# UC Davis
## UC Davis Electronic Theses and Dissertations

**Title**

Molecular Simulations of Advanced Alloys Under Extreme Conditions

**Permalink**

https://escholarship.org/uc/item/8xd689kb

**Author**

Stimac, Jared Cole

**Publication Date**

2022

Peer reviewed|Thesis/dissertation

Molecular Simulations of Advanced Alloys Under Extreme Conditions

By

JARED STIMAC
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemical Engineering

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

---

Jeremy Mason, Chair

---

Ambar Kulkarni

---

Nir Goldman

Committee in Charge

2022

i

# Abstract

Computational models can support materials development by identifying the key factors that affect material properties and by guiding the search for optimal chemical and processing conditions. However, the success of this venture assumes the ability to accurately model material structures and the relationship of those structures to material properties. State-of-the-art tools in high-performance computing and machine learning are continually improving the performance of these models, thereby furthering their integration into the process of developing better materials including advanced alloys.

This dissertation includes three projects that use atomic simulations to support the development of advanced alloys for applications in extreme conditions. First is the construction of a novel framework for machine learning potentials (MLPs). MLPs could dramatically accelerate simulations of atomic systems while providing the accuracy of electronic structure techniques through the use of supervised regression algorithms. MLPs do still have a higher computational expense than empirical potentials though, both during their construction and for every evaluation of the potential energy. With the purpose of reducing these costs and alleviating the necessity for enormous training data sets, our framework for producing MLPs combines an efficient implementation of a sparse Gaussian process algorithm with a novel set of descriptors for atomic environments. Second, molecular dynamics is used to investigate energy storage and heat evolution during high-strain-rate deformation of the refractory metal Ta. This is encapsulated in a quantity known as the Taylor–Quinney coefficient, which is critical to models of material failure in conditions where direct experimental measurement of the temperature is infeasible. Other than developing a phenomenological model for the energy stored in the material, this chapter identifies that a significant amount of the energy is stored in the form of point defects. Third, molecular dynamics is used to study the defect structures that evolve in irradiated materials in the low temperature and high radiant flux regime. The algorithm used involves the successive insertion of Frenkel pairs and relaxation of the simulation cell, and allowed the study of Fe, equiatomic CrCoNi, and a fictitious metal with identical bulk properties to the CrCoNi up to the equivalent of 2.0 displacements per atom (dpa). Several areas requiring further research are identified, including the mechanisms by which Shockley partials develop in FCC metals at low dpa and robust ways to measure point defect concentrations in heavily-damaged FCC materials.

# Acknowledgements

Throughout the duration of completing this Ph.D., I have found myself utterly grateful to be working with my advisor, Prof. Jeremy Mason, who gave tremendous and consistent effort supporting me and who was always looking out for what was in my best interest. I will not forget the degree to which he was paramount during this experience.

I would like to acknowledge the other members of the Mason research group who also supported my efforts through advice, listening through some of my fumbling presentations, and being sources of inspiration–not just as researchers–but also as genuinely good people. Thank you Ozan Ericok, Erdem Eren, Sakura Kawano, Chad Serrao, Conner Winkeljohn, Ethan Suwandi, Nickie O'Shea, Kavan Joshi, and Enver Kapan.

Another mentor who I am happy to recognize is Dr. Vasily Bulatov, who has provided me with great opportunities to grow as a researcher in this field and who has provided a wealth of valuable expertise and guidance. I would also like to thank Dr. Nicolas Bertin for being incredibly helpful in my collaborations with Lawrence Livermore National Laboratory.

I want to recognize Dr. Laina Geary, my research mentor while attending the University of Nevada, Reno. Thank you, Dr. Geary, for giving me opportunities to get involved in academic research and for the immense support in pursuing this degree.

Finally, I would like to thank my friends and family for the continual support and patience throughout this degree; in particular, my parents and my brother Jacob. I could not have gotten through this without you.

# Contents

# Chapter 1

# Introduction

*"All models are wrong; some are useful"*

-George E.P. Box

The process of developing novel high-performance metal alloys is remarkably challenging because of the large design space and continual difficulties associated with competing properties. Depending on the application, a suitable alloy providing an acceptable trade-off of competing properties, manufacturability, and cost could be lost in the combinatorial space of potential compositions. The alloy development space, in a manner similar to a plethora of other materials problems, poses fundamental scientific questions that are laborious and expensive to address experimentally. Many such questions are well suited to be addressed using computational models and simulations. John von Neumann and his colleagues at Los Alamos National Laboratory provided some of the earliest demonstrations of the utility of computational material simulations following the end of World War II with the inception of Monte Carlo (MC) techniques [1]. von Neumann's MC simulations directly led to well-recognized scientific breakthroughs in nuclear physics relating to neutron diffusion [2]. Since then, computational models and simulations have become increasingly integral to materials development processes, providing insight into the connections between microstructural features and their effect on material properties and thereby guiding the search for optimal chemical and processing conditions. Indeed, this is the basis for the Integrated Computational Materials Engineering (ICME) paradigm, in which computational models are integrated into and used to accelerate multiple stages of the materials development pipeline [3].

An example application that could benefit from the ICME methodology is the development of structural components for next-generation nuclear reactors that are intended to provide a consistent and low-carbon baseline power supply during the upcoming century [4]. All of the concept Generation IV nuclear reactors

would require materials to operate for much longer periods at higher temperatures and higher radiation doses than the current light water reactors, most of which were constructed in the 1970s and 80s [5, 6]. Whereas current light water reactors have structural components that are exposed to temperatures around 300 °C and up to 80 displacements per atom (dpa, the standard measure of radiation damage) over a 40 year operation life, several of the Generation IV reactor concepts would be exposed to temperatures of up to 1000 °C and 200 dpa over the course of 60 years [7]. Advancing the properties of metal alloys to be able to withstand these fierce conditions is one of the most prominent challenges facing the realization of such nuclear reactors.

Computational models and simulations provide a means of investigating prospective alloys' strength, radiation resistance, and corrosion susceptibility for this and other applications in extreme environments. Metallic strengthening mechanisms in an alloy of interest can be studied through deformation simulations. When a metal is plastically deformed, most of the work that is imparted to the material is dissipated as heat. The remaining fraction of the work is stored in the material in the form of defects, commonly referred to as the "stored energy of cold work" or "latent energy". If the deformation occurs at a high strain rate, the evolved heat is unable to dissipate to the surroundings, resulting in a potentially significant increase in the metal's temperature. Since the mechanical properties of metals are temperature dependent, it is important to accurately describe both the heat evolution and dissipation during continuum-scale simulations.

The ratio of evolved heat to plastic work was first experimentally investigated by Taylor and Quinney nearly century ago [8] using the torsion of copper rods to high strains. This ratio is commonly referred to as the Taylor–Quinney Coefficient (TQC), and according to their measurements is initially close to 0.9 before asymtotically approaching 1.0 with increasing strain. The TQC is ultimately a function of the microstructural evolution of the material [9] and can depend on the alloy species, strain, strain rate [10], deformation orientation [11], and infrequently the loading mode [12]. The total stored cold work depends on the configuration and distribution of evolved defects that can include dislocations, point defects, grain boundaries, stacking faults, and twins. Although molecular dynamics (MD) simulations have not frequently been used to study the stored energy of cold work[13], the atomic-scale resolution can help to resolve the relevant defects and provide complete information about the partitioning of energy into thermal and potential contributions, reducing any systematic errors involved in evaluating the TQC.

Another property that must be considered specifically for prospective structural alloys in nuclear reactors is, of course, radiation resistance. Passing high-energy particles (typically neutrons) can transfer kinetic energy to an atom of the metal lattice, initiating a collision cascade that displaces the surrounding atoms, produces a thermal spike, and generates a high concentration of point defects [14]. As the density of point defects increases with further radiation damage, they can initiate processes that sharply degrade the

2

material properties. Specifically, point defect concentration and migration are heavily implicated in radiation hardening, radiation-induced segregation and precipitation, void swelling, and radiation-induced creep [7]. MD simulations have been used in recent decades to simulate and study the cascade collision events initiated by impacts from high-energy radiation particles [15, 16, 14]. While the same simulations could in principle be used to directly investigate the degrading effects listed above, in practice it is difficult to simulate a sufficient number of overlapping collision cascades to reach the required point defect concentrations.

Multi-principal element alloys (MPEAs) are a promising class of metal alloys for this and related applications. Prior to the initial identification of MPEAs (or high-entropy alloys) in the early 2000s, alloy development mainly involved adding small amounts of one or more secondary elements to a single primary element that defined the properties of the underlying lattice. The development of MPEAs has dramatically expanded the domain of alloys to consider though, with much of the recent research being into alloys where three or more elements are in near-equal concentrations and form a single solid solution (a solid-phase homogeneous mixture). The resulting alloys could be ideal for structural applications in extreme environments, with reported properties including an unusual combination of strength and ductility [17, 18], corrosion resistance [19], and resistance to radiation damage [20]. For example, the CrMnFeCoNi Cantor alloy has demonstrated a remarkable combination of yield strength and fracture toughness that is already competitive with all other state-of-the-art alloys developed to date [18]. Furthering our understanding of the aspects of MPEAs that are responsible for their exceptional properties is essential to being able to tune their composition and structure and achieve further improvements.

The modeling of MPEAs, along with almost every other area of science and engineering, is benefiting from the recent and widespread application of machine learning approaches. The application of data science techniques to address materials science problems in particular is already considered to be a distinct subfield of "materials informatics" [21]. Well-established data science methods for pattern recognition, regression, dimension reduction, and classification have become ubiquitous here partly because materials science investigations continually produce large quantities of high-dimensional data. A major objective of material informatics is to expedite the materials design process using *inverse strategies*, or methods that use machine learning models to identify a small subset of candidate material morphologies and compositions possessing a desirable combination of macroscopic properties for a given application. However, as pointed out by Jain et al. [22], developing successful inverse strategies will require efficient feedback loops that use current simulations and experimental approaches as *forward strategies*. Specifically, forward simulations and experiments provide the experience and theory that is necessary to construct inverse models, and they allow the inverse models to be validated and directed. In addition to supporting the search for highly effective inverse models, machine learning has already shown success in other materials modeling problems including bringing

simulations at disparate length and time scales [23].

On the higher-resolution end of the range of computational techniques are *ab initio* electronic structure calculations that rely on a knowledge of fundamental electron interactions, with reasonable approximations for those parts of the interactions that are not well understood. The computational resources that are available today limit such methods to small numbers of atoms, on the order of $10^3$. On the other end of this range are continuum models that can capture macro-scale phenomena to varying degrees of accuracy. These approaches are driven by parameterized constitutive relationships, making accurate parameterization critical to the extraction of meaningful information regarding material properties. There is great interest in the computational materials modeling community in developing methods that bridge these time and length scales, ideally using more accurate and more expensive models to parameterize or train less expensive ones. This could ultimately enable the *ab initio* techniques to even inform the behaviors that emerge on the macro-scale.

One of the most successful examples of using machine learning algorithms for scale-bridging has been the machine learning interatomic potentials for atomistic simulations. Referred to as machine learning potentials (MLPs), these functions interpolate *ab initio* calculations for the potential energy of an atomic configuration using regression algorithms. The construction of an MLPs initially requires a computationally-convenient representation of an atom's local environment. This is usually an array of scalars known as *descriptors* that is invariant to the symmetries of the physical system and from which the machine learning algorithm can extract important features related to the potential energy. Next, a regression function is constructed on the space of descriptors to map a local atomic environment to a potential energy. The success of machine learning in this area is a consequence of machine learning effectively being the nonparametric regression of functions in high-dimensional spaces, and particularly complex functions that do not necessarily have analytical definitions. There are several MLPs already reported in the literature that display DFT-level accuracy with orders of magnitude lower computational cost [24, 25, 26].

Recent work by Rosenbrock et al. [25], provides an apt example of the capabilities of MLPs for metallic systems, while also the highlighting the differences between two state-of-the-art MLP frameworks, namely, Gaussian Approximation Potential (GAP) and the Moment Tensor Potential (MTP). Both MLP models were able to fit a potential for a Ag-Pd system and accurately predict DFT-calculated forces, energies, virial stresses, and phonon dispersion curves for a wide array of compositions. The superior computational efficiency of MTP made it the only suitable choice for calculating a temperature-composition phase diagram for the alloy with acceptable resolution. On the other hand, GAP demonstrated the greater transferability (ability to make predictions for configurations far away from the training examples) in that it gave more reasonable pathways between configurations. Understanding which MLP frameworks are best suited for a

specific application and developing new MLP frameworks is a currently active area of research.

The subsequent chapters of this dissertation describe three projects involving MD simulations that are motivated by the development of alloys under extreme conditions and, to varying degrees, have aspects involved with bringing model length and time scales. Chapter 2 describes a novel framework for constructing machine learning interatomic potentials (MLP). This framework, like others in the literature, bridges electronic structure calculations and MD simulations; the interatomic potential energies that determine the atomic trajectories are informed by, and trained on, quantum mechanical calculations. Chapter 3 describes a MD study of high-strain rate compression of single crystal Ta. This study reports TQC values up to a total strain of 1.0 and also develops two models for the stored energy, one that could be utilized in continuum-level deformation simulations and a second that reveals the surprisingly high energy contribution of point defects. Finally, Ch. 4 describes atomistic simulations of radiation damage in a CrCoNi MPEA in the low temperature and high radiant flux regime. Several areas requiring further research are identified, including the mechanisms by which Shockley partials develop in FCC metals at low dpa and robust ways to measure point defect concentrations in heavily-damaged FCC materials.

# Chapter 2

# Machine learning potential

## 2.1 Introduction

Atomistic simulations, encompassing both Monte Carlo (MC) techniques and molecular dynamics (MD), are well-established tools in a variety of scientific fields, from biochemistry to materials science. All of these techniques rely on the Born–Oppenheimer approximation, i.e., they assume that the electronic degrees of freedom relax on a much shorter time scale than that of atomic motion, allowing the atoms to be treated as classical point particles. MD simulations then apply the classical equations of motion, finding the atomic trajectories by integrating Newton's second law. While these are not the true trajectories because of the inherently chaotic dynamics of many body systems [27], they approximate the behavior of physical systems well enough in practice. MC simulations instead sample the configuration space (the space such that a configuration of all atoms corresponds to a single point) in a way that does not involve time integration but is nevertheless consistent with the canonical Gibbs measure.

Repeated evaluations of the interatomic potential energies and the forces acting on all the atoms are a critical part of all atomistic simulations. The potential energy can be viewed as a hypersurface over the configuration space, known as the potential energy surface (PES), on which the electrons are always assumed to be in a ground-state as a consequence of the Born-Oppenheimer approximation. Methods to evaluate the interatomic potential energy, commonly referred to in the literature as "potentials" or "force fields", generally reduce the dimension of the problem by fitting the PES as a sum of functions defined only on the local atomic environments of individual atoms. These functions are specific both to the atomic species and to the material, and the accuracy of the underlying assumptions and the fitting procedure itself fundamentally limits the accuracy of atomistic simulations and the apparent properties that emerge.

There are ongoing efforts to develop empirical interatomic potentials using only elementary functions that can reasonably describe a variety of physical systems based on the types of atomic bonding present. In metallic systems the embedded atom method (EAM) [28] and modified embedded atom method (MEAM) [29] have been widely successful. This approach involves dividing the potential energy $E_i$ of an atom $i$ into a sum of pair-wise interactions and a term that accounts for the energy cost of embedding that atom into the local electron cloud [30]:

$$E_i = \frac{1}{2} \sum_j V(r_{ij}) + F(\bar{\rho}_i) \tag{2.1}$$

where $V$ is a function of the pair-wise atomic distance $r_{ij}$ from atom $i$ to a neighboring atom $j$, and F is the embedding function over the local electron density $\bar{\rho}_i$. The local electron density is in turn defined as [30]:

$$\bar{\rho}_i = \sum_j \rho_p(r_{ij}) \tag{2.2}$$

where $\rho_p(r_{ij})$ is the contribution to the local electron density from neighboring atom $j$. The refinement introduced by the MEAM potential includes an angular term in the electron density function.

Despite the general success of these methods, the systematic errors introduced by empirical interatomic potentials are generally unknown, and there are many questions that call for more accurate potential energy evaluations. The simulated behavior of dislocations (which govern the deformation behavior of crystalline metal alloys) can be sensitive to even minor perturbations to the potential. For example, Xu et al. used three different potentials to study screw dislocation behavior in W and found that only one reproduced the *ab initio* results [31]. This sensitivity would seem to encourage the exclusive use of high-fidelity potentials, but the desire for accuracy needs to be balanced with other trade offs. In particular, if the potential is too expensive, limitations on simulation size and duration could preclude studies of the intended phenomenon. For example, dislocation core radii are on the scale of 1 nm [32], so simulations of the stress field for a single dislocation must exceed this, lest non-physical effects of the dislocation interacting with its periodic image will occur.

The questionable accuracy of empirical potentials and the requirement for efficient evaluation of the potential energy has motivated a variety of attempts over the last two decades to use machine learning to develop higher accuracy potentials. These methods are commonly referred to as "machine learning potentials" (MLPs), and are motivated by the idea that the number of physical assumptions made about the PES can be reduced by performing a small number of electronic structure calculations and interpolating the rest of the surface. Setting aside the training data, MLPs consist of two primary components: the descriptors of a local atomic environment and a machine learning algorithm. The descriptors are sets of real numbers

that are intended to represent the configuration of neighboring atoms within a sphere around a central atom. The boundary of this sphere marks a cutoff radius, with atoms outside the cutoff radius assumed to have a negligible effect on the potential energy of the central atom. These descriptors (or "fingerprints") should be invariant to the various symmetries inherent to the PES, including translational and rotational symmetries and permutations of the labels of atoms of the same species. Respecting these symmetries is essential for the fitting of the PES to be tractable for the machine learning algorithm. Once the descriptors have been evaluated, a machine learning algorithm, trained on electronic structure data usually provided by density function theory (DFT) calculations, predicts the central atom's potential energy.

MLPs initially appeared in the literature around 1995 when Skinner and Broughton proposed that neural networks could be used to interpolate the PES [33]; however, widespread exploration of this idea did not begin until Behler and Parrinello introduced Neural Network Potentials (NNPs) in 2007 [26]. A central assumption made by NNPs and subsequent MLPS is that the total potential energy $E$ of the system can be partitioned into a sum of atomic contributions:

$$E = \sum_i E_i. \tag{2.3}$$

NNPs predict an atom's contribution $E_i$ to the potential energy by using feedforward neural networks (FNN) trained to interpret the descriptors proposed by Behler and Parrinello, referred to as "symmetry functions". These symmetry functions are of two varieties. The radial symmetry functions $G_i^1$ characterize the pairwise distances between an atom $i$ and the neighboring atoms $j$ by means of a sum over Gaussians [26]:

$$G_i^1(\boldsymbol{r}_{ij}|\eta, r_s) = \sum_{i \neq j} e^{-\eta(r_{ij} - r_s)^2} f_c(r_{ij}) \tag{2.4}$$

where $r_s$ and $\eta$ are parameters that control the mean and variance of the Gaussian. The radial character of the environment is resolved using a set of values $\eta$ to construct a set of $G_i^1$. The $f_c$ is a cuttoff function that ensures that the potential energy contribution smoothly decays to zero as a neighboring atom is displaced beyond the cutoff radius $r_c$.

The symmetry functions $G_i^2$ instead characterize the central angles between pairs of neighboring atoms $i$ and $j$, and is a "three-body" descriptor in the sense that it requires knowledge of the set of all atom triplets. They are defined as:

$$G_i^2(\boldsymbol{r}_{ij}, \boldsymbol{r}_{ik}|\zeta, \lambda) = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos\theta_{ijk})^\zeta \ e^{-\zeta\left(r_{ij}^2 + r_{ik}^2 + r_{jk}^2\right)} f_c(r_{ij}) f_c(r_{ik}) f_c(r_{jk}) \tag{2.5}$$

8

where $\theta_{ijk} = (\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{ik})/(r_{ij}r_{ik})$ are the central angles between pairs of neighboring atoms $j$ and $k$. A set of $\zeta$ values are used to resolve the space of central angles in the same way that $\eta$ is used with the $G^1$. The parameter $\lambda$ controls the domain of the cosine term.

Training a NNP involves determining the proper sets of symmetry functions (usually by a process of guess and check) and the values of weights and biases for the FNN. FNNs contain a predefined number of "hidden-layers" that transform the input features using non-linear activation functions in a way that allows for highly-flexible function fitting. The $E_i$ returned by the FNN for each of the atoms are summed as specified by Eq. 2.3 to find the total potential energy of the system, and this is compared with the DFT-derived total potential energies for each example in the training set. Training the NNP consists of reducing the discrepancy using standard machine learning procedures to optimize the weights and biases of the FNN, with backpropagation being the most widely employed [34].

Several other descriptor formulations and MLP frameworks have been developed following the early successes of NNPs. The Gaussian Approximation Potential (GAP) developed by Barktók et al. [35] uses a set of descriptors referred to as the smooth overlap of atomic positions (SOAP) descriptors. These begin by constructing a Gaussian-smeared atomic neighbor density function $\rho$ for a local atomic environment:

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{i=N} \exp(-\alpha|\boldsymbol{r} - \boldsymbol{r}_i|^2) \tag{2.6}$$

a function of an arbitrary displacement vector $\boldsymbol{r} \in \mathbb{R}^3$ from the central atom defining the local atomic environment to any point within the environment. This function can be described as a sum of Gaussians centered on each of the neighboring atoms, defined by their displacement vectors from the central atom $\boldsymbol{r}_i$. The parameter $\alpha$ controls the smoothness of the density and is important for the numerical conditioning and $N$ is the number of atoms of the same atomic type. The function $\rho(\boldsymbol{r})$ describes an atomic environment in a way that is invariant to permutations of the atomic labels due to the commutativity of addition. This function is also invariant to translations, being defined relative to the position $\boldsymbol{r}_i$ of the central atom, but is not rotationally invariant. The atomic density functions $\rho(\boldsymbol{r})$ and $\rho'(\boldsymbol{r})$ of two different local atomic environments can be compared in a rotationally invariant way by taking the inner product of the two averaged over all possible relative rotations $\tilde{R}$. The result is a similarity kernel:

$$k(\rho, \rho') = \int \left[ \int \rho(\boldsymbol{r})\rho'(\tilde{R}\boldsymbol{r})d\boldsymbol{r} \right]^n d\tilde{R} \tag{2.7}$$

where $n \geq 2$ is an adjustable parameter. This integration is not done explicitly; instead, the atomic neighbor density function is expanded as a linear combination of products of spherical harmonics and orthonormal

radial basis functions with coefficients $c_{nlm}$. The properties of the spherical harmonics under rotations allows the integration in Eq. 2.7 to be simplified to:

$$K(\rho, \rho') = \frac{k(\rho, \rho')}{\sqrt{k(\rho, \rho)k(\rho', \rho')}} = \sum_{nn'l} p_{nn'l}p'_{nn'l} \tag{2.8}$$

where the $p_{nn'l}$ are the coefficients of the rotationally invariant power spectrum:

$$p_{nn'l} = \sum_{m=-l}^{m=l} c^*_{nlm}c_{n'lm}. \tag{2.9}$$

The SOAP descriptors have been extended to systems with multiple atomic species. This is done by defining a Gaussian-smeared atomic neighbor density function and a corresponding basis function expansion for each atomic species [36].

At the heart of GAP is a sparse version of Gaussian process regression on the space of all local atomic environments, and rather than explicitly constructing a set of descriptors on this space, some versions of GAP use the similarity kernel $K$ as a rough way to measure the proximity of two local atomic environments. This type of algorithm will be discussed in greater detail in the later sections of this chapter.

Since the literature on MLPs is vast and rapidly growing, only a few of the other prominent MLP frameworks will be briefly mentioned here. The Spectral Neighbor Analysis Potential (SNAP) uses linear regression over the bispectrum coefficients of the atomic neighbor density function $\rho$ to predict the atomic energies $E_i$ [37, 38], and has the advantage of a simple training procedure. The Moment Tensor Potential (MTP) also uses linear regression, but over an overcomplete set of invariant polynomials of the relative atomic coordinates [39, 40, 41]. The second generation of the Chebyshev Interaction Model for Efficient Simulation (ChIMES) developed by Lindsey et al. in 2017 [42] fits many-body atomic interactions with Chebyshev polynomial expansions via a force matching procedure [43, 44], and also benefits from a simple linear fitting procedure. ChIMES is capable of fitting an arbitrary order of $n$-body correlations, which has been carried out up to four-body interactions [45, 46]. Further in-depth discussion on the history and current state of MLPs can be found in the 2016 perspective on MLPs from Jörge Behler [47] and in other reviews Refs. [48, 49, 50].

The most common approach to constructing descriptors, used by roughly half of the aforementioned frameworks, is to represent local atomic environments using only two- and three-body information. However, this was proven by Pozdnyakov et al. [51] to be insufficient to distinguish all possible atomic environments up to the relevant symmetry operations. The proof is by counterexample where they identify a situation in which two different environments have the same two- and three-body information but could have distinct

potential energies. Related situations apply to four-body information, and it is unclear whether there is any higher-order information (short of the $n$-body information for an environment of $n$ atoms) that could be included to ensure systematic errors are avoided. The existence of distinct environments with degenerate representations is therefore a widespread possibility and could limit the effectiveness of MLPs which use such descriptors. There are comparatively few MLPs that do not suffer from this issue (e.g., [39, 52]).

Despite such concerns about the descriptors, MLPs have been widely successful. They have been used in a plethora of atomistic simulations for a variety of material systems and are well-recognized as being capable of achieving DFT-level accuracy for far smaller computational cost. That said, they are still typically 2-3 orders of magnitude more costly than empirical potentials [53]. This brings us to the directions of future MLP development, three of which are commonly discussed. The first is to improve the MLP's accuracy of fitting the PES, and more specifically the ability to generalize from the electronic structure data in the training set. Many MLP frameworks are already capable of fitting a PES with an accuracy greater than that to which DFT calculations are trusted (~0.5 eV/atom [54]), so in principle more fundamental electronic structure methods such as coupled-cluster calculations could be used instead [53]. This accuracy does not necessarily extend to local atomic configurations outside of the training set though, about which few uncertainty guarantees are currently available. The second direction is to reduce computational cost. As mentioned, MLPs are still far more computationally intensive than empirical potentials, limiting the size and scale of possible simulations. Improvements in this area could involve further development of either the descriptors or the machine learning algorithm. The third direction is with regard to the efficiency and simplicity of training. Ideally, MLPs would be fast enough to construct that they could even be used as surrogate functions for small pieces of the PES, e.g., around a particular reaction pathway.

More formally, the MLP is regarded as a function $\bar{\phi}$ from local atomic environments to atomic potential energies that is constructed by the composition of two other functions. The first function is a map $f : \mathcal{R}_{\mathcal{N}} \to Z$ from local atomic environments to sets of descriptors. The space defining the local atomic environment is $\mathcal{R}_{\mathcal{N}} = \prod_N B_{r_c}$, or the $N$-fold direct product of 3-balls surrounding the central atom. A point is this space defines the set of displacement vectors $\boldsymbol{r}_j = \boldsymbol{x}_j - \boldsymbol{x}_i$ from the central atom $i$ to the $N$ neighboring atoms $j$ within a specified cutoff radius $r_c$, where any atoms beyond the cutoff radius are assumed to contribute negligibly to the potential. The space of sets of descriptors $Z \subset \mathbb{R}^m$ is usually constructed as a region in $m$-dimensional Euclidean space, where $m$ is an adjustable number of descriptors. The second function is a map $g : Z \to \mathbb{R}$ from sets of descriptors to atomic potential energies $E_i \in \mathbb{R}$ and is usually constructed by means of a machine learning algorithm. The relationship of these three functions is indicated schematically in Fig. 2.1. The utility of this formal perspective of the structure of MLPs is that it clarifies that the choice of a set of descriptors and of a regression algorithm can be independent, and that the functions realizing
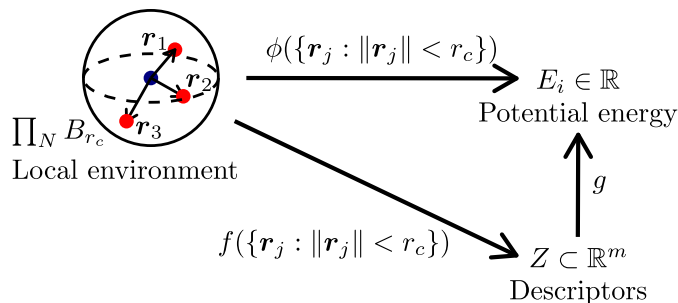
Figure 2.1: General structure of machine learning potentials.

these choices should be subject to different requirements.

The objective of this chapter is to develop our framework for constructing MLPs. Many elements of our approach resemble those used by GAP, and this is not an accident. There are numerical and theoretical reasons why it is advantageous to construct the descriptors with the expansion coefficients of the atomic density function, and the use of Gaussian processes regression in principle allows for robust estimates of the uncertainty in the predicted atomic potential energies. That said, GAP is well known to to be highly computationally expensive, and our belief is that this related to certain choices that were made during the development of GAP that had unforeseen consequences. Section 2.2 derives a novel set of atomic descriptors, referred to as the $\xi$ descriptors, that provably contain the same information as the bispectrum cofficients used in more recent implements of GAP but that are 1-2 orders of magnitude faster to evaluate depending on the implementation. Section 2.5 reviews the theory of Gaussian processes, and of the specific sparsification method that our framework uses. This method is now standard in the literature, and is reported to give more accurate approximations to the true posterior distribution than the one used by GAP. Section 2.6 benchmarks the resulting framework by fitting the forces and energies of a DFT data set for Cu and compares the results with contemporary alternatives in the literature.

## 2.2 Atomic descriptors

Let $X$ be a point cloud in $\mathbb{R}^3$, and $Y_i$ be the intersection of $X$ with $B_{r_c}(\boldsymbol{x}_i)$, a ball of radius $r_c$ centered on $\boldsymbol{x}_i \in X$. $Y_i$ is referred to as the local atomic configuration centered at $\boldsymbol{x}_i$. This implicitly defines the set $\mathcal{R}$ of all local atomic configurations as

$$\mathcal{R} = \bigcup_{N=0}^{\infty} \mathcal{R}_N$$

$$\mathcal{R}_N = \{(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) : \|\boldsymbol{r}_i\| < r_c \text{ for all } i\}$$

where $\boldsymbol{r}_j = \boldsymbol{x}_j - \boldsymbol{x}_i$. Let there additionally be a function $\phi : \mathcal{R} \to \mathbb{R}$ called the interatomic potential that maps a local atomic configuration to a real quantity known as the interatomic potential energy. This can be used to define dynamics on the points in $X$ in such a way that their motion resembles that of the atoms in a physical system. The process of numerically integrating the dynamical relations to solve for the point trajectories is known as a molecular dynamics simulation.

The function $\phi$ obeys a number of physical symmetries. First, the isotropy of space suggests that $\phi$ should be invariant to the simultaneous action of the orthogonal group $\mathrm{O}(3)$ on all the $\boldsymbol{r}_i$. Second, the indistinguishability of atoms of the same element suggests that $\phi$ should be invariant to permutations of the indices $1 \leq i \leq N$ of the $\boldsymbol{r}_i$, i.e., to the action of the symmetric group $\mathrm{S}_N$.

Suppose that there is a known set of local atomic configurations $Y_i$ and associated potential energies $\phi(Y_i)$, and that one would like to approximate $\phi$ by constructing a regression function $\bar{\phi}$ from this data. $\bar{\phi}$ can be required to obey the same physical symmetries as $\phi$ by constructing the regression function as the composition

$$\bar{\phi} = g \circ f$$

where $f$ identifies points if and only if they are in the same orbits of the orthogonal group and symmetric group, and the regression analysis is performed by $g$.

