Interatomic machine learning potentials for aluminium: 
application to solidification phenomena

Noel Jakse,1 Johannes Sandberg,1,2,3 Anthony Saliou,1 Philippe Jarry,4 Emilie Devijver,5 Thomas Voigtmann,2,3 Jürgen Horbach,6 and Andreas Meyer2,7

1Université Grenoble Alpes, CNRS, 
Grenoble INP, SIMaP F-38000 Grenoble, France

2Institut für Materialphysik im Weltraum, 
Deutsches Zentrum für Luft- und Raumfahrt (DLR), 51170 Köln, Germany

3Department of Physics, Heinrich-Heine-Universität Düsseldorf, 
Universitätsstraße 1, 40225 Düsseldorf, Germany

4C-TEC, Parc Economique Centr’alp, 725 rue Aristide Bergès, 
CS10027, Voreppe 38341 cedex, France

5Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, F-38000 Grenoble, France

6Institut für Theoretische Physik II, 
Heinrich-Heine-Universität Düsseldorf, 
Universitätsstraße 1, 40225 Düsseldorf, Germany

7Institut Laue-Langevin (ILL), 38042 Grenoble, France
Abstract

In studying solidification process by simulations on the atomic scale, modeling crystal nucleation or amorphisation require the construction of interatomic interactions that are able to reproduce the properties of both the solid and the liquid states. Taking into account rare nucleation events or structural relaxation under deep undercooling conditions requires much larger length scales and longer time scales than those achievable by ab initio molecular dynamics (AIMD). This problem is addressed by means of classical MD simulations using a high dimensional neural network potential trained on a relevant set of configurations generated by AIMD and applied to elemental aluminium. The dataset contains various crystalline structures and liquid states at different pressures, including their time fluctuations in a wide range of temperatures. The resulting potential is shown to be efficient to reproduce the basic structural, dynamics and thermodynamic quantities in the liquid and undercooled states with an accuracy close to the ab initio one, without the need to include neither explicitly the forces nor all kind of configurations in the training procedure. The early stages of the crystallization is further investigated on a scale much larger than the ab initio allowing us to unravel features of the homogeneous nucleation mechanisms.
I. INTRODUCTION

Apart from steel alloys, aluminium and its alloys represent the most used and attractive structural metallic materials due to their specific properties such as low weight, low energy cost of remelting, and the possibility of almost complete recycling. Therefore, these materials represent a major axis of the energy transition [1]. An intimate understanding of its condensed phase properties is founded upon a description of the atomic level structure and dynamics, and requires an accurate representation of chemical bonding [2]. This is of utmost importance in order to tackle phenomena such as phase changes and solidification process during which a liquid morphs into a solid either by crystallization or amorphization [3, 4], showing eventually a change in electronic structure as a metal-to-semiconductor transition [5, 6]. First principles approaches, essentially through the Density Functional Theory (DFT) [7, 8], represent the dedicated framework especially with the breakthrough provided by *ab initio* molecular dynamics (AIMD) simulations [9] in combining atomic dynamics with DFT. Despite its enormous success in many complex chemical bonding environments [2], DFT implementations are limited to a few hundred atoms over time scales less than 1 ns [10, 11] on actual large-scale supercomputing facilities, impeding its use for phenomena at length and time scales typical of solidification [12].

The desire to bridge typical scales of the electronic structure to those of the properties under investigation has led to deriving interatomic potentials with semi-empirical functional forms that average out or otherwise model electronic degrees of freedom. For metallic materials of interest here, various approaches were proposed, starting in the second half of the 20th Century with pair-potentials based on a nearly-free electron gas description of the electronic structure [13] using simple model within the pseudopotential theory (PT) [15–18]. It was later acknowledged that it was impossible for pair potentials to describe on the same footing the structure, dynamic, and thermodynamic properties in the liquid and solid state [14], with inherent mechanical instability under shear for crystals. Many-body approaches such as the Embedded-Atom Model (EAM) [19, 20], modified EAM (MEAM) [21] of current widespread use as well as the Reactive Force Field (ReaxFF) [22], just to name a few among many others [23], can be considered as successful in this respect. Fitting the parameters of
these potentials is most often oriented towards describing crystalline phases and transitions between some of them [25, 26], more rarely taking a full account of the liquid state [24]. This leads to a lack of transferability [14] and a limited ability to tackle phenomena involving several phases such as crystal nucleation [12].

Over the last decade, impressive progress was made in designing potentials from electronic structure calculations using supervised Machine Learning (ML) methods [27–33]. There are now standard libraries for the ML training [34] that can be used in combination with molecular dynamics (MD) simulation packages [35] such as LAMMPS [36] or in combination with workflow softwares such as ASE [37]. On-the-fly ML force field methods have been also proposed [38] and implemented directly into ab initio codes in order to bypass most of electronic-structure calculation steps [39]. Different ML techniques have been used, ranging from simple linear regression (LR) methods such as the spectral neighbor analysis potential (SNAP) method [31, 42] to highly non-linear regression methods using High Dimensional Neural Networks [34, 40] (HDNN) or Kernel Regression (KR) [41, 43, 44]. The designed potentials reach in general an accuracy close to the ab initio calculations from which the database was formed, with exceptional results for the description of relative stability between crystalline phases [46] and defects [26]. However, approaches taking full account of liquid and crystalline states remain scarce [6, 46] and are often limited to the objective of a good description of the melting point. The main reason for this stems from the fact that the chosen ab initio configurations should cover all situations, as ML techniques may become less reliable outside the training domain [27]. It becomes even more crucial for crystal nucleation occurring under deep undercooling conditions with a strong evolution of the liquid structure with respect to that above melting, showing an increasing icosahedral [10] ordering and structural heterogeneity [48] triggering homogeneous nucleation [49].

Machine-learning potentials for aluminium were designed very recently to describe essentially the properties of the solid states [50, 51] and the melting temperature [50]. Two different approaches were put forward respectively with a Gaussian kernel regression [50] and a deep NN [51] with a dataset built from configurations extracted from AIMD simulations at various temperatures. ML potentials were initially trained using the DFT energies starting from the work of Behler and Parrinello on bulk Si [40]. It was subsequently pointed out that the
learning process could benefit from a wealth of additional information if the three components of the force and six components of the stress per atom are taken into account 27, 33, while one has only a single energy value per simulated configuration. Still in some works, only the forces have been used for the training showing that properties like the vibrational properties in the solid states can be reproduced, but they remain insufficient to get full account of thermodynamic quantities 44, 50. Interestingly, a very recent ML potential for bulk Si, shown to be able to unravel the early stages of crystallization, was trained only using the energies 6. Thus the question whether additional information enhances the training or not is still an open issue, given also the fact that the relative importance of the energy, forces and stresses for estimating the Mean-Square Error (MSE) or the Root-MSE (RMSE) introduce two additional free training parameters 27. Moreover, the question of the transferability of a ML potential taking into account both the liquid and solid phases as mentioned above, remains essentially unexplored for aluminium and is of importance when dealing with solidification phenomena.

The aim of the present work is to develop a ML potential for pure aluminium dedicated to the description of condensed phases, namely liquid and solid states for temperatures up to 8000 K and pressures up to 300 GPa. For this purpose, we have developed a HDNN on the basis of Behler and Parrinello’s approach 27, 40. The latter was trained on a data set generated by DFT-based simulations for the main crystalline structures and liquid states covering the targeted pressure and temperature domain, including their time fluctuations by an appropriate sampling of phase space trajectories obtained by AIMD. Training the HDNN solely on the basis of energies, the description of the structure, dynamics and thermodynamics is shown to be accurate in the investigated domain. The resulting potential is applied here to solidification processes, namely amorphisation and/or early stages of crystal nucleation, allowing us to unravel features of the homogeneous nucleation mechanisms.

The remaining part of the paper is organized as the following. In Sec. II, the specific features of the HDNN, the training procedure as well as the basic assessment of the potential on independent DFT and experimental thermodynamic data. Sec. III is devoted to the test of the potential’s accuracy in describing some structural properties, the dynamics of the liquid state in the investigated pressure-temperature domain, as well as the homogeneous
nucleation. Finally, in Sec. IV the main outcomes of the work are given.

