

## Neural network interatomic potential for laser-excited materials

Pascal Plettenberg <sup>1</sup>✉, Bernd Bauerhenne <sup>1</sup> & Martin E. Garcia <sup>1</sup>

Data-driven interatomic potentials based on machine-learning approaches have been increasingly used to perform large-scale, first-principles quality simulations of materials in the electronic ground state. However, they are not able to describe situations in which the electrons are excited, like in the case of material processing by means of femtosecond laser irradiation or ion bombardment. In this work, we propose a neural network interatomic potential with an explicit dependency on the electronic temperature. Taking silicon as an example, we demonstrate its capability of reproducing important physical properties with first-principles accuracy and use it to simulate laser-induced surface modifications on a thin film at time and length scales that are impossible to reach with first-principles approaches like density functional theory. The method is general and can be applied not only to other laser-excited materials but also to condensed and liquid matter under non-equilibrium situations in which electrons and ions exhibit different temperatures.

<sup>1</sup>Theoretical Physics and Center for Interdisciplinary Nanostructure Science and Technology (CINsaT), University of Kassel, Heinrich-Plett-Strasse 40, 34132 Kassel, Germany. ✉email: [plettenberg@uni-kassel.de](mailto:plettenberg@uni-kassel.de)

Interatomic potentials (or force fields) are essential quantities for describing the mechanical properties of materials with sub-nanometer scale resolution. Many numerical methods such as Monte-Carlo or molecular dynamics (MD) simulations, widely used in statistical physics, materials science and soft-matter studies, rely on interatomic potentials<sup>1,2</sup>. In particular, MD simulations of large atomic structures are crucial for many different research fields<sup>3,4</sup>. In order to calculate the atomic energies and forces during the simulation, first-principle (or ab-initio) methods like density functional theory (DFT) can be applied<sup>5</sup>. DFT describes the necessary underlying quantum mechanical effects very accurately, but simulations based on DFT are limited to at most 1000 atoms and to picosecond timescales<sup>6–8</sup>. On larger spatial and temporal scales, however, the atomic interactions need to be treated classically by using an interatomic potential, which is either constructed in a data-driven way or is derived from physical assumptions and approximations.

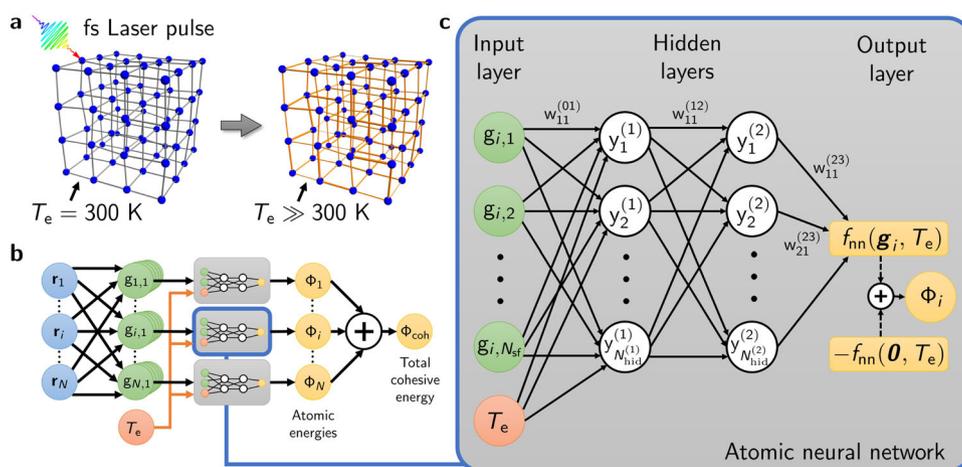
In the last years, interatomic potentials based on machine-learning (ML) methods like neural networks have become more and more popular. They allow to perform highly efficient atomistic simulations with first-principles accuracy and therefore have gained huge interest among researchers<sup>9–11</sup>. In general, ML potentials are used to perform large-scale simulations of materials that are assumed to be in the electronic ground state, which is a reasonable approximation for many thermodynamical processes at low temperatures (i.e., much lower than the Fermi temperature). However, this assumption fails in the case of laser-excited materials.

Ultrashort laser pulses with a pulse duration in the order of a few femtoseconds ( $1 \text{ fs} = 10^{-15} \text{ s}$ ) are essential for many fields such as biological imaging<sup>12</sup>, surgery<sup>13</sup>, or material processing<sup>14</sup>. Among others, they are used for surface texturing with applications in the microelectronics and photovoltaic industry<sup>15</sup>. A femtosecond laser pulse primarily interacts with the electrons and leaves the ions almost unaffected, leading to an extreme non-equilibrium state of hot electrons at an increased temperature  $T_e$  and “cold” ions (see Fig. 1a). The potential energy surface (PES) generated by the hot electrons differs significantly from the one resulting when electrons are at the ground state or very low temperatures. After the laser excitation, the electronic temperature decreases and equilibrates

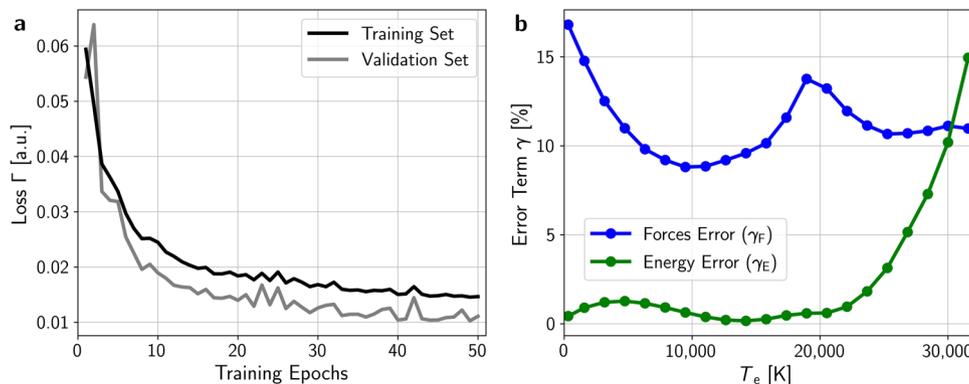
with the ionic temperature on a picosecond timescale due to the electron–phonon coupling. Therefore, interatomic potentials for simulations of laser-excited solids have to depend explicitly on the electronic temperature to account for both, changes in interatomic bonding and the electronic temperature decrease. Since standard ML potentials from the literature do not fulfill this requirement, they are not suitable for this specific task even if they are trained on a dataset generated at high  $T_e$ . In the past, interatomic potentials with an explicit  $T_e$ -dependency were only constructed empirically based on physical approximations<sup>16–18</sup>. Therefore, they lack the advantages of modern ML potentials like increased flexibility and, in particular, ab-initio accuracy.

Highly accurate  $T_e$ -dependent ML potentials could strongly increase our theoretical understanding of laser-excited materials. For example, they could be used to simulate ultrafast non-thermal phenomena like structural phase transitions<sup>19–22</sup> and transient phonon effects<sup>23–26</sup>, which result from bond softening, bond hardening, or lattice instabilities<sup>27–29</sup> caused by laser-induced changes to the PES. These effects occur within the first picoseconds after the laser excitation, when the electrons can still be treated as if their temperature was suddenly increased to a high constant value. On longer timescales, however, the  $T_e$ -dependency of the interatomic potential becomes even more important because the electronic temperature cannot be assumed to be constant anymore due to the electron–phonon coupling. Only with  $T_e$ -dependent interatomic potentials it can be possible to perform large- and ultralarge-scale MD simulations treating laser-induced bond changes and incoherent lattice heating via electron–phonon coupling on the same accuracy level<sup>30</sup>. Further utilizing the benefits of ML potentials could provide insights on interesting long-term effects like the spontaneous formation of laser-induced periodic surface structures<sup>31–33</sup>, which is still not fully understood.

