# Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces

Jörg Behler and Michele Parrinello

*Department of Chemistry and Applied Biosciences, ETH Zurich, USI-Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland*

(Received 27 September 2006; published 2 April 2007)

The accurate description of chemical processes often requires the use of computationally demanding methods like density-functional theory (DFT), making long simulations of large systems unfeasible. In this Letter we introduce a new kind of neural-network representation of DFT potential-energy surfaces, which provides the energy and forces as a function of all atomic positions in systems of arbitrary size and is several orders of magnitude faster than DFT. The high accuracy of the method is demonstrated for bulk silicon and compared with empirical potentials and DFT. The method is general and can be applied to all types of periodic and nonperiodic systems.

DOI: 10.1103/PhysRevLett.98.146401 PACS numbers: 71.15.Pd, 61.50.Ah, 82.20.Kh

The reliability of molecular dynamics (MD) or Monte Carlo (MC) simulations depends crucially on the accuracy of the underlying potential-energy surface (PES). *Ab initio* methods based on density-functional theory [1] (DFT) provide accurate PESs for many systems, but they are computationally very demanding and even on the most advanced platforms *ab initio* MD simulations are limited to tens of picoseconds and a few thousand atoms. This is the reason for the continuing popularity of empirical potentials which provide fast access to energy and forces. However the construction of reliable empirical potentials is a difficult and lengthy process which usually relies on fitting the parameters of a guessed, physically motivated simple functional form for the interaction potential. This can lead to qualitatively wrong results when used in circumstances in which the assumed functional form is not appropriate. The database used in the fitting can include experimental or theoretical data and even the forces obtained in an *ab initio* MD run [2–4].

In this Letter we present a generalized neural-network (NN) method for constructing DFT-based PESs which have *ab initio* accuracy and are capable of describing all types of bonding. The method overcomes the limitations that have so far restricted the use of NNs to low-dimensional PESs [5,6]. This is achieved by combining NN precision and flexibility with a PES representation that is inspired by empirical potentials. The resulting many-body potentials are a function of all atomic coordinates and can be used in systems of arbitrary size. We apply our ideas to the construction of an NN-based many-body potential for bulk silicon. Constructing an empirical potential for Si that is valid across the phase diagram has proven to be a frustrating challenge for conventional empirical potentials. Our potential works well in the solid semiconducting and in the liquid metallic phases. In addition we can reproduce the small energy differences between the different high-pressure phases of crystalline Si.

Neural networks are biology-inspired algorithms that provide an accurate tool for the representation of arbitrary functions. Given a number of points in which the value of the function is known, the parameters of the NN are optimized in order to reproduce the input data in a "training" process and then used to evaluate the function elsewhere. For the representation of PESs DFT calculations are generally used to provide the training data set. Once trained, the atomic coordinates are given to the NN and the potential energy, from which also forces can be calculated analytically, is received [5,6].

The structure of a simple NN as it has hitherto been used to represent PESs is shown schematically in Fig. 1 for a two-dimensional PES. In the nodes of the input layer the two generalized coordinates $G_i^1$ and $G_i^2$ that determine the energy of configuration $i$ are provided. The node in the output layer yields the associated energy $E_i$. In between the

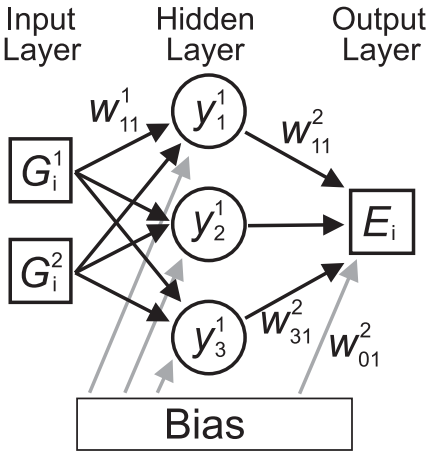


FIG. 1. Example of a standard neural network employed for fitting potential-energy surfaces [5,6]. The node in the output layer yields the energy $E_i$, which in this case depends on the values of the two input nodes, $G_i^1$ and $G_i^2$. In between the input and the output layer there is a hidden layer with three nodes represented by the circles. The arrows correspond to the 13 weight parameters $w_{ij}^k$, which connect node $j$ in layer $k$ with node $i$ in layer $k-1$. The bias node is used to adapt the nonlinearity region of the activation functions. The functional form of this small network is given in Eq. (1).

