SYSTEMATIC DESIGN OF METALLIC FUNCTIONALLY GRADED MATERIALS & STRUCTURES

A Dissertation

by

MARSHALL DONALD ALLEN

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Richard J. Malak
Co-Chair of Committee,	Raymundo Arróyave
Committee Members,	Ibrahim Karaman
	Vinayak Krishnamurthy
Head of Department,	Guillermo Aguilar

May 2025

Major Subject: Mechanical Engineering

ABSTRACT

As performance demands on metallic structures increase, multi-material solutions like compositionally graded alloys (CGAs) are essential for achieving localized performance objectives where a single alloy is insufficient. Metal additive manufacturing (AM) provides unique potential for producing complex CGA structures due to its spatial control over composition. To realize effective multi-material design, sophisticated methods are required to handle the intricate alloy processing landscape. This dissertation presents a framework to address this intricate landscape for systematic materials design of CGAs.

First, this work presents a comprehensive materials information schema to connect alloy processing to the resultant structure and properties. At the core of this framework is the proposed Alloy Topology-Linked informAtion Schema (ATLAS), a graph-based schema for instantiating materials graph databases for systematic materials design of CGAs. Unlike previous approaches, ATLAS graphs not only enable the design of 'two-terminal' gradient paths between two terminal ('end point') alloys, but also complex 'multi-terminal' gradients between an arbitrary number of terminal alloys. Multi-terminal gradient design is achieved through a novel formulation of the CGA design problem as a minimum Steiner tree problem in graphs. Additionally, the use of several other graph algorithms is also proposed for efficient, targeted experimental assessment of the alloy processing state space.

Furthermore, the Tree-based Material Adjacency Propagation (TreeMAP) algorithm is proposed to map multi-terminal CGAs into discretized 3D structures, linking material processing with spatial performance targets. This algorithm enables the connection of alloy processing states with properties and performance in corresponding discretized part geometries, fully enabling systematic materials design of CGAs for structure-driven performance constraints and objectives.

Lastly, a proposed framework integrates ATLAS, multi-terminal gradient design, and TreeMAP, establishing a first-of-its-kind approach for systematic CGA design. This work concludes by demonstrating this systematic approach by designing a compositionally graded structure using

real materials, starting with the terminal alloys, then the multi-terminal gradient to join them, and, ultimately, the compositionally graded structure itself. This demonstration shows the framework's ability to enable optimized, location-specific structural performance beyond traditional single-alloy structures. Overall, this integrated framework extends the materials design and metal AM state-of-the-art, expanding the capabilities of metallic materials for demanding engineering applications.

DEDICATION

"Accept whatever befalls you,

and in times of humiliation be patient.

For gold is tested in the fire,

and those found acceptable, in the furnace of humiliation.

Trust in him, and he will help you;

make your ways straight, and hope in him."

Sirach 2:4-6

I dedicate this work to my wife Johanna;

to my parents, Don and Molly;

and to my Lord, Christ Jesus, the way, the truth, and the life.

ACKNOWLEDGMENTS

I extend my gratitude to my advisors, Dr. Richard Malak and Dr. Raymundo Arróyave, for their unwavering mentorship and critical feedback throughout the course of this work. Their contributions to my development as a researcher have been invaluable. I would also like to thank Dr. Vinayak Krishnamurthy for his thoughtful feedback, and Dr. Ibrahim Karaman for profound insight into the details and challenges of experimental materials science research.

I am also grateful to the students and faculty within the engineering and systems design community at Texas A&M University for their collaborative input and constructive discussions. Additionally, I owe special thanks to Dr. Tanner Kirk for his pivotal guidance as I joined the Design Systems Laboratory and began my graduate research during the COVID-19 pandemic; his collaboration fundamentally shaped my trajectory.

Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Dr. Richard J Malak Jr. [advisor] and Dr. Vinayak Krishnamurthy of the J. Mike Walker '66 Department of Mechanical Engineering as well as Dr. Raymundo Arróyave [co-advisor] and Dr. Ibrahim Karaman of the Department of Materials Science and Engineering.

The equilibrium and Scheil-Gulliver solidification data used in Chapter 6, and the Hf-Mo-Nb-Ta-V-W data were provided by Brent Vela. The property regressor model used in Chapter 6 was provided by Vahid Attari. The material property prediction scripts used for Chapter 6 were provided by Brent Vela and James Hanagan.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by NSF through the project NRT-DESE: Data-Enabled Discovery and Design of Energy Materials (D3EM) under the award number NSF-DGSE-1545403 in addition to a fellowship from the J. Mike Walker '66 Department of Mechanical Engineering at Texas A&M University.

NOMENCLATURE

Abbreviations

AM	Additive Manufacturing
ATLAS	Alloy Topology-Linked informAtion Schema
BCC	Body Centered Cubic
BFS	Breadth-First Search
CAD	Computer-Aided Design
CALPHAD	CALculation of PHAse Diagrams
CGA	Compositionally Graded Alloy
CNC	Computer Numerical Control
CSV	Comma-Separated Values
CTE	Coefficient of Thermal Expansion
DFT	Density Functional Theory
FCC	Face Centered Cubic
FGM	Functionally Graded Material
GDB	Graph Database
ICME	Integrated Computational Materials Engineering
KNN	k-Nearest Neighbors
LP-DED	Laser-Powder Directed Energy Deposition
LPG	Labeled Property Graph
ML	Machine Learning
MCC	Maximally Connected Component
MMTO	Multi-Material Topology Optimization

MM-MJ TO	Multi-Material Multi-joint Topology Optimization
MOBHEA	Thermo-Calc High Entropy Alloy Mobility
MPEA	Multi-Principal Element Alloy
MST	Minimum Spanning Tree
NIMPLEX	NIM simPLEX
NoSQL	Not only SQL
PRM*	Optimal Probabilistic Roadmap
PSP	Process-Structure-Property
RAMP	Rational Approximation of Material Properties
RDBMS	Relational Database Management Systems
RDF	Resource Description Framework
RRT	Rapidly-exploring Random Tree
RRT*FN	Optimal Rapidly-exploring Random Tree Fixed Nodes
SPG	Minimum Steiner tree Problem in Graphs
SQL	Structured Query Language
STEP-NC	Standard for the Exchange of Product model data compliant Numerical Control
STL	STereoLithography
TCHEA	Thermo-Calc High Entropy Alloy
TO	Topology Optimization
TreeMAP	Tree-based Material Adjacency Propagation
t-SNE	t-Distributed Stochastic Neighbor Embedding
UMAP	Uniform Manifold Approximation and Projection

Mathematical Symbols

B	Bulk modulus
c	Cost function relating an edge or path to a quality metric
d	Dimensionality of the alloy state space
e	Edge in a graph
E	Set of all edges in a graph
G	A graph of nodes and edges
G_{A}	ATLAS graph
G_{AM}	ATLAS materials graph
G_{P}	Spatial discretized part graph
Kou′	Normalized Kou cracking susceptibility criterion
l	Euclidean length of a path
L	Set of leaf nodes in a tree
m_c	Coalescent material parameter
M_{creep}	Creep merit index
N_T	Number of terminal nodes
p_i	Arbitrary graph property key
P	Set of all graph property keys
S	Shear modulus
T	Temperature
v	Node in a graph
V_T	Set of terminal nodes
V	Set of all nodes in a graph
W	Set of all edge weights in a graph

x	Alloy composition
z	Vertical build position in an additively-manufactured structure
α	Normalized index of an alloy in a gradient path
$\mathbb{R}^n$	Real Coordinate n-space
σ	Continuous 1D path in the alloy state space
σ_{min}	Feasible 1D path that minimizes the cost function
σ_{y}	Yield strength
τ	Graph tree
τ_{CGA}	Compositionally graded alloy graph tree
τ_{min}	Steiner minimal tree

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

1. INTRODUCTION

1.1 Research Overview

With performance demands continually increasing for metallic structures, multi-material solutions, such as compositionally graded alloys (CGAs) have the potential to meet localized performance objectives that are beyond the capabilities of any single alloy composition. As a result, material design efforts must also account for these multi-material solutions. Metal AM techniques in particular have drawn increased attention to CGAs due to metal AM's capabilities for controlling spatial variation of material composition throughout a single build.

Despite their performance appeal, there are significant research gaps limiting the design and implementation of CGAs. Particularly, existing materials design methods are ill-suited for designing CGAs specialized for arbitrarily many spatial performance optimization criteria. Additionally, mapping such gradients into corresponding complex physical structures further complicates the problem. The intertwined challenges of these dual research gaps are the primary focus of this work.

Realizing the potential of multi-material alloy design first requires advanced material design strategies that can navigate the complexities of the alloy design landscape. The most fundamental design challenge is avoiding the formation of deleterious brittle phases along a compositional gradient [9, 10]. Perhaps the most transformative effort for accelerated materials design in this era is Integrated Computational Materials Engineering (ICME), an approach to systematically integrate materials data in computational tools alongside engineering product performance analysis and manufacturing-process simulation [11]. The ultimate goal of this data integration is systematic materials design [3]. To achieve this goal, one must first scientifically deduce the relationships between the three-link chain of processing, structure, properties, and performance. Then, engineering design, which is driven by performance, can be accomplished by mapping targeted structural performance criteria back to the necessary processing. The application of this systematic materials

design paradigm to CGAs is critical for the effective design and implementation of these materials.

In addition to the materials design challenges with CGAs, there are also significant challenges with structural design using these complex materials. Existing multi-material topology optimization approaches lack a variety of considerations necessary for realistically implementing CGAs in structural designs. These shortcomings include limited or nonexistent interfacial considerations, inaccurate property models, incompatible assumptions, and overly simplistic material mixing constraints.

In light of these research gaps, this dissertation presents a novel framework for the systematic design of CGAs and related CGA structures by linking processing to performance for these advanced, complex materials. First, this work proposes the Alloy Topology-Linked informAtion Schema (ATLAS), a graph-based information schema for materials design of compositionally graded alloys. The ATLAS integrates the separate benefits of prior materials design methods while also providing novel capabilities for data analysis and accelerated alloy development. Using graph path-finding algorithms with a graph database structured according to the ATLAS, one can design alloy gradient paths between two terminal alloys that avoid the issue of deleterious phase formation. Furthermore, a novel formulation of the CGA design problem, as a minimum Steiner tree problem in graphs, enables the design of multi-terminal alloy gradients that avoid the same issues, extending beyond the two-terminal gradients given by prior methods. This minimum Steiner tree problem formulation is presented and demonstrated using real materials settings. To extend systematic CGA design to the performance of compositionally graded structures, this work also proposes a Tree-based Material Adjacency Propagation (TreeMAP) algorithm to map any multi-terminal gradient, represented as a Steiner tree, into a fixed, discretized physical structure. This algorithm is described and demonstrated with several synthetic and real alloy systems and corresponding structures. Furthermore, key considerations for this algorithm are discussed and illustrated. Lastly, the capabilities of a combined framework composed of the ATLAS, Steiner tree gradient design formulation, and TreeMAP algorithm are illustrated by designing three terminal alloys, a corresponding multi-terminal CGA, and a compositionally graded structure using realistic

materials data. It is shown that the final structure possesses locally optimized material performance beyond the capabilities of any single composition in the same alloy system. Overall, this framework enables the systematic design of compositionally graded alloys and structures that expand the capabilities of metal AM and materials design while also providing noteworthy insights into the alloy processing state space for the broader field of alloy design.

This first chapter begins with a high-level discussion of the motivation for this work. Then, a review of compositionally graded alloys and relevant manufacturing methods is presented. Afterward, the chapter transitions to an introduction to materials design and outlines the significant materials design challenges when developing CGAs, detailing current methods and their limitations. Afterward, an overview of design methods for multi-material structures is given, noting that these design methods lack critical considerations for designing with CGAs. Lastly, the chapter concludes with a description of the research objectives and contributions of this dissertation.

1.2 Motivation

Two fundamental gaps in the systematic design of CGAs and CGA structures motivated this work. The entangled relationship between these research gaps further complicates each other. The first gap is identifying CGAs with more than two terminal ('end point') alloy compositions. The two predominant computational materials design approaches for CGAs were developed to identify CGAs with two terminal compositions. These approaches were first presented in [9] and [12] with recent notable refinements in [13] and [14, 15], respectively. Note that a two-terminal gradient path can be called one-dimensional since a path can be represented as a continuous function $\sigma : [0, 1] \rightarrow (\mathbb{R}^n)$; the composition along a path is described by a single path index variable $\alpha \in [0, 1]$ that scales with Euclidean distance l along the path in $(\mathbb{R}^n)$ from one terminal alloy ($\alpha = 0$) to another ($\alpha = 1$). This is illustrated in Fig. 1.1a for a two-terminal gradient between terminal alloys x_1 and x_2. However, structures may have more than two spatial performance criteria for materials design. This concept is illustrated for a stylistic gas turbine blade in Fig. 1.2. Designing terminal compositions for fewer simultaneous performance criteria enables further material specialization, reducing the limiting effect of property trade-offs that motivate multi-material design in the first

place. Thus, the gas turbine blade example shown in Fig. 1.2 illustrates that compositionally graded structures may benefit from the use of many unique terminal alloy compositions specialized for each (or at least fewer) of the respective spatial performance objectives in the structure.

The second gap is the placement of such multi-terminal CGAs in corresponding physical structures. In manufacturing, CGAs have been demonstrated with uniaxial variation in composition in Cartesian [9], cylindrical [16, 17], or spherical coordinates [17]. These demonstrations are notable, but while acceptable for parts using only two terminal alloys with a unidirectional gradient interface and simple geometric features, the material placement becomes more complex when three or more terminal alloys are considered simultaneously, and complex geometry is introduced. An illustration is shown in Fig. 1.1 comparing these challenging design conditions to current capabilities. Figure 1.1a shows a one-dimensional compositionally graded alloy, typically represented as a path between two terminal compositions. This path must fundamentally avoid compositions expected to form deleterious brittle phases in order to be manufacturable. Translating this gradient path into physical space involves correlating the compositions along the CGA path with a physical coordinate, such as the build height z in metal AM. Figure 1.1b shows how this simple mapping can be applied to arbitrary structures, given the placement of the terminal alloy compositions.

To illustrate the compounding complexity of these dual research gaps, Fig. 1.1c shows that the existing approach of identifying and mapping a 1D path breaks down when an additional terminal alloy composition is added, indicating the non-trivial challenge of both representing a compositional gradient between more than two terminal alloy compositions, as well as mapping such a gradient onto a structure in physical space. If piecewise two-terminal paths are used to connect each unique pair of terminal alloys, the challenge of identifying at what composition(s) they should connect and where the paths meet these compositions in the physical structure is nontrivial. Thus, these gaps are not addressed with current methods for either gradient material design or spatial material placement, inhibiting the application of CGAs to complex geometries with more than two competing property objectives.

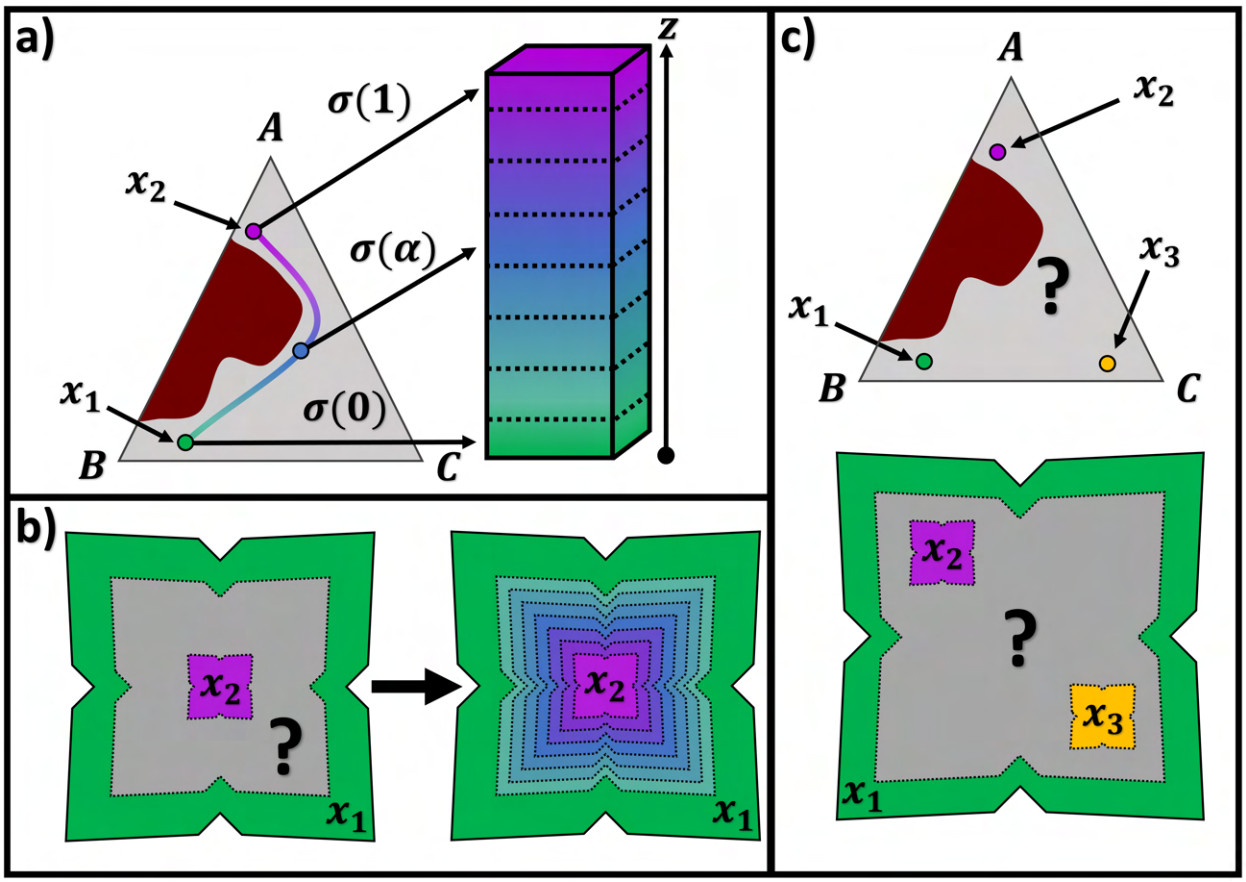

Figure 1.1: An illustration showing a) the design of a one-dimensional compositionally graded alloy path, σ, between compositions x_1 and x_2 in a synthetic ternary alloy system consisting of elements **A**, **B**, and **C**, with the maroon region indicating deleterious compositions. This compositional gradient can be easily mapped into physical space by correlating compositions along the gradient, $\sigma(\alpha), \alpha \in [0, 1]$, to a corresponding physical coordinate, such as build height z. Similarly, b) demonstrates how a one-dimensional gradient can be applied to a structure with arbitrary geometry by segmenting it into encompassing shells or layers and performing a similar mapping. Finally, c) illustrates the non-trivial extension of compositional grading to multiple compositions in both materials design and physical mapping, exemplified by the three terminal alloys x_1, x_2, and x_3.

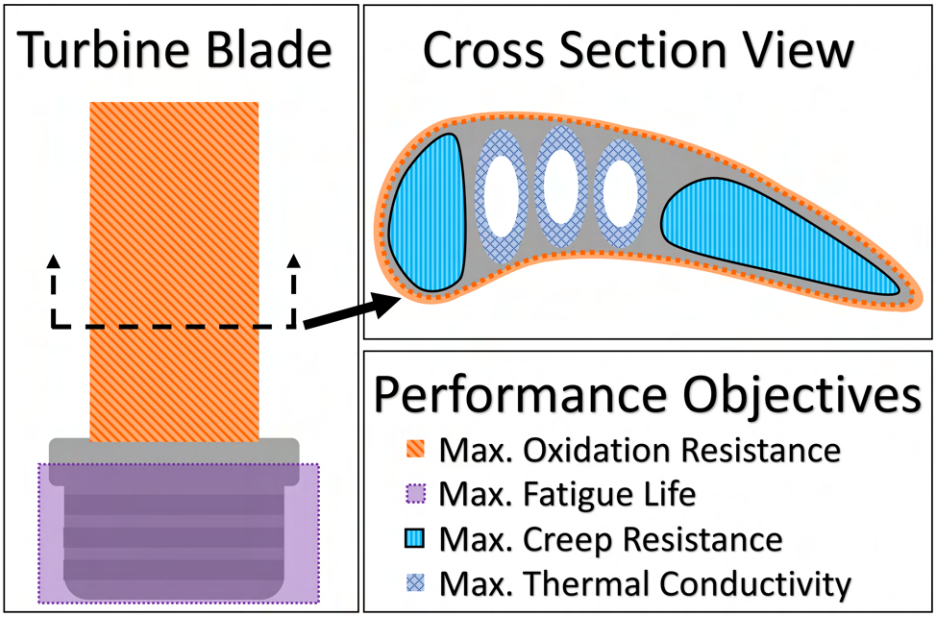

Figure 1.2: Structural design problem for a stylistic gas turbine blade with four distinct spatial material performance objectives including maximal oxidation resistance on the high-temperature oxidation susceptible blade surface, maximal fatigue life in the fir tree root which experiences the peak stress, maximal creep resistance at the interior of the blade (particularly towards the high-temperature regions on the leading and trailing edge), and maximal thermal conductivity in proximity to internal cooling channels for improved heat transfer. Adapted from the prior work in [1].

1.3 Compositionally Graded Alloys

Traditional approaches to alloy design are fundamentally restricted by the trade-offs between properties, which can result in alloy design campaigns where the problem is over-constrained and no satisfactory alloy composition exists. For example, Vela et al. filtered 53,124 candidate alloys in the W-Mo-V-Ta-Nb-Al refractory multi-principal-element alloy (MPEA) composition space for a variety of design constraints ultimately resulting in a mere 0.54% of compositions passing [18]. Thus, as the performance demands for new alloys increase, it is evident that there will be an increasingly limited number of alloys that possess all desired characteristics simultaneously, if any at all.