In their first versions, early ML potentials based on neural networks could only be applied to small molecules due to the fast growth of the network size for larger systems and the challenging task of finding appropriate symmetric coordinates<sup>34–37</sup>. However, the introduction of atom-centered symmetry functions and the high-dimensional neural network architecture opened the possibility to handle arbitrarily large structures, in principal<sup>38,39</sup>. These



**Fig. 1 Modeling laser-excited materials with a  $T_e$ -dependent neural network interatomic potential.** **a** A femtosecond laser pulse excites the electrons of an atomic structure to high temperatures  $T_e \gg 300 \text{ K}$ . For a few picoseconds, the ions remain cold until the energy starts to transfer from the electrons to the lattice due to electron–phonon coupling. **b** The Cartesian coordinates of the atoms are described by symmetry functions  $\mathbf{g}_i$ , which are passed to the neural network potential together with the electronic temperature  $T_e$ . The network is evaluated individually for every atom in the structure and the results are summed to obtain the total cohesive energy  $\Phi_{\text{coh}}$ . **c** The atomic neural network in the case of two hidden layers. The input values are the  $N_{\text{sf}}$  symmetry function values for atom  $i$  and the electronic temperature  $T_e$ . The network function  $f_{\text{nn}}$  is also evaluated for  $\mathbf{g}_i = \mathbf{0}$  and subtracted from the output value to obtain the atomic cohesive energy  $\Phi_i$ .



**Fig. 2 Training curve and prediction errors.** **a** The loss function  $\Gamma$  evaluated on the training and validation set during the training process. **b** The relative prediction errors  $\gamma_E$  and  $\gamma_F$  for the total cohesive energy and the atomic forces as a function of the electronic temperature  $T_e$  averaged over all configurations occurring in the MD simulations from the training and validation set.

kinds of ML potentials were shown to reach ab-initio accuracy for many different systems<sup>40–42</sup> and in contrast to classical interatomic potentials, they do not require human experts to derive complex system-specific interaction terms.

Since the introduction of high-dimensional neural network potentials, several modifications to the model were suggested, e.g., extensions to multicomponent systems<sup>43,44</sup>, magnetic systems<sup>45</sup>, or the inclusion of non-local long-range interactions<sup>46–48</sup>. Furthermore, interatomic potentials based on other ML algorithms than neural networks were proposed, e.g., Gaussian approximation potentials<sup>49,50</sup>, spectral neighbor analysis potentials<sup>51,52</sup>, or moment tensor potentials<sup>53,54</sup>. Convolutional and graph neural networks have also been used to construct interatomic potentials<sup>55–58</sup>.

Despite these many modifications and extensions, ML potentials were still not able to describe laser-excited systems so far. In this work, we propose the first ML potential that includes a dependency on the electronic temperature  $T_e$  and is therefore applicable to large-scale simulations of materials excited by femtosecond laser-pulses. Our model, which is based on the high-dimensional neural network potential, is exemplarily applied to silicon (Si) but the method is general and could easily be transferred to other materials. We show that our derived interatomic potential can describe many important physical properties of laser-excited and unexcited Si with ab-initio accuracy and furthermore, the successful application of the model in MD simulations involving more than 160,000 Si atoms is demonstrated in this work.

## Results

**Neural network potential for laser-excited silicon.** We propose an interatomic potential that can be applied in MD simulations of laser-excited materials. Thereby, we describe the PES as a function  $\Phi(\mathbf{r}_1, \dots, \mathbf{r}_N, T_e)$  of the atomic positions  $\mathbf{r}_i$  and the electronic temperature  $T_e$ . Furthermore, we consider a configuration with  $N$  identical atoms. The total potential energy  $\Phi$  of this configuration can be split up into the total cohesive energy  $\Phi_{\text{coh}}$  and the Helmholtz free energy of the isolated atoms  $\Phi_0$ :

$$\Phi(\mathbf{r}_1, \dots, \mathbf{r}_N, T_e) = \Phi_{\text{coh}}(\mathbf{r}_1, \dots, \mathbf{r}_N, T_e) + N \cdot \Phi_0(T_e). \quad (1)$$

We designed a ML model that predicts  $\Phi_{\text{coh}}$  from  $\mathbf{r}_1, \dots, \mathbf{r}_N$ , and  $T_e$  and is able to learn this relationship from an ab-initio training dataset. As a useful example, we explicitly construct the model for laser-excited Si. In order to obtain the total potential energy  $\Phi$ , we calculate  $\Phi_0(T_e)$  using an existing polynomial model for Si from the literature (see “Methods”).

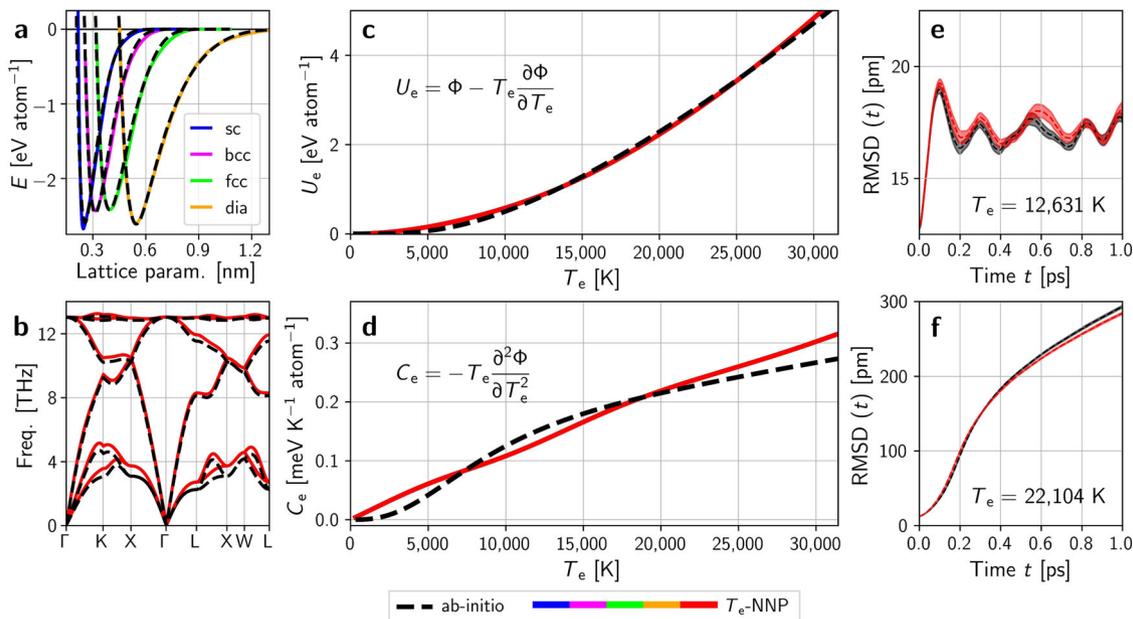
Our model for  $\Phi_{\text{coh}}$  is a  $T_e$ -dependent neural network interatomic potential ( $T_e$ -NNP). It is based on the high-dimensional neural

network architecture introduced by Behler and Parrinello in 2007<sup>38</sup>, in which the potential energy is treated as a sum of local atomic contributions. A feed-forward neural network is trained to predict the atomic energy from special symmetric coordinates describing the atomic environment. We extend this concept to the description of laser-excited systems by implementing a  $T_e$ -dependency in the form of an additional input node to the neural network, which is depicted schematically in Fig. 1b (high-dimensional neural network architecture) and Fig. 1c (atomic neural network).

Our aim is to train the model to predict the cohesive energy and the atomic forces for structures occurring in large-scale MD simulations of laser-excited Si with ab-initio accuracy. Therefore, we train the  $T_e$ -NNP on a specially designed dataset that consists of atomic configurations obtained from several ab-initio calculations of a small supercell containing a Si thin film. By choosing this geometry, we are not only taking into account bulk but also surface effects. Using  $T_e$ -dependent DFT, we perform MD simulations of the thin film at different constant electronic temperatures from a wide range of values. Furthermore, the training dataset contains artificial compressions and expansions of the structure to also include atomic configurations with extreme local pressures and densities. For all calculations, the Si atoms are initialized in the diamond-like crystal structure.

