Identifying early stage precipitation in large-scale atomistic simulations of superalloys

You may also be interested in:

Thermodynamic interpretation of reactive processes in Ni–Al nanolayers from atomistic simulations
Luis Sandoval, Geoffrey H Campbell and Jaime Marian

First-principles theory of 250 000-atom microstructure
C Wolverton

Modelling defects in Ni–Al with EAM and DFT calculations
F Bianchini, J R Kermode and A De Vita

Topological fingerprints for intermetallic compounds for the automated classification of atomistic simulation data
T Schablitzki, J Rogal and R Drautz

Effects of temperature gradient and elastic strain on spinodal decomposition and microstructure evolution of binary alloys
Yongsheng Li, Yuxuan Pang, Xingchao Wu et al.

Kinetics of heterogeneous grain boundary precipitation of Ni3Al in nickel alloy
C Hin

A combined cluster dynamics/kinetic Monte Carlo model for precipitate nucleation in interstitial/substitutional alloys
J Ženíšek, J Svoboda, E Kozeschnik et al.

Solute clustering in Al–Mg binary alloys
D Zhang and R C Picu

Identifying early stage precipitation in large-scale atomistic simulations of superalloys

Eric Schmidt and Paul D Bristowe

Department of Materials Science and Metallurgy, University of Cambridge, Cambridge CB3 0FS, United Kingdom

E-mail: pdb1000@cam.ac.uk

Abstract

A method for identifying and classifying ordered phases in large chemically and thermally disordered atomistic models is presented. The method uses Steinhardt parameters to represent local atomic configurations and develops probability density functions to classify individual atoms using naïve Bayes. The method is applied to large molecular dynamics simulations of supersaturated Ni-20 at% Al solid solutions in order to identify the formation of embryonic γ'-Ni$_3$Al. The composition and temperatures are chosen to promote precipitation, which is observed in the form of ordering and is found to occur more likely in regions with above average Al concentration producing 'clusters' of increasing size. The results are interpreted in terms of a precipitation mechanism in which the solid solution is unstable with respect to ordering and potentially followed by either spinodal decomposition or nucleation and growth.

Keywords: atomistic simulation, supervised learning, γ' precipitation, superalloys

(Some figures may appear in colour only in the online journal)

1. Introduction

Classical atomistic simulations involving more than a million atoms are now commonplace, in one extreme case with up to 4.125×10^{12} particles [1]. They are used to probe the structure

and kinetics of materials on a scale that is usually not feasible from first principles and can access phenomena such as intergranular fracture, surface adsorption, dislocation dynamics, grain growth and precipitation. Most of these processes are straightforward to understand using standard visualisation or animation techniques even when the computational cell is large. This is not the case for early stage precipitation where the nuclei form on the sub-nanometre scale and are embedded deep within a matrix material. Identifying where and when the nuclei form, as well as their size, shape and distribution is of fundamental importance in the theory of phase transformations. As the precipitates grow and affect the macroscopic properties of a material, such as creep and fracture strength, they become of increasing technological interest. A well-known example is the nucleation and growth of γ'-Ni$_3$Al precipitates in Ni-base superalloys [2]. Here the L1$_2$ ordered precipitates (γ') impede the motion of dislocations in the γ (fcc) solid solution matrix thereby strengthening the alloy such that it can be used in extreme environments such as those found in a turbine engine [3]. In the early stage of precipitation in a Ni–Al alloy the precipitates are thought to be roughly spherical and then as they coarsen they become cuboidal [4, 5]. Cubes form to minimise the precipitate/matrix lattice mismatch and create coherency.

In this work we are concerned with the embryonic stage of the precipitation where small ordered clusters form acting as precursors for the later decomposition process on a large scale. While precipitation on the large scale is quite well understood there are still uncertainties about the process on the atomic level shortly after annealing and during quenching. Generally precipitation depends on the temperature treatment and the concentration profile. It has been clear for some time that chemical clustering and ordering are interdependent during precipitation [6]. This interdependency can be explained with a graphical method based on geometrical considerations of the free energy curve of the observed phases [7]. Using the free energy curves the method distinguishes between heterogeneous ordering (the solid solution is *metastable* with respect to ordering) and homogeneous ordering (the solid solution is *unstable* with respect to ordering). Heterogeneous ordering is followed by spinodal decomposition whereas homogeneous ordering is followed by spinodal decomposition or nucleation and growth depending on the second derivative of the free energy with respect to the concentration. This approach successfully explains experimental observations such as the *conditional spinodal* which is a spinodal decomposition process that requires prior ordering [8]. The method also agrees with theoretical predictions based on mean field and phase field theory [9–12]. Studies using atom probe and electron microscopy show that the precipitation process is very fast and already takes place during the quenching process. In this case precipitation forms interconnected fluctuations of the concentration profile suggesting chemical ordering and spinodal decomposition [13–15]. However, because of the time constraints in the experiments, it is not perfectly clear in which order chemical ordering and spinodal decomposition really happens. This assumes that the observed Ni-based alloys are systems that are in fact *unstable* with respect to ordering. An objective of this study is therefore to shed light on precipitation at the atomic level using classical molecular dynamics (MD), investigating the process during quenching.

In this work we study the initial stages of precipitation by simulating binary Ni-20 at% Al single crystal systems in the $\gamma' + \gamma$ two-phase region, which are unstable with respect to ordering. For the simulations we use an empirical potential [16] of the embedded atom method (EAM) type [17]. In order to distinguish chemically ordered from disordered regions we describe the local neighbourhood of individual atoms and the resulting geometrical 'bonds' with Steinhardt parameters [18]. Additionally we make these Steinhardt parameters sensitive to the chemical environment and use them as input for a naïve Bayes classifier that we train. We demonstrate that the classifier is highly robust, identifying ordered regions

correctly under significant chemical and thermal disorder. More importantly, we observe ordering in all our simulations, particularly in regions with above average Al concentration. We interpret this finding as confirmation of the proposed homogeneous ordering which precedes possible spinodal decomposition for the simulated system.

2. Theory

Identifying local atomic configurations via numerical algorithms requires a representation of those configurations. Of importance in this case is a representation that is robust enough to capture rotation symmetries of configurations subject to significant noise. Current ways of representing atomic configurations include bond angle cosines, so-called Steinhardt parameters, bond orders, Voronoi analysis and common neighbours [19]. These methods can be used to distinguish between liquids, simple crystal structures (fcc, bcc, hcp, diamond cubic) and various kinds of defects [19–21].

For this work Steinhardt parameters [18] are chosen since they are known to provide a good basis for distinguishing various lattice types. Once parameters are found that capture the local configuration corresponding distributions can be obtained, e.g. for the fcc lattice type. Here the Steinhardt parameters will be made sensitive to the chemical surroundings, similar to a previous study on connectivity in damaged ceramics [22]. Then a decision rule is needed to classify local atomic configurations and for this naïve Bayes is chosen. The combination of both the normalised parameter distributions and the decision is referred to here as the 'classifier'.

