Quantification of Elastic Incompatibilities at Triple Junctions via Physics-Based Surrogate Models

by

Aaditya Rau

B.S. Engineering Mechanics, Johns Hopkins University (2022)

Submitted to the Center for Computational Science and Engineering
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN COMPUTATIONAL SCIENCE AND ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2024

Authored by:	Aaditya Rau Center for Computational Science and Engineering May 21, 2024
Certified by:	Raul A. Radovitzky J. C. Hunsaker Professor of Aeronautics and Astronautics, Thesis Supervisor
Accepted by:	Youssef Marzouk Professor of Aeronautics and Astronautics Co-Director, Center for Computational Science and Engineering

Quantification of Elastic Incompatibilities at Triple Junctions via Physics-Based Surrogate Models

by

Aaditya Rau

Submitted to the Center for Computational Science and Engineering
on May 21, 2024 in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN COMPUTATIONAL SCIENCE AND ENGINEERING

ABSTRACT

Stresses at grain boundaries resulting from elastic incompatibilities have long been known to drive the premature failure and loss of desirable macroscopic properties in polycrystalline materials. As a result, there have been significant efforts in the field of grain boundary engineering to understand the sources of grain boundary incompatibilities in polycrystals and potential mitigation strategies through microstructure manipulation. Thus, understanding the relationship between grain incompatibility and failure is important for the practical use of polycrystalline materials.

Surrogate models based on machine learning methods have gained broad popularity due to their ability to furnish a functional, albeit approximate, description of complex phenomena. The goal of this thesis is to predict quantitative metrics of incompatibility from various triple junction configurations using a surrogate model. High-fidelity finite element simulations of a cubic-crystal triple junction under hydrostatic extension were used to generate a synthetic dataset for training the surrogate model. A set of J integrals computed around microcracks placed along the triple junction boundaries were used to quantify the elastic incompatibilities between the grains. A multi-layer perceptron network was trained using the grain rotation angles and J integrals as the feature and label data respectively. We demonstrate that the trained network establishes an accurate functional dependence between the triple junction angles and the J integrals. We use the surrogate model to efficiently sweep the configuration space and create contour maps of the largest stress intensification at the triple junction as a function of the grain rotation angles. Furthermore, we show that the surrogate model can be utilized to identify the most and least compatible triple junction configurations via optimization. These configurations are then compared to those identified as favorable through the theory of coincident site lattices.

Thesis supervisor: Raul A. Radovitzky
Title: J. C. Hunsaker Professor of Aeronautics and Astronautics

Acknowledgments

I would first like to thank my advisor, Raúl Radovitzky, for giving me the opportunity to study at a world-class institution like MIT as part of his group. I am thankful for his mentorship, his technical insights, and positive attitude. Thank you, Raúl, for helping me become a better mechanician, scientist, and academic.

Furthermore, I would like to thank Prof. Chris Schuh for his feedback on this work, particularly on how to connect it to the materials science and grain boundary engineering communities.

Next, I would like to thank the students of the RR Group, who have helped make working in computational mechanics exciting, interesting, and fun. In particular, special thanks go to Christopher Quinn, who provided me with the interpolation code that was crucial for computing the *J*-integrals that are fundamental to this work. Next, I would like to thank Théo Rulko for his companionship and for answering all my software questions. I would like to thank Daniel Pickard for his great sense of humor, his passion for scholarship, and his insistence on having fun while doing mechanics. Finally, I appreciate the energy that Madison Hobbs, Jake Sonandres, and Summer Hoss have contributed to the group.

I would like to thank my greatest supporters: my family, both in the U.S. and abroad. In particular, I cannot thank my parents enough for all they have done to support my education. Their love and wisdom have been emboldening at every step, in academics and beyond. Finally, I would like to thank my brother, Advaith, who has been the best sibling that I could ask for.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Role of Grain Incompatibilities in Material Failure

The role of grain boundaries as sources of incompatibility stresses resulting from crystal anisotropy and grain misorientation has long been established. These stresses can severely compromise the macroscopic properties of polycrystals, such as ductility, by inducing premature intergranular fracture originating at grain boundaries [1–9]. Examples of such grain boundary fractures are given in Figure 1.1. These failures are in contrast to single crystalline structures, where the desirable properties of many materials are realized to their maximum extent due to the absence of internal grain boundary constraints. However, multiple reasons preclude the use of single crystals alone in engineering applications: conventional materials synthesis and processing methods can only produce polycrystalline microstructures [10]; manufacturing single crystals is often expensive, time-consuming, and impractical for high-throughput applications [11–14]; polycrystallinity often imparts desirable properties, such as so-called grain boundary strengthening in metals [15–18].

In recognition of these challenges, the field of "grain boundary engineering" (GBE) has emerged. A key objective of GBE is to manipulate microstructures using materials processing techniques to address the aforementioned limitations on polycrystals [4, 10, 20–24]. Detailed chronologies of the development of GBE have been provided by Randle [23] and Watanabe [10]. Techniques utilized in GBE include changes to the processing, morphology, and/or chemistry of materials to achieve desirable microstructures. An example of GBE through processing is the work by Schuh on copper-based shape memory alloys, which showed that the crystallographic texture produced by melt-casting enabled large recoverable strains without failure [25]. Recent work on superelastic zirconia polycrystals [26, 27] demonstrates the use of morphology to reduce cracking. By creating smaller zirconia samples with oligocrystalline microstructures, the martensitic transformation incompatibilities were mitigated; the resulting zirconia samples withstood 50 to 100 load cycles without fracture [27] and achieved entirely recoverable strains on the order of 10% [26]. For comparison, zirconia polycrystals often crack within a few cycles and at strains below 2%. Experimental demonstrations of both the improved ultimate strain and cycle life of grain boundary-engineered zirconia are given in Figure 1.2. Nanocrystalline alloys are another class of

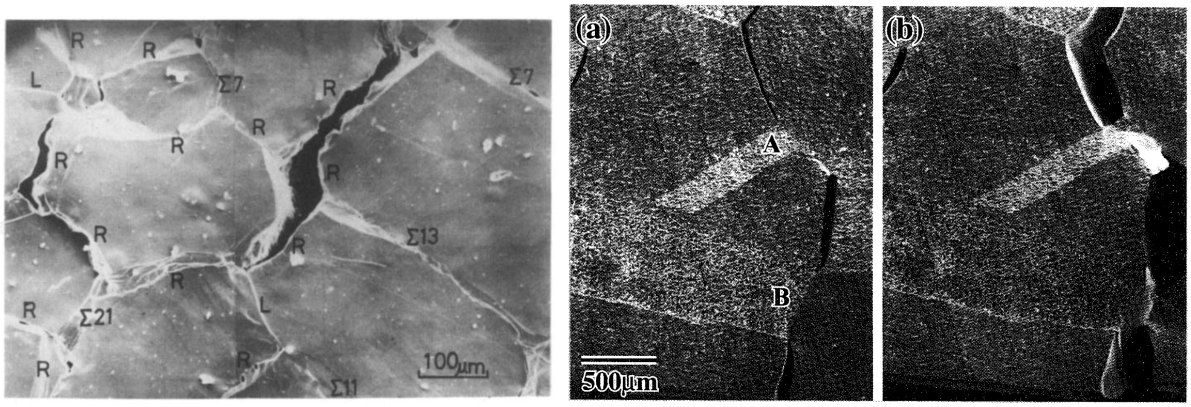

Figure 1.1: Examples of grain boundary fracture. The left is from a polycrystal of Fe-0.8 at.% Sn [5], the right is a copper polycrystal [19].

materials that have benefitted from grain boundary engineering [28]. While these materials have superlative properties including strength [29–32], hardness [33–36], and wear resistance [31, 37, 38], these grain boundary structures are typically unstable. However, these structures can be stabilized by alloying a nanocrystalline metal with a species that preferentially segregates to the grain boundary [39–42].

Grain boundary incompatibilities are clearly determined by the grain boundary configurations, through the anisotropy mismatch between adjacent grains. It is, therefore, of great importance to develop methodologies that relate microstructural features and the resulting grain boundary incompatibilities. Through extensive and continued experimental work [3, 7, 9, 43–49] our understanding of grain boundary incompatibilities and how to obviate them through GBE has improved immensely. Furthermore, significant analytical [50–61] and numerical [62–71] work has been conducted to study grain boundary incompatibilities for a variety of materials and grain boundary geometries.

In the specific context of grain boundary incompatibility, analytical works that describe the stress and strain fields at grain interfaces have primarily focused on bicrystals [50, 54–57] and triple junctions [51, 52, 58, 59], although other simple configurations have been considered [53, 60, 61]. These analyses solve boundary value problems of linear anisotropic elastostatics to characterize the stress concentration associated with grain incompatibility. For completeness, in the absence of body forces, these equations are as follows:

$$\begin{aligned} \varepsilon &= \frac{1}{2}\left(\nabla \mathbf{u} + \nabla \mathbf{u}^{\top}\right), \\ \nabla \cdot \sigma &= \mathbf{0}, \\ \sigma &= \mathbb{C} : \varepsilon, \end{aligned} \tag{1.1}$$

where $\mathbf{u}$ is the displacement field, ε is the infinitesimal strain field, σ is the stress, and $\mathbb{C}$ are the elastic moduli. In general, $\mathbb{C}$ contains 21 independent constants, however simplifications to $\mathbb{C}$ can be derived for certain crystal structures through imposing crystal symmetry [72]. However, grain boundary interfaces introduce discontinuities in the solution to (1.1). Thus, analytical

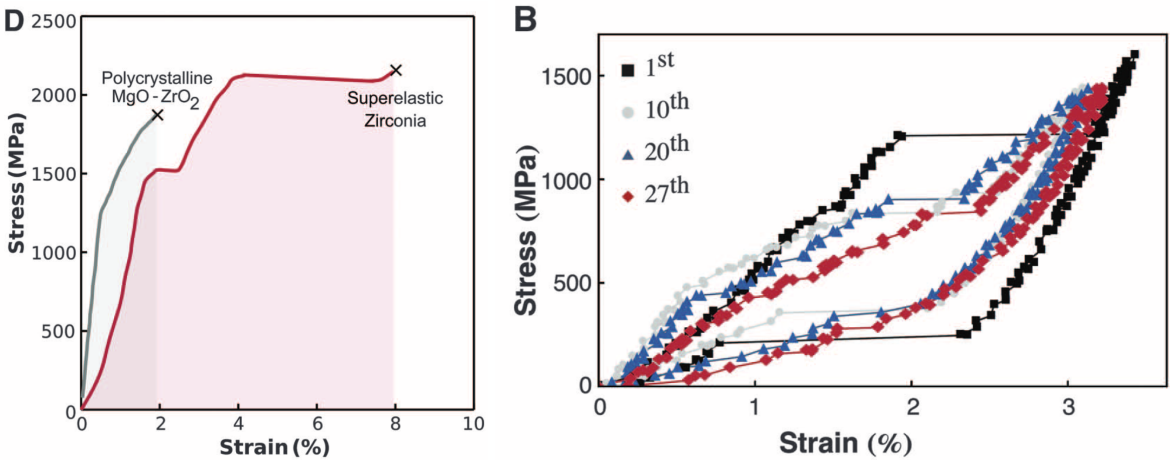

Figure 1.2: Stress-strain curve for both standard polycrystalline zirconia and grain boundary-engineered zirconia, showing the premature fracture of the zirconia polycrystal (left). Stress-strain curves for a cyclically loaded grain boundary-engineered zirconia pillar, demonstrating stable, repeatable behavior after many cycles. Both figures are from [26].

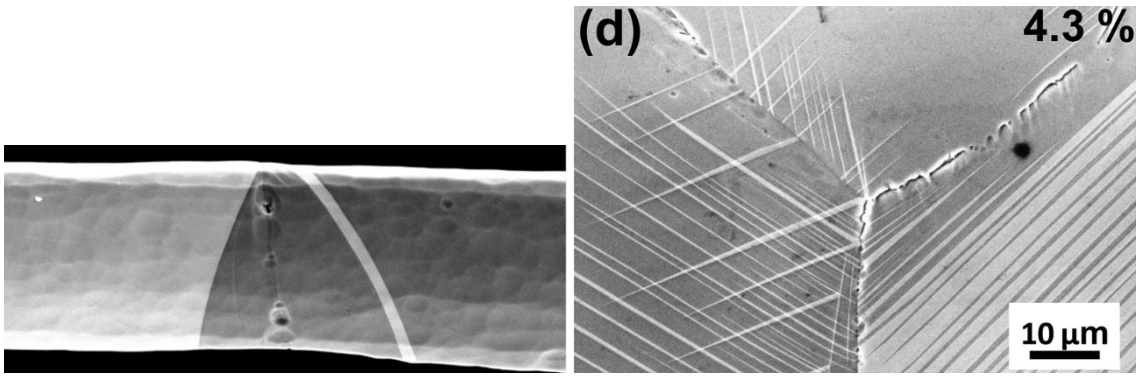

Figure 1.3: Bicrystals (left) and triple junctions (right) are two common classifications of grain boundaries found in polycrystals [6].

study of the mechanical fields in bodies with interfaces requires a mathematical description of the jump conditions at the interface. These jump conditions are often derived though analysis of the the configuration shown in Figure 1.4, which depicts a region consisting of two materials with interface normal $\mathbf{n}$. If the displacement, strain, and stress fields in each material are denoted as $\mathbf{u}^{(i)}, \varepsilon^{(i)}, \sigma^{(i)}$ $(i = 1, 2)$, then physical considerations induce the following restrictions on these fields when evaluated at the interface[1]:

$$\begin{aligned} \text{Displacement continuity} &\implies \begin{cases} \mathbf{u}^{(1)} = \mathbf{u}^{(2)} \\ \varepsilon^{(1)} - \varepsilon^{(2)} = \frac{1}{2}\left(\mathbf{a} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{a}\right) \end{cases} \\ \text{Force equilibrium} &\implies \sigma^{(1)}\mathbf{n} = \sigma^{(2)}\mathbf{n}. \end{aligned} \tag{1.2}$$

The first condition enforces that the interface remain bonded through the deformation, i.e. that no cracks form. Displacement continuity also implies the second condition, which enforces that the out-of-plane strain components cannot jump across the interface. The third condition on the stress enforces that the stress components normal to the interface remain constant. It is worth mentioning that alternate formulations of the jump conditions (1.2) have been derived [76, 77].

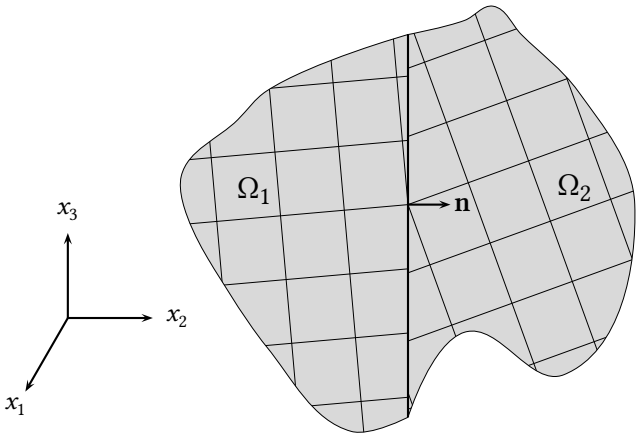

Figure 1.4: An internal interface within a body such as those found at grain boundaries, where properties and orientations can vary discontinuously across the interface $\partial\Omega_1 \cap \partial\Omega_2$. The cross-hatching indicates the different material orientations across the interface.

The relations (1.2) were used by Gemperlova et al [50] to identify the so-called compatibility stresses induced by the elastoplastic deformation of anisotropic bicrystals. Their analysis contemplates two half-spaces welded at an interface $x_2 = 0$ (i.e. the vector $\mathbf{n}$ in Figure 1.4 and equations (1.2) is $\mathbf{e}_2$), each of which possesses both different elastic compliances and a different

[1]While this thesis is primarily concerned with infinitesimal deformations, large deformation analogs of (1.2) can be derived. In particular, the large deformation version of the strain condition $(1.2)_2$ is known as the Hadamard jump condition and is foundational to the study of materials that develop internal interfaces in response to external stimuli, such as the formation of martensite in shape memory materials [73, 74] and domains in ferroelectrics [75].

set of slip systems that govern the plastic response. By recourse to the superposition of a "compatibility" stress and strain field onto the stresses and strains in each material and an application of the constitutive law, they derive the following system for the non-zero compatibility stresses τ_{ij}:

$$\begin{pmatrix} \tau_{11} \\ \tau_{33} \\ \tau_{13} \end{pmatrix} = -\mathrm{sgn}(x_2) \begin{pmatrix} s^+_{11} & s^+_{13} & s^+_{15} \\ s^+_{13} & s^+_{33} & s^+_{35} \\ s^+_{15} & s^+_{35} & s^+_{55} \end{pmatrix}^{-1} \begin{pmatrix} \Delta e_{11} \\ \Delta e_{33} \\ \Delta e_{13} \end{pmatrix}, \tag{1.3}$$

where s^+_{ij} is the sum of the elastic compliances (in Voigt notation) of both materials and Δe_{ij} is the jump in incompatibility strain across the interface. Using the system (1.3), specific solutions were obtained for cubic bicrystals rotated about the $\langle 100 \rangle$ and $\langle 110 \rangle$ axes. A similar analysis of elastic bicrystals with cubic anisotropy was conducted by Liu et al. [57], who identified via optimization that a bicrystal arrangement of $\langle 100 \rangle$ versus $\langle 111 \rangle$ led to the largest incompatibility stress for any cubic material, i.e. independently of the specific elastic constants. Richeton and Berbenni adopted a different approach by using Field Dislocation Mechanics [78], which augments the standard field equations of solid mechanics with additional fields related to so-called geometrically necessary dislocations. Using this theory, they decomposed the incompatibility stress in general anisotropic bicrystals into elastic and plastic contributions [55] and computed the incompatibility stress in a cubic bicrystal as a function of a single misorientation angle [56]. Evans [58] calculated the stress intensification at grain boundaries due to thermal anisotropy mismatch by introducing defects at the boundary and approximating the mechanical response of the grains as isotropic. The analysis identified triple junctions as the main site for microcrack nucleation and estimated the critical crack size for microfracture based on material properties, defect size distribution, and grain boundary energy. Tvergaard and Hutchinson [59] used the singularity analysis technique from Williams [79] to characterize the stress singularity at triple junctions in cubic and orthotropic crystals subject to mechanical and thermal loading. In this technique, the radial and circumferential displacement components of the displacement are assumed to have a separable form:

$$(u_r, u_\theta) = k r^{\lambda+1} (\tilde{u}_r, \tilde{u}_\theta), \tag{1.4}$$

where k is an "amplitude factor," λ is the singularity exponent, and $\tilde{u}_r(\theta), \tilde{u}_\theta(\theta)$ encode the angular variation of the displacement field components. Substitution of the above displacement into the strain-displacement relation and the constitutive relation for each grain results in a system of differential equations, which is supplemented with boundary conditions arising from the jump conditions (1.2) and the symmetry attendant to the problem under their consideration. Through this analysis, they identified the strength of the triple junction singularity (as measured by λ) as a function of grain anisotropy and misorientation. Vakaeva et al. [52] derived an approximate stress distribution around triple junctions in the presence of pores with three-fold symmetry using Goursat-Kolosov [80] stress potential functions, but did not investigate the role of anisotropy in the resulting stress distribution.

Analytical results not only provide insights and quantitative metrics of incompatibility but also a parameterization of these metrics on the problem inputs, which can include the grain

boundary configuration, material properties, and applied loading. This functional dependence enables the identification of configurations that mitigate incompatibilities with simple optimization, e.g. [57]. However, analytical models are often subject to significant simplifications of material responses and geometries. By contrast, computational approaches have emerged as an important tool to quantify complex stress fields at the microstructural scale in bicrystals [62–64], triple junctions [65, 66, 69], and general polycrystals [70, 71, 81–88]. Early computational work regarding polycrystals can be found in the seminal work of Sir G.I. Taylor on plastic deformation in metals [89]. In particular, Sir Taylor considers the effect of applied loads on the development of plastic slip in polycrystals. Using "Mallock's equation-solving machine," an early analog computer [90], Sir Taylor identifies the active slip planes for crystal aggregates under isochoric deformation. Although this analysis does not consider explicit interactions between grains, it represents an early effort towards the use of computing to understand the response of polycrystals.

Typically, finite element methods have been used to quantitatively describe the stress and strain fields along grain boundaries in micromechanics problems. Early applications of the finite element method to grain-scale problems often studied incompatibilities in multi-phase materials such as steels [91], tungsten carbides [92, 93], as well as other metal alloys [94, 95]. In these cases, elastic incompatibilities are induced by property differences, i.e. stiffness, between phases. Meyers and Ashworth [96] conducted a set of plane strain simulations of bicrystals to identify the extent to which elastic incompatibilities can be used to explain the dependence of the yield stress in metals on the grain size. In their simulations, the orientations of the constituent grains in the bicrystal were fixed to be [100] and [111]. By including a thin, compliant layer between the bicrystals to model grain boundary segregation or denudation and varying the domain size, the shear strain at the interface was observed to decrease with grain size. Based on this result, Meyers and Ashworth suggest that grain boundary segregation and denudation are physical mechanisms that affect the size dependence of yield strength in metals. It is worth mentioning that in these early computational efforts, feasible finite element simulations were constrained to be two-dimensional and utilized coarse meshes, which precluded the ability to obtain high-resolution stress and strain fields.

Improved computational resources enabled the simulation of three-dimensional problems with more complex geometries and improved resolution. In the study of Peralta et al [62], three-dimensional finite element simulations of copper bicrystals in both twist and tilt grain boundaries were conducted. Upon subjecting the bicrystals to uniaxial tension perpendicular to the grain boundary plane, they found that the average values of the stresses in the interior of the bicrystal were in agreement with the predictions of Gemperlova et al [50], but that there were large stress gradients near the free boundary. In the aforementioned study of Tvergaard and Hutchinson [59], they also performed simulations to characterize the critical defect size for grain-boundary microfracture initiated at triple junctions. They quantified the level of incompatibility at a triple junction using finite element simulations combined with the calculation of the J-integral [97] around a single crack placed along one of the triple junction boundaries. They demonstrated that elastic anisotropy mismatch can further decrease the critical defect size for microfracture initiation beyond thermal expansion mismatch alone.

Computational methods have been used to analyze the mechanical fields over a large set of grain orientations that are not accessible experimentally due to challenges related to synthesis, processing, and/or cost. Optimal configurations are typically identified by recourse to parametric sweeps over the input space [63, 98–103]. These sweeps typically result in significant computational costs, however through improved sampling techniques, i.e. Monte Carlo methods, the portions of the input space requiring sampling can be reduced [101, 104]. In the study of Wang et al [63] on elastic and transformation incompatibilities in superelastic zirconia, multiple bicrystal configurations were simulated under uniaxial compression using finite element analysis. Through these simulations, the examined orientations could be classified into those that induced large elastic incompatibilities and those with small elastic compatibilities but large phase-transformation-related incompatibilities. Finite element analysis has also been used in the microstructural design of ferroelectrics [102]. It is worth mentioning that in addition to crystallographic orientation, microstructure optimization incorporating variables such as volume fractions of various phases and the inclusion of fibers or other inclusions has been considered [98, 100], in addition to approaches that integrate the microstructural problem within a multi-scale model [99].