 

input and the output layer are one or more "hidden layers," each with a certain number of nodes. All nodes in each layer are connected to the nodes in the adjacent layers by real-valued weight parameters, which initially are chosen randomly. For a given set of coordinates the output of the NN is then given by the expression

$$E_i = f_a^2\left[w_{01}^2 + \sum_{j=1}^{3} w_{j1}^2 f_a^1\left(w_{0j}^1 + \sum_{\mu=1}^{2} w_{\mu j}^1 G_i^\mu\right)\right]. \quad (1)$$

Here, $w_{ij}^k$ is the weight parameter connecting node $j$ in layer $k$ with node $i$ in layer $k-1$, and $w_{0j}^k$ is a bias weight that is used as an adjustable offset for the activation functions $f_a^k$. Activation functions are typically nonlinear functions that introduce the capability to fit nonlinear functions into the NN [5,6]. In the present work the hyperbolic tangent has been used as an activation function in the hidden layers, and a linear function for the output layer. Since the weight parameters initially are chosen randomly, the output of the NN does not correspond to the correct total energy, but since the latter is known for a set of points from DFT calculations, an error function can be constructed and minimized to optimize the weight parameters in an iterative way. The optimized set of weights obtained can then be used to calculate the potential energy for a new set of coordinates.

This NN structure has several disadvantages that hinder its application to high-dimensional PESs. Since all weights are generally different, the order in which the coordinates of a configuration are fed into the NN is not arbitrary, and interchanging the coordinates of two atoms will change the total energy even if the two atoms are of the same type. Another limitation related to the fixed structure of the network is the fact that a NN optimized for a certain number of degrees of freedom, i.e., number of atoms, cannot be used to predict energies for a different system size, since the optimized weights are valid only for a fixed number of input nodes. Thus, in order to represent PESs useful for all system sizes, a new NN topology has to be introduced.

The main idea is to represent the total energy $E$ of the system as a sum of atomic contributions $E_i$, an approach that is typically also used in empirical potentials

$$E = \sum_i E_i. \quad (2)$$

The general structure of this new network topology is shown schematically in Fig. 2 for a system consisting of three atoms and all associated degrees of freedom. The $\{R_i^\alpha\}$ represent the Cartesian coordinates $\alpha$ of atom $i$. In a first step these coordinates are transformed into a set of symmetry function values $\{G_i^\mu\}$ for each atom $i$. These symmetry function values describe the energetically relevant local environment of each atom and are subsequently used as input for the NN. They depend on the positions of all atoms in the system, as indicated by the dotted arrows.

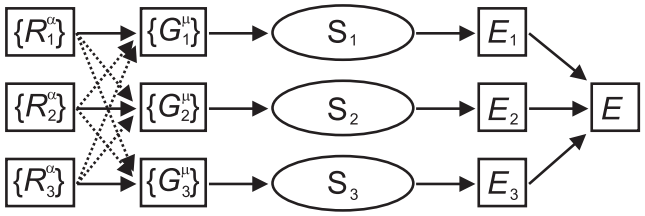


FIG. 2. Structure of the neural network as applied in this Letter to a system containing three atoms. The Cartesian coordinates of atom $i$ are given by $R_i^\alpha$. These are transformed to a set of $\mu$ symmetry function values $G_i^\mu$ describing the local geometric environment of atom $i$, which depends on the positions of all atoms in the system as indicated by the dotted arrows. The symmetry function values of atom $i$ then enter the subnet $S_i$ yielding the energy contribution $E_i$ of atom $i$ to the total energy of the system $E$. The structure of the subnets corresponds to the neural network shown in Fig. 1.

