universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Investigation of the phase transition from low- to high chalcocite in nanorods using High-Dimensional Neural Network Potentials

verfasst von / submitted by

Hauke Fischer, BSc BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2021 / Vienna, 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 876
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Physik UG2002
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Christoph Dellago

Zusammenfassung

Kupfer(I)-sulfid ist ein häufig untersuchtes Material, das für potenzielle Anwendungen in Solarzellen erforscht wird. Es hat einen fest-fest Phasenübergang, bei dem die Schwefelatome im festen Zustand verweilen, während die Kupferatome innerhalb des Schwefelgitters mobil werden. Der Mechanismus des Phasenübergangs wird mit MD Simulationen von Nanokristallen untersucht. Dafür wird ein High-Dimensional Neural Network Potenial (HDNNP, ein machine learning Algorithmus) entwickelt – aufbauend auf einem bereits vorhandenem HDNNP für das Material unter periodischen Randbedingungen – das die Simulationen mit ab initio Genauigkeit durchführt. Das HDNNP berechnet Kräfte und Energien sehr viel schneller als reguläre ab initio Methoden und kann daher für große und lange Simulationen verwendet werden. Mit dem HDNNP werden Simulationen von Nanokristallen, unter besonderer Berücksichtigung des Einflusses der Oberfläche, durchgeführt. Es wird entdeckt, dass sich sowohl die Diffusion der Kupferatome, als auch die lokalen Umgebungen der Atome (radiale Verteilungsfunktionen) nur minimal in Oberflächennähe ändern, es aber zu sehr viel stärkeren Änderungen dieser Größen im Zentrum des Materials kommt. Es wird weiter entdeckt, dass die Energie / Atom Änderung für den Phasenübergang für Kristalle mit prozentual mehr Atomen in Oberflächennähe sinkt, und das der Phasenübergang auch in umgekehrte Richtung simuliert werden kann. Bei Letzterem stellt sich heraus, dass alle entdeckten Veränderungen im umgekehrten Phasenübergang genau rückgängig gemacht werden. Zuletzt zeigt sich, dass es beim Phasenübergang zu einer systematischen Änderung im Trägheitstensor kommt, die so interpretiert werden kann, dass die Homogenität der Massendichte während der Transformation steigt.

Abstract

Copper(I) sulfide has been studied for possible applications in photo voltaic cells. It exhibits a solid-solid phase transition, in which the sulfur atoms remain on a rigid lattice, but the copper atoms gain high mobility. An accurate High-Dimensional Neural Network Potential (a machine learning algorithm) is developed, based on a previous HDNNP for bulk chalcocite, that is capable of MD simulations of nano crystals with ab initio quality with a low fraction of the usual computational cost. This potential is used to drive simulations of nano crystals, with a particular focus on the influence of the surface on the material. The simulations show that the diffusion of copper atoms and the local environments of atoms (quantified by the radial distribution functions) change only marginally close to the surface compared to atoms far removed from the surface. Similarly, crystals with a higher fraction of atoms at the surface / atoms far away from the surface show a lower potential energy step / atom. Furthermore, it is discovered that the reverse phase transition can also be simulated and that all changes precisely reverse. A less related discovery is a systematic change in the inertia tensor that can be interpreted as an increase in the homogeneity of the mass-density.

Danksagung

Ich möchte mich zu Beginn bei all denen, die mich bei meiner Masterarbeit unterstützt haben, bedanken:
An erster Stelle bei meinem Betreuer Univ.-Prof. Mag. Dr. Christoph Dellago, der mir das Thema vorgeschlagen hat und bereit war sich mehrmals über mehrere Stunden mit mir zu treffen, um Resultate zu besprechen und mir mehrmals Tipps gegeben hat, wie ich meine Untersuchungen fortsetzen soll, die wesentlich für den Erfolg der Arbeit waren.
Ich bedanke mich auch bei Dr. Andreas Singraber, der bei den genannten Treffen immer dabei war, wertvolle Ratschläge gegeben hat und mir besonders bei auftretenden technischen Problemen mehrmals geholfen hat. Ohne Ihre vorherige Arbeit zu neuronalen Netzen, insbesondere zu „bulk“ Chalkosin, wäre meine Masterarbeit nicht möglich gewesen und ihre Werke waren sehr oft richtungsweisend für mich.
Zuletzt möchte ich mich bei meiner Verlobten Nathalia Haider bedanken, die mir enorm viele alltäglichen Aufgaben abgenommen hat, damit ich mich uneingeschränkt auf die Fertigstellung meiner Masterarbeit konzentrieren konnte. Ohne Sie hätte ich länger gebraucht, mich ungesünder ernährt und mehr Stress gehabt.

Table of Contents

1 Introduction

Copper(I) sulfide (Cu_2S), also known as chalcocite, is a commonly mined mineral. Despite being a very profitable copper ore, it is a semiconductor which has been studied for possible applications in photovolatic cells [1–4].
Bulk chalcocite exhibits a solid-solid phase transition at around $T_c = 103.5°C$ at atmospheric pressure: Below this temperature it is in the low chalcocite (LC) phase with a monoclinic units cell consisting of 144 atoms. The positions of the sulfur atoms are approximately on a distorted hexagonal lattice ($\beta = 116.35°$), while the copper atoms are fixed at interstices in between the sulfur atoms [5, 6]. Above T_c it transforms into the high chalcocite (HC) phase, in which the positions of the sulfur atoms change slightly to be on a perfectly hexagonal lattice, while the copper atoms gain high mobility and can move inside the sulfur sub-lattice. At this point, the copper atoms are in the liquid phase and only their distribution at different sites inside a much smaller hexagonal unit cell (six atoms, two Cu_2S units) can be determined [7–10].
Over the course of the last few years experiments were conducted to gain insight into the details of this phase transition [11, 12]. By observing nanorods with electron microscopy the fluctuations between the low- and high chalcocite phase could be observed with atomistic resolution. Furthermore, the time scale of this transformation was measured using X-ray spectroscopy.

The aim of this master thesis is to complement these experimental findings with molecular dynamics simulation to gain further insight into the mechanism of this phase transition. Due to the complexity of the phase transition and the underlying structure, empirical potentials are likely incapable of capturing all the details necessary. A much more reliable approach is to calculate energies and forces directly form an ab initio method, like density functional theory (DFT) combined with the Born-Oppenheimer approximation [13].
However, due to the enormous computational cost of such calculations, large scale simulations using DFT are simply infeasible.
In recent years, machine learning methods have come to the rescue to make the much higher accuracy of ab initio methods accessible at a very small fraction of the cost of doing DFT simulations.

Specifically, in 2007 Joerg Behler and Michele Parrinello presented a general method capable of representing any Potential Energy Surface (PES) with neural networks (NNs), a machine learning algorithm [14]. Their neural networks are *"trained"* with reference data to be able to predict energies and forces from DFT, without having to solve the underlying equations anymore. In this context, NNs are referred to as High-Dimensional Neural Network Potentials (HDNNPs) and the HDNNPs can be evaluated roughly five orders of magnitude faster than DFT calculations.
This thesis started with reference data for bulk chalcocite for the PBE functional [15] that was presented in a previous paper [16]. Over the course of this thesis, the data set was systematically extended to include surface configurations of chalcocite, allowing for a HDNNP that is able to simulate nano crystals. This allowed simulations of the phase transition in nano crystals comparable to the ones investigated in the experiments mentioned. A general introduction to NNs followed by the details of the HDNNP method is given in the next section.

2 Theoretical background of High-Dimensional Neural Network Potentials

2.1 A general introduction to neural networks

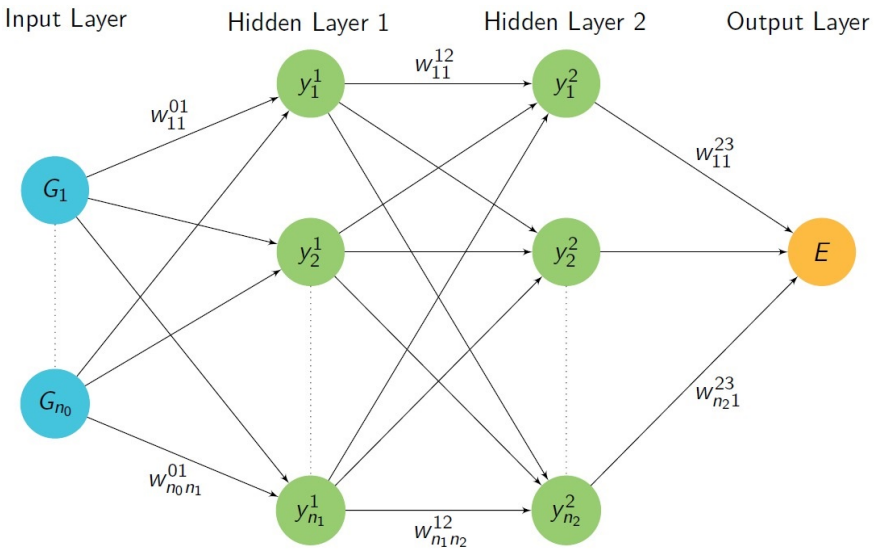


Fig. 1: This is a simple feed-forward network which consists of an input layer, two hidden layers and an output layer. Its topology is n_0 x n_1 x n_2 x 1, where n_0 stands for the number of nodes in the input layer etc. Only some arrows have been labeled.

This section will give a brief overview of neural networks (NNs). The case of HDNNPs and their application to physics will be explained in the next section.
Neural networks (NN) are algorithms inspired by biological neural networks found in brains [17]. They are capable of approximating arbitrary functions [18]. Given a set of reference data for which the values of a function are known, the free parameters of the NN are adjusted in so called "training epochs" until the NN is able to reproduce the reference data well. Once the training is over, it can be evaluated at any point. The whole procedure of training a NN can be thought of as a generalization of curve fitting for functions from a finite dimensional space to another. The training process will be explained in chapter 3. First the occurring notation is introduced and a brief explanation of how a NN is evaluated once it has been trained is given.
Assume that the NN shown in Fig. 1 has been trained to approximate a function $E_{ref}(G_1,...,G_{n_0})$. If the values $(G_1,...,G_{n_0})$ are given to the NN, it will calculate its predicted value of E_{ref}, labeled E, in the following manner:
The values $G_1,...,G_{n_0}$ correspond to the values of nodes (or artificial neurons, they are drawn as circles) in the input layer. In the first step the values of nodes in hidden layer 1 y_j^1 will be calculated:

$$y_j^1 = f_1(b_j^1 + \sum_{i=1}^{n_0} w_{ij}^{01} G_i). \tag{2.1.1}$$

The numbers w_{ij}^{01} are called weights or synapses and are drawn as arrows. The value of w_{ij}^{01} will determine how strongly G_i will affect the outcome of y_j^1. A bias weight b_j^1 is added to allow shifting the usually non-linear activation function f_1. Whereas weights can be adjusted when a NN is trained, activation functions are fixed and a number of different ones have been suggested in the literature, such as tanh(x) or log(1+e^x) [19, 20]. The second step is analogous to the first step:

$$y_j^2 = f_2(b_j^2 + \sum_{i=1}^{n_1} w_{ij}^{12} y_i^1). \tag{2.1.2}$$

Finally the value of E is calculated from the value of nodes in the second hidden layer:

$$E = f_3(b^3 + \sum_{i=1}^{n_2} w_{i1}^{23} y_i^2). \tag{2.1.3}$$

Since the value of nodes in a layer are only affected by the preceding layer, these kind of networks are called feed-forward networks.
There are a few things worth noting. If the goal is to approximate a function from $\mathbb{R}^m$ to $\mathbb{R}^n$, the input layer has to contain m nodes and the output layer n nodes. The number of hidden layers, and nodes inside those layers, is in principle arbitrary, but as a rule of thumb more hidden layers with more nodes will enable the NN to better fit complicated functions. There are many areas where NNs are used, they include speech recognition, computer vision, medical image analysis and many more [21–23].

2.2 An introduction to High-Dimensional Neural Network Potentials

This section will present the HDNNP method from Behler and Parrinello following their original paper [14]. A naive approach of training a neural network that is able to predict the energy and forces at each time step in a simulation could be the following: If we have a system with N atoms, we could give all $3N$ Cartesian coordinates to the NN and and let the NN output the total energy of the system. Taking partial derivatives of that energy with respect to the Cartesian coordinates would also provide the forces necessary. There are a number of problems with that attempt. The topology (number of layers, and nodes per layer) of a NN is fixed. A NN that has $3N$ input nodes could only be used to run a simulation with precisely N particles. Another problem is that using Cartesian coordinates will not allow the NN to respect physical symmetries. If a system is described by two equivalent coordinate systems (one is simply rotated or shifted with respect to the other one) the Cartesian coordinates will be different, and the NN would probably give slightly different results for these two coordinate systems. Similarly, exchanging two identical atoms and therefore changing which coordinate is given to which node will likely lead to inconsistencies.
There are two steps to resolve these problems. First, the total energy of the system is written as a sum of per atom contributions:

$$E_{Total} = \sum_{i=1}^{N} E_i. \tag{2.2.4}$$

The terms E_i have no physical meaning, but merely serve as a mathematical tool. Only E_{Total} is the physical energy of the system.
Instead of calculating the total energy from all $3N$ coordinates, a separate standard NN will be evaluated for each atom that only yields its contribution to the total energy. Atoms of the same type are allocated the same NN. Once all contributions have been calculated, they are summed to yield the total energy of the system. Thus, extending the system by one atom simply means doing one more evaluation of a NN.
Secondly, the question of what the input layer for each NN looks like is addressed.
Instead of giving Cartesian coordinates to the NN directly, a fixed number of symmetry functions is calculated for each atom. Symmetry functions are descriptors of the local environment of an atom that only depend on relative distances and angels, therefore being invariant under rotations and shifts. The value of those symmetry functions is given to the NNs. Their method is summarized in Fig. 2. We will continue with a detailed discussion of the symmetry functions used throughout this thesis.

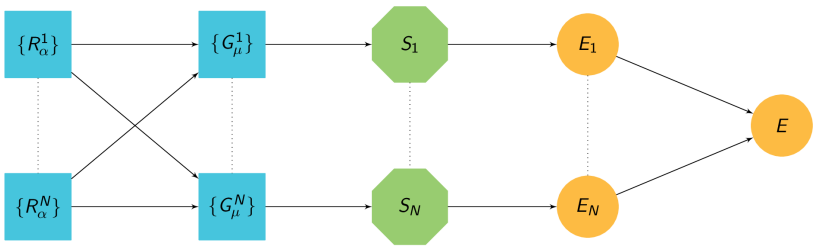


Fig. 2: This is a schematic overview of the method of Behler and Parrinello. For atom i, $\{R_\alpha^i\}$ denotes its Cartesian coordinates. The corresponding symmetry functions $\{G_\mu^i\}$ are directly calculated from the Cartesian coordinates of all other atoms in the simulation. These values are passed to the sub-net S_i of its standard NNP (e.g the hidden layers). The evaluation results in the value E_i. Once this procedure has been repeated for all atoms, the individual energy contributions are summed, yielding E. For atoms of the same type, the symmetry functions as well as the sub-nets are identical.

2.3 Symmetry functions

Throughout this thesis three types of symmetry functions, radial and angular (narrow and wide), were used

$$G_i^{\text{radial}} = \sum_{j\neq i} e^{-\eta(r_{ij}-r_s)^2} f_c(r_{ij}), \tag{2.3.5}$$

$$G_i^{\text{ang.n}} = 2^{1-\zeta} \sum_{\substack{j,k\neq i \\ j<k}} (1+\lambda\cos\theta_{ijk})^\zeta e^{-\eta(r_{ij}^2+r_{ik}^2+r_{jk}^2)} f_c(r_{ij}) f_c(r_{ik}) f_c(r_{jk}), \tag{2.3.6}$$

$$G_i^{\text{ang.w}} = 2^{1-\zeta} \sum_{\substack{j,k\neq i \\ j<k}} (1+\lambda\cos\theta_{ijk})^\zeta e^{-\eta(r_{ij}^2+r_{ik}^2)} f_c(r_{ij}) f_c(r_{ik}). \tag{2.3.7}$$

The distance between atom i and j is given by r_{ij}, whereas $\cos\theta_{ijk}$ is given by $\frac{\vec{R}_{ij}\cdot\vec{R}_{ik}}{r_{ij}\cdot r_{ik}}$, with $\vec{R}_{ij} = \vec{R}_i - \vec{R}_j$ and $\vec{R}_{ik} = \vec{R}_i - \vec{R}_k$.
Here f_c denotes a cutoff function that takes the value $f_c(0) = 1$, monotonously decreases, and smoothly goes to zero at the cutoff radius. The radial symmetry function only depends on relative distances r_{ij} between atom i and j.
One term corresponds to a Gaussian multiplied by the cutoff function, whose width can be controlled by the parameter η and whose center is determined by r_s. The resulting shape for a couple of different values of r_s and η is shown in Fig. 3.

Angular symmetry functions have two additional parameters $\zeta \in \mathbb{N}$ and $\lambda \in \{-1, 1\}$. Fig. 4 shows a single term of $G_i^{\text{ang.w}}$ for different values of λ and ζ. Note that the terms of $G_i^{\text{ang.n}}$ only differ by a factor of $e^{-\eta r_{jk}^2} f_c(r_{jk})$ from the terms of $G_i^{\text{ang.w}}$. Therefore the shape is similar, but it is narrower overall, since the additional factor is <1 for $r_{jk} > 0$.

In our case, the cutoff function is given by

$$f_c(r) = \begin{cases} x^3(x(15-6x)-10)+1, & r \leq r_c \\ 0, & r > r_c \end{cases} \tag{2.3.8}$$

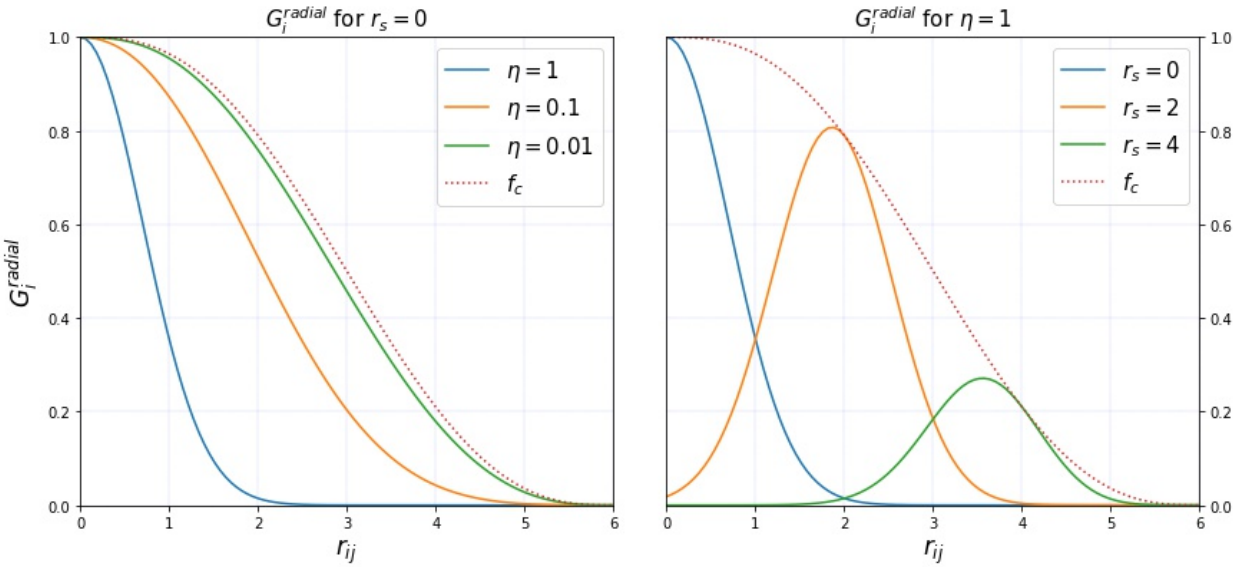


Fig. 3: On the left a single term of G_i^{radial} is plotted for different values of η and $r_s = 0$, while on the right $\eta = 1$ is fixed and the values of r_s are varied. The cutoff radius has been set to $r_c = 6$.

where $x = \frac{r}{r_c}$. This is a simple polynomial with $f(0) = 1$ that monotonously decreases and fulfills $f_c(r_c) = f'(r_c) = 0$. A plot of $f_c(r)$ is included in Fig. 3. Other cutoff functions can be used, but a polynomial was employed for its low computational cost. This cutoff function was successfully used to train a HDNNP for bulk chalcocite [16] before.

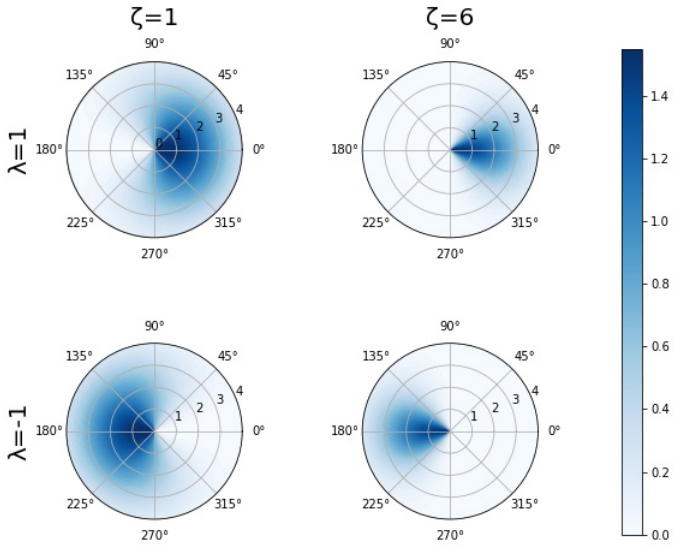


Fig. 4: One term of $G_i^{\text{ang.w}}$ is shown for $r_{ij} \in [0,4]$, $r_{ik} = 2$ constant, $\eta = 0.001$ and $r_c = 6$. Increasing ζ makes $G_i^{\text{ang.w}}$ narrower, while changing λ from 1 to -1 results in shifting the values from 0° to 180°.

Even though the HDNNPs only directly calculate the total energy of the system, forces acting on individual particles can be calculated analytically:

$$\vec{F}_i = -\vec{\nabla}_i E^{\text{pot}} = -\sum_{j=1}^{N} \vec{\nabla}_i E_j \tag{2.3.9}$$

$$= -\sum_{j=1}^{N} \sum_{i=1}^{N_j} \frac{\partial E_j}{\partial G_{j,i}} \vec{\nabla}_i G_{j,i}. \tag{2.3.10}$$

Here $\vec{\nabla}_i$ refers to the partial derivatives w.r.t. the Cartesian coordinates of atom i, N is the total number of atoms and N_j the number of symmetry functions of atom j. The following chapter will be focused on the technical details and algorithms for training HDNNPs.

3 Training of HDNNPs

In order to train HDNNPs to represent the potential energy surface of a system, a sufficiently large reference data set is needed. The reference data set contains a number of structures for which the energies and forces have been calculated from a reference method such as DFT. This data set can be obtained systematically. The chapter on extending the data set for chalcocite to include surfaces will cover the method that was used in great detail. In this section we assume that the data set is already available.
E_{Total} is a function of all symmetry functions of all atoms. Let $G_{i,j}$ denote the j'th symmetry function of atom i. If we have M reference structures we know E_{Total} for M different combinations of $G_{i,j}$. To avoid a large number of indices, we simply define G^k as the set of values $G_{i,j}$ corresponding to the reference structure k. That means our data set is given by $D = \{(E_{Total}(G^k), G^k) : k \in \{1,...,M\}\}$. Training the HDNNP simply means finding weights s.t. the root mean square error is minimized to the extent that overfitting is avoided.

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(E_{\text{Total}}(G^i) - E(G^i))^2}, \tag{3.0.11}$$

where $E(G^i)$ is calculated by the HDNNP. Forces can also be taken into account, as will be demonstrated in the next section. An iterative minimization algorithm such as stochastic gradient descent [24] or an extended Kalman filer [16] is used for the minimization of the RMSE. In the context of machine learning, one such iteration is typically called a training epoch. Before describing the variant of an extended Kalman that was used for training in this thesis, a brief explanation of what overfitting is and how it was addressed is given.