II. COMPUTATIONAL BACKGROUND

A. Constructing a ML potential

In the last three decades, many potentials of pure Al and its alloys for the use in atomistic simulations been developed [1 2 52 57 59 60 62 65], using the Morse potential ansatz [52], the PT method [60], embedded-atom method (EAM) [2 54 56 59 63], the modified EAM (MEAM) [1 21 65], and many-body approaches such as COMB3 [64]. However, only very few studies have employed the ML approach [44, 50, 51], none of them taking systematically into account the properties of the liquid state. In this work, we derive a ML-based potential for Al that accurately describes the properties of Al melts and solidification processes in Al.

Among the various approaches put forward to design ML potentials [28 31], the choice was made to set up a high-dimensional neural network built in a similar way to the one proposed by Behler and Parrinello [40] and Zhang et al. [66]. This approach has been proven successful for pure silicon [6] as well as water [46] to model reliably both their properties in the liquid and solid states. As a detailed description can be found for instance in the tutorial review by Behler [27] among others, the focus is mainly made on the specificity of our scheme. The main part consists in a supervised learning task from a relevant sample of atomic configurations with known energy generated by AIMD in various crystalline and liquids structures in the desired temperature and pressure domain. In the corresponding portion of Potential Energy Landscape (PEL), each configurational energy is considered as a sum of individual atomic energies determined from their local atomic environment within a cut-off radius $r_S$ often extending beyond the first-neighbor atomic shell, and taken here to be 6.4 Å for Al. This decomposition allows us to train $N$ Neural Networks (NN), each of them being assigned to an individual local atomic environment. The NN are then combined to recover the energy of the whole configuration of $N$ atoms.

The individual NN is defined by the same network topology for given atomic species. It specifies the number of neurons formally named here $y_{i,j}^{N,t}$ and their connectivity through
the weights, $w_{i,j}^{N,l}$, and called a Multi-Layer Perceptron (MLP). The weights associated with each node pair are optimized during the learning process by a Feed-Forward technique \[67\]. Thus, each of the $M$ layers within the neural network consists of sets of nodes receiving multiple inputs from the previous layer and passing outputs to the next layer. Here a fully connected network is used, in which every output of a layer is an input for every neuron in the next layer. The corresponding mathematical description is as follows: the input signals are linearly combined before being activated by function $f$ to give each output $y_i^{N,l}$ of a given fully connected layer $l$ as

$$y_i^{N,l} = f\left(\sum_{j=1}^{M_{l-1}} w_{i,j}^{N,l} y_j^{N,l-1} + b_i^{N,l}\right), \quad (1)$$

where $M_l$ refers to the size of the $l$-th layer, i.e. the number of neurons in the layer. Note that positive weights enhance connections while negative weights tend to inhibit them. Most of the activation functions are chosen to have a range in either $[0, 1]$ or $[-1, 1]$ and modulate the amplitude of the output. The activation function $f$ is applied element-wise and is taken as the hyperbolic tangent form $f(x) = \tanh(x)$. Back-propagation is used to update the network weights and their gradients.

The input layer of a NN takes values representative of one local atomic environment in the form of a feature vector whose dimension is then equal to the number of its nodes. The feature vector is built on the basis of Behler-Parrinello (BP) descriptors \[40\] to represent the radial and angular arrangements of atoms in the local structure using Gaussian symmetry functions having the translational and rotational invariance. For aluminium, the number of components of the BP feature was chosen to be 22, comprising of 12 radial and 10 angular components, as described in more detail in the Supplementary Information (SI) file. Then, the NN architecture for aluminum is $22 \times 10 \times 10 \times 1$ with, as a common case, 2 hidden layers of 10 nodes each.

The NN was coded using Keras module from the TensorFlow Python package \[68\] in the regression mode. NNs of all $N$ atoms of a configuration are then associated using the ADD module to form the HDNN which is obviously invariant to permutation of atoms. The HDNN is then trained on the single DFT energy of the whole configuration.
FIG. 1. Schematic representation of the feed-forward neural network built as a densely connected multi-layer perceptron. The input layer $x_i^{N,0}$ will be fed with the atomic BP feature of the training set, and the output is the atomic energy. The neural network contains two hidden layers with superscript 1 and 2, respectively. The two layers are composed of 10 neurons each. The weights $w_{ijk}^{N,l}$ and bias $b_i^{N,l}$ are optimized during the training (see text).

B. Building the dataset

Designing an appropriate dataset is the crucial and demanding step for the construction of the NN interatomic potential. It was built from AIMD simulations that were partly taken from our previous works [10,69–71] and extended here to have a better representation of the undercooled liquid region at high temperature, the crystalline fcc configurations up to the melting point at zero pressure, and crystalline fcc, bcc, and hcp up to 300 GPa with and without defects. The different thermodynamic states and structures as well as the number of configurations sampled in each case are given in Tables SII and SIII in the Supplementary Information file. In total, 24300 configurations of $N = 256$ atoms were gathered in the database that enabled us to cover solid and liquid states at ambient pressure as well as liquid samples at temperatures up to 8000 K and pressures up to 300 GPa. Non-equilibrium
trajectories in the undercooled region were also generated to take into account crystal nucleation and solidification processes in the ML fitting procedure.

For the sake of self-consistency, the main technical details of the AIMD simulations are recalled here. They were performed by means of the Vienna Ab initio Simulation Package (VASP) [72]. The local density approximation (LDA) [73, 74] within projected augmented plane-waves was applied to all simulations with a plane-wave cutoff of 241 eV. For the liquid states, only the Γ-point is used while for crystalline states, the Γ-centered grid of k-points in the irreducible part of the Brillouin zone was set to $2 \times 2 \times 2$ following the Monkhorst–Pack scheme [76, 77]. All the simulations were performed with $N = 256$ atoms placed in a cubic simulation box (except for the hcp crystal where an orthorombic box was used) with standard periodic boundary conditions (PBC). Newton’s equations of motion were solved numerically with Verlet’s algorithm in the velocity form with a time step of 1.5 fs, and phase-space trajectories were constructed within the canonical ensemble (NVT), by means of a Nosé thermostat to control the temperature $T$. The temperature evolution in the undercooled states was obtained by quenching the system stepwise down to 600 K with a temperature step of 50 K. For each temperature, the simulation cell was resized according to the experimental density [78] and the run was continued for 30 ps before performing the next quench, resulting in an average cooling rate of $3.3 \times 10^{12}$ K/s. The calculated pressures for all the temperatures studied here were less than $\pm 1$ GPa generally, so that on average a quasi constant pressure during the quenching is observed. For temperatures ranging from $T = 1000$ K to 600 K, the run was continued for equilibration during a time up to 200 ps. A similar procedure was applied for heating the fcc crystal from 10 K to 900 K.

Several aspects deserve attention in the perspective of building the ML potential. First of all, the choice of the exchange and correlation (XC) functional for the electronic structure calculations is crucial. As the ML potential may reach an accuracy similar to the DFT calculations, it will mirror the ability of the XC functional used in predicting the properties, at least in the thermodynamic domain inside which it was trained. This has guided our choice of the Local Density Approximation (LDA) functional given the fact that the Generalized-Gradient-Approximation (GGA) overestimate the atomic volume [80]. Moreover, it was shown in our previous contributions that the LDA gives a good description of the liquid structure [10].
More importantly, atomic transport properties such as the selfdiffusion coefficient, which are very sensitive to the details of the potentials, are well reproduced within the LDA compared to state-of-the-art experimental data [79, 81]. Such a good agreement with experiments was very recently confirmed on the dynamic structure factors as well as the structural relaxation times extracted from the self-intermediate scattering function [71]. For high pressures, it was shown [82] that the difference between LDA and GGA [83] is negligible in describing the pressure-density phase diagram of aluminium up to pressures as high as 10 TPa.