In order to determine the models generalization capability and to prevent overfitting, we also monitor its accuracy on a distinct validation dataset. It contains similar ab-initio calculations carried out at intermediate  $T_e$ 's between the  $T_e$ -steps of the training set (see “Methods” for details on the training and validation set). Figure 2a shows the evolution of the training and validation set error  $\Gamma$  (see Eq. (10)) during the training process. Despite some small fluctuations, the validation set error does not increase during the training process, which would be an indicator for overfitting.

**Performance validation.** As a first approach to validate the accuracy of the trained neural network interatomic potential, we analyze the relative prediction errors of the total cohesive energy and the atomic forces occurring during the MD simulations from the training and validation set as a function of the electronic temperature, as depicted in Fig. 2b. Thereby, we use the relative root-mean-square errors  $\gamma_E$  and  $\gamma_F$ , as defined in Eq. (10). The energy error is below 2% for  $T_e < 25,000$  K and only increases significantly for extremely high electronic temperatures far above the non-thermal melting threshold at  $T_e = 17,052$  K<sup>59</sup>. The relative prediction error for the atomic forces is between 8 and 12% at most temperatures and only increases at very low electronic



**Fig. 3 Physical properties reproduced by the  $T_e$ -dependent neural network interatomic potential.** **a** The Helmholtz free cohesive energy per atom at  $T_e = 12,631$  K as a function of the lattice parameter for different crystal structures. **b** The phonon band structure of bulk diamond Si at  $T_e = 12,631$  K. **c** The internal energy  $U_e$  of the electrons as a function of  $T_e$  for bulk diamond Si. **d** The specific heat  $C_e$  of the electrons as a function of  $T_e$  for bulk diamond Si. **e** The root-mean-square displacement (RMSD) of the Si atoms as a function of time after the laser excitation at  $T_e = 12,631$  K. **f** The root-mean-square displacement (RMSD) of the Si atoms as a function of time after the laser excitation at  $T_e = 22,104$  K. The width of the curves in **(e)** and **(f)** corresponds to twice the standard error of the mean over multiple simulation runs. For all figures, ab-initio results are shown in black, while results obtained with the  $T_e$ -NNP are colored.

temperatures as well as in the range slightly above the non-thermal melting threshold.

Although measuring the performance of the model in terms of the energy and forces errors is a reasonable approach, low prediction errors do not guarantee a physically meaningful description of the PES. Therefore, we also evaluate the models performance by calculating different atomic and electronic properties and comparing the results to ab-initio data in a qualitative way. First, we calculate the cohesive energy per atom as a function of the lattice parameter for different ideal crystal structures. We perform these calculations at various  $T_e$ 's and on configurations similar to those in the training dataset (diamond-like structure) as well as on configurations not contained in the training dataset (sc, fcc, and bcc). The results are shown exemplarily for  $T_e = 12,631$  K in Fig. 3a (see Supplementary Fig. 1 for results at other electronic temperatures). For all considered electronic temperatures and ideal crystal structures, the cohesive energy curves are almost perfectly reproduced by the  $T_e$ -NNP. Furthermore, the cohesive energies are correctly predicted to approach to zero for large interatomic distances. Note that these kind of structures are not included in the training dataset and that the correct modeling of the underlying physics is achieved by a special neural network architecture (see “Methods”).

Besides the cohesive energies, we also calculate the phonon band structure of bulk Si in the diamond-like crystal structure at various  $T_e$ 's. Again, the ab-initio-results are reproduced with high accuracy, as can be seen in Fig. 3b for  $T_e = 12,631$  K (see Supplementary Fig. 2 for results at other electronic temperatures). Both the optical and acoustic phonon modes are well described. Furthermore, our model correctly predicts a bond softening for increasing electronic temperatures indicated by decreasing phonon frequencies. Note that the overall accuracy in the description of this property is significantly higher than that of most classical  $T_e$ -dependent interatomic potentials for Si from the literature<sup>16,17</sup>. Only the interatomic potential by Bauerhenne

et al.<sup>18</sup> achieves a comparable accuracy (see ref. <sup>30</sup> for an extensive comparison of literature models).

In order to test if the  $T_e$ -NNP is also able to reproduce electronic properties, we further determine the internal energy  $U_e$  and the specific heat  $C_e$  of the electrons for bulk diamond Si using the thermodynamic relations

$$U_e = \Phi - T_e \frac{\partial \Phi}{\partial T_e}, \quad C_e = -T_e \frac{\partial^2 \Phi}{\partial T_e^2}. \quad (2)$$

In Fig. 3c, d, these two properties are plotted against the electronic temperature. While the internal energy is almost perfectly reproduced, the deviations from the DFT results are slightly higher for the specific heat. Nevertheless, the overall description of this property is qualitatively correct and the  $T_e$ -NNP is indeed capable of reproducing electronic properties, which is remarkable since the electronic temperature  $T_e$  is the only information on the electrons that is passed to the model.

Next, we use the trained  $T_e$ -NNP to perform MD simulations on laser-excited bulk Si at various electronic temperatures. Thereby, the Si atoms are initiated in the ideal diamond-like crystal structure and periodic boundary conditions are applied in all spatial directions. The root-mean-square displacement RMSD( $t$ ) of the Si atoms is shown as a function of time during the first picosecond after the laser excitation at  $T_e = 12,631$  K (Fig. 3) and  $T_e = 22,104$  K (Fig. 3f). The corresponding results at  $T_e = 15,789$  K and  $T_e = 18,947$  K are provided in Supplementary Fig. 3. The simulations are repeated several times and the width of the curves corresponds to twice the standard error of the mean over all simulation runs. For temperatures below the non-thermal melting threshold, the RMSD( $t$ ) oscillates and indicates phonon antisqueezing, while for higher temperatures, it increases monotonously and a non-thermal melting of the crystal structure can be observed. Both of these effects are reproduced correctly by the neural network interatomic potential.

Finally, we calculate the melting temperature of Si by simulating the coexistence of liquid and crystal using a simulation setup described in ref. <sup>30</sup>. We determine the melting temperature

$$T_m(p) = (1263 \pm 2) \text{ K} - (25 \pm 2) \text{ K GPa}^{-1} \times p \quad (3)$$

near zero pressure  $p$ . This value agrees with  $T_m(p) = (1300 \pm 50) \text{ K} - 58 \text{ K GPa}^{-1} \times p$  from DFT using the local density approximation (LDA)<sup>60</sup>. Thus, the  $T_e$ -dependent neural network interatomic potential correctly reproduces the melting temperature given by the reference method used to generate its training data. Note that this value differs significantly from the experimental value  $T_m(p) = (1687 \pm 5) \text{ K} - 58 \text{ K GPa}^{-1} \times p$ <sup>61,62</sup>. However, this only affects the timescale for the recrystallization of the material at  $t > 100$  ps after the laser-excitation. Nevertheless, we propose a possible approach to correct the melting temperature in the Discussion.

**Application to large-scale molecular dynamics simulation.** We perform large-scale MD simulations using our derived  $T_e$ -dependent interatomic potential to demonstrate its practical applicability for systems that cannot be treated with DFT anymore. We consider a simulation cell containing a Si thin film with more than 160,000 atoms, which is more than 500 times the number of atoms contained in the supercell used for the training dataset. Furthermore, we insert two different kinds of spherically shaped surface impurities centered on the top layer of the thin film: A hill and a sink, each with a diameter of 2 nm. We simulate the effect of the laser excitation on these structures separately and compare the results for various pulse energies.

In addition to performing MD simulations on larger length scales, we also increase the simulation time to 10 ps. On this timescale, the energy is transferred from the electronic to the ionic system and thereby,  $T_e$  decreases while the ionic temperature  $T_i$  increases. We explicitly consider the electron-phonon coupling in our simulations using an extended two-temperature model MD (TTM-MD) setup (see “Methods”). The method strictly requires a  $T_e$ -dependent interatomic potential that provides a physically meaningful specific heat of the electrons with  $C_e \geq 0$ . The  $T_e$ -NNP is fulfilling this requirement as can be seen in Fig. 3d.