2.1. Steinhardt parameters

Steinhardt parameters are calculated by selecting individual atoms and their surroundings up to a threshold distance or threshold number of neighbours N [18]. These thresholds determine the possible resolution of the atomic configurations that are identified. Including all atoms in the calculation of the Steinhardt parameters, i.e. using an infinite cut-off radius, yields a bulk description of the system [21].

The Steinhardt parameters q_l utilised in this paper are given by

$$q_l \equiv \sqrt{\frac{4\pi}{2l+1}\sum_{m=-l}^{l}|\bar{Q}_{l,m}|^2}\ , \tag{1}$$

where

$$\bar{Q}_{l,m} \equiv \frac{1}{N}\sum_{n=1}^{N}Y_l^m(\theta_n,\phi_n). \tag{2}$$

Y_l^m represents spherical harmonic functions of degree l and order $m \in [-l,\ l]$ with polar angle $\theta \in [0,2\pi]$ and azimuth angle $\phi \in [0,\ \pi]$. The definition in equation (2) provides a local average for fixed m and l and the quantities $\bar{Q}_{l,m}$ can be interpreted as vectors for a fixed l. This makes q_l in equation (1) the corresponding norm which is calculated for each atom in the current atomic assembly. Note that the q_l parameters are only a subset of the Steinhardt parameters.

Besides the usefulness of the Steinhardt parameters in identifying crystal structures they have also been used in MD simulations as collective variables biasing given potentials for Umbrella sampling or Metadynamics simulations [23, 24].

2.2. Classification of atoms with naïve Bayes

The classification of atoms here makes use of distinct distributions of Steinhardt parameters for chemically ordered and disordered systems. This technique has been used before to distinguish between fcc, bcc and hcp lattice types or structures [18, 21, 25]. Normalising these distributions yields conditional probability density functions (PDFs) P(neighbourhood | phase) which can be used for classification using the Bayesian framework. Given a certain phase the conditional PDF P(neighbourhood | phase) returns a probability for an atom based on its neighbourhood. Note that when we refer to 'phase' in the context of classification we mean the local atomic arrangement and not the phase in the thermodynamic limit. The advantage of a Bayesian approach for the classification of atomic structures is that it is not limited to Steinhardt parameters. Given that one has a PDF for each phase of interest, the following Bayesian decision rule can be used for classification:

$$\text{Phase} = \text{argmax}_{\text{phase} \in \text{Phases}} P(\text{neighbourhood}|\text{phase}) P(\text{phase}), \tag{3}$$

where the neighbourhood here is a list of Steinhardt parameters for each atom based on its neighbourhood. Then a value for P(phase) has to be chosen, the prior probability of a phase. Due to a lack of knowledge about which phase should be more likely to appear a priori, we assume a so-called uniform prior P(phase) $= 1/$number of phases. This makes equation (3) the naïve Bayes decision rule. In order to distinguish between chemically ordered and disordered configurations for the same lattice type, we slightly modify the calculation of the Steinhardt parameters. Based on the chemical element of the current atom being processed this modification selects the surroundings using three filters: (1) 'all' elements, (2) only 'like' elements and (3) only 'unlike' elements. This selection is done after the neighbourhood has been searched using either a cut-off radius or maximum number of nearest neighbours. This scheme results in two additional Steinhardt parameters per atom and l. To distinguish them they are denoted as $q_{l,\text{all}}$, $q_{l,\text{like}}$ and $q_{l,\text{unlike}}$. Here $q_{l,\text{like}}$ and $q_{l,\text{unlike}}$ are particularly useful for identifying the γ' phase.

3. Method

In order to apply the naïve Bayes decision rule for the classification of atomic configurations we need to obtain the conditional PDFs. The 'classifier' here is then the combination of conditional PDFs and the naïve Bayes decision rule. These conditional PDFs should be very accurate under significant chemical and thermal noise and therefore care needs to be taken in their construction. For this we use, in the present case, single crystal examples of γ Ni-20 at% Al and perfectly stoichiometric γ', as described section 3.1. Since the purpose of the developed conditional PDFs is to differentiate between chemical order and disorder during annealing we include a cross-validation step to test the resulting classifier. The cross-validation process is addressed in section 3.2.

3.1. Obtaining conditional PDFs

To produce conditional PDFs, which yield a robust classifier, we use different degrees of lattice distortion for the given phases to produce 'samples' with which to train. Two methods were used to obtain samples for the classification of atoms into different 'phases': (1) annealing in MD and (2) using artificial noise created with a random number generator. Both methods introduce distortions into ideal crystals. We found that the latter method is surprisingly successful and we will therefore focus on that one in the following.

Table 1. Comparison between temperature (T) and the degree of distortion (η) for the γ' phase in terms of the root mean square displacement (RMSD) averaged over all frames. The model used contained $10 \times 10 \times 10$ unit cells and the RMSD values are in fractional coordinates with respect to the supercell.

η	RMSD	T (K)	RMSD
10^{-5}	0.001 70	0.1	0.000 73
10^{-4}	0.005 41	500	0.006 99
10^{-3}	0.016 88	1500	0.017 00

The Steinhardt parameters were calculated after the neighbourhood search. Two common criteria for the neighbour selection are the cutoff radius and the number of neighbours. Testing both, we found that for phases having a fcc lattice a fixed number of 18 neighbours (i.e. including all first and second nearest neighbours), leads to clearer parameter distributions when considering a range of temperatures up to 1500 K.

In order to obtain a classifier which is stable against chemical disorder and able to efficiently identify γ' particles, we found it sufficient to create one conditional PDF for γ' and one for γ, for each species filter. Each conditional PDF is created incorporating three different degrees of lattice distortion, $\eta = 10^{-5}$, 10^{-4} and 10^{-3}. These lattice distortions were used to calculate the entries Δx_i of the diagonal covariance matrix for the multivariate normal distribution as $\Delta x_i = \eta \cdot a_i$, with a_i the length of the unit cell in dimension i. In this case the same η was applied to all spatial dimensions. The unit cells of both γ' and γ are cubic and hence Δx_i is the same in all directions. This leads to an increase in the distortion of the ideal lattice position with increasing η, thus simulating higher thermal fluctuations. This approach may seem quite simple but has been found to work surprisingly well and permits quick generation of multiple classifiers combining different phases and distortions. It is particularly useful in cases where the relevant empirical potentials for use in MD simulations are unavailable. Table 1 compares η with temperature via the root mean square displacement (RMSD) averaged over all frames for γ'. It is seen that the maximum distortion correlates with a temperature of about 1500 K.

