A neural network potential-energy surface for the water dimer based on environment-dependent atomic energies and charges

Tobias Morawietz, Vikas Sharma, and Jörg Behler

Citation: The Journal of Chemical Physics **136**, 064103 (2012); doi: 10.1063/1.3682557
View online: http://dx.doi.org/10.1063/1.3682557
View Table of Contents: http://aip.scitation.org/toc/jcp/136/6
Published by the American Institute of Physics

A neural network potential-energy surface for the water dimer based on environment-dependent atomic energies and charges

Tobias Morawietz, Vikas Sharma, and Jörg Behler[a)]
Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

(Received 4 November 2011; accepted 18 January 2012; published online 9 February 2012)

Understanding the unique properties of water still represents a significant challenge for theory and experiment. Computer simulations by molecular dynamics require a reliable description of the atomic interactions, and in recent decades countless water potentials have been reported in the literature. Still, most of these potentials contain significant approximations, for instance a frozen internal structure of the individual water monomers. Artificial neural networks (NNs) offer a promising way for the construction of very accurate potential-energy surfaces taking all degrees of freedom explicitly into account. These potentials are based on electronic structure calculations for representative configurations, which are then interpolated to a continuous energy surface that can be evaluated many orders of magnitude faster. We present a full-dimensional NN potential for the water dimer as a first step towards the construction of a NN potential for liquid water. This many-body potential is based on environment-dependent atomic energy contributions, and long-range electrostatic interactions are incorporated employing environment-dependent atomic charges. We show that the potential and derived properties like vibrational frequencies are in excellent agreement with the underlying reference density-functional theory calculations. © *2012 American Institute of Physics*. [doi:10.1063/1.3682557]

I. INTRODUCTION

Due to its high biological and chemical relevance water has been studied in much greater detail than any other liquid.[1] Nevertheless, a profound understanding of liquid water still remains challenging for both theory and experiment. Theoretical studies using molecular dynamics (MD) or Monte Carlo (MC) simulations[2] can in principle provide valuable insights into the unique properties of water at the atomic level.[3] However, in order to perform such simulations a reliable representation of the interatomic potential is required.

A tremendous number of water potentials has been developed to date.[3] Most of them are based on simple analytical expressions whose parameters are fitted in order to reproduce experimental data of the bulk liquid. Popular examples for empirical water potentials are, among many others, the SPC/E,[4] TIP4P,[5] and TIP5P[6] models. Such potentials are fast to evaluate and thus enable to study very large systems, but they only have a limited transferability and often fail to reproduce properties of water under conditions, which differ from those used in the parameterization.

A complementary approach is to employ *ab initio* calculations to describe the interaction of water molecules. For instance, in *ab initio* MD (AIMD) simulations, the energy and atomic forces are evaluated "on-the-fly" from first principles.[7,8] While yielding accurate results, the high computational costs of AIMD techniques only allow the investigation of rather small systems at short time scales. Moreover, only very efficient methods like density-functional theory (DFT) can be employed to solve the electronic structure problem.

An alternative approach, which aims to combine the efficiency of empirical water models and the accuracy of electronic structure methods, is to develop potentials fitted to data obtained from *ab initio* calculations prior to the actual simulation.[9–11] Here, only a moderate number of electronic structure calculations is needed for representative points of the potential-energy surface (PES), which in general is a complicated high-dimensional function yielding the energy and forces if the atomic positions are provided. Once constructed, such "*ab initio* potentials" can be evaluated much faster than the underlying electronic structure method. Still, many water potentials of this type employ rigid water monomers, i.e., the monomer degrees of freedom are kept fixed and are not explicitly considered in MD simulations.[12] This approximation reduces the dimensionality of the PES significantly and allows to use larger time steps since the high frequency OH modes do not need to be integrated. Nevertheless, freezing the monomers can be a drastic approximation, which might have a strong impact on the simulation results and prevents to study a number of interesting processes involving the dissociation and recombination of water molecules. In recent years, some fully flexible *ab initio* potentials for water have been proposed, which take the deformation of the monomers into account,[13] and also a few reactive force fields for water have been reported[14] to address these problems.

The quality of potentials fitted to *ab initio* data strongly depends on the functional form. Often rather simple functional forms based on physical considerations are employed containing only a few parameters. Consequently, their capability to reproduce the complicated topology of

a)Electronic mail: joerg.behler@theochem.ruhr-uni-bochum.de.

ab initio PESs with high precision is limited. Still, the numerical accuracy could be improved by employing more flexible functions. A promising new method for the construction of highly accurate PESs is based on artificial neural networks (NNs). Inspired by biological NNs, this class of algorithms is used primarily in the fields of pattern recognition and data classification.[15,16] However, in the last two decades NNs have been increasingly employed to function fitting. A particularly interesting example for such applications in chemistry is the construction of potential-energy surfaces by NNs, which has been reviewed by several authors.[17–19] Unlike potentials with physically-motivated functional forms, NNs have a high flexibility, which enables to represent even most complicated functions, and it has been formally proven that in principle any real-valued function, like PESs, can be fitted with arbitrary accuracy.[20]

A large number of NN PESs has been published in recent years for a variety of molecules with up to twelve degrees of freedom[21–27] and also to describe molecule-surface interactions.[28–37] Still, surprisingly, to date no NN potential for fully flexible water clusters or liquid water has been published. The only full-dimensional NN PES for water has been proposed by Manzhos *et al.* for a single water molecule in the gas phase.[38] In all other NN potentials reported for water, the monomer geometries have been frozen. For instance, Gassner *et al.* used NNs to represent the three-body interaction in the H_2O-Al^{3+}-H_2O system.[39] No *et al.* developed a NN potential for the water dimer,[40] which has been extended by Cho, No, and Scheraga to improve the TIP4P model by adding polarization effects implicitly.[41] Finally, Popelier and coworkers have focused on the explicit description of the electrostatic polarization contributions to the PES for clusters containing up to six water molecules.[42,43] They used a set of NNs to reproduce *ab initio* electrostatic multipole moments of a central water molecule as a function of the environment. Still, they have applied NNs just to the electrostatic energy contribution, while no NN total energy surface has been reported.

In general, NNs offer many advantages for the construction of a reliable potential for water. They are numerically very accurate and can be derived using any electronic structure method like, e.g., DFT or higher level post-Hartree Fock methods. They do not require any knowledge about the functional form of the PES, and they are generally "reactive," i.e., they provide the energy and forces as a function of the atomic positions only. No bonds need to be specified, and accordingly water molecules can dissociate and recombine.

In the present work, we explore the applicability of the high-dimensional NN method proposed by Behler and Parrinello in 2007 (Ref. 44) to the water dimer as a first step towards the construction of a potential for liquid water. DFT is used as the reference method. Following an extension of this approach to multicomponent systems by Artrith, Morawietz, and Behler,[45] long-range electrostatic interactions are taken into account by environment-dependent atomic charges. Additionally, in the present work a screening of the electrostatic interactions at short interatomic distances is introduced, which further improves the quality of the potential.

The reliability of the obtained NN potential is investigated by comparing total energies, forces, and atomic charges with reference DFT calculations. Further, binding energies of the water dimer as well as vibrational frequencies and potential curves of several stationary points on the NN PES are examined. Finally, we demonstrate that the NN potential is able to generate MD trajectories that are of essentially the same quality as those obtained in AIMD simulations.

II. NEURAL NETWORK POTENTIALS

A. High-dimensional neural network potentials

To date, NNs have been mainly used to construct low-dimensional PESs of small molecules with up to twelve degrees of freedom. In order to develop a NN potential for water, the NN potential must be applicable to extended systems with a significantly larger number of atoms. In recent years, essentially two approaches to construct high-dimensional NN PESs have been proposed. Both are based on expanding the total energy of an N-atom system as a sum of individual atomic energy contributions E_i, which depend on the local chemical environments and are expressed by a separate NN for each atom,

$$E = \sum_{i=1}^{N} E_i. \tag{1}$$

In the NN scheme of Smith and coworkers[46,47] the atomic environments are described using chains of atoms, and a variable number of neighboring atoms is taken into account by using atomic NNs of adjustable size. In the approach of Behler and Parrinello[44] for each atom a standard feed-forward NN can be used, and the atomic environments are given by a set of many-body symmetry functions for each atom.