The main concern of this section is the map $f : \mathcal{R} \to Z$ from the space of local atomic configurations to a space $Z$ with coordinates known as descriptors. This map would ideally have several additional properties. First, the second derivatives of $f$ with respect to the $\boldsymbol{r}_i$ should be continuous everywhere in $B_{r_c}$, including on the boundary. This is required for the dynamics on $X$ to be well-defined as atoms enter and leave the local atomic environment. Second, $f$ should minimize the maximum distortion throughout the space as indicated by, e.g., the singular value spectrum of the Jacobian matrix. Regression algorithms usually constrain the set of possible functions by assuming that some property of the function is roughly uniform on the space, and are more accurate when this is so. Third, one would prefer to be able to always distinguish points on different orbits of the orthogonal and symmetric groups with as few coordinates as possible. This is because the computational complexity of regression algorithms often increases as $\mathcal{O}(d^2)$ or $\mathcal{O}(d^3)$, where $d$ is the number of coordinates.

The motivation for the approach developed here is the observation that the atomic density function

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i) \tag{2.10}$$

allows reconstruction of a local atomic environment up to the action of $\mathrm{S}_N$, but is not invariant to the

13

action of O(3). Given a complete orthonormal basis for $L^2$ (square-integrable) functions on $B_{r_c}$, $\rho(\boldsymbol{r})$ could alternatively be represented by an infinite sequence of expansion coefficients. Certain functions of the expansion coefficients are invariant to the action of O(3), and set of such functions could perhaps be used to construct a suitable map $f$.

The recent literature [51, 55] suggests that the main difficulty with this approach is verifying that the resulting descriptors contain enough information to allow a local atomic environment to be reconstructed up to the action of $S_N$ and O(3). If it does not, $\bar{\phi}$ could be constrained to assign certain physically-distinct local atomic environments the same potential energies, introducing systematic errors that cannot be reduced by any choice of $g$. That said, there are indications in the literature that four-body information (indicating the positions of three neighboring atoms relative to the central atom) is often sufficient in practice [51, 56]. Our intention with this section is therefore to simply construct a set of descriptors that include all four-body information, have continuous derivatives up to second order, are invariant to the required isometries, and are as efficient as possible to evaluate.

## 2.3 Complete orthogonal systems

Following the motivation in Sec. 2.2, this section effectively constructs canonical sets of functions of the expansion coefficients of $\rho(\boldsymbol{r})$ that are invariant to the action of O(3), though by a different procedure than initially proposed. Specifically, $\rho(\boldsymbol{r})$ is not expanded over a complete orthonormal basis for $L^2$ functions on $B_{r_c}$, but over a complete orthonormal basis for O(3)-invariant $L^2$ functions on $B_{r_c}$. The resulting expansion coefficients then have all the required invariance properties without further manipulation, and the suitability of the map $f$ is reduced to the question of the number of descriptors that is sufficient and the efficiency of their evaluation.

### 2.3.1 Functions on $B_{r_c}$

While there are many complete orthogonal bases for $L^2$ functions on $B_{r_c}$, the one considered here is based on eigenfunctions of the Helmholtz equation

$$\nabla^2 \Psi(r, \Omega) = -k^2 \Psi(r, \Omega)$$

where $\Omega = \{\theta, \phi\}$ and $\theta$ and $\phi$ are the polar and azimuthal angles in spherical coordinates. The Laplacian can be written as

$$\nabla^2 = \nabla_r^2 + \frac{1}{r^2} \nabla_\Omega^2$$

where

$$\nabla_r^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) \qquad\qquad \nabla_\Omega^2 = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial}{\partial\phi}.$$

Separating variables with $\Psi(r,\theta,\phi) = R(r)\Omega(\theta,\phi)$ gives

$$\nabla_r^2 R(r) + \left[k^2 - \frac{l(l+1)}{r^2}\right]R(r) = 0 \tag{2.11}$$

and

$$\nabla_\Omega^2\Omega(\theta,\phi) + l(l+1)\Omega(\theta,\phi) = 0. \tag{2.12}$$

The solutions to the second equation are the spherical harmonics, defined as

$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{im\phi} \tag{2.13}$$

with integer $l \geq 0$ and $-l \leq m \leq l$, where the Condon–Shortley phase is included in the definition of the associated Legendre polynomials $P_l^m(\cos\theta)$. It will be useful below that $\hat{I}Y_l^m(\Omega) = (-1)^l Y_l^m(\Omega)$ where $\hat{I}$ is the inversion operator. The solutions to the first equation are the spherical Bessel functions of the first kind $j_l(kr)$. Further details of this derivation can be found in the standard references (e.g., Sec. 9.4 of Ref. [57]).

The functions $\Psi_{lm}(r,\Omega;k) = j_l(kr)Y_l^m(\Omega)$ form a complete orthogonal system for $L^2$ functions on $\mathbb{R}^3$. The orthogonality condition

$$\int_0^{2\pi}\int_0^\pi\int_0^\infty \Psi_{l'm'}^*(r,\Omega;k')\Psi_{lm}(r,\Omega;k)r^2\sin\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi = \delta(k'-k)\delta_{l'l}\delta_{m'm}\frac{\pi}{2k^2}$$

is a consequence of those for the $j_l(kr)$ and the $Y_l^m(\Omega)$. Completeness follows by observing that the eigenfunctions of the Helmholtz equation in Cartesian coordinates are the complex exponentials $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$, that these constitute a complete basis for $L^2$ functions on $\mathbb{R}^3$, and that they are related to the $\Psi_{lm}(r,\Omega;k)$ by the plane wave expansion [58]

$$\exp(i\boldsymbol{k}\cdot\boldsymbol{r}) = 4\pi\sum_{l=0}^\infty\sum_{m=-l}^l i^l Y_l^{m*}(\hat{\boldsymbol{k}})j_l(kr)Y_l^m(\hat{\boldsymbol{r}})$$

where, e.g., $r = \|\boldsymbol{r}\|$ and $\hat{\boldsymbol{r}} = \boldsymbol{r}/r$. That is, the $\Psi_{lm}(r,\Omega;k)$ are related to a complete orthogonal system by a change of basis corresponding to a change from Cartesian to spherical coordinates, and are therefore

complete as well.

This basis is restricted to one for $L^2$ functions $f(r, \Omega)$ on $B_{r_c}$ subject to the boundary condition $f(r_c, \Omega) = 0$ by requiring that $k = u_{nl}/r_c$, where $u_{nl}$ is the $(n+1)$st nonzero root of $j_l(x)$. The further boundary conditions $\partial f/\partial r|_{r_c} = 0$ and $\partial^2 f/\partial r^2|_{r_c} = 0$ are enforced by replacing $j_l(u_{nl}r/r_c)$ with the functions $g_{nl}(r)$ defined in Appendix B.1. Letting $\Psi_{lm}^n(r, \Omega) = g_{nl}(r)Y_l^m(\Omega)$, the orthogonality condition becomes

$$\int_0^{2\pi} \int_0^\pi \int_0^{r_c} \Psi_{l'm'}^{n'*}(r, \Omega)\Psi_{lm}^n(r, \Omega)r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi = \delta_{n'n}\delta_{l'l}\delta_{m'm}. \tag{2.14}$$

That the $\Psi_{lm}^n(r, \Omega)$ are a complete basis for $L^2$ functions subject to the specified conditions is a consequence of the Hilbert projection theorem [59]. Specifically, the basis elements $\Psi_{lm}(r, \Omega; k)$ are projected into the appropriate subspace of the Hilbert space of $L^2$ functions and orthogonalized by a Gram–Schmidt procedure, giving the $\Psi_{lm}^n(r, \Omega)$ as a basis for this subspace.

At this point, it is useful to observe that the $\Psi_{lm}^n(r, \Omega)$ are eigenfunctions of the angular momentum operators $\mathbf{L}^2$ and $L_z$ [60]. That is, they behave as irreducible tensors of rank $l$ under the action of SO(3), or

$$\Psi_{lm}^n(r, \Omega') = \sum_{m'} \Psi_{lm'}^n(r, \Omega)D_{m'm}^l(\alpha, \beta, \gamma)$$

where $D_{m'm}^l(\alpha, \beta, \gamma)$ is a Wigner $D$-function obeying the conventions described in Sec. 1.4.1 of Ref. [60], and $\Omega$ and $\Omega'$ are the spherical coordinates of a single unit vector in the original and rotated coordinate systems. Notably, since $D_{00}^0(\alpha, \beta, \gamma)$ is a constant, the functions $\Psi_{00}^n(r, \Omega)$ are invariant to the action of SO(3). Moreover, they are a complete basis for SO(3)-invariant $L^2$ functions on $B_{r_c}$ with the prescribed boundary conditions; recall that

$$f(r, \Omega') = \exp(-i\omega\hat{\boldsymbol{n}} \cdot \mathbf{L})f(r, \Omega)$$

describes the effect of a rotation of the coordinate system by an angle $\omega$ about an axis $\hat{\boldsymbol{n}}$ on the function $f(r, \Omega)$, as indicated in Sec. 4.5.1 of Ref. [60]. If $f(r, \Omega)$ is SO(3)-invariant, then this implies

$$[1 - \exp(-i\omega\hat{\boldsymbol{n}} \cdot \mathbf{L})]f(r, \Omega) = 0.$$

In particular, this needs to hold for infinitesimal rotations since they are the generators of the rotation group. Taking the first-order Taylor expansion of the exponential about $\omega = 0$ gives the condition $\hat{\boldsymbol{n}} \cdot \mathbf{L}f(r, \Omega) = 0$ for all $\hat{\boldsymbol{n}}$. This is only possible if every component of the angular momentum of $f(r, \Omega)$ vanishes. That is, the only SO(3)-invariant functions are precisely the linear combinations of the eigenfunctions of $\mathbf{L}^2$ with $l = 0$. Since $\hat{I}\Psi_{00}^n(r, \Omega) = \Psi_{00}^n(r, \Omega)$, these also form a basis for O(3)-invariant functions.

16

That said, the $\Psi_{00}^n(r, \Omega)$ are not particularly useful for representing the atomic density function of a local atomic environment given in Eq. 2.10, since expanding $\rho(r, \Omega)$ over this basis only retains information about the distances of nearby atoms from the central one. Let $A_\Psi$ be the function space spanned by the $\Psi_{00}^n(r, \Omega)$, and $\hat{P}_\Psi$ the operator that projects a function onto $A_\Psi$. Since $\hat{P}_\Psi$ is a linear operator, $\hat{P}_\Psi \rho(\boldsymbol{r})$ is the same as $\sum_i \hat{P}_\Psi \delta(\boldsymbol{r} - \boldsymbol{r}_i)$. Moreover, since $\hat{P}_\Psi \delta(\boldsymbol{r} - \boldsymbol{r}_i)$ is invariant to the action of O(3), it is equivalent to $\hat{P}_\Psi \delta(\boldsymbol{r} - r_i \hat{\boldsymbol{z}})$. That is, $\hat{P}_\Psi \rho(\boldsymbol{r})$ is equivalent to $\hat{P}_\Psi \sum_i \delta(\boldsymbol{r} - r_i \hat{\boldsymbol{z}})$, or as claimed above, only retains information about the histogram of the $r_i$. This is clearly not sufficient to reconstruct a local atomic environment.

## 2.3.2 Functions on $B_{r_c}^2$

Instead, one could consider pairs of atoms in the local atomic environment, and expand this over a complete orthonormal basis for $L^2$ functions on $B_{r_c}^2$. Such a basis is provided by the irreducible tensor product of the $\Psi_{l_0 m_0}^{n_0}(\boldsymbol{r})$ and $\Psi_{l_1 m_1}^{n_1}(\boldsymbol{r}')$ (Sec. 3.1.7 of Ref. [60]), or

$$\Phi_{l_0 l_1 l m}^{n_0 n_1} = \sum_{m_0 m_1} C_{l_0 m_0 l_1 m_1}^{lm} \Psi_{l_0 m_0}^{n_0}(\boldsymbol{r}) \Psi_{l_1 m_1}^{n_1}(\boldsymbol{r}')$$

where $C_{l_0 m_0 l_1 m_1}^{lm}$ is a Clebsch–Gordan coefficient. Many of the properties of the $\Phi_{l_0 l_1 l m}^{n_0 n_1}$ are the same as those of the bipolar spherical harmonics as defined in Sec. 5.16.1 of Ref. [60]. In particular, the $\Phi_{l_0 l_1 l m}^{n_0 n_1}$ transform as an irreducible tensor of rank $l$ under the action of SO(3). When $l = 0$, the properties of the Clebsh–Gordan coefficients and the $Y_l^m(\Omega)$ allow the definition above to be simplified to

$$\Phi_{l_0}^{n_0 n_1} = \frac{(-1)^{l_0}}{\sqrt{2l_0 + 1}} \sum_{m_0} \Psi_{l_0 m_0}^{n_0 *}(\boldsymbol{r}) \Psi_{l_0 m_0}^{n_1}(\boldsymbol{r}').$$

By the same reasoning as in Sec. 2.3.1, the $\Phi_{l_0}^{n_0 n_1}$ are a complete orthonormal basis for SO(3)-invariant $L^2$ functions on $B_{r_c}^2$ that satisfy the prescribed boundary conditions. Since $\hat{I} \Phi_{l_0}^{n_0 n_1} = \Phi_{l_0}^{n_0 n_1}$, these also form a basis for O(3)-invariant functions. As such, they can be used to expand the atomic pair density function

$$\rho(\boldsymbol{r}, \boldsymbol{r}') = \sum_{i=1}^N \sum_{j=1}^N \delta(\boldsymbol{r} - \boldsymbol{r}_i) \delta(\boldsymbol{r}' - \boldsymbol{r}_j). \tag{2.15}$$

Let $A_\Phi$ be the function space spanned by the $\Phi_{l_0}^{n_0 n_1}$, and $\hat{P}_\Phi$ the operator that projects a function onto $A_\Phi$. Then $\hat{P}_\Phi \rho(\boldsymbol{r}, \boldsymbol{r}')$ is equivalent to the projection of a modified atomic pair density function where, for each of the $\delta(\boldsymbol{r} - \boldsymbol{r}_i) \delta(\boldsymbol{r}' - \boldsymbol{r}_j)$, the coordinate system is rotated to orient $\boldsymbol{r}_i$ along the $z$-axis and $\boldsymbol{r}_j$ in the $xz$-plane with positive $x$ coordinate. That is, $\hat{P}_\Phi \rho(\boldsymbol{r}, \boldsymbol{r}')$ only retains information about the histogram of triples

$\{r_i, r_j, \hat{\boldsymbol{r}}_i \cdot \hat{\boldsymbol{r}}_j\}$[1]. This is known to not be sufficient for $f : \mathcal{R} \to Z$ to distinguish atomic environments up to isometry, with counterexamples available in the literature [51]. Note that restricting the second summation in Eq. 2.15 to $j > i$ makes $\hat{P}_\Phi \rho(\boldsymbol{r}, \boldsymbol{r}')$ variant to the action of $\mathrm{S}_N$, and using $j \neq i$ instead makes $\hat{P}_\Phi \rho(\boldsymbol{r}, \boldsymbol{r}')$ unable to describe environments with $N < 2$ atoms.

### 2.3.3 Functions on $B_{r_c}^3$

One could perhaps be discouraged at this point, without a visible proof in the literature that $n$-point histograms are sufficient to reconstruct all local atomic configurations up to the action of $\mathrm{S}_N$ and O(3) for any $n < N$. As stated in Sec. 2.2 though, four-body information (or 3-point histograms) is already sufficient for MLPs to reproduce electronic structure calculations of the atomic potential energies with an accuracy greater than that to which DFT calculations are trusted [53]. The question of whether there is a set of descriptors that is complete in the sense of all local atomic configurations being reconstructable is therefore set aside as the subject of a future study, and this section proceeds with the derivation of an efficiently computable set of four-body descriptors.

The construction begins by defining the irreducible tensor products of $\Phi_{l_1 l_2 l' m'}^{n_1 n_2}$ and $\Psi_{l_0 m_0}^{n_0}(\boldsymbol{r})$ as the functions

$$\chi_{l_0 l_1 l_2 l' l m}^{n_0 n_1 n_2} = \sum_{m_0 m'} \sum_{m_1 m_2} C_{l_0 m_0 l' m'}^{lm} C_{l_1 m_1 l_2 m_2}^{l' m'} \Psi_{l_0 m_0}^{n_0}(\boldsymbol{r}) \Psi_{l_1 m_1}^{n_1}(\boldsymbol{r}') \Psi_{l_2 m_2}^{n_2}(\boldsymbol{r}'').$$

When $l = 0$, the properties of the Clebsh–Gordan coefficients allow this to be simplified to

$$\chi_{l_0 l_1 l_2}^{n_0 n_1 n_2} = \frac{(-1)^{l_0}}{\sqrt{2l_0 + 1}} \sum_{m_0 m_1 m_2} C_{l_1 m_1 l_2 m_2}^{l_0 m_0} \Psi_{l_0 m_0}^{n_0 *}(\boldsymbol{r}) \Psi_{l_1 m_1}^{n_1}(\boldsymbol{r}') \Psi_{l_2 m_2}^{n_2}(\boldsymbol{r}'') \tag{2.16}$$

$$= \sum_{m_0 m_1 m_2} \begin{pmatrix} l_2 & l_1 & l_0 \\ m_2 & m_1 & m_0 \end{pmatrix} \Psi_{l_0 m_0}^{n_0}(\boldsymbol{r}) \Psi_{l_1 m_1}^{n_1}(\boldsymbol{r}') \Psi_{l_2 m_2}^{n_2}(\boldsymbol{r}'')$$

where the second version introduces a $3jm$ symbol [60] and makes the symmetry of the $\chi_{l_0 l_1 l_2}^{n_0 n_1 n_2}$ more apparent. Notice that the Clebsh–Gordan coefficient and therefore the $\chi_{l_0 l_1 l_2}^{n_0 n_1 n_2}$ vanish unless $l_0$, $l_1$ and $l_2$ satisfy the usual triangle conditions. These functions form a complete orthonormal basis for SO(3)-invariant $L^2$ functions on $B_{r_c}^3$ that satisfy the prescribed boundary conditions. Orthonormality is a consequence of the corresponding property of the tripolar spherical harmonics in Sec. 5.16.2 of Ref. [60] and the $g_{nl}(r)$ in B.1. Since $\hat{I}\chi_{l_0 l_1 l_2}^{n_0 n_1 n_2} = (-1)^{l_0 + l_1 + l_2} \chi_{l_0 l_1 l_2}^{n_0 n_1 n_2}$, the subset of the $\chi_{l_0 l_1 l_2}^{n_0 n_1 n_2}$ where $L = l_0 + l_1 + l_2$ is even is the corresponding basis for O(3)-invariant functions.

---

[1] This can also be shown by applying the spherical harmonic addition theorem to the definition of the $\Phi_{l_0}^{n_0 n_1}$.

Let $\hat{P}_\chi$ and $\hat{P}'_\chi$ respectively be the projection operators onto the function space $A_\chi$ spanned by the $\chi^{n_0 n_1 n_2}_{l_0 l_1 l_2}$ and the subspace $A'_\chi$ where $L$ is even. The functions

$$\rho(\boldsymbol{r}, \boldsymbol{r}', \boldsymbol{r}'') = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i)\delta(\boldsymbol{r}' - \boldsymbol{r}_j)\delta(\boldsymbol{r}'' - \boldsymbol{r}_k)$$

are S$_N$ invariant, and $\hat{P}_\chi \rho(\boldsymbol{r}, \boldsymbol{r}', \boldsymbol{r}'')$ is SO(3) invariant. By the same reasoning as in Sec. 2.3.2, this projection is equivalent to that of a corresponding function where, for each of the $\delta(\boldsymbol{r} - \boldsymbol{r}_i)\delta(\boldsymbol{r}' - \boldsymbol{r}_j)\delta(\boldsymbol{r}'' - \boldsymbol{r}_k)$, the coordinate system is rotated to orient $\boldsymbol{r}_i$ along the $z$-axis and $\boldsymbol{r}_j$ in the $xz$-plane with positive $x$ coordinate. That is, $\hat{P}_\chi \rho(\boldsymbol{r}, \boldsymbol{r}', \boldsymbol{r}'')$ is equivalent to a histogram of oriented triples of points, and $\hat{P}_{\chi'} \rho(\boldsymbol{r}, \boldsymbol{r}', \boldsymbol{r}'')$ to a histogram of unoriented triples of points.. The coefficients

$$\xi^{n_0 n_1 n_2}_{l_0 l_1 l_2} = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \chi^{n_0 n_1 n_2*}_{l_0 l_1 l_2}(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k) \tag{2.17}$$

of the projection of $\rho(\boldsymbol{r}, \boldsymbol{r}', \boldsymbol{r}'')$ onto a linear combination of the $\chi^{n_0 n_1 n_2}_{l_0 l_1 l_2}$ where $L$ is even therefore constitute an infinite set of four-body descriptors with the required isometry invariances.

## 2.4  Evaluating the $\xi^{n_0 n_1 n_2}_{l_0 l_1 l_2}$

The $\xi^{n_0 n_1 n_2}_{l_0 l_1 l_2}$ will need to be evaluated frequently during any MD simulation employing an MLP constructed to use these descriptors. The purpose of this section is to provide a more computationally efficient alternative to directly evaluating Eq. 2.17, and to construct an explicit sequence from which the required number of descriptors can be drawn.

Initially inserting Eq. 2.16 into Eq. 2.17, exchanging the order of summations, and converting the sum of products into a product of sums gives

$$\xi^{n_0 n_1 n_2}_{l_0 l_1 l_2} = \frac{(-1)^{l_0}}{\sqrt{2l_0 + 1}} \sum_{m_0 m_1 m_2} C^{l_0 m_0}_{l_1 m_1 l_2 m_2} \left[ \sum_i \Psi^{n_0}_{l_0 m_0}(\boldsymbol{r}_i) \right] \left[ \sum_j \Psi^{n_1*}_{l_1 m_1}(\boldsymbol{r}_j) \right] \left[ \sum_k \Psi^{n_2*}_{l_2 m_2}(\boldsymbol{r}_k) \right]. \tag{2.18}$$

The properties of the Clebsch–Gordan coefficients indicate that

$$C^{l_0 m_0}_{l_1 m_1 l_2 m_2} = \frac{1}{2}\left[ C^{l_0 m_0}_{l_1 m_1 l_2 m_2} + (-1)^{l_0 + l_1 + l_2} C^{l_0 - m_0}_{l_1 - m_1 l_2 - m_2} \right] = \frac{1}{2}\left( C^{l_0 m_0}_{l_1 m_1 l_2 m_2} + C^{l_0 - m_0}_{l_1 - m_1 l_2 - m_2} \right)$$

where the second equality is a consequence of $L$ being required to be even. Inserting this into Eq. 2.18, relabeling the summation indices, and using the property that $\Psi^n_{l-m}(\boldsymbol{r}) = (-1)^m \Psi^{n*}_{lm}(\boldsymbol{r})$ inherited from the

19

spherical harmonics gives

$$\xi_{l_0 l_1 l_2}^{n_0 n_1 n_2} = \frac{(-1)^{l_0}}{\sqrt{2l_0+1}} \sum_{m_0 m_1 m_2} C_{l_1 m_1 l_2 m_2}^{l_0 m_0} \frac{1}{2} \left\{ \left[ \sum_i \Psi_{l_0 m_0}^{n_0}(\boldsymbol{r}_i) \right] \left[ \sum_j \Psi_{l_1 m_1}^{n_1 *}(\boldsymbol{r}_j) \right] \left[ \sum_k \Psi_{l_2 m_2}^{n_2 *}(\boldsymbol{r}_k) \right] \right.$$
$$\left. + \left[ \sum_i \Psi_{l_0 m_0}^{n_0 *}(\boldsymbol{r}_i) \right] \left[ \sum_j \Psi_{l_1 m_1}^{n_1}(\boldsymbol{r}_j) \right] \left[ \sum_k \Psi_{l_2 m_2}^{n_2}(\boldsymbol{r}_k) \right] \right\} \quad (2.19)$$

where the property that the Clebsch–Gordan coefficients are nonzero only if $m_0 + m_1 + m_2 = 0$ was used to simplify the signs. Using the definition of the $\Psi_{lm}^n(r,\Omega)$ in Sec. 2.3.1, a single one of the summations over $\Psi_{lm}^n(\boldsymbol{r}_i)$ is explicitly written as

$$\sum_i \Psi_{lm}^n(\boldsymbol{r}_i) = \frac{1}{\sqrt{2\pi}} \sum_i g_{nl}(r_i) \tilde{P}_l^m(\cos\theta_i) \cos(m\phi_i) + \frac{i}{\sqrt{2\pi}} \sum_i g_{nl}(r_i) \tilde{P}_l^m(\cos\theta_i) \sin(m\phi_i)$$
$$= c_{lm}^n + i s_{lm}^n \quad (2.20)$$

where $i$ is the imaginary unit, the dependence of $c_{lm}^n$ and $s_{lm}^n$ on the relative positions of the neighboring atoms in the local atomic environment is left implicit, and

$$\tilde{P}_l^m(\cos\theta) = \sqrt{\frac{2l+1}{2} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta)$$

is a normalized associated Legendre function. Finally, inserting Eq. 2.20 into Eq. 2.19, expanding the terms inside the braces, and simplifying leads to

$$\xi_{l_0 l_1 l_2}^{n_0 n_1 n_2} = \frac{(-1)^{l_0}}{\sqrt{2l_0+1}} \sum_{m_0 m_1 m_2} C_{l_1 m_1 l_2 m_2}^{l_0 m_0} \left[ c_{l_0 m_0}^{n_0} \left( c_{l_1 m_1}^{n_1} c_{l_2 m_2}^{n_2} - s_{l_1 m_1}^{n_1} s_{l_2 m_2}^{n_2} \right) \right.$$
$$\left. + s_{l_0 m_0}^{n_0} \left( c_{l_1 m_1}^{n_1} s_{l_2 m_2}^{n_2} + s_{l_1 m_1}^{n_1} c_{l_2 m_2}^{n_2} \right) \right]. \quad (2.21)$$

There are two properties of Eq. 2.21 that are particularly significant. The first is that the $\xi_{l_0 l_1 l_2}^{n_0 n_1 n_2}$ for even $L$ are real quantities, and can be calculated using only real quantities on floating point units without resorting to complex numbers. The second is that once the quantities $c_{lm}^n$ and $s_{lm}^n$ defined in Eq. 2.20 are calculated for a given local atomic environment for all of the required sets of indices $\{n, l, m\}$ in $\mathcal{O}(N)$ operations, calculating the $\xi_{l_0 l_1 l_2}^{n_0 n_1 n_2}$ requires an additional number of operations that is constant in $N$. This is clearly preferable to the $\mathcal{O}(N^3)$ operations that the direct evaluation of Eq. 2.18 requires.

The ordering of the $\xi^{n_0 n_1 n_2}_{l_0 l_1 l_2}$ is set by lexicographically ordering the tuples of integers $(q, r, s, t, u, v)$ where

$$0 \leq q \qquad\qquad 0 \leq r \leq q \qquad\qquad 0 \leq s \leq r$$

$$0 \leq t \leq s \qquad\qquad 0 \leq u \leq q - r \qquad\qquad 0 \leq v \leq q - r - u.$$

These are converted to tuples of indices $(n_0, n_1, n_2, l_0, l_1, l_2)$ using the following equations:

$$n_0 = r - s \qquad\qquad n_1 = s - t \qquad\qquad n_2 = t$$

$$l_0 = q - r - u \qquad\qquad l_1 = q - r - v \qquad\qquad l_2 = u + v$$

The rationale for this ordering is that the oscillation frequency of the $\chi^{n_0 n_1 n_2}_{l_0 l_1 l_2}$ roughly increases with $q$, that it is practical to separately consider the sums $\sum_i n_i = r$ and $\sum_i l_i = 2(q - r)$ of the radial and angular indices due to the constraints on the $l_i$ (including that $L$ be even), and that reverse lexicographic ordering is otherwise convenient.

There are computational reasons why it is desirable to use the smallest possible number of descriptors sufficient for the intended purpose. This encourages the removal of any $\xi^{n_0 n_1 n_2}_{l_0 l_1 l_2}$ from the resulting sequence that are clearly dependent on the preceding terms. Consider that the invariance of the $3jm$ symbols in the second version of Eq. 2.16 to all permutations of the columns when $L$ is even implies that the $\chi^{n_0 n_1 n_2}_{l_0 l_1 l_2}$ are invariant to all permutations of the sets $\{n_i, l_i, r_i\}$. This means that the $\xi^{n_0 n_1 n_2}_{l_0 l_1 l_2}$ defined in Eq. 2.17 are invariant to all permutations of the sets $\{n_i, l_i\}$, or explicitly that

$$\xi^{n_0 n_1 n_2}_{l_0 l_1 l_2} = \xi^{n_0 n_2 n_1}_{l_0 l_2 l_1} = \xi^{n_1 n_0 n_2}_{l_1 l_0 l_2} = \xi^{n_1 n_2 n_0}_{l_1 l_2 l_0} = \xi^{n_2 n_0 n_1}_{l_2 l_0 l_1} = \xi^{n_2 n_1 n_0}_{l_2 l_1 l_0}.$$

While is it possible that the sequence indexing could be revised to include only one such equivalent term, the resulting scheme would be unwieldy. A simpler solution is to simply generate the tuples $(q, r, s, t, u, v)$ in lexicographic order, and as they are converted to the tuples $(n_0, n_1, n_2, l_0, l_1, l_2)$, to exclude any that differ by an arbitrary permutation of the sets $\{n_i, l_i\}$ from a previously-generated tuple.

Let $N_b$ be a user-defined upper bound on the number of neighboring atoms in a local atomic environment. The number of terms of the resulting infinite sequence of the $\xi^{n_0 n_1 n_2}_{l_0 l_1 l_2}$ that should be used for the descriptor set is unclear, but will certainly depend on $N_b$. A lower bound is $3(N_b - 1)$ since this is the number of independent variables required to specify the relative locations of the neighboring atoms up to the action of O(3), and an upper bound is $N_b(N_b - 1)(N_b - 2)$ since there are six independent variables required to specify each unique triple of neighboring atoms up to the action of O(3). This still leaves an interval of considerable

size, and while the accuracy of many MLPs appears to improve with increasing numbers of descriptors, the computational expense of training and evaluating the MLPs increases with the number of descriptors as well. The appropriate balance of these concerns depends on the user's requirements.

## 2.5   Gaussian Processes

**Notation**

*Vectors and matrices*

- vectors are bold (e.g., $\boldsymbol{y}$)
- matrices are capitalized (e.g., $K$)
- block matrices are given carets (e.g., $\widehat{K}$).

*Operations*

- matrix determinant: $|K|$
- matrix transpose: $K^T$
- matrix trace: $\text{tr}(K) = \sum_i k_{ii}$
- vector p–norms: $\|\boldsymbol{x}\|_p = (\sum_i x_i^p)^{1/p}$; by default $\|\boldsymbol{x}\| = \|\boldsymbol{x}\|_2$

Gaussian processes (GPs) are powerful machine learning tools used to perform supervised regression and classification. This section reviews the concepts necessary for a general understanding of Gaussian process regression (GPR), including of the specific implementation used in the MLP framework developed here. GPs are appealing for regression for several reasons, including the robust uncertainty estimates that they can provide and the small number of assumptions about the form of the fitted function; it is for this last reason that GPs are classified as nonparametric models. This is beneficial for the MLP application because it is not obvious what the form of the function relating the descriptor space to the potential energy should be and a nonparametric model limits the possibility of human bias. Further description of GPs is provided in the text *Gaussian Processes for Machine Learning* [61].

