

Extended Abstract Track

Galois features: Nearly-complete invariants on symmetric matrices

Editors: List of editors' names

Abstract

In this extended abstract, we summarize results from our recent work (Authors, 2024), in which we provide a mathematical formulation for learning functions on symmetric matrices that are invariant with respect to the action of permutations by conjugation. To achieve this, we construct $O(n^2)$ invariant features derived from generators for the field of rational functions on $n \times n$ symmetric matrices that are invariant under joint permutations of rows and columns. We obtain these generators using an argument from Galois theory. We show that these invariant features can separate all distinct orbits of symmetric matrices except for a measure zero set; such features can be used to universally approximate invariant functions on almost all weighted graphs. We empirically demonstrate the feasibility of our approach in a molecular properties regression problem.

Keywords: Graph learning, invariant machine learning, Galois theory, invariant theory

1. Introduction

Many machine learning (ML) applications come equipped with intrinsic symmetry, and incorporating that knowledge into the design of models has often boosted the performance and efficiency. As a result, the success of symmetry-informed models has spurred interest in group invariant and equivariant classes of functions. These developments have been particularly relevant for ML applied to scientific domains, such as molecular chemistry and physics, where known symmetries are imposed by physical law Zhang et al. (2023).

Here we consider the problem of learning a permutation-invariant function on (node and edge) weighted graphs¹. In this problem the underlying invariant theory is hard; indeed, computing a complete set of generating invariants is at least as hard as graph isomorphism (solvable in quasi-polynomial time (Babai, 2016); unknown if solvable in polynomial time).

Motivated by computational scalability, we propose a relaxation of universal approximation. At the cost of throwing out a closed, measure zero set of “bad” point clouds (respectively, weighted graphs), we provide an approach to these learning problems that has a universal approximation guarantee on the remaining “good” weighted graphs.

Our method is to look for computationally tractable invariants that contain the same information as a set of generators for the field of invariant rational functions on the data. It is known that such field generators separate orbits away from a Zariski-closed (and thus Euclidean-closed and measure zero) “bad” subset; thus the proposed invariants do as well. A standard argument based on the Stone-Weierstrass theorem then implies that the proposed invariants are universally approximating away from the “bad” subset. For an n -node weighted graph we extract $O(n^2)$ invariant features. Since such a graph has $O(n^2)$ weights, this is tight information-theoretically.

1. In the the full paper Authors (2024) we also consider the closely related problem of learning invariant functions on d -dimensional point clouds of n points, that is, invariant functions with respect both to permutations and orthogonal transformations using $O(nd)$ features. We refer the reader to the full paper for those results.

Extended Abstract Track

2. Related work

Due to space constraints we refer the reader to the full paper for a thorough discussion on related work. We identify several related lines of research including:

Invariant theory. Describing the class of functions invariant under a group action is historically the project of invariant theory (see e.g.: [Kemper et al. 2022](#)).

Machine learning on weighted graphs. Graph neural networks (GNNs) are a popular ML tool for learning functions on graphs. GNNs are designed to satisfy permutation invariance (for graph-level output), where the choice of enforcing invariance leads to a trade-off between efficiency and expressivity (cf.: [Huang et al. 2024](#)). The expressivity of GNNs has been extensively studied using the graph isomorphism test and comparison with the Weisfeiler-Lehman (WL) algorithms ([Xu et al., 2018](#); [Morris et al., 2019](#)) —a hierarchy of combinatorial graph invariants. To mitigate such issues, more powerful GNNs have been proposed, including higher-order GNNs, subgraph-based GNNs, and spectral GNNs, yet they typically incur higher computation costs due to higher-order tensor operations or additional preprocessing. By using intractable high-order tensors, universality results are established in [Maron et al. \(2019\)](#); [Keriven and Peyré \(2019\)](#). Beyond the Weisfeiler-Lehman hierarchy, expressivity of GNNs can be studied using equivariant polynomials ([Puny et al., 2023](#)) or graph spectral invariants ([Lim et al., 2022](#)).