We find that the laser-excited thin film undergoes a period of expansion and shrinking, which is more pronounced for higher pulse energies. During the expansion, the thickness of the thin film increases by a rate between 1.9% for a pulse energy of 0.1 eV and 11.1% for 0.5 eV. It reaches its maximum value between 4 to 5 ps after the laser excitation. The time evolution of the thickness of the thin film is shown in Supplementary Fig. 4. After 10 ps, the surface of the thin film gets slightly coarser and for pulse energies greater than 0.3 eV, both kinds of surface impurities vanish. Figure 4 shows the surface of both structures at  $t = 0$  ps,  $t = 5$  ps, and  $t = 10$  ps after a laser excitation with a pulse energy of 0.35 eV. The corresponding results at  $t = 10$  ps after laser excitations with other pulse energies (0.1, 0.3, and 0.5 eV) are shown in Supplementary Fig. 5.

## Discussion

The  $T_e$ -dependent neural network interatomic potential can be evaluated very efficiently compared to DFT and since it scales linearly with the number of atoms, it can be applied to arbitrary system sizes. At the same time, we demonstrated that it is able to achieve ab-initio accuracy both in MD simulations as well as in reproducing important physical properties like the phonon band structure or the specific heat of the electrons.

We also showed that the  $T_e$ -dependent neural network interatomic potential is correctly reproducing the melting temperature of Si given by the underlying reference DFT method

used to generate the training data. However, this value differs from the experimental result by more than 300 K. A possible approach to force the neural network interatomic potential to reproduce the experimental melting temperature would be to replace the training data for atomic configurations at an ionic temperature  $T_i$  near the melting temperature with data generated by a different method that is known to give a more accurate melting temperature, e.g., the random phase approximation<sup>63</sup>. Both methods could be combined in a weighted sum with  $T_i$ -dependent weights to ensure a smooth transition between them. An alternative approach would be to combine the different reference methods in a hierarchical fashion using a multi-fidelity approach<sup>64</sup>.

A transfer of our model to other laser-excited materials could be accomplished with only a few changes. First, the symmetry function parameters would have to be readjusted for the new dataset (see “Methods”). Second, an expression for the energy  $\Phi_0(T_e)$  of the isolated atoms is required, but this could for example be done with a simple polynomial model fitted to easily generateable ab-initio data. Finally, our model could be directly adopted for the training on a sufficiently large and diverse ab-initio dataset of other laser-excited materials including varying electronic temperatures.

In conclusion, we want to emphasize that our model is easier to derive compared to existing empirically determined  $T_e$ -dependent interatomic potentials due to its data-driven approach and furthermore, it is more flexible in terms of possible extensions. For example, in early stages of the laser excitation (10–100 fs) the electrons have not thermalized to a Fermi distribution yet. In order to better describe this situation, the  $T_e$ -input-node could be replaced by multiple input nodes representing a suitable discretization of a non-equilibrium electron distribution. We seek to investigate this possibility in a future work.

## Methods

**$T_e$ -dependent high-dimensional neural network potential.** Our model is a  $T_e$ -dependent neural network interatomic potential ( $T_e$ -NNP) based on the high-dimensional architecture introduced by Behler and Parrinello in 2007<sup>38</sup> and is depicted in Fig. 1b. It treats the cohesive energy  $\Phi_{\text{coh}}$  as a sum of atomic contributions  $\Phi_i$ ,

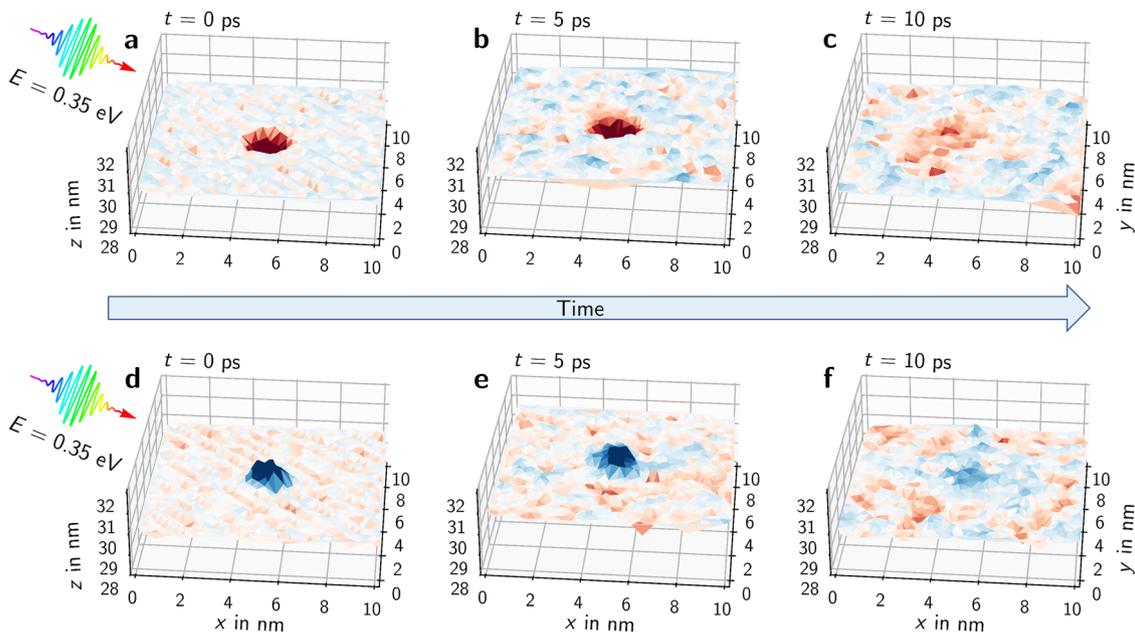
$$\Phi_{\text{coh}}(\mathbf{r}_1, \dots, \mathbf{r}_N, T_e) = \sum_{i=1}^N \Phi_i(\mathbf{g}_i(\mathbf{r}_1, \dots, \mathbf{r}_N), T_e). \quad (4)$$

Thereby,  $\mathbf{g}_i$  is a set of atom-centered symmetry functions describing the environment of atom  $i$  in terms of the atomic coordinates  $\mathbf{r}_1, \dots, \mathbf{r}_N$ .

A feed-forward neural network is trained to predict  $\Phi_i$  for a single atom from the symmetry functions  $\mathbf{g}_i$  and the electronic temperature  $T_e$  (see Fig. 1c). The same network, which from now on will be referred to as the atomic neural network, is evaluated individually for every atom in the structure and the results are summed up to obtain the total cohesive energy. Training the model can be carried out by comparing the network output with reference DFT data and applying gradient-based optimization. Once the network is trained, it can be applied to large-scale atomistic simulations at low computational costs compared to DFT.

To implement the dependency of  $\Phi_{\text{coh}}$  on the electronic temperature  $T_e$ , we introduce an additional input node to the atomic neural network that represents  $T_e$ . With this modification, it is possible to train the model on the physical relevant atomic configurations appearing in MD simulations at different  $T_e$ 's. Thus, the network output  $\Phi_i$  is a function of the symmetry functions  $\mathbf{g}_i$  describing the environment of atom  $i$  (they are, in turn, functions of the atomic positions  $\mathbf{r}_i$ ) and the electronic temperature  $T_e$ . For numerical stability, we scale these input values to the range [0, 1].

We use the standard mathematical expression of a feed-forward neural network, i.e., a sequence of matrix multiplications combined with a non-linear activation function  $f_{\text{act}}$ . In most neural network applications, a trainable bias vector is also added to each hidden layer. Since the influence of the bias on the prediction accuracy is rather small, we increase the physical meaning of the model and its capability to describe physical properties like the cohesive energy per atom for different crystal structures by introducing a special bias architecture. It is based on the observation that all symmetry functions will be zero if every interatomic distance is greater than the cutoff radius. In this case of very large interatomic distances the cohesive energy per atom approaches to zero. However, biases



**Fig. 4 Large-scale MD simulation using the  $T_e$ -dependent neural network interatomic potential.** The top layer of two different Si thin films (**a-c** Sink-like surface defect, **d-f** Hill-like surface defect) is shown at three times (0 ps, 5 ps, 10 ps) after a laser excitation with a pulse energy of 0.35 eV. Both surface defects vanish within 10 ps after the laser excitation. The MD simulations were performed using the  $T_e$ -NNP on a supercell containing more than 160,000 atoms.

generally lead to a non-zero network output in the case of only zeros as inputs. Therefore, we omit the biases of the hidden layers completely and only introduce a fixed non-trainable bias to the output layer to compensate for the non-zero  $T_e$ -input-node. This bias corresponds to the network output evaluated for  $\mathbf{g}_i = \mathbf{0}$  and is subtracted from the original output value. In the case of two hidden layers the atomic potential energy can thus be written as

$$\Phi_i(\mathbf{g}_i, T_e) = f_{\text{nn}}(\mathbf{g}_i, T_e) - f_{\text{nn}}(\mathbf{0}, T_e) \quad (5)$$

with

$$f_{\text{nn}}(\mathbf{g}_i, T_e) = \sum_{l=1}^{N_{\text{hid}}^{(2)}} w_{l1}^{(23)} \cdot f_{\text{act}}^{(2)} \left( \sum_{k=1}^{N_{\text{hid}}^{(1)}} w_{kl}^{(12)} \cdot f_{\text{act}}^{(1)} \left( \sum_{j=1}^{N_{\text{sf}}} w_{jk}^{(01)} \cdot \mathbf{g}_{ij} + w_{N_{\text{sf}}+1,k}^{(01)} \cdot T_e \right) \right). \quad (6)$$

Here,  $N_{\text{hid}}^{(p)}$  is the number of nodes in the  $p$ -th layer,  $f_{\text{act}}^{(p)}$  is the non-linear activation function applied in the  $p$ -th layer,  $\mathbf{W}^{(pq)}$  is the weight matrix connecting layer  $p$  with layer  $q$  and  $N_{\text{sf}}$  is the number of symmetry functions.

In order to calculate the total potential energy of a system using Eq. (1) and to correctly reproduce the derivative of  $\Phi$  with respect to  $T_e$ , one also has to calculate the energy  $\Phi_0(T_e)$  of the isolated atoms. In this work, we use an existing model for Si from the literature, which is based on polynomials<sup>18</sup>.

**Atom-centered symmetry functions.** We use the concept of atom-centered symmetry functions<sup>39</sup> in order to describe the atomic environments in the form of input values for the atomic neural network. However, we propose a different functional form that can be calculated very efficiently because it does not involve the repeated expensive computation of the exponential function like commonly used functions<sup>10</sup>. The functions are inspired by the functional form of the  $T_e$ -dependent interatomic potential for Si proposed by Bauerhenne et al.<sup>18</sup>. Our radial symmetry functions are constructed as sums of two-body terms to describe the radial environment of atom  $i$  and are defined by

$$g_i^{\text{rad}} = \sum_{j=1}^N f_c(R_{ij}, \eta^{\text{rad}} R_c^{\text{rad}}), \quad (7)$$

where  $R_{ij}$  is the distance from atom  $i$  to its  $j$ -th neighboring atom and  $f_c(R_{ij}, R_c)$  is the cutoff function controlling the atomic interaction range,

$$f_c(R_{ij}, R_c) = \begin{cases} \left(1 - \frac{R_{ij}}{R_c}\right)^2, & R_{ij} \leq R_c \\ 0, & R_{ij} > R_c \end{cases}. \quad (8)$$

Here, different values of  $\eta^{\text{rad}}$  with  $\eta^{\text{rad}} \in \mathbb{R}$ ,  $0 < \eta^{\text{rad}} \leq 1$  are used, which means that the function is evaluated for different cutoff radii  $R_c = \eta^{\text{rad}} R_c^{\text{rad}}$  up to a fixed maximum cutoff radius  $R_c^{\text{rad}}$ . Furthermore, we also take three-body interactions

**Table 1 Parameters of the radial symmetry functions.**

Index	1	2	3	4	5	6	7
$\eta^{\text{rad}} R_c^{\text{rad}}$ [nm]	0.265	0.295	0.326	0.357	0.388	0.419	0.450
Index	8	9	10	11	12	13	
$\eta^{\text{rad}} R_c^{\text{rad}}$ [nm]	0.481	0.512	0.542	0.573	0.604	0.635	

into account by using angular symmetry functions defined as

$$g_i^{\text{ang}} = 2^{1-\zeta} \sum_{j,k \neq i}^{\text{all}} \left(1 + \lambda \cos \theta_{ijk}\right)^\zeta \cdot f_c(R_{ij}, \eta^{\text{ang}} R_c^{\text{ang}}) \cdot f_c(R_{ik}, \eta^{\text{ang}} R_c^{\text{ang}}), \quad (9)$$

where  $\zeta \in \mathbb{N}$  and  $\lambda = \pm 1$ . Moreover,  $\theta_{ijk} = \arccos(\mathbf{r}_{ij} \cdot \mathbf{r}_{ik} / R_{ij} \cdot R_{ik})$  is the angle centered at atom  $i$  with respect to two neighboring atoms  $j$  and  $k$ . The function  $G_i^{\text{ang}}$  essentially corresponds to the functional form proposed in ref. <sup>38</sup>, except that the radial parts are replaced by the cutoff function  $f_c$ . In general, the cutoff radii  $\eta^{\text{ang}} R_c^{\text{ang}}$  used for the angular symmetry functions are chosen differently from those used for the radial symmetry functions because different interatomic distances may play a role for two- and three-body interactions.

The parameters  $\eta^{\text{rad}}$ ,  $\eta^{\text{ang}}$ ,  $\zeta$  and  $\lambda$  in Eqs. (7) and (9) are varied to cover the whole atomic environment up to the maximum cutoff radii  $R_c^{\text{rad}}$  and  $R_c^{\text{ang}}$ . The number  $N_{\text{sf}}$  of used symmetry functions differing in these parameters determines the resolution of the atomic environment description. The exact parameters we use in this work are summarized in Tables 1 and 2.

Note that the parameters of the symmetry functions have to be treated as hyperparameters which are specific to the material that is investigated and are not learned by the model during training. Therefore, they have to be adjusted when transferring the model to other materials. However, the space of possible solutions can be reduced by restricting the parameters  $\eta^{\text{rad}}$  and  $\eta^{\text{ang}}$  to be equidistant. In this way, only the spacing between these parameters as well as the maximum cutoff radii  $R_c^{\text{rad}}$  and  $R_c^{\text{ang}}$  have to be optimized using standard hyperparameter optimization techniques. An alternative method to find the symmetry function parameters is to minimize the linear correlation between a fixed number of symmetry functions evaluated on the training dataset. This technique is used in this study to find the  $\zeta$  parameters for the angular symmetry functions.

**Training and optimization.** In this work, we use an atomic neural network with two hidden layers and  $N_{\text{hid}}^{(1)} = N_{\text{hid}}^{(2)} = 50$ . Furthermore, we use the tanh activation function after the first hidden layer and the Gaussian error linear unit (GELU)<sup>65</sup> after the second hidden layer. Especially the choice of the activation functions strongly influences the quality of the reproduction of physical properties. For

**Table 2 Parameters of the angular symmetry functions.**

Index	$r_l^{\text{ang}} R_c^{\text{ang}}$ [nm]	$\lambda$	$\zeta$ -values
14–18	0.318	+1	1, 2, 3, 4, 5
19–23		−1	1, 5, 7, 9, 10
24–28	0.335	+1	1, 2, 3, 4, 5
29–33		−1	1, 5, 7, 9, 12
34–38	0.392	+1	1, 2, 3, 5, 7
39–43		−1	1, 4, 6, 9, 16
44–48	0.429	+1	1, 3, 6, 8, 11
49–53		−1	1, 5, 8, 13, 22
54–58	0.466	+1	1, 5, 7, 9, 13
59–63		−1	1, 5, 9, 16, 26
64–68	0.503	+1	1, 5, 7, 10, 14
69–73		−1	1, 6, 13, 21, 30

example, we find that using the exponential linear unit (ELU)<sup>66</sup> results in anomalies in the cohesive energy curves at small lattice parameters.

