

# Machine learning meets chemical physics

Cite as: J. Chem. Phys. **154**, 160401 (2021); <https://doi.org/10.1063/5.0051418>

Submitted: 25 March 2021 • Accepted: 29 March 2021 • Published Online: 23 April 2021

 Michele Ceriotti,  Cecilia Clementi and  O. Anatole von Lilienfeld

## COLLECTIONS

Paper published as part of the special topic on [Machine Learning Meets Chemical Physics](#)



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

### [Chemical physics software](#)

The Journal of Chemical Physics **155**, 010401 (2021); <https://doi.org/10.1063/5.0059886>

### [Perspective on integrating machine learning into computational chemistry and materials science](#)

The Journal of Chemical Physics **154**, 230903 (2021); <https://doi.org/10.1063/5.0047760>

### [Machine learning for interatomic potential models](#)

The Journal of Chemical Physics **152**, 050902 (2020); <https://doi.org/10.1063/1.5126336>



Chemical Physics Reviews

First Articles Now Online!

READ NOW >>>



# Machine learning meets chemical physics

Cite as: *J. Chem. Phys.* **154**, 160401 (2021); doi: [10.1063/5.0051418](https://doi.org/10.1063/5.0051418)

Submitted: 25 March 2021 • Accepted: 29 March 2021 •

Published Online: 23 April 2021



View Online



Export Citation



CrossMark

Michele Ceriotti,<sup>1,a)</sup>  Cecilia Clementi,<sup>2,b)</sup>  and O. Anatole von Lilienfeld<sup>3,c)</sup> 

## AFFILIATIONS

<sup>1</sup>Laboratory of Computational Science and Modeling, IMX, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

<sup>2</sup>Department of Physics, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany

<sup>3</sup>Faculty of Physics, University of Vienna, Kolingasse 14-18, 1090 Vienna, Austria

**Note:** This paper is part of the JCP Special Topic on Machine Learning Meets Chemical Physics.

<sup>a)</sup>Author to whom correspondence should be addressed: [michele.ceriotti@epfl.ch](mailto:michele.ceriotti@epfl.ch)

<sup>b)</sup>Electronic mail: [cecilia.clementi@fu-berlin.de](mailto:cecilia.clementi@fu-berlin.de)

<sup>c)</sup>Electronic mail: [anatole.vonlilienfeld@univie.ac.at](mailto:anatole.vonlilienfeld@univie.ac.at)

## ABSTRACT

Over recent years, the use of statistical learning techniques applied to chemical problems has gained substantial momentum. This is particularly apparent in the realm of physical chemistry, where the balance between empiricism and physics-based theory has traditionally been rather in favor of the latter. In this guest Editorial for the special topic issue on “Machine Learning Meets Chemical Physics,” a brief rationale is provided, followed by an overview of the topics covered. We conclude by making some general remarks.

Published under license by AIP Publishing. <https://doi.org/10.1063/5.0051418>

## I. INTRODUCTION

Machine learning is making its way into all fields of science, and chemical physics is no exception. This special topic collects several contributions that showcase the level to which data-driven methodologies have become intertwined with the practice of this discipline. From the construction of interatomic potentials and of models of atomic-scale properties to the accelerated sampling of rare events and the construction of coarse-grained (CG) descriptions of molecular interactions, there is no corner of computational chemistry and materials science that has not benefited from the incorporation of machine-learning techniques.

Several general trends emerge from the articles that are published in this special issue, which documents the evolution of the field since the publication of the collection on “Data enabled theoretical chemistry” in 2018.<sup>1</sup> One is the coming of age of the discipline: exploratory studies and benchmarks have been increasingly replaced by an effort to optimize and study systematically the interplay of data-driven and physics-inspired approaches, with a particular focus on the descriptors that are used to represent an atomistic configuration. The art of building a set of reference structures for training has also become more standardized, merging with techniques used to sample structural landscapes, active learning, and uncertainty quantification. Thanks to these technical

advances, models of the potential energy have become more accurate, transferable, and easy to build and are often combined with advanced simulation techniques to study problems of greater complexity and sophistication to an extent that was not possible with either empirical force fields or *ab initio* methods. The connection with coarse-graining approaches extends the length and time scales of systems that can be treated with machine-learning potentials. What is more, statistical learning is being applied to other atomic-scale properties beyond energies and forces, such as polarizabilities or nuclear magnetic resonance (NMR)-shifts, as well as to the prediction of ingredients of an electronic-structure calculation. Several articles report successful attempts to predict, or use as inputs, matrix elements of a Hamiltonian or the electron density—further blurring the lines between the numerics of chemical physics-based approximations, and machine learning. The exchange of concepts between the two fields is more intense than ever and is maybe the main driver of the fast-paced progress, as also manifested by the many and important papers in this special issue.