Enforcing symmetries in machine learning models. More generally, invariant and equivariant machine learning is a very active research area, where researchers incorporate symmetries into the design of machine learning models. There are many different approaches, including orbit averaging, frame averaging, representation theory, group convolutions, weight sharing, canonicalization, and invariant theory (a list of relevant references is provided at the end). The present work considers the invariant theory approach.

3. Invariant functions on symmetric matrices

Suppose $X = (X_{ij})$ is a real symmetric $n \times n$ matrix, and $\mathcal{S}(n)$ is the vector space of all such matrices. The symmetric group \mathfrak{S}_n on n points acts on $\mathcal{S}(n)$ by

$$\pi X := P_\pi X P_\pi^\top, \tag{1}$$

where P_π is the permutation matrix corresponding to the permutation π . This is the action corresponding to node reordering in a graph. Specifically, an undirected weighted graph G on n nodes can be represented by a symmetric $n \times n$ matrix X , known as the adjacency matrix, where X_{ij} describes the edge weight between nodes i and j . (The diagonal entries X_{ii} can be viewed as node weights.) When X is binary and the diagonal is zero, we recover the class of simple undirected graphs.

The order of the nodes is not an intrinsic property of the graph, but rather a choice; therefore, the representation of the graph as a matrix is not unique. The orbit of X under the action of \mathfrak{S}_n consists of all adjacency matrices corresponding to the graph G . In other words, we can identify the space of weighted graphs with $\mathcal{S}(n)/\mathfrak{S}_n$, where $\mathcal{S}(n)/\mathfrak{S}_n$ is the quotient of $\mathcal{S}(n)$ by the group action.

Here we give a set of $O(n^2)$ (specifically, $\binom{n+1}{2} + 1$) invariant functions that are easy to compute in practice and that *almost* universally approximate invariant functions on $\mathcal{S}(n)$.

Extended Abstract Track

For $k = 1, \dots, n$, and for $\ell = 1, \dots, \binom{n}{2}$ let

$$\begin{aligned} f_k^d(X) &= \text{the } k\text{th largest of the numbers } X_{11}, \dots, X_{nn}, \quad (d \text{ for “diagonal”}) \\ f_\ell^o(X) &= \text{the } \ell\text{th largest of the numbers } X_{ij}, \quad 1 \leq i < j \leq n, \quad (o \text{ for “off-diagonal”}) \\ f^*(X) &= \sum_{i \neq j} X_{ii} X_{ij}. \end{aligned} \tag{2}$$

Theorem 1 *There is a closed, \mathfrak{S}_n -invariant, measure zero set $B \subset \mathcal{S}(n)$ (the “bad set”), such any continuous, \mathfrak{S}_n -invariant function $\mathcal{S}(n) \setminus B \rightarrow \mathbb{R}$ defined on the complement of B can be uniformly approximated on any compact subset by a multi-layer perceptron that takes the f 's defined above (2) as inputs. The “bad set” B is defined by the vanishing of some polynomials; in other words, it is an algebraic subvariety of $\mathcal{S}(n)$.*

The argument follows a standard pattern in machine learning of translating an orbit-separation result (Proposition 3.1 below) into a universal approximation result via the Stone-Weierstrass theorem.

Proposition 3.1 *There is a closed, \mathfrak{S}_n -invariant, measure zero set $B \subset \mathcal{S}(n)$ such that any two distinct orbits of \mathfrak{S}_n in the complement of B are separated by some f .*

The proof uses the main idea from (Thiéry, 2000, Theorem 11.2). We provide a proper proof in the full paper. Informally the idea is the following: the f^d 's and f^o 's are invariants of a bigger group acting on $\mathcal{S}(n)$ with the action of permutations that permute diagonal and off-diagonal elements independently. Specifically, the f^d 's and f^o 's are separating invariant functions for the action of the (bigger) direct product group $\Gamma := \mathfrak{S}_n \times \mathfrak{S}_{n(n-1)/2}$ on $\mathcal{S}(n)$.

Consider the field $K := \mathbb{R}(X_{ij: 1 \leq i < j \leq n})$ of rational functions on $\mathcal{S}(n)$. The subfield of K consisting of functions invariant with respect to Γ is $K^\Gamma := \mathbb{R}(f_k^d: 1 \leq k \leq n, f_\ell^o: 1 \leq \ell \leq \binom{n}{2})$.

