Deep Learning for Molecular Orbitals

Daniel S. King,[†] Daniel Grzenda,[‡] Ray Zhu,[¶] Nathaniel Hudson,[‡,§] Ian Foster,[‡,§] and Laura Gagliardi[*,‖]

[†]*Department of Chemistry, University of Chicago, Chicago IL 60637*

[‡]*Department of Computer Science, University of Chicago, Chicago IL 60637*

[¶]*Pritzker School of Molecular Engineering, Chicago IL 60637*

[§]*Data Science and Learning Division, Argonne National Laboratory, Lemont IL 60439*

[‖]*Department of Chemistry, Pritzker School of Molecular Engineering, James Franck Institute, Chicago Center for Theoretical Chemistry, University of Chicago, Chicago IL 60637*

E-mail: lgagliardi@uchicago.edu

Abstract

The advancement of deep learning in chemistry has resulted in state-of-the-art models that incorporate an increasing number of concepts from standard quantum chemistry, such as orbitals and Hamiltonians. With an eye towards the future development of these deep learning approaches, we present here what we believe to be the first work focused on assigning labels to orbitals, namely energies and characterizations, given the real-space descriptions of these orbitals from standard electronic structure theories such as Hartree-Fock. In addition to providing a foundation for future development, we expect these models to have immediate impact in automatizing and interpreting the results of advanced electronic structure approaches for chemical reactivity and spectroscopy.

1 Introduction

The development of deep learning techniques in quantum chemistry is remarkable and has been the subject of various reviews.[1–5] However, despite the novel nature of these techniques, accuracy has often increased in these methods through the development of "physics informed" architectures[6] that integrate ideas from classical quantum chemistry. For example, Aldossary and coworkers outline four "generations" of machine-learned potentials[2] with later generations (e.g., ANI-1[7] and PhysNet[8]) incorporating the concept of partial charges from standard force field development to capture long-range effects. A similar progression has taken place in the application of deep learning to electronic structure theory, with state-of-the-art approaches such as SchnOrb,[9] OrbNet,[10,11] DeepH,[12] and the recent model of Ceriotti and coworkers[13] making explicit reference to the concepts of orbitals and Hamiltonians, being reminiscent of semiempirical models such as AM1[14] and PM3.[15]

Despite this work, relatively little focus has been given to predicting properties (e.g., energies or character) of orbitals themselves. It is evident why this might be the case: orbital properties are (with some exceptions[16]) vague, unobservable properties that only serve as intermediates in the calculation of the energy and other chemical features. Nevertheless, in the same way that much effort has gone into the development of atomic charges[7,8,17]—also inherently unobservable quantities that only serve as useful intermediates for calculation—the properties of orbitals themselves may also be worth considering. We outline several motivations.

The first and most pressing motivation is that much as partial charges are necessary for capturing long-range effects, orbitals are needed to calculate the kinetic energy. Indeed, although Kohn-Sham density functional theory appears at face value to do away with the molecular orbital picture entirely by means of the Hohenberg-Kohn theorem,[18] it still must evaluate the kinetic energy in an orbital-driven fashion identical to Hartree-Fock.[19] This fact remains the principal reason for the $\mathcal{O}(N^3)$ scaling of (local) KS-DFT, and developing

kinetic energy functionals to avoid this cost, enabling an "orbital-free" DFT, remains an active area of research.[20] Nevertheless, a straightforward way to avoid this scaling in KS-DFT would simply be to predict the Kohn-Sham orbitals themselves without calculation. Building models to understand orbital properties such as the energy stands as an important intermediary goal in this direction.

Secondly, orbital features—although intermediate quantities—are generally of some use. The foremost example is that of the HOMO and LUMO energies, which correspond to the ionization potential and electron affinity of a system by means of Koopmans' theorem.[21,22] Indeed, it is this area which has seen the most work in predicting orbital characteristics,[23–27] with physical quantities such as ionization potentials and excitation energies being the focus of these approaches. Additionally, orbital properties (e.g., shape, size, and energy) are critical in how chemists understand and discuss chemistry: molecular orbitals serve as a fundamental and transferable concept for describing the behavior of molecules, and if models are to ever understand chemistry they will probably need to understand molecular orbitals themselves. Moreover, understanding the shape of molecular orbitals has practical purposes for calculations: methods such as the complete active space self-consistent field theory (CASSCF) method depend on an "active space" of orbitals and electrons which rely on human-annotated labels (e.g., π, π^*, bonding, antibonding).[28] Automating these human-annotated labels has the potential to accelerate greatly the calculation and interpretation of these highly accurate methods.