A drawback to physics-based computational models is the requirement of a separate simulation for each set of input parameters. As a result, the microstructure optimization problem is challenging due to the lack of a parametric functional dependence on the problem inputs. Surrogate models based on machine learning methods can approximate the response of physical systems and have become valuable in diverse fields [105–121]. Generally, surrogate models are a class of high-dimensional interpolants that can be used to develop an approximate functional representation of an unknown relationship. This functional relationship can be developed algorithmically by "learning" from previously acquired data in a process called "training". Trained surrogate models furnish, in effect, a functional relationship between feature and label sets which is much cheaper to evaluate than the original physics model. In cases where experimental data is sparse, synthetic data from numerical simulations can be used for generating the label set. The typical label generation procedure in this case involves evaluating a high-fidelity, expensive numerical model on a large sample of the feature set, encompassing a suitable range of input parameters. The resulting outputs are collected with the associated features to form the training dataset for the surrogate model. Furthermore, surrogate models enable the solution of inverse problems. For example, Peurifoy et al [122] demonstrated that the analytical properties of a multilayer perceptron network can be utilized to obtain gradients, which were used to solve inverse design problems involving nanophotonics. A similar approach was adopted in [118] for the inverse design of optical meta-surfaces. Similar approaches have also been used in the materials design community, where gradient descent has been used to identify novel compositions and chemical structures [123, 124] with optimal properties. A different approach was taken in [120], which extended the physics-informed neural network methodology to solve inverse problems by supplementing the training loss function with an additional term that penalizes deviations of the model output from the desired set of values.

The specific application of machine learning to grain boundary engineering has seen multiple recent successes. In the realm of classification models, where the model output is discrete, Mangal and Holm [125, 126] used a random forest method to identify whether grains were "stress

hotspots" in both face-centered cubic [125] and hexagonal close-packed [126] polycrystals based on data from crystal plasticity simulations. A similar approach was taken by Zhang et al. [127] to predict whether grain boundaries in magnesium polycrystals would develop strain incompatibilities or stress concentrations. In the category of regression models, where the model output is continuous, Frankel et al. [128] employed crystal plasticity simulations to generate stress-strain responses for steel oligocrystals and predicted these responses utilizing a hybrid network architecture with features extracted from images of synthetic microstructures. Donegan et al. [129] followed a similar approach in a study of thermally-induced incompatibility stresses. Random forest regression models were used in the work of Guziewski et al. [104] to predict the energy and strength of grain boundaries in silicon carbide using a combination of molecular dynamics simulations, crystallographic data, and physical chemistry information. A neural network was used with experimental digital image correlation data by Vieira and Lambros [130] to learn the components of the plastic strain field at grain boundaries in steel as a function of the angle between the grain boundary and the applied loading direction. Liu et al. [131] employed a network trained on crystal plasticity simulations to identify microstructures that optimized chosen properties.

1.2 Thesis Objectives and Approach

We find a number of deficiences in the approaches outlined above in the setting of the microstructure optimization problem pertinent to grain boundary engineering; analytical methods are often restricted to simple geometries and loading, and classical numerical approaches do not furnish a parametric relationship between the grain boundary configuration and the incompatibilities. Although surrogate models generally do provide such a relationship, classification networks that produce binary output cannot be subject to standard optimization methods to minimize incompatibilities. Regression networks realize these aims, providing a functional relationship between the model inputs and outputs that is continuous and can thus be optimized.

The goal of this thesis is to develop and apply a framework that combines physics-based simulations of grain incompatibility with surrogate modeling to predict quantitative metrics of incompatibility from grain boundary configurational data, for the purposes of identifying optimal grain orientations. As microfracture typically nucleates from triple junctions in polycrystals, we return to the triple junction model studied by Hutchinson and Tvergaard [59] to characterize grain incompatibilities. Concretely, we aim to predict the incompatibilities at a triple junction composed of cubic crystals rotated about the [001] axis by arbitrary angles through the training of a neural network. These three angles will be adopted as the feature set. Regarding the label set, we adopt the J integral [97] as a quantitative metric of incompatibility. In particular, we consider three values of J computed around three microcracks placed radially along the triple junction interfaces in order to analyze the stress intensification arising from each of the triple junction boundaries. The training dataset was formed by conducting 100,000 finite element simulations of the described triple junction. The TensorFlow [132] and Keras [133] packages were used to train a multi-layer perceptron network to predict each J-integral given the grain rotation angles.

We show that the model trained on this dataset can predict the value of J for the three microcracks with high accuracy on the validation set. We then demonstrate that the trained model can be optimized with respect to the feature set to identify configurations that minimize or maximize grain incompatibilities. In Chapter 2, the physics-based model of grain incompatibility is presented. Firstly, the boundary value problem used to simulate the triple junction behavior is presented. Some salient features of the finite element solution to this problem are shown. Then, the J integral is introduced and its numerical implementation is discussed and then verified. In Chapter 3, the fundamentals and training procedure for the multi-layer perceptron network used in this thesis are then presented. The performance of the network is then assessed on the training and validation sets as a verification of the chosen network architecture. In Chapter 4, the computational framework is applied concretely to the problem of quantifying elastic incompatibilities at triple junctions using the surrogate model. Demonstrations of forward predictions from the trained surrogate model are presented. Then, the integration of the surrogate model into an optimization routine to identify the most and least compatible configurations is presented, and the ensuing configurations are compared to those predicted from the crystallographic theory of coincident site lattices. Chapter 5 presents the summary and main conclusions of this thesis.

Chapter 2

Physics-Based Model of a Triple Junction with Microcracks

This chapter summarizes the physics-based model of grain boundary incompatibility adopted in this thesis. Full-field and high-resolution simulations of the elastic response of the triple junction are required for the surrogate model to accurately predict incompatibilities. To this end, the finite element method is adopted, as implemented in the research code ∑MIT [134]. In typical analyses of triple junctions, the grains are taken to be conforming. This assumption, which is seldom physical, introduces a stress singularity [59]. This singularity poses challenges from a numerical perspective, as obtaining the mechanical fields at the triple junction requires extremely high resolution. A more reasonable approach is the one proposed by Tvergaard and Hutchinson [59], where microcracks are placed at the triple junction. These microcracks lead to asymptotic stress fields around each crack tip, whose strength can be characterized by the J-integral of fracture mechanics [97]. This enables the quantification of the level of incompatibilities at each of the three individual interfaces comprising the triple junction.

In this section, the boundary value problem used to model the triple junction is presented along with representative finite element solutions of a few grain boundary configurations. Then, theoretical and physical results regarding J are presented and the numerical implementation in ∑MIT is described. The numerical implementation is verified in two problems. The first problem uses the method of manufactured solutions using the analytical results for a semi-infinite crack [97]. The second problem verifies that the adopted numerical calculation reproduces the path-independence properties of J.

2.1 Statement of Boundary Value Problem

Training the surrogate model involves conducting high-resolution finite element simulations of the stress field at a triple junction to generate a synthetic dataset. The geometry of the triple junction is illustrated in Figure 2.1. The triple junction is modeled using a disk split equally into three portions, each representing a grain. This geometry represents the region local to a triple junction in the hexagonal array studied by Tvergaard and Hutchinson [59]. A similar 3D cylin-

drical model was employed by Li et al. to compute the stress and strain distributions around triple junctions [65]. Following the use of grain boundary defects by Evans [58] and Tvergaard and Hutchinson [59], three microcracks of length a were concurrently placed radially along the triple junction boundaries. Micrographs of microstructures indicate that triple junctions generally contain an initial flaw or pore [58, 135–145]. In particular, morphologies similar to the trifurcated geometry adopted herein have been shown to occur in experiments [146–152] and atomistic simulations of crack nucleation at triple junctions [153]. The disk was subjected to hydrostatic extension by prescribing a radial displacement $\bar{u} = R\varepsilon_{rr}$ on the boundary of the disk, where R is the radius of the disk and ε_{rr} is a radial strain. Elastic incompatibilities are introduced by rotating each grain about its [001] axis by an angle θ. The triple junction configuration is thus characterized by the tuple $(\theta_1, \theta_2, \theta_3)$, which we also denote by θ_i, where $i = 1, 2, 3$.

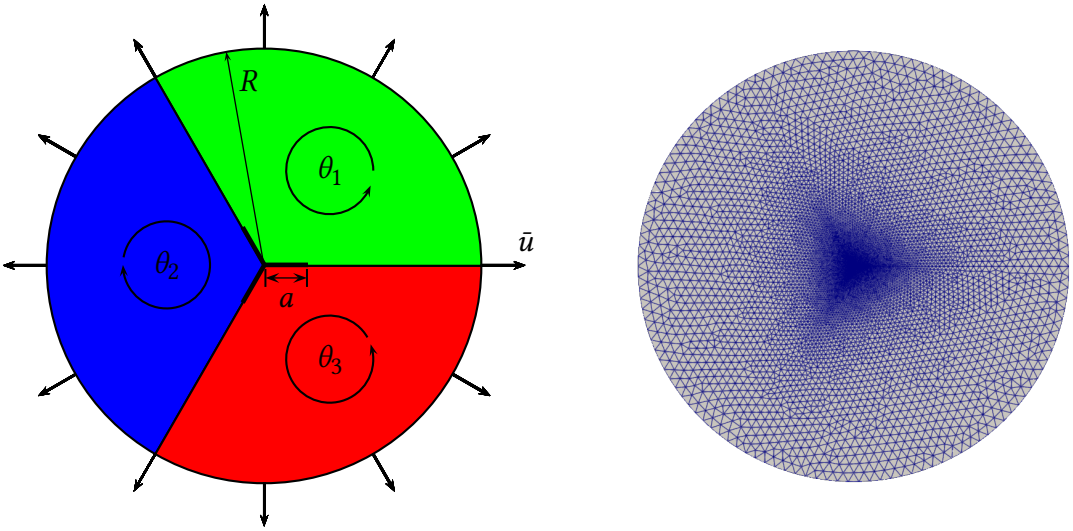

Figure 2.1: Schematic of the triple junction geometry with radial extension condition (left), and the mesh of this geometry used in the finite element simulations (right).

For concreteness, the mechanical response was simulated using the model for cubic materials proposed by Kambouchev et al. [154]. This finite deformation elasticity model satisfies the so-called polyconvexity property of the strain energy function W. This property affords theoretical guarantees regarding the existence of solutions to the nonlinear elasticity boundary value problem [155, 156]. Within the linear theory considered in this thesis, the constitutive model reduces to the standard linear elasticity model $\sigma = \mathbb{C} : \varepsilon$, where the elasticity tensor $\mathbb{C}$ for cubic materials is given by [157]:

$$\mathbb{C} = C_{12}1 \otimes \mathbf{1} + 2C_{44}\mathbb{I}^{\text{sym}} + (C_{11} - C_{12} - 2C_{44})\,(\mathbf{a} \otimes \mathbf{a} \otimes \mathbf{a} \otimes \mathbf{a} + \mathbf{b} \otimes \mathbf{b} \otimes \mathbf{b} \otimes \mathbf{b} + \mathbf{c} \otimes \mathbf{c} \otimes \mathbf{c} \otimes \mathbf{c})\,, \quad (2.1)$$

where C_{11}, C_{12}, and C_{44} are the cubic elastic constants, $\mathbf{1}$ is the second-order identity tensor, $\mathbb{I}^{\text{sym}}$ is the fourth-order symmetric identity tensor, and $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are the unit vectors defining the cubic

lattice. The degree of anisotropy of cubic crystals can be characterized by its anisotropy or Zener ratio $A = \frac{2C_{44}}{C_{11}-C_{12}}$, where $A = 1$ corresponds to isotropy [158] and materials with A significantly different from $A = 1$ are strongly anisotropic. The model proposed by Kambouchev [154] is suitable for cubic materials with $\frac{1}{2} \leq A \leq 1$, a condition which is satisfied, for instance, by a set of transition metals with a body-centered cubic crystal structure. In this thesis, we adopt the material properties of chromium for definiteness. The elastic constants of chromium, in addition to other materials that satisfy the aforementioned condition on A, are given in Table 2.1. It bears emphasis that while other classes of anisotropy could be considered in this framework, the cubic model was used for simplicity. Finally, it is worth mentioning that the nonlinear, consistent, elastic tangent moduli are provided in [154], although these moduli reduce to (2.1) in the linear theory. Access to these moduli enables the implicit solution of the boundary value problem described in the previous section. Implicit solution methods enable converged solutions that accurately resolve the crack tip fields, which will be used to compute the J-integral, as described in Section 2.3.

Table 2.1: The elastic properties and anisotropy ratios of some cubic materials and compounds that satisfy $\frac{1}{2} \leq A \leq 1$ [158, 159]

Material	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	A
Cr	350	57.8	101	0.6913
Mo	460	176	110	0.7746
W	521	201	160	1.0000
V	228	119	42.6	0.7817
Nb	246	134	28.7	0.5125
AgBr	56.3	33.0	7.2	0.6180
NaCl	48.7	12.4	12.6	0.6942

2.2 Finite Element Solution

The boundary value problem described in Section 2.1 was solved using the finite element method implemented in the $\sum$MIT research code. In particular, after effecting the finite element discretization, the resulting linear system was then solved using the PETSc [160] library. A sample solution showing the displacement and stress fields for the case $\theta_i = 0$ is given in Figure 2.2. As a result of the fine mesh used for the problem, the fields are well-resolved.

Hydrostatic stress contours of various triple junction configurations are given in Figure 2.3. An illustration of the effect of the grain orientation on the finite element solution can be seen in Figure 2.3; changes in the triple junction orientation cause the crack-tip singularity to rotate, demonstrating the asymmetric loading incurred by grain misorientation.

The hydrostatic stress field is shown near one of the crack tips in Figure 2.4. Far from the crack tips, the stress field becomes ostensibly hydrostatic, which is consistent with Saint Venant's principle.

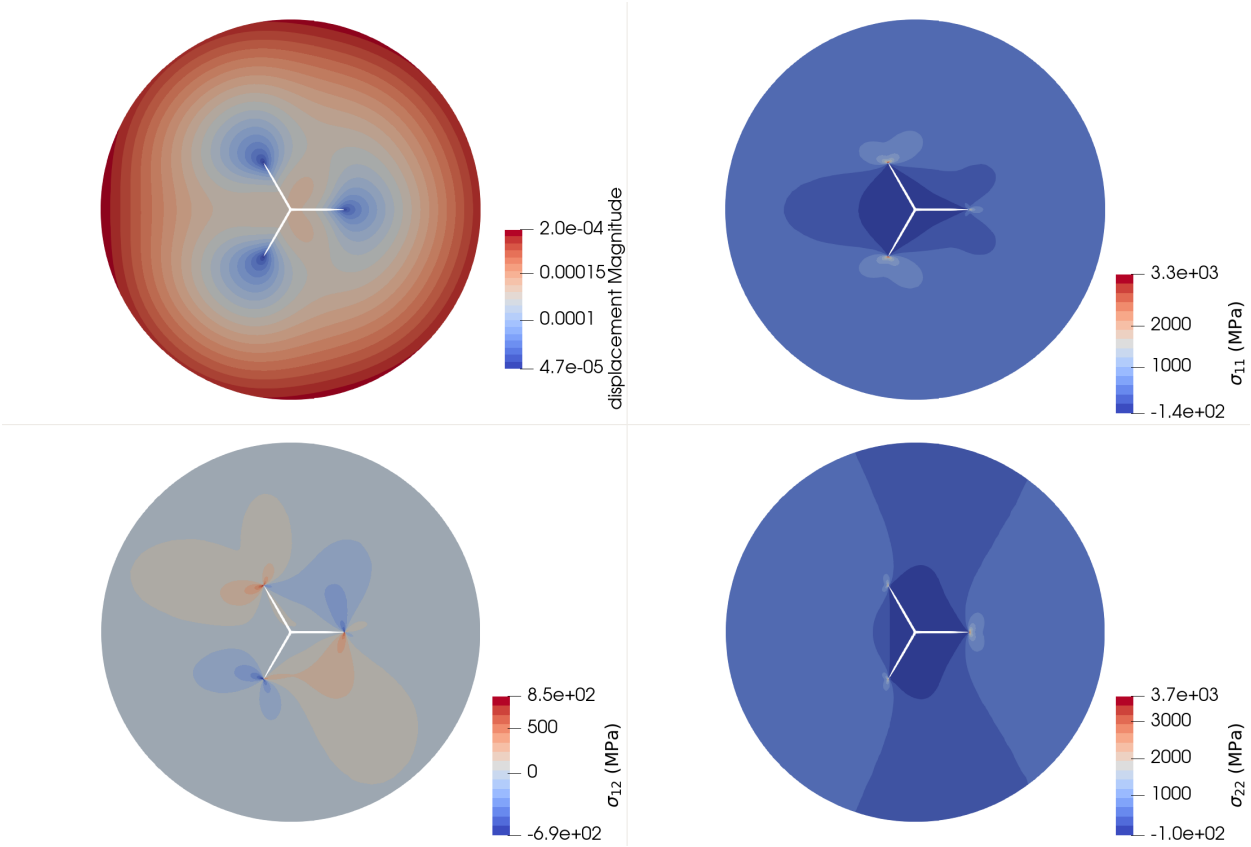

Figure 2.2: Displacement magnitude and stress contours for the case $(\theta_1, \theta_2, \theta_3) = (0, 0, 0)$, in the region local to the microcracks. A warp factor of 10 has been applied to aid in visualization.

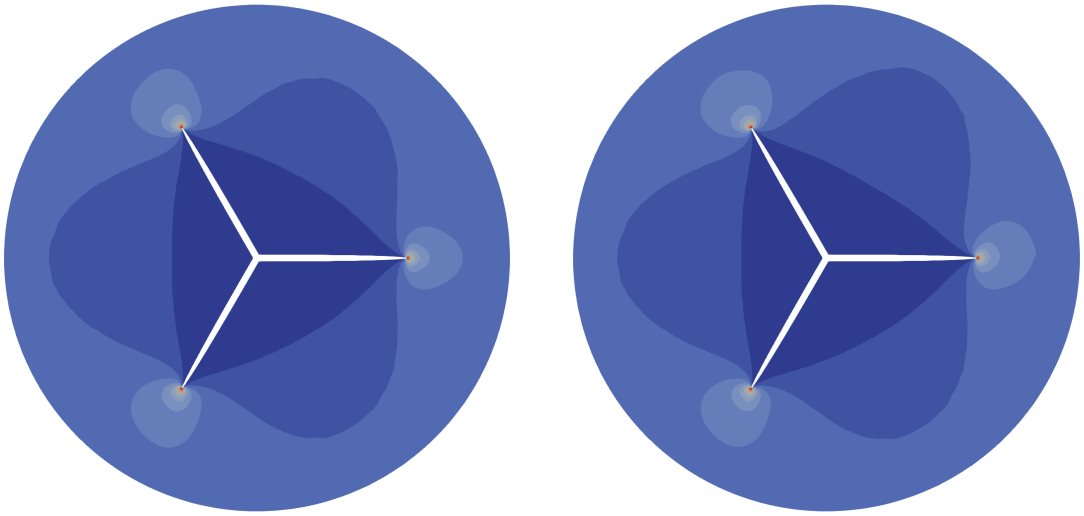

Figure 2.3: The region local to the triple junction shown on the warped geometry, for triple junction configurations $(\theta_1, 0, 0)$ with $\theta_1 = 0^\circ$, and 60°.

2.3 Characterization of Stress Intensification: The J-integral

The mathematical statement of solid mechanics involves force and moment balances, supplemented with compatibility relations and constitutive assumptions. In various circumstances, it has been useful to introduce additional quantities and related conservation laws to help describe the effect of cracks, defects, and internal material surfaces [161]. In this section, we describe the celebrated J-integral [97], which is of great importance in fracture mechanics[1]. In this work, J will be used to quantify elastic incompatibilities between the grains at the triple junction.

2.3.1 Definition and Interpretation of J

Boundary value problems involving cracks and notches are commonplace in solid mechanics due to both theoretical and practical interest. However, mathematical analysis of the mechanical fields can be complicated by the crack-tip singularity present in this class of problems. The J-integral provides a robust mathematical formalism to analyze the mechanics of notches and cracks and is defined as follows:

$$J = \lim_{\Gamma \to 0} \oint_{\Gamma} \mathbf{n} \cdot (W\mathbf{I} - \boldsymbol{\sigma} \nabla \mathbf{u})\, \mathbf{q} ds = \lim_{\Gamma \to 0} \oint_{\Gamma} \mathbf{n} \cdot \mathbf{H}\mathbf{q} ds, \tag{2.2}$$

[1]It is worth mentioning that in addition to J, there are many other examples of so-called conservation integrals [97, 162–169]. Many of these quantities have interpretations in the context of Noether's theorem [170] on invariant variational principles [171].

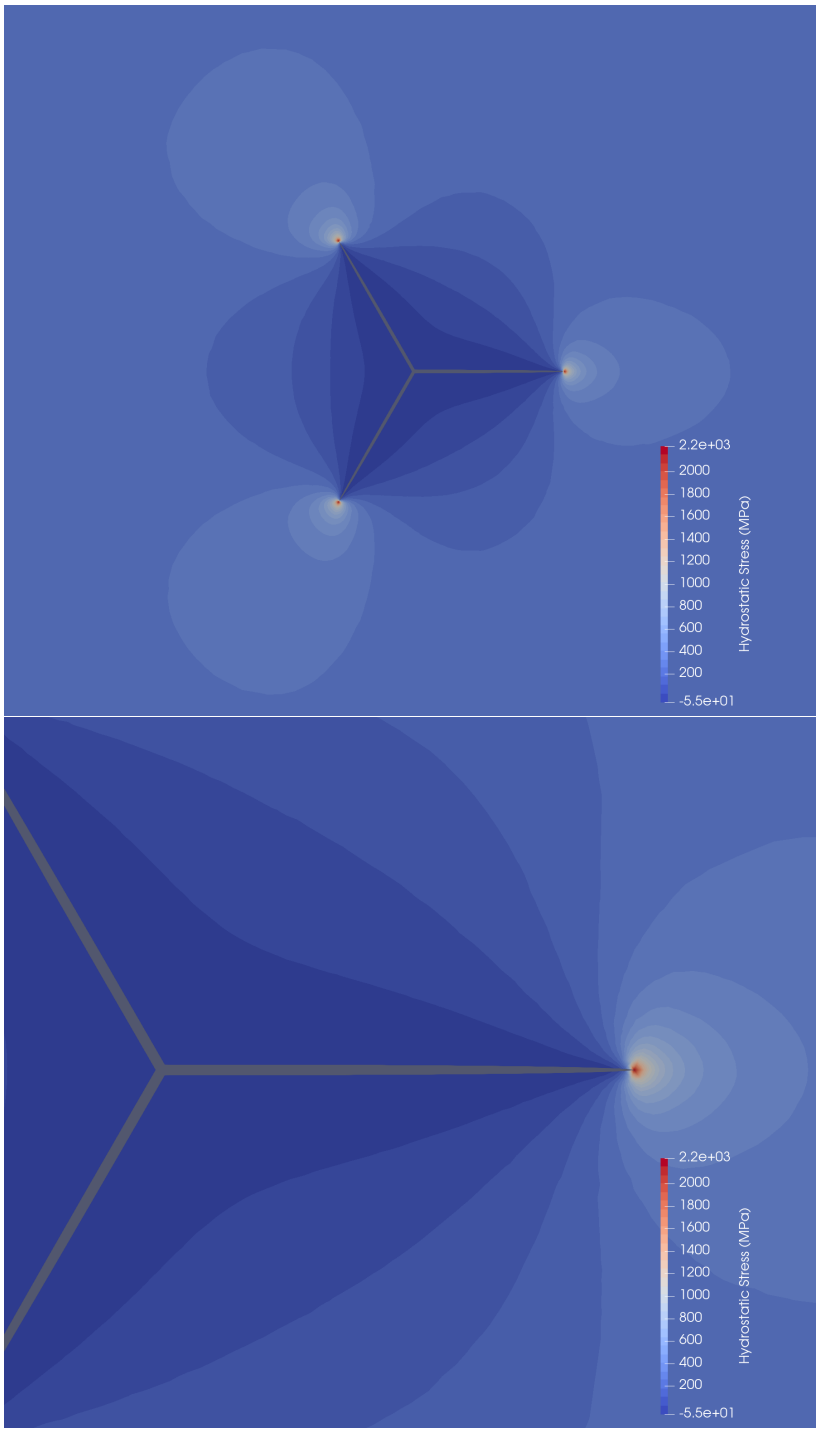

Figure 2.4: Contours of hydrostatic stress for the entire domain and local to the crack tip.

where $\mathbf{n}$ is the normal vector to the contour, W is the strain energy density, $\boldsymbol{\sigma}$ is the stress tensor, $\nabla\mathbf{u}$ is the displacement gradient, $\mathbf{q}$ is the so-called virtual crack extension direction, and Γ is a counter-clockwise contour that shrinks onto the crack tip.