Secondly, in the perspective of performing MD simulations, care has to be taken in describing not only average thermodynamic properties but also the fluctuations around the mean value, especially in order to capture the features of local basins of the PEL [84]. This requires the sampling of a large number of configurations along AIMD phase space trajectories. Therefore, in the present work, for each of the considered thermodynamic states (see Tables SII and SIII of the SI file), 1000 configurations were generated on AIMD production runs over 40 ps.

Finally, as mentioned in the introduction, the question whether including the additional information of the forces or even the stresses in addition to the energies improves the learning process and the accuracy of the potential is still an open issue. Given the recent success in designing a ML potential for bulk Si [6] solely based on energies in the training, this strategy was also followed here.

C. Training the Neural Network

The supervised training is carried out using as input the BP feature vectors describing local atomic environments in each configuration. AIMD energies of these configurations are used to find the optimal set of weights and biases. The complete dataset of configurations is firstly randomized and scaled using the standard scaler of Scikit-learn, i.e. centering the feature components about their mean and normalize them according to their standard deviation. It is then split into a training set of 80% of the data and a test set containing the 20% remaining part. In the training set 20% of the data are retained to create validation sets. They are used (i) for a cross-validation procedure to estimate the performance of various NN
architectures through the MSE, and (ii) to monitor the MSE on the validation data during the learning process to detect overfitting. For a given architecture, the optimization is performed using the training data without the validation set, and terminating when the validation error starts to increase. Reduction of the noise of the MSE during training is obtained by including a callback with a stepwise reduction of the learning rate. Simultaneously, a $L_2$ norm regularization is performed to reduce the model complexity, and thus to prevent overfitting. Once trained, the weights and biases are stored in a format compatible with the LAMMPS HDNNP pair-style [35].

This training stage is repeated with various NN architectures to find the optimal one capturing at best the functional dependence of the data. Evaluation of the MSE is carried out through a stochastic gradient descent minimization using the Adam optimization algorithm [67] giving a measure of the loss with a learning rate starting at 0.01 and reducing most of the time to 0.0001 during the training, $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\varepsilon = 10^{-8}$. The early-stopping was performed with maximum loss variation of $10^{-6}$ and a patience of 45 epochs. The typical duration of the training period was about 10000 to 15000 epochs. The least MSE loss is obtained for an architecture of 10 neurons in the first and second hidden layers. A typical evolution of loss and the validation loss is shown in Fig. 2(a). A cross-validation performed over 5 independent trainings gives a RMSE of $(3.1 \pm 0.5)$ meV on the per atom energy. Figure 2(b) displays the predictive ability of the model on the unseen data of the test set, with a high quality over the whole range of the energies.

The predictive ability of the HDNN is illustrated in Fig. 3 on the three forces components extracted from AIMD configurations of a simulation at 1500 K over 2 ps. This shows that the forces can be predicted with a good accuracy without being explicitly part of the learning process, and thus is in favor of the supervised learning strategy based only on the energy of the configurations.

The HDNN is further tested on the prediction of the energy as a function of time. Consecutive configurations of AIMD simulation of liquids were considered at $T = 600$ K in the undercooled region, $T = 950$ K in the vicinity of the melting point, at 1500 K far above the melting point at zero pressure, and $T = 8000$ K just above the melting line for a pressure of 322 GPa, as show in Fig. 4. Energy fluctuations are very well reproduced for
FIG. 2. (a) Evolution of the MSE losses as a function of number of epochs for the training and validation sets. (b) Test-Predict curve showing the quality of the prediction on the test set for the optimized architecture $22 \times 10 \times 10 \times 1$. The red solid line represents the known output of the energies and the blue dots the values predicted against the known ones. The main panel corresponds to thermodynamic states at low pressures ($< 5$ GPa) and temperatures between 10 K and 1500 K containing either crystalline and liquid configurations. The inset corresponds to energies in the full range of temperatures and pressures (see text).

all the temperatures, even for the extreme values for which the sampling is scarce, as they correspond to the tail of the Gaussian distribution of energy fluctuations. Probably the most impressive agreement is that of the simulation at $T = 8000$ K where the energy range of the fluctuation is as large as $0.25 \text{eV/atom}$ and still very well predicted. This demonstrates the quality of the ML potential.

D. Molecular dynamics simulation

Classical MD simulations were carried out using the LAMMPS package [36]. These simulations were performed in different ensembles, namely the canonical ensemble ($NVT$, constant temperature, volume, and number of atoms), the isobaric-isothermal ensemble
FIG. 3. Test-Predict curves for the three components of the predicted forces against those extracted from the AIMD configurations at $T = 1500$ K over 2 ps. The red solid line represents the three AIMD reference force components and the blue dots the values predicted from the corresponding configurations.

$(NPT$, constant temperature, pressure, and number of atoms), and the isobaric-isoenthalpic ensemble $(NPH$, constant pressure, enthalpy, and number of atoms). Temperature and pressure were kept constant via the Nose-Hoover thermostat and barostat [85, 86], respectively. In all simulations, PBC in the three spatial directions were employed. The integration of the equations of motion was done via Verlet’s algorithm in the velocity form, choosing a time step of 1 fs. The use of our HDNN potential in LAMMPS was possible through the library-based implementation of high-dimensional neural network potentials by Singraber et al. [35].

Simulations at constant volume with $N \approx 10000$ atoms are done to directly compare structural and dynamic properties to those obtained from AIMD and to evaluate the accuracy of the potential with respect to the temperature dependence of the pressure $P$. For the solid states the target temperature is obtained by heating the perfect crystal at $5 \times 10^{12}$ K/s. After a typical equilibration time of 20 ps, the pair-correlation functions, the diffusion coefficients, and thermodynamic quantities are calculated over a production time ranging from 100 ps to 1 ns depending on the thermodynamic state under consideration. Additional $NPT$ simulations are carried out with a similar procedure for the melt-quenching of bulk systems down to the glass transition temperature or the stepwise heating of the crystal up to the melting temperature to determine the enthalpy-temperature diagram as well as to simulate homogeneous nucleation processes along isotherms.
FIG. 4. Energy per atom as a function of time from AIMD simulation of liquids and predicted by the HDNN potential at $T = 600$ K in the undercooled region, $T = 950$ K in the vicinity of the melting point, at 1500 K far above the melting point at zero pressure, and $T = 8000$ K just above the melting line for a pressure of 322 GPa.
Liquid-solid interfaces (LSI) simulations are performed for the purpose of determining the melting line by the two-phase coexistence. The procedure follows the approaches proposed in Refs. [87, 89, 101, 106] and is similar to the protocol used in Ref. [24]. A simulation cell containing around $N = 43000$ atoms was set up with an initial crystalline configuration with shape $28 \times 7 \times 7$ primitive cells on which PBC applied to the three directions of space. Starting at zero pressure, this system is heated and equilibrated at constant pressure to a temperature of $50\, \text{K}$ below a guess of the melting temperatures. Half of the simulation cell in the $x$ direction is further heated and maintained at a temperature of approximately $1300\, \text{K}$ until a complete melting is observed. The liquid part was then cooled down and equilibrated at a temperature $50\, \text{K}$ above the guess, thus creating a solid-liquid coexistence containing two crystal-melt interfaces due to the PBC. The simulation of the entire system is pursued in the isobaric-isoenthalpic ensemble so that the temperature of the LSI is an internal parameter free to evolve toward a steady state corresponding to the thermodynamic melting temperature if both phases survive. The simulation is continued for $1\, \text{ns}$, and the average melting temperature is determined on the last $100\, \text{ps}$ when a steady position of the two interfaces is observed. If a complete melting or solidification occurs, the procedure is started over again with a refined guess of the melting temperature. This procedure is repeated with subsequent higher pressures by first shrinking the volume of the whole simulation cell from the coexistence configuration at the preceding pressure and then increasing the temperature at constant pressure to a new guess of the melting line.