Once the samples for the phases were generated, the Steinhardt parameters using all three chemical 'filters' (all, like and unlike chemical elements) were calculated for $l = 4$ and $l = 6$ which are particularly useful for fcc, bcc and hcp structures [18]. Since their values are limited by definition to be between 0 and 1 they were multiplied by a factor of 10 to improve processing within the developed software. Then normalised histograms (the conditional PDFs) were created for each phase and degree of distortion and filter. Thus we obtain three normalised histograms for a single phase with a given degree of distortion, one for each filter 'all', 'like' and 'unlike'. Since we want to create only two sets of conditional PDFs, one set for each phase containing conditional PDFs for the three filters, we need to combine the normalised histograms with different degrees of distortion. Here we used the same weighting for all degrees of distortion. Thus the conditional PDF for the 'all' filter for γ' corresponds to

$$P(\text{data} \mid \gamma') = \sum_{i=0}^{N} w_{\eta_i} P(\text{data} \mid \gamma'_{\eta_i}), \tag{4}$$

where N is the number of degrees of distortions, w_{η_i} the weight and $P(\text{data} \mid \gamma'_{\eta_i})$ is the conditional PDF for distortion of degree η_i. Experimenting with different combinations of distortions reveals that using only high degrees of distortion are detrimental to the overall accuracy of classification. On the other hand using only low degrees of distortion leads to
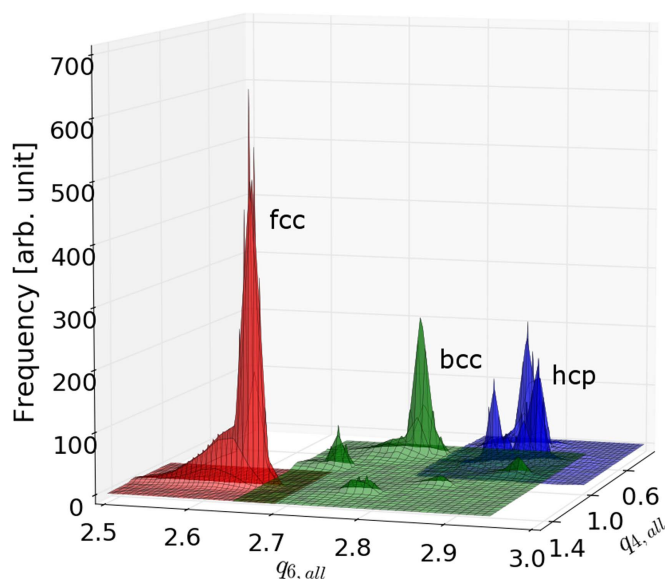
Figure 1. Conditional probability density functions for the fcc, bcc and hcp structures as a function of the Steinhardt parameters q_l ($l = 4$ and $l = 6$) shown in red, green and blue respectively.

very poor results at high distortions even though the accuracy of classification at low temperatures is good. It is found that the optimal solution is to use either an exponential function for the weights if the lattice distortions are chosen on a linear scale or to use an exponential distribution of distortions directly, as done in the present work for selecting the degrees of distortion. Hence the weights w_{η_i} here are all equal to 1. Choosing distortions on a linear scale here means, for example, carrying out MD simulations with a linear sequence of temperatures. A weighting function which decays with increasing degree of distortion appeals intuitively since we are interested in finding a phase at high temperatures, for example γ', which we know quite well from low temperatures. Hence we require a strong representation of the distribution of Steinhardt parameters from low distortions with also some corrections to that distribution at high distortions. Three conditional PDFs generated this way are shown in figure 1 for fcc, bcc and hcp structures using the same degrees of distortion as are used for γ' and γ. The figure demonstrates, with its separated peaks, why Steinhardt parameters are so successful at distinguishing between these structures even at high temperatures.

3.2. Cross-validation

In order to test the reliability of the developed classifier we created two kinds of atomic models where we embedded a spherical γ' precipitate of radius 10 Å in both pure fcc Ni and γ Ni-20 at% Al. These models are chosen since they reflect what we will be looking for in our search for ordering in our γ single crystals annealed over a long period. It is expected that the classifier does better for the Ni + γ' system than for the $\gamma + \gamma'$ system since there is less uncertainty regarding the chemical ordering present. The classification of these atomic models is then combined with binary classification scores that quantify what previously was only loosely defined as 'accuracy'. For this we use parameters quite common in statistics, the so-called 'true positive' (t_p), 'true negative' (t_n), 'false positive' (f_p), 'false negative' (f_n),
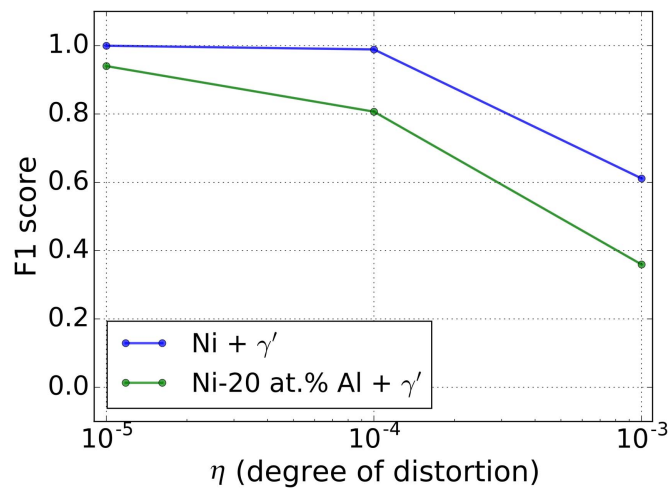
Figure 2. F1 score measuring the ability of the classifier to correctly identify γ' within pure Ni and γ' Ni-20 at% Al, for three different levels of lattice distortions η. The F1 score is computed comparing two frames of the same degree of distortion.
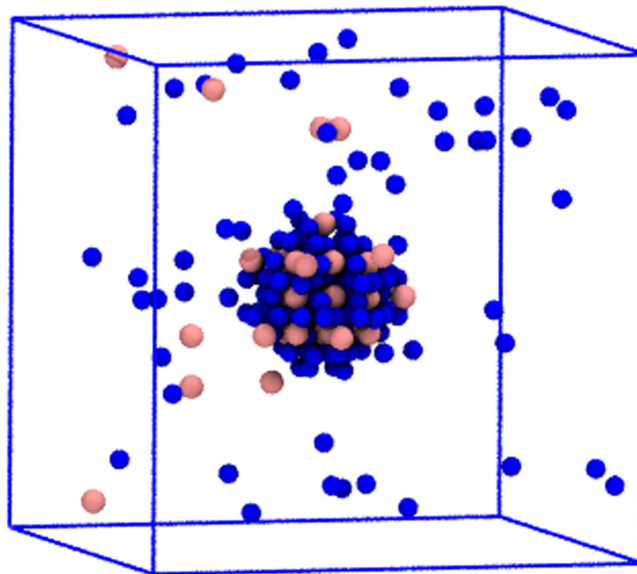

Figure 3. Atoms (Ni blue and Al pink) identified as γ'-like in a model of γ' Ni-20 at% Al containing a pre-existing spherical γ' precipitate. The system has been annealed using MD at 1500 K for 20 ps.

'precision' (P), 'recall' (R) and 'F1 score'. In short 'true positive' refers to the case where the classifier correctly identifies an atom as belonging to a phase we are looking for. On the other hand 'false positive' refers to the case that the classifier falsely identifies an atom as belonging to our phase. The other parameters t_n and f_n behave accordingly. The remaining parameters are combinations of the previous four parameters aimed at creating a score which yields a

balanced image of the classifier's performance, i.e. only taking t_p as a measure to be maximised could lead to a classifier which simply predicts all atoms to be of the phase we are looking for causing a large f_p value. The 'recall' R is defined as $R = t_p/(t_p + f_n)$ and 'precision' as $P = t_p/(t_p + f_p)$ which are combined in the 'F1 score' as $F1 = 2RP/(R + P)$. Hence the aim is to maximise the F1 score, which is limited to the interval between 0 and 1, to obtain a good classifier.

