
OrbNet-Spin: Quantum Mechanics Informed Geometric Deep Learning For Open-shell Systems

Beom Seok Kang

Chemistry and Chemical Engineering
California Institute of Technology
bkang@caltech.edu

Mohammadamin Tavakoli

Computing and Mathematical Sciences
California Institute of Technology
amint@caltech.edu

Vignesh C Bhethanabotla

Chemistry and Chemical Engineering
California Institute of Technology
vbhethan@caltech.edu

William A Goddard III

Chemistry and Chemical Engineering
California Institute of Technology
wag@caltech.edu

Anima Anandkumar

Computing and Mathematical Sciences
California Institute of Technology
anima@caltech.edu

Abstract

Modern quantum chemical methods involve a trade-off between accuracy and computational cost/complexity. As an alternative, deep learning methods are used as shortcuts to create accurate predictions with small computational complexity. Such models proved to be highly effective in predicting closed-shell systems (where all electrons are paired). However, little focus has been given to predicting open-shell systems (where there are unpaired electrons), despite their importance in describing species like radicals and reaction intermediates. We introduce OrbNet-Spin, built upon OrbNet-Equi, a geometric- and quantum-aware deep learning model for representing chemical systems at the electronic structure level. OrbNet-Spin incorporates a spin-polarized treatment into the underlying semiempirical quantum mechanics orbital featurization and adjusts the model architecture accordingly while maintaining the geometrical constraints. OrbNet-Spin can accurately describe both closed and open-shell electronic structures. We validate OrbNet-Spin's performance using the QMSpin dataset of open-shell carbenes, achieving a mean absolute error below the chemical accuracy in both singlet and triplet carbenes.

1 Introduction

Understanding molecular properties is vital for applications in drug design, materials optimization, and atmospheric chemistry [24, 13, 5]. Quantum chemical methods like density functional theory (DFT) offer high accuracy but are computationally prohibitive for large molecular systems [3, 28, 15]. In contrast, semiempirical approaches are computationally efficient but sacrifice accuracy [7].

AI-based predictive systems are emerging as high-throughput alternatives for predicting molecular systems of arbitrary size [5, 6]. These models, when trained, are orders of magnitude faster than quantum chemical simulations, while still delivering reasonably accurate predictions. However, the use of these models is often restricted to the chemistry of their training data or their architectural capabilities. For instance, while several models have been developed for closed-shell molecular

systems [10, 8, 17, 18, 26, 31] (where all orbitals are fully paired), there are relatively few successful predictive approaches for open-shell systems. In contrast to closed-shell systems, open-shell systems feature unpaired electrons and accurately characterizing these systems requires accounting for both the spatial and spin degrees of freedom in the calculated orbitals [27]. Open-shell systems play a critical role in various areas of chemistry, such as describing radicals, reaction intermediates, transition metal complexes, and other applications involving unpaired electrons [12, 11].

Cheng et al. [1] developed a machine learning method called molecular-orbital-based machine learning (MOB-ML) that utilizes Fock (**F**), exchange (**K**), and Coulomb (**J**) matrices, and the localized molecular orbitals to represent molecular structures. This representation is used as the input of a Gaussian Process Regression (GPR) to perform molecule-level predictions. MOB-ML utilizes the restricted open-shell Hartree-Fock (ROHF) method. In general, MOB-ML shows accurate predictions with several different datasets, for both closed-shell and open-shell. Unke et al. [32] introduced SpookyNet, a deep learning model that constructs force fields via self-attention. It represents a molecular structure as a combination of atomic numbers, Cartesian coordinates, total charge, and total angular momentum. Although SpookyNet shows good generalization to open- and closed-shell systems, its performance drops when applied to unknown conformations. Finally, Tavakoli et al. [30, 29] developed a chemically informed predictive model specifically for restricted open-shell reactions. Their model is data-driven and is designed to predict radical interactions, rather than just single-molecule representations.