The average location of each solid-liquid interface and the local structural analysis were determined using the common-neighbor analysis (CNA) [90] with the indexing of Faken and Jonsson [91] and a bond-based algorithm as implemented in the OVITO software [92] where a uniform cut-off radius corresponding to the first minimum of the pair-correlation function of the liquid is applied to create bonds between pairs of particles. The CNA classifies pairs around each atom by sets of three indices: the first index represents the number of nearest-neighbors common to this pair, the second index corresponds to the number of nearest-neighbor bonds among the shared neighbors, and the third index indicates the longest chain of bonded atoms among them. For instance, $421$ and $422$ bonded pairs are characteristic of close packed structures fcc and hcp, respectively. The occurrence of $444$ and $666$ pairs, with
specific proportions, signals the presence of bcc ordering. The degree of five-fold symmetry is obtained from the proportion of 555, 554 and 433 pairs, which represent perfect (555) and distorted FFS based motifs. The system at the interface is considered as solid at a given distance if at least 50% of the atoms have a crystalline environment according to the CNA.

III. RESULTS AND DISCUSSION

In a first step, the optimized HDNN potential is accessed on the local structure and dynamics though the pair-correlation function \( g(r) \) and the mean-square displacement \( R^2(t) \), respectively. The function \( g(r) \) gives the probability of finding a particle \( j \) at distances \( r_{ij} \) relative to a particle \( i \) located at the origin, and reads:

\[
g(r_{ij}) = \frac{N}{V} \frac{n(r_{ij})}{4\pi r_{ij}^2 \Delta r}.
\]  

\( n(r) \) represents the mean number of particles \( j \) in a spherical shell of radius \( r \) and thickness \( \Delta r \) centered on particle \( i \). Finally, an average of \( g(r) \) over all \( N \) particle \( i \) of the simulation box is performed. Integrating \( 4\pi r^2 \rho g(r) \), with \( \rho = N/V \) up to the first minimum \( g(r) \) gives access to the mean coordination number. Figure 5 displays the curves of \( g(r) \) from NVT simulations. For all the thermodynamic states an excellent match with AIMD simulations is seen. A quantitative estimation of the departure was obtained by calculating the MSE between the classical MD and AIMD curves for each temperature. The MSE ranges from \( 6^{-4} \) typically in the case of the liquid states tot \( 1.2 \times 10^{-2} \) in the case of fcc solid states. The larger MSE for the solid might come from the fact that even a very small position shift of \( g(r) \) can induce a significant departure as peaks are sharp and narrow. Comparison with existing experiments was already done in our previous work \([10]\) and is not repeated here due to the very good agreement found here. This is all the more true for the average coordination numbers with a departure that does not exceed 0.2 (see Table I). A comparison with the most widely used EAM \([2]\) and MEAM \([1]\) potentials is shown in Fig. S1 in the Supplementary Information File for the same thermodynamic states. They are shown to perform less well than the HDNN potential, as assessed by a \( t \)-statistics, and especially for the high pressures.

Beside the local structural properties, dynamic properties represent a stringent test as
FIG. 5. Pair-correlation function for various temperatures and pressures (a) for low temperature liquid and solid states at room pressure and (b) for high temperature and pressure liquid states. Curves for 600 K, 800 K, 950 K, 2000 K, 4500 K, 8000 K (340 GPa) and 8000 K (240 GPa) are shifted upwards by an amount of 4, 8, 12, 1.5, 3, 4.5 and 6, respectively. The solid lines are results with the HDNN potential and the dashed lines with corresponding colors are those of the AIMD simulations.

| $T$ (K) | 600 | 950 | 1500 | 2000 | 4500 | 8000 (H) | 8000 (L) |
|---------|-----|-----|------|------|------|---------|---------|
| $P$ (GPa) | 0   | 0   | 0    | 30   | 110  | 240     | 340     |
| $N_C$ | 11.93 | 11.57 | 11.10 | 12.22 | 12.70 | 12.70   | 12.65   |
| $D$ ($\text{Å}^2/\text{ps}$) | 0.20 | 0.72 | 1.89 | 0.44 | 0.86 | 1.06    | 1.57    |
|       | 0.14 | 0.62 | 1.61 | 0.37 | 0.68 | 0.92    | 1.34    |

TABLE I. Pressure $P$, coordination number $N_C$, diffusion coefficient $D$ for selected temperatures. Values in second rows are from the AIMD.

they are even more sensitive to the details of the potentials. Among these, diffusion plays an important role in the solidification process \cite{12, 93} and was evaluated here through the
FIG. 6. Mean-square displacement for various temperatures and pressures (a) for high temperature and pressure liquid states and (b) for low temperature liquid and solid states at room pressure. The solid lines correspond to the HDNN potential and the dashed lines with corresponding colors are the AIMD simulations.

mean-square displacement (MSD)

$$R^2(t) = \frac{1}{N} \sum_{l=1}^{N} \left\langle \left( r_l(t + t_0) - r_l(t_0) \right)^2 \right\rangle_{t_0},$$

where $r_l(t)$ denotes the position of atom $l$ at time $t$ and $N$ is the number of atoms. In addition to the mean over all atoms, an averaging over time origins $t_0$ as indicated by the angular brackets is performed. The self-diffusion coefficient $D$ is determined from the slope of the linear behavior at long times of the MSD. In Fig. 6, the MSD is shown for temperatures in the stable and undercooled liquid states at ambient pressure as well as for temperatures along the melting line for pressures up to 340 GPa. The overall trends of AIMD curves are well reproduced by the HDNN potential for all temperatures and pressures, with a systematic slight overestimation of the diffusion coefficient as can be seen in Table I. The MSD curves show a ballistic regime at very short times $t < 0.05$ ps, followed by a diffusive regime at long times. For the lower temperatures at ambient pressures and at high pressures a well-known caging effect [93] takes place after the ballistic motion and delays the diffusive regime, which is well predicted by the HDNN potential with respect to the AIMD.

The real predictive character of the HDNN potential is assessed for $T = 950$ K, $T = 2000$ K and $T = 8000$ K at the lowest pressure as these thermodynamic states were not included in
the training, but the accuracy is similar to the other states for the pair-correlation function and the diffusion. This is all the more true for the pressure for all states shown in Table I with a departure less than 7% at the highest ones, which is remarkable since the forces were not included in the training.

Important quantities for solidification phenomena are the latent heat of fusion \[94, 95\] as well as the densities of the solid and liquid phases at the melting temperature \(T_M\). Its determination requires the calculation of the enthalpy difference between the liquid and solid branches at the melting temperature \(T_M\). The temperature evolution of the enthalpy at ambient pressure for the solid and liquid branches of the HDNN potential are drawn in Fig. 7. The simulation is started with a perfect fcc crystal at \(T = 300\) K and heated stepwise with a temperature step of 50 K with an average heating rate of \(10^{12}\) K/s. At each temperature, a simulation is performed over 50 ps (25 ps equilibration and 25 ps production) during which an average value of the enthalpy is calculated. The increase of temperature is repeated until a dynamic melting is observed, that occurred at \(T = 1250\) K. The latter value is noticeably higher than the thermodynamic melting temperature \(T_M = 970\) K obtained from LSI simulations due to overheating effects. For the liquid branch, the simulations are started at \(T = 1600\) K with an equilibrated configuration after the heating process. The same procedure as for the solid branch is followed but with a step-wise cooling down to 300 K that we called the slow cooling. Above \(T = 1250\) K, the difference in the enthalpy from the heating and cooling processes is negligible, indicating that there is no reminiscence of the crystalline state. Below 600 K the liquid undergoes a partial crystallization during cooling. Then, the cooling procedure for the liquid branch is repeated with a higher cooling rate of \(10^{13}\) K/s to avoid crystallization, and a glass transition is seen at \(T_G = 378\) K inferred from a crossover between the liquid and glassy branches as shown in Fig. 7. From the liquid and solid branches we find an enthalpy of melting of 11.67 kJ/mol which compares reasonably well with the experimental value of 13.34 kJ/mol \[96\]. Taking the numerical derivative of the solid branch yields a value of the specific heat at constant pressure, \(C_P\), of 0.99 J/g/K which is also in good agreement with the experimental value of 0.91 J/g/K. For the liquid a value of 1.158 J/g/K is obtained, which is in the range of experimental data between 1.03 J/g/K and 1.18 J/g/K close to the latest assessed values of 1.127 J/g/K \[96\]. The specific heat is a
FIG. 7. Enthalpy as a function of temperature for the solid (from heating) and liquid branches (from cooling) at ambient pressure as described in the text. The red lines indicate the slope of the amorphous and the supercooled states as a guide for the eyes for the crossover between the liquid and amorphous regimes, and marked by the vertical dotted line. The vertical dashed line marks the melting temperature obtained from the LSI simulations.

typical derivative quantity that depends on the fluctuations of the enthalpy \cite{85, 86}. The very good agreement is a strong indication that including the time fluctuations from AIMD is a fruitful strategy to describe at least the thermodynamics.