3.1 Overfitting

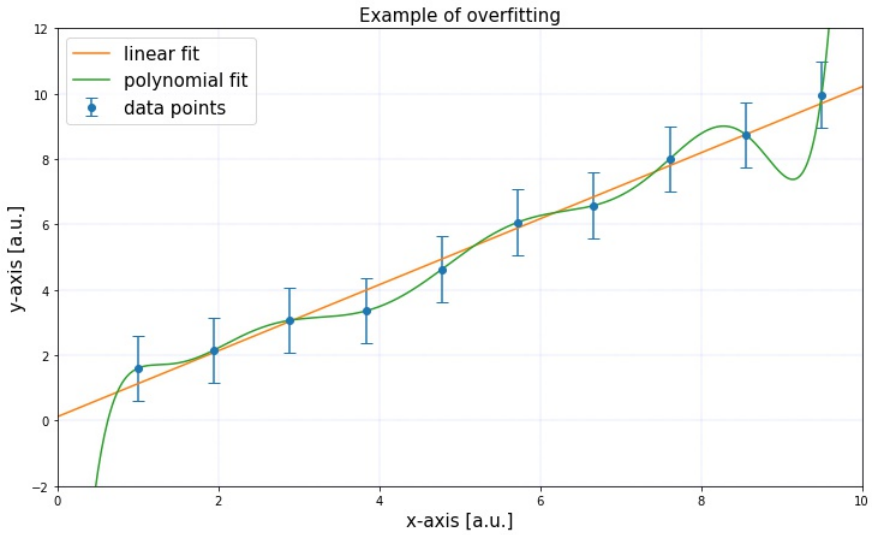


Fig. 5: Measurement data with some errors are fitted linearly and with a polynomial of degree nine. Assume that it is known that the true value of y is a linear function of x, e.g. from a law of physics. While the RMSE of the polynomial is zero and therefore certainly lower than the RMSE of the straight line, it will still be much worse at predicting other pairs (x,y) that have not been measured yet. Therefore, when the true relationship of y and x is not known, a method is needed that detects overfitting.

Overfitting means that a particular data set is fitted too closely and the fit may therefore fail to generalize well to data it has not been fitted to. The cause can be that the number of free parameters

of the model used exceeds the number necessary to capture the underlying structure well. Fig. 5 shows an example of overfitting.
A common method to avoid overfitting is the early stopping method [25], which was utilized for training. Before starting an iterative optimization algorithm, the data set is randomly split into two parts: A training data set, and a test data set with a ratio of roughly 10:1. The NN is only fitted to the training data set and has no knowledge of the test data. After every training epoch the RMSE is calculated separately for the training- and test data set. If the NN is improving to generalize, both errors should decrease with an increasing number of epochs. However, if overfitting happens, the RMSE of the training set will decrease further, while the error on the test data will start to increase. This is the point when the NN starts learning the specific details of the training set at the cost of getting worse at generalizing. The training of the NN should be stopped, because it has reached the point at which it is best able to generalize to unknown data. An illustration of the early stopping method is given in Fig. 6.

Fig. 6: Until training epoch 20, the RMSE for the test and training data decreases. While the RMSE for the training data keeps dropping, the RMSE of the test data eventually start to increase. From this point on the NN is learning the details of the training data set and becomes worse at generalizing. Training should be stopped.

3.2 Multistream extended Kalman Filter

A multistream variant of the extended Kalman Filter was used for training that has been introduced in [16]. This section will closely follow the given reference and is a summary of the parts relevant to the training process. The notation that has been chosen in this paper will be used unaltered.
Kalman Filters [26] are algorithms that estimate the unknown state of a linear dynamical system from a series of measurement data, taking statistical noise into account. A generalization to non-linear systems is given by the extended Kalman filter [27].
In the context of NN training, if the training patterns are given by $\{(\{\boldsymbol{x}_a\},\{\boldsymbol{y}_a^{\text{ref}}\})|\ a=1,...,M\}$, where M is the total number of patterns and $(\{\boldsymbol{x}_a\},\{\boldsymbol{y}_a^{\text{ref}}\})$ the input / output pairs, a can be used as an artificial time parameter t. With this parameter, the estimated weights (when updating the weights w.r.t. structure a corresponding to t) can be written was $\mathbf{w}(t)=[w_1(t),...,w_n(t)]^T$, which is the estimated state of the system at time t. The goal of training the NN is to minimize the root-mean-square error

$$\text{RMSE}=\sqrt{\frac{1}{M}\sum_{a=1}^{M}(\boldsymbol{y}_a^{\text{ref}}-\boldsymbol{y}_a)^2}, \tag{3.2.12}$$

which can be shown to be achieved by following the following algorithm:
At the beginning of training, $\mathbf{w}(0)$ is typically initialized randomly. The first step of one iteration is to calculate the error w.r.t. the current training pattern

$$\boldsymbol{\xi}(t)=\boldsymbol{y}^{\text{ref}}-\boldsymbol{y}(t), \tag{3.2.13}$$

where $\boldsymbol{y}(t)=[y_1(\boldsymbol{w}(t),\boldsymbol{x}_a),...,y_m(\boldsymbol{w}(t),\boldsymbol{x}_a)]^T$ is the NNs prediction of the values of the m output nodes corresponding to training pattern a and $\boldsymbol{y}^{\text{ref}}=[y_1^{\text{ref}}(\boldsymbol{x}_a),...,y_m^{\text{ref}}(\boldsymbol{x}_a)]^T$ the reference values. In the next step, the derivative matrix $\mathbf{H}$ is calculated

$$H_{ij}=\frac{\partial}{\partial w_i}\xi_j(t), \tag{3.2.14}$$

and afterwards the Kalman gain matrix $\mathbf{K}$ via

$$\mathbf{K}(t)=\mathbf{P}(t)\mathbf{H}(t)\Big(\frac{1}{\eta(t)}\mathbf{I}+\mathbf{H}^T(t)\mathbf{P}(t)\mathbf{H}(t)\Big)^{-1}, \tag{3.2.15}$$

where $\eta(t)$ denotes the learning rate and $\mathbf{P}$ the error covariance matrix.
For $\eta(t)$ an exponential increase can be used:

$$\eta(t)=\min(\eta_0 e^{\frac{t}{\tau_\eta}},\eta_{\max}). \tag{3.2.16}$$

Finally the weights and error covariance matrix are updated

$$\mathbf{w}(t+1)=\mathbf{w}(t)+\mathbf{K}(t)\boldsymbol{\xi}(t) \tag{3.2.17}$$

$$\mathbf{P}(t+1)=\mathbf{P}(t)-\mathbf{K}(t)\mathbf{H}^T(t)\mathbf{P}(t)+\mathbf{Q}(t) \tag{3.2.18}$$

and one iteration is concluded. $\mathbf{Q}$ has been introduced as artificial process noise which can be used to prevent being trapped in local minima [28,29]. One iteration through the whole data set is called one epoch and training a NN typically consists of many training epochs, which will eventually minimize the root-mean-square error. Additionally, the error covariance matrix needs to be set at the beginning of training, and is given by

$$\boldsymbol{P}(0)=\varepsilon^{-1}\mathbf{I}, \tag{3.2.19}$$

with ε between 10^{-2} and 10^{-3}. Similarly, the artificial process noise is given by

$$\mathbf{Q}(t) = q(t)\mathbf{I}, \tag{3.2.20}$$

where $q(t)$ can be constant between 10^{-6} and 10^{-2} or decrease over time such as

$$q(t) = \max\left(q_0 e^{-\frac{t}{\tau_q}}, q_{min}\right). \tag{3.2.21}$$

To be able to use the above algorithm, the error in 3.2.13 needs to be specified. Rather than using all the information contained in one training pattern at once, individual forces and energies are treated independently. This gives more control of the relative importance of force and energy training. Thus the errors are given by

$$\boldsymbol{\xi}(t) = [E_a^{\text{ref}} - E_a(\mathbf{w}(t))], \tag{3.2.22}$$

if an energy is chosen for a weight update, while it is

$$\boldsymbol{\xi}(t) = [\beta(F_{ai,\alpha}^{\text{ref}} - F_{ai,\alpha}(\mathbf{w}(t)))], \tag{3.2.23}$$

in case a single force in the α direction is chosen. The parameter β has been introduced and can be used to decide the relative importance of force to weight updates. Iterating this algorithm will minimize

$$\text{RMSE} = \sqrt{\sum_{a=1}^{M}(E_a^{\text{ref}} - E_a)^2 + \beta^2 \sum_{a=1}^{M}\sum_{i=1}^{N_a}(\vec{F}_{ai}^{\text{ref}} - \vec{F}_{ai})^2}, \tag{3.2.24}$$

where N_a is the total number of atoms. Note that the computational cost of this algorithm is dominated by the size of the error covariance matrix $\mathbf{P}$, which is of dimension n x n, where n is the total number of weights. In case a HDNNP is trained for more than one atom species, the computational cost can be lowered by treating different atom species independently. This is achieved by changing the errors to

$$\boldsymbol{\xi}^{\chi}(t) = \left[\frac{N_a^{\chi}}{N_a}(E_a^{\text{ref}} - E_a(\mathbf{w}(t)))\right], \tag{3.2.25}$$

in case of an energy update and

$$\boldsymbol{\xi}^{\chi}(t) = \left[\beta\frac{N_{ai}^{\chi}}{N_{ai}}(F_{ai,\alpha}^{\text{ref}} - F_{ai,\alpha}(\mathbf{w}(t)))\right], \tag{3.2.26}$$

in case of a force update. The index χ is introduced to refer to different atom species, N_a^{χ} is the number of atoms in structure a belonging to species χ, N_{ai}^{χ} is the number of local neighbors with element χ of atom i in structure a inside the cutoff radius and N_{ai} is the total number of atoms inside the cutoff radius. Similarly, equations 3.2.14-3.2.18 receive an index χ and whenever a weight update is done for the parameters of one atom species, only its weights are considered in the algorithm. This effectively reduces the dimensions of the matrices involved in the algorithm. Finally, the artificial time parameter t can be chosen differently to take the information of multiple training patterns into account simultaneously. For this purpose the error in equation 3.2.13 is rewritten as

$$\boldsymbol{\xi}(t) = [\boldsymbol{\xi}_1^T(t), ..., \boldsymbol{\xi}_s^T(t)]^T, \tag{3.2.27}$$

where s is the number of training patterns (or streams) used at time step t and $\boldsymbol{\xi}_a(t)$ the error vector for structure a. Changing the parameter s has an effect on the speed of the algorithm (since some parts of the algorithm can be parallelized) and potentially the quality of the fit (more information is included during each weight update for larger s). Together with the adaptations for multiple atom species, this is the algorithm that was used for HDNNP training.

4 A very brief introduction to density functional theory

A full introduction to DFT is beyond the scope of a master thesis, yet still the most important ideas are briefly summarized. The aim of DFT is to determine the quantum mechanical ground state of a system consisting of a number of electrons. Typically, the electrons are assumed to be in an electrostatic field created by nuclei in fixed positions (decoupling the motion of electrons from the nuclei is referred to as Born-Oppenheimer approximation [13] and justified by the fact that electrons move much faster than nuclei). While in principle the ground state can be determined by solving the Schrödinger equation, the equations of DFT can be solved with a significantly lower computational cost.
Rather than trying to determine the wave function of the system, DFT is based on the insight that all properties of the system can in principle be derived from the electron density, which only depends on three Cartesian coordinates independent of the number of electrons.
The theoretical foundation is given by the two Hohenberg-Kohn-theorems (HK1 and HK2) [32,33]. The energy of the electronic system is given by

$$E = T + U + V, \tag{4.0.28}$$

namely the kinetic energy, the potential energy from electron-electron interactions, as well as from electron-nuclei interactions. Their first theorem states that the electron density ρ uniquely determines the external potential V and vice versa. Therefore, also the Hamilton operator is uniquely determined by 4.0.28, and all other properties follow from the Schrödinger equation. In particular, this means that the energy of the system (as well as other observables) is a functional of the electron density $E = E[\rho]$ (hence the name density functional theory).
Their second theorem states that the energy functional is minimized by the electron density of the ground state, namely

$$E[\rho] \geq E[\rho_{\text{groundstate}}] \quad \forall \rho. \tag{4.0.29}$$

This means that finding the ground state electron density can be achieved by minimizing the energy functional. However, the exact form of $E[\rho]$ is in general not known.
A further key insight is that a system of N interacting electrons can be rewritten as a system of N non-interacting electrons moving in an effective potential with the same electron density (and therefore the same ground state properties by HK1). Thus, minimizing the energy functional can be shown to be equivalent to solving the Kohn-Sham equations of non-interacting particles [34]:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(r)\right)\psi_i(r) = E_i\psi_i(r), \tag{4.0.30}$$

where the electron density is given by

$$\rho(r) = \sum_i^N |\psi_i(r)|^2, \tag{4.0.31}$$

and the effective potential by

$$V_{\text{eff}}(r) = V(r) + e^2\int \frac{\rho(x)}{|x-r|}dx + \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(r)}. \tag{4.0.32}$$

The exchange-correlation energy functional $E_{xc}[\rho]$ has been introduced, which unlike the first two terms cannot be derived from classical physics, and is the only unknown part of the Kohn-Sham equations. There are many different approximations to E_{xc}, which depend on the application in mind. Assuming that a good approximation of E_{xc} is known, equations 4.0.30 - 4.0.32 can be solved in an iterative way: An initial guess of ρ is made and used to calculate V_{eff}. This is used to solve equations 4.0.30 and 4.0.31 to obtain an improved estimate of ρ. This procedure is iterated

until it converges.
Finally, the total energy of the system can be calculated via

$$E = \sum_{i}^{N} E_i - \frac{e^2}{2} \int dr' \int dr \frac{\rho(r')\rho(r)}{|r'-r|} + E_{\mathrm{xc}}[\rho] - \int \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta \rho(r)} \rho(r) dr. \tag{4.0.33}$$

Lastly, it should be noted that not only energies were obtained from DFT, but also the forces acting on the nuclei, following the already mentioned Born-Oppenheimer approximation. While the energy as a function of the coordinates of the nuclei $E(\mathbf{R})$ can be calculated with DFT, the forces are obtained by taking the negative gradient with respect to $\mathbf{R}$. Thus, the nuclei are treated classically, while the electrons are treated as quantum objects. This finishes our very short introduction to DFT and the next chapters will present the procedure and the results of the computations done for this thesis.

5 Extending the DFT data set

The approximation to the exchange-correlation energy functional used was the PBE functional [15]. All reference data was calculated with the Vienna Ab initio Simulation Package (VASP) version 5.4.4 [35] on the Vienna Scientific Cluster (VSC-4). This high performance cluster has about 800 nodes, each equipped with two processors of type Intel Xeon Platinum 8174 Processor @ 3.1 GHz (24 cores).
The HDNNP training was done with a previously presented C++ library n2p2 (neural network potential package) [16, 36] that includes an interface to the molecular dynamics software LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator) [37], which was used for simulations.
A previously developed data set for bulk copper sulfide containing 3965 configurations [16] was the starting point to developing a HDNNP that can reliably predict forces and energies of nano crystals in the low- and high chalcocite phase, as well as the transition regime. Since the existing data set was already able to cover the phase transition for the bulk material, this work only focused on adding surface configurations.
To ensure consistency between the already existing data and the newly generated data, identical VASP settings, including a plane wave cutoff of 600 eV and a 2 x 2 x 2 Monkhorst-Pack point-grid, were used.
First, a summary of the two HDNNP comparison method will be given, which was chosen to extend the data set.

5.1 The two HDNNP comparison method

When developing a new data set to create a HDNNP, a method is needed to decide which DFT structures are necessary to make the HDNNP robust and ready for large scale simulations. Since DFT calculations are computationally very expensive, it is desirable to make as few calculations as possible. The two HDNNP comparison method [38, 39] was used to detect structures that are not well represented in the already existing data set. The idea is rather simple and illustrated in Fig. 7: If two HDNNPs with different parameters are trained with the same reference data, one of them is used to run a MD simulation. The second HDNNP will also compute its energy and force predictions of the structures occurring in the simulation.
For structures where no similar reference data is available, the prediction of both HDNNPs can differ wildly. An energy (and/or force) difference threshold is chosen, and energies and forces for structures exceeding this threshold are calculated with the reference method (e.g. VASP in our case) and added to the reference data set. Both HDNNPs are retrained with the new reference data. Now their predictions will not differ as much as before anymore. An improvement has been made. One of them can be used to run a new simulation and the procedure is iterated. When the differences are small for all structures that are likely to occur in large scale simulations the HDNNP is ready for use.

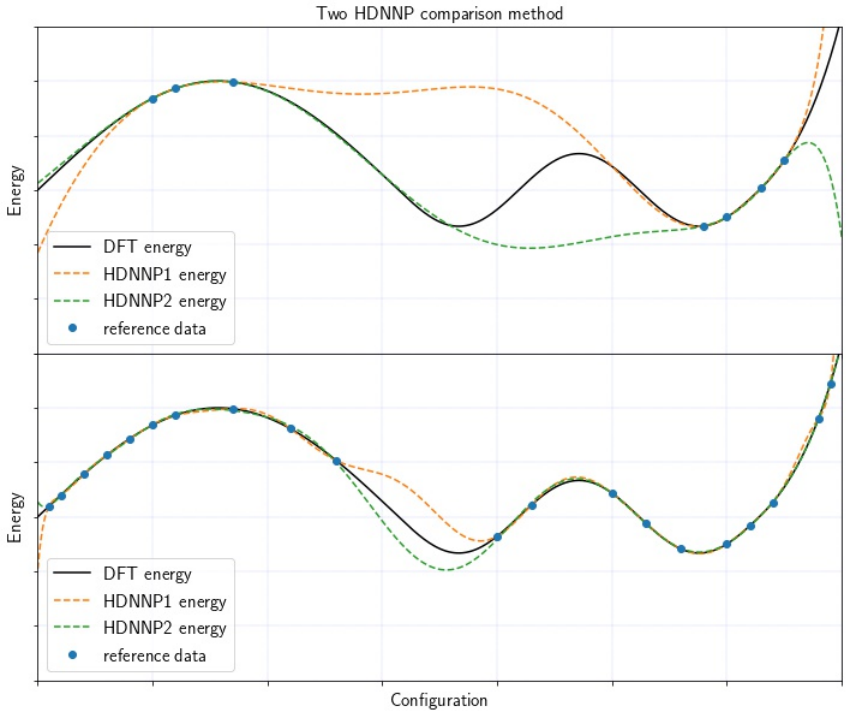


Fig. 7: This is an illustration of the two HDNNP comparison method. At the top two HDNNPs have been trained with very little reference data. For configurations where little reference data is available, their predicted energies can differ wildly. If data is added where the differences are high, the HDNNPs become much more reliable (bottom).

5.2 The reference data set

The simulations that one would eventually like to do consist of up to a couple of thousand atoms. However, DFT calculations are so elaborate that structures with a few thousand atoms would be far too costly. Thus, the reference data only contains much smaller structures of up to 144 atoms (one unit cell). This is justified by the fact that even in large geometries, atoms are only affected by their local neighborhood, which are only a couple of dozen atoms in the case of copper sulfide. Still, calculating the energy and forces for a single structure containing 144 atoms takes about 45 minutes on one node (48 cores) on the VSC-4. Since thousands of structures are needed to create a reliable HDNNP, the total computational cost amounts to a couple of hundred thousand core hours. Considering that small structures are needed to represent the surfaces of large structures, the question arises which structures are suitable for this task. A plausible starting point was to begin with the surfaces of the low chalcocite unit cell.

A total of three starting structures (slabs) were used: The atom arrangement was always taken to be the perfect LC unit cell, while the simulation box was extended in different directions, namely the x-direction, the y-direction and the direction of the tilted box vector (by roughly a factor two of the original size). Surfaces facing the void effectively behave as physical surfaces, which is usually not true under periodic boundary conditions. Fig. 8 shows the three starting configurations that were labeled Slab X, Slab Y and Slab Z (the name Slab Z was only used for convenience, even though it might be misleading).

The further procedure was as follows: From the ideal structures 100 structures each (for a total of 300 structures) were created through random displacements of the atoms by a maximum of 0.3 $\AA$, and energy and forces calculated with VASP. This, together with all the available bulk reference data, was used to train two different initial HDNNPs with different topologies (one had 27x27 nodes for the hidden layers, the other one 25x25). The symmetry functions used were identical to the ones previously used for the bulk HDNNP and a detailed list is provided in the appendix 11.2. Next, MD simulations were performed in the NVT ensemble, applying the two HDNNP compa-

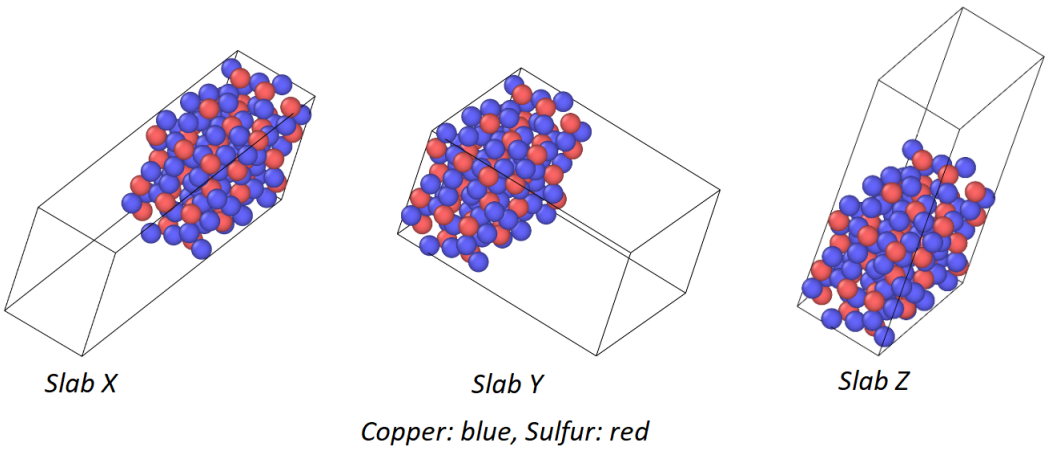


Fig. 8: These are the starting configurations for which DFT calculations were done.

rison method. The first simulations were done with a thermostat set to 350 K. Once structures for all slabs were selected (for a total of 1912), energy and forces were again computed by VASP and two new HDNNPs with all available data were trained. One iteration of selecting structures for all slabs, computing them and training new HDNNPs is called a run, and a sub run is only considering structure selection for one of the slab configurations. Subsequent runs were performed at different temperatures up to 700 K and typically consisted of less structures than the first run, since the HDNNPs improved quickly. The full list of all runs (also including slabs not mentioned so far) (15) and sub runs (23) is very extensive and therefore provided in the appendix 11.1. The RMSEs of the HDNNPs on the training (test) data set - approximately $1.0{\cdot}10^{-3}$eV/atom for energies and $6.0{\cdot}10^{-2}$ eV/$\AA$ for the forces - was always close the the RMSEs of the previously presented bulk HDNNP ($1.06{\cdot}10^{-3}$ ($1.08{\cdot}10^{-3}$) eV/atom for energies, and $5.54{\cdot}10^{-2}$ ($5.53{\cdot}10^{-2}$) eV/$\AA$ for forces).

After four runs a very short (0.1 ps, T=400 K) test MD simulation of a cylinder with roughly 4000 atoms was done, to check how often the HDNNP had to extrapolate (each time issuing an extrapolation warning (EW)). This number was very large, and looking at how far the HDNNP had to extrapolate (i.e. looking at the value symmetry functions took when extrapolating compared to the minimum/maximum values contained in the already available data set) clearly revealed that the three mentioned slabs are not sufficient to provide the reference data needed for nano crystals.
More precisely, MD simulations of all slabs were done up to T=700 K and their maximum/minimum values for all symmetry functions determined and compared to the maximum/minimum values of the symmetry functions of the cylinder. The values of the cylinder were sometimes a few orders of magnitude lower than the ones in the slab simulations, clearly demonstrating the need for more reference structures. A word of caution: An initial mistake that was made was not equilibrating the cylinder. Unequilibrated structures take extreme values for symmetry functions, which does not point to an insufficient HDNNP. After a short equilibration, the extrapolations of the cylinder were far less extreme, but could still not be covered by the slabs.
A number of short MD simulations up to roughly 700 K were done with different small geometries, inspecting their symmetry function distribution. In principle, geometries that are not slabs, but surrounded by surfaces (i.e. no periodicity at all) tend to have the lowest values, but are too expensive to be solved by VASP. This is because the electron density has to be computed over the full simulation box, even in places where it is empty. (A structure with no periodicity in a sufficiently large simulation box of 160 atoms took roughly six hours to be solved, compared to 45 minutes for the slabs)
A good compromise was to take slightly smaller geometries and keeping only one direction periodic, thus effectively introducing manageable geometries with edges and curvature. Fig. 9 shows

the last two structures that were used to produce reference data. The procedure for producing refe-

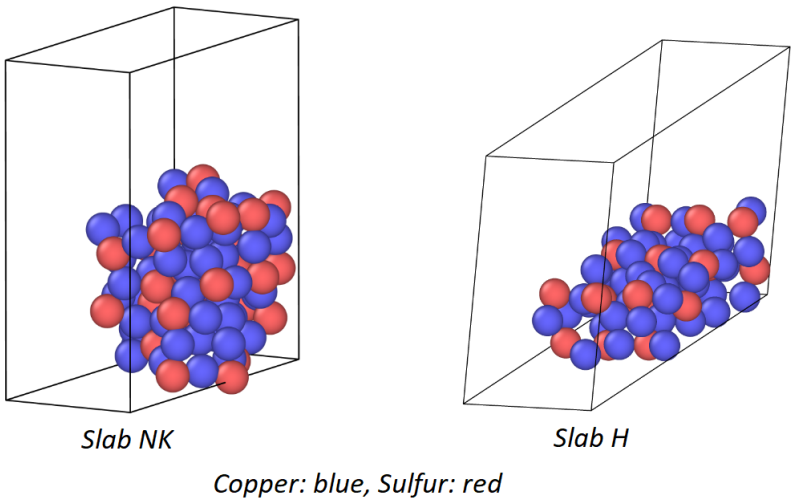


Fig. 9: These are two additional structures that include edges and curvature at higher temperatures. Slab NK was cut out off the unit cell and has dimensions 1.0 nm x 1.1 nm x 1.117 nm. It has a total of 93 atoms, namely 62 copper atoms and 31 sulfur atoms. Its x period (1.0 nm) is slightly off compared to the x period of the unit cell (1.5246 nm) and it was equilibrated before using it as a starting configuration. On the other hand, Slab H has the same x period as the unit cell and is simply the unit cell cut in half w.r.t. the z- direction. It consists of 72 atoms, 24 of which are sulfur atoms and 48 are copper atoms.

rence data of Slab NK and Slab H was completely analogous to the other slabs. There is one thing worth mentioning:
Slab NK was added before Slab H, and there was originally no intention to add Slab H. This was because, once Slab NK data was added to the reference data set, the two HDNNP comparison method selected *more* Slab X-Z structures than before, despite having comparably low RMSEs. Since there were a few odd things about Slab NK (its x-period is slightly different from the true x-period of low chalcocite, and it was trained with a time step of 10 fs) another structure was added to test if Slab NK lead to inconsistencies in the data set and could be replaced by a better structure. However, adding substantially more reference data for all five structures resolved all strange behavior and ultimately all geometries were kept. The resulting HDNNP with the full reference data set is consistent on all geometries, as well as cylinders.
The next section will cover all reliability tests that have been made for the final HDNNP.