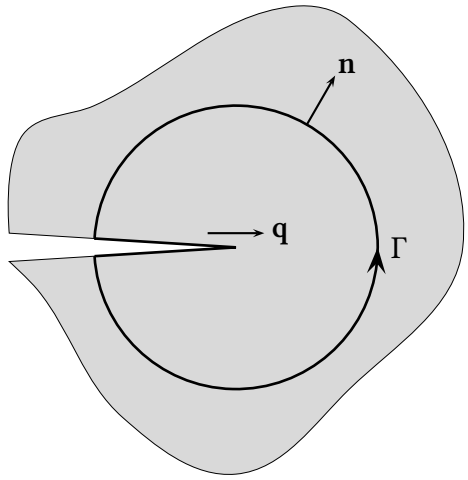

Figure 2.5: A sharp crack tip with the crack extension direction $\mathbf{q}$, contour Γ, and associated normal vector $\mathbf{n}$ shown.

It is worth mentioning that $\mathbf{H}$ is the transpose of Eshelby's elastic energy-momentum tensor [172]. Following the explanation by Eshelby, the integral (2.2) can be interpreted as the energy per unit area required to displace the crack-tip singularity by a unit distance in the direction $\mathbf{q}$. In linear elastic fracture mechanics, this interpretation is equivalent to the energy release rate [157, 173]. Importantly, J is also directly related to the stress intensification factor at the crack tip [157] and therefore serves as an excellent metric of elastic incompatibilities at triple junctions. Under certain conditions, the integral (2.2) is path-independent and the limit can be removed [174]. Furthermore, in the case that $\mathbf{q} = \mathbf{e}_1$ and the integral is path-independent, the expression given by (2.2) reduces to the original result provided by Rice [97, 157]:

$$J = \oint_{\Gamma} W dx_2 - t_i \frac{\partial u_i}{\partial x_1} ds. \tag{2.3}$$

In the case that Γ encircles a region containing no singularity, $J = 0$, which is the conservation law associated with the J-integral [171]. For linear, isotropic, homogeneous solids, the J-integral can be explicitly expressed in terms of the three *stress intensity factors* $K_{\mathrm{I}}, K_{\mathrm{II}}, K_{\mathrm{III}}$, corresponding to opening, in-plane sliding, and anti-plane sliding modes [97, 157]:

$$J = \frac{1-\nu^2}{E}\left(K_{\mathrm{I}}^2 + K_{\mathrm{II}}^2\right) + \frac{1+\nu}{E} K_{\mathrm{III}}^2, \tag{2.4}$$

where ν is Poisson's ratio and E is Young's modulus. The equation (2.4) will be used when verifying the computational framework in Section 2.3.3.

2.3.2 Calculation of J

In this thesis, we choose to compute J directly via the line integral form (2.2). First, Γ is restricted to be a circle with radius r centered at the crack tip. Substituting this form of Γ into (2.2) yields:

$$J = \int_0^{2\pi} [\mathbf{n} \cdot \mathbf{Hq}]_{\Gamma(t)} \|\Gamma'(t)\| dt = \int_0^{2\pi} [\mathbf{n} \cdot \mathbf{Hq}]_{\Gamma(t)} r dt, \tag{2.5}$$

where t is the parameterization, which in the case of circular contours coincides with the angle from the horizontal axis. Then, t is discretized at Q integration points. For the corresponding points along the contour, the finite element solution is evaluated to obtain W, σ, and $\nabla \mathbf{u}$. If the integration points do not coincide with the nodes of the finite element mesh, these values are obtained through interpolation of the nodal values using the finite element shape functions. These fields are then used as input into the Riemann sum:

$$J \approx \sum_{q=1}^{Q} [\mathbf{n} \cdot \mathbf{Hq}]_{\Gamma(t_q)} r \Delta t, \tag{2.6}$$

where $\Delta t = 2\pi/Q$. This integration procedure is illustrated in Figure 2.6. It is worth mentioning that domain integral forms of the J-integral are commonplace in finite element calculations [157, 174], although direct evaluation of the contour integral (2.2) has been done in experimental fracture mechanics [175–177]. In Section 2.3.3, it will be shown that the line integral calculation presented herein can be used to compute J accurately and reproduces the path-independence property in the case that the body is homogeneous.

2.3.3 Verification of J-Integral Calculation

In this section, we verify that the numerical calculation of the J-integral described in 2.3.2 reproduces both the analytical value of J for a standard problem in fracture mechanics, that of the finite crack in an infinite body and that for a homogeneous material, J is path-independent.

In the original paper of Rice [97], the J-integral for a narrow crack of length $2a$ in an infinite plate is given as:

$$J = \frac{\pi(1-\nu^2)}{E} \sigma_\infty^2 a. \tag{2.7}$$

where σ_∞ is the far-field loading. A diagram of this geometry is given in Figure 2.7. Near the crack tip, the displacement field can be found analytically via complex variable method [178, 179] and is given by:

$$u_1(r,\theta) = \frac{K_I}{2\mu}\sqrt{\frac{r}{2\pi}} \cos\left(\frac{\theta}{2}\right)\left(\kappa - 1 + 2\sin^2\left(\frac{\theta}{2}\right)\right), \tag{2.8}$$

$$u_2(r,\theta) = \frac{K_I}{2\mu}\sqrt{\frac{r}{2\pi}} \sin\left(\frac{\theta}{2}\right)\left(\kappa + 1 - \cos^2\left(\frac{\theta}{2}\right)\right), \tag{2.9}$$

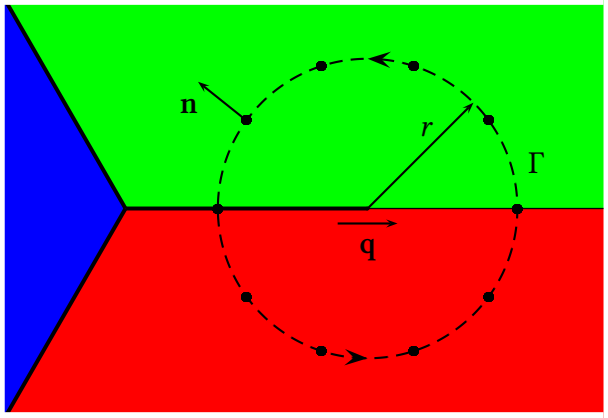

Figure 2.6: A schematic of the contour integration procedure shown around one of the microcracks (left).

where $K_I = \sigma_\infty \sqrt{\pi a}$ is the Mode I stress intensity factor, μ is the shear modulus, $\kappa = 3 - 4\nu$, and r, θ are the polar coordinates of a point in a system centered at the crack tip. It bears emphasis that according to the displacement field (2.8), the crack is only subject to Mode I loading, thus the J integral can be found by substituting $K_I = \sigma_\infty \sqrt{\pi a}, K_{II} = K_{III} = 0$ into (2.4) to obtain (2.7).

This example furnishes a verification case for the numerical implementation described in Section 2.3.2 and is done via the method of manufactured solutions. We restrict the analysis to the rectangular domain $\Omega = [0, 1] \times [-\frac{1}{2}, \frac{1}{2}]$ with a line crack along $0 \leq x_1 \leq \frac{1}{2}, x_2 = 0$. On the boundary, the displacement is constrained to be equal to the analytical displacement field (2.8). For demonstration purposes, the parameters used in this simulation were $E = 50$ Pa, $\nu = 0.45$, and $\sigma_\infty = 0.1$ Pa. Then, J was computed for varying mesh sizes and polynomial orders to investigate the convergence properties of the numerical integration method (2.6). Furthermore, the convergence of the finite element solution $\mathbf{u}_h$ was monitored using the $\mathscr{H}_1$ norm on the error $\mathbf{e} = \mathbf{u} - \mathbf{u}_h$. The finite element solution and contours of the stress are given in Figure 2.8.

The errors in both J and $\mathbf{u}_h$ as a function of mesh size and polynomial order are given in Figure 2.9. Clearly, the convergence of these two quantities are closely related, as the rates of convergence are very similar. It is worth mentioning that the reduced rate of convergence is consistent with finite element error estimates in the presence of cracks, where the stress singularity and resulting large strain-gradient can slow convergence [180]. It can be seen that the errors in the finite element solution are smaller than those in the value of J. However, for most mesh sizes examined and all polynomial orders, the relative error in J is of order 10^{-2} or smaller, indicating that the numerical implementation can compute J sufficiently accurately for the purposes of the current work.

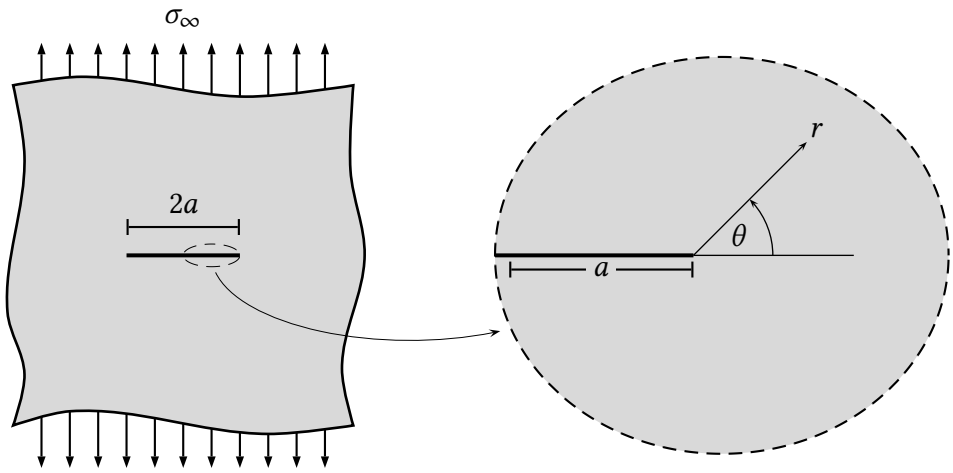

Figure 2.7: Crack in an infinite body subjected to uniaxial remote stress σ_∞.

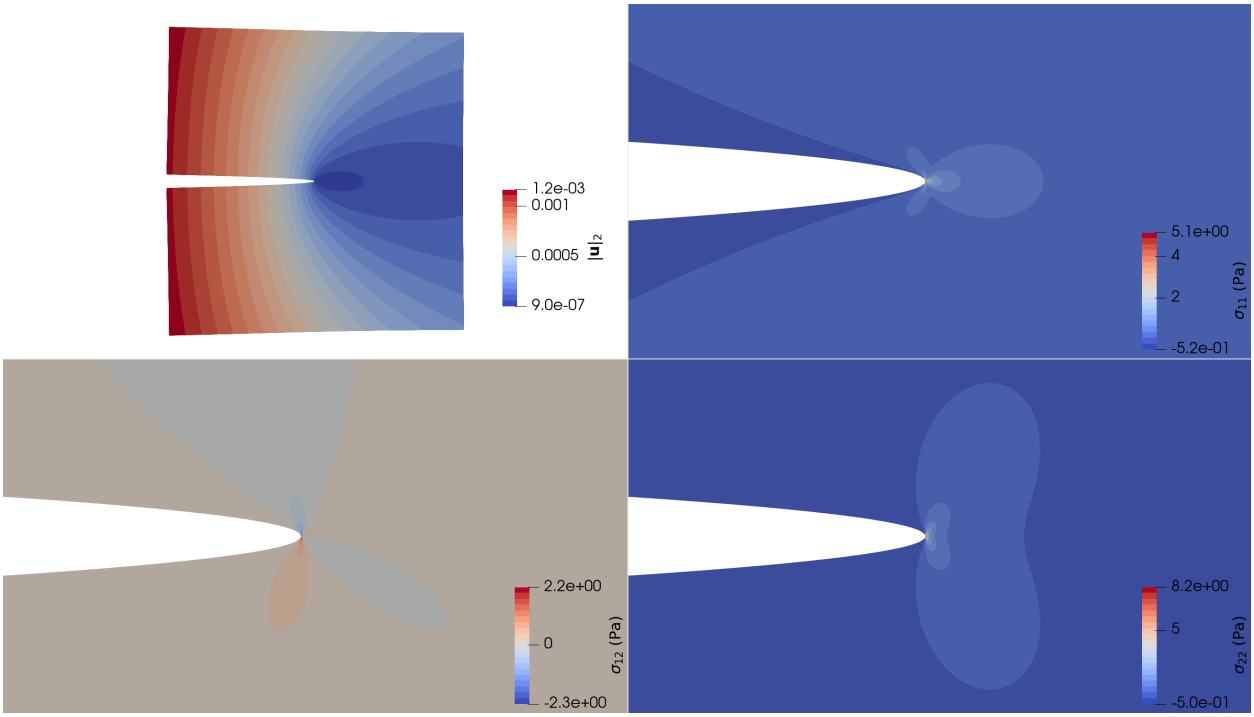

Figure 2.8: Displacement and stress contours associated with the semi-infinite crack problem shown on the warped geometry with a warping factor of 10 for clarity.

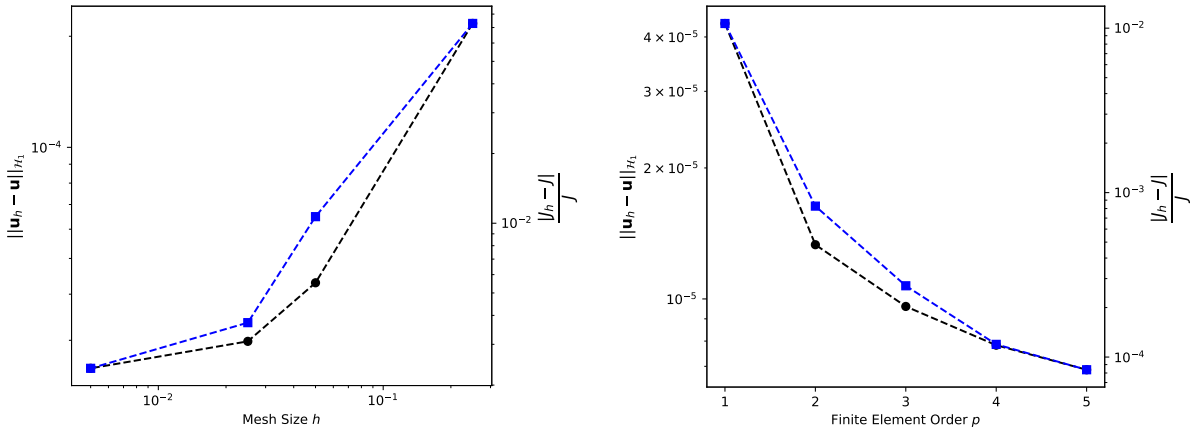

Figure 2.9: Errors in the finite element solution and the J integral as a function of mesh size (left) and polynomial order (right). The black dots correspond to errors in $\mathbf{u}_h$ and the blue dots correspond to errors in J_h.

It has been established the J-integral is path-independent for homogeneous bodies under static loading. We choose to demonstrate this property using the triple junction geometry, where the configuration $(\theta_1, \theta_2, \theta_3) = (0, 0, 0)$ corresponds to a homogeneous body. In particular, we verified the path-independence property by computing J_i for varying contour radii r. We also verified the convergence of the computed numerical values of J_i by varying the number of integration points Q. Figure 2.10 shows the impact of these parameters on the calculated J integrals. For convenience, these results are expressed using the following non-dimensional parameters:

$$r^* = \frac{r}{a}, \ \Delta s^* = \frac{2\pi r}{Q}\frac{1}{h}, \tag{2.10}$$

where h is the characteristic length scale of the mesh, r^* is the ratio of the contour size to the crack length, and Δs^* is the ratio between the arclength step in the discretization of Γ and the mesh length scale h. The three J integrals for different $r^*, \Delta s^*$ are given in Figure 2.10. For each microcrack, the values for J approach a comparable value for all values of r^* as Δs^* decreases. Furthermore, J_2 and J_3 approach similar values, which can be expected due to the $x_1 - x_3$ symmetry of the boundary value problem. The error between the largest and smallest values of J for each microcrack is of order 0.5%. While further mesh refinement could decrease this error, this level of error over the examined range of $r^*, \Delta s^*$ suggests that J is reasonably path-independent and that the current mesh is suitable for the present work. Furthermore, this demonstrates that the numerical implementation described in Section 2.3.2 preserves this property of J to a reasonable level of precision.

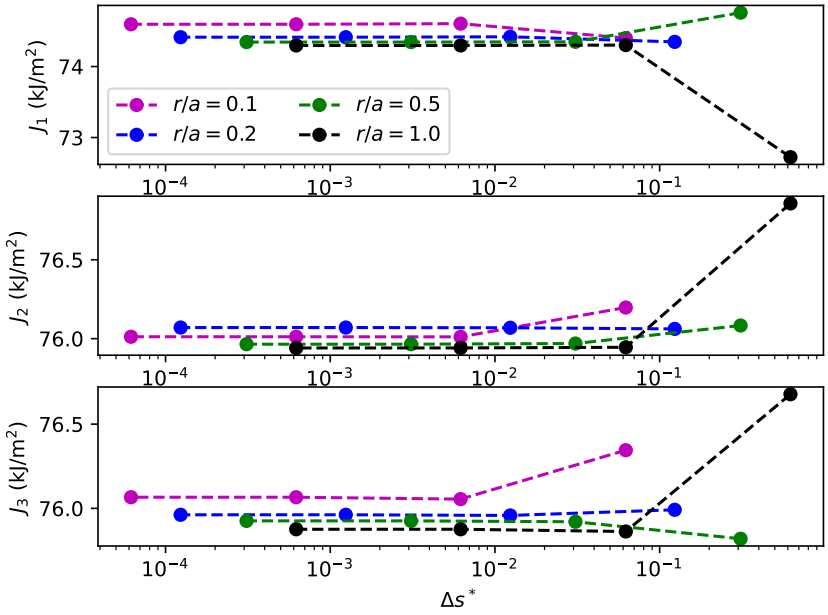

Figure 2.10: The three J-integrals at the triple junction expressed as a function of the nondimensional contour size r^* and discretization size Δs^*.

Chapter 3

Surrogate Models Using Neural Networks

In this chapter, we summarize important concepts related to multi-layer perceptron (MLP) networks. We first review the general structure of an MLP, then describe the training procedure that computes the weights of the MLP. We show that the adopted network architecture is able to produce accurate predictions on the validation set. Finally, we explain how this architecture can be utilized for inverse problems related to optimization of the surrogate model outputs.

3.1 Multi-Layer Perceptron Networks

In the current problem, we seek to predict the J_i given a triple junction configuration parameterized by θ_i. In general, this relationship is expected to be nonlinear. Furthermore, we allow each J integral to be potentially affected by all three of the grain rotation angles. As will be shown in Chapter 4, this dependence occurs in the boundary value problem considered in Chapter 2. The multi-layer perceptron (MLP) network, a type of supervised learning model, affords a powerful regression method that can naturally account for this nonlinear coupling of θ_i and J_i.

An MLP is a nonlinear function that maps an input vector $X \in \mathbb{R}^n$ to an output vector $Y \in \mathbb{R}^m$. At a high level, MLPs compose nonlinear functions and affine transformations of the features to obtain the outputs. More concretely, this mapping is achieved through a successive composition of "layers," which map an intermediate input $U \in \mathbb{R}^k$ to an intermediate output $V \in \mathbb{R}^q$ through the following set of operations; U is affinely transformed using a matrix of weights $W \in \mathbb{R}^{k \times q}$ and a vector of offsets, also known as biases, $W_0 \in \mathbb{R}^q$. The resulting transformed quantity is then acted upon by a pre-defined "activation function" f, which is applied component-wise. Typical choices of f include the sigmoid and rectified linear unit (ReLU) functions:

$$f_{\text{sigmoid}}(x) = \frac{1}{1 + e^{-x}},$$
$$f_{\text{ReLU}}(x) = \max(0, x).$$

This set of operations can be written succinctly as:

$$V = f\left(W^\top U + W_0\right). \tag{3.1}$$

By setting the outputs of one layer equal to the inputs of the next, the network can approximate nonlinear functions. It has been shown that neural networks with sufficiently smooth activation functions can approximate certain functions to arbitrary accuracy provided sufficiently many intermediate, also known as "hidden," layers are used [181]. Due to this property, MLPs are also known as "universal approximators".

In the current problem, the first layer accepts $U = (\theta_1, \theta_2, \theta_3)$ as an input, while the last layer produces the approximation $V = (\hat{J}_1, \hat{J}_2, \hat{J}_3)$ as an output. We denote the weights and biases for layer ℓ as W^ℓ and W_0^ℓ respectively, for $\ell = 1, \dots, L.$, i.e. for a network with L layers. Additionally, as the activation function can be different for each layer, the activation functions for the L layers are denoted as f^ℓ, $\ell = 1, \dots, L$. For ease of notation, the range of ℓ will not be stated explicitly. Thus, we write the functional form of the neural network as:

$$\hat{J}_i = \mathrm{NN}(\theta_i; W^\ell, W_0^\ell), \tag{3.2}$$

to indicate that the neural network is a function of the features, parameterized by W^ℓ and W_0^ℓ. A schematic of the MLP architecture used in this thesis is presented in Figure 3.1.

$$\{\theta_1, \theta_2, \theta_3\} \longrightarrow f^1\left(W^1 V_0 + W_0^1\right) \longrightarrow f^2\left(W^2 V_1 + W_0^2\right) \longrightarrow \cdots \longrightarrow \{\hat{J}_1, \hat{J}_2, \hat{J}_3\}$$

Figure 3.1: Diagram of the MLP network used in this thesis.