Figure 1 illustrates the structure of a high-dimensional NN potential of the Behler Parrinello type. Typically, the atomic configuration is given by a set of Cartesian coordinates $\mathbf{R}_i = (X_i, Y_i, Z_i)$. Unfortunately, these are not a good choice as input for the atomic NNs, because the numerical values of Cartesian coordinates change with rotation and translation of the system. In order to obtain a functional relation between the

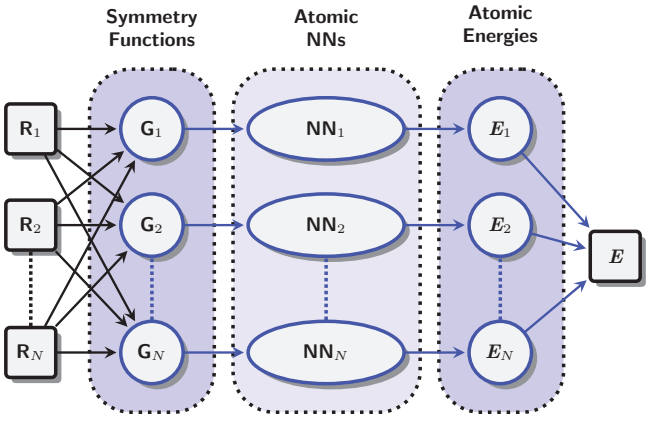

FIG. 1. Illustration of the high-dimensional neural network (NN) scheme[44] for an N atom system: The Cartesian coordinates $\mathbf{R}_i = (X_i, Y_i, Z_i)$ are transformed into vectors of symmetry functions $\mathbf{G}_i$ describing the local environment of each atom. The symmetry function values represent the inputs for atomic NNs providing the atomic energy contributions E_i to the total energy.

structure and energy, a set of coordinates needs to be found, which is rotationally and translationally invariant. Therefore, the Cartesian coordinates $\mathbf{R}_i$ are transformed into so-called symmetry functions G_i, which are many-body functions depending only on the relative positions of all atoms in the local environment.[48] A cutoff function f_c is employed to truncate the radial extension of the symmetry functions.[44] The symmetry function values are then used as input for the atomic NNs, which perform a mapping between the atomic environments and atomic energies E_i. For a given element, the atomic NNs are identical, which takes the invariance of the NN output with respect to a permutation of the atoms into account. Whenever two atoms of the same element switch order, the lines in Fig. 1 are exchanged correspondingly, while the total energy remains unchanged. Furthermore, the high-dimensional NN scheme can be applied to systems with any number of atoms. If an atom is added, another atomic NN is inserted into the scheme of Fig. 1, and whenever an atom is removed, the corresponding NN is deleted.

The structure of an atomic NN, which determines the functional relation between the energy of an atom and its local environment, is shown in Fig. 2. The basic units of an NN are artificial neurons, also called nodes, which are arranged in layers. The nodes in the input layer represent the symmetry function coordinates describing the local environment of the corresponding atom. The node in the output layer yields the atomic energy. The layers in between are called hidden layers. They have no physical meaning, but they define the functional form of the NN. In feed-forward NNs, the information flows only in one direction, from the input to the output nodes. Nodes in adjacent layers are connected via weight parameters, which are the fitting parameters of the NN. They are shown as arrows in Fig. 2. For instance, the weight a_{ij}^{kl} connects node i in layer k with node j in layer $l = k + 1$.

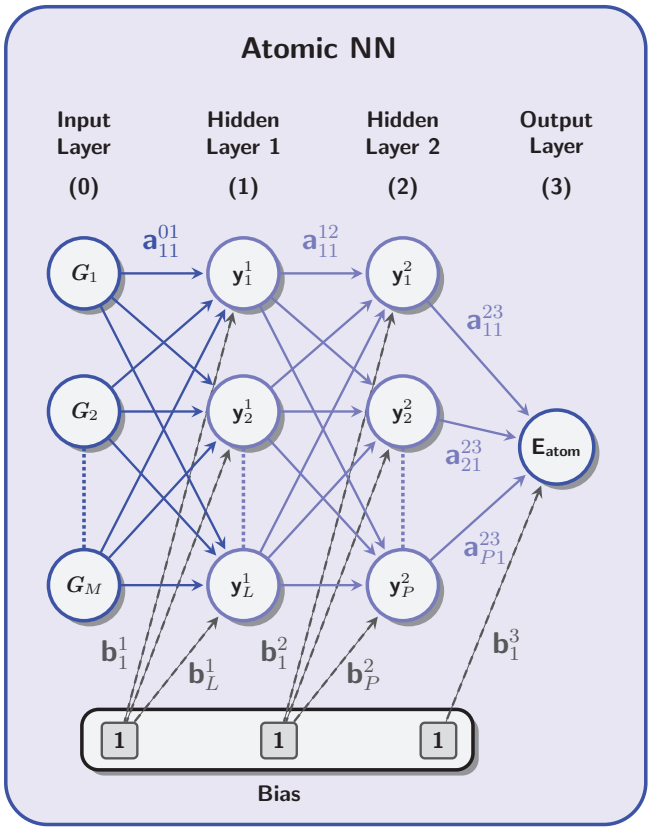

FIG. 2. Example for a feed-forward neural network (NN) used to construct the atomic energy contributions: The atomic energy E_{atom} in the output layer depends on the set of M symmetry function values $\{G_j\}$ provided in the input layer. In the present example there are two hidden layers with L and P nodes, respectively, which define the functional form of the NN. Nodes in adjacent layers are connected by weight parameters shown as arrows, which are the fitting parameters of the NN. The analytic expression for an atomic NN is given in Eq. (2).

To obtain the value y_i^j of node i in layer j, the values of the nodes in the proceeding layer $j-1$ are multiplied by the corresponding connecting weight parameters. Then, all products are added and a non-linear function is applied. This "activation function" provides the capability to represent arbitrary functions. An additional bias weight b_i^j is used to shift the sum to the non-linear region of the activation function. In this work, we have used hyperbolic tangent $f(x) = \tanh(x)$ activation functions for the nodes in the hidden layers and a linear function $f(x) = x$ for the output node. The analytic expression of the atomic NN shown in Fig. 2 containing M input nodes and two hidden layers with L and P nodes, respectively, is then given by

$$E_{\text{atom}} = f_1^3\left(b_1^3 + \sum_{k=1}^{P} a_{k1}^{23} \cdot f_k^2\left(b_k^2 + \sum_{j=1}^{L} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{M} a_{ij}^{01} \cdot G_i\right)\right)\right). \tag{2}$$

The weight parameters of the atomic NNs can be optimized iteratively by minimizing the error of a set of reference energies from electronic structure calculations. Any electronic structure method can be used to calculate the reference data. The reference total energies can be used directly to determine the NN weights, a partitioning of the energy into individual atomic energy contributions is not necessary. In addition to the energy also atomic forces can be used for the optimization process,[26,49,50] which provide $3 \times N$ additional pieces of information per structure. In the NN, the force component F_{α_k} acting on atom k in direction $\alpha = (x, y, z)$ is given by the negative gradient of the energy with respect to α_k

$$F_{\alpha_k} = -\frac{\partial E}{\partial \alpha_k} = -\sum_{i=1}^{N} \frac{\partial E_i}{\partial \alpha_k} = -\sum_{i=1}^{N} \sum_{j=1}^{M_i} \frac{\partial E_i}{\partial G_{i,j}} \frac{\partial G_{i,j}}{\partial \alpha_k}. \tag{3}$$

M_i is the number of symmetry functions describing the environment of atom i. The derivative $\partial E_i / \partial G_{i,j}$ is defined by the architecture of the atomic NNs, while the derivative $\partial G_{i,j} / \partial \alpha_k$ is given by the functional form of the symmetry functions. Several examples for suitable symmetry functions are given in Ref. 48. High-dimensional NN potentials of the Behler Parrinello scheme have been reported for a number of materials like silicon,[51,52] sodium,[53] carbon,[54,55] and copper.[56]