5.3 The final HDNNP and reference data evaluation

For a final and thorough analysis of the reference data, two HDNNPs with a slightly different hidden layer (HL) topology (30x30, 32x32) were trained and a number of tests were made. In both cases 90% of the entire reference data set was used for training. (e.g. 90% from all 3965 bulk configurations, as well as the newly generated 10577 surface configurations, for a total of 14542). The remaining 10% of structures (1454) were used to detect possible overfitting. A list of the relevant HDNNPs parameters used for training is provided in the appendix 11.2. The first test is of course given by the RMSEs on the test and training data set:

HL topology	RMSE train-/ test data	
	E/atom [meV]	Forces [meV/Å]
30 x 30	0.945 / 1.02	58.2 / 59.5
32 x 32	0.935 / 1.02	58.4 / 58.6

Comparing these values with the RMSEs on the bulk HDNNP as presented in [16] shows an only marginal increase of the RMSE for forces (55.4 / 53.4), and a small decrease for energies (1.06 / 1.08). Thus, the symmetry functions that were used for bulk training can also be used for representing surfaces with a comparable RMSE. We proceed with showing the performance of a two HDNNP comparison on the structures in the data set. Five short MD simulations (10 ps, with a time step of 1 fs) were done with the smaller HDNNP, and energy and force prediction compared with the larger HDNNP. The results are shown in Fig. 10.
ΔF_{mean} is far below the RMSE for all structures without exception, which is very reassuring, since this is the value that matters for MD simulations. Overall, the HDNNPs have slightly higher differences for Slab H and Slab NK, which is likely due to the greater amount of reference data available for the quite similar structures Slab X-Z.
The more important question is how well the HDNNPs perform on geometries of interest: nanorods. As a test, a short MD simulation (shown in Fig. 11) of a cylinder was done, the two HDNNP method applied and extrapolation warnings collected. Energy and force differences are extremely similar to the slightly better results of Slab X-Z and hardly any extrapolation warnings are left: The HDNNP seems to be ready for use.

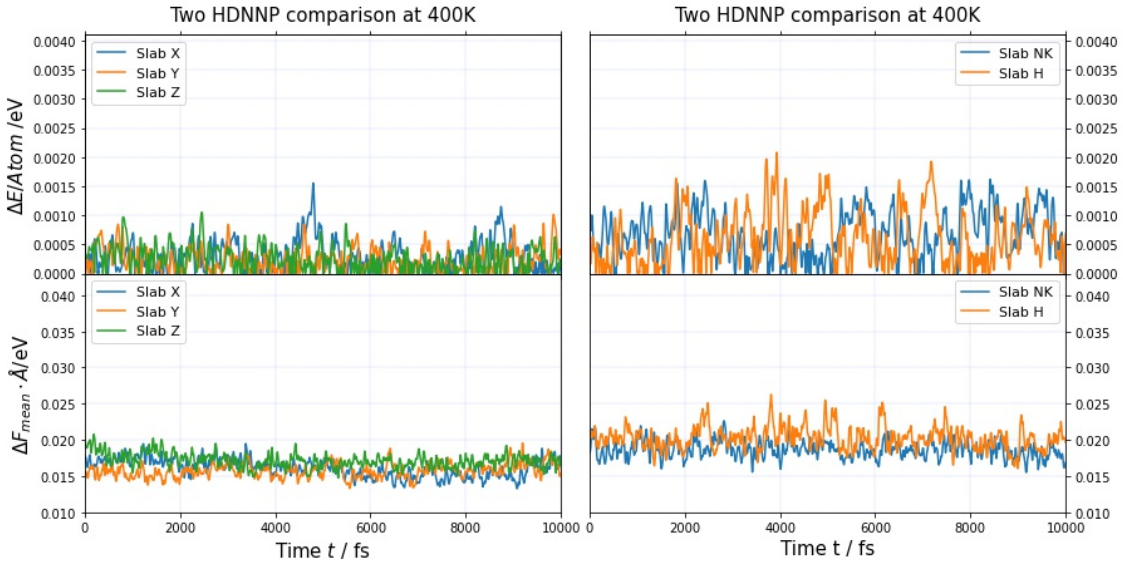


Fig. 10: This is a two HDNNP comparison between the two final HDNNPs. The MD simulations were done at 400 K.

As a final test, the radial distribution functions of a simulation done with VASP were compared to a simulation driven by the HDNNP. The excellent match in Fig. 12 shows that the HDNNP closely reproduces the reference method.

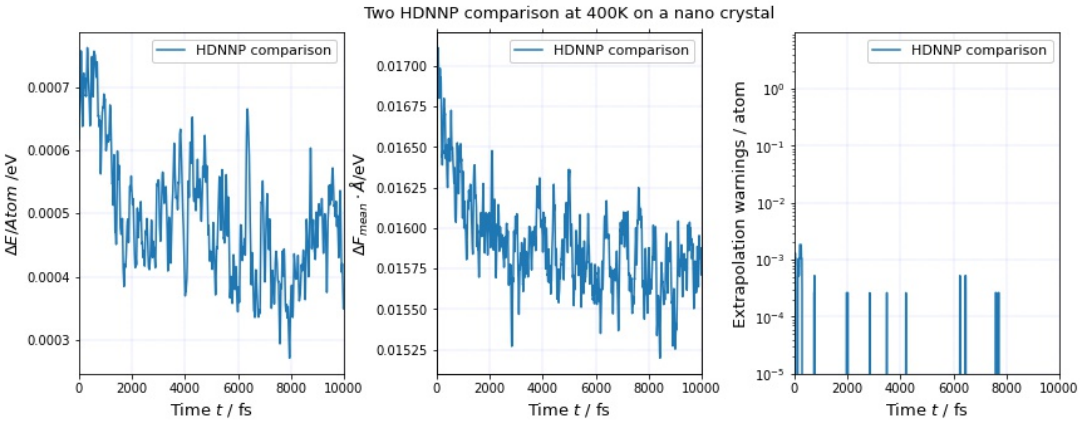


Fig. 11: A short MD simulation of a cylinder with a diameter of 3.5 nm, 6 nm length and 3792 particles was done. No periodicity along the symmetry axis of the cylinder was turned on, which means that the ends of the cylinder also acted as surfaces and edges were included. This is a difficult and realistic test for the HDNNPs. The extrapolation warnings (EW)/ atom were obtained every fs by taking the total number of EWs at this femto second and dividing it by the number of atoms (3792).

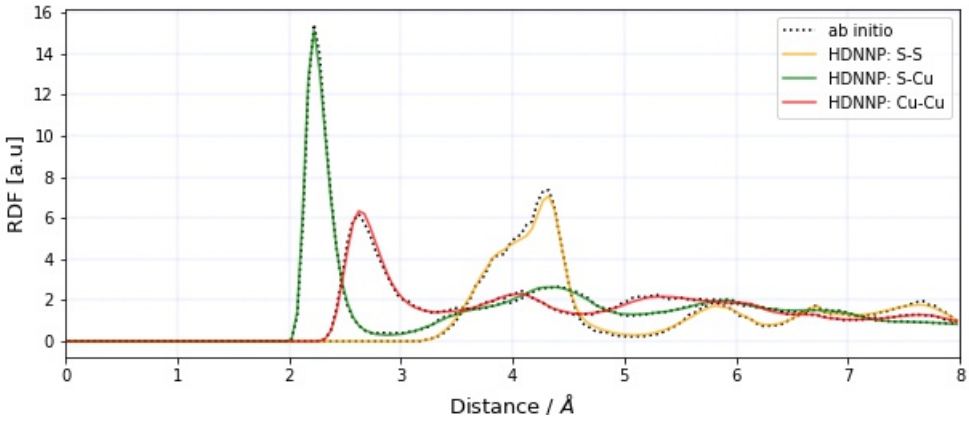


Fig. 12: The final positions of atoms after a 2n s long equilibration of Slab X at 300 K was used as a starting configuration for an ab initio MD simulation (2 ps with a time step of 1 fs directly done with VASP) and another 2 ns MD simulation with the HDNNP. Both simulations had identical time step and thermostat settings (300 K). From the 2 ns run, one snapshot was extracted every ps to calculate the radial distribution functions. The RDFs have been calculated with the python library MDTraj [40].

A brief justification of why Slab X was chosen to be compared to the reference method is given: Evidently, Slab X-Z all showed very similar performance, thus all of them are considered to be roughly equivalent and choosing Slab X among these three had no particular reason. It was chosen over Slab H and Slab NK, because the energy and force differences of the cylinder suggest that Slab X-Z are structures are used more to simulate the cylinder than Slab H and Slab NK structures. Since the HDNNPs appear to be reliable, the slightly smaller one (HL topology of 30 x 30) was used to run large scale simulations of different nanorods. The extrapolation warnings for eight long simulations show that the HDNNP can also handle long simulations of cylinders with varying curvature, length and temperature as well as during the phase transition; they are included in the appendix 11.3. Before focusing on actual simulations, we investigate if the HDNNP can also be trusted in the liquid phase.

5.4 Beyond solid copper sulfide - prospects of the liquid phase

After the successful creation of a HDNNP for copper sulfide with reference data that contain structures up to a temperature of 700 K, a possible question is: How well does the HDNNP handle the liquid phase, and how much effort would it be to extend the reference data to encompass it? To gain some insight, one of the configurations (Slab NK with 93 atoms) was heated up to 1500 K. A visual inspection of the trajectory revealed that the material is clearly liquid at this temperature. The liquid Slab NK configuration was used to start a short MD simulation at 1500 K, with 10000 time steps. A total of 24738 extrapolation warnings were accumulated during this simulation, which amounts to 0.0266 EW / atom / time step on average.
The energy and forces were again compared with the second HDNNP to investigate the consistency of the predictions. The average differences on all configurations in the simulation were

$$\Delta E/\mathrm{Atom} = 0.922\mathrm{meV}$$
$$\Delta F = 0.0379\frac{eV}{\mathring{A}}.$$

These results are better than expected, since both values are below the RMSEs (0.945 (1.02) meV, 0.0582 (0.0595) $\frac{eV}{\mathring{A}}$ for the training (test) data) .
Finally, all maximum and minimum values symmetry functions took in the resulting trajectory were computed with the tool nnp-scaling from the already mentioned n2p2 library and compared to the maximum / minimum values symmetry functions take on the full reference data set. This allows for a visual inspection of how far the HDNNP had to extrapolate in the liquid phase and also gives a clue of how much more reference data would be needed. However, since different symmetry functions can take very different values and a total of 138 were used, the information was compressed in a suitable way:
For all of the 138 symmetry functions the following values were calculated

$$\alpha_i := \frac{\mathrm{Reference}_{min,i}}{\mathrm{Liquid}_{min,i}} \tag{5.4.34}$$

$$\beta_i := \frac{\mathrm{Liquid}_{max,i}}{\mathrm{Reference}_{max,i}}, \quad i = 1,...138, \tag{5.4.35}$$

where $\mathrm{Liquid}_{min,i}$ is the lowest value symmetry function i took on the 1500 K simulation, $\mathrm{Reference}_{min,i}$ is the the lowest value occurring in the entire reference data set etc. A value of $\alpha_i > 1$ means that the HDNNP had to extrapolate below the lowest value in the reference data, while $\beta_i > 1$ corresponds to an extrapolation above the highest value. Values below one mean no extrapolation happened. These values are shown in Fig. 13

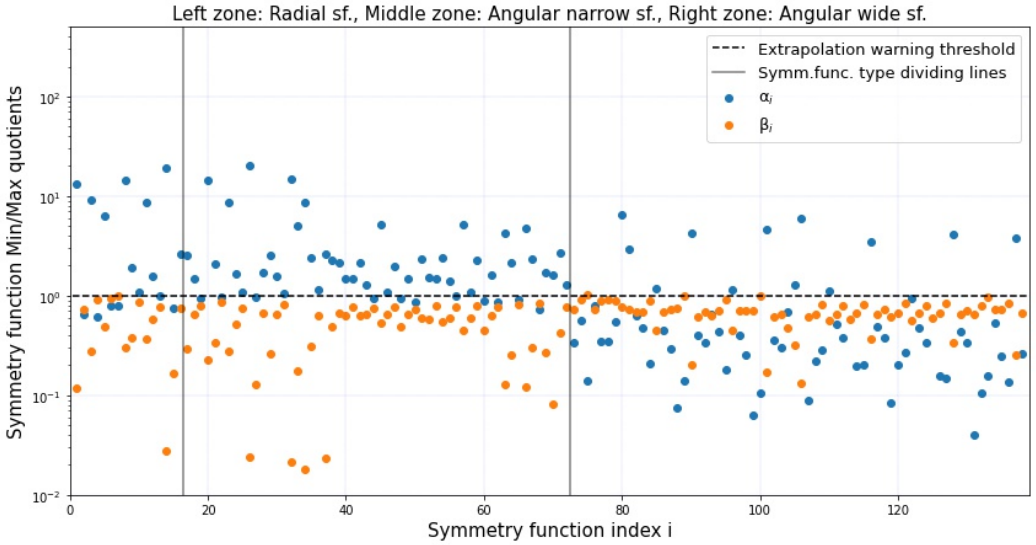


Fig. 13: This figure shows the extrapolations of the HDNNP for a short (10 ps) simulation of Slab NK at 1500 K, where the material is evidently liquid. Additionally, the symmetry functions are organized into the three different types presented earlier: Dots left of the first pale black line belong to radial symmetry function, between the first and second line they belong to the angular narrow symmetry functions, and after the second line they belong to the angular wide type of symmetry function. The symmetry function indices are precisely in the same order as presented in Table 10 in the appendix.

Extrapolations above the maximum are very rare, which can easily be explained by the fact that bulk configurations have the highest density and most local neighbors on average. Since symmetry functions are defined as sums over the local neighbors, bulk structures lead to the highest values. On the other hand, a large number of symmetry functions had to extrapolate below the minimum, in some cases a factor of over 20 below the minimum in the reference data was reached. While the values for ΔE/Atom and ΔF look very promising, Fig. 13 suggests that a substantial number of additional structures would be needed in the reference data set to reliably represent the liquid phase.
These results suggest that the HDNNP could still be improved to cover the full phase diagram and should not be trusted in the liquid phase. Nonetheless, training was stopped at this point, since the phase transition of interest is hundreds of degrees below the melting point.

6 First set of simulations

With the HDNNP in hand, an initial eight simulations of nanorods (cylinders) were carried out to investigate the phase transition for different sizes. All simulations were done with n2p2 library combined with its interface to lammps. Additionally, the python library MDAnalysis [41, 42] has been used extensively for various analyses on the resulting trajectories.
A time step of 1fs was used with a Nosé-Hoover thermostat continuously ramping up the temperature at a rate of 10 K / ns in the NVT ensemble.

6.1 Starting configurations

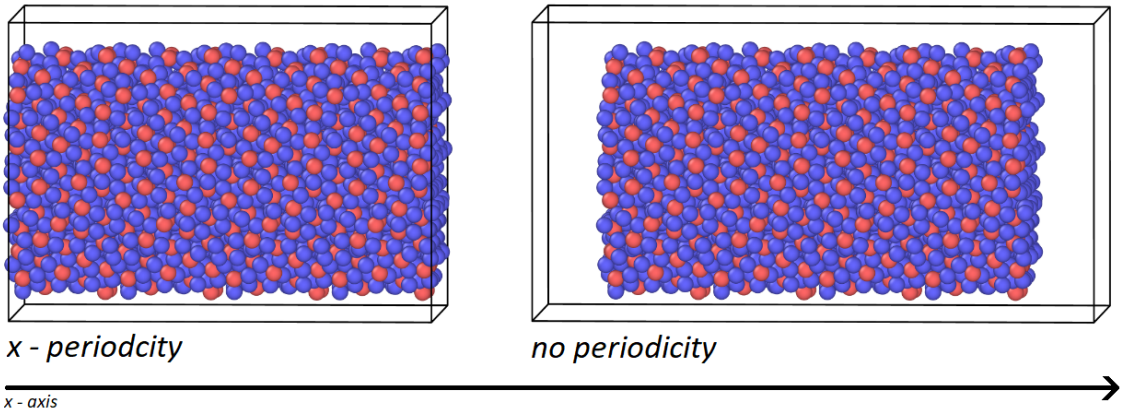


Fig. 14: This is a short illustration between the distinction of periodic structures (left) and non-periodic structures (right). In the first case periodic boundary conditions apply to the x - axis, while in the latter case fixed boundaries were used.

There are two major factors that distinguished the individual simulations from each other: Atom number and x - periodicity as shown in Fig. 14. For the periodic structures the length is an exact multiple of the natural x period. The largest structures contained approximately 4000 atoms. Since a 12 ns simulation of 3800 atoms on 480 cores took already about 100h in real time, cylinders with more atoms were not simulated. The smallest structures have around 1000 atoms. Originally smaller structures were simulated as well, but they deformed at higher temperatures, which makes it very difficult to track the copper diffusion; hence, they were avoided.
Table 1 contains all the details of the structures that were used to run simulations.
To create the initial geometries, the ideal low chalcocite unit cell was replicated in all directions (with lammps) to a geometry of a few hundred thousand atoms. From this large piece of Cu2S, cylinders were cut with a short python program. Such a procedure typically does not lead to a perfect 2:1 ratio of copper:sulfur atoms. In order to compensate the deviation of Cu2S, the radii were varied slightly to find geometries where the ratio is off by at most 10 copper atoms. From the resulting structures, copper atoms were either removed from, or placed slightly above the surface to achieve a perfect 2:1 ratio. This can be done in a similar way as classifying surface atoms.
The precise locations at the surface were always chosen with a random number generator.
It should be remarked that this can lead to serious problems. If copper atoms are directly added to the surface (and not above the surface) this can lead to a local explosion at the surface, where atoms violently fly away from the cylinder. Typically, at the very beginning of the MD the number of extrapolation warnings are higher until the cylinders are in a stable configuration, but remain low from this point on (including during phase transitions).

# atoms	periodic	diameter / nm	length / nm	surf / vol ratio	surf / vol ratio 2
1062	✓	2.6	3.0492	0.293	0.720
2124	✓	2.6	6.0984	0.293	0.720
3069	✓	3.6	4.5738	0.219	0.554
3816	✓	3.49	6.0984	0.219	0.572
1065		2.3	3.9	0.396	0.839
1974		3.0	4.18	0.336	0.745
2952		3.1	6.0	0.299	0.703
4089		3.0	8.7	0.292	0.693

Table 1: These are the details of the starting geometries. The surface / volume ratio was calculated by counting all atoms that have a distance of up to 2 $\AA$ from the surface and dividing it by the total number of atoms. Similarly, the surface / volume ratio 2 counts the number of atoms that are at most 6 $\AA$ far away from the surface. These numbers were calculated as well, because geometries with a significantly different number of atoms, but a similar diameter, had remarkably similar copper diffusion and energy curves (which can be seen for example in Fig. 15) and a surface dependence was assumed. The algorithms used are explained in the appendix 11.4.

Except for the periodic structures, which had x - periodicity enabled, all other boundaries were fixed. Thus, atoms could not influence each other across boundaries, but the simulation box was still large enough to ensure that no atoms were lost. The two following subsections will show the results of all simulations.

6.2 Thermodynamic results periodic structures

The periodic structures were heated up from T=320 K to T=450 K over a period of 13 ns. Some of the structures were equilibrated prior to the actual simulation to allow the material to find a stable surface configuration. Information on these pre-simulations can be found in the appendix 11.5. Furthermore, keeping one direction periodic effectively exerts pressure on the material since it cannot expand in this direction. The average pressures over the entire simulations were

# atoms	average pressure / bar
1062	167.79
2124	140.31
3069	527.51
3816	527.22

Table 2: This is the average pressure on the nanorods over the entire simulation.

In order to evaluate if a phase transition happened, first the mean square displacement (MSD), copper diffusivity and energy / atom were calculated over the entire trajectory. For each diffusion coefficient, a sub trajectory of 500 ps length was extracted from the entire trajectory (the length was chosen to make the standard error of the diffusion coefficent low). The resulting data was fitted linearly, and the slope divided by six taken as the diffusion coefficent. Error bars have been plotted as well and show the standard error of the slope of the linear fit (divided by six), which in case of a one dimensional linear regression is given by [43, 44]

$$\frac{\frac{\sum(y_i-\tilde{y}_i)^2}{n-2}}{\sum(x_i-x)^2}, \tag{6.2.36}$$

where n is the total number of know data point pairs (x_i, y_i), x is the average value of the x_i and $\tilde{y}_i$ is the predicted value for y_i by the linear fit. The resulting errors are so tiny, that upper and lower cap have no distance to each other on the scale shown. This might lead to skepticism, thus a graphic showing all fits for a 12 ns simulation (including diffusion with respect to zones as explained in later chapters) was added to the appendix (Fig. 34), which indeed shows an excellent agreement for all 96 fits in that trajectory. The resulting data is displayed in Fig. 15.

First of all, the two larger structures show very clear signs of a phase transition at a temperature of about 410 K. The mean square displacement shows a clear bent at this temperature, and also looking at the energy / atom and diffusivity reveals a drastic increase.
However, the situation is very different for the two smaller geometries. While the diffusivity increased as well, it is not evident that this happened due to a phase transition. There is no sign of an energy step, nor did the diffusivity increase rapidly. Also, despite the increase, the diffusivity always stays well below the results for the larger structures.
The data might be explained by atoms gaining more energy and mobility because they are heated, in the absence of a phase transition. This certainly requires an explanation, which has been discovered throughout a second set of simulations (the following chapter), and will be given at the appropriate point.
A very important observation is that the MSD and diffusivity for the two smaller structures is nearly identical, the same holds true for the two larger structures (and in the case of the latter two the temperature of the phase transition). How can this be? This observation lead to calculating the surface / volume ratio of these structures, which indeed shows that the pairs of structures behaving so similarly indeed have an almost identical surface / volume ratio (Table 1).

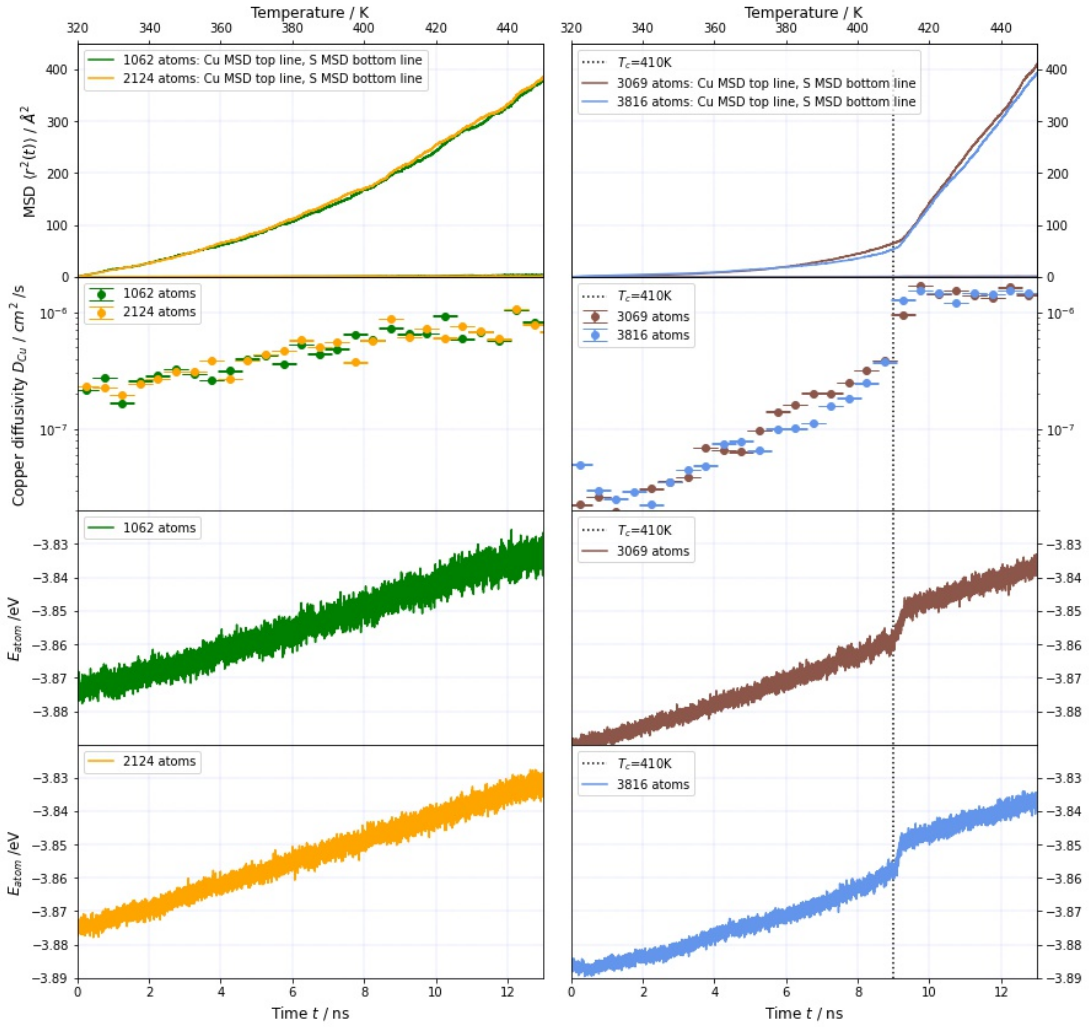


Fig. 15: This is the thermodynamic data that was calculated to investigate the phase transitions of the periodic nanorods.