Due to these increasing performance requirements, there is growing interest in CGAs, a type of functionally graded material (FGM), where the alloy composition is varied spatially to meet competing performance requirements in different regions of a single part [9]. Recently, there have been significant research efforts exploring design paradigms where alloy composition itself is a design variable to be optimized alongside the part geometry for holistically optimized structural performance, such as DARPA's Multiobjective Engineering and Testing of Alloy Structures (METALS) initiative [19]. Since the property requirements for alloys often stem from localized performance demands in engineering structures, it is hypothesized that, by locally tailoring properties, part designs that leverage CGAs can exceed the performance of traditional single-material designs. Furthermore, CGAs enable additional capabilities beyond typical alloys, such as embedding information in metal structures with selective placement of magnetic compositions [20]. Spatial distribution of alloy compositions can also be leveraged to tailor the placement of materials with significant considerations concerning material cost or criticality [21, 22].

While there are significant motivations for manufacturing CGAs, there are several noteworthy challenges in the design and manufacturing of these materials. The primary challenge for CGAs is avoiding the formation of deleterious brittle phases at intermediate compositions [23, 24, 25, 26, 27, 10]. Sharp discontinuities in material properties such as the coefficient of thermal expansion (CTE), which describes material expansion/contraction with temperature changes, can also

be problematic at interfaces between dissimilar alloys [28]. These discontinuities ultimately result in stress concentrations, which can cause crack initiation, especially if brittle intermetallic phases are present. Despite significant efforts to address the manufacturing challenges in recent years, these materials are still in their infancy. When successful CGAs are identified, few demonstrations extend beyond manufacturing trivial prismatic shapes.

1.3.1 Manufacturing Methods

1.3.1.1 Dissimilar Welding

Historically, the primary approach for implementing multiple dissimilar alloys in a monolithic structure was through dissimilar welding, often using discrete intermediate alloy compositions with the two dissimilar alloys to be joined compared to their compatability with one another [29]. For example, Sun and Ion [28] note that silver and chromium have poor laser weldability. However, titanium and chromium have good laser weldability, and titanium and silver have fair weldability. Thus, titanium could be used as an interlayer to improve the material compatibility from poor to good/fair across the joint.

While this approach has historically been used to join dissimilar alloys, dissimilar joints produced with this approach do not fully constitute CGAs, as they are typically comprised of three discrete alloy compositions. This characteristic also makes these joints more prone to stress concentrations due to the sharp spatial variation in alloy composition and properties. These significant discrete compositional changes indicate that these joints lack the continuous gradual change in composition and properties typical to CGAs. By comparison, metal AM, discussed in the following subsection, offers far greater spatial compositional control with gradual variation.

1.3.1.2 Additive Manufacturing

Metal additive manufacturing (AM) is an approach to manufacturing where three-dimensional (3D) objects with near-net shapes are manufactured, typically by joining material layer-by-layer, directly from 3D geometric models. The general process of metal AM typically involves creating a computer-aided design (CAD) file of a 3D geometric object, which is then converted into a

STereoLithography (STL) file to represent the surfaces of the part. Then, the STL file is input into a slicing software, which slices the 3D model into thin layers and generates the tool path instructions for the AM machine to build the part layer-by-layer. The machine is given these build instructions and the relevant feedstock to manufacture the part, which is then post-processed, inspected, and tested as needed. Metal AM can be used to achieve a variety of benefits, including light-weighting, assembly simplification, cost & lead time reduction, distributed manufacturing, novel materials, and the capability to manufacture complex geometries [30]. As a result, significant research efforts across many disciplines are needed to advance and implement metal AM further.

Most metallic AM systems fit into three general categories: powder bed, powder feed, and wire-fed systems [31]. Powder bed systems operate by spreading metal powder across a powder bed in a controlled environmental chamber using a rake/roller; once each part slice is sintered in the new top layer of powder using a laser or electron beam, the bed is lowered, and a new layer is spread as the process repeats. Meanwhile, powder feed systems direct powders through one or more nozzles onto the build surface into a focused laser or electron beam. Each layer is created by moving the build surface or deposition head as the powder is sintered. In contrast, powder bed methods can produce fine-resolution features compared to other methods. These powder feed systems typically have larger build volumes than powder bed systems and can be used to add or repair features on existing parts. Lastly, wire feed systems deposit wire onto a substrate using either an electron beam, laser beam, or plasma arc. These wire-fed systems are typically used for the largest build volumes and highest deposition rates but poor surface finish. It is worth noting that there are also hybrid approaches [32, 33, 34] which merge elements of these categories, as well as novel techniques such as drop-on-demand molten metal jetting [35, 36], which have not yet been implemented broadly.

Metal AM techniques are particularly well-suited for manufacturing CGAs due to the capability to customize the spatial variation in alloy composition and resultant properties in a single build. For example, Laser-Powder Directed Energy Deposition (LP-DED), a type of powder feed metal AM system, has accelerated the development of CGAs due to its capability to alloy in situ by

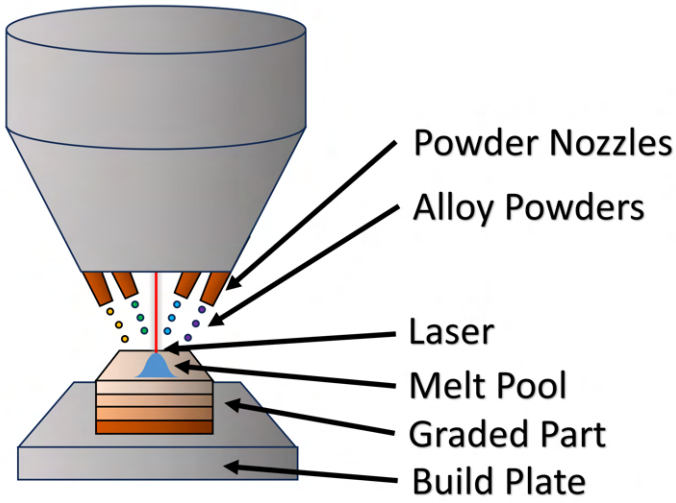

Figure 1.3: A schematic of the Laser-Powder Directed Energy Deposition (LP-DED) process being used to manufacture a compositionally graded alloy part by varying the ratio of metal alloy powders deposited from different nozzles into the melt pool on a layer-by-layer basis. Reprinted with permission from Allen at al. (2024) [2].

varying the amounts of various metal powders sprayed by different nozzles into the melt pool [37]. With this capability, LP-DED can change the alloy composition on a layer-by-layer basis to create CGAs [38, 39, 40, 41]. An example of this process for LP-DED is illustrated in Fig. 1.3. There is also growing research demonstrating the manufacturing of CGAs using powder bed [17, 42, 43], wire-fed [44, 45, 46], and hybrid powder-wire [32, 33, 34] metal AM systems.

From an experimental perspective, the metal AM of CGAs is also useful for characterizing the alloy state space, given by all possible alloy compositions and processing conditions. By manufacturing and characterizing CGAs, instead of single alloy compositions at a time, different regions of the alloy state space can be efficiently assessed with less material and fewer individual builds [47, 48, 42]. This approach to experimental assessment and exploration could be further improved when used in conjunction with autonomous experimentation [49], to continually accelerate the rate of development for novel alloys. Ultimately, the development of CGAs strengthens the value proposition of metal AM, enabling simplified supply chains with an emphasis on domestic manufacturing and jobs; CGAs unlock a wider variety of use cases and enhanced performance capabilities for additively manufactured metallic structures.

1.4 Materials Design

As the rate of material advancement continues to accelerate with the improvement of computational material science methodologies, largely driven by machine learning (ML), artificial intelligence (AI), and the increasing amount of materials data available, the field of materials science is poised to fuel the so-called Fourth Industrial Revolution, also called Industry 4.0. Since the Stone Age, humanity's technological capabilities have been defined by enabling materials. In the present day, modern computational methods like ICME are rapidly decreasing the timescale for these transformative material breakthroughs compared to the rest of history.

The modern ICME effort to integrate materials data across processing, manufacturing, and product performance analysis is an outgrowth of the systems approach to computational materials design first articulated with the seminal 1997 publication by Greg Olson [3]. In this work, Olson described the linear linkages between the four primary elements in materials science and engineering, namely 1) processing, 2) structure, 3) properties, and 4) performance, which form a three-link chain that is crucial for conducting systematic materials design. In this chain, the scientific, deductive cause-and-effect logic flows from processing through to the resultant material performance, while the targeted, inductive goal-means perspective of engineering flows in the opposite direction. This chain and these opposing flows are illustrated in Fig. 1.4. Materials design can proceed inversely by capturing the necessary information to deduce cause-and-effect modalities across this chain in computational tools, inducing the necessary processing for targeted properties and performance. The implementation of this systems view of materials design in research has led to countless advancements across various types of materials [50, 51, 52, 53, 54].

Due to the significant capabilities of metal AM and its growing implementation in research and industry, there is a need to integrate materials data into a holistic informatics approach that can be linked with engineering product performance analysis and manufacturing-process simulation for ICME in metal AM for single and multi-material development and implementation. Several computational methodologies have been proposed for the design of CGAs. However, these methods have notable limitations. Furthermore, due to the materials design challenges and very recent

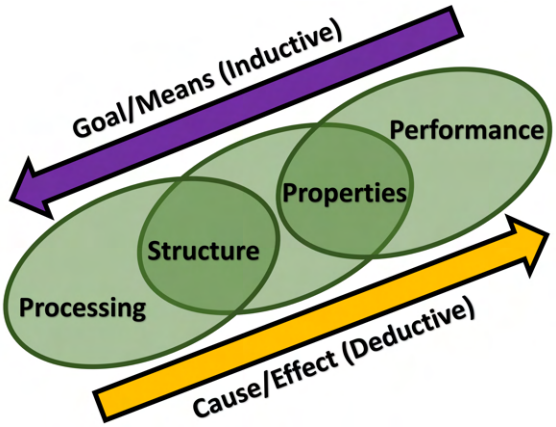

Figure 1.4: A schematic, adapted from work by Olson [3], of the Process-Structure-Property-Performance three-link chain model which undergirds systematic materials design, illustrating the deductive flow of scientific logic and inductive flow of engineering design in opposite directions.

development of materials design methodologies, existing methods for designing multi-material structures are inadequate for joining dissimilar alloys. This section will a) outline the existing methods that link processing, structure, & properties for computational materials design of CGAs and their limitations, and b) discuss current methods that link properties to performance for the design of multi-material structures and their limitations.

1.4.1 Multi-Materials Design Methods & Limitations

Hofmann et al. first proposed using ternary phase diagrams for CGA design to design CGAs to avoid the formation of deleterious phases. These ternary phase diagrams can be modeled using the CALculation of PHAse Diagrams (CALPHAD) method and visualized in 2D [9]. The resultant phase diagrams can be used as maps to visually discern if a feasible two-terminal gradient (a path between two terminal 'endpoint' alloys) exists and, if so, to manually design a manufacturable CGA path between these alloys. This approach has been further augmented with Scheil-Gulliver simulations [55, 56, 57, 58] and cracking susceptibility metrics to address the additional concern of solidification cracking [59]. This augmented methodology has recently been released in the open-source software MaterialsMap [13]. While this approach could be used for manually planning gradients with arbitrarily many terminal alloy compositions or 'many-terminal gradients', this has

not yet been demonstrated in the literature.

Meanwhile, Kirk et al. introduced a computational methodology [12, 60] based on a robotic motion planning algorithm, RRT*FN [61, 62]. This methodology used the CALPHAD method to perform equilibrium thermodynamic calculations to train a machine learning classifier as the obstacle model for RRT*FN to automatically optimize a two-terminal gradient in composition for a given objective function while avoiding the formation of deleterious phases, with several experimental demonstrations [63, 64]. This approach was later augmented with a sampling procedure that considers all subspaces of the alloy design space in the gradient design process [14]. Most recently, on-the-fly CALPHAD sampling was proposed to reduce the overall computational cost of this approach [15]. Overall, these methods have been successfully used to identify two-terminal, one-dimensional compositional gradients through multidimensional alloy spaces to join an arbitrary pair of terminal alloys.

While these methods enable materials design of CGAs, they have noteworthy limitations. For example, while terminal alloys could be selected based on properties and compatibility visualized on ternaries, the ternary map approach does not consider alloys with more than three active elements at a time. This limits the overall design space, potentially isolating desirable alloys that could be joined using additional elements. Manual selection of gradient compositions in visualized ternaries is also suboptimal compared to automated optimization methods such as path planning. Furthermore, when applying robotic path planning to this problem, the designer lacks insight into the feasible design space, as the algorithm automatically performs its search in high-dimensional space. This can make it difficult to determine whether a manufacturable path exists between terminal alloys without running the RRT*FN design algorithm. Additionally, if a path is not generated after numerous iterations, it remains unclear whether a viable path is absent or if further iterations are required. Moreover, these approaches assume that a path linking two terminal alloys is sufficient for achieving the desired properties, limiting the degrees of freedom for varying alloy composition to control properties.

Notably, recent efforts have reduced the computational cost of RRT-based gradient path plan-

ning using on-the-fly CALPHAD calculations [15]. While this approach is efficient in its use of CALPHAD sampling, particularly in high-order systems, the method shares similar issues as the prior work using RRT*FN. First, the resultant gradient path is tailored for one specific gradient design problem. Second, trial and error must be used to identify compatible terminal alloys in a given state space. Third, the user still lacks insight into the high-dimensional feasible design space. Lastly, this framework also has the inherent limitation of RRT planning algorithms: the inability to determine whether a solution exists when the algorithm fails to identify one.

In addition to these limitations, the materials design process for CGAs often begins with a material bias, intending to identify intermediate compositions to form a defect-free joint between two specific terminal alloys. However, this can unnecessarily constrain the problem if the design intent is actually to achieve a part with specific performance, regardless of material composition. The author acknowledges that there are cases in which designers seek to join specific alloys due to existing material availability, standards, and longstanding use in specific applications. Nevertheless, this processing-biased approach assumes that the properties of a defect-free gradient between these terminal alloys will satisfy the design requirements, if a defect-free gradient even exists in the first place. Removing this constraint enables materials design to align with a systems approach, mapping desired performance back to the required processing steps [3].

1.4.2 Structural Design Methods & Limitations

Due to challenges in designing and manufacturing CGAs [10], there has been little research on methods for designing parts that specifically use these advanced materials. Existing multi-material design approaches have a variety of fundamental flaws that hinder their application to the design of compositionally graded alloy structures.

Topology optimization and generative design techniques are often used with AM due to their capability to generate complex, optimized structures that are difficult and/or expensive to manufacture with traditional methods. Multi-material topology optimization (MMTO) is an area of growing research that seeks to extend the capabilities of topology optimization (TO) by adding material composition as a design variable. With this additional degree of freedom, MMTO algo-

rithms can generate designs with heightened performance compared to their single material TO counterparts [65, 4]. Designs generated by MMTO are becoming increasingly relevant due to the design freedom afforded both geometrically and compositionally by the growing adoption of multi-material-capable AM machines. However, current MMTO methods lack joining considerations when designing with dissimilar alloys, accounting for potential formation of deleterious phases and the need for compositional grading. For example, Fig. 1.5 highlights the significant amounts CALPHAD-predicted deleterious phases if using simple linear one-dimensional gradients between common alloys seen in MMTO literature [4] such as Ti-6Al-4V, aluminum alloys, and 316 stainless steel. Even if no compositional gradient is used, diffusion between these alloys during high-temperature melting and reheating during AM will result in the blending of these compositions and corresponding phase formation at the interface. Issues with these gradients are also well-documented in experimental literature [10].

Some recent work has sought to rectify this gap by imposing a joining interlayer between dissimilar alloy regions in so-called multi-material multi-joint topology optimization or MM-MJ TO [66, 67], but this oversimplifies the joining problem in dissimilar metals with respect to both properties and geometry. Regarding properties, a discrete interface between dissimilar alloys is likely to produce brittle intermetallics [10]. Dissimilar metal welding [29] can be used to improve an interface between different alloys, but dissimilar welds are not truly CGAs. Additionally, a gradient will yield nontrivial property changes across the joint interface, which can violate the property assumptions of these MM-MJ TO methods. Secondly, with respect to geometry, more physical space in the part may be required than the single layer dedicated to the joint in MM-MJ TO, depending on the number of gradient interlayers needed to preserve manufacturable phases and feasible properties, including the management of thermal stresses from adjacent dissimilar alloys which typically have different values for the coefficient of thermal expansion (CTE).

Elsewhere in the literature, there are FGM-specific TO methods, but these efforts are primarily concerned with functionally graded composites and metamaterials [68, 69, 70, 71, 72]. The properties of these materials, many of which follow the general rule of mixtures, are typically simple

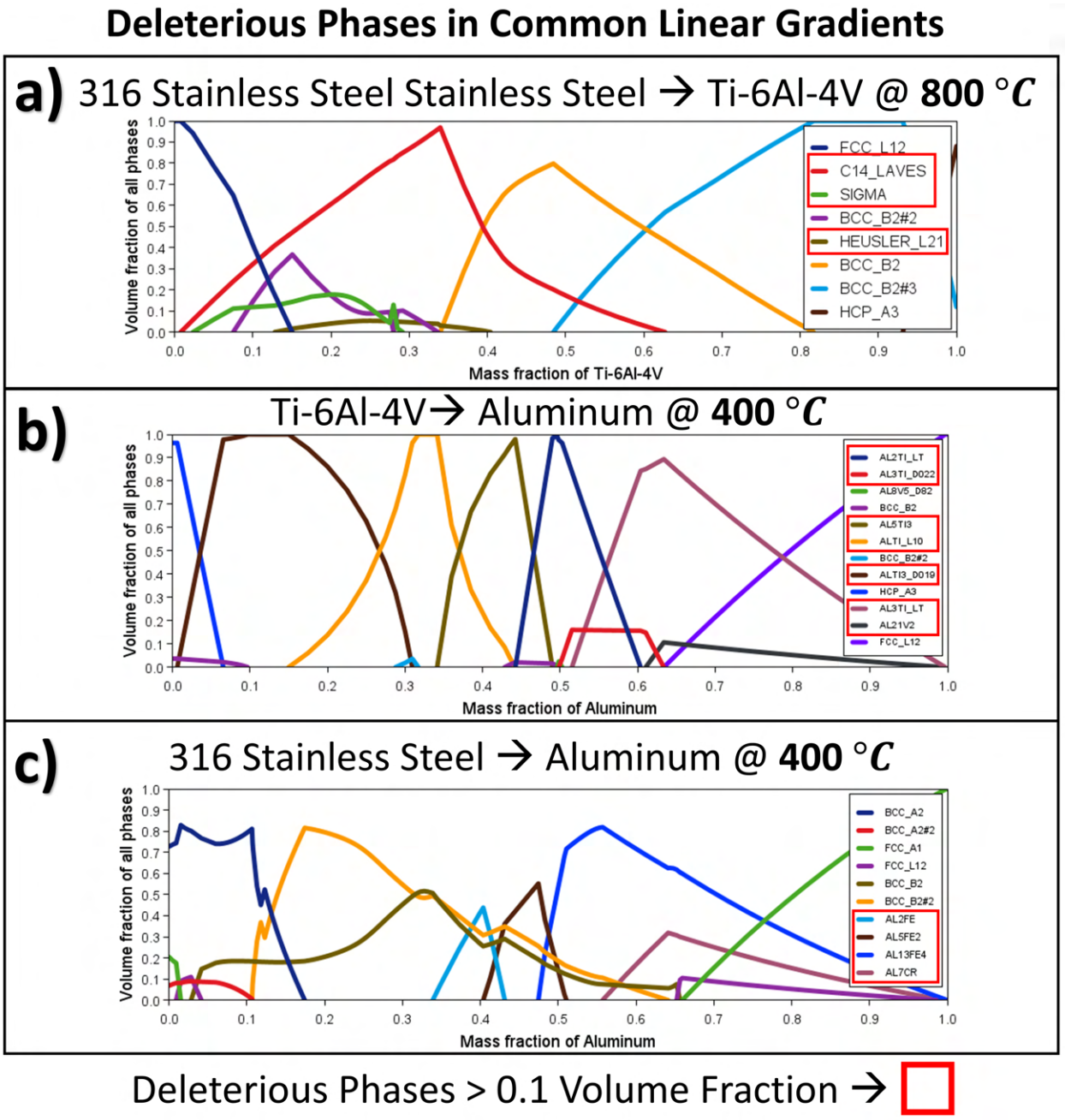

Figure 1.5: Illustrations showing significant amounts of deleterious phases predicted to form for linear pairwise gradients between common alloys used in MMTO literature [4], namely: a) 316 stainless steel & Ti-6Al-4V, b) Ti-6Al-4V & pure Al, and c) 316 stainless steel & pure Al. These predictions were made using material-to-material equilibrium phase simulations in Thermo-Calc [5] version 2023b using the TCHEA database version 6 [6].

by comparison to CGAs. For example, Silva et al. recently used an interpolation scheme using the Rational Approximation of Material Properties (RAMP) method with Hashin-Shtrikman bounds [73] under the assumptions of material isotropy and homogeneity [72]. This approach is not generally applicable for CGAs, which can have nontrivial changes in microstructure including phases and grain morphology across the material gradient. These changes have a corresponding nontrivial impact on the resultant properties [47, 48, 42]. Silva et al. acknowledge this in their discussion regarding their material model noting that the assumptions used are better suited for polymers.

Additionally, functionally graded composites and metamaterials can typically be blended without significant restrictions in comparison to CGAs. Functionally graded composites involve heterogeneous mixtures between two classes of materials, such as metals and ceramics or polymers and ceramics, which can be blended in any proportion without affecting the microstructure of any of the constituent materials. In the case of metamaterials, typically, one chemical material composition is used; compatibility relates to the physical arrangement and geometry of the structural unit cells. Metamaterial blending can be constrained by enforcing a lower bound to ensure a valid structure [71]. However, for CGAs, one must consider the nontrivial constraints due to deleterious phase formation and cracking susceptibility at specific ratios of alloying elements. Thus, existing FGM-specific TO methods have a variety of characteristics, including property calculation and mixing constraints, which are ill-suited for design with CGAs.