Regarding densities, the HDNN potential gives a value of 0.0587 Å⁻³ and 0.0547 Å⁻³, respectively in the solid and liquid at its melting point $T_M = 970$ K, giving rise to a density change of 0.004 Å⁻³. These values compare well with the respective experimental values \cite{78, 96, 97} of 0.0573 Å⁻³ and 0.05306 Å⁻³ with a density change of 0.0042 Å⁻³. At the experimental melting temperature, the calculated density change remains essentially unchanged, and the densities in both phases depart only by 2% with respect to the measurements.

The melting curve of aluminium was measured \cite{98, 99} up to 80 GPa using diamond anvil
FIG. 8. Melting curves for aluminium. The blue circles obtained from our LSI simulations with the HDNN are compared to the experimental data [98–100], the AIMD two-phase approach [107]. The Inset highlights the lower pressure range up to 25 GPa.

cells (DAC) and even higher at 125 GPa by means of shock experiments [100]. In Fig. 8, the results obtained from the HDNN are compared to these experimental data [98–100], the ab initio based equation of states (EOS) [82] as well as the AIMD two-phase approach [107] for which the GGA for the XC functional and 512 atoms were used. At ambient pressure, the HDNN potential yields a value of \( T_M = 970 \) K which overestimates the experimental one of 933 K by 5%. This is also the case for the AIMD [107] to a lesser extent recalling that the GGA was used and overestimates the atomic volume [80, 107]. With increasing pressure, the melting curve from the HDNN potential slightly underestimates the experiments as well as the EOS. By using two different sizes and shapes, negligible influence on the determination of the melting curve was found, confirming earlier results on pure Zr [24]. Noticeably, the reliability of the present potential on the melting line up 200 GPa is then assessed, even if high pressure thermodynamic states included in the training set are really scarce.

Finally, the homogeneous nucleation is investigated and depicted in Fig. 9. From the fast
cooling liquid branch drawn in Fig. 7 a configuration is taken at 600 K and the simulation is restarted in the \( NPT \) ensemble at ambient pressure. At such a high degree of undercooling \( \Delta T = (T_M - T) / T_M = 0.38 \) an extremely fast nucleation process is observed \[94, 95\]. Given the smallness of the simulation box containing here 10976 atoms, a genuine statistics to obtain the nucleation rate is out of reach \[12, 24\]. However, from IS configurations collected before the onset of nucleation occurring at about 50 ps, a CNA analysis \[91\] was performed. Inspection of embryos with atoms having a crystalline ordering (either fcc, hcp or bcc) showed that they dissolve back to the liquid with a size less than 90 atoms. The latter value does not represent \textit{per se} the size critical nucleus but rather a lower limit. As expected, the main crystalline phase during the growth was identified as fcc, as can be seen in Fig. 9, but a significant hcp ordering also appears during the nucleation and remains as stacking fault after complete solidification of the simulation box.

It is an open question in general, whether the homogeneous nucleation process follows the Landau Theory in which the bcc precursor is favored in the early stages of crystal nucleation \[108\] or the Ostwald step rule \[109\] according to which a primary crystal phase could be different from the the fcc one. Our findings show that aluminium follows a single step process with an onset of homogeneous nucleation showing emerging embryos with a fcc ordering. The resulting nuclei grow in a rather patchy shape with a small amount of hcp stacking fault defects. This nucleation the scenario is different from the Lennard-Jones case \[110, 111\] which follows the Landau theory and the Ostwald step rule. The present results confirm very recent molecular dynamics simulations \[3, 113\].

IV. CONCLUSION

In the present work, we developed a machine learning potential for pure aluminium by means of a high dimensional neural network on the basis of the Behler-Parrinello approach \[35, 40\]. The ML potential is devoted to the description of condensed phases, namely liquid and solid states at ambient pressure as well as those at pressures up to 300 GPa with resulting temperatures as high as 8000 K. A crucial point was the training of the potential with a data set generated by DFT-based simulations not only to cover the targeted domain of
FIG. 9. Homogeneous nucleation of deeply undercooled aluminium along the $T = 600$ K isotherm. (a) Time evolution of the energy. (b) to (f) snapshots of the simulation as various times during the nucleation. Only the atoms with crystalline ordering in the sense of the common-neighbor analysis are drawn: fcc (green); hcp (red); bcc (blue).

thermodynamic states for a question of transferability but also to consider for each of them in a physical meaningful manner the time fluctuations by an appropriate sampling of phase space trajectories obtained by \textit{ab initio} molecular dynamics. This allows to include in the training the relevant accessible microstates of the considered thermodynamic states. Another approach based on metadynamics was shown to be efficient in selecting the relevant configurations to train the neural network \cite{[6]}.

The HDNN potential thus obtained was shown to be efficient in reproducing the structural, dynamics as well as thermodynamic quantities in the liquid, undercooled and crystalline
states at ambient pressures as well as in the liquid state at high pressure up to 300 GPa, including the melting line. One important outcomes is that a reliable ML potential could be obtained without including explicitly the forces in the training, a procedure that was adopted recently for the nucleation of silicon \[6\]. The early stages of the homogeneous crystal nucleation was further investigated on a scale much larger than what is possible from the \textit{ab initio} molecular dynamics but with a similar accuracy. Results show that aluminium follows a single step nucleation process with an emerging fcc ordering and hcp stacking fault defects, confirming recent works using large scale molecular dynamics \[3, 113\].

Finally, the fact that the HDNN potential keeps a good accuracy even in domains where the thermodynamic states in the training set are scarce opens up a research line based upon active learning for regression approaches to reduce efficiently the training set. Dynamics properties such as the diffusion coefficients considered here are sensitive to the details of the potential and should be introduced in the training procedure in a more direct way than through the choice of the XC functional in the DFT calculations. This would represent a real step forward in designing ML potentials.

**ACKNOWLEDGMENTS**

We acknowledge the CINES and IDRIS under Project No. INP2227/72914, as well as CIMENT/GRICAD for computational resources. This work was performed within the framework of the Centre of Excellence of Multifunctional Architectured Materials “CEMAM” ANR-10-LABX-44-01 funded by the “Investments for the Future” Program. This work has been partially supported by MIAI@Grenoble Alpes (ANR-19-P3IA-0003). Fruitful discussions within the French collaborative network in high-temperature thermodynamics GDR CNRS 3584 (TherMatHT) are also acknowledged. J. S. acknowledges funding from the German Academic Exchange Service (DAAD) through the DLR-DAAD programme, grant No. 509.

[1] J. R. Davis, *Aluminum and Aluminum Alloys*, ASM International, p351, (2001).

[2] J. Hafner J. Comput. Chem. 29 2044 (2008).
[3] C. P. Royall and S. R. Williams, Phys. Rep. 560, 1 (2015).

[4] K. F. Kelton A. L. Greer, *Nucleation in Condensed Matter: Applications in Materials and Biology* (Pergamon, 2010).

[5] N. Jakse and A. Pasturel, Phys. Rev. Lett. 999, 2 (2007).

[6] L. Bonati and M. Parrinello, Phys. Rev. Lett. 121, 265701 (2018).