A GP is a generalization of the Gaussian probability distribution, and is formally defined as a collection of random variables $\boldsymbol{f} = (f_1, f_2, ..., f_d)$ such that any finite subset of variables follows a joint multivariate Gaussian distribution [61] (the reason for the nonstandard notation will be clear below). The probability density function (PDF) of a multivariate Gaussian distribution is often written as

$$P(\boldsymbol{f}) = 2\pi^{-\frac{d}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left[-(\boldsymbol{f} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{f} - \boldsymbol{\mu})/2\right] \tag{2.22}$$
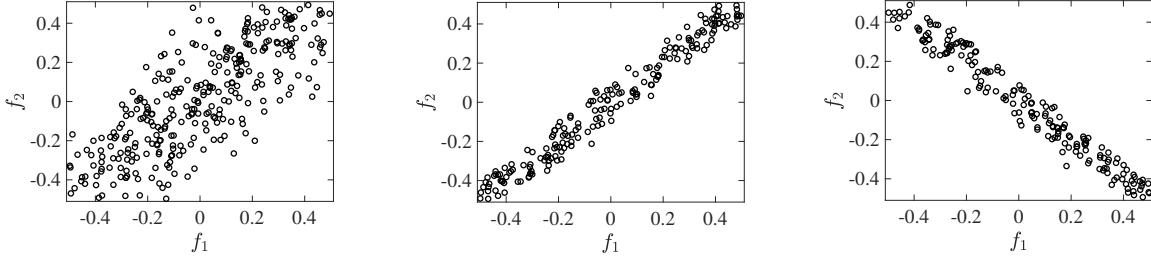
22

Figure 2.2: Jointly sampled random variables $f_1$ and $f_2$ from a multivariate Gaussian distribution. Each of the three figures show the results for different values of the covariance $\sigma_{12}^2$, with $\sigma_{12} = \{0.985, 0.999, -0.999\}$ on the left, middle, and right, respectively.

where $d$ is the dimension of the random variable, $\boldsymbol{\mu}$ is the mean vector, and $\Sigma$ is the covariance matrix. If a multivariate random variable $\boldsymbol{f}$ is sampled from a multivariate Gaussian distribution (i.e., $\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$), then any individual element $f_i$ will be distributed according to a univariate Gaussian distribution (or a normal distribution). This univariate distribution has a mean $\mu_i$ given by the corresponding element of the mean vector and a variance $\sigma_{ii}^2$ given by the $i$th diagonal element of $\Sigma$. The off-diagonal elements of $\Sigma$ (i.e., $\sigma_{ij}$ for all $i \neq j$) are the corresponding covariances between pairs of random variables. The magnitude and sign of these covariances provides information as to what is expected when jointly sampling two variables. As shown in Fig. 2.2, a large positive covariance between two random variables $f_1$ and $f_2$ suggests that larger values for $f_1$ correspond to larger values for $f_2$. A negative covariance instead implies that larger values for $f_1$ correspond to smaller values for $f_2$, and a zero-valued covariance indicates that the random variables are uncorrelated.

There are several ways to interpret GPR, and this section will follow the function-space interpretation. The objective of any regression model is usually to infer a continuous function $f(\boldsymbol{x})$ that maps a set of inputs $\boldsymbol{x}_i$ to a set of function values $f_i$. The relevance of GPs to regression begins with the idea that the (infinite) set of function values $f_i$ over the (infinite) set of points $\boldsymbol{x}_i$ comprising the domain of $f(\boldsymbol{x})$ be modeled as random variables obeying a joint multivariate Gaussian distribution. This construction would not be useful if the function values $f_i$ were independently distributed, but since $f(\boldsymbol{x})$ is assumed to be continuous the function values $f_i$ and $f_j$ associated with nearby points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ in the underlying space should be positively correlated, i.e., they should have a positive covariance. The GP then defines a probability distribution of functions $P[f(\boldsymbol{x})]$ where $\mu_i$ indicates the expectation value of $f_i$ at $\boldsymbol{x}_i$, $\sigma_{ii}$ indicates the uncertainty in the function value $f_i$ at $\boldsymbol{x}_i$, and $\sigma_{ij}$ indicates whether the function values $f_i$ and $f_j$ at distinct points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are positively or negatively correlated.

Recognizing that the set of $\boldsymbol{x}_i$ is infinite, that $\boldsymbol{\mu}$ is an infinite vector of expectation values of $f(\boldsymbol{x})$, and that $\Sigma$ is an infinite matrix of variances and covariances of the function values, the motivation for the

function-space interpretation is now clear. This views a GP as defining a probability distribution $P[f(\boldsymbol{x})]$ over functions $f(\boldsymbol{x})$ completely specified by a mean function $m(\boldsymbol{x})$ and covariance function $k(\boldsymbol{x}, \boldsymbol{x}')$, where $m(\boldsymbol{x})$ is the expectation value of $f(\boldsymbol{x})$ at $\boldsymbol{x}$ and $k(\boldsymbol{x}, \boldsymbol{x}')$ is the covariance between two points $\boldsymbol{x}$ and $\boldsymbol{x}'$ (not necessarily distinct). Of course, it is impractical to explicitly construct the space of functions on which this probability distribution is defined; instead, the probability distribution of functions is defined implicitly by $m(\boldsymbol{x})$ and $k(\boldsymbol{x}, \boldsymbol{x}')$.

Machine learning models are necessarily trained on observations, and GPR performs the training using Bayesian inference. This involves conditioning a prior distribution on a set of observations to construct a posterior distribution, i.e., an initial probability distribution of possible regression functions is updated to reflect the increased knowledge provided by the observations about the underlying system. The resulting posterior mean function $m_p(\boldsymbol{x})$ is often used to make predictions of $f(\boldsymbol{x})$, and the posterior covariance function $k_p(\boldsymbol{x}, \boldsymbol{x}')$ can be used to quantify the uncertainty of those predictions.

The training process requires an explicit choice of initial covariance function for the prior distribution. A frequent simplifying assumption is that the covariance function does not depend on the absolute position $\boldsymbol{x}$ but only on the relative position $\boldsymbol{x} - \boldsymbol{x}'$, and covariance functions with this property are referred to as being stationary. Since the prior distribution needs to be constructed before seeing any data, the prior distribution usually has a large and uniform variance to reflect the initial lack of knowledge about $f$. The training data used to condition the prior distribution includes input values $X = \{\boldsymbol{x}_i\}_{i=1}^{i=n}$ and corresponding output values $\boldsymbol{y} = \{y_i\}_{i=1}^{i=n}$ that are often assumed to be noisy measurements such that $y_i = f(x_i) + \epsilon_i$ with $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$. Conditioning on the training data effectively removes all functions from the prior distribution that do not pass through these data points, at least to within a tolerance determined by the assumed noise variance $\sigma^2$. Using the standard GP model described above, the posterior mean and posterior covariance functions have the following explicit form [61]:

$$m_p(\boldsymbol{x}) = K_{\boldsymbol{x}n}(\sigma^2 I + K_{nn})^{-1}\boldsymbol{y} \tag{2.23}$$

$$k_p(\boldsymbol{x}, \boldsymbol{x}') = k(\boldsymbol{x}, \boldsymbol{x}') - K_{\boldsymbol{x}n}(\sigma^2 I + K_{nn})^{-1}K_{\boldsymbol{x}'n}^T \tag{2.24}$$

where $K_{\boldsymbol{x}n}$ is a row vector containing the values of the prior covariance function $k(\boldsymbol{x}, \boldsymbol{x}')$ evaluated for $\boldsymbol{x}$ and the $n$ training points in $X$, and $K_{nn}$ is a covariance matrix with element $k_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j)$ in the $i$th row and $j$th column evaluated using the prior covariance function for all pairs of the $n$ training points in $X$. An example of a GP prior distribution and the posterior distribution after conditioning on the training data is displayed in Fig. 2.3.

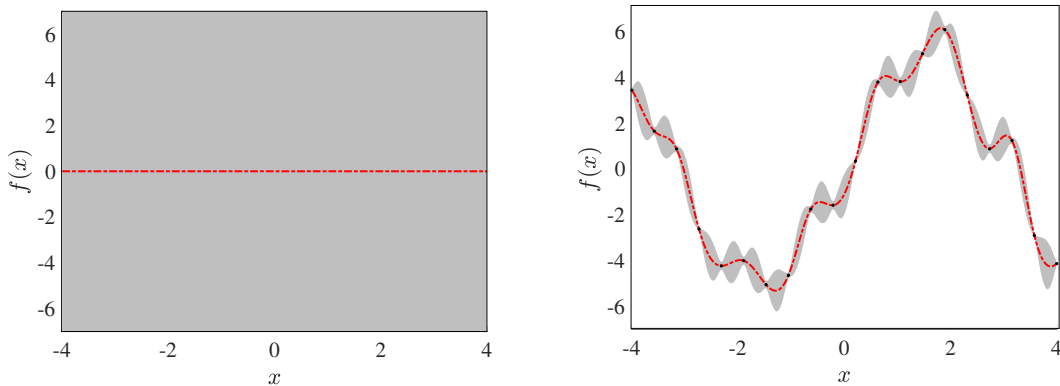While the application of Eqs. 2.23 and 2.24 appears to be straightforward, they depend sensitively on the

Figure 2.3: GP prior distribution (left) and posterior distribution (right) for a latent function $f(x)$. The red dotted line is the mean function in both distributions and the gray area represents a 95% confidence interval specified by the covariance function. The black circle markers are sampled training points. The the function used for the covariance of the prior is the 5/2-Matérn kernel.

choice of the prior covariance function and the noise variance $\sigma^2$. The entire difficulty of the training process consists of finding the prior covariance function and noise variance that return an optimal posterior distribution as defined by an appropriate objective function. This is handled by assuming a specific parametric form for $k(\boldsymbol{x}, \boldsymbol{x}')$, identifying the set of quantities $\boldsymbol{\theta}$ defining $k(\boldsymbol{x}, \boldsymbol{x}')$ and the noise variance as the hyperparameters, and applying standard numerical optimization techniques to find the optimal $\boldsymbol{\theta}$. The objective function, called the log-marginal likelihood, is the logarithm of the probability density $P(f)$ integrated over the subset of functions $f$ that pass through the training points. The training procedure therefore maximizes (in a measure-theoretic sense) the probability that a function sampled from the prior distribution would be consistent with the training points. The the log-marginal likelihood is given by [61]:

$$\log p(\boldsymbol{y}|X, \boldsymbol{\theta}) = -\frac{1}{2}\boldsymbol{y}^T(\sigma^2 I + K_{nn})^{-1}\boldsymbol{y} - \frac{1}{2}\log|\sigma^2 I + K_{nn}| - \frac{n}{2}\log 2\pi \qquad (2.25)$$

where the first term is a penalty that encourages the posterior distribution to pass through the data points, the second term is complexity penalty that helps to prevent overfitting, and the third term is a normalization term related to the underlying Gaussian distribution.

It follows that the choice of parametric form for the prior covariance function is crucial to constructing a GPR model. Covariance functions, also referred to as kernel functions, are utilized in a wide class of machine learning models [62]. One of the most commonly used covariance functions is the squared exponential (or
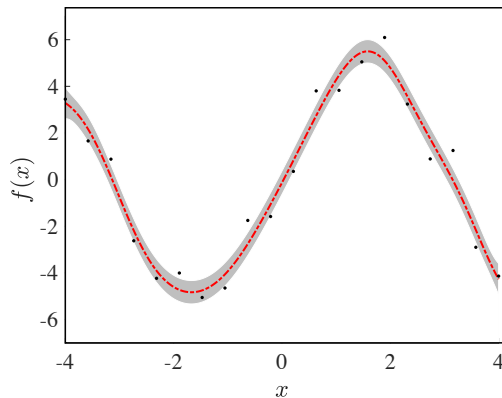
Figure 2.4: The posterior distribution for a latent function $f(x)$ after training with the data used to fit the model in Fig. 2.3, but using a squared exponential covariance function. The red dotted line is the mean function and the gray area represents a 95% confidence interval specified by the covariance function. The black circle markers are sampled training points.

the radial basis function) that has the following form [61]:

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_v^2 \exp\left(\frac{-\rho^2}{2}\right) \tag{2.26}$$

where $\rho = \|\boldsymbol{x} - \boldsymbol{x}'\|/l^2$ is a scaled distance between two points in the domain of $f(\boldsymbol{x})$, $l$ is a length-scale parameter that regulates how rapidly $f(\boldsymbol{x})$ is able to vary with respect to $\boldsymbol{x}$, and $\sigma_v^2$ regulates the initial uncertainty in the values of $f(\boldsymbol{x})$. The simplicity of this kernel is appealing for many applications, especially if that application requires modifying the covariance function. For example, the procedure for including derivative information into the model requires taking derivatives of the covariance function (described in the next section), and this operation is quite straightforward for the squared exponential. However, since the squared exponential is smooth (infinitely differentiable), each function in the posterior distribution will be smooth as well, and therefore could be written down as a Taylor series expansion using only derivative information in an arbitrarily small region of the domain. It is unclear that it is desirable or even physically reasonable that perfect knowledge of, e.g., the derivatives of an interatomic potential energy in the neighborhood of an arbitrary atomic configuration could be used to reconstruct the interatomic potential energy function for all atomic configurations. That is, it is unclear that the prior distribution should be constrained to only analytic functions when there is no evidence that nature inherently prefers such functions.

The MLP framework developed here instead uses the 5/2-Matérn kernel:

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_v^2 \left[1 + \sqrt{5}\rho + \frac{5}{3}\rho^2\right] \exp\left(-\sqrt{5}\rho\right) \tag{2.27}$$

26

which is only twice differentiable, relaxing the smoothness assumption while still enforcing that atomic forces and elastic constants derived from the MLP be continuous functions of the atomic coordinates. The squared exponential covariance function is actually a limiting case of a one-parameter family of Matérn kernels with adjustable smoothness. The effect of increasing the smoothness by using the squared exponential instead of the 5/2-Matérn kernel can be seen by comparing Fig. 2.5 to Fig. 2.3.

### 2.5.1 Sparse Gaussian Processes

Despite their attractive theoretical properties, it is intractable to use GPs for regression problems involving large datasets (i.e., $n \sim 10,000$ training examples) on a typical workstation. This difficulty is related to the $\mathcal{O}(n^3)$ computational complexity and $\mathcal{O}(n^2)$ memory requirements of finding the matrix inverses in Eqs. 2.23 and 2.24 on each training step. This has encouraged the development of a variety of approximate methods, called sparse GP methods, to reduce the order of the scaling [63]. These approximate the GP posterior using a smaller set of $m \ll n$ *inducing points* (also referred to as support variables or pseudo inputs), the locations of which are formally included among the hyperparameters that define the prior distribution and in this sense *induce* the posterior distribution [64]. The inducing points are used to construct a reduced-rank approximation $Q_{nn}$ to the covariance matrix of training points $K_{nn}$ via the Nystöm approximation. By exploiting knowledge of the reduced rank, $Q_{nn}$ can be inverted in only $\mathcal{O}(nm^2)$ time using the Woodbury matrix inversion identity (see App. A).

The sparse method used here is the Variational Free Energy (VFE) method formulated by Titsias [64]. The VFE approach constructs a lower bound $\mathcal{LB}$ to the log-marginal likelihood that, in addition to measuring the likelihood of the observed data given the prior distribution, also measures the difference between the true and approximate posterior distributions. Therefore, using $\mathcal{LB}$ as the objective function not only ensures that the posterior distribution fits the data, but also that the full GP is reasonably approximated [65]. One drawback to this approach is that it can overestimate the noise variance $\sigma^2$ [65], so enforcing smaller values of this hyperparameter could be necessary to enhance fitting. Another drawback of the VFE is that the $\mathcal{LB}$ is prone to local minima that make finding the global optimum more difficult [65], though this can in principle be overcome with careful choice of the optimization routine. This work uses Powell's Derivative Free Optimizers (PDFO) [66], which iteratively construct quadratic approximations to the objective function to determine where to sample next, and which in our experience, out-perform gradient-based optimization methods for this application. The $\mathcal{LB}$ has the following form [64]:

$$\mathcal{LB} = -\frac{1}{2} \boldsymbol{y}^T Q_{nn}^{-1} \boldsymbol{y} - \frac{1}{2} \log |Q_{nn}| - \frac{n}{2} \log 2\pi - \frac{r}{2\sigma^2} \tag{2.28}$$
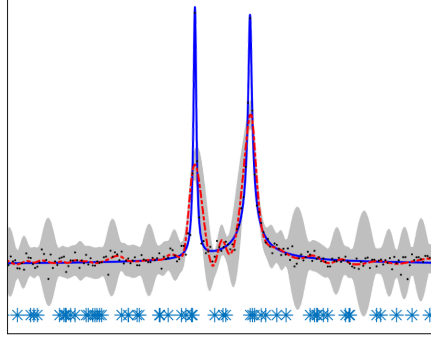
Figure 2.5: Fitting a 1D function (blue line) with the VFE sparse GP. The model prediction is the red dotted line, the black circle markers are training points with added noise, the grey area represents a 95% confidence interval evaluated using the posterior covariance, and the blue stars at the bottom mark where inducing points have been selected.

where $Q_{nn} = K_{nm}K_{mm}^{-1}K_{nm}^T + \sigma^2 I$, $r = \text{tr}(K_{nn} - K_{nm}K_{mm}^{-1}K_{nm}^T)$, and $K_{nm}$ is the covariance matrix between the training points and the inducing points. The corresponding sparse posterior mean and covariance functions are [64]

$$m_p^q(\boldsymbol{x}) = K_{\boldsymbol{x}m}(\sigma^2 K_{mm} + K_{nm}^T K_{nm})^{-1}K_{nm}^T\boldsymbol{y} \tag{2.29}$$

$$k_p^q(\boldsymbol{x},\boldsymbol{x}') = k(\boldsymbol{x},\boldsymbol{x}') - K_{\boldsymbol{x}m}[\sigma^2(\sigma^2 K_{mm} + K_{nm}^T K_{nm})^{-1} - K_{mm}^{-1}]K_{\boldsymbol{x}'m}^T \tag{2.30}$$

An example of a sparse model fitting a 1D function is shown in Fig. 2.5. The function was purposely constructed to be difficult for GPR methods, and yet the sparse model is still able to approximate the function reasonably well. The presence of the function's sharp peaks forces the length scale hyperparameters of the prior covariance function to be small (the functions in the prior distribution can vary quickly), requiring a relatively dense sampling of the space to constrain the uncertainty in the posterior distribution. Notice that the posterior variance is reduced only where there are inducing points, and that there is a lower bound on the posterior variance that reflects the magnitude of the sampling error.

The selection of inducing points is a subtle and important part of the algorithm. While the inducing points do not have to be chosen from the training set, they usually are to reduce the complexity of the optimization problem. Here, the set of inducing points is constructed by a greedy (suboptimal but with low computational cost) algorithm to build an $\epsilon$-net on the training set. An $\epsilon$-net is a subset of a point cloud point in a general-dimensional space such that every point in the cloud is within a specified distance $\epsilon$ of one or more points the subset (the subset is referred to as the net). The procedure for constructing the $\epsilon$-net begins by picking the first point to add to the set at random, and then iteratively adding the point that is furthest away (as defined by the $L^2$ norm) in descriptor space from any point already in the set. This process
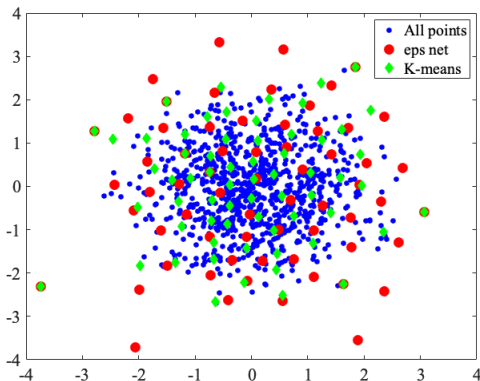
Figure 2.6: Representative subsets of a two-dimensional point cloud as constructed using either a greedy $\epsilon$-net or the centroids of clusters identified by a $k$-means algorithm.

continues until the distance to the most distant point falls below a specified parameter $\epsilon$. Fig. 2.6 shows an example of such an $\epsilon$-net built for a two-dimensional point cloud; an alternative method for building a representative subset, as the centroids of the clusters identified by the $k$-means clustering algorithm, is shown for comparison. Two advantages of the $\epsilon$-net are that points in the subset are more evenly distributed, avoiding clustering of inducing points in regions already well-covered by the training set, and that the $\epsilon$-net preferentially includes points on the boundary of the point cloud, reducing the uncertainty of the posterior distribution when the model is asked to extrapolate.

### 2.5.2 Including atomic forces

There are several modifications that need to be made to the standard GP to align it with the requirements of the MLP application. The training data for the MLP model as provided by electronic structure calculations includes atomic forces and total potential energies of the simulation cell. Atomic forces are incorporated into the GPR model by formulating a joint (or multivariate) GP in the same way as does the GAP [67]. This not only allow the atomic forces to be included in the training data, but allows the forces to be directly evaluated for arbitrary local atomic configurations. If the PES can be well described by a GP, then the forces necessarily can be well described by a GP as well; this is a consequence of the forces being defined by the negative gradient of the PES, and of the property of GPs that applying a linear operator results in another GP. For simplicity of notation, the derivation below treats the forces as gradients and then corrects for the sign by taking the negative of the force observation data $\boldsymbol{y}^f$.

Constructing the joint GP requires formulating the covariances and mixed covariances for the force GP and energy GP priors. Let $k^{ee}(\boldsymbol{x}, \boldsymbol{x}')$ be the covariance function for the energies (the 5/2-Matérn kernel) where $\boldsymbol{x}$ and $\boldsymbol{x}'$ are two points in the space of descriptors. Let $O$ be a fixed Cartesian coordinate system as
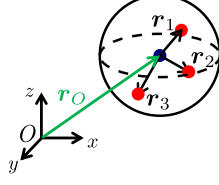
Figure 2.7: The relationship of the relevant coordinate systems. $O$ is the coordinate system with respect to which the forces are calculated. The displacement vectors $\boldsymbol{r_k}$ to the neighboring atoms in local atomic environments are defined with respect to the position of the central atom.

indicated in Fig. 2.7. The covariance between the force on an atom with local environment $\boldsymbol{x}$ and the energy of an atom with local environment $\boldsymbol{x}'$ is given by [67]:

$$k^{fe}(\boldsymbol{x}, \boldsymbol{x}') = \frac{\partial k^{ee}(\boldsymbol{x}, \boldsymbol{x}')}{\partial r_{O,k}} \tag{2.31}$$

where $r_{O,k}$ is the $k$th Cartesian coordinate of the atom with local environment $\boldsymbol{x}$, with the same coordinate system $O$ used to define the $k$th component of the observed force. Similarly, the covariance between both environments' forces is [67]:

$$k^{ff}(\boldsymbol{x}, \boldsymbol{x}') = \frac{\partial^2 k^{ee}(\boldsymbol{x}, \boldsymbol{x}')}{\partial r_{O,k} \partial r'_{O,q}} \tag{2.32}$$

where $r'_{O,q}$ is the $q$th Cartesian coordinate of the atom with local environment $\boldsymbol{x}'$, and the same coordinate system $O$ is again used to define the $q$th component of the observed force.

Equations 2.31 and 2.32 require derivatives of the descriptors with respect to the coordinates of the central atom in coordinate system $O$, but $k^{ee}(\boldsymbol{x}, \boldsymbol{x}')$ is an explicit function of the descriptors, and the descriptors are explicit functions of spherical coordinates of the neighboring atoms with respect to the central atom. The calculation of these derivatives therefore requires three coordinate transformations. Let $r^s_{p,\zeta}$ be the $\zeta$th component of the spherical coordinates of the $p$th neighboring atom relative to the central atom, and $r_{p,k}$ be the Cartesian coordinates of the $p$th neighboring atom relative to the central atom, and $r^s_{p,\zeta}$ be the $\zeta$th component of the spherical coordinates of the $p$th neighboring atom relative to the central atom. The first transformation involves multiplying by the Jacobian of the descriptors $x_d$ with respect to the relative spherical coordinates $r^s_{p,\zeta}$, the second involves multiplying by the Jacobian of the relative spherical coordinates $r^s_{p,\zeta}$ with respect to the relative Cartesian coordinates $r_{p,k}$, and the third involves multiplying by the Jacobian of the relative Cartesian coordinates $r_{p,k}$ with respect to the absolute Cartesian coordinates $r'_{O,q}$. That is:

$$\frac{\partial k^{ee}(\boldsymbol{x}, \boldsymbol{x}')}{\partial r_{O,k}} = \sum_{dp\zeta} \frac{\partial k^{ee}(\boldsymbol{x}, \boldsymbol{x}')}{\partial x_d} \frac{\partial x_d}{\partial r^s_{p,\zeta}} \frac{\partial r^s_{p,\zeta}}{\partial r_{p,k}} \frac{\partial r_{p,k}}{\partial r_{O,k}} \tag{2.33}$$

With the covariance functions so defined, the joint GP model is constructed by organizing the covariances

of the forces and energies into a block matrix such that:

$$\widehat{K} = \left[ \begin{array}{c|c} K^{ff} & K^{fe} \\ \hline K^{ef} & K^{ee} \end{array} \right] \tag{2.34}$$

Similarly, the training data is organized as a block matrix:

$$\widehat{\boldsymbol{y}} = \left[ \begin{array}{c} -\boldsymbol{y}^f \\ \hline \boldsymbol{y}^e \end{array} \right] \tag{2.35}$$

Replacing the covarinace matrices with block matrices, the sparse model Eqs. 2.28-2.30 become:

$$\mathcal{LB} = -\frac{1}{2}\widehat{\boldsymbol{y}}^T \widehat{Q}_{nn}^{-1} \widehat{\boldsymbol{y}} - \frac{1}{2}\log|\widehat{Q}_{nn}| - \frac{n}{2}\log 2\pi - \frac{1}{2}\tilde{r} \tag{2.36}$$

$$m_p^q(\boldsymbol{x}) = \widehat{K}_{\boldsymbol{x}m}(\widehat{K}_{mm} + \widehat{K}_{nm}^T \Lambda^{-1} \widehat{K}_{nm})^{-1} \widehat{K}_{nm}^T \Lambda^{-1} \widehat{\boldsymbol{y}} \tag{2.37}$$

$$k_p^q(\boldsymbol{x}, \boldsymbol{x}') = k(\boldsymbol{x}, \boldsymbol{x}') - \widehat{K}_{\boldsymbol{x}m}[(\widehat{K}_{mm} + \widehat{K}_{nm}^T \Lambda^{-1} \widehat{K}_{nm})^{-1} - \widehat{K}_{mm}^{-1}]\widehat{K}_{m\boldsymbol{x}} \tag{2.38}$$

where $\widehat{Q}_{nn} = \widehat{K}_{nm}\widehat{K}_{mm}^{-1}\widehat{K}_{nm}^T + \Lambda$, $\tilde{r} = \mathrm{tr}[\Lambda^{-1}(\widehat{K}_{nn} - \widehat{K}_{nm}\widehat{K}_{mm}^{-1}\widehat{K}_{nm}^T)]$, and $\Lambda$ is a diagonal matrix with entries $\sigma_e^2$ or $\sigma_f^2$ for the energy and force noise variances respectively [67]:

$$\Lambda = \begin{bmatrix} \sigma_f^2 & & \\ & \ddots & \\ & & \sigma_e^2 \end{bmatrix} \tag{2.39}$$

This makes the expected noise for the two GPs independent, which is sensible since the atomic forces and energies should not necessarily have functionally-related uncertainties or measurement errors.

### 2.5.3  Including total energies of atomic configurations

Electronic structure theory does not naturally provide per atom potential energies, the energy instead being a property of the electron density. This means that the GP for the per atom potential energies must be replaced with one for the total potential energy of a simulation cell. Since the summation that relates the per atom energies to the total energy is a linear operation (the same as the derivative relating potential energies to forces), the total energies can also be assumed to be described by a GP. All that remains is to define the covariances and mixed covariances accordingly.

For brevity, a configuration of atoms in an electronic structure calculation is referred to as a *snap* (short for snapshot) and is denoted as $\mathcal{S}$. By linearity of summation, the covariance between the total potential energies of snaps $\mathcal{S}$ and $\mathcal{S}'$ is simply the sum of the per atom atomic energy covariances for all pairs of atoms

between the snaps [67]:

$$k^{EE}(\mathcal{S}, \mathcal{S}') = \sum_{\substack{\boldsymbol{x}_j \in \mathcal{S} \\ \boldsymbol{x}'_g \in \mathcal{S}'}} k^{ee}(\boldsymbol{x}_j, \boldsymbol{x}'_g) \tag{2.40}$$

In the same way, a mixed covariance between the total energy of snap $\mathcal{S}$ and the force on an atom with local environment $\boldsymbol{x}'$ is

$$k^{Ef}(\mathcal{S}, \boldsymbol{x}') = \sum_{\boldsymbol{x}_j \in \mathcal{S}} k^{ef}(\boldsymbol{x}_j, \boldsymbol{x}') \tag{2.41}$$

Since equations 2.40 and 2.41 only involve summations, the snap covariance matrices can be constructed from those involving atomic energies by matrix operations [67]:

$$K^{EE} = L K^{ee} L^T \tag{2.42}$$

$$K^{Ef} = L K^{ef} \tag{2.43}$$

where $L$ is a matrix whose rows correspond to snaps in the training set and columns correspond to *all* atomic environments in all the snaps in the training set. The $l_{ij}$ element of this matrix is one if the $j$th atomic environment is in the $i$th snap and is zero otherwise.

The final complication is that the equations for the sparse model often involve covariances with inducing points that are not associated with any particular snap. This means that the definition of the $\widehat{K}_{mm}$ remains unchanged, and the definition of $\widehat{K}_{nm}$ changes to:

$$\widehat{K}_{nm} = \left[ \begin{array}{c|c} K^{ff}_{nm} & K^{fe}_{nm} \\ \hline K^{Ef}_{nm} & K^{Ee}_{nm} \end{array} \right] \tag{2.44}$$

using what is believed to be obvious notation. This version of $\widehat{K}_{nm}$ should be used for all occurrences in Eqs. 2.36-2.38. The block matrix of the training data needs to be similarly revised to:

$$\widehat{\mathbf{y}} = \left[ \begin{array}{c} -\boldsymbol{y}^f \\ \hline \boldsymbol{y}^E \end{array} \right] \tag{2.45}$$

where $\boldsymbol{y}^E$ is a vector containing the total energies of all snaps in the training data. The last required modification is to change the associated noise hyperparameter of the total energy GP (in accordance with the central limit theorem) to:

$$\sigma^2_{E,j} = N_j \sigma^2_e \tag{2.46}$$

where $N_j$ is the number of atoms in snap $j$ and the resulting $\sigma^2_{E,j}$ appear on the lower diagonal of $\Lambda$.
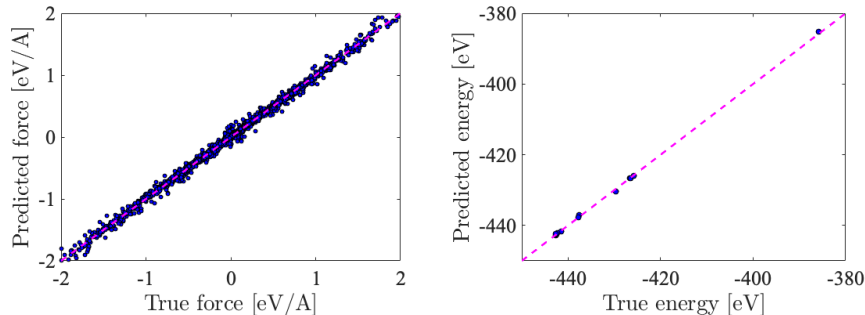
Figure 2.8: Test set values for atomic forces (left) and total energies (right) of Cu systems. The DFT-calculated ground truth is on the horizontal axis, and the values predicted by the MLP framework are on the vertical axis.