The cost function for the gradient-based training process is based on the error function used in ref. 18 and is defined by

$$\Gamma = \frac{1}{2} (\gamma_E^2 + \gamma_F^2) = \frac{1}{2} \cdot \left( \frac{\sum_t (\Phi_{\text{coh}}^{(t)} - E^{(t)})^2}{\sum_t (E^{(t)})^2} + \frac{\sum_t \sum_i |\nabla_{\mathbf{r}_i^{(t)}} \Phi_{\text{coh}}^{(t)} - \mathbf{F}_i^{(t)}|^2}{\sum_t \sum_i |\mathbf{F}_i^{(t)}|^2} \right). \quad (10)$$

Here,  $E$  is the reference ab-initio total Helmholtz free cohesive energy and  $\mathbf{F}_i$  is the ab-initio force acting on atom  $i$ . The particular form of this relative error function is selected to guarantee that the network learns to predict the energies and forces in both situations, near equilibrium as well as at high electronic temperatures, with similar relative accuracy despite strongly varying magnitudes of the predicted quantities. For one weight update, the cost function is evaluated as the mean over multiple training examples indexed by  $t$  within a batch of size  $N_{\text{batch}}$  (which we set to 20). During one training epoch, the batches are selected randomly from the training set without repetition until the model has processed every example of the training set (minibatch learning). For this stochastic gradient-based optimization of the cost function we further use the Adam algorithm<sup>67</sup> with the learning rate  $\alpha = 0.001$  and the exponential decay rates  $\beta_1 = 0.9$  and  $\beta_2 = 0.999$  for the first and second moment estimates. The neural network is trained for 50 epochs with an integrated early stopping mechanism which is activated when the validation set error does not decrease for ten epochs in a row. However, this condition was not fulfilled during our experiments.

**Details on DFT calculations and MD simulations.** All ab-initio calculations in this work, including those contained in the training dataset, are carried out using our in-house  $T_e$ -dependent DFT program CHIVES (Code for Highly excited Valence Electron Systems)<sup>7,68,69</sup>. CHIVES yields a very good agreement with experiment<sup>70</sup> as well as established DFT codes, including Wien2k and ABINIT, and is more than 200 times faster than ABINIT at the same degree of accuracy<sup>71</sup>.

The dataset used to train the  $T_e$ -NNP was originally generated in ref. 18 for the construction of a classical  $T_e$ -dependent interatomic potential for laser-excited Si. It contains ab-initio calculations of a Si thin film at 316 K as well as at ten higher electronic temperatures ranging from 3158 to 31,578 K in steps of 3158 K  $\approx$  10 mHa. Further ab-initio calculations of the thin film at intermediate electronic temperatures from 1579 to 29,999 K in equally sized steps are used as the validation dataset.

At each electronic temperature in the training and validation dataset, three different types of ab-initio calculations are performed:

1. MD simulation: The laser-induced dynamics are simulated by suddenly increasing  $T_e$  from room temperature to a constant value. The duration of the MD simulations is set to 1 ps with a step size of 2 fs resulting in 501 different thin film configurations with corresponding energies and forces for each electronic temperature.
2. Artificial compression: The Si thin film is compressed stepwise and the corresponding energies and forces are calculated to also include atomic configurations with high local atomic densities in the training dataset, which are likely to occur after a laser-excitation. In this way, 25 additional data points are generated for each electronic temperature.
3. Artificial expansion: Since laser-excitations can also lead to high local negative pressures with decreased local atomic densities, mechanical expansions of the thin film are simulated and the corresponding energies and forces are calculated. In this way, 501 data points are added to the dataset for all  $T_e < 25,000$  K. For higher temperatures, the thin film is already naturally expanding during the MD simulation.

All in all, both the training and the validation dataset contain more than 9000 different atomic configurations of the thin film with corresponding labels for the cohesive energy as well as the three force components for each atom. We decided to use a comparatively small percentage of the total available data for training to emphasize the data efficiency of our model.

The thin film used in this study has a thickness of 5.3 nm and consists of 320 atoms. The simulation supercell is constructed by arranging  $2 \times 2 \times 10$  conventional unit cells. Since CHIVES uses periodic boundary conditions in all spatial directions, a vacuum is inserted on top of the thin film by doubling the supercell size in  $z$ -direction in order to simulate surface effects. The Si atoms are prepared in the ideal diamond-like crystal structure and initialized to room temperature using an Andersen thermostat<sup>72</sup>.

Among others, the model performance is evaluated by comparing the atomic root-mean-square displacement during MD simulations of bulk Si with results from ab-initio calculations (see Fig. 3e, f). Note that these configurations are not contained in the training or validation dataset. For the bulk MD simulations at  $T_e = 12,631$  K and  $T_e = 15,789$  K, we use a simulation cell containing 640 atoms in total and consisting of  $4 \times 4 \times 5$  conventional unit cells. We repeat the simulation 10 times. For  $T_e = 18,947$  K and  $T_e = 22,104$  K, the simulation cell contains 288 atoms ( $3 \times 3 \times 4$  conventional unit cells) and we repeat the simulation 40 times. The simulation duration and step size is chosen similarly to the ab-initio calculations of the thin film. The bulk MD simulations are first performed using CHIVES and then repeated with the  $T_e$ -NNP using the same initializations.

**Details on large-scale MD simulations.** The large-scale MD simulations are carried out on a simulation cell containing a Si thin film with a thickness of 29.7 nm. It is constructed by arranging  $19 \times 19 \times 56$  conventional unit cells. Again, in order to simulate surface effects despite of periodic boundary conditions, we further insert a vacuum of 50 nm on top of and below the thin film. The surface impurities are spherically shaped and both have a diameter of 2 nm. The exact number of atoms contained in the simulation cells are 161,836 (hill) and 161,633 (sink). We perform MD simulations of the laser excitation at various pulse energies in the range from 0.1 to 0.5 eV with a step size of 0.05 eV. The simulation duration is set to 10 ps with a step size of 1 fs.

As described previously, the electron–phonon coupling has to be taken into account if the simulation time exceeds a few picoseconds. In this case, the electronic temperature cannot be treated as constant anymore. Therefore, we perform the MD simulations in the frame of the two-temperature model (TTM-MD)<sup>73</sup> as formulated by Ivanov and Zhigilei<sup>74</sup>. In this combined atomistic-continuum approach, the electrons with their individual temperature  $T_e$  are treated in a continuum, while the ions are modeled by a classical ground state interatomic potential. Furthermore, the electron–phonon coupling is included by performing a velocity scaling. However, since we are using a  $T_e$ -dependent interatomic potential to also take non-thermal effects into account, we deploy the extended MD simulation setup introduced by Bauerhenne<sup>30</sup>, which includes the electron–phonon coupling as well as the usage of a  $T_e$ -dependent interatomic potential. Thereby, we use the electron–phonon coupling constant for Si from ref. 75.

The TTM-MD scheme has been shown to yield an accurate description of laser-induced mechanical relaxation processes in materials leading to spallation, ablation, formation of nanoparticles, and nanostructuring of surfaces and thin films<sup>76–79</sup>.

## Data availability

The datasets analyzed in this work are available from the corresponding author on reasonable request.

## Code availability

The  $T_e$ -NNP is implemented in the Python programming language using the Tensorflow library<sup>80</sup>. The code is available from the corresponding author on reasonable request.