For each atom in the system there is now a "standard" NN (cf. Fig. 1), which we call subnet $S_i$ and which after the weight optimization yields the energy contribution $E_i$ to the total energy $E$. Summing these energy contributions then finally yields the total energy of the system. To ensure the invariance of the total energy with respect to the interchanging of two atoms the structure of all subnets and the values of the weight parameters are constrained to be identical in each $S_i$.

The crucial point is the introduction of a new type of symmetry function. While other types of symmetry functions have been used before [5], in our approach the symmetry function values of each atom reflect the local environment that determines its energy; i.e., two structures with different energies must yield different sets of symmetry function values, while identical local environments must give rise to the same set. Furthermore, the symmetry function values must be invariant with respect to a rotation or translation of the system. Finally, the number of symmetry functions must be independent of the coordination of the atom, because the coordination number of an atom can change in a MD simulation, while the structure of the subnets must not be changed if the NN is to remain applicable generally.

Symmetry functions can be constructed from atomic positions in a way similar to empirical potentials. But while in the latter case these terms are used to construct directly the total energy of the system, in the case of the NN they are used only to describe the structure. The assignment of the energies to the structures is done in a second step by the NN.

In order to define the energetically relevant local environment we employ a cutoff function $f_c$ of the interatomic distance $R_{ij}$, which has the form

$$f_c(R_{ij}) = \begin{cases} 0.5 \times \left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1\right] & \text{for } R_{ij} \le R_c, \\ 0 & \text{for } R_{ij} > R_c. \end{cases} \quad (3)$$

At interatomic separations larger than the cutoff $R_c$ this function yields zero value and slope. The cutoff has to be sufficiently large to include several nearest neighbors, and in the present Letter a cutoff of 6 Å has been used.

Radial symmetry functions are constructed as a sum of Gaussians with the parameters $\eta$ and $R_s$,

$$G_i^1 = \sum_{j\neq i}^{\text{all}} e^{-\eta(R_{ij}-R_s)^2} f_c(R_{ij}). \quad (4)$$

The summation over all neighbors $j$ ensures the independence of the coordination number.

Angular terms are constructed for all triplets of atoms by summing the cosine values of the angles $\theta_{ijk} = \frac{\mathbf{R_{ij}}\cdot\mathbf{R_{ik}}}{R_{ij}R_{ik}}$ centered at atom $i$, with $\mathbf{R_{ij}} = \mathbf{R_i} - \mathbf{R_j}$,

$$G_i^2 = 2^{1-\zeta} \sum_{j,k\neq i}^{\text{all}} (1 + \lambda\cos\theta_{ijk})^{\zeta} \times e^{-\eta(R_{ij}^2+R_{ik}^2+R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}), \quad (5)$$

with the parameters $\lambda$ $(= +1, -1)$, $\eta$, and $\zeta$. The multiplication by the three cutoff functions and by the Gaussian ensures a smooth decay to zero in the case of large interatomic separations. We note that the $G_i^\mu$ in Eqs. (4) and (5) depend on all atomic positions inside the cutoff radius and thus represent "many-body" terms. Several functions of each type with different parameter values are used. The choice of symmetry functions and their parameters is not unique nor does it need to be, and many types of functions can be used, as long as the set of function values is suitable for describing the environment of an atom.

To demonstrate the capability of the method we calculated the PES of bulk silicon using DFT in the local density approximation (LDA). The system used for the optimization of the NN parameters contains 64 atoms yielding 64 atomic environments per calculation. The calculations were carried out employing the plane-wave pseudopotential method as implemented in PWSCF [7]. A cutoff of 20 Ry was applied in combination with an ultrasoft pseudopotential [8]. A mesh of $3 \times 3 \times 3$ $k$ points was used. To improve the convergence of the metallic phases a Fermi smearing of 0.1 eV was employed.

Since the functional form of the NN has no physical motivation, the construction of an optimized NN requires special care. The structures used to train the NN [9] were initially taken from crystal structures including high-pressure phases [10] and MD simulations at different pressures and temperatures. Starting from this data set a series of fits was generated employing different NN topologies, i.e., numbers of hidden layers and nodes per hidden layer. The best fits can then be used to optimize the NN in a self-consistent way by performing MD, hybrid Monte Carlo [11,12], and metadynamics [13,14] runs based on these fits and subsequently recalculating several hundred representative structures with DFT. If the root mean square error (RMSE) is larger than the error of the fit, the DFT calculations are added to the training set and new fits determined, which are used to generate more structures, and so forth.