## II. SUMMARY OF AREAS COVERED

The number and variety of contributions collected in this special topic testify to the activity and excitement surrounding the application of machine learning to chemical physics

problems. We give a brief overview of the main subject areas that are represented and provide glimpses of the state of the art presented.

### A. Representations and models for atomic-scale learning

The accuracy of a machine-learning scheme to predict structure–property relations at the atomic scale depends on the interplay between the descriptors used to represent the structures and the regression technique used to associate them with the target properties. Several papers in this special issue try to understand better the interaction between these components and how they determine the performance of the model. In Ref. 2, Bilbrey *et al.* rationalize the neural-network models of the energy of water oligomers by developing topological descriptors of the connectivity of different structures and using them to characterize the dataset and the performance of different models. The choice of input representation is recognized as a crucial ingredient: both Jinnouchi *et al.*<sup>3</sup> and Low *et al.*<sup>4</sup> analyze quantitatively the role of descriptors in the construction of potentials—the former discussing silicon and magnesium oxide and the latter focusing on the prediction of the melting point of ionic liquids. Onat *et al.* take a more abstract approach, investigating the response of different representations to perturbations of the atomic positions.<sup>5</sup>

Representations being so fundamental, it is no surprise that considerable activity concerns the optimization of the input representations: Li *et al.* do so by using pair distribution functions to build atom-centered neural-network potentials,<sup>6</sup> while Casier *et al.* demonstrate the effectiveness of a simple principal component analysis (PCA) compression of the input features to improve the performance of a neural network.<sup>7</sup> The computational efficiency is not less important than the accuracy of the model, and the two are not necessarily in opposition, as shown by Christensen *et al.*,<sup>8</sup> who present FCHL19, a numerically improved variation of the original FCHL18 representation. Grisafi and Ceriotti combine local environment descriptors based on symmetrized atom-density correlations with the ability to describe long-range electrostatics,<sup>9</sup> while Nigam *et al.* provide an efficient scheme to increase the body order of such atom correlations to obtain more descriptive features, providing remarkable accuracy even with the simplest linear models.<sup>10</sup> Finally, as discussed in Ref. 11, Christiansen *et al.* introduce an image-based representation that is specifically developed for reinforcement learning algorithms—complementing the features dedicated to statistical property regression that make up the bulk of those discussed in this issue.

### B. Potentials for materials and molecules

By and large, training models that are capable of predicting energy and forces of atomistic systems—both gas-phase molecules and condensed phases—is the most mature and widespread application of machine learning in atomistic simulations since it directly caters to molecular dynamics applications. Several papers in this issue present the construction of potential energy surfaces (PESs) for molecular systems, pushing the boundaries of the size and complexity of the system being studied. Song *et al.*<sup>12</sup> study a relatively simple  $\text{OH} + \text{HO}_2 \rightarrow \text{O}_2 + \text{H}_2\text{O}$  reaction, focusing on reducing the number of reference quantum chemistry calculations. Dral *et al.*,<sup>13</sup> on the other hand, use a hierarchy of PES trained on different levels of

theory to obtain high accuracy with only few high-end energy evaluations. Bowman and collaborators apply permutation-invariant polynomials to fit the PES of a 15-atom molecule,<sup>71</sup> while Sugisawa *et al.*<sup>14</sup> build a Gaussian process model for a protonated imidazole dimer, corresponding to a 51-dimensional PES. Glick *et al.* use a pairwise neural network to achieve high accuracy in the description of intermolecular terms,<sup>15</sup> while Metcalf *et al.* tackle directly the problem of predicting interaction energies by learning terms computed by symmetry-adapted perturbation theory decomposition.<sup>16</sup> In Ref. 17, Saucedo *et al.* compare gradient-domain machine learning with conventional force fields to achieve a more efficient implementation of molecular PES. In the condensed phase, the focus is on transferability. Rowe *et al.*<sup>18</sup> present an extremely robust potential for carbon, while George *et al.* discuss how one can simultaneously improve the accuracy of vibrational frequency predictions and the transferability of machine-learning potentials.<sup>19</sup> Sinz *et al.*, in Ref. 20, apply wavelet scattering transform to build potentials for both molecular and condensed phase systems, which can maintain high levels of accuracy even when working in an extrapolative regime.