Predictive models achieve greater accuracy by incorporating a system’s physical and geometrical properties [10, 8, 17]. Leveraging invariance or equivariance to spatial transformations (e.g., SE(3)) and physics-based representations, such as the Schrödinger Equation, enhances model performance. OrbNet-Equi is a QM- and geometry-aware deep learning model that maps molecular electronic structure to physical quantities [18]. Qiao et al. [18] represents electronic structures using atomic orbitals features obtained from the low-cost semiempirical method GFN1-xTB. This model has demonstrated strong potential, achieving accuracy comparable to the DFT/B3LYP level of theory for closed-shell molecules, as seen with the QM9 dataset [19].

Despite its potential, the OrbNet-Equi model cannot encode electron spin into its physics or geometry-aware operations. Therefore, its use is limited to closed-shell systems. Here we introduce OrbNet-Spin, built upon OrbNet-Equi, to enable an integrated representation of closed-shell and open-shell molecules. We evaluate the performance of OrbNet-Spin and compare it to similar models, using the QMSpin dataset [22, 23, 2] with quantum chemical information generated using the spGFN1-xTB semiempirical method [16]. Our evaluations show that OrbNet-Spin outperforms other deep learning models. Furthermore, OrbNet-Spin is the only model that achieved errors lower than the chemical accuracy (1 kcal/mol) on the standard test sets of the QMSpin dataset while enabling chemical simulations with significantly higher throughput.

2 Method

OrbNet-Equi is a geometry- and physics-informed model for predicting molecular energies using atomic orbital (AO) features, which represent the electronic structure [18]. AO features -derived from quantum chemical matrices during SCF convergence- form graph-structured data. The graph undergoes SE(3) equivariant message-passing and atom-wise decoding operations which then can be mapped to property space via a property-specific equivariant pooling operation.

While AO features can be derived from any method, using OrbNet to map features from cheaper methods (e.g., semi-empirical) to more accurate, costly methods (e.g., DFT) is advantageous. For instance, with the quantum mechanical information provided from GFN1-xTB [9], OrbNet-Equi successfully predicts the chemical properties of the small organic molecules in the QM9 dataset [19], calculated with DFT/B3LYP level of theory, achieving a mean absolute error (MAE) of 0.081 kcal/mol for the total energy.

OrbNet-Equi featurizes electronic structures of closed-shell systems using the Fock (**F**), density (**P**), the Hamiltonian (**H**), and overlap (**S**) matrices, calculated using the GFN1-xTB method. Therefore, the input tensor is $\mathbf{T} = [\mathbf{F}, \mathbf{P}, \mathbf{H}, \mathbf{S}]$. However, such representation is not sufficient to describe open-shell systems, as the Fock and density matrices split accordingly to spin degrees of freedom.

To develop OrbNet-Spin that can handle open-shell input representations, one needs to take into account different spin orbitals. In closed-shell systems, the net spin (S) is zero. Therefore, all

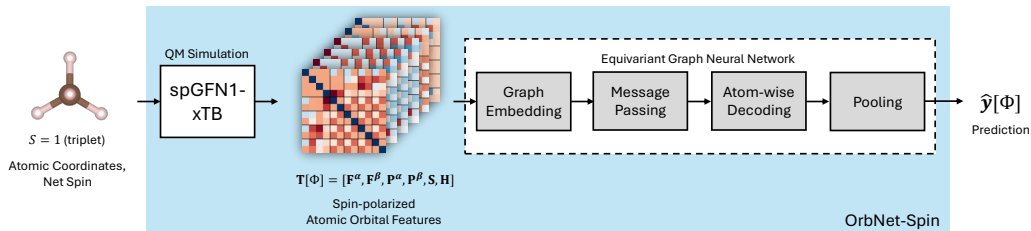


Figure 1: Schematic of the overall process of using the OrbNet-Spin framework to predict physical properties using atomic orbital features generated from quantum mechanics (QM) simulations.

closed-shell molecules are “singlets”, which means the spin multiplicity is 1 ($2S + 1 = 1$). For open-shell molecules with two unpaired electrons, where each electron contributes $+1/2$, the net spin is $S = +1/2 + 1/2 = 1$. They have spin multiplicity of 3 ($2 \times 1 + 1 = 3$), and are called “triplets”.