Given these motivations, we present what we believe to be the first work focused on assigning labels to orbitals themselves given input from *ab initio* calculations. Specifically, we aim to develop functions:

$$f(\phi) \rightarrow l_\phi, \tag{1}$$

where l_ϕ is a label corresponding to either an orbital energy (e.g., from KS-DFT) or a

human-annotated description (e.g., bonding or antibonding). We develop new datasets for both of these tasks, the QMOrb dataset, consisting of full Molden files[29] for the entire QM9 dataset,[30] and the TMOrb dataset, consisting of generated metal-ligand bonding and antibonding orbitals for a wide range of transition metals. Additionally, we compare two models for achieving these separate tasks: (a) a grid-based approach in which orbital properties are computed from their real space values on a voxel grid, and (b) a force-field like approach in which orbital properties are computed from their projection onto different atomic basis functions. We compare these approaches and present initial results.

2 Data

As stated, the QMOrb dataset consists of Molden file outputs for the entire QM9 database. In this initial work we have carried out restricted Hartree-Fock (RHF) calculations in the minimal STO-3G basis in PySCF.[31] In all, this dataset consists of over seven million orbitals. In this preliminary manuscript, we use only a fraction of these orbitals for training, looking at a random selection of 3000 molecules in the QM9 dataset, resulting in an orbital dataset of only 159,350 orbitals. The labels for this dataset consist of the orbital energies from RHF STO-3G. We are now in the process of generating similar data with double-zeta and triple-zeta basis sets.

The TMOrb dataset consists of constructed transition metal orbitals that are characteristic of four generally human-annotated labels: metal valence orbitals, ligand valence orbitals, metal-ligand bonding orbitals, and metal-ligand antibonding orbitals. Orbitals of these types are generated and assigned labels through the following process:

1. Project the restricted open-shell Hartree-Fock (ROHF)[32] Fock matrix F separately into the space of metal valence orbitals and ligand (i.e., non-metal) valence orbitals, then diagonalize the Fock matrix in these subspaces to form the canonical metal valence

orbitals $\{\phi_M\}$ and ligand valence orbitals $\{\phi_L\}$. These orbitals are given the labels "metal valence" and "ligand valence," respectively.

2. Looping over the space of $\{\phi_M\}$ and $\{\phi_L\}$, project the Fock operator into the space of one ligand orbital ϕ_L and one metal orbital ϕ_M, then diagonalize within this space. If there is a stabilization of the bonding orbital energy greater than 0.05 Hartree with respect to the initial orbitals, the canonical orbitals within this space are added to the dataset and given the labels "metal-ligand bonding" and "metal-ligand antibonding," respectively.

Geometries for the procedure above were taken from the large dataset of Kulik and coworkers of 4865 first-row transition metal octahedral complexes.[33] Calculations were undertaken with ROHF in a modified version of the "minao" basis in PySCF, generated by taking the first contracted functions from cc-pVTZ,[34–36] and excluding the addition of the 4s orbital from calculations. Converged calculations resulted in a final dataset of 3702 first-row transition metal complexes. As there are naturally many more ligand orbitals than metal-ligand bonding and antibonding orbitals, the metal-ligand bonding and antibonding orbitals were added to the dataset with an equal number of randomly selected ligand metal orbitals for a total dataset of 41,952 orbitals: 10,488 each of metal valence, ligand valence, metal-ligand bonding, and metal-ligand antibonding.

