The many-body expansion combined with neural networks

Kun Yao, John E. Herr, and John Parkhill

Citation: The Journal of Chemical Physics **146**, 014106 (2017); doi: 10.1063/1.4973380
View online: http://dx.doi.org/10.1063/1.4973380
View Table of Contents: http://aip.scitation.org/toc/jcp/146/1
Published by the American Institute of Physics

The many-body expansion combined with neural networks

Kun Yao, John E. Herr, and John Parkhill
Department of Chemistry, University of Notre Dame du Lac, 251 Nieuwland Science Hall, Notre Dame, Indiana 46556, USA

(Received 23 September 2016; accepted 13 December 2016; published online 4 January 2017)

Fragmentation methods such as the many-body expansion (MBE) are a common strategy to model large systems by partitioning energies into a hierarchy of decreasingly significant contributions. The number of calculations required for chemical accuracy is still prohibitively expensive for the *ab initio* MBE to compete with force field approximations for applications beyond single-point energies. Alongside the MBE, empirical models of *ab initio* potential energy surfaces have improved, especially non-linear models based on neural networks (NNs) which can reproduce *ab initio* potential energy surfaces rapidly and accurately. Although they are fast, NNs suffer from their own curse of dimensionality; they must be trained on a representative sample of chemical space. In this paper we examine the synergy of the MBE and NN's and explore their complementarity. The MBE offers a systematic way to treat systems of arbitrary size while reducing the scaling problem of large systems. NN's reduce, by a factor in excess of 10^6, the computational overhead of the MBE and reproduce the accuracy of *ab initio* calculations without specialized force fields. We show that for a small molecule extended system like methanol, accuracy can be achieved with drastically different chemical embeddings. To assess this we test a new chemical embedding which can be inverted to predict molecules with desired properties. We also provide our open-source code for the neural network many-body expansion, Tensormol. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4973380]

I. INTRODUCTION

The many-body expansion (MBE) lies at the heart of a multitude of computational methods being developed in the realm of *ab initio* theory and force fields. In insulators the high-order terms of the MBE decay rapidly with distance, which makes this type of approximation useful for low-scaling, high-accuracy models of liquids, solids, and biological molecules.[1–11] However an *ab initio* MBE is orders of magnitude more costly than a classical force field. The main limitation comes from the combinatorial growth of effort by including increasing order many-body terms.

In chemistry neural networks (NNs) are growing in popularity to predict molecular properties.[12–20] In neural networks (NNs), the value of a neuron in the hidden layers and output layer is determined by the value of the neurons in the previous layer via the function

$$y_{i,m} = f(b^i_m + \sum_n w^{i,m}_{i-1,n} y^{i-1}_n), \tag{1}$$

where $y_{i,m}$ is the value of the mth neuron in the ith layer, f is the activation function, usually $f(x) = \max(0, x)$ or $\tanh(x)$, b^i_m is the learned bias for the neuron, y^{i-1}_n is the value of the nth neuron in the $i-1$th layer, and $w^{i,m}_{i-1,n}$ is the learned weight that connects these two neurons. With the function above, the input values are non-linearly transformed in the hidden layers to approximate the target value. However NNs have their own limitations, their input must have a constant shape; moreover they must be trained on a representative number of samples. Many previous works using NNs to fit energy surfaces have been highly accurate due to implicit inclusion of many-body effects, but this requires a training set which represents the size of the system to be modeled, and chemical space grows exponentially with molecular size. This curse of dimensionality in the training set is the main barrier to the creation of a universal NN force field with very high accuracy. The purpose of this paper is to show that the MBE provides a very natural and accurate way to alleviate this curse of dimensionality, by balancing the inclusion order of many-body effects with accuracy, while retaining the generality and efficiency of a NN. It is worth mentioning that there are other approaches that can generalize a neural network trained on small systems to large systems.[13,14,21]

Force fields based on the many-body expansion are growing in popularity.[3,6,9,22] Under the MBE scheme, the total energy of an system can be expanded as the sum of the many-body terms,

$$E_{total} = \sum_i^{N_i} E_i + \sum_{i<j}^{N_{ij}} \Delta E_{ij} + \sum_{i<j<k}^{N_{ijk}} \Delta E_{ijk} + \cdots, \tag{2}$$

where E_i is the monomer energy of molecule i, ΔE_{ij} is the two-body energy between molecules i and j, and ΔE_{ijk} is the three-body energy between molecules i, j, k, and so on up to the order of the expansion. The energy of each many-body term is defined by subtracting the energy of lower order many-body terms from the total energy. Explicitly, the two-body energy is given by

$$\Delta E_{ij} = E_{ij} - E_i - E_j \tag{3}$$

and the three-body energy is given by

$$\Delta E_{ijk} = E_{ijk} - \Delta E_{ij} - \Delta E_{ik} - \Delta E_{jk} - E_i - E_j - E_k, \tag{4}$$

where E_{ij} is the energy of the dimer of molecules i and j, and E_{ijk} is the energy of the trimer of molecules i, j, and k. High

order terms are more costly calculations and the error of the MBE is often balanced with the error of the underlying model chemistry at third order[23–25] so long as care is taken to correct for basis set superposition error (BSSE).[26,27] An electrostatically embedded MBE (EE-MBE) has also been proposed as a means to improve the accuracy.[9,28] Others have suggested a many-body expansion scheme of overlapping-fragments as a way to improve upon the accuracy of the energies.[3,29,30]

Statistical models from machine learning are becoming popular chemical models. Examples include fitted potential energy surfaces[14,16,31,32] with atom-centered symmetry functions,[13,21,33,34] and with permutation invariant polynomials.[10,15,18,35–39] Permutationally invariant polynomials have been used to express the many-body energies of water clusters[10,39] and water-methane clusters[40] with great success. Also, machine learning has been used to predict properties, such as atomization energies, HOMO and LUMO eigenvalues, ionization potentials, force constants, dielectric constants,[41–49] quantum transport coefficients,[50] and nuclear magnetic resonance parameters.[17] It has also been used to construct kinetic energy functionals[51–53] and to design new materials.[47,54–57]

To our knowledge, there are few works that combine neural networks with the MBE and those have focused on elemental solids. The closest work predicted the many-body energy of $Si_n (n = 1, 2, \ldots, 7)$[58] clusters. Bartók used machine learning techniques based on Bayesian inference to correct density functional theory (DFT) one-body and two-body energies for water.[59] Gastegger and co-workers used the high-dimensional NN approach of Behler and Parinello[13,33] as an atom-wise fragmentation method for reconstructing the total energy of *trans*-alkanes up to 10 000 carbon atoms.[60] Others have used a genetic algorithm as a means to determine the most relevant many-body interaction types in a cluster expansion method of model Hamiltonians.[61,62] Manzhos and co-workers have used NNs to represent the component functions of a high dimensional model representation. By using NNs, they were able to fit the component functions of an m-th order mode term simultaneously, improving the fit over previous methods, while increasing the order of mode terms that can be used practically, and using fewer *ab initio* points.[63–67] In this paper we learn the many-body energies of condensed phase liquid methanol within mE_h accuracy. We show that one can use the MBE for methanol clusters of a thousand molecules without significant computational expense on typical Graphical Processing Unit (GPU) workstations. We also present a novel chemical embedding, which has the advantage that it is invertible to ball-and-stick geometries, asses it as a descriptor to learn the MBE, and propose it as a useful tool for inverse-design.