[7] M. C. Payne, M. P. Teter, D. C. Allan, T. Arias, and J. Joannopoulos, Rev. Mod. Phys. 64, 1045 (1992).

[8] K. Burke, J. Chem. Phys. 136, 150901 (2012).

[9] R. Car and M. Parrinello, Phys. Rev. Lett. 55, 2471 (1985).

[10] N. Jakse and A. Pasturel, Sci. Rep. 3, 3135 (2013).

[11] A. Pasturel and N. Jakse, J. Appl. Phys. 123, (2018).

[12] Sosso, G. C. et al., Chem. Rev. 116, 7078 (2016).

[13] J. Hafner, *From Hamiltonians to Phase Diagrams* (Springer-Verlag Berlin Heidelberg, 1987).

[14] D. K. Belashchenko, Physics-Uspekhi 56, 1176 (2013).

[15] N. W. Ashcroft Pkys. Lett. 23 48 (1966).

[16] I. M. Wills and W. A. Harrison, Phks. Rev. B 28 4363 (1983).

[17] J. A. Moriarty, Phks. Rev. B 42 1609 (1990).

[18] N. Jakse and J.-L. Bretonnet, J. Phys.: Condens. Matter 7, 3803 (1995).

[19] M.S. Daw and M.I. Baskes, Phys. Rev. B 29, 6443 (1984).

[20] M.S. Daw, S.M. Foiles, M.I. Baskes, Mater. Sci. Rep. 9, 251 (1993).

[21] M.I. Baskes, Phys. Rev. B 46, 2727 (1992).

[22] H. S. Huang, L. Q. Ai, A. C. T. van Duin, M. Chen, and Y. J. Lü, J. Chem. Phys. 151, 094503 (2019).

[23] D.G. Pettifor, *Bonding and Structure of Molecules and Solids* (Oxford University Press, Oxford, 1996).

[24] S. Becker, E. Devijver, R. Molinier, and N. Jakse, Phys. Rev. B 102, 104205 (2020).

[25] H. Zong, G. Pilania, X. Ding, G. J. Ackland, and T. Lookman, npj Comput. Mater. 4, 48 (2018).
[26] A. M. Goryaeva, C. Lapointe, C. Dai, J. Dérès, J.-B. Maillet, and M.-C. Marinica, Nat. Commun. 11, 4691 (2020).
[27] J. Behler, Int. J. Quantum Chem. 115, 1032 (2015).
[28] J. Behler, J. Chem. Phys. 145, 170901 (2016).
[29] R. Ramprasad, R. Batra1, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, npj Comput. Mater. 3, (2017).
[30] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, npj Comput. Mater. 5, 83 (2019).
[31] A. M. Goryaeva, J. B. Maillet, and M. C. Marinica, Comput. Mater. Sci. 166, 200 (2019).
[32] T. Mueller, A. Hernandez, and C. Wang, J. Chem. Phys. 152, (2020).
[33] M. R. G. Marques, J. Wolff, C. Steigemann and M. A. L. Marques, Phys. Chem. Chem. Phys. 21, 6506 (2019).
[34] A. Singraber T. Morawietz, J. Behler, and Christoph Dellago, J. Chem. Theory Comput. 15, 3075 (2019).
[35] A. Singraber, J. Behler, and C. Dellago, J. Chem. Theory Comput. 15, 1827 (2019).
[36] S. J. Plimpton, J. Comp. Phys. 117, 1 (1995); http://www.lammps.sandia.gov.
[37] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dulak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolbjer, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson et al., J. Phys. Condens. Matter 29, 273002 (2017).
[38] Z. Li, J.R. Kermode, A. De Vita, Phys. Rev. Lett. 114, 096405 (2015).
[39] R. Jinnouchi, F. Karsai, and G. Kresse, Phys. Rev. B 100, 014105 (2019).
[40] J. Behler and M. Parrinello, Phys. Rev. Lett. 98, 146401 (2007).
[41] A. P. Bartók, G. Csányi, Int. J. Quantum Chem. 115, 1051 (2015).
[42] A. Thompson, L. Swiler, C. Trott, S. Foiles, G. Tucker, J. Comp. Phys. 285, 316 (2015).
[43] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. 104, 136403 (2010).
[44] V. Botu, R. Batra, J. Chapman, R. Ramprasad, J. Phys. Chem. C 121, 511 (2017).
[45] G. C. Sosso et al., Phys. Rev. B - Condens. Matter Mater. Phys. 86, 1 (2012).
[46] T. Morawietz et al., Proc. Natl. Acad. Sci. U. S. A. 113, 8368 (2016).
[47] A. P. Bartők, J. Kermode, N. Bernstein, G. Csányi, Phys. Rev. X 8, 41048 (2018).
[48] A. Pasturel and N. Jakse, npj Computational Materials 3, 33 (2017).
[49] J. Russo and H. Tanaka. J. Chem. Phys. 145, 211801, (2016).
[50] I. Kruglov, O. Sergeev, A. Yanilkin, and A. R. Oganov, Sci. Rep. 7, (2017).
[51] A. S. Bochkarev, A. van Roekeghem, S. Mossa, and N. Mingo, Phys. Rev. Mater. 3, 093803 (2019).
[52] L. A. Girifalco, and V. G. Weizer, Physical Review, 114, 687 (1959).
[53] K. W. Jacobsen, J. K. Norskov, and M. J. Puska, Physical Review B, 35, 7423 (1987).
[54] J. Mei and J. W. Davenport, Phys. Rev. B 46, 21 (1992).
[55] F. Ercolessi, and J. B Adams, Europhys. Letters 26, 583 (1994).
[56] Y. Mishin, D. Farkas, M.J. Mehl, and D.A. Papaconstantopoulos, Phys. Rev. B 59, 3393 (1999).
[57] J. B. Sturgeon, and B. B. Laird (2000), Phys. Rev. B 62, 14720 (2000).
[58] B.-J. Lee, J.-H. Shim, and M. I. Baskes, Phys. Rev. B, 68, 144112 (2003).
[59] X.-Y. Liu, F. Ercolessi, and J. B. Adams, Modell. Sim. Mat. Sci. Eng., 12, 665 (2004).
[60] E. B. El Mendoub, R. Albaki, I. Charpentier, J.-L. Bretonnet, J.-F. Wax, and N. Jakse, J. Non-Crystal Solids 353, 3475 (2007).
[61] M. I. Mendelev, M. J. Kramer, C. A. Becker, and M. Asta, Philosophical Magazine, 88, 1723 (2008).
[62] J. M. Winey, A. Kubota, and Y. M. Gupta, Modell. Simul. Mat. Sci. Eng. 17, 055004 (2009).
[63] V. V. Zhakhovskii, N. A. Inogamov, Y. V. Petrov, S. I. Ashitkov, and K. Nishihara, Appl. Surf. Sci., 255, 9592 (2009).
[64] K. Choudhary, T. Liang, A. Chernatynskiy, Z. Lu, A. Goyal, S. R. Phillpot, and S. B. Sinnott, J. Phys.: Condens. Matter, 27, 015003 (2015).
[65] M.I. Pascuet, and J.R. Fernández J. Nucl. Materials, 467, 229 (2015).
[66] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Phys. Rev. Lett. 120, 143001 (2018).
[67] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning Data Mining, Inference, and Prediction, 2nd Edition, Springer Series in Statistics (Springer 2008).
[68] the Python package can be found here : https://www.tensorflow.org
[69] N. Jakse and A. Pasturel, J. Phys. Condens Matter 25, 28510 (2013).
[70] N. Jakse and T. Bryk, J. Chem. Phys. 151, 034506 (2019).
[71] F. Demmel, L. Hennet, and N. Jakse, Sci. Rep. (Nature) 11, 11815 (2021).
[72] G. Kresse and J. Furthmuller, Comput. Mater. Sci. 6, 15 (1996).
[73] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. 45, 566 (1980).
[74] J. P. Perdew and A. Zunger, Phys. Rev. B 23, 5048 (1981).
[75] G. Kresse and D. Joubert, Phys. Rev. B 59, 1758 (1999).
[76] H. J. Monkhorst and J. D. Pack, Phys. Rev. B 13, 5188 (1976).
[77] P. E. Blöchl, O. Jepsen and O. K. Andersen, Phys. Rev. B 49, 16223 (1994).
[78] M. J. Assael, et al., J. Phys. Chem. Ref. Data 35, 285 (2006).
[79] F. Demmel, D. Szubrin, W. C. Pilgrim, and C. Morkel, Phys. Rev. B 84, 014307 (2011).
[80] D. Alfè, Phys. Rev. B 68, 064423 (2003).
[81] F. Kargl, H. Weis, T. Unruh, and A. Meyer, J. Phys.: Conf. Ser. 340, 012077 (2012).
[82] T. Sjostrom, S. Crockett, and S. Rudin, Phys. Rev. B 94, 144101 (2016).
[83] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B 46, 6671 (1992).
[84] D. J. Wales, Energy Landscapes (Cambridge University Press, 2003).
[85] M. P. Allen, D. J. Tildesley. Computer simulation of liquids, Oxford Science Publication (1989).
[86] B. Smit and D. Frenkel. Understanding molecular simulations, 2nd ed., Academic Press, San Diego (2002).
[87] D. Y. Sun, M. Asta, and J. J. Hoyt, Phys. Rev. B 69, 024108 (2004).
[88] J. R. Morris, C. Z. Wang, K. M. Ho, and C. T. Chan, Phys. Rev. B 49, 3109 (1994).
[89] J. R. Morris and X. Y. Song, J. Chem. Phys. 116, 9352 (2002).
[90] J. D. Honeycutt, and H. C. Andersen, J. Phys. Chem. 91, 4950 (1987).
[91] D. Faken and H. Jonsson. Comput. Mat. Sci. 2, 279 (1994).
[92] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO - the Open Visualization Tool, Modelling Simul. Mater. Sci. Eng. 18, 015012 (2010).
[93] K. Binder and W. Kob, Glassy Materials and Disordered Solids (World Scientific Publishing, Singapore, 2005).
[94] D. M. Herlach, S. Binder, P. Galenko, J. Gegner, D. Holland-Moritz, S. Klein, M. Kolbe, and T. Volkmann, Metall. Mater. Trans. A Phys. Metall. Mater. Sci. 46, 4921 (2015).
[95] J. Orava and A. L. Greer, J. Chem. Phys. 140, 214504 (2014).
[96] M. Leitner et al., Metall. Mater. Trans. A Phys. Metall. Mater. Sci. 48, 3036 (2017).
[97] G. Simmons and H. Wang, Single Crystal Elastic Constants and Calculated Aggregate Properties. A Handbook 2nd Edition (The MIT Press, Cambridge, 2009).
[98] R. Boehler and M. Ross, Earth Planet. Sci. Lett. 153, 223 (1997).
[99] A. Hänström and P. Lazor, J. Alloys Compd. 305, 209 (2000).
[100] J. W. Shaner, J. M. Brown, and R. G. McQueen, High Pressure in Science and Technology (North Holland, Amsterdam, 1984).
[101] T. Zykov-Timan, R. E. Rozas, J. Horbach, and K. Binder, J. Phys.: Condens. Matter 21, 464102 (2009).
[102] T. Zykov-Timan, J. Horbach, and K. Binder, J. Chem. Phys. 133, 014705 (2010).
[103] P. Kuhn and J. Horbach, Phys. Rev. B 87, 014105 (2013).
[104] R. Benjamin and J. Horbach, J. Chem. Phys. 143, 014702 (2015).
[105] R. Rozas, A. Demirag, P. G. Toledo, and J. Horbach, J. Chem. Phys. 145, 064515 (2016).
[106] R. E. Rozas, L. G. MacDowell, P. G. Toledo, and J. Horbach, J. Chem. Phys. 154, 184704 (2021).
[107] J. Bouchet et al., Phys. Rev. B - Condens. Matter Mater. Phys. 80, 1 (2009).
[108] Alexander, S. & McTague, J. P. Should all crystals be bcc? Landau theory of solidification and crystal nucleation. Phys Rev Lett 41, 702–705 (1978).
[109] Ostwald, W. The formation and changes of solids (Translated from German). Z. Phys. Chem. 22, 289–330 (1897).
[110] ten Wolde, P. R., Ruiz-Montero, M.J. and Frenkel D., Phys Rev Lett 75, 2714–2717 (1995).
[111] ten Wolde, P. R., Ruiz-Montero, M. J. & Frenkel, D. Numerical calculation of the rate of crystal nucleation in a Lennard-Jones system at moderate undercooling. J. Chem. Phys. 104, 9932–9947 (1996).
[112] Mahata, A., Zaeem, M. A. & Baskes, M. I. Understanding homogeneous nucleation in solidification of aluminum by molecular dynamics simulations. Model. Simul. Mater. Sci. Eng.
[113] S. Becker, E. Devijver, R. Molinier, and N. Jakse, arXiv:2109.06797 (unpublished 2021).
V. BEHLER-PARRINELLO SYMMETRY FUNCTIONS