Table 2.1: Test set RMSEs of evaluated forces and energies of a Cu data set for our framework (labeled $\xi$-GP) and others reported in Ref. [24].

|  | $\xi$-GP | GAP | MTP | NNP | SNAP | qSNAP |
|---|---|---|---|---|---|---|
| Energy [meV/atom] | 4.54 | 0.56 | 0.52 | 1.68 | 0.87 | 1.16 |
| Force [eV/A] | 0.04 | 0.02 | 0.01 | 0.06 | 0.08 | 0.05 |

## 2.6   Resulting Framework

Using the atomic descriptors described in Sec. 2.2 with the sparse Gaussian process described in Sec. 2.5 produces a functioning MLP framework that can fit a single element PES using atomic forces and total potential energies of simulation cells. Fortunately, Ref. [24] reports the accuracy and computational cost of fitting DFT-calculated forces and total energies for five single element systems with several prominent MLPs. They included a data set containing a collection of *ab initio* MD snapshots with low and high temperature trajectories, vacancies and surfaces, and distorted structures. The results of using our model to fit the Cu data provided by Ref. [24] can be seen in Fig. 2.8. The root mean squared error (RMSE) of the evaluated forces and energies are 0.0370 eV/A and 4.54 meV/atom, respectively.

While the reported test set errors of our framework are exceedingly low, it is relevant to compare its performance to other frameworks in the literature for the same Cu data set. As shown in Table 2.1) ours predicts forces with higher accuracy than NNP, SNAP, and qSNAP. However, our model's predicted energies are less accurate than those of any of the other models. Several straightforward ideas about how the accuracy could be improved are discussed in the conclusions of this dissertation.

# Chapter 3

# Energy storage under high-rate compression of single crystal tantalum

The following chapter is comprised of a manuscript that was a joint effort of J. C. Stimac, N. Bertin, J. K. Mason, and V. V. Bulatov [68].

## 3.1  Introduction

The Taylor-Quinney coefficient (TQC), or the coefficient of thermo-mechanical conversion, is named for two scientists who observed that only a fraction of the mechanical work performed on a plastically deformed material is converted into heat and concluded that the remainder must be stored in the material [8]. The TQC is specifically defined as the fraction of mechanical work converted into heat, with the difference between one and the TQC indicating the extent of energy storage. The original experiments by Taylor and Quinney imposed significant torsion and compression on copper rods and found that this ratio began around 0.85 or 0.9 before converging to 1.0 with increasing strain. As a consequence, the TQC is frequently approximated as a constant around 0.9. Although its nature was unknown at the time, Taylor and Quinney convincingly related the stored energy to an increase in the flow stress (plastic strength), a phenomenon that would eventually become known as strain hardening. This suggested that perhaps the stored energy itself constituted a material state variable that evolved during plastic deformation and defined the material strength. Less than a year after publishing the original paper, Taylor further developed this connection by proposing (simultaneously with Orowan and Polanyi) that crystal plasticity results from the motion of

distinct line-shaped defects called dislocations [69]. Following that ground-breaking discovery nearly 90 years ago, it has subsequently been established that the stored energy reflects changes in the defect microstructure, e.g., the distributions of dislocations, point defects, grain boundaries, stacking faults, deformation twins, etc. For most deformation conditions of interest though, both strain hardening and energy storage are mainly consequences of the multiplication and storage of dislocations.

From a theoretical standpoint, it has been proposed that the TQC is related to the configurational entropy. An early review made note of estimated contributions to this entropy change from dislocations and point defects to estimate the free energy change during a cold-work process [70]. Later, the TQC became a quantity of interest in thermodynamic theories of plasticity [71], including the thermodynamic dislocation theory (TDT) [72, 73] which introduces an effective temperature to quantify the configurational disorder of non-equilibrium dislocation systems. This later framework was recently used to derive thermodynamic constraints on the TQC [74]. Recent work by Zubelewics also proposed a novel interpretation of the TQC as a measure of the ability of the material to raise the configurational entropy enough to untangle the dislocation network and enable dislocation motion [75].

Experimentally, direct measurement of the TQC requires delicate accounting of the heat released by the material during the prescribed straining conditions. The measurement accuracy is hampered at low strain rates by the typically small amounts of heat released and the difficulty of thermally isolating the system from its environment. High strain-rate deformation can release substantially more heat and suffers from fewer thermal losses to the environment, but can make thermal measurements themselves more difficult. For example, measurements of the dissipated thermal energy by means of infrared detection have been found to be rather sensitive to the calibration procedure, with updates to this procedure able to change the apparent TQC by as much as a factor of two [76]. During material tests performed at most extreme deformation rates, e.g., at the National Ignition Facility [77, 78], uncertainty in specimen temperature can reach hundreds of Kelvin making extraction of the material's intrinsic thermo-mechanical properties quite difficult. Despite (or perhaps due to) our continually expanding and improving experimental capabilities, the range of TQC values appearing in the literature is not narrowing, and if anything has expanded to the entire interval between 0.1 and 1.0. Reference [12] provides a review of the recent literature on this increasingly controversial subject.

Accurate knowledge of the TQC is desirable since it allows the prediction of the temperature change of a material that is plastically deforming in adiabatic conditions, e.g., at high strain rates. Given that a material's mechanical properties can be temperature dependent, the TQC can appear as a critical parameter in continuum-level models of plastic deformation [79, 13]. The value of the TQC can be relevant to other simulations of material behavior as well, with the stored energy of cold work known to be implicated in the onset of recrystallization and the Baushinger effect [80].

Since it remains effectively impossible to partition mechanical work between released heat and stored energy in experimental tests performed at the most extreme deformation conditions, computer simulations offer an alternative means to gain much needed insights into energy storage. Unlike high strain rate experiments, numerical simulations can provide complete information about the simulated material response and allow mechanical work, released heat and stored energy to all be computed and monitored incrementally at every time step. Reference [80] was perhaps the first simulation to explore energy storage due to dislocation multiplication during plastic deformation, and used a discrete dislocation dynamics (DDD) method. However, given that it was a highly abridged two-dimensional version of DDD with several critical parameter values adjusted in an ad hoc fashion to compensate for the missing third dimension, the values of the TQC reported there do not represent any real material or deformation condition. More recently, molecular dynamics (MD) simulations of dislocation motion in single crystals and of plasticity in nanocrystals [13] were used to calculate the TQC of aluminum, copper, iron, and tantalum under high strain-rate shearing in isothermal-isobaric conditions. Based on these simulations, the authors concluded that only extreme dislocation multiplication up to densities of $10^{16}$ m$^{-2}$ or extreme grain refinement down to diameters of 10 nm could account for a TQC lower than 1.0 at the deformation rates accessible to MD simulations ($\sim 10^8$ s$^{-1}$). The authors then argued that grain refinement is the more likely of the two mechanisms to account for appreciable energy storage in practice. That said, their simulations of nanograined samples had grain sizes that were orders of magnitude smaller than what is usually observed in experimental specimens, artificially inflating the apparent contribution of grain refinement to energy storage. Furthermore, their simulations of dislocated single crystals initially contained only a single edge dislocation dipole, effectively precluding dislocation multiplication as a means of energy storage.

In this paper, large-scale MD simulations of high-rate deformation of single crystal tantalum in isothermal-isochoric conditions are performed and analyzed up to a true strain of 1.0 to observe and quantify the role of dislocation multiplication as a mechanism of energy storage. Tantalum (Ta) has long been recognized for its toughness and resistance to corrosion [81], resulting in its use in a variety of high strain-rate applications. For this reason, Ta is used as a model material to study body-centered cubic (BCC) metals under extreme strains and strain rates. The MD simulations and the methods used to monitor the thermo-mechanical response of simulated single crystals are described in Sec. 3.2. Our simulation results on the TQC and energy storage in Ta for the specified deformation protocols are given in Sec. 3.2. The relative utility of the differential and integral forms of the TQC for understanding the materials response, a phenomenological model for both the differential and integral TQC as functions of plastic strain, and a model that describes the energy stored in the defect microstructure as a function of only the dislocation density and point defect concentration are all discussed in Sec. 3.4. Conclusions on the basis of our observations are offered in Sec. 3.5.
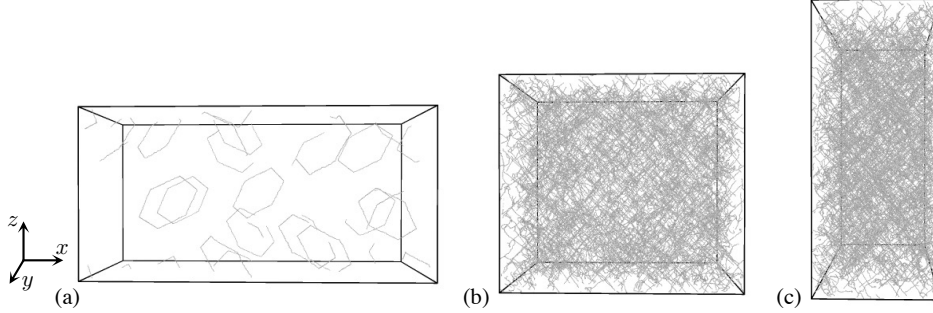
Figure 3.1: Simulation volume and dislocation network at true plastic strains of (a) 0.0, (b) 0.4, and (c) 1.0. The compression axis $x$ points to the right in all three frames.

## 3.2 Methods

All simulations were carried out using the LAMMPS software [82] with the embedded atom method (EAM) interatomic potential for Ta of Li, et al. [83] and a 10 fs time step. A single-crystal rectangular prism of Ta containing about 33 million atoms was generated with a 2:1:1 aspect ratio along the [100], [010] and [001] crystallographic directions. Periodic boundary conditions were enforced in all directions, and the crystal was seeded with an equal number of hexagonal dislocation loops for each of the four $\frac{1}{2}\langle 111\rangle$ Burgers vectors of the body-centered cubic lattice as is visible in Fig. 3.1a. The crystal was then subject to uniaxial compression along the $x$-dimension (the longest dimension aligned with the [100] direction), and the displacements along the $y$- and $z$-dimensions enforced that the deformed crystal remain a tetragonal prism of constant volume. The temperature was maintained at 300 K by means of a Langevin thermostat. All simulation settings were identical or similar to those reported in Ref. [84], the only difference being the use of isothermal-isochoric rather than isothermal-isobaric conditions. This minor change was made largely for computational convenience since the crystal volume was fixed to maintain near-zero pressure during plastic flow (crystal plasticity is naturally isochoric).

The crystal was subjected to two different dislocation seed configurations and deformation protocols to investigate the extent to which the TQC depends on the initial dislocation density and loading conditions. The first simulation (A) initialized the crystal with 12 hexagonal dislocation loops that were 40 lattice constants in diameter and subjected the crystal to a constant compressive strain rate along the $x$-axis of $2 \times 10^8$ s$^{-1}$ throughout the entire simulation. As discussed further below, this caused a stress overshoot to develop as a consequence of the initial few dislocations requiring time to multiply to a level sufficient to relieve the stress by means of their motion. The second simulation (B) initialized the crystal with 16 hexagonal dislocation loops that were 80 lattice constants in diameter, twice that of the first simulation. The compressive strain rate along the $x$-axis was linearly ramped from 0 to $4 \times 10^8$ s$^{-1}$ over an interval of
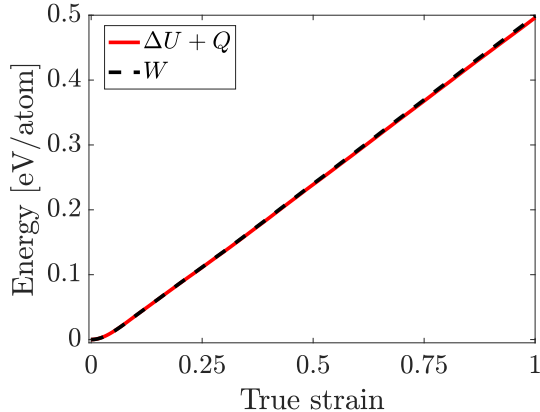
Figure 3.2: Verification that the ramped simulation obeys the first law of thermodynamics. The total mechanical work $W$ done on the crystal is equal to $\Delta U + Q$, or the sum of the change in the internal energy $\Delta U$ and the heat $Q$ exchanged with the thermostat.

2 ns to a total strain of 0.4, and afterwards held constant at $4 \times 10^8$ s$^{-1}$. As a result of ramping the strain rate and initializing a greater number of larger-sized dislocation loops, simulation B did not exhibit a stress overshoot.

The amount of external mechanical work $dW$ performed on the crystal in a given time step was computed as

$$dW = \sum_i \sigma_{ii} A_i dL_i \tag{3.1}$$

or as the sum of the products of the diagonal components $\sigma_{ii}$ of the internal stress tensor, the areas $A_i$ of the corresponding faces, and the incremental displacements $dL_i$ of the faces in their normal directions (the shear components all vanished since the simulation volume remained strictly tetragonal). The mechanical work, the internal energy $U$ of the entire crystal, and the thermal energy $\delta Q$ exchanged with the Langevin thermostat were continuously recorded during the simulation. Figure 3.2 shows that, to a high degree of accuracy and over the entire simulation, the total mechanical work $W$ performed on the crystal (the integral of Eq. 3.1) is equal to the sum of the change in the internal energy $\Delta U$ and the heat $Q$ exchanged with the thermostat. Our partitioning of mechanical work into dissipated heat and stored energy should be accurate to the extent that this energy balance holds.

In their original paper, Taylor and Quinney computed the work-to-heat conversion ratio by dividing the heat released by the total mechanical work, the latter computed as the area under the true stress–true strain curve [8]. It has subsequently become more common to use the plastic work instead of the total work, the rationale being that the elastic component is reversible and therefore never converted into heat. This is schematically illustrated on the left side of Fig. 3.3 where plastic work only begins to accrue after the onset
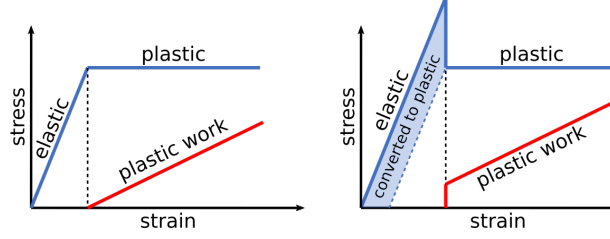
Figure 3.3: Stress-strain schematics that indicate the relationship of yield point phenomenon and the transformation of elastic work to plastic work at low strains.

of yielding. This distinction has no bearing on the results and observations in the original paper since, at the low deformation rates employed there, the elastic work constituted a negligible fraction of the total work. However, at the high compressive strain rates employed in our MD simulations the crystal experienced much higher stresses on the order of several GPa (and even higher during the stress overshoot in simulation A). Thus, a substantial fraction of the total work is stored in the form of elastic strain energy. Since our concern is with the part of the total work that is irreversibly stored in the defect microstructure, here we follow the prevailing practice and use only the plastic work to calculate the TQC.

The plastic work is found by subtracting the elastic energy (the energy that could be recovered by unloading the crystal) from the total work. The equation for the elastic energy for cubic crystals in the $\langle 100 \rangle$ frame reduces to

$$
\begin{aligned}
E_{\text{el}} &= \frac{V}{2} \sum_{ijkl} S_{ijkl} \sigma_{ij} \sigma_{kl} \\
&= \frac{V}{2} [S_{11}(\sigma_{11}^2 + \sigma_{22}^2 + \sigma_{33}^2) + 2S_{12}(\sigma_{11}\sigma_{22} + \sigma_{11}\sigma_{33} + \sigma_{22}\sigma_{33}) + S_{44}(\sigma_{12}^2 + \sigma_{13}^2 + \sigma_{23}^2)] \quad (3.2)
\end{aligned}
$$

where $V$ is the crystal volume, $\sigma_{ij}$ are the internal stress tensor components, and $S_{ijkl}$ and $S_{ij}$ are the elastic compliance components in tensor and Voigt notations. The elastic compliance components were found from the elastic stiffness matrix components $C_{ij}$ by the standard equations

$$
\begin{aligned}
S_{11} &= \frac{C_{11} + C_{12}}{(C_{11} - C_{12})(C_{11} + 2C_{12})} \\
S_{12} &= \frac{-C_{12}}{(C_{11} - C_{12})(C_{11} + 2C_{12})} \\
S_{44} &= \frac{1}{C_{44}}
\end{aligned}
\quad (3.3)
$$

and the elastic stiffness components were measured in a separate calculation using the same interatomic potential, temperature and pressure as in simulations A and B.

The Langevin thermostat employed in our simulations did not strictly maintain the temperature at 300 K,
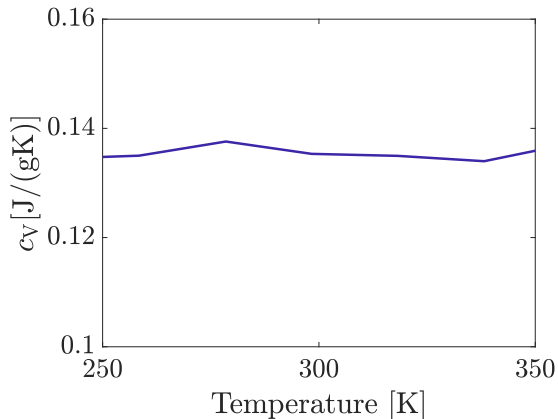
Figure 3.4: Constant-volume specific heat capacity of Ta over the relevant temperature range. The blue trend trend interpolates observations measured over 1 ps intervals.

and instead allowed the temperature to increase by a few Kelvin. Therefore, a precise calculation of the fraction of mechanical work expended in forming defect structures requires knowledge of the heat capacity of Ta for the selected interatomic potential [83]. The specific heat capacity at constant volume was measured by maintaining the simulation cell at constant volume, slowly incrementing the target temperature over the interval 250 K $\leq T \leq$ 350 K, and continuously observing the heat exchanged with the thermostat. The results are reported in Fig. 3.4 where the specific heat capacity at constant volume apprears to stay constant at 0.135 J/(gK) in the relevant interval, coincidentally agreeing with the $c_p = 0.14$ J/(gK) reported in Ref. [13] for a different interatomic potential for Ta.

## 3.3  Results

The stress in simulation A (with a constant strain rate) is shown in Fig. 3.5a and exhibits an overshoot to an upper yield point followed by a substantial drop to a lower yield point, with the flow stress afterwards remaining nearly constant for the rest of the simulation. Particularly at high strain rates, the initial transient in such stress-strain curves is generally attributed to delayed dislocation multiplication in response to a rising stress. The initially low dislocation density limits the amount of deformation that can be accommodated via dislocation motion, resulting in the majority of the mechanical work being converted into elastic energy. On reaching the upper yield point, dislocations have multiplied sufficiently for their motion to produce plastic strain at a rate that equals and even exceeds the externally-imposed strain rate; this overproduction of plastic strain causes the subsequent stress drop. As the stress drops, the elastic energy accumulated during the initial stress rise is rapidly converted into plastic work (shaded area in the schematic on the right of Fig. 3.3), most of which is immediately released as heat. This temporarily increased the temperature to 307 K
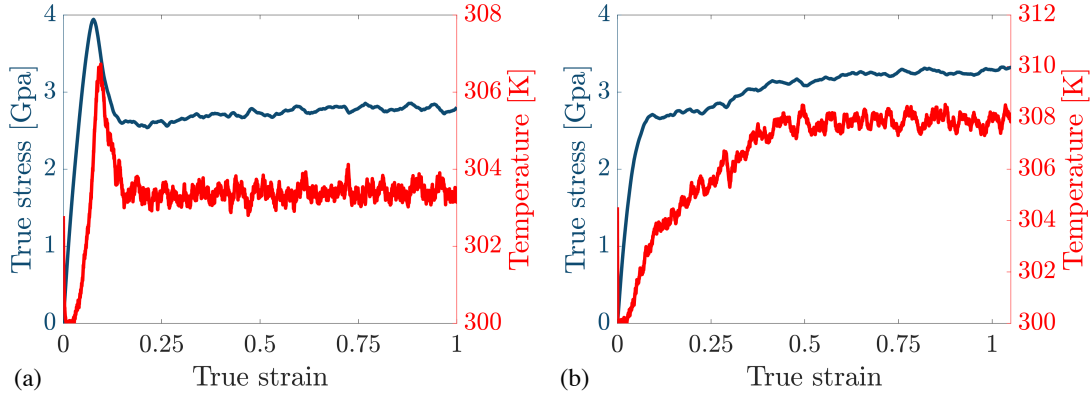
Figure 3.5: True stress and absolute temperature as a function of true strain for (a) simulation A and (b) simulation B. While the temperature is not constant, the Langevin thermostat keeps the temperature to within a few Kelvin of the target temperature, with the difference being roughly proportional to the instantaneous rate of plastic work.

despite the use of a Langevin thermostat intended to maintain the temperature constant at 300 K. The thermostat relaxation time was set to 10 ps as a compromise between the desires for a constant simulation temperature and to avoid excessively perturbing the dynamics from the random atomic kicks and friction forces that the thermostat employs. The majority of the heat released during the simulation is still exchanged with the thermostat, with the minority responsible for the observed temperature variations being accounted for by adding the term $C_V dT(t)$ to the incremental heat exchanged with the reservoir $dQ(t)$. The specific heat capacity 0.135 J/(gK) (Fig. 3.4) and mass density 16.55 g/cm$^{-3}$ were specific to the Ta model used in this study. This correction affects our computed TQC only during the initial yield transient and becomes insignificant past the lower yield point where the temperature settles at about 304 K and stays very nearly constant through the rest of the simulation.

The intent with simulation B was to suppress the initial upper-lower yield transient that caused the rapid conversion of elastic energy into plastic work after the upper yield point in simulation A, thus possibly affecting the value of the TQC. This was achieved (Figure 3.5b) by seeding a greater initial dislocation density and linearly increasing (ramping) the strain rate. The temperature in simulation B similarly rises monotonically with strain, though to a higher value than in simulation A as is consistent with the higher steady-state strain rate.

The differential TQC is defined as the ratio of the incremental heat released to incremental plastic work performed during the time step starting at time $t$, or [79, 85]

$$\beta_{\text{diff}}(t) = \frac{dQ(t) + C_V dT(t)}{dW_p(t)} \tag{3.4}$$
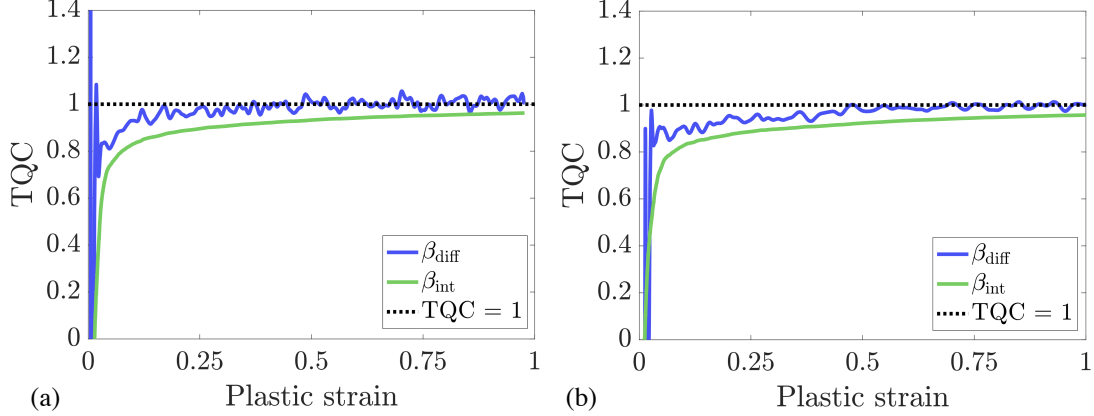
42

Figure 3.6: Differential (blue) and integral (green) TQC as calculated by Eqs. 3.4 and 3.5 for (a) constant (simulation A) and (b) ramped (simulation B) strain-rate compression.

where $dW_p(t) = dW(t) - dE_{el}(t)$ is the incremental plastic work defined by Eqs. 3.1 and 3.2 and the remaining quantities are taken directly from the standard output of the MD simulation. The integral TQC is instead defined as the ratio of total heat released by time $t$ and the total plastic work done on the crystal by the same time $t$:

$$\beta_{int}(t) = \frac{Q(t) + C_V \Delta T(t)}{W_p(t)}. \tag{3.5}$$

While $\beta_{diff}$ is more informative as a measure of the instantaneous partitioning of mechanical work along the deformation path, $\beta_{int}$ is more commonly reported of the two in the literature, perhaps for the reason that it is easier to measure experimentally. As was explained in [85] where the two parameters were first referred to as "differential" and "integral", the integral TQC is not equal to the integral of the differential TQC. Instead, the following equality holds:

$$\beta_{int}(t) = \int_0^t \beta_{diff}(t') \frac{dW_p(t')}{W_p(t)} \tag{3.6}$$

in which the contribution of $\beta_{diff}$ computed at time $t'$ to $\beta_{int}$ is weighted by the fractional plastic work performed during $[t', t' + dt']$ relative to that during $[0, t]$.

Figure 3.6 shows the computed differential and integral TQC values as blue and green lines for simulations A and B. The differential TQC is susceptible to high-frequency fluctuations attributed to uncertainties in the small values of the denominator in Eq. 3.4 computed over short time intervals, especially during the initial stress rise where only a small fraction of the mechanical work is converted into plastic work; such uncertainties can even cause $\beta_{diff}$ to be ostensibly higher than one over a few time intervals. Indeed, fluctuations of this kind are all too familiar in MD simulations with finite numbers of atoms. For the systems of the size

considered here, a smoothing operation is applied to the differential TQC that entails a moving average in a 20 ps window iteratively applied for 10 passes. The rapid increase of the integral and differential TQC during and after yielding indicates that initially a large fraction of plastic work is expended in generating new crystal defects (storage), and that this fraction gradually decreases with increasing plastic strain as the system approaches a state of stationary flow.

## 3.4 Discussion

Simulations A and B are generally consistent with each other and with other similar large-scale simulations reported elsewhere [84, 86, 87]. As has been observed in all simulations performed at comparable strain rates, the flow stress reaches a plateau at true strains in the range of 0.3 to 0.5 (Fig. 3.5). The flow stress in simulation B has nearly plateaued when the end of ramping is reached at a true strain of 0.4 and remains nearly constant from a true strain of ~0.5 to the end of the simulation. The plateau stress is higher than in simulation A because the final strain rate of $4 \times 10^8$ s$^{-1}$ in simulation B is twice as high as that in simulation A.

### 3.4.1 Comparison between TQC measures

Simulations A and B confirm earlier observations that, when subjected to straining under conditions that remain constant for a sufficiently long time, a plastic crystal can attain a state of stationary flow [84]. In such a stationary state the dislocation microstructure itself becomes stationary but only in a statistical sense; although mutually balanced, dislocation multiplication and annihilation continue unabated resulting in the incessant rebuilding and dissolution of a dynamically evolving dislocation network. Since the total defect density eventually reaches a stationary state, $\beta_{\mathrm{diff}}$ must be 1.0 in stationary flow—to the degree the defect density is unvarying—since all of the plastic work must therefore be dissipated as heat. The heat is generated primarily by dislocation motion converting the work of the Peach-Kohler force into the energy of lattice vibrations (phonons). Further plastic work is converted into phonons as a result of energy-reducing dislocation reactions such as dislocation annihilation (and to a lesser extent the formation of dislocation junctions). In stationary flow, the energy released as heat through dislocation annihilation is compensated by the additional defect energy accrued in the crystal by dislocation multiplication, leaving the net defect energy invariant.

This is fully consistent with data presented in Fig. 3.6 showing that, indeed, on reaching the true strain where the flow stress reaches its plateau, $\beta_{\mathrm{diff}}$ reaches and remains close to 1.0 within the fluctuations expected in our simulations. Although not obvious from Fig. 3.6, the average of $\beta_{\mathrm{diff}}$ is actually slightly below 1.0 as a

consequence of the point defect density continuing to rise past the apparent start of stationary flow, discussed further in Sec. 3.4.3. By contrast, $\beta_{\text{int}}$ never reaches but only asymptotically approaches 1.0 from below. This is a direct consequence of the relationship in Eq. 3.6 since, if $\beta_{\text{diff}}$ is smaller than 1.0 at any point and does not rise above 1.0, $\beta_{\text{int}}$ necessarily remains smaller than 1.0 for all strains. As occurred in simulations A and B, there is often appreciable energy storage at the early stages of straining that is integrated into $\beta_{\text{int}}$, making this parameter rather insensitive to the subsequent straining trajectory. For this reason it is tempting to advocate, after Ref. [13], that $\beta_{\text{diff}}$ is the more informative of the two parameters since it provides information about the instantaneous state of the system. On the other hand, $\beta_{\text{diff}}$ can be difficult or even impossible to extract during experiments, whereas $\beta_{\text{int}}$ gives a simple measure of thermo-mechanical conversion under straining. Regardless, their relationship in Eq. 3.6 makes clear that the two quantities contain precisely the same information, simply in different forms. Perhaps a more pertinent issue is that a value of the TQC (whether differential or integral) obtained in one experiment or simulation is generally not applicable to experiments or simulations performed under different conditions, depending as they do on the material's initial defect microstructure as well as on the entire thermo-mechanical trajectory followed during the test.

Asymptotic convergence notwithstanding, our simulations indicate that $\beta_{\text{int}}$ remains markedly lower than 1.0 up to rather large plastic strains in situations where dislocation multiplication is the primary energy storage mechanism. For instance, $\beta_{\text{int}}$ is still below 0.9 at a plastic strain of 0.5 for both simulations in Fig. 3.6. This is at variance with Ref. [13] which found that $\beta_{\text{int}}$ quickly converged to 1.0 after plastic yield in MD simulations involving dislocations. They concluded that energy storage due to dislocation multiplication could not explain experimental observations of $\beta_{\text{int}}$ less than 1.0, and that some other energy storage mechanism must be responsible. It is significant to this discussion that the simulations in Ref. [13] used an initial condition consisting of a single edge dislocation dipole, effectively precluding dislocation multiplication which is the key energy storage mechanism in crystal plasticity. By comparison, dislocation multiplication takes place naturally in our simulations in a statistically representative ensemble of interacting dislocations. Indeed, our results suggest the opposite conclusion, that energy storage due to dislocation multiplication is substantial and should be accounted for when assessing thermo-mechanical conversion during high-rate straining experiments. As discussed in more detail in Sec. 3.4.3, the total energy additionally stored in the defect microstructure in simulations A and B is 17 and 22 meV/atom, respectively, or roughly 5% of the total plastic work performed on the crystals in the same simulations. Assuming that the specific heat capacity remains unchanged at 0.135 J/(gK), this amounts to about 80 K of additional temperature rise if all of the stored energy were to be released as heat. By comparison, an estimate given in Ref. [13] for the amount of energy that could be stored by a grain refinement mechanism is only $2.5 \times 10^3$ J/m$^3$ =

$0.28 \times 10^{-3}$ meV/atom, five orders of magnitude lower than what is observed in the two simulations reported here and hence entirely negligible on the energy scales involved with high-rate plasticity.