3 Methods

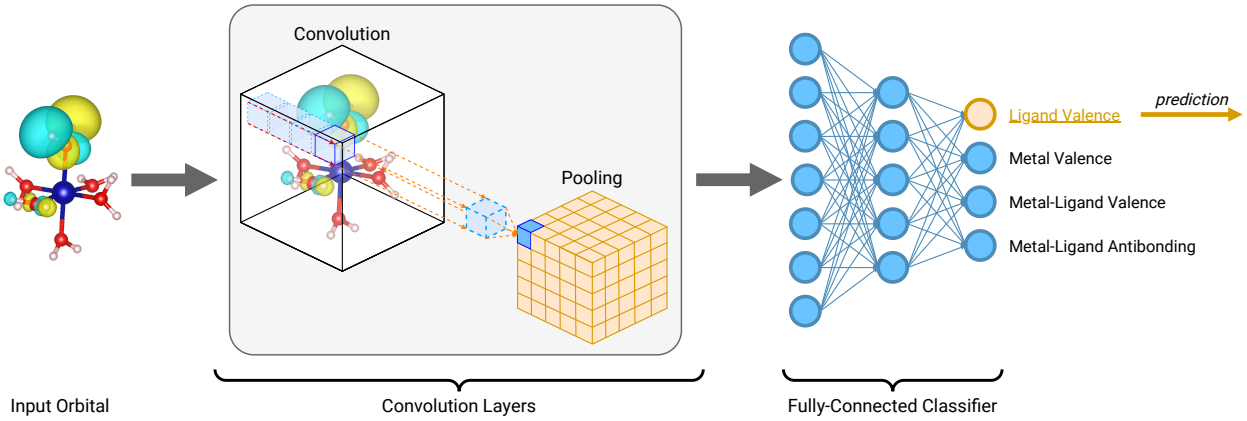

Figure 1: A sketch of our model which learns orbital features from voxel CUBE data. The large CUBE files are fed into convolutional channels which are used to predict metal, ligand, bonding, and antibonding character.

Two broad methods are considered for the learning of orbital labels. Figure 1 shows the first of these methods, in which orbital properties are predicted from their featurization via the standard CUBE format generally used to visualize molecular orbitals. This data consists of vectors (f_i, x, y, z), where f_i is the value of the i^{th} orbital ϕ_i at the coordinate (x, y, z) on a $80 \times 80 \times 80$ grid spanning the molecule. This is the most straightforward approach, as data of this type is already visualized directly by users of electronic structure codes to assign qualitative labels (e.g., bonding or antibonding). Additionally, the voxel-like nature of the input data allows for the use of standard convolutional techniques. For these data, we train a 3D convolutional neural network with an architecture composed of two 3D convolutional layers followed by three fully-connected layers. The standard Rectified Linear Unit (ReLU) nonlinear activation function is used between layers. Model training is performed by using the Adam optimization algorithm[37] with a learning rate of $\eta = 10^{-5}$.

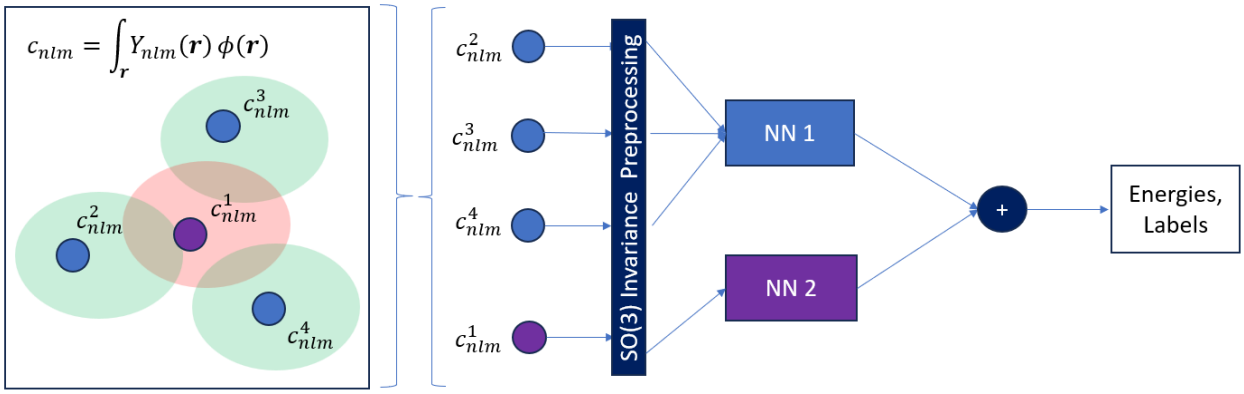

Figure 2: A sketch of our model which learns orbital features from projections of molecular orbitals onto atom-centered basis functions with different radial diffusion (n), angular momentum (l), and Cartesian direction (m) characteristics. These features are then preprocessed and fed into separate feed-forward networks for energy prediction.