II. METHODS

Studies have shown that the MBE converges rapidly for van der Waals and water clusters.[26,68–71] Convergence is relatively slow for metallic or covalent interactions,[70,72] although schemes have been proposed to improve the accuracy of the MBE on covalent systems.[2,3] We chose methanol for its strong hydrogen bonding and partition our many-body fragments as single methanol molecules and include many-body terms up to third order. Training and test geometries are drawn using the OpenMD software from an AMBER molecular dynamics trajectory[73–76] of 108 methanol molecules at 330 K and an *ab initio* trajectory of three methanol molecules at 500 K. We note that atomic connectivities remained constant throughout both trajectories, restricting ourselves to non-reactive sampling. The total data set includes 844 800 sample methanol molecules for learning one-body energies, 74 240 methanol dimers for two-body energies, and 36 864 methanol trimers for three-body energies. 20% of the total data set is used for testing.

RI-MP2 with the cc-pVTZ basis is used to calculate all of the training and testing data for the many-body energies. The integral precision and SCF convergence criteria were as tight as possible[77,78] and BSSE using the k-mer centered basis set approach (*k-CBS*)[69] was applied. All of the *ab initio* calculations are done with the *Q-Chem* package.[79] Previous studies have suggested using a cutoff distance to reduce the number of many-body terms needed to be calculated to reduce the cost of a MBE treatment.[28,68] We also found that both the two-body and three-body energies negligibly different from limiting formulas (dipole-dipole interaction and Axilrod-Teller potential) at a cutoff of 10 Å for a cluster with 108 methanol molecules as shown in Figure SI-1 of the supplementary material and our dimers and trimers were generated within this cutoff of 10 Å.

Choosing the chemical embedding for the system as the input to the NN has a great effect on performance. Many different chemical descriptors have been proposed, including the Coulomb matrix (CM),[80–83] symmetry functions,[13,14] bispectrum,[84,85] permutation invariant polynomials,[15,18] metric fingerprints,[86–88] and the radial distribution Fourier series,[89] which is based on the electronic density and is similar to a descriptor our group has used in the past for learning kinetic functionals.[53] Systematic comparison of different descriptors is beyond the scope of this paper and we choose the Coulomb matrix (CM) as our input for the neural networks for its simplicity and we show that it is capable of the task. The formula of Coulomb matrix can be written as[80]

$$M_{ij} = \begin{cases} 0.5 Z_i^{2.4} & i = j, \\ \frac{Z_i Z_j}{|R_i - R_j|} & i \neq j, \end{cases} \tag{5}$$

where i and j are the ith and jth atom of the molecule, Z_i and Z_j are the charges of the atoms, and R_i and R_j are the positions of the atoms. Since our data set only contains methanol, the diagonal elements of the CM are constant in our scheme and we only include the off-diagonal elements as an input to the network. While the diagonalized CM is permutationally invariant, the CM itself is not; therefore, in this study we augmented our training data with all the permutations of hydrogen atoms on carbon and all the permutations of methanol molecules in the dimer and trimer to learn the permutation invariance. Similar data augmentation techniques have been widely used in image recognition to achieve translation and rotation invariance.[90] As shown in the Figs. SI-2 and SI-3 of the supplementary material, the permutation invariance is learned with satisfying accuracy. The permutation invariance can also be achieved by averaging the result of all the possible permutations.

We experimented with a novel chemical embedding, which we call the depth map (D-map). The purpose of

this descriptor is not to improve over the accuracy of the CM but rather to have an input which provides reasonable energies and inverts directly to molecular geometries. If networks could accurately learn from an invertible input, they could also become useful tools for molecular inverse-design. Similar types of NN inputs have been used in the area of 3D detection and object recognition.[91–94] An example of D-map can be seen in Figure 1. It is simply a depth of field image of a ball-and-stick structure. A simple routine to calculate one is given in the supplementary material. To enforce translational invariance and rotational invariance about two axes, we define the center of the D-map as the Cartesian center of the three methanol molecules, and we rotate each sample to ensure that the oxygen from each methanol molecule is equidistant from the plane of the D-map, which further helps to provide as much information about the dipoles as possible. We augmented our data by adding rotations about the axis normal to the plane of the D-map to train for rotation invariance about this axis. For a simple molecule like methanol, a canonical method to enforce invariance of the D-map is simple, but for larger and more complex molecules this will not be so straightforward. Given the usefulness of this input in molecular design we were curious how well it could be used to predict energies and will compare it to the CM in the results. Generative Adversarial Networks (GANs) have since been studied extensively for their ability to hallucinate authentic looking images.[95–100] We trained a GAN on the D-map to produce hallucinated images of methanol trimers and discuss the utility this provides.

For the CM input, fully connected neural networks were used to learn the many-body energies. The one-body energy was learned by the neural network with one hidden layer (with 50 000 neurons). The two-body energy was learned by the neural network with two hidden layers (with 10 000 neurons in the first layer and 5000 neurons in the second layer). A neural network with three hidden layers was applied to learn the three-body energy, with 1000 neurons, 2000 neurons, and 2000 neurons in the first, second, and third layers, respectively. For the depth map, a combination of five convolutional layers, followed by three fully connected hidden layers, was used. The five convolutional layers have 64, 128, 256, 256, and 128 filters, respectively, with a filter size of 5×5, 4×4, 3×3, 3×3, and 3×3, respectively. The three fully connected layers each contain 1024 neurons. A rectified linear unit (ReLU) $f(x) = \max(0, x)$ was chosen as the activation function of each neuron[101] and $L2$ regularization[102] was applied to prevent over fitting of each NN. The bias and weights are initialized randomly and are optimized to decrease the difference between the neural network output value and the target value using stochastic gradient descent[103] with analytic gradients provided by backpropagation.[90,104] *Cuda-Convnet*[90] was used to train and evaluate the neural networks. Our scheme has since been implemented, including the calculation of forces, into our own software package, Tensormol.