### 3.4.2 Phenomenological model

As discussed in Sec. 3.1, Taylor and Quinney conjectured in their original paper that the stored energy constituted a material state variable [8]. They did this without knowledge of the defect microstructure, and while our understanding of such defects has developed extensively in the intervening years, there remains an appealing simplicity to modeling the evolution of the energy stored in the defect microstructure without concern for the specifics of the defect evolution processes (a subject that will be explored in Sec. 3.4.3).

Let $E_{\text{stored}} = U - U_{\text{ref}} - E_{\text{el}} - C_V \Delta T$ be defined as the energy stored in the defect microstructure, where $\Delta T$ is the deviation from the reference temperature (300 K) and $U_{\text{ref}} = N_{\text{at}} E_{\text{coh}}$ where $N_{\text{at}}$ is the number of atoms in the crystal and $E_{\text{coh}}$ is the per-atom energy in a perfect crystal at the reference temperature. Suppose that there is a limiting value of $E_{\text{stored}}$ in a material that is continuously deformed in unchanged straining conditions for a sufficiently long time; this value, indicated by $E_\infty$, is a material function that could in principle be computed or measured as a function of temperature, pressure and straining rate. The quantity $E_\infty - E_{\text{stored}}$ is then the remaining capacity of the material to store energy in the defect microstructure given a prescribed deformation protocol. Suppose further that the material stores in the defect microstructure a fraction of the incremental plastic work that is proportional to the remaining energy storage capacity, or

$$dE_{\text{stored}}/dW_p = \zeta(E_\infty - E_{\text{stored}}), \tag{3.7}$$

where $\zeta$ is the constant of proportionality (conjectured to be a material constant). The solution of this differential equation for the stored energy as a function of plastic work is

$$E_{\text{stored}}(\epsilon_p) = E_\infty - (E_\infty - E_0) \exp[-\zeta W_p(\epsilon_p)] \tag{3.8}$$

where $E_{\text{stored}}$ and $W_p$ are both taken to be functions of plastic strain $\epsilon_p$ and $E_0$ is the energy stored in the defect microstructure at the beginning of deformation ($\epsilon_p = 0$). The definition in Eq. 3.4 then implies the following explicit form for the differential TQC:

$$\beta_{\text{diff}}(\epsilon_p) = 1 - dE_{\text{stored}}/dW_p$$
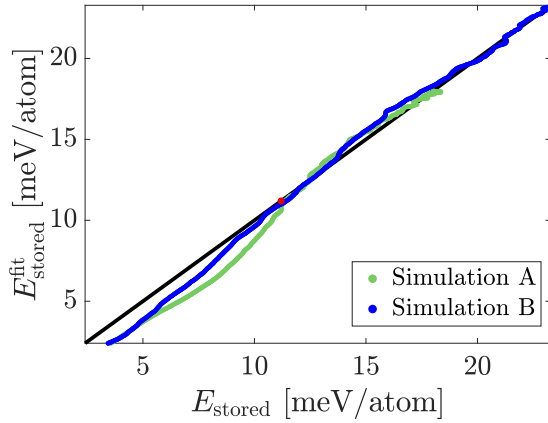$$= 1 - \zeta(E_\infty - E_0) \exp[-\zeta W_p(\epsilon_p)]. \tag{3.9}$$

Figure 3.7: Comparison of the predictions of Eq. 3.8, $E_{\text{stored}}^{\text{fit}}$, with the energy stored in the defect microstructure, $E_{\text{stored}}$, for both simulations. The red marker roughly indicates the point where steady-state plastic flow begins in simulations A and B. The overall RMSE is 0.434 meV/atom.

Finally, substituting this into Eq. 3.6 and performing the integration there gives the following explicit form for the integral TQC:

$$\beta_{\text{int}}(\epsilon_p) = 1 - (E_\infty - E_0)\frac{1 - \exp[-\zeta W_p(\epsilon_p)]}{W_p(\epsilon_p)}. \tag{3.10}$$

Provided both $E_\infty$ and $E_0$ are known, the simple model expressed in Eqs. 3.9 and 3.10 involves a single fitting parameter $\zeta$ defining the rate of energy storage as a function of plastic strain. Here a nonlinear least square procedure is used to fit to the stored energy data $E_{\text{stored}}^{\text{sim}}$ extracted from the two MD simulations and both $\zeta$ and $E_\infty$ regarded as fitting parameters.

The resulting values for simulation A are $\zeta = 12.9$ and $E_\infty = 18.1$ meV/atom, and for simulation B are $\zeta = 8.78$ and $E_\infty = 24.1$ meV/atom. Inserting these values into Eq. 3.8 and using the value of the initial energy stored in the defect microstructure for $E_0$ gives the values of $E_{\text{stored}}$ plotted in Fig. 3.7. The accuracy of the model noticeably improves for both simulations after steady-state plastic flow is reached (indicated by the red marker in the figure), likely as a consequence of the model not being able to capture the complex processes of defect multiplication around the yield point. The same parameters were used to evaluate the predictions of Eqs. 3.9 and 3.10 for $\beta_{\text{diff}}$ and $\beta_{\text{int}}$, with the results shown in Fig. 3.8. The model agrees reasonably well with the values of TQC extracted directly from the two MD simulations, with the most significant deviations occurring right after yielding for plastic strains of less than 0.1; this is the same interval where the fluctuations in the observed values of the TQC are largest.

Fundamentally, the TQC reflects the complexity of microstructural evolution [9] and should depend on the straining rate [10], crystal orientation [11], and even loading mode [12]. A phenomenological model of the kind described here, while not resolving any relevant details of the microstructure evolution, offers
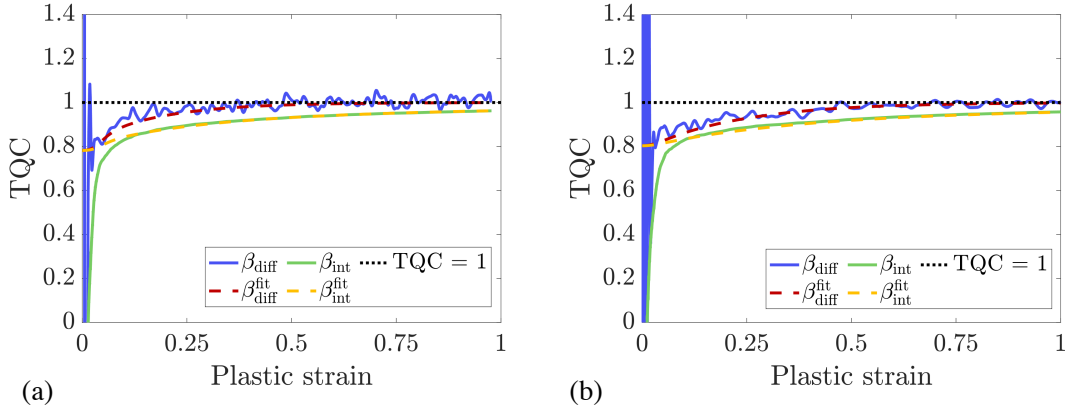
47

Figure 3.8: Comparison of the predictions of Eqs. 3.9 and 3.10 with the values of $\beta_{\text{diff}}$ and $\beta_{\text{int}}$ plotted in Fig. 3.6 for (a) simulation A and (b) simulation B.

an approximate quantitative description that goes beyond the current practice where TQC is treated as a material constant. Indeed, at the expense of one (or perhaps two) additional adjustable parameters, such a model is likely to predict with better accuracy the rate of heat release and the resulting rise in temperature as a function of plastic strain. Again assuming that parameter $\zeta$ is a material constant, the kinetics of energy storage could be described over complex straining trajectories that entail variations in pressure, temperature and loading mode. In such situations, the asymptotic stored energy $E_\infty$ could be viewed as a moving target to which the instantaneous energy $E_{\text{stored}}$ relaxes, e.g., according to Eq. 3.7. Given the dependence of material response on temperature, this would considerably reduce experimental uncertainties about material behavior under extreme straining conditions.

### 3.4.3 Stored energy model

Having defined stored energy as a measurable macroscopic quantity, the purpose of this section is to identify microscopic defects that actually contribute to $E_{\text{stored}}$. Here we show that, although the majority of the stored energy is attributed to dislocations, the dislocation contribution alone does not account for all of the energy storage in our MD simulations. It is significant that an early review on the TQC included the contributions of point defects, along with dislocations, when estimating the stored energy of cold work [70]. Indeed, our MD simulations indicate that point defects including vacancies, interstitials, and their clusters contribute substantially to the energy stored in the defect microstructure. The concentration of point defects generally increases with increasing straining rates and dislocation density. The defects are produced largely as debris in the wake of moving screw dislocations dragging cross-kinks and/or jogs formed by dislocation intersections [88] [89]. Moreover, in the absence of internal interfaces or other sinks, point
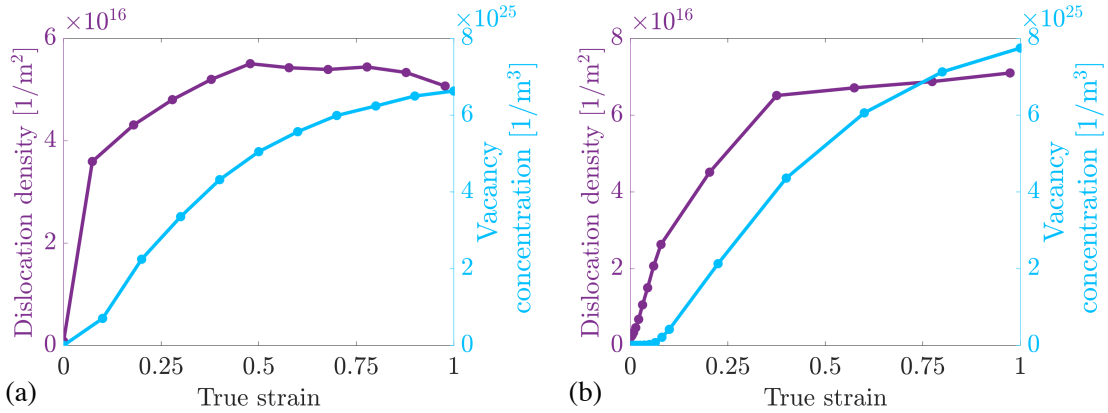
48

Figure 3.9: Dislocation density and vacancy concentrations as functions of true strain for (a) simulation A and (b) simulation B.

defect recombination and absorption by dislocations are the two most likely mechanisms of point defect removal in the crystal bulk.

The copious point defect debris in our high strain-rate simulations is mostly single vacancies, though single interstitials and interstitial clusters also occur. The BCC Defect Analysis (BDA) method [90] was used to count the vacancies produced in our simulations. The same BDA method is inaccurate in counting interstitials, even though we were able to visually identify multiple interstitials and interstitial clusters using various functionalities of the OVITO software [91]. Visually, interstitials appear to be less numerous than vacancies, but their contribution to energy storage should not be discounted because the formation energy of a single interstitial in the Ta model employed here is 7.54 eV, nearly three times that of a vacancy at 2.75 eV. Here we shall assume that energy stored in the interstitials is directly proportional to the energy stored in the vacancies at all times.

Shown in Fig. 3.9a is a plot of the vacancy concentration in simulation A as a function of strain along with the evolution of the dislocation density extracted using DXA [92, 93]. At a strain of ~0.5 where the dislocation density appears to have saturated, the vacancy concentration is still rising and has not fully saturated even by the end of simulation A at a strain of 1.0. The same trend can be seen in Fig. 3.9b pertaining to simulation B. This asynchronicity could reflect the fact that, while the rate of debris production depends only on dislocation density, the rate of debris removal should also depend on the point defect concentration itself. Thus, for the rate of debris removal to balance the rate of debris production by dislocations, the debris may need to reach sufficiently high concentrations that are not fully attained even at a strain of 1.0. Assuming that at any given time the energy of stored interstitials is proportional to the energy of stored vacancies, the contribution of the debris to the stored energy is modeled by multiplying the
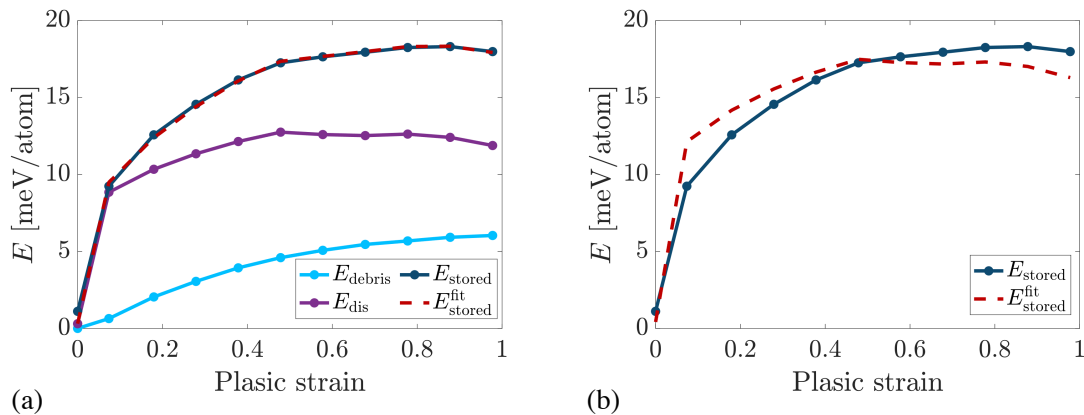
49

Figure 3.10: Stored energy model for simulation A as a function of true plastic strain. Including point defects in the fitting (a) gives $\chi = 1.26$ and $\xi = 1.73$, whereas excluding point defects (b) gives $\chi = 1.72$.
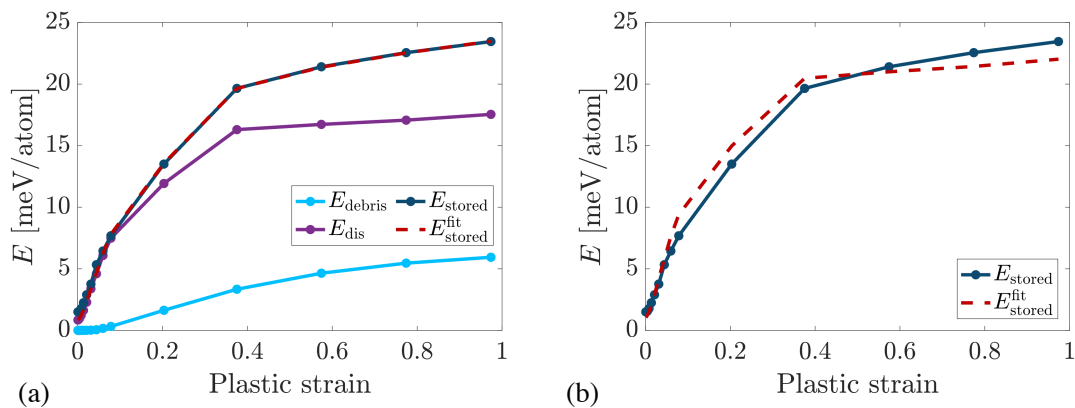


Figure 3.11: Stored energy model for simulation B as a function of true plastic strain. Including point defects in the fitting (a) gives $\chi = 1.40$ and $\xi = 1.55$, whereas excluding point defects (b) gives $\chi = 1.75$.

energy stored in vacancies by a constant factor $\xi > 1$:

$$E_{\text{debris}} = \xi N_{\text{vcy}} E_{\text{vcy}} \tag{3.11}$$

where $E_{\text{vcy}}$ is the vacancy formation energy and $N_{\text{vcy}}$ is the number of vacancies counted using the BDA algorithm.

The energy stored in the dislocation network is modeled following Ref. [94]:

$$E_{\text{dis}} = \eta \chi \frac{\mu b^2}{4\pi} \rho \ln\left(\frac{1}{r_c \sqrt{\rho}}\right) \tag{3.12}$$

which includes the linear dependence of the elastic energy on the dislocation density $\rho$ and a logarithmic term that accounts for the screening and core cutoff $r_c$ of the dislocations' stress fields. Here $\mu$ is the shear modulus, $b = 0.286$ nm is the Burgers vector magnitude, and $\chi$ and $\eta$ are parameters on the order of unity that are discussed below and in more detail in the Appendix. Although this model was initially introduced to describe the energy associated with arrays of infinite straight dislocations, it has been recently shown to accurately capture the energetics of complex and realistic dislocation networks [94].

To more accurately account for the energy stored in dislocations, Eq. 3.12 is adjusted from Ref. [94] as follows. First, a factor $\eta$ is introduced to correct the isotropic model proposed in Ref. [94] for the elastic anisotropy of the Ta model used in our MD simulations. As described in the Appendix, this is estimated to be $\eta \simeq 1.35$ at 300 K. Second, a factor $\chi$ is used as a fitting parameter to account for variations in the stored energy as a consequence of the precise geometric arrangement of the dislocation lines. Although the physical meaning of $\chi$ is subtle, as was shown in Ref. [94] and is further discussed in the Appendix, physically reasonable values of this parameter are bound to the interval between 1.0 and $1/(1 - \nu) \simeq 1.5$. The cutoff radius is set to $r_c = 0.5b$ as a coarse estimate for dislocation core energies reported for the same model of Ta elsewhere [95]. More details on the parameters $\eta$, $\chi$, $r_c$, and their relationships are given in the Appendix C.

Assuming that the point defect debris and the dislocations are the only contributions to the energy of the defect microstructure gives

$$E_{\text{stored}} = E_{\text{dis}} + E_{\text{debris}} \tag{3.13}$$

where the coefficients $\xi$ and $\chi$ in Eqns. 3.11 and 3.12, respectively, are treated as fitting parameters to describe the variations in the stored energy data $E_{\text{stored}}$ extracted from our MD simulations. Results for the fit obtained with Eq. 3.13 are shown in Fig. 3.10 a for simulation A, with $\chi = 1.26$ and $\xi = 1.73$ providing a remarkably good fit to the energy storage data. The value of $\xi > 1$ indicates that the contribution to the stored energy from the unknown population of interstitials is a significant fraction of that of the vacancies,

while the value of $\chi = 1.26$ falls within the expected interval defined elsewhere [94]. Shown in Fig. 3.10b for comparison is a best fit at $\xi = 0$ ignoring the contribution of the point defects and accounting only for the dislocation network energy. This results in a best fit value $\chi = 1.72$ which is outside the acceptable range of 1.0 to 1.5 for this parameter. In comparison to Fig. 3.10a, the agreement between the "dislocations only" model and the MD data is rather poor which additionally supports our assessment that the dislocation network alone cannot reasonably account for the total stored energy. Thus, we find that point defects provide a significant mechanism for energy storage, with their contribution potentially half that of the (very dense) dislocation network at strains of 1.0 in Fig. 3.10a.

Similar results were obtained for simulation B in Fig. 3.11a for which the best fit values are $\chi = 1.40$ and $\xi = 1.55$. Accounting only for the dislocation network energy in Fig. 3.11b instead yields a higher value of $\chi = 1.75$ which, just as in the case of simulation A, falls outside the physically admissible range. That the best fit parameters have slightly different values for simulations A and B likely reflects differences in their straining trajectories and ultimate straining rates.

## 3.5    Conclusion

Taylor and Quinney defined a thermo-mechanical conversion factor (TQC) as a measure of the energy stored in a material undergoing plastic deformation. This continues to be a quantity of enduring interest that can be and has been experimentally measured at low and moderately high deformation rates. That said, fully dynamic high-rate simulations of the kind reported here can be used to quantify thermomechanical conversion under extreme deformation conditions where experimental measurements are difficult or impossible.

Since the original study of Taylor and Quinney, advances in microscopy and material characterization have offered a variety of ways to observe and quantify changes in the defect microstructure that contributes to energy storage during mechanical deformation. Likewise, the atomistic simulations reported here allow one to observe the mechanisms of energy storage during high-rate deformation *in silico* and in arbitrary detail.

Our results support the following conclusions:

- Under compressive deformation conditions held unchanged for sufficient strains, single crystal tantalum is predicted to asymptotically approach a state of steady flow in which the defect microstructure becomes statistically stationary. Energy storage ceases and the TQC approaches 1.0 in the same asymptotic limit.

- The asymptotic limit $E_\infty$ of stored energy is a measure of the material's energy storage capacity and

can be computed and tabulated as a function of the straining conditions. It is used here to define a phenomenological model of energy storage kinetics potentially applicable to complex deformation histories.

- Dislocation multiplication is a major energy storage mechanism in the high-rate deformation conditions studied here. However, point defect debris is also produced in copious quantities and contributes as much as 1/2 of the energy stored in the dislocation network. Dislocation multiplication and debris production together account for the TQC staying measurably below 1.0 up to large compressive strains.

- Although minor, energy storage due to debris production continues well beyond the strains at which dislocation multiplication ceases to contribute to energy storage.

## Author contributions

JCS and NB performed all simulations. JKM and VVB developed the models and concepts. All authors analyzed simulation results and wrote the paper.

# Chapter 4

# Dependence of simulated radiation damage on crystal structure and atomic misfit in metals

The following chapter is comprised of a manuscript that was a joint effort of C. Serrao, J. C. Stimac, and J. K. Mason.

## 4.1 Introduction

Structural components in nuclear fission reactors need to be engineered to withstand decades of exposure to radiation and elevated temperatures [96, 97]. Materials in the next generation of fission reactors and future fusion reactors will be subject to substantially higher radiation dosages and temperatures and likely highly corrosive environments as well [5, 6]. Irradiation by exposure to high-energy particles displaces atoms and damages the crystalline structure of pure metals and alloys [14, 98], resulting in a variety of microstructural changes. Specifically with respect to the degrading effects of radiation, there are five considerations for structural metals in reactor environments: radiation hardening (low temperature), radiation-induced segregation and precipitation, void swelling, radiation-induced creep, and helium embrittlement (high temperature) [99, 7]. The three intermediate temperature effects are usually the most relevant in practice, and are often observed simultaneously since they are all strongly associated with the underlying ability of point defects generated by radiation to migrate through the lattice [100, 101].

Radiation-induced segregation of substitutional solutes is generally attributed to the inverse Kirkendall

effect where, as vacancies migrate to and annihilate on sinks, the different atomic species have different migration rates in the opposite direction [102]. The segregation extent is governed by differences in diffusive mobility, with undersized solutes generally being enriched and oversized solutes being depleted in the vicinity of the sinks [103, 101]. Void swelling is a serious engineering concern that entails an increase in the volume of the material by as much as 1% per dpa [104] from the nucleation and growth of voids in the bulk. This process is driven by an imbalance in the concentrations of vacancies and interstitials, with annihilation of high-mobility interstitials on dislocations and other sinks leaving behind a relative excess of vacancies that precipitate as voids [100, 105]. While a variety of mechanisms have been proposed for irradiation creep, the two main mechanisms are believed to be the climb of favorably-oriented dislocations by the stress-induced preferred absorption of point defects and the glide of dislocations enabled by climb over obstacles [106, 100].

Multi-principal component alloys (MPEAs) [107, 108] (often called high-entropy alloys or compositionally complex alloys) consist of a few to several atom types in solid solution with nominally equi-atomic concentrations. This emerging class of metals could perform well as structural materials in irradiated environments, with initial evidence showing exceptional mechanical properties [109, 18] and higher resistance to radiation damage [110, 111] than traditional metals. One experimental and computational study of the effects of radiation on a CrCoNi MPEA showed decreases in the relative disorder, the number of defects, and large defect clusters compared to irradiated samples of pure Ni and a NiFe binary alloy [110]. The authors attributed these effects to decreasing dislocation mobility with increasing number of atomic species. Lu et al. [112] studied void swelling in several Ni-based alloys including pure Ni, NiFe, CrCoNi and two quinary alloys with and without prior nanoindentation, and found that nanoindentation improved resistance to swelling by increasing the density of defects like dislocations, stacking faults, and twin boundaries that promote vacancy annihilation. Curiously, Velişa et al. [113] found that CrCoNi only showed superior irradiation resistance relative to NiCr only at temperatures near and below 300 K though. The reason for this is not well established, but is believed to be related to the specifics of the chemical short range order (SRO) that developed in the two alloys.

The most direct way to simulate radiation damage at the atomic level uses molecular dynamics (MD) simulations of collision cascades [98, 14]. This involves assigning a large initial velocity to the primary knock-on atom (PKA) to mimic a passing neutron or other high-energy particle transferring kinetic energy to the lattice. The PKA recoils, displacing many of the surrounding atoms from their lattice sites and converting the initial kinetic energy into a thermal spike with sufficient energy to facilitate the regeneration of the crystalline lattice and the recombination of many, but not all, of the generated point defects. The resulting interstitials and vacancies are respectively distributed on the periphery and the interior of the affected zone, increase the point defect concentration in the material, and are directly responsible for the most visible

degrading effects of radiation.

Mass conservation requires that the interstitials and vacancies generated by a single collision cascade occur as Frenkel pairs. The accumulated number of such Frenkel pairs divided by the number of atoms in the material is known as the displacements per atom (dpa), and is the standardized measure of the extent of radiation damage in crystalline materials [114]. Early work by Kinchin and Pease modeled atoms as hard-spheres that exhibit elastic collisions during collision cascades and laid the theoretical foundation for radiation effects in crystalline materials [115]. If the energy transferred to an atom exceeded a material-specific threshold value, then the atom was said to have been displaced from its lattice site. While energies below the threshold could still displace the atom, it would return to its lattice site after the initial perturbation. The theory developed by Kinchin and Pease was further developed in the work of Norgett, Robinson, and Torrens (hereafter referred to as the NRT model) who added additional terms to account for energy lost to ionization and for the affects of inelastic collisions [116]. Several more recent studies concluded that the NRT model overestimates the number of defects generated by a collision cascade and neglects the mixing from atomic replacements though, spurring a number of proposed refinements [98, 14]. These are significant for the reason that accurate damage models that can reliably predict dpa are essential to reliably compare radiation damage resistance among various materials.

Full-scale atomic simulations of collision cascades have been performed for several decades now, and are useful to uncover the evolution of radiation-induced primary damage [15, 117]. However, the use of MD for this application is subject to several limitations. As Ref. [98] points out, interatomic potentials cannot capture the effects of deviations from the Born-Oppenheimer approximation when excited electronic states are induced. The other main limitations are the time and length scales that can reasonably be achieved with modern computational resources. MD simulations are usually no longer than a few nanoseconds, limiting the overall radiation dose that can reasonably be achieved by full cascade simulations without the events overlapping in time; Refs. [118, 119] further discuss these limitations. As a result of the high computational cost to reach appreciable dpa, the doses investigated in MD simulations have been historically been less than 1.0 dpa. Although these are useful to understand defect creation at low doses, structural materials in a nuclear reactor core can experience as much as 80 dpa over a 40 year service life [7]

A variety of strategies have been used to circumvent the computational limitations imposed by cascade simulations, one of which involves the direct insertion of a high density of Frenkel pairs [120, 121, 122]. Such Frenkel pair accumulation (FPA) techniques forgo the dynamics of time-resolved high-energy atomic collisions stemming from the primary knock-on event. Instead, they intermittently introduce Frenkel pairs by randomly displacing atoms from their lattice positions, usually followed by some form of equilibration and time integration. These methods benefit from a clearly defined radiation dose, based on the the number

of displaced atoms, and dose rate, based on the ratio of displaced atoms to the simulated time. For example, Chartier et al. [123] used a FPA procedure to model irradiation of $UO_2$ and found the steady-state dislocation density to be in good agreement with experiments. A major limitation of FPA procedures though is the absence of any effects related to the thermal spike [119], particularly providing the thermal energy necessary for diffusion and clustering of defects. This notably includes the recombination of point defects when vacancies and interstitials collide. Analysis of FPA simulations should be done with careful consideration of this limitation to avoid the potential for nonphysical extrapolation.

Recent work by Derlet and Dudarev [118] introduced a variant of the FPA method that further streamlines the process of sampling irradiated microstructures. Known as the creation-relaxation algorithm (CRA), this differs from preceding FPA methods in that there are no time integrated dynamics. CRA simulations randomly select atoms and displace them with random directions and magnitudes, just as other FPA methods do, but always follow this with an energy minimization procedure. The entire simulation involves repeating this process for a specified number of displacements, with the canonical dpa equal to the number of displaced atoms divided by the total number of atoms in the system. Dudarev and Derlet applied the CRA to BCC Fe systems of a variety of sizes and reported good agreement between full cascade simulations [124, 125] and the CRA for interstitial density as a function of dpa, at least up to a linear rescaling of both the independent and dependent variables. The need for rescaling is likely related to the CRA effectively being performed at zero Kelvin, as with other FPA methods [126]. The thermally-driven diffusion of point defects that is prevalent at high temperatures and substantially contributes to microstructure evolution of irradiated materials is negligible in such conditions. That said, the CRA can be viewed as simulating radiation damage in conditions where thermally-driven diffusion is active but negligible compared to other mass transport mechanisms. For example, Dudarev and Derlet observed evolution of the microstructure resulting from fluctuating atomic-level stresses in the simulation cell as defects were introduced.

The CRA has since been used to investigate radiation effects in materials other than BCC Fe. One such study used the CRA to simulate irradiation of a NiFe system doped with C and evaluated the ability of C interstitials to decrease radiation damage [127]. Several others applied the CRA to W to investigate the relationships between dpa and specific physical parameters at relatively high doses (above 1.0 dpa), either alone or with the assistance of other computational or experimental methods. Parameters that were investigated include thermal conductivity [128], hydrogen embrittlement and tritium concentration [129], and radiation induced structural evolution [130].

This paper investigates the microstuctures of highly irradiated Fe, equi-atomic CrCoNi, and a fictitious metal with identical bulk properties to the CrCoNi composed of a single atom type referred to as an A-atom. The main motivation for including the Fe and A-atom systems is to establish points of comparison for the

investigation of the reported radiation resistance of CrCoNi. The CRA is used to simulate the irradiation of all three systems up to a final dose of 2.0 dpa. The details of the implemented CRA, as well as an analysis of the experimental dose rates and temperatures for which it is likely relevant, are described in Sec. 4.2. The same section also outlines our methods for identifying material defects including dislocations, vacancies, interstitials, and stacking faults, and a model for the energy stored in those defects. The results and a discussion of the simulations are included in Sec. 4.3, and Sec. 4.4 draws conclusions to inform further research in this area.

## 4.2  Methods

### 4.2.1  CRA simulations

All molecular dynamics simulations were performed using the LAMMPS software [82]. Orthorombic simulation cells were used with periodic boundary conditions for all cell faces, and the number of atoms remained constant. No temperature or time-steps were defined for any simulation. After constructing the initial space-filling single crystals, the volume of the simulation cell was relaxed using a potential energy minimization and fixed thereafter. The present study examined three material systems: BCC Fe, equi-atomic FCC Cr-CoNi, and FCC A-atom designed to reproduce the bulk properties of CrCoNi using a single fictitious atom type [131]. Comparison with an A-atom model more directly allows identification of the effects caused by chemical short range order (SRO) and lattice distortion (LD) which are widely implicated in the enhanced physical properties of MPEAs [18].

The BCC Fe simulation consisted of a simulation volume of 40 units cells along each of the three dimensions, and with two atoms in each unit cell contained a total of $128,000$ atoms. This simulation used the Mendelev-II embedded atom method (EAM) potential [132]. The FCC CrCoNi simulation consisted of 32 unit cells along each of the three dimensions, and with four atoms in each unit cell contained a total of $131,072$ atoms. All three atomic species were represented with equal concentrations, and were initially distributed uniformly at random throughout the simulation cell. The propose of reducing the number of unit cells for the FCC systems was to make the number of atoms as close as possible to that in the Fe simulation. The equiatomic CrCoNi simulation utilized the EAM potential developed by Li et al. [133]. The FCC A-atom simulation used the EAM potential offered by Jian et al. [134] but was otherwise identical to the CrCoNi simulation.