B. Long-range electrostatic interactions

In the high-dimensional NN approach described above, the energy of each atom depends on its local chemical environment, i.e., on the positions of all neighboring atoms inside a sphere with a cutoff radius R_c. This cutoff represents a convergence parameter, and if chosen sufficiently large, very accurate potentials can be obtained. However, in multicomponent systems charge transfer can occur, leading to long-range electrostatic interactions, and truncating these interactions may reduce the accuracy of the NN potential. Interactions, which decay rapidly with increasing interatomic distance, e.g., electrostatic interactions between higher order multipoles, are implicitly considered by the atomic NNs if an appropriate cutoff is used, but monopole-monopole interactions are rather long-ranged and cannot always be neglected. Recently, Artrith, Morawietz, and Behler have thus proposed an extension of the high-dimensional NN approach, which incorporates electrostatic interactions obtained from environment-dependent atomic charges.[45] The total energy of a system is then written as a sum over a short-range part corresponding to the Behler Parrinello scheme and an electrostatic part,

$$E_{\text{tot}} = E_{\text{short}} + E_{\text{elec}}. \tag{4}$$

The resulting high-dimensional NN potential for multicomponent systems is shown schematically in Fig. 3. Two sets of atomic NNs are employed. One NN per atom represents the atomic energies depending on the local environment, which

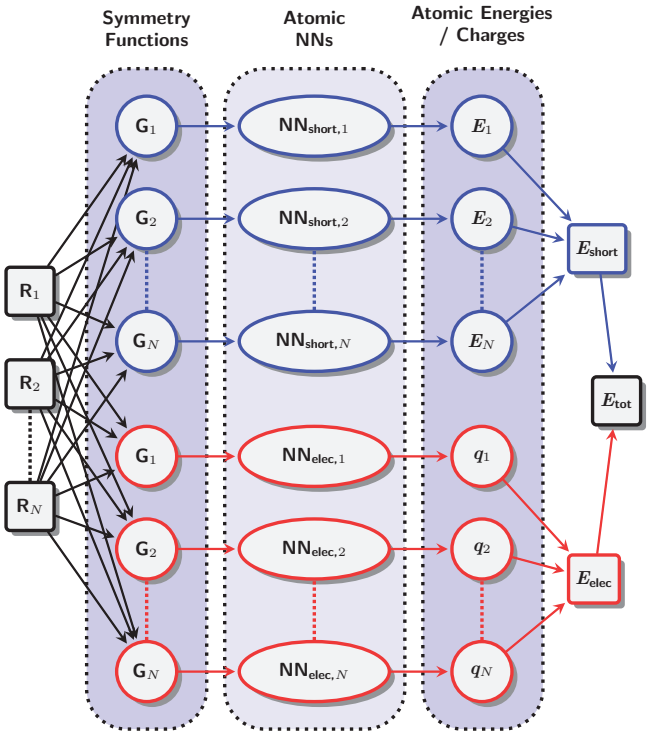

FIG. 3. Scheme of the high-dimensional neural network (NN) potential for multicomponent systems:[45] The total energy E_{tot} of the system is obtained from a short-range contribution E_{short}, which is a sum of atomic energies E_i, and a long-range electrostatic contribution E_{elec}, which is calculated from atomic charges q_i. Both, the short-range atomic energies and the atomic charges, depend on the local chemical environments of the atoms.

in sum yield the short-range energy. A second NN is used for each atom to construct the environment-dependent atomic charges. In the present work, these NNs are trained to reproduce reference charges from DFT calculations. The electrostatic energy can then be calculated employing standard methods like an Ewald summation.[57] As a first example, this extended high-dimensional NN approach for multicomponent systems has been applied to construct a NN potential for zinc oxide.[45]

In the reference DFT calculations the short-range energy is not directly accessible. Hence, in order to determine the weight parameters of the short-range atomic NNs, the electrostatic contribution to the total energy must first be removed from the DFT reference data. For this purpose, the electrostatic energies and forces computed from the atomic charges are subtracted from the respective total DFT values. However, for the separation of the forces into a short-range and an electrostatic part it has to be taken into account that the atomic charges are not fixed, but depend on the chemical environments in a complex way, and this dependence is not available from DFT calculations.

In general, the electrostatic force component $F_{\alpha_k}^{\text{elec}}$ is given by the negative gradient of the electrostatic energy with respect to coordinate α_k

$$F_{\alpha_k}^{\text{elec}} = -\frac{\partial}{\partial \alpha_k} E_{\text{elec}} = -\sum_i \sum_{j>i} \frac{\partial}{\partial \alpha_k} \frac{q_i q_j}{R_{ij}}. \tag{5}$$

By making use of the product rule, Eq. (5) can be expanded into

$$\begin{aligned} F_{\alpha_k}^{\text{elec}} &= -\sum_i \sum_{j>i} \left(q_i q_j \frac{\partial}{\partial \alpha_k} R_{ij}^{-1} + \frac{q_j}{R_{ij}} \frac{\partial q_i}{\partial \alpha_k} + \frac{q_i}{R_{ij}} \frac{\partial q_j}{\partial \alpha_k} \right) \\ &= -\sum_{j \neq k} q_k q_j \frac{\alpha_j - \alpha_k}{R_{kj}^3} - \sum_i \sum_{j>i} \frac{1}{R_{ij}} \left(q_j \frac{\partial q_i}{\partial \alpha_k} + q_i \frac{\partial q_j}{\partial \alpha_k} \right). \end{aligned} \tag{6}$$

The first term is the usual expression for the electrostatic force between two fixed point charges. The two remaining terms, the derivatives of the atomic charges q_i and q_j with respect to the coordinate α_k, could in principle be obtained from finite differences. However, this approach would require a significantly larger number of reference DFT calculations making the construction of the NN PES very costly. We have thus chosen a more efficient approach by separating the fitting of the short-range and the long-range electrostatic part of the NN potential. First, only the atomic NNs representing the charges are fitted. For this fit the derivatives of the charges with respect to the atomic positions do not need to be known. Having established a functional relation between the atomic charges and the atomic positions in this way, the analytic derivatives $\partial q_i / \partial \alpha_k$ are available. The electrostatic NN is then employed to compute the electrostatic energy and force contributions for each structure according to Eq. (7). The required derivatives of the charges with respect to the atomic positions are given by the NN architecture and the definition of the

employed symmetry functions,

$$\frac{\partial q_i}{\partial \alpha_k} = \sum_{j=1}^{M_i} \frac{\partial q_i}{\partial G_{i,j}} \frac{\partial G_{i,j}}{\partial \alpha_k}. \quad (7)$$

Finally, after subtracting the electrostatic energy and force components from the DFT reference total energies and forces the short-range NN is trained in order to reproduce the remaining short-range part of the potential.

The explicit inclusion of long-range electrostatic interactions improves the description of the PES for large interatomic distances, as the electrostatic part of the potential is not truncated. However, for short interatomic distances the corrugation of the short range part of the potential is increased with respect to the rather smooth DFT total energy surface. The problem is that the electrostatic part of the potential contains cusps for small interatomic distances. Removing these cusps from the total energy surface introduces cusps of opposite sign in the short-range part. An accurate fit of the short range part of the PES therefore becomes significantly more challenging than without the explicit inclusion of electrostatics. This problem can be solved by screening the electrostatic interactions between close atoms. This is possible without loss of accuracy, because the short range part is able to represent all physical interactions within the cutoff radius regardless of their physical origin. We have thus introduced a screening function

$$f_{\text{screen}} = \begin{cases} \frac{1}{2}\left[1 - \cos\left(\frac{\pi R_{ij}}{R_s}\right)\right] & \text{for } R_{ij} \le R_s \\ 1 & \text{for } R_{ij} > R_s. \end{cases} \quad (8)$$

Multiplying the electrostatic energy contribution of each pair of charges by this function the electrostatic energy contribution is smoothly switched on until beyond the screening radius R_s the full electrostatic interaction is taken into account. For the electrostatic energy we obtain now

$$E_{\text{elec}}^{\text{screen}} = \sum_i \sum_{j>i} f_{\text{screen}} \frac{q_i q_j}{R_{ij}}, \quad (9)$$

and for the electrostatic force we have

$$\begin{aligned} F_{\alpha_k}^{\text{elec,screen}} = &- \sum_{j\neq k} q_k q_j \frac{\alpha_j - \alpha_k}{R_{kj}^3} \cdot f_{\text{screen}} \\ &- \sum_i \sum_{j>i} \frac{1}{R_{ij}} \left(q_j \frac{\partial q_i}{\partial \alpha_k} + q_i \frac{\partial q_j}{\partial \alpha_k} \right) \cdot f_{\text{screen}} \\ &- \sum_{j\neq k} \frac{q_k q_j}{R_{kj}} \frac{1}{2} \sin\left(\frac{\pi R_{kj}}{R_s}\right) \frac{\pi}{R_s} \frac{\alpha_k - \alpha_j}{R_{kj}}. \end{aligned} \quad (10)$$

Utilizing a screening function, the short range part is much smoother, and the numerical accuracy of the NN fit is improved. The screening function, the electrostatic energy of two point charges $q_1 = e$ and $q_2 = -e$ and the screened electrostatic energy of these charges are plotted in Fig. 4.