1.5 Contributions

In light of the current research gaps in literature, the primary contribution of this research is *an automated, reusable, dimensionally scalable, and topologically interpretable framework linking processing, structure, processing, and performance for the systematic design of compositionally graded alloys and structures.* Existing methods for materials design of compositionally graded alloys are limited in their trade-off between either the reusable and topologically interpretable method of ternary phase diagram maps [9, 58, 13] or the automated and scalable method of motion planning algorithms [12, 60, 64, 14]. In this work, an approach is considered *topologically interpretable* if the alloy design space can be decomposed into distinct regions or 'islands' that are

separate from one another. This is true of the ternary phase diagram maps because any separate regions of feasibility can be observed visually and interpreted by the user in 2D using the ternary diagrams. While both of these prior methods have been shown to successfully connect processing, structure, and properties for designing manufacturable gradients free of deleterious phases, there is a need for a dimensionally scalable approach where both the design space is topologically interpretable and the gradient design process can be automated. Meanwhile, current structural design methods, which can be used to link properties to performance, lack the necessary capabilities or processing/structure considerations to design realistic, manufacturable CGAs. Thus, there is also a broader need to connect the full process-structure-property-performance chain for true systematic design of compositionally graded structures. Addressing materials design and structural placement of CGAs simultaneously also yields the opportunity to solve the intertwined challenges of designing multi-terminal CGAs and corresponding structures.

1.5.1 Scalable, Reusable, & Topologically Interpretable Data Schema

In light of prior materials design efforts [9, 58, 13, 12, 60, 64, 14], it is evident that this data representation should be scalable, reusable, & topologically interpretable to advance CGA design efforts. The specific need for reusability indicates a formal data schema that enables effective storage and repeated querying of alloy datasets. Common methods for storage and querying interconnected datasets include relational database management systems (RDBMS), and, more recently, graph databases (GDBs) [74].

In this work, it will be shown that GDBs possess the necessary characteristics to support a scalable, reusable, & topologically interpretable framework compared to RDBMS. Various types of GDBs will be compared, highlighting the relevant benefits of the labeled property graph (LPG) representation over other methods. This LPG representation uses the classical graph model [75] with the additional capability to store node- and edge-wise metadata. A novel LPG schema, which captures relevant topological information by linking nodal alloy compositions with edges, will be presented. Demonstrations will illustrate that this *scalable & reusable* data representation enables *topologically interpretable* materials data analysis for alloys of arbitrary complexity, and topolog-

ical partitioning of the alloy design space with benefits beyond compositionally graded alloys.

1.5.2 Systematic Design of Multi-terminal Compositionally Graded Alloys

While representing the alloy design space as a graph requires more upfront data structuring compared to methods such as robotic path planning with RRT*FN [12, 60, 64, 14], once a graph is constructed, its structure can be used for extremely efficient data traversal. A variety of algorithms can be used to automatically identify the minimum weight path in graphs as a one-dimensional gradient path between two terminal alloy compositions. To put this work into context, several relevant methods for two-terminal path finding are identified.

Ultimately, this work extends beyond two-terminal gradient paths to multi-terminal gradients using graph trees. To accomplish this for materials design, a formulation of the gradient design problem as a minimal Steiner tree problem in graphs [76] is presented, for which a solution can be approximated using heuristic methods [77, 78]. This work demonstrates how this Steiner tree formulation can be used to automatically design multi-terminal CGAs between any number of terminal alloy compositions. Furthermore, for materials exploration and experimentation, it is shown that maximally spanning trees can be used similarly to design a CGA between every composition in a connected graph or subgraph for multi-material experimental assessment.

1.5.3 Conformal Placement of Compositionally Graded Alloys in Engineering Structures

For a complete materials design paradigm, CGAs must be linked with realistic physical engineering structures with specific performance in mind. To this end, a graph algorithm is proposed that can map an arbitrarily complex multi-terminal CGA to a discretized 3D geometric model, where the discrete locations of the terminal alloys are initially labeled. First, the necessary inputs and outputs are identified and discussed. Then, the algorithm is presented, and considerations for a valid input are outlined. Next, a synthetic 2D example is used as a basic illustration of the algorithm in use, and key characteristics and assumptions of the algorithm are discussed. Then, a practical example is used to demonstrate the algorithm in use by designing a real engineering structure based on realistic materials data. Lastly, the scalability of the algorithm is discussed.

1.5.4 Design of Compositionally Graded Alloy Structures

With a formal information schema linking processing, structure, and properties for CGAs, a formulation for multi-terminal CGA design, and a method for mapping arbitrarily complex CGAs to a corresponding discretized structure/part, the groundwork is in place for systematic design of compositionally graded alloy structures. Thus, this work will lastly demonstrate the combined capabilities of the presented information schema, multi-terminal CGA design approach, and conformal gradient placement algorithm. This unified framework will be used for a full systematic design demonstration using real materials to design a complex, realistic compositionally graded structure. A complex discretized 3D geometric model with targeted spatial performance objectives and constraints will be defined. These performance criteria will be used to optimize the design of three novel terminal alloys. Then, a multi-terminal CGA will be designed to join these three terminals. This gradient will be subsequently mapped onto the 3D discretized physical part geometry. Lastly, the resultant CGA and corresponding part will be shown to avoid deleterious phases and meet the initial spatial performance criteria which would not be possible using a single material.

Lastly, the use of cracking susceptibility indicator metrics has been shown to be a relevant consideration for CGAs [59]. However, there is no specific constraint on these metrics that has been shown to categorically prevent cracking. Thus, to better incorporate manufacturability considerations, a novel cost function for minimizing the maximum constraint violation is proposed for minimizing the worst-case cracking susceptibility across a gradient path without imposing an arbitrary hard cutoff. This cost function is implemented and demonstrated for a realistic materials dataset as part of the aforementioned design demonstration. The final design will be shown to trade off between the worst-case crack susceptibility at the expense of larger change in composition across the gradient tree compared to the typical case of minimizing the change in composition.

2. MATERIALS INFORMATION SCHEMA FOR METAL ADDITIVE MANUFACTURING*

This chapter details the proposed information schema that serves as the foundation of this work. This schema was created to format alloy materials data to be reusable, topologically interpretable, scalable to high dimensions, and compatible with automated motion planning/data traversal algorithms. To establish the rationale behind the proposed schema, the first section enumerates the requirements. Then, potential concepts are introduced and evaluated with respect to these requirements. Afterward, a solution is selected, and a formal information schema is proposed to meet the initial requirements. Lastly, the implementation of this schema is discussed, and its use with several real materials datasets is demonstrated.

2.1 Requirements Analysis

To address the primary challenge of deleterious phase formation in CGA design, Hoffman et al. [9] originally proposed ternary phase diagrams to be used as visual maps to identify and avoid compositions expected to form these undesirable phases. These ternary phase diagrams can be modeled using the CALculation of PHAse Diagrams (CALPHAD) method [79], representing the predicted phases for all combinations of three components, either pure elements or alloys, at a fixed temperature. The CALPHAD method uses thermodynamic principles to predict the phase fractions in a material given models of phase free energies and a thermodynamic state defined by composition, temperature, and pressure. Typically, atmospheric pressure is used, though there are growing efforts to leverage AM in the vacuum of space [80, 81], and electron beam AM methods require the use of a vacuum [82].

Since this initial work by Hoffman et al., two primary methods for CGA design have been developed, also based on CALPHAD-predicted phase information. The first method is a more direct continuation of the prior work, and visually uses ternary phase diagrams to designing CGAs by

*Contains material reprinted with permission from M. Allen, R. Arróyave, and R. Malak, "A Graph Database Schema for Metal Additive Informatics," *Volume 2B: 44th Computers and Information in Engineering Conference (CIE)*, Aug. 2024. [2].

visualizing the feasible and infeasible regions of the alloy state space [55, 56, 58, 59]. The second method uses robotic path planning algorithms with an obstacle model trained on CALPHAD data [12, 63, 60, 64, 14, 83]. The obstacle model is used to classify if an unseen composition is in the feasible or infeasible regions of the alloy state space. These methods relate processing, structure, and properties across the continuum of the alloy state space by identifying the phase *structure* expected to form during the *processing* conditions of metal AM, and constrained the design space based on the expected presence of brittle phases and other undesirable *properties*.

These methods have several benefits which should be included in the requirements for a novel AM information schema. First, the visual method gives designers an intuition regarding the topology of the feasible design space, and the ternary diagrams function as reusable maps for grading other pairs of terminal alloy compositions in the same state space. Meanwhile, the path planning method is scalable to any number of dimensions, automatically identifying gradient paths, and optimizing the gradient path for a user-prescribed objective function.

These methods also have several limitations that a novel AM information schema should avoid. With the visual ternary diagram method, the paths are manually determined, which limits the ability of this method to tailor gradient composition to quantifiable design objectives. This can be addressed by enabling automated path planning/traversal of the alloy state space. Meanwhile, the path planning method only provides an output applicable to the given problem definition and must be re-run for each case when the constraints, cost function, and/or terminal alloys change. Ideally, an information schema that is easily reusable for different design conditions would allow more efficient allocation of computational and human resources for design. The path-planning method also does not provide information regarding the topology of the design space beyond three dimensions. It also does not conclusively determine whether terminal alloys are compatible, since an incomplete path result could indicate that more iterations are needed to find a continuous path. To put it more broadly, the path planning method does not explicitly identify which regions in the design space can be joined in a CGA.

While it is desirable to use computational methods such as CALPHAD when linking process-

ing, structure, and properties in metal AM, it is practical to implement data from real manufacturing experiments to improve design fidelity. While an information schema for CGA design must be capable of representing fixed, standardized compositions (likely with known PSP linkages) for applications in industry, this schema should also allow users to explore and develop new alloys and CGAs with the artificial intelligence (AI), machine learning (ML), & data analytics methods of Industry 4.0. Exploring these novel alloy compositions with current methods will require initial, inexpensive CALPHAD modeling of expected material properties and phases. However, this predicted data should later be assessed with experimental characterization for increased fidelity at compositions of interest. Recent work by Ko et al. proposed a framework for PSP causal analytics in AM in which they point out the advantages for knowledge accumulation by integrating closed-loop iteration between both experimental efforts and multi-scale modeling & simulation [84]. By closing the loop between experimental efforts and modeling/simulation, such an approach allows for a continuous PSP-learning process where experimental data collection iteratively improves the mapping across processing, structure, and properties. This rapid learning capability enables the exploration of novel alloy compositions with increasing fidelity due to experimental feedback.

It is worth reemphasizing here that there is a specialized interest in compositional grading for experimental exploration of the alloy state space using compositional gradients [47, 48, 42]. Furthermore, an approach has been proposed recently for designing CGAs with customizable property changes along the build [60], which could be used for targeted assessment of property variation across the alloy state space. Due to the novel alloys often used to join dissimilar materials, a novel information schema that can be used to design compositional gradients for experimental assessment would enable a closed-loop continuous PSP-learning process with experimental results to validate or modify the cheaper computational predictions provided by CALPHAD. Thus, the requirements of data storage & reusability, topological interpretability, and dimensional scalability can be augmented with the additional requirement that the schema should be easily updated with new knowledge. This updated knowledge could come from sources including improved computational methods, literature, or self-driven experimental studies. This would enable the information

Table 2.1: Informatics schema requirements for the design of compositionally graded alloys

ID #	Requirement
1	Can store & reuse data
2	Must capture alloy state space topology
3	Needs to relate alloys for automated path planning
4	Scalable to any # of dimensions
5	Easily updated with new knowledge

schema to aid in the design of multi-material experiments, while also benefiting from them for an increasingly accurate representation of the alloy state space.

Considering these prior CGA design methods and CGA experimental considerations, the requirements for an ideal informatics approach can be summarized. The ideal information schema would possess both the data storage/reusability and design space interpretation capabilities of the visual CGA design method, combined with the design automation and dimensional scalability of the path planning method. Lastly, the informatics approach should be able to close the loop between experimental efforts and modeling/simulation to enable a continuous PSP-learning process to update our informatics approach with new, experimentally-assessed information. These requirements are summarized in Table 2.1.

Ultimately, to implement an informatics approach that can meet these requirements, various database representations that leverage relationships between data entries were investigated due to the basic requirement for the storage and reuse of data. To this end, relational database management systems (RDBMS) and graph databases (GDBs) were specifically investigated and compared due to their potential for addressing the listed requirements.

2.2 Concept Evaluation

2.2.1 RDBMS vs. Graph Databases

While RDBMS are a longstanding method for storing data with relationships, GDB platforms such as Neo4j [85] have grown in popularity more recently. GDBs fall under the "Not only SQL" (NoSQL) category of non-relational databases, while relational databases are considered struc-

tural databases, and are typically used with structured query language (SQL). NoSQL databases emerged in the last decade primarily from web services with large user bases such as Amazon, Facebook, and Google [86]. RDBMS are best for either 1) small and frequent read/write transactions or 2) large batch transactions with rare write accesses. They are not ideal for large-scale read/write workloads. As a result, large web service companies have developed their own NoSQL databases and workflows [86].

GDB methods provide noteworthy benefits compared to RDBMS under certain conditions. First and foremost, RDBMS are limited to smaller datasets with limited relationships, being ill-suited for large, highly-interconnected networks such as hyperlinks in the World Wide Web or social media communities [87, 88, 89]. Due to this low scalability property of RDBMS, when they increase in size to a certain point, they must be partitioned across systems, reducing data availability. Additionally, RDBMS are tabular in structure, relying on relationship tables to define the connections between data entries in the database records. Since this relational format requires specific structuring of the tabular data to fit with the relational scheme at the outset, RDBMS have a more rigid arrangement by comparison to GDBs and cannot be easily adapted to new relationship schemes or additional relationships [88]. For example, the GDB platform Neo4j exhibits significantly faster and more scalable query times compared to RDBMS for operations including finding immediate connections by edge type, retrieving features of immediate connections, and finding all vertices with specific label(s) [87, 88]. Other demonstrations illustrate GDB superiority over RDBMS on relationship traversal queries and character data queries, especially as the dataset scales [90]. Regarding reliability & data accessibility, RDBMS have the ACID traits (atomicity, consistency, isolation, and durability), while NoSQL databases provide the weaker BASE traits (Basically Available, Soft state, and Eventually consistent) at the expense of relaxed ACID traits [89, 86]. The ACID traits ensure reliable and accurate transactional processing, while the BASE traits are more concerned with database scalability and availability. While comparing GDBs and RDBMS, it is worth noting that RDBMS have some advantages in compactness and security. First, RDBMS in MySQL require slightly less storage than GDBs in Neo4j [90]. Secondly, due to their

longstanding use, RDBMS have much more security support, while GDBs have limited security features [88]. Thus, the choice between RDBMS and GDBs comes down to the specific use case and desired features.

As mentioned previously, some significant challenges with current materials informatics approaches to CGA design are the limitations regarding knowledge of the design space topology, particularly in more than three dimensions, and the need to store and reuse PSP information across the alloy state space. Thus, it is noteworthy that GDBs can function as living repositories for novel multi-material exploration and design due to their rich capabilities for representing data relationships/topology and high performance with data traversal queries compared to traditional RDBMS. A GDB could be initialized with existing knowledge from *ab initio* methods, CALPHAD software, and/or experimentation, but then grown and adapted over time with data from additional simulations, improved computational methods, and ongoing experimentation due to their flexibility.

On the experimental side, the GDB approach is notably valuable for designing CGAs. Due to the adaptability of GDBs, alongside their utility for design space exploration, GDBs overall show potential for facilitating and benefiting from high throughput multi-material experimentation. A GDB-based approach to multi-material design and data management could allow researchers to identify and develop CGAs to experimentally characterize gradients through the design space, explore desirable property changes, and validate manufacturability across critical compositions. These experiments could also be used to explore and compare the performance of different CGA paths. There is significant potential for this to extend experimental efficiency on its own, but even more, if leveraged alongside ongoing autonomous experimentation efforts for alloy development [91]. GDBs are significantly valuable due to their ability to design and benefit from these CGA experiments. By using the increased data output from multi-material experiments to overwrite or augment existing data, a GDB could rapidly become better tailored and validated for the manufacturing method(s) in use, building a better understanding of the design space topology, and informing future experiments and designs for both CGAs and single materials. Due to the significant potential of GDBs to meet the requirements outlined in Table 2.1, this type of database

is explored further in the next section in order to select a specific type of GDB for the proposed information schema.

2.2.2 Graph Database Selection

There are three primary types of GDBs as follows: 1) Labeled Property Graphs (LPGs), 2) Hypergraphs, and 3) triple stores or Resource Description Frameworks (RDFs) [74]. LPGs are based on the classical graph model $G = (V, E)$ where V is a set of nodes (also called vertices), and E is a set of corresponding edges (also called relationships) that compose the graph G [75]. Nodes are data entities, while edges are tuples recording the in and out nodes of the edge where $v_{in}, v_{out} \in V$, thus representing a relationship or interaction between nodes. Edges can be either directed or undirected. An undirected edge represents a connection in both directions, while a directed edge is a connection only in the specific order of the vertex pair as listed in V. The LPG model expands on the classical graph model using labels and properties that can be associated with individual nodes or edges. Labels differentiate between different classes of nodes and edges. Properties are key-value pairs that identify a property name and its corresponding value at a specific node or edge [92].

Hypergraphs are similar to LPGs, except edges can connect any number of nodes instead of just two. In this context, an edge can also be referred to as a hyperedge. While this expressiveness can be useful, Hypergraphs are not available in common GDB platforms. Additionally, due to their unique use of hyperedges, common graph algorithms may need to be altered to work with the hypergraph structure.

RDFs are more distinct compared to LPGs and Hypergraphs, as they use a subject-predicate-object triplet to store data. They are used primarily for applications related to the Semantic Web but have been applied for conceptual modeling in manufacturing [93]. Notably, RDFs can't store properties or labels alongside subjects/objects (similar to nodes) and relationships (similar to edges). Thus, their representational power is limited. To mimic this capability of LPGs & Hypergraphs in an RDF, one would have to define a subject corresponding to a label/property and define a relationship that relates that label/property to the desired object. Their edge-centric triplet structure

also results in potentially lower data traversal speeds in comparison to the node-centric structure of LPGs [94].

Considering the initial requirements enumerated in Tab 2.1, an information schema based on the LPG format is best equipped to achieve these requirements. First, an LPG can inherently store and reuse data, which can be accessed in a variety of ways through the use of graph queries. Second, the inherent graph representation with nodes and edges in an LPG database allows one to represent the topological information regarding the alloy state space, which can also relate alloys for graph traversal/path planning queries. Meanwhile, the label and property expressiveness of LPGs also allows for streamlined integration of data, regardless of dimensionality. GDBs are also inherently flexible, and they can easily be updated with new knowledge by modifying the schema or adding, modifying, or replacing the label/property data stored on the nodes and edges. By performing this investigation of LPG databases, their benefits for metal AM informatics are apparent due to their expressive capabilities, availability, flexibility, and efficiency with data storage and queries compared to other methods. Neo4j, a common LPG database software, allows for straightforward implementation of LPGs natively or through an API via a common platform such as Python [95]. NetworkX, a default Python library, allows implementation of LPGs directly in Python [7]. In the following section, a novel schema is defined, based on the LPG format, to meet the requirements for CGA design.

2.3 Proposed Information Schema

To create an LPG database for a specific application, it is important to define a graph schema. Fundamentally, a graph database structures a set of given data following a graph schema [96]. This schema describes the logical structure of the graph database, particularly the types of nodes and edges that a database following the schema can contain. While this is an important component of developing the initial implementation, the flexibility of LPGs allows for this schema to be adapted and modified as needed over time.

2.3.1 Alloy Topology-linked Information Schema

Based on the initial requirements in Table 2.1, the following Alloy Topology-Linked informAtion Schema (ATLAS) is proposed based on the LPG format. Figure 2.1 shows the formal hierarchical data model for the ATLAS, specifically tailored for the design of CGAs. This data model represents the alloy state space topologically with $Alloy$ nodes and $JOINS\_TO$ edges. Here, $Alloy$ nodes specifically represent unique compositions; a single processing condition is assumed since the current practice for manufacturing CGAs typically targets a single set of processing parameters to simplify the manufacturing process [63]. The $JOINS\_TO$ edges relate pairs of similar nodal $Alloy$ compositions that *have the potential* to be physically joined in a CGA. A given $Alloy$ node pair can be considered similar and assigned a $JOINS\_TO$ edge based on compositional similarity, experimental gradient assessment, or other criteria. In this work, compositional similarity is used, based on local grid-neighboring compositions in a grid sampling of the alloy state space. This enables complete coverage of the alloy state space with minimal edges, which is important since many graph queries scale with the number of edges present. Additionally, given the recent work on the generation of traversal graphs in simplicial composition spaces [97], sampling and edge generation for simplicial grids is much more efficient than random sampling and edge assignment based on k-nearest neighbors or distance. Together, these $Alloy$ nodes and $JOINS\_TO$ edges form the 'ATLAS materials graph', G_{AM}, at the core of the proposed ATLAS.

The other three node types, $Element$, $Phase$, and $ConstrainedSubgraph$, as well as their associated edge types, are included to filter the $Alloy$ nodes and increase graph query efficiency. These nodes are essentially the most relevant properties for filtering $Alloy$ nodes for graph queries, but can be defined as nodes to improve the efficiency of filtering the alloy state space using these properties since they enable efficient label filtering and edge traversals on large datasets. These properties were selected since they are expected to be the most commonly used for filtering the dataset. Due to the flexibility of GDBs, this schema can also be readily adapted in the future if other filtering criteria become relevant. In the proposed schema, $Element$ nodes and their associated $HAS\_ELEMENT$ edges linking them to $Alloy$ nodes allow for quick filtering of alloys

based on the desired alloy systems. $Phase$ nodes and their associated HAS_PHASE edges linking them to $Alloy$ nodes allow for similar filtering based on the desired alloy phase constitution. Lastly, the $ConstrainedSubgraph$ nodes and their associated $IN\_SUBGRAPH$ edges to $Alloy$ nodes allow for topologically partitioning of the alloy state space based on different constraints. Constrained subgraphs are identified and associated with $Alloy$ nodes based on their membership in the graph's maximally connected components (MCCs) [98] when the $Alloy$ nodes are filtered based on a specific set of constraints. For an arbitrary graph $G = (V, E)$, the n maximally connected components are subgraphs $G_i = (V_i', E_i'), i \in [0, 1, ..., n]$ where 1) G_i is connected for all i, and 2) for all vertices $u \in V$ and $u \notin V_i'$ no vertex $v \in V_i'$ exists where $(u, v) \in E$ [98]. $Alloy$ node membership within an MCC subgraph can be used for reachability queries between subsets of $Alloy$ nodes and topological partitioning of the alloy state space for query efficiency, network analysis, and adaptive graph refinement.