This work circumvented the computational limitations associated with full collision cascade simulations by using the creation relaxation algorithm (CRA) [118] to simulate the effects of irradiation by the repeated

**Algorithm 1** Creation Relaxation Algorithm

---
1: $n := 0$                                     ▷ number of displaced atoms
2: **while** $n/N < \phi_{\mathrm{dpa}}$ **do**                       ▷ $N$ is total number of atoms
3:      Randomly select an atom $i$ uniformly over all atoms
4:      Move atom $i$ to a randomly selected point within the simulation cell
5:      **while** Potential energy increase exceeds a threshold **do**
6:          Move atom $i$ to a randomly selected point within the simulation cell
7:      **end while**
8:      Relax the system using energy minimization
9:      $n \leftarrow n + 1$
10: **end while**

---

introduction of Frenkel pairs [135, 123]; the basic CRA is described in Alg. 1. All three experiments used the CRA to simulate radiation damage up to 2.0 dpa, requiring a total of $256,000$ atomic displacements in the Fe system and $262,144$ displacements in both the CrCoNi and A-atom systems. The potential energy minimization conducted after each atomic displacement consisted of two steps. First, the position of every atom was randomly perturbed by a small Gaussian-distributed displacement, and second, a standard gradient-based energy minimization was applied to the atomic positions to reduce the system energy. The purpose of the small perturbations in the first step was to help the system to escape shallow local energy minima as would naturally occur in a thermal system. The perturbation magnitude was fixed after systematically exploring the relationship between the perturbation magnitude and the relaxed potential energy of a smaller BCC Fe system of $2,000$ atoms. This involved displacing 100 atoms with energy minimization but without perturbations between displacements. After all the displacements, the atoms were subjected to repeated cycles of perturbations and energy minimization, with the resulting potential energy profiles for a given standard deviation of the perturbation magnitude shown in Fig. 4.1 as a function of cycle number. Standard deviations of 0.005 and 0.01 angstroms were found to rapidly result in a relatively stable potential energy minimum, with larger displacements often introducing additional defects and smaller displacements unable to reliably allow the system to escape shallow potential energy minima. A perturbation magnitude of 0.01 angstroms was chosen for the subsequent simulations for the reason that it reached the potential energy minimum in the fewest number of cycles.

Since the intention was for the perturbations to make the system more closely resemble a thermal one, a natural question is what would be the corresponding temperature for a given perturbation magnitude. The equipartition theorem implies that the potential energy per atom relative to that in an ideal crystal should be $(3/2)k_{\mathrm{B}}T$ for a thermal system; equating this with the average potential energy increase per atom resulting from a single perturbation gives an equivalent temperature of ~30 K for all three systems. This implies that the additional energy supplied by the perturbations is consistent with the assumption
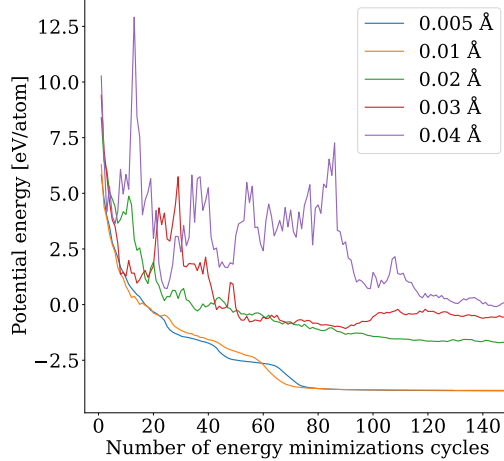
Figure 4.1: The absolute potential energy per atom of a highly defected BCC Fe system as a function of minimization cycles. Perturbation magnitudes of 0.005 and 0.01 angstroms allow the system to quickly reach a deep potential energy minimum.

that any thermally-driven diffusion of point defects is insignificant compared to alternative mass transport mechanisms.

### 4.2.2 Applicable temperatures and dose rates

As mentioned in Sec. 4.1, the microstructures that develop when using the CRA are a direct product of relaxations of the atomic-level stress fields, and do not necessarily represent microstructures that develop in conditions with appreciable thermally-driven diffusion. We characterize the regime of physical conditions for which the CRA is applicable by a simple argument that compares the rates of atomic transport by thermally-driven diffusion and ballistic displacements from collision cascades. The main criterion for applicability is that the rate of ballistic displacements per atom $K_0$ be much greater than the interstitial hopping rate $\gamma_{\mathrm{diff}}$ (assumed to be the fastest thermally-driven diffusion event). In such conditions, any effects of thermally-driven diffusion should be negligible compared to those resulting from ballistic displacements and the subsequent stress relaxation. The interstitial hopping rate, as described by transition state theory, obeys the Arrhenius relation $\gamma_{\mathrm{diff}} = \nu_0 \exp\left(-E_{\mathrm{a}}/k_b T\right)$ where $\nu_0$ is the high-temperature limit for the site-hopping frequency, $E_{\mathrm{a}}$ is the activation energy, $k_b$ is the Boltzmann constant, and $T$ is the temperature. The random walk diffusion model suggests that $\nu_0$ be approximated by $6D_0/\lambda^2$ where $D_0$ is the diffusivity prefactor and $\lambda$ is the distance separating two interstitial sites [136].

Using the values for the diffusivity and activation energy of CrCoNi given by Ref. [137] leaves only the temperature and $K_0$ as independent variables. We now define a measure of applicability $\mathcal{A} = \log_{10}(K_0/\gamma_{\mathrm{diff}})$ such that a factor of one increase in $\mathcal{A}$, indicates an order of magnitude increase in the ratio of $K_0$ to $\gamma_{\mathrm{diff}}$.
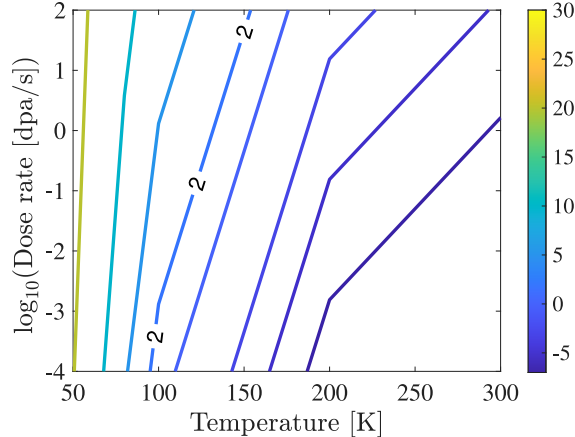
60

Figure 4.2: Contour plot of the applicability measure $\mathcal{A} = \log_{10}(K_0/\gamma_{\mathrm{diff}})$ as a function of the logarithm of the dose rate and temperature. The calculations used the diffusivity and activation energy of CrCoNi given by Ref. [137].

A contour plot of $\mathcal{A}$ for CrCoNi is included as Fig. 4.2 and indicates that applicability increases at lower temperatures and higher dose rates; moreover, the dependence on the temperature is much more significant than the dose rate, a consequence of the exponential scaling in the Arrhenius relation for the diffusivity. Across the range of dose rates considered, the temperatures are in the cryogenic range for a tolerance of applicability $\mathcal{A} \geq 2$. This low temperature constraint is entirely consistent with the equivalent temperature resulting from the perturbations in the atomic positions described in Sec. 4.2.1.

### 4.2.3 Identifying defects

It is difficult to precisely identify crystalline defects in a material that has suffered extensive radiation damage. Certainly there is still an underlying lattice, but conventional defect models assume that the defect is isolated, or equivalently, that the surrounding atoms occupy well-defined lattice positions. This is not true for the high defect concentrations that can occur in irradiated materials, and identification of reference lattice sites is further complicated in MPEAs like CrCoNi where the reference lattice is already perturbed by atomic size differences.

The numbers and types of dislocations in the simulations were determined using OVITO's dislocation extraction algorithm (DXA) [91]; the algorithm requires a trial circuit length and a value for circuit stretchability which were set to 14 and 9 respectively. Surprisingly, the DXA appeared to be relatively robust to extensive disruption of the crystalline lattice from the CRA, with very few isolated dislocation segments appearing in the networks reported below in Sec. 4.3.2.

The number of vacancies were determined using OVITO's Wigner-Seitz (WS) analysis. This constructs Voronoi cells around the lattice sites in an initial crystal structure and checks the atom occupancy of each
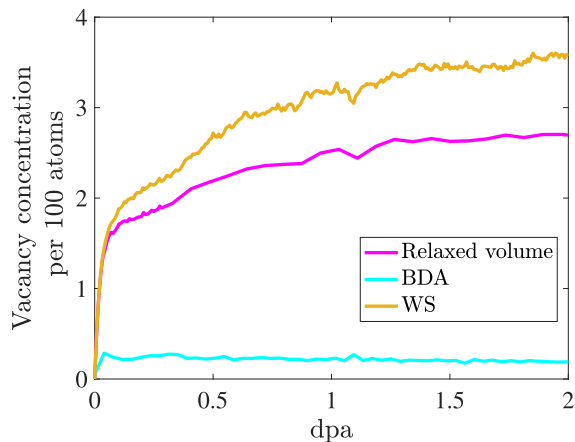
Figure 4.3: Three different estimates for the number of vacancies in BCC Fe as a function of dpa.

Voronoi cell in subsequent time steps. A Voronoi cell that does not contain any atoms is regarded as indicating the presence of a vacancy, though this approach does not precisely define the vacancy location. The WS analysis was compared with two other approaches to estimating vacancy concentration in the BCC Fe systems, the first of which was the BCC defect analysis (BDA) [90]. The second approach involved evaluating the relaxed volume of the simulation box. Inserting a single Frenkel pair produces an interstitial and a lattice site occupied by a vacancy, resulting in the expansion of a crystal subject to zero traction boundary conditions. Since the simulations were conducted at constant volume, the insertion of Frenkel pairs in the CRA instead elevated the system pressure. The equivalent volume change was evaluated by relaxing the simulation cell subject to a zero traction boundary condition, and enabled the number of vacancies to be estimated by assuming that the volume change from inserting a Frenkel pair remained constant throughout the simulation. Figure 4.3 compares the three methods and shows that the WS analysis agrees much better with with the vacancy concentration estimated from the relaxed volume than from the BDA. The factor by which the WS and relaxed volume method differ can be interpreted as the reduction in volume change per effective Frenkel pair insertion with increasing dpa. The poor performance of the BDA is likely a consequence of extensive radiation damage making it difficult to identify vacancies based on features of local atomic environments without a clear underlying crystal lattice.

Interstitials are the defect type that is most difficult to identify in our simulations. The multiplicity of local atomic configurations that can occur for non-isolated interstitials means that template-based approaches would be difficult or even infeasible to implement, and examining the expected number of atoms contained within a surface that passes only through crystalline material (the analogue of a Burgers circuit) often fails because of the obstructions to constructing such surfaces. In the absence of a canonical alternative, our approach involves constructing a distribution of atomic volumes as estimated by the Voronoi polyhedra.

The compressive stresses around an interstitial reduce the volumes of the interstitial and of the surrounding atoms in a characteristic way when the interstitial is in an otherwise perfect crystal. The distribution of atomic volumes in our simulations is decomposed with a K-means algorithm into a superposition of peaks, one for each characteristic atomic volume surrounding an isolated interstitial. Part of the utility of K-means is that the boundaries that define the locations the peaks can be dynamically updated from one time step to the next; this is necessary to account for the shifting and broadening of the peaks as the structure reaches higher defect concentrations. Comparing the numbers of atoms assigned to each peak with the corresponding numbers of atoms for an isolated interstitial gives an estimate for the number of interstitials in our simulations, though the accuracy is expected to decrease as the peaks broaden and overlap with increasing radiation damage.

Stacking fault (SF) densities in the FCC CrCoNi and A-atom systems are relatively simple to evaluate by comparison. Since the local atomic structure around atoms belonging to a SF appears to be HCP, all that is necessary for a reasonable estimate is to count the number atoms classified as HPC by OVITO's common neighbor analysis and to convert this to an equivalent SF area using lattice geometry.

The Warren-Cowley parameters are used to quantify the type and degree of chemical short range order (SRO) and are defined as [138, 139]

$$\alpha_{ij} = 1 - p_{ij}/c_j \tag{4.1}$$

where $p_{ij}$ is the probability that an atomic type $j$ is present in the first nearest-neighbor shell of an atomic type $i$, and $c_j$ is the overall concentration of atomic type $j$. Positive values of $\alpha_{ij}$ means there is a repulsion between species $i$ and $j$, while negative values indicate attraction.

### 4.2.4 Energy model

This section develops a model to partition the energetic contributions of the various types of material defects to the overall potential energy change as a function of dpa. The potential energy stored in defects is defined as $E_{\mathrm{stored}} = E_{\mathrm{pe}} - E_{\mathrm{coh}} - E_{\mathrm{elas}}$, where $E_{\mathrm{pe}}$ is the potential energy, $E_{\mathrm{coh}}$ is the cohesive energy, and $E_{\mathrm{elas}}$ is the elastic strain energy. This stored energy is modeled as a sum over defect contributions:

$$E_{\mathrm{fit}} = E_{\mathrm{dis}} + E_{\mathrm{vcy}} + E_{\mathrm{int}} - E_{\mathrm{sro}} \tag{4.2}$$

which includes terms for the contributions of dislocations, vacancies, interstitials, and chemical short range order, respectively. The contribution of stacking faults to $E_{\mathrm{fit}}$ was found to be negligible.

The energy of a well-developed dislocation network in an elastically isotropic material is well described

Table 4.1: Material properties used as parameters to evaluate the dislocation energy.

| | $\mu$ [Gpa] | $b$ [A] | $a_0$ [A] | $\nu$ |
|---|---|---|---|---|
| Fe | 49.5 | 2.49 | 2.87 | 0.373 |
| CrCoNi | 37.0 | 2.47 | 3.50 | 0.414 |
| A-atom | 43.0 | 2.47 | 3.50 | 0.403 |

by the equation [94]:

$$E_{\mathrm{dis}} = \chi \frac{\mu b^2}{4\pi} \rho \ln \left( \frac{1}{r_c \sqrt{\rho}} \right) \tag{4.3}$$

where $\mu$ is the shear modulus, $b$ is the Burgers vector, $r_c$ is a cutoff radius of the dislocation core stress field, and $\rho$ is the dislocation density. The value of $r_c$ is approximated as the lattice parameter, $a_0$. $\chi$ is a parameter that accounts for the overall dislocation character, and depending on the material and the development of the network is expected to be in the interval $1.0 \leq \chi \leq 1/(1 - \nu)$ where $\nu$ is Poisson's ratio. The elastic constants that are necessary to evaluate Eq. 4.3 were found in Refs. [134, 132], and the material parameters used in the model are included in table 4.1.

The energetic contributions of the point defects are modeled as:

$$E_{\mathrm{vcy}} = \beta e_{\mathrm{vcy}} N_{\mathrm{vcy}} \tag{4.4}$$

$$E_{\mathrm{int}} = \gamma e_{\mathrm{int}} N_{\mathrm{int}}. \tag{4.5}$$

$e_{\mathrm{vcy}}$ and $e_{\mathrm{int}}$ are the formation energies of an isolated vacancy or interstitial as found by inserting a single vacancy or interstitial into an otherwise perfect crystal. The values of these formation energies are reported in table 4.2. After minimizing the potential energy and measuring the potential energy change $\Delta E$ relative to the perfect crystal, the isolated point defect formation energies were defined as:

$$e_{\mathrm{vcy}} = \Delta E_{\mathrm{remove}} + E_{\mathrm{coh}} \tag{4.6}$$

$$e_{\mathrm{int}} = \Delta E_{\mathrm{add}} - E_{\mathrm{coh}}. \tag{4.7}$$

$N_{\mathrm{vcy}}$ and $N_{\mathrm{int}}$ are the numbers of vacancies and interstitials, and $\beta$ and $\gamma$ are fitting parameters to account for deviations from the isolated point defect energies that occur as the density of point defects increases with dpa.

Finally, $E_{\mathrm{sro}}$ is the energy associated with the chemical short range order relative to a random solid solution, and was evaluated for the CrCoNi MPEA system using molecular statics calculations. Starting with an initial configuration, the types of all the atoms were randomly reassigned to one of the three

Table 4.2: Isolated point defect formation energies.

|        | $e_{\mathrm{vcy}}$ [eV] | $e_{\mathrm{int}}$ [eV] |
|--------|-------------------------|-------------------------|
| Fe     | 1.71                    | 4.01                    |
| CrCoNi | 1.49                    | 2.45                    |
| A-atom | 1.62                    | 3.85                    |

constituent elements and the potential energy of the structure was minimized. This procedure was repeated five times for each configuration to account for statistical variations, and the the average potential energy change relative to the initial structure was included in the model for each dpa for which the fitting was conducted.

The resulting energy model contains the three adjustable parameters $\chi$, $\beta$, and $\gamma$. Fitting the model to each of the three systems involved first constructing the bounds on the dislocation energy that would be realized by the minimum and maximum allowed values for $\chi$. The dislocation energy was then fixed at the lower bound, and a least-squares technique was used to find the values of $\beta$ and $\gamma$ that minimized the difference between $E_{\mathrm{stored}}$ and $E_{\mathrm{fit}}$ over the entire interval up to 2.0 dpa. Repeating this procedure with the dislocation energy fixed at the upper bound allowed the construction of corresponding intervals for the predicted contributions of the vacancy and interstitial terms. As discussed below in Sec. 4.3.5, the largest contribution to the model error is likely the estimated numbers of point defects $N_{\mathrm{vcy}}$ and $N_{\mathrm{int}}$.

## 4.3  Results and Discussion

### 4.3.1  Total energy and pressure

The successive generation of Frenkel pairs considerably increases the potential energy of the simulated systems, particularly in the absence of thermally-driven point defect migration and recombination. As displayed in Fig. 4.4, the energies of all three systems roughly reach steady states by 0.5 dpa, suggesting the activity of another recovery mechanism that offsets the energy increase of additional Frenkel pairs. The convergence of the system potential energy does not indicate that the microstructure has reached a steady state though; a slow but continual increase in pressure past 0.5 dpa that is most visible for the Fe system in Fig. 4.5 implies that at least some features of the microstructure continue to evolve up to much higher dpa.

The supplied energy that persists through the structural relaxations can be partitioned into local material defects and an overall elastic energy—there is no kinetic energy in the absence of atomic velocities. The increasing elastic energy is a direct consequence of the introduction of Frenkel pairs; consider that displacing an internal atom to an external surface of a crystal increases the crystal volume by one atomic volume.
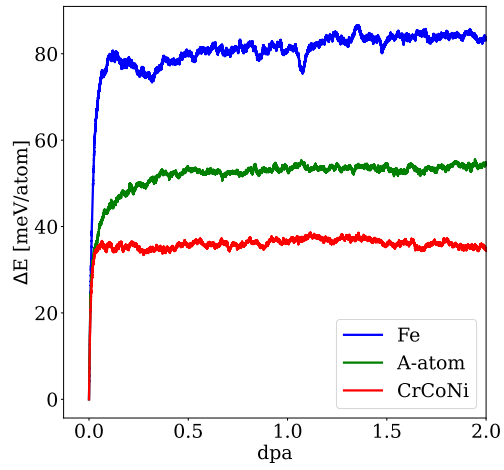
65

Figure 4.4: The change in the energy of the simulation cell as a function of dpa for all three material systems.
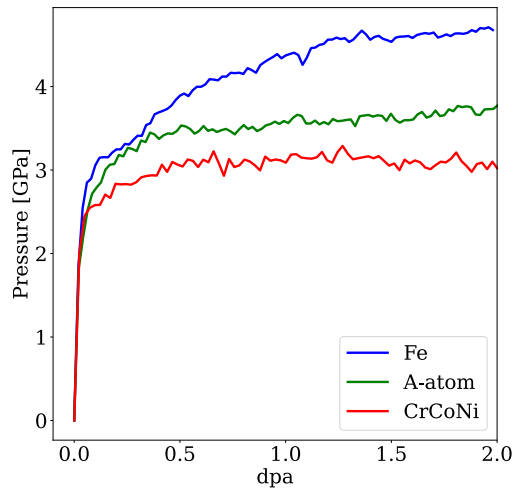


Figure 4.5: The pressure of the simulation cell as a function of dpa for all three material systems.

The simulations are performed at fixed volume though, meaning that the boundary conditions can only be satisfied by subjecting the crystal to an increasing pressure to maintain a net zero volumetric strain. The elastic energy is defined as the elastic work that would need to be performed on the relaxed system in a zero pressure configuration to return it to the required volume. This can be calculated from the material's elastic constants and the volume of the system in the relaxed configuration, with the later evaluated by allowing the simulation box to expand while minimizing the potential energy. The percent of the supplied energy that resides as elastic energy is negligible in all situations, being 3%, 0.2% and 0.3% for the Fe, CrCoNi and A-atom systems, respectively. The overwhelming majority of the supplied energy therefore resides in the form of defects, specifically dislocations, interstitials, and vacancies (the contribution of stacking faults was found to be negligible in Sec. 4.2.4).

### 4.3.2 Dislocations

While the dislocation networks that developed in the three systems are clearly distinct, the dislocation densities of all three initially rapidly increased before falling to steady state values. Figure 4.6 shows the dislocation networks in the Fe (left), CrCoNi (middle), and A-atom (right) systems at 0.5 dpa (top) and 2.0 dpa (bottom) as found by the DXA, with the network density visibly lower at higher dpa for the Fe and A-atom systems. A more quantitative analysis of the distribution of dislocation types as a function of dpa is provided in Fig. 4.7, where the dislocation density for the Fe system is visibly lower than that for both the FCC systems at all dpa. This can be explained by the main dislocation production mechanism in irradiated materials; isolated self-interstitials precipitate as interstitial disks, forming dislocations loops that evolve and eventually develop into a larger network. In Fe this is known to produce more mobile $1/2\langle 111\rangle$ and less mobile $\langle 100\rangle$ loops, with the population of the former being greater at low temperatures [140, 141]. The $1/2\langle 111\rangle$ dislocations begin as isolated loops, but appear to be mobile enough to migrate and react once the dislocation density and internal stresses reach critical values around 0.8 dpa, reducing the overall dislocation density as a dislocation network is formed.

The expectation was that the dislocation networks in the CrCoNi and A-atom systems would begin as isolated extrinsic Frank loops resulting from the precipitation of interstitials on {111} planes. It is well established experimentally that these sessile Frank eventually unfault to form a glissle dislocation network containing both perfect dislocations and Shockley partials, though the precise mechanism by which this occurs continues to be a subject of study [142, 143, 144]. What was unexpected in Fig. 4.7 is that this unfaulting process should begin at the outset, with the population of Shockley partials visibly exceeding that of any other dislocation type in Fig. 4.7 even for very low dpa. There is a measurable population of
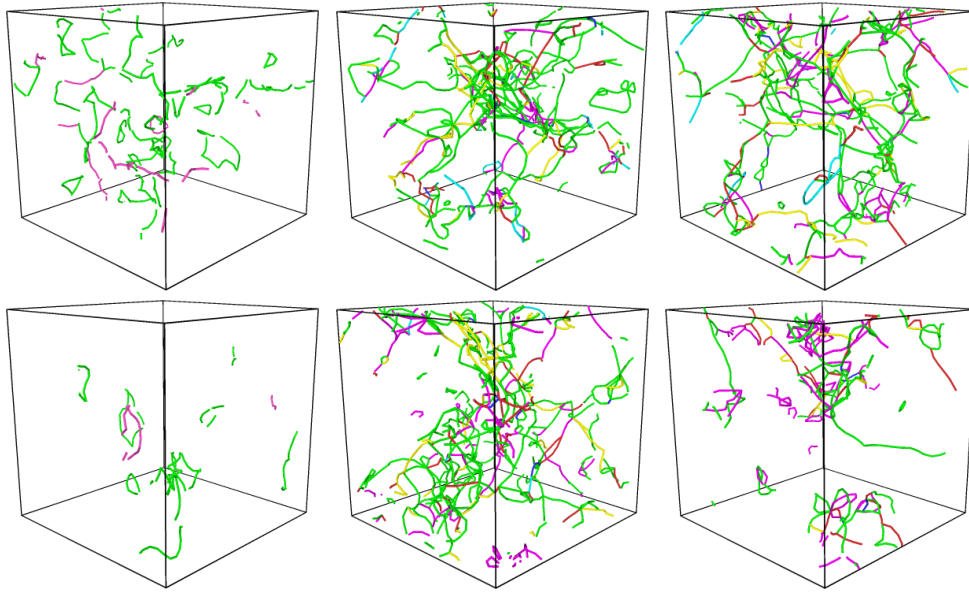
Figure 4.6: Dislocation networks at 0.5 dpa (top) and 2.0 dpa (bottom) for Fe (left), CrCoNi (middle), and A-atom (right) systems where color indicates dislocation types. For the Fe system green are $1/2\langle 111\rangle$ and purple are $\langle 100\rangle$ dislocations. For the CrCoNi and A-atom systems green are $1/6\langle 112\rangle$ Shockley partials, purple are $1/6\langle 110\rangle$ stair-rods, yellow are $1/3\langle 100\rangle$ Hirth, light blue are $1/3\langle 111\rangle$ Frank, and dark blue are $1/2\langle 110\rangle$ perfect dislocations.
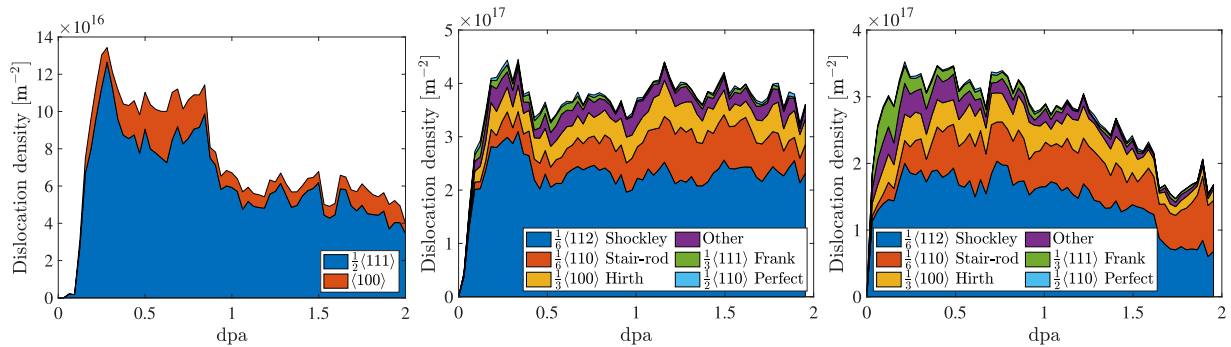


Figure 4.7: The dislocation density and dislocation types for Fe (top left), CrCoNi (top right), and A-atom (bottom) as a function of dpa.

68

extrinsic Frank loops that slowly decreases up to around 1.0 dpa, but the rapid growth of the population of Shockley partials well before this point suggests that there is some other mechanism by which Shockley partials are being generated. The details of the dislocation formation mechanisms at the earliest stages of radiation damage will require a dedicated study that is relegated to future work though. The generally increasing populations of sessile stair-rod dislocations with dpa is an expected result of the maturation of a glissile dislocation network mainly composed of Shockely partials. Finally, it is significant that the dislocation network in the A-atom system continues to undergo substantial change even up to 1.6 dpa, well beyond the 0.5 dpa at which the potential energy converged in Fig. 4.4. The invariance of the potential energy even as the dislocation density is reduced by nearly half requires a corresponding increase in the population of other defect types, a point that will be significant in the following.

### 4.3.3   Stacking faults

As expected, stacking faults (SFs) were only observed in the FCC CrCoNi and A-atom systems. The SF configurations for these two systems at 2.0 dpa are shown in Fig. 4.8, and the SF density as a function of dpa is reported in Fig. 4.9. The very high SF densities observed in these systems are consistent with the low reported values for their SF energies [134] and with the high density of partial dislocations in Fig. 4.7. It is interesting that the SF density for the A-atom system was higher than for the CrCoNi system up to around 0.8 dpa, but by 2.0 dpa the situation had reversed with the SF density in the A-atom system falling by almost a factor of three. This roughly correlates with the decrease in the density of Shockley partials in Fig. 4.7, though the correspondence is not exact. If the decrease in the SF density is in fact driven by the maturation of the dislocation network, then an important question is why a similar decrease in the density of Shockley partials and SFs was not observed in the CrCoNi system. Any explanation should likely involve the chemical SRO since this is (by design) the main difference between the CrCoNi and A-atom systems. Possible mechanisms include atomic rearrangements following the formation of the SF lowering the energy of the SF relative to the pristine state and increasing the unfaulting barrier, or the fluctuations in the local stacking fault energy decreasing the difficulty of cross slip and increasing the complexity of the dislocation network. If true, either of these mechanisms would help to explain why the dislocation density for the CrCoNi system does not fall as quickly with increasing dpa as for the A-atom system.

The elevated SF density observed here could have significant implications for the mechanical strength of irradiated CrCoNi since SFs can act as barriers to dislocation motion that increase plastic strength. Recent work by Richie et al. used *in situ* transmission electron microscopy to investigate CrCoNi under cryogenic conditions and noted that a high density of SFs and extensive cross slip likely contributed to the superior
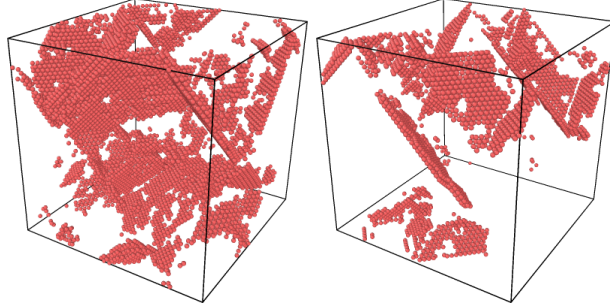
Figure 4.8: Stacking faults as identified by locally HCP coordinated atoms in CrCoNi (left) and A-atom (right) systems at 2.0 dpa
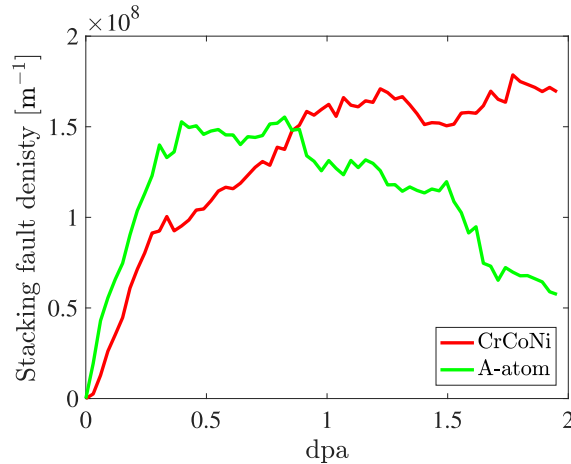


Figure 4.9: The stacking fault density as a function of dpa for the CrCoNi and A-atom systems.

mechanical properties of the alloy [145]. SFs have also been suggested to decrease radiation-induced void swelling of CrCoNi by alleviating internal stress build up [112].