3.2 Network Training Procedure

As stated in the previous section, the weights W^ℓ and biases W_0^ℓ parameterize the network. These serve as the fitting parameters to be tuned in the network training process. Specifically, the weights and biases are chosen such that $\mathrm{NN}(\theta_i; W^\ell, W_0^\ell)$ best approximates the known label data, i.e. as the solution to an optimization problem that minimizes the error between the predictions of the model and the feature data. This requires a scalar measure of network prediction error, which is provided by the so-called loss function, denoted herein as $\mathscr{L}$. A common loss function, and the one used in this thesis, is the mean-squared-error (MSE) loss, which has the following form for the current MLP:

$$\mathscr{L}(\theta_i; W^\ell, W_0^\ell) = \frac{1}{3}\sum_{i=1}^{3} \frac{1}{N_{\text{train}}} \sum_{k=1}^{N_{\text{train}}} \left(J_i^k - \mathrm{NN}\left(\theta_i; W^\ell, W_0^\ell\right)\right)^2. \tag{3.3}$$

It is worth mentioning that (3.3) can be appended by an L_2 regularization of the form $\lambda\left(\|W^\ell\|_2^2 + \|W_0^\ell\|_2^2\right)$, where $\|\bullet\|_2$ is the L_2 norm. This regularization term is added to improve the predictive accuracy of the surrogate model on out-of-data samples [182], however as will be demonstrated in Section 3.3, this regularization was not necessary in the current work. It bears emphasis that the loss function, like the network itself, is a function of the angles θ_i, but can be manipulated through changes

to the parameters W^ℓ, W_0^ℓ. Thus, the network training process can be conceived of as fixing the N_{train} sets of θ_i to be those in the training feature data, and performing the optimization with respect to W^ℓ, W_0^ℓ to obtain the optimal $W^{\ell*}, W_0^{\ell*}$, i.e.:

$$W^{\ell*}, W_0^{\ell*} = \arg\min_{W^\ell, W_0^\ell} \mathscr{L}(\theta_i; W^\ell, W_0^\ell). \tag{3.4}$$

Given the composition of layers needed to obtain an accurate interpolant, the optimization problem (3.4) is generally non-convex and nonlinear. An immediate consequence of these characteristics is that $W^{\ell*}, W_0^{\ell*}$ depend on the starting point of the optimization process. This can be easily addressed by sampling multiple initial guesses and simply choosing the solution with the best performance. More fundamentally, the size of the optimization problem associated with (3.4) poses a significant computational challenge. In practice, modern MLPs can contain hundreds of thousands to millions of trainable parameters. This renders second-order (or higher-order) methods such as Newton-Raphson impractical due to the requisite matrix inversion [183]. While quasi-Newton schemes have occasionally been used [184–186], the standard optimization algorithm for training is based on the first-order gradient descent update:

$$W^{\ell,k+1} = W^{\ell,k} - \alpha \frac{\partial \mathscr{L}\left(\theta_i; W^{\ell,k}, W_0^{\ell,k}\right)}{\partial W^\ell}, \tag{3.5}$$

where k is the current iterate of gradient descent, and α is the step size. It is worth mentioning that α is often called the "learning rate" in the machine learning community. While α can be computed through a line search at each iterate, in this thesis α was taken as constant. The update for W_0^ℓ is done analogously to (3.5). First-order methods such as (3.5) require fewer derivative calculations than higher-order methods, at the cost of a reduced convergence rate.

However, the gradient descent update (3.5) still poses multiple computational challenges. Firstly, naive computation of $\frac{\partial \mathscr{L}}{\partial W^\ell}$ is generally intractable due to layer composition, which causes $\mathscr{L}$ to depend on each of the components of the weights in a highly nonlinear and complicated manner. However, these gradients can be efficiently computed by recourse to the so-called back-propagation procedure, which is the successive application of the chain rule through the layers of the network. The gradient of the loss function with respect to the weights of an arbitrary layer ℓ can be expressed as follows by an application of the chain rule:

$$\begin{aligned} \frac{\partial \mathscr{L}}{\partial W^\ell} &= \frac{\partial Z^\ell}{\partial W^\ell} \left(\frac{\partial \mathscr{L}}{\partial Z^\ell}\right)^\top \\ &= V^{\ell-1} \left(\frac{\partial \mathscr{L}}{\partial Z^\ell}\right)^\top, \end{aligned}$$

where $Z^\ell = W^\ell V^{\ell-1} + W_0^\ell$ is the intermediate quantity computed by the network before the activation function f^ℓ is applied. Through once again applying the chain rule consecutively, the

loss function gradient can be expanded further:

$$\begin{aligned}\frac{\partial \mathscr{L}}{\partial Z^\ell} &= \frac{\partial V^\ell}{\partial Z^\ell}\frac{\partial Z^{\ell+1}}{\partial V^\ell}\frac{\partial V^{l+1}}{\partial Z^{\ell+1}}\cdots\frac{\partial Z^L}{\partial V^{L-1}}\frac{\partial V^L}{\partial Z^L}\frac{\partial \mathscr{L}}{\partial V^L} \\ &= f^{\ell\prime}(Z^\ell)W^{\ell+1}f^{\ell+1\prime}(Z^{\ell+1})\ldots W^L f^{L\prime}(Z^L)\frac{\partial \mathscr{L}}{\partial V^L}.\end{aligned} \tag{3.6}$$

This expansion allows for the expression of the derivative of the loss function with respect to any of the training parameters to be expressed in terms of quantities from previous layers. Thus, the training process involves "forward" passes that simply evaluate $\mathscr{L}$ with fixed weights and biases on different portions of the training set, and "error backpropagation" that evaluates the gradient (3.6) and updates the weights according to (3.5).

Despite the approach to tractably compute gradients of the loss function described above, the size of the optimization problem still poses challenges — the matrix multiplications present in (3.6) render the calculation of the gradient impractical for all trainable parameters in the case of large networks. Furthermore, the gradient descent update (3.5) with constant step size is not guaranteed to avoid local minima in the loss function, which can hinder the accuracy of the interpolant. The stochastic gradient descent (SGD) algorithm is typically used to circumvent both of these challenges [187]. SGD can be applied to so-called "separable" objective functions f:

$$f(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} f_i(\mathbf{x}).$$

It is obvious that the MSE loss function (3.3) is of this form. In the SGD algorithm, the gradient $\nabla f(\mathbf{x})$ is approximated by an unbiased estimator $\tilde{\nabla} f(x)$. For example, this can be taken as the gradient of one of the terms in the sum, i.e. by $\nabla f_k(\mathbf{x})$ for some randomly chosen index k. More typically, so-called batches are used, which approximate the true gradient by the average of some $\nabla f_i(\mathbf{x})$. This leverages additional information about $f(\mathbf{x})$ to improve convergence and regularize stochasticity associated with the random update. The computational cost of SGD is significantly lower than standard gradient descent as it requires fewer derivative calculations. Furthermore, the randomized nature of SGD allows the algorithm to avoid local minima, as the terms $f_i(\mathbf{x})$ do not necessarily share optimizers with $f(\mathbf{x})$. Interestingly, the noise provided by the SGD update also improves the predictive accuracy of the surrogate model by avoiding the minimizers of $\mathscr{L}$ arising from the specific selection of training data.

As the SGD algorithm uses less information than standard gradient descent, it has a slower convergence rate. However, it is worth mentioning that SGD typically converges in two "stages," the first of which is rapid convergence to a neighborhood of a stationary point, followed by a lack of convergence within a so-called noise ball. The size of the noise ball is controlled, in part, by the step size α. Thus, once convergence stalls, the optimal solution is typically chosen as the best solution visited within the noise ball. A schematic demonstrating the convergence behavior of SGD is given in Figure 3.2.

Modern network training approaches further improve upon the SGD algorithm outlined above through the use of the "AdaGrad" [188] algorithm, which uses an adaptive learning rate for each

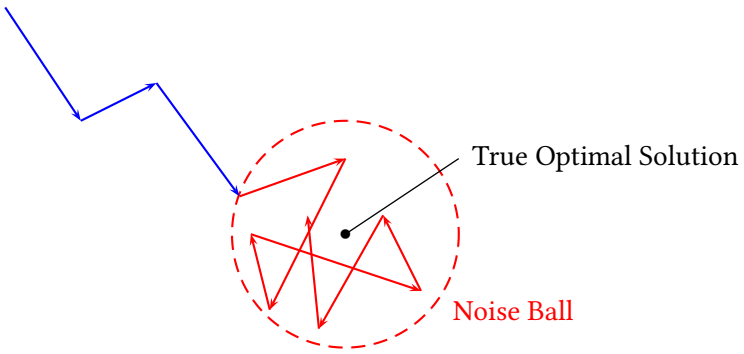

Figure 3.2: Schematic demonstrating convergence behavior of SGD algorithm. Initial iterates quickly reduce the objective but subsequent iterates remain trapped in a "noise ball" where convergence stalls.

of the optimization variables. This, in effect, acts as a preconditioner for the gradient descent update, which accounts for variance in scale across the different optimization variables. This is similar to the well-known invariance of Newton-Raphson under affine transformations of the optimization variable space. The performance of AdaGrad is further improved through the inclusion of "momentum," where the descent direction in the update includes information from prior gradients. Momentum can be useful for escaping local minima and accelerating convergence in favorable descent directions. This augmentation of AdaGrad is called Adam [189] and has become standard for modern machine-learning problems based on its empirically good performance. However, it is worth mentioning that Adam does not share the theoretical guarantees of GD, SGD, or AdaGrad — it has been shown to diverge even for convex problems [190].

Standard practice in cases where the total dataset is sufficiently large is to randomly partition the dataset into training, testing, and validation datasets [191]. The training dataset is usually taken to be the largest portion and is used to fit the training parameters. The validation dataset is used to estimate the error of the model during training, which will be explained in more detail in the following paragraph. Finally, the testing dataset is used post-training to evaluate the performance of the model, i.e. not "seen" during the training process at all. An additional measure to evaluate the accuracy of the trained model is cross-validation. A standard example of cross-validation is K-fold cross-validation, where the training data set is split into K parts. During training, $K-1$ parts are used as the training data set and compared to the remaining part. This is done multiple times with different parts removed from the training set, and the average loss is reported in order to monitor the model performance more accurately. $K = 5, 10$ are typical values that balance computational cost, the error in predictions, and the variability of predictions [191].

In the current work, the TensorFlow [132] and Keras [133] packages were used to instantiate the neural network. The Adam optimizer [189] was used during training to learn the network parameters. Additional information regarding the MLP architecture is given in Table 3.1. Here, "epochs" is the maximum number of iterations of the optimization algorithm performed before stopping. However, following the discussion of SGD above, the "patience" parameter refers to the number of subsequent iterations after which the training is stopped if no improvement is observed. The learning rate and batch size have the same meaning as that described above. These

"hyperparameters" were found to lead to reasonable accuracy of the trained model. However, for general problems, hyperparameter tuning is often employed to improve the training and performance of the surrogate model.

Table 3.1: Hyperparameters used during the model training procedure.

Parameter	Value
Layers	[256,256,256,3]
Activation Functions	[ReLU, ReLU, ReLU, Linear]
Learning Rate	0.001
Batch Size	256
Max Epochs	2000
Patience	10

The validity and accuracy of the surrogate model are also dependent on the quality of the feature set and the chosen network architecture. Error analyses of neural networks [192] identify these two components of the network training problem as central to the performance of the trained model. A method to evaluate whether the feature set and network architecture accurately model the label data is to compare the training and validation loss as a function of dataset size [121]. One possibility is that the training loss is significantly lower than the validation loss. This suggests that either the feature set does not contain sufficient information to describe the statistical variations in the label set, or that the network has too many trainable parameters and is suffering from overfitting. Another possibility is that the training and validation losses are comparable to each other, but remain large. In this case, the model did not learn meaningful trends from the training data and is said to be underfit. These cases are schematically demonstrated in Figure 3.3. This analysis will be used with the synthetic data produced by the physics-based model to validate our choice of network architecture in Section 3.3.

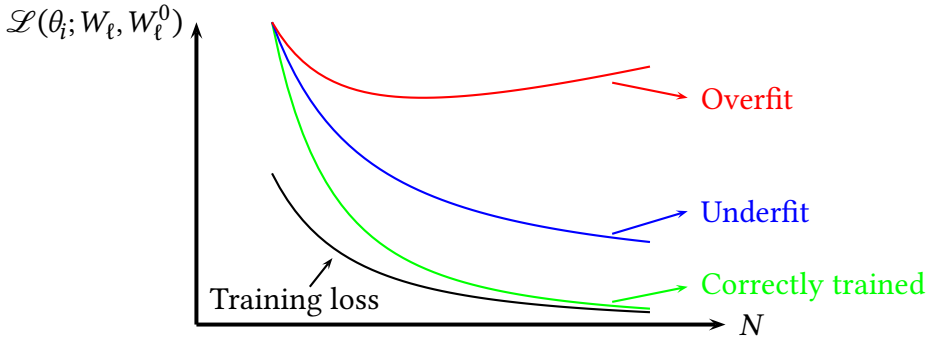

Figure 3.3: Schematic of the possible outcomes when training the MLP as the dataset size is increased, classified based on the behavior of the validation loss (green, blue, red) compared to the training loss (black).

3.3 Network Training

In order to form the synthetic dataset for training the MLP, a set of feature data is needed. In this case, the three angles θ_i were all sampled independently from a uniform distribution on $[0, 2\pi]$ 100,000 times each. Then, the finite element simulation described in Section 2.1 was conducted for each set of θ_i. Using the integration method described in Section 2.3, the three J-integrals were post-processed from the finite element solution and collected to form the label data portion of the synthetic dataset. In machine learning, the distribution of the label data is important in statistical analyses of neural networks [191]. Thus, histograms of the J_i produced by the physics-based model are provided in Figure 3.4. It is interesting to note that the three histograms are approximately identical, due to both the rotational symmetry of the problem and the uniform sampling of the angles. However, it is worth mentioning that the histograms are not identical due to finite sample size; if more angles were to be sampled, these histograms would approach each other to a greater degree.

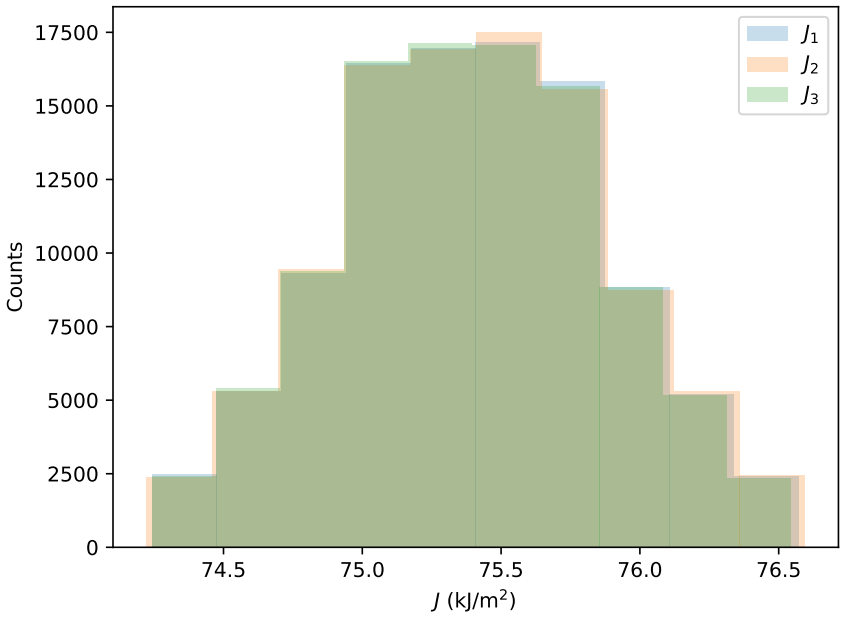

Figure 3.4: Histograms of the values of J for each crack.

Following the discussion on error analysis of neural networks from Section 3.1, the training and validation losses were computed at the end of the training of five different networks for a range of training dataset sizes; multiple trials are used to account for the random initialization of weights and biases at the start of training. The mean and standard deviation of the training and validation loss over the five trials are plotted in Figure 3.5. As the dataset size increases, the training and validation losses can be seen to decrease in tandem, which indicates that the

chosen feature set is representative of the variations in the label set. This also indicates that the model has not been overfitted. Finally, the low values of the loss function ($\mathcal{O}(10^{-4})$) relative to the size of the labels at the largest dataset size indicate that underfitting has not occurred. Figure 3.5 demonstrates that the surrogate model is sufficiently accurate for the purposes of this work, obviating the need for further sweeps over the network parameter space to improve the performance of the model.

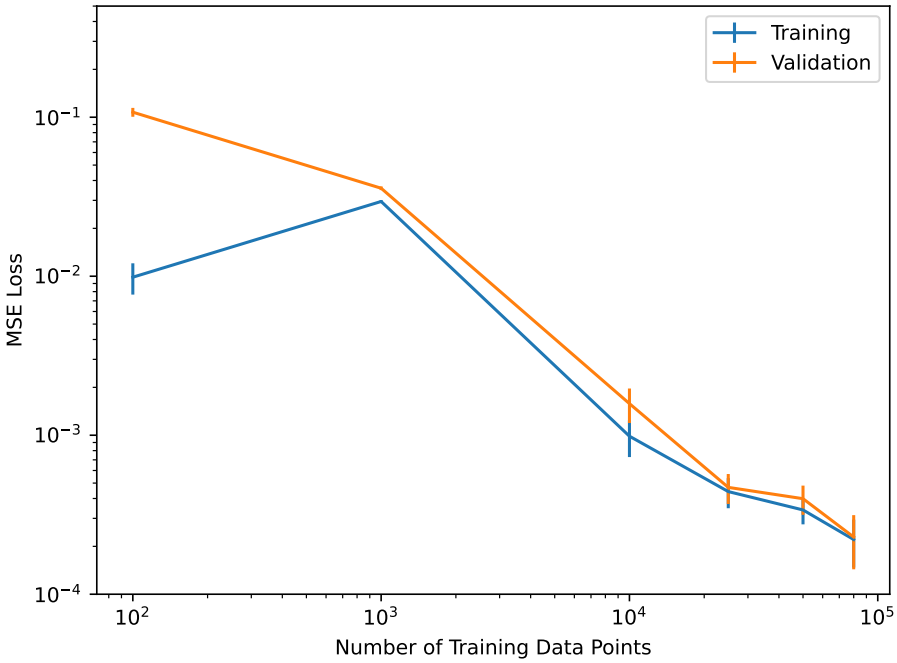

Figure 3.5: Training and validation loss for increasing amounts of training data.

3.4 Inverse Problem Solution Procedure

The trained surrogate model provides *forward* prediction capabilities, i.e. the prediction of the three J-integrals given the grain boundary configuration. However, as stated in Section 1, one of the core aims of grain boundary engineering is to determine which grain configurations lead to either favorable or deleterious properties. While inverting the forward relationship is generically challenging in the most general setting, the explicit functional form furnished by the surrogate model enables the inversion of this relationship for the purposes of optimizing the stress intensification at the triple junction. This can be done following the procedure described in the work of Peurifoy et al. [122], where the input parameters of the MLP are adjusted to either optimize or obtain a desired output. In the current work, the inverse problem is finding the angles θ_i that optimize a given scalar measure of the J_i. For demonstration purposes, we take this function to

be the ∞-norm $\|J\|_\infty$:

$$\|J\|_\infty = \max\{|J_1|, |J_2|, |J_3|\}. \tag{3.7}$$

$\|J\|_\infty$ can be interpreted as related to the largest stress intensification among all three crack tips. Note that given the trained network, $\|J\|_\infty$ can be evaluated for any feasible set of input angles. Then, using the trained surrogate model, the angles that minimize $\|J\|_\infty$ can also be found through gradient descent:

$$\theta_i^{k+1} = \theta_i^k - \alpha \frac{\partial \|J\|_\infty}{\partial \theta_i}. \tag{3.8}$$

This is done by instantiating a new model with the same weights as the original forward model, but with the loss function set to the objective (3.7). It bears emphasis that this optimization problem is in contrast to that associated with training the network; during training, the θ_i are the fixed data and W^ℓ, W_0^ℓ are iteratively tuned to minimize the loss function, and the functional relationship is written as (3.2). In this case, W^ℓ, W_0^ℓ are held fixed and the θ_i are varied to minimize (3.7), thus the functional relationship in the inverse problem can be stated as $\hat{J}_i = \mathrm{NN}(W^\ell, W_0^\ell; \theta_i)$. The sign of the update in (3.8) simply needs to be changed to maximize $\|J\|_\infty$.

The gradients $\frac{\partial \|J\|_\infty}{\partial \theta_i}$ present in (3.8) can also be efficiently computed by recourse to backpropagation. This can be done as follows

$$\frac{\partial \|J\|_\infty}{\partial \theta_i} = \sum_{\alpha=1}^{3} \frac{\partial \|J\|_\infty}{\partial J_\alpha} \frac{\partial J_\alpha}{\partial \theta_i}, \tag{3.9}$$

where each derivative $\frac{\partial J_\alpha}{\partial \theta_i}$ can be computed by a chain rule expansion similar to (3.6), as the functional form of J_i is obtained from the training process.

Chapter 4

Quantification of Elastic Incompatibilities at Triple Junctions

In this section, we combine the components of the computational framework presented in Chapter 2 and Chapter 3 both efficiently sweep configurational space and to identify optimal grain boundary configurations. First, we perform a set of forward calculations that demonstrate both the accuracy of the model as well as its capability to furnish fast and direct evaluations of the stress intensification resulting from elastic incompatibilities. Then, using the method of Peurifoy et al [122], the surrogate model is then used to solve the inverse problem of finding the triple junction configurations that minimize and maximize the incompatibility-induced stress intensification. Then, the theory of coincident site lattices (CSLs), an important concept in GBE [4, 10, 20, 193], is summarized and used to compare the angles identified via optimization with those configurations considered favorable due to good atomic fit.

4.1 Forward Predictions Using Trained Surrogate Model

For a further illustration of the network accuracy, we compare the physics-based and surrogate models for a subset of the feature space. For simplicity, we consider the one-dimensional subsets of the feature space $(\theta_1, 0, 0)$, $(0, \theta_2, 0)$, and $(0, 0, \theta_3)$ where $\theta_1, \theta_2, \theta_3 \in [0, \pi/2]$ respectively (the restriction to $[0, \pi/2]$ is due to the cubic symmetry of the grains). Figure 4.1 shows the comparison between the surrogate model and finite element calculation. In these three cases, the surrogate model accurately reproduces the values produced by the finite element simulation, with a low relative error. This is consistent with the small loss function value obtained during training. While the relative errors are small, it is worth mentioning that the model appears to slightly underestimate J_1 and overestimate J_3, at least over this subset of the feature space. These errors could be reduced by increasing the complexity of the network, i.e. increasing the number of nodes or layers. However, due to standard results in machine learning analysis regarding the bias-variance tradeoff [192, 194, 195], this increase in model complexity may introduce noise in the predictions that do not accurately represent the physical model.