Altogether, the $ConstrainedSubgraph$, $Phase$, $Element$, and $Alloy$ nodes with their corresponding $IN\_SUBGRAPH$, HAS_PHASE, $JOINS\_TO$, and $HAS\_ELEMENT$ compose the ATLAS. A graph instantiated according to the ATLAS using any of the additional node or edge types beyond $Alloy$ and $JOINS\_TO$ is referred to based on the schema hierarchy as an 'ATLAS graph', G_{A}. The edges in this schema are undirected except for a special case for the $JOINS\_TO$ edges which will be discussed later in this chapter. The hierarchical distinction between an ATLAS materials graph G_{AM} and an encompassing ATLAS graph G_{A} is made since a G_{A} is ideal for accumulating large datasets for repeated global filtering based on multiple, changing criteria. Meanwhile, the use of an ATLAS materials graph G_{AM} is practical for smaller scale, directed alloy design efforts without the need for repeated global data filtering. After small scale design efforts with a G_{AM}, the graph data could be accumulated into an overarching ATLAS graph G_{A}. Here, G_{AM} functions as the core topological graph representation of the alloy state space, particularly when the $Alloy$ nodes and $JOINS\_TO$ edges are filtered into MCCs based on specific design constraints. Inversely, an ATLAS materials graph G_{AM} (or any subgraph thereof) can also be extracted from an overarching ATLAS graph G_{A} to reduce the necessary computational

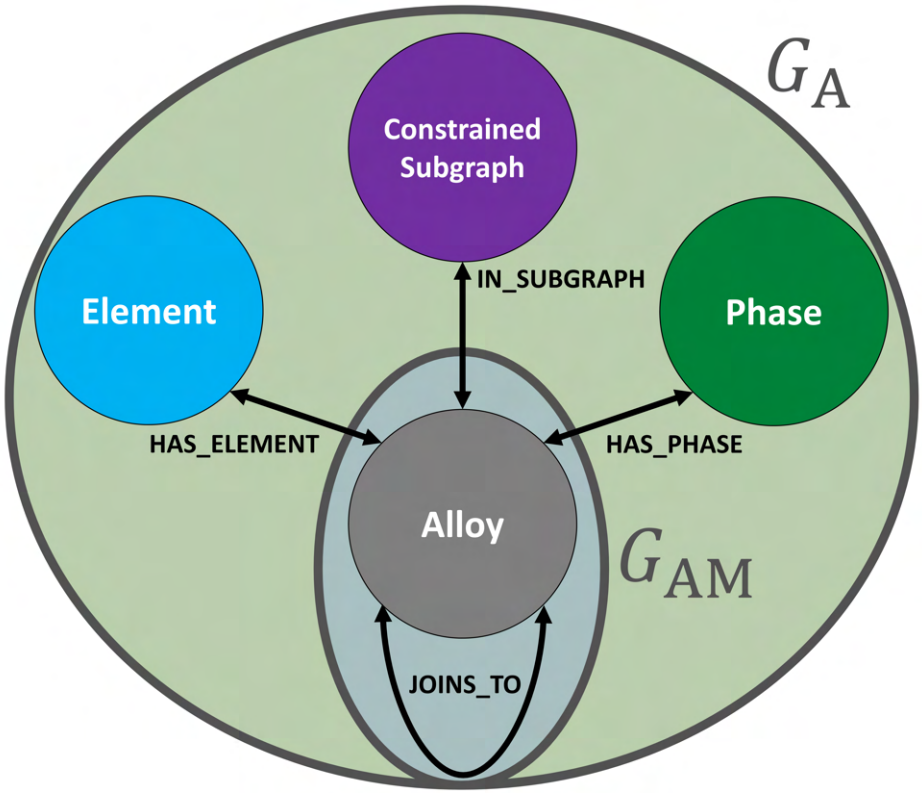

Figure 2.1: Proposed hierarchical Alloy Topology-Linked informAtion Schema (ATLAS) to encode process-structure-property relationships for metal additive informatics. This formal data model represents the types of nodes and edges which may be present in a graph database constructed according to the schema, with the ATLAS graph (G_{A}) and the ATLAS materials graph (G_{AM}) as hierarchical partitions within the schema. Adapted from the prior work in [2].

resources for storing data locally as well as performing queries and operations locally for alloy design.

2.3.2 Labels

The primary node labels for the ATLAS are simply the labels designating node type, namely $Alloy$, $Element$, $ConstrainedSubgraph$, and $Phase$. These labels designate the type of each node in accordance with the schema, with each node only possessing one of these labels. The same conditions are true for the edges, with labels $JOINS\_TO$, $HAS\_ELEMENT$, $IN\_SUBGRAPH$, and HAS_PHASE.

The $Alloy$ nodes have additional labels that can be assigned. First, labels such as $Modeled$ and

$ExpAssessed$ can be applied to indicate alloys that phases or properties for a given composition have been computationally modeled or acquired through experimental assessment, respectively. These labels could be modified specifically denote the methods used for such assessments. Meanwhile, $Feasible$ and $Infeasible$ labels can be used to designate compositions which pass (or do not pass) specific design constraints, particularly when working directly with G_{AM}. These feasibility labels enable categorical filtering of the feasible design space for CGA design when using the compact ATLAS materials graph with minimal filtering queries.

2.3.3 Properties

Properties enable the storage and use of data values at the node and edge level within an LPG. Here, relevant node- and edge-level properties for alloy design will be discussed. The properties outlined in this discussion are by no means an exhaustive list, and due to the flexibility of GDBs, one could easily subtract from, add to, or modify these properties when implementing the ATLAS.

2.3.3.1 Units

For an LPG, it is critical to define the units of all numerical graph properties. Since lists with heterogeneous data types cannot be stored in common LPG database platforms like Neo4j [95], this eliminates the option of storing a list with a numerical property and a unit string at the same property key. There are several practical ways to store the relevant units for interpreting the numerical properties in an LPG database. The first method is that units could be defined using corresponding node- and edge-level properties. For example, for an arbitrary numerical graph property key p_i, the corresponding graph property key p_i_Units can be defined, which stores the corresponding property units as a string. This allows units to be defined at the node/edge level, which enables more customization, but also results in a greater number of properties stored in the graph which store redundant information.

Alternatively, these units can be defined separately. One method for doing this would be to store the units in a separate data structure, such as a Python [99] dictionary with key-value pairs of graph property keys and their associated units. This table can be referenced when a user is

placing data into an ATLAS graph. This is a very simple approach but does not inherently define the units within the graph. Another approach, focused on storing this information in the graph, would be to define a single node, labeled $Units$. For all properties in an ATLAS graph, with keys $p_i \in P$, this $Units$ node can have a corresponding property key p_i_Units which is assigned a string denoting the units used for that property (if the property has units). The assigned unit should then be used for all instances of that numerical graph property in the graph. For example, if there is a numerical property $density$, used in the ATLAS graph, the $Units$ node can store a corresponding property such as $density\_Units =$ ['g/cm$\hat{3}$']. This approach is more complex than the Python dictionary approach since it involves instantiating a new nodes withing the GDB, but it allows the unit information to persist with the graph.

2.3.3.2 Alloy

The $Alloy$ nodes serve as the core of the ATLAS, each representing a unique alloy composition. This alloy composition is denoted for each node using the properties a) $composition$, which stores the alloy composition as a list of numbers in units such as mole fraction or mass fraction, and b) $alloySystem$, which indicates the elements corresponding to the list values in $composition$. For example, a nodal $Alloy$ with composed of 0.1 Cr, 0.83 Fe, and 0.7 Ni in mole fraction (mol/mol) would have property keys & values of $composition = [0.02, 0.91, 0.07]$ [mol/mol] and $alloySystem =$ ['Cr', 'Fe', 'Ni']. Note that for consistency, element strings should be listed in a consistent order, such as alphabetical order. While some properties of $Alloy$ nodes can be extended to the other nodes in the schema for efficient filtering of large datasets as mentioned previously, this information can be stored on the $Alloy$ nodes themselves as an $ATLAS$ materials graph for a compact format when dealing with smaller datasets. Specialized properties can be defined for any node in the proposed schema, including $Alloy$ nodes. For example, a nodal alloy composition predicted to be single phase with LP-DED based on CALPHAD could have the property $singlePhase\_LPDED\_CALPHAD =$ 'BCC'.

In cases where multiple property assessments have been done, the assessment type(s) can be stored as a homogeneous list of strings for both phases and physical properties. For example, at

a given alloy composition, computational modeling and experimental assessment/characterization can both be performed. In this case, the property assessment types can be stored under the property name $assessments\_p_i$, with a value such as $assessments_E =$ ['DFT', 'LPDED'] for a nodal $Alloy$ composition that has been computationally assessed for yield strength with density functional theory ('DFT') [100] but also experimentally assessed by manufacturing a sample with LP-DED and characterizing its properties [101] ('LPDED'). Associated material properties can be indicated in a couple of ways. One method is using a modifier to the property key. For example, for an arbitrary property key p_i, property keys for the various assessments could be denoted $p_i\_assessment$. For example, the experimentally assessment property key could be denoted p_i_LPDED while the DFT-assessed property could be denoted p_i_DFT. The second, more compact approach is simply to expand a given property p_i into a list. For example, property $E = [199.8, 202.5]$ where 199.8 [GPa] is the DFT-predicted modulus of elasticity and 202.5 [GPa] is the experimentally measured modulus.

For storing properties across temperature, this could be accomplished in several ways. One approach would be to store a corresponding identifier linking a property model which predicts phase fraction given a composition and temperature using a graph property key $p_i\_ModelID$. For example, the property key $E\_ModelID$ could denote the index for an elastic modulus model for an $Alloy$ node with property E, where the model is stored in a different data structure alongside the ATLAS graph. Alternatively, a property key listing corresponding temperatures p_i_Temps could be correlated to a property key p_i reporting the yield strength or another material property at those temperatures. For example, yield strength predictions using the strengthening model for refractory BCC high entropy alloys by Maresca and Curtin [102] for an $Alloy$ at multiple temperatures could be stored as $YS\_MarescaCurtin = [601.25, 402.20, 101.16]$ [MPa] and $YS\_MarescaCurtin\_Temps = [25, 500, 1000]$ $[^\circ C]$.

For an alloy with a single set of phase data, these can be stored simply with a $phaseFractions$ property storing a list of the numerical phase fractions, along with a $phaseNames$ property storing the corresponding phase names in order. An $assessmentPhase$ property can be used to designate

the source of the phase information, such as $assessmentPhase$ = 'CALPHAD'. If there are a single set of phases for a given composition, they can be stored in $Alloy$ node properties, but if temperature-dependent phases or phases from different assessments must be stored, they are more easily stored using the HAS_PHASE edges and $Phase$ nodes, which will be discussed later in this chapter.

2.3.3.3 Element

The $Element$ nodes have the primary properties of $elementName$ and $elementSymbol$ which each store a string designating the unique corresponding periodic element for a given $Element$ node. For example, an $Element$ node for the element iron would have $elementName$ = 'Iron' and, correspondingly, $elementSymbol$ = 'Fe'. Additional information from the periodic table, such as $atomicNumber$ and $molarMass$. For iron these would be stored as $atomicNumber = 26$ and $molarMass = 55.845$ [g/mol]. The cost and raw material criticality index [21, 22] of a specific element could also be stored as property keys $cost$ and $criticalityIndex$.

A given $Element$ node serves as the global representative of the corresponding periodic element as indicated by the $elementName$ nodal property. For example, to represent iron, one $Element$ node would be instantiated with the relevant metadata as discussed in the prior paragraph, including $elementName$ = 'Iron'. All references to iron would be made using the $HAS\_ELEMENT$ edge type between the relevant $Alloy$ nodes and the $Element$ node with $elementName$ = 'Iron'. This will be discussed in a later section on the $HAS\_ELEMENT$ edge type.

2.3.3.4 ConstrainedSubgraph

The $ConstrainedSubgraph$ nodes represent MCC subgraphs in the ATLAS materials graph, G_{AM}, when $Alloy$ nodes are filtered based on specific constraints. An example of the procedure for constraint filtering and MCC identification is shown in Fig. 2.2. These $ConstrainedSubgraph$ nodes store the index of the MCC they represent, when the MCCs are arranged in descending order based on the number of nodes they contain. The $ConstrainedSubgraph$ nodes can store

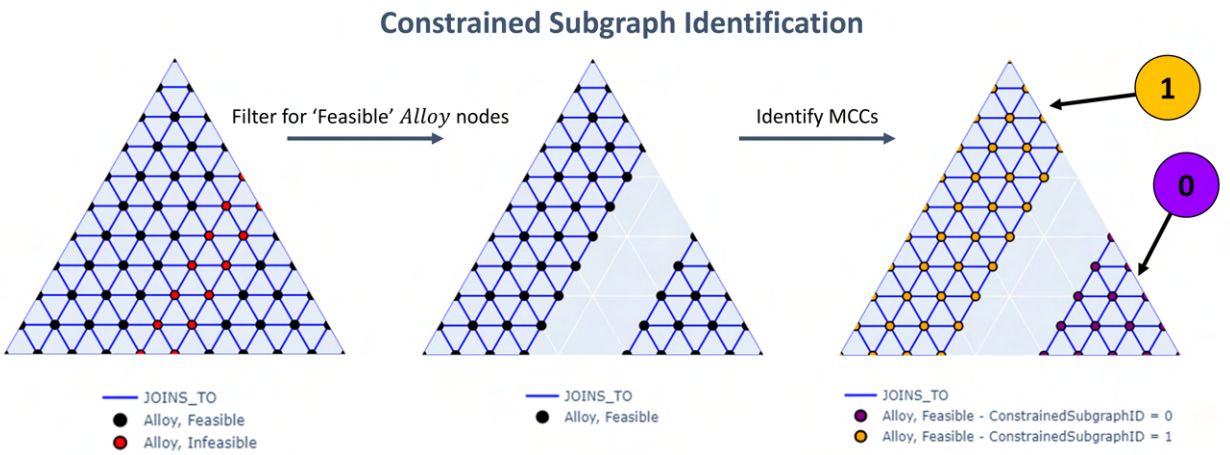

Figure 2.2: Illustration of the procedure for identifying constrained subgraphs in an ATLAS materials graph, G_{AM}, of a synthetic ternary alloy system. In this example, the constraint is that $Alloy$ nodes must have the label $Feasible$. Once the alloy nodes are filtered based on the constraint, the MCCs of the remaining graph can be easily identified with a breadth-first search (BFS). The constrained subgraphs are the MCCs resulting from this specific constraint filtering, which are indexed in descending order by the number of nodes they contain, as shown with the two constrained subgraphs highlighted in the ternary graph plot on the right.

the corresponding index as $constrainedSubgraphID$. The constraints can be stored with keys specific to the type(s) of constraints used. The constraint conditions define the nodes that will be retained after the filtering. Constraints for specific labels can be stored as $labelConstraints$ which stores a list of the required $Alloy$ node labels, such as $labelConstraints =$ ['Feasible'] for the example shown in Fig. 2.2.

Constraints for specific $Alloy$ node properties can be stored as $propertiesMax$ and $propertiesMin$, with corresponding property keys $propertiesMaxVals$ and $propertiesMinVals$. Here, $propertiesMax$ denotes the specific $Alloy$ nodal property keys, and $propertiesMaxVals$ with stores the maximum cutoff values for each respective property. Inversely, $propertiesMin$ stores a list of the property names for any minimum constraints, and $propertiesMinVals$ with stores the corresponding list of minimum cutoff values for each property. For example, a constraint where the elastic modulus must be above 98.0 GPa, and the yield strength must be between 210.0 and 360.0 Mpa would be represented as $propertiesMin =$ ['elasticModulus', 'YS'] with $propertiesMinVals = [98.0, 210.0]$ and $propertiesMax =$ ['YS'] with $propertiesMaxVals = [360.0]$. Meanwhile, nodal $Alloy$ graph property constraints which are not numerical can also be stored. For example, constraints on categorical/string properties can be stored as $propertiesCategorical$ with a list of categorical $Alloy$ property keys, and $propertiesCategoricalVals$ with a list of corresponding strings required to pass the constraint. For example, $propertiesCategorical =$ ['singlePhaseLPDED_CALPHAD'], $propertiesCategoricalVals =$ ['BCC'] could be used to constrain the alloy state space, keeping compositions which were predicted by CALPHAD to be single phase BCC if manufactured with LP-DED.

2.3.3.5 Phase

$Phase$ nodes have one terminal at an $Alloy$ node and the other at an $Phase$ node. These $Phase$ nodes have the primary property $phaseName$ which stores a string designating the unique phase name for the corresponding $Phase$ node. For example, a $Phase$ node for the σ-phase would be the only $Phase$ node with $phaseName =$ 'sigma'. Additional phase information such as $prototype$ can also be stored. For σ-phase these would be stored as $prototype =$ 'sigma-CrFe'.

Similar to $Element$ nodes, a given $Phase$ node serves as the global representative of the corresponding phase, as indicated by the $phaseName$ nodal property. All $Alloy$ nodes referencing a particular phase can do so with the HAS_PHASE edge type between those $Alloy$ nodes and the relevant $Phase$ node. This will be discussed in a later section on the HAS_PHASE edge

type.

2.3.3.6 JOINS_TO

The $JOINS\_TO$ edges have an $Alloy$ node at each terminal. These edges serve alongside $Alloy$ nodes as the core of the ATLAS schema and the two aspects of an ATLAS materials graph. These edges represent *the potential* for pairs of $Alloy$ nodes with similar chemistry to be physically joined in a compositional gradient, and enable graph traversal of the alloy state space. This traversal can be used for BFS identification of MCCs for evaluating the alloy state space subject to design constraints, as well as queries such as path planning for designing compositional gradients.

Since the $JOINS\_TO$ edges enable path planning, they can be used to store relevant cost function information, which can be used as edge weights for algorithms to find the so-called 'shortest'path in a weighted graph. Since the notion of 'shortest'primarily relates to considerations regarding geometric distance, in this work we will refer to this instead as the minimum-weight path in a weighted graph. Several cost functions have been used for applying robotic motion-planning techniques to CGA design, including cost functions for minimizing: compositional path length [12, 63], distance from undesirable phase regions [12, 63], property lack-of-monotonicity [60], and constraint violation [64]. These cost functions can be computed and stored as properties on the $JOINS\_TO$ edges as $cost\_c$, such as $cost\_CompositionalLength$ for the value of the compositional path length cost function on a given $JOINS\_TO$ edge. If one is using a direction-specific cost function, such as property lack-of-monotonicity, the $JOINS\_TO$ edges must be defined as directed edges, since the corresponding property change must be identified as either increasing or decreasing across the edge.

2.3.3.7 HAS_ELEMENT

The $HAS\_ELEMENT$ edges have one terminal at an $Alloy$ node and the other at an $Element$ node. These edges denote that the composition of a given $Alloy$ node contains a given element corresponding to an $Element$ node. Each $Alloy$ node has a unique $HAS\_ELEMENT$ edge to all $Element$ nodes corresponding to the elements present in that alloy's composition. This

edge can store information such as the amount of each element present as $moleFraction$ or $weightFraction$. For example, an $Alloy$ node with property $composition = [0.10, 0.83, 0.07]$ [mol/mol] and $alloySystem =$ ['Cr', 'Fe', 'Ni'] would have $HAS\_ELEMENT$ edges to the three $Element$ nodes with the corresponding property $symbol =$ 'Cr', 'Fe', and 'Ni'. These $HAS\_ELEMENT$ edges would have the corresponding property values $moleFraction = 0.10$, 0.83, and 0.07 [mol/mol], respectively.

2.3.3.8 IN_SUBGRAPH

The $IN\_SUBGRAPH$ edges simply denote a given $Alloy$ nodes membership in an MCC when the ATLAS materials graph G_{AM} is constrained based on a certain set of criteria. Each $Alloy$ node can have up to one $IN\_SUBGRAPH$ edge to a $ConstrainedSubgraph$ node per unique set of constraints. If that $Alloy$ node does not pass a given constraint, then it is not present in any MCC for that constraint and thus does not have an $IN\_SUBGRAPH$ edge for that constraint. There are no necessary properties for edges with this $IN\_SUBGRAPH$ label.

2.3.3.9 HAS_PHASE

If multiple sets of phase data for an $Alloy$ node are required due to different temperatures, assessment types, etc., this data can be stored on the HAS_PHASE edges. This is because each $Alloy$ nodes has a unique HAS_PHASE edge to a $Phase$ node corresponding to every phase present in that alloy. The assessment types for phases can be stored as a list with the key $assessmentsPhase$. For example, $assessmentsPhase =$ ['CALPHAD', 'LPDED-XRD'] for a nodal $Alloy$ composition with phases predicted from CALPHAD as well as from an experimental LP-DED print that was characterized using X-ray diffraction (XRD) [103] and Rietveld analysis [104, 105, 106] to determine phase fractions. The corresponding $phaseFraction$ property can store a list of the phase volume fractions corresponding to the list of assessment techniques. These phase fractions are specific to the nodal $Phase$ terminal of the HAS_PHASE edge, which has a $phaseName$ property to indicate the corresponding phase for that node and all associated HAS_PHASE edges.