Each spatial orbital accommodates two spin orbitals, and calculations in the spin basis can be performed using either restricted or unrestricted treatments. In the restricted case, spatial orbitals are constrained to be singly or doubly occupied, while in the unrestricted case, orbitals can relax their spin degrees of freedom based on the initial guess in the SCF method. Figure 2 illustrates the energy levels for different spin and spatial orbitals. In this work, we use the unrestricted SCF procedure for spin-polarized calculations in the GFN1-xTB, known as the spGFN1-xTB method [16]. For closed-shell systems, only the energy levels of spatial orbitals need consideration, allowing the Roothaan and density equations to be formulated with respect to spatial orbitals. In contrast, for unrestricted open-shell systems, different spin orbitals occupy different energy levels, requiring two Roothaan equations for opposite spins, as follows:

$$\mathbf{F}^\alpha \mathbf{C}^\alpha = \mathbf{S} \mathbf{C}^\alpha \epsilon^\alpha, \quad \mathbf{F}^\beta \mathbf{C}^\beta = \mathbf{S} \mathbf{C}^\beta \epsilon^\beta. \quad (1)$$

Here, α and β denote different spins, \mathbf{C} is the matrix of orbital coefficients, and ϵ is the diagonal matrix of orbital energies. Hence, for unrestricted open-shell systems, there are two Fock matrices for the different spins [27], rather than a single Fock matrix from spatial orbital interactions. Similarly, two density matrices are obtained for the different spins, given by:

$$P_{\mu\nu}^\alpha = \sum_a^{N^\alpha} C_{\mu a}^\alpha (C_{\nu a}^\alpha)^*, \quad P_{\mu\nu}^\beta = \sum_a^{N^\beta} C_{\mu a}^\beta (C_{\nu a}^\beta)^*, \quad (2)$$

where P and C are the individual elements of the density matrix and the orbital coefficients matrix, respectively, the subscripts represent the position of elements, and N^α and N^β are the number of electrons with α spin and β spin, respectively [27].

The calculations of \mathbf{F} and \mathbf{P} for different spins also apply to closed-shell molecules, where matrices for different spins are identical: $\mathbf{F}^\alpha = \mathbf{F}^\beta = \mathbf{F}$ and $\mathbf{P}^\alpha = \mathbf{P}^\beta = \frac{1}{2}\mathbf{P}$. Thus, we use the AO features $\mathbf{T} = [\mathbf{F}^\alpha, \mathbf{F}^\beta, \mathbf{P}^\alpha, \mathbf{P}^\beta, \mathbf{S}, \mathbf{H}]$. This representation allows us to consistently describe both closed-shell and open-shell systems, with distinctions made by the different spin matrices. This featurization generalizes the OrbNet-Equi framework to be able to process open-shell molecules, as well as closed-shell molecules. We name this framework OrbNet-Spin. Note that technically any SCF method that solves the Roothaan equation can be used in this framework, but we use spGFN1-xTB method [16] to generate AO features inexpensively. The method is implemented using the tblite package [4]. Furthermore, for training, we apply the delta-learning strategy, which is to learn the difference between the label energies and the energies calculated using the cheap method, spGFN1-xTB, in this case, [20]. Since for creating AO features we also get the energies estimated using spGFN1-xTB, the calculation of the energy differences requires no additional cost.

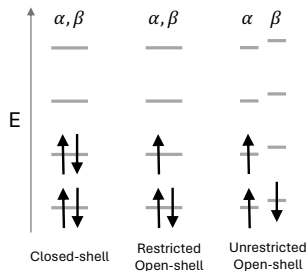


Figure 2: Illustration of energy levels (E) of orbitals and electrons occupying the orbitals. α and β represent different spins.