The field extension $K^\Gamma \subset K$ is a Galois extension. Therefore, the fundamental theorem of Galois theory implies that if we construct a set of invariants that are only fixed by the desired group, then those invariants generate the field of invariant functions; therefore, by Rosenlicht (1956), they uniquely identify all orbits except for those in a proper algebraic subvariety (the “bad set” B of the theorem). Here, the desired group is $\mathfrak{S}_n \subset \Gamma$ carrying the action of permutation by conjugation.

Now we consider the action of \mathfrak{S}_n by (1), in which the X_{ii} 's and the X_{ij} 's are permuted consistently. What we want are the invariants for the action of that subgroup \mathfrak{S}_n of Γ where the entries in a pair $(\sigma, \tau) \in \Gamma = \mathfrak{S}_n \times \mathfrak{S}_{n(n-1)/2}$ are induced by the same underlying permutation of the index set.

This is where f^* enters. For an element (σ, τ) of Γ to leave f^* invariant, it must send terms of the form $X_{ii} X_{ij}$ to other such terms. So the pair (σ, τ) of permutations must preserve the “incidence relation” (i.e., the relation of sharing an index) between the X_{ii} 's and the X_{ij} 's ($i \neq j$). Now σ (acting on the X_{ii} 's) tells us what the underlying permutation of the indices must be, and the fact that the pair (σ, τ) preserves this incidence relation tells us that τ must be induced by the same underlying permutation of the indices. (This point is argued carefully in the proof in Authors (2024).) Thus the only elements of Γ that leave f^* invariant are precisely those belonging to \mathfrak{S}_n .

Extended Abstract Track

4. Experimental results

We consider predicting the properties of the molecules given the molecular structure². The properties are invariant with respect to rotations and reflections of the molecular graph, and permuting the nucleus labels. We use the standard dataset QM7b (Blum and Reymond, 2009; Montavon et al., 2013a). It consists of 7,211 molecules with 14 regression targets. For each molecule, the input feature X is an $n \times n$ symmetric Coulomb Matrix (CM) (Rupp et al., 2012)—a chemical descriptor of molecule. Specifically, X is an explicit invariant function of the 3D coordinates $\{R_i\}_{i=1}^n, R_i \in \mathbb{R}^3$ of the nuclei, and the nuclear charges $\{Z_i\}_{i=1}^n, Z_i \in \mathbb{R}$ (see details in Authors 2024).

In practice rather than computing the f ’s from the input X by sorting, we use the sets $\{X_{ii}\}$ and $\{X_{ij}\}$ as inputs to a DeepSet³: *DeepSet for Conjugation Invariance* (DS-CI)

$$\text{DS-CI}(X) = \text{MLP}_c \left(\text{DeepSet}_1(\{f_k^d(X)\}_{k=1,\dots,n}), \text{DeepSet}_2(\{f_\ell^o(X)\}_{\ell=1,\dots,n(n-1)/2}), \text{MLP}_3(f^*(X)) \right).$$

We also consider an extension dubbed DS-CI+. This extension uses a “binary expansion” (Montavon et al., 2013b, Appendix B) of the invariant features f , which is a common preprocessing step in molecular regression.

We use an 80/10/10 train-validation-test split on QM7b and repeat the experiment over 10 random data splits. We use Mean Absolute Error (MAE) as the loss function and Adam optimizer with initial learning rate 0.01. We train the models for at most 1000 epochs and report the test accuracy at the model checkpoint with the best validation accuracy. Table 1 shows that our lightweight models based on the proposed invariant features achieve competitive performance.

Table 1: Mean absolute error (MAE) on the test set over 10 random data splits (80/10/10 for train/validation/test sets). The results for the Kernel Ridge Regression (KRR) and Deep Tensor Neural Network (DTNN) with the same data split ratio are taken from (Wu et al., 2018, Table 10).