For storing phases across temperature, this could be accomplished similarly to storing $Alloy$ material properties across temperature. One could store a key/index linking a phase model which indicates the fraction of a given phase for a specific composition and temperature using a property key like $phaseModelID$. Alternatively, a property list of temperatures $phaseTempsAssessment$ could be correlated to a property list of the phase volume fraction $phaseFractionAssessment$ at those temperatures. For example, CALPHAD equilibrium phase predictions at multiple temperatures could be stored as $phaseFractionEqCALPHAD = [0.20, 0.25, 0.98]$ [mol/mol] and $phaseTempsEqCALPHAD = [25, 500, 1000]$ $[°C]$.

2.3.4 Graph Label & Property Summary

Overall, the LPG data representation enables the ability to store relevant data on the nodes and edges in the graph, particularly for filtering the nodal $Alloy$ compositions based on design constraints. The proposed ATLAS includes several node types with corresponding edge types that can be used to represent the alloy state space for materials design and efficiently filter the state space as the number of nodal $Alloy$ compositions increases. These node and edge types are designated by a primary label associated with each respective node or edge in the graph. A variety of additional labels and properties were discussed for both nodes and edges. In summary, the node labels and the potential corresponding properties outlined in this work, with examples, are summarized in Table 2.2. Similarly, the primary edge labels and the potential corresponding properties outlined in this work, with examples, are summarized in Table 2.3. Due to the flexibility of GDBs, these properties and labels can be easily adapted in the future as research in this field continues to progress.

Table 2.2: Summary of the ATLAS node labels & property types with examples

Node Labels	Property Key(s)	Example(s)
$Units$[1]	p_i_Units	$density\_Units$ ='g/cm^3'
$Alloy$[2] $Feasible$, $Infeasible$, $Modeled$, $ExpAssessed$	p_i	$YS = 361.18$ $singlePhase\_LPDED\_CALPHAD$ = 'BCC'
	$composition$ & $alloySystem$	= [0.02, 0.91, 0.07] & = ['Cr', 'Fe', 'Ni']
	$assessments\_p_i$ & $p_i\_Assessment$ or p_i	$assessments\_E$ =['DFT', 'LPDED'] & $E\_DFT = 19.98$ $E\_LPDED = 20.25$ or $E = [19.98, 20.25]$
	$p_i\_ModelID$ & p_i	$E\_ModelID = 25$ & E ='model'
	p_i & p_i_Temps	$YS\_MarescaCurtin = [691.25, 489.20, 111.16]$ & $YS\_MarescaCurtin\_Temps = [25, 500, 1000]$
	$phaseFractions$ & $phaseNames$	= [0.22, 0.22, 0.32, 0.24] & = ['BCC', 'Cr2O3', 'FCC', 'FCC_L12']
	$assessmentPhase$	= 'CALPHAD'
$Element$	$elementName$	= 'Iron'
	$elementSymbol$	= 'Fe'
	$atomicNumber$	= 26
	$molarMass$	= 55.845
	$cost$	= 120
	$criticalityIndex$	= 5
$ConstrainedSubgraph$	$constrainedSubgraphID$	= 1
	$labelConstraints$	= ['Feasible', 'ExpAssessed']
	$propertiesMax$ & $propertiesMaxVals$	= ['YS'] & = [360.0]
	$propertiesMin$ & $propertiesMinVals$	= ['elasticModulus', 'YS'] & = [98.0, 210.0]
	$propertiesCategorical$ & $propertiesCategoricalVals$	= ['singlePhase_LPDED_CALPHAD'] & = ['BCC']
$Phase$	$phaseName$	= 'sigma'
	$prototype$	= 'sigma-CrFe'

[1] Optional

[2] Necessary for ATLAS materials graph

Table 2.3: Summary of the ATLAS edge labels & property types with examples

Edge Labels	Property Key(s)	Example(s)
$JOINS\_TO$[1]	$cost\_c$	$cost\_CompositionalLength = 0.05$
$HAS\_ELEMENT$	$moleFraction$	$moleFraction = 10$
	$weightFraction$	$weightFraction = 16$
$IN\_SUBGRAPH$	N/A	N/A
HAS_PHASE	$assessmentsPhase$ & $phaseFraction$	$assessmentsPhase =$ ['CALPHAD', 'LPDED-XRD'] & $phaseFraction = [0.20, 0.25]$
	$phaseModelID$ & $phaseFraction$	$phaseModel = 12$ & $phaseFraction =$'model'
	$phaseFractionAssessment$ & $phaseTempsAssessment$	$phaseFractionEqCALPHAD = [0.20, 0.25, 0.98]$ & $phaseTempsEqCALPHAD = [25, 500, 1000]$

[1] Necessary for ATLAS materials graph

2.3.5 Efficient Graph Filtering

As mentioned prior, the three labeled node types, $Element$, $Phase$, and $ConstrainedSubgraph$, as well as their associated edges, were included to filter the $Alloy$ nodes and increase graph query efficiency since they enable efficient label filtering and edge traversals on large datasets. As an example, $Alloy$ nodes could all be filtered by elements based on the nodal property $alloySystem$. However, this may become cumbersome as the amount of alloy data increases for sufficiently large alloy design campaigns in the future. This property filtering can be done efficiently for the much smaller amount of $Element$ nodes (at most, the 118 chemical elements in the periodic table), $Phase$ nodes, or $ConstrainedSubgraph$ nodes, compared to the potential hundreds of thousands of nodal $Alloy$ compositions [107] based on current efforts or even more with the increasing adoption of parallel computing [108] and high-throughput experimentation [108, 109].

For example, finding all alloys with specific elements, such as Cr, Fe, and Ni, can be done by finding the corresponding $Element$ nodes via a property search for corresponding values of $elementSymbol$ or $elementName$. Then, the union of $Alloy$ nodes on the $HAS\_ELEMENT$

edges from the identified $Element$ nodes would yield all nodal $Alloy$ compositions with Cr, Fe, and Ni. If alloys containing only Cr, Fe, and Ni are desired, then of the $Alloy$ nodes selected, only those with three $HAS\_ELEMENT$ edges are kept.

2.4 Schema Implementation

2.4.1 Graph Data and Analysis Tools

There are several methods for implementing the ATLAS with a materials dataset. This schema can be implemented for CALPHAD, experimental, and/or any other relevant materials data for multiple compositions in the alloy state space. To implementing the ATLAS with a dedicated GDB platform, one can use a platform that supports the LPG format such as Neo4j [95]. Neo4j is a popular platform for LPG databases, with either local- or cloud-hosted options. It also has wide adoption and support as well as competitive computational performance among LPG platforms [110]. These characteristics make Neo4j well-suited for hosting large datasets in an ATLAS graph format, especially in a collaborative context due to its features for cloud-hosting, security, and data backup. Additionally, the Neo4j Python API can be used to implement code for creating and querying Neo4j GDBs for increased flexibility with Python-based computational workflows.

Initializing a Neo4j database with the specified graph schema can be accomplished through several methods, including importing comma-separated values (CSV) data for nodes and edges, or using the Python API to interact with the database and import data. The CSV approach is easier for manually checking data and ensuring input formats are correct, while the approach using Python or other APIs is easier for integrating GDBs into automated workflows. Once the data is uploaded for labeled $Alloy$ compositions, the $Element$, $Phase$, and $Subgraph$ nodes with their associated edges can be added in accordance with the schema as needed. Then, one can use the Python API to perform GDB queries on the implemented GDB.

Meanwhile, graph programming libraries, such as the Python library NetworkX [7], can be used for straightforward implementation of an ATLAS materials graph for alloy design. This is best suited for working with smaller datasets that can be used in local memory. In practical

use, datasets could be extracted from a GDB instance and then converted to NetworkX for local workflows, or, inversely, graphs could assembled in NetworkX from tabular data for local work or formatting, and then uploaded and/or merged with a GDB instance. NetworkX is a common library for graph data analysis, which enables the storage of node and edge *attributes* as a generic, combined abstraction of the labels and properties in an LPG context. This library has a multitude of relevant graph queries, including shortest path searches and a BFS method for identifying MCCs.

2.4.2 Data Collection & Initialization

As mentioned previously, compositional similarity in this work is based on local grid-neighboring nodal $Alloy$ compositions in a grid sampling of the alloy state space. These grid-neighboring $Alloy$ composition pairs are all assigned a $JOINS\_TO$ edge. This enables complete coverage of the alloy state space with relatively few edges, as the computation time for graph traversal queries relates to the number of edges in a given graph. For example, alternative graph methods for path planning, such as optimal probabilistic roadmaps (PRM*) [62] typically have greater edge connectivity among nodes, which improves path optimality at the expense of query time as the number of nodes and edges increase. While the sampling-based graph approach using optimal probabilistic roadmaps (PRM*) has probabilistically complete and asymptotically optimal guarantees for path planning, these rely on high density random sampling. Under these conditions, the construction of a PRM* would scale poorly with current methods, and the vastly increased number of nodes and edges would negatively impact the performance of many graph queries. Grid sampling and graph generation is much more efficient than random sampling and edge assignment based on k-nearest neighbors or distance given the recent work on the generation of grids and traversal graphs in simplicial composition spaces [97]. The compositional similarity of nodal $Alloy$ compositions is assumed for grid-neighboring compositions in a grid sampling of the alloy state space.

A simplicial graph of nodal $Alloy$ compositions, defined by an array of nodal compositions and corresponding grid-neighbor edges, can be efficiently generated using the NIM simPLEX library (NIMPLEX) [97]. The nodal compositions can be assessed with high-throughput CALPHAD calculations using open-source tools such as pycalphad [111] and OpenCalphad [112], or

commercial software such as Thermo-Calc [5]. This provides an initial modeled assessment of a space-covering sampling of compositions in the alloy state space, which can be refined with future modeling and experiments. With an initial set of $Alloy$ nodes and $JOINS\_TO$ edges, the ATLAS materials graph is in place. Additional data regarding elements, phases, and constrained subgraphs can be added at the $Alloy$ node level as needed, or by using the full ATLAS schema and using the $Element$, $ConstrainedSubgraph$, and $Phase$ nodes to define an ATLAS graph.

2.5 Two-terminal Gradient Design

Previous work on two-terminal CGA design focused on the use of sampling-based path planning methods, particularly RRT*FN, due to guarantees of probabilistic completeness and asymptotic optimality [12, 63, 64, 60, 14]. Graph path-finding searches were avoided primarily due to the expectation that they would be prohibitively expensive to construct for a gradient design problem, with additional concerns about space complexity. However, efforts using sampling-based planners primarily use high-throughput CALPHAD calculations for a blanket sampling of the chemistry-temperature state space to gather data for training the necessary obstacle models. This blanket computation is a significant portion of the computational expense involved with constructing a graph of the alloy state space.

The remaining computationally expensive step for graph construction is instantiating the edges, which has been efficiently addressed for simplicial grid sampling in the recently introduced NIMPLEX library by Krajewski et al. as mentioned previously [97]. Thus, by leveraging an initial simplicial grid sampling of the chemistry-temperature state space with high-throughput CALPHAD calculations, one can instantiate an ATLAS graph with little to no more computational cost than preparing the obstacle models for prior methods, while unlocking the benefits of graph path planning methods, such as Dijkstra's algorithm [113] or A* [114].

These graph path planning methods guarantee optimality and completeness, and increased efficiency in the case of A* if an admissible heuristic is used. Notably, Dijkstra's algorithm has recently been shown to achieve *universal optimality* when paired with a specific heap data structure that leverages the working set property, meaning that the algorithm performs at the highest

level of efficiency with respect to running time and number of comparisons on any graph topology [115].

The concern regarding the space complexity of graph-based path planning is addressed by this work in two ways. First, in-memory storage is not necessary for GDBs, which can be primarily kept in disk storage or on the cloud [95]. For graphs called into RAM for path planning, the space complexity can be reduced by leveraging the applicable constrained subgraphs, since only nodal $Alloy$ compositions within the same constrained subgraphs are reachable. This can significantly reduce the number of edges and nodes called into RAM, in addition to the use of simplicial grid graphs instead graphs with dense nodes and edges such as PRM*.

2.6 Demonstrations

This section discusses two demonstrations to illustrate the use of the ATLAS in practice with real materials datasets. First, a simple illustration of an ATLAS materials graph is shown for the FeNiCr ternary system. A simple ternary is used first due to the straightforward visualization of the ATLAS materials graph's topological representation of the alloy state space alongside exemplary labels and properties for the $Alloy$ nodes and $JOINS\_TO$ edges. Then, the scalability and topological interpretability of the proposed schema in higher dimensions is shown using an ATLAS materials graph in a 5-element FeNiCrTiCo alloy state space, which would not be interpretable with prior methods.

2.6.1 FeNiCr Alloy System

A simple illustration of an ATLAS materials graph is shown in detail for the FeNiCr ternary in Fig. 2.3. The ternary diagram on the left shows $Alloy$ nodes, sampled with a compositional grid spacing of 0.05 mole fraction [mol/mol], and grid-neighboring $Alloy$ nodes were connected with a $JOINS\_TO$ edge. The compositions were then labeled based on a binary k-nearest neighbors (KNN) classifier (with parameter $k = 3$) using equilibrium phase data from Thermo-Calc's [5] TCHEA database version 4 [116] for the FeNiCr alloy system. In this training data, compositions were labeled 'infeasible' if they contained >1 mole fraction [mol/mol] of any deleterious phase

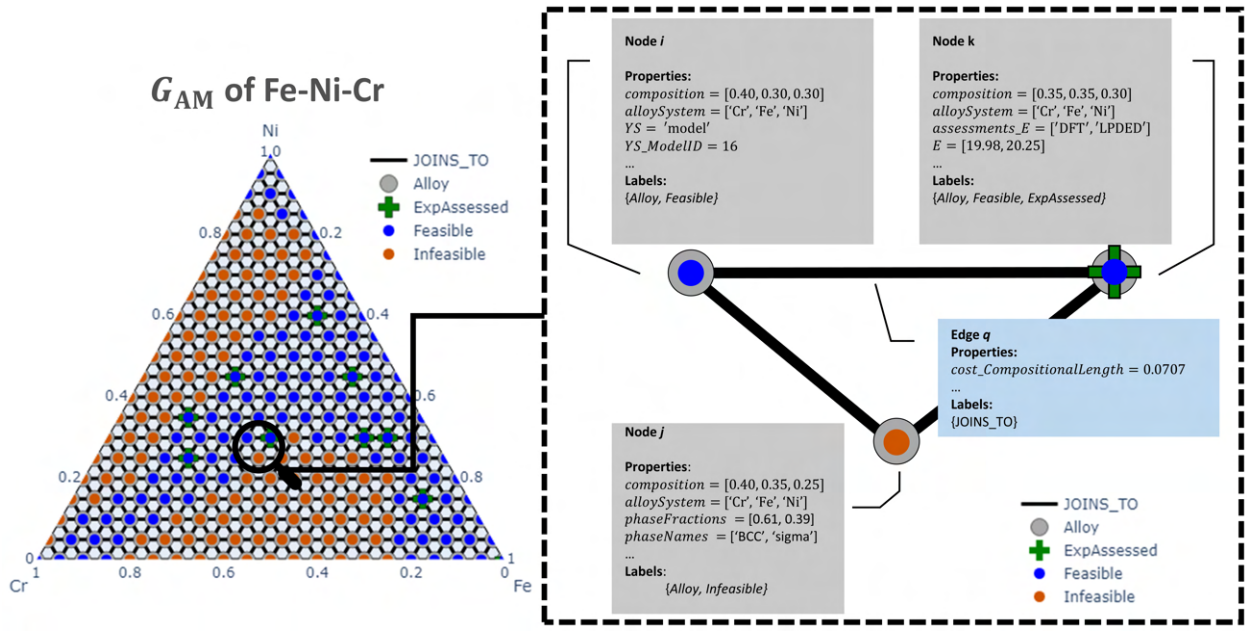

Figure 2.3: ATLAS materials graph, G_{AM} containing $Alloy$ nodes and $JOINS\_TO$ edges for the FeNiCr alloy design space. Note that the $Alloy$ nodes and the $JOINS\_TO$ edges can have labels and properties as part of a Labeled Property Graph. The $JOINS\_TO$ edge properties can be used to store cost functions for path finding, such as Euclidean compositional distance.

from the manufacturing temperature of 1100 $^{\circ}C$ down to 300 $^{\circ}C$ to emulate the constraint used in an experimental CGA study in the FeNiCr system [63]. The $Alloy$ nodes in G_{AM} were assigned either an $Infeasible$ or $Feasible$ label based on the output of the KNN classifier at their corresponding composition. Fig. 2.3 shows the resultant gray $Alloy$ nodes, with different markers based on their auxiliary labels, where $Infeasible$ is indicated with an orange dot, $Feasible$ with a blue dot.

Meanwhile, the label $ExpAssessed$, indicated with a green cross, illustrates how $Alloy$ nodes can be labeled according to the experimental assessment of the alloy state space. The $ExpAssessed$-labeled $Alloy$ nodes in this demonstration were added as a visual example of the multi-fidelity capabilities of an ATLAS graph database to incorporate data from multiple sources. The diagram on the right side of Fig. 2.3 shows example properties and labels for the three nodes and one edge within the magnified region of the graph, where the node i has a corresponding yield strength property model while node k has experimentally assessed and DFT-modeled material properties.

The $JOINS\_TO$ edge highlighted in the magnified region of the graph in Fig. 2.3 includes an example of the common Euclidean compositional distance cost function for CGA design using path planning [12, 64, 14], stored as an edge property.

The proposed ATLAS enables a variety of capabilities, which should be noted for this simple example before moving on to higher-dimensional alloy state spaces. First, it facilitates holistic materials design exploration by identifying MCCs in G_{AM} based on arbitrary label and property constraints enforced on the $Alloy$ nodes. In the G_{AM} shown in Fig. 2.3, there is only one graph MCC when applying the constraint that $Alloy$ nodes must have the $Feasible$ label. This constrained subgraph is shown in Fig. 2.4. Here, it is visually evident that an MCC is a connected subgraph, that is, within a single MCC every $Alloy$ node is reachable from any other $Alloy$ node by traversing a series of $JOINS\_TO$ edges. This is critical information for CGA design and was interpreted visually in 2D ternary diagrams with previous methods. However, in this work, we will show how this can be extended to higher dimensions since it is now accomplished using an automated graph query.

For CGA design, this partitioning can be used for the systematic selection of terminal alloy compositions based on desired performance characteristics. The compatibility of terminal alloys can be easily known by comparing their subgraph membership based on the current design constraints. If terminal alloys are within the same constrained subgraph, the existence of a gradient path is guaranteed, which was not possible with the prior robotic path planning approach. This schema also enables efficient gradient alloy design by traversing the $JOINS\_TO$ edges with existing graph path finding methods such as Dijkstra's algorithm [113] or A* [114].

While the ternary example shown in Fig. 2.3 involves a simple $Feasible$ label constraint on the $Alloy$ nodes for CGA design, this general approach to filtering and topological partitioning of the alloy state space is applicable in the broader context of alloys design. For example, the design space could be topologically partitioned based on continuous compositional regions with a specific phase (or phases) to assess compositional regions of the high-dimensional alloy state space with respect to phase stability. This could be done for any set of constraints, enabling topologically-

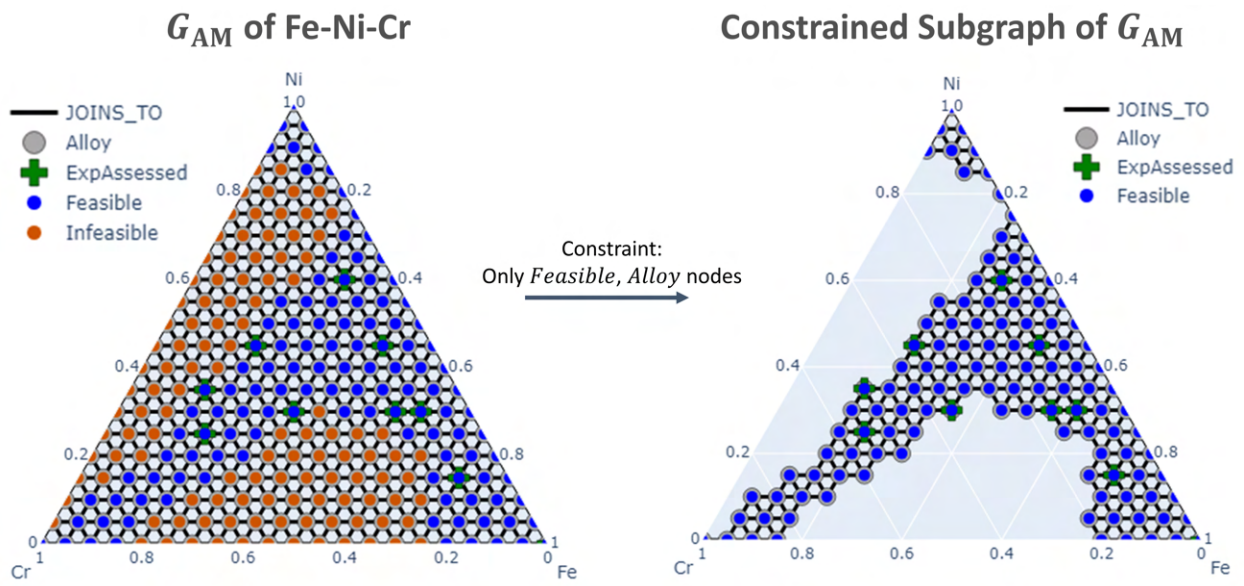

Figure 2.4: Illustration showing the extraction of a constrained subgraph of the ATLAS materials graph, G_{AM} containing $Alloy$ nodes and $JOINS\_TO$ edges for the FeNiCr alloy design space. In this case, the constraint is that $Alloy$ nodes must also have the $Feasible$ label. In this case there is only one MCC, shown on the right. This partitioning guarantees that pairwise paths exist on the $JOINS\_TO$ edges for any set of $Alloy$ nodes within the subgraph.

interpretable partitioning of the alloy state space in arbitrary dimensions. The following subsection will show how these partitions can be visualized, even for alloy chemistries with more than three elements.