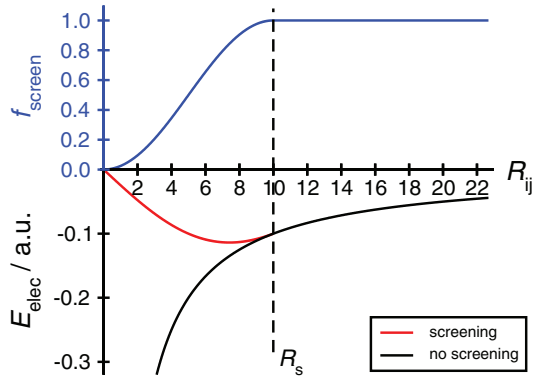

FIG. 4. Plot of the screening function f_{screen} defined in Eq. (8) (upper panel), and of the electrostatic energy E_{elec} of a pair of point charges $q_1 = e$ and $q_2 = -e$ with and without screening (lower panel). The screening range is defined by a screening radius R_s of 10 Bohr here.

III. RESULTS AND DISCUSSION

A. Reference calculations

As reference method for training the NN potential for the water dimer we have chosen DFT. The exchange-correlation energy has been described by the Perdew-Burke-Ernzerhof (PBE) functional.[58] All calculations have been carried out using the FHI-aims package[59] utilizing a basis set of numeric atomic orbitals. The basis set size and the integration grids have been carefully converged in order to achieve an error below 2 meV per monomer for energy differences. In principle, the use of atom-centered basis sets can give rise to basis set superposition errors (BSSEs) leading to artificially overestimated binding energies. We have confirmed that for the chosen basis set ("tier 2") the BSSE is less than 2 meV/H_2O and can safely be neglected. Therefore, no counterpoise correction has been applied. The reference atomic charges have been derived from the DFT calculations employing a Hirshfeld partitioning scheme.[60] We note that our approach is equally applicable to any other reference electronic structure method, and similar fitting results can be expected using wave-function-based methods, since the overall complexity of first-principles PESs is similar for different electronic structure methods, while only subtle details may differ.

B. Systematic construction of the training set

The NN potential for the water dimer has been constructed using our NN program RuNNer.[61] The atomic environments of the oxygen and hydrogen atoms have been defined by using a cutoff radius of 10 Å. For oxygen and hydrogen, 16 and 19 symmetry functions, have been used to describe the positions of the atoms in the chemical environment, respectively. For fitting the charges and short range atomic energy contributions, various NNs with different numbers of hidden layers and nodes per layer have been tested. The initial weight parameters have been initialized randomly and have then been normalized applying the scheme of Nguyen and Widrow.[62] Before subtracting the electrostatic energies and forces from the total DFT values they were screened employing a screening radius of $R_s = 5.3$ Å.

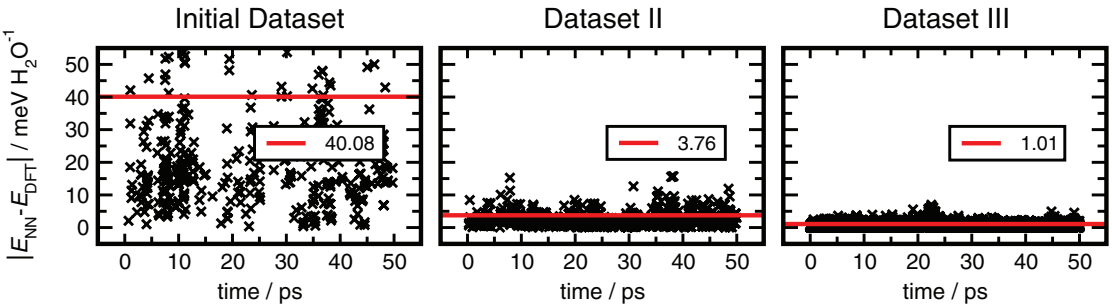

FIG. 5. Illustration of the systematic improvement of the neural network (NN) potential. Shown is the difference between the DFT and NN energies for structures obtained in molecular dynamics simulations at 300 K based on preliminary NN potentials. The red line displays the root mean squared error (RMSE) of the NN energies with respect to DFT in meV/H_2O. The NN potential fitted using the initial data set generates many structures with large deviations between the DFT and NN energies. If these structures are included in the training set to refine the fit, an improved NN potential is obtained ("Dataset II"), which is much closer to the DFT potential-energy surface and has a significantly reduced RMSE of 3.76 meV/H_2O. Including additional structures obtained using this NN potential, the RMSE can be further reduced to 1.01 meV/H_2O. Dataset II contains 375 structures and dataset III 1283 structures in addition to the initial dataset of 3852 structures.

An important aspect of constructing NN PESs is the choice of the reference configurations for the electronic structure calculations. For low-dimensional systems like a single water molecule, the PES can be mapped systematically for all degrees of freedom employing a dense grid of points. However, this soon becomes unfeasible with increasing dimensionality of the PES, and already for the 12-dimensional PES of the water dimer a more efficient scheme needs to be found. For instance, considering only five grid points per degree of freedom would lead to $5^{12} \approx 250\,000\,000$ structures, which have to be computed. Moreover, a systematic scan of the whole PES would produce a large number of high-energy structures, which are not accessible in the chemically relevant temperature range.

In order to focus the accuracy of the NN PES on the important part of the configuration space, we employ a self-consistent approach to construct the training set for the NN following in spirit the work of Collins[63] and Raff.[23] First, a preliminary NN potential is fitted to an initial set of structures. Additional relevant structures can then be found by applying this preliminary NN to suggest new configurations in MD simulations or geometry optimizations starting from random dimers. These new structures can then be recalculated by DFT and a structure is added to the training set if the NN energy error is significant larger then the average fitting error of the whole dataset. Then, the improved fit is used to generate more structures, which are again compared to DFT and so on.

Still, this approach can be very time-consuming, because the DFT energies need to be known for all points to decide if they should be added to the training set. The number of DFT calculations can be drastically reduced by identifying points in poorly represented regions of the configuration space using only the NN.[56] This can be done by first recalculating the energies of candidate structures employing further, independently fitted NN potentials, which have a similar quality as the potential used to suggest the structures. If all NN potentials predict very similar energies for a given structure, it is likely to be similar to a point already present in the training set. If the structure is far from any point used to construct the NN potentials, the predictions of the NN potentials will be very different because of the highly flexible functional form of the NNs. Such structures should then be selected for DFT calculations to systematically increase the training set until a NN potential is obtained, which is able to predict structures with the desired accuracy.