### 4.3.4 Short range ordering

The CrCoNi system, initialized as a random solid solution, did develop a minor degree of SRO during the CRA simulation. While not visually apparent this ordering was statistically significant as indicated by the WC SRO parameters reported in Fig. 4.10. All of the WC parameters stabilized by 0.5 dpa, with Cr displaying an increased likelihood to neighbor Co and both Cr and Co less likely to neighbor other atoms of the same species. It is interesting that this degree of order developed despite the randomizing effects of atomic displacements in the CRA, and is significant that the order generally agrees with what has been found experimentally [146] and with more accurate DFT-based Monte Carlo simulations [147]. Finally, despite the magnitude of the SRO being relatively small, the contribution toward lowering the potential energy of the system is still substantial; as will be discussed in Sec. 4.3.5 below, the magnitude of the $E_{\mathrm{sro}}$ can be as much
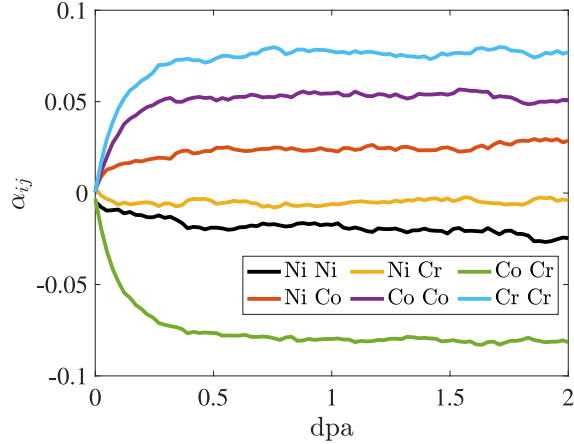
Figure 4.10: The Warren-Cowley SRO parameters as a function of dpa for the CrCoNi system.

as 30% of the magnitude of $E_{\text{fit}}$.

### 4.3.5 Point defects and energy balance

The results discussed up to this point have included dislocations and stacking faults, i.e., defects for which there are standard identification techniques and about which we have general confidence. The same is not true for point defects though, for the identification of which there are arguably no methods in the literature (including the ones used here) that work reliably at high defect concentrations. This section not only reports the nominal point defect concentrations as calculated using the methods described in Sec. 4.2, but highlights the inconsistency of these estimates with respect to the stored energy. This is particularly concerning since the point defect balance equations that describe the evolution of point defect concentrations are the foundations of our understanding of damage development in irradiated materials [100, 97], and it is unclear whether it is currently possible to evaluate the independent variables in these equations either by experiment or simulation. Our sincere hope is that this will be identified as an area requiring the further attention of the research community in the future.

One expectation is that the concentrations of vacancies and intertsitials should be equal at low dpa and increase linearly with the number of displacements. This is because both are initially produced in equal amounts by Frenkel pair insertions and the defect density is not high enough for there to be appreciable annihilation. While this is satisfied for the point defect concentrations reported in Fig. 4.11, the vacancy and interstitial concentrations quickly diverge with increasing dpa just as in the CRA simulations of Ref. [118]. This is conventionally believed to indicate that the more mobile interstitials are clustering to form dislocation loops or are annihilating on previously-existing loops once a threshold density is reached, with the less mobile vacancies remaining in the bulk. Although our simulations only considered single crystal systems, interstitials
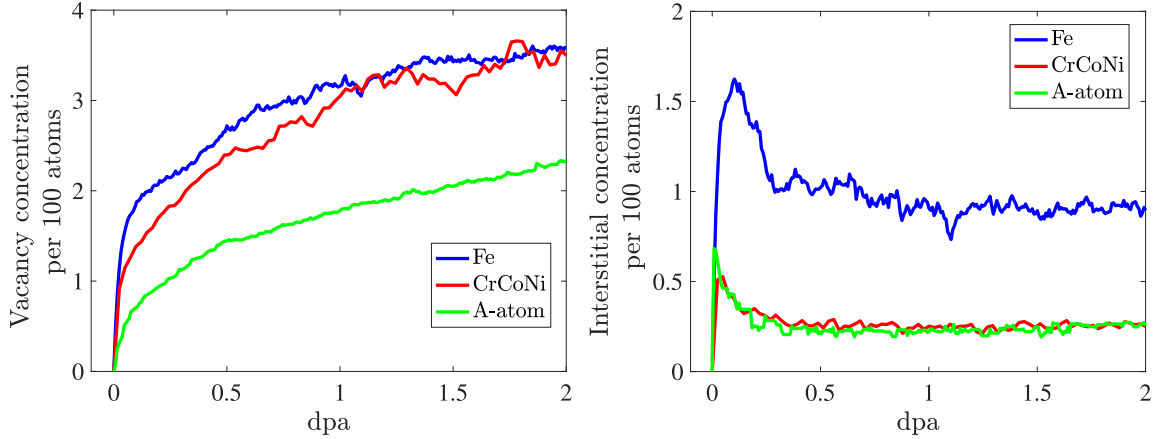
Figure 4.11: Estimated vacancy and interstitial concentrations as a function of dpa for all three material systems.

can also migrate to and annihilate on other types of sinks (e.g., grain boundaries, phase boundaries, voids) in more general materials. The nominal vacancy concentrations of the Fe and CrCoNi systems were comparable at all dpa, and were consistently ~50% greater than that of the A-atom system. It is suspicious that the vacancy concentrations of the two FCC systems as found by the WS method continue to increase all the way to 2.0 dpa though, considering that the pressures of these systems have already converged by 0.5 dpa in Fig. 4.5; along with the results of the energy fitting reported below, this suggests that the WS method increasingly overestimates the vacancy concentrations of the FCC systems at higher dpa. Confusingly, the same does not seem to be true for the Fe system for which the vacancy concentration, pressure, and energy model results are all consistent. The reason for this discrepancy is unknown, but highlights the need for more robust ways to identify vacancies (or a more precise definition of what constitutes a vacancy) in highly damaged structures.

The results of fitting the energy model $E_{\text{fit}}$ of Sec. 4.2.4 to $E_{\text{stored}}$ for all three structures are reported in Fig. 4.12, with the values of the fitting parameters $\beta$ and $\gamma$ given in Table 4.3. The model works remarkably well for the Fe system, with $E_{\text{fit}}$ reproducing all of the general trends of $E_{\text{stored}}$ and many of the smaller features as well. The fitted values of $\beta$ and $\gamma$ for the Fe system are also physically reasonable, being slightly below one as is necessary for the clustering of point defects to be energetically favorable. This is not the case for the CrCoNi and A-atom systems though; both of these systems exhibit a $E_{\text{fit}}$ that continually increases with dpa rather than converging around 0.5 dpa, and $\gamma$ values that are well above the physically reasonable bound of 1.0. The second observation in particular suggests that the interstitial concentration is severely underestimated, with the model compensating for an unreasonably low value of $N_{\text{int}}$ in Eq. 4.5 by elevating the value of $\gamma$. Supposing from the Fe system that $\gamma$ should be ~0.9, the interstitial concentration in the FCC systems appears to be underestimated by a factor of two to three at high dpa. While it is
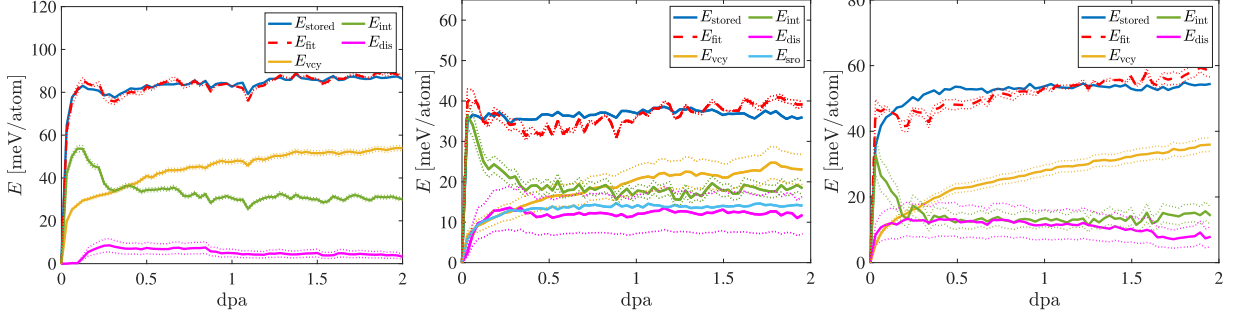
Figure 4.12: Fit of the stored energy model described in Sec. 4.2.4 for the Fe (left), CrCoNi (center), and A-atom (right) systems. The dotted trend lines are bounds derived from the estimated bounds on the dislocation network energy.

Table 4.3: Parameters fit by the energy model that account for deviations of energy from isolated defects. $\beta$ is the coefficient for vacancies and $\gamma$ is for the interstitials.

|  | $\beta$ | $\gamma$ |
| --- | --- | --- |
| Fe | 0.879 | 0.840 |
| CrCoNi | 0.454 | 2.90 |
| A-atom | 0.970 | 1.49 |

true that the steady-state interstitial concentrations should not necessarily be comparable in BCC and FCC systems, Fig. 4.11 is also consistent with the interstitial concentrations in FCC systems being systematically underestimated. The source of the error is likely that the difference in atomic volumes of interstitials and atoms of the crystalline lattice is less pronounced in FCC than BCC systems, with interstitials in BCC Fe often adopting split-dumbbell configurations.

If the vacancy concentrations are relatively accurate and the errors in the interstitial concentrations are of a similar magnitude for the CrCoNi and A-atom systems, then the gap between the vacancy and interstitial concentrations would be significantly larger for the CrCoNi than for the A-atom system throughout the simulations, even at very low dpa. If this is true, then that would have important implications for the effects of SRO and lattice distortions on the development of radiation damage. Based on a conventional understanding of the point defect balance equations [100, 97], prior observations of CrCoNi's low-temperature radiation resistance [113] imply that the gap between the vacancy and interstitial concentrations should be smaller for CrCoNi, encouraging point defect recombination and suppressing vacancy precipitation. While this is superficially inconsistent with our results, observe that a higher sustained vacancy concentration does not necessarily lead to a higher susceptibility to voids. Specifically, there is evidence that SRO and lattice distortions can increase the magnitude of energetic well depths for vacancies [148], effectively decreasing the vacancy-vacancy and vacancy-void capture radii and allowing higher vacancy concentrations to be sustained as compared to the A-atom system. Given the magnitude of the apparent errors in the estimated vacancy

and interstitial concentrations though, there is not strong evidence for this conclusion at present.

There are other qualifications relating to the CRA's limitations that should be made regarding the observed differences in point defect concentrations across the three material systems. One important factor absent in these simulations is the thermal spike associated with a collision cascade. MPEAs have been found to have a shorter mean-free electron path and a lower thermal conductivity than traditional alloys, and it has been proposed that this could prolong the thermal spike and increase the extent of defect recombination [149, 150]. It is also worth considering the potential effects of nanotwinning, which has been experimentally observed to be an important cryogenic strengthening mechanism for CrCoNi [145]. While the CRA did not induce any visible nanotwinning, this could be related to the relative sizes of the simulation cell and the critical nucleation event. Nanotwins would not only affect the development of the dislocation network, but have shown the ability to capture or promote the transport of point defects to sinks in irradiated Cu [151].

## 4.4 Conclusion

CRA simulations were conducted to investigate differences in the development of irradiated microstructures in BCC Fe, FCC CrCoNi, and FCC A-atom systems in the low temperature and high dose rate regime up to 2.0 dpa. The CrCoNi system developed the highest overall dislocation density and exhibited the ability to maintain that dislocation density even as the dislocation networks in the Fe and A-atom systems matured and simplified. The higher stacking fault density in CrCoNi is likely related to the consistently higher density of partial dislocations relative to the other systems. The successive insertion of Frenkel pairs entailed by the CRA mildly increased the degree of chemical short range order in CrCoNi relative to the initial random solid solution in a way that is consistent with other modeling and experimental studies. A model was developed for the energy stored in the material defects, and strongly suggests that the interstitial concentrations for the FCC systems as estimated from the distribution of Voronoi cell volumes are lower than the actual effective concentrations by a factor of two to three. It is also possible that the Wigner-Sietz method of identifying vacancies slightly overestimates the vacancy concentration for the FCC systems at high defect concentrations. The same methods of point defect identification seem to be much more reliable for the BCC Fe system though, with the energy model closely following the simulation results. Given the overwhelming importance of point defects to the development of radiation damage and the intense interest in CrCoNi and other FCC MPEAs for nuclear applications, our results reveal a critical need to either develop more robust ways to measure point defect concentrations in heavily-damaged FCC materials, or perhaps to reevaluate what is meant by a point defect in such materials.

through the Advancing Scientific Careers to Enhance Nuclear Technologies (ASCENT) program at UC Davis.

## Author contributions

CS and JCS performed all the simulations and analysis. JKM conceptualized the project and developed the methodology. All authors collaborated to writing, reviewing, and editing.

## Competing interests

The authors declare no competing interests.

# Chapter 5

# Conclusion and future work

This dissertation describes a novel framework for constructing MLPs in addition to two studies that use atomistic simulations to study the behavior of metal alloys in extreme conditions. While the intention was to use the MLP framework in the latter two studies, they were being conducted simultaneously with the MLP development. This practical limitation is the main reason why the simulations of Ta and CrCoNi resorted to using EAM potentials previously reported in the literature. Naturally, there are several areas in which this research could continue in the future.

Regarding the MLP framework described in Ch. 2, there are several straightforward modifications that should improve the balance of accuracy and computational efficiency. The first is to explore alternatives to the $\epsilon$-net procedure to construct the set of inducing points. While the $\epsilon$-net procedure is inexpensive and robust, equivalent performance could likely be achieved with fewer inducing points (and therefore higher efficiency) provided that the inducing points are appropriately distributed. The second is to allow the hyperparamter length-scales of the GP to vary independently for each of the descriptors. This is expected to improve the accuracy by allowing for a more flexible regression function given a fixed number of inducing points, though this will need to be done carefully to limit the increased computational cost of training with more hyperparameters. The descriptors could also be modified to include information about additional many-body interactions, and research in this direction is ongoing.

In addition to modifications of features of the MLP framework that are already functioning, there are other features that would ideally be implemented in the near future. These include:

- Extending the framework to be capable of including multiple atomic species would be necessary for simulations of more complex systems, for example, MPEAs. This could theoretically be accomplished by weighting the Dirac delta functions in Eq. 2.10 with weights that depend on the species of the

atom. By defining several sets of weights and evaluating the descriptors independently for each set, the Gaussian process would receive information both about the atomic species and positions. This approach was suggested as part of the GAP framework [35], but there is still considerable uncertainty in the literature about the construction of optimal sets of weights.

- Extending the GP to include fitting observed components of the stress tensor could be important depending on the phenomenon of interest. Accurate predictions of mechanical responses of alloys undergoing deformation is an example application where accurate stresses are essential. The system stress components can be incorrect in model systems with periodic boundary conditions even with an interatomic potential that perfectly reproduces atomic forces [152, 153]. A GP over snap stress components could be added to the already-constructed joint model in an similar manner to that described for the atomic forces in Sec. 2.5.2, and as is fully described in Ref. [67].

- Using the posterior variance of the GP to develop an active learning scheme would improve the simplicity and efficiency of training. The variance acts as a measure of uncertainty for all force and energy evaluations. If the uncertainty of the atomic forces exceeds a pre-specified tolerance, further electronic structure calculations could be performed specifically for the problematic atomic configurations, and the results added to the model data. While the earliest mention of this idea of which we know is in Ref. [35], this has been recently implemented in another Gaussian process based MLP framework that is referred to as Fast Learning of Atomic Rare Events (FLARE) [154].

As for further study related to TQC and the stored energy of cold work, which was the focus of Ch. 3, a potentially interesting and worthy avenue to pursue is testing and extending both stored energy models to other metallic systems, including polycrystalline ones. Kositski et al. [13] conducted MD simulations of multiple grains which suggested that grain coarsening could act as a significant energy source, and further investigating that contribution to the total stored energy would be worthwhile. Additionally, the parameters of the phenomenological model could be tabulated for various materials under specific straining conditions, making the model of more practical use for estimating the temperature increase in continuum scale simulations of deformation.

Regarding the study on radiation-induced damage in CrCoNi (Ch. 4), this work and other simulations of radiation damage would benefit immeasurably from more robust ways to identify point defects in heavily damaged materials. At a minimum this would allow for more confident assessments regarding the imbalance between the vacancy and interstitial concentration which is believed to be the main driver for void swelling and other types of radiation damage. Considering the fundamental motivation for studying MPEAs, exploring whether radiation resistance could be improved by varying the composition of the primary elements or

by adding additional elements would be impactful, and could be done with the Frenkel pair accumulation method described here.

# Appendix A

# Matrix inversions and determinants

A critical part of the sparse Gaussian process is the application of the Woodbury matrix inversion identity (also referred to as the Sherman-Morison-Woodbury formula) to the low-rank matrix $\widehat{Q}_{nn}$ to reduce the computational scaling from $\mathcal{O}(n^3)$ to $\mathcal{O}(nm^2)$. One formulation of the Woodbury identity is [155]:

$$\left(A + UBU^T\right)^{-1} = A^{-1}\left[I - U\left(B^{-1} + U^T A^{-1} U\right)^{-1} U^T A^{-1}\right] \tag{A.1}$$

Notice that the LHS is of the same form as $\widehat{Q}_{nn} = \Lambda + \widehat{K}_{nm}\widehat{K}_{mm}^{-1}\widehat{K}_{nm}^T$. This formula has well-documented stability issues, though the stability can be improved by adjusting the definitions of $B$ and $U$. Specifically, let $B := R^T \widehat{K}_{mm}^{-1} R$ where $R$ is upper triangular matrix constructed from the economic QR decomposition $\widehat{K}_{nm} = QR$, $U := Q$, and $A := \Lambda$.

Apart from the inversion of $\Lambda$ (a diagonal matrix), some of the other matrices that need to be inverted during the training of a sparse GP can be poorly conditioned. A common way to address these issues is to add "jitter" as discussed in detail in technical report Ref. [64]. This entails improving the conditioning of a matrix $K$ by adding a small multiple of the identity matrix before taking the inverse. That is, $K^{-1}$ is approximated as $(K + \epsilon I)^{-1}$, with the Cholesky decomposition often used to solve for the inverse where possible. However, since the smallest value of $\epsilon$ required to make $K$ positive definite and apply the Cholesky decomposition could be excessively large, our implementation generally used the Moore-Penrose pseudo-inverse method [156].

Finally, for the matrix determinant of $\widehat{Q}_{nn}$ required to evaluate the lower bound to the log-marginal likelihood in Eq. 2.36, the following identity is used [157]:

$$|\widehat{Q}_{nn}| = |\widehat{K}_{mm} + \widehat{K}_{nm}^T \Lambda \widehat{K}_{nm}||\widehat{K}_{mm}^{-1}||\Lambda|. \tag{A.2}$$

# Appendix B

# Atomic descriptors

## B.1    Definition of the $g_{nl}$

Let $f_{nl}(r)$ be a linear combination of spherical Bessel functions of the first kind

$$f_{nl}(r) = a_{nl}j_l(u_{nl}r/r_c) + b_{nl}j_l(u_{n+1l}r/r_c) \tag{B.1}$$

where $a_{nl}$ and $b_{nl}$ are constants, $j_l(r)$ is the spherical Bessel function of the first kind of order $l$, $u_{nl}$ is the $(n+1)$st nonzero root of $j_l(r)$, and $r_c$ is the cutoff radius. Since $f_{nl}(r_c) = 0$ by design, the objective is to find $a_{nl}$ and $b_{nl}$ such that $f'_{nl}(r_c) = 0$ and $f''_{nl}(r_c) = 0$. Using the differentiation rules for the spherical Bessel functions

$$j'_l(x) = j_{l-1}(x) - \frac{l+1}{x}j_l(x) \qquad j'_l(x) = \frac{l}{x}j_l(x) - j_{l+1}(x)$$

in, e.g., Sec. 14 of Ref. [57], $f'_{nl}(r)$ and $f''_{nl}(r)$ can be shown to vanish at $r = r_c$ if the coefficients in Eq. B.1 satisfy

$$a_{nl} = \frac{u_{n+1l}}{j_{l+1}(u_{nl})}c_{nl} \qquad b_{nl} = -\frac{u_{nl}}{j_{l+1}(u_{n+1l})}c_{nl}$$

for an arbitrary constant $c_{nl}$. This constant is constrained by requiring that $f_{nl}(r)$ be appropriately normalized, or

$$\int_0^{r_c} f_{nl}(r)f_{nl}(r)r^2 \mathrm{d}r = 1.$$

Using the relevant orthogonality relation for the spherical Bessel functions [57]

$$\int_0^{r_c} j_l(u_{n'l}r/r_c)j_l(u_{nl}r/r_c)r^2 \mathrm{d}r = \delta_{n'n}\frac{r_c^3}{2}[j_{l+1}(u_{nl})]^2$$

leads to

$$f_{nl}(r) = \left( \frac{1}{r_c^3} \frac{2}{u_{nl}^2 + u_{n+1l}^2} \right)^{1/2} \left[ \frac{u_{n+1l}}{j_{l+1}(u_{nl})} j_l \left( u_{nl} \frac{r}{r_c} \right) - \frac{u_{nl}}{j_{l+1}(u_{n+1l})} j_l \left( u_{n+1l} \frac{r}{r_c} \right) \right]$$

as an explicit equation for the $f_{nl}(r)$.

For any integer $l \geq 0$, the $g_{nl}(r)$ for $n \geq 0$ are a set of orthonormal functions derived by applying the Gram–Schmidt process to the $f_{nl}(r)$. Let $g_{0l}(r) = f_{0l}(r)$. The process then involves constructing $g_{nl}(r)$ for $n > 0$ given $f_{nl}(r)$ and $g_{n'l}(r)$ for $0 \leq n' \leq n-1$. Observe that $g_{n'l}(r)$ contains components of all $f_{n''l}(r)$ for $0 \leq n'' \leq n'$, and therefore terms involving $j_l(u_{n''l}r/r_c)$ for all $0 \leq n'' \leq n'+1$. Since $f_{nl}(r)$ only contains terms involving $j_l(u_{nl}r/r_c)$ and $j_l(u_{n+1l}r/r_c)$, $f_{nl}(r)$ is already orthogonal to all $g_{n'l}(r)$ for $0 \leq n' \leq n-2$ by the orthogonality of the spherical Bessel functions of the first kind. The only remaining step is to subtract the projection of $f_{nl}(r)$ onto $g_{n-1l}(r)$, or

$$h_{nl}(r) = f_{nl}(r) - g_{n-1l}(r) \int_0^{r_c} f_{nl}(r) g_{n-1l}(r) r^2 \mathrm{d}r. \tag{B.2}$$

$h_{nl}(r)$ is orthogonal to all $g_{n'l}(r)$ for $l \leq n' \leq n-1$, but is not yet normalized. Let

$$d_{nl} = \int_0^{r_c} h_{nl}(r) h_{nl}(r) r^2 \mathrm{d}r \tag{B.3}$$

be the squared magnitude of $h_{nl}(r)$. The desired orthonormal $g_{nl}(r)$ is

$$g_{nl}(r) = h_{nl}(r)/\sqrt{d_{nl}}. \tag{B.4}$$

Explicitly evaluating the integral

$$\int_0^{r_c} f_{nl}(r) g_{n-1l}(r) r^2 \mathrm{d}r = -\sqrt{\frac{e_{nl}}{d_{n-1l}}}$$

where the constants $e_{nl}$ are defined as

$$e_{nl} = \frac{u_{n-1l}^2 u_{n+1l}^2}{(u_{n-1l}^2 + u_{nl}^2)(u_{nl}^2 + u_{n+1l}^2)}$$

allows Eqs. B.2, B.3 and B.4 to be solved to derive the recursion relations

$$d_{nl} = 1 - \frac{e_{nl}}{d_{n-1l}} \qquad g_{nl}(r) = \left[ f_{nl}(r) + \sqrt{1 - d_{nl}} g_{n-1l}(r) \right]/\sqrt{d_{nl}}.$$

These recursion relations can be initialized with $d_{0l} = 1$ and $g_{0l}(r) = f_{0l}(r)$ for any $0 \leq l$.

81

# Appendix C

# Dislocation network energy model

This section provides additional details on the model used for the evaluation of the dislocation network energy in our Ta simulations. Following Ref. [94], the energy of a complex dislocation network in an isotropic medium is well described by the expression

$$E_{\text{dis}} = C \frac{\mu b^2}{4\pi} \rho \ln\left(\frac{1}{r_c \sqrt{\rho}}\right) \qquad (\text{C.1})$$

where $\mu$ is the shear modulus, $b$ is the Burgers vector magnitude, $\rho$ is the dislocation density, and $r_c$ is the core cutoff radius. $C$ is a coefficient that depends on the distribution of the dislocation line character angles, with values ranging from $C = 1.0$ for an array of pure screw dislocations to $C = 1/(1 - \nu)$ for an array of edge dislocations, where $\nu$ is Poisson's ratio. Reference [94] examined the validity of Eq. C.1 for complex dislocation networks using discrete dislocation dynamics (DDD) simulations in FCC crystals and showed that this simple model is accurate with the value of $C$ reflecting the dislocation character angle averaged over all lines in the network.

Here the validity of the same analytical model is examined for dislocation networks in Ta. For this purpose we first extracted the dislocation networks attained at several strain values in simulation B using the DXA method [92, 93]. Energies of the extracted dislocation networks were found by summing up the pairwise interaction energies of all dislocation segments in the networks as in Ref. [94]. Since real space solutions for such interaction energies are only available for elastically isotropic crystals, isotropic elasticity was assumed with $\mu = 55$ GPa and $\nu = 0.339$. The core energy contributions were computed for the same networks using a simple core model described in Ref. [94] and added to the elastic energies. The so-computed network energies containing both isotropic elastic and core contributions are plotted in Fig. C.1a as a function of the dislocation density. On the same plot the best fit to the analytical energy model in Eq. C.1 is also shown as a solid blue line with $C$ as the only fitting parameter. This simple model captures the energy
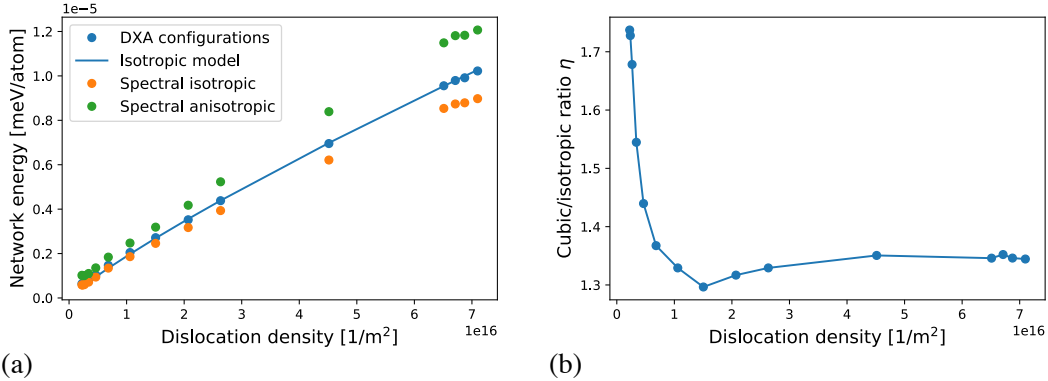
Figure C.1: (a) Evaluation of the dislocation network energy as a function of the dislocation density using different methods and the isotropic model in Eq. C.1. (b) Evolution of the empirical anisotropic correction factor $\eta$ defined as the ratio of the energies computed for cubic and isotropic elasticity using the spectral method.

of our complex BCC networks quite well. The best fit was obtained for $C = 1.1$ which is consistent with the average character angle of $\theta \simeq 20^o$ computed from the fully developed dislocation networks in our MD simulations; this average angle reflects the prevalence of screw dislocations in our configurations.

Despite seemingly working well, the simple analytical model in Eq. C.1 needs to be corrected to be applicable to the dislocation networks produced by our MD simulations. First, variations in the core energy with dislocation character angle are subtle and, as was shown in Ref. [95], cannot possibly be described using a single value of the core cutoff parameter $r_c$. However, this inaccuracy only enters through the slowly-varying logarithmic term. Fixing the core radius at $r_c = 0.5b$ is generally consistent with the core energy values calculated at 0 K in Ref. [95], and all remaining variations are subsumed into the single fitting parameter $\chi$. For simplicity, the same parameter $\chi$ also accounts for the character angle dependence of the network energy. We expect $\chi$ to be in the range $1 \leq \chi \leq 1/(1 - \nu)$, similar to the previously-used fitting parameter $C$.

A second and more substantial correction is required to account for the elastic anisotropy of our atomistic Ta model which has a Zener's anisotropy ratio of $2C_{44}/(C_{11} - C_{12}) = 2.1$ at 300 K. This could make the approximation of isotropic elasticity previously used to validate Eq. C.1, and in particular the $\mu b^2/(4\pi)$ energy factor, significantly inaccurate. This inaccuracy is accounted for using an anisotropy correction factor $\eta$. Computing $\eta$ would involve computing the energies of the same dislocation networks in a crystal with cubic symmetry. Since there are no analytical solutions for the interaction energies of dislocation segments in cubic crystals, we resort to a spectral method in which the energy of the elastic field induced

83

by a dislocation network is computed by directly integrating the strain energy density over the simulation volume [158].

The method entails summing up Fourier components of the elastic strain energy obtained using the elastic Green's function available in an analytical form for solids of arbitrary symmetry in the $k$-space. To obtain the correction factor, this spectral method is used to compute and compare energies of the same dislocation networks with full cubic anisotropy and in the isotropic approximation. The spectral method is approximate, with an accuracy limited by the spacing of the grid used by the Fast Fourier Transform (FFT) to map the field values into $k$-space. The spectral method was used to compute both the cubic and the isotropic energies for a number of dislocation networks extracted from simulation B for grids containing $128^3$, $256^3$ and $512^3$ points. These grids are likely too coarse to fully capture the variations of the strain fields near the dislocation cores, as indicated by the results for the $256^3$ grid shown in Fig. C.1a (orange and green symbols). Compared to the energies computed using the more accurate real-space method of Ref. [94] (blue symbols), the network energies computed with the spectral method in the isotropic approximation are consistently lower, a consequence of the smearing of the elastic strain field on the regular grid. Nevertheless, we observe that the ratio of the cubic to the isotropic network energies is nearly independent of the grid size. At higher dislocation densities where the networks are fully developed, the ratio of the anisotropic and isotropic energies is found to converge to a value $\eta \simeq 1.35$ (Fig. C.1b). This is introduced as the correction factor in

$$E_{\text{dis}} = \eta \chi \frac{\mu b^2}{4\pi} \rho \ln \left( \frac{1}{r_c \sqrt{\rho}} \right), \tag{C.2}$$

and is subsequently used to fit the above equation to the energy storage data extracted directly from our MD simulations.

# Bibliography

[1] Pamela Rugen and Barbara Callahan. An overview of Monte Carlo, a fifty year perspective. *Human and Ecological Risk Assessment: An International Journal*, 2(4):671–680, 1996.

[2] Avneet Sood, R Arthur Forster, Billy Joe Archer, and Robert Currier Little. Neutronics calculation advances at Los Alamos: Manhattan Project to Monte Carlo. *Nuclear Technology*, 207(sup1):S100–S133, 2021.

[3] John Allison, Dan Backman, and Leo Christodoulou. Integrated computational materials engineering: a new paradigm for the global materials profession. *JOM*, 58(11):25–27, 2006.

[4] Massachusetts Institute of Technology. The future of nuclear energy in a carbon-constrained world, 2018.

[5] Generation IV International Forum. A technology roadmap for generation iv nuclear energy systems, 2002.

[6] Generation IV International Forum. Gif 2020 annual report, 2020.

[7] Steven J Zinkle and GS Was. Materials challenges in nuclear energy. *Acta Materialia*, 61(3):735–758, 2013.

[8] Geoffrey Ingram Taylor and H. Quinney. The latent energy remaining in a metal after cold working. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 143(849):307–326, 1934.

[9] J. C. Nieto-Fuentes, S. Osovski, A. Venkert, and D. Rittel. Reassessment of the dynamic thermomechanical conversion in metals. *Physical Review Letters*, 123:255502, Dec 2019.

