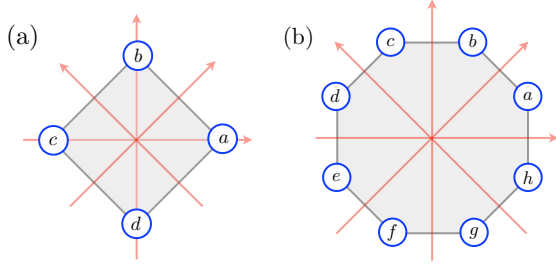


FIG. 6. The two basic types of neighboring groups with (a) 4 and (b) 8 sites. These groups form invariant blocks in the representations of the neighborhood classical field.

is: $8 = A_1 \oplus A_2 \oplus B_1 \oplus B_2 \oplus 2E$. Importantly, there are two doublet E IRs. The corresponding coefficients are

$$\begin{aligned}
 f^{A_1} &= Q_a + Q_b + Q_c + Q_d + Q_e + Q_f + Q_g + Q_h, \\
 f^{A_2} &= Q_a - Q_b + Q_c - Q_d + Q_e - Q_f + Q_g - Q_h, \\
 f^{B_1} &= Q_a - Q_b - Q_c + Q_d + Q_e - Q_f - Q_g + Q_h, \\
 f^{B_2} &= Q_a + Q_b - Q_c - Q_d + Q_e + Q_f - Q_g - Q_h, \\
 f_1^E &= Q_a - Q_e, & f_2^E &= Q_c - Q_g \\
 f_1^{E'} &= Q_b - Q_f, & f_2^{E'} &= Q_d - Q_h.
 \end{aligned} \tag{67}$$

Since for most point groups, the dimension of the IRs is often very small, and IR of the same transformation properties appears many times in the decomposition of the vector \vec{U} representation of the neighborhood, we label the IR index as $\Gamma = (\mathbb{T}, r)$, where \mathbb{T} denotes the symmetry type of the IR, and r enumerates the multiple occurrence of this symmetry in the decomposition. Using the above formulas for the different neighborhood blocks, we thus decompose the $\{Q_j\}$ variables into five different IRs $f^{(A_1, r)}$, $f^{(A_2, r)}$, $f^{(B_1, r)}$, $f^{(B_2, r)}$, and $\mathbf{f}^{(E, r)} = (f_1^{(E, r)}, f_2^{(E, r)})$.

Now that we have all the IR components, we can now use Eq. (33) to compute the bispectrum coefficients. To this end, we first list in Table II the decomposition of tensor product for the D_4 group. The nonzero Clebsch-Gordan (CG) coefficients of the tensor products can be found in, e.g. Ref. [108]. The bispectrum coefficients can be classified according to the tensor-product table. First, we list the coefficients involving only the singlets:

$$b_{r, r', r''}^{(A_1, A_1, A_1)} = f^{(A_1, r)} f^{(A_1, r')} f^{(A_1, r'')}, \tag{68a}$$

$$b_{r, r', r''}^{(A_1, A_2, A_2)} = f^{(A_1, r)} f^{(A_2, r')} f^{(A_2, r'')}, \tag{68b}$$

$$b_{r, r', r''}^{(A_1, B_1, B_1)} = f^{(A_1, r)} f^{(B_1, r')} f^{(B_1, r'')}, \tag{68c}$$

$$b_{r, r', r''}^{(A_1, B_2, B_2)} = f^{(A_1, r)} f^{(B_2, r')} f^{(B_2, r'')}, \tag{68d}$$

$$b_{r, r', r''}^{(A_2, B_1, B_2)} = f^{(A_2, r)} f^{(B_1, r')} f^{(B_2, r'')}. \tag{68e}$$

There are four different types of bispectrum coefficients involving the doublet. Their expressions can be simplified

	A_1	A_2	B_1	B_2	E
A_1	A_1	A_2	B_1	B_2	E
A_2		A_1	B_2	B_1	E
B_1			A_1	A_2	E
B_2				A_1	E
E					$A_1 \oplus A_2 \oplus B_1 \oplus B_2$

TABLE II. Direct products of irreducible representations of the D_4 point group.

fied using the Pauli matrices $\sigma_{1,2,3}$:

$$\begin{aligned}
 b_{r, r', r''}^{(A_1, E, E)} &= f^{(A_1, r)} \left(f_1^{(E, r')} f_1^{(E, r'')} + f_2^{(E, r')} f_2^{(E, r'')} \right) \\
 &= f^{(A_1, r)} \mathbf{f}^{(E, r')} \cdot \mathbf{f}^{(E, r'')}, \tag{69a}
 \end{aligned}$$

$$\begin{aligned}
 b_{r, r', r''}^{(A_2, E, E)} &= f^{(A_2, r)} \left(f_1^{(E, r')} f_2^{(E, r'')} - f_2^{(E, r')} f_1^{(E, r'')} \right) \\
 &= f^{(A_2, r)} \mathbf{f}^{(E, r')} \cdot (-i\sigma_2) \cdot \mathbf{f}^{(E, r'')}, \tag{69b}
 \end{aligned}$$

$$\begin{aligned}
 b_{r, r', r''}^{(B_1, E, E)} &= f^{(B_1, r)} \left(f_1^{(E, r')} f_1^{(E, r'')} - f_2^{(E, r')} f_2^{(E, r'')} \right) \\
 &= f^{(B_1, r)} \mathbf{f}^{(E, r')} \cdot \sigma_3 \cdot \mathbf{f}^{(E, r'')}, \tag{69c}
 \end{aligned}$$

$$\begin{aligned}
 b_{r, r', r''}^{(B_2, E, E)} &= f^{(B_2, r)} \left(f_1^{(E, r')} f_2^{(E, r'')} + f_2^{(E, r')} f_1^{(E, r'')} \right) \\
 &= f^{(B_2, r)} \mathbf{f}^{(E, r')} \cdot \sigma_1 \cdot \mathbf{f}^{(E, r'')}, \tag{69d}
 \end{aligned}$$

While bispectrum provides a complete description of the neighborhood within a cutoff, a formal descriptor based on bispectrum requires a large number of b coefficients, which makes it infeasible practically. Besides, the various coefficients b are not independent of each other. To simplify the calculation, our approach here is to use the power spectrum p_r^Γ supplemented by some of the bispectrum coefficients to obtain an equivalent, but more efficient, descriptor.

C. Reference coefficient for irreducible representations

The bispectrum provides a systematic method to obtain invariants which contain crucial information regarding the relative “phases” between different IRs. However, the set of all bispectrum coefficients listed in Eqs. (68) is obviously over-complete. For example, since $f_r^{(A_1)}$ is already an invariant itself, the $b_{r, r', r''}^{(A_1, A_1, A_1)}$ is redundant. Here we propose a novel method to retain the phase information based on the idea of *reference* coefficients for irreducible representations of the site-symmetry group. Importantly, this method allows us to significantly reduce the number of feature variables required to reconstruct the environment configuration module the site symmetry.

To demonstrate the idea of reference coefficients, we consider bispectrum $b^{(A_1, \Gamma, \Gamma)}$, where $\Gamma = A_2, B_1$, or B_2 is one of the singlet IR. For convenience, we define the phase of the singlet coefficient as

$$f^{(\Gamma, r)} = \sqrt{p_r^\Gamma} \eta_r^\Gamma, \tag{70}$$

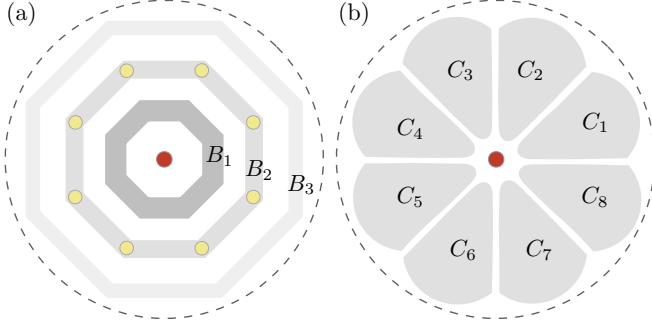


FIG. 7. Construction of the reference irreducible representations from (a) a particular neighboring group, and (b) symmetry-related clusters.

Since the A_1 part is already invariant under symmetry operations, the invariance of $b^{(A_1, \Gamma, \Gamma)}$ is equivalent to the invariance of the following product

$$f^{(\Gamma, r)} f^{(\Gamma, r')} = \sqrt{p_r^\Gamma p_{r'}^\Gamma} \eta_{r, r'}^\Gamma \quad (71)$$

where we have defined the relative phase of two expansion coefficients as

$$\eta_{r_1, r_2}^\Gamma = \eta_{r_1}^\Gamma \eta_{r_2}^\Gamma. \quad (72)$$

In addition to the power spectrum coefficients p^Γ , the relative phase η_{r_1, r_2}^Γ is a crucial invariant encoded in the bispectrum. However, the relative phases are not independent of each other. Indeed, from its definition it is straightforward to show that

$$\eta_{r_1, r_2}^\Gamma \eta_{r_2, r_3}^\Gamma \eta_{r_3, r_1}^\Gamma = 1. \quad (73)$$

To derive the set of truly independent phase coefficients, we introduce the phase η_*^Γ of a reference expansion coefficient $f^{(\Gamma, *)}$ of IR- Γ , and define the relative phase between a given coefficient $f^{(\Gamma, r)}$ and the reference one as $\eta_{*, r}^\Gamma$. The relative phase between two expansion coefficients of the same IR Γ is then given by

$$\eta_{r_1, r_2}^\Gamma = \eta_{*, r_1}^\Gamma \eta_{*, r_2}^\Gamma. \quad (74)$$

This relation thus allow us to reconstruct all the $b^{(A_1, \Gamma, \Gamma)}$ bispectrum coefficients.

It is worth noting that the purpose of the reference coefficient of a given IR is to construct invariants which retain the relative phase between a given expansion coefficient and the reference one. Consequently, the amplitude of the reference representations is completely irrelevant, as long as it is nonzero. Also importantly, there is no unique procedure to obtain these reference expansion coefficients $f^{(\Gamma, *)}$. Fig. 7 shows two examples of computing the reference coefficients from the neighborhood configuration. One approach is simply to use the decomposition of a particular 8-site neighbor block, e.g. the B_2 neighbors in Fig. 7(a), as the reference. However,

the small size of such an 8-site block might result in the undesirable situation of vanishing expansion coefficients for some IRs. While there is no general method to avoid such pathological situation, one can try to minimize such probability, which is already very small in most cases, by using weighted results from several neighbor blocks to derive the reference expansion coefficients.

Another approach, shown in Fig. 7(b), is based on symmetry-related clusters built from the neighborhood sites. First, a cluster of sites, say C_1 , is introduced either based on geometrical consideration, or simply randomly. Applying symmetry operations of the point group to C_1 then generates all clusters C_K that are related to C_1 by the site symmetry. Next we define the average of dynamical variables for each cluster: $Q_K = (1/M) \sum_{i \in C_K} Q_i$, where M is the number of sites in each cluster. Obviously, these cluster-based variables $\{Q_K\}$ form an 8-dimensional reducible representation. Its decomposition using Eq. (67) thus gives rise to a set of references coefficients $f^{(\Gamma, *)}$ for all five IRs of the D_4 group.

As discussed above, the amplitude of the reference expansion coefficients is unimportant for the descriptor, as the essential information is carried by their phase. In the following, we use the phase of the reference coefficients to define a new set of feature variables $\{g_r^{(\Gamma, r)}\}$ which are manifestly invariant under symmetry operations. Importantly, these coefficients supplemented by a few additional normalized bispectrum factor form a complete descriptor such that the local environment can be faithfully reconstructed from them. First, since expansion coefficients of the A_1 IR is already an invariant, we define $g^{(A_1, r)} = f^{(A_1, r)}$. For other singlet IRs $\Gamma = A_2, B_1$, and B_2 , the new coefficient is defined as

$$g^{(\Gamma, r)} = f^{(\Gamma, r)} \eta_*^\Gamma = \sqrt{p_r^\Gamma} \eta_{*, r}^\Gamma \quad (75)$$

where $\eta_{*, r}^\Gamma$ is the relative phase introduced in Eq. (74). This coefficient is obviously an invariant of the symmetry group. Importantly, the bispectrum coefficients in Eq. (68a)–(68d) can be readily expressed in terms of these invariant g coefficients:

$$b_{r, r', r''}^{(A_1, \Gamma, \Gamma)} = g^{(A_1, r)} g^{(\Gamma, r')} g^{(\Gamma, r'')}, \quad (76)$$

where $\Gamma = A_1, A_2, B_1, B_2$. We still need to consider the special bispectrum $b^{(A_2, B_1, B_2)}$ in Eq. (68e), which encodes the relative phases between expansion coefficients of the three 1D IRs. However, it is easy to show that expression in Eq. (68e) can be expressed as

$$b_{r, r', r''}^{(A_2, B_1, B_2)} = g^{(A_2, r)} g^{(B_1, r')} g^{(B_2, r'')} b_*^{(A_2, B_1, B_2)}, \quad (77)$$

where we have introduced a normalized reference bispectrum

$$b_*^{(A_2, B_1, B_2)} = \eta_*^{(A_2)} \eta_*^{(B_1)} \eta_*^{(B_2)}, \quad (78)$$

It can be readily checked that this normalized bispectrum of the reference coefficient is invariant under symmetry operation. Eqs. (76) and (77) indicate that all

bispectrum coefficients of Eq. (68) can be restored using invariant coefficients $g^{(\Gamma,r)}$ and $b_*^{(A_2,B_1,B_2)}$.

Next we consider the bispectrum coefficients in Eq. (69) that involve the doublet E IRs. We define a similar normalized doublet vector $\epsilon_* = (\epsilon_*^x, \epsilon_*^y)$ from the reference coefficients of the doublet IR:

$$\epsilon_*^x = f_1^{(E,*)}/|f_1^{(E,*)}|, \quad \epsilon_*^y = f_2^{(E,*)}/|f_2^{(E,*)}|. \quad (79)$$

The x and y components have amplitude $|\epsilon_*^{x,y}| = 1$. It can be readily checked that this 2-component vector ϵ_* transforms as a doublet IR under the D_4 group. Consequently, ϵ_* can be used to build invariant coefficients using the the bispectrum formula in Eq. (69). Specifically, for any given doublet vector $\mathbf{f}^{(E,r)}$, we introduce two invariant coefficients $\mathbf{g}^{(E,r)} = (g_1^{(E,r)}, g_2^{(E,r)})$ defined as

$$\begin{aligned} g_1^{(E,r)} &= \epsilon_* \cdot \mathbf{f}^{(E,r)} / \sqrt{2} \\ &= (\epsilon_*^x f_1^{(E,r)} + \epsilon_*^y f_2^{(E,r)}) / \sqrt{2}, \end{aligned} \quad (80a)$$

$$\begin{aligned} g_2^{(E,r)} &= \eta_*^{(B_1)} \epsilon_* \cdot \boldsymbol{\sigma}_3 \cdot \mathbf{f}^{(E,r)} / \sqrt{2} \\ &= \eta_*^{(B_1)} (\epsilon_*^x f_1^{(E,r)} - \epsilon_*^y f_2^{(E,r)}) / \sqrt{2}. \end{aligned} \quad (80b)$$

This can be readily inverted to give

$$f_1^{(E,r)} = \epsilon_*^x (g_1^{(E,r)} + \eta_*^{(B_1)} g_2^{(E,r)}) / \sqrt{2}, \quad (81a)$$

$$f_2^{(E,r)} = \epsilon_*^y (g_1^{(E,r)} - \eta_*^{(B_1)} g_2^{(E,r)}) / \sqrt{2}. \quad (81b)$$

By substituting the above expressions for $f_{1,2}^{(E,r)}$ into Eqs. (69), we can express the bispectrum coefficients in terms of the invariant g coefficients

$$b_{r,r',r''}^{(A_1,E,E)} = g^{(A_1,r)} g^{(E,r')} \cdot g^{(E,r'')}, \quad (82a)$$

$$\begin{aligned} b_{r,r',r''}^{(A_2,E,E)} &= b_*^{(A_2,B_1,B_2)} b_*^{(B_2,E,E)} \\ &\quad \times g^{(A_2,r)} g^{(E,r')} \cdot (-i\boldsymbol{\sigma}_2) \cdot g^{(E,r'')}, \end{aligned} \quad (82b)$$

$$b_{r,r',r''}^{(B_1,E,E)} = g^{(B_1,r)} g^{(E,r')} \cdot \boldsymbol{\sigma}_1 \cdot g^{(E,r'')}, \quad (82c)$$

$$b_{r,r',r''}^{(B_2,E,E)} = b_*^{(B_2,E,E)} g^{(B_2,r)} g^{(E,r')} \cdot \boldsymbol{\sigma}_3 \cdot g^{(E,r'')}. \quad (82d)$$

Here we have defined another important normalized reference bispectrum coefficient

$$b_*^{(B_2,E,E)} = \eta_*^{(B_2)} \epsilon_*^x \epsilon_*^y. \quad (83)$$

which encodes the relative phase between the reference B_2 and E coefficients.

To summarize, a complete description of the local environment is given by the following set of invariant coefficients:

$$\left\{ g^{(A_1,r)}, g^{(A_2,r)}, g^{(B_1,r)}, g^{(B_2,r)}, g^{(E,r)}, b_*^{(A_2,B_1,B_2)}, b_*^{(B_2,E,E)} \right\}. \quad (84)$$

We integrate this descriptor with a six-layer NN to develop a ML energy model following the framework shown

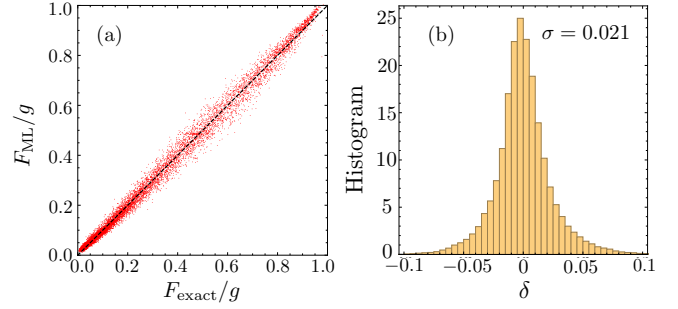


FIG. 8. (a) Electron forces predicted from ML model with the bispectrum descriptor versus exact solutions for test dataset of the Holstein model with $g = 1.5t$. Here the forces are normalized by the coupling constant g , hence are the same as the on-site electron density n_i . (b) Histogram of the force errors $\delta = (F_{ML} - F_{exact})/g$.

in Fig. 1. The NN is trained from exact solution of the Holstein model Eq. (60) on a 30×30 square lattice. The ML predicted normalized forces, shown in Fig. 8, agree very well with the exact values, as evidenced by the rather small standard deviation from the histogram of prediction errors.

It is worth noting that, thanks to the regularity of the lattice geometry and the simplicity of the scalar classical field, decent force prediction can be obtained using a ML model even without a descriptor, namely directly using an array of the scalar variables $\{Q_j\}$ in the neighborhood as the input. However, the site symmetry is an approximate symmetry in such naive approaches, though the accuracy of the symmetry could be improved by increasing the size of the training dataset. On the other hand, the employment of the lattice descriptor ensures that the ML force field model preserves the lattice symmetry.

V. EXAMPLE: DYNAMICS OF COOPERATIVE JAHN-TELLER COUPLING

As a second example of the type-I models, we consider the Jahn-Teller (JT) coupling between the e_g electrons and the distortion modes of local MnO_6 octahedron in manganites. The JT-coupling, along with the DE mechanism, are important for the physics of colossal magnetoresistance and polaron liquids [98, 99, 111, 112]. Here the local distortion of the tetrahedron is characterized by three modes. The symmetric breathing mode Q_i^1 is essentially the same as the Holstein phonons discussed in the previous Section IV. Here we are interested in the dynamics of the asymmetric normal modes that are described by a e_g doublet $\mathbf{Q}_i = (Q_i^x, Q_i^z)$. Here the Q_i^x component has the symmetry of $(x^2 - y^2)$, and Q_i^z has the symmetry of $(3z^2 - r^2)$. Here we consider the two-orbital electron Hamiltonian with the JT coupling on a

square lattice [98, 99]

$$\hat{\mathcal{H}} = - \sum_{\langle ij \rangle} \sum_{\alpha\beta=a,b} \left(t_{ij}^{\alpha\beta} \hat{c}_{i\alpha}^\dagger \hat{c}_{j\beta} + \text{h.c.} \right) \quad (85)$$

$$- g \sum_i \left[Q_i^x \left(\hat{c}_{ia}^\dagger \hat{c}_{ib} + \hat{c}_{ib}^\dagger \hat{c}_{ia} \right) + Q_i^z \left(\hat{c}_{ia}^\dagger \hat{c}_{ia} - \hat{c}_{ib}^\dagger \hat{c}_{ib} \right) \right],$$

Here $\hat{c}_{i\alpha}^\dagger/\hat{c}_{i,\alpha}$ are creation/annihilation operators of electron at site- i with an e_g orbital-index $\alpha = a, b$, corresponding to orbitals $d_{x^2-y^2}$ and $d_{3z^2-r^2}$, respectively. The superscript x, z indicates that the corresponding Q -mode couples to the x and z pseudo-spin of the orbitally degenerate e_g electrons. Because of the orbital degrees of freedom, the electron hopping coefficients are anisotropic: $t_{ij}^{aa} = -\sqrt{3}t_{ij}^{ab} = -\sqrt{3}t_{ij}^{ba} = 3t_{ij}^{bb} = t$ for $\langle ij \rangle$ along the x -direction, and $t_{ij}^{aa} = \sqrt{3}t_{ij}^{ab} = \sqrt{3}t_{ij}^{ba} = 3t_{ij}^{bb} = t$ for $\langle ij \rangle$ along the y -direction.

The adiabatic dynamics of the JT phonons is described by a similar Langevin equation

$$\mu \frac{d^2 \mathbf{Q}_i}{dt^2} + \lambda \frac{d\mathbf{Q}_i}{dt} = - \frac{\partial \mathcal{V}}{\partial \mathbf{Q}_i} - \frac{\partial \langle \hat{\mathcal{H}} \rangle}{\partial \mathbf{Q}_i} + \boldsymbol{\eta}_i(t). \quad (86)$$

Here μ is the effective mass, λ is the dissipation coefficient, and $\boldsymbol{\eta}_i(t)$ represent stochastic thermal forces. We have also included the classical elastic energy \mathcal{V} of the JT phonons [98, 99]. Again, the time-consuming part, which has to be carried out at every time-step, is the calculation of the electron forces:

$$\mathbf{F}_i^{\text{elec}} = - \frac{\partial \langle \hat{\mathcal{H}} \rangle}{\partial \mathbf{Q}_i} \quad (87)$$

$$= g \left(\langle \hat{c}_{ia}^\dagger \hat{c}_{ib} + \hat{c}_{ib}^\dagger \hat{c}_{ia} \rangle, \langle \hat{c}_{ia}^\dagger \hat{c}_{ia} - \hat{c}_{ib}^\dagger \hat{c}_{ib} \rangle \right).$$

The two components of the force are given by the on-site two-point correlation function $\rho_{i\alpha,i\beta} = \langle \hat{c}_{i,\beta}^\dagger \hat{c}_{i,\alpha} \rangle$, which can be obtained by exact diagonalization for the JT Hamiltonian above. Next we outline the bispectrum descriptors for the JT doubles, which can be combined with a learning model to develop an effective classical energy, as outlined in Fig. 1 for the adiabatic dynamics of the JT phonons.

First, we note that the JT doublet \mathbf{Q}_i has the transformation property of an e_g doublet under the cubic point group, e.g. O_h . Since here we consider the two-dimensional model, which is relevant for applications of CMR manganites, the two components of \mathbf{Q}_i transform independently of each other under the in-plane D_4 point group of the square lattice. As shown in Table I, while the Q_i^z mode transforms with the A_1 symmetry, the Q_i^x mode behaves as the B_1 IR. Consequently, the Q_i^z components are essentially a scalar field, and their decomposition can be obtained following exactly the same procedure for the Holstein model outlined in Sec. IV B, and the corresponding coefficients are given by Eqs. (66) and (67). On the other hand, the fact that the Q_i^x mode transforms as an B_1 IR mixes the transformations of the lattice sites with those of Q^x amplitude. Take the four-site neighboring

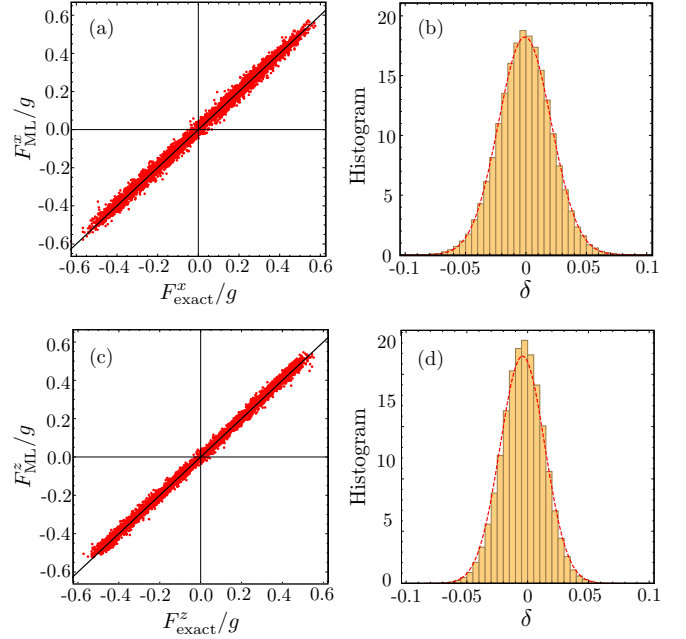


FIG. 9. Electron forces predicted from ML model with the bispectrum descriptor versus exact solutions for test dataset of the Jahn-Teller model with $g = 1.5t$ for the (a) Q^x and (c) Q^z components. Here the forces are normalized by the coupling constant g , hence are the same as the on-site electron density n_i . The corresponding histograms of the force errors $\delta = (F_{\text{ML}} - F_{\text{exact}})/g$ are shown in panel (b) and (d), respectively.

group in Fig. 6(a) as an example, under the $C_4(z)$ rotation, the four Q^2 modes acquire an additional -1 sign in addition to the lattice rotation:

$$\begin{aligned} Q_a^x &\rightarrow -Q_b^x, & Q_b^x &\rightarrow -Q_c^x \\ Q_c^x &\rightarrow -Q_d^x, & Q_d^x &\rightarrow -Q_a^x \end{aligned}$$

To account for this additional -1 sign, the expansion coefficients of the corresponding IRs are given by

$$\begin{aligned} f^{A_1} &= Q_a^x - Q_b^x + Q_c^x - Q_d^x, \\ f^{B_1} &= Q_a^x + Q_b^x + Q_c^x + Q_d^x, \\ f_1^E &= Q_a^x - Q_c^x, & f_2^E &= Q_d^x - Q_b^x. \end{aligned} \quad (88)$$

Similar expressions can be obtained for the 8-site neighboring block.

$$\begin{aligned} f^{A_1} &= Q_a^x - Q_b^x - Q_c^x + Q_d^x + Q_e^x - Q_f^x - Q_g^x + Q_h^x, \\ f^{A_2} &= Q_a^x + Q_b^x - Q_c^x - Q_d^x + Q_e^x + Q_f^x - Q_g^x - Q_h^x, \\ f^{B_1} &= Q_a^x + Q_b^x + Q_c^x + Q_d^x + Q_e^x + Q_f^x + Q_g^x + Q_h^x, \\ f^{B_2} &= Q_a^x - Q_b^x + Q_c^x - Q_d^x + Q_e^x - Q_f^x + Q_g^x - Q_h^x, \\ f_1^E &= Q_a^x - Q_e^x, & f_2^E &= Q_g^x - Q_c^x \\ f_1^{E'} &= Q_b^x - Q_f^x, & f_2^{E'} &= Q_h^x - Q_d^x. \end{aligned} \quad (89)$$

Given the expansion coefficients of both the Q^x and Q^z phonons, bispectrum coefficients, with contributions from both modes, can be obtained as outlined in

Sec. IV B. The implementation can similarly be simplified using reference IR method. Combining the descriptor with a neural network, an effective energy model is trained by datasets from exact diagonalization of the JT Hamiltonian on 30×30 lattices. The ML predicted forces for both components versus the exact values are shown in Fig. 9 along with the histograms of the prediction errors.

VI. EXAMPLE: ADIABATIC DYNAMICS OF ITINERANT MAGNETS

As a final example, in this Section we demonstrate the descriptors for the spin dynamics of itinerant magnets, which are a representative example of the type-II models. Explicitly, we consider the following single-band s-d model [56, 57]

$$\hat{\mathcal{H}} = -t \sum_{\langle ij \rangle} \left(\hat{c}_{i\alpha}^\dagger \hat{c}_{j\alpha} + \text{h.c.} \right) - J \sum_i \mathbf{S}_i \cdot \hat{c}_{i\alpha}^\dagger \boldsymbol{\sigma}_{\alpha\beta} \hat{c}_{i\beta}, \quad (90)$$

where $\hat{c}_{i\alpha}^\dagger/\hat{c}_{i\alpha}$ are creation/annihilation operators of electron with spin $\alpha = \uparrow, \downarrow$ at site i , $\langle ij \rangle$ indicates the nearest neighbors, t is the electron hopping constant, J_H is the local Hund's rule coupling between electron spin and local magnetic moment \mathbf{S}_i due to the localized d or f electrons. Here repeated indices α, β imply summation. The s-d Hamiltonian offers a fundamental description of the electron-spin interaction; it has been widely used to model the dynamics of magnetic textures, such as domain-walls and skyrmions, under the influence of conducting electrons. The strong coupling $J \gg t$ limit of the s-d Hamiltonian, also known as the double-exchange model, exhibits intriguing phase-separated states in which ferromagnetic clusters with low electron density are mixed with half-filled antiferromagnetic domains [56]. Such electronic phase separation is crucial to the emergence of novel material functionalities such as colossal magnetoresistance and high- T_c superconductivity.

The adiabatic dynamics of the classical spins is governed by the stochastic Landau-Lifshitz-Gilbert (LLG) equation [113, 114]

$$\frac{d\mathbf{S}_i}{dt} = \gamma \mathbf{S}_i \times \left(\frac{\partial \langle \hat{\mathcal{H}} \rangle}{\partial \mathbf{S}_i} + \boldsymbol{\eta}_i \right) - \lambda \mathbf{S}_i \times \left(\mathbf{S}_i \times \frac{\partial \langle \hat{\mathcal{H}} \rangle}{\partial \mathbf{S}_i} \right), \quad (91)$$

where $\boldsymbol{\eta}_i(t)$ is a stochastic local field of zero mean, and λ is the damping coefficient. The computational overhead is dominated by the calculation of the electron contribution to the exchange forces:

$$\mathbf{H}_i^{\text{elec}} = - \frac{\partial \langle \hat{\mathcal{H}} \rangle}{\partial \mathbf{S}_i} = J \boldsymbol{\sigma}_{\alpha\beta} \langle \hat{c}_{i\alpha}^\dagger \hat{c}_{i\beta} \rangle. \quad (92)$$

As demonstrated in Fig. 1 and in previous works [72, 73], the calculation of this electronic force can be speed up with the use of ML energy models. Here we present details of the descriptors for the case of dynamical spins, or the vector field.

The local spins in the standard s-d model are often assumed to have a fixed length. On the other hand, Eq. (90) can also be viewed as an effective Hamiltonian for the spin-density wave (SDW) obtained from a mean-field treatment of the interacting electron Hamiltonian. In the case of the SDW, the spin length itself is a dynamical variable, and the Gilbert damping in Eq. (91) has to be replaced by a Langevin-type dissipation force [62]. Here we consider the general case of descriptor for a vector field $\mathbf{S}(\mathbf{r}_i) = (S_i^x, S_i^y, S_i^z)$ with variable spin lengths.

The internal symmetry of the vector field is the $\text{SO}(3)$ rotation group, which is locally isomorphic to the $\text{SU}(2)$ group. The IRs of the rotation group is labeled by the angular momentum quantum number j which can be either integers or half-integers. The 3-component vector $\mathbf{S} = (S^x, S^y, S^z)$ is already an IR of dimension 3 of the rotation group, which is equivalent to the angular momentum $j = 1$ representation. In order to construct the bispectrum coefficients of the $\text{SO}(3)$ group, we consider the tensor product of two spins $\mathbf{S}_j \otimes \mathbf{S}_k$ in the neighborhood. Since this can also be viewed as the tensor product $(j = 1) \otimes (j = 1)$, from the theory of angular momentum addition, it can be decomposed to a direct sum of $j = 0$ (scalar), $j = 1$ (vector), and $j = 2$ (rank-2 traceless symmetric tensor).

The scalar component ($j = 0$) is simply the inner product of the two vectors

$$\mathbf{p}_{jk} = \mathbf{S}_j \cdot \mathbf{S}_k, \quad (93)$$

This scalar invariant is the same as the generalized power spectrum coefficient Eq. (44) in Sec. III, which is a special case of the bispectrum coefficients obtained from the trivial representation and two $j = 1$ IRs. The invariants \mathbf{p}_{jk} with $j \neq k$ represents the rotation-invariant spin-spin correlation, sometimes also called a bond variable. For the case of SDWs, one also needs to consider the on-site invariant $\mathbf{p}_{jj} = |\mathbf{S}_j|^2$, which corresponds to the spin length.

The $j = 1$ IR in the decomposition of $\mathbf{S}_j \otimes \mathbf{S}_k$ is the vector product $\mathbf{S}_j \times \mathbf{S}_k$, which can be combined with a third vector to form a scalar invariant under rotation

$$\mathbf{b}_{jkl} = \mathbf{S}_i \times \mathbf{S}_k \cdot \mathbf{S}_l. \quad (94)$$

These are the bispectrum coefficients Eq. (45) constructed from three $j = 1$ IRs of the $\text{SO}(3)$ group. This scalar invariant \mathbf{b}_{jkl} , also called the scalar spin chirality, provides a measure of the non-coplanarity of the triplet spins. Importantly, according to the bispectrum theory of group representations, the collection of all scalars \mathbf{p}_{jk} and \mathbf{b}_{jkl} provides a faithful representation of the neighborhood that is invariant with respect to $\text{SO}(3)$ rotations. On the other hand, this is obviously an over-complete representation. For example, the relative orientation between three spins (ijk) can be completely specified by three bond variables \mathbf{p}_{ij} , \mathbf{p}_{ik} , and \mathbf{p}_{jk} ; consequently, the scalar chirality \mathbf{b}_{ijk} is redundant. This also means that an equally faithful description of three spins can be attained from \mathbf{p}_{ij} , \mathbf{p}_{ik} , and \mathbf{b}_{ijk} . However, the effectiveness

of the descriptor depends on the choice of the feature variables, and we find that explicit inclusion of the scalar chirality is very efficient for describing magnetic states with non-coplanar spins.

A. Atom-centered symmetry functions

As a first example, we discuss the symmetry functions based on the rotation-invariant building blocks discussed above. Following Eqs. (53)–(56) in Sec. III D, we define the following symmetry functions associated with a center spin at \mathbf{r}_i . First, since there is no on-site singlet (trivial) IR \mathbf{f}_j for the spin model, there is no G_{2a} . The only nontrivial two-body symmetry function is

$$G_{2b}(\{\xi_m\}) = \sum_{j \neq i} F'_2(R_{ij}; \{\xi_m\})(\mathbf{S}_i \cdot \mathbf{S}_j). \quad (95)$$

And there are also two 3-body invariants involving a pair of spins (jk) along with the center site- i :

$$G_{3a}(\{\xi_m\}) = \sum_{jk \neq i} F_3(R_{ij}, R_{ik}, R_{jk}; \theta_{ijk})(\mathbf{S}_j \cdot \mathbf{S}_k), \quad (96)$$

$$G_{3b}(\{\xi_m\}) = \sum_{jk \neq i} F'_3(R_{ij}, R_{ik}, R_{jk}; \theta_{ijk})(\mathbf{S}_i \cdot \mathbf{S}_j \times \mathbf{S}_k). \quad (97)$$

We note that these symmetry functions can also be applied to describe local spin environment in disordered atomic systems. For example, it could be used to develop a ML model for the dynamics of spin glasses, as exemplified by the dilute magnetic alloys such as CuMn. The interaction between the randomly distributed magnetic atoms in spin glass is mediated by conducting electrons of the metallic matrix. Conventionally, this effective spin-spin interaction is modeled by integrating out electrons beforehand, giving rise to the Ruderman-Kittel-Kasuya-Yosida (RKKY) pair interaction at weak coupling. The above magnetic ACSF descriptor combined with a learning model can provide large-scale dynamical simulations of spin-glass alloys with a better accuracy even for strong electron-strong coupling. Another application is to combine the magnetic ACSF descriptor with molecular dynamics for the simulations of magnetic molecules in, e.g. ferrofluids [115, 116].

Here we apply the descriptor to the case of s-d model on a square lattice. A two-parameter function of the form Eq. (50) is used for both F_2 and F'_2 . Specifically we use $F_2(R; d, w) = \exp[-(R - d)^2/w^2]$ to extract spin correlations at lattice sites in a ring of width $\sim \xi_2$ at a distance ξ_1 from the center site; a hard cutoff at R_c is used. On the other hand, since all triplet angles θ_{ijk} are pre-defined on a lattice, we ignore the angular dependence of the F_3 function. Instead of Eq. (52), we use the following parameterization:

$$F_3(R_{ij}, R_{ik}, R_{jk}) = e^{-\frac{(R_{ij}-d)^2 + (R_{ik}-d)^2}{w^2}} e^{-\frac{(R_{jk}-d')^2}{w'^2}}.$$

In both 2- and 3-body symmetry functions, a finite width w is used to ensure overlaps between consecutive rings,

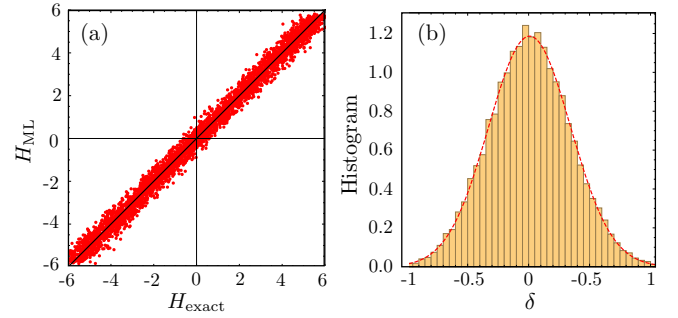


FIG. 10. (a) Exchange forces predicted by the ML model with the ACSF descriptor versus the exact solution from the exact diagonalization for the s-d model with exchange coupling $J = 6t$. (b) Histogram of the force prediction error.

thus avoiding the spurious symmetry due to independence of rings. Combining the ACSF descriptor with a NN model for the s-d model, Fig. 10 shows the ML predicted exchange forces versus the exact values. Although the ML prediction overall agrees well with the exact calculation, a rather large error was obtained with the ACSF descriptor, especially compared with the prediction error of the ML model with the bispectrum descriptor and the same NN structure to be discussed below. This is partly due to the fact that, from the representation theory point of view, the ACSF descriptor is mostly dominated by the fully symmetric A_1 representation, as also discussed in Sec. III D. Contributions from other nontrivial IRs could be implicitly included in ACSF of higher order. For example, consider the 4-body symmetry functions (not included in our implementation)

$$G_4 = \sum_{jkl \neq i} F_4(R_{ij}, \dots)(\mathbf{S}_i \cdot \mathbf{S}_j)(\mathbf{S}_k \cdot \mathbf{S}_l). \quad (98)$$

This is a summation of the A_1 IR of four-spin variables. As shown in II, this could result from the direct product of two-spin IR of either A_2 , B_1 , B_2 , or E symmetries, which cannot be captured by either of the two-spin symmetry functions in Eq. (95) and (96).

B. Bispectrum

Compared with the ACSF, the bispectrum coefficients based on the group-theoretical method provides a more systematic approach to build the descriptor. Following the discussion in Sec. III C, we consider a vector $\vec{U} = \{\mathbf{p}_{jk}, \mathbf{b}_{jkl}\}$ consisting of the bond and scalar chirality variables, which are already invariant with respect to the internal $SO(3)$ rotations. This vector, which is a high-dimensional representation of the point group associated with a given center site, is then decomposed into the IRs. As discussed above, the collection of all such variables is an over-complete representation. A proper downsizing to avoid too much overlap is required, also for practical reasons. To this end, we restrict ourselves to three types of invariants shown in Fig. 11: (i) bond

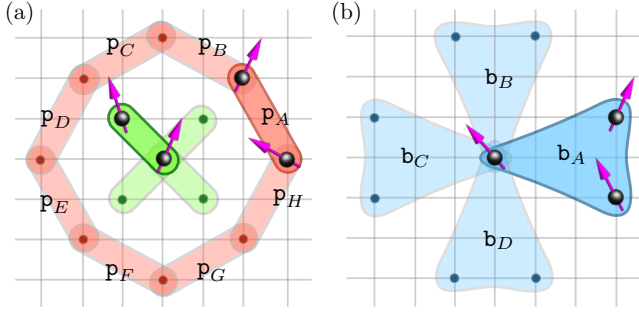


FIG. 11. Examples of (a) bond variables \mathbf{p}_{jk} and (b) spin scalar chirality \mathbf{b}_{ijk} that are related by the point group symmetry of the lattice. These 2-spin and 3-spin correlations are also invariants of the internal $\text{SO}(3)$ symmetry, and are building blocks for computing the bispectrum coefficients of the point group.

variables \mathbf{p}_{ij} between center site and another site- j in the neighborhood, (ii) bond-variables \mathbf{p}_{jk} between two sites that are different from the center, and (iii) scalar chirality variables from triplets (ijk) that include the center spin. Notably, they correspond to those used in the symmetry functions G_{2b} , G_{3a} , and G_{3b} , respectively, discussed in the previous subsection.

Even with this simplification, there are still a large number of these local variables. Fortunately, as discussed above, the vector representation for $\vec{\mathcal{U}}$ is block-diagonalized with each block consisting of spin-pairs or triplets with a fixed distance from the center site. For the case of square lattice, that means each block has a dimension of 4 or 8; there are also 12-member block, but these are trivial union of the dim-4 and dim-8 blocks. Examples of these blocks that are close on themselves under lattice rotation/reflection are shown in Fig. 11. The decomposition of these blocks is similar to that discussed in Sec. IV B for the scalar field. For the dimension-4 blocks, such as the four scalar chirality variables shown in Fig. 11(b), we have $4 = A_1 \oplus B_1 \oplus E$:

$$\begin{aligned} f^{A_1} &= \mathbf{b}_A + \mathbf{b}_B + \mathbf{b}_C + \mathbf{b}_D, \\ f^{B_1} &= \mathbf{b}_A - \mathbf{b}_B + \mathbf{b}_C - \mathbf{b}_D, \\ f^E &= (\mathbf{b}_A - \mathbf{b}_B, \mathbf{b}_C - \mathbf{b}_D). \end{aligned} \quad (99)$$

And the decomposition of a dim-8 block is illustrated by the eight off-center bond variables in Fig. 11(a): $8 = A_1 \oplus A_2 \oplus B_1 \oplus B_2 \oplus 2E$. The corresponding expansion coefficients are

$$\begin{aligned} f^{A_1} &= \mathbf{p}_A + \mathbf{p}_B + \mathbf{p}_C + \mathbf{p}_D + \mathbf{p}_E + \mathbf{p}_F + \mathbf{p}_G + \mathbf{p}_H, \\ f^{A_2} &= \mathbf{p}_A - \mathbf{p}_B + \mathbf{p}_C - \mathbf{p}_D + \mathbf{p}_E - \mathbf{p}_F + \mathbf{p}_G - \mathbf{p}_H, \\ f^{B_1} &= \mathbf{p}_A - \mathbf{p}_B - \mathbf{p}_C + \mathbf{p}_D + \mathbf{p}_E - \mathbf{p}_F - \mathbf{p}_G + \mathbf{p}_H, \\ f^{B_2} &= \mathbf{p}_A + \mathbf{p}_B - \mathbf{p}_C - \mathbf{p}_D + \mathbf{p}_E + \mathbf{p}_F - \mathbf{p}_G - \mathbf{p}_H, \\ f^E &= (\mathbf{p}_A - \mathbf{p}_E, \mathbf{p}_C - \mathbf{p}_G), \\ f^{E'} &= (\mathbf{p}_B - \mathbf{p}_F, \mathbf{p}_D - \mathbf{p}_H). \end{aligned} \quad (100)$$

Applying these decompositions to all blocks in the neighborhood, one obtains the coefficients of all IRs $\{\mathbf{f}^\Gamma\}$ from

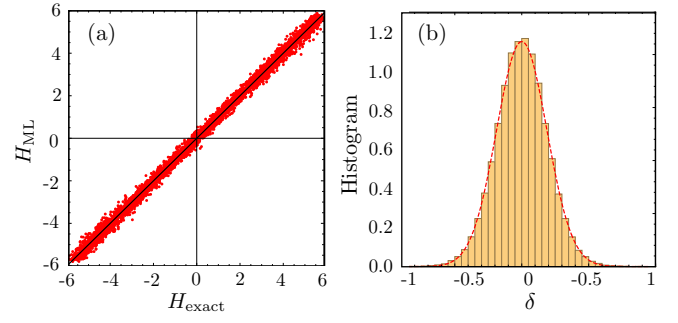


FIG. 12. (a) Exchange forces predicted by the ML model with the bispectrum descriptor versus the exact solution from the exact diagonalization for the s-d model with exchange coupling $J = 6t$. (b) Histogram of the force prediction error.

the reducible representation of bond and scalar chirality variables $\vec{\mathcal{U}} = \{\mathbf{p}_{jk}, \mathbf{b}_{jkl}\}$ within the cutoff radius. Here the basis functions of a given irrep are arranged into a vector $\mathbf{f}^\Gamma = (f_1^\Gamma, f_2^\Gamma, \dots, f_{n_r}^\Gamma)$, where Γ labels the symmetry of the IR, and r enumerates the multiple occurrence of Γ in the decomposition of $\vec{\mathcal{U}}$.

Feature variables that are invariant under the point group symmetry are given by the bispectrum coefficients computed from these expansion coefficients. Details of the calculation of the bispectrum with the aid of the reference IRs can be found in Sec. IV B for the case of square lattice. Here we outline the procedure from a different perspective. First, the power spectrum $p^\Gamma = |\mathbf{f}^\Gamma|^2$ is obviously invariant under discrete symmetry operations of the point group. However, power-spectrum descriptor contains spurious symmetries since it does not take into account the fact that different IR need to transform consistently, instead of independently, under symmetry operations. For example, the angle $\cos \theta_{12} = (\mathbf{f}^{(E,1)} \cdot \mathbf{f}^{(E,2)}) / (|\mathbf{f}^{(E,1)}| |\mathbf{f}^{(E,2)}|)$ between the vectors of two doublet IR characterizes the relative orientation of the two doublets, and is also an invariant of the point group. Consequently, the relative phases between different IRs should also be included in the descriptor in addition to the power spectrum [87].

The reference IR $\mathbf{f}^{(\Gamma,*)}$ discussed in Sec. IV C offers a way to properly incorporate the “phase” of the IRs into the descriptor. These reference coefficients 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 a irrep as the projection of its basis functions onto the reference basis: $\eta^\Gamma \equiv \mathbf{f}^\Gamma \cdot \mathbf{f}_{\text{ref}}^\Gamma / (|\mathbf{f}^\Gamma| |\mathbf{f}_{\text{ref}}^\Gamma|)$. The feature variables of the descriptor are then the collection of power spectrum coefficients and the relative phases: $\{p^\Gamma, \eta^\Gamma\}$. The various steps in the process of obtaining the descriptor are summarized in the following

$$\mathcal{C}_i \rightarrow \{\mathbf{p}_{jk}, \mathbf{b}_{jkl}\} \rightarrow \{\mathbf{f}^\Gamma\} \rightarrow \{p^\Gamma, \eta^\Gamma\}$$

The invariant feature variables characterizing the neighborhood spins, are then forwarded to the neural network which produces the local energy at its output node. This

means the local energy associated with spin \mathbf{S}_i depends on its neighborhood through the effective coordinates: $\varepsilon(\mathcal{C}_i) = \varepsilon(\{p^\Gamma, \eta^\Gamma\})$, which obviously preserves both the SO(3) spin-rotational symmetry and the discrete lattice symmetry.

A six-layer NN model is constructed and trained using PyTorch [117–122]. The training dataset consists of 3500 snapshots of spins and local exchange forces, obtained from exact diagonalization of a 30×30 lattice. Fig. 12(a) shows components of local exchange forces \mathbf{H}_i predicted by our trained NN model versus the exact results on test datasets. The difference $\delta = H_{\text{ML}} - H_{\text{exact}}$ is well described by a Gaussian distribution with a rather small mean-square error of $\sigma^2 = 0.035$, as shown in Fig. 12(b). Interestingly, the histogram of the deviation δ implies that the statistical error of the ML model can be interpreted as an effective or artificial temperature in Langevin dynamics.

VII. SUMMARY AND DISCUSSION

In this work, we present a numerical framework of utilizing machine learning methods for multi-scale dynamical modeling of condensed matter systems with emergent dynamical classical fields. These classical degrees of freedom could arise from the coupling to lattice dynamics, or magnetic moments of localized d or f electrons. They could also represent the collective electron behaviors, as exemplified by the order-parameter field in a symmetry breaking phase, of interacting electrons. The slow adiabatic dynamics of the emergent classical fields is often dominated by the electrons or quasi-particles, which are assumed to be in quasi-equilibrium of the instantaneous Hamiltonian parameterized by the classical variables. As in the quantum or *ab initio* molecular dynamics methods, accurate simulation of the dynamical classical fields requires solving the electronic structure problem at every time-step. Motivated by the success of ML-enabled large scale quantum MD simulations, we propose a similar approach for condensed matter systems in which the complex dependence of the local energy on the neighborhood classical field is encoded in a ML energy model.

The two important components of the ML energy model are the descriptor for characterizing the local classical field configuration, and the learning model used to encode the dependence on the local environment. Several learning models developed in the context of quantum MD can also be used for the effective energy model of the condensed-matter systems. Among the various ML models, the deep-learning neural network (NN) is perhaps the most versatile and accurate. The descriptor is crucial for properly incorporating symmetry of the system into the ML energy model. The so-called feature variables, which are input to the learning model, must be invariant with respect to symmetry transformations of the electron Hamiltonian. While a large number of descriptors have been proposed for ML-MD methods, the theory of de-

scriptor for classical fields of condensed matter models has yet to be developed.

We discuss common features of the descriptor of classical fields of electronic lattice models, and formulate a general theory by first distinguishing two types of models depending on the absence or presence of an internal symmetry for the classical fields. Several specific approaches to derive a descriptor have been discussed. First, a general descriptor is given by the ordered eigenvalues of the correlation matrix of the neighborhood classical fields, which is similar in spirit to the Weyl or Coulomb matrix descriptor used to characterize the atomic environment. Another approach, also motivated by ML-models for quantum MD simulation, is the generalization of the atom-centered symmetry functions which incorporates the internal symmetry of the classical fields.

The majority of our effort focuses on the group theoretical method which offers systematic and controlled approach to build fundamental invariants of the symmetry group. In this approach, the local classical fields, which form a high-dimensional representation of the site-symmetry point group, is first decomposed into the irreducible representations. Fundamental invariants are given by the bispectrum coefficients of three IRs, which are similar to the scalar or triple product of three vectors. To cope with the issue due to the large number of the over-complete bispectrum invariants, we propose a simplification method based on the concept of the reference IRs. Instead of keeping all the bispectrum coefficients, both the amplitude and the relative “phase” of each IR can be faithfully retained via an inner product with the reference IR.

Finally, we demonstrate the implementation of the various descriptors on well-known lattice models including the Holstein and Jahn-Teller model, and the s-d Hamiltonian for itinerant magnets. The classical field in the former case corresponds to local structural distortions. In particular, the scalar field in the Holstein-type models offers the simplest example to illustrate the working of the lattice descriptor. On the other hand, the s-d model characterized by a vector classical field is used to demonstrate the construction of a descriptor with an independent internal symmetry.

Our work laid the foundation for applying ML methods to multi-scale dynamical modeling in condensed matter systems. Contrary to ML-based MD methods which is an ongoing active research field by itself, the goal here is to model the adiabatic dynamics of classical fields under the influence of quasi-equilibrium electrons. The capability of going beyond empirical methods for large-scale dynamical simulations of such classical fields has numerous implications in condensed matter physics. For example, one particularly important application is the accurate dynamical modeling of topological defects of multi-component classical fields, which are prevalent in condensed matter systems. Notable examples include vortices in superconductivity and skyrmions in itinerant magnetism.

Moreover, complex inhomogeneous electronic states are ubiquitous in correlated electron systems. Not only are these mesoscopic textures of fundamental importance in correlated electron physics, they also play a crucial role in the emergence of novel macroscopic functionalities. For example, complex mixed-phase states are prevalent in colossal magnetoresistant materials and several high- T_c superconductors also exhibit intriguing stripe or checkerboard patterns. Accurate modeling of these complex nanoscale textures is thus of paramount importance in the engineering of these novel material functionalities. However, large-scale simulations of such electronic textures so far are based on empirical or phenomenological

models, mostly because of the extreme difficulty for the multi-scale dynamical modeling of such systems. We believe that the ML force field approach along with the proper descriptor outlined in this work will be an indispensable tool to enable large-scale dynamical simulations of complex patterns in correlated electron materials.

ACKNOWLEDGMENTS

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.

-
- [1] S. V. Kalinin, B. G. Sumpter, and R. K. Archibald, Big-deep-smart data in imaging for guiding materials design, *Nat. Mater.* **14**, 973 (2015).
 - [2] V. Dunjko and H. J. Briegel, Machine learning & artificial intelligence in the quantum domain: a review of recent progress, *Rep. Prog. Phys.* **81**, 074001 (2018).
 - [3] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, *Rev. Mod. Phys.* **91**, 045002 (2019).
 - [4] A. Radovic, M. Williams, D. Rousseau, M. Kagan, D. Bonacorsi, A. Himmel, A. Aurisano, K. Terao, and T. Wongjirad, Machine learning at the energy and intensity frontiers of particle physics, *Nature* **560**, 41 (2018).
 - [5] S. D. Sarma, D.-L. Deng, and L.-M. Duan, Machine learning meets quantum physics, *Phys. Today* **72**, 48 (2019).
 - [6] D. Baron, Machine Learning in Astronomy: a practical overview, *arXiv:1904.07248* (2019).
 - [7] D. Morgan and R. Jacobs, Opportunities and Challenges for Machine Learning in Materials Science, *Annu. Rev. Mater. Res.* **50**, 71 (2020).
 - [8] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, Machine learning for molecular and materials science, *Nature* **559**, 547 (2018).
 - [9] E. Bedolla, L. C. Padierna, and R. Castaneda-Priego, Machine learning for condensed matter physics, *J. Phys.: Condens. Matter* **33**, 053001 (2021).
 - [10] M. Meuwly, Machine Learning for Chemical Reactions, *Chem. Rev.* **121**, 10218 (2021).
 - [11] J. A. Keith, V. Vassilev-Galindo, B. Cheng, S. Chmiela, M. Gastegger, K.-R. Müller, and A. Tkatchenko, Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems, *Chem. Rev.* **121**, 9816 (2021).
 - [12] J. G. Greener, S. M. Kandathil, L. Moffat, and D. T. Jones, A guide to machine learning for biologists, *Nat. Rev. Mol.* **23**, 40 (2022).
 - [13] M. W. Libbrecht and W. S. Noble, Machine learning applications in genetics and genomics, *Nat. Rev. Genetics* **16**, 321 (2015).
 - [14] A. Ourmazd, Science in the age of machine learning, *Nat. Rev. Phys.* **2**, 342 (2020).
 - [15] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning, *Phys. Rev. Lett.* **108**, 058301 (2012).
 - [16] J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, Finding Density Functionals with Machine Learning, *Phys. Rev. Lett.* **108**, 253002 (2012).
 - [17] F. Brockherde, L. Vogt, L. Li, M. E. Tuckerman, K. Burke, and K.-R. Müller, Bypassing the Kohn-Sham equations with machine learning, *Nat. Commun.* **8**, 872 (2017).
 - [18] K. T. Schütt, M. Gastegger, A. Tkatchenko, K.-R. Müller, and R. J. Maurer, Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions, *Nat. Commun.* **10**, 5024 (2019).
 - [19] S. Wang, K. Fan, N. Luo, Y. Cao, F. Wu, C. Zhang, K. A. Heller, and L. You, Massive computational acceleration by using neural networks to emulate mechanism-based biological models, *Nat. Commun.* **10**, 4354 (2019).
 - [20] M. Tsubaki and T. Mizoguchi, 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).
 - [21] M. Bürkle, U. Perera, F. Gimbert, H. Nakamura, M. Kawata, and Y. Asai, Deep-Learning Approach to First-Principles Transport Simulations, *Phys. Rev. Lett.* **126**, 177701 (2021).
 - [22] L. Huang and L. Wang, Accelerated Monte Carlo simulations with restricted Boltzmann machines, *Phys. Rev. B* **95**, 035105 (2017).
 - [23] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, Self-learning Monte Carlo method, *Phys. Rev. B* **95**, 041101(R) (2017).
 - [24] J. Schmidhuber, Deep Learning in Neural Networks: An Overview, *Neural Netw.* **61**, 85 (2015).
 - [25] Y. LeCun, Y. Bengio, and G. Hinton, Deep Learning, *Nature* **521**, 436 (2015).
 - [26] G. Cybenko, Approximations by superpositions of sigmoidal functions, *Math. Control Signals Syst.* **2**, 303 (1989).
 - [27] K. Hornik, M. Stinchcombe, H. White, Multilayer feed-forward networks are universal approximators. *Neural*

- Netw. **2**, 359 (1989).
- [28] A. Barron, Universal approximation bounds for superpositions of a sigmoidal function, *IEEE Trans. Inf. Theory* **39**, 930 (1993).
 - [29] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, *Phys. Rev. Lett.* **98**, 146401 (2007).
 - [30] A. P. Bartók, M. C. Payne, R. Kondor, G. Csányi, Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons, *Phys. Rev. Lett.* **104**, 136403 (2010).
 - [31] Z. Li, J. R. Kermode, and A. De Vita, Molecular Dynamics with On-the-Fly Machine Learning of Quantum-Mechanical Forces, *Phys. Rev. Lett.* **114**, 096405 (2015).
 - [32] V. Botu, R. Batra, J. Chapman, and R. Ramprasad, Machine Learning Force Fields: Construction, Validation, and Outlook, *J. Phys. Chem. C* **121**, 511 (2017).
 - [33] Y. Li, H. Li, F. C. Pickard IV, B. Narayanan, F. G. Sen, M. K. Chan, S. K. Sankaranarayanan, B. R. Brooks, and B. Roux, Machine Learning Force Field Parameters from Ab Initio Data, *J. Chem. Theory Comput.* **13**, 4492 (2017).
 - [34] J. S. Smith, O. Isayev, and A. E. Roitberg, ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost, *Chem. Sci.* **8**, 3192 (2017).
 - [35] L. Zhang, J. Han, H. Wang, R. Car, and Weinan E, Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics, *Phys. Rev. Lett.* **120**, 143001 (2018).
 - [36] J. Behler, Perspective: Machine learning potentials for atomistic simulations, *J. Chem. Phys.* **145**, 170901 (2016).
 - [37] V. L. Deringer, M. A. Caro, and G. Csányi, Machine learning interatomic potentials as emerging tools for materials science, *Adv. Mater.* **31**, 1902765 (2019).
 - [38] F. Noé, A. Tkatchenko, K.-R. Müller, and C. Clementi, Machine Learning for Molecular Simulation, *Annu. Rev. Phys. Chem.* **71**, 361 (2020).
 - [39] R. T. McGibbon, A. G. Taube, A. G. Donchev, K. Siva, F. Hernández, C. Hargus, K.-H. Law, J. L. Klepeis, and D. E. Shaw, Improving the accuracy of Moller-Plesset perturbation theory with neural networks, *J. Chem. Phys.* **147**, 161725 (2017).
 - [40] H. Suwa, J. S. Smith, N. Lubbers, C.D. Batista, G.-W. Chern, and K. Barros, Machine learning for molecular dynamics with strongly correlated electrons, *Phys. Rev. B* **99**, 161107 (2019).
 - [41] T. Mueller, A. Hernandez, and C. Wang, Machine learning for interatomic potential models, *J. Chem. Phys.* **152**, 050902 (2020).
 - [42] E. Dagotto, *Nanoscale phase separation and colossal magnetoresistance* (Berlin, Springer 2002).
 - [43] E. Dagotto, Complexity in strongly correlated electronic systems, *Science* **309**, 257 (2005).
 - [44] A. Moreo, S. Yunoki, and E. Dagotto, Phase separation scenario for manganese oxides and related materials, *Science* **283**, 2034 (1999).
 - [45] N. Mathur and P. Littlewood, Mesoscopic texture in manganites, *Phys. Today* **1**, 25 (2003).
 - [46] S. A. Kivelson, E. Fradkin, and V. J. Emery, Electronic liquid-crystal phases of a doped Mott insulator, *Nature* **393**, 550 (1998).
 - [47] J. M. Tranquada, B. J. Sternlieb, J. D. Axe, Y. Nakamura, and S. Uchida, Evidence for stripe correlations of spins and holes in copper oxide superconductors, *Nature* **375**, 561 (1995).
 - [48] T. Hanaguri, C. Lupien, Y. Kohsaka, D.-H. Lee, M. Azuma, M. Takano, H. Takagi, and J. C. Davis, A ‘checkerboard’ electronic crystal state in lightly hole-doped $\text{Ca}_{2-x}\text{Na}_x\text{CuO}_2\text{Cl}_2$, *Nature* **430**, 1001 (2004).
 - [49] M. Vershinin, S. Misra, S. Ono, Y. Abe, Y. Ando, A. Yazdani, Local ordering in the pseudogap state of the high- T_c superconductor $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$, *Science* **303**, 1995 (2004).
 - [50] S. H. Pan, J. P. O’Neal, R. L. Badzey, C. Chamon, H. Ding, J. R. Engelbrecht, Z. Wang, H. Eisaki, S. Uchida, A. K. Gupta, K.-W. Ng, E. W. Hudson, K. M. Lang, and J. C. Davis, Microscopic electronic inhomogeneity in the high- T_c superconductor $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+x}$, *Nature* **413**, 282 (2001).
 - [51] K. M. Lang, V. Madhavan, J. E. Hoffman, E. W. Hudson, H. Eisaki, S. Uchida, and J. C. Davis, Imaging the granular structure of high- T_c superconductivity in underdoped $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$, *Nature* **415**, 412 (2002).
 - [52] F. Chen, M. Xu, Q. Q. Ge, Y. Zhang, Z. R. Ye, L. X. Yang, J. Jiang, B. P. Xie, R. C. Che, M. Zhang, A. F. Wang, X. H. Chen, D. W. Shen, J. P. Hu, and D. L. Feng, Electronic identification of the parental phases and mesoscopic phase separation of $\text{K}_x\text{Fe}_{2-y}\text{Se}_2$ superconductors, *Phys. Rev. X* **1**, 021020 (2011).
 - [53] C. Zener, Interaction between the d -Shells in the Transition Metals. II. Ferromagnetic Compounds of Manganese with Perovskite Structure, *Phys. Rev.* **82**, 403 (1951).
 - [54] P. W. Anderson and H. Hasegawa, Considerations on Double Exchange, *Phys. Rev.* **100**, 675 (1955).
 - [55] P. -G. de Gennes, Effects of Double Exchange in Magnetic Crystals, *Phys. Rev.* **118**, 141 (1960).
 - [56] S. Yunoki, J. Hu, A. L. Malvezzi, A. Moreo, N. Furukawa, and E. Dagotto, Phase Separation in Electronic Models for Manganites, *Phys. Rev. Lett.* **80**, 845 (1998).
 - [57] S. Zhang and Z. Li, Roles of Nonequilibrium Conduction Electrons on the Magnetization Dynamics of Ferromagnets, *Phys. Rev. Lett.* **93**, 127204 (2004).
 - [58] E. Fradkin, *Field Theories of Condensed Matter Physics* (Cambridge, Cambridge University Press, 2013).
 - [59] G. Grüner, The dynamics of charge-density waves, *Rev. Mod. Phys.* **60**, 1129 (1988).
 - [60] E. Fradkin, S. A. Kivelson, and J. M. Tranquada, Theory of intertwined orders in high temperature superconductors, *Rev. Mod. Phys.* **87**, 457 (2015).
 - [61] G. Grüner, The dynamics of spin-density waves, *Rev. Mod. Phys.* **66**, 1 (1994).
 - [62] G.-W. Chern, K. Barros, Z. Wang, H. Suwa, and C. D. Batista, Semiclassical dynamics of spin density waves, *Phys. Rev. B* **97**, 035120 (2018).
 - [63] R. Iftimie, P. Minary, and M. E. Tuckerman, *Ab initio* molecular dynamics: Concepts, recent developments, and future trends, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6654 (2005).
 - [64] D. Marx and J. Hutter, *Ab initio molecular dynamics: basic theory and advanced methods* (Cambridge University Press, Cambridge, 2009).
 - [65] A. Onuki, *Phase transition Dynamics* (Cambridge, Cambridge Univ. Press 2002).
 - [66] A. Bray, Theory of phase-ordering kinetics, *Adv. Phys.* **43**, 357 (1994).

- [67] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, *Rev. Mod. Phys.* **68**, 13 (1996).
- [68] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Electronic structure calculations with dynamical mean-field theory, *Rev. Mod. Phys.* **78**, 865 (2006).
- [69] S. R. White, Density matrix formulation for quantum renormalization groups, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [70] U. Schollwoeck, The density-matrix renormalization group, *Rev. Mod. Phys.* **77**, 259 (2005).
- [71] D. Ceperley and B. Alder, Quantum Monte Carlo, *Science* **231**, 555 (1986).
- [72] P. Zhang, P. Saha, and G.-W. Chern, Machine learning dynamics of phase separation in correlated electron magnets, *arXiv:2006.04205* (2020).
- [73] P. Zhang and G.-W. Chern, Arrested Phase Separation in Double-Exchange Models: Large-Scale Simulation Enabled by Machine Learning, *Phys. Rev. Lett.* **127**, 146401 (2021).
- [74] S. Zhang, P. Zhang, and G.-W. Chern, Anomalous phase separation and hidden coarsening of superclusters in the Falicov-Kimball model, *arXiv:2105.13304* (2021).
- [75] L. M. Falicov and J. C. Kimball, Simple Model for Semiconductor-Metal Transitions: SmB_6 and Transition-Metal Oxides, *Phys. Rev. Lett.* **22**, 997 (1969).
- [76] J. K. Freericks and V. Zlatić, Exact dynamical mean-field theory of the Falicov-Kimble model, *Rev. Mod. Phys.* **75**, 1333 (2003).
- [77] L. M. Ghiringhelli, J. Vybiral, S. V. Levchenko, C. Draxl, and M. Scheffler, Big Data of Materials Science: Critical Role of the Descriptor, *Phys. Rev. Lett.* **114**, 105503 (2015).
- [78] A. P. Bartók, R. Kondor, and G. Csányi, On representing chemical environments, *Phys. Rev. B* **87**, 184115 (2013).
- [79] L. Himanen, M.O.J. Jäger, E. V. Morooka, F. F. Canova, Y. S. Ranawat, D. Z. Gao, P. Rinke, A. S. Foster, DScribe: Library of descriptors for machine learning in materials science, *Comput. Phys. Commun.* **247**, 106949 (2020).
- [80] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, Bond-orientational order in liquids and glasses, *Phys. Rev. B* **28**, 784 (1983).
- [81] J. Behler, Atom-centered symmetry functions for constructing high-dimensional neural network potentials, *J. Chem. Phys.* **134**, 074106 (2011).
- [82] A. V. Shapeev, Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials, *Multiscale Model. Simul.* **14**, 1153 (2016).
- [83] R. Drautz, Atomic cluster expansion for accurate and transferrable interatomic potentials, *Phys. Rev. B* **99**, 014104 (2019).
- [84] K. Hansen, F. Biegler, R. Ramakrishnan, W. Pronobis, O. A. von Lilienfeld, K.-R. Müller, A. Tkatchenko, Machine learning predictions of molecular properties: Accurate many-body potentials and nonlocality in chemical space, *J. Phys. Chem. Lett.* **6**, 2326 (2015).
- [85] F. Faber, A. Lindmaa, O.A.v. Lilienfeld, R. Armiento, Crystal Structure Representations for Machine Learning Models of Formation Energies, *Int. J. Quantum Chem.* **115** 1094 (2015).
- [86] H. Huo and M. Rupp, Unified Representation of Molecules and Crystals for Machine Learning, *arXiv:1704.06439v3* (2018).
- [87] J. Ma, P. Zhang, Y. Tan, A. W. Ghosh, and G.-W. Chern, Machine learning electron correlation in a disordered medium, *Phys. Rev. B* **99**, 085118 (2019).
- [88] X.-G. Wen, *Quantum Field Theory of Many-body Systems* (Oxford University Press, 2010).
- [89] G. Kotliar and A. E. Ruckenstein, New functional integral approach to strongly correlated Fermi systems: The Gutzwiller approximation as a saddle point, *Phys. Rev. Lett.* **57**, 1362 (1986).
- [90] N. Lanatá, Y. Yao, X. Deng, V. Dobrosavljević, and G. Kotliar, Slave Boson theory of orbital differentiation with crystal field effects: Application to UO_2 , *Phys. Rev. Lett.* **118**, 126401 (2017).
- [91] P. C. Hohenberg and B. I. Halperin, Theory of dynamic critical phenomena, *Rev. Mod. Phys.* **49**, 435 (1977).
- [92] A. Weisse, G. Wellein, A. Alvermann, and H. Fehske, The kernel polynomial method, *Rev. Mod. Phys.* **78**, 275 (2006).
- [93] K. Walter, Density functional and density matrix method scaling linearly with the number of atoms, *Phys. Rev. Lett.* **76**(17), 3168.
- [94] E. Prodan, and K. Walter, Nearsightedness of electronic matter, *Proc. Natl. Acad. Sci. U.S.A.* **102**.33 (2005): 11635-11638.
- [95] S. Mazumdar and S. N. Dixit, Coulomb Effects on One-Dimensional Peierls Instability: The Peierls-Hubbard Model, *Phys. Rev. Lett.* **51**, 292 (1983).
- [96] J. E. Hirsch, Effect of Coulomb Interactions on the Peierls Instability, *Phys. Rev. Lett.* **51**, 296 (1983).
- [97] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Soliton excitations in polyacetylene, *Phys. Rev. B* **22**, 2099 (1980).
- [98] Z. Popovic and S. Satpathy, Cooperative Jahn-Teller Coupling in the Manganites, *Phys. Rev. Lett.* **84**, 1603 (2000).
- [99] C. Sen, G. Alvarez, H. Aliaga, and E. Dagotto, Colossal magnetoresistance observed in Monte Carlo simulations of the one- and two-orbital models for manganites, *Phys. Rev. B* **73**, 224441 (2006).
- [100] G.-W. Chern, K. Barros, C. D. Batista, J. Kress, and G. Kotliar, Mott transition in a metallic liquid – Gutzwiller molecular dynamics simulations, *Phys. Rev. Lett.* **118**, 226401 (2017).
- [101] H. Weyl, *The Classical Groups, Their Invariants and Representations* (Princeton University Press, Princeton, NJ, 1946).
- [102] R. Kakarala, A group theoretic approach to the triple correlation, *IEEE Workshop on higher order statistics*, pp. 28–32 (1993).
- [103] R. Kondor, A novel set of rotationally and translationally invariant features for images based on the non-commutative bispectrum, *arXiv:0701127*. (2007).
- [104] R. Kakarala, Completeness of bispectrum on compact groups, *arXiv:0902.0196* (2009).
- [105] M. Mameresh, *Group Theory and Its Application to Physical Problems* (Dover, New York, 1962).
- [106] T. Inui, Y. Tanabe, Y. Onodera, *Group Theory and Its Applications in Physics* Springer Series in Solid-State Sciences (Berlin, Springer Verlag 1990).
- [107] D. A. Varshalovich, A. N. Moskalev, and V. K. Kher-

- sonskii, *Quantum Theory of Angular Momentum* (World Scientific, Singapore, 1987).
- [108] K. Rykhlinskaya and S. Fritzsche, Generation of Clebsch-Gordan coefficients for the point and double groups, *Comput. Phys. Commun.* **174**, 903 (2006).
 - [109] T. Holstein, Studies of polaron motion: Part I. The molecular-crystal model, *Ann. Phys.* **8**, 325 (1959).
 - [110] J. Zhong and H.-B. Schüttler, Polaronic anharmonicity in the Holstein-Hubbard model, *Phys. Rev. Lett.* **69**, 1600 (1992).
 - [111] A. J. Millis, B. I. Shraiman, R. Mueller, Dynamic Jahn-Teller Effect and Colossal Magnetoresistance in $\text{K}_{1-x}\text{Sr}_x\text{MnO}_3$, *Phys. Rev. Lett.* **77**, 175 (1996).
 - [112] R. Maezono and N. Nagaosa, Jahn-Teller effect and electron correlation in manganites, *Phys. Rev. B* **67**, 064413 (2003).
 - [113] W. F. Brown, Jr., Thermal fluctuations of a single-domain particle, *Phys. Rev.* **130**, 1677 (1963).
 - [114] V. P. Antropov, S. V. Tretyakov, and B. N. Harmon, Spin dynamics in magnets: Quantum effects and numerical simulations, *J. Appl. Phys.* **81**, 3961 (1997).
 - [115] N. A. Clark, Ferromagnetic ferrofluids, *Nature* **504**, 229 (2013).
 - [116] A. Mertelj, D. Lisjak, M. Drofenik, M. Copic, Ferromagnetism in suspensions of magnetic platelets in liquid crystal, *Nature* **504** 237 (2013).
 - [117] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimeshain, L. Antiga, and A. Desmaison, PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 2019* (pp. 8024-8035).
 - [118] V. Nair and G. E. Hinton, in *Proceedings of the 27th International Conference on International Conference on Machine Learning, ICML'10* (Omnipress, USA, 2010) pp. 807-814.
 - [119] J. Barron. Continuously differentiable exponential linear units. *arXiv:1704.07483* (2017).
 - [120] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga and A. Lerer, Automatic differentiation in pytorch, 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. (2017).
 - [121] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pp. 1026-1034. 2015.
 - [122] D. P. Kingma and J. Ba, Adam: A Method for Stochastic Optimization, *arXiv:1412.6980* (2014).