The water dimer configurations in the initial dataset are based on structures reported in the literature. 2378 structures with fixed monomer geometries were taken from the supporting information of Ref. 64. This dataset contains 2510 structures and was first used by Mas *et al.*[65] to construct an *ab initio* water dimer potential. After recomputing all structures by DFT, 132 high-energy structures with a water dimer binding energy larger than 1 eV have been removed from the dataset as at this energy the dimer would immediately dissociate. Another 1026 structures were derived from the 10 stationary points reported by Torheyden and Jansen[66] by varying the monomer-monomer distance. To ensure the correct behavior at the dissociation limit 448 monomer structures, generated by scanning the three-dimensional monomer PES on a regular grid, have been included. Employing these monomer and dimer structures, a first NN PES has been constructed, and the dataset has then been extended in the self-consistent way described above. For this purpose MD simulations have been performed using the MD program TINKER[67] and the NN potential. In Fig. 5 typical errors for structures visited in the MD simulations are shown for the preliminary potential and two improved training data sets. It can be seen that the root mean squared error (RMSE) for structures not included in the training set improves rapidly, from 40 meV/H_2O for the initial preliminary potential to only about 1 meV/H_2O in dataset III.

Apart from structures obtained from MD simulations also configurations close to the stationary points on the dimer PES have been included in the training dataset. To improve the intramolecular forces a new set of monomer structures, computed on a denser grid, was added. The final self-consistent dataset contains 5617 monomer and 12 006 dimer structures comprising 88 887 atomic charges and 266 661 force components. In general, for all fits the dataset have been split into a training set (90%) used to update the NN weights and an independent test set (10%) to check the transferability of the potential. We finally obtained the best fit for the atomic charges

TABLE I. Root mean squared errors (RMSE) and normalized root mean squared errors (nRMSE) of the energies (meV/H_2O), forces (meV/Å), and charges ($10^{-3}e$) of the water monomer and dimer structures in the training set. The test set values are given in brackets.

		RMSE	nRMSE
H_2O	Energies	0.44 (0.43)	0.18% (0.18%)
	Forces	3.02 (2.94)	0.15% (0.14%)
	Charges O	0.09 (0.08)	0.99% (0.94%)
	Charges H	0.10 (0.10)	1.12% (1.12%)
$(H_2O)_2$	Energies	1.05 (1.19)	1.39% (1.51%)
	Forces	10.14 (10.57)	1.62% (1.72%)
	Charges O	0.30 (0.42)	1.30% (1.86%)
	Charges H	0.30 (0.36)	2.06% (2.42%)

1meV = 0.023 kcal/mol

with two hidden layers and 60 nodes in each layer. The short range energies and forces are fitted by NNs with two hidden layers, each of which contains 70 nodes.

C. Accuracy of the neural network potential

The RMSE values of the energies, forces, and charges of the monomer and dimer configurations are shown in Table I. The DFT total energies are fitted with an extremely high accuracy of about 1 meV per monomer, the forces have an error of approximately 10 meV/Å, and the charges are reproduced to within 0.0003 e. The RMSE of structures in the training and the test set is basically the same, which indicates that the NN PES covers all relevant parts of the configuration space and the potential is also applicable to structures not included in the fitting process. The fitting error of the dimer energies of about 1 meV/H_2O is even smaller than the convergence level of the reference DFT calculations, which confirms that by replacing the DFT calculations by the NN PES simulation results would be essentially the same.

In order to compare the accuracy of quantities with different units, it is useful to investigate also the normalized RMSE (nRMSE). It is obtained by dividing the RMSE by the standard deviation σ of the data set,

$$\text{nRMSE} = \frac{\text{RMSE}}{\sigma}. \tag{11}$$

A nRMSE of 100% or larger means that there is no correlation at all between the predicted and the reference values. For a perfect fit with no deviation from the reference values the nRMSE is 0%. Table I shows that all quantities fitted by the NN have a similar small nRMSE with values of only a few percent or less.

Figure 6 further illustrates the very close agreement between DFT and NN dimer energies over the entire range of values for the training set (black symbols) as well as the test structures (red symbols). All points in the upper panel are very close to the line with a slope of 45° corresponding to a perfect correlation. In the bottom panel it can be seen that low energy configurations as well as distorted high-energy structures far away from the minimum are equally well represented.

Finally, we have investigated the effect of the screening of the electrostatic interactions for small interatomic distances. In Fig. 7 the dimer structures have been grouped according to the oxygen-oxygen distance R_{OO}. For each group of structures the RMSE of the total forces is shown with and without applying the screening function. It can be seen that for large R_{OO} the RMSE values are very similar, while for small R_{OO} the error of the forces without screening is about twice as large. This reflects the higher corrugation of the short-range part of the potential, if the cusps of the electrostatic interaction are not removed by the screening function. The same result has been found for the data set of monomer structures, which is also shown for comparison. Here, the atoms are necessarily very close, and the force RMSE is clearly larger without applying the screening function.

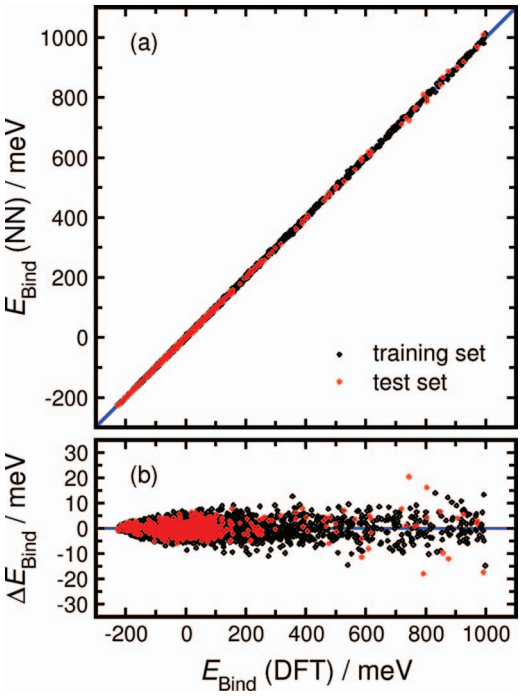

FIG. 6. Comparison of the neural network (NN) and DFT binding energies of the water dimer configurations in the training set (black diamonds) and the test set (red diamonds) obtained by removing two times the energy of the optimized water monomer (a). In panel (b) the binding energy difference between DFT and the NN potential is shown.

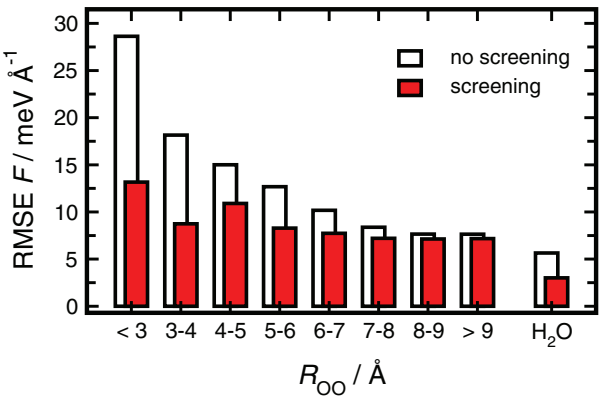

FIG. 7. Comparison of a fit with (red bars) and without (white bars) screening of the electrostatic interactions at small interatomic distances. The histogram shows the root mean squared errors (RMSE) of the forces averaged over dimer structures within several oxygen–oxygen distance intervals R_{OO}. Additionally, the RMSEs for the monomer structures are shown.