Before any deeper explanations are looked for, we look at what happened when periodicity was entirely removed.

6.3 Thermodynamic results non-periodic structures

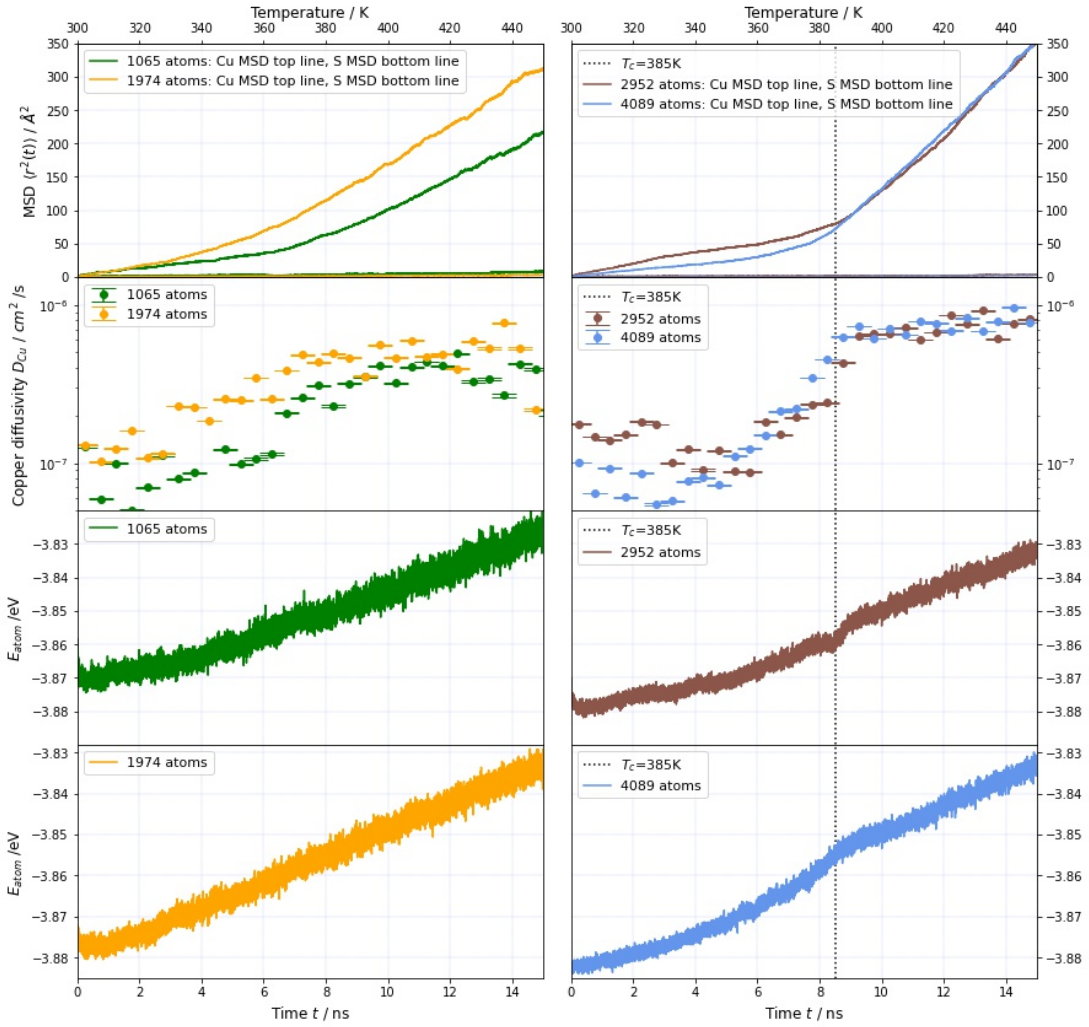


Fig. 16: This is the thermodynamic data that was calculated to investigate the phase transitions of the non periodic nanorods.

The non-periodic structures were heated from 300 K to 450 K over a period of 15 ns. When periodicity is completely removed, the results are a little bit different. First of all, the average pressure is now zero, and therefore much more comparable to atmospheric pressure than for the periodic structures. Looking at the two larger structures, the phase transition is not as sharply visible as before, and the values of the diffusion constant always stay below 10^{-6} cm^2 / s, unlike in the periodic case. It is also evident that the temperature of the phase transition has dropped to about 385 K and is hence 25 K lower than with periodic boundary conditions. This is probably explained by the already mentioned much higher pressure of over 500 bars in the latter case.

The two smaller structures do not show a clear sign of a phase transition. It is important to note that the two larger structures show nearly identical MSD and copper diffusion at higher temperatures, while this is not true for the small structures. Looking at the surface / volume ratio reveals once again that the two large structures have a nearly identical surface / volume ratio, while this is not true for the small structures in this case. Next, the radial distribution functions were calculated at different temperatures to check if the structure of the material has changed, revealing a phase transition.

6.4 Radial distribution functions for the first set of simulations

First of all we look at how the radial distribution functions (RDFs) changed from low temperatures to high temperatures in all simulations. Any major change would reveal a restructuring of the atoms.

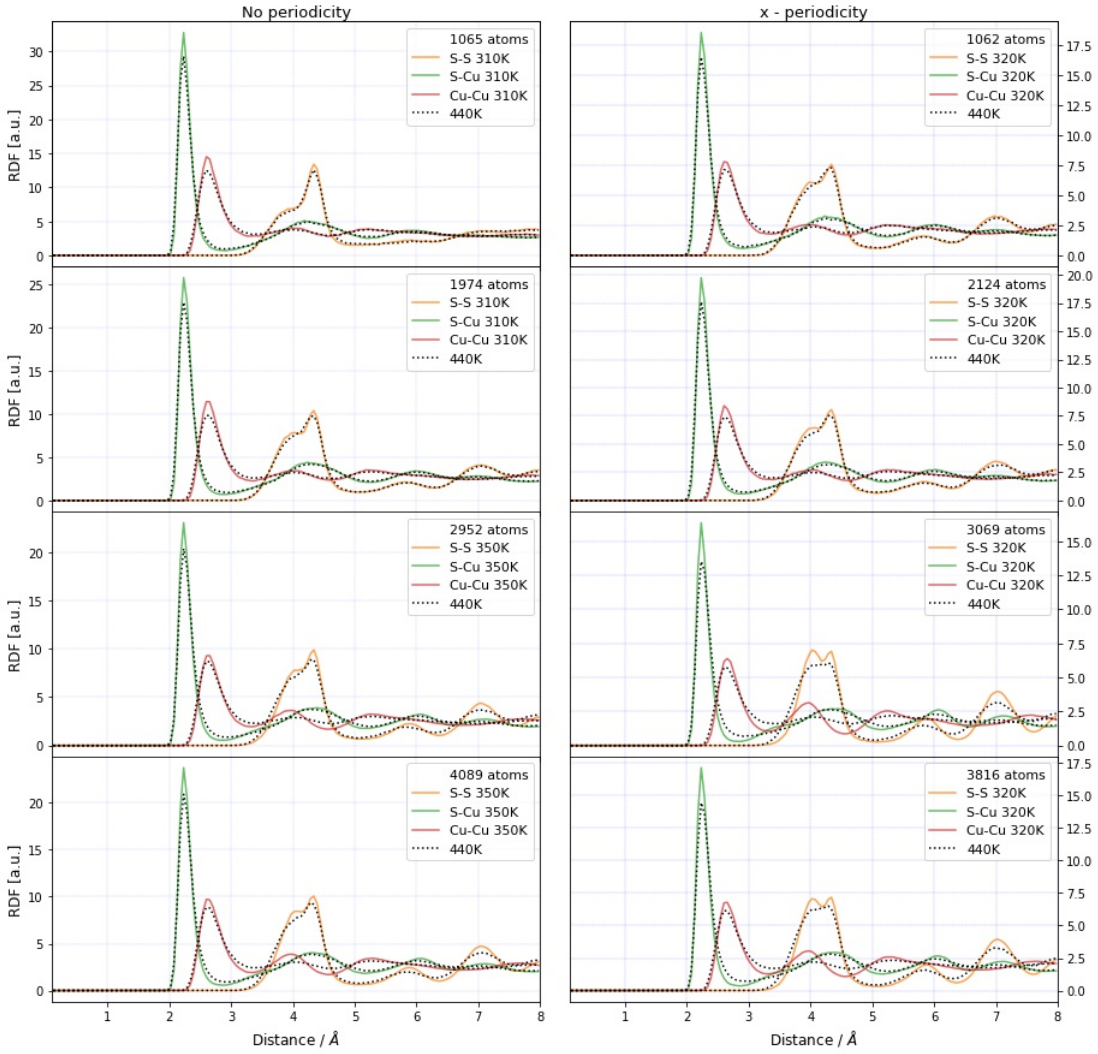


Fig. 17: This is a comparison of radial distribution functions at low and high temperatures for all simulations. Non-periodic structures are shown on the left, periodic structures on the right. All RDFs have been averaged over a period of 25 ps.

The results can be summarized as follows:
For the larger structures (2952, 4089, 3069, 3816 atoms) there is a clear change in all RDFs, which is more pronounced for the periodic structures. This is particularly visible at the second peak of the Cu-Cu RDF, which gets notably lower, as well as the two peaks of the S-S RDF at roughly 4 $\AA$ which are flattened. This is further evidence that in these cases a phase transition did indeed happen, while it remains to be answered why it is less pronounced for non-periodic structures. Only small changes are visible for all other structures and it will later become evident that there are reasons why indeed no phase transition happened.
Lastly, we check that the transformation of the RDFs for the two large periodic structures is indeed visible at the onset of the energy ramp at 410 K, by making a comparison over a very short time interval (300 ps, which amounts to a temperature difference of 3 K):

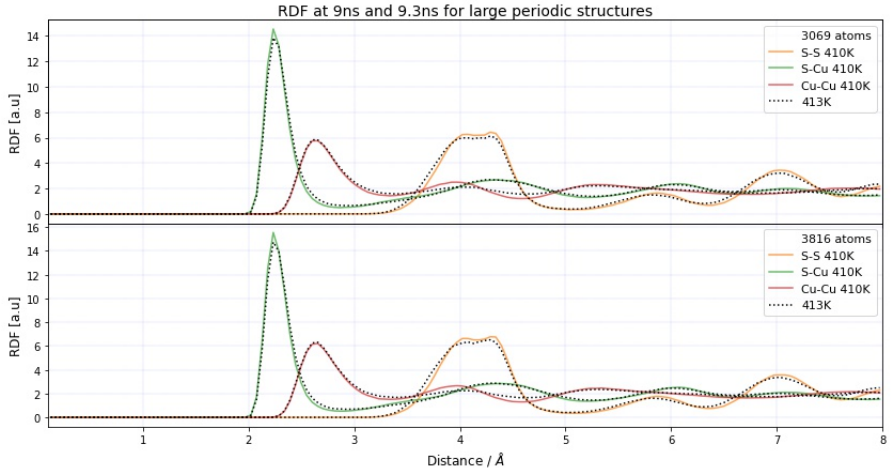


Fig. 18: The RDFs at 9 and 9.3 ns (410 / 413 K) are displayed at the temperature of the phase transition (they were averaged of 25 ps). The flattening of the two sulfur peaks at 4 $\mathring{A}$ as well as changes in the Cu-Cu RDF are visible in a very short temperature range, demonstrating that the transformation happened during the phase transition. Plotting RDFs at different temperatures below 400 K revealed no noticeable changes.

Even the flattening of the two peaks at 4 $\mathring{A}$ of the S-S RDF can be noticed, suggesting that a small restructuring of the sulfur grid is indeed observed (as expected).
The first set of simulations raised important questions: Why is the phase transition less pronounced for the two large non-periodic structures and more importantly, why is it not observed for the small ones? The key insight that drove the further investigation was the similarity in MSD and diffusion coefficients in structures with nearly identical surface / volume ratio, as well as the similarity in phase transition temperatures.

7 Second set of simulations

# atoms	periodic	diameter / nm	length / nm	surf / vol ratio	surf / vol ratio 2
1023	✓	3.6	1.5246	0.219	0.554
3816		3.5	6.039	0.275	0.657
4167	✓	4.2	4.5738	0.183	0.489
4248	✓	2.6	12.1968	0.293	0.720

Table 3: Details of starting geometries for the second set of simulations

# atoms	average pressure / bar
1023	484.41
4167	693.34
4248	148.66

Table 4: This is the average pressure on the nanorods from 380-450 K. For the non-periodic structure it is zero.

The hypothesis that was the foundation of the second set of simulations was that the absence or presence of a phase transition, the copper diffusion, as well as how strongly pronounced it is, depends on the surface / volume ratio. To test it, a geometry with 1023 atoms and the same surface / volume ratio as the two large periodic structures of the last section was simulated. If the hypothesis is true, a very similar phase transition should be visible for this geometry. Furthermore, a non-periodic structure with 3816 atoms, but a lower surface / volume ratio than all other non-periodic structures, was simulated. The changes caused by the phase transition should be stronger than for all other non-periodic structures, if the hypothesis holds. Two more periodic structures with roughly 4000 atoms were simulated. One of them has the same surface / volume ratio as the small periodic structures (1000, 2000 atoms) that did not show a phase transition (hence a phase transition was not expected), and the other one has the lowest surface / volume ratio of all structures. The simulations were started at 380 K.

First we look at the energy curves and the change in radial distribution functions in Fig. 19. The results are indeed as expected, there is a clear phase transition for the structure with only 1023 atoms, no phase transition for the structure with 4248, but a very high surface / volume ratio, a strongly pronounced phase transition for the structure with the highest surface / volume ratio (4167 atoms) and the highest energy step of all non-periodic structures for the last structure. Looking at the RDFs confirms this statement.

Note that the 1023 and 3816 atom structures were slowly heated prior to the shown simulation, while the other two were started from an unequilibrated state (e.g. directly started form a cut cylinder). The reason for this pre-heating is interesting in itself and will therefore be discussed (together with the full thermodynamic data of the pre-heating phase) in a separate chapter 9.2.

A total of seven phase transitions from low to high chalcocite were available at this point and used for a comprehensive analysis. This analysis lead to much insight and will therefore be done in its own chapter. The absence of the phase transition in some simulations will be investigated in a later chapter 9.

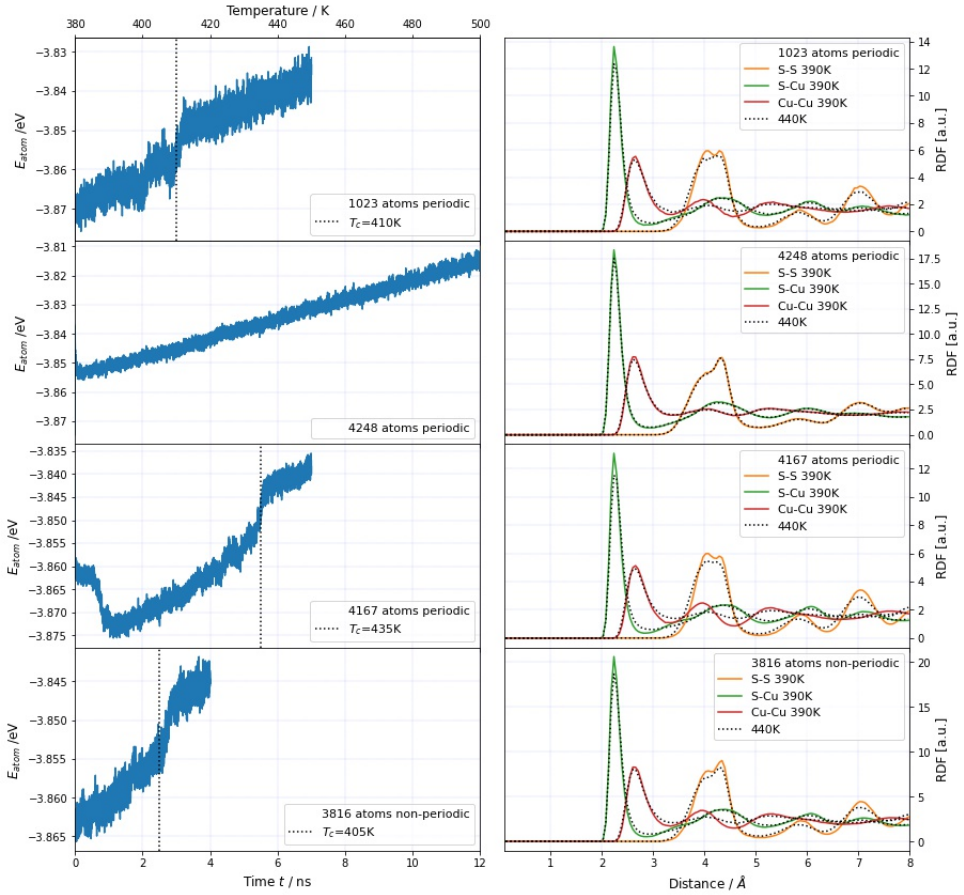


Fig. 19: This figures shows all energy curves and a RDF comparison for the second set of simulations.

8 Analysis of the phase transition

8.1 Analysis with respect to zones

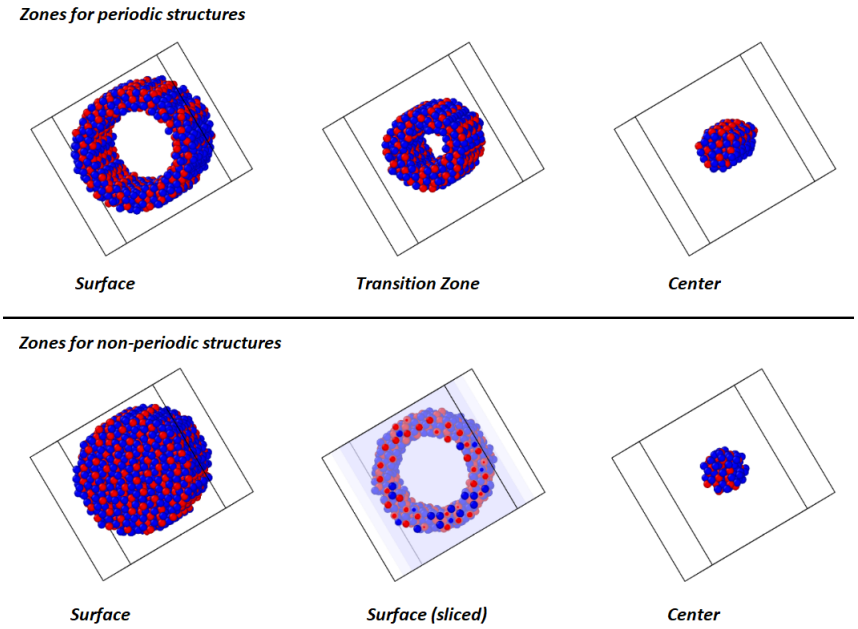


Fig. 20: The top row shows what zones look like for the periodic case. The first two pictures of the second row show the surface of a non-periodic structure. In this case a whole cylinder mantle is needed, since the ends of the cylinder act as surfaces as well. The Surface (sliced) picture is actually the exact same object as the Surface picture for non-periodic structures, but visually sliced open (the planes were it was sliced are visible), to reveal that it is actually hollow inside. The transition zone for a non-periodic structure is of course identical to the surface, but smaller, and the center is a solid cylinder.

From this point on the behavior of atoms at the surface and non-surface was investigated much more closely for all structures where a phase transition had been observed. For this purpose, zones for cylinders were defined: center, transition zone and surface, taking into account that periodic and non-periodic structures have to be treated differently. An illustration of these zones can be seen in Fig. 20 and the algorithm used to calculate them can be found in the appendix 11.4. We start by looking at the diffusion with respect to zones. Note that atoms can change their zones and thus the zones were recalculated every 500 ps. Similarly, diffusion coefficients were calculated every 500 ps. The MSD has been reset to zero when new zones were calculated to avoid correlations with the the rest of the trajectory. A visual example of this fitting procedure w.r.t. zones as well as its accuracy is given in the appendix (Fig. 34).
The MSD w.r.t. zones was originally calculated as well, but discarded since it is not well defined. Particles change their zones, thus they should be updated regularly. However, each time zones are updated there are strong discontinuities in the MSD, since the MSD of particles entering zones is abruptly added (or subtracted if they leave a zone). Resetting the MSD for those particles to zero will again lead to discontinuities; these troubles are solved by restricting investigations to the copper diffusion coefficients.

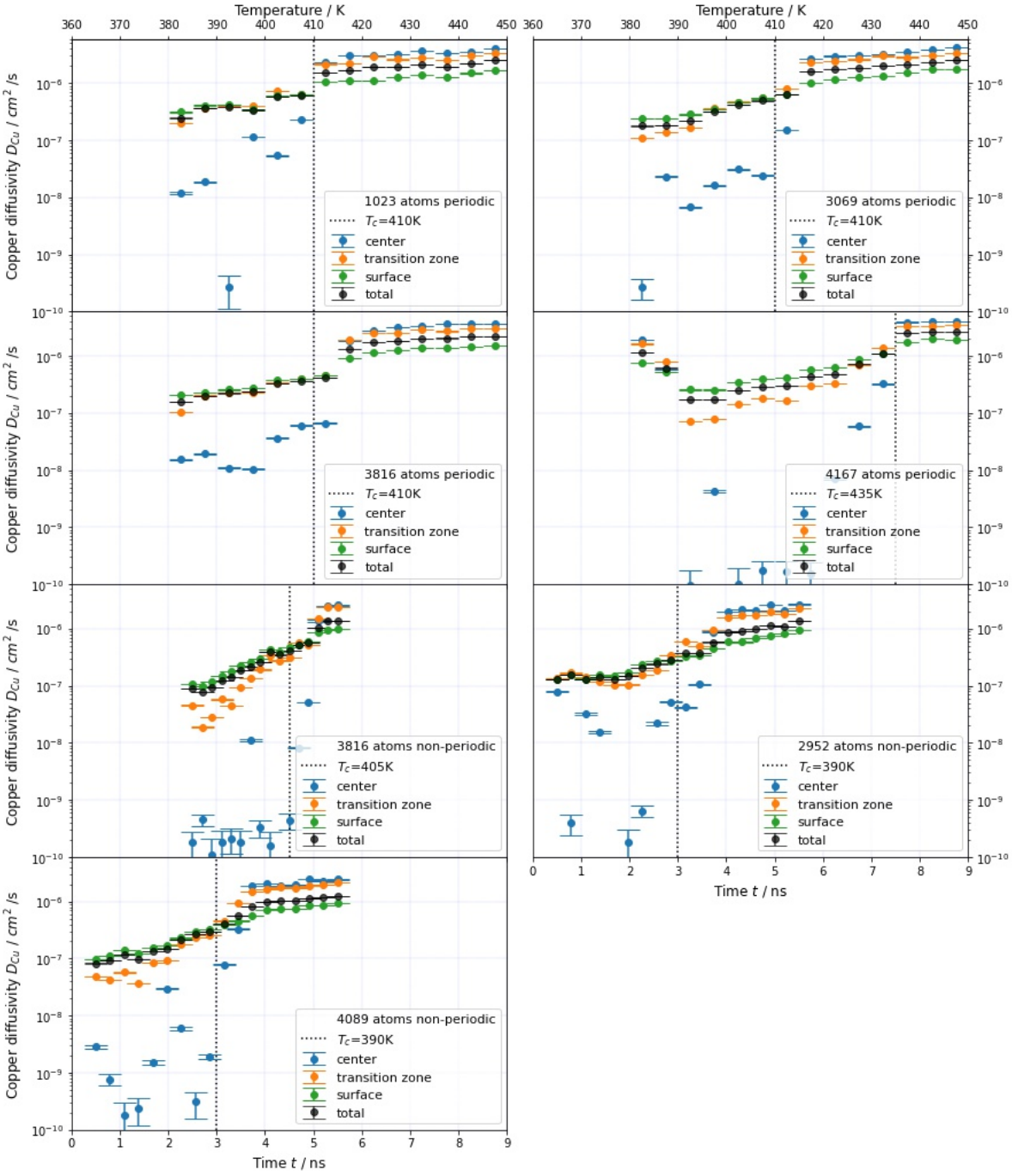


Fig. 21: This is the copper diffusion w.r.t. zones for all of the simulations where a phase transition was observed. The vertical dashed lines indicate the transition temperature, determined by the energy steps in the previous figures 19, 15, 16.