The Behler-Parrinello BP feature vectors represent the local environment of individual atoms $i$ within a defined cut-off radius $R_c$ through a cut-off function:

$$f_c(r_{ij}) = \begin{cases} 
0.5 \left( \cos\left(\frac{\pi r_{ij}}{R_c}\right) + 1 \right) & \text{if } r_{ij} \leq R_c, \\
0 & \text{if } r_{ij} > R_c,
\end{cases}$$  \hspace{1cm} (4)

where $r_{ij}$ refers to the distance between a central atom $i$ and a neighboring atom $j$. The components of the BP feature vector are determined from a Gaussian radial symmetry function

$$G_i^2 = \sum_{j=1}^{N_{\text{atoms}}} \exp[-\eta(r_{ij} - R_s)^2]f_c(r_{ij}),$$  \hspace{1cm} (5)

where $\eta$ and $R_s$ are respectively the width and the shift, as well as an angular symmetry function:

$$G_i^5 = 2^{1-\zeta} \sum_{j \neq i} \sum_{k > j} (1 + \lambda \cos(\theta_{jik}))^\zeta \exp(-\eta(r_{ij}^2 + r_{ik}^2))f_c(r_{ij})f_c(r_{ik}),$$  \hspace{1cm} (6)

where $\theta_{jik}$ represents the angle between $r_{ij}$ and $r_{ik}$, and $\zeta$, $\eta$ and $\lambda$ are parameters with given values. Using a cutoff radius $R_c = 6.8$ Å, Table SII lists the parameter sets used for each component of the BP feature vector.

VI. DATASET FOR THE TRAINING OF THE HDNN POTENTIAL

Tables SIII and SIV gather all thermodynamics states that have been simulated by AIMD to generates configuration for the training procedure of the HDNN potential. For each state, after an equilibration at a target temperature a phase space trajectory was produced from which a sample of configurations was randomly extracted to include in the data set.
|    | $R_s$(Å) | $\eta$(Å$^{-2}$) | $\eta$(Å$^{-2}$) | $\zeta$ | $\lambda$ |
|----|----------|------------------|------------------|--------|---------|
| 1  | 6.105    | 1.85             |                  |        |         |
| 2  | 5.087    | 1.85             |                  |        |         |
| 3  | 4.070    | 1.85             |                  |        |         |
| 4  | 3.052    | 1.85             |                  |        |         |
| 5  | 2.035    | 1.85             |                  |        |         |
| 6  | 1.017    | 1.85             |                  |        |         |
| 7  | 0        | 1.85             |                  |        |         |
| 8  | 0        | 0.66             |                  |        |         |
| 9  | 0        | 0.26             |                  |        |         |
| 10 | 0        | 0.1              |                  |        |         |
| 11 | 0        | 0.04             |                  |        |         |
| 12 | 0        | 0.0013           |                  |        |         |
| 13 | 0.0013   | 1                |                  |        |         |
| 14 | 0.0013   | 2                |                  |        |         |
| 15 | 0.0013   | 4                |                  |        |         |
| 16 | 0.0013   | 16               |                  |        |         |
| 17 | 0.0013   | 64               |                  |        |         |
| 18 | 0.0013   | 1                |                  | -1     |         |
| 19 | 0.0013   | 2                |                  | -1     |         |
| 20 | 0.0013   | 4                |                  | -1     |         |
| 21 | 0.0013   | 16               |                  | -1     |         |
| 22 | 0.0013   | 64               |                  | -1     |         |