Received: 16 December 2022; Accepted: 4 August 2023;

Published online: 15 August 2023

## References

1. Becker, C. A., Tavazza, F., Trautt, Z. T. & Buarque de Macedo, R. A. Considerations for choosing and using force fields and interatomic potentials in materials science and engineering. *Curr. Opin. Solid State Mater. Sci.* **17**, 277–283 (2013).
2. Handley, C. M. & Behler, J. Next generation interatomic potentials for condensed systems. *Eur. Phys. J. B* **87**, 1–16 (2014).
3. Durrant, J. D. & McCammon, J. A. Molecular dynamics simulations and drug discovery. *BMC Biol.* **9**, 71–80 (2011).

4. Hollingsworth, S. A. & Dror, R. D. Molecular dynamics simulation for all. *Neuron* **99**, 1129–1143 (2018).
5. Jones, R. O. Density functional theory: its origins, rise to prominence, and future. *Rev. Mod. Phys.* **87**, 897–923 (2015).
6. Silvestrelli, P. L., Alavi, A., Parrinello, M. & Frenkel, D. Ab initio molecular dynamics simulation of laser melting of silicon. *Phys. Rev. Lett.* **77**, 3149–3152 (1996).
7. Zijlstra, E. S., Kalitsov, A., Zier, T. & Garcia, M. E. Squeezed thermal phonons preclude nonthermal melting of silicon as a function of fluence. *Phys. Rev. X* **3**, 011005 (2013).
8. Thapa, R., Ugwumadu, C., Nepal, K., Tremblay, J. & Drabold, D. Ab initio simulation of amorphous graphite. *Phys. Rev. Lett.* **128**, 236402 (2022).
9. Deringer, V. L., Caro, M. A. & Csányi, G. Machine learning interatomic potentials as emerging tools for materials science. *Adv. Mater.* **31**, 1902765 (2019).
10. Behler, J. Four generations of high-dimensional neural network potentials. *Chem. Rev.* **121**, 10037–10072 (2021).
11. Dhaliwal, G., Nair, P. B. & Singh, C. V. Machine learned interatomic potentials using random features. *Npj Comput. Mater.* **8**, 7 (2022).
12. Xu, C. & Wise, F. Recent advances in fibre lasers for nonlinear microscopy. *Nat. Photon.* **7**, 875–882 (2013).
13. Chung, S. H. & Mazur, E. Surgical applications of femtosecond lasers. *J. Biophotonics* **2**, 557–572 (2009).
14. Sugioka, K. & Cheng, Y. Ultrafast lasers—reliable tools for advanced materials processing. *Light Sci. Appl.* **3**, e149 (2014).
15. Phillips, K. C., Gandhi, H. H., Mazur, E. & Sundaram, S. K. Ultrafast laser processing of materials: A review. *Adv. Opt. Photon.* **7**, 684–712 (2015).
16. Shokeen, L. & Schelling, P. K. Thermodynamics and kinetics of silicon under conditions of strong electronic excitation. *J. Appl. Phys.* **109**, 073503 (2011).
17. Darkins, R., Ma, P.-W., Murphy, S. T. & Duffy, D. M. Simulating electronically driven structural changes in silicon with two-temperature molecular dynamics. *Phys. Rev. B* **98**, 024304 (2018).
18. Bauerhenne, B., Lipp, V. P., Zier, T., Zijlstra, E. S. & Garcia, M. E. Self-learning method for construction of analytical interatomic potentials to describe laser-excited materials. *Phys. Rev. Lett.* **124**, 085501 (2020).
19. Cavalleri, A. et al. Femtosecond structural dynamics in VO<sub>2</sub> during an ultrafast solid-solid phase transition. *Phys. Rev. Lett.* **87**, 237401 (2001).
20. Collet, E. et al. Laser-induced ferroelectric structural order in an organic charge-transfer crystal. *Science* **300**, 612–615 (2003).
21. Sciaini, G. et al. Electronic acceleration of atomic motions and disordering in bismuth. *Nature* **458**, 56–59 (2009).
22. Buzzi, M., Först, M., Mankowsky, R. & Cavalleri, A. Probing dynamics in quantum materials with femtosecond X-rays. *Nat. Rev. Mater.* **3**, 299–311 (2018).
23. Johnson, S. L. et al. Directly observing squeezed phonon states with femtosecond X-ray diffraction. *Phys. Rev. Lett.* **102**, 175503 (2009).
24. Cheng, T. K. et al. Mechanism for displacive excitation of coherent phonons in Sb, Bi, Te, and Ti<sub>2</sub>O<sub>3</sub>. *Appl. Phys. Lett.* **59**, 1923–1925 (1991).
25. Hase, M., Kitajima, M., Constantinescu, A. M. & Petek, H. The birth of a quasiparticle in silicon observed in time-frequency space. *Nature* **426**, 51–54 (2003).
26. Gamaly, E. G. & Rode, A. V. Physics of ultra-short laser interaction with matter: From phonon excitation to ultimate transformations. *Prog. Quantum Electron.* **37**, 215–323 (2013).
27. Recoules, V., Clérouin, J., Zérah, G., Anglade, P. M. & Mazevet, S. Effect of intense laser irradiation on the lattice stability of semiconductors and metals. *Phys. Rev. Lett.* **96**, 055503 (2006).
28. Grigoryan, N. S., Zier, T., Garcia, M. E. & Zijlstra, E. S. Ultrafast structural phenomena: Theory of phonon frequency changes and simulations with code for highly excited valence electron systems. *J. Opt. Soc. Am. B* **31**, 22–27 (2014).
29. Fritz, D. M. et al. Ultrafast bond softening in bismuth: Mapping a solid’s interatomic potential with X-rays. *Science* **315**, 633–636 (2007).
30. Bauerhenne, B. *Materials Interaction with Femtosecond Lasers: Theory and Ultra-large-scale Simulations of Thermal and Nonthermal Phenomena* (Springer, 2021).
31. Varlamova, O., Costache, F., Reif, J. & Bestehorn, M. Self-organized pattern formation upon femtosecond laser ablation by circularly polarized light. *Appl. Surf. Sci.* **252**, 4702–4706 (2006).
32. Reif, J., Varlamova, O., Varlamov, S. & Bestehorn, M. The role of asymmetric excitation in self-organized nanostructure formation upon femtosecond laser ablation. *AIP Conf. Proc.* **1464**, 428–441 (2012).
33. Bonse, J. & Gräf, S. Ten open questions about laser-induced periodic surface structures. *Nanomaterials* **11**, 3326 (2021).
34. Blank, T. B., Brown, S. D., Calhoun, A. W. & Doren, D. J. Neural network models of potential energy surfaces. *J. Chem. Phys.* **103**, 4129–4137 (1995).
35. Gassner, H., Probst, M., Lauenstein, A. & Hermansson, K. Representation of intermolecular potential functions by neural networks. *J. Phys. Chem. A* **102**, 4596–4605 (1998).
36. Lorenz, S., Groß, A. & Scheffler, M. Representing high-dimensional potential-energy surfaces for reactions at surfaces by neural networks. *Chem. Phys. Lett.* **395**, 210–215 (2004).
37. Manzhos, S., Wang, X., Dawes, R. & Carrington, T. A nested molecule-independent neural network approach for high-quality potential fits. *J. Phys. Chem. A* **110**, 5295–5304 (2006).
38. Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
39. Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **134**, 074106 (2011).
40. Behler, J., Martoňák, R., Donadio, D. & Parrinello, M. Metadynamics simulations of the high-pressure phases of silicon employing a high-dimensional neural network potential. *Phys. Rev. Lett.* **100**, 185501 (2008).
41. Artrith, N. & Behler, J. High-dimensional neural network potentials for metal surfaces: a prototype study for copper. *Phys. Rev. B* **85**, 045439 (2012).
42. Gastegger, M. & Marquetand, P. High-dimensional neural network potentials for organic reactions and an improved training algorithm. *J. Chem. Theory Comput.* **11**, 2187–2198 (2015).
43. Artrith, N., Morawietz, T. & Behler, J. High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide. *Phys. Rev. B* **83**, 153101 (2011).
44. Morawietz, T., Sharma, V. & Behler, J. A neural network potential-energy surface for the water dimer based on environment-dependent atomic energies and charges. *J. Chem. Phys.* **136**, 064103 (2012).
45. Eckhoff, M. & Behler, J. High-dimensional neural network potentials for magnetic systems using spin-dependent atom-centered symmetry functions. *Npj Comput. Mater.* **7**, 170 (2021).
46. Ghasemi, S. A., Hofstetter, A., Saha, S. & Goedecker, S. Interatomic potentials for ionic systems with density functional accuracy based on charge densities obtained by a neural network. *Phys. Rev. B* **92**, 045131 (2015).
47. Xie, X., Persson, K. A. & Small, D. W. Incorporating electronic information into machine learning potential energy surfaces via approaching the ground-state electronic energy as a function of atom-based electronic populations. *J. Chem. Theory Comput.* **16**, 4256–4270 (2020).
48. Ko, T. W., Finkler, J. A., Goedecker, S. & Behler, J. A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer. *Nat. Commun.* **12**, 398 (2021).
49. Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
50. Bartók, A. P., Kermode, J., Bernstein, N. & Csányi, G. Machine learning a general-purpose interatomic potential for silicon. *Phys. Rev. X* **8**, 041048 (2018).
51. Thompson, A. P., Swiler, L. P., Trott, C. R., Foiles, S. M. & Tucker, G. J. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *J. Comput. Phys.* **285**, 316–330 (2015).
52. Wood, M. A. & Thompson, A. P. Extending the accuracy of the snap interatomic potential form. *J. Chem. Phys.* **148**, 241721 (2018).
53. Shapeev, A. V. Moment tensor potentials: a class of systematically improvable interatomic potentials. *Multiscale Model. Simul.* **14**, 1153–1173 (2016).
54. Podryabinkin, E. V. & Shapeev, A. V. Active learning of linearly parametrized interatomic potentials. *Comput. Mat. Sci.* **140**, 171–180 (2017).
55. Schütt, K. et al. SchNet: a continuous-filter convolutional neural network for modeling quantum interactions. *Adv. Neural Inf. Process. Syst.* **30**, 991–1001 (2017).
56. Barry, M. C., Wise, K. E., Kalidindi, S. R. & Kumar, S. Voxelized atomic structure potentials: predicting atomic forces with the accuracy of quantum mechanics using convolutional neural networks. *J. Phys. Chem. Lett.* **11**, 9093–9099 (2020).
57. Gasteiger, J., Becker, F. & Günnemann, S. Gemnet: Universal directional graph neural networks for molecules. *Adv. Neural Inf. Process. Syst.* **34**, 6790–6802 (2021).
58. Batzner, S. et al. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453 (2022).
59. Zier, T., Zijlstra, E. S. & Garcia, M. E. Silicon before the bonds break. *Appl. Phys. A* **117**, 1–5 (2014).
60. Alfé, D. & Gillian, M. J. Exchange-correlation energy and phase diagram of Si. *Phys. Rev. B* **68**, 205212 (2003).
61. Yamaguchi, K. & Itagaki, K. Measurement of high temperature heat content of silicon by drop calorimetry. *J. Therm. Anal. Calorim.* **69**, 1059–1066 (2002).
62. Jayaraman, A., Klement, W. & Kennedy, G. C. Melting and polymorphism at high pressures in some groups IV elements and III-V compounds with the diamond/zincblende structure. *Phys. Rev.* **130**, 540–547 (1963).
63. Dorner, F., Sukurma, Z., Dellago, C. & Kresse, G. Melting Si: beyond density functional theory. *Phys. Rev. Lett.* **121**, 195701 (2018).
64. Pilania, G., Gubernatis, J. E. & Lookman, T. Multi-fidelity machine learning models for accurate bandgap predictions of solids. *Comput. Mater. Sci.* **129**, 156–163 (2018).