The surrogate model can now be used to provide a complete understanding of the influence of

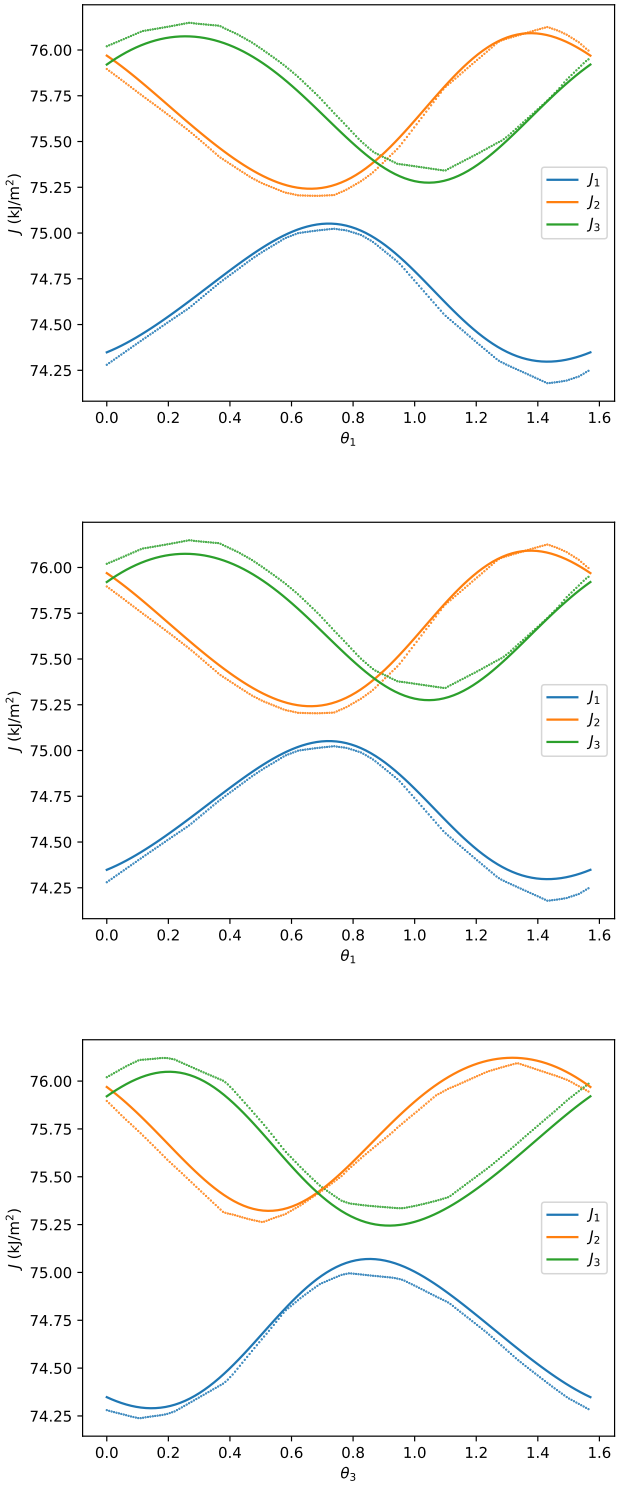

Figure 4.1: The predictions of the surrogate model (dots) compared to the values from the finite element simulation (lines).

the triple junction configuration on elastic incompatibilities. To this end, Figure 4.2 shows contour maps of the largest stress intensification of the three cracks as a function of the triple junction, as measured by (3.7). Concretely, we consider the 2D parameter spaces $(\theta_1, \theta_2, 0)$, $(\theta_1, 0, \theta_3)$, and $(0, \theta_2, \theta_3)$, where here $\theta_1, \theta_2, \theta_3 \in [0, 2\pi]$ as appropriate. In all three contour plots, $\|J\|_\infty$ is approximately $\pi/2$-periodic in the features, as expected from the material symmetry of cubic crystals. Furthermore, it is interesting to note that the contour plots 4.2c and 4.2a show qualitatively similar contour patterns. As with the histograms in Figure 3.4, this is once again due to the symmetry present in the triple junction geometry. It bears emphasis that repeated evaluation of the surrogate model to produce these visualizations took a practically insignificant amount of compute time.

4.2 The Inverse Problem: Identification of Optimal Configurations

4.2.1 Identification of Extremal Configurations

To illustrate the convergence process, Figure 4.3 shows the results of the optimization procedure for a minimization and maximization problem with a shared initial guess. Figures 4.3a and 4.3b show that when minimizing $\|J\|_\infty$, the gradient descent algorithm manipulates the angles to set all of the J integrals equal to each other. Once they are equal, the largest J integral alternates between the three cracks due to finite step size, causing the J_i to oscillate. Interestingly, the angles can be seen to evolve near the end of the iteration, even as the resulting change in $\|J\|_\infty$ is small. This indicates that in certain configurations, the minimum value of $\|J\|_\infty$ is somewhat insensitive to perturbations in θ_i. For the maximization problem, J_3 simply increases until it plateaus; the tradeoff between the cracks observed in the minimization problem does not occur. Accordingly, the angles also cease to change noticeably within a few hundred iterations.

We now seek to identify all favorable or unfavorable configurations, which correspond to the minimizers and maximizers of $\|J\|_\infty$ respectively. This was done by conducting the optimization procedure described above for 100 random initial guesses each for the minimization and maximization problems separately. After the iteration concluded, the optimal set of θ_i in either case was collected. These sets are shown in the full configurational space in Figure 4.5. Projections of these angles into $\theta_1 - \theta_3$ and $\theta_2 - \theta_3$ space are given in Figure 4.6. In these plots, the angles were further restricted to be contained in $[0, \pi/2]$ using the modulo function, as this was found to make the structure of the extremal configurations more apparent. For the minimization problem, the configurations appear to approximately form lines in both $\theta_1 - \theta_2$ and $\theta_2 - \theta_3$ spaces. This can be related to the aforementioned insensitivity of $\|J\|_\infty$ to changes in θ_i. Along these lines, the configurations share a similar value of $\|J\|_\infty$. The maximization problem, however, shows three distinct clusters, each of which likely corresponds to the configuration that sets one of the three J integrals as the largest. Based on the projections in Figure 4.6, it can be seen that some of the configurations lie on the lines $\theta_1 = \pi/2 - \theta_3$ and $\theta_2 = \pi/2 - \theta_3$, which represent the symmetric boundaries (the $\pi/2$ arises due to crystal symmetry considerations), are some of the bicrystallo-

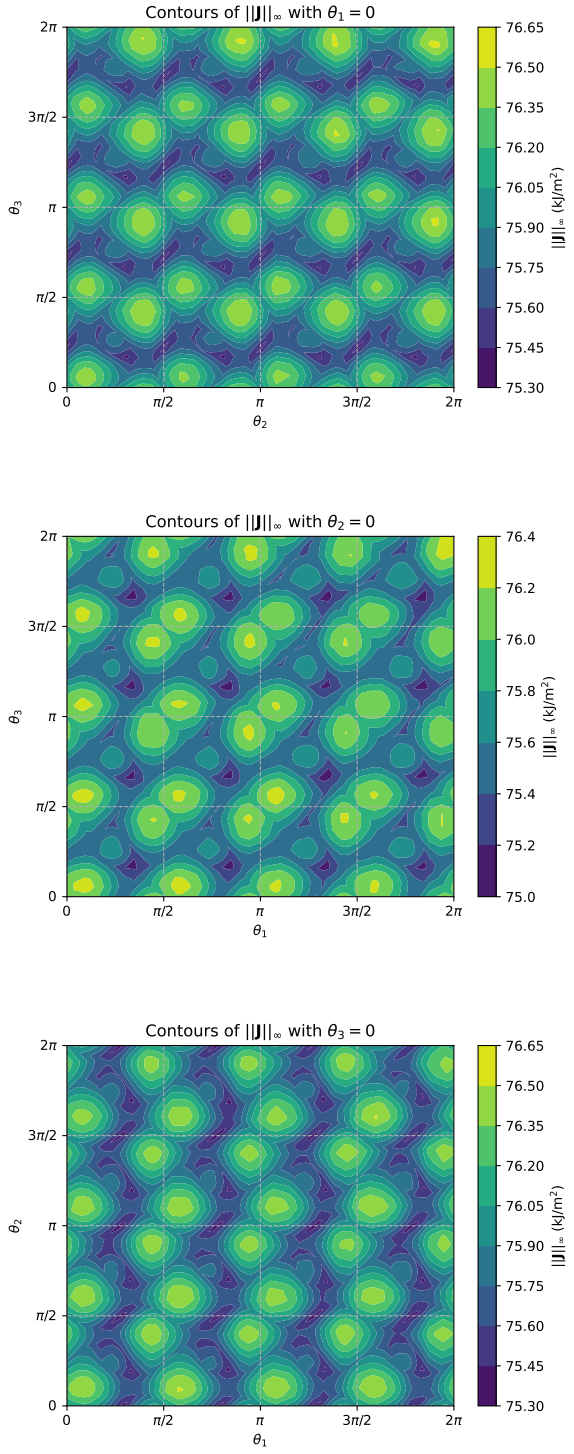

Figure 4.2: Contour maps of $\|J\|_\infty$ for various two-dimensional subsets of the feature space.

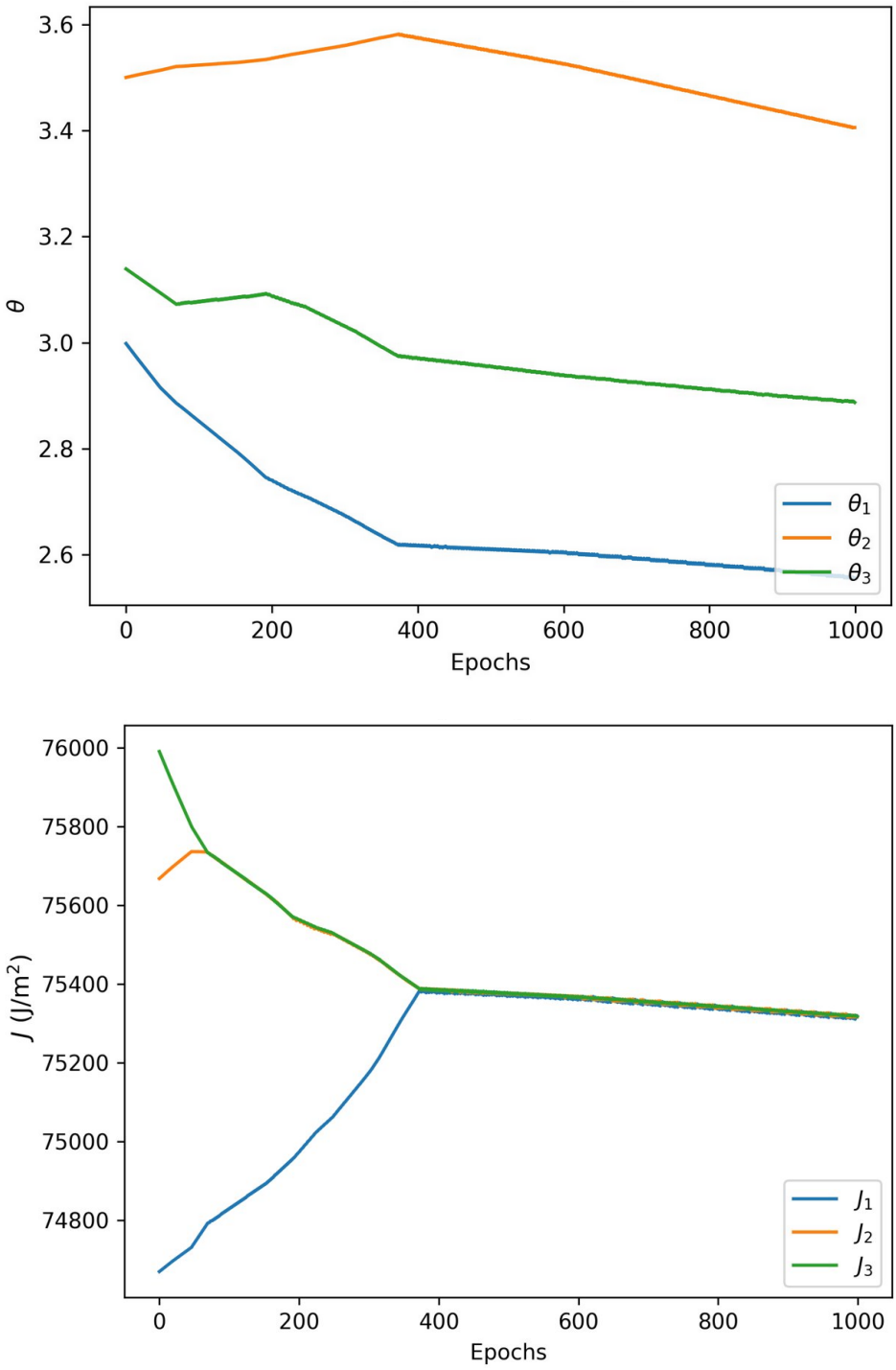

Figure 4.3: The angles and J-integrals obtained during the gradient descent when minimizing $\|J\|_\infty$.

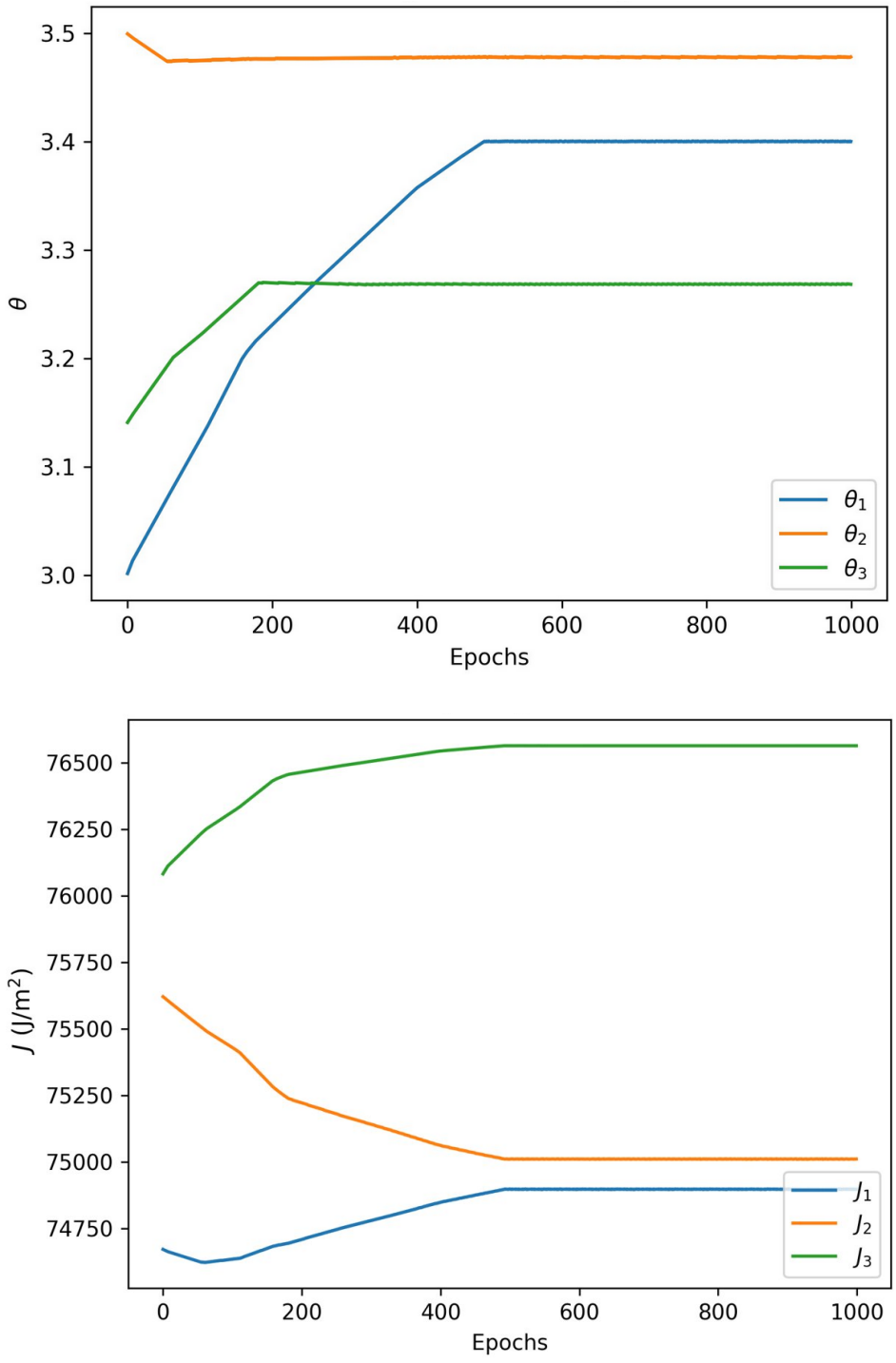

Figure 4.4: The angles and J-integrals obtained during the gradient descent when maximizing $\|J\|_\infty$.

graphies of interest in grain boundary engineering. Based on the identified configurations, the symmetric minimizing configurations correspond to small or large rotations of the lattice, while the maximizing configurations are found near the moderate rotations of the lattice

4.2.2 Application of Coincident Site Lattice Theory

Atomistic considerations have been used in grain boundary engineering to identify favorable configurations. An important example of these considerations is furnished by the theory of coincident site lattices (CSLs). Given the lattices of adjacent grains (e.g. cubic, orthorhombic, triclinic, etc.), CSL theory seeks to find rotations of one grain relative to the other such that a fraction of the atoms in both lattices spatially coincide. Examples of CSLs for cubic lattices are given in Figure 4.7.

If n is the reciprocal of the fraction of coincident atoms in a unit cell of the CSL lattice taken as a subset of the parent lattice, the configuration is denoted as Σn, as demonstrated in Figure 4.8. Here, the CSL lattice is cubic, and there are four non-CSL points and four CSL points within one unit cell. However, as the unit cell only contains a quarter of each corner lattice point, the unit cell in total contains five lattice points (in the plane), out of which a total of one is a CSL point, thus $n = 5$ in this case. In cases where multiple configurations share n but have different rotation axes, additional letters are used to distinguish between these configurations, i.e. Σna. CSL grain boundaries are expected to have low energy due to good atomic fit, and CSL analysis has been used experimentally to characterize the probability of failure in materials [196, 197].

A description of CSL configurations requires some measure of misalignment between grains, which is usually taken to be the *misorientation* angle and is computed as follows: given rotations $\mathbf{Q}_1, \mathbf{Q}_2$ of each grain, the rotation of grain 1 relative to grain 2 is given by $\Delta\mathbf{Q}_{12} = \mathbf{Q}_2\mathbf{Q}_1^{-1}$. Then, the associated misorientation axis $\mathbf{r}$ and angle Θ is computed as follows:

$$\begin{aligned} \cos(\Theta) &= \frac{\mathrm{tr}(\mathbf{Q}_{12}) - 1}{2}, \\ \mathbf{r} &= \frac{1}{\sin(\Theta)} \mathrm{vec}\left[\frac{1}{2}\left(\mathbf{Q}_{12} - \mathbf{Q}_{12}^{\top}\right)\right], \end{aligned} \tag{4.1}$$

where vec is the operation that maps skew-symmetric tensors to the associated vector. However, as the component lattices are invariant under symmetry operations within the associated point group, the above equations do not yield a unique value of Θ. The *disorientation* angle Θ^* remedies this non-uniqueness and is defined through the following optimization problem:

$$\Theta^* = \min_{\mathbf{R}_1 \in \mathscr{P}_1} \min_{\mathbf{R}_2 \in \mathscr{P}_2} \Theta\left(\mathbf{R}_2\mathbf{Q}_2\left(\mathbf{R}_1\mathbf{Q}_1\right)^{-1}\right), \tag{4.2}$$

where $\mathscr{P}_1, \mathscr{P}_2$ are the respective symmetry groups of each lattice, and $\Theta(\bullet)$ indicates calculation of the angle through $(4.1)_1$. The optimization problem (4.2) applies all possible symmetry operations to find the smallest possible angle to describe the grain configuration. These symmetry operations are established and commonplace in the crystal physics literature, e.g. [198].

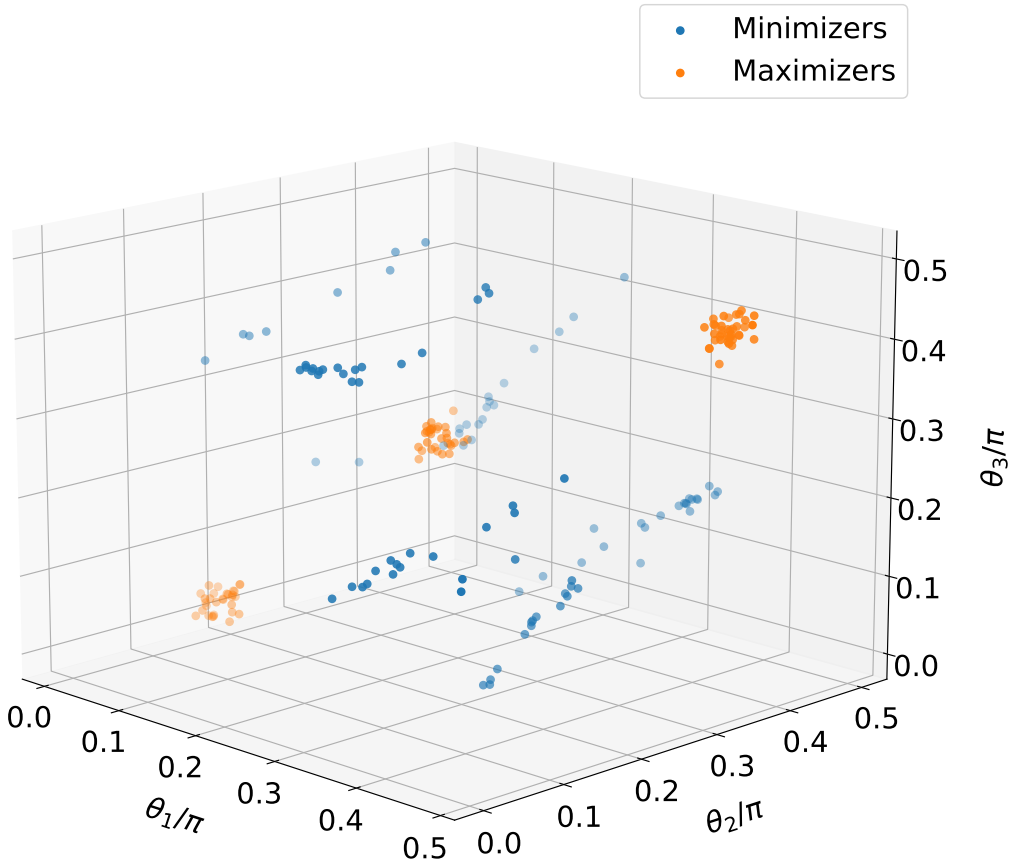

Figure 4.5: 3D scatter plot of the extremal angles obtained after minimizing and maximizing $\|J\|_\infty$.