### C. Machine learning for meso-scale models

Inspired by the success in the learning of accurate potential energy functions for atomistic systems from quantum mechanical calculations, similar tools have been used also to learn effective models at reduced resolutions. In particular, in Ref. 21, Wang *et al.* use a kernel based approach to learn a coarse-grained (CG) force field and illustrate the method on the molecular dynamics simulation of two peptides. In the same spirit, dual graph convolution neural networks are used by Ruza *et al.* in Ref. 22 to design temperature transferable CG force fields of ionic liquids. The reverse problem, that is, backmapping from a CG representation to an atomistic description, has also been tackled with machine learning: an approach based on generative adversarial networks has been proposed<sup>23</sup> for backmapping CG macromolecules. The key to the success of machine learned CG force fields lies in the flexible representation of the many-body terms. This is also illustrated by the work of Boattini *et al.*<sup>24</sup> on the modeling of interaction potentials between elastic spheres through symmetry functions. The representation of complex molecular systems as a function of just one or a few collective coordinates for the study of rare events can also be seen as a type of coarse-graining (or model reduction), and machine-learning methods have been applied in this domain as well. Rabben *et al.*<sup>25</sup> show how a neural network can be used to represent dynamical systems by the linear Koopman operator for the study of rare events. Efficient model reduction of complex chemical reactions can also be performed by combining neural networks with multiscale modeling.<sup>26</sup> Additionally, analytical forms for the classical free energy functional of fluids can be obtained by using an “Equation Learning Network,” as presented in Ref. 27. Finally, hydration free energy can be learned with a kernel-based approach, as shown by Rauer and Bereau.<sup>28</sup> These authors also examine how the database bias affects the results.

### D. Machine learning meets electronic structure

Substantial progress has been reported regarding the use of machine learning for the study of quantum properties of molecules or materials. In Ref. 29, Fabrizio *et al.* have successfully developed

and applied machine learning to correct the electronic potential energy dependency on the electron number for optimal tuning of long-range corrected functionals. Molecular dipole moment predictions based on a combination of machine-learned atomic partial charges and atomic dipole moments has been studied by Veit *et al.*<sup>30</sup> The machine learning based modeling of dielectric constants in crystals has been introduced,<sup>31</sup> as well as kinetic energy density fitting for orbital-free density functional theory (DFT), comparing linear to Gaussian process regression (GPR).<sup>32</sup> Deep neural networks have been used by Westermayr and Marquetand to predict UV absorption spectra throughout chemical compound space (SchNarc)<sup>33</sup> as well as by Gastegger *et al.* for molecular wave functions in minimal basis sets<sup>34</sup> and by Qiao *et al.* for predicting single determinant properties throughout compound space using symmetry-adapted atomic-orbital features (OrbNet).<sup>35</sup> Modeling Frenkel Hamiltonian parameters for accelerated exciton dynamics<sup>36</sup> has benefited from machine learning, as well as electron correlation models based on frozen core approximations.<sup>37</sup>

### E. Sampling structural landscapes

The longstanding problem of sampling the high-dimensional configurational space of complex molecular systems can also be tackled by machine learning, for example, in the form of “active learning” schemes that use existing information to propose new data-points and streamline the construction of training sets. In this context, Wang and Tiwary<sup>38</sup> analyze their recently proposed RAVE algorithm that combines machine learning and molecular dynamics for enhanced sampling. Dutta and Sengupta<sup>39</sup> use a Bayesian approach combined with expectation maximization to learn free energy surfaces and transition states of high dimensional systems while Mancini *et al.*<sup>72</sup> use an evolutionary algorithm to reach for low-lying conformers. Active learning is used also by Zhai *et al.*<sup>40</sup> to obtain optimal structural datasets for the training of multibody potential energy functions with quantum accuracy. In similar vein, Karabin and Perez<sup>41</sup> create diverse structural datasets for learning interatomic potentials with an entropy-maximization approach. In Ref. 42, the data for the design of Gaussian Process Regression (GPR) models are generated through an adaptive sampling procedure, while in Ref. 43, the data generation is combined with a GPR for the construction of potential energy surfaces. Virtual reality is also used in combination with neural networks to generate data for the training of potential energy surfaces.<sup>44</sup> Schran *et al.*<sup>45</sup> identify the relevant configurations for model training and control the generalization error by using committee models. In Ref. 46, Lindsey *et al.* use cluster analysis and Shannon information theory for the generation of robust training sets for machine learned force-fields. The active learning approach itself is examined in Ref. 47 to evaluate the suitability of a model to propose structural candidates with the desired properties.