65. Hendrycks, D. & Gimpel, K. Gaussian error linear units (GELUs). Preprint at <https://arxiv.org/abs/1606.08415> (2016).
66. Clevert, D.-A., Unterthiner, T. & Hochreiter, S. Fast and accurate deep network learning by exponential linear units (ELUs). Preprint at <https://arxiv.org/abs/1511.07289> (2015).
67. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. Preprint at <https://arxiv.org/abs/1412.6980> (2014).
68. Zijlstra, E. S., Huntemann, N., Kalitsov, A., Garcia, M. E. & Von Barth, U. Optimized Gaussian basis sets for Goedecker-Teter-Hutter pseudopotentials. *Model. Simul. Mater. Sci. Eng.* **17**, 015009 (2009).
69. Zijlstra, E. S., Kalitsov, A., Zier, T. & Garcia, M. E. Fractional diffusion in silicon. *Adv. Mater.* **25**, 5605–5608 (2013).
70. Waldecker, L. et al. Coherent and incoherent structural dynamics in laser-excited antimony. *Phys. Rev. B* **95**, 054302 (2017).
71. Zijlstra, E. S. et al. Femtosecond-laser-induced bond breaking and structural modifications in silicon, TiO<sub>2</sub>, and defective graphene: an ab initio molecular dynamics study. *Appl. Phys. A* **114**, 1–9 (2014).
72. Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* **72**, 2384–2393 (1980).
73. Anisimov, S. et al. Electron emission from metal surfaces exposed to ultrashort laser pulses. *Zh. Eksp. Teor. Fiz* **66**, 375–377 (1974).
74. Ivanov, D. S. & Zhigilei, L. V. Combined atomistic-continuum modeling of short-pulse laser melting and disintegration of metal films. *Phys. Rev. B* **68**, 064114 (2003).
75. Sadasivam, S., Chan, M. K. Y. & Darancet, P. Theory of thermal relaxation of electrons in semiconductors. *Phys. Rev. Lett.* **119**, 136602 (2017).
76. Wu, C. & Zhigilei, L. V. Microscopic mechanisms of laser spallation and ablation of metal targets from large-scale molecular dynamics simulations. *Appl. Phys. A* **114**, 11–32 (2014).
77. Shih, C.-Y. et al. Two mechanisms of nanoparticle generation in picosecond laser ablation in liquids: the origin of the bimodal size distribution. *Nanoscale* **10**, 6900–6910 (2018).
78. Ivanov, D. S. et al. Experimental and theoretical investigation of periodic nanostructuring of Au with ultrashort UV laser pulses near the damage threshold. *Phys. Rev. Appl.* **4**, 064006 (2015).
79. Ivanov, D. S. et al. The mechanism of nanobump formation in femtosecond pulse laser nanostructuring of thin metal films. *Appl. Phys. A* **92**, 791–796 (2008).
80. Abadi, M. et al. Tensorflow: a system for large-scale machine learning. in *12th USENIX Symposium on Operating Systems Design and Implementation* 265–283 (2016).

## Acknowledgements

The computations for this work were performed on the Lichtenberg High Performance Computer (HHLR), TU Darmstadt, and on the computing cluster of the IT Servicecenter (ITS), University of Kassel. Support from the Deutsche Forschungsgemeinschaft (DFG) through grant GA 465/27-1 is acknowledged.

## Author contributions

M.E.G. devised the main conceptual idea and the project. M.E.G., B.B., and P.P. conceived the research plan. P.P. designed and trained the ML model and wrote the related software. B.B. generated the reference DFT dataset and wrote the software for integrating the ML model in MD simulations. P.P. and B.B. wrote the software for performance evaluation of the ML model and performed the large-scale MD simulations. P.P. analyzed the data produced with the ML model. All three authors discussed the results and wrote the manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43246-023-00389-w>.

**Correspondence** and requests for materials should be addressed to Pascal Plettenberg.

**Peer review information** *Communications Materials* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Milica Todorović and Aldo Isidori. A peer review file is available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023