MAE ↓	Atomization PBE0	Excitation ZINDO	Absorption ZINDO	HOMO ZINDO	LUMO ZINDO	1st excitation ZINDO	Ionization ZINDO
KRR	9.3	1.83	0.098	0.369	0.361	0.479	0.408
DS-CI (Ours)	12.849±0.757	1.776±0.069	0.086±0.003	0.401±0.017	0.338±0.048	0.492±0.058	0.422±0.012
DTNN	21.5	1.26	0.074	0.192	0.159	0.296	0.214
DS-CI+ (Ours)	7.650±0.399	1.045±0.030	0.069±0.005	0.172±0.009	0.119±0.005	0.160±0.011	0.189±0.011
MAE ↓	Affinity ZINDO	HOMO KS	LUMO KS	HOMO GW	LUMO GW	Polarizability PBE0	Polarizability SCS
KRR	0.404	0.272	0.239	0.294	0.236	0.225	0.116
DS-CI (Ours)	0.404±0.047	0.302±0.009	0.225±0.01	0.329±0.016	0.213±0.008	0.255±0.015	0.114±0.008
DTNN	0.174	0.155	0.129	0.166	0.139	0.173	0.149
DS-CI+ (Ours)	0.122±0.002	0.169±0.007	0.135±0.007	0.183±0.005	0.139±0.004	0.139±0.005	0.088±0.004

2. Code available at TBA

3. Since the sets we consider for this application are 1-dimensional, other machine learning techniques are also available (for instance, turning sets of scalars into histograms by binning and then using models defined on 1-dimensional probability distributions such as histogram regressions Irpino and Verde (2015); Dias and Brito (2015)).

Extended Abstract Track

References

- Helmer Aslaksen, Shih-Ping Chan, and Tor Gulliksen. Invariants of s_4 and the shape of sets of vectors. *Applicable Algebra in Engineering, Communication and Computing*, 7: 53–57, 1996.
- Benjamin Aslan, Daniel Platt, and David Sheard. Group invariant machine learning by fundamental domain projections. In *NeurIPS Workshop on Symmetry and Geometry in Neural Representations*, pages 181–218. PMLR, 2023.
- Authors. Learning functions on symmetric matrices and point clouds via lightweight invariant features. 2024.
- László Babai. Graph isomorphism in quasipolynomial time. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 684–697, 2016.
- Beatrice Bevilacqua, Fabrizio Frasca, Derek Lim, Balasubramaniam Srinivasan, Chen Cai, Gopinath Balamurugan, Michael M Bronstein, and Haggai Maron. Equivariant subgraph aggregation networks. *arXiv preprint arXiv:2110.02910*, 2021.
- Benjamin Bloem-Reddy, Yee Whye, et al. Probabilistic symmetries and invariant neural networks. *Journal of Machine Learning Research*, 21(90):1–61, 2020.
- L. C. Blum and J.-L. Reymond. 970 million druglike small molecules for virtual screening in the chemical universe database GDB-13. *J. Am. Chem. Soc.*, 131:8732, 2009.
- Ben Blum-Smith and Soledad Villar. Machine learning and invariant theory. *Notices of the American Mathematical Society*, 70(8):1205–1213, 2022.
- Jan Böker, Ron Levie, Ningyuan Huang, Soledad Villar, and Christopher Morris. Fine-grained expressivity of graph neural networks. *Advances in Neural Information Processing Systems*, 36, 2024.
- Giorgos Bouritsas, Fabrizio Frasca, Stefanos P Zafeiriou, and Michael Bronstein. Improving graph neural network expressivity via subgraph isomorphism counting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
- Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.
- Zhengdao Chen, Lei Chen, Soledad Villar, and Joan Bruna. Can graph neural networks count substructures? *Advances in neural information processing systems*, 33:10383–10395, 2020.
- Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pages 2990–2999. PMLR, 2016.
- Taco S Cohen, Mario Geiger, and Maurice Weiler. A general theory of equivariant CNNs on homogeneous spaces. *Advances in neural information processing systems*, 32, 2019.