III. RESULTS

Figure 2 shows the comparison of the one-body, two-body, and three-body energies calculated using MP2 and our neural network for the test samples reserved from the original data set. The neural networks give close agreement with MP2 such that errors in the underlying model chemistry would be the limiting factor of MBE-NN. The Mean Absolute Error (MAE) and Mean Signed Error (MSE) of the energy of each many-body terms over a test set independent of the training data are shown in Table I. The MSE is balanced in the sense that each order of the expansion adds comparable microhartree errors to a total energy, as we will discuss in results which follow.

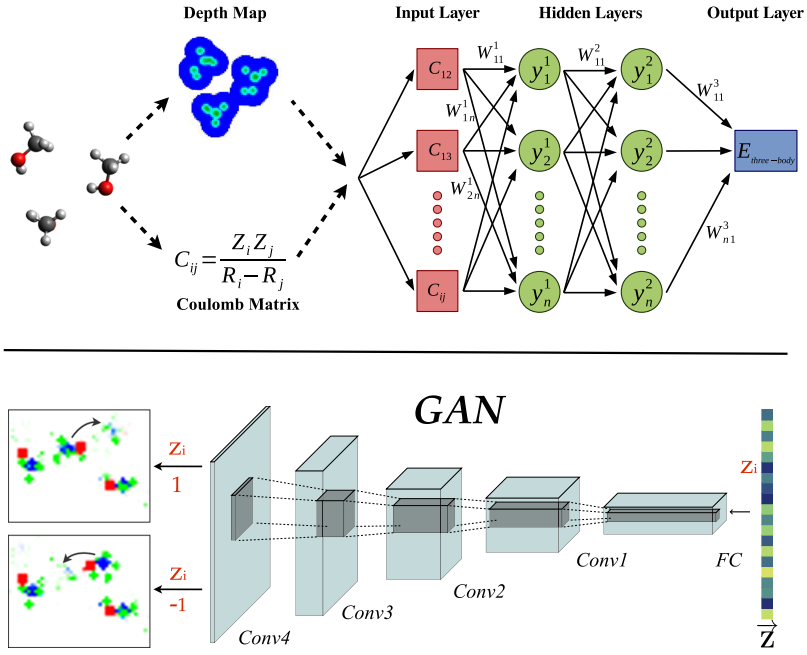

FIG. 1. Top panel: The fragment energy of each N-Body is calculated by embedding the geometry into either the Coulomb matrix or depth map, and evaluating the output of a neural net with several hidden layers and one output. Bottom panel: Generative adversarial network (GAN) scheme. A z-vector is transformed and passed through convolutional hidden layers to generate a hallucinated depth map.

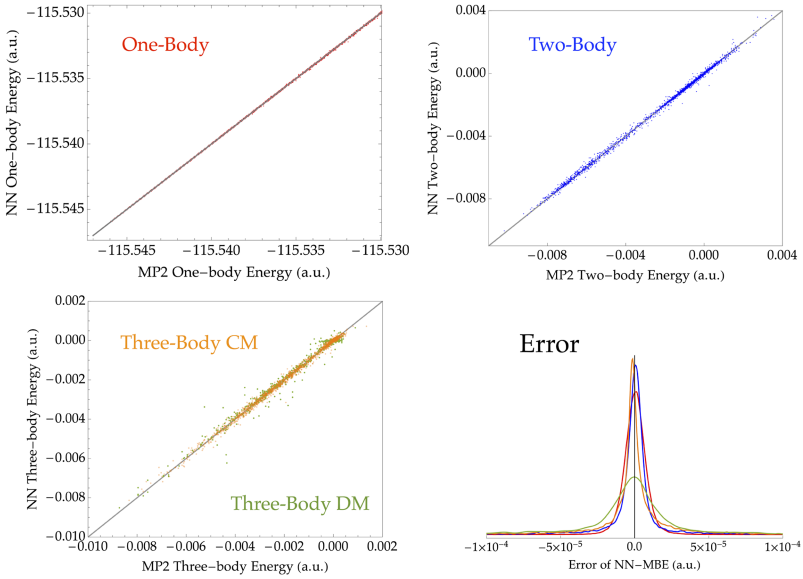

FIG. 2. The top left, top right, and bottom left panels show plots of the one-body, two-body, and three-body energies calculated from MP2 (x-axis) and our neural network (y-axis), respectively. For three-body energy, the result of using the Coulomb matrix (CM) as the input is shown in orange and the result of using the depth map (DM) as the input is shown in green. The bottom right panel shows the histogram of the errors of the many-body energy terms predicted by the neural network. The one-body and two-body errors are shown in red and blue, respectively.

The higher MAE of the higher-body terms is a predictable consequence of our design principle that the ratio of total error to wall-hours should be kept roughly constant at each order of expansion. The three-body energy is a surface of much higher dimension than the two body energy, and this causes difficulty in three ways: a need for more training data, a larger number of invariances, and network capacity. However, we use much less three body data because it is more expensive and less significant. The histogram of errors of all the many-body terms are also shown in Figure 2 and appears uncorrelated which is supported by our later observation that the error per-methanol is essentially insensitive to system size.

The three-body network trained on the D-map also provides reasonable energies. Comparing the three-body energy plots for the CM and the D-map, in most cases the D-map appears to do nearly as well as the CM; however, for a handful of cases the D-map makes significantly poorer predictions. We note this tends to happen when all or part of an oxygen atom becomes eclipsed by the methyl group. Furthermore, we also note the D-map tends to make poorer predictions for energies which are near zero. The distribution of errors remains normally distributed. The D-map does provide a few advantages over the CM: It provides a low dimensional encoding of the space of methanol geometries. It also has constant shape regardless of chemical input and suffers from fewer problems with invariance for a simple molecule like methanol.

As expected the CM makes more accurate predictions than the D-map; however, our aim with the D-map was to provide a chemical embedding which can easily be mapped back to the original geometry of the system. For example this can be used to predict molecules, which maximize a desired property directly without searching chemical space. To this end we then trained a GAN, based on Radford *et al.*,[97] using the D-map by separating element types into separate color channels, to produce hallucinated images of methanol trimers. An example of hallucinated D-map can be seen in the bottom panel of Figure 1. The GAN maps a random z-vector to the data space of the D-map. By varying one element of the z-vector, we were able to control image generation to tune properties. The examples in Figure 1 are from the same z-vector with one varied element z_i, which rotates one of the methanol ends over end. Further work is required to improve the generation of samples and to explore the possibility of this technique generalizing to more interesting properties. For the remainder of this paper we employ the Coulomb embedding for the MBE.

Table II gives the relative energies of the three minimal energy geometries, chair, bowl, and chain[105] of a methanol trimer. All methods shown in Table II get the ordering of these three geometries correct. Compared with MP2, our many-body expansion neural network (NN-MBE) has an error within 10%, which is small compared to Hartree-Fock and B3LYP, which are both significantly more costly. Another

TABLE I. The MAE and MSE (microhartree) of one-body energy, two-body energy, and three-body energy with Coulomb matrix input and three-body energy with depth map input predicted by neural network. We calculate a rate of error as MAE (microhartree)/wall-hours of RI-MP2 in Q-Chem required to generate the training data.

Error	1-body	2-body	3-body (CM)	3-body (DM)
MSE	0.24	0.90	−1.16	−3.96
MAE	5.99	15.6	20.0	39.0
Rate	0.005	0.008	0.005	0.009

TABLE II. Relative energies (mE_h) of three minimal energy geometries of methanol trimer.

Geometry	MP2[a]	MBE-NN	HF[a]	B3LYP[a]
Chair	0	0	0	0
Bowl	1.50	1.40	1.75	2.02
Chain	4.54	4.10	2.56	5.29

[a]MP2, HF, and B3LYP energies are extrapolated to a complete basis.

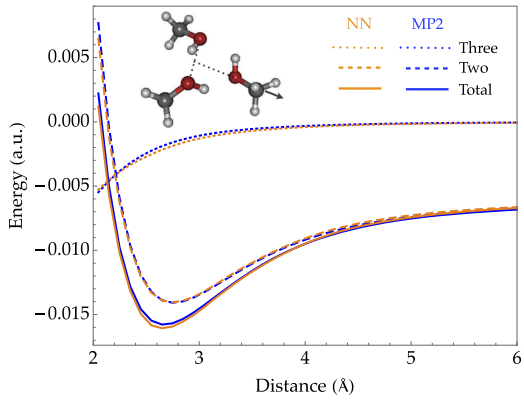

FIG. 3. Dashed line, dotted line, and solid line show the change of two-body energy, three-body energy, and total energy when one methanol is pulled away from the other two. Energies calculated by MP2-MBE and NN-MBE are shown in blue curve and orange curve, respectively.

important feature is the smoothness of predicted surfaces. Figure 3 plots the change of the three-body energy, two-body energy, and total energy when one methanol in a methanol trimer is pulled away from the other two. The agreement is best near the bonding minimum and long distance. Relatively fewer training cases in this region are sampled by the MD trajectories.

The relative energies of five random $(MeOH)_{20}$ clusters are shown in Figure 4 to assess the errors inherent to our MBE model and the error due to the NN. Compared with MP2 energies, the MAE of the MBE using MP2 (MP2-MBE) is 0.10 mE_h per molecule, and for the NN-MBE the MAE is 0.11 mE_h per molecule. Remarkably there is no degradation of accuracy involved in using the NN-MBE, despite massive speedup. Instead the method is limited by the quality of the model chemistry it is built on and the accuracy of the MBE itself. Further, we compare the relative energies obtained

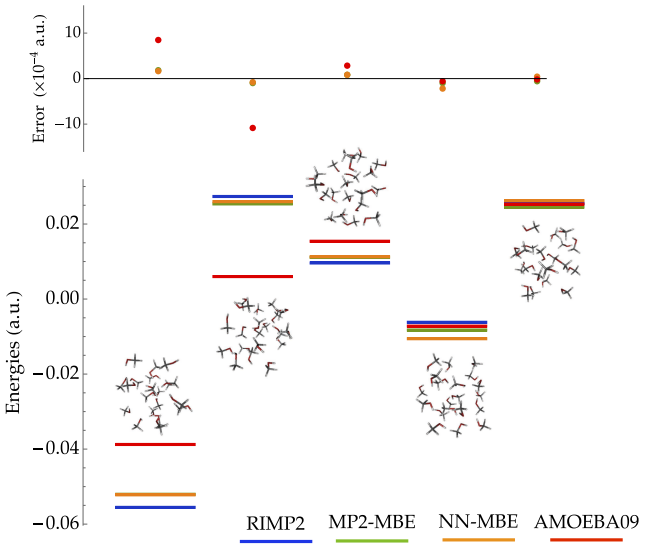

FIG. 4. Bottom panel: relative energies of five $MeOH_{20}$ clusters. Energies calculated by MP2, MP2-MBE, NN-MBE, and AMOEBA09 are shown in blue, green, orange, and red lines, respectively. Top panel: The energy difference relative to MP2 energies is shown for MP2-MBE, NN-MBE, and AMOEBA09 in green, orange, and red dots, respectively.

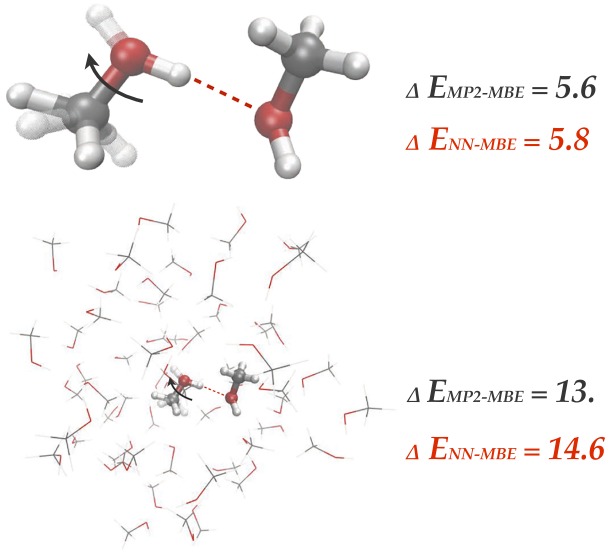

FIG. 5. The top panel (without solvation shell) and bottom panel (with solvation shell) show the energy changes of breaking a hydrogen bond between two methanol by rotating one methanol by 180° around the C–O bond. The units of the energy are kcal/mol. The solvation shell influences the energy change significantly and the neural network predicts the energy change with an accuracy of 1 kcal/mol.

using the AMOEBA09 force field in the TINKER software package[106–108] and obtained an MAE of 0.45 mE_h per molecule, showing that for methanol, our method provides more accuracy.

Proper treatment of solvent effects is crucial for describing most chemical processes. The top panel of Figure 5 shows the energy change of breaking the hydrogen bond between two methanols when the solvation shell is not included. MP2-MBE predicts the energy change to be 5.6 kcal/mol and NN-MBE gives 5.8 kcal/mol. When the solvation shell (with a radius of 10 Å) is included, as shown in the bottom panel, the energy change dramatically increased to 13.3 kcal/mol, which shows the large influence of solvent effects. The NN-MBE predicts the energy change with solvation shell to be 14.6 kcal/mol, 1.3 kcal/mol larger than the MP2-MBE result. Considering the speed up of the NN-MBE, discussed below, and its accuracy, the scheme shows promise for condensed phase phenomena.

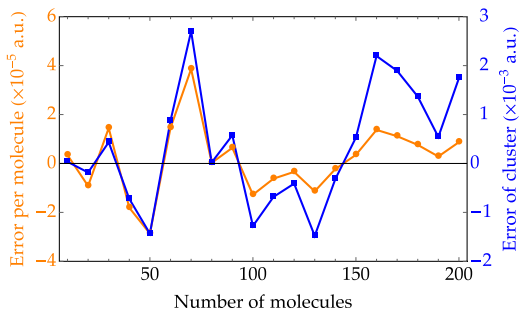

FIG. 6. Error per MeOH (left orange axis) and total error per cluster (right blue axis) of methanol clusters with different size. Error is defined as the difference between NN-MBE energy and MP2-MBE energy. The total error of per cluster which includes up to 200 molecules is in the range of 3 mE_h and the error per molecule decreases with system size.

We also investigated the error of NN-MBE as a function of system size. Figure 6 shows the error per molecule and the error per cluster of the NN-MBE (with respect to MP2-MBE) with an increasing number of molecules in the cluster. The error per cluster stays within the range of 3 mE_h and the error per molecule reaches a maximum at 70 units and shows signs of sub-extensive behavior. Figure SI-5 of the supplementary material provides the total wall time comparison of the NN-MBE and MP2-MBE, showing that the NN-MBE offers a speed up of more than 2×10^6 relative to MP2-MBE without any type of optimization.

IV. DISCUSSION AND CONCLUSIONS

We have shown that a NN-MBE can be used to calculate the energy of methanol clusters with a speed up in the millions with respect to the MP2-MBE. The error of the NN-MBE is within mE_hs, which is similar to the error of MP2-MBE with expansion up to three-body terms. The histogram of the errors of the NN-MBE displays Gaussian shape, which makes the error per molecule decrease with the increase of system size. The satisfying accuracy and huge speed up enable the NN-MBE to treat large system with *ab initio* accuracy, which would otherwise be impossible, such as treating solvation shell effects in *ab initio* calculations. The Coulomb matrix is not invariant to permutations and even though we have shown that permutation invariance can be learned by augmenting the training samples with all of the possible permutations, it is still not perfectly invariant. Our current study focused on methanol up to third order. Theoretically, with our code Tensormol, the NN-MBE scheme can be applied to any system with arbitrary order; however, learning the energy of high order terms takes more computational effort since it is a more challenging task for the neural network and requires more training set.

We introduced a new descriptor, the D-map, which is invertible with the geometry of a system. The D-map was able to predict the three-body energies reasonably well and provides several advantages of its own. We then showed that a generative adversarial network could be trained on the D-map to provide hallucinated images, which are tunable and should be useful for inverse molecular design.

SUPPLEMENTARY MATERIAL

See supplementary material for more details of the test calculations, learning the permutation invariance, and cost/accuracy comparison with AMOEBA09. Our Tensormol code has been open-sourced and the link is also provided.

ACKNOWLEDGMENTS

We thank the University of Notre Dame's College of Science and Department of Chemistry and Biochemistry for generous start-up funding and Nvidia corporation for some processors used in this work.

[1] K. Raghavachari and A. Saha, "Accurate composite and fragment-based quantum chemical models for large molecules," Chem. Rev. **115**, 5643–5677 (2015).
[2] N. J. Mayhall and K. Raghavachari, "Many-overlapping-body (mob) expansion: A generalized many body expansion for nondisjoint monomers in molecular fragmentation calculations of covalent molecules," J. Chem. Theory Comput. **8**, 2669–2675 (2012).
[3] R. M. Richard and J. M. Herbert, "A generalized many-body expansion and a unified view of fragment-based methods in electronic structure theory," J. Chem. Phys. **137**, 064113 (2012).
[4] J. Liu and J. M. Herbert, "Pair-pair approximation to the generalized many-body expansion: An alternative to the four-body expansion for *ab initio* prediction of protein energetics via molecular fragmentation," J. Chem. Theory Comput. **12**, 572 (2016).
[5] S. Wen, K. Nanda, Y. Huang, and G. J. Beran, "Practical quantum mechanics-based fragment methods for predicting molecular crystal properties," Phys. Chem. Chem. Phys. **14**, 7578–7590 (2012).
[6] G. J. Beran, "Approximating quantum many-body intermolecular interactions in molecular clusters using classical polarizable force fields," J. Chem. Phys. **130**, 164115 (2009).
[7] A. Saha and K. Raghavachari, "Dimers of dimers (dod): A new fragment-based method applied to large water clusters," J. Chem. Theory Comput. **10**, 58–67 (2013).
[8] M. S. Gordon, D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, "Fragmentation methods: A route to accurate calculations on large systems," Chem. Rev. **112**, 632–672 (2012).
[9] E. E. Dahlke and D. G. Truhlar, "Electrostatically embedded many-body expansion for large systems, with applications to water clusters," J. Chem. Theory Comput. **3**, 46–53 (2007).
[10] G. R. Medders, V. Babin, and F. Paesani, "A critical assessment of two-body and three-body interactions in water," J. Chem. Theory Comput. **9**, 1103–1114 (2013).
[11] O. A. von Lilienfeld and A. Tkatchenko, "Two-and three-body interatomic dispersion energy contributions to binding in molecules and solids," J. Chem. Phys. **132**, 234109 (2010).
[12] K. J. Jose, D. Beckett, and K. Raghavachari, "Vibrational circular dichroism spectra for large molecules through molecules-in-molecules fragment-based approach," J. Chem. Theory Comput. **11**, 4238–4247 (2015).
[13] J. Behler and M. Parrinello, "Generalized neural-network representation of high-dimensional potential-energy surfaces," Phys. Rev. Lett. **98**, 146401 (2007).
[14] J. Behler, "Neural network potential-energy surfaces in chemistry: A tool for large-scale simulations," Phys. Chem. Chem. Phys. **13**, 17930–17955 (2011).
[15] B. Jiang and H. Guo, "Permutation invariant polynomial neural network approach to fitting potential energy surfaces," J. Chem. Phys. **139**, 054112 (2013).
[16] C. M. Handley and P. L. Popelier, "Potential energy surfaces fitted by artificial neural networks," J. Phys. Chem. A **114**, 3371–3383 (2010).
[17] J. Cuny, Y. Xie, C. J. Pickard, and A. A. Hassanali, "*Ab initio* quality nmr parameters in solid-state materials using a high-dimensional neural-network representation," J. Chem. Theory Comput. **12**, 765–773 (2016).
[18] Z. Zhang and D. H. Zhang, "Effects of reagent rotational excitation on the $H + CHD_3 \rightarrow H_2 + CD_3$ reaction: A seven dimensional time-dependent wave packet study," J. Chem. Phys. **141**, 144309 (2014).
[19] W. Koch and D. H. Zhang, "Communication: Separable potential energy surfaces from multiplicative artificial neural networks," J. Chem. Phys. **141**, 021101 (2014).
[20] J. Chen, X. Xu, and D. H. Zhang, "Communication: An accurate global potential energy surface for the $OH + CO \rightarrow H + CO_2$ reaction using neural networks," J. Chem. Phys. **138**, 221104 (2013).
[21] J. S. Smith, O. Isayev, and A. E. Roitberg, "ANI-1: An extensible neural network potential with DFT accuracy at force field computational cost," preprint arXiv:1610.08935 (2016).
[22] P. Pinski and G. Csanyi, "Reactive many-body expansion for a protonated water cluster," J. Chem. Theory Comput. **10**, 68–75 (2013).
[23] S. S. Xantheas, "*Ab initio* studies of cyclic water clusters $(H_2O)_n$, n = 1-6. II. Analysis of many-body interactions," J. Chem. Phys. **100**, 7523 (1994).
[24] S. S. Xantheas, "Cooperativity and hydrogen bonding network in water clusters," Chem. Phys. **258**, 225–231 (2000).
[25] A. D. Kulkarni, V. Ganesh, and S. R. Gadre, "Many-body interaction analysis: Algorithm development and application to large molecular clusters," J. Chem. Phys. **121**, 5043 (2004).