However, the CUBE-based approach shown in Figure 1 presents challenges, mainly those of (a) storing the CUBE data, and (b) overcoming the problem of rotational invariance, which is not well-captured by the voxel format. Thus we also propose the alternative approach presented in Figure 2, in which atom-centered features of the orbitals c^{α}_{nlm} are obtained via projection of the orbitals onto atom-centered basis functions of different radial diffusion (n), angular momentum (l), and cartesian direction (m):

$$c^{\alpha}_{nlm} = \int_{\mathbf{r}} Y^{\alpha}_{nlm}(\mathbf{r})\phi(\mathbf{r})d\mathbf{r} \tag{2}$$

where α represents a single atom in the molecule. These inputs are then fed into a preprocessing layer in which the different components of the angular momentum are summed over to construct atom-centered rotationally invariant features d^{α}_{nl} (i.e., the "SO(3) Preprocessing" layer in Figure 2):

$$d^{\alpha}_{nl} = \sum_{m} (c^{\alpha}_{nlm})^2 \tag{3}$$

These features are then fed into element-specific feed-forward networks to predict the molec-

ular orbital energy:

$$E(\phi) = \sum_{\alpha} f^{\xi(\alpha)}(d_{nl}^{\alpha}(\phi)) \tag{4}$$

where $\xi(\alpha)$ is the atomic number of atom α.

This approach is inspired by developed methods for the machine learning of density functionals,[38,39] mainly the work of Dick and Fernandez-Serra.[38] A total of 14 basis functions are employed for each of $l \in \{s, p, d\}$, resulting in a total of 37 d_{nl}^{α} features for each atom.[38,39] However, we note that here it is much more efficient to implement these models compared to density functional applications as the molecular orbitals are described naturally in an atomic orbital basis, whereas the density projections must be found via quadrature.

4 Results

We present first results of the atomic feature model in estimating molecular orbital energies and then results in labeling molecular orbitals. Note that these are just initial results to introduce the procedures, larger models trained on the entirety of the dataset are ongoing in our group.