Next we take the above definition for the F1 score and apply it to our two cross-validation cases of a spherical γ' precipitate embedded in both pure Ni and γ Ni-20 at% Al. The F1 score to correctly identify γ' is calculated for each model for each degree of distortion comparing two configurations. The resulting behaviour is shown in figure 2. We can see for both cases that the F1 score decreases almost linearly with increasing degree of distortion. We also see that the F1 score for embedding γ' in pure Ni is always higher than for the embedding γ' in γ. This is to be expected since chemical disorder is not present in pure Ni, increasing the certainty of encountering γ'.

Also, we have annealed an atomistic model ($14 \times 14 \times 14$ unit cells) containing a spherical γ' precipitate of radius 10 Å embedded in γ Ni-20 at% Al for 20 ps at 1500 K. The MD was performed using the LAMMPS code [26] and a NPT barostat. This is to test the classifier's performance with respect to actual thermal distortions during MD simulations. The resulting classified final snapshot is shown in figure 3. As expected, we find some random single atoms together with the spherical γ' precipitate that we inserted in the initial configuration. The single atoms could be false positives or actually γ'-like atoms caused by the random arrangements within γ phase. The γ' sphere we find has smaller radius than the expected 10 Å. We understand this in terms of the 18 nearest neighbours contributing to the Steinhardt parameters for each atom and in terms of the classification process rather than a consequence of MD relaxation. Approaching the γ/γ' interface the value of the 'like' and 'unlike' Steinhardt parameters change due to a change in the local chemical symmetry. The classification then assigns a phase to both 'like' and 'unlike' filters for each atom, allowing for four distinct filter label combinations. If the neighbourhood for 'like' and 'unlike' filters is identified as γ' the atom is identified as γ', and similarly for γ. If on the other hand, both filters return both γ' and γ then the atom neighbourhood is in some intermediate state. The atoms in the interface between the γ' precipitate and the γ' host is a mix of γ and the intermediate states. This can be interpreted as a slight bias of the classifier towards γ over γ', which here is desirable for reducing the false positives of γ'. Though the MD relaxation certainly contributes to the specific form of the interface observed in figure 3, it is unlikely to be the main reason for the apparent reduction in the precipitate size since this effect persisted in this work when testing a range of classifiers on ideal and MD relaxed structures.

4. Results - ordering within a single crystal

A single crystal of γ Ni-20 at% Al was annealed at 1300, 1400 and 1500 K using MD as implemented in LAMMPS [26] with a NPT barostat. The model contained 108 000 atoms with thermal vacancies but no pre-existing γ' precipitate. The simulations were performed for 100 ns with time steps of 2 fs and used the Ni–Al EAM-type potentials of [16]. The number of thermal vacancies (9 at 1300 K, 23 at 1400 K and 49 at 1500 K) was determined using literature values ($H_{f,Ni} = 1.56$ eV, $H_{f,Al} = 0.775$ eV, $S_{f,Ni} = 3.3\,k_B$, $S_{f,Al} = 2.6\,k_B$) [27] of the vacancy formation enthalpies H_f and entropies S_f of pure Ni and Al as follows:
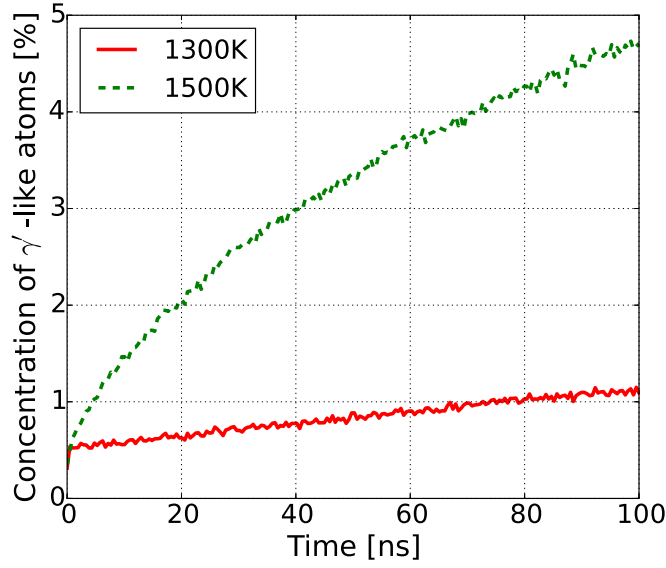
Figure 4. Concentration of atoms identified as γ'-like for annealed single crystals of 108 000 atoms over a period of 100 ns at 1300 and 1500 K.

$$N_{\text{vac}} = N_{\text{atoms}} \times \exp\left(-\frac{c_{\text{Al}}H_{\text{f,Al}}}{k_{\text{B}}T} - \frac{c_{\text{Ni}}H_{\text{f,Ni}}}{k_{\text{B}}T} + c_{\text{Al}}S_{\text{f,Al}} + c_{\text{Ni}}S_{\text{f,Ni}}\right), \tag{5}$$

where N_{atoms} is the total number of atoms in the model and c_{Al} and c_{Ni} are the concentrations of Al and Ni respectively. Note that for H_{f} and S_{f} the literature values were averaged when multiple values were available. After the simulations completed the Steinhardt parameters for $l = 4$ and $l = 6$ were calculated for all frames for all three filters ('all', 'like' and 'unlike').