The above figure shows the change in copper diffusion for all simulations where a clear phase transition from low to high chalcocite occurred. There are indeed some very clear patterns that can be recognized:
The copper diffusion in the low chalcocite phase is far lower in the center compared to the surface. This abruptly changes at the phase transition where copper atoms in the center gain very high mobility, surpassing the mobility at the surface. The more important insight turns out to be that copper atoms at the surface have only a very small change in mobility during the phase transition. Two simulations seem to contradict this observation; the 4167 and 2952 atoms structures have a higher or nearly identical copper diffusion in the center compared to the surface at the beginning of the simulation. A very detailed analysis will later demonstrate that they actually started in the high chalcocite phase and transitioned to the low chalcocite phase at the beginning of the simulation,

thus the pattern is consistent 9.1.
Another interesting observation that was made was that the height of the energy step / atom as shown in Fig. 22 also strongly depends on the geometry of the material. To get an estimate of the dependence of this step on the surface / volume ratio, the low and high chalcocite energy curves were fitted linearly and the difference at $t = 0$ of these fits was taken to be the height of the energy step. The curves and details of the fits can all be found in the appendix 11.6. The resulting figure 22 should only be taken to be a qualitative observation, since the errors are difficult to estimate. This can be understood by looking at them directly 11.6. The energy curves were not always straight lines and the results depend on what (x,y) value range is used to fit any of the phases. Nonetheless, the figure below strongly suggests that increasing the surface / volume ratio will result in a lower energy step.

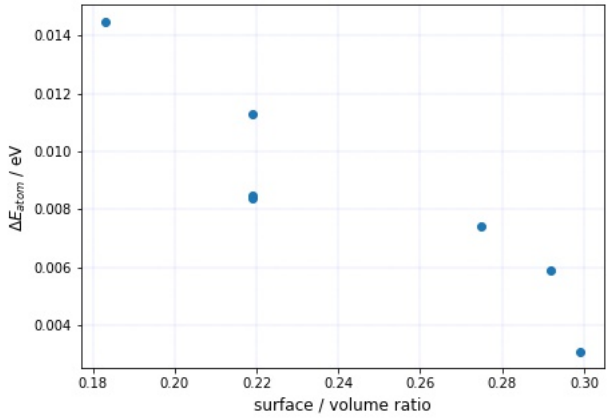


Fig. 22: This is the energy step / atom for all simulations where a phase transition from LC to HC was observed.

Combining these two observations (the diffusion does not change much at the surface and the energy step seems to be lower at the surface) lead to the idea that perhaps the geometry of the surface changes very little throughout this phase transition, which would explain the other similarities at the surface in the two different phases. To test this hypothesis, the cylinder was again investigated w.r.t. the different zones, but this time the change in radial distribution functions was analyzed. For the following investigation surface, transition zone and center have been treated as independent geometric objects and the radial distribution functions have been calculated for all of these zones, neglecting atoms in other zones. The RDFs before and after the phase transition for one of the structures is shown in Fig. 23. Indeed, the changes at the surface are very minor compared to changes in the center or the transition zone.

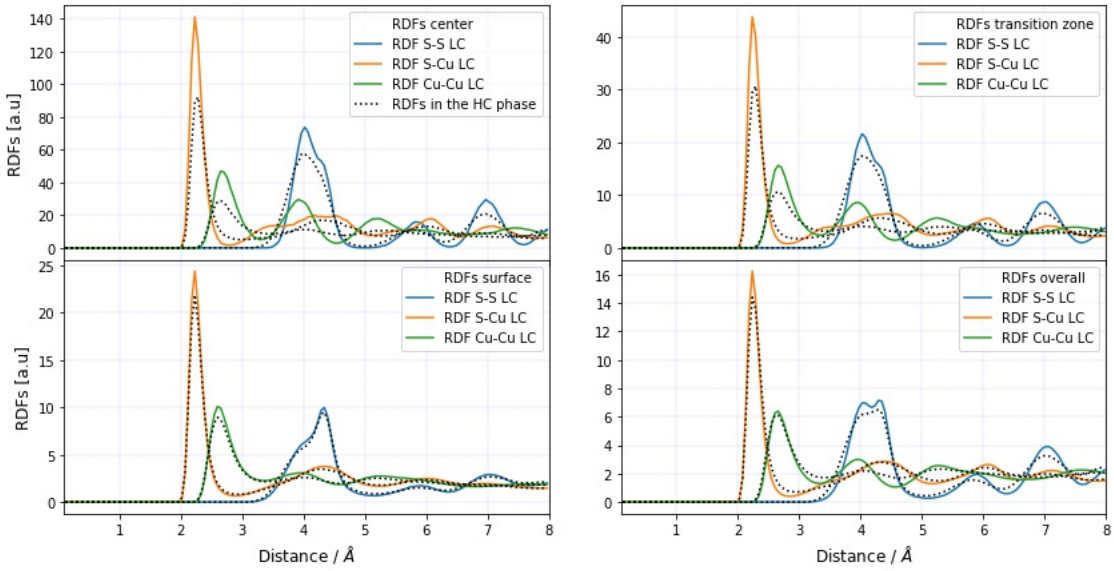


Fig. 23: This are the radial distribution functions of the periodic cylinder with 4.2 nm diameter and 4167 atoms. The LC RDFs were obtained by averaging over 200 ps at 400 K, while the HC RDFs were calculated at 440 K.

For a solid quantification of this phenomenon, we define the relative change of a function $f_t(x)$ from time $t = t_0$ to $t = t_1$ as

$$\frac{\|f_{t_0} - f_{t_1}\|_1}{\|f_{t_0}\|_1} = \frac{\int |f_{t_0}(x) - f_{t_1}(x)|dx}{\int |f_{t_0}(x)|dx}.$$

Here, t_0 stands for a time when the simulation was in the low chalcocite phase, and t_1 for a time when the simulation was in the high chalcocite phase. The functions f are given by the RDFs, and x is the distance going from zero to eight $\mathring{A}$. A total of 150 bins were used in each case, and the integrals were calculated with the trapezoidal rule.
Only the relative change w.r.t. the 1-norm was calculated, since the empty simulation volume was not always the same for different simulations and a relative value makes the change for different RDFs more comparable.

# Atoms	Periodic	Temperature RDFs LC / K	Temperature RDFs HC / K
2952		360	410
4089		360	410
3816		370	415
3069	✓	380	440
3816	✓	380	440
1023	✓	380	430
4167	✓	400	440

Table 5: These are the details for computing the RDFs in the HC and LC phase. The table shows at which temperature the RDFs have been calculated. All of them have been averaged over 200 ps.

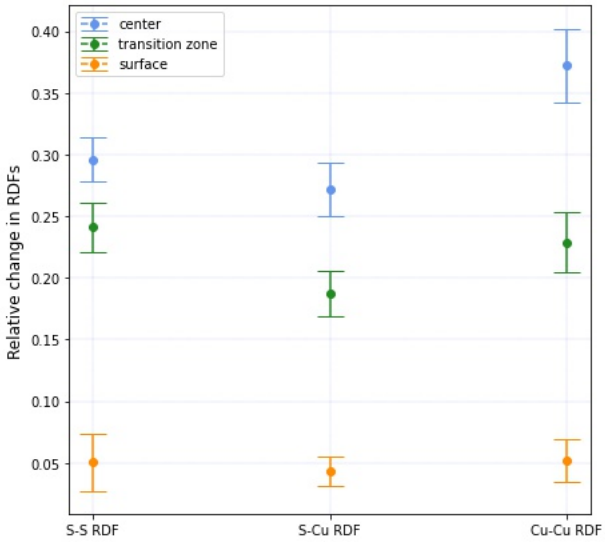


Fig. 24: An RDF comparison was calculated for all seven phase transitions from LC to HC. The average value is given by the dots, and the error is given by the standard deviation of the seven results.

As can be seen in Fig. 24, the local neighborhood for atoms at the surface change very little at the surface compared to the transition zone and the center. The changes in the center are strongest. This is a convincing geometric explanation of why atoms at the surface have a smaller energy step and why their diffusion does not change much: They stay roughly in the same position, which costs very little potential energy and cannot lead to a different diffusion. It also provides a clear answer of why the phase transition of the non-periodic structures in the first set of simulations were less pronounced: They had a much higher surface / volume ratio than the periodic structures.
Naturally, the question arises why the surface does not change much? This, however, is left as an open question beyond the scope of this thesis.

So far, a coherent picture of the different behavior of surface atoms and bulk atoms could be developed. There is one more aspect to the phase transition that seems to be quite unrelated, but was nonetheless discovered and should therefore be stated: The inertia tensor was computed for all of the simulations that showed a phase transition, and a consistent change was discovered. The results and an interpretation are given in the following subsection.

8.2 Analysis with respect to the inertia tensor

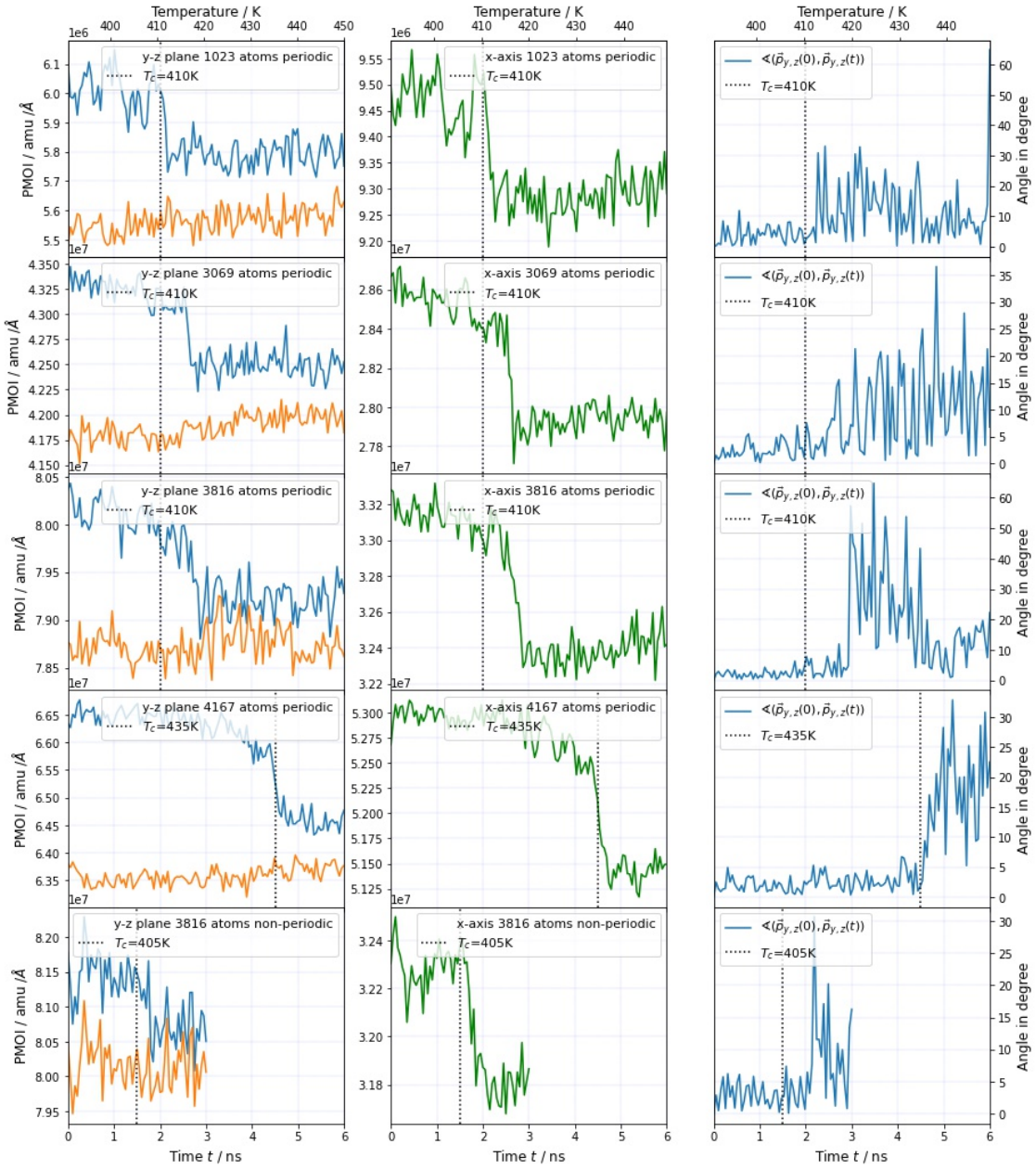
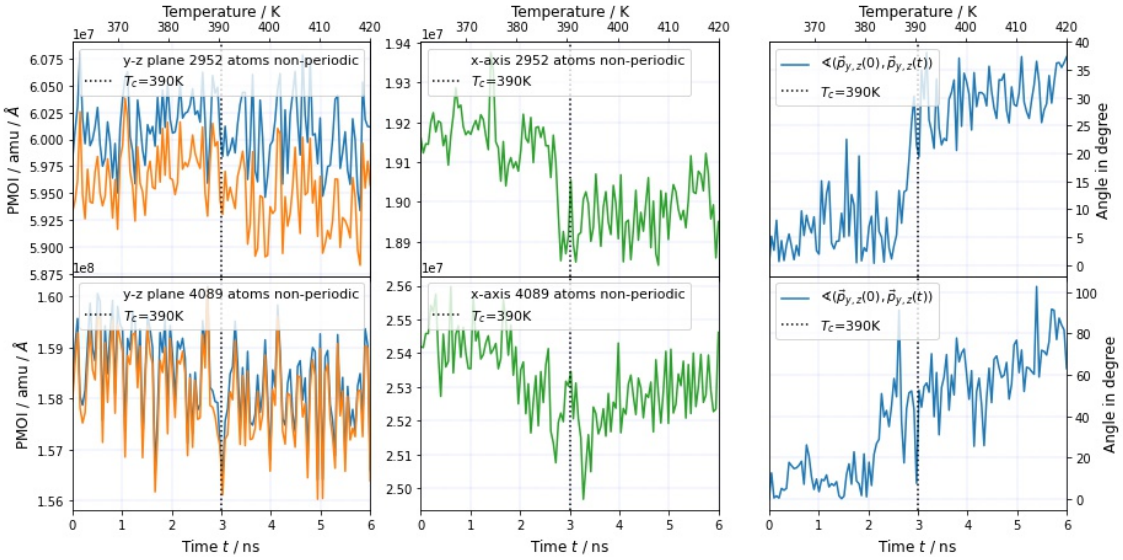


Fig. 25: This figure shows the principal moments of inertia (PMOIs) for five of the phase transition, as well as the rotation of the principal axes in the y-z plane. All quantities have been computed every 50 ps. The vertical dashed lines indicate the transition temperature, determined by the energy steps in the previous figures 19, 15, 16.

There are two quantities of interest when computing the inertia tensor: the principal moments of inertia (PMOI) and the principal axes. Computing the principal axes showed that one of them is always aligned with the x-direction (with a difference of at most 2°), which is parallel to the height axis of the cylinders. The other two are therefore always in the y-z plane, in which the cylinders have their radial symmetry. It was observed in all simulations that the PMOIs corresponding to the y-z plane come very close to each other at the phase transition and the principal axes in the y-z plane start rotating rapidly. A few things have to be mentioned. The first is that the inertia tensor is symmetric, thus the principal axes are orthogonal to each other. Since the x-axis is fixed, when one of the principal axis in the y-z plane starts rotating, the other one has to rotate as well by the same angle, to stay orthogonal to the other ones. This is the reason why only one angle is shown in Fig.

Secondly, the principal axes were labeled by the magnitude of the PMOIs. There is one slightly non trivial issue:
The algorithms used to compute the principal axes sometimes converge to an eigenvector in the opposite direction (which is of course a legitimate eigenvector!), but this would result in an angle of 180 degrees compared to the last time step. Tracking this angle is of course meaningless. To be able to still track the angle, the number of flips $f(t)$ were counted over time (it was set to $f(0) = 0$ and each time the condition $\sphericalangle(\vec{p}_{y,z}(t), \vec{p}_{y,z}(t - 50\text{ ps})) > 90^\circ$ was satisfied, it was set to $f(t) = f(t - 50\text{ ps}) + 1$). For the cases where $f(t)$ is even (no overall flip) $\sphericalangle(\vec{p}_{y,z}(t), \vec{p}_{y,z}(0))$ was directly calculated from the returned eigenvectors, in the case where it is odd 180° were subtracted from the angle (and the absolute value taken, since only subtracting 180° returns the negative angle).
The remaining results for the two non-periodic structures that showed only a weak phase transitions are added below. The vertical dashed lines indicates the transition temperature, determined by the energy steps in Fig. 16.



The change in the PMOIs for these two structures is far less visible compared to the other simulations, which fits well with the already established conclusion that structures with more surface have a smaller change in geometry overall. (Those are the two structures with the most surface that still showed a clear phase transition)
Finally, the results can be interpreted physically: The cylinders have a radial symmetry in the y-z plane, thus, if the mass was distributed completely homogeneously there would be no way to tell which direction in the y-z plane corresponds to any of the principal axes, since they would be equivalent! The LC phase apparently has a slightly inhomogeneous mass density, thus there are two well defined directions for the principal axes in the y-z plane. However, since the unit cell of the HC phase is more symmetric, the mass distribution becomes so homogenous that it becomes more and more difficult to distinguish between any of the two principal axes in the y-z plane. This is the reason they appear to be rapidly rotating after the phase transition. The next chapters are devoted to shining light on all simulations that did not show a phase transition, as well as investigating the reverse transition: From HC to LC.

9 Explaining the absence of phase transition for some geometries

An important question that remains unanswered is: Why was the phase transition only observed for some geometries? The reason was actually discovered at some point and will be clarified in this chapter. Going back to the second set of simulations shown in Fig. 19 it was mentioned that the 3816 and 1023 atom structures were pre-heated. The reason for this is that both of them started with such a high copper diffusivity (without pre-heating), that the simulation was restarted at a lower temperature, because it was suspected that they go to the HC phase directly. Below Fig. 26, a detailed explanation together with all thermodynamic data is given. Note that the data shown below is complete, there are no hidden equilibrations; all simulations directly started from the cut cylinder.

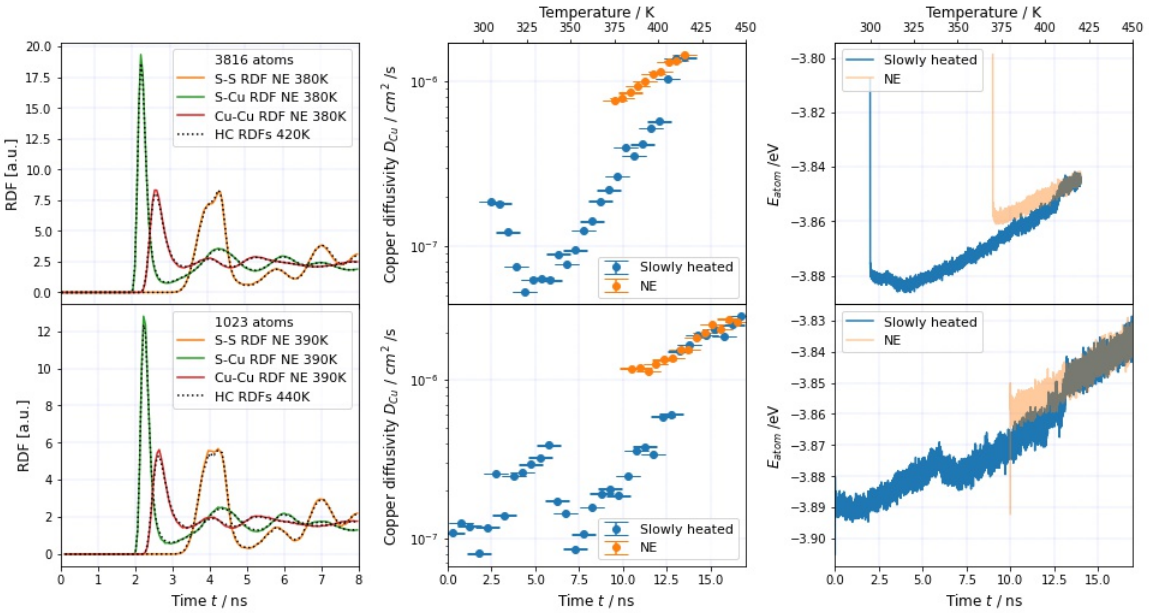


Fig. 26: The abbreviation NE for non-equilibrated has been used in the above figure to refer to the original simulation that started from an unequilibrated cylinder at 370 K (top) or 380 K (bottom). There are a couple of strong pieces of evidence that NE started from the HC phase: First of all, comparing the RDFs at 380 K / 390 K from the NE structures to the pre-heated simulations at 420 K (e.g. after its phase transition to HC) shows a clear overlap (left). The NE structures are evidently in HC far below the usual phase transition temperature. The same conclusion can be reached by looking at the copper diffusion that is far too high for LC and matches the copper diffusion of the pre-heated simulations after their phase transition to HC. Lastly, looking at the energy curves again confirms that the energy of the original simulations was at a HC level. It is important to point out that energies were computed every ps and the energy level for the top row was *never* at the level of LC (again there is no hidden equilibration). All of this strongly suggests that the structures indeed directly went into the HC phase.

The above figure seems very unintuitive. The cylinders are cut from the ideal unit cell of the LC structure, why do they directly go to HC? Looking at the energy curves more carefully shows that they start at a far too high level (this was observed for all simulations). A plausible explanation is that the starting configuration is unrealistic for both: the HC and LC phase, because cutting a cylinder from a bulk structure leads to a completely unrealistic and highly energetic surface. Evidently, a visual inspection of all trajectories showed that the atoms close to the surface rearrange significantly at the very beginning of a simulation. The original cylinders are unrealistic for either

phase and the material is not in a physically well-defined phase; the atoms need to rearrange until a realistic configuration is found. If the energy is too high, this turns out to be the HC phase.
One more thing is disturbing: The pre-heated structures also started with an energy level that is far too high for the LC phase, did they also start in HC? The answer turns out to be yes; very unintuitively *all* simulations started in the HC phase, *some* of them transitioned to LC after some time such that a transition from LC to HC was observed at a later stage in the simulation. This is the reason why the phase transition was not observed for all simulations: They already started in HC and never went to LC. Such a claim needs strong evidence, which will be presented in the next chapter. The following chapter will be full of technical details making a case for each simulation and can be skipped if one is willing to believe this claim. However, it will also show that trends that have been discovered for the transition from LC to HC exactly reverse for the opposite phase transition, which makes the conclusions of the previous chapters actually even more convincing.