2.6.2 FeNiCrCo & FeNiCrTi Alloy Systems

2.6.2.1 Topological Partitioning & Analysis

One of the primary benefits of an ATLAS graph is this capability to find and partition the alloy state space using arbitrary constraints, particularly in high dimensions. Using this capability alongside methods for dimensionality reduction for visualization enables visual topological interpretation of high dimensional alloy design state spaces. Several dimensionality reduction techniques have been used in literature for visualizing the state space for multi-principal element alloys (MPEAs), which are compositionally complex alloys without a single principal alloy component [117]. These methods include t-Distributed Stochastic Neighbor Embedding (t-SNE)

[118, 119, 120] and Uniform Manifold Approximation and Projection (UMAP) [121, 18, 101].

Recently, it has been shown that, compared to UMAP, t-SNE has a tendency to skew compositions with a principal alloy component towards the boundary of the projection, resulting in overcrowding [122]. Thus, in this work, UMAPs are used for dimensionality-reduced visualizations of alloy state spaces with more than three elements. The UMAP dimensionality reduction technique is an unsupervised machine learning method for non-linear dimensionality reduction, which first assumes that the input data is uniformly distributed on a locally-connected Riemannian manifold [121]. Then, by using local manifold approximations, the UMAP method iteratively optimizes a dimensionality-reduced fuzzy topological representation to emulate the original high-dimensional data. When used with regular simplicial grid samplings of the alloy chemistry space, UMAP projections are hypocycloidal shapes with d cusps, where d is the number of elemental components in the alloy chemistry space. With UMAP, similar compositions will appear closer in the 2D projection, with purer compositions closer to the cusps and compositions with more configurational entropy (MPEAs) closer to the center.

To demonstrate this capability for a real alloy state space, the topological partitioning of a 5-element alloy state space is shown using UMAPs in Fig. 2.5. This state space is comprised of the FeNiCrCo & FeNiCrTi quaternary alloy systems sampled in a grid at 5 at. %, which overlap at the FeNiCr ternary. First, the UMAP for the entire alloy state space is shown in the top left, illustrating the specific stiffness, E/ρ, at different feasible $Alloy$ node compositions. While this is useful for understanding the breadth of compositions present in a material design space with more than three degrees of freedom [18], the overlap between dissimilar UMAP points visually conceals partitions in the overall topology of the design space. This ultimately inhibits visual identification of continuous regions of feasible alloy compositions for CGA design.

The five UMAPs in the bottom left of Fig. 2.5 show the 5 largest (by number of nodes) constrained subgraphs. The topological distinctness of these subgraph regions is not apparent when revisiting the original UMAP projection of the entire state space. Since, by definition, these subgraphs are connected graphs, at least one path exists between any pair of nodes within a given

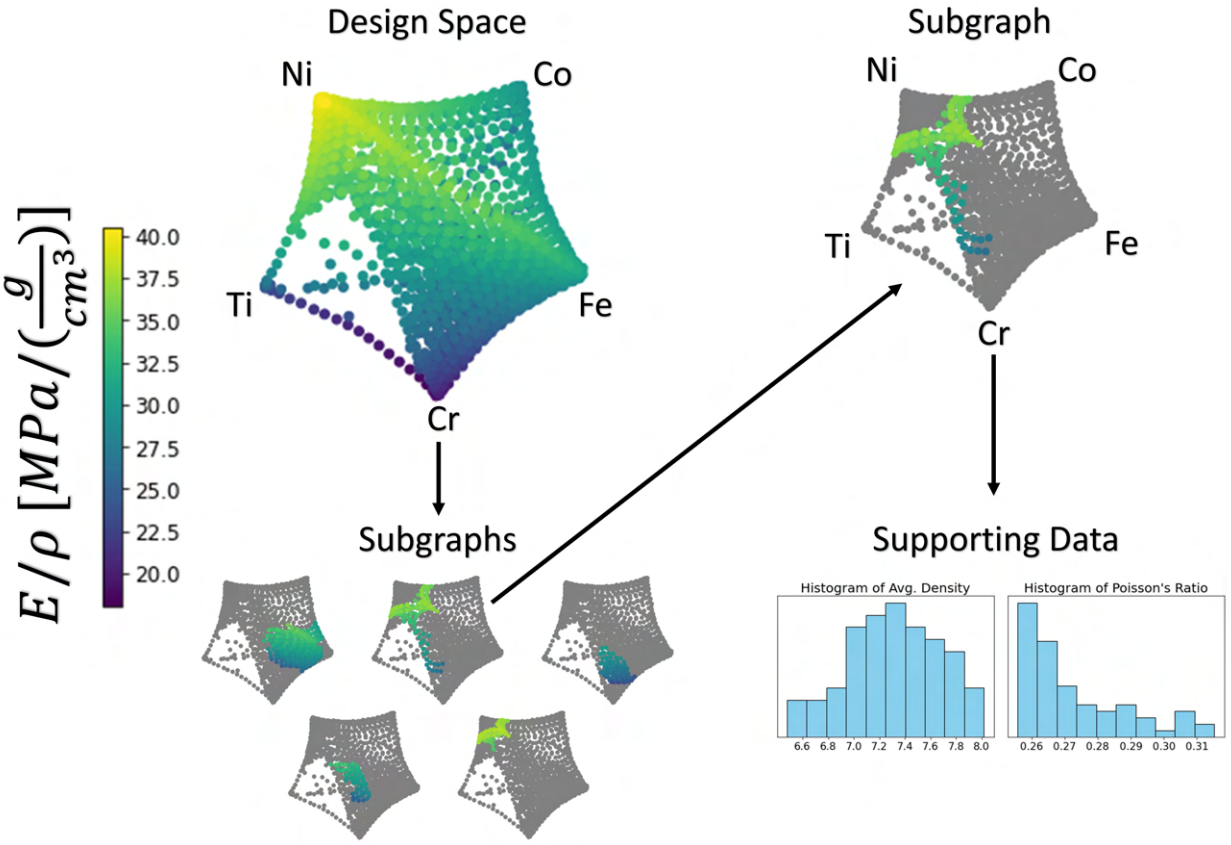

Figure 2.5: UMAP dimensionality-reduced illustration of the combined FeNiCrCo & FeNiCrTi quaternary state spaces, constrained for feasibility, showing the specific stiffness of the feasible nodal *Alloy* compositions. The alloy state space can be broken down into maximally connected component subgraphs, of which the largest 5 are shown. Further analysis of subgraphs can reveal specific information of interest such as property trade-offs.

subgraph. This is also called reachability in the context of graphs. Inversely, no path exists between any pair of nodes in two separate subgraphs. Since these subgraphs are mutually exclusive of one another, they can be analyzed as independent CGA design spaces. Thus, the question of reachability between two or more terminal $Alloy$ compositions can be answered before gradient path planning entirely. This eliminates trial and error when looking for compatible alloys in alloy state spaces with three or more dimensions, and was the driving motivation for computing the MCCs for constrained ATLAS materials graphs and including the identified subgraphs in the formal definition of the ATLAS.

Further analyses can be applied to the constrained subgraphs, such as observing the ranges of properties and compositions within each subgraph. In the case of the selected subgraph shown in detail in the top right of Fig. 2.5, the histograms of Poisson's ratio and density provide additional information about the property trade-offs that are achievable within the continuous region in the constrained alloy state space, represented discretely by the constrained subgraph $Alloy$ nodes. In this case, the range and distribution for the properties density and Poisson's ratio are shown for the nodal $Alloy$ compositions in the selected constrained subgraph. Analysis of the compositions within this subgraph also reveal that it only extends through portions of the FeNiCrTi quaternary, and does not include any compositions containing Co.

These constrained subgraphs can be recomputed based on different potential design constraints, which can be repeated to iteratively determine the best problem formulation that does not overconstrain the design space. If targeted properties or compositions are not present within the same constrained subgraph, this can motivate the addition of more alloying components in the ATLAS graph or modifying the constraints. This constrained subgraph partitioning and analysis enables the consideration of continuously manufacturable regions of the alloy space as distinct entities. Essentially, each of these subgraphs is a distinct design space for CGAs. This scalable, topologically interpretable approach is a critical advancement for the design and manufacturing of CGAs. These distinct design spaces can now be partitioned in any dimension, and their characteristics can be visualized and analyzed for systematic materials design.

2.7 Chapter Summary

The proposed Alloy Topology-Linked informAtion Schema (ATLAS) addresses the key requirements identified from the literature for an informatics schema to design compositionally graded alloys. Using a labeled property graph data structure, ATLAS captures the topology of the alloy state space in an ATLAS materials graph of $Alloy$ nodes and $JOINS\_TO$ edges, which can be further interpreted by applying subsequent nodal filters based on materials design requirements. In particular, the proposed schema enables the identification of constrained subgraphs, delineating separate regions of the constrained alloy state space that can be evaluated in isolation as design spaces for continuous compositional gradation. These subgraphs also allow a designer to identify if a given set of terminal alloys is compatible before expending the effort to identify a specific compositional gradient. Additionally, this schema can flexibly integrate additional data, serving as a living repository of materials data that can be refined over time with data from various sources including experimentation and improved models and simulation. This schema can be applied to small-scale problems in local memory using the ATLAS materials graph, and deployed for the storage and use of expansive datasets using an ATLAS graph that incorporates all aspects of the schema for improved scalability. On its own, this first-of-its-kind schema enables significant insight into the alloy state space for all types of alloy design, enabling alloy data storage and refinement over time to support diverse workflows. Furthermore, ATLAS serves as the basis upon which this work proposes methods for multi-terminal CGA design as well as material-to-structural mapping for arbitrarily complex CGAs and structures in the upcoming chapters.

3. MULTI-TERMINAL GRADIENT DESIGN USING GRAPHS*

While prior work formulated the design of CGAs as a two-terminal path planning problem [12, 64, 60, 63], the challenge of designing multi-terminal compositional gradients with more than two terminal alloys has not been considered. Although piecewise two-terminal path planning could be used to assemble a multi-terminal gradient, this naive procedure is eclipsed by heuristic approaches specific to multi-terminal route planning, which is typically considered as a minimum Steiner tree problem. Implementing a graph representation of the alloy state space enables the novel application of graph algorithms to address this design challenge by solving the minimum Steiner tree problem in graphs (SPG). This chapter introduces a novel formulation of the multi-terminal gradient design problem as an SPG alongside an ATLAS materials graph. Then, existing solvers are discussed, and a common heuristic approximation method is selected for its availability and compatability with the widespread NetworkX Python library [7] for graph data processing. Thus, this library can be used for processing ATLAS graphs as well as performing multi-terminal gradient design using the minimum Steiner tree problem formulation. The capability of this formulation for multi-terminal CGA design is demonstrated in an ATLAS materials graph of a real 6-element alloy state space. Then, the use of minimally spanning trees for designing CGAs for experimental assessment is introduced, and a computational example is used to illustrate a spanning CGA and corresponding physical mapping to an experimental sample using a real materials dataset.

3.1 Multi-terminal Gradient Design

Using a graph representation of the alloy state space enables the computational design of CGAs between more than two terminal alloy compositions simultaneously. Specifically, graph trees containing an arbitrary number of terminal alloys can serve as a representation of multi-terminal CGAs. Designing these trees enables simultaneous consideration of all terminal alloy compo-

*Contains material reprinted with permission from M. Allen, R. Arróyave, and R. Malak, "A Graph Algorithm for the Design of Functionally Graded Metal Components," *Volume 3A: 50th Design Automation Conference (DAC)*, Aug. 2024. [1] and M. Allen, R. Arróyave, and R. Malak, "A Graph Database Schema for Metal Additive Informatics," *Volume 2B: 44th Computers and Information in Engineering Conference (CIE)*, Aug. 2024. [2].

sitions, which can be seen as a generalization of two-terminal path planning. In the context of graphs, the problem of optimizing a tree containing a given set of terminal nodes (here, terminal nodal $Alloy$ composition) and a cost function is referred to as the minimal Steiner tree problem in graphs [76]. By formulating the multi-terminal CGA design problem as a minimum Steiner tree problem in graphs, we can formally design optimal multi-terminal CGAs.

3.2 The General Steiner Tree Problem in Graphs

The Steiner tree problem in is a classic NP-hard problem, and one of the most studied problems in combinatorial optimization [123, 124]. The SPG is a generalization of two other classic optimization problems of 1) shortest paths and 2) minimum spanning trees, and can be defined as follows: given an undirected graph as a pair of nodes/vertices and edges, $G = (V, E)$, a set of terminal nodes $V_T \subseteq V$, and a positive cost function $c(v_j, v_k) \forall (v_j, v_k) \in E$, find the connected subgraph of G with the minimum cost that contains all terminal nodes V_T. This can be formulated as an optimization problem shown below:

$$
\begin{aligned}
\text{Minimize:} \quad & \sum_{(v_j, v_k) \in E} c(v_j, v_k) \cdot q_{jk} \\
\text{Subject to:} \quad & q_{jk} \in \{0, 1\} \quad \forall \, (v_j, v_k) \in E \\
& \sum_{v \in \tau} q_{jk} \geq 1 \quad \forall \, v_k \in V_T \\
& \tau \subseteq G
\end{aligned}
$$

The objective is to minimize the total cost c of connecting the terminal nodes through a tree τ, which is an acyclic subgraph of the graph G, where q_{vw} is a decision variable that takes binary values (0 or 1) and indicates whether an edge (v, w) is included in the tree. The total cost is the sum of costs associated with the selected edges for the tree. Each terminal alloy composition must be connected in the tree by requiring at least one edge existing between the nodes in τ and each

terminal node in V_T. By solving this problem, a Steiner tree τ_{min} thus connects all specified terminals with the minimum required cost. The goal is to find the Steiner tree τ_{min} which minimizes the chosen cost metric while connecting the specified terminals.

3.3 Multi-terminal Gradient Design Problem Formulation

This Steiner tree approach can be seen as a graph-based multi-terminal generalization of prior two-terminal path planning approaches to the computational design of CGAs. In this specific application for CGA design, the ATLAS materials graph $G_{\text{AM}} = (V_A, E_J)$ has $Alloy$ nodes, V_A, and $JOINS\_TO$ edges, E_J, represent the alloy state space. This materials graph can be filtered for feasibility based on manufacturing and design constraints such that only the graph of feasible $Alloy$ nodes V_{A-F} and their shared $JOINS\_TO$ edges E_{J-F} remain in order to represent the feasible region within the alloy state space.

To formulate a multi-terminal gradient design problem, the SPG can then be defined in terms of the feasible, constrained ATLAS materials graph, $G_{\text{AM-F}} = (V_{A-F}, E_{J-F})$. Here, the prescribed terminal nodes $V_T \subseteq V_{A-F}$ can be defined as the terminal nodal $Alloy$ compositions to be joined in a CGA. The positive cost function $c(v_j, v_k) \; \forall \; (v_j, v_k) \in E_{J-F}$ can take a variety of forms, including cost functions used in two-terminal CGA design literature such as compositional distance [12], property monotonicity [60], and distance from constraint regions [12, 63, 64].

The multi-terminal gradient design problem can then be similarly defined as an SPG optimization problem:

$$\text{Minimize:} \quad \sum_{(v_j,v_k)\in E_{J-F}} c(v_j, v_k) \cdot q_{jk}$$

$$\text{Subject to:} \quad q_{jk} \in \{0,1\} \quad \forall\, (v_j, v_k) \in E_{J-F}$$

$$\sum_{v\in\tau} q_{jk} \geq 1 \quad \forall\, v_k \in V_T$$

$$\tau \subseteq G_{\text{AM-F}}$$

Here, the objective is to minimize the total cost c, such as Euclidean compositional distance c_{d}, of connecting the terminal nodal $Alloy$ compositions V_T through a tree τ, which is an acyclic subgraph of $Alloy$ nodes and $JOINS\_TO$ edges from the constrained materials graph $G_{\text{AM-F}}$. The total cost is the sum of the $JOINS\_TO$ edge costs in the tree τ. The resultant Steiner tree represents a multi-terminal gradient optimized for all terminal alloy compositions. Figure 3.1 is a dimensionally-reduced projection of a Steiner tree for 5 terminal $Alloy$ nodes in a feasible ATLAS materials graph of the 6-element HfMoNbTaVW alloy state space filtered based on CALPHAD-predicted phases and solidification range. Note that the compatability of the terminal alloy compositions can be predetermined by assessing their membership in the constrained subgraphs, which would be a time consuming process of trial and error using prior methods.

Arguably, the most straightforward cost function is Euclidean compositional distance, c_{d}, that optimizes the gradient to minimize the amount of intermediate composition layers required for manufacturing. This is accomplished by minimizing unnecessary compositional variation along the path. While barycentric coordinates are often used for alloy state spaces, the compositions should be represented in terms of all Euclidean elemental bases to enforce consistent lengths across the alloy state space when computing this cost function. In this computation, any balance element coordinate must be included to avoid skewing the compositional distances. This computation is functionally equivalent for Euclidean and barycentric coordinates if the simplicial vertices are

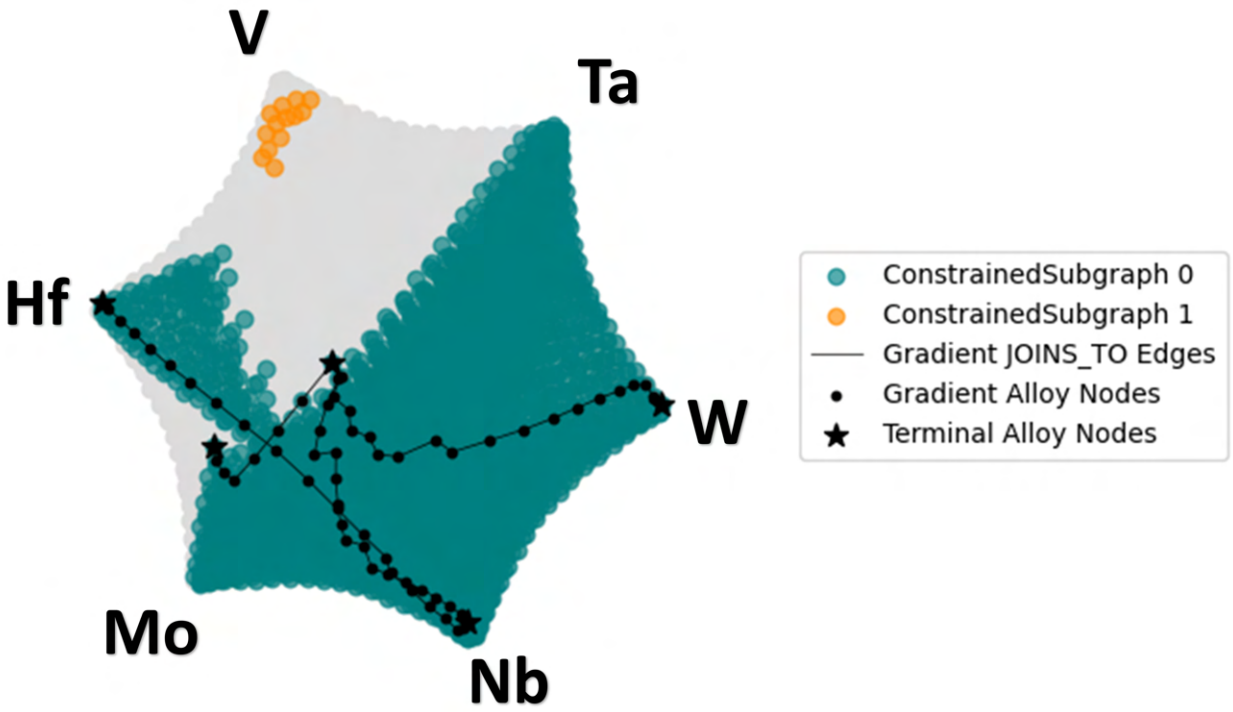

Figure 3.1: UMAP projection of a multi-terminal compositional gradient design problem, with a solution acquired using the SPG for five arbitrary terminal alloys in a feasible ATLAS materials graph, $G_{\text{AM-F}}$ for the 6D HfMoNbTaVW alloy state space.

pure elements, which is not always the case. The Euclidean compositional distance cost c_{d} can be computed for all edges $e_i \in E_J$ as:

$$c_{\text{d}}(e_i) = \|\mathbf{x_j} - \mathbf{x_k}\|_2 \tag{3.1}$$

where c_{d} is the weight for an arbitrary graph edge $e_i = (j, k) \in E$, x_j and x_k are the elemental Euclidean compositions at the $Alloy$ nodes v_j and v_k respectively, and $\|\cdot\|_2$ represents the Euclidean norm.

3.3.1 Solver Algorithms

Due to the wide significance of the SPG across many fields, a multitude of solver algorithms have been proposed with varied performance characteristics and solution quality. The development of exact algorithms is, notably, still limited, with an exact solver developed by Polzin & Vahdati-

Daneshmand [125, 126] in the early 2000s remaining largely unchallenged throughout the efforts of the last two decades [124]. As a result, current exact methods are not able to solve as many instances of the SPG to optimality, nor are they able to find solutions as efficiently when compared to the solver by Polzin & Vahdati-Daneshmand. Unfortunately, this solver is not publicly available.

Meanwhile, there are a variety of heuristic techniques [77, 78, 127] to efficiently solve for a Steiner tree which are not exact. These methods typically have a bounded approximation within some factor of the cost for a true Steiner tree. One widespread heuristic algorithm for the SPG by Kou, Markowsky, & Berman [77] guarantees a solution with a total weight no more than $2\left(1-\frac{1}{l_{\text{min}}}\right)$ times that of a Steiner minimal tree, where l_{min} is the minimum number of leaves in the optimal tree. Note that for the case with two terminal nodes, $l_{\text{min}} = 2$, the total weight is equal to the minimal weight because the two-terminal problem is a minimum path search, which can be solved to optimality. This solver has a worst-case time complexity of $O(|V_T||V|^2)$. Later, an implementation of this algorithm by Mehlhorn improved the running time to $O(|E| + |V|log|V|)$ [78].