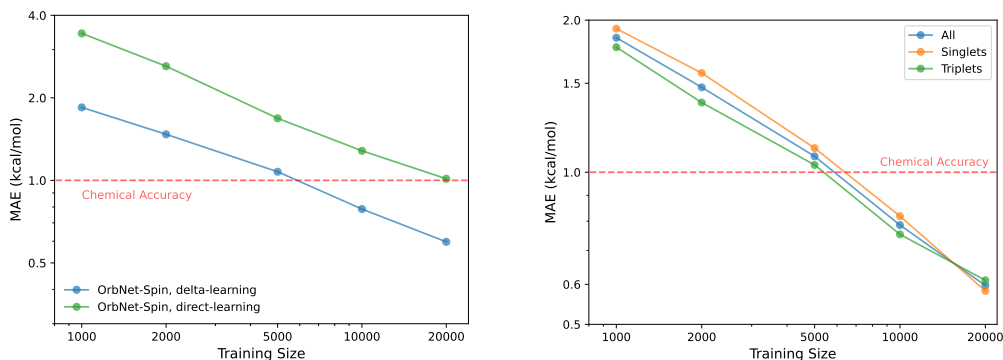


Figure 3: Learning curves of delta-learning and direct-learning (left) and delta-learning for singlets, triplets, and all (right). For training size N , the training set contains $N/2$ singlets and $N/2$ triplets.

3 Results

We train OrbNet-Spin using the QMSpin dataset [23, 22], which includes 4,981 singlet-optimized and 7,835 triplet-optimized carbene geometries, providing 25,632 energy points (12,816 singlet energies and 12,816 triplet energies). The geometries are optimized via restricted open-shell B3LYP/def2-TZVP, and the energies are obtained at MRCISD+Q-F12/cc-pVDZ-F12, which is orders of magnitude more computationally expensive than DFT methods [1].

Table 1: Mean absolute error (MAE) in kcal/mol on total molecular energy when trained with 10,000 molecules in total on different training sets of the QMSpin dataset.

		OrbNet-Spin (Singlets Only)	OrbNet-Spin (Singlets + Triplets)
Test Data	All	4.604	0.786
	Singlets	0.662	0.819
	Triplets	8.545	0.753

OrbNet-Spin is evaluated on singlet and triplet systems through two training configurations: (1) a model trained with 10,000 singlets (singlets-only) and validated on 500 singlets, and (2) a model trained with 5,000 singlets and 5,000 triplets (singlets + triplets) and validated on 500 singlets and 500 triplets. Training, validation, and test sets are selected randomly, ensuring no shared geometries. Both models are tested on three sets: (1) a combined set of 2,316 singlets and triplets (all), (2) 2,316 singlets, and (3) 2,316 triplets.

As Table 1 shows, OrbNet-Spin can learn from the combined electronic structures of singlets and triplets. Specifically, the MAE is significantly improved when using both singlets and triplets as training data. Additionally, the singlet + triplet model performs on par with the singlets-only model

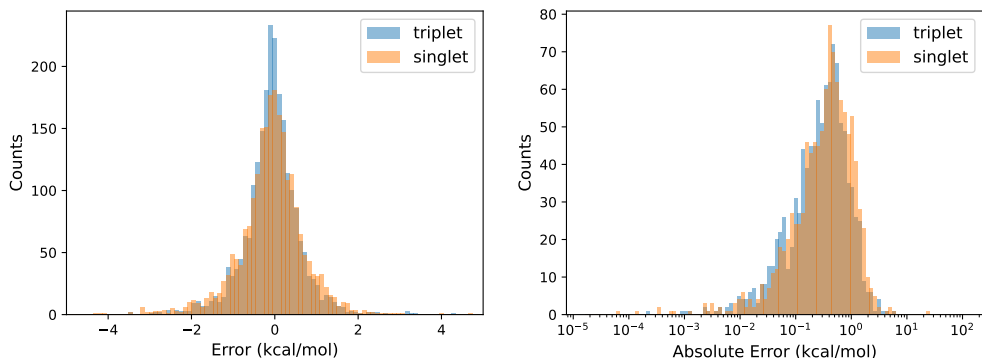


Figure 4: Error distribution histogram (left) and absolute error log distribution histogram (right) for OrbNet-Spin predictions on the QMSpin test set, trained with 10,000 singlets and 10,000 triplets.

when testing on singlets data. This indicates that OrbNet-Spin can interpret the spin-polarized AO features to accurately predict the energies of different spin states. Furthermore, displayed in Figure 3, delta-learning consistently exhibits lower errors at all training sizes. For the QMSpin dataset training, the delta-learning surpasses the chemical accuracy with a 10,000 training size, whereas direct-learning requires $> 20,000$ training size. In addition, we observe small differences between the errors of singlets and triplets for different training sizes, implying that OrbNet-Spin can describe the different spin states equally well. We also present the error distribution histograms in Figure 4—the prediction error (MAE) of OrbNet-Spin centers at zero, showing a near-Gaussian error distribution.