[26]J. F. Ouyang, M. W. Cvitkovic, and R. P. Bettens, "Trouble with the many-body expansion," J. Chem. Theory Comput. **10**, 3699–3707 (2014).
[27]S. Boys and F. Bernardi, "The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors," Mol. Phys. **19**, 553–566 (1970).
[28]E. E. Dahlke and D. G. Truhlar, "Electrostatically embedded many-body correlation energy, with applications to the calculation of accurate second-order møller-plesset perturbation theory energies for large water clusters," J. Chem. Theory Comput. **3**, 1342–1348 (2007).
[29]R. M. Richard and J. M. Herbert, "Many-body expansion with overlapping fragments: Analysis of two approaches," J. Chem. Theory Comput. **9**, 1408–1416 (2013).
[30]K. U. Lao, K.-Y. Liu, R. M. Richard, and J. M. Herbert, "Understanding the many-body expansion for large systems. II. Accuracy considerations," J. Chem. Phys. **144**, 164105 (2016).
[31]K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, "Quantum-chemical insights from deep tensor neural networks," preprint arXiv:1609.08259 (2016).
[32]S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. Schütt, and K.-R. Müller, "Machine learning of accurate energy-conserving molecular force fields," preprint arXiv:1611.04678 (2016).
[33]J. Behler, "Atom-centered symmetry functions for constructing high-dimensional neural network potentials," J. Chem. Phys. **134**, 074106 (2011).
[34]R. Z. Khaliullin, H. Eshet, T. D. Kühne, J. Behler, and M. Parrinello, "Nucleation mechanism for the direct graphite-to-diamond phase transition," Nat. Mater. **10**, 693–697 (2011).
[35]X. Xin, J. Chen, and D. H. Zhang, "Global potential energy surface for the $H + CH_4 \leftrightarrow H_2 + CH_3$ reaction using neural networks," Chin. J. Chem. Phys. **27**, 373 (2014).
[36]K. Shao, J. Chen, Z. Zhao, and D. H. Zhang, "Communication: Fitting potential energy surfaces with fundamental invariant neural network," J. Chem. Phys. **145**, 071101 (2016).
[37]J. Li and H. Guo, "Permutationally invariant fitting of intermolecular potential energy surfaces: A case study of the Ne–C_2H_2 system," J. Chem. Phys. **143**, 214304 (2015).
[38]J. Li, J. Chen, Z. Zhao, D. Xie, D. H. Zhang, and H. Guo, "A permutationally invariant full-dimensional *ab initio* potential energy surface for the abstraction and exchange channels of the $H + CH_4$ system," J. Chem. Phys. **142**, 204302 (2015).
[39]G. R. Medders, A. W. Götz, M. A. Morales, P. Bajaj, and F. Paesani, "On the representation of many-body interactions in water," J. Chem. Phys. **143**, 104102 (2015).
[40]R. Conte, C. Qu, and J. M. Bowman, "Permutationally invariant fitting of many-body, non-covalent interactions with application to three-body methane–water–water," J. Chem. Theory Comput. **11**, 1631–1638 (2015).
[41]M. Rupp, "Machine learning for quantum mechanics in a nutshell," Int. J. Quantum Chem. **115**, 1058–1073 (2015).
[42]K. Hansen, F. Biegler, R. Ramakrishnan, W. Pronobis, O. A. von Lilienfeld, K.-R. Müller, and A. Tkatchenko, "Machine learning predictions of molecular properties: Accurate many-body potentials and nonlocality in chemical space," J. Phys. Chem. Lett. **6**, 2326–2331 (2015).
[43]G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, "Machine learning of molecular electronic properties in chemical compound space," New J. Phys. **15**, 095003 (2013).
[44]G. Pilania, C. Wang, X. Jiang, S. Rajasekaran, and R. Ramprasad, "Accelerating materials property predictions using machine learning," Sci. Rep. **3**, 2810 (2013).
[45]S. A. Ghasemi, A. Hofstetter, S. Saha, and S. Goedecker, "Interatomic potentials for ionic systems with density functional accuracy based on charge densities obtained by a neural network," Phys. Rev. B **92**, 045131 (2015).
[46]K. Schütt, H. Glawe, F. Brockherde, A. Sanna, K. Müller, and E. Gross, "How to represent crystal structures for machine learning: Towards fast prediction of electronic properties," Phys. Rev. B **89**, 205118 (2014).
[47]R. Olivares-Amaya, C. Amador-Bedolla, J. Hachmann, S. Atahan-Evrenk, R. S. Sánchez-Carrera, L. Vogt, and A. Aspuru-Guzik, "Accelerated computational discovery of high-performance materials for organic photovoltaics by means of cheminformatics," Energy Environ. Sci. **4**, 4849–4861 (2011).
[48]X. Ma, Z. Li, L. E. Achenie, and H. Xin, "Machine-learning-augmented chemisorption model for CO_2 electroreduction catalyst screening," J. Phys. Chem. Lett. **6**, 3528–3533 (2015).
[49]V. Ediz, A. C. Monda, R. P. Brown, and D. J. Yaron, "Using molecular similarity to develop reliable models of chemical reactions in complex environments," J. Chem. Theory Comput. **5**, 3175–3184 (2009).
[50]A. Lopez-Bezanilla and O. A. von Lilienfeld, "Modeling electronic quantum transport with machine learning," Phys. Rev. B **89**, 235411 (2014).
[51]J. C. Snyder, M. Rupp, K. Hansen, L. Blooston, K.-R. Müller, and K. Burke, "Orbital-free bond breaking via machine learning," J. Chem. Phys. **139**, 224104 (2013).
[52]J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, "Finding density functionals with machine learning," Phys. Rev. Lett. **108**, 253002 (2012).
[53]K. Yao and J. Parkhill, "Kinetic energy of hydrocarbons as a function of electron density and convolutional neural networks," J. Chem. Theory Comput. **12**, 1139–1147 (2016).
[54]J. Hachmann, R. Olivares-Amaya, S. Atahan-Evrenk, C. Amador-Bedolla, R. S. Sánchez-Carrera, A. Gold-Parker, L. Vogt, A. M. Brockway, and A. Aspuru-Guzik, "The harvard clean energy project: Large-scale computational screening and design of organic photovoltaics on the world community grid," J. Phys. Chem. Lett. **2**, 2241–2251 (2011).
[55]J. Hachmann, R. Olivares-Amaya, A. Jinich, A. L. Appleton, M. A. Blood-Forsythe, L. R. Seress, C. Roman-Salgado, K. Trepte, S. Atahan-Evrenk, and S. Er, "Lead candidates for high-performance organic photovoltaics from high-throughput quantum chemistry–the harvard clean energy project," Energy Environ. Sci. **7**, 698–704 (2014).
[56]G. Hautier, C. C. Fischer, A. Jain, T. Mueller, and G. Ceder, "Finding natures missing ternary oxide compounds using machine learning and density functional theory," Chem. Mater. **22**, 3762–3767 (2010).
[57]V. Ediz, J. L. Lee, B. A. Armitage, and D. Yaron, "Molecular engineering of torsional potentials in fluorogenic dyes via electronic substituent effects," J. Phys. Chem. A **112**, 9692–9701 (2008).
[58]M. Malshe, R. Narulkar, L. Raff, M. Hagan, S. Bukkapatnam, P. Agrawal, and R. Komanduri, "Development of generalized potential-energy surfaces using many-body expansions, neural networks, and moiety energy approximations," J. Chem. Phys. **130**, 184102 (2009).
[59]A. P. Bartók, M. J. Gillan, F. R. Manby, and G. Csányi, "Machine-learning approach for one-and two-body corrections to density functional theory: Applications to molecular and condensed water," Phys. Rev. B **88**, 054104 (2013).
[60]M. Gastegger, C. Kauffmann, J. Behler, and P. Marquetand, "Comparing the accuracy of high-dimensional neural network potentials and the systematic molecular fragmentation method: A benchmark study for all-trans alkanes," J. Chem. Phys. **144**, 194110 (2016).
[61]G. L. Hart, V. Blum, M. J. Walorski, and A. Zunger, "Evolutionary approach for determining first-principles hamiltonians," Nat. Mater. **4**, 391–394 (2005).
[62]V. Blum, G. L. W. Hart, M. J. Walorski, and A. Zunger, "Using genetic algorithms to map first-principles results to model hamiltonians: Application to the generalized ising model for alloys," Phys. Rev. B **72**, 165113 (2005).
[63]S. Manzhos and T. Carrington, "A random-sampling high dimensional model representation neural network for building potential energy surfaces," J. Chem. Phys. **125**, 084109 (2006).
[64]S. Manzhos and T. Carrington, "Using neural networks to represent potential surfaces as sums of products," J. Chem. Phys. **125**, 194105 (2006).
[65]S. Manzhos and T. Carrington, "Using neural networks, optimized coordinates, and high-dimensional model representations to obtain a vinyl bromide potential surface," J. Chem. Phys. **129**, 224104 (2008).
[66]S. Manzhos, K. Yamashita, and T. Carrington, Jr., "Fitting sparse multi-dimensional data with low-dimensional terms," Comput. Phys. Commun. **180**, 2002–2012 (2009).
[67]S. Manzhos, R. Dawes, and T. Carrington, "Neural network-based approaches for building high dimensional and quantum dynamics-friendly potential energy surfaces," Int. J. Quantum Chem. **115**, 1012–1020 (2015).
[68]J. Cui, H. Liu, and K. D. Jordan, "Theoretical characterization of the $(H_2O)_{21}$ cluster: Application of an n-body decomposition procedure," J. Phys. Chem. B **110**, 18872–18878 (2006).

[69]U. Góra, R. Podeszwa, W. Cencek, and K. Szalewicz, "Interaction energies of large clusters from many-body expansion," J. Chem. Phys. **135**, 224102 (2011).
[70]A. Hermann, R. P. Krawczyk, M. Lein, P. Schwerdtfeger, I. P. Hamilton, and J. J. Stewart, "Convergence of the many-body expansion of interaction potentials: From van der waals to covalent and metallic systems," Phys. Rev. A **76**, 013202 (2007).
[71]G. R. Medders and F. Paesani, "Many-body convergence of the electrostatic properties of water," J. Chem. Theory Comput. **9**, 4844–4852 (2013).
[72]B. Paulus, K. Rosciszewski, N. Gaston, P. Schwerdtfeger, and H. Stoll, "Convergence of the *ab initio* many-body expansion for the cohesive energy of solid mercury," Phys. Rev. B **70**, 165106 (2004).
[73]C. J. Fennell and J. D. Gezelter, "Is the Ewald summation still necessary? Pairwise alternatives to the accepted standard for long-range electrostatics," J. Chem. Phys. **124**, 234104 (2006).
[74]M. Lamichhane, J. D. Gezelter, and K. E. Newman, "Real space electrostatics for multipoles. I. Development of methods," J. Chem. Phys. **141**, 134109 (2014).
[75]M. Lamichhane, K. E. Newman, and J. D. Gezelter, "Real space electrostatics for multipoles. II. Comparisons with the Ewald sum," J. Chem. Phys. **141**, 134110 (2014).
[76]D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, J. Swails, W. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, D. A. Case, R. M. Betz, and P. Kollman, *Amber 2016* (University of California, San Francisco, 2016).
[77]R. M. Richard, K. U. Lao, and J. M. Herbert, "Aiming for benchmark accuracy with the many-body expansion," Acc. Chem. Res. **47**, 2828–2836 (2014).
[78]R. M. Richard, K. U. Lao, and J. M. Herbert, "Understanding the many-body expansion for large systems. I. Precision considerations," J. Chem. Phys. **141**, 014108 (2014).
[79]Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. Gill, and M. Head-Gordon, "Advances in molecular quantum chemistry contained in the Q-Chem 4 program package," Mol. Phys. **113**, 184–215 (2015).
[80]M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, "Fast and accurate modeling of molecular atomization energies with machine learning," Phys. Rev. Lett. **108**, 058301 (2012).
[81]K. Hansen, G. Montavon, F. Biegler, S. Fazli, M. Rupp, M. Scheffler, O. A. Von Lilienfeld, A. Tkatchenko, and K.-R. Muller, "Assessment and validation of machine learning methods for predicting molecular atomization energies," J. Chem. Theory Comput. **9**, 3404–3419 (2013).
[82]G. Montavon, K. Hansen, S. Fazli, M. Rupp, F. Biegler, A. Ziehe, A. Tkatchenko, A. V. Lilienfeld, and K.-R. Müller, "Learning invariant representations of molecules for atomization energy prediction," in *Advances in Neural Information Processing Systems*, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc., 2012), Vol. 25, pp. 440–448.
[83]F. Faber, A. Lindmaa, O. A. von Lilienfeld, and R. Armiento, "Crystal structure representations for machine learning models of formation energies," Int. J. Quantum Chem. **115**, 1094–1101 (2015).
[84]A. P. Bartók and M. C. Payne, "Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons," Phys. Rev. Lett. **104**, 136403 (2010).
[85]A. P. Bartók, R. Kondor, and G. Csányi, "On representing chemical environments," Phys. Rev. B **87**, 184115 (2013).
[86]A. Sadeghi, S. A. Ghasemi, B. Schaefer, S. Mohr, M. A. Lill, and S. Goedecker, "Metrics for measuring distances in configuration spaces," J. Chem. Phys. **139**, 184118 (2013).
[87]L. Zhu, M. Amsler, T. Fuhrer, B. Schaefer, S. Faraji, S. Rostami, S. A. Ghasemi, A. Sadeghi, M. Grauzinyte, and C. Wolverton, "A fingerprint based metric for measuring similarities of crystalline structures," J. Chem. Phys. **144**, 034203 (2016).
[88]B. Schaefer and S. Goedecker, "Computationally efficient characterization of potential energy surfaces based on fingerprint distances," J. Chem. Phys. **145**, 034101 (2016).
[89]O. A. von Lilienfeld, R. Ramakrishnan, M. Rupp, and A. Knoll, "Fourier series of atomic radial distribution functions: A molecular fingerprint for machine learning models of quantum chemical properties," Int. J. Quantum Chem. **115**, 1084–1093 (2015).
[90]A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc., 2012), Vol. 25, pp. 1097–1105.
[91]S. Song and J. Xiao, "Sliding shapes for 3D object detection in depth images," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, 6-12 September 2014, Proceedings, Part VI*, edited by D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars (Springer International Publishing, Cham, 2014), pp. 634–651.
[92]A. Shrivastava and A. Gupta, "Building part-based object detectors via 3D geometry," in The IEEE International Conference on Computer Vision (ICCV), 2013.
[93]K. Lai, L. Bo, X. Ren, and D. Fox, "Detection-based object labeling in 3D scenes," in *IEEE International Conference on Robotics and Automation (ICRA), 2012* (IEEE, 2012), pp. 1330–1337.
[94]B.-s. Kim, S. Xu, and S. Savarese, "Accurate localization of 3D objects from RGB-D data using segmentation hypotheses," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013.
[95]I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, edited by Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (Curran Associates, Inc., 2014), Vol. 27, pp. 2672–2680.
[96]E. L. Denton, S. Chintala, A. Szlam, and R. Fergus, "Deep generative image models using a laplacian pyramid of adversarial networks," in *Advances in Neural Information Processing Systems*, edited by C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett (Curran Associates, Inc., 2015), Vol. 28, pp. 1486–1494.
[97]A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," e-print arXiv:1511.06434 (2015).
[98]D. J. Im, C. D. Kim, H. Jiang, and R. Memisevic, "Generating images with recurrent adversarial networks," e-print arXiv:1602.05110 (2016).
[99]D. Yoo, N. Kim, S. Park, A. S. Paek, and I. Kweon, "Pixel-level domain transfer," e-print arXiv:1603.07442 (2016).
[100]T. Salimans, I. J. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training gans," e-print arXiv:1606.03498 (2016).
[101]G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for lvcsr using rectified linear units and dropout," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2013* (IEEE, 2013), pp. 8609–8613.

[102] T. Hastie, R. Tibshirani, J. Friedman, and J. Franklin, "The elements of statistical learning: Data mining, inference and prediction," Math. Intell. **27**, 83–85 (2005).
[103] K. C. Kiwiel, "Convergence and efficiency of subgradient methods for quasiconvex minimization," Math. Prog. **90**, 1–25 (2001).
[104] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," Nature **323**, 533–536 (1986).
[105] S. Kazachenko, S. Bulusu, and A. J. Thakkar, "Methanol clusters $(CH_3OH)_n$: Putative global minimum-energy structures from model potentials and dispersion-corrected density functional theory," J. Chem. Phys. **138**, 224303 (2013).
[106] Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, "Polarizable atomic multipole-based AMOEBA force field for proteins," J. Chem. Theory Comput. **9**, 4046–4063 (2013).
[107] P. Ren, C. Wu, and J. W. Ponder, "Polarizable atomic multipole-based molecular mechanics for organic molecules," J. Chem. Theory Comput. **7**, 3143–3161 (2011).
[108] J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, and T. Head-Gordon, "Current status of the AMOEBA polarizable force field," J. Phys. Chem. B **114**, 2549–2564 (2010).