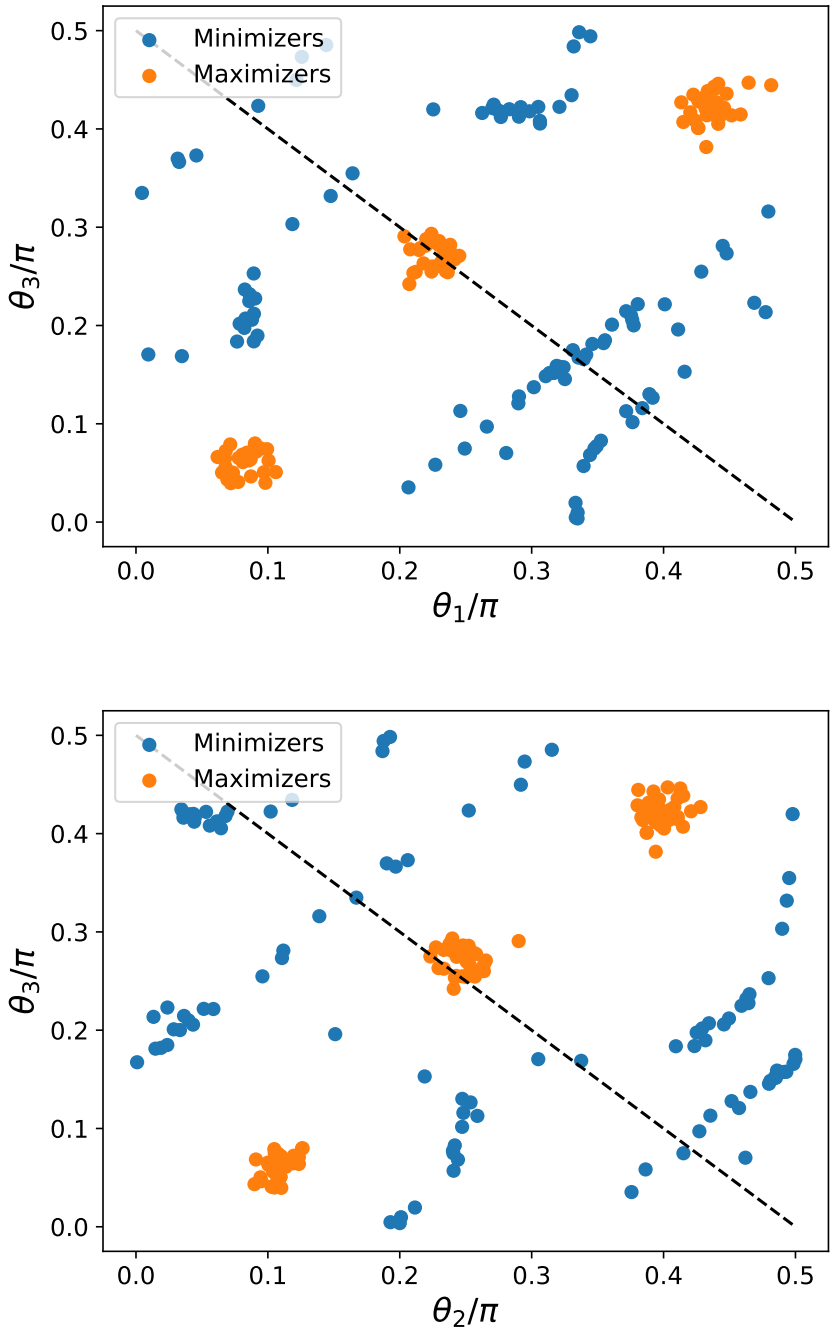

Figure 4.6: 2D projections of the configurations identified through optimization. The black, dashed line indicates the symmetric grain boundary configurations.

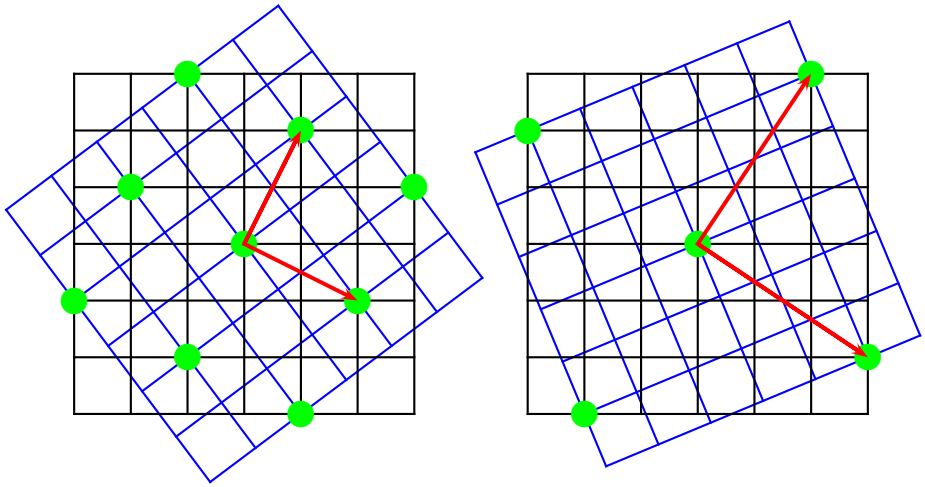

Figure 4.7: Two CSL configurations for cubic lattices, Σ5 (left) and Σ13*a* (right) attained by in-plane rotations of 36.86° and 22.61° respectively. The coincident points are denoted using green dots and the lattice vectors of the CSL are shown in red.

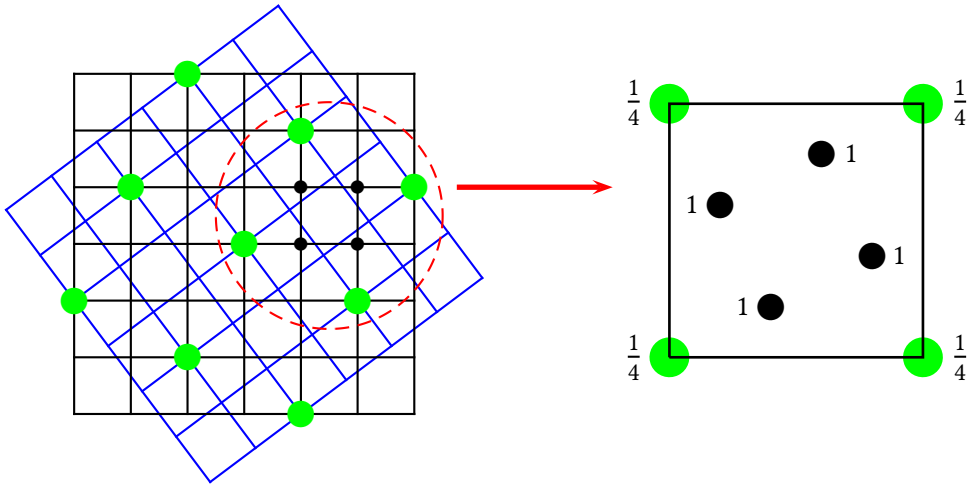

Figure 4.8: Subset of lattice points from one of the parent lattices, showing that in the portion of the parent lattice (black) contained in the CSL unit cell (green), exactly one lattice point is a member of the CSL, following the standard "counting" rules for the number of atoms in a unit cell. An analogous calculation could be performed with the blue lattice, with identical results.

Upon calculation of the disorientation angle for the configurations identified via gradient descent, they are compared to the CSL angles from [199] with axis [001] in Figure 4.9. The configurations that maximize $\|J\|_\infty$ are mostly contained in the lower left-hand portion of the figure, where there are no CSL boundaries, indicating that this region has poor atomic fit between the adjacent grains. In contradistinction, the configurations that minimize $\|J\|_\infty$ are generally spread over the upper right-hand region of the plot, and many configurations are near a CSL configuration for either one, or both, grain boundaries. In particular, many minimizing configurations are close to the $\Sigma 17$ configuration with disorientation angle 28.07°. However, comparatively fewer minimizing configurations are found at the $\Sigma 13$ and $\Sigma 5$ configurations at 22.61° and 36.86° respectively, which have a greater fraction of coincident atoms. Thus, to a limited extent, the grain configurations expected to be favorable from atomistic considerations are reproduced by the surrogate model.

Also prominent in Figure 4.9 is the lack of exact coincidence between the CSL sites and the ones identified from the inverse design. Thus, beyond the qualitative discussion of CSL sites in the previous paragraph, several criteria have been proposed in the literature to identify the maximum possible angular deviation of a grain boundary from a CSL boundary while still being classified as one. Some standard selection criteria are of the form:

$$\Theta_x = \theta_0 \Sigma_x^{-p}, \tag{4.3}$$

where $\theta_0 \approx 15°$ is the low-angle boundary limit, and p is an exponent between $1/2$ and 1. The commonly used Brandon [200] and Palumbo-Aust [201] criteria correspond to $p = 1/2, 5/6$ respectively. In the context of triple junctions, Frary and Schuh [202] proposed the combination of the "sigma combination rule" and the "deviation limit rule":

$$\begin{aligned} \Sigma_1 \Sigma_2 &= d^2 \Sigma_3, \\ \theta_{\max} &\leq \theta_1 + \theta_2. \end{aligned} \tag{4.4}$$

where Σ_i is the CSL value of the i^{th} grain boundary, d is a common divisor of Σ_1 and Σ_2, $\theta_{\max}$ is the largest deviation of any of the three triple junction boundary from the CSL orientation, and θ_1, θ_2 are the deviations of the other two boundaries. The deviation limit rule affords an extension of the sigma combination rule to the case of triple junctions with deviations from the ideal CSL configurations. These different criteria were evaluated on the configurations obtained via inverse design to quantitatively understand the degree of deviation from the ideal configurations. According to the Brandon criterion, 95% and 100% of the minimizers and maximizers are within the acceptable deviation from their nearest CSL configuration respectively. For the Palumbo-Aust criterion, these percentages are 46% and 99%, which is both due to the clustering of maximizers in the low-angle region, as well as the stricter nature of the Palumbo-Aust criterion as compared to the Brandon criterion. The Palumbo-Aust criterion was adopted to create the uncertainty bands for the CSL configurations shown in Figure 4.9 as the lightly-shaded blue regions. However, when using the augmented combination rule, only 19% of the maximizers satisfied both conditions, while none of the minimizers did. The failure to satisfy the augmented criterion was often due to a failure in the sigma combination rule. One potential reason for this failure is that

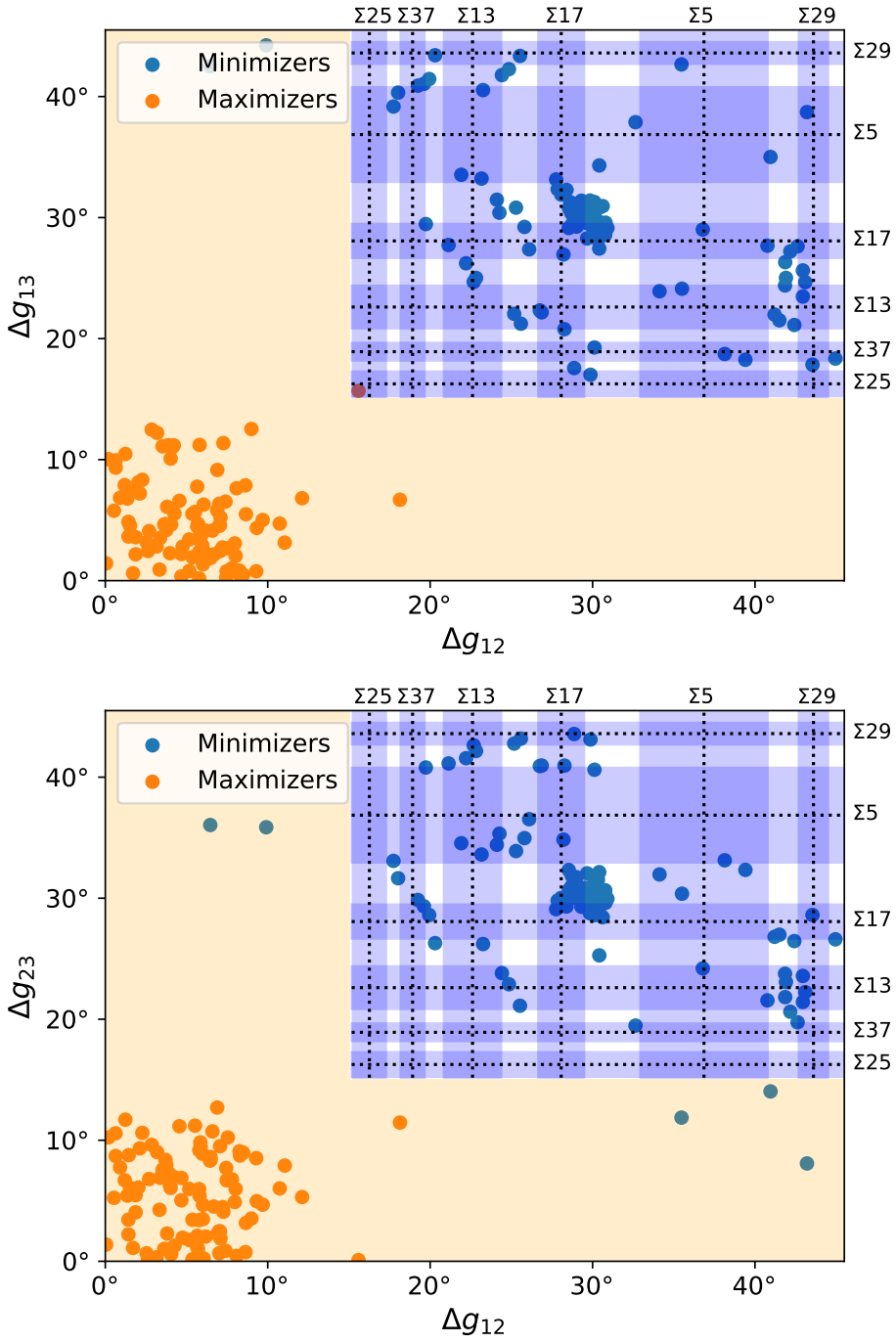

Figure 4.9: The misorientations computed from the optimized configurations as compared to the CSL configurations shown in dashed lines. The shaded orange region indicates the so-called low angle boundaries, and the blue bands represent the acceptable angular deviation as given by the Palumbo-Aust criterion 4.4.

the inclusion of microcracks renders the incompatibility mostly local to the crack tip, which is generally dominated by the two grains on either side. As a result, the triple junction connectivity that is used to derive (4.4) is not maintained in the geometry under current consideration.

Chapter 5

Conclusions

The goal of this thesis is to develop and apply a framework that combines physics-based simulations of grain incompatibility with surrogate modeling to predict quantitative metrics of incompatibility from grain boundary configurational data, for the purposes of identifying optimal grain orientations.

To this end, we developed a physics-based model of grain incompatibilities at a triple junction. The choice of triple junctions was influenced by analytical studies [58, 59] that identify triple junctions as the most likely sites of microfracture initiation in polycrystals. In particular, a triple junction embedded in a disk and subjected to hydrostatic extension was adopted as the model geometry and loading. The resulting stress and strain fields were obtained using the finite element method. Grain incompatibilities were introduced by recourse to rotations of each grain about its [001] axis. Following the approach of Tvergaard and Hutchinson [59], microcracks were introduced along the triple junction boundaries, and the J-integral of fracture mechanics was used to quantify the level of grain incompatibility. Accordingly, the J-integral was implemented in $\sum$MIT as a post-processing step through directly discretizing and evaluating the line integral definition of J.

We then designed a surrogate model to predict the J-integrals for each microcrack as a function of the grain boundary rotation angles. The weights and the biases of the model were calculated using a synthetic dataset produced by the finite element simulations, where the grain rotation angles and J-integrals are the feature and label sets respectively. To form the dataset, 100,000 finite element simulations were conducted. Using the TensorFlow and Keras packages, a multi-layer perceptron was trained using this synthetic dataset.

Then, the physics-based continuum model was verified. The calculation of J was first verified using the method of manufactured solutions with the K-field associated to a plate with a semi-infinite crack. In the case of a homogeneous triple junction, the numerical implementation of J was shown to reproduce path-independence over contours with widely varying size. Regarding the surrogate model, the suitability of the adopted network architecture and training data was assessed by computing the training and validation loss as a function of training dataset size. This demonstrated that the surrogate model was able to achieve very low normalized error and that the network was not subject to overfitting, as the training and validation losses were comparable.

The surrogate model was also shown to accurately reproduce the J-integrals for one-dimensional subsets of the feature space.

We then demonstrated the surrogate model in two contexts. First, we swept over configuration space to create contour maps of the largest energy release rate as a function of the grain boundary angles. It bears emphasis that such a parameter sweep is enabled by the fast and direct evaluation afforded by surrogate models. These contour plots also showed that the learned interpolant accurately reproduces the expected cubic symmetry of the crystal system considered in this thesis. The surrogate model was then also used for the inverse problem of identifying the configurations that minimize or maximize the largest J-integral across the three microcracks. This optimization was done using the method of Peurifoy et al [122], which utilizes the analytical properties of the surrogate model for gradient descent updates. These configurations were then compared to the CSL boundaries commonly used in grain boundary engineering. While some of the minimizers did coincide with CSL angles, the optimization identified many of the maximizing configurations as low-angle boundaries. Application of CSL criteria from the literature suggests that the two materials at each microcrack interface present the dominant physical effect during microstructure optimization.

References

[1] S Miyazaki, Tadashi Kawai, and K Otsuka. "Study of fracture in Cu-Al-Ni shape memory bicrystals". In: *Le Journal de Physique Colloques* 43.C4 (1982), pp. C4–813. DOI: 10.1051/jphyscol:19824133.

[2] S Miyazaki, T Kawai, and K Otsuka. "On the origin of intergranular fracture in β phase shape memory alloys". In: *Scripta Metallurgica* 16.4 (1982), pp. 431–436. DOI: 10.1016/0036-9748(82)90167-3.

[3] Hiroaki Kurishita, Shunichi Kuba, Haruyoshi Kubo, and Hideo Yoshinaga. "Misorientation dependence of grain boundary fracture in molybdenum bicrystals with various ⟨110⟩ twist boundaries". In: *Transactions of the Japan institute of metals* 26.5 (1985), pp. 332–340.

[4] Tadao Watanabe. "Grain boundary design for the control of intergranular fracture". In: *Materials Science Forum*. Vol. 46. Trans Tech Publ. 1989, pp. 25–48.

[5] Tadao Watanabe. "The impact of grain boundary character distribution on fracture in polycrystals". In: *Materials Science and Engineering: A* 176.1 (1994), pp. 39–49. ISSN: 0921-5093. DOI: 10.1016/0921-5093(94)90957-1.

[6] S. M. Ueland, Y. Chen, and C. A. Schuh. "Oligocrystalline Shape Memory Alloys". In: *Adv. Funct. Mater.* 22.10 (2012), pp. 2094–2099. DOI: 10.1002/adfm.201103019.

[7] Stian M Ueland and Christopher A Schuh. "Grain boundary and triple junction constraints during martensitic transformation in shape memory alloys". In: *Journal of Applied Physics* 114.5 (2013), p. 053503. DOI: 10.1063/1.4817170.

[8] Isabel R. Crystal and Christopher A. Schuh. "Grain-size effect on cracking accumulation in yttria-doped zirconia ceramics during cyclic martensitic transformations". In: *Acta Materialia* 209 (2021), p. 116789. ISSN: 1359-6454. DOI: 10.1016/j.actamat.2021.116789.

[9] Jun-Yuan Zheng, J.Q. Ran, and M.W. Fu. "Constitutive modeling of multiscale polycrystals considering grain structures and orientations". In: *International Journal of Mechanical Sciences* 216 (2022), p. 106992. ISSN: 0020-7403. DOI: 10.1016/j.ijmecsci.2021.106992.

[10] Tadao Watanabe. "Grain boundary engineering: historical perspective and future prospects". In: *Journal of materials science* 46.12 (2011), pp. 4095–4115.

[11] Seung-Eek Eagle Park and Wesley Hackenberger. “High performance single crystal piezoelectrics: applications and issues”. In: *Current Opinion in Solid State and Materials Science* 6.1 (2002), pp. 11–18.

[12] Jayse Langdon and Arumugam Manthiram. “A perspective on single-crystal layered oxide cathodes for lithium-ion batteries”. In: *Energy Storage Materials* 37 (2021), pp. 143–160. ISSN: 2405-8297. DOI: https://doi.org/10.1016/j.ensm.2021.02.003.

[13] Suk-Joong L Kang, Ji-Hoon Park, Seok-Young Ko, and Ho-Yong Lee. “Solid-state conversion of single crystals: the principle and the state-of-the-art”. In: *Journal of the American Ceramic Society* 98.2 (2015), pp. 347–360.

[14] Said Abdallah, Shafahat Ali, Salman Pervaiz, and Sathish Kannan. “Conventional Machining of Single Crystal Metals and Super Alloys: A Review”. In: *Journal of Manufacturing Science and Engineering* 144.9 (Apr. 2022), p. 090801. ISSN: 1087-1357. DOI: 10.1115/1.4053807. eprint: https://asmedigitalcollection.asme.org/manufacturingscience/article-pdf/144/9/090801/6872133/manu_144_9_090801.pdf.

[15] E.O. Hall. “The Deformation and Ageing of Mild Steel: III Discussion of Results”. In: *Proceedings of the Physical Society B* 64.9 (1951), pp. 747–753. DOI: 10.1088/0370-1301/64/9/303.

[16] NJ Petch. “The cleavage strength of polycrystals”. In: *J. Iron Steel Inst.* 174 (1953), pp. 25–28.

[17] Z. C. Cordero, B. E. Knight, and C. A. Schuh. “Six decades of the Hall-Petch effect - a survey of grain-size strengthening studies on pure metals”. In: *International Materials Reviews* 61.8 (2016), pp. 495–512. DOI: 10.1080/09506608.2016.1191808.

[18] J.E. Flinn, D.P. Field, G.E. Korth, T.M. Lillo, and J. Macheret. “The flow stress behavior of OFHC polycrystalline copper”. In: *Acta Materialia* 49.11 (2001), pp. 2065–2074. ISSN: 1359-6454. DOI: https://doi.org/10.1016/S1359-6454(01)00102-1.

[19] Tadao Watanabe. “A General Approach to the Control of Grain Boundary Fracture in Brittle Materials (Brittle Fracture)”. In: *Proceedings of the Asian Pacific Conference on Fracture and Strength and International Conference on Advanced Technology in Experimental Mechanics 1.01. 203.* The Japan Society of Mechanical Engineers. 2001, pp. 126–133.

[20] Tadao Watanabe. “An approach to grain boundary design for strong and ductile polycrystals”. In: *Res mechanica* 11.1 (1984), pp. 47–84.

[21] Tadao Watanabe and Sadahiro Tsurekawa. “The control of brittleness and development of desirable mechanical properties in polycrystalline systems by grain boundary engineering”. In: *Acta Materialia* 47.15 (1999), pp. 4171–4185. ISSN: 1359-6454. DOI: 10.1016/S1359-6454(99)00275-X.

[22] P. Davies and V. Randle. “Grain boundary engineering and the role of the interfacial plane”. In: *Materials Science and Technology* 17.6 (2001), pp. 615–626. DOI: 10.1179/026708301101510384. eprint: 10.1179/026708301101510384.

[23] Valerie Randle. “Grain boundary engineering: an overview after 25 years”. In: *Materials science and technology* 26.3 (2010), pp. 253–261.

[24] G Palumbo, PJ King, PC Lichtenberger, KT Aust, and U Erb. “Grain boundary structure control for intergranular stress-corrosion resistance”. In: *MRS Online Proceedings Library (OPL)* 238 (1991).

[25] Nihan Tuncer and Christopher A Schuh. “Melt-cast microfibers of Cu-based shape memory alloy adopt a favorable texture for superelasticity”. In: *Scripta Materialia* 117 (2016), pp. 46–50.

[26] Alan Lai, Zehui Du, Chee Lip Gan, and Christopher A Schuh. “Shape memory and superelastic ceramics at small scales”. In: *Science* 341.6153 (2013), pp. 1505–1508.

[27] Zehui Du et al. “Size effects and shape memory properties in ZrO2 ceramic micro-and nano-pillars”. In: *Scripta Materialia* 101 (2015), pp. 40–43.