Mehlhorn's solver implementation begins with a connected undirected weighted graph $G = (V, E, W)$ and a set of terminal nodes $V_T \subseteq V$. Then, an undirected weight graph $G_1 = (V_1, E_1, W_1)$ is constructed from G such that $V_1 = V_T$ and that for every edge edge $v_{1,j}, v_{1,k} \in E_1$, the corresponding edge weight $w_{jk} \in W_1$ equals the total cost of the minimum cost path from v_j to v_k. The solver algorithm then proceeds by finding the minimum spanning tree (MST) τ_1 of G_1. Then, G_S is created by replacing all edges in τ_1 with the corresponding path of nodes and edges in the original graph G. Lastly, the MST of G_S is found, and, if necessary, edges are deleted from this tree until all the leaves L satisfy $L \subseteq V_T$, resulting in a Steiner minimal tree τ_{min}.

The Mehlhorn solver was selected for this work, including the demonstration in Fig 3.1, due to its widespread implementation and compatability with the common Python library NetworkX [7]. This enables the use of NetworkX for both processing ATLAS materials graphs and constrained subgraphs in memory, as well as solving for τ_{min}. A few critical characteristics of the output of this solver should be noted. First, for the leaf nodes L of the returned Steiner tree τ_{min}, the condition

$L \subseteq V_T$ is always true, where V_T is the set of terminal nodes. The second characteristic is that τ_{min} is always a tree, and, thus, there is only one path between any two $Alloy$ nodes in the Steiner tree. In the following chapter, a new algorithm will be introduced which leverages this tree structure to sequentially place CGAs of arbitrary complexity within discretized physical structures.

3.4 Gradient Design for Efficient Experimental Assessment

The concept of multi-terminal gradients as graph trees can be taken even further for experimental assessment of the alloy state space with CGAs. For this application, *spanning trees* could be a useful method for materials design. Fundamentally, a tree τ in a given graph G is a *spanning tree* if τ 1) does not contain any cycles, and 2) contains all nodes in G. If G is an edge-weighted graph, then a tree τ is a *minimum spanning tree* (MST) if the cost (sum of edge weights) of τ is the minimum cost of all spanning trees.

In the context of a constrained subgraph $G_{\text{CSG}} = (V_{\text{CSG}}, E_{\text{CSG}})$ in an ATLAS materials graph G_{AM}, a spanning tree or MST of G_{CSG} would capture all of the subgraph's nodal $Alloy$ composition within a tree. This tree would contain all sampled $Alloy$ compositions from a targeted region of the alloy state space in a unified multi-terminal 'spanning gradient'. Here, for a given constrained subgraph $G_{\text{CSG}} = (V_{\text{CSG}}, E_{\text{CSG}})$, we define a spanning gradient as a connected subgraph $G_{\text{span}} = (V_{\text{span}}, E_{\text{span}})$ with $V_{\text{span}} = V_{\text{CSG}}$. The alloy state space could be assessed for specific phases and/or properties by constraining an ATLAS materials graph based on these conditions and using spanning gradients. Experimental assessment can be further targeted on the set of nodal $Alloy$ compositions on or near the boundary of a constrained subgraph. As discussed prior, constraint uncertainty is a significant challenge due to the inherent uncertainty of both computational modeling and compositional control during manufacturing, especially when alloying in-situ [63, 83].

There are several well known greedy algorithms for computing MSTs in polynomial time [75], namely Kruskal's algorithm [8] and the Jarník-Prim algorithm [128, 129]. The Jarník-Prim algorithm begins at an arbitrary node v_{start} from a given connected graph, initializing a list of nodes that have been visited. It proceeds to the node along the lowest-cost edge from v_{start}, adding this node to the visited list. The algorithm proceeds sequentially by selecting the lowest cost edge from

the aggregated nodes where the other terminal is unvisited (so that a cycle is not created) until no edges to unvisited nodes remain. Meanwhile, Kruskal's algorithm begins with all graph nodes, and sequentially adds the lowest cost edge which does not create a cycle. This procedure is repeated $(n-1)$ times, where n is the number of nodes in the graph. This is logical since trees possess $(n-1)$ edges by definition [75].

With the spanning gradient represented as a tree, which is a type of planar graph, the material gradient can be directly embedded into the plane such that the edges do not overlap, but only intersect at their shared nodal terminals. By defining volumetric dimensions for this planar gradient representation, one could design a corresponding branched arboreal structure to manufacture all nodal $Alloy$ composition in the tree with metal AM for experimental assessment. Alternatively, the CGA could be partitioned into smaller trees which can be manufactured as simpler arboreal structures, or two-terminal gradient paths which can be fabricated as prismatic test samples. Partitioning into smaller trees or paths would be particularly useful when the spanning gradient is extremely large. These arboreal or prismatic samples could then be used characterize targeted topological regions of the alloy state space in any number of dimensions with more experimental efficiency than sequential single material trials. Since the objective here is to assess the alloy state space, additional cost functions could be considered, such as the inverse of the Euclidean compositional distance in order to maximize the compositional distance traversed through the state space to select the $JOINS\_TO$ edges which traverse as many compositions as possible. Figure 3.2a shows a minimal spanning tree of the second constrained subgraph for the Hf-Mo-Nb-Ta-V-W alloy state space using this inverse Euclidean composition distance as the cost function. Meanwhile, Figure 3.2b illustrates a 2D embedding of the spanning CGA tree mapped to an exemplary arboreal structure that could be manufactured to assess the feasible region represented by the constrained subgraph.

3.5 Chapter Summary

In this chapter, the graph representation from ATLAS was used as the groundwork for defining the multi-terminal CGA problem as a Steiner tree problem in graphs. This problem formulation

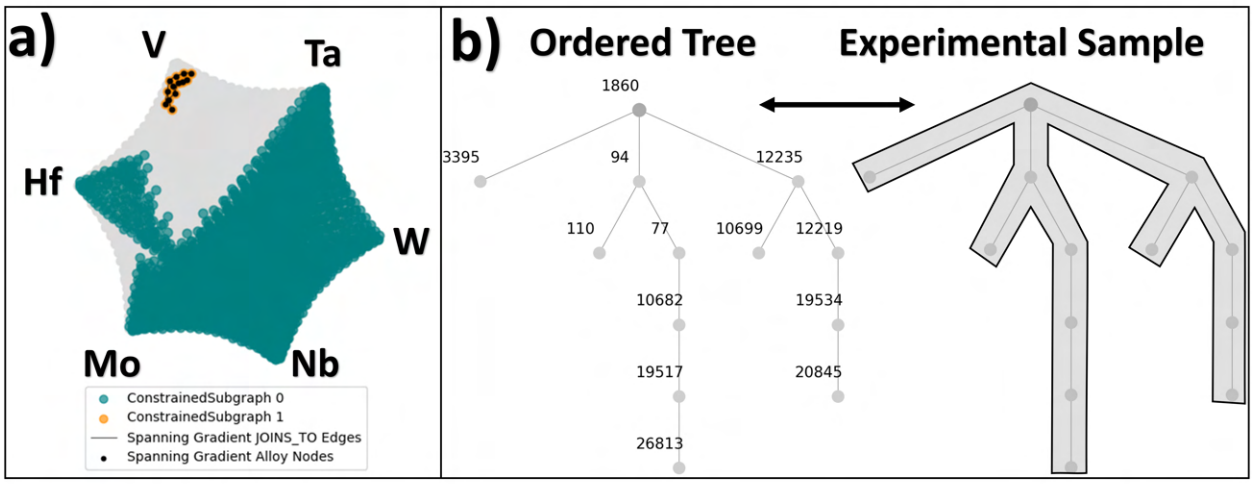

Figure 3.2: Illustration of a) a UMAP projection of the MST for the second largest constrained subgraph in a feasible ATLAS materials graph of the Hf-Mo-Nb-Ta-V-W alloy state space. Meanwhile, b) shows a corresponding 2D embedding of the MST labeled by global node index in the ATLAS materials graph and rooted at the highlighted $Alloy$ node (1860). The arrow indicates mapping the CGA to a physical arboreal structure which could be use to asses the region of the alloy state space represented by the second constrained subgraph. This experiment could be conducted with a single build instead of 13 single-alloy builds for each nodal $Alloy$ composition in the tree, which would also not indicate information about the interfaces between the compositions. The NetworkX [7] implementation of Kruskal's algorithm [8] was used to identify the MST.

enables the optimization of a CGA for a given objective function simultaneously based on all terminal nodal $Alloy$ compositions, unlike naive approaches such as piecewise path planning. In this context, a CGA is represented by a graph tree extracted from a relevant ATLAS materials graph using the proposed formulation. An accessible and straightforward heuristic solver [78] was discussed and used on a straightforward material example, while recognizing the ongoing work to develop general and accessible exact solvers. Here, the heuristic solver was used on real materials data with five terminal alloy compositions, which also provided an example of the output representation of the CGA as a graph tree. This tree representation of CGAs further lays the groundwork for systematic materials design, with the remaining gap being the capability of mapping from the CGA materials onto a corresponding physical structure with quantifiable performance characteristics.

The notion of gradients as graph trees was further extended to the experimental assessment of the alloy state space. Particularly, minimum spanning trees (MSTs), essentially a special type of minimum Steiner tree, include every composition in a given constrained subgraph into a single gradient represented as a graph tree. The structural graph property that trees are also always planar graphs was used to define a straightforward mapping from the material gradient tree structure to an arboreal structure in physical space, which could be used for efficient experimental assessment of the alloy state space compared to experiments focused on single material test samples. On its own, this novel formulation of the CGA design problem enables the design and representation of multi-terminal CGAs, extending far beyond the current literature which essentially focuses on two-terminal CGAs exclusively. This formulation will be used to design CGAs with arbitrarily many terminals, which lays the groundwork for mapping between CGAs and structures of arbitrary complexity.

4. AN ALGORITHM FOR THE DESIGN OF STRUCTURES WITH MULTI-TERMINAL COMPOSITIONALLY-GRADED ALLOYS*

It is critical to link the design of CGAs with the design of corresponding CGA structures for systematic materials design. In the previous chapter, the formulation of multi-terminal CGA design as a minimum Steiner tree problem in graphs was proposed and demonstrated. One specific property of graph trees is particularly notable in this context; within a graph tree, there is only one path between any given pair of nodes. Thus, there are specific subsets of shared compositions where different two-terminal CGA paths within the tree overlap. This property can be leveraged for mapping these multi-terminal CGAs into corresponding physical structures. Thus, this chapter presents a novel graph algorithm to perform a mapping procedure between a multi-terminal CGA, represented as a graph tree, and a corresponding discretized 3D structural geometry, linking the materials design of CGAs with corresponding structures. In this chapter, the fundamental design problem is described, and the inputs, outputs, assumptions, and constraints for the proposed algorithm are outlined. Then, the proposed graph algorithm is introduced, and a simple synthetic demonstration is provided to illustrate its operation. Then, the algorithm's soundness, and completeness are discussed. Next, a demonstration with a real alloy system and practical structural application is shown to illustrate the use of the proposed algorithm in engineering practice. Lastly, the scalability and performance characteristics of the proposed algorithm are discussed.

4.1 Problem Definition

The objective for this algorithm is to design a CGA structure through controlled spatial mapping of the material gradient into the physical domain. The resultant structure should be manufacturable, meaning that the adjacent compositions in the part should be compatible as indicated by their nodal $Alloy$ adjacency via $JOINS\_TO$ edges in the corresponding ATLAS materials

*Contains material reprinted with permission from M. Allen, R. Arróyave, and R. Malak, "A Graph Algorithm for the Design of Functionally Graded Metal Components," *Volume 3A: 50th Design Automation Conference (DAC)*, Aug. 2024. [1].

graph and, consequently, their adjacency in a multi-terminal CGA represented as a graph tree. This section will detail the necessary inputs, outputs, constraints, and assumptions for the proposed algorithm to perform controlled spatial CGA mapping to design a valid CGA structure as described.

4.1.1 Inputs

Due to the acyclic nature of graph trees, there is only one simple path (where all vertices in the path are distinct) between any given pair of nodes in a tree. This property can be leveraged to procedurally map a CGA, given as a graph tree $\tau_{\text{CGA}} = (V_{\text{CGA}}, E_{\text{CGA}})$ with an arbitrary number of nodal $Alloy$ terminals $V_T \subseteq V_{\text{CGA}}$, into a discretized structural geometry. This graph tree representation of a CGA, τ_{CGA}, can be designed using the SPG formulation of the multi-terminal CGA design problem in a constrained ATLAS materials graph, $G_{\text{AM-F}}$, such that the resulting $\tau_{\text{min}} = \tau_{\text{CGA}}$. Here, τ_{CGA} is the CGA between the terminal nodal $Alloy$ compositions used for the SPG. This CGA is optimized according to a prescribed cost function over the $JOINS_TO$ edges used as the edge weights for the SPG.

Since the primary design consideration in CGAs is the controlled spatial variation in alloy composition, a mapping algorithm between the material and physical domains must enforce such a variation of composition across a corresponding part volume. To accomplish this systematically, a volumetrically-discretized part is used to assign material compositions in distinct spatial subregions of the part volume. The algorithm must traverse this discretization to apply the material gradient based on the required step-wise material adjacency represented in τ_{CGA}. While any volumetric discretization method can be used, it is ideal for the discretization to relate to the spatial resolution of the manufacturing method being used, such as a voxelization based on the layer thickness and hatch spacing in LP-DED. Alternatively, if conformal AM tool paths will be used for fabrication, the structure could be correspondingly discretized in part using conformal surfaces with relevant offset steps.

For valid CGA placement, the alloy composition assigned to each distinct volumetric cell, and the adjacency between cells considering their alloy composition is critical. Since the volumetri-

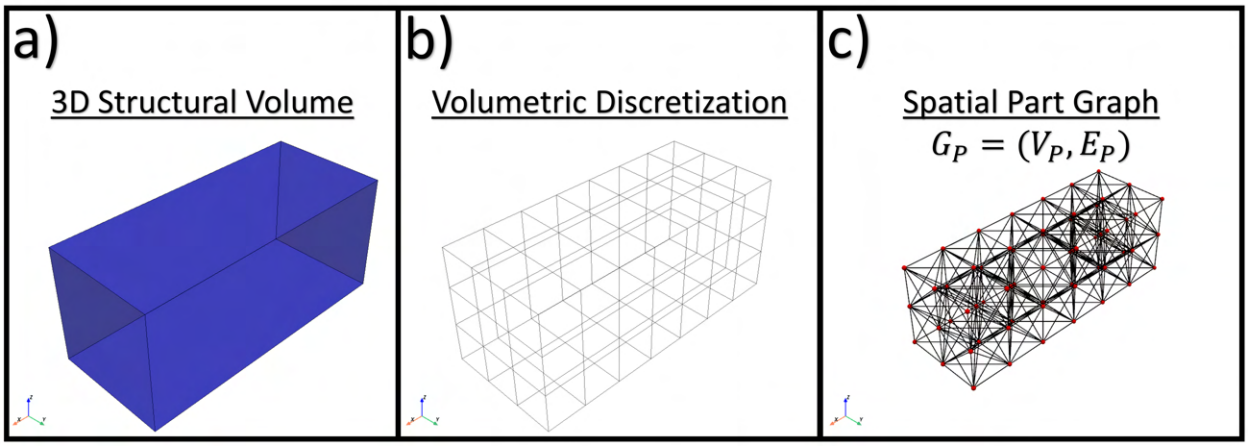

Figure 4.1: Illustration of the spatial part graph creation process for an example 3D structure with dimensions $(X = 7, Y = 3, Z = 3)$ showing a) the initial 3D structural volume, b) a volumetric discretization of the structure into a tetrahedral voxelization/mesh with a spacing of one unit along each axis, and c) the spatial part graph $G_P = (V_P, E_P)$ with cell nodes V_P corresponding to the discrete cells, and edges E_P for nodal cell pairs that share a vertex in the voxelization/mesh. Here the voxels are plotted at the cell center coordinates of their corresponding discrete volumetric cell.

cally discretized structural representation will be traversed to place a compositional gradient, it can be represented by a spatial part graph $G_P = (V_P, E_P)$ to enable efficient traversal procedures. This can be done by defining a node for each discrete volumetric cell, $v_{P,i} \in V_P \; \forall \; i \in [1, 2, ..., n]$, with edges, $e_{P,j} \in E_P \; \forall \; j \in [1, 2, ..., m]$, between the nodes for physically adjacent cell pairs. The notion of physical adjacency is dependent on the discretization method used, but as an example, for a tetrahedral voxelization/mesh, cells which share a vertex could be considered physically adjacent. This notion of taking a 3D structural model and using it to create a volumetrically-discretized structure with a corresponding spatial part graph is illustrated in Fig. 4.1. With these inputs in place, there is now a gradient representation, τ_{CGA}, capable of representing an arbitrarily complex multi-terminal gradient, as well as a representation for efficient traversal of a corresponding, discretized physical structure, $G_P = (V_P, E_P)$.

The remaining information critical for placing the CGA is where the process should start and end. First, the starting point(s) of the gradient mapping process are defined by labeling cells in the discretized structure and their corresponding spatial cell nodes $v_{P,i} \in V_P$ with any number

of the materials in τ_{CGA}. Typically, this will be the terminal alloys. Since the spatial placement of the terminals is dependent on one another and the necessary gradient steps for joining, it is convenient to define the spatial placement of $(N_T - 1)$ of the N_T terminal $Alloy$ compositions $v_T, i \in V_T \, \forall \, i \in [1, 2, ..., N_T]$ in the initial labeling. The placement of the unused terminal will then be established in the gradient placement process. Note that only up to one terminal composition should be left unlabeled for the terminals that are also leaf nodes in τ_{CGA}; if an unused terminal is a leaf node, there is no reason for the algorithm to leverage the branch of the gradient tree necessary for linking that alloy with the others if it is not placed initially. The reasoning for using at least $(N_T - 1)$ terminal alloys for labeling is described further in the following paragraph.

The final missing piece of information is an order to the mapping procedure in conjunction with the initially labeled discretized structure and the corresponding gradient tree τ_{CGA}. For this procedure, we introduce a so-called 'coalescent material' parameter, m_c. This parameter is used to re-root τ_{CGA} with the tree's nodal $Alloy$ compositions ordered by the number of $JOINS\_TO$ edge steps from the root m_c, in ascending order. The result is an ordered gradient tree τ_{CGA} with m_c called the coalescent material because the placement of this $Alloy$ is last, resulting in the final coalescence of the CGA in the physical domain. If $(N_T - 1)$ terminal alloys are used in the initial labeling of G_P, and the unused terminal node is a leaf node, then m_c must be set to the unused terminal nodal $Alloy$ composition to ensure it is placed in the gradient labeling process. To reiterate, leaving one terminal alloy unused in the initial labeling is convenient since the spatial placement is dependent on the location and distribution of the other terminal $Alloy$ compositions, the discretization, and the required gradient steps between these terminals in τ_{CGA}. Thus, leaving one terminal alloy out of the initial labeling leaves it dependent on the fixed placement of the other terminal compositions and the various gradient steps required for the multi-terminal CGA. This approach of using a terminal alloy as m_c is favored due to the assumption that the use of the terminal alloys should be prioritized. Thus, when used in this way, m_c is essentially the dependent terminal alloy; TreeMAP first places the required gradient layers with the remaining discrete cells being labeled with m_c. This concept will be illustrated later in this chapter.

Overall, these inputs result in a ordered sequence of gradient compositions in τ_{CGA} from the leaves to the root m_c, an efficient means of traversing physically adjacent volumes in the part structure with G_P, and a defined start for placing the CGA given by the initial labeling of $Alloy$ compositions in the discretized structure and corresponding G_P.

4.1.2 Output

The desired output is a valid compositionally graded structure. Here, this will be returned as a fully-labeled spatial part graph G_P, with material labels that can be mapped back to the corresponding cells in the discretization to achieve a manufacturable CGA structural design. Again, in this case, a valid structure denotes that adjacent pairs of spatial cell nodes, $(v_{P,a}, v_{P,b}) \in E_P$, should be labeled with either identical or adjacent nodal $Alloy$ compositions from τ_{CGA}. If the resultant part respects the adjacency of τ_{CGA}, then the output can be considered valid, since the compositions $V_{\text{CGA}} \in \tau_{\text{CGA}}$ pass the feasibility constraint(s) used for filtering the ATLAS materials graph for a given design problem.

4.2 Tree-based Material Adjacency Propagation Algorithm

We now have 1) an ordered sequence of compositions in τ_{CGA} from the leaves to the root m_c, 2) an efficient representation for traversing physically adjacent volumetric cells in the part structure with G_P, and 3) a defined start for placing adjacent gradient compositions given by the initial labeling of $Alloy$ compositions in the discretized structure and corresponding G_P. With these inputs, the following Tree-based Material Adjacency Propagation (TreeMAP) algorithm is proposed to map the CGA into the physical structure. This algorithm leverages the convergence of all pairwise paths through the tree structure to unify the initially-placed compositions based on the multi-terminal gradient in τ_{CGA}. Specifically, this is accomplished by stepping through the ordered tree in ascending order, skipping the leaf nodes since they are terminal compositions which must be initially defined within the structure. The algorithm starts with the farthest non-leaf $Alloy$ node from m_c in τ_{CGA}. Then, the algorithm finds the child nodal composition(s) of the current node in the ordered tree τ_{CGA}. Next, it finds the set of all cells labeled with the child composition(s),

then finds the set of all unlabeled cells neighboring that set. These unlabeled neighbors are labeled with the composition of the current nodal composition from τ_{CGA}. This process repeats until the final composition in the ordered tree, m_c, is reached. Since there may be additional unused regions within the structure, all remaining unlabeled voxels are labeled with m_c. This procedure for the proposed TreeMAP algorithm is shown in detail as pseudocode in Algorithm 1.

Algorithm 1: TreeMAP

Data:

- $\tau_{\text{CGA}} = (V_{\text{CGA}}, E_{\text{CGA}})$: Ordered compositional gradient tree rooted at m_c.
- $G_P = (V_P, E_P)$: Spatial part graph representing the discretized structure.
- Initial labeling of alloys from τ_{CGA} in G_P (typically terminal alloys, excluding m_c if a terminal alloy is used as m_c).