In order to confirm whether chemical ordering occurs beyond thermally induced false classifications, the *global measure* of the number of atoms classified as γ' at 1300 K and 1500 K was calculated and shown in figure 4. The following effects can be seen: (1) for both temperatures the curves show a trend, (2) the concentration of γ'-like atoms increases with time and (3) the concentration of γ'-like atoms is higher at higher temperature. Additionally, all curves start from non-zero amounts of γ' because there will always be a small percentage of atoms that are, by chance, in a γ' environment initially. The exact thermal contribution to false classifications is not known, but we think it is fair to assume that it would not lead to a systematic increase or decrease, as observed in figure 4, and rather contribute in the form of white noise, which is also observed and found to be small in comparison. After observing that there is a global increase in the γ' content it is of interest to understand if any spatially localised structures emerge. Clusters of γ' precipitate can be constructed by linking γ'-like atoms that are less than 4 Å apart to include 2nd nearest neighbours. Decreasing this distance leads to an increase in the frequency of small clusters and a decrease in the frequency of large clusters, whereas an increase in clustering distance produces the opposite effect. However, the changes are small and do not affect the overall shape of the cluster size distribution. In figures 5(a)–(c) the resulting cluster size distributions before and after 100 ns annealing at 1300, 1400 and 1500 K are shown. The cluster sizes appear to follow near-exponential distributions (as seen in figure 5(c)) that shift to higher frequencies with increasing temperature.
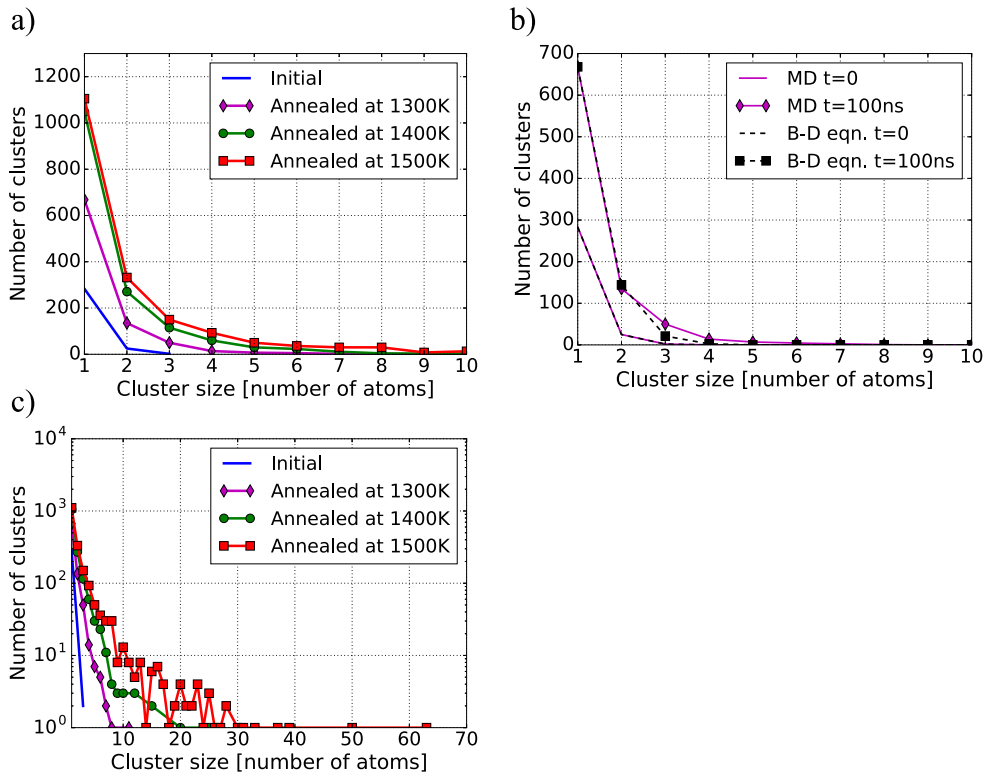
a)

b)

c)

Figure 5. (a) linear and (c) log scale size distributions of γ' clusters before and after 100 ns annealing. Shown are the initial distribution (blue) and the final distributions at 1300 K (magenta), 1400 K (green) and 1500 K (red). (b) Linear cluster-size distribution as shown in figure 5(a) for the single crystal of γ' Ni-20 at% Al annealed at 1300 K using MD together with fitted functions which solve the Becker–Doehring (B–D) equation.

The kinetics of cluster formation and evolution during phase transformations can be described by Becker–Doehring theory [28]. This theory derives the diffusion equation for the temporal change in cluster-size distribution. Therefore the simulated evolution of the cluster-size distributions from MD can be tested by solving the Becker–Doehring equation using the calculated initial boundary conditions (see appendix for details). Tested were a constant diffusion coefficient and a diffusion coefficient that varies with cluster size. It was found that a constant diffusion coefficient best describes the simulations with values of 0.0029 atoms ns^{-1}, 0.0052 atoms ns^{-1} and 0.0057 atoms ns^{-1} at 1300 K, 1400 K and 1500 K respectively. The derived model at 1300 K is compared to the distribution simulated with MD, see figure 5(b). Fits to the data for 1400 and 1500 K look similar and are generally good though the formation of larger clusters tends to be underestimated. This is likely due to the quite simple assumptions of constant and linear functions for the diffusion coefficient.