### F. Applications

As anticipated in the Introduction of this Editorial, one of the clearest signals of the maturation of the field is the fact that several contributions are not about the development of new machine-learning methods but rather on their use to perform atomistic studies of diverse and relevant systems of interest. Important contributions include acetaldehyde in atmospheric processes,<sup>48</sup> mechanical

properties of solid platinum,<sup>49</sup> copper clusters at zinc surfaces,<sup>50</sup> energy transfer in vibrationally excited CO,<sup>51</sup> Si crystal growth,<sup>52</sup> boron cage effects on Nd-Fe-B crystal structure's stability,<sup>53</sup> the structure of chalcogen overlayers on Au-surfaces,<sup>54</sup> shear induced ordering in colloidal systems,<sup>55</sup> and self-diffusion<sup>56</sup> as well as structural and thermodynamic properties<sup>57</sup> in Lennard-Jones fluids. Several contributions demonstrate how machine learning can contribute to elucidate the behavior of water. From the dielectric constant in the supercritical fluid<sup>58</sup> and the temperature dependence of nuclear quantum effects in the liquid<sup>59</sup> all the way to the structure and thermodynamics of the liquid/vapor interface,<sup>60</sup> machine learning brings first-principles accuracy to the determination of the properties of one of the most studied systems in chemical physics. Methodological work has also been contributed: Daly and Hernandez report on the prediction of organismal viability from sparse data,<sup>61</sup> Houchins and Viswanathan describe a calculator for cathode optimization of Li-ion batteries,<sup>62</sup> Tran *et al.* describe multi-fidelity methods for uncertainty quantification and Bayesian optimization for materials design of ternary random alloys,<sup>63</sup> and Muraro *et al.* report on a combined machine learning and quantum chemistry protocol to account for radical scavenging activity of bio-active molecules.<sup>64</sup> A different twist on interatomic potentials is given by Liu *et al.*,<sup>65</sup> who use machine learning to rationalize the parameter space of Buckingham potentials for silica.

Even though the majority of contributions are directly associated with atomistic modeling, some papers stand out for the original way they incorporate data-driven techniques into other realms of chemical physics: Chang and Medford classify and predict the energetics of biomass reactions,<sup>66</sup> while Deng *et al.* solve inverse problems in quantum dynamics by Bayesian optimization.<sup>70</sup> Namba *et al.* optimize the parameters of an experimental setup to align molecules using laser beams.<sup>67</sup> Hassan uses neural networks for the inverse design of nanoparticles.<sup>68</sup> Finally, Kratz and Kierfeld use image recognition to improve the speed and accuracy of pendant drop tensiometry,<sup>69</sup> demonstrating the level at which machine learning has also become valuable for experimental applications.

### III. CONCLUSIONS

The application of machine-learning techniques to chemical physics has grown past the point of proof of principle. From the more established applications to theoretical chemistry and simulations to more recent experiments, the work collected in this issue shows that data-driven approaches—although not yet routine—are now part of the tools of the trade, and their use in applications has become quite natural, and can no longer be called novel. The growing understanding of how machine-learning methods should be adapted to the specific requirements of the field is making them more effective and easy to use. We observe a trend to combine them with electronic-structure theory, and more in general with physics-based approaches, getting the best features out of the two paradigms. Machine learning and chemical physics have met, and it seems like they will stay together for the foreseeable future. The enormous potential of this union has been already extensively demonstrated, but we do not think it has yet been fully realized. We look forward with anticipation to the next conceptual advances and exciting applications, and we expect to read about many of these on the pages of this journal.

## ACKNOWLEDGMENTS

We would like to thank our own respective research groups and all the authors who contributed to this issue, the journal editors and the reviewers who ensured its excellent scientific quality, and the editorial staff who assisted throughout its preparation.

## REFERENCES

- M. Rupp, O. A. von Lilienfeld, and K. Burke, "Guest editorial: Special topic on data-enabled theoretical chemistry," *J. Chem. Phys.* **148**, 241401 (2018).
- J. A. Bilbrey, J. P. Heindel, M. Schram, P. Bandyopadhyay, S. S. Xantheas, and S. Choudhury, "A look inside the black box: Using graph-theoretical descriptors to interpret a continuous-filter convolutional neural network (CF-CNN) trained on the global and local minimum energy structures of neutral water clusters," *J. Chem. Phys.* **153**, 024302 (2020).
- R. Jinnouchi, F. Karsai, C. Verdi, R. Asahi, and G. Kresse, "Descriptors representing two- and three-body atomic distributions and their effects on the accuracy of machine-learned inter-atomic potentials," *J. Chem. Phys.* **152**, 234102 (2020).
- K. Low, R. Kobayashi, and E. I. Izgorodina, "The effect of descriptor choice in machine learning models for ionic liquid melting point prediction," *J. Chem. Phys.* **153**, 104101 (2020).
- B. Onat, C. Ortner, and J. R. Kermode, "Sensitivity and dimensionality of atomic environment representations used for machine learning interatomic potentials," *J. Chem. Phys.* **153**, 144106 (2020).
- L. Li, H. Li, I. D. Seymour, L. Koziol, and G. Henkelman, "Pair-distribution-function guided optimization of fingerprints for atom-centered neural network potentials," *J. Chem. Phys.* **152**, 224102 (2020).
- B. Casier, S. Carniato, T. Miteva, N. Capron, and N. Sisourat, "Using principal component analysis for neural network high-dimensional potential energy surface," *J. Chem. Phys.* **152**, 234103 (2020).
- A. S. Christensen, L. A. Bratholm, F. A. Faber, and O. Anatole von Lilienfeld, "FCHL revisited: Faster and more accurate quantum machine learning," *J. Chem. Phys.* **152**, 044107 (2020).
- A. Grisafi and M. Ceriotti, "Incorporating long-range physics in atomic-scale machine learning," *J. Chem. Phys.* **151**, 204105 (2019).
- J. Nigam, S. Pozdnyakov, and M. Ceriotti, "Recursive evaluation and iterative contraction of  $N$ -body equivariant features," *J. Chem. Phys.* **153**, 121101 (2020).
- M.-P. V. Christiansen, H. L. Mortensen, S. A. Meldgaard, and B. Hammer, "Gaussian representation for image recognition and reinforcement learning of atomistic structure," *J. Chem. Phys.* **153**, 044107 (2020).
- Q. Song, Q. Zhang, and Q. Meng, "Revisiting the Gaussian process regression for fitting high-dimensional potential energy surface and its application to the OH + HO<sub>2</sub> → O<sub>2</sub> + H<sub>2</sub>O reaction," *J. Chem. Phys.* **152**, 134309 (2020).
- P. O. Dral, A. Owens, A. Dral, and G. Csányi, "Hierarchical machine learning of potential energy surfaces," *J. Chem. Phys.* **152**, 204110 (2020).
- H. Sugisawa, T. Ida, and R. V. Krems, "Gaussian process model of 51-dimensional potential energy surface for protonated imidazole dimer," *J. Chem. Phys.* **153**, 114101 (2020).
- D. P. Metcalf, A. Koutsoukas, S. A. Spronk, B. L. Claus, D. A. Loughney, S. R. Johnson, D. L. Cheney, and C. D. Sherrill, "Approaches for machine learning intermolecular interaction energies and application to energy components from symmetry adapted perturbation theory," *J. Chem. Phys.* **152**, 074103 (2020).
- Z. L. Glick, D. P. Metcalf, A. Koutsoukas, S. A. Spronk, D. L. Cheney, and C. D. Sherrill, "AP-Net: An atomic-pairwise neural network for smooth and transferable interaction potentials," *J. Chem. Phys.* **153**, 044112 (2020).
- H. E. Saucedo, M. Gastegger, S. Chmiela, K.-R. Müller, and A. Tkatchenko, "Molecular force fields with gradient-domain machine learning (GDML): Comparison and synergies with classical force fields," *J. Chem. Phys.* **153**, 124109 (2020).
- P. Rowe, V. L. Deringer, P. Gasparotto, G. Csányi, and A. Michaelides, "An accurate and transferable machine learning potential for carbon," *J. Chem. Phys.* **153**, 034702 (2020).
- J. George, G. Hautier, A. P. Bartók, G. Csányi, and V. L. Deringer, "Combining phonon accuracy with high transferability in Gaussian approximation potential models," *J. Chem. Phys.* **153**, 044104 (2020).
- P. Sinz, M. W. Swift, X. Brumwell, J. Liu, K. J. Kim, Y. Qi, and M. Hirn, "Wavelet scattering networks for atomistic systems with extrapolation of material properties," *J. Chem. Phys.* **153**, 084109 (2020).
- J. Wang, S. Chmiela, K.-R. Müller, F. Noé, and C. Clementi, "Ensemble learning of coarse-grained molecular dynamics force fields with a kernel approach," *J. Chem. Phys.* **152**, 194106 (2020).
- J. Ruza, W. Wang, D. Schwalbe-Koda, S. Axelrod, W. H. Harris, and R. Gómez-Bombarelli, "Temperature-transferable coarse-graining of ionic liquids with dual graph convolutional neural networks," *J. Chem. Phys.* **153**, 164501 (2020).
- W. Li, C. Burkhart, P. Polińska, V. Harmandaris, and M. Doxastakis, "Backmapping coarse-grained macromolecules: An efficient and versatile machine learning approach," *J. Chem. Phys.* **153**, 041101 (2020).
- E. Boattini, N. Bezem, S. N. Punnathanam, F. Smallegange, and L. Filion, "Modeling of many-body interactions between elastic spheres through symmetry functions," *J. Chem. Phys.* **153**, 064902 (2020).
- R. J. Rabben, S. Ray, and M. Weber, "ISOKANN: Invariant subspaces of Koopman operators learned by a neural network," *J. Chem. Phys.* **153**, 114109 (2020).
- W. Yang, L. Peng, Y. Zhu, and L. Hong, "When machine learning meets multiscale modeling in chemical reactions," *J. Chem. Phys.* **153**, 094117 (2020).
- S.-C. Lin, G. Martius, and M. Oettel, "Analytical classical density functionals from an equation learning network," *J. Chem. Phys.* **152**, 021102 (2020).
- C. Rauer and T. Berreau, "Hydration free energies from kernel-based machine learning: Compound-database bias," *J. Chem. Phys.* **153**, 014101 (2020).
- A. Fabrizio, B. Meyer, and C. Corminboeuf, "Machine learning models of the energy curvature vs particle number for optimal tuning of long-range corrected functionals," *J. Chem. Phys.* **152**, 154103 (2020).
- M. Veit, D. M. Wilkins, Y. Yang, R. A. DiStasio, and M. Ceriotti, "Predicting molecular dipole moments by combining atomic partial charges and atomic dipoles," *J. Chem. Phys.* **153**, 024113 (2020).
- K. Morita, D. W. Davies, K. T. Butler, and A. Walsh, "Modeling the dielectric constants of crystals using machine learning," *J. Chem. Phys.* **153**, 024503 (2020).
- S. Manzhos and P. Golub, "Data-driven kinetic energy density fitting for orbital-free DFT: Linear vs Gaussian process regression," *J. Chem. Phys.* **153**, 074104 (2020).
- J. Westermayr and P. Marquetand, "Deep learning for UV absorption spectra with SchNarc: First steps toward transferability in chemical compound space," *J. Chem. Phys.* **153**, 154112 (2020).
- M. Gastegger, A. McSloy, M. Luya, K. T. Schütt, and R. J. Maurer, "A deep neural network for molecular wave functions in quasi-atomic minimal basis representation," *J. Chem. Phys.* **153**, 044123 (2020).
- Z. Qiao, M. Welborn, A. Anandkumar, F. R. Manby, and T. F. Miller, "OrbNet: Deep learning for quantum chemistry using symmetry-adapted atomic-orbital features," *J. Chem. Phys.* **153**, 124111 (2020).
- A. Farahvash, C.-K. Lee, Q. Sun, L. Shi, and A. P. Willard, "Machine learning Frenkel Hamiltonian parameters to accelerate simulations of exciton dynamics," *J. Chem. Phys.* **153**, 074111 (2020).
- Y. Ikabata, R. Fujisawa, J. Seino, T. Yoshikawa, and H. Nakai, "Machine-learned electron correlation model based on frozen core approximation," *J. Chem. Phys.* **153**, 184108 (2020).
- Y. Wang and P. Tiwary, "Understanding the role of predictive time delay and biased propagator in RAVE," *J. Chem. Phys.* **152**, 144102 (2020).
- P. Dutta and N. Sengupta, "Expectation maximized molecular dynamics: Toward efficient learning of rarely sampled features in free energy surfaces from unbiased simulations," *J. Chem. Phys.* **153**, 154104 (2020).
- Y. Zhai, A. Caruso, S. Gao, and F. Paesani, "Active learning of many-body configuration space: Application to the Cs<sup>+</sup>-water MB-nrg potential energy function as a case study," *J. Chem. Phys.* **152**, 144103 (2020).
- M. Karabin and D. Perez, "An entropy-maximization approach to automated training set generation for interatomic potentials," *J. Chem. Phys.* **153**, 094110 (2020).

- <sup>42</sup>M. J. Burn and P. L. A. Popelier, "Creating Gaussian process regression models for molecular simulations using adaptive sampling," *J. Chem. Phys.* **153**, 054111 (2020).
- <sup>43</sup>G. Schmitz, E. L. Klinting, and O. Christiansen, "A Gaussian process regression adaptive density guided approach for potential energy surface construction," *J. Chem. Phys.* **153**, 064105 (2020).
- <sup>44</sup>S. Amabilino, L. A. Bratholm, S. J. Bennie, M. B. O'Connor, and D. R. Glowacki, "Training atomic neural networks using fragment-based data generated in virtual reality," *J. Chem. Phys.* **153**, 154105 (2020).
- <sup>45</sup>C. Schran, K. Brezina, and O. Marsalek, "Committee neural network potentials control generalization errors and enable active learning," *J. Chem. Phys.* **153**, 104105 (2020).
- <sup>46</sup>R. K. Lindsey, L. E. Fried, N. Goldman, and S. Bastea, "Active learning for robust, high-complexity reactive atomistic simulations," *J. Chem. Phys.* **153**, 134117 (2020).
- <sup>47</sup>Z. del Rosario, M. Rupp, Y. Kim, E. Antono, and J. Ling, "Assessing the frontier: Active learning, model accuracy, and multi-objective candidate discovery and optimization," *J. Chem. Phys.* **153**, 024112 (2020).
- <sup>48</sup>S. Käser, O. T. Unke, and M. Meuwly, "Isomerization and decomposition reactions of acetaldehyde relevant to atmospheric processes from dynamics simulations on neural network-based potential energy surfaces," *J. Chem. Phys.* **152**, 214304 (2020).
- <sup>49</sup>J. Chapman and R. Ramprasad, "Predicting the dynamic behavior of the mechanical properties of platinum with machine learning," *J. Chem. Phys.* **152**, 224709 (2020).
- <sup>50</sup>M. L. Paleico and J. Behler, "Global optimization of copper clusters at the ZnO(1010) surface using a DFT-based neural network potential and genetic algorithms," *J. Chem. Phys.* **153**, 054704 (2020).
- <sup>51</sup>J. Chen, J. Li, J. M. Bowman, and H. Guo, "Energy transfer between vibrationally excited carbon monoxide based on a highly accurate six-dimensional potential energy surface," *J. Chem. Phys.* **153**, 054310 (2020).
- <sup>52</sup>L. Miao and L.-W. Wang, "Liquid to crystal Si growth simulation using machine learning force field," *J. Chem. Phys.* **153**, 074501 (2020).
- <sup>53</sup>D.-N. Nguyen, D.-A. Dao, T. Miyake, and H.-C. Dam, "Boron cage effects on Nd-Fe-B crystal structure's stability," *J. Chem. Phys.* **153**, 114111 (2020).
- <sup>54</sup>D.-J. Liu, J. W. Evans, P. M. Spurgeon, and P. A. Thiel, "Structure of chalcogen overlayers on Au(111): Density functional theory and lattice-gas modeling," *J. Chem. Phys.* **152**, 224706 (2020).
- <sup>55</sup>J. Piękałski, W. Rządowski, and A. Z. Panagiotopoulos, "Shear-induced ordering in systems with competing interactions: A machine learning study," *J. Chem. Phys.* **152**, 204905 (2020).
- <sup>56</sup>J. P. Allers, J. A. Harvey, F. H. Garzon, and T. M. Alam, "Machine learning prediction of self-diffusion in Lennard-Jones fluids," *J. Chem. Phys.* **153**, 034102 (2020).
- <sup>57</sup>G. T. Craven, N. Lubbers, K. Barros, and S. Tretiak, "Machine learning approaches for structural and thermodynamic properties of a Lennard-Jones fluid," *J. Chem. Phys.* **153**, 104502 (2020).
- <sup>58</sup>R. Hou, Y. Quan, and D. Pan, "Dielectric constant of supercritical water in a large pressure-temperature range," *J. Chem. Phys.* **153**, 101103 (2020).
- <sup>59</sup>Y. Yao and Y. Kanai, "Temperature dependence of nuclear quantum effects on liquid water via artificial neural network model based on scan meta-GGA functional," *J. Chem. Phys.* **153**, 044114 (2020).
- <sup>60</sup>O. Wohlfahrt, C. Dellago, and M. Sega, "Ab initio structure and thermodynamics of the RPBE-D3 water/vapor interface by neural-network molecular dynamics," *J. Chem. Phys.* **153**, 144710 (2020).
- <sup>61</sup>C. A. Daly and R. Hernandez, "Optimizing bags of artificial neural networks for the prediction of viability from sparse data," *J. Chem. Phys.* **153**, 054112 (2020).
- <sup>62</sup>G. Houchins and V. Viswanathan, "An accurate machine-learning calculator for optimization of Li-ion battery cathodes," *J. Chem. Phys.* **153**, 054124 (2020).
- <sup>63</sup>A. Tran, J. Tranchida, T. Wildey, and A. P. Thompson, "Multi-fidelity machine-learning with uncertainty quantification and Bayesian optimization for materials design: Application to ternary random alloys," *J. Chem. Phys.* **153**, 074705 (2020).
- <sup>64</sup>C. Muraro, M. Polato, M. Bortoli, F. Aiolfi, and L. Orian, "Radical scavenging activity of natural antioxidants and drugs: Development of a combined machine learning and quantum chemistry protocol," *J. Chem. Phys.* **153**, 114117 (2020).
- <sup>65</sup>H. Liu, Y. Li, Z. Fu, K. Li, and M. Bauchy, "Exploring the landscape of Buckingham potentials for silica by machine learning: Soft vs hard interatomic forcefields," *J. Chem. Phys.* **152**, 051101 (2020).
- <sup>66</sup>C. Chang and A. J. Medford, "Classification of biomass reactions and predictions of reaction energies through machine learning," *J. Chem. Phys.* **153**, 044126 (2020).
- <sup>67</sup>T. Namba, M. Yoshida, and Y. Ohtsuki, "Machine-learning approach for constructing control landscape maps of three-dimensional alignment of asymmetric-top molecules," *J. Chem. Phys.* **153**, 024120 (2020).
- <sup>68</sup>S. A. Hassan, "Artificial neural networks for the inverse design of nanoparticles with preferential nano-bio behaviors," *J. Chem. Phys.* **153**, 054102 (2020).
- <sup>69</sup>F. S. Kratz and J. Kierfeld, "Pendant drop tensiometry: A machine learning approach," *J. Chem. Phys.* **153**, 094102 (2020).
- <sup>70</sup>Z. Deng, I. Tutunnikov, I. S. Averbukh, M. Thachuk, and R. V. Krems, "Bayesian optimization for inverse problems in time-dependent quantum dynamics," *J. Chem. Phys.* **153**, 164111 (2020).
- <sup>71</sup>P. Houston, R. Conte, C. Qu, and J. M. Bowman, "Permutationally invariant polynomial potential energy surfaces for tropolone and H and D atom tunneling dynamics," *J. Chem. Phys.* **153**, 024107 (2020).
- <sup>72</sup>G. Mancini, M. Fusè, F. Lazzari, B. Chandramouli, and V. Barone, "Unsupervised search of low-lying conformers with spectroscopic accuracy: A two-step algorithm rooted into the island model evolutionary algorithm," *J. Chem. Phys.* **153**, 124110 (2020).