Result: Fully-labeled spatial part graph G_P with valid CGA placement.

foreach *Alloy node* $v_{CGA} \in V_{CGA}$ *in ascending order of depth* ***skipping leaf nodes*** **do**
  Let $V_{\text{CGA, child}}$ be the set of children nodes of parent of v_{CGA} in τ_{CGA} Identify all spatial graph nodes currently labeled with the compositions of $V_{\text{CGA, child}}$, $V_{P,L} \subset V_P$;
  $V_{P,U} \leftarrow \texttt{GetUnlabeledNeighbors}(V_{P,L}, G_P)$;
  Assign the material label given by v_{CGA} to $V_{P,U}$
$V_{P,U} \leftarrow \texttt{GetUnlabeledNodes}(G_P)$;
Label $V_{P,U}$ with the composition m_c;
return Fully-labeled spatial part graph G_P;

In summary, this algorithm enforces the material adjacency required to map the materials domain into the physical domain to implement a valid compositionally graded structural design. Using a graph representation of the discretized part geometry enables quick graph traversal for identifying adjacent, unlabeled volumetric cells and propagating the adjacent gradient composition per the gradient, τ_{CGA}.

4.3 Synthetic 2D Example

To illustrate the operation of this algorithm, a simple 2D example is provided in Fig. 4.2, showing the steps of the proposed TreeMAP algorithm. Here, an initial, discretized and labeled 2D structure and CGA tree τ_{CGA} have been defined. For simplicity, the spatial part graph for the structure is not shown. Also, note that in this example, volumetric cells sharing a vertex are considered adjacent in G_P. First, the $Alloy$ node 'E' is skipped since it is a leaf node. Then, for the node 'D', the only child is 'E'. All unlabeled neighboring cells of spatial part nodes labeled with 'E' are found, and labeled as 'D'. Next, 'C' is skipped because it is a leaf node. Now, for 'B', there are two child nodal $Alloy$ compositions, 'D' and 'C'. The unlabeled neighboring cells of 'D' and 'C' are found, and labeled with material 'B'. Lastly, the remaining unlabeled cells are found and labeled with $m_c =$ 'A'. While not used in the initial labeling, $m_c =$ 'A' is placed in the structure utilizing all remaining cells after placing the required gradient compositions ('D' and 'B'). The specific placement of 'A' is a direct result of both the initial placement of the other terminal nodal $Alloy$ compositions, and the subsequent gradient materials required by τ_{CGA}. The resultant solution provided by TreeMAP is clearly valid as all adjacent cells in the structure contain either the same material, or an adjacent material from τ_{CGA}.

4.3.1 Assumptions

The first assumption for the TreeMAP algorithm is that it can not alter the volumetric discretization of the structure. This is due to the assumption that the discretization relates to the spatial resolution of the manufacturing method being used, such as a voxelization based on the layer thickness and hatch spacing in LP-DED as mentioned prior. As a result, TreeMAP altering the discretization could negatively impact the utility of a given design if it can no longer be manufactured with such an altered discretization. Thus, the discretization is determined by the user and treated as fixed by TreeMAP.

It is also assumed that the terminal alloys should be the primary compositions used, and, consequently, the proposed TreeMAP algorithm does not distribute the gradient across the available,

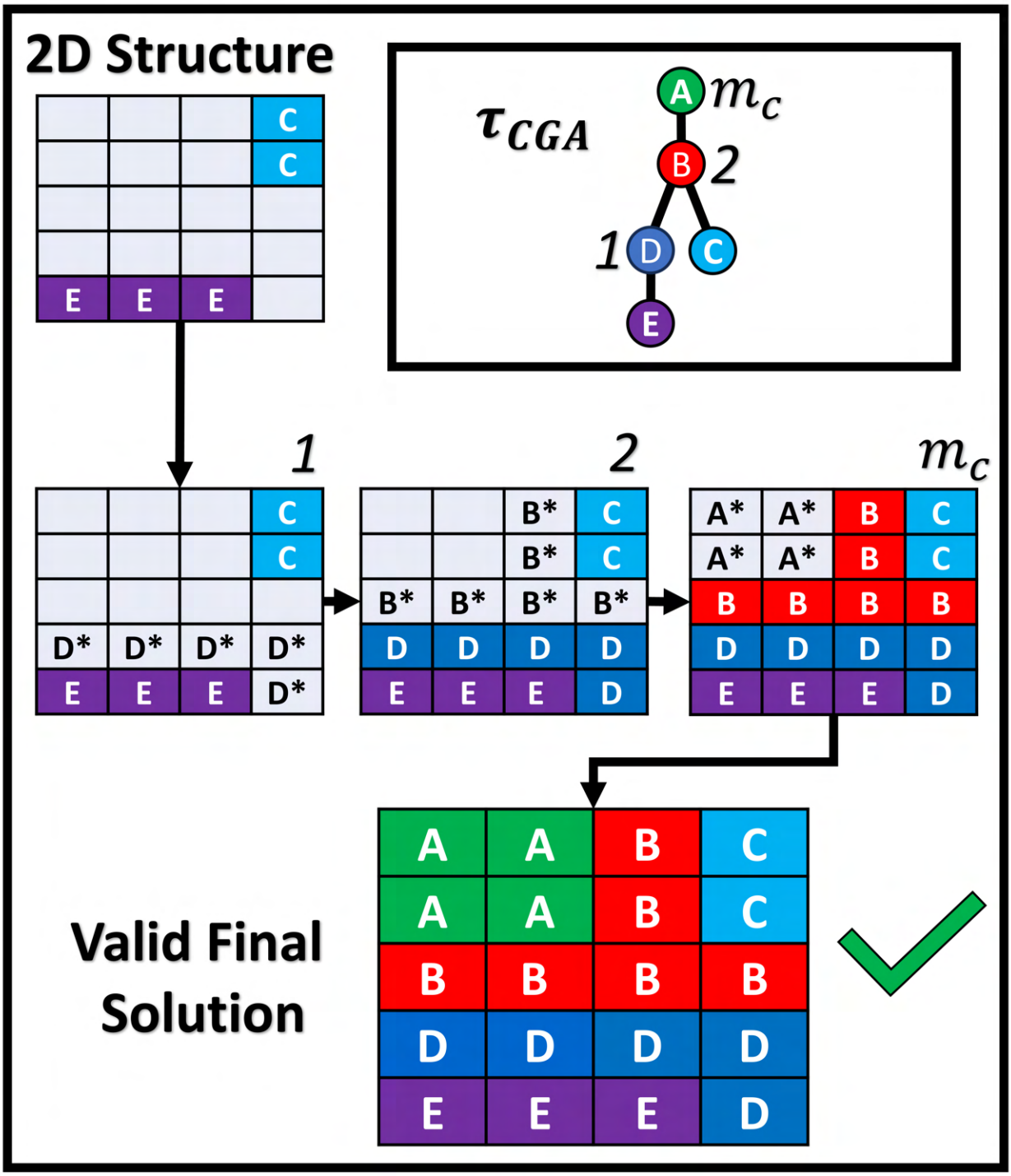

Figure 4.2: Synthetic 2D demonstration of the TreeMAP algorithm. The steps are illustrated directly in the discretized structure for simplicity. However, in practice, these material labels would first be applied in G_P, and then mapped back to the discretized structure after the TreeMAP algorithm is complete.

unlabeled cells. This is due to the fact that many gradients use conventional alloy terminals which have widely available, cost-effective feedstock, while achieving gradient compositions can be more costly. Additionally, in-situ alloying for intermediate compositions is less precise than depositing a single feedstock. Even if the terminals are custom-designed alloy compositions, they are typically selected due to superior properties in the alloy state space and should thus be prioritized over intermediate gradient compositions, which likely possess inferior properties by comparison.

If desired stepwise changes in composition or properties are desired, this can be achieved by post-processing the compositional gradient given by τ_{CGA}, adding, removing, or adjusting nodal alloy compositions as necessary prior to running TreeMAP. However, all adjustments should be done while considering the feasibility of the altered gradient for the design problem at hand, such as avoiding deleterious phases. Additionally, the altered CGA should still retain a graph tree structure.

Furthermore, due to the assumption of terminal alloy primacy, the remaining unlabeled cells at the end of the placement procedure are labeled with m_c. This allows the distribution of an unused terminal alloy composition (if one is used as m_c), and effectively allows the user to let TreeMAP find the largest volume for placing m_c in the discretized structure once the necessary intermediate gradient compositions are all placed.

4.3.2 Soundness & Completeness

Since the TreeMAP algorithm does not exert any control over the volumetric discretization. As a result, a given discretized and labeled structure must allocate sufficient discrete spatial steps to place the necessary corresponding discrete compositional steps between all nodal $Alloy$ compositions from τ_{CGA} initially placed/labeled in the structure and corresponding G_P. For example, if there are three intermediate compositions between two hypothetical alloy compositions x_α and x_β in τ_{CGA}, but, at minimum, there are only 2 spatial steps in G_P separating the subset of cells labeled x_α and the subset of cells labeled x_β, then the input is invalid. To alleviate this, the distance between compositions can be considered during the initial placement of compositions in the discretized structure.

To fully explain this constraint on the input partially-labeled G_P, let $S_{\textbf{part}}$ be the set of shortest unweighted paths in G_P between part node pairs $(k_i, w_i) \in V_P$ with different material labels, $k_{i,m} \neq w_{i,m}$. The paths for $S_{\textbf{part}}$ can be calculated efficiently using a parallelized path search algorithm such as Δ-stepping [130]. Meanwhile, with τ_{CGA} as the gradient tree with nodal alloy compositions $v_{\text{CGA}} \in V_{\text{CGA}}$, let $S_{\textbf{material}}$ be the set of shortest unweighted paths in G_{AM} between material node pairs (p_i, q_i) (where $p_i, q_i \in V_{\text{CGA}}$) corresponding to S_{part} such that $k_{i,m} = p_i, w_{i,m} = q_i$. Essentially, $S_{\textbf{material}}$ contains a material path from the gradient tree corresponding to the spatial path stored in $S_{\textbf{part}}$ between dissimilar material cells in G_P. Lastly, recall that a valid solution is a fully labeled part graph G_P where the material labels of neighboring nodes $V_P \in G_P$ correspond to the same or adjacent nodal $Alloy$ compositions in the gradient τ_{CGA}. Figure 4.3 illustrates this case where an invalid structure is given due to initial material labels being located too close together for the placement of the required intermediate compositions from the CGA tree. In general terms, the shortest number of discrete spatial steps between two materials in the initial structure should be greater than or equal to the number of material steps required between those two materials in the gradient τ_{CGA}. If this condition is violated and the input is not valid, it could be resolved by using a finer structural discretization (assuming a finer level of detail is realistic), labeling with nodal $Alloy$ compositions that require fewer intermediate steps, labeling regions that are further apart, or increasing the size of the intermediate region(s) of the structure.

TreeMAP iterates through a loop once for each $Alloy$ node in the gradient, $V_{\text{CGA}} \in \tau_{\text{CGA}}$. Each iteration traverses to the set of one-step neighboring unlabeled nodes $V_{P,U} \in G_P$ starting from the set of nodes with a given material label $V_{P,L} \in G_P$. These unlabeled neighbors are assigned the current material for the iteration, v_{CGA}. Since the algorithm proceeds in ascending order based on tree depth from the root m_c, there is no material for which an iteration is completed without a child node material from τ_{CGA} present at one or more part nodes in $V(P)$. Since every possible step is taken to the unlabeled neighbors of $V_{P,L}$, one step is made along each shortest path $s_{\textbf{part, i}} \in S_{\textbf{part}}$ at each iteration, one step for each $Alloy$ node in τ_{CGA}. Considering $S_{\textbf{material}}$, we can count the material nodes for each $s_{\textbf{material, i}} \in S_{\textbf{material}}$ as the number of steps made along each corresponding

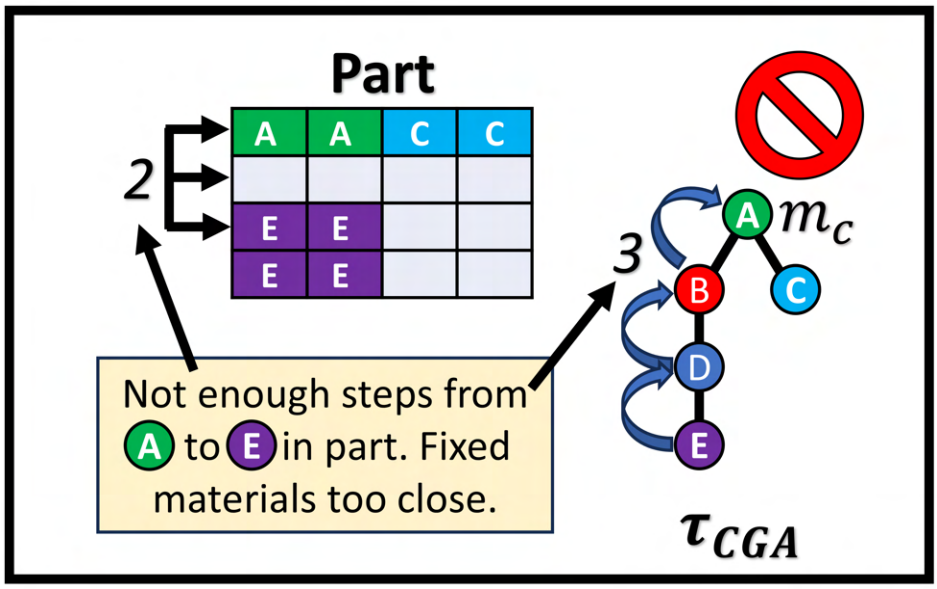

Figure 4.3: Illustrated example of the algorithm where the CGA steps and initial material labeling are incompatible. The G_P generated for this case would be considered an invalid input to the TreeMAP algorithm.

$s_{\textbf{part, i}} \in S_{\textbf{part}}$ throughout the algorithm. For a valid solution, as defined, the algorithm must at least place all material labels $s_{\textbf{part, i}}$ along the shortest corresponding part path $s_{\textbf{part, i}}$ between dissimilar materials. Since a) the number of material nodes for each $s_{\textbf{material, i}} \in S_{\textbf{material}}$ correspond to the number of steps made by the algorithm along each corresponding $s_{\textbf{part, i}} \in S_{\textbf{part}}$ over its course, one composition placed for each step of one neighboring, unlabeled spatial node $v_{P,U} \in V_{P,U}$, and b) a valid solution must obey the material adjacency in τ_{CGA}, then TreeMAP will produce a valid solution if and only if a valid input structure is given as G_P, where $S_{\textbf{part, i}} \geq S_{\textbf{material, i}} \; \forall \; i \in \{1, ...n\}$. Thus, the algorithm is sound under this condition.

Furthermore, if a valid input structure is passed with the corresponding G_P, then a solution exists. Fundamentally, TreeMAP only places the same or adjacent nodal $Alloy$ compositions from τ_{CGA} as it traverses G_P. Thus, in these cases with a valid input structure, TreeMAP will always return a valid solution. Thus, in addition to being sound, the TreeMAP algorithm is also complete.

4.4 Subsea Mechanical Connector Case Study

4.4.1 Problem Setup

To illustrate the practical use of this algorithm, we applied it to design a multi-material mechanical connector for subsea oil and gas production systems as a demonstration and validation step. The geometry for this design problem is a stylized 3D mechanical connector, modeled using CAD software (Solidworks® version 2024), resulting in the STL shown in Fig. 4.7a. The design goals in this case study are to balance the needs for a) corrosion resistance to external seawater, b) wear and corrosion resistance to internal chemicals and multi-phase flows, and c) minimizing the unnecessary use of costly materials. In light of these design goals, we identified the three following relevant (approximated) alloy compositions in the Fe-Ni-Cr-Nb-Mo quinary alloy system:

1. $Fe_5Ni_{60}Cr_{20}Mo_{15}$ [mole %], Hastelloy C-276
2. $Fe_{10}Ni_{70}Cr_{15}Nb_5$ [mole %], Inconel 718 (IN718)
3. $Fe_{70}Ni_{15}Cr_{15}$ [mole %], Stainless Steel 316

4.4.2 Topological Analysis

Using ATLAS, we can determine if these terminal compositions are compatible before attempting to design a specific gradient by investigating the nodal $Alloy$ membership in constrained subgraphs. For this problem, an ATLAS materials graph was generated using a simplicial grid sampling with a spacing of 0.05 mole fraction using the NIMPLEX library [97]. The feasibility of the nodal $Alloy$ compositions were determined with a CALPHAD surrogate classifier trained on phase data acquired from ThermoCalc equilibrium calculations for the FeNiCrNbMo quinary alloy system. Figure 4.4 shows a UMAP projection highlighting the filtered, feasible nodal $Alloy$ compositions. The ThermoCalc equilibrium calculations were made using Thermo-Calc's TCHEA4 database [116, 131], and the surrogate k-nearest neighbors classifier models were trained to classify compositions as deleterious when > 0.1 mole fraction of any deleterious phase is present from 2900 K down to 1400 K, with a separate model trained for each subsystem present in the

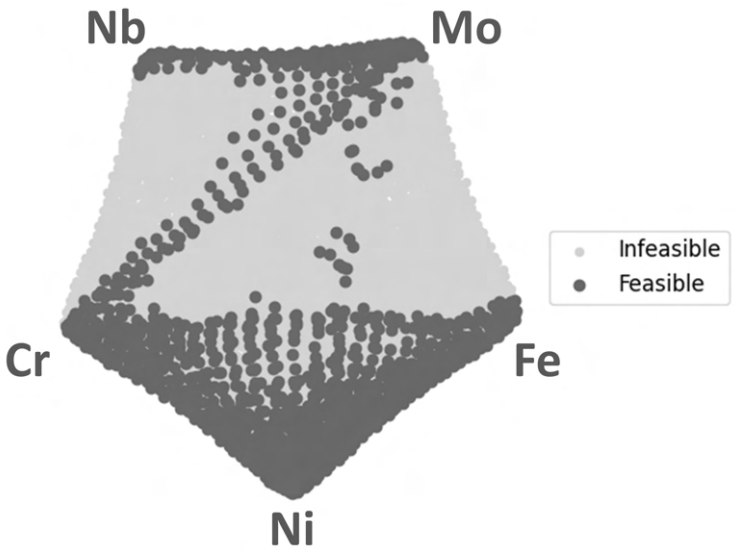

Figure 4.4: Uniform Manifold Approximation and Projection (UMAP) plot showing an embedded 2-dimensional embedded representation of the infeasible nodal $Alloy$ compositions (light grey) and feasible nodal $Alloy$ compositions (dark gray) in the 5-component FeNiCrNbMo composition space.

FeNiCrNbMo quinary for improved accuracy. These models were used to label the grid sampled compositions as feasible or infeasible. The resulting constrained ATLAS materials graph, $G_{\text{AM-F}}$, was used to identify constrained subgraphs. The constrained subgraphs revealed that, under the prescribed feasibility conditions, all terminal alloys are compatible. The nodal $Alloy$ compositions for the subgraph containing these terminal nodes are highlighted in Fig. 4.5, with specific emphasis on the terminal alloy compositions. Thus, since the terminal alloys belong to the same feasible constrained subgraph, one can confidently proceed with designing a specific compositional gradient for these terminal alloys.

4.4.3 Multi-terminal Gradient Design

Using the existing heuristic solver by Mehlhorn [78], a compositional gradient tree was designed for the identified terminal alloys, τ_{CGA}, that connects Hastelloy C-276, Inconel 718, and Stainless Steel 316. Here, a brute force node augmentation step was used to ensure a tree with more than two leaf nodes was acquired. This node augmentation step was done by repeatedly running the heuristic algorithm while randomly adding a single node from the same constrained

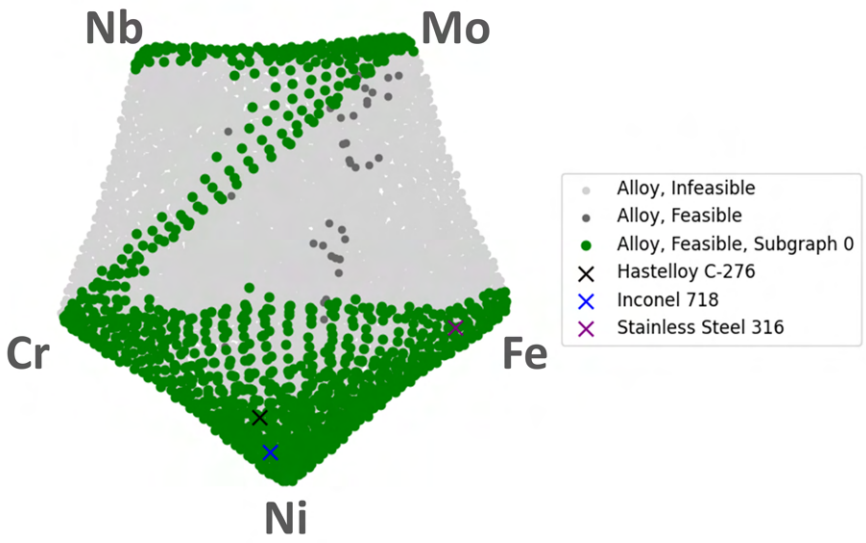

Figure 4.5: Uniform Manifold Approximation and Projection (UMAP) plot showing an embedded 2-dimensional embedded representation of the infeasible nodal $Alloy$ compositions (light grey), feasible nodal $Alloy$ compositions (dark gray), and the nodal $Alloy$ compositions from the constrained subgraph containing the terminal alloys in the 5-component FeNiCrNbMo composition space.

subgraph as an additional terminal, and retaining the tree with the best cost after repeating this procedure for all unseen nodes in the constrained subgraph. While this scales poorly with higher dimensionality and larger graphs, for this demonstration it ensures a three-leaf-node branched tree structure to best illustrate TreeMAP in operation. The topological structure of the resultant CGA is visualized in Fig. 4.6 alongside a unique nodal $Alloy$ composition identifier, '*Material_ID*', for each composition.

4.4.4 Structural Discretization, Labeling, & Spatial Part Graph

In this case, Hastelloy C-276 is highly desirable for corrosion resistance at seawater-facing surfaces, Inconel 718 is desirable for wear and corