




Multilayer atomic cluster expansion for semilocal interactions

Anton Bochkarev ^{1,*}, Yury Lysogorskiy ¹, Christoph Ortner,² Gábor Csányi,³ and Ralf Drautz ^{1,†}

¹ICAMS, Ruhr-Universität Bochum, Bochum, Germany

²Department of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z2

³Engineering Laboratory, University of Cambridge, Cambridge CB2 1PZ, United Kingdom



(Received 14 March 2022; revised 18 July 2022; accepted 20 July 2022; published 1 November 2022)

Traditionally, interatomic potentials assume local bond formation supplemented by long-range electrostatic interactions when necessary. This ignores intermediate-range multiatom interactions that arise from the relaxation of the electronic structure. Here, we present the multilayer atomic cluster expansion (ml-ACE) that includes collective, semi-local multiatom interactions naturally within its remit. We demonstrate that ml-ACE significantly improves fit accuracy and efficiency compared to a local expansion on selected examples and provide physical intuition to understand this improvement.

DOI: [10.1103/PhysRevResearch.4.L042019](https://doi.org/10.1103/PhysRevResearch.4.L042019)

Recent years have seen tremendous progress in modeling atomic interactions [1–7]. State-of-the-art machine learning potentials interpolate reference data from high-throughput electronic structure calculations with errors on the order of meV/atom [8–10]. Commonly, the energy or other atomic quantities are represented as a function of the local atomic environment enclosed within a cutoff radius centered on each atom. Contributions to the energy from electrostatics cannot be partitioned into local atomic environments and methods to incorporate long-range interactions efficiently have been developed [11], including self-consistent models that mimic the charge transfer of the underlying electronic structure calculations [12,13].

However, electronic structure calculations contain contributions that evade a local chemical description and cannot be captured by long-range electrostatic models either, even if self-consistent charge transfer is included. We introduce the term “semi-local” for interactions that reach significantly beyond the local atomic environment but are not directly associated to long-range charge transfer or directed bond formation. Semi-local interactions are ubiquitous in density functional theory (DFT) and arise from the relaxation of the electronic structure, yet they have not been discussed in the context of machine learning potentials. Examples are the change of interaction in small clusters with size that approach bulk interactions only slowly; intraatomic occupation changes upon hybridization, such as the promotion of electrons in carbon from *s* to *p* states, that alter the carbon bonding characteristics; variation in atomic hybridization in different

atomic environments that induces metal-insulator transitions with consequences for the decay of the density matrix and related bond formation; electronic states along one-dimensional chains that can extend far beyond the local chemical environment.

We build our analysis of semi-local interactions on a local description of the electronic structure and expand the DFT energy with respect to modifications of the density matrix [14–17]

$$E = E_0 + \text{tr}(\mathbf{H}\Delta\rho) + \text{tr}(\mathbf{J}\Delta\rho\Delta\rho) + \dots, \quad (1)$$

with $E_0 = E[\rho_0]$, the Hamiltonian matrix \mathbf{H} , and the density matrix $\rho = \rho_0 + \Delta\rho$. For making contact with interatomic interactions we assume orbitals α, β, \dots , that are localized on atoms i, j, \dots , and density matrix elements $\rho_{i\alpha j\beta} = \langle i\alpha | \hat{\rho} | j\beta \rangle$. The spectrally resolved density matrix $n_{i\alpha j\beta}(E_F) = \frac{d\rho_{i\alpha j\beta}}{dE}(E_F)$ is linked to the Hamiltonian through the generalized moments theorem [7,18]

$$\begin{aligned} \int E^N n_{i\alpha j\beta}(E) dE &= \langle i\alpha | \hat{H}^N | j\beta \rangle \\ &= \sum_{k\gamma l\delta} H_{i\alpha k\gamma} H_{k\gamma l\delta} H_{l\delta \dots} \dots H_{\dots j\beta}, \end{aligned} \quad (2)$$

with $H_{i\alpha j\beta} = \langle i\alpha | \hat{H} | j\beta \rangle$ and where the orbitals are orthonormal and complete. This enables to reconstruct the density matrix from products of Hamiltonian matrix elements of varying order [7,19,20],

$$\rho_{i\alpha j\beta} = \chi_1 H_{i\alpha j\beta} + \chi_2 \sum_{k\gamma} H_{i\alpha k\gamma} H_{k\gamma j\beta} + \dots, \quad (3)$$

where the response functions $\chi_N = \chi_N(E_F)$ depend on the Fermi level. In the tight-binding approximation off-diagonal Hamiltonian matrix elements $H_{i\alpha j\beta}$ with $i \neq j$ depend only weakly on electronic redistribution, while the diagonal elements $H_{i\alpha i\alpha}$ follow the effective one-particle potential and adjust to optimize energy for hybridization and charge transfer [15]. The detailed change $\Delta H_{i\alpha i\alpha}$ is a function of the local

*anton.bochkarev@rub.de

†ralf.drautz@rub.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

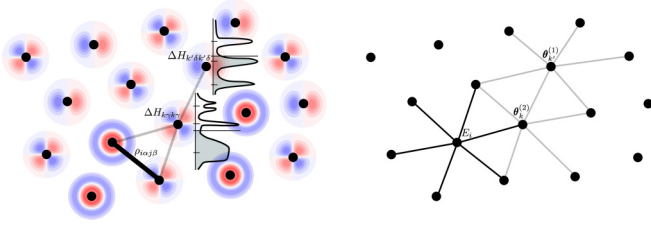


FIG. 1. Semi-local interactions in electronic structure calculations. Density matrix element $\rho_{i\alpha j\beta}$ is modified by $\Delta H_{k\gamma k\gamma}$. Changes in the onsite levels are a function of the local environment that depends on the onsite levels of further distant atoms (left). Abstraction in ml-ACE. The energy E_i of atom i depends on state of neighboring atoms through indicator field $\theta_k^{(2)}$ on atom k that depends on further indicator fields $\theta_{k'}^{(1)}$ (right).

atomic environment of atom i , which may be understood from Eq. (2) and applied to the local density of states $n_{i\alpha i\alpha}(E)$. For example, in metals often the charge transfer is negligible and $\Delta H_{i\alpha i\alpha}$ adjusts to variations in the local density of states for constant electron count.

Therefore, electronic relaxation $\Delta H_{k\gamma k\gamma}$ on atom k affects the density matrix to lowest order as

$$\Delta \rho_{i\alpha j\beta} \propto H_{i\alpha k\gamma} \Delta H_{k\gamma k\gamma} H_{k\gamma j\beta}. \quad (4)$$

The off-diagonal Hamiltonian elements decay rapidly with distance and local neighbors k have the strongest effect on the bond $i-j$, cf. Refs. [7,15,20]. The onsite Hamiltonian matrix elements on atom k , in turn, are determined variationally. This involves the density of states $n_{k\gamma k\gamma}(E)$ that depend on the environment of atom k , which includes atoms k' further distant from the bond $i-j$, and so on. Therefore, the bond $i-j$ is modified by a decaying cascade of modifications on neighbors of neighbors, resulting in semi-local interactions as illustrated in Fig. 1. We show the first moment of the atomic density of states $\mu_i^{(1)} = \sum_{\alpha} H_{i\alpha i\alpha}$ of a linear chain of Cu atoms in Fig. 2.

We present a general framework for integrating semi-local interactions from electronic structure calculations efficiently into machine-learned potentials. Initially, atoms condense information about their local environment into an indicator field.

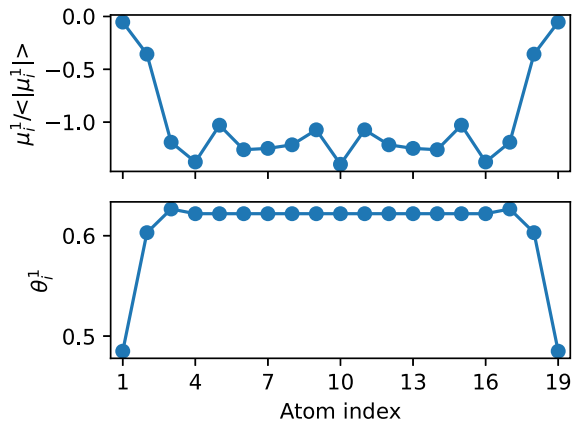


FIG. 2. Normalized first moment of atomic density of states along 19-atom linear Cu chain (upper panel). Indicator of the first layer of a two layer ACE (lower panel).

Next atoms combine indicator fields of their neighbors into their own indicator field, see Fig. 1. This is repeated and with each layer information from atoms at larger distances is incorporated, mimicking the electronic structure relaxation cascade in self-consistent calculations and enabling the description of collective interactions that extend multiple times beyond the local cutoff radius.

The atomic cluster expansion (ACE) [6,21–23] provides a complete and efficient local representation of the energy as a sum over atomic contributions

$$E_i = \mathcal{F}(\varphi_i^{(1)}, \dots, \varphi_i^{(P)}), \quad (5)$$

where \mathcal{F} is a general nonlinear function. Each atomic property $\varphi_i^{(p)}$ is given by a linear expansion

$$\varphi_i^{(p)} = \sum_v \tilde{c}_v^{(p)} A_{iv}, \quad (6)$$

with expansion coefficients $\tilde{c}_v^{(p)}$ and multiatom basis functions A_{iv} .

Other variables than the atomic positions, for example, charges or magnetic moments, may be taken into account [24]. To this end, the state σ is introduced that collects all necessary variables and comprises edges and vertices. We include semi-local interactions by extending the state by an indicator field θ . The field is general and it may comprise several scalar, vectorial or tensorial elements. For an expansion on atom i , the state of a neighboring atom j is defined as $\sigma_{ji} = (z_j, \mathbf{r}_{ji}, \theta_j)$, while the state of atom i is given by $\sigma_{ii} = (z_i, \theta_i)$, where z_i, \mathbf{r}_i denote, respectively, the chemical element and position of atom i . These variables are absorbed into a complete set of single-particle single-bond basis functions $\phi_v(\sigma_{ji})$, respectively, from which the atomic base is computed

$$A_{iv} = \sum_j \phi_v(\sigma_{ji}), \quad (7)$$

and $A_{iv}^{(0)} = \phi_v^{(0)}(\sigma_{ii})$. The products of the atomic base $A_{iv} = A_{iv_0}^{(0)} \prod_n A_{iv_n}$ of various orders form a complete set of basis functions and enable Eq. (6) to represent any atomic function of σ , including scalar, vectorial, or tensorial objects [24].

Often one chooses product basis functions of the form

$$\phi_v(\sigma_{ji}) = e_{\kappa}(z_j) R_{nl}(r_{ji}) Y_l^m(\hat{\mathbf{r}}_{ji}) T_k(\theta_j), \quad (8)$$

where v collects indices and the functions T_k are complete in the space spanned by the indicator fields [24]. Indicator fields $\theta_j^{(n)}$ and expansion coefficients $\tilde{c}_v^{(p,n)}$ then depend on layer n and enter ACE as $\varphi_i^{(p,n)} = \sum_v \tilde{c}_v^{(p,n)} A_{iv}(\theta^{(n)})$. The indicator field in layer $n+1$ is obtained from a nonlinear function as Eq. (5)

$$\theta_i^{(n+1)} = \mathcal{H}(\varphi_i^{(1,n)}, \dots, \varphi_i^{(P,n)}). \quad (9)$$

Layer n_{\max} is the output layer, while the input layer $n=0$ is initialized via $T_k(\theta^{(0)}) = 1$, see Fig. 1. The indicator fields in general carry particular symmetries, most importantly covariance under rotation and inversion, which requires symmetrization [21,24]. We denote the resulting model the *multilayer atomic cluster expansion* (ml-ACE).

By adding layers of indicator fields ml-ACE becomes high-dimensional and sparse basis sets are required for converging

ml-ACE parametrizations. In the present work this is accomplished by physical intuition and hierarchical analysis. The further a layer is away from the final layer, the smaller its impact will be on the energy, i.e., the details are often lost in the distance. This means that the complexity of the indicator field needs to be varied across the layers for best performance and efficient convergence. Because of the layered structure of ml-ACE, gradients may be obtained efficiently and the computational expense for energy and force evaluation is linear or less with the number of layers, as some basis functions can be reused between layers, in particular the spherical harmonics (see Appendix A).

The ml-ACE may be adapted to represent various message passing networks architectures. In fact, general message passing networks may be obtained as special cases of ml-ACE, see Batatia *et al.* [25] and Nigam *et al.* [26]. For example, Thomas *et al.* [27] and the closely related NEQUIP [28] may be cast in the form of particular low-order ACE on each layer. In turn, ml-ACE injects the completeness of the ACE basis into message passing network architectures and can therefore be used to extend oftentimes empirical message passing networks to make them systematically convergent.

In simple tight-binding models the differences of the onsite levels are fixed and their shifts may be characterized by the first moment of the atomic density of states $\mu_i^{(1)} = \sum_{\alpha} H_{i\alpha i\alpha}$ using Eq. (2). For a basic version of ml-ACE we, therefore, take a single indicator variable, which in addition we assume to be rotationally invariant, and that is represented by ACE as Eq. (6). We expect a linear change of the energy for small indicator fields, Eq. (4). To make contact with traditional neural networks, we further choose $\theta_i = \mathcal{H}(\varphi_i) = \tanh(\varphi_i)$ and T_k as Chebyshev polynomials of the first kind.

We demonstrate the performance of the ml-ACE with examples of two distinct cases, namely, small metallic Cu clusters and ten small organic molecules. In a small cluster of N atoms to which one more atom is added, the extra atom can significantly change the interaction between all atoms in the cluster, implying that $(N + 1)$ -body interactions are necessary, and $(N + 2)$ -body interactions when a further atom is added. One expects a slow convergence to bulk interactions only as $N^{-1/3}$, simply as in a compact cluster the number of surface atoms with a significantly modified density of states is proportional to $N^{2/3}$. We employ a dataset of nearly 70 000 small Cu clusters containing 2 to 25 atoms. Energies and forces of the cluster configurations were computed with FHI-AIMS [30,31] using the tight basis set. Most of the clusters were generated by randomly placing atoms inside a sphere and ensuring that the distance between any pair of atoms is not smaller than 80% of the bulk nearest-neighbor distance. In addition, cuts from various crystalline bulk phases were used as well as geometries obtained with empirical potentials. For clusters of up to four atoms the positions were varied systematically to sample the complete configuration space. The distribution of size and energies of the clusters is shown in Fig. 3(a). All the cluster structures were distinctly different, none of the clusters were relaxed, and in particular, we did not use data along molecular dynamics trajectories to avoid correlations. We randomly selected 10% of the dataset for training and the remaining 90% of the clusters for testing.

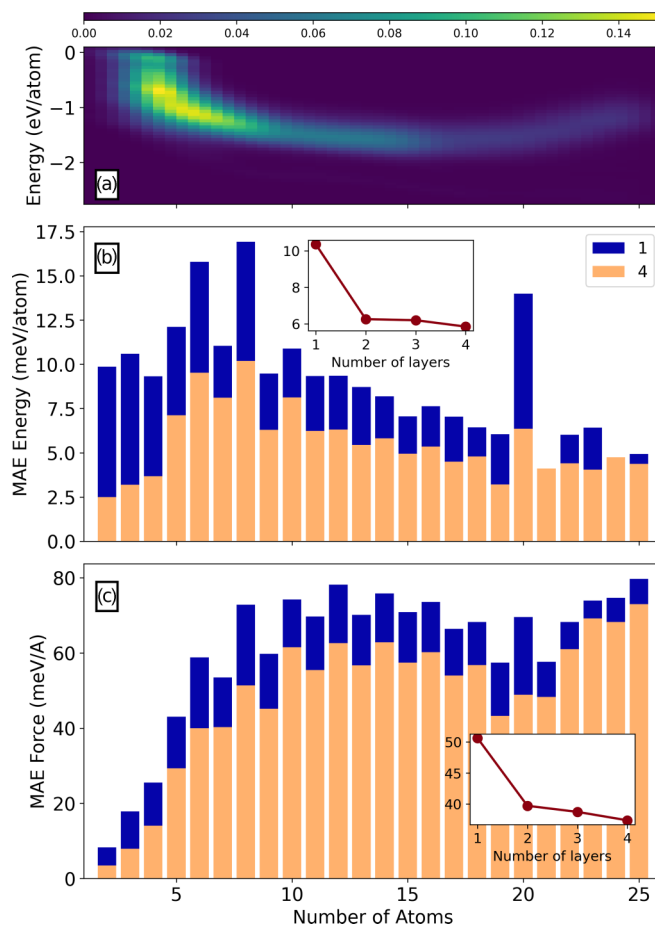


FIG. 3. (a) Distribution in Cu clusters test set. Color shows clusters with given energy and size. (b,c) Test MAE with ml-ACE with various number of layers, where 1 denotes a standard nonlayer ACE. Insets demonstrate convergence of the ml-ACE predictions for entire test set.

Figure 3 shows the convergence of mean average error (MAE) of ml-ACE for a single, scalar invariant indicator variable. A major improvement of nearly a factor of 2 is achieved compared with the standard single layer ACE model. A further increase in the model depth yields smaller yet consistent improvements, in accordance with the expected rapid convergence with the number of layers [20]. Small clusters with up to four atoms and larger clusters with about 15 or more atoms show errors of only a few meV. The description of small clusters with up to four atoms benefits from the accurate description of the many-body potentials up to body-order four enabled by ml-ACE. Larger clusters from about 15 atoms benefit from the accurate description of semi-local interactions. The largest errors are observed for cluster sizes between about five to ten atoms.

We illustrate the correlation between indicator field and electronic structure in Fig. 2, where we show the first moment of the atomic density of states along a 19-atom linear chain of Cu atoms and compare this to the values of the scalar indicator field of a two-layered ACE model. The indicator field picks up the largest deviations at the boundary of the chain. (Linear chains were not part of the training set.)

TABLE I. Performance of ACE models with various number of layers for the ethanol molecule from revised MD17 dataset [29].

# of layers	MAE E, meV	MAE F, meV/Å
1	1.9	9.8
2	1.0	6.2
3	0.8	5.4
4	0.7	4.4

Extended interactions also contribute to bond formation in small molecules. Table I shows the convergence of ml-ACE models for an ethanol molecule from the revised MD17 dataset [29,32,33]. Adding a second layer gives the biggest performance gain, while every further layer results in smaller yet consistent improvements. Thus, for all other molecules we trained four-layer ACE models and the corresponding performance metrics are summarized in Table II (see Appendix A for model details). The resulting models outperform most machine learning potentials [32,34,35]. Here we only compare to models trained on the revised MD17 dataset and Table II shows a comparison to the closely related linear ACE model and the best performing equivariant NEQUIP [28]. For other related models see Refs. [35–38]. Performance improvement of the ml-ACE via semi-local information is evidenced by comparing to linear ACE models. These models contain an order of magnitude more independent basis functions yet they are outperformed by ml-ACE for all molecules. On the other hand, NEQUIP models improve over ml-ACE for almost all molecules with varying performance differences. The largest absolute difference in force MAE (~ 6.5 meV/Å) is observed for the aspirin molecule, while the smallest difference is observed for benzene (0.1 meV/Å) with the ml-ACE model

TABLE II. MAE for energies (E , meV) and forces (F , meV/Å) evaluated on the test set for the four-layer ACE models and compared to the linear ACE models [35] and NEQUIP [28] (we show values corresponding to $l = 3$).

Molecule		ml-ACE (this work)	Linear ACE	NEQUIP
Aspirin	E	4.7	6.1	2.3
	F	14.9	17.9	8.5
Azobenzene	E	2.3	3.6	0.7
	F	7.7	10.9	3.6
Benzene	E	0.02	0.04	0.04
	F	0.2	0.5	0.3
Ethanol	E	0.7	1.2	0.4
	F	4.4	7.3	3.4
Malonaldehyde	E	1.3	1.7	0.8
	F	8.6	11.1	5.2
Naphthalene	E	0.7	0.9	0.2
	F	3.6	5.1	1.2
Paracetamol	E	3.2	4.0	1.4
	F	10.7	12.7	6.9
Salicylic acid	E	1.5	1.8	0.7
	F	7.7	9.3	4.0
Toluene	E	0.8	1.1	0.3
	F	4.3	6.5	1.6
Uracil	E	0.6	1.1	0.4
	F	4.0	6.6	3.2

being slightly more accurate. This discrepancy can be understood from the structures of the molecules. Charge density is nearly uniform in the benzene molecule, and therefore, small variations in the atomic environments during dynamics do not introduce significant charge redistribution and the semi-local information provided by a scalar invariant indicator is sufficient. On the contrary, in the aspirin molecule charge distribution is nonuniform and changes during dynamics [39]. This leads to a variation of s - p hybridization on the atoms and nonuniform onsite level modifications for s and p orbitals. The NEQUIP potential benefits from propagating additional non-scalar equivariant information across the layers which scalar invariant indicators of the current implementation of ml-ACE are unable to capture (see Appendix B for more details).

To conclude, we introduce ml-ACE that efficiently captures semi-local interactions in self-consistent schemes such as DFT. We show that the indicator fields from ml-ACE can be understood from physical and chemical intuition and note that message passing networks may be cast in the form of ml-ACE. We demonstrate ml-ACE numerically for the special case of a single scalar invariant per atom. This invariant ml-ACE allows us to reduce errors for small metallic clusters and molecules to about 50% as compared to single-layer models. Employing only a single scalar invariant is also the main limitation of the numerical implementation presented here. Including several indicator variables with nonscalar rotational characteristics, corresponding to p , d , and so on onsite Hamiltonian elements are the necessary next steps for reducing the remaining errors further. As the focus of our work is on semi-local interactions, our implementation further neglected long-range interactions due to charge transfer. Both contributions, equivariant indicator fields associated to p - and d -valent onsite levels as well as charge transfer need to be taken into account for our next implementation of ml-ACE.

ACKNOWLEDGMENTS

R.D. acknowledges funding through the German Science Foundation (DFG), Projects No. 405621217 and No. 403582885. Y.L. acknowledges funding through the German Science Foundation (DFG), Project No. 405602047. R.D. acknowledges computational resources of research center ZGH, Ruhr-University Bochum, Germany.

APPENDIX A: MODEL AND FITTING DETAILS

Multilayer ACE models are trained via minimizing a loss function of the following form

$$\mathcal{L} = (1 - \kappa) \sum_{n=1}^{N_{\text{struct}}} w_n^{(E)} \left(\frac{E_n^{\text{ACE}} - E_n^{\text{ref}}}{n_{\text{at},n}} \right)^2 + \kappa \sum_{n=1}^{N_{\text{struct}}} \sum_{i=1}^{n_{\text{at},n}} w_{ni}^{(F)} (\mathbf{F}_{ni}^{\text{ACE}} - \mathbf{F}_{ni}^{\text{ref}})^2, \quad (\text{A1})$$

where κ is a trade-off between energy and force contribution, N_{struct} is the number of structures employed in the parametrization, $n_{\text{at},n}$ the number of atoms in structure n , and $w_n^{(E)}$ and $w_{ni}^{(F)}$ are the per-structure and per-atom weights for the energy and force residuals, which were set to 1 for

TABLE III. Details of the potential configurations used in this work.

System	# of layers	cutoff, Å	κ	# functions/element	Time per force call, ms/atom
Cu clusters	1	8	0.1	993	0.41 (0.22)
	2			1421	0.78
	3			1911	1.06
	4			2401	1.36
Ethanol	1	4	0.99	1207	1.09 (0.14)
	2			1363	1.04
	3			2269	1.33
	4			2982	1.67
Aspirin	4	5	0.99	2982	0.96
Naphthalene				2922	0.80
Azobenzene	4	4	0.99	2982	0.92
Malonaldehyde				2982	1.52
Salicylic acid				2982	1.10
Toluene				2922	0.85
Benzene				2922	0.97
Paracetamol				2906	1.03
Uracil				2906	1.44

every structure and normalized by the number of structures and atoms, respectively. For small molecules, we select structures for fitting according to the first split from the revMD17 dataset [29]. Multilayer ACE models were implemented and fitted within the TENSORPOTENTIAL package [23] and Table III summarizes the model hyperparameters used for each system. The table also contains the timings for each model obtained using a 2.50-GHz Intel Xeon Gold 6248 CPU for a single

molecule or the biggest Cu cluster. We note, however, that this implementation is intended for efficient training of the models on large datasets and is not optimized for fast evaluation of a single structure. As a reference, the performance of the PACE code [22] is given in parentheses where applicable for single-layer ACE. A comparable performance improvement can be expected for ml-ACE models when implemented in PACE. Our ml-ACE models are several times faster than those reported in Kovács *et al.* [35] for the same level of accuracy, including the linear ACE models.

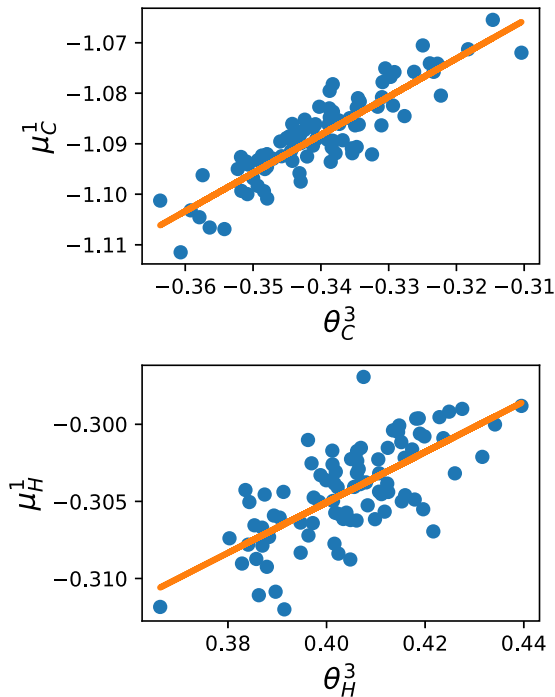


FIG. 4. Correlation between first moment $\mu_i^{(1)}$ and indicator value $\theta_i^{(3)}$ for carbon (top) and hydrogen (bottom) atoms in the benzene molecule. Values are computed for 15 randomly selected molecules from the training set.

APPENDIX B: SMALL MOLECULES

For understanding the difference in performance of the multilayer ACE models for different molecules we con-

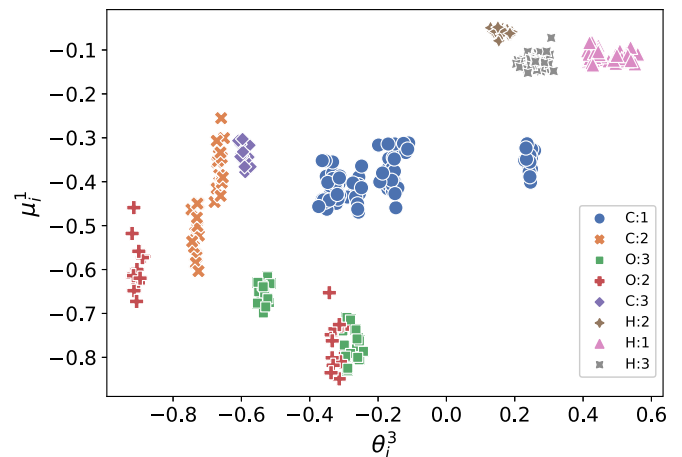


FIG. 5. Correlation plot between first moment $\mu_i^{(1)}$ and indicator value $\theta_i^{(3)}$ for atoms in the aspirin molecule. Values are computed for 15 randomly selected molecules from the training set. Index in the legend denotes the group of an atom, where 1 is benzene ring, 2 the acetyl group, and 3 the carboxyl group.

sider two cases—the best performing model for the benzene molecule and the worst performing model for the aspirin molecule. Figure 4 shows the correlation between values of the scalar indicator $\theta_i^{(3)}$ of the four-layer ACE model and the first moment $\mu_i^{(1)}$ of the atomic DOS for each atom type computed for 15 randomly selected molecules from the training set using FHI-AIMS [30,31] with tight settings. As expected, the electronic distribution in benzene is rather uniform and does not change significantly during dynamics, which is illustrated by the narrow window of values of the first moment. Thus, the invariant scalar indicator is able to

capture these small differences via propagating the semi-local information through the local atomic environments, which is illustrated by the linear correlation between values of $\mu_i^{(1)}$ and $\theta_i^{(3)}$. Figure 5 shows the correlation for 15 randomly selected aspirin molecules, recomputed with FHI-AIMS with tight settings. Data are grouped according to their type and position in a particular functional group. However, unlike Fig. 4, no clear correlation between value of $\theta_i^{(3)}$ and $\mu_i^{(1)}$ is observed. This implies that in this case the scalar indicator provides incomplete information leading to inferior model performance.

-
- [1] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [2] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, Without the Electrons, *Phys. Rev. Lett.* **104**, 136403 (2010).
- [3] A. V. Shapeev, Moment tensor potentials: a class of systematically improvable interatomic potentials, *Multiscale Model. Simul.* **14**, 1153 (2016).
- [4] M. J. Willatt, F. Musil, and M. Ceriotti, Feature optimization for atomistic machine learning yields a data-driven construction of the periodic table of the elements, *Phys. Chem. Chem. Phys.* **20**, 29661 (2018).
- [5] G. P. P. Pun, R. Batra, R. Ramprasad, and Y. Mishin, Physically informed artificial neural networks for atomistic modeling of materials, *Nat. Commun.* **10**, 2339 (2019).
- [6] R. Drautz, Atomic cluster expansion for accurate and transferable interatomic potentials, *Phys. Rev. B* **99**, 014104 (2019).
- [7] R. Drautz and D. G. Pettifor, Valence-dependent analytic bond-order potential for transition metals, *Phys. Rev. B* **74**, 174117 (2006).
- [8] J. Behler and G. Csányi, Machine learning potentials for extended systems: a perspective, *Eur. Phys. J. B* **94**, 142 (2021).
- [9] G. L. W. Hart, T. Mueller, C. Toher, and S. Curtarolo, Machine learning for alloys, *Nat. Rev. Mater.* **6**, 730 (2021).
- [10] F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi, and M. Ceriotti, Physics-inspired structural representations for molecules and materials, *Chem. Rev.* **121**, 9759 (2021).
- [11] A. Grisafi and M. Ceriotti, Incorporating long-range physics in atomic-scale machine learning, *J. Chem. Phys.* **151**, 204105 (2019).
- [12] T. W. Ko, J. A. Finkler, S. Goedecker, and J. Behler, A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer, *Nat. Commun.* **12**, 398 (2021).
- [13] X. Xie, K. A. Persson, and D. W. Small, Incorporating electronic information into machine learning potential energy surfaces via approaching the ground-state electronic energy as a function of atom-based electronic populations, *J. Chem. Theory Comput.* **16**, 4256 (2020).
- [14] W. M. C. Foulkes and R. Haydock, Tight-binding models and density-functional theory, *Phys. Rev. B* **39**, 12520 (1989).
- [15] M. Finnis, *Interatomic Forces in Condensed Matter* (Oxford University Press, New York, 2003).
- [16] R. Drautz and D. G. Pettifor, Valence-dependent analytic bond-order potential for magnetic transition metals, *Phys. Rev. B* **84**, 214114 (2011).
- [17] R. Drautz, T. Hammerschmidt, M. Čák, and D. G. Pettifor, Bond-order potentials: derivation and parameterization for refractory elements, *Modell. Simul. Mater. Sci. Eng.* **23**, 074004 (2015).
- [18] F. Cyrot-Lackmann, On the electronic structure of liquid transitional metals, *Adv. Phys.* **16**, 393 (1967).
- [19] D. G. Pettifor, New Many-Body Potential for the Bond Order, *Phys. Rev. Lett.* **63**, 2480 (1989).
- [20] J. Thomas, H. Chen, and C. Ortner, *Arch Rational Mech Anal* (2022), doi:10.1007/s00205-022-01809-w.
- [21] G. Dusson, M. Bachmayr, G. Csányi, R. Drautz, S. Etter, C. van der Oord, and C. Ortner, Atomic cluster expansion: Completeness, efficiency and stability, [arXiv:1911.03550v3](https://arxiv.org/abs/1911.03550v3).
- [22] Y. Lysogorskiy, C. v. d. Oord, A. Bochkarev, S. Menon, M. Rinaldi, T. Hammerschmidt, M. Mrovec, A. Thompson, G. Csányi, C. Ortner, and R. Drautz, Performant implementation of the atomic cluster expansion (PACE) and application to copper and silicon, *npj Comput Mater* **7**, 97 (2021).
- [23] A. Bochkarev, Y. Lysogorskiy, S. Menon, M. Qamar, M. Mrovec, and R. Drautz, Efficient parametrization of the atomic cluster expansion, *Phys. Rev. Mater.* **6**, 013804 (2022).
- [24] R. Drautz, Atomic cluster expansion of scalar, vectorial, and tensorial properties including magnetism and charge transfer, *Phys. Rev. B* **102**, 024104 (2020).
- [25] I. Batatia, S. Batzner, D. P. Kovács, G. N. C. Simm, R. Drautz, C. Ortner, B. Kozinsky, and G. Csányi, A unified understanding of equivariant interatomic potentials, [arXiv:2205.06643](https://arxiv.org/abs/2205.06643).
- [26] J. Nigam, G. Fraux, and M. Ceriotti, Unified theory of atom-centered representations and graph convolutional machine-learning schemes, *J. Chem. Phys.* **156**, 204115 (2022).
- [27] N. Thomas, T. Smidt, S. Kearnes, L. Yang, L. Li, K. Kohlhoff, and P. Riley, Tensor field networks: Rotation- and translation-equivariant neural networks for 3d point clouds, [arXiv:1802.08219](https://arxiv.org/abs/1802.08219).
- [28] S. Batzner, T. E. Smidt, L. Sun, J. P. Mailoa, M. Kornbluth, N. Molinari, and B. Kozinsky, Se(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials, *Nat. Commun.* **13**, 2453 (2022).
- [29] A. Christensen and O. A. von Lilienfeld, Revised MD17 dataset, *Materials Cloud Archive* 2020.82 (2020), doi:10.24435/materialscloud:wy-kn.

- [30] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Ab initio molecular simulations with numeric atom-centered orbitals, *Comput. Phys. Commun.* **180**, 2175 (2009).
- [31] V. Havu, V. Blum, P. Havu, and M. Scheffler, Efficient o(n) integration for all-electron electronic structure calculation using numerically tabulated basis functions, *J. Comput. Phys.* **228**, 8367 (2009).
- [32] A. S. Christensen and O. A. von Lilienfeld, On the role of gradients for machine learning of molecular energies and forces, *Machine Learning: Science and Technology* **1**, 045018 (2020).
- [33] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, Machine learning of accurate energy-conserving molecular force fields, *Sci. Adv.* **3**, e1603015 (2017).
- [34] A. S. Christensen, L. A. Bratholm, F. A. Faber, and O. Anatole von Lilienfeld, Feh1 revisited: Faster and more accurate quantum machine learning, *J. Chem. Phys.* **152**, 044107 (2020).
- [35] D. P. Kovács, C. V. D. Oord, J. Kucera, A. E. A. Allen, D. J. Cole, C. Ortner, and G. Csányi, Linear atomic cluster expansion force fields for organic molecules: Beyond rmse, *J. Chem. Theory Comput.* **17**, 7696 (2021).
- [36] B. Anderson, T.-S. Hy, and R. Kondor, Cormorant: Covariant molecular neural networks, [arXiv:1906.04015](https://arxiv.org/abs/1906.04015).
- [37] K. T. Schütt, O. T. Unke, and M. Gastegger, Equivariant message passing for the prediction of tensorial properties and molecular spectra, *CoRR* **abs/2102.03150**, [arXiv:2102.03150](https://arxiv.org/abs/2102.03150).
- [38] M. Haghghatlari, J. Li, X. Guan, O. Zhang, A. Das, C. J. Stein, F. Heidar-Zadeh, M. Liu, M. Head-Gordon, L. Bertels, H. Hao, I. Leven, and T. Head-Gordon, Newtonnet: A newtonian message passing network for deep learning of interatomic potentials and forces, [arXiv:2108.02913](https://arxiv.org/abs/2108.02913).
- [39] C. Hauf, A.-A. Hernandez Salvador, M. Holtz, M. Woerner, and T. Elsaesser, Phonon driven charge dynamics in polycrystalline acetylsalicylic acid mapped by ultrafast x-ray diffraction, *Struct. Dyn.* **6**, 014503 (2019).