9.1 Four examples of phase transitions from HC to LC to HC

Because of the large amount of simulations not all transitions from HC to LC were analyzed in depth. This subsection will show four simulations that started in HC, transitioned to LC and again transitioned back to HC. Despite showing that these four simulations started in HC, it will also become clear that all trends reverse for the opposite phase transition. The following two pages show four figures. All of them have precisely the same format:

In the left column, the energy curves with its two phase transitions and the diffusion w.r.t. zones can be seen. The dotted lines show the phase transition temperatures. The second column shows a comparison of RDFs; the top picture shows that the RDFs indeed change significantly after the transition from HC to LC and the bottom picture shows that the RDFs close to the beginning of the simulation and close to the end of the simulation are roughly identical; indeed the first phase transition is precisely reversed. The last row shows the principal values of the moment of inertia which move away from each other in the *y*-*z* plane after the first phase transition (lowering its mass homogeneity), and find their original distance after the second phase transition. The RDFs have been averaged over 200 ps. In all cases the energy curves are shown with a resolution of 1 ps, none of them dropped to the LC level at the beginning of the simulation, suggesting that they directly equilibrated to HC. (Again, all data below is complete, there are no hidden equilibrations)

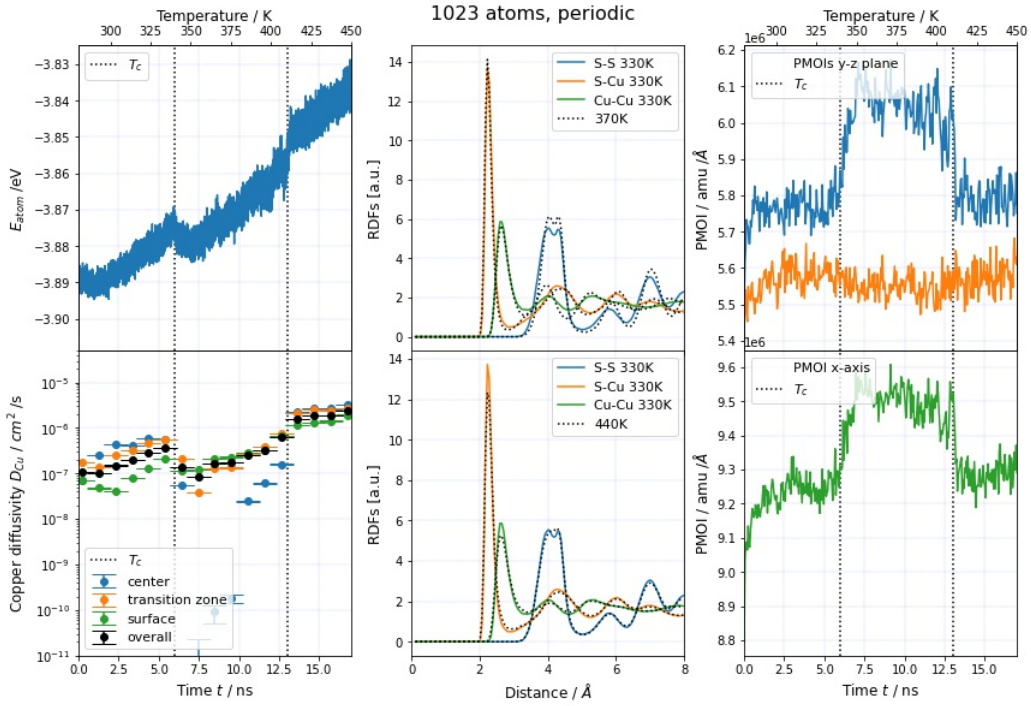


Fig. 27: A detailed explanation is given in the second paragraph of 9.1.

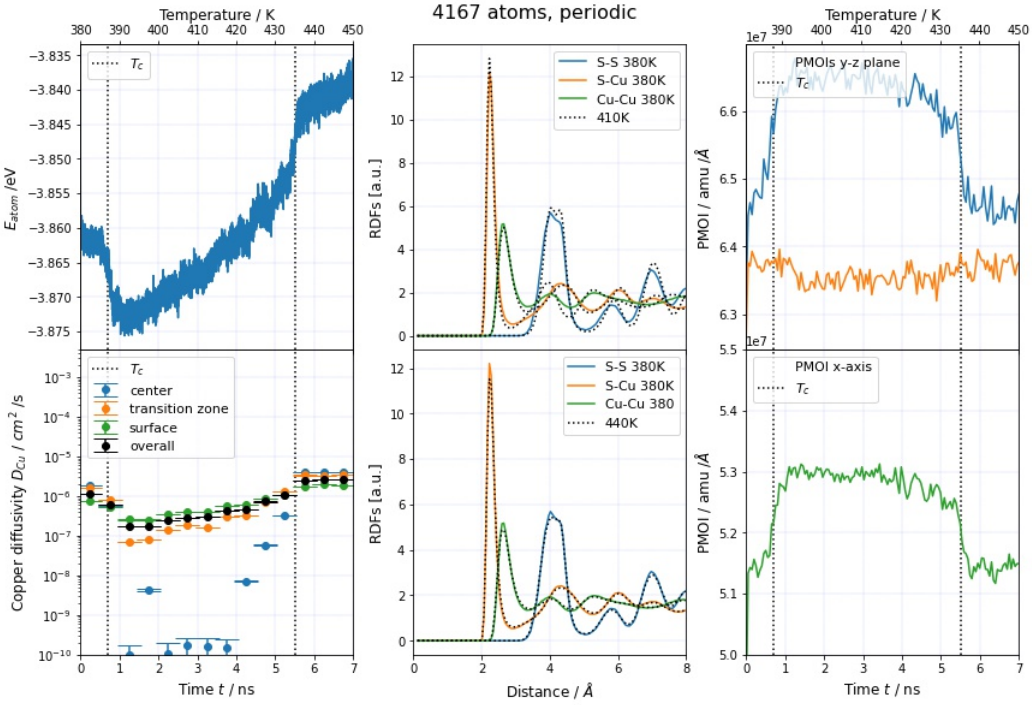


Fig. 28: A detailed explanation is given in the second paragraph of 9.1.

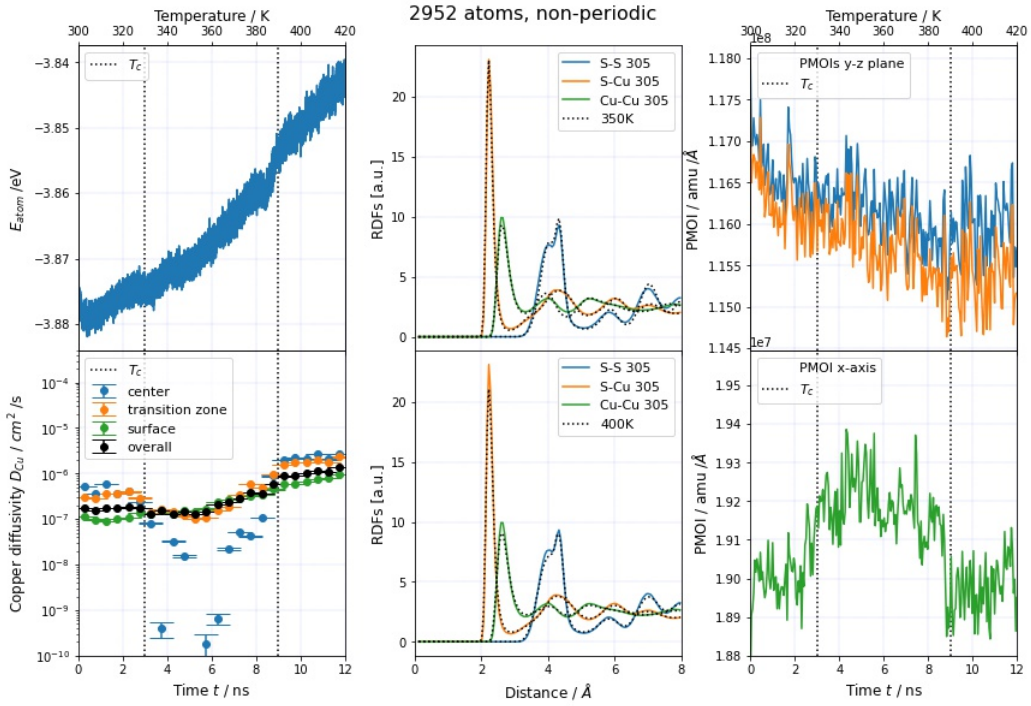


Fig. 29: A detailed explanation is given in the second paragraph of 9.1.

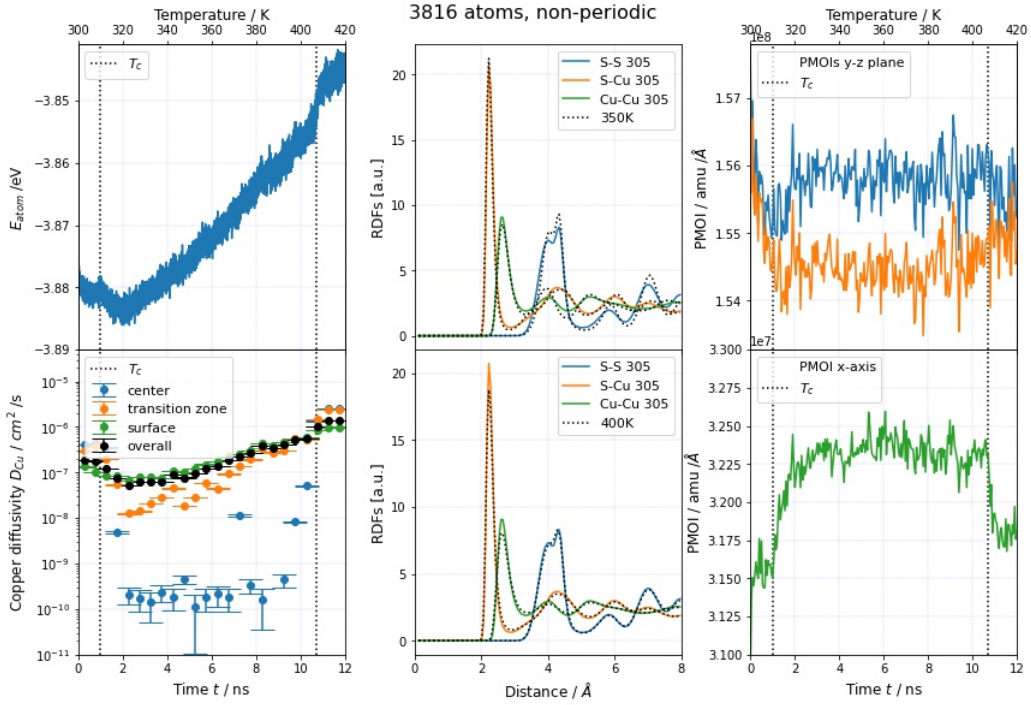
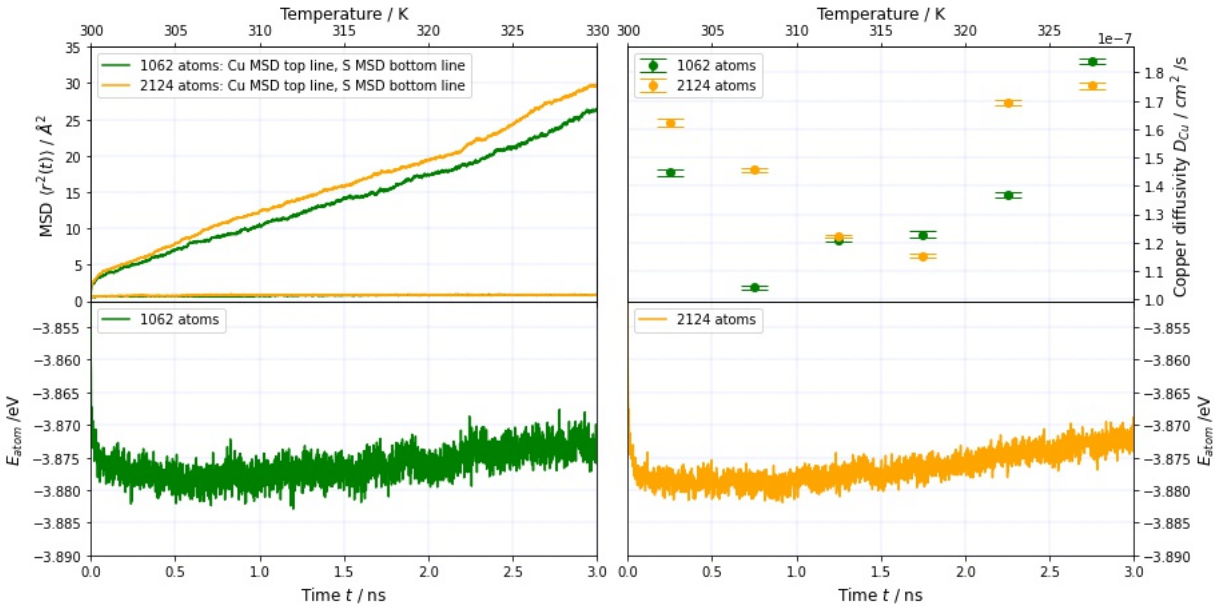


Fig. 30: A detailed explanation is given in the second paragraph of 9.1.

In this section we have demonstrated that four more structures started in HC, additionally we could confirm that the phase transition from HC to LC reverses all changes. The next subsection will demonstrate that the 1062, 2124, 1065 and 1974 structures from the first set of simulations - as well as the 4248 atom structure of the second set of simulations - were in the HC phase, which is why they could not transition to HC anymore.

9.2 Explaining the absence of some phase transitions for all remaining structures

Showing that the 1062 and 2124 atoms structures (periodic) were indeed in HC was done in great detail. First of all, the simulations were restarted from an unequilibrated state at a temperature of 300 K and heated up to 330 K:



Looking at the energy curves starting from an unequilibrated state suggests that no phase transition happened. Comparing the copper diffusion at 335 K with the simulations shown in Fig. 15 shows roughly the same value, indicating that no prior phase transition happened during the original equilibration. To provide a much more thorough examination, a simulation of the smaller geometry (1062 atoms) was started at a much lower temperature of 200 K and heated up to 350 K. The intention was that the phase transition might be seen at lower temperatures.

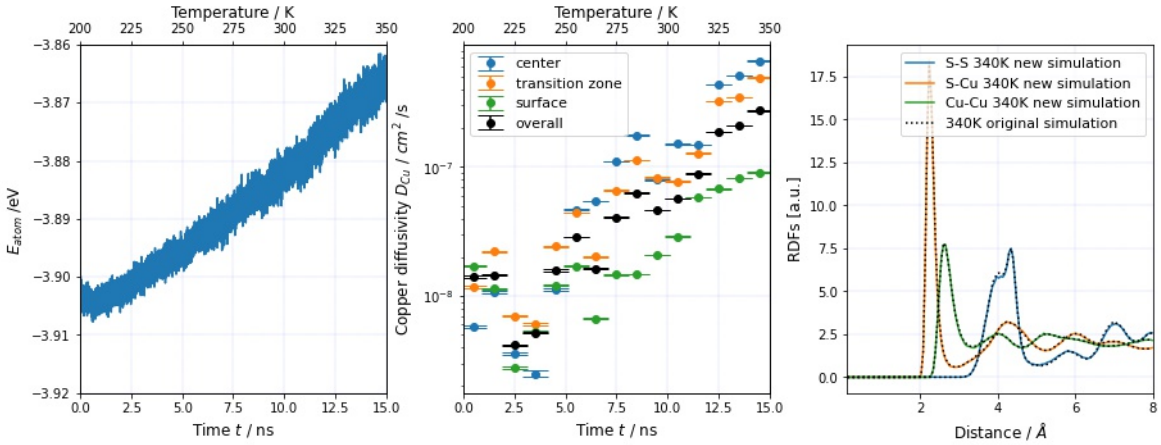


Fig. 31: This is the simulation of the 1062 atoms periodic structure that was repeated at lower temperatures. The energy curve does not indicate a phase transition. On the right the RDFs of the restarted simulation at 340 K are compared to the RDFs at 340 K of the simulation presented in the chapter First set of simulations. They overlap perfectly, and the materials are thus in the same phase.

First, the RDFs of the new simulation are compared to the RDFs of the original simulation at 340 K. The perfect overlap shows that they are indeed in the same phase, e.g. if it can be demonstrated that

the restarted simulation is in HC, so was the original simulation. Looking at the copper diffusion with respect to zones clearly shows that for higher temperatures the diffusion is much higher in the center, which we know to be an indicator of the HC phase. However, it is very interesting to point out that up to roughly 210 K the diffusion in the center is *lower* than at the surface. Does that contradict the conclusions we found so far?
The answer is no: There are reasons to believe that the material actually started in the LC phase and transitioned to the HC phase at a later stage in the simulation. Since the surface / volume ratio is higher than for all cylinders were a phase transition was observed, only a very tiny energy step is expected that might not be visible at all. A careful investigation of the RDFs in Fig. 32 confirms that the sulfur lattice has changed between 210 K and 280 K (pointing to a phase transition, together with the rapid increase of copper diffusivity in the center), but no noticeable changes are seen between 280 K and 340 K in Fig. 33.
The comparison at higher temperatures (280 K and 340 K) was added, because the changes of RDFs in Fig. 32 are less pronounced than in all previous examples given. While this might be justified by the weaker change of geometries in materials with more surface, perhaps the changes are only due to an increase of the temperature of the material (in the absence of a phase transition). Testing that there is a noticeable change in the temperature range where the diffusivity increases drastically, while the material does not show similar transformations in other temperature ranges, makes a phase transition the more plausible explanation.

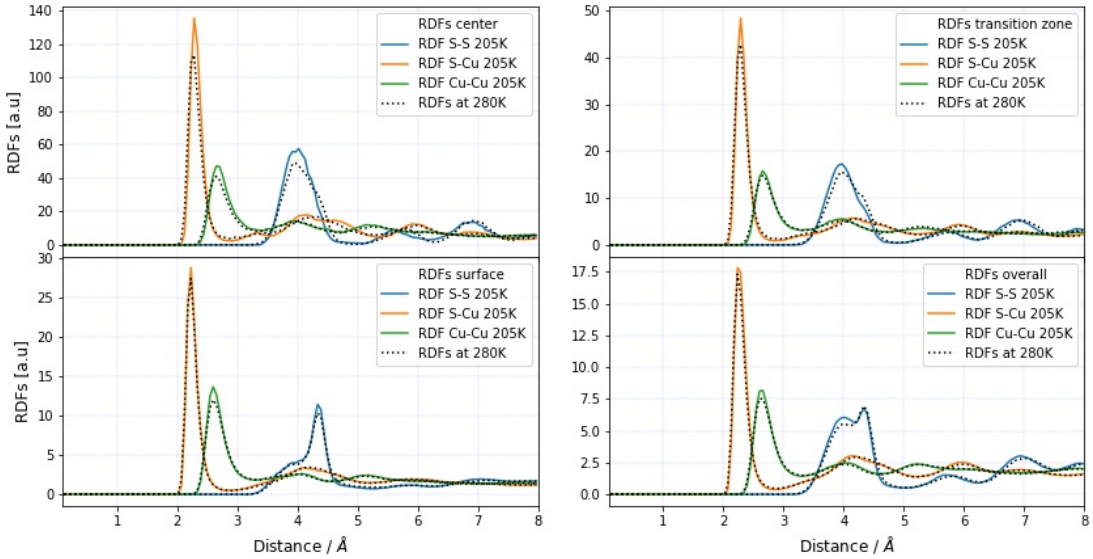


Fig. 32: This is an RDF comparison of the restarted simulation with 1062 atoms. Changes of the sulfur lattice are visible in the center and transition zone.

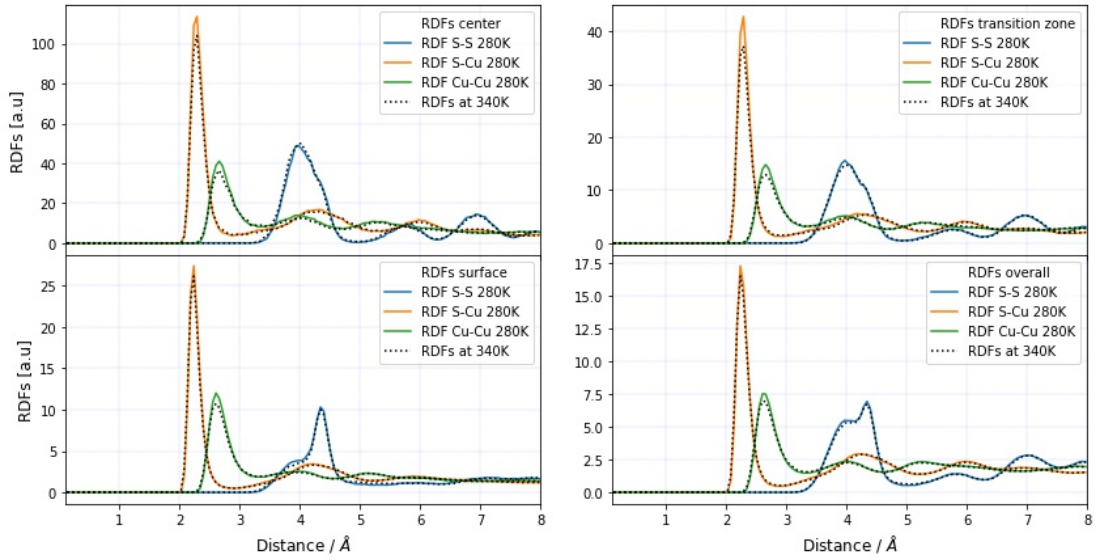
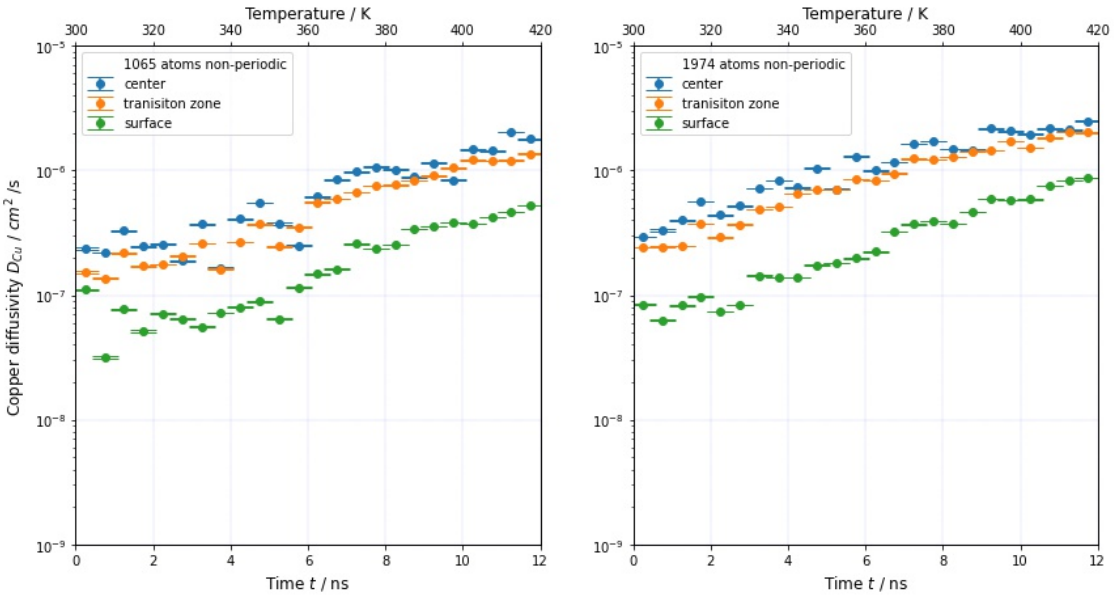
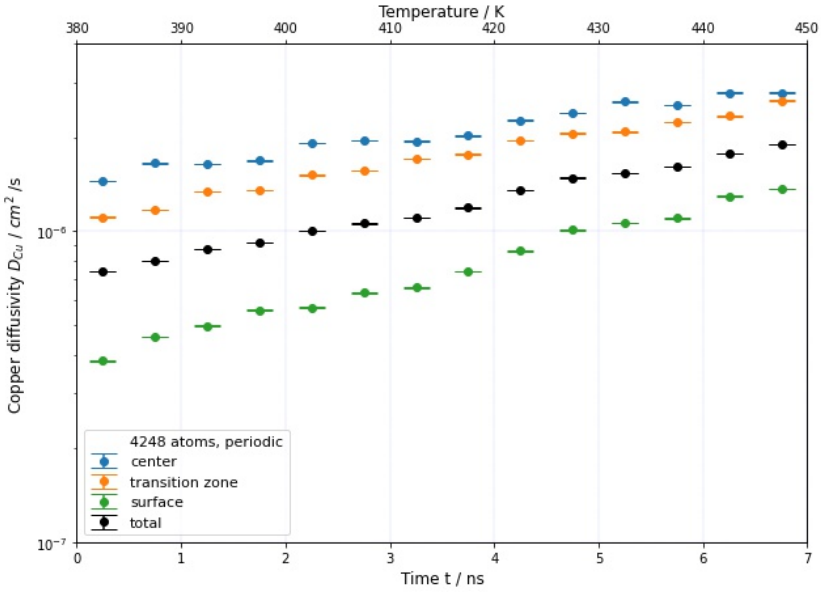


Fig. 33: This is an RDF comparison of the restarted simulation with 1062 atoms. At higher temperatures the sulfur lattice appears to be unchanged.