Extended Abstract Track

- Sónia Dias and Paula Brito. Linear regression model with histogram-valued variables. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 8(2):75–113, 2015.
- Bryn Elesedy and Sheheryar Zaidi. Provably strict generalisation benefit for equivariant models. In *International Conference on Machine Learning*, pages 2959–2969. PMLR, 2021.
- Fabrizio Frasca, Beatrice Bevilacqua, Michael M Bronstein, and Haggai Maron. Understanding and extending subgraph gns by rethinking their symmetries. *arXiv preprint arXiv:2206.11140*, 2022.
- Fabian Fuchs, Daniel Worrall, Volker Fischer, and Max Welling. Se (3)-transformers: 3d roto-translation equivariant attention networks. *Advances in neural information processing systems*, 33:1970–1981, 2020.
- Mario Geiger and Tess Smidt. e3nn: Euclidean neural networks. *arXiv preprint arXiv:2207.09453*, 2022.
- Ben Gripaios, Ward Haddadin, and Christopher G Lester. Lorentz-and permutation-invariants of particles. *Journal of Physics A: Mathematical and Theoretical*, 54(15):155201, 2021.
- Ward Haddadin. Invariant polynomials and machine learning. *arXiv preprint arXiv:2104.12733*, 2021.
- William L Hamilton. *Graph representation learning*. Morgan & Claypool Publishers, 2020.
- Ningyuan Huang, Ron Levie, and Soledad Villar. Approximately equivariant graph networks. *Advances in Neural Information Processing Systems*, 36, 2024.
- Ningyuan Teresa Huang and Soledad Villar. A short tutorial on the Weisfeiler-Lehman test and its variants. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8533–8537. IEEE, 2021.
- Ningyuan Teresa Huang, Soledad Villar, Carey Priebe, Da Zheng, Chengyue Huang, Lin Yang, and Vladimir Braverman. From local to global: Spectral-inspired graph neural networks. In *NeurIPS 2022 Workshop: New Frontiers in Graph Learning*, 2022.
- Antonio Irpino and Rosanna Verde. Linear regression for numeric symbolic variables: a least squares approach based on wasserstein distance. *Advances in Data Analysis and Classification*, 9(1):81–106, 2015.
- Sékou-Oumar Kaba, Arnab Kumar Mondal, Yan Zhang, Yoshua Bengio, and Siamak Ravanbakhsh. Equivariance with learned canonicalization functions. In *International Conference on Machine Learning*, pages 15546–15566. PMLR, 2023.
- Gregor Kemper, Artem Lopatin, and Fabian Reimers. Separating invariants over finite fields. *Journal of Pure and Applied Algebra*, 226(4):106904, 2022.
- Nicolas Keriven and Gabriel Peyré. Universal invariant and equivariant graph neural networks. *Advances in Neural Information Processing Systems*, 32, 2019.

Extended Abstract Track

- Risi Kondor and Shubhendu Trivedi. On the generalization of equivariance and convolution in neural networks to the action of compact groups. In *International conference on machine learning*, pages 2747–2755. PMLR, 2018.
- Risi Kondor, Zhen Lin, and Shubhendu Trivedi. Clebsch–gordan nets: a fully fourier space spherical convolutional neural network. *Advances in Neural Information Processing Systems*, 31, 2018.
- Derek Lim, Joshua David Robinson, Lingxiao Zhao, Tess Smidt, Suvrit Sra, Haggai Maron, and Stefanie Jegelka. Sign and basis invariant networks for spectral graph representation learning. In *The Eleventh International Conference on Learning Representations*, 2022.
- Haggai Maron, Heli Ben-Hamu, Nadav Shamir, and Yaron Lipman. Invariant and equivariant graph networks. In *International Conference on Learning Representations*, 2018.
- Haggai Maron, Ethan Fetaya, Nimrod Segol, and Yaron Lipman. On the universality of invariant networks. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 4363–4371. PMLR, 09–15 Jun 2019. URL <https://proceedings.mlr.press/v97/maron19a.html>.
- Grégoire Montavon, Matthias Rupp, Vivekanand Gobre, Alvaro Vazquez-Mayagoitia, Katja Hansen, Alexandre Tkatchenko, Klaus-Robert Müller, and O Anatole von Lilienfeld. Machine learning of molecular electronic properties in chemical compound space. *New Journal of Physics*, 15(9):095003, 2013a. URL <http://stacks.iop.org/1367-2630/15/i=9/a=095003>.
- Grégoire Montavon, Matthias Rupp, Vivekanand Gobre, Alvaro Vazquez-Mayagoitia, Katja Hansen, Alexandre Tkatchenko, Klaus-Robert Müller, and O Anatole Von Lilienfeld. Machine learning of molecular electronic properties in chemical compound space. *New Journal of Physics*, 15(9):095003, 2013b.
- Christopher Morris, Martin Ritzert, Matthias Fey, William L Hamilton, Jan Eric Lenssen, Gaurav Rattan, and Martin Grohe. Weisfeiler and leman go neural: Higher-order graph neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 4602–4609, 2019.
- Christopher Morris, Yaron Lipman, Haggai Maron, Bastian Rieck, Nils M. Kriege, Martin Grohe, Matthias Fey, and Karsten Borgwardt. Weisfeiler and leman go machine learning: The story so far. *Journal of Machine Learning Research*, 24(333):1–59, 2023. URL <http://jmlr.org/papers/v24/22-0240.html>.
- Peter J Olver. Invariants of finite and discrete group actions via moving frames. *Bulletin of the Iranian Mathematical Society*, 49(2):1–12, 2023.
- Omri Puny, Matan Atzmon, Heli Ben-Hamu, Ishan Misra, Aditya Grover, Edward J Smith, and Yaron Lipman. Frame averaging for invariant and equivariant network design. *arXiv preprint arXiv:2110.03336*, 2021.