4.1 Predicting Orbital Energies

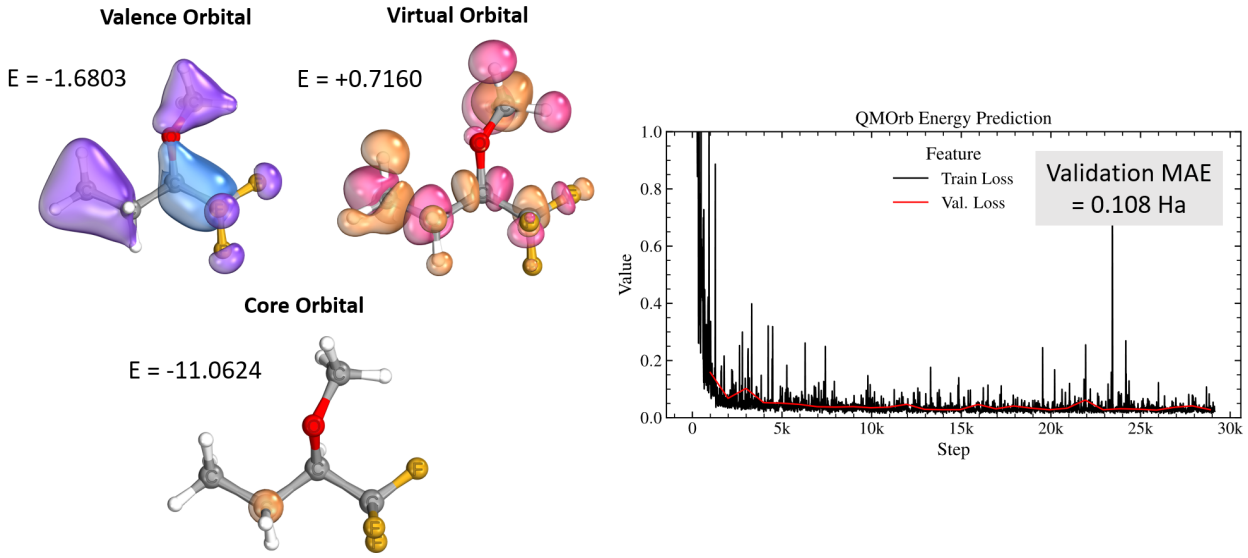

Figure 3: Results of training the element-based model (Equation 4) on predicting orbital energies in the QMOrb dataset. Left: examples of different types of molecular orbitals in the QMOrb dataset: core, valence doubly occupied, and virtual. Right: Validation mean absolute error of the model as a function of the training step.

The results of the atomic feature model (Equation 4) in estimating the energies of QM9 molecular orbitals are shown in Figure 3. Each of the 5 element networks (one for each of C,H,O,N,F in QM9) consists of a deep feed-forward network of six layers with 20 nodes each employing the Gaussian error linear unit (GELU) activation function,[40] resulting in a model with roughly 12k parameters. In addition, separate models are constructed for predicting the energies of doubly occupied and virtual orbitals, bringing the total number of parameters to 24k. Constructing separate models for doubly occupied and virtual orbitals is necessary due to symmetry breaking in HF.

The training curve, optimized with the Adam optimizer[37] with both learning rate (or step size) and weight decay set to 10^{-3}, is shown in Figure 3. At training end, a mean absolute error of 0.108 Ha is achieved on the holdout validation set of a randomly selected 20% of the data. While the performance of this preliminary model (with only a small fraction of the total data and no interaction between atoms, see Equation 4) is evidently not useful for

predictive techniques, it shows that the featurization can efficiently capture the differences between occupied, virtual, and core orbitals on a diverse set of molecules.

4.2 Predicting Orbital Characteristics

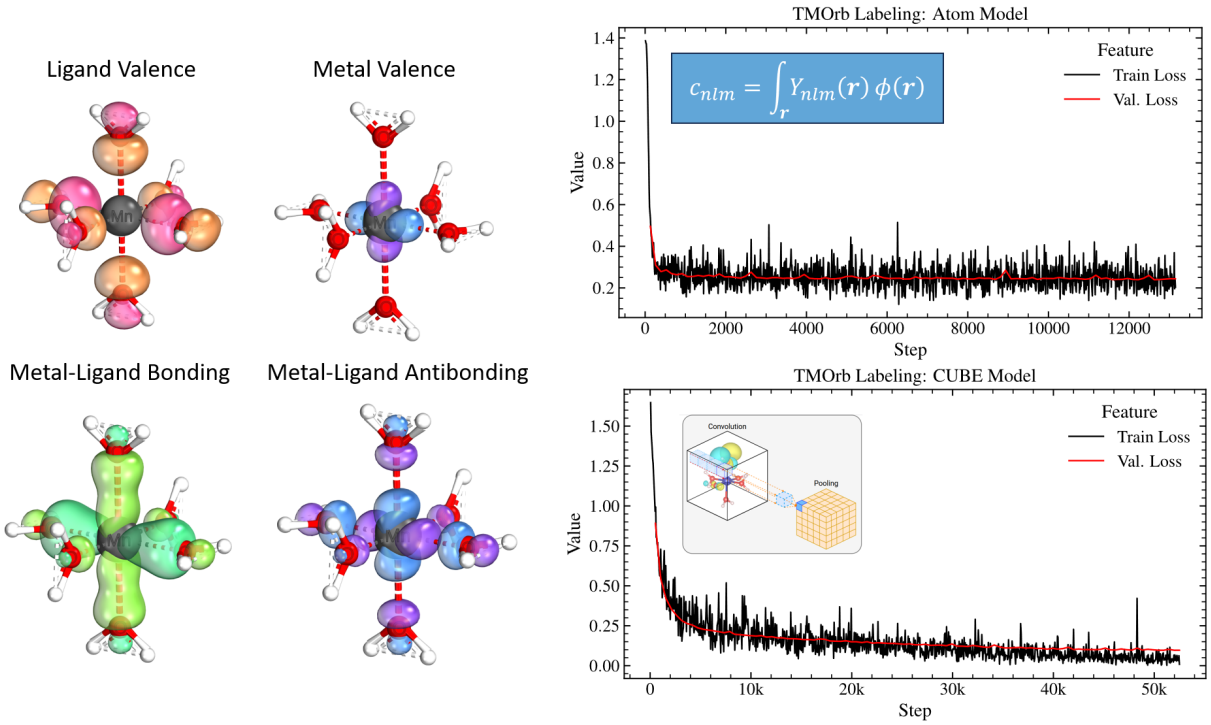

Figure 4: Results of model training on the task of assigning orbital labels to different types of constructed metal-ligand orbitals. Left: The four different types of orbital labels available in the training data. Right: Training accuracies on the holdout validation set of data for both the atom-based approach (top) and grid-based approach (bottom).

Next, we show the performance of our models in labeling different types of molecular orbitals for a wide array of octahedral transition metal complexes; examples of each category for the complex $Mn(OH_2)_6$ are shown on the left of Figure 4. As one can see, the constructed ligand orbital and metal orbital cleanly rotate into a bonding and antibonding orbital pair. The performance of both the atom-based and grid-based models in correctly labeling these different categories is shown on the right of Figure 4.

For this task, we employed a simpler version of the atom-based model which consists of a

single network that is shared between all transition metal atoms, with no networks for non-transition metal atoms (and with one less layer than the QMOrb model, resulting in a model with about 2000 parameters). This featurization is centered on the transition metal, and thus provides an efficient basis for identifying bonding and antibonding character compared to the grid-based approach. A prediction accuracy of 90% is achieved in a small number of training steps, again showing the capability of this featurization in discriminating between different types of orbitals.

The grid-based model also achieves good results on this task. While evidently a less efficient featurization of the input data (i.e., with a lack of rotational invariance and uncentered on the transition metal), the model surpasses the accuracy of the atom-based approach, achieving a validation accuracy of 97%. This result suggests that while the atom-centered approach may be more efficient, the grid-based approach is ultimately more expressive in capturing orbital character, at least with the current preliminary implementations of these models.

5 Conclusion and Future Work

Given the growing status of deep learning in quantum chemistry, models are needed that understand the concept and characteristics of molecular orbitals. Towards this goal, we have here presented what we believe to be the first work focused on assigning labels, specifically energy and bonding character, to molecular orbitals from electronic structure calculations. We have presented two new datasets for this task: QMOrb, which consists of Molden output files from RHF calculations for the entire QM9 set of organic molecules, and TMOrb, which consists of thousands of automatically constructed bonding and antibonding orbitals for a diverse set of first-row octahedral transition metal complexes.

For these two tasks of (a) learning orbital energies, and (b) learning orbital characteristics,

we have presented two complementary approaches: (a) a grid-based approach, which learns orbital labels from their characterization by CUBE files (i.e., a real-space voxel description), and (b) an atom-centered approach that predicts orbital labels from their projection onto different atom-centered basis functions. These two approaches provide complementary trade-offs, with the atom-centered approach serving as a more efficient and rotationally invariant description, while the grid-based approach is likely to be more expressive given the real-space nature of the orbitals. This distinction is marked in the different performance of these models on the classification task, with the atom-centered approach reaching high accuracy (90% accuracy) more quickly while the CUBE-based approach is ultimately more accurate (97% accuracy).

For the predictions of orbital energies, the atom-centered approach achieves a mean absolute error of 0.108 Ha, which while not useful for practical applications, shows that it can broadly identify between different sets of important orbital types (e.g., valence, virtual, and core orbitals). We emphasize that this error is a preliminary result, and future efforts will be focused on training the model by using the entire dataset of over 7M orbitals and scaling up the model by including interactions between atomic features. Additionally, the performance of the CUBE-based approach for energy prediction will be tested and compared. However, we hope that these current results have shown this field to be a promising region of development.

This approach, once finalized, will be helpful to perform quantum chemistry calculations, especially multireference ones, on large data sets of chemical reactions and electronic excitations. We expect that the training of these larger models will take 3-4 months of further development.

6 Acknowledgment

This work is supported as part of the Catalyst Design for Decarbonization Center, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under award no. DE-SC0023383. Additionally, the authors thank the Research Computing Center (RCC) at the University of Chicago for access to computational resources.

References

(1) Dral, P. O. Quantum Chemistry in the Age of Machine Learning. *J. Phys. Chem. Lett.* **2020**, *11*, 2336–2347.

(2) Aldossary, A.; Campos-Gonzalez-Angulo, J. A.; Pablo-Garcia, S.; Leong, S. X.; Rajaonson, E. M.; Thiede, L.; Tom, G.; Wang, A.; Avagliano, D.; Aspuru-Guzik, A. In Silico Chemical Experiments in the Age of AI: From Quantum Chemistry to Machine Learning and Back. *ChemRxiv* **2024**, doi:10.26434/chemrxiv–2024–1v269.

(3) Westermayr, J.; Marquetand, P. Machine Learning for Electronically Excited States of Molecules. *Chem. Rev.* **2021**, *121*, 9873.

(4) Huang, B.; Von Lilienfeld, O. A. Ab Initio Machine Learning in Chemical Compound Space. *Chem. Rev.* **2021**, *121*, 10001–10036.

(5) Keith, J. A.; Vassilev-Galindo, V.; Cheng, B.; Chmiela, S.; Gastegger, M.; Muller, K.-R.; Tkatchenko, A. Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems. *Chem. Rev.* **2021**, *121*, 9816–9872.

(6) Cuomo, S.; Di Cola, V. S.; Giampaolo, F.; Rozza, G.; Raissi, M.; Piccialli, F. Scientific Machine Learning Through Physics–Informed Neural Networks: Where We Are and What's Next. *J. Sci. Comput.* **2022**, *92*, 88.

(7) Nebgen, B.; Lubbers, N.; Smith, J. S.; Sifain, A. E.; Lokhov, A.; Isayev, O.; Roitberg, A. E.; Barros, K.; Tretiak, S. Transferable Dynamic Molecular Charge Assignment Using Deep Neural Networks. *J. Chem. Theory Comput.* **2018**, *14*, 4687–4698.

(8) Unke, O. T.; Meuwly, M. PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges. *J. Chem. Theory Comput.* **2019**, *15*, 3678–3693.

(9) Schütt, K. T.; Gastegger, M.; Tkatchenko, A.; Müller, K.-R.; Maurer, R. J. Unifying Machine Learning and Quantum Chemistry With a Deep Neural Network for Molecular Wavefunctions. *Nat. Commun.* **2019**, *10*, 5024.

(10) Qiao, Z.; Welborn, M.; Anandkumar, A.; Manby, F. R.; Miller, T. F. OrbNet: Deep Learning for Quantum Chemistry Using Symmetry-Adapted Atomic-Orbital Features. *J. Chem. Phys.* **2020**, *153*, 124111.

(11) Christensen, A. S.; Sirumalla, S. K.; Qiao, Z.; O'Connor, M. B.; Smith, D. G. A.; Ding, F.; Bygrave, P. J.; Anandkumar, A.; Welborn, M.; Manby, F. R.; Miller III, T. F. OrbNet Denali: A Machine Learning Potential for Biological and Organic Chemistry With Semi-Empirical Cost and DFT Accuracy. *arXiv:2107.00299 [physics]* **2021**, arXiv: 2107.00299.

(12) Gong, X.; Li, H.; Zou, N.; Xu, R.; Duan, W.; Xu, Y. General Framework for E (3)-Equivariant Neural Network Representation of Density Functional Theory Hamiltonian. *Nat. Commun.* **2023**, *14*, 2848.

(13) Cignoni, E.; Suman, D.; Nigam, J.; Cupellini, L.; Mennucci, B.; Ceriotti, M. Electronic Excited States From Physically Constrained Machine Learning. *ACS Cent. Sci.* **2024**, *10*, 637–648.

(14) Dewar, M. J.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.

(15) Stewart, J. J. Optimization of Parameters for Semiempirical Methods IV: Extension of MNDO, AM1, and PM3 to More Main Group Elements. *J. Mol. Model.* **2004**, *10*, 155–164.

(16) Krylov, A. I. From Orbitals to Observables and Back. *J. Chem. Phys.* **2020**, *153*.

(17) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8*, 527–541.

(18) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864.

(19) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*; John Wiley & Sons, 2013; Google-Books-ID: k4R6cf7I7q0C.