The advantage of MD simulations is that they enable the atomic level dynamics of ordering to be investigated using *local measures*. For example, to resolve the effect of composition on ordering we can examine the concentration profiles of Al (c_{Al}) and γ' ($c_{\gamma'}$) as a function of time. This can be done as a projection along one axis since the effects are similar along other axes. Figure 6 shows c_{Al} and $c_{\gamma'}$ along the x-axis as a function of time at 1300 and
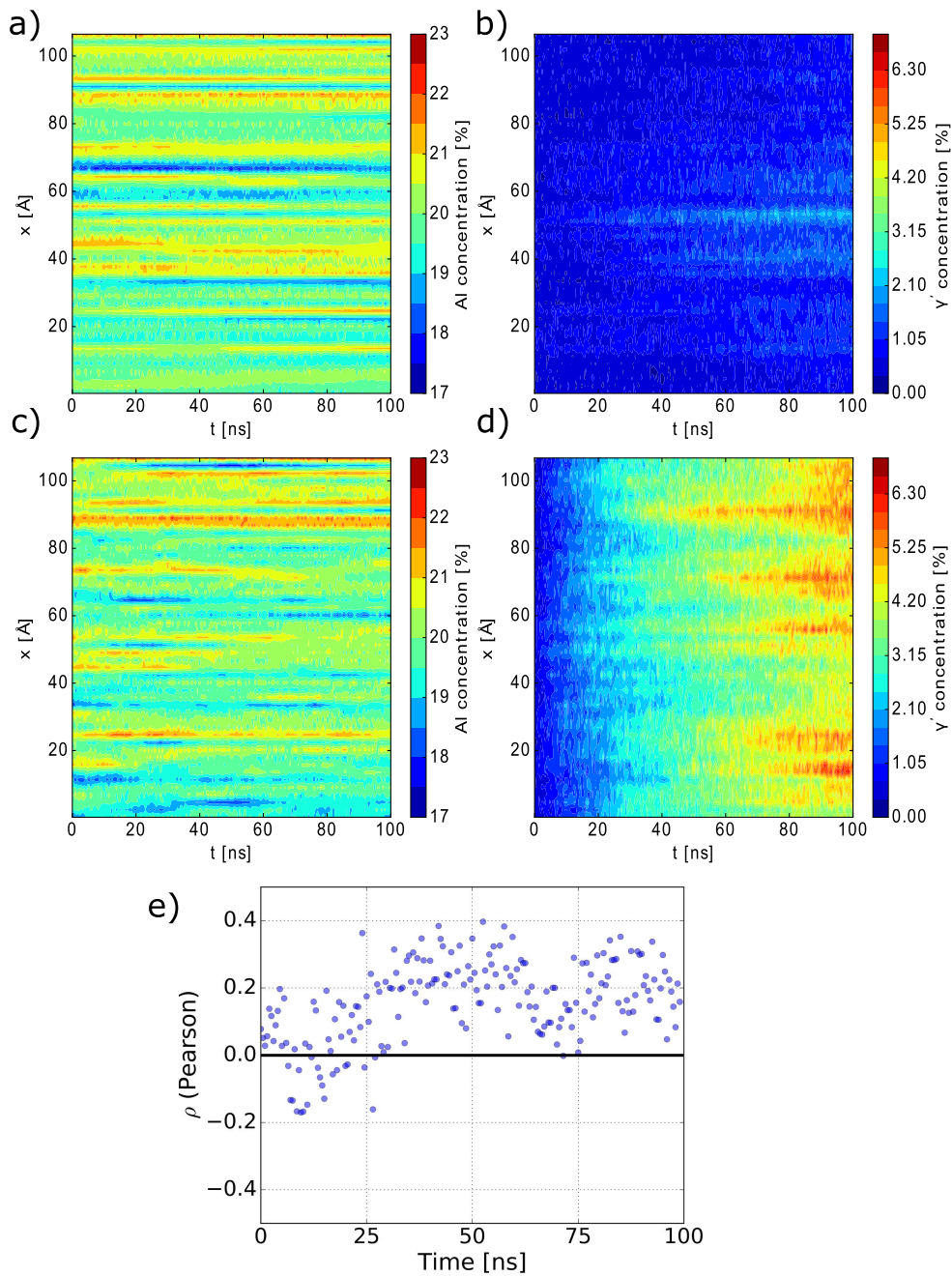
Figure 6. c_{Al} (left (a) and (c)) and $c_{\gamma'}$ (right (b) and (d)) concentration profiles as a function of annealing time along the x-axis at 1300 K (top (a) and (b)) and 1500 K (middle (c) and (d)). Part (e) displays the Pearson correlation coefficient between the initial c_{Al} and the $c_{\gamma'}$ at 1500 K, with $c_{\gamma'}$ varying over the simulation time.
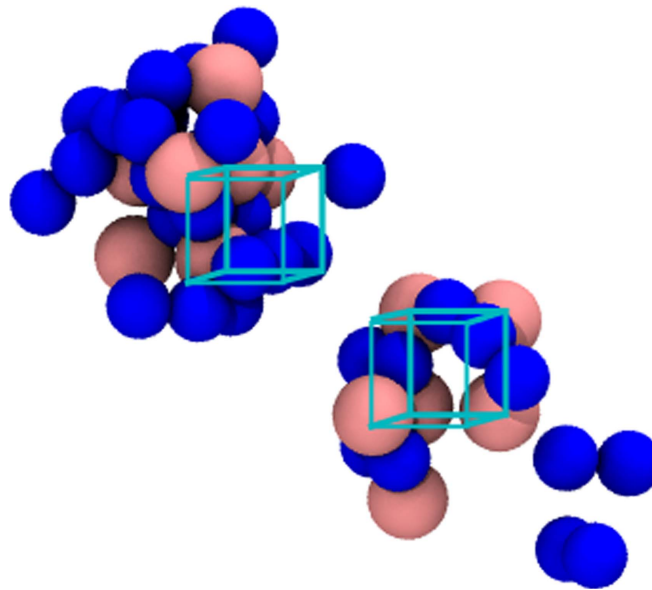
Figure 7. Two clusters of atoms are shown which are classified as γ'-like after annealing with MD for 100 ns at 1500 K. Pink atoms represent Al and blue Ni. The superimposed cubes indicate γ' unit cells which are in an antiphase orientation.

1500 K. The bin width along x is taken to be 2 Å. The c_{Al} profiles (figures 6(a) and (c)) show hill and valley type behaviour along x which is expected since the initial distributions were created randomly. The γ' profiles (figures 6(b) and (d)) show an increase in concentration with time which becomes localised along x during the simulation period. This is particularly evident at 1500 K. Furthermore, the increase in γ' seems to be correlated with above average Al concentrations in the calculated profiles.

In order to study the correlation between c_{Al} and $c_{\gamma'}$ along the dimension x we use the Pearson correlation coefficient, which is defined as $\rho = \text{cov}(Y_1, Y_2)/\sigma_{Y_1}\,\sigma_{Y_2}$ with $\text{cov}(Y_1, Y_2)$ the covariance of the random variables Y_1 and Y_2 and σ the standard deviation. In figure 6(e) ρ is shown as the correlation between the initial c_{Al} and the time dependent $c_{\gamma'}$. It can be seen that ρ fluctuates around a value of 0.16 (at 1500 K) which indicates a small positive correlation. Although not shown, a similar trend can be found when using the time dependent c_{Al} instead of fixing to the initial profile.

The kinetics of the ordering process produces clusters such as shown in figure 7. In this example two clusters of γ'-like atoms are shown. It can be seen that the clusters contain characteristic $\langle 100 \rangle$ rows of Al atoms. The formation of these rows aids the ordering process. In figure 7 one can also see that the two clusters are in antiphase orientation, as indicated by the superimposed cubes connecting Al atoms along the edges of the γ' unit cell.

5. Discussion

In the previous sections we classified atoms according to the geometry and chemistry of their neighbourhood. This led us to find systematic changes on the atomic scale as well in the concentration profiles of γ' and the formation of clusters. In this section we are going to discuss our method and the results in more detail and how they relate to phase transformations

in Ni-based superalloys in particular and the differentiation of phases in atomistic simulations in general.

5.1. Classification

The problem of classifying local atomic configurations was approached in this work using a supervised learning method. Classifications can be performed using a large variety of algorithms such as artificial neural networks, support vector machines and others [29]. In this work we chose to use naïve Bayes. This is because we wanted to have direct access to the conditional PDFs and the option of extending the basis of recognisable phases without having to re-train the classifier. The application of naïve Bayes using a uniform prior for the phases is justified by the task itself, to find individual atoms with γ'-like surroundings. If for example an iterative update of the prior given the neighbourhood probabilities had been applied, a subsequent smoothing would remove details we are interested in, such as small clusters.

The downside of the current approach is that effects from extended defects such as grain boundaries are inseparable from the matrix leading to convoluted signals and therefore convoluted conditional PDFs. Hence with the current approach a separation into individual classes, such as of host and defect, is not straightforward. However, a defect such as an antiphase boundary may still be discernable even though the respective atoms may not be directly classified as such. But this shortcoming can be addressed as will be shown in a forthcoming study. Despite being quite simple the current approach of using conditional PDFs based on distorted ideal forms of the phases of interest with naïve Bayes turns out to be quite robust. We have demonstrated that it is quite capable of accurately distinguishing between γ' and γ. This is promising particularly when one is interested in quickly studying the behaviour of phases in given simulations without having to run additional calculations to generate samples. Furthermore, this method should be directly extendable to other 'features' of local atomic configurations than Steinhardt parameters.