Table 2: Mean absolute error (MAE) in kcal/mol on total molecular energy when trained with 20,000 molecules (10,000 singlets and 10,000 triplets) for different models.

		SpookyNet	MOB-ML	OrbNet-Spin (ours)
Test Data	All	1.568 [32]	0.872 [1]	0.597
	Singlets	-	0.623 [1]	0.582
	Triplets	-	1.120 [1]	0.611

Next, we compare the performance of the OrbNet-Spin to the related work discussed above. We use the same training size as mentioned in Cheng et al. [1] and Unke et al. [32] which is 10,000 singlets and 10,000 triplets. While SpookyNet trained a single model using the combined set (without using the QM features), MOB-ML trained two separate models for singlets and triplets.

As it is shown in Table 2, OrbNet-Spin demonstrates the best performance, solely achieving mean absolute errors (MAEs) below the threshold of chemical accuracy (1 kcal/mol) for both singlet and triplet states. This is particularly impressive given that the MRCISD+Q-F12/cc-pVDZ-F12 level of theory is used for the dataset, which is significantly more accurate than DFT. More importantly, MOB-ML uses restricted open-shell Hartree-Fock (ROHF) with aug-cc-pVTZ basis to generate its features [1]. On the other hand, we use the cheaper method of GFN1-xTB. This demonstrates that our model not only provides much higher throughput compared to MOB-ML but also successfully maps less accurate input representations to the MRCISD+Q-F12/cc-pVDZ-F12 level of theory. Overall, OrbNet-Spin stands out as the only model that can describe molecules of different spin states with a single model, while maintaining the error below the chemical accuracy.

Lastly, our method enables energy evaluation orders of magnitude faster than traditional approaches. It consists of two steps: (1) AO feature creation, which takes milliseconds on a regular CPU, and (2) neural network inference, with an average time of 0.367 ms per data point (batch size of 64) and 8.458 ms without batching using an RTX 4090. Thus, OrbNet-Spin’s inference time ranges from a few milliseconds to ~ 10 ms per data point, significantly outperforming MRCISD+Q-F12 (hours to days) per convergence [14, 25, 21].

4 Conclusion

We introduced OrbNet-Spin, an extension of OrbNet-Equi with spin-polarized AO featurization to handle both open-shell and closed-shell systems. This method differentiates molecules with varying total spins using distinct Fock and density matrices. By training OrbNet-Spin using the QMSpin dataset, we demonstrate its ability to describe singlet and triplet molecules, achieving mean absolute errors below chemical accuracy and outperforming SpookyNet and MOB-ML. We believe OrbNet-Spin can be extended to other open-shell systems, such as radicals and reaction intermediates. Additionally, it could be utilized for high-throughput prediction of various other physical properties. These promising directions are left for future work.

Finally, we stress the one critical limitation of our model over other atomistic ML models, which is that whether the model can process data depends also on the convergence of SCF using the semiempirical method. Atomistic ML models take atomic numbers and coordinates as inputs and hence can process any data. However, our neural network functions with quantum chemical matrices obtained from QM simulations, but if the SCF procedure fails to converge, the features cannot be generated, making it impossible to acquire the prediction.