(20) Mi, W.; Luo, K.; Trickey, S.; Pavanello, M. Orbital-Free Density Functional Theory: An Attractive Electronic Structure Method for Large-Scale First-Principles Simulations. *Chem. Rev.* **2023**, *123*, 12039–12104.

(21) Koopmans, T. Über Die Zuordnung Von Wellenfunktionen Und Eigenwerten Zu Den Einzelnen Elektronen Eines Atoms. *physica* **1934**, *1*, 104–113.

(22) Phillips, J. Generalized Koopmans' Theorem. *Phys. Rev.* **1961**, *123*, 420.

(23) Pereira, F.; Xiao, K.; Latino, D. A.; Wu, C.; Zhang, Q.; Aires-de Sousa, J. Machine Learning Methods to Predict Density Functional Theory B3LYP Energies of HOMO and LUMO Orbitals. *J. Chem. Inf. Model.* **2017**, *57*, 11–21.

(24) Pinheiro, G. A.; Mucelini, J.; Soares, M. D.; Prati, R. C.; Da Silva, J. L.; Quiles, M. G. Machine Learning Prediction of Nine Molecular Properties Based on the SMILES Representation of the QM9 Quantum-Chemistry Dataset. *J. Phys. Chem. A* **2020**, *124*, 9854–9866.

(25) Westermayr, J.; Maurer, R. J. Physically Inspired Deep Learning of Molecular Excitations and Photoemission Spectra. *Chem. Sci.* **2021**, *12*, 10755–10764.

(26) Chen, K.; Kunkel, C.; Cheng, B.; Reuter, K.; Margraf, J. T. Physics-Inspired Machine Learning of Localized Intensive Properties. *Chem. Sci.* **2023**, *14*, 4913–4922.

(27) Gaul, C.; Cuesta-Lopez, S. Machine Learning for Orbital Energies of Organic Molecules Upwards of 100 Atoms. *Phys. Status Solidi B Basic Res.* **2024**, *261*, 2200553.

(28) Veryazov, V.; Malmqvist, P. Å.; Roos, B. O. How to Select Active Space for Multiconfigurational Quantum Chemistry? *Int. J. Quantum Chem.* **2011**, *111*, 3329–3338.

(29) Schaftenaar, G.; Noordik, J. H. Molden: A pre-and post-processing program for molecular and electronic structures. *J. Comput. Aided Mol. Des.* **2000**, *14*, 123–134.

(30) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; Von Lilienfeld, O. A. Quantum Chemistry Structures and Properties of 134 Kilo Molecules. *Sci. Data* **2014**, *1*, 1–7.

(31) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K.-L. PySCF: The Python-Based Simulations of Chemistry Framework. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1340.

(32) Rittby, M.; Bartlett, R. J. An open-shell spin-restricted coupled cluster method: Application to ionization potentials in nitrogen. *J. Chem. Phys.* **1988**, *92*, 3033–3036.

(33) Liu, F.; Duan, C.; Kulik, H. J. Rapid Detection of Strong Correlation With Machine Learning for Transition-Metal Complex High-Throughput Screening. *J. Phys. Chem. Lett.* **2020**, *11*, 8067–8076.

(34) Dunning Jr, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(35) Woon, D. E.; Dunning Jr, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum Through Argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

(36) Balabanov, N. B.; Peterson, K. A. Systematically Convergent Basis Sets for Transition Metals. I. All-Electron Correlation Consistent Basis Sets for the 3d Elements Sc–Zn. *J. Chem. Phys.* **2005**, *123*.

(37) Kingma, D. P.; Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* **2014**,

(38) Dick, S.; Fernandez-Serra, M. Machine Learning Accurate Exchange and Correlation Functionals of the Electronic Density. *Nat. Commun.* **2020**, *11*.

(39) King, D. S.; Truhlar, D. G.; Gagliardi, L. Machine-Learned Energy Functionals for Multiconfigurational Wave Functions. *J. Phys. Chem. Lett.* **2021**, *12*, 7761–7767.

(40) Hendrycks, D.; Gimpel, K. Gaussian Error Linear Units (Gelus). *arXiv:1606.08415 [cs.LG]* **2016**, Preprint.