5.2. Clustering and ordering

As we have seen in figures 6(b) and (d), vacancy assisted ordering clearly takes place in the simulated single crystal of Ni-20 at% Al. Using a semi-grand canonical ensemble approach for modelling precipitation in alloys [30] we estimate that the two-phase $\gamma' + \gamma$ region of the phase diagram lies between 17.6 at% Al and 23.5 at% Al, putting the simulated systems in this work right into the middle. This position in the two-phase region provides the driving force for the observed ordering. It may be surprising that the amount of γ' found increases faster to higher levels with increasing temperature even though the driving force should be higher at lower temperatures. This can be explained by the atomic mobilities, which are higher at higher temperatures allowing for faster sampling of configuration space and accelerated ordering and the higher concentration of thermal vacancies. However, longer simulations would be expected to result in higher γ' concentrations in the order 1300 K > 1400 K > 1500 K. At 1400 K after a long time we would expect 40.6 at% γ' to be present using the lever rule. Comparing figures 6(b) and (d) with (a) and (c) shows that ordering is correlated with above average Al concentrations but in a non-deterministic way. The non-deterministic behaviour may stem from the motion of the vacancies themselves and their limited supply and mobility. In figure 6(e) we see that the initial c_{Al} profile is positively correlated with the time dependent $c_{\gamma'}$ profile. The correlation is found to increase from 1300 to 1500 K, where the significance of ordering increases as well. Hence we think that the observed ordering can be interpreted as homogeneous ordering [7] which can precede or

appear concomitantly with Al concentration profile changes and therefore ordering the phase without decomposition. Possible continuations of the observed process could be spinodal decomposition, also called 'conditional spinodal' [8], or nucleation and growth [7, 11, 12]. This leads us to believe that it is the onset of spinodal decomposition that is seen here, which is consistent with experimental observations [13–15]. In order to clearly confirm spinodal decomposition much longer simulations would be required.

In addition to the concentration profiles, we simulated cluster-size distributions and modelled them using the Becker–Doehring theory. We created clusters by linking γ'-like atoms within 4 Å of one another. This way of constructing clusters possibly explains the slightly overestimated cluster frequency compared to that predicted by the theory. This is because the theory assumes an attachment/detachment rate for a given cluster size implying that their shape is not too dissimilar. This, however, is subject to the definition of a cluster. In the present case the clusters can be quite spongy, as found previously in simulations of ordering in a bcc binary alloy [31]. More compact clusters were observed during simulations of liquid-solid transitions [23, 32]. We suggest that by modifying the function used for the diffusion coefficient a better fit may be possible and this would be useful for predicting of the development of the cluster-size distribution for times longer than simulated with MD. Also, a detailed investigation of the vacancy-cluster interaction for the observed ordering would be of interest.

Lastly, we note that the formation of rows of Al atoms seems to be a relevant, possibly rate limiting, step in the precipitation of γ' clusters since the global Al concentration is lower than the equilibrium γ' concentration. We anticipate that continued ordering will depend strongly on the interaction of clusters as they grow and coalesce and that antiphase clusters, as shown in figure 7, will probably slow down the overall growth process.

6. Conclusion

In this study we applied Steinhardt parameters using naïve Bayes in a supervised learning scheme to classify local atomic configurations in annealed single crystals of γ Ni-20 at% Al obtained from large MD simulations. The ordered structure γ'-Ni$_3$Al (L1$_2$) was characterised with Steinhardt parameters sensitive to the chemical surroundings, separating them into 'like' and 'unlike' according to the species neighbouring a given atom.

The method developed allows for accurate differentiation between the γ and γ' phases under significant chemical and thermal distortion. It has enabled us to gain insight into the evolution of local atomic arrangements in disordered Ni-20 at% Al systems using MD up to 1500 K and in doing so observe homogeneous ordering and the formation of γ' clusters in regions where there is above average Al concentration indicating the beginning of spinodal decomposition. The simulated cluster size distributions increase both in time and with temperature. This evolution obeys the Becker–Doehring theory which when parameterised only slightly underestimates the cluster frequencies at large cluster sizes.

Lastly, the method used to generate a classifier based on artificial lattice distortions of ideal versions of the phases of interest was found to be quick and reliable. It should be useful for a range of phases other than γ or γ' which need to take into account the chemical surrounding.

Acknowledgments

Appendix

Becker–Doehring theory and the cluster-size distribution evolution

Here we model the cluster-size distribution on the basis of Becker–Doehring theory as described in [28]. Adopting the same notation as in [28] we describe the cluster-size distribution $Z_{n,t}$ as dependent on time t and cluster size n. The central model equation is then

$$\partial_t Z_{n,t} = \partial_n(q_n O_n \partial_n Z_{n,t}) + \partial_n\left(N_n q_n O_n Z_{n,t} \partial_n\left(\frac{1}{N_n}\right)\right), \tag{A.1}$$

where q_n is the probability of an atom attaching to a cluster of size n, O_n is the cluster surface, N_n is the equilibrium cluster-size distribution and $\partial_n \equiv \partial/\partial n$. Neglecting the second term one obtains the well-known heat equation $\partial_t Z_{n,t} = \partial_n(D_n \partial_n Z_{n,t})$ which is equivalent to

$$\partial_t Z_{n,t} = c\partial_n Z_n + d\partial_n^2 Z_n, \tag{A.2}$$

where $D_n = q_n O_n$ assuming time independence and $c = \partial_n D_n$ and $d = D_n$.

Equation (A.2) can be solved numerically given the observed cluster-size distributions from MD using implicit Euler or 'implicit diffusion', [33]. The problem is discretised with $n \in \{0, N\}$ and hence $N + 1$ grid points. Δn is 1 in all the cases considered. In the indices of Z and D the quantity Δn is replaced by 1 for sake of clarity. The initial condition $Z_{n,0}$ is taken from the cluster-size distribution from the MD simulations after identifying and clustering γ'-like atom configurations.

The boundary condition $Z_{0,t}$ set knowing that the number of smallest clusters will be subject to change over time $Z_{1,t} = g(t)$, decreasing or increasing depending on the atomic interactions. Depending on whether D_n is constant or not with respect to n one can choose two discrete forms of equation (A.2). When $D_n = D = $ constant we obtain the discrete form of equation (A.2) as

$$\frac{Z_{n,t}}{\Delta t} = \frac{1}{\Delta t}Z_{n,t+\Delta t} + d(Z_{n-1,t+\Delta t} - 2Z_{n,t+\Delta t} + Z_{n+1,t+\Delta t}) \tag{A.3}$$

with the boundary condition using $Z_{0,t+\Delta t} = g(t + \Delta t) = g_1$ for $n = 1$

$$\frac{1}{\Delta t}Z_{1,t+\Delta t} + \frac{d}{(\Delta n)^2}(g_1 - 2Z_{1,t+\Delta t} + Z_{2,t+\Delta}) = \frac{Z_{1,t}}{\Delta t}$$

and $Z_{N,t+\Delta t} = 0$ for $n = N-1$

$$\frac{1}{\Delta t}Z_{t+\Delta t}(n_{N-1}) + \frac{d}{(\Delta n)^2}(Z_{t+\Delta t}(n_{N-2}) - 2Z_{t+\Delta t}(n_{N-1})) = \frac{Z_t(n_{N-1})}{\Delta t}.$$