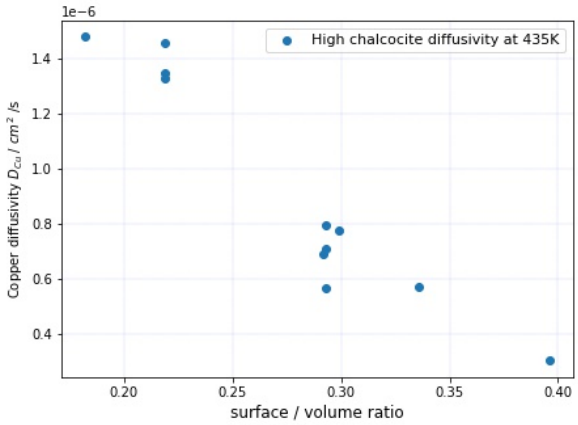
Summarizing, the simulation is very likely in the HC phase at 320 K and a comparison of RFDs of this simulation with the original simulation starting at 320 K confirms that it was in the HC phase as well. Furthermore, it looks like a phase transition happened at a much lower temperature, but was significantly more difficult to detect due to the large surface / volume ratio.
Since the 2124 atom structure showed nearly identical overall copper diffusion and MSD, it is not a far stretch to conclude that it was also simulated in the HC phase. For the two non-periodic structures (1065, 1974 atoms) of the first set of simulation only the copper diffusion with respect to zones was checked:



Looking at the much higher diffusion in the center compared to the surface is a good indicator of the structures being in the HC phase throughout the entire simulation. Due to the enormous amount of data and the very coherent overall explanation no further details were checked for these two structures. Lastly, we show the copper diffusion of the 4248 structure of the second set of simulations, which also strongly suggests that the structure was in the HC phase throughout the simulation:



At this point, it seems that *all* cylinders that did not show a phase transition started in the HC phase. We saw that in at least four of the other cases a phase transition from HC to LC happened at some point, allowing for a later phase transition from LC to HC. The energy curves of simulations that were not addressed in this chapter in detail can be found in the appendix 11.5, 11.6. They also suggest that the three simulations that were not discussed transitioned from HC to LC close to the beginning of the simulation. We conclude this chapter by noting that all simulations were in the HC phase at 435 K, which allows to plot the overall diffusion in the HC phase at 435 K with respect to the surface / volume ratio for a large number of data points:



The fitting errors for the copper diffusion are of course very low and not shown (since they do not appear to be relevant); the data points were directly extracted from the presented simulations and do not necessarily accurately present the copper diffusion at this temperature for a very long simulation without ramping the temperature. Nonetheless, the decreasing trend of the overall copper diffusion for materials with more surface is clearly visible and compatible with the low change of copper diffusion at the surface of nanorods. For more precise information of the data points Table 6 can be studied.

			Cu Diffusivity D_{Cu} / cm^2 /s $\cdot 10^6$	
# Atoms	Periodic	surf / vol ratio	LC 395 K	HC 435 K
4167	✓	0.182	0.147	1.48
1023	✓	0.219	0.283	1.33
3069	✓	0.219	0.244	1.35
3816	✓	0.219	0.169	1.46
3816		0.275	0.338	NA
4089		0.292	NA	0.689
1062	✓	0.293	NA	0.708
2124	✓	0.293	NA	0.567
4248	✓	0.293	NA	0.794
2952		0.299	NA	0.777
1974		0.336	NA	0.572
1065		0.396	NA	0.302

Table 6: These are the details for the diffusion coefficients w.r.t. surface / volume ratio.

10 Conclusion

The research carried out for this Master's Thesis was full of complicated challenges, in particular, working out the details of why many of the simulations did not show a phase transition was an arduous process. Clarifying the reason for all oddities discovered resulted in a large number of coherent conclusion that can be drawn from the research:
First of all, it was demonstrated that the symmetry functions that were originally used to train a HDNNP for bulk copper sulfide are fully capable of describing surfaces as well. A trustworthy HDNNP has been developed that can be used to simulate nano crystals of copper sulfide in both phases, as well as the transition regime. A number of new aspects of the phase transition were discovered:
The copper diffusion at the surface is higher in the LC phase and lower in the HC compared to the center and the transition zone. It changes only moderately at the surface. Similarly, the height of the potential energy step decreases with and increasing surface / volume ratio. Both of these phenomena are rooted in the much smaller geometric change of the surface of the material during the phase transition, compared to the center. This results in the additional conclusion that the overall copper diffusion in the HC phase depends on the surface / volume ratio as demonstrated in the last chapter.
Moreover, the lower energy steps for atoms at the surface are likely the reason why the phase transition from LC to HC start at the surface and spreads to the center (this was observed experimentally [11]): Since these atoms need less energy to transition to the new phase, a nucleation event is likely to happen at the surface first. This logic can also be reversed: For the transition from HC to LC atoms in the center release the most energy, and will therefore tend to transition into LC first (which was also observed in the experiment).
Next, it was accidentally discovered that a simulation of the reverse phase transition is indeed possible, allowing to study it in both directions. This is not trivial, as this was not achieved for bulk copper sulfide [16]. It was verified that all changes from LC to HC are reversed for the opposite phase transition, adding another consistency check for all drawn conclusions. Additionally, it was discovered that structures that are directly cut from the bulk LC phase tend to have far too high surface energies. In one case 9.2 starting the simulation at a much lower temperature probably lead to the structure equilibrating to the LC phase directly.
The last discovery was that there is a consistent change in the inertia tensor, suggesting that the mass density is highly homogeneous in the HC phase, but more inhomogeneous in the LC phase.
Lastly, the phase transition temperatures in the simulations were anywhere between 390 to 435 K, which is substantially higher than the temperatures observed in experiments [12]. This temperature shift likely arises from the PBE functional, that is itself just an approximation to the true exchange-correlation energy functional.

While some contributions towards understanding the phase transition could be made, a full explanation of why the phase transition happens in the way it does is still out of sight, and will require a substantial amount of additional research.
For future simulations, it could be attempted to equilibrate structures at much lower temperatures to go to the LC phase directly. Additionally, working with cylinders with a diameter of at least 3.5 nm makes changes much more detectable, while going below 3 nm should perhaps be avoided or done with all the difficulties presented in mind.

11 Appendix

11.1 Details of all runs

The following tables provide a detailed list of all runs that were done. The runs are enumerated in chronological order, i.e. run 1 is the first run, run 15 the last run etc. Energies and forces for structures of different slabs belonging to the same run were computed simultaneously and two new HDNNPs were only trained when all structures of a corresponding run were available. All MD simulations for the two HDNNP comparison method were done in the NVT ensemble.

Run	# of structures	Creation process
1	300	100 random displacements (≤ 0.3 Å) of atoms in the ideal structure for Slab X, Y, Z.
2	1912	Slab X: MD: 10000 time steps (0.5 fs) at 350 K. Selection threshold (E/atom, forces): 3 meV, 200 meV/Å. 635 structures selected.
		Slab Y: MD: 7000 time steps (0.5 fs) at 350 K. Selection threshold (E/atom, forces): 1.5 meV, 200 meV/Å. 562 structures selected.
		Slab Z: MD: 800 time steps (0.5 fs) at 350 K. Selection threshold (E/atom, forces): 1 meV, 200 meV/Å. 715 structures selected.
3	1002	Slab X: MD: 10000 time steps (10 fs) from 150-450 K. Selection threshold (E/atom, forces): 1.9 meV, 170 meV/Å. 323 structures selected.
		Slab Y: MD: 10000 time steps (10 fs) from 150-450 K. Selection threshold (E/atom, forces): 1.1 meV, 160 meV/Å. 320 structures selected.
		Slab Z: MD: 10000 time steps (10 fs) from 150-450 K. Selection threshold (E/atom, forces): 1.1 meV, 160 meV/Å. 359 structures selected.
Continued on the next page		

Run	# of structures	Creation process
4	1609	Slab X: MD: 10000 time steps (10 fs) from 450-600 K. Selection threshold (E/atom, forces): 2.05 meV, 200 meV/$\AA$. 528 structures selected.
		Slab Y: MD: 10000 time steps (10 fs) from 450-600 K. Selection threshold (E/atom, forces): 2.25 meV, 200 meV/$\AA$. 509 structures selected.
		Slab Z: MD: 10000 time steps (10 fs) from 450-600 K. Selection threshold (E/atom, forces): 2.6 meV, 260 meV/$\AA$. 572 structures selected.
5	100	100 random displacements ($\leq$0.3 $\AA$) of atoms in the ideal structure for Slab NK.
6	964	Slab X: MD: 10000 time steps (10 fs) from 500-700 K. Selection threshold (E/atom, forces): 1.6 meV, 200 meV/$\AA$. 354 structures selected.
		Slab Y: MD: 10000 time steps (10 fs) from 500-700 K. Selection threshold (E/atom, forces): 1.3 meV, 200 meV/$\AA$. 291 structures selected.
		Slab Z: MD: 10000 time steps (10 fs) from 500-700 K. Selection threshold (E/atom, forces): 1.3 meV, 160 meV/$\AA$. 319 structures selected.
7	468	Slab NK: MD: 5000 time steps (10 fs) at 300 K. Selection threshold (E/atom, forces): 1.4 meV, 200 meV/$\AA$. 468 structures selected.
8	379	Slab NK: MD: 5000 time steps (10 fs) from 300-500 K. Selection threshold (E/atom, forces): 4 meV, 200 meV/$\AA$. 379 structures selected.
9	100	100 random displacements ($\leq$0.3 $\AA$) of atoms in the ideal structure of Slab H.
10	647	Slab NK: MD: 5000 time steps (0.5 fs) at 500 K. Selection threshold (E/atom, forces): 2.2 meV, 210 meV/$\AA$. 647 structures selected.
Continued on the next page		

Run	# of structures	Creation process
11	1354	Slab H: MD: 4000 time steps (0.5 fs) at 300 K. Selection threshold (E/atom, forces): 1.4 meV, 200 meV/$\AA$. 447 structures selected.
		Slab X: MD: 5000 time steps (0.5 fs) at 500 K. Selection threshold (E/atom, forces): 1.1 meV, 200 meV/$\AA$. 296 structures selected.
		Slab Y: MD: 5000 time steps (0.5 fs) at 500 K. Selection threshold (E/atom, forces): 1.2 meV, 260 meV/$\AA$. 268 structures selected.
		Slab Z: MD: 5000 time steps (0.5 fs) at 500 K. Selection threshold (E/atom, forces): NA 343 structures selected.
12	334	Slab H: MD: 10000 time steps (0.5 fs) from 300-500 K. Selection threshold (E/atom, forces): 2.5 meV, 200 meV/$\AA$. 334 structures selected.
13	565	Slab H: MD: 10000 time steps (0.5 fs) at 500 K. Selection threshold (E/atom, forces): 1.3 meV, 200 meV/$\AA$. 565 structures selected.
14	501	Slab H: MD: 10000 time steps (0.5 fs) from 500-700 K. Selection threshold (E/atom, forces): 2.1 meV, 200 meV/$\AA$. 501 structures selected.
15	342	Slab NK: MD: 5000 time steps (0.5 fs) from 500-700 K. Selection threshold (E/atom, forces): 1.9 meV, 200 meV/$\AA$. 342 structures selected.

Table 7: Details of all runs.

This extensive list is concluded with a brief summary of the total number of structures for each geometry.

Structure	Total number of structures
Slab X	2236
Slab Y	2050
Slab Z	2408
Slab NK	1936
Slab H	1947
All slabs	10577
Bulk	3965
All structures	14542

Table 8: Summary.

11.2 HDNNP parameters

This part of the appendix provides a list of HDNNP parameters used for training the final HDNNPs. Note that, except for the neural network topologies, all parameters are identical to the ones previously used for bulk chalcocite [16]. The formulas on this page were also taken from this reference. The softplus activation function $f_a(x)$=$\log(1+e^x)$ was used.
Additionally, data set normalization was used to make the fitting process independent of the physical units used. The energies were normalized according to

$$E_a^* = \frac{E_a - N_a\langle e\rangle}{\sigma_e}, \tag{11.2.37}$$

$$\langle e\rangle = \frac{1}{N}\sum_{a=1}^{N}\frac{E_a}{N_a} \tag{11.2.38}$$

$$\sigma_e = \sqrt{\frac{1}{N-1}\sum_{a=1}^{N}\left(\frac{E_a}{N_a} - \langle e\rangle\right)^2}, \tag{11.2.39}$$

where $\langle e\rangle$ is the mean energy per atom, and σ_e the standard deviation. For the scaled units $\langle e^*\rangle = 0$ and $\sigma_{e^*} = 1$ holds. The energy conversion factor is thus given by $c_{\text{energy}} = \frac{1}{\sigma_e}$. Similarly, a length conversion factor $c_{\text{length}} = \frac{\sigma_F}{\sigma_e}$ was introduced with

$$\sigma_F = \sqrt{\frac{\sum_{a=1}^{N}\sum_{i=1}^{N_a}\vec{F}_{\text{ai},\alpha}^2}{3\sum_{a=1}^{N}N_a - 1}}. \tag{11.2.40}$$

Hence, the scaled forces are given by $F^* = \frac{c_{\text{energy}}}{c_{\text{length}}}F$ and we find $\sigma_{F^*} = 1$.

In addition symmetry function normalization was used to avoid large input neuron values saturating activation functions. The normalization scheme chosen was

$$G^* = \frac{G - \langle G\rangle}{\sigma_G}, \tag{11.2.41}$$

for each symmetry function, where G is the unscaled symmetry function, $\langle G\rangle$ its mean and σ_G its standard deviation (of structures in the reference data set). Only scaled values were used during training and for simulations.

The following table contains all parameters of the Kalman Filter.

Parameter	Value
β	10.0
ε	0.01
q_0	0.01
τ_q	2.302
q_{min}	10^{-6}
η_0	0.01
τ_η	2.302
η_{max}	1
# streams	48

Table 9: Kalman Filter parameters used for HDNNP training.

The symmetry function values of the training data set were always kept in memory to speed up the process. Due to the extensive data set roughly 110 GB of ram are needed for efficient training, which exceeds the amount available on a regular node on the VSC 4. Nodes with additional ram had to be requested, since increasing the number of streams (and thus the number of nodes needed, and the ram provided) lead to poor RMSEs. The section is concluded with a table containing all symmetry function parameters.

Element	Neighbor 1	Neighbor 2	$\eta \cdot Å$	λ	ζ	$r_c/Å$	$r_s/Å$
Radial symmetry functions							
S	S	-	2.0	-	-	6.0	1.5
S	S	-	2.0	-	-	6.0	2.0
S	S	-	2.0	-	-	6.0	2.5
S	Cu	-	2.0	-	-	6.0	1.5
S	Cu	-	2.0	-	-	6.0	2.0
S	Cu	-	2.0	-	-	6.0	2.5
S	Cu	-	2.0	-	-	6.0	3.0
Cu	S	-	2.0	-	-	6.0	1.5
Cu	S	-	2.0	-	-	6.0	2.0
Cu	S	-	2.0	-	-	6.0	2.5
Cu	S	-	2.0	-	-	6.0	3.0
Cu	S	-	2.0	-	-	6.0	4.0
Cu	S	-	2.0	-	-	6.0	5.0
Cu	Cu	-	2.0	-	-	6.0	2.0
Cu	Cu	-	2.0	-	-	6.0	2.5
Cu	Cu	-	2.0	-	-	6.0	3.0
Narrow angular symmetry functions							
S	S	S	0.2222	-1.0	1.0	6.0	0.0
S	S	S	0.2222	1.0	1.0	6.0	0.0
S	S	S	0.2222	1.0	6.0	6.0	0.0
S	S	S	0.04082	-1.0	1.0	6.0	0.0
S	S	S	0.04082	1.0	1.0	6.0	0.0
S	S	S	0.04082	-1.0	6.0	6.0	0.0
S	S	S	0.04082	1.0	6.0	6.0	0.0
S	S	S	0.01653	-1.0	1.0	6.0	0.0
S	S	S	0.01653	1.0	1.0	6.0	0.0
S	S	S	0.01653	-1.0	6.0	6.0	0.0
S	S	S	0.01653	1.0	6.0	6.0	0.0

Continued on the next page

Element	Neighbor 1	Neighbor 2	$\eta \cdot \AA$	λ	ζ	$r_c/\AA$	$r_s/\AA$
S	S	Cu	0.2222	-1.0	1.0	6.0	0.0
S	S	Cu	0.2222	1.0	1.0	6.0	0.0
S	S	Cu	0.2222	1.0	6.0	6.0	0.0
S	S	Cu	0.04082	-1.0	1.0	6.0	0.0
S	S	Cu	0.04082	1.0	1.0	6.0	0.0
S	S	Cu	0.04082	-1.0	6.0	6.0	0.0
S	S	Cu	0.04082	1.0	6.0	6.0	0.0
S	S	Cu	0.01653	-1.0	1.0	6.0	0.0
S	S	Cu	0.01653	1.0	1.0	6.0	0.0
S	S	Cu	0.01653	-1.0	6.0	6.0	0.0
S	S	Cu	0.01653	1.0	6.0	6.0	0.0
S	Cu	Cu	0.2222	1.0	1.0	6.0	0.0
S	Cu	Cu	0.04082	-1.0	1.0	6.0	0.0
S	Cu	Cu	0.04082	1.0	1.0	6.0	0.0
S	Cu	Cu	0.04082	-1.0	6.0	6.0	0.0
S	Cu	Cu	0.04082	1.0	6.0	6.0	0.0
S	Cu	Cu	0.01653	-1.0	1.0	6.0	0.0
S	Cu	Cu	0.01653	1.0	1.0	6.0	0.0
S	Cu	Cu	0.01653	-1.0	6.0	6.0	0.0
S	Cu	Cu	0.01653	1.0	6.0	6.0	0.0
Cu	S	S	0.2222	-1.0	1.0	6.0	0.0
Cu	S	S	0.2222	1.0	1.0	6.0	0.0
Cu	S	S	0.2222	1.0	6.0	6.0	0.0
Cu	S	S	0.04082	-1.0	1.0	6.0	0.0
Cu	S	S	0.04082	1.0	1.0	6.0	0.0
Cu	S	S	0.04082	-1.0	6.0	6.0	0.0
Cu	S	S	0.04082	1.0	6.0	6.0	0.0
Cu	S	S	0.01653	-1.0	1.0	6.0	0.0
Cu	S	S	0.01653	1.0	1.0	6.0	0.0
Cu	S	S	0.01653	-1.0	6.0	6.0	0.0
Cu	S	S	0.01653	1.0	6.0	6.0	0.0
Cu	S	Cu	0.04082	-1.0	1.0	6.0	0.0
Cu	S	Cu	0.04082	1.0	1.0	6.0	0.0
Cu	S	Cu	0.04082	1.0	6.0	6.0	0.0
Cu	S	Cu	0.01653	-1.0	1.0	6.0	0.0
Cu	S	Cu	0.01653	1.0	1.0	6.0	0.0
Cu	S	Cu	0.01653	1.0	6.0	6.0	0.0
Cu	Cu	Cu	0.04082	-1.0	1.0	6.0	0.0
Cu	Cu	Cu	0.04082	1.0	1.0	6.0	0.0
Cu	Cu	Cu	0.04082	-1.0	6.0	6.0	0.0
Cu	Cu	Cu	0.04082	1.0	6.0	6.0	0.0
Cu	Cu	Cu	0.01653	-1.0	1.0	6.0	0.0
Cu	Cu	Cu	0.01653	1.0	1.0	6.0	0.0
Cu	Cu	Cu	0.01653	-1.0	6.0	6.0	0.0
Cu	Cu	Cu	0.01653	1.0	6.0	6.0	0.0

Continued on the next page

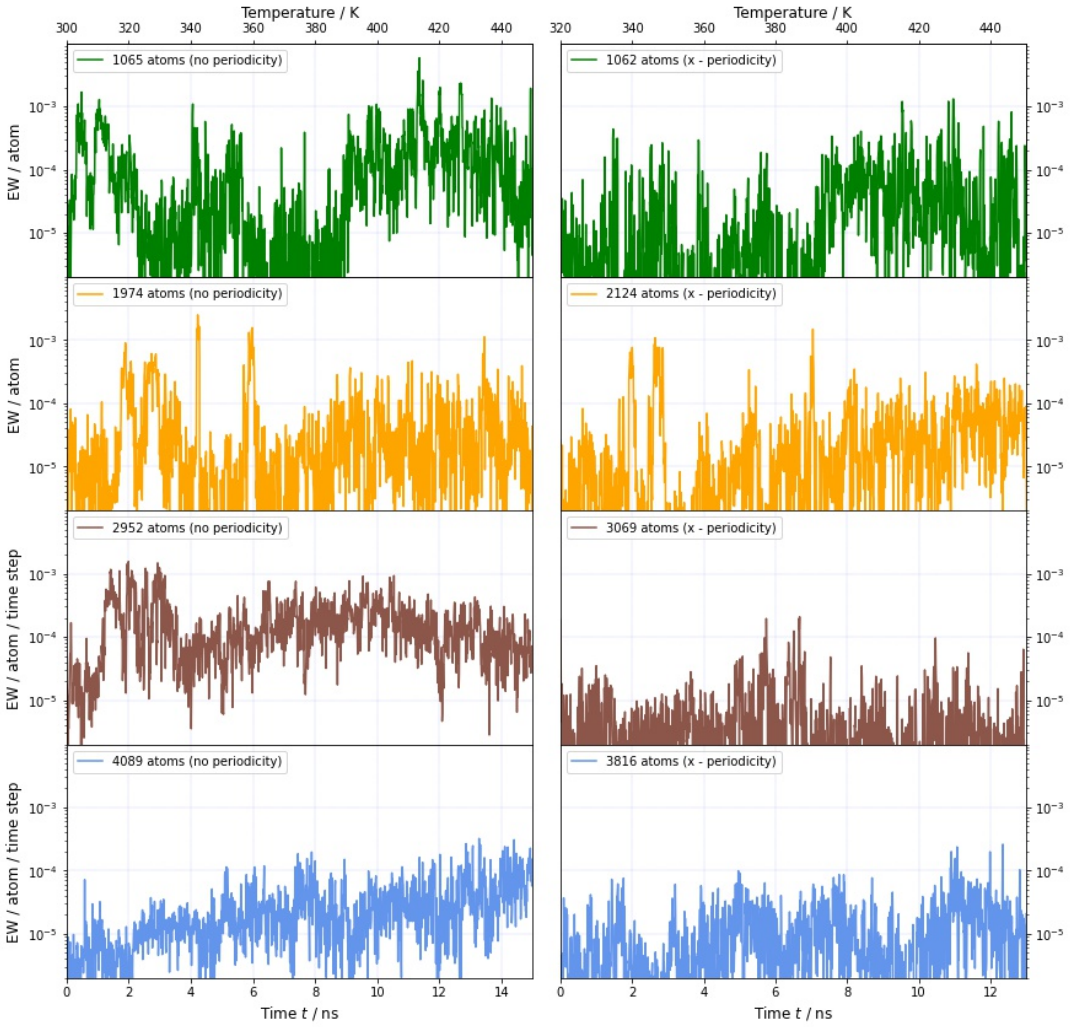
Element	Neighbor 1	Neighbor 2	$\eta \cdot \AA$	λ	ζ	$r_c/\AA$	$r_s/\AA$
Wide angular symmetry functions							
S	S	S	0.2222	-1.0	1.0	6.0	0.0
S	S	S	0.2222	1.0	1.0	6.0	0.0
S	S	S	0.2222	1.0	6.0	6.0	0.0
S	S	S	0.04082	-1.0	1.0	6.0	0.0
S	S	S	0.04082	1.0	1.0	6.0	0.0
S	S	S	0.04082	-1.0	6.0	6.0	0.0
S	S	S	0.04082	1.0	6.0	6.0	0.0
S	S	S	0.01653	-1.0	1.0	6.0	0.0
S	S	S	0.01653	1.0	1.0	6.0	0.0
S	S	S	0.01653	-1.0	6.0	6.0	0.0
S	S	S	0.01653	1.0	6.0	6.0	0.0
S	S	Cu	0.2222	-1.0	1.0	6.0	0.0
S	S	Cu	0.2222	1.0	1.0	6.0	0.0
S	S	Cu	0.2222	-1.0	6.0	6.0	0.0
S	S	Cu	0.2222	1.0	6.0	6.0	0.0
S	S	Cu	0.04082	-1.0	1.0	6.0	0.0
S	S	Cu	0.04082	1.0	1.0	6.0	0.0
S	S	Cu	0.04082	-1.0	6.0	6.0	0.0
S	S	Cu	0.04082	1.0	6.0	6.0	0.0
S	S	Cu	0.01653	-1.0	1.0	6.0	0.0
S	S	Cu	0.01653	1.0	1.0	6.0	0.0
S	S	Cu	0.01653	-1.0	6.0	6.0	0.0
S	S	Cu	0.01653	1.0	6.0	6.0	0.0
S	Cu	Cu	0.2222	-1.0	1.0	6.0	0.0
S	Cu	Cu	0.2222	1.0	1.0	6.0	0.0
S	Cu	Cu	0.2222	1.0	6.0	6.0	0.0
S	Cu	Cu	0.04082	-1.0	1.0	6.0	0.0
S	Cu	Cu	0.04082	1.0	1.0	6.0	0.0
S	Cu	Cu	0.04082	-1.0	6.0	6.0	0.0
S	Cu	Cu	0.04082	1.0	6.0	6.0	0.0
S	Cu	Cu	0.01653	-1.0	1.0	6.0	0.0
S	Cu	Cu	0.01653	1.0	1.0	6.0	0.0
S	Cu	Cu	0.01653	-1.0	6.0	6.0	0.0
S	Cu	Cu	0.01653	1.0	6.0	6.0	0.0
Cu	S	S	0.2222	-1.0	1.0	6.0	0.0
Cu	S	S	0.2222	1.0	1.0	6.0	0.0
Cu	S	S	0.2222	1.0	6.0	6.0	0.0
Cu	S	S	0.04082	-1.0	1.0	6.0	0.0
Cu	S	S	0.04082	1.0	1.0	6.0	0.0
Cu	S	S	0.04082	-1.0	6.0	6.0	0.0
Cu	S	S	0.04082	1.0	6.0	6.0	0.0
Cu	S	S	0.01653	-1.0	1.0	6.0	0.0
Cu	S	S	0.01653	1.0	1.0	6.0	0.0
Cu	S	S	0.01653	-1.0	6.0	6.0	0.0
Cu	S	S	0.01653	1.0	6.0	6.0	0.0

Continued on the next page

Element	Neighbor 1	Neighbor 2	$\eta \cdot \AA$	λ	ζ	$r_c/\AA$	$r_s/\AA$
Cu	S	Cu	0.2222	-1.0	1.0	6.0	0.0
Cu	S	Cu	0.2222	1.0	1.0	6.0	0.0
Cu	S	Cu	0.2222	1.0	6.0	6.0	0.0
Cu	S	Cu	0.04082	-1.0	1.0	6.0	0.0
Cu	S	Cu	0.04082	1.0	1.0	6.0	0.0
Cu	S	Cu	0.04082	-1.0	6.0	6.0	0.0
Cu	S	Cu	0.04082	1.0	6.0	6.0	0.0
Cu	S	Cu	0.01653	-1.0	1.0	6.0	0.0
Cu	S	Cu	0.01653	1.0	1.0	6.0	0.0
Cu	S	Cu	0.01653	-1.0	6.0	6.0	0.0
Cu	S	Cu	0.01653	1.0	6.0	6.0	0.0
Cu	Cu	Cu	0.2222	-1.0	1.0	6.0	0.0
Cu	Cu	Cu	0.2222	1.0	1.0	6.0	0.0
Cu	Cu	Cu	0.04082	-1.0	1.0	6.0	0.0
Cu	Cu	Cu	0.04082	1.0	1.0	6.0	0.0
Cu	Cu	Cu	0.04082	-1.0	6.0	6.0	0.0
Cu	Cu	Cu	0.04082	1.0	6.0	6.0	0.0
Cu	Cu	Cu	0.01653	-1.0	1.0	6.0	0.0
Cu	Cu	Cu	0.01653	1.0	1.0	6.0	0.0
Cu	Cu	Cu	0.01653	-1.0	6.0	6.0	0.0
Cu	Cu	Cu	0.01653	1.0	6.0	6.0	0.0

Table 10: Symmetry function parameters.