In total about 9000 DFT energies were calculated, 8200 of which were used for optimizing the NN and 800 as an independent test set to investigate the predictive capability of the NN for structures not included in the optimization set. The RMSE of the optimization set is typically 4–5 meV per atom, the RMSE of the test set 5–6 meV. For the NN atomic forces we found a RMSE of about 0.2 eV/Å with respect to DFT. The subnet employed consists typically of 2 hidden layers, each of which has about 40 nodes. In total 48 symmetry functions, i.e., input nodes, with different values of $\eta$, $R_s$, and $\zeta$ have been used resulting in a few thousand fitting parameters for the NN.

As a first test of the NN potential we calculated the energy vs volume curves for the different crystal structures of silicon [10]. It is well known that empirical potentials are not able to describe the correct energetic sequence of the various phases [15] while DFT is in good agreement with the experimental data [10]. The NN potential accurately reproduces the curves and the transition pressures of DFT. To test the ability of the NN potential to describe also disordered structures we calculated the radial distribution function (RDF) of a silicon melt at 3000 K. The result is shown in Fig. 3 and compared to other potentials of varying form and complexity [15–17]. The MD simulations were run for 20 ps (8 ps in the case of DFT [18]). The RDF obtained from the NN is very close to the DFT data, while there are significant deviations for the empirical potentials. The origin of the small difference between DFT and the NN is probably due to the fact that in the *ab initio* MD only the $\Gamma$ point has been used to sample the Brillouin zone,

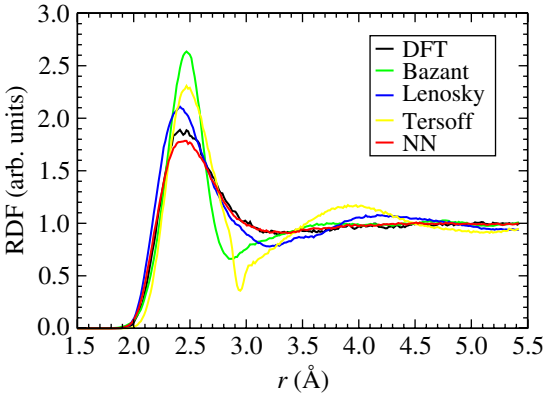


FIG. 3 (color online). Radial distribution function (RDF) of a silicon melt at 3000 K as obtained using a cubic 64 atom cell ($a = 20.526$ bohr). The curves shown were obtained from the Bazant [17,19], the Lenosky [15,19], the Tersoff [16,20], a neural network (NN) potential, and from density-functional theory (DFT) [18].

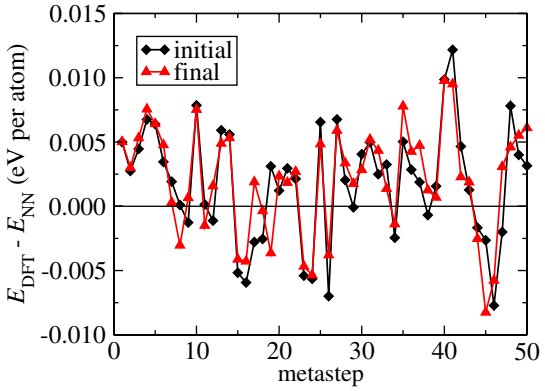


FIG. 4 (color online). Difference between the energies predicted by the neural network (NN) and recalculated energies obtained from density-functional theory (DFT) for the initial and final structures in each step of a metadynamics simulation [14] of bulk silicon. Each metastep involves a molecular dynamics simulation of 2 ps. The metadynamics simulation for the 64 atom cell starts from the $\beta$-tin structure with a pressure of 15 GPa at 300 K.

while the NN is trained to reproduce the energetics of a converged $k$ point mesh.