If on the other hand D_n varies with n, equation (A.2) can be represented as 'implicit diffusion', [33]

$$\frac{1}{\Delta t}(Z_{n,t+1} - Z_{n,t}) = \frac{c}{\Delta n}(Z_{n+1,t+1} - Z_{n,t+1}) + \frac{d}{\Delta n^2}(Z_{n-1,t+1} - 2Z_{n,t+1} + Z_{n+1,t+1}),$$

$$\text{(A.4)}$$

with boundary condition for $n = 1$

$$\frac{1}{\Delta t}Z_{1,t} + \frac{dg_1}{\Delta n^2} = \left(\frac{c}{\Delta n} + \frac{2d}{\Delta n^2} + \frac{1}{\Delta t}\right)Z_{1,t+1} - \left(\frac{c}{\Delta n} + \frac{d}{\Delta n^2}\right)Z_{2,t+1},$$

and for

$$n = N - 1$$

$$\frac{1}{\Delta t}Z_{N-1,t} = -\frac{d}{\Delta n^2}Z_{N-2,t+1} + \left(\frac{c}{\Delta n} + \frac{2d}{\Delta n^2} + \frac{1}{\Delta t}\right)Z_{N-1,t+1}.$$

In order to describe the cluster-size distribution (equation (A.2)), the two finite difference equations for D and D_n have to be solved using equations (A.3) and (A.4) respectively.

References

[1] Eckhardt W and Heinecke A 2013 SuperMUC boosts the largest molecular dynamics simulation by 4X in number of particles *InSide* **11** http://inside.hlrs.de/_old/htm/Edition_01_13/article_06.html

[2] Caron P and Khan T 1983 Improvement of creep strength in a nickel-base single-crystal superalloy by heat treatment *Mater. Sci. Eng.* **61** 173–84

[3] Pollock T M and Argon A S 1992 Overview no. 95 creep resistance of CMSX-3 nickel base superalloy single crystals *Acta Metall. mater.* **40** 1–30

[4] Ricks R A, Porter A J and Ecob R C 1983 The growth of γ' precipitates in nickel-base superalloys *Acta Met.* **31** 43–53

[5] Doi M, Miyazaki T and Wakatsuki T 1984 The effect of elastic interaction energy on the morphology of γ' precipitates in nickel-based alloys **67** 247–53

[6] Soffa W A and Laughlin D E 1982 Recent experimental studies of continuous transformations in alloys: an overview *Proceedings of an International Conference on Solid-Solid Phase Transformations, AIME* pp 159–83

[7] Soffa W A and Laughlin D E 1989 Decomposition and ordering processes involving thermodynamically first-order order → disorder transformations *Acta Metall.* **37** 3019–28

[8] Allen S M and Cahn J W 1976 Mechanisms of *phase* transformations within the miscibility gap of Fe-rich Fe–Al alloys *Acta Metall.* **24** 425–37

[9] Cahn J W 1961 On spinodal decomposition *Acta Metall.* **9** 795–801

[10] Hillert M 1961 A solid-solution model for inhomogeneous systems *Acta Metall.* **9** 525–35

[11] Khachaturyan A, Lindsey T and Morris J 1988 Theoretical investigation of the precipitation of δ' in Al–Li *Metall. Trans.* A **19A** 249–58

[12] Chen L and Khachaturyan A 1991 Computer simulation of structural transformations during precipitation of an ordered intermetallic *phase Acta Metall. Mater.* **39** 2533–51

[13] Babu S S, Miller M K, Vitek J M and David S A 2001 Characterization of the microstructure evolution in a nickel base superalloy during continuous cooling conditions *Acta Mater.* **49** 4149–60

[14] Viswanathan G B, Banerjee R, Singh A, Nag S, Tiley J and Fraser H L 2011 Precipitation of ordered phases in metallic solid solutions: a synergistic clustering and ordering process *Scr. Mater.* **65** 485–8

[15] Tan X P, Mangelinck D, Perrin-Pellegrino C, Rougier L, Gandin C-A, Jacot A, Ponsen D and Jaquet V 2014 Spinodal decomposition mechanism of γ' precipitation in a single crystal ni-based superalloy *Metall. Mater. Trans.* A **45A** 4725–30

[16] Mishin Y 2004 Atomistic modeling of the γ and γ'-phases of the Ni–Al system *Acta Mater.* **52** 1451–67

[17] Daw M S and Baskes M I 1983 Semiempirical, quantum mechanical calculation of Hydrogen embrittlement in metals *Phys. Rev. Lett.* **50** 1285–8

[18] Steinhardt P J, Nelson D R and Ronchetti M 1983 Bond-orientational order in liquids and glasses *Phys. Rev.* B **28** 784–805
[19] Stukowski A 2012 Structure identification methods for atomistic simulations of crystalline materials *Model. Simul. Mater. Sci. Eng.* **20** 45021
[20] Ackland G J and Jones A P 2006 Applications of local crystal structure measures in experiment and simulation *Phys. Rev.* B **73** 54104
[21] Lechner W and Dellago C 2008 Accurate determination of crystal structures based on averaged local bond order parameters *J. Chem. Phys.* **129** 114707
[22] Archer A, Foxhall H R, Allan N L, Gunn D S D, Harding J H, Todorov I T, Travis K P and Purton J A 2014 Order parameter and connectivity topology analysis of crystalline ceramics for nuclear waste immobilization *J. Phys.: Condens. Matter* **26** 485011
[23] Auer S and Frenkel D 2004 Numerical prediction of absolute crystallization rates in hard-sphere colloids *J. Chem. Phys.* **120** 3015–29
[24] Quigley D and Rodger P M 2008 Metadynamics simulations of ice nucleation and growth *J. Chem. Phys.* **128** 154518
[25] Gades H and Mitus A C 1991 Local structures in computer-generated liquids and glasses: Classification of three-dimensional patterns *Physica* A **176** 297–324
[26] Plimpton S 1995 Fast parallel algorithms for short-range molecular dynamics *J. Comput. Phys.* **117** 1–19
[27] Rogal J, Divinski S V, Finnis M W, Glensk A, Neugebauer J, Perepezko J H, Schuwalow S, Sluiter M H F and Sundman B 2014 Perspectives on point defect thermodynamics *Phys. Status Solidi* **251** 97–129
[28] Christian J W 1975 *Theory of Transformation in Metals and Alloys* (Oxford: Pergamon)
[29] Kotsiantis S B, Zaharakis I D and Pintelas P E 2006 Machine learning: a review of classification and combining techniques *Artif. Intell. Rev.* **26** 159–90
[30] Sadigh B, Erhart P, Stukowski A, Caro A, Martinez E and Zepeda-Ruiz L 2012 Scalable parallel Monte Carlo algorithm for atomistic simulations of precipitation in alloys *Phys. Rev.* B **85** 184203
[31] Athènes M, Bellon P, Martin G and Haider F 1996 A Monte-Carlo study of B2 ordering and precipitation via vacancy mechanism in b.c.c. lattices *Acta Mater.* **44** 4739–48
[32] Rein ten W P, Ruiz-Montero M J and Frenkel D 1996 Numerical calculation of the rate of crystal nucleation in a Lennard–Jones system at moderate undercooling *J. Chem. Phys.* **104** 9932
[33] Strang G 2012 *Computational Science and Engineering* (Wellesley MA: Wellesley-Cambridge Press)