11.3 Extrapolation warnings for the first eight simulations



Those are the extrapolation warnings of all simulations presented in 6. The sum of all EWs was checked every 10000 time steps. The fraction EW / atom / time step is defined by

$$\frac{\text{EWs in 10000 time steps}}{\text{\#of atoms} * 10000}.$$

11.4 Details of various algorithms

First, the algorithm that decided the number of atoms at the surface is discussed. This is followed by a description of the algorithm that was used to repeatedly decide which atoms belong to any of the three zones: center, transition zone and surface. To decide the number of atoms at the surface only the starting configuration of the simulation was used. This number can in principle change during the simulation.
In case of a periodic structure, an atom is considered to be at the surface if it fulfills the following condition:

$$R - 0.2\,\text{Å} \leq r \leq R,$$

where r is its distance from the cylinders height axis, and R its radius. For non periodic structures a little bit more care is needed to find the right surface conditions: an atom is counted to be at the surface if the condition $\neg($ 11.4.42 $\wedge$ 11.4.43) is true:

$$r \leq R - 0.2\,\text{Å} \tag{11.4.42}$$

$$x_{min} + 0.2\,\text{Å} \leq x \leq x_{max} - 0.2\,\text{Å}, \tag{11.4.43}$$

where x_{min}/x_{max} are the lowest/ highest x - values of all atoms. The condition (11.4.42 $\wedge$ 11.4.43) simply means that the atom is found in a cylinder inside the actual cylinder with a radius that is 2 Å smaller, and 2 Å have been cut off from both ends of the larger cylinder. Not being inside the smaller cylinder is equivalent to being at most 2 Å away from the surface.

The algorithm for repeatedly deciding which zone an atom is in is very similar:
For a periodic structure, an atom is in the center if

$$r < \frac{R}{3}, \tag{11.4.44}$$

it is in the transition zone if

$$r < \frac{2R}{3}, \tag{11.4.45}$$

and the particle does *not* fulfill 11.4.44. If none of these conditions is fulfilled it is counted to be on the surface. For non-periodic geometries the algorithm is again slightly more complicated:
A particle is in the center if it fulfills the following two conditions:

$$r < \frac{R}{3} \tag{11.4.46}$$

$$\frac{x_{min} + x_{max}}{2} - \frac{L}{6} < x < \frac{x_{min} + x_{max}}{2} + \frac{L}{6}, \tag{11.4.47}$$

where L is the length of the cylinder. Atoms in the transition zone are defined by fulfilling the condition $\neg($ 11.4.46 $\wedge$ 11.4.47) and the two conditions

$$r < \frac{2R}{3} \tag{11.4.48}$$

$$\frac{x_{min} + x_{max}}{2} - \frac{L}{3} < x < \frac{x_{min} + x_{max}}{2} + \frac{L}{3}. \tag{11.4.49}$$

Atoms that are neither in the center, nor in the transition zone are counted to be on the surface.
This subsection is concluded by showing an illustration of the fitting procedure that was used to calculate diffusion coefficients w.r.t. zones:

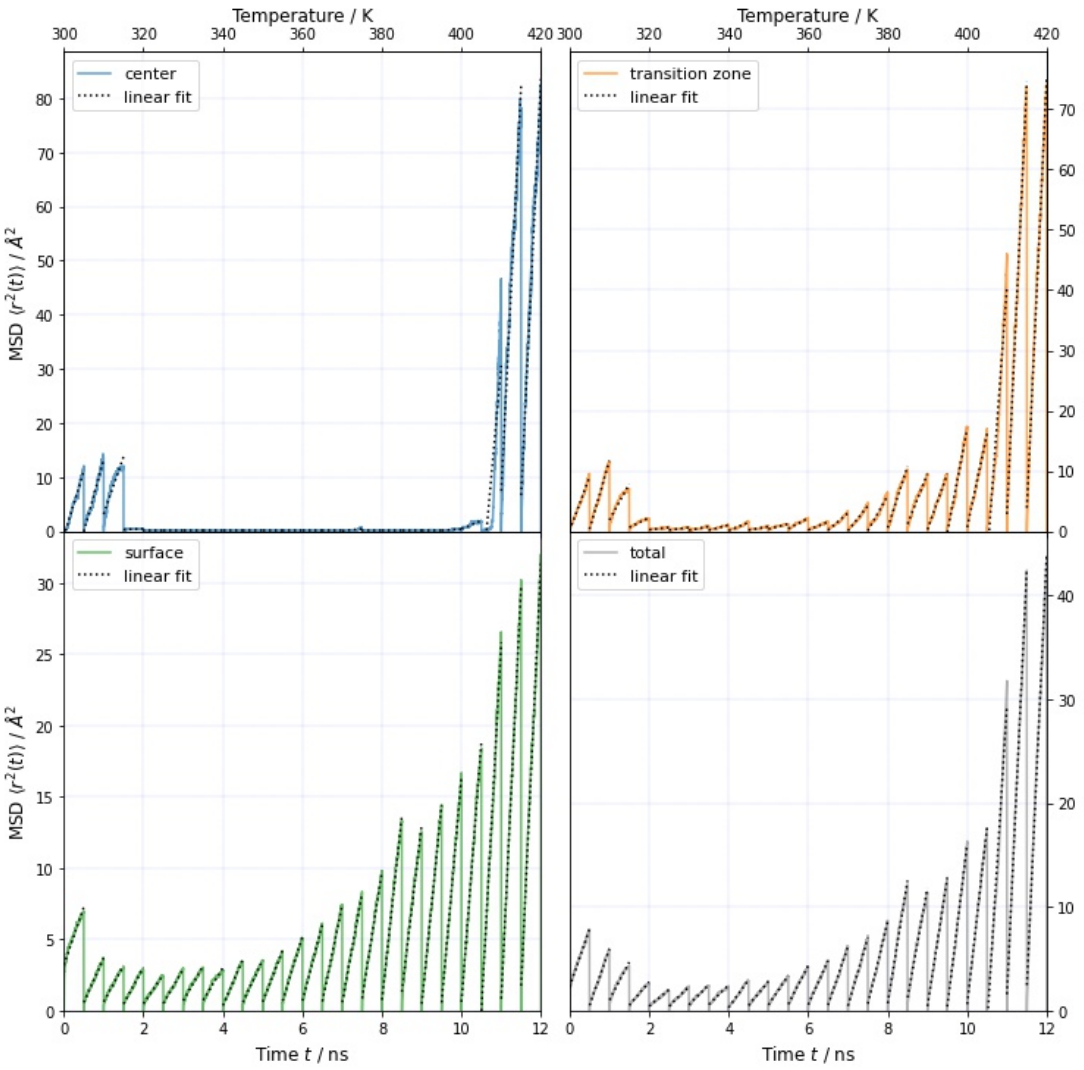


Fig. 34: This figure shows all linear fits to calculate the diffusion coefficients w.r.t. zones from the entire trajectory of the non-periodic cylinder of the second set of simulations (3.5 nm diameter, 6.039 nm length, 3816 atoms). To avoid correlations from the previous trajectory, the MSD has been reset to zero after every 500 ps, and each of the 500 ps intervals has been used to calculate a diffusion coefficient.

11.5 Equilibrations prior to simulations

Some of the structures of the first set of simulations had been equilibrated prior to the actual simulation. The pre-heating simulations of the second set of simulations can be found elsewhere 9.1. We show the energy curves and for the case of the two smallest periodic structures also the copper diffusion (since it started unreasonably high in the actual simulation, suggesting a prior phase transition).

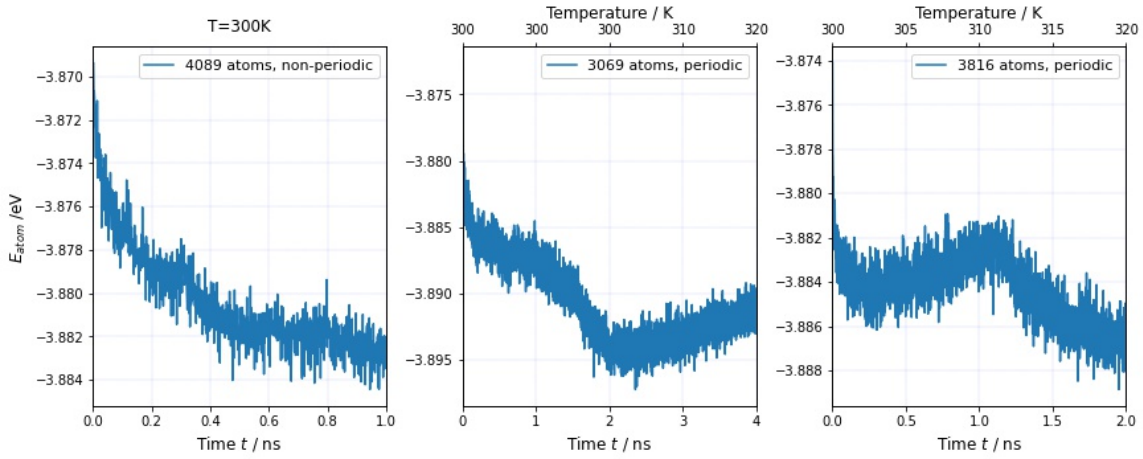


Fig. 35: Those are the energy curves for three of the equilibrations. All energies start at a far too high value and especially the right curve suggests a phase transition from high to low chalcocite.

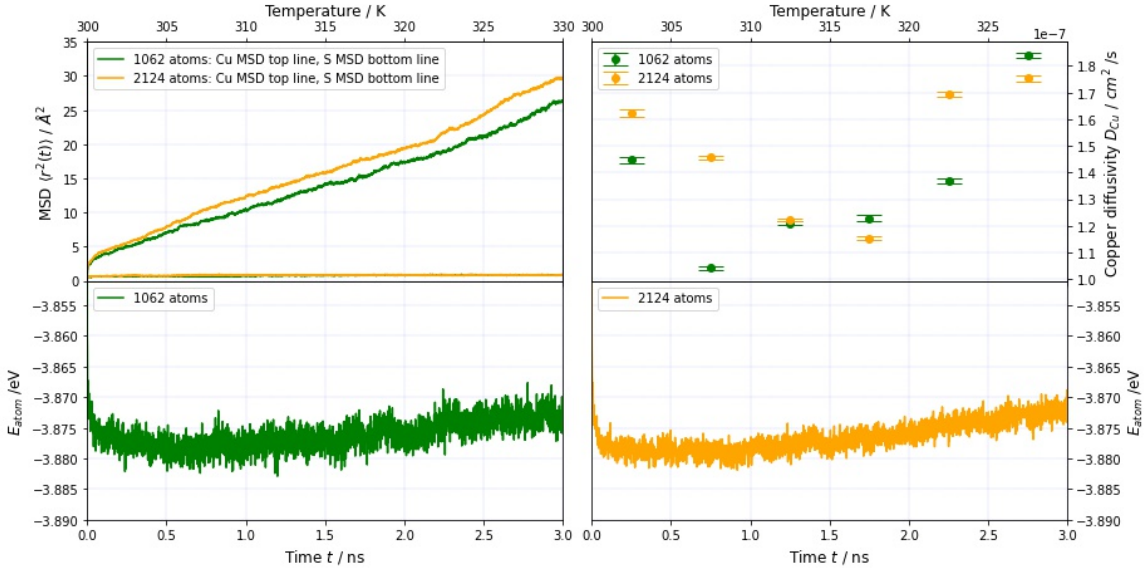
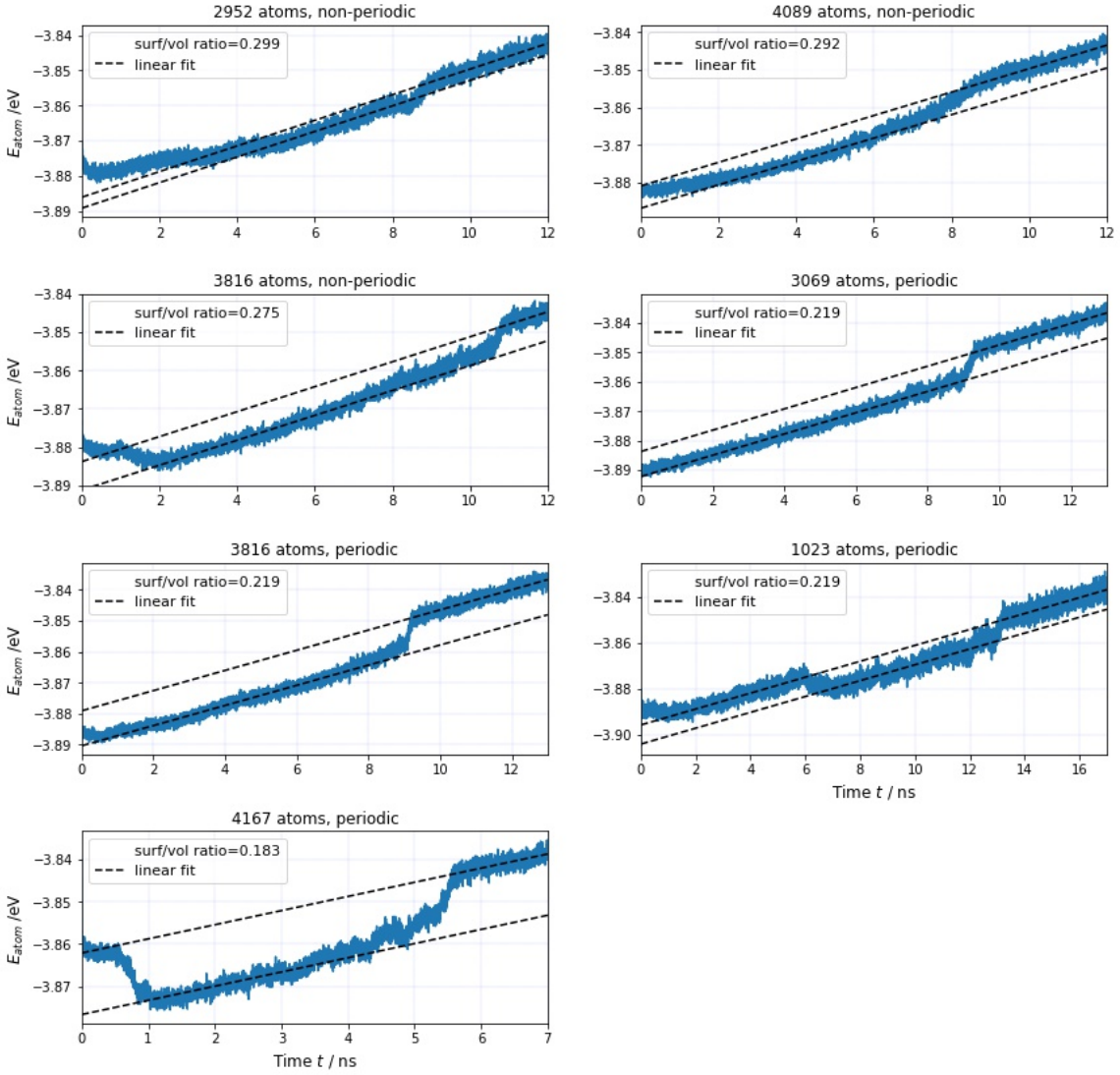


Fig. 36: This is the thermodynamic data for the two smallest periodic structure when started from an unequilibrated geometry at 300 K. They again start at too high energies.

11.6 Energy step surface ratio dependence

The fits for all energy curves are shown. Since the energy curves were not always straight in any of the phases, choosing a range of values for the fit strongly influenced the resulting fit. Thus, the results are only a rough estimate. The ranges that were used for fitting are shown in the table below.



# Atoms	Periodic	surf / vol ratio	range fit 1 / ns	range fit 2 / ns
2952		0.299	4.5-8	11.211-end
4089		0.292	3-6	10-end
3816		0.275	2-8	11.653-end
3069	✓	0.219	1-9	11-end
3816	✓	0.219	1-8.5	10-end
1023	✓	0.219	6.65-11.5	15-end
4167	✓	0.183	0.8-3.745	6.2-end

References

[1] Claudia Coughlan, Maria Ibanez, Oleksandr Dobrozhan, Ajay Singh, Andreu Cabot, and Kevin M Ryan. Compound copper chalcogenide nanocrystals. *Chemical reviews*, 117(9):5865–6109, 2017.

[2] Poulomi Roy and Suneel Kumar Srivastava. Nanostructured copper sulfides: synthesis, properties and applications. *CrystEngComm*, 17(41):7801–7815, 2015.

[3] Mona Lotfipour, Tony Machani, Daniel P Rossi, and Katherine E Plass. α-chalcocite nanoparticle synthesis and stability. *Chemistry of Materials*, 23(12):3032–3038, 2011.

[4] Yue Wu, Cyrus Wadia, Wanli Ma, Bryce Sadtler, and A Paul Alivisatos. Synthesis and photovoltaic application of copper (i) sulfide nanocrystals. *Nano letters*, 8(8):2551–2555, 2008.

[5] Howard T Evans. Crystal structure of low chalcocite. *Nature Physical Science*, 232(29):69–70, 1971.

[6] Howard T Evans. The crystal structures of low chalcocite and djurleite. *Zeitschrift für Kristallographie-Crystalline Materials*, 150(1-4):299–320, 1979.

[7] Lin-Wang Wang. High chalcocite cu 2 s: A solid-liquid hybrid phase. *Physical review letters*, 108(8):085703, 2012.

[8] MJ Buerger and Bernhardt J Wuensch. Distribution of atoms in high chalcocite, cu2s. *Science*, 141(3577):276–277, 1963.

[9] Bernhardt J Wuensch and MJ Buerger. The crystal structure of chalcocite, cu2s. *Mineral Soc Am Spec Paper*, 1:164–170, 1963.

[10] Ryoichi Sadanaga, Masaaki Ohmasa, and Nobuo Morimoto. On the statistical distribution of copper ions in the structure of β-chalcocite. *Mineralogical Journal*, 4(4):275–290, 1965.

[11] Haimei Zheng, Jessy B Rivest, Timothy A Miller, Bryce Sadtler, Aaron Lindenberg, Michael F Toney, Lin-Wang Wang, Christian Kisielowski, and A Paul Alivisatos. Observation of transient structural-transformation dynamics in a cu2s nanorod. *Science*, 333(6039):206–209, 2011.

[12] Jessy B Rivest, Lam-Kiu Fong, Prashant K Jain, Michael F Toney, and A Paul Alivisatos. Size dependence of a temperature-induced solid–solid phase transition in copper (i) sulfide. *The Journal of Physical Chemistry Letters*, 2(19):2402–2406, 2011.

[13] Max Born and Robert Oppenheimer. Zur quantentheorie der molekeln. *Annalen der physik*, 389(20):457–484, 1927.

[14] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical review letters*, 98(14):146401, 2007.

[15] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

[16] Andreas Singraber, Tobias Morawietz, Joerg Behler, and Christoph Dellago. Parallel multistream training of high-dimensional neural network potentials. *Journal of chemical theory and computation*, 15(5):3075–3092, 2019.

[17] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[18] Balázs Csanád Csáji et al. Approximation with artificial neural networks. *Faculty of Sciences, Etvs Lornd University, Hungary*, 24(48):7, 2001.

[19] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 315–323, 2011.

[20] Bekir Karlik and A Vehbi Olgac. Performance analysis of various activation functions in generalized mlp architectures of neural networks. *International Journal of Artificial Intelligence and Expert Systems*, 1(4):111–122, 2011.

[21] Dan Ciregan, Ueli Meier, and Jürgen Schmidhuber. Multi-column deep neural networks for image classification. In *2012 IEEE conference on computer vision and pattern recognition*, pages 3642–3649. IEEE, 2012.

[22] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[23] Jon Russell. Google's alphago ai wins three-match series against the world's best go player. *TechCrunch. https://techcrunch. com/2017/05/24/alphago-beats-planets-best-human-go-player-ke-jie/(25 May 2017)*, 2017.

[24] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.

[25] Rich Caruana, Steve Lawrence, and C Lee Giles. Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. In *Advances in neural information processing systems*, pages 402–408, 2001.

[26] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. 1960.

[27] Gerald L Smith, Stanley F Schmidt, and Leonard A McGee. *Application of statistical filter theory to the optimal estimation of position and velocity on board a circumlunar vehicle*. National Aeronautics and Space Administration, 1962.

[28] Felix Heimes. Extended kalman filter neural network training: experimental results and algorithm improvements. In *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 98CH36218)*, volume 2, pages 1639–1644. IEEE, 1998.

[29] Emad W Saad, Danil V Prokhorov, and Donald C Wunsch. Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks. *IEEE Transactions on neural networks*, 9(6):1456–1470, 1998.

[30] Eric A Wan, Rudolph Van Der Merwe, and Simon Haykin. The unscented kalman filter. *Kalman filtering and neural networks*, 5(2007):221–280, 2001.

[31] Lee A Feldkamp and Gintaras V Puskorius. A signal processing framework based on dynamic neural networks with application to problems in adaptation, filtering, and classification. *Proceedings of the IEEE*, 86(11):2259–2277, 1998.

[32] Mel Levy. On the simple constrained-search reformulation of the hohenberg–kohn theorem to include degeneracies and more (1964–1979). *International Journal of Quantum Chemistry*, 110(15):3140–3144, 2010.

[33] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.

[34] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.

[35] Georg Kresse and J Furthmüller. Vienna ab-initio simulation package (vasp). *Vienna: Vienna University*, 2001.

[36] Andreas Singraber, Joerg Behler, and Christoph Dellago. Library-based lammps implementation of high-dimensional neural network potentials. *Journal of chemical theory and computation*, 15(3):1827–1840, 2019.

[37] Lammps. http://lammps.sandia.gov. Accessed: 2020-10-31.

[38] Nongnuch Artrith and Jörg Behler. High-dimensional neural network potentials for metal surfaces: A prototype study for copper. *Physical Review B*, 85(4):045439, 2012.

[39] Jörg Behler. Constructing high-dimensional neural network potentials: A tutorial review. *International Journal of Quantum Chemistry*, 115(16):1032–1050, 2015.

[40] Robert T. McGibbon, Kyle A. Beauchamp, Matthew P. Harrigan, Christoph Klein, Jason M. Swails, Carlos X. Hernández, Christian R. Schwantes, Lee-Ping Wang, Thomas J. Lane, and Vijay S. Pande. Mdtraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophysical Journal*, 109(8):1528 – 1532, 2015.

[41] Richard J Gowers, Max Linke, Jonathan Barnoud, Tyler John Edward Reddy, Manuel N Melo, Sean L Seyler, Jan Domanski, David L Dotson, Sébastien Buchoux, Ian M Kenney, et al. Mdanalysis: a python package for the rapid analysis of molecular dynamics simulations. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2019.

[42] Naveen Michaud-Agrawal, Elizabeth J Denning, Thomas B Woolf, and Oliver Beckstein. Mdanalysis: a toolkit for the analysis of molecular dynamics simulations. *Journal of computational chemistry*, 32(10):2319–2327, 2011.

[43] Steven M Crunk and Xinh Huynh. On minimizing the standard error of the slope in simple linear regression. *Open Journal of Statistics*, 2014, 2014.

[44] Sanford Weisberg. *Applied linear regression*, volume 528. John Wiley & Sons, 2005.