**TABLE SII.** List of parameters for the radial (number 1 to 12) and angular (number 13 to 22) components of the BP Feature with a cut-off of 6.8 Å.
FIG. S10. Pair-correlation functions obtained using the Embedded Atom Model (dashed lines) and Modified Embedded Atom Model (dotted lines) potentials, with solid lines showing the corresponding \( g(r) \) obtained from AIMD simulations. Lines of same color correspond to the same temperature and system volume. The plots for \( T = 600, 800, \) and 950 K are shifted upwards by 4, 8, and 12. Likewise the plots for \( T = 2000, 4500, 8000 \) K (HP), and 8000 K (LP) are shifted by 1.5, 3, 4.5, and 6. The two 8000 K lines correspond to high (360 GPa) and low (250 GPa) pressure.

Additional AIMD at \( T = 950 \) K and \( P = 0 \) GPa, 1500 K and 2.5 GPa, 2000 K and 28 GPa, and 8000 K and 227 GPa were performed but were not included in the data set for the sake of testing the predictive ability of the HDNN potential. The database contains in total 24300, which amounts to 6 220 800 atoms.

VII. EAM AND MEAM PAIR-CORRELATION FUNCTIONS

To get a sense of how our HDNNPP compares to other widely used potentials, we have run simulations using the Embedded Atom Model (EAM) [2] and Modified Embedded Atom Model (MEAM) [1] potentials both known as performing well in the liquid and solid states (see Ref. [3] and references therein). These simulations were performed identically to the ones used to obtain the pair-correlation functions for our HDNNPP, and for the same set of temperatures and densities (see the main text). Figure S10 shows \( g(r) \) obtained from these simulations, along with the corresponding AIMD ones. For the sake of clarity, the results of the HDNN potential are not shown since they match very closely the AIMD curves.
| Structure | $T$ (K) | $P$ (GPa) | Trajectory (ps) | Sample size |
|-----------|---------|-----------|-----------------|-------------|
| fcc       | 10      | 0         | 40              | 1000        |
| fcc       | 300     | 0         | 40              | 1000        |
| fcc       | 400     | 0         | 40              | 1000        |
| fcc       | 500     | 0         | 40              | 1000        |
| fcc       | 600     | 0         | 40              | 1000        |
| fcc       | 700     | 0         | 40              | 1000        |
| fcc       | 800     | 0         | 40              | 1000        |
| fcc       | 10      | 1         | 10              | 100         |
| fcc       | 10      | 10        | 10              | 100         |
| fcc       | 10      | 100       | 10              | 100         |
| hcp       | 10      | 0         | 10              | 100         |
| hcp       | 10      | 10        | 10              | 100         |
| hcp       | 10      | 100       | 10              | 100         |
| hcp       | 10      | 200       | 10              | 100         |
| hcp       | 10      | 300       | 10              | 100         |
| bcc       | 10      | 0         | 10              | 100         |
| bcc       | 10      | 10        | 10              | 100         |
| bcc       | 10      | 100       | 10              | 100         |
| bcc       | 10      | 200       | 10              | 100         |
| bcc       | 10      | 300       | 10              | 100         |

TABLE SIII. Characteristics of the data set built from AIMD simulations. Are given the structure of the simulation (fcc, hcp, bcc, and liquid), the temperature $T$, the pressure $P$, the time span of the AIMD trajectory from which the configuration are sampled, the sample size, namely the number of configurations randomly extracted from the trajectory. The pressures less that 1 GPa were indicated as 0.
| Structure | $T$ (K) | $P$ (GPa) | Trajectory (ps) | Sample size |
|-----------|---------|-----------|----------------|-------------|
| Liquid    | 500     | 0         | 40             | 1000        |
| Liquid    | 600     | 0         | 40             | 1000        |
| Liquid    | 650     | 0         | 40             | 1000        |
| Liquid    | 700     | 0         | 40             | 1000        |
| Liquid    | 750     | 0         | 40             | 1000        |
| Liquid    | 800     | 0         | 40             | 1000        |
| Liquid    | 1000    | 0         | 40             | 1000        |
| Liquid    | 1100    | 0         | 40             | 1000        |
| liquid    | 1250    | 0         | 40             | 1000        |
| liquid    | 1350    | 0         | 40             | 1000        |
| liquid    | 1500    | 0         | 40             | 1000        |
| liquid    | 1600    | 0         | 40             | 1000        |
| liquid    | 1700    | 0         | 40             | 1000        |
| liquid    | 3100    | 56        | 40             | 1000        |
| liquid    | 4500    | 107       | 40             | 1000        |
| liquid    | 8000    | 320       | 40             | 1000        |

**TABLE SIV.** Characteristics of the data set built from AIMD simulations (continued). Same caption as Table SIII.

It is worth mentioning that above 20 GPa we were not able to achieve MD simulations with the EAM potential. At ambient pressure, both EAM and MEAM reproduce well the AIMD simulations even if noticeable can bee seen. At larger pressures and temperature the agreement worsen showing that the EAM is not transferable while the MEAM still gives reasonable results.

The mean squared errors for EAM and MEAM stand in Table SIII. We also provide the p-values to test if errors are similar to NNP using a t-test. We treat the errors at different radii as independent to get a sample of MSE, that we use to compare the means. Using a
TABLE SV. Mean-square error of the pair-correlation function $g(r)$ obtained via the EAM and MEAM potentials, and evaluated against the one obtained from AIMD, with corresponding $p$-values (noted here as $p$) as explained in the text.

| $T$ (K) | 300 | 600 | 800 | 950 |
|--------|-----|-----|-----|-----|
| MSE (EAM) | 0.0297 | 0.0041 | 0.0048 | 0.0032 |
| MSE (MEAM) | 0.036 | 0.0065 | 0.0064 | 0.0018 |
| $p$ (EAM) | $5.49 \times 10^{-6}$ | $3.59 \times 10^{-4}$ | $4.48 \times 10^{-6}$ | $3.71 \times 10^{-6}$ |
| $p$ (MEAM) | $5.48 \times 10^{-7}$ | $2.20 \times 10^{-13}$ | $3.98 \times 10^{-9}$ | $3.78 \times 10^{-10}$ |

| $T$ (K) | 1500 | 2000 | 4500 | 8000 (H) | 8000 (L) |
|--------|-----|-----|-----|---------|---------|
| MSE (EAM) | 0.0028 | 0.012 | N/A | N/A | N/A |
| MSE (MEAM) | 0.00086 | 0.066 | 0.0091 | 0.015 | 0.0095 |
| $p$ (EAM) | $4.24 \times 10^{-5}$ | $1.51 \times 10^{-6}$ | N/A | N/A | N/A |
| $p$ (MEAM) | $5.96 \times 10^{-7}$ | $9.63 \times 10^{-16}$ | $1.04 \times 10^{-11}$ | $7.44 \times 10^{-11}$ | $5.92 \times 10^{-11}$ |

level $\alpha = 0.05$, we reject the hypothesis of equality and NNP is proved to perform better.
[1] B.-J. Lee, J.-H. Shim, and M. I. Baskes, Phys. Rev. B, 68, 144112 (2003).

[2] M. I. Mendelev, M. J. Kramer, C. A. Becker, and M. Asta, Philosophical Magazine, 88, 1723 (2008).

[3] Mahata, A., Zaeem, M. A., and Baskes, M. I. Model. Simul. Mater. Sci. Eng. 26, (2018).