[28] Christopher A Schuh and Ke Lu. “Stability of nanocrystalline metals: The role of grain-boundary chemistry and structure”. In: *MRS Bulletin* 46.3 (2021), pp. 225–235.

[29] Herbert Gleiter. “Nanocrystalline materials”. In: *Advanced Structural and Functional Materials: Proceedings of an International Seminar Organized by Deutsche Forschungsanstalt für Luft-und Raumfahrt (DLR), Köln, June 1991*. Springer. 1991, pp. 1–37.

[30] Herbert Gleiter. “Nanostructured materials: basic concepts and microstructure”. In: *Acta materialia* 48.1 (2000), pp. 1–29.

[31] KS Kumar, H Van Swygenhoven, and S Suresh. “Mechanical behavior of nanocrystalline metals and alloys”. In: *Acta materialia* 51.19 (2003), pp. 5743–5774.

[32] E Botcharova, J Freudenberger, and L Schultz. “Mechanical and electrical properties of mechanically alloyed nanocrystalline Cu–Nb alloys”. In: *Acta materialia* 54.12 (2006), pp. 3333–3341.

[33] Andrew J Detor and Christopher A Schuh. “Tailoring and patterning the grain size of nanocrystalline alloys”. In: *Acta Materialia* 55.1 (2007), pp. 371–379.

[34] Jason R Trelewicz and Christopher A Schuh. “The Hall–Petch breakdown in nanocrystalline metals: a crossover to glass-like deformation”. In: *Acta Materialia* 55.17 (2007), pp. 5948–5958.

[35] Jason R Trelewicz and Christopher A Schuh. “The Hall–Petch breakdown at high strain rates: optimizing nanocrystalline grain size for impact applications”. In: *Applied Physics Letters* 93.17 (2008).

[36] Meyers, M.A. and Mishra, A. and Benson, D.J. “Mechanical properties of nanocrystalline materials”. In: *Progress in Materials Science* 51 (2006), pp. 427–556. DOI: 10.1016/j.pmatsci.2005.08.003.

[37] WL Li, NR Tao, Z Han, and K Lu. “Comparisons of dry sliding tribological behaviors between coarse-grained and nanocrystalline copper”. In: *Wear* 274 (2012), pp. 306–312.

[38] Timothy J Rupert and Christopher A Schuh. “Sliding wear of nanocrystalline Ni–W: structural evolution and the apparent breakdown of Archard scaling”. In: *Acta Materialia* 58.12 (2010), pp. 4137–4148.

[39] Andreas Michels, CE Krill, H Ehrhardt, R Birringer, and DT Wu. “Modelling the influence of grain-size-dependent solute drag on the kinetics of grain growth in nanocrystalline materials”. In: *Acta materialia* 47.7 (1999), pp. 2143–2152.

[40] K Lücke and K Detert. “A quantitative theory of grain-boundary motion and recrystallization in metals in the presence of impurities”. In: *Acta Metallurgica* 5.11 (1957), pp. 628–637.

[41] John W Cahn. “The impurity-drag effect in grain boundary motion”. In: *Acta metallurgica* 10.9 (1962), pp. 789–798.

[42] K Boylan, D Ostrander, U Erb, G Palumbo, and KT Aust. “An in-situ tem study of the thermal stability of nanocrystalline NiP”. In: *Scripta metallurgica et materialia* 25.12 (1991), pp. 2711–2716.

[43] R.E Hook and J.P Hirth. “The deformation behavior of isoaxial bicrystals of Fe-3%Si”. In: *Acta Metallurgica* 15.3 (1967), pp. 535–551. ISSN: 0001-6160. DOI: https://doi.org/10.1016/0001-6160(67)90087-9.

[44] S. Zaefferer, J.-C. Kuo, Z. Zhao, M. Winning, and D. Raabe. “On the influence of the grain boundary misorientation on the plastic deformation of aluminum bicrystals”. In: *Acta Materialia* 51.16 (2003), pp. 4719–4735. ISSN: 1359-6454.

[45] Nousha Kheradmand, Horst Vehoff, and Afrooz Barnoush. “An insight into the role of the grain boundary in plastic deformation by means of a bicrystalline pillar compression test and atomistic simulation”. In: *Acta Materialia* 61.19 (2013), pp. 7454–7465. ISSN: 1359-6454. DOI: 10.1016/j.actamat.2013.08.056.

[46] Nousha Kheradmand, Alain Franz Knorr, Michael Marx, and Yun Deng. “Microscopic incompatibility controlling plastic deformation of bicrystals”. In: *Acta Materialia* 106 (2016), pp. 219–228. ISSN: 1359-6454. DOI: 10.1016/j.actamat.2016.01.006.

[47] I. Tiba, T. Richeton, C. Motz, H. Vehoff, and S. Berbenni. “Incompatibility stresses at grain boundaries in Ni bicrystalline micropillars analyzed by an anisotropic model and slip activity”. In: *Acta Materialia* 83 (2015), pp. 227–238. ISSN: 1359-6454. DOI: 10.1016/j.actamat.2014.09.033.

[48] Shao-Shi Rui, Li-Sha Niu, Hui-Ji Shi, Shaolou Wei, and C. Cem Tasan. “Diffraction-based misorientation mapping: A continuum mechanics description”. In: *Journal of the Mechanics and Physics of Solids* 133 (2019), p. 103709. ISSN: 0022-5096. DOI: 10.1016/j.jmps.2019.103709.

[49] Rui Zhang, Zhutian Xu, Linfa Peng, Xinmin Lai, and M.W. Fu. “Intragranularly misoriented grain boundary evolution affected by local constraints and grain size in micro-scale deformation of ultra-thin metallic sheets”. In: *International Journal of Plasticity* 157 (2022), p. 103377. ISSN: 0749-6419. DOI: 10.1016/j.ijplas.2022.103377.

[50] J Gemperlova, V Paidar, and F Kroupa. “Compatibility stresses in deformed bicrystals”. In: *Czechoslovak Journal of Physics B* 39.4 (1989), pp. 427–446.

[51] Morton E. Gurtin and Lallit Anand. “Nanocrystalline grain boundaries that slip and separate: A gradient theory that accounts for grain-boundary stress and conditions at a triple-junction”. In: *Journal of the Mechanics and Physics of Solids* 56.1 (2008). Bridging scales in mechanics - ”Where are the bottom and the top?” The Needleman-Tvergaard Solid Mechanics Symposium, pp. 184–199. ISSN: 0022-5096. DOI: 10.1016/j.jmps.2007.09.001.

[52] A.B. Vakaeva, S.A. Krasnitckii, A.M. Smirnov, M.A. Grekov, and M.Yu. Gutkin. “Stress Concentration and Distribution at Triple Junction Pores of Three-Fold Symmetry in Ceramics”. In: *Reviews on Advanced Materials Science* 57.1 (2018), pp. 63–71. DOI: doi:10.1515/rams-2018-0048.

[53] Wenbin Liu, Yangyang Cheng, Haonan Sui, Jiaqi Fu, and Huiling Duan. “Microstructure-based intergranular fatigue crack nucleation model: Dislocation transmission versus grain boundary cracking”. In: *Journal of the Mechanics and Physics of Solids* 173 (2023), p. 105233. ISSN: 0022-5096. DOI: 10.1016/j.jmps.2023.105233.

[54] J. Qu and J.L. Bassani. “Cracks on bimaterial and bicrystal interfaces”. In: *Journal of the Mechanics and Physics of Solids* 37.4 (1989), pp. 417–433. ISSN: 0022-5096. DOI: 10.1016/0022-5096(89)90022-7.

[55] T. Richeton and S. Berbenni. “Effects of heterogeneous elasticity coupled to plasticity on stresses and lattice rotations in bicrystals: A Field Dislocation Mechanics viewpoint”. In: *European Journal of Mechanics - A/Solids* 37 (2013), pp. 231–247. ISSN: 0997-7538. DOI: 10.1016/j.euromechsol.2012.06.010.

[56] T. Richeton and S. Berbenni. “From bicrystals to spherical inclusions: A superposition method to derive analytical expressions of stress fields in presence of plastic strain gradients”. In: *International Journal of Solids and Structures* 51.3 (2014), pp. 794–807. ISSN: 0020-7683. DOI: 10.1016/j.ijsolstr.2013.11.005.

[57] Kai Liu and Marcel H.F. Sluiter. “Stresses at grain boundaries: The maximum incompatibility stress in an infinitely extended elastic bicrystal under uniaxial loading”. In: *Scripta Materialia* 234 (2023), p. 115570. ISSN: 1359-6462. DOI: 10.1016/j.scriptamat.2023.115570.

[58] A.G. Evans. “Microfracture from thermal expansion anisotropy—I. Single phase systems”. In: *Acta Metallurgica* 26.12 (1978), pp. 1845–1853. ISSN: 0001-6160. DOI: 10.1016/0001-6160(78)90097-4.

[59] Viggo Tvergaard and John W. Hutchinson. “Microcracking in Ceramics Induced by Thermal Expansion or Elastic Anisotropy”. In: *Journal of the American Ceramic Society* 71.3 (1988), pp. 157–166. DOI: 10.1111/j.1151-2916.1988.tb05022.x. eprint: https://ceramics.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1151-2916.1988.tb05022.x.

[60] Fariborz Ghahremani, John W. Hutchinson, and Viggo Tvergaard. “Three-Dimensional Effects in Microcrack Nucleation in Brittle Polycrystals”. In: *Journal of the American Ceramic Society* 73.6 (1990), pp. 1548–1554. DOI: 10.1111/j.1151-2916.1990.tb09795.x. eprint: https://ceramics.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1151-2916.1990.tb09795.x.

[61] Rajeev Ahluwalia, Turab Lookman, and Avadh Saxena. “Elastic Deformation of Polycrystals”. In: *Phys. Rev. Lett.* 91 (5 2003), p. 055501. DOI: 10.1103/PhysRevLett.91.055501.

[62] P. Peralta, A. Schober, and C. Laird. “Elastic stresses in anisotropic bicrystals”. In: *Materials Science and Engineering: A* 169.1 (1993), pp. 43–51. ISSN: 0921-5093. DOI: 10.1016/0921-5093(93)90597-8.

[63] Zhiyi Wang, Alan Lai, Christopher A. Schuh, and Raúl Radovitzky. “Phase transformation and incompatibility at grain boundaries in zirconia-based shape memory ceramics: a micromechanics-based simulation study”. In: *Journal of Materials Science* 57.24 (June 2022), pp. 11132–11150. ISSN: 1573-4803. DOI: 10.1007/s10853-022-07324-3.

[64] C.R. Chen, S.X. Li, and Z.G. Wang. “Characteristics of strain and resolved shear stress in a bicrystal with the grain boundary perpendicular to the tensile axis”. In: *Materials Science and Engineering: A* 247.1 (1998), pp. 15–22. ISSN: 0921-5093. DOI: 10.1016/S0921-5093(97)00858-7.

[65] S. X. Li, D. B. Ren, W. P. Jia, C. R. Chen, X. W. Li, and Z. G. Wang. “On the stress distribution around a triple junction”. In: *Philosophical Magazine A* 80.8 (2000), pp. 1729–1741. DOI: 10.1080/01418610008219080.

[66] C.R. Chen, S.X. Li, J.L. Wen, and W.P. Jia. “Finite element analysis about effects of stiffness distribution on stresses and elastic strain energy near the triple junction in a tricrystal”. In: *Materials Science and Engineering: A* 282.1 (2000), pp. 170–176. ISSN: 0921-5093. DOI: 10.1016/S0921-5093(99)00760-1.

[67] Yaowu Zhao and Robert Tryon. “Automatic 3-D simulation and micro-stress distribution of polycrystalline metallic materials”. In: *Computer Methods in Applied Mechanics and Engineering* 193.36 (2004), pp. 3919–3934. ISSN: 0045-7825. DOI: 10.1016/j.cma.2004.02.015.

[68] Masayuki Kamaya, Yoshihiro Kawamura, and Takayuki Kitamura. "Three-dimensional local stress analysis on grain boundaries in polycrystalline material". In: *International Journal of Solids and Structures* 44.10 (2007), pp. 3267–3277. ISSN: 0020-7683. DOI: 10.1016/j.ijsolstr.2006.09.020.

[69] A. Fallahi and A. Ataee. "Effects of crystal orientation on stress distribution near the triple junction in a tricrystal gamma-TiAl". In: *Materials Science and Engineering: A* 527.18 (2010), pp. 4576–4581. ISSN: 0921-5093. DOI: 10.1016/j.msea.2010.03.040.

[70] Ushasi Roy and Min Zhou. "A computational framework for predicting the fracture toughness of metals as function of microstructure". In: *Journal of the Mechanics and Physics of Solids* 142 (2020), p. 103955. ISSN: 0022-5096. DOI: 10.1016/j.jmps.2020.103955.

[71] Ushasi Roy, David L. McDowell, and Min Zhou. "Effect of grain orientations on fracture behavior of polycrystalline metals". In: *Journal of the Mechanics and Physics of Solids* (2021), p. 104384. ISSN: 0022-5096. DOI: 10.1016/j.jmps.2021.104384.

[72] John Frederick Nye. *Physical properties of crystals: their representation by tensors and matrices*. Oxford university press, 1985.

[73] J.M. Ball and R.D. James. "Fine phase mixtures as minimizers of energy". In: *Archive for Rational Mechanics Analysis* 100.1 (1987), pp. 13–52. DOI: 10.1007/BF00281246.

[74] Kaushik Bhattacharya. "Theory of martensitic microstructure and the shape-memory effect". In: *Available from author: bhatta@ co. caltech. edu* (1998).

[75] Y. C. Shu and K. Bhattacharya. "Domain patterns and macroscopic behaviour of ferroelectric materials". In: *Philosophical Magazine B* 81.12 (2001), pp. 2021–2054. DOI: 10.1080/13642810108208556. eprint: 10.1080/13642810108208556.

[76] John Dundurs. "Boundary conditions at interfaces". In: *Micromechanics and Inhomogeneity: The Toshio Mura 65th Anniversary Volume*. Springer, 1990, pp. 109–114.

[77] J. Dundurs and D. A. Sotiropoulos. "On stress conditions and computations at bi-material interfaces". In: *Computational Mechanics* 21.4 (May 1998), pp. 300–305. ISSN: 1432-0924. DOI: 10.1007/s004660050306.

[78] Amit Acharya. "A model of crystal plasticity based on the theory of continuously distributed dislocations". In: *Journal of the Mechanics and Physics of Solids* 49.4 (2001), pp. 761–784. ISSN: 0022-5096. DOI: 10.1016/S0022-5096(00)00060-0.

[79] M. L. Williams. "The stresses around a fault or crack in dissimilar media". In: *Bulletin of the Seismological Society of America* 49.2 (Apr. 1959), pp. 199–204. ISSN: 0037-1106. DOI: 10.1785/BSSA0490020199. eprint: https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/49/2/199/5303393/bssa0490020199.pdf.

[80] GV Kolosov. "The use of complex diagrams and the theory of functions of a complex variable in the theory of elasticity". In: *Leningrad–Moscow: ONTI* (1935).

[81] Z. Zhao, R. Radovitzky, and A. Cuitino. “A study of surface roughening in f.c.c. metals using Direct Numerical Simulation”. In: *Acta Materialia* 52.20 (2004), pp. 5791–5804. DOI: 10.1016/j.actamat.2004.08.037.

[82] Z. Zhao, S. Kuchnicki, R. Radovitzky, and A. Cuitiño. “Influence of in-grain mesh resolution on the prediction of deformation textures in fcc polycrystals by crystal plasticity FEM”. In: *Acta Materialia* 55.7 (2007), pp. 2361–2373. DOI: 10.1016/j.actamat.2006.11.035.

[83] Z. Zhao, M. Ramesh, D. Raabe, A. Cuitiño, and R. Radovitzky. “Investigation of Three-Dimensional Aspects of Grain-Scale Plastic Surface Deformation of an Aluminum Oligocrystal”. In: *International Journal of Plasticity* 24 (2008), 2278:2297. DOI: http://dx.doi.org/10.1016/j.ijplas.2008.01.002.

[84] A. Jerusalem, M. Dao, S. Suresh, and R. Radovitzky. “Three-dimensional model of strength and ductility of polycrystalline copper containing nanoscale twins”. In: *Acta Materialia* 56 (2008), pp. 4647–4657. DOI: http://dx.doi.org/10.1016/j.actamat.2008.05.033.

[85] D. Gonzalez, I. Simonovski, P.J. Withers, and J. Quinta da Fonseca. “Modelling the effect of elastic and plastic anisotropies on stresses at grain boundaries”. In: *International Journal of Plasticity* 61 (2014), pp. 49–63. ISSN: 0749-6419. DOI: 10.1016/j.ijplas.2014.03.012.

[86] Xiaoxian Zhang and Fionn P.E. Dunne. “3D CP-XFEM modelling of short crack propagation interacting with twist/tilt nickel grain boundaries”. In: *Journal of the Mechanics and Physics of Solids* 168 (2022), p. 105028. ISSN: 0022-5096. DOI: 10.1016/j.jmps.2022.105028.

[87] Ricardo A. Lebensohn, Anand K. Kanjarla, and Philip Eisenlohr. “An elasto-viscoplastic formulation based on fast Fourier transforms for the prediction of micromechanical fields in polycrystalline materials”. In: *International Journal of Plasticity* 32-33 (2012), pp. 59–69. ISSN: 0749-6419. DOI: 10.1016/j.ijplas.2011.12.005.

[88] P.D. Zavattieri and H.D. Espinosa. “Grain level analysis of crack initiation and propagation in brittle materials”. In: *Acta Materialia* 49 (2001), pp. 4291–4311. DOI: 10.1016/S1359-6454(01)00292-0.

[89] G.I. Taylor. “Plastic strain in metals”. In: *Journal of the Institute of Metals* 62 (1938), pp. 307–324.

[90] R. R. M. Mallock and Charles Galton Darwin. “An electrical calculating machine”. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 140.841 (1933), pp. 457–483. DOI: 10.1098/rspa.1933.0081. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1933.0081.

[91] B. Karlsson and B.O. Sundström. “Inhomogeneity in plastic deformation of two-phase steels”. In: *Materials Science and Engineering* 16.1 (1974), pp. 161–168. ISSN: 0025-5416. DOI: https://doi.org/10.1016/0025-5416(74)90150-5.

[92] Bernt O Jaensson and Bengt O Sundström. “Determination of Young’s modulus and poisson’s ratio for WC-Co alloys by the finite element method”. In: *Materials science and engineering* 9 (1972), pp. 217–222.

[93] Bengt O Sundström. “Elastic-plastic behaviour of WC-Co analysed by continuum mechanics”. In: *Materials Science and Engineering* 12.5-6 (1973), pp. 265–276.

[94] Jindrich Jinoch, Sreeramamurthy Ankem, and Harold Margolin. “Calculations of stress-strain curve and stress and strain distributions for an $\alpha - \beta$ Ti-8Mn alloy”. In: *Materials Science and Engineering* 34.3 (1978), pp. 203–211. ISSN: 0025-5416. DOI: https://doi.org/10.1016/0025-5416(78)90052-6.

[95] S. Ankem and H. Margolin. “Modeling Deformation in Two-Phase Alloys”. In: *JOM* 38.4 (Apr. 1986), pp. 25–29. ISSN: 1543-1851. DOI: 10.1007/BF03257865.

[96] Marc A. Meyersm and E. Ashworth. “A model for the effect of grain size on the yield stress of metals”. In: *Philosophical Magazine A* 46.5 (1982), pp. 737–759. DOI: 10.1080/01418618208236928. eprint: 10.1080/01418618208236928.

[97] J.R. Rice. “A path-independent integral and approximate analysis of strain concentrations by notches and cracks”. In: *Journal of Applied Mechanics* 35 (1968), pp. 379–386.

[98] Surya R. Kalidindi, J. Houskamp, G. Proust, and H. Duvvuru. “Microstructure Sensitive Design with First Order Homogenization Theories and Finite Element Codes”. In: *Textures of Materials - ICOTOM 14*. Vol. 495. Materials Science Forum. Trans Tech Publications Ltd, July 2005, pp. 23–30. DOI: 10.4028/www.scientific.net/MSF.495-497.23.

[99] David L. McDowell. “Simulation-assisted materials design for the concurrent design of materials and products”. In: *JOM* 59.9 (Sept. 2007), pp. 21–25. ISSN: 1543-1851. DOI: 10.1007/s11837-007-0111-7.

[100] David T. Fullwood, Stephen R. Niezgoda, Brent L. Adams, and Surya R. Kalidindi. “Microstructure sensitive design for performance optimization”. In: *Progress in Materials Science* 55.6 (2010), pp. 477–562. ISSN: 0079-6425. DOI: 10.1016/j.pmatsci.2009.08.002.

[101] Arash Dehghan Banadaki, Mark A. Tschopp, and Srikanth Patala. “An efficient Monte Carlo algorithm for determining the minimum energy structures of metallic grain boundaries”. In: *Computational Materials Science* 155 (2018), pp. 466–475. ISSN: 0927-0256. DOI: 10.1016/j.commatsci.2018.09.017.

[102] Yongmei Zhang and Liangliang Liu. “Computational design of microstructures of textured ferroelectric ceramics by phase field simulation”. In: *Computational Materials Science* 159 (2019), pp. 24–31. ISSN: 0927-0256. DOI: 10.1016/j.commatsci.2018.11.051.

[103] Matthew Guziewski, Arash D. Banadaki, Srikanth Patala, and Shawn P. Coleman. “Application of Monte Carlo techniques to grain boundary structure optimization in silicon and silicon-carbide”. In: *Computational Materials Science* 182 (2020), p. 109771. ISSN: 0927-0256. DOI: 10.1016/j.commatsci.2020.109771.

[104] Matthew Guziewski, David Montes de Oca Zapiain, Rémi Dingreville, and Shawn P. Coleman. "Microscopic and Macroscopic Characterization of Grain Boundary Energy and Strength in Silicon Carbide via Machine-Learning Techniques". In: *ACS Applied Materials & Interfaces* 13.2 (2021), pp. 3311–3324. DOI: 10.1021/acsami.0c15980.

[105] Alexander IJ Forrester, Neil W Bressloff, and Andy J Keane. "Optimization using surrogate models and partially converged computational fluid dynamics simulations". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 462.2071 (2006), pp. 2177–2204.

[106] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. "Physics-informed machine learning". In: *Nature Reviews Physics* 3.6 (June 2021), pp. 422–440. ISSN: 2522-5820. DOI: 10.1038/s42254-021-00314-5.

[107] Shengze Cai, Zhicheng Wang, Sifan Wang, Paris Perdikaris, and George Em Karniadakis. "Physics-informed neural networks for heat transfer problems". In: *Journal of Heat Transfer* 143.6 (2021), p. 060801.

[108] Khemraj Shukla, Ameya D Jagtap, James L Blackshire, Daniel Sparkman, and George Em Karniadakis. "A physics-informed neural network for quantifying the microstructural properties of polycrystalline nickel using ultrasound data: A promising approach for solving inverse problems". In: *IEEE Signal Processing Magazine* 39.1 (2021), pp. 68–77.

[109] Ehsan Haghighat, Ali Can Bekar, Erdogan Madenci, and Ruben Juanes. "Deep learning for solution and inversion of structural mechanics and vibrations". In: *Modeling and Computation in Vibration Problems, Volume 2*. 2053-2563. IOP Publishing, 2021, pp. 1–17. ISBN: 978-0-7503-3487-7. DOI: 10.1088/978-0-7503-3487-7ch1.