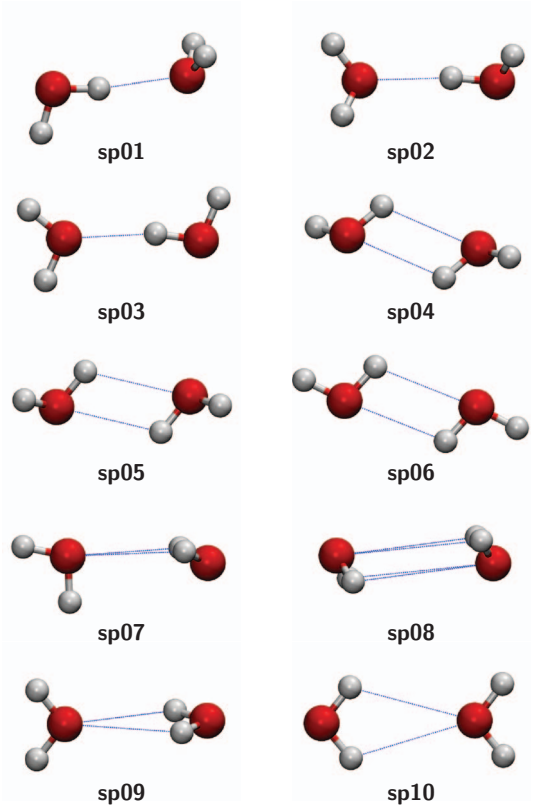

FIG. 8. Stationary points (sp) on the water dimer potential-energy surface according to Smith *et al.*[68]

D. Stationary points on the dimer potential-energy surface

In order to further investigate the quality of the NN potential, ten low-lying stationary points on the water dimer PES (shown in Fig. 8) have been located and analyzed. These structures, which comprise of the global minimum, transition states, and higher-order saddle points with up to three imaginary frequencies (cf. Table II), have been first reported by Smith *et al.*[68] Anderson and Tschumper have examined all stationary points with ten different exchange correlation functionals including the PBE functional by calculating harmonic vibrational frequencies.[69]

TABLE II. Minimum-energy oxygen-oxygen distances (Å) of 10 stationary points (sp) of the water dimer optimized with DFT and the neural network (NN) potential, respectively. N_i is the order of the stationary point. The numbers in parentheses show the absolute error of the NN potential with respect to DFT. The DFT values have been taken from Ref. 69.

sp no.	Symmetry	N_i	DFT	NN
1	C_s	0	2.898	2.897 (0.001)
2	C_1	1	2.916	2.914 (0.002)
3	C_s	2	2.924	2.924 (0.000)
4	C_i	1	2.848	2.853 (0.005)
5	C_2	1	2.848	2.847 (0.001)
6	C_{2h}	3	2.804	2.806 (0.002)
7	C_s	2	2.996	2.988 (0.008)
8	C_{2h}	3	3.330	3.336 (0.006)
9	C_{2v}	1	3.042	3.042 (0.000)
10	C_{2v}	2	3.247	3.247 (0.000)

TABLE III. Relative energies (meV) of ten stationary points (sp) of the water dimer optimized with DFT and the neural network (NN) potential. The DFT values have been taken from Ref. 69. The numbers in parentheses show the absolute error of the NN potential with respect to DFT. The NN energies are shown for fits with two different symmetry function cutoff radii of $R_c = 6$ Å and $R_c = 10$ Å, respectively.

sp no.	DFT	NN (10 Å)	NN (6 Å)
1	0.0	0.0 (–)	0.0 (–)
2	28.6	28.8 (0.2)	26.8 (1.8)
3	36.4	35.3 (1.1)	29.4 (7.0)
4	46.0	45.5 (0.5)	44.4 (1.6)
5	61.1	60.8 (0.3)	58.0 (3.1)
6	75.9	76.6 (0.7)	74.3 (1.6)
7	98.0	97.7 (0.3)	96.0 (2.0)
8	162.6	162.0 (0.6)	159.3 (3.3)
9	91.9	92.7 (0.8)	93.9 (2.0)
10	133.5	133.9 (0.4)	127.7 (5.8)

The stationary points on the PES have been located by performing a transformation of the Cartesian coordinates and forces into internal coordinates and forces. The water dimer PES can be defined by five Euler angles[70] describing the orientation of the monomers with respect to each other, the oxygen-oxygen distance and the monomer angle and bonds. Each structure was then optimized using the L-BFGS algorithm,[71] while keeping those Euler angles fixed, which correspond to the particular symmetry of the

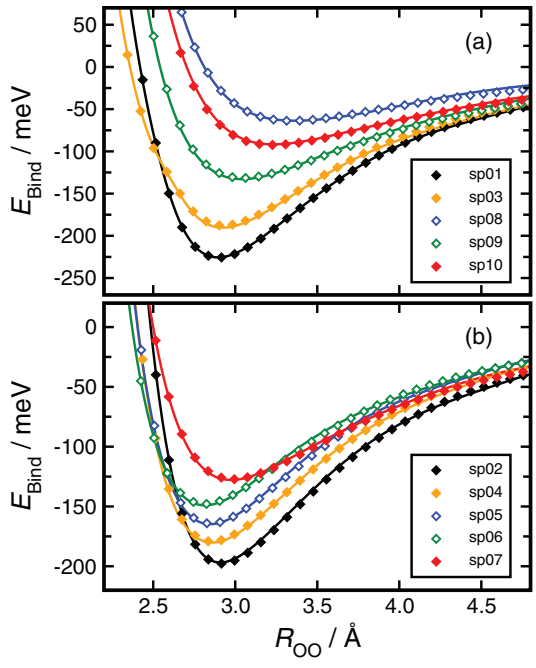

FIG. 9. Binding energy as a function of the oxygen-oxygen distance R_{OO} for ten stationary points (sp) on the water dimer potential-energy surface. The curves have been obtained by relaxing the structures with the neural network (NN) potential. The NN values (lines) have been recomputed by DFT (diamonds). The structures in plot (a) have been fully relaxed, in plot (b) only the monomer degrees of freedom have been optimized for given R_{OO} values to maintain the symmetry of the stationary points.

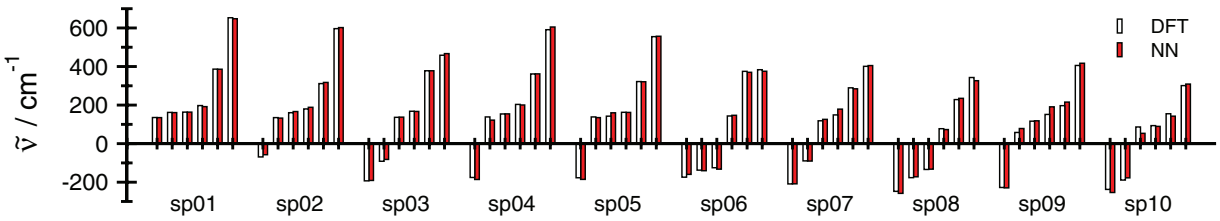

FIG. 10. Comparison of the vibrational frequencies (modes 1–6) of ten stationary points (sp) on the DFT and the neural network (NN) potential-energy surface. Imaginary frequencies are represented by negative numbers. The DFT values have been taken from Ref. 69.

stationary point. The resulting geometries are very close to those reported by Anderson and Tschumper. The error of the oxygen-oxygen distance is in the order of 10^{-3} Å, as reported in Table II. Table III shows the relative binding energies of all investigated geometries. The NN potential is able to reproduce the correct energetic ordering of the structures and the binding energies are within 1 meV or less of the DFT values for all stationary points. The absolute binding energy of the global minimum is –225.6 meV, compared to –225.0 meV obtained with DFT. Additionally, we have investigated the effect of a reduced cutoff radius of the symmetry functions of 6 Å. The results are also given in Table III. Even with this reduced cutoff still a reasonable agreement with DFT could be obtained.

Further, we have examined the accuracy of the NN PES for the stationary points if the oxygen-oxygen distance is modified. Figure 9 shows the binding energy curves as a function of the intermonomer distance obtained by relaxing all stationary points at a fixed oxygen-oxygen distance using the NN potential and recomputing the resulting structures by DFT. Although none of these DFT points has been used for training the NN PES, the NN values are in almost perfect agreement with the recomputed DFT energies.

Besides a reliable description of the binding energies also the gradient of the PES should be accurate, since the atomic forces are the key quantity in MD simulations. Vibrational frequencies strongly depend on the curvature of the PES. The comparison of vibrational frequencies therefore represents a sensitive test for the capability of the NN potential to reproduce the reference DFT forces. Below, we compare the vibrational frequencies of all ten stationary points computed with the NN potential with the PBE-DFT values provided by Anderson and Tschumper in the supporting information of Ref. 69. The harmonic vibrational frequencies have been determined by diagonalizing the Hessian matrix. Its matrix elements were obtained via finite differences of gradients since

TABLE IV. Vibrational frequencies (cm^{-1}) of ten stationary points (sp) on the DFT and the neural network (NN) potential-energy surface of the water dimer (MRE = mean relative error of the NN frequencies with respect to the DFT values). The DFT values have been taken from Ref. 69.