[10] D. Rittel, A. Bhattacharyya, B. Poon, J. Zhao, and G. Ravichandran. Thermomechanical characterization of pure polycrystalline tantalum. *Materials Science and Engineering: A*, 447(1):65–70, 2007.

[11] D. Rittel, M.L. Silva, Benny Poon, and G. Ravichandran. Thermomechanical behavior of single crystalline tantalum in the static and dynamic regime. *Mechanics of Materials*, 41:1323–1329, 12 2009.

[12] D. Rittel, L.H. Zhang, and S. Osovski. The dependence of the Taylor–Quinney coefficient on the dynamic loading mode. *Journal of the Mechanics and Physics of Solids*, 107:96–114, 2017.

[13] Roman Kositski and Dan Mordehai. Employing molecular dynamics to shed light on the microstructural origins of the Taylor–Quinney coefficient. *Acta Materialia*, 205:116511, 2021.

[14] Kai Nordlund, Steven J Zinkle, Andrea E Sand, Fredric Granberg, Robert S Averback, Roger Stoller, Tomoaki Suzudo, Lorenzo Malerba, Florian Banhart, William J Weber, et al. Improving atomic displacement and replacement calculations with physically realistic damage models. *Nature Communications*, 9(1):1–8, 2018.

[15] Kai Nordlund. Molecular dynamics simulation of ion ranges in the 1–100 kev energy range. *Computational Materials Science*, 3(4):448–456, 1995.

[16] Kai Nordlund, Janne Wallenius, and L Malerba. Molecular dynamics simulations of threshold displacement energies in Fe. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 246(2):322–332, 2006.

[17] Ruopeng Zhang, Shiteng Zhao, Jun Ding, Yan Chong, Tao Jia, Colin Ophus, Mark Asta, Robert O Ritchie, and Andrew M Minor. Short-range order and its impact on the CrCoNi medium-entropy alloy. *Nature*, 581(7808):283–287, 2020.

[18] Easo P George, Dierk Raabe, and Robert O Ritchie. High-entropy alloys. *Nature Reviews Materials*, 4(8):515–534, 2019.

[19] Paulius Malinovskis, Stefan Fritze, Lars Riekehr, Linus von Fieandt, Johan Cedervall, David Rehnlund, Leif Nyholm, Erik Lewin, and Ulf Jansson. Synthesis and characterization of multicomponent (CrNbTaTiW) C films for increased hardness and corrosion resistance. *Materials & Design*, 149:51–62, 2018.

[20] Fredric Granberg, Flyura Djurabekova, Emil Levo, and Kai Nordlund. Damage buildup and edge dislocation mobility in equiatomic multicomponent alloys. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 393:114–117, 2017.

[21] Jeffrey M Rickman, Turab Lookman, and Sergei V Kalinin. Materials informatics: From the atomic-level to the continuum. *Acta Materialia*, 168:473–510, 2019.

[22] Avni Jain, Jonathan A. Bollinger, and Thomas M. Truskett. Inverse methods for material design. *AIChE Journal*, 60(8):2732–2740, 2014.

[23] Frederic E Bock, Roland C Aydin, Christian J Cyron, Norbert Huber, Surya R Kalidindi, and Benjamin Klusemann. A review of the application of machine learning and data mining approaches in continuum materials mechanics. *Frontiers in Materials*, 6:110, 2019.

[24] Yunxing Zuo, Chi Chen, Xiangguo Li, Zhi Deng, Yiming Chen, Jörg Behler, Gábor Csányi, Alexander V. Shapeev, Aidan P. Thompson, Mitchell A. Wood, and Shyue Ping Ong. Performance and cost assessment of machine learning interatomic potentials. *The Journal of Physical Chemistry A*, 124(4):731–745, 2020.

[25] Conrad W Rosenbrock, Konstantin Gubaev, Alexander V Shapeev, Livia B Pártay, Noam Bernstein, Gábor Csányi, and Gus LW Hart. Machine-learned interatomic potentials for alloys and alloy phase diagrams. *npj Computational Materials*, 7(1):1–9, 2021.

[26] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical Review Letters*, 98:146401, 2007.

[27] Robert D Skeel. What makes molecular dynamics work? *SIAM Journal on Scientific Computing*, 31(2):1363–1378, 2009.

[28] Murray S Daw, Stephen M Foiles, and Michael I Baskes. The embedded-atom method: a review of theory and applications. *Materials Science Reports*, 9(7-8):251–310, 1993.

[29] Bohumir Jelinek, Sebastien Groh, Mark F Horstemeyer, Jeffery Houze, Seong-Gon Kim, Gregory J Wagner, Amitava Moitra, and Michael I Baskes. Modified embedded atom method potential for Al, Si, Mg, Cu, and Fe alloys. *Physical Review B*, 85(24):245102, 2012.

[30] Arthur F Voter. The embedded atom method. *Intermetallic Compounds: Principles*, 1:77, 1994.

[31] Ke Xu, Liang-Liang Niu, Shuo Jin, Xiaolin Shu, Hongxian Xie, Lifang Wang, and Guang-Hong Lu. Atomistic simulations of screw dislocations in BCC tungsten: From core structures and static properties to interaction with vacancies. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 393:174–179, 2017.

[32] Nicolas Bertin, Ryan B Sills, and Wei Cai. Frontiers in the simulation of dislocations. *Annual Review of Materials Research*, 50(SAND-2020-3747J; LLNL-JRNL-809834), 2020.

[33] AJ Skinner and JQ Broughton. Neural networks in computational materials science: Training algorithms. *Modelling and Simulation in Materials Science and Engineering*, 3(3):371, 1995.

[34] Arjen Van Ooyen and Bernard Nienhuis. Improving the convergence of the back-propagation algorithm. *Neural Networks*, 5(3):465–471, 1992.

[35] Albert P. Bartók, Risi Kondor, and Gábor Csányi. On representing chemical environments. *Physical Review B*, 87:184115, 2013.

[36] Sandip De, Albert P Bartók, Gábor Csányi, and Michele Ceriotti. Comparing molecules and solids across structural and alchemical space. *Physical Chemistry Chemical Physics*, 18(20):13754–13769, 2016.

[37] A.P. Thompson, L.P. Swiler, C.R. Trott, S.M. Foiles, and G.J. Tucker. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *Journal of Computational Physics*, 285:316–330, 2015.

[38] Spencer Wyant, Andrew Rohskopf, and Asegun Henry. Machine learned interatomic potentials for modeling interfacial heat transport in Ge/GaAs. *Computational Materials Science*, 200:110836, 2021.

[39] Alexander V. Shapeev. Moment tensor potentials: A class of systematically improvable interatomic potentials. *Multiscale Modeling and Simulation*, 14(3):1153–1173, 2016.

[40] Ivan S Novikov, Konstantin Gubaev, Evgeny V Podryabinkin, and Alexander V Shapeev. The MLIP package: moment tensor potentials with MPI and active learning. *Machine Learning: Science and Technology*, 2(2):025002, 2020.

[41] II Novoselov, AV Yanilkin, AV Shapeev, and EV Podryabinkin. Moment tensor potentials as a promising tool to study diffusion processes. *Computational Materials Science*, 164:46–56, 2019.

[42] Rebecca K. Lindsey, Laurence E. Fried, and Nir Goldman. ChIMES: A force matched potential with explicit three-body interactions for molten carbon. *Journal of Chemical Theory and Computation*, 13(12):6222–6229, 2017.

[43] Rebecca K Lindsey, Laurence E Fried, and Nir Goldman. Application of the ChIMES force field to nonreactive molecular systems: Water at ambient conditions. *Journal of Chemical Theory and Computation*, 15(1):436–447, 2018.

[44] Rebecca K Lindsey, Nir Goldman, Laurence E Fried, and Sorin Bastea. Chemistry-mediated Ostwald ripening in carbon-rich C/O systems at extreme conditions. *Nature Communications*, 13(1):1–7, 2022.

[45] Huy Pham, Rebecca Lindsey, Laurence Fried, and Nir Goldman. Reducing optimal training set design with many-body repulsive potentials for high accuracy density-functional tight binding models. *Bulletin of the American Physical Society*, 2022.

[46] Cong Huy Pham, Rebecca K Lindsey, Laurence E Fried, and Nir Goldman. Calculation of the detonation state of hn3 with quantum accuracy. *The Journal of Chemical Physics*, 153(22):224102, 2020.

[47] Jörg Behler. Perspective: Machine learning potentials for atomistic simulations. *The Journal of Chemical Physics*, 145(17):170901, 2016.

[48] Oliver T Unke, Stefan Chmiela, Huziel E Sauceda, Michael Gastegger, Igor Poltavsky, Kristof T Schütt, Alexandre Tkatchenko, and Klaus-Robert Müller. Machine learning force fields. *Chemical Reviews*, 121(16):10142–10186, 2021.

[49] Tim Mueller, Alberto Hernandez, and Chuhong Wang. Machine learning for interatomic potential models. *The Journal of Chemical Physics*, 152(5):050902, 2020.

[50] Y Mishin. Machine-learning interatomic potentials for materials science. *Acta Materialia*, 214:116980, 2021.

[51] Sergey N. Pozdnyakov, Michael J. Willatt, Albert P. Bartók, Christoph Ortner, Gábor Csányi, and Michele Ceriotti. Incompleteness of atomic structure representations. *Physical Review Letters*, 125:166001, Oct 2020.

[52] Geneviève Dusson, Markus Bachmayr, Gábor Csányi, Ralf Drautz, Simon Etter, Cas van der Oord, and Christoph Ortner. Atomic cluster expansion: Completeness, efficiency and stability. *Journal of Computational Physics*, 454:110946, 2022.

[53] Huziel E Sauceda, Stefan Chmiela, Igor Poltavsky, Klaus-Robert Müller, and Alexandre Tkatchenko. Molecular force fields with gradient-domain machine learning: Construction and application to dynamics of small molecules with coupled cluster forces. *The Journal of chemical physics*, 150(11):114102, 2019.

[54] Jeremy N Harvey. On the accuracy of density functional theory in transition metal chemistry. *Annual Reports Section "C" (Physical Chemistry)*, 102:203–226, 2006.

[55] Behnam Parsaeifard and Stefan Goedecker. Manifolds of quasi-constant soap and acsf fingerprints and the resulting failure to machine learn four-body interactions. *The Journal of Chemical Physics*, 156(3):034302, 2022.

[56] Behnam Parsaeifard, Deb Sankar De, Anders S Christensen, Felix A Faber, Emir Kocer, Sandip De, Jörg Behler, O Anatole von Lilienfeld, and Stefan Goedecker. An assessment of the structural resolution of various fingerprints commonly used in machine learning. *Machine Learning: Science and Technology*, 2(1):015018, mar 2021.

[57] G.B. Arfken, H.J. Weber, and F.E. Harris. *Mathematical Methods for Physicists: A Comprehensive Guide.* Elsevier Science, 2013.

[58] R.G. Newton. *Scattering Theory of Waves and Particles.* Theoretical and Mathematical Physics. Springer Berlin Heidelberg, 2013.

[59] N. Young. *An Introduction to Hilbert Space.* Cambridge mathematical textbooks. Cambridge University Press, 1988.

[60] D.A. Varshalovich, A.N. Moskalev, and V.K. Khersonskii. *Quantum Theory Of Angular Momemtum.* World Scientific Publishing Company, 1988.

[61] C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning.* MIT Press, 2006.

[62] Yuichi Motai. *Data-variant Kernel Analysis.* John Wiley & Sons, 2015.

[63] Haitao Liu, Yew-Soon Ong, Xiaobo Shen, and Jianfei Cai. When gaussian process meets big data: A review of scalable GPs. *IEEE Transactions on Neural Networks and Learning Systems*, 31(11):4405–4423, 2020.

[64] Michalis Titsias. Variational learning of inducing variables in sparse gaussian processes. In David van Dyk and Max Welling, editors, *Proceedings of the Twelth International Conference on Artificial Intelligence and Statistics*, volume 5 of *Proceedings of Machine Learning Research*, pages 567–574, Hilton Clearwater Beach Resort, Clearwater Beach, Florida USA, 16–18 Apr 2009. PMLR.

[65] Matthias Bauer, Mark van der Wilk, and Carl Edward Rasmussen. Understanding probabilistic sparse gaussian process approximations. *Advances in Neural Information Processing Systems*, 29, 2016.

[66] Michael JD Powell. The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*, 26, 2009.

[67] Albert P Bartók and Gábor Csányi. Gaussian approximation potentials: A brief tutorial introduction. *International Journal of Quantum Chemistry*, 115(16):1051–1057, 2015.

[68] J.C. Simac, N. Bertin, J.K. Mason, and V.V. Bulatov. Energy storage under high-rate compression of single crystal tantalum. *Acta Materialia*, page 118253, 2022.

[69] G. I. Taylor. The mechanism of plastic deformation of crystals. Part I. Theoretical. *Proceedings of the Royal Society of London Series A*, 145(855):362–387, July 1934.

[70] M.B. Bever, D.L. Holt, and A.L. Titchener. The stored energy of cold work. *Progress in Materials Science*, 17:5–177, 1973.

[71] VL Berdichevsky. On thermodynamics of crystal plasticity. *Scripta Materialia*, 54(5):711–716, 2006.

[72] JS Langer, Eran Bouchbinder, and Turab Lookman. Thermodynamic theory of dislocation-mediated plasticity. *Acta Materialia*, 58(10):3718–3732, 2010.

[73] JS Langer. Statistical thermodynamics of crystal plasticity. *Journal of Statistical Physics*, 175(3):531–541, 2019.

[74] Charles Ka Cheong Lieou and Curt Allan Bronkhorst. Thermomechanical conversion in metals: dislocation plasticity model evaluation of the Taylor–Quinney coefficient. *Acta Materialia*, 202, 10 2020.

[75] Aleksander Zubelewicz. Century-long Taylor–Quinney interpretation of plasticity-induced heating re-examined. *Scientific reports*, 9(1):9088–7, 2019.

[76] D. Macdougall. Determination of the plastic work converted to heat using radiometry. *Experimental Mechanics*, 40:298–306, 09 2000.

[77] DE Fratanduono, M Millot, DG Braun, SJ Ali, A Fernandez-Pañella, CT Seagle, J-P Davis, JL Brown, Y Akahama, RG Kraus, et al. Establishing gold and platinum standards to 1 terapascal using shockless compression. *Science*, 372(6546):1063–1068, 2021.

[78] Richard G Kraus, Russell J Hemley, Suzanne J Ali, Jonathan L Belof, Lorin X Benedict, Joel Bernier, Dave Braun, RE Cohen, Gilbert W Collins, Federica Coppari, et al. Measuring the melting curve of iron at super-earth core conditions. *Science*, 375(6577):202–205, 2022.

[79] J.J. Mason, A.J. Rosakis, and G. Ravichandran. On the strain and strain rate dependence of the fraction of plastic work converted to heat: an experimental study using high speed infrared detectors and the kolsky bar. *Mechanics of Materials*, 17(2):135–145, 1994.

[80] AA Benzerga, Y Bréchet, A Needleman, and E Van der Giessen. The stored energy of cold work: Predictions from discrete dislocation plasticity. *Acta Materialia*, 53(18):4765–4779, 2005.

[81] W. Köck and P. Paschen. Tantalum—processing, properties and applications. *JOM - Journal of the Minerals, Metals and Materials Society*, 41(10):33–39, October 1989.

[82] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. 't in Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications*, 271:108171, 2022.

[83] Youhong Li, Donald J. Siegel, James B. Adams, and Xiang-Yang Liu. Embedded-atom-method tantalum potential developed by the force-matching method. *Physical Review B*, 67:125101, Mar 2003.

[84] Luis A Zepeda-Ruiz, Alexander Stukowski, Tomas Oppelstrup, and Vasily V Bulatov. Probing the limits of metal plasticity with molecular dynamics simulations. *Nature*, 550(7677):492–495, 2017.

[85] Daniel Rittel. On the conversion of plastic work to heat during high strain rate deformation of glassy polymers. *Mechanics of Materials*, 31(2):131–139, 1999.

[86] Luis A Zepeda-Ruiz, Alexander Stukowski, Tomas Oppelstrup, Nicolas Bertin, Nathan R Barton, Rodrigo Freitas, and Vasily V Bulatov. Atomistic insights into metal hardening. *Nature Materials*, 20(3):315–320, 2021.

[87] Nicolas Bertin, LA Zepeda-Ruiz, and VV Bulatov. Sweep-tracing algorithm: in silico slip crystallography and tension-compression asymmetry in BCC metals. *Materials Theory*, 6(1):1–23, 2022.

[88] Peter M Anderson, John P Hirth, and Jens Lothe. *Theory of Dislocations.* Cambridge University Press, 2017.

[89] Jaime Marian, Wei Cai, and Vasily V Bulatov. Dynamic transitions from smooth to rough to twinning in dislocation motion. *Nature Materials*, 3(3):158–163, 2004.

[90] Johannes J Möller and Erik Bitzek. BDA: A novel method for identifying defects in body-centered cubic crystals. *MethodsX*, 3:279–288, 2016.

[91] Alexander Stukowski. Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool. *Modelling and Simulation in Materials Science and Engineering*, 18(1):015012, 2009.

[92] Alexander Stukowski and Karsten Albe. Extracting dislocations and non-dislocation crystal defects from atomistic simulation data. *Modelling and Simulation in Materials Science and Engineering*, 18(8):085001, 2010.

[93] Alexander Stukowski. A triangulation-based method to identify dislocations in atomistic models. *Journal of the Mechanics and Physics of Solids*, 70:314–319, 2014.

[94] Nicolas Bertin and Wei Cai. Energy of periodic discrete dislocation networks. *Journal of the Mechanics and Physics of Solids*, 121:133–146, 2018.

[95] N Bertin, W Cai, S Aubry, and VV Bulatov. Core energies of dislocations in BCC metals. *Physical Review Materials*, 5(2):025002, 2021.

[96] Steven J Zinkle and Jeremy T Busby. Structural materials for fission & fusion energy. *Materials today*, 12(11):12–19, 2009.

[97] Gary S Was. *Fundamentals of radiation materials science: metals and alloys*. Springer, 2016.

[98] Kai Nordlund, Steven J Zinkle, Andrea E Sand, Fredric Granberg, Robert S Averback, Roger E Stoller, Tomoaki Suzudo, Lorenzo Malerba, Florian Banhart, William J Weber, et al. Primary radiation damage: A review of current understanding and models. *Journal of Nuclear Materials*, 512:450–479, 2018.

[99] Werner Schilling and Hans Ullmaier. *Physics of Radiation Damage in Metals*, chapter 9, pages 180–241. John Wiley & Sons, Ltd, 2006.

[100] LK Mansur. Theory and experimental background on dimensional changes in irradiated alloys. *Journal of Nuclear Materials*, 216:97–123, 1994.

[101] Alan J Ardell and Pascal Bellon. Radiation-induced solute segregation in metallic alloys. *Current Opinion in Solid State and Materials Science*, 20(3):115–139, 2016.

[102] TR Anthony. Solute segregation in vacancy gradients generated by sintering and temperature changes. *Acta Metallurgica*, 17(5):603–609, 1969.

[103] AJ Ardell. Radiation-induced solute segregation in alloys. In *Materials Issues for Generation IV Systems*, pages 285–310. Springer, 2008.

[104] FA Garner, MB Toloczko, and BH Sencer. Comparison of swelling and irradiation creep behavior of FCC-austenitic and BCC-ferritic/martensitic alloys at high neutron exposure. *Journal of Nuclear Materials*, 276(1-3):123–142, 2000.

[105] Stephen Taller, Gerrit VanCoevering, Brian D Wirth, and Gary S Was. Predicting structural material degradation in advanced nuclear reactors with ion irradiation. *Scientific reports*, 11(1):1–14, 2021.

[106] JR Matthews and MW Finnis. Irradiation creep models–an overview. *Journal of Nuclear Materials*, 159:257–285, 1988.

[107] Brain Cantor, ITH Chang, P Knight, and AJB Vincent. Microstructural development in equiatomic multicomponent alloys. *Materials Science and Engineering: A*, 375:213–218, 2004.

[108] J-W Yeh, S-K Chen, S-J Lin, J-Y Gan, T-S Chin, T-T Shun, C-H Tsau, and S-Y Chang. Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes. *Advanced Engineering Materials*, 6(5):299–303, 2004.

[109] Bernd Gludovatz, Anton Hohenwarter, Dhiraj Catoor, Edwin H Chang, Easo P George, and Robert O Ritchie. A fracture-resistant high-entropy alloy for cryogenic applications. *Science*, 345(6201):1153–1158, 2014.

[110] Fredric Granberg, K Nordlund, Mohammad W Ullah, Ke Jin, Chenyang Lu, Hongbin Bei, LM Wang, F Djurabekova, WJ Weber, and Y Zhang. Mechanism of radiation damage reduction in equiatomic multicomponent single phase alloys. *Physical review letters*, 116(13):135504, 2016.

[111] Zhouran Zhang, David EJ Armstrong, and Patrick S Grant. The effects of irradiation on CrMnFeCoNi high-entropy alloy and its derivatives. *Progress in Materials Science*, 123:100807, 2022.

[112] Chenyang Lu, Tai-Ni Yang, Ke Jin, Gihan Velisa, Pengyuan Xiu, Qing Peng, Fei Gao, Yanwen Zhang, Hongbin Bei, William J Weber, et al. Irradiation effects of medium-entropy alloy nicocr with and without pre-indentation. *Journal of Nuclear Materials*, 524:60–66, 2019.

[113] G Velişa, Z Fan, ML Crespillo, H Bei, WJ Weber, and Y Zhang. Temperature effects on damage evolution in ion-irradiated nicocr concentrated solid-solution alloy. *Journal of Alloys and Compounds*, 832:154918, 2020.

[114] ASTM E521-16. Standard practice for investigating the effects of neutron radiation damage using charged-particle irradiation, 2016.

[115] GH Kinchin and RS Pease. The displacement of atoms in solids by radiation. *Reports on progress in physics*, 18(1):1, 1955.

[116] MJ Norgett, MT Robinson, and I Mcalder Torrens. A proposed method of calculating displacement dose rates. *Nuclear engineering and design*, 33(1):50–54, 1975.

[117] Keith M Beardmore and Niels Grønbech-Jensen. Efficient molecular dynamics scheme for the calculation of dopant profiles due to ion implantation. *Physical Review E*, 57(6):7278, 1998.

[118] P. M. Derlet and S. L. Dudarev. Microscopic structure of a heavily irradiated material. *Physical Review Materials*, 4:023605, Feb 2020.

[119] Elton Chen, Chaitanya Deo, and Remi Dingreville. Reduced-order atomistic method for simulating radiation damage in metals. *Journal of Physics: Condensed Matter*, 32, 10 2019.

[120] Jean-Paul Crocombette, Alain Chartier, and William J Weber. Atomistic simulation of amorphization thermokinetics in lanthanum pyrozirconate. *Applied physics letters*, 88(5):051912, 2006.

[121] VI Belko and A Yu Kuznetsov. Frenkel pair accumulation in ion-and electron-irradiated sic. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 248(1):77–82, 2006.

[122] Alain Chartier, Tomokazu Yamamoto, Kazuhiro Yasuda, Constantin Meis, and Syo Matsumura. Frenkel pair accumulation induced crystallization of amorphous mgal2o4. *Journal of nuclear materials*, 378(2):188–192, 2008.

[123] A Chartier, Claire Onofri, L Van Brutzel, Ch Sabathier, O Dorosh, and J Jagielski. Early stages of irradiation induced dislocations in urania. *Applied Physics Letters*, 109(18):181902, 2016.

[124] J Byggmästar, F Granberg, and K Nordlund. Effects of the short-range repulsive potential on cascade damage in iron. *Journal of Nuclear Materials*, 508:530–539, 2018.

[125] Fredric Granberg, Jesper Byggmästar, and Kai Nordlund. Cascade overlap with vacancy-type defects in fe. *The European Physical Journal B*, 92(7):1–7, 2019.

[126] Alain Chartier and M-C Marinica. Rearrangement of interstitial defects in alpha-fe under extreme condition. *Acta Materialia*, 180:141–148, 2019.

[127] Guojia Ge, Feida Chen, Xiaobin Tang, Hai Huang, Jiwei Lin, Shangkun Shen, and Jing Gao. Effects of interstitial carbon on the radiation tolerance of carbon-doped nife binary alloys from atomistic simulations. *Nuclear Materials and Energy*, 24:100785, 2020.

[128] Daniel R Mason, Abdallah Reza, Fredric Granberg, and Felix Hofmann. Estimate for thermal diffusivity in highly irradiated tungsten using molecular dynamics simulation. *Physical Review Materials*, 5(12):125407, 2021.

[129] Daniel R Mason, Fredric Granberg, Max Boleininger, Thomas Schwarz-Selinger, Kai Nordlund, and Sergei L Dudarev. Parameter-free quantitative simulation of high-dose microstructure and hydrogen retention in ion-irradiated tungsten. *Physical Review Materials*, 5(9):095403, 2021.

[130] Daniel R Mason, Suchandrima Das, Peter M Derlet, Sergei L Dudarev, Andrew J London, Hongbing Yu, Nicholas W Phillips, David Yang, Kenichiro Mizohata, Ruqing Xu, et al. Observation of transient

and asymptotic driven structural states of tungsten exposed to radiation. *Physical Review Letters*, 125(22):225503, 2020.

[131] Céline Varvenne, Aitor Luque, Wolfram G Nöhring, and William A Curtin. Average-atom interatomic potential for random alloys. *Physical Review B*, 93(10):104201, 2016.

[132] MI Mendelev, S Han, DJ Srolovitz, GJ Ackland, DY Sun, and M Asta. Development of new interatomic potentials appropriate for crystalline and liquid iron. *Philosophical magazine*, 83(35):3977–3994, 2003.

[133] Qing-Jie Li, Howard Sheng, and Evan Ma. Strengthening in multi-principal element alloys with local-chemical-order roughened dislocation pathways. *Nature Communications*, 10(1):1–11, 2019.

[134] Wu-Rong Jian, Zhuocheng Xie, Shuozhi Xu, Yanqing Su, Xiaohu Yao, and Irene J Beyerlein. Effects of lattice distortion and chemical short-range order on the mechanisms of deformation in medium entropy alloy cocrni. *Acta Materialia*, 199:352–369, 2020.

[135] A. Chartier, C. Meis, J.-P. Crocombette, W. J. Weber, and L. R. Corrales. Molecular dynamic simulation of disorder induced amorphization in pyrochlore. *Physical Review Letters*, 94:025505, Jan 2005.

[136] Samuel M Allen, Robert W Balluffi, and W Craig Carter. *Kinetics of materials*. John Wiley & Sons, 2005.

[137] Shijun Zhao, Yuri Osetsky, and Yanwen Zhang. Preferential diffusion in concentrated solid solution alloys: Nife, nico and nicocr. *Acta Materialia*, 128:391–399, 2017.

[138] JM Cowley. An approximate theory of order in alloys. *Physical Review*, 77(5):669, 1950.

[139] Flynn Walsh, Mark Asta, and Robert O Ritchie. Magnetically driven short-range order can explain anomalous measurements in crconi. *Proceedings of the National Academy of Sciences*, 118(13), 2021.

[140] Jaime Marian, Brian D Wirth, and J Manuel Perlado. Mechanism of formation and growth of¡ 100¿ interstitial loops in ferritic materials. *Physical review letters*, 88(25):255507, 2002.

[141] K Arakawa, T Amino, and H Mori. Direct observation of the coalescence process between nanoscale dislocation loops with different burgers vectors. *Acta Materialia*, 59(1):141–145, 2011.

[142] DS Gelles. A frank loop unfaulting mechanism in FCC metals during neutron irradiation. In *Dislocation Modelling of Physical Systems*, pages 158–162. Elsevier, 1981.

[143] David Rodney. Molecular dynamics simulation of screw dislocations interacting with interstitial frank loops in a model FCC crystal. *Acta Materialia*, 52(3):607–614, 2004.

[144] Tomoko Kadoyoshi, Hideo Kaburaki, Futoshi Shimizu, Hajime Kimizuka, Shiro Jitsukawa, and Ju Li. Molecular dynamics study on the formation of stacking fault tetrahedra and unfaulting of frank loops in FCC metals. *Acta materialia*, 55(9):3073–3080, 2007.

[145] Qingqing Ding, Xiaoqian Fu, Dengke Chen, Hongbin Bei, Bernd Gludovatz, Jixue Li, Ze Zhang, Easo P George, Qian Yu, Ting Zhu, et al. Real-time nanoscale observation of deformation mechanisms in crconi-based medium-to high-entropy alloys at cryogenic temperatures. *Materials Today*, 25:21–27, 2019.

[146] FX Zhang, Shijun Zhao, Ke Jin, H Xue, G Velisa, H Bei, R Huang, JYP Ko, DC Pagan, JC Neuefeind, et al. Local structure and short-range order in a nicocr solid solution alloy. *Physical review letters*, 118(20):205501, 2017.

[147] Jun Ding, Qin Yu, Mark Asta, and Robert O Ritchie. Tunable stacking fault energies by tailoring local chemical order in CrCoNi medium-entropy alloys. *Proceedings of the National Academy of Sciences*, 115(36):8919–8924, 2018.

[148] Spencer L Thomas and Srikanth Patala. Vacancy diffusion in multi-principal element alloys: The role of chemical disorder in the ordered lattice. *Acta Materialia*, 196:144–153, 2020.

[149] Yanwen Zhang, G Malcolm Stocks, Ke Jin, Chenyang Lu, Hongbin Bei, Brian C Sales, Lumin Wang, Laurent K Béland, Roger E Stoller, German D Samolyuk, et al. Influence of chemical disorder on energy dissipation and defect evolution in concentrated solid solution alloys. *Nature Communications*, 6(1):1–9, 2015.

[150] Yanwen Zhang, Yuri N Osetsky, and William J Weber. Tunable chemical disorder in concentrated alloys: defect physics and radiation performance. *Chemical Reviews*, 122(1):789–829, 2021.

[151] Y Chen, K Y Yu, Y Liu, S Shao, H Wang, MA Kirk, J Wang, and X Zhang. Damage-tolerant nanotwinned metals with nanovoids under radiation environments. *Nature Communications*, 6(1):1–8, 2015.

[152] Manuel J Louwerse and Evert Jan Baerends. Calculation of pressure in case of periodic boundary conditions. *Chemical Physics Letters*, 421(1-3):138–141, 2006.

[153] Nir Goldman, Laurence E Fried, and Lucas Koziol. Using force-matched potentials to improve the accuracy of density functional tight binding for reactive conditions. *Journal of chemical theory and computation*, 11(10):4530–4535, 2015.

[154] Jonathan Vandermause, Steven B Torrisi, Simon Batzner, Yu Xie, Lixin Sun, Alexie M Kolpak, and Boris Kozinsky. On-the-fly active learning of interpretable bayesian force fields for atomistic rare events. *npj Computational Materials*, 6(1):1–11, 2020.

[155] Elizabeth L Yip. A note on the stability of solving a rank-p modification of a linear system by the sherman–morrison–woodbury formula. *SIAM Journal on Scientific and Statistical Computing*, 7(2):507–513, 1986.

[156] Eliakim H Moore. On the reciprocal of the general algebraic matrix. *Bull. Am. Math. Soc.*, 26:394–395, 1920.

[157] Eric R Ziegel. Matrix algebra from a statistician's perspective. *Technometrics*, 40(2):164, 1998.

[158] Nicolas Bertin. Connecting discrete and continuum dislocation mechanics: A non-singular spectral framework. *International Journal of Plasticity*, 122:268–284, 2019.