References

- [1] Lixue Cheng, Jiace Sun, J. Emiliano Deustua, Vignesh C. Bhethanabotla, and Thomas F. Miller. Molecular-orbital-based machine learning for open-shell and multi-reference systems with kernel addition Gaussian process regression. *The Journal of Chemical Physics*, 157(15):154105, October 2022. ISSN 0021-9606, 1089-7690. doi: 10.1063/5.0110886.
- [2] Lixue Cheng, Jiace Sun, J. Emiliano Deustua, Vignesh C. Bhethanabotla, and Thomas F. Miller III. Criegee, h10 chain, small radicals, water bond dissociation, and qmspin energy datasets with mob features for mob-ml(ka-gpr), 2022. URL <https://data.caltech.edu/records/20200>.
- [3] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. Challenges for density functional theory. *Chemical reviews*, 112(1):289–320, 2012.
- [4] Sebastian Ehlert. awvwgk/tblite: Light-weight tight-binding framework. <https://github.com/awvwgk/tblite>, 2024.
- [5] Nikita Fedik, Roman Zubatyuk, Maksim Kulichenko, Nicholas Lubbers, Justin S Smith, Benjamin Nebgen, Richard Messerly, Ying Wai Li, Alexander I Boldyrev, Kipton Barros, et al. Extending machine learning beyond interatomic potentials for predicting molecular properties. *Nature Reviews Chemistry*, 6(9):653–672, 2022.
- [6] David Fooshee, Aaron Mood, Eugene Gutman, Mohammadamin Tavakoli, Gregor Urban, Frances Liu, Nancy Huynh, David Van Vranken, and Pierre Baldi. Deep learning for chemical reaction prediction. *Molecular Systems Design & Engineering*, 3(3):442–452, 2018.
- [7] Gilles Frison and Gilles Ohanessian. A comparative study of semiempirical, ab initio, and dft methods in evaluating metal–ligand bond strength, proton affinity, and interactions between first and second shell ligands in zn-biomimetic complexes. *Journal of computational chemistry*, 29(3):416–433, 2008.
- [8] Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs. In *International Conference on Learning Representations*, 2020. URL <https://openreview.net/forum?id=B1eWbxStPH>.
- [9] Stefan Grimme, Christoph Bannwarth, and Philip Shushkov. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-block elements ($z = 1-86$). *Journal of Chemical Theory and Computation*, 13(5):1989–2009, 2017. doi: 10.1021/acs.jctc.7b00118. URL <https://doi.org/10.1021/acs.jctc.7b00118>. PMID: 28418654.
- [10] Peter Bjørn Jørgensen, Karsten Wedel Jacobsen, and Mikkel N. Schmidt. Neural message passing with edge updates for predicting properties of molecules and materials, 2018. URL <https://arxiv.org/abs/1806.03146>.
- [11] Wolfgang Kaim and Alexa Paretzki. Interacting metal and ligand based open shell systems: Challenges for experiment and theory. *Coordination Chemistry Reviews*, 344:345–354, 2017. ISSN 0010-8545. doi: <https://doi.org/10.1016/j.ccr.2016.12.008>. URL <https://www.sciencedirect.com/science/article/pii/S0010854516303903>. Chemical Bonding: "State of the Art".
- [12] Anna I Krylov. The quantum chemistry of open-shell species. *Reviews in Computational Chemistry*, 30:151–224, 2017.
- [13] Laure Mamy, Dominique Patureau, Enrique Barriuso, Carole Bedos, Fabienne Bessac, Xavier Louchart, Fabrice Martin-Laurent, Cecile Miege, and Pierre Benoit. Prediction of the fate of organic compounds in the environment from their molecular properties: a review. *Critical reviews in environmental science and technology*, 45(12):1277–1377, 2015.
- [14] David Mester and Mihaly Kallay. A simple range-separated double-hybrid density functional theory for excited states. *Journal of Chemical Theory and Computation*, 17(2):927–942, 2021.