The accuracy of the NN energies has been checked in various way. Here we report the result of a metadynamics simulation [14] which involves strongly disordered nonequilibrium structures due to the change in shape and volume of the cell during the simulation. The details of the simulation are given in Fig. 4. The comparison of the NN predictions to recalculated DFT energies shows that in each metastep the potential is accurately described.

Compared with empirical potentials the number of DFT calculations required to optimize the NN parameters is rather large because of the very flexible functional form. This, however, has the advantage that no modifications to the NN are required if new DFT data are included. The accuracy of the NN is limited only by that of the training data. Here we have used DFT and an approximate exchange-correlation functional, but the NN is by no means restricted to fit DFT potentials. The method is general and can be applied to all types of periodic systems such as crystals, liquids, and surfaces as well as to nonperiodic systems. A limitation of the NN is the lack of extrapolation capability to structures very different from the structures included in the training set. We thus do not expect our current parametrization for bulk silicon to yield good results for silicon clusters. However, the potential can be systematically improved by properly extending the training set. Because of the cutoff applied to the symmetry functions, long-range interactions are not included in the present implementation. Inclusion of these effects is nevertheless possible by adding the corresponding potential terms. For a 64 atom system the NN is currently about 5 orders of magnitude faster than the DFT calculations, and in contrast to DFT the NN scales linearly with system size and is easily parallelized. An extension to multicomponent systems is straightforward, but requires the incorporation of the corresponding cross terms in the symmetry functions and an extended set of DFT energies.

In summary, we have introduced a fast way to represent high-dimensional PESs based on neural networks taking into account the positions of all atoms in systems of arbitrary size, which significantly extends the applicability of NNs to study chemical processes. The capability of the method, which is intended for long MD and Monte Carlo simulations of large systems, has been demonstrated for bulk silicon. For all phases examined, the accuracy of the description of all investigated properties is basically the same as that of the underlying DFT data and is superior to all tested empirical potentials.

We thank S. Lorenz, A. Groß, and M. Scheffler for providing their advanced neural-network code [5]. J. B. thanks S. Lorenz and M. Scheffler for sharing their experience with the NN methodology. The authors acknowledge computing time allocated by the CSCS and financial support by the DFG and the SNF (No. 200021-109673).

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[2] A. Laio *et al.*, Science **287**, 1027 (2000).
[3] F. Ercolessi and J. B. Adams, Europhys. Lett. **26**, 583 (1994).
[4] S. Izvekov *et al.*, J. Chem. Phys. **120**, 10 896 (2004).
[5] S. Lorenz, A. Groß, and M. Scheffler, Chem. Phys. Lett. **395**, 210 (2004).
[6] T. B. Blank *et al.*, J. Chem. Phys. **103**, 4129 (1995).
[7] S. Baroni *et al.*, http://www.pwscf.org
[8] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).
[9] The current code is an extended version of the NN code by S. Lorenz, A. Groß, and M. Scheffler.
[10] A. Mujica *et al.*, Rev. Mod. Phys. **75**, 863 (2003).
[11] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).
[12] M. E. Clamp *et al.*, J. Comput. Chem. **15**, 838 (1994).
[13] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. U.S.A. **99**, 12 562 (2002).
[14] R. Martoňák, A. Laio, and M. Parrinello, Phys. Rev. Lett. **90**, 075503 (2003).
[15] T. J. Lenosky *et al.*, Model. Simul. Mater. Sci. Eng. **8**, 825 (2000).
[16] J. Tersoff, Phys. Rev. B **38**, 9902 (1988).
[17] M. Z. Bazant, E. Kaxiras, and J. F. Justo, Phys. Rev. B **56**, 8542 (1997).
[18] T. D. Kühne, M. Krack, F. R. Mohamed, and M. Parrinello, Phys. Rev. Lett. **98**, 066401 (2007).
[19] S. Goedecker, Comput. Phys. Commun. **148**, 124 (2002).
[20] DL_POLY_2, Version 2.16, W. Smith *et al.*, CCLRC Daresbury Laboratory, 2006.