Mode	sp01	sp02	sp03	sp04	sp05	sp06	sp07	sp08	sp09	sp10	
1	136	69i	193i	176i	177i	174i	209i	248i	228i	238i	DFT
	135	57i	191i	186i	185i	159i	208i	258i	229i	254i	NN
2	162	135	91i	138	138	138i	89i	177i	58	189i	DFT
	161	133	82i	122	135	140i	90i	172i	78	178i	NN
3	164	161	136	154	142	125i	119	134i	117	86	DFT
	164	167	138	154	160	132i	126	132i	119	53	NN
4	198	180	168	204	163	143	149	78	151	93	DFT
	192	188	167	200	162	147	179	73	191	89	NN
5	386	311	378	362	322	375	289	228	197	155	DFT
	386	318	378	362	321	370	284	235	216	143	NN
6	653	597	459	591	555	384	401	343	406	301	DFT
	648	602	467	605	557	375	405	326	417	309	NN
7	1590	1585	1589	1588	1580	1575	1583	1599	1585	1589	DFT
	1600	1595	1597	1597	1573	1568	1582	1605	1581	1588	NN
8	1610	1609	1617	1590	1608	1598	1603	1600	1609	1606	DFT
	1613	1612	1620	1598	1611	1611	1617	1606	1614	1608	NN
9	3548	3581	3608	3656	3665	3694	3704	3708	3707	3710	DFT
	3559	3583	3611	3657	3656	3691	3706	3705	3710	3710	NN
10	3702	3711	3720	3672	3678	3697	3708	3710	3714	3719	DFT
	3701	3711	3722	3674	3668	3700	3710	3716	3719	3717	NN
11	3784	3786	3792	3793	3798	3815	3793	3806	3795	3806	DFT
	3779	3784	3788	3783	3787	3803	3776	3797	3778	3798	NN
12	3806	3818	3828	3793	3799	3815	3814	3808	3814	3816	DFT
	3799	3812	3823	3786	3792	3813	3812	3798	3802	3802	NN
MRE	0.5%	2.6%	1.3%	2.0%	1.8%	2.0%	2.7%	2.0%	6.4%	5.6%	

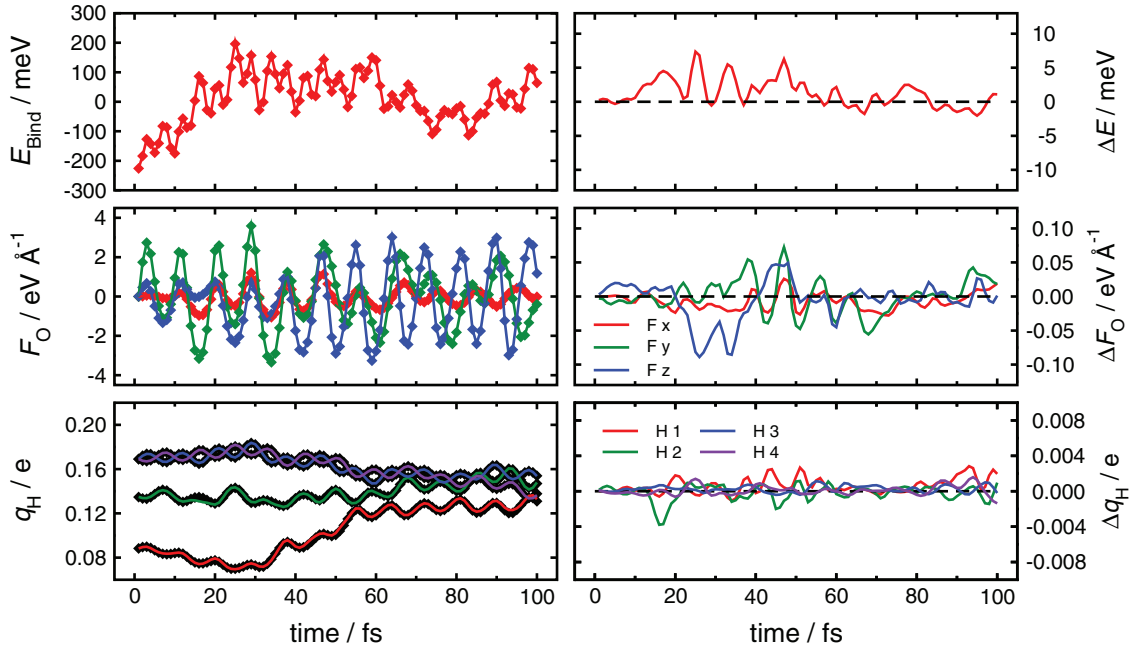

FIG. 11. Comparison of DFT (diamonds) and neural network (lines) energies, forces, and charges along an *ab initio* molecular dynamics trajectory of a water dimer at 300 K. The binding energy E_{Bind} is the interaction energy of the two water monomers obtained by removing twice the energy of the optimized monomer. The agreement of the forces is illustrated using the components of the force acting on the oxygen atom of the first monomer. The charges are compared for the four hydrogen atoms. The left panels show the values, the right panels show the difference between the NN and DFT values.

analytic second derivatives, although in principle accessible, have not yet been implemented in the RuNNer-Code. Figure 10 shows the six lowest vibrational modes for all structures obtained with the NN potential and the DFT values. The corresponding values are listed in Table IV. The number of imaginary frequencies is consistent with DFT and all saddle points have the right order. For the global minimum, frequencies deviate on average by about 0.5% from the DFT values. Also higher stationary points show a good agreement with errors of a few percent at most.

E. Molecular dynamics

As a final test, we have carried out molecular dynamics simulations for the water dimer. As even tiny numerical differences between the NN and DFT forces inevitably result in different trajectories after some simulation time, it is not possible to directly compare trajectories run independently using the NN PES and DFT. Still, both the NN potential and DFT yield qualitatively the same results. Since many physical quantities can be determined as ensemble averages from long molecular dynamics trajectories, it is important to confirm that the structures visited in NN-based and DFT-based trajectories are similar.

Figure 11 shows the first 100 fs of an *ab initio* MD trajectory of the water dimer in the canonical ensemble (NVT) at 300 K. The DFT binding energies of the dimer, the force acting on the oxygen atom of one of the molecules, and the Hirshfeld charges of the four hydrogen atoms are plotted as diamond symbols in the left panels. These quantities have then been recalculated by the NN and they are shown as lines in the same panels. The agreement is excellent. For a closer investigation, also the differences between the NN and DFT numbers are shown in the right panels of Fig. 11. Similar tests have also been carried out for the opposite order, i.e., recalculating the DFT energies, forces, and charges for trajectories based on the NN potential. Also in this case results of the same quality have been found.

IV. CONCLUSIONS

In summary, we have reported the first full-dimensional NN potential for the water dimer. It has been constructed employing the high-dimensional NN approach of Behler and Parrinello to describe the short-range atomic interactions. Further, the long-range electrostatic energy contributions have been considered explicitly by using environment-dependent atomic charges. We have found that the NN potential is able to represent very accurately the underlying reference PBE-DFT energies, forces, and charges. The NN is able to predict these quantities with about the same quality also for structures not included in the training set. Binding energies and vibrational frequencies of 10 stationary points on the NN PES are in close agreement with published DFT data.

The NN potential is reactive, i.e., the energies, forces, and charges are given as a function of the atomic positions only. No classification of atoms is required and no bonds need to be specified. The potential is thus in principle able to describe the dissociation and recombination of water molecules. The method is fully general and not restricted to the water dimer, but is equally applicable also to larger water clusters and liquid water. However, further reference calculations are needed to extend the range of validity of the potential in case of larger systems. Work in this direction is currently in progress. It is

well-known that DFT has a limited accuracy for the description of water. Using the presented scheme, similar NN potentials can also be constructed using higher-level quantum chemical methods.

ACKNOWLEDGMENTS

We thank the Deutsche Forschungsgemeinschaft (DFG) (Emmy Noether program), the Fonds der Chemischen Industrie, and the Research Department Interfacial Systems Chemistry (IFSC) of the Ruhr-Universität Bochum for financial support. Discussions with the FHI-aims team are gratefully acknowledged.