Extended Abstract Track

- Omri Puny, Derek Lim, Bobak Kiani, Haggai Maron, and Yaron Lipman. Equivariant polynomials for graph neural networks. In *International Conference on Machine Learning*, pages 28191–28222. PMLR, 2023.
- Siamak Ravanbakhsh, Jeff Schneider, and Barnabas Poczos. Equivariance through parameter-sharing. In *International conference on machine learning*, pages 2892–2901. PMLR, 2017.
- Maxwell Rosenlicht. Some basic theorems on algebraic groups. *American Journal of Mathematics*, 78(2):401–443, 1956.
- Matthias Rupp, Alexandre Tkatchenko, Klaus-Robert Müller, and O Anatole Von Lilienfeld. Fast and accurate modeling of molecular atomization energies with machine learning. *Physical review letters*, 108(5):058301, 2012.
- Erik Henning Thiede, Wenda Zhou, and Risi Kondor. Autobahn: Automorphism-based graph neural nets, 2021. URL <https://arxiv.org/abs/2103.01710>.
- Nicolas M Thiéry. Algebraic invariants of graphs; a study based on computer exploration. *ACM SIGSAM Bulletin*, 34(3):9–20, 2000.
- Soledad Villar, David W Hogg, Kate Storey-Fisher, Weichi Yao, and Ben Blum-Smith. Scalars are universal: Equivariant machine learning, structured like classical physics. *Advances in Neural Information Processing Systems*, 34:28848–28863, 2021.
- Soledad Villar, David W Hogg, Weichi Yao, George A Kevrekidis, and Bernhard Schölkopf. Towards fully covariant machine learning. *Transactions on Machine Learning Research*, 2023a.
- Soledad Villar, Weichi Yao, David W Hogg, Ben Blum-Smith, and Bianca Dumitrascu. Dimensionless machine learning: Imposing exact units equivariance. *Journal of Machine Learning Research*, 24(109):1–32, 2023b.
- Maurice Weiler, Patrick Forré, Erik Verlinde, and Max Welling. Coordinate independent convolutional networks—*isometry and gauge equivariant convolutions on riemannian manifolds*. *arXiv preprint arXiv:2106.06020*, 2021.
- Zhenqin Wu, Bharath Ramsundar, Evan N Feinberg, Joseph Gomes, Caleb Geniesse, Aneesh S Pappu, Karl Leswing, and Vijay Pande. Moleculenet: a benchmark for molecular machine learning. *Chemical science*, 9(2):513–530, 2018.
- Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *International Conference on Learning Representations*, 2018.
- Xuan Zhang, Limei Wang, Jacob Helwig, Youzhi Luo, Cong Fu, Yaochen Xie, Meng Liu, Yuchao Lin, Zhao Xu, Keqiang Yan, et al. Artificial intelligence for science in quantum, atomistic, and continuum systems. *arXiv preprint arXiv:2307.08423*, 2023.