- [15] Aaron Mood, Mohammadamin Tavakoli, Eugene Gutman, Dora Kadish, Pierre Baldi, and David L Van Vranken. Methyl anion affinities of the canonical organic functional groups. *The Journal of organic chemistry*, 85(6):4096–4102, 2020.
- [16] Hagen Neugebauer, Benedikt Bädorf, Sebastian Ehlert, Andreas Hansen, and Stefan Grimme. High-throughput screening of spin states for transition metal complexes with spin-polarized extended tight-binding methods. *Journal of Computational Chemistry*, 44(27):2120–2129, July 2023. ISSN 1096-987X. doi: 10.1002/jcc.27185. URL <http://dx.doi.org/10.1002/jcc.27185>.
- [17] Zhuoran Qiao, Matthew Welborn, Animashree Anandkumar, Frederick R. Manby, and Thomas F. Miller. OrbNet: Deep learning for quantum chemistry using symmetry-adapted atomic-orbital features. *The Journal of Chemical Physics*, 153(12):124111, September 2020. ISSN 0021-9606, 1089-7690. doi: 10.1063/5.0021955.
- [18] Zhuoran Qiao, Anders S. Christensen, Matthew Welborn, Frederick R. Manby, Anima Anandkumar, and Thomas F. Miller. Informing geometric deep learning with electronic interactions to accelerate quantum chemistry. *Proceedings of the National Academy of Sciences*, 119(31):e2205221119, August 2022. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.2205221119.
- [19] Raghunathan Ramakrishnan, Pavlo O. Dral, Matthias Rupp, and O. Anatole Von Lilienfeld. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific Data*, 1(1):140022, August 2014. ISSN 2052-4463. doi: 10.1038/sdata.2014.22.
- [20] Raghunathan Ramakrishnan, Pavlo O. Dral, Matthias Rupp, and O. Anatole von Lilienfeld. Big data meets quantum chemistry approximations: The -machine learning approach. *Journal of Chemical Theory and Computation*, 11(5):2087–2096, 2015. doi: 10.1021/acs.jctc.5b00099. URL <https://doi.org/10.1021/acs.jctc.5b00099>. PMID: 26574412.
- [21] Masaaki Saitow, Yuki Kurashige, and Takeshi Yanai. Multireference configuration interaction theory using cumulant reconstruction with internal contraction of density matrix renormalization group wave function. *The Journal of chemical physics*, 139(4), 2013.
- [22] Max Schwilk, Diana N. Tahchieva, and O. Anatole von Lilienfeld. Large yet bounded: Spin gap ranges in carbenes, 2020. URL <https://arxiv.org/abs/2004.10600>.
- [23] Schwilk, Max, Tahchieva, Diana N., and von Lilienfeld, O. Anatole. The qmspin data set: Several thousand carbene singlet and triplet state structures and vertical spin gaps computed at mrcisd+q-f12/cc-pvdz-f12 level of theory, 2020. URL <https://archive.materialscloud.org/record/2020.0051/v1>.
- [24] Jie Shen and Christos A Nicolaou. Molecular property prediction: recent trends in the era of artificial intelligence. *Drug Discovery Today: Technologies*, 32:29–36, 2019.
- [25] Toru Shiozaki, Gerald Knizia, and Hans-Joachim Werner. Explicitly correlated multireference configuration interaction: Mrci-f12. *The Journal of chemical physics*, 134(3), 2011.
- [26] Guillem Simeon and Gianni de Fabritiis. TensorNet: Cartesian tensor representations for efficient learning of molecular potentials, 2023. URL <https://arxiv.org/abs/2306.06482>.
- [27] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry - Introduction to Advanced Electronic Structure Theory*. Dover Publications, 1989.
- [28] Mohammadamin Tavakoli, Aaron Mood, David Van Vranken, and Pierre Baldi. Quantum mechanics and machine learning synergies: graph attention neural networks to predict chemical reactivity. *Journal of Chemical Information and Modeling*, 62(9):2121–2132, 2022.
- [29] Mohammadamin Tavakoli, Yin Ting T Chiu, Pierre Baldi, Ann Marie Carlton, and David Van Vranken. Rmechdb: A public database of elementary radical reaction steps. *Journal of chemical information and modeling*, 63(4):1114–1123, 2023.
- [30] Mohammadamin Tavakoli, Pierre Baldi, Ann Marie Carlton, Yin Ting Chiu, Alexander Shmakov, and David Van Vranken. Ai for interpretable chemistry: predicting radical mechanistic pathways via contrastive learning. *Advances in Neural Information Processing Systems*, 36, 2024.

- [31] Mohammadamin Tavakoli, Ryan J Miller, Mirana Claire Angel, Michael A Pfeiffer, Eugene S Gutman, Aaron D Mood, David Van Vranken, and Pierre Baldi. Pmechdb: A public database of elementary polar reaction steps. *Journal of Chemical Information and Modeling*, 64(6): 1975–1983, 2024.
- [32] Oliver T. Unke, Stefan Chmiela, Michael Gastegger, Kristof T. Schütt, Huziel E. Saucedo, and Klaus-Robert Müller. SpookyNet: Learning force fields with electronic degrees of freedom and nonlocal effects. *Nature Communications*, 12(1):7273, December 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-27504-0. URL <https://www.nature.com/articles/s41467-021-27504-0>.