[1] F. Franks, in *Water: A Matrix of Life*, 2nd ed. (Royal Society of Chemistry, Cambridge, 2000).
[2] M. P. Allen and D. J. Tildesley, in *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
[3] B. Guillot, J. Mol. Liq. **101**, 219 (2002).
[4] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[5] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[6] M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).
[7] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[8] D. Marx and J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods* (Cambridge University Press, Cambridge, 2009).
[9] G. C. Schatz, Rev. Mod. Phys. **61**, 669 (1989).
[10] T. Hollebeek, T. S. Ho, and H. Rabitz, Annu. Rev. Phys. Chem. **50**, 537 (1999).
[11] O. Engkvist, P. O. Astrand, and G. Karlstrom, Chem. Rev. **100**, 4087 (2000).
[12] K. Szalewicz, C. Leforestier, and A. van der Avoird, Chem. Phys. Lett. **482**, 1 (2009).
[13] Y. M. Wang and J. M. Bowman, Chem. Phys. Lett. **491**, 1 (2010).
[14] J. C. Fogarty, H. M. Aktulga, A. Y. Grama, A. C. T. van Duin, and S. A. Pandit, J. Chem. Phys. **132**, 174704 (2010).
[15] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation* (Addison-Wesley, Reading, 1996).
[16] C. M. Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press, Oxford, 1995).
[17] C. M. Handley and P. L. A. Popelier, J. Phys. Chem. A **114**, 3371 (2010).
[18] J. Behler, Chem. Modelling **7**, 1 (2010).
[19] J. Behler, Phys. Chem. Chem. Phys. **13**, 17930 (2011).
[20] K. Hornik, Neural Networks **4**, 251 (1991).
[21] F. V. Prudente, P. H. Acioli, and J. J. S. Neto, J. Chem. Phys. **109**, 8801 (1998).
[22] D. F. R. Brown, M. N. Gibbs, and D. C. Clary, J. Chem. Phys. **105**, 7597 (1996).
[23] L. M. Raff, M. Malshe, M. Hagan, D. I. Doughan, M. G. Rockley, and R. Komanduri, J. Chem. Phys. **122**, 084104 (2005).
[24] S. Manzhos and T. Carrington, Jr., J. Chem. Phys. **125**, 194105 (2006).
[25] S. Manzhos and T. Carrington, Jr., J. Chem. Phys. **129**, 224104 (2008).
[26] A. Pukrittayakamee, M. Malshe, M. Hagan, L. M. Raff, R. Narulkar, S. Bukkapatnum, and R. Komanduri, J. Chem. Phys. **130**, 134101 (2009).
[27] M. Malshe, L. M. Raff, M. G. Rockley, M. Hagan, P. M. Agrawal, and R. Komanduri, J. Chem. Phys. **127**, 134105 (2007).
[28] T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, J. Chem. Phys. **103**, 4129 (1995).
[29] S. Lorenz, A. Groß, and M. Scheffler, Chem. Phys. Lett. **395**, 210 (2004).
[30] S. Lorenz, M. Scheffler, and A. Groß, Phys. Rev. B **73**, 115431 (2006).
[31] J. Behler, B. Delley, S. Lorenz, K. Reuter, and M. Scheffler, Phys. Rev. Lett. **94**, 036104 (2005).
[32] J. Behler, S. Lorenz, and K. Reuter, J. Chem. Phys. **127**, 014705 (2007).
[33] C. Carbogno, J. Behler, A. Groß, and K. Reuter, Phys. Rev. Lett. **101**, 096104 (2008).
[34] C. Carbogno, J. Behler, K. Reuter, and A. Groß, Phys. Rev. B **81**, 035410 (2010).
[35] J. Behler, K. Reuter, and M. Scheffler, Phys. Rev. B **77**, 115421 (2008).
[36] D. A. R. S. Latino, R. P. S. Fartaria, F. F. M. Freitas, J. Aires-De-Sousa, and F. M. S. S. Fernandes, J. Electroanal. Chem. **624**, 109 (2008).
[37] J. Ludwig and D. G. Vlachos, J. Chem. Phys. **127**, 154716 (2007).
[38] S. Manzhos, X. G. Wang, R. Dawes, and T. Carrington, Jr., J. Phys. Chem. A **110**, 5295 (2006).
[39] H. Gassner, M. Probst, A. Lauenstein, and K. Hermansson, J. Phys. Chem. A **102**, 4596 (1998).
[40] K. T. No, B. H. Chang, S. Y. Kim, M. S. Jhon, and H. A. Scheraga, Chem. Phys. Lett. **271**, 152 (1997).
[41] K. H. Cho, K. T. No, and H. A. Scheraga, J. Mol. Struct. **641**, 77 (2002).
[42] C. M. Handley and P. L.A. Popelier, J. Chem. Theory Comput. **5**, 1474 (2009).
[43] C. M. Handley, G. I. Hawe, D. B. Kell, and P. L. A. Popelier, Phys. Chem. Chem. Phys. **11**, 6365 (2009).
[44] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
[45] N. Artrith, T. Morawietz, and J. Behler, Phys. Rev. B **83**, 153101 (2011).
[46] A. Bholoa, S. D. Kenny, and R. Smith, Nucl. Instrum. Methods Phys. Res. B **255**, 1 (2007).
[47] E. Sanville, A. Bholoa, R. Smith, and S. D. Kenny, J. Phys.: Condens. Matter **20**, 285219 (2008).
[48] J. Behler, J. Chem. Phys. **134**, 074106 (2011).
[49] J. B. Witkoskie and D. J. Doren, J. Chem. Theory Comput. **1**, 14 (2005).
[50] H. M. Le and L. M. Raff, J. Phys. Chem. A **114**, 45 (2010).
[51] J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, Phys. Rev. Lett. **100**, 185501 (2008).
[52] J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, Phys. Status Solidi B **245**, 2618 (2008).
[53] H. Eshet, R. Z. Khaliullin, T. D. Kühne, J. Behler, and M. Parrinello, Phys. Rev. B **81**, 184107 (2010).
[54] R. Z. Khaliullin, H. Eshet, T. D. Kühne, J. Behler, and M. Parrinello, Phys. Rev. B **81**, 100103 (2010).
[55] R. Z. Khaliullin, H. Eshet, T. D. Kühne, J. Behler, and M. Parrinello, Nature Mater. **10**, 693 (2011).
[56] N. Artrith and J. Behler, Phys. Rev. B **85**, 045439 (2012).
[57] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[58] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[59] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comput. Phys. Commun. **180**, 2175 (2009).
[60] F. L. Hirshfeld, Theor. Chim. Acta **44**, 129 (1977).
[61] J. Behler, RuNNer - A Neural Network Code for High-Dimensional Potential-Energy Surfaces, Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum.
[62] D. H. Nguyen and B. Widrow, IEEE Control Syst. Mag. **3**, 18 (1990).
[63] J. Ischtwan and M. A. Collins, J. Chem. Phys. **100**, 8080 (1994).
[64] R. Bukowski, K. Szalewicz, G. C. Groenenboom, and A. van der Avoird, J. Chem. Phys. **128**, 094313 (2008).
[65] E. M. Mas, R. Bukowski, K. Szalewicz, G. C. Groenenboom, P. E. S. Wormer, and A. van der Avoird, J. Chem. Phys. **113**, 6687 (2000).
[66] M. Torheyden and G. Jansen, Mol. Phys. **104**, 2101 (2006).
[67] J. W. Ponder, TINKER, Software Tools for Molecular Design version 5.0, Washington University School of Medicine, Saint Louis, MO, 2009, http://dasher.wustl.edu/tinker.
[68] B. J. Smith, D. J. Swanton, J. A. Pople, H. F. Schaefer, and L. Radom, J. Chem. Phys. **92**, 1240 (1990).
[69] J. A. Anderson and G. S. Tschumper, J. Phys. Chem. A **110**, 7268 (2006).
[70] M. Bouten, Physica **42**, 572 (1969).
[71] D. C. Liu and J. Nocedal, Math. Program. **45**, 503 (1989).