[110] Danial Amini, Ehsan Haghighat, and Ruben Juanes. "Physics-Informed Neural Network Solution of Thermo-Hydro-Mechanical Processes in Porous Media". In: *Journal of Engineering Mechanics* 148.11 (2022), p. 04022070. DOI: 10.1061/(ASCE)EM.1943-7889.0002156. eprint: https://ascelibrary.org/doi/pdf/10.1061/%28ASCE%29EM.1943-7889.0002156.

[111] Danial Amini, Ehsan Haghighat, and Ruben Juanes. "Inverse modeling of nonisothermal multiphase poromechanics using physics-informed neural networks". In: *Journal of Computational Physics* 490 (2023), p. 112323. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2023.112323.

[112] Jan Niklas Fuhg, Christoph Böhm, Nikolaos Bouklas, Amelie Fau, Peter Wriggers, and Michele Marino. "Model-data-driven constitutive responses: Application to a multiscale computational framework". In: *International Journal of Engineering Science* 167 (2021), p. 103522. ISSN: 0020-7225. DOI: 10.1016/j.ijengsci.2021.103522.

[113] Jan Niklas Fuhg, Amélie Fau, Nikolaos Bouklas, and Michele Marino. "Enhancing phenomenological yield functions with data: Challenges and opportunities". In: *European Journal of Mechanics - A/Solids* 99 (2023), p. 104925. ISSN: 0997-7538. DOI: 10.1016/j.euromechsol.2023.104925.

[114] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. “Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations”. In: *arXiv preprint arXiv:1711.10561* (2017).

[115] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. “Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations”. In: *arXiv preprint arXiv:1711.10566* (2017).

[116] Andrea Rovinelli, Michael D. Sangid, Henry Proudhon, and Wolfgang Ludwig. “Using machine learning and a data-driven approach to identify the small fatigue crack driving force in polycrystalline materials”. In: *npj Computational Materials* 4.1 (2018), p. 35. ISSN: 2057-3960. DOI: 10.1038/s41524-018-0094-7.

[117] Karthik Duraisamy, Gianluca Iaccarino, and Heng Xiao. “Turbulence Modeling in the Age of Data”. In: *Annual Review of Fluid Mechanics* 51.1 (2019), pp. 357–377. DOI: 10.1146/annurev-fluid-010518-040547. eprint: 10.1146/annurev-fluid-010518-040547.

[118] Raphaël Pestourie, Youssef Mroueh, Thanh V. Nguyen, Payel Das, and Steven G. Johnson. “Active learning of deep surrogates for PDEs: application to metasurface design”. In: *npj Computational Materials* 6.1 (Oct. 2020), p. 164. ISSN: 2057-3960. DOI: 10.1038/s41524-020-00431-2.

[119] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”. In: *Journal of Computational Physics* 378 (2019), pp. 686–707.

[120] Yuyao Chen, Lu Lu, George Em Karniadakis, and Luca Dal Negro. “Physics-informed neural networks for inverse problems in nano-optics and metamaterials”. In: *Optics express* 28.8 (2020), pp. 11618–11633.

[121] Hernan J. Logarzo, German Capuano, and Julian J. Rimoli. “Smart constitutive laws: Inelastic homogenization through machine learning”. In: *Computer Methods in Applied Mechanics and Engineering* 373 (2021), p. 113482. ISSN: 0045-7825. DOI: 10.1016/j.cma.2020.113482.

[122] John Peurifoy et al. “Nanophotonic particle simulation and inverse design using artificial neural networks”. In: *Science Advances* 4.6 (2018), eaar4206. DOI: 10.1126/sciadv.aar4206. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.aar4206.

[123] Rafael Gómez-Bombarelli et al. “Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules”. In: *ACS Central Science* 4.2 (2018). PMID: 29532027, pp. 268–276. DOI: 10.1021/acscentsci.7b00572. eprint: 10.1021/acscentsci.7b00572.

[124] Chun-Teh Chen and Grace X Gu. “Generative deep neural networks for inverse materials design using backpropagation and active learning”. In: *Advanced Science* 7.5 (2020), p. 1902607.

[125] Ankita Mangal and Elizabeth A. Holm. “Applied machine learning to predict stress hotspots I: Face centered cubic materials”. In: *International Journal of Plasticity* 111 (2018), pp. 122–134. ISSN: 0749-6419. DOI: 10.1016/j.ijplas.2018.07.013.

[126] Ankita Mangal and Elizabeth A. Holm. “Applied machine learning to predict stress hotspots II: Hexagonal close packed materials”. In: *International Journal of Plasticity* 114 (2019), pp. 1–14. ISSN: 0749-6419. DOI: 10.1016/j.ijplas.2018.08.003.

[127] Sheng Zhang et al. “Predicting grain boundary damage by machine learning”. In: *International Journal of Plasticity* 150 (2022), p. 103186. ISSN: 0749-6419. DOI: 10.1016/j.ijplas.2021.103186.

[128] A.L. Frankel, R.E. Jones, C. Alleman, and J.A. Templeton. “Predicting the mechanical response of oligocrystals with deep learning”. In: *Computational Materials Science* 169 (2019), p. 109099. ISSN: 0927-0256. DOI: 10.1016/j.commatsci.2019.109099.

[129] Sean P. Donegan, Navneet Kumar, and Michael A. Groeber. “Associating local microstructure with predicted thermally-induced stress hotspots using convolutional neural networks”. In: *Materials Characterization* 158 (2019), p. 109960. ISSN: 1044-5803. DOI: 10.1016/j.matchar.2019.109960.

[130] R. B. Vieira and J. Lambros. “Machine Learning Neural-Network Predictions for Grain-Boundary Strain Accumulation in a Polycrystalline Metal”. In: *Experimental Mechanics* 61.4 (Jan. 2021), pp. 627–639. DOI: 10.1007/s11340-020-00687-1.

[131] Ruoqian Liu, Abhishek Kumar, Zhengzhang Chen, Ankit Agrawal, Veera Sundararaghavan, and Alok Choudhary. “A predictive machine learning approach for microstructure optimization and materials design”. In: *Scientific Reports* 5.1 (June 2015), p. 11551. ISSN: 2045-2322. DOI: 10.1038/srep11551.

[132] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015.

[133] François Chollet et al. *Keras*. 2015.

[134] The ∑MIT Development Group. *∑MIT, A scalable computational framework for large-scale simulation of complex mechanical response of materials*. 2023.

[135] Tatsuya Okada et al. “Grain-boundary sliding and its accommodation at triple junctions in aluminum and copper tricrystals”. In: *Materials Transactions* 60.1 (2019), pp. 86–92.

[136] Hsuan-Ling Hsu, Hsuan Lee, Chi-Wei Wang, Chenju Liang, and Chih-Ming Chen. “Impurity evaporation and void formation in Sn/Cu solder joints”. In: *Materials Chemistry and Physics* 225 (2019), pp. 153–158. ISSN: 0254-0584. DOI: https://doi.org/10.1016/j.matchemphys.2018.12.036.

[137] I.A. Ovid’ko, A.G. Sheinerman, and N.V. Skiba. “Elongated nanoscale voids at deformed special grain boundary structures in nanocrystalline materials”. In: *Acta Materialia* 59.2 (2011), pp. 678–685. ISSN: 1359-6454. DOI: https://doi.org/10.1016/j.actamat.2010.10.005.

[138] Ming Ke, Walter W. Milligan, Stephen A. Hackney, John E. Carsley, and Elias C. Aifantis. "HREM Study of Fracture and Deformation Behavior of Nanostructured Thin Films". In: *MRS Proceedings* 308 (1993), p. 565. DOI: 10.1557/PROC-308-565.

[139] Hyo-Jong Lee et al. "Characterization of the crystallographic microstructure of the stress-induced void in Cu interconnects". In: *Applied Physics Letters* 92.14 (2008).

[140] JA Querin, JA Schneider, and MF Horstemeyer. "Analysis of micro void formation at grain boundary triple points in monotonically strained AA6022-T43 sheet metal". In: *Materials Science and Engineering: A* 463.1-2 (2007), pp. 101–106.

[141] D.X. Li, D.H. Ping, H.Q. Ye, X.Y. Qin, and X.J. Wu. "HREM study of the microstructure in nanocrystalline materials". In: *Materials Letters* 18.1 (1993), pp. 29–34. ISSN: 0167-577X. DOI: https://doi.org/10.1016/0167-577X(93)90051-X.

[142] R.C. Hugo, H. Kung, J.R. Weertman, R. Mitra, J.A. Knapp, and D.M. Follstaedt. "In-situ TEM tensile testing of DC magnetron sputtered and pulsed laser deposited Ni thin films". In: *Acta Materialia* 51.7 (2003), pp. 1937–1943. ISSN: 1359-6454. DOI: https://doi.org/10.1016/S1359-6454(02)00599-2.

[143] K.S. Kumar, S. Suresh, M.F. Chisholm, J.A. Horton, and P. Wang. "Deformation of electrodeposited nanocrystalline nickel". In: *Acta Materialia* 51.2 (2003), pp. 387–405. ISSN: 1359-6454. DOI: https://doi.org/10.1016/S1359-6454(02)00421-4.

[144] Junya Inoue, Yosuke Fujii, and Toshihiko Koseki. "Void formation in nanocrystalline Cu film during uniaxial relaxation test". In: *Acta Materialia* 56.17 (2008), pp. 4921–4931. ISSN: 1359-6454. DOI: https://doi.org/10.1016/j.actamat.2008.06.008.

[145] Jia Li, Qihong Fang, and Youwen Liu. "Void formation of nanocrystalline materials at the triple junction of grain boundaries". In: *Materials Research Express* 1.1 (2014), p. 015013.

[146] M.A. Charpagne et al. "Slip localization in Inconel 718: A three-dimensional and statistical perspective". In: *Acta Materialia* 215 (2021), p. 117037. ISSN: 1359-6454. DOI: https://doi.org/10.1016/j.actamat.2021.117037.

[147] Shigeaki Kobayashi, Masafumi Terakado, Kousei Ebata, Haruhi Tsutsui, Takateru Yamamuro, and Sadahiro Tsurekawa. "Low-angle grain boundary engineering based on texture control for enhancement of corrosion and fatigue resistance in 436L type ferritic stainless steel". In: *Journal of Materials Science* 59.3 (2024), pp. 1010–1026. ISSN: 1573-4803. DOI: 10.1007/s10853-023-09227-3.

[148] Shunyu Yao et al. "Effects of various grain boundary engineering processing on microstructure and corrosion behaviors of 304 stainless steel analyzed with a fractal model". In: *Journal of Materials Research and Technology* 25 (2023), pp. 13–24. ISSN: 2238-7854. DOI: https://doi.org/10.1016/j.jmrt.2023.05.175.

[149] Shigeaki Kobayashi, Ryosuke Kobayashi, and Tadao Watanabe. "Control of grain boundary connectivity based on fractal analysis for improvement of intergranular corrosion resistance in SUS316L austenitic stainless steel". In: *Acta Materialia* 102 (2016), pp. 397–405. ISSN: 1359-6454. DOI: https://doi.org/10.1016/j.actamat.2015.08.075.

[150] Shigeaki Kobayashi, Tatsuya Maruyama, Sakae Saito, Sadahiro Tsurekawa, and Tadao Watanabe. “In situ observations of crack propagation and role of grain boundary microstructure in nickel embrittled by sulfur”. In: *Journal of Materials Science* 49.11 (2014), pp. 4007–4017. ISSN: 1573-4803. DOI: 10.1007/s10853-014-8056-z.

[151] S. Tsurekawa S. Kobayashi and T. Watanabe. “Structure-dependent triple junction hardening and intergranular fracture in molybdenum”. In: *Philosophical Magazine* 86.33-35 (2006), pp. 5419–5429. DOI: 10.1080/14786430600672711. eprint: https://doi.org/10.1080/14786430600672711.

[152] Shigeaki Kobayashi, Sadahiro Tsurekawa, and Tadao Watanabe. “Roles of structure-dependent hardening at grain boundaries and triple junctions in deformation and fracture of molybdenum polycrystals”. In: *Materials Science and Engineering: A* 483-484 (2008). 14th International Conference on the Strength of Materials, pp. 712–715. ISSN: 0921-5093. DOI: https://doi.org/10.1016/j.msea.2006.12.161.

[153] M. S. Wu, K. Zhou, and A. A. Nazarov. “Crack nucleation at disclinated triple junctions”. In: *Phys. Rev. B* 76 (13 2007), p. 134105. DOI: 10.1103/PhysRevB.76.134105.

[154] Nayden Kambouchev, Javier Fernandez, and Raul Radovitzky. “A polyconvex model for materials with cubic symmetry”. In: *Modelling and Simulation in Materials Science and Engineering* 15.5 (2007), p. 451.

[155] J.M. Ball. “Convexity conditions and existence theorems in nonlinear elasticity”. In: *Arch. Rat. Mech. Anal.* 63 (1977), pp. 337–403.

[156] J. M. Ball. “Constitutive inequalities and existence theorems in nonlinear elastostatics”. In: *Nonlinear Analysis and Mechanics, Heriot-Watt Symposium*. 1977, pp. 187–241.

[157] L. Anand and S. Govindjee. *Continuum Mechanics of Solids*. Oxford Graduate Texts. Oxford University Press, 2020. ISBN: 9780198864721.

[158] John Price Hirth and Jens Lothe. *Theory of dislocations*. eng. 2nd ed. New York: Wiley, 1982. ISBN: 0471091251; 9780471091257.

[159] Nayden Kambouchev, J. Fernandez, and Raul Radovitzky. “Polyconvex Model for Materials with Cubic Anisotropy”. In: (June 2005).

[160] S. Balay et al. *PETSc Users Manual*. Tech. rep. ANL-95/11 - Revision 3.13. Argonne National Laboratory, 2020.

[161] M. E. Gurtin. *Configurational Forces as Basic Concepts of Continuum Physics*. Springer-Verlag, 1999.

[162] B. Budiansky and J. R. Rice. “Conservation Laws and Energy-Release Rates”. In: *Journal of Applied Mechanics* 40.1 (Mar. 1973), pp. 201–203. ISSN: 0021-8936. DOI: 10.1115/1.3422926. eprint: https://asmedigitalcollection.asme.org/appliedmechanics/article-pdf/40/1/201/5452980/201_1.pdf.

[163] Colin R. Meyer, John W. Hutchinson, and James R. Rice. "The Path-Independent M Integral Implies the Creep Closure of Englacial and Subglacial Channels". In: *Journal of Applied Mechanics* 84.1 (Oct. 2016), p. 011006. ISSN: 0021-8936. DOI: 10.1115/1.4034828. eprint: https://asmedigitalcollection.asme.org/appliedmechanics/article-pdf/84/1/011006/5973442/jam_084_01_011006.pdf.

[164] K Ohji, K Ogura, and S Kubo. "Transactions, Japanese Society for Mechanical Engineers". In: *Transactions, Japanese Society for Mechanical Engineers* 42 (1976), pp. 350–358.

[165] JD Landes and JA Begley. "A fracture mechanics approach to creep crack growth". In: *Astm Stp* 590 (1976), pp. 128–148.

[166] KM Nikbin, GA Webster, and CE Turner. "Relevance of nonlinear fracture mechanics to creep cracking". In: *Cracks and fracture*. ASTM International, 1976.

[167] D. A. Wagner and J. C. Simo. "Fracture parameter for thermoinelasticity". In: *International Journal of Fracture* 56.2 (1990), pp. 1573–2673. DOI: 10.1007/BF00015598.

[168] J. C. Simo and T. Honein. "Variational Formulation, Discrete Conservation Laws, and Path-Domain Independent Integrals for Elasto-Viscoplasticity". In: *Journal of Applied Mechanics* 57.3 (Sept. 1990), pp. 488–497. ISSN: 0021-8936. DOI: 10.1115/1.2897050. eprint: https://asmedigitalcollection.asme.org/appliedmechanics/article-pdf/57/3/488/5461000/488_1.pdf.

[169] Satya N. Atluri. "Path-independent integrals in finite elasticity and inelasticity, with body forces, inertia, and arbitrary crack-face conditions". In: *Engineering Fracture Mechanics* 16.3 (1982), pp. 341–364. ISSN: 0013-7944. DOI: https://doi.org/10.1016/0013-7944(82)90113-8.

[170] E. Noether. "Invariante Variationsprobleme". ger. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918 (1918), pp. 235–257.

[171] J. K. Knowles and Eli Sternberg. "On a class of conservation laws in linearized and finite elastostatics". In: *Archive for Rational Mechanics and Analysis* 44.3 (Jan. 1972), pp. 187–211. ISSN: 1432-0673. DOI: 10.1007/BF00250778.

[172] J. D. Eshelby. "The elastic energy-momentum tensor". In: *Journal of Elasticity* 5.3 (Nov. 1975), pp. 321–335. ISSN: 1573-2681. DOI: 10.1007/BF00126994.

[173] Ted Anderson. *Fracture Mechanics: Fundamentals and Applications*. CRC Press, 2017.

[174] C. F. Shih, B. Moran, and T. Nakamura. "Energy release rate along a three-dimensional crack front in a thermally stressed body". In: *International Journal of Fracture* 30.2 (Feb. 1986), pp. 79–102. ISSN: 1573-2673. DOI: 10.1007/BF00034019.

[175] Chenzhuo Li, Xinyue Wei, Meng Wang, Mokhtar Adda-Bedia, and John M. Kolinski. "Crack tip kinematics reveal the process zone structure in brittle hydrogel fracture". In: *Journal of the Mechanics and Physics of Solids* 178 (2023), p. 105330. ISSN: 0022-5096. DOI: https://doi.org/10.1016/j.jmps.2023.105330.

[176] Ariel Livne, Eran Bouchbinder, Ilya Svetlizky, and Jay Fineberg. "The near-tip fields of fast cracks". In: *Science* 327.5971 (2010), pp. 1359–1363.

[177] Meng Wang, Mokhtar Adda-Bedia, John M Kolinski, and Jay Fineberg. "How hidden 3D structure within crack fronts reveals energy balance". In: *Journal of the Mechanics and Physics of Solids* 161 (2022), p. 104795.

[178] H.M. Westergaard. "Bearing Pressures and Cracks". In: *Journal of Applied Mechanics* 6 (1939), pp. 49–53.

[179] Martin H Sadd. *Elasticity: theory, applications, and numerics*. Academic Press, 2009.

[180] J. T. Oden and John O'Leary. "Some Remarks on Finite Element Approximations of Crack Problems and an Analysis of Hybrid Methods". In: *Journal of Structural Mechanics* 6.4 (1978), pp. 415–436. ISSN: 03601218. DOI: 10.1080/03601217808907347.

[181] Kurt Hornik. "Approximation capabilities of multilayer feedforward networks". In: *Neural Networks* 4.2 (1991), pp. 251–257. ISSN: 0893-6080. DOI: 10.1016/0893-6080(91)90009-T.

[182] Rich Caruana, Steve Lawrence, and C. Giles. "Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping". In: *Advances in Neural Information Processing Systems*. Ed. by T. Leen, T. Dietterich, and V. Tresp. Vol. 13. MIT Press, 2000.

[183] Roberto Battiti. "First-and second-order methods for learning: between steepest descent and Newton's method". In: *Neural computation* 4.2 (1992), pp. 141–166.

[184] Rudy Setiono and Lucas Chi Kwong Hui. "Use of a quasi-Newton method in a feedforward neural network construction algorithm". In: *IEEE Transactions on Neural Networks* 6.1 (1995), pp. 273–277.

[185] Donald Goldfarb, Yi Ren, and Achraf Bahamou. "Practical quasi-newton methods for training deep neural networks". In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 2386–2396.

[186] B Robitaille, B Marcos, M Veillette, and G Payre. "Modified quasi-Newton methods for training neural networks". In: *Computers & chemical engineering* 20.9 (1996), pp. 1133–1140.

[187] Léon Bottou. "Large-scale machine learning with stochastic gradient descent". In: *Proceedings of COMPSTAT'2010: 19th International Conference on Computational StatisticsParis France, August 22-27, 2010 Keynote, Invited and Contributed Papers*. Springer. 2010, pp. 177–186.

[188] John Duchi, Elad Hazan, and Yoram Singer. "Adaptive subgradient methods for online learning and stochastic optimization." In: *Journal of machine learning research* 12.7 (2011).

[189] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2017. arXiv: 1412.6980 [cs.LG].

[190] Sashank J Reddi, Satyen Kale, and Sanjiv Kumar. “On the convergence of ADAM and beyond”. In: *arXiv preprint arXiv:1904.09237* (2019).

[191] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., 2001.

[192] Yoram Reich and S.V. Barai. “Evaluating machine learning models for engineering problems”. In: *Artificial Intelligence in Engineering* 13.3 (1999), pp. 257–272. ISSN: 0954-1810. DOI: 10.1016/S0954-1810(98)00021-1.

[193] Valerie Randle. “The coincidence site lattice and the ‘sigma enigma’”. In: *Materials Characterization* 47.5 (2001), pp. 411–416.

[194] Stuart Geman, Elie Bienenstock, and René Doursat. “Neural Networks and the Bias/Variance Dilemma”. In: *Neural Computation* 4.1 (Jan. 1992), pp. 1–58. ISSN: 0899-7667. DOI: 10.1162/neco.1992.4.1.1. eprint: https://direct.mit.edu/neco/article-pdf/4/1/1/812244/neco.1992.4.1.1.pdf.

[195] Steve Lawrence, Ah Chung Tsoi, and Andrew D Back. “Function approximation with neural networks and local methods: Bias, variance and smoothness”. In: *Australian conference on neural networks*. Vol. 1621. Citeseer. 1996.

[196] A Creuziger and WC Crone. “Grain boundary fracture in CuAlNi shape memory alloys”. In: *Materials Science and Engineering: A* 498.1 (2008), pp. 404–411. DOI: 10.1016/j.msea.2008.08.039.

[197] P Lin, KT Aust, G Palumbo, and U Erb. “Influence of grain boundary character distribution on sensitization and intergranular corrosion of alloy 600”. In: *Scripta Metallurgica et materialia* 33.9 (1995).

[198] Michael John Preston. Musgrave. *Crystal acoustics; introduction to the study of elastic waves and vibrations in crystals [by] M. J. P. Musgrave*. English. Holden-Day San Francisco, 1970, xv, 288 p.

[199] H. Grimmer, W. Bollmann, and D. H. Warrington. “Coincidence-site lattices and complete pattern-shift in cubic crystals”. In: *Acta Crystallographica Section A* 30.2 (1974), pp. 197–207. DOI: https://doi.org/10.1107/S056773947400043X. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1107/S056773947400043X.

[200] DG Brandon. “The structure of high-angle grain boundaries”. In: *Acta metallurgica* 14.11 (1966), pp. 1479–1484.

[201] G. Palumbo and K.T. Aust. “Structure-dependence of intergranular corrosion in high purity nickel”. In: *Acta Metallurgica et Materialia* 38.11 (1990), pp. 2343–2352. ISSN: 0956-7151. DOI: https://doi.org/10.1016/0956-7151(90)90101-L.

[202] M. Frary and C.A. Schuh. “Combination rule for deviant CSL grain boundaries at triple junctions”. In: *Acta Materialia* 51.13 (2003), pp. 3731–3743. ISSN: 1359-6454. DOI: https://doi.org/10.1016/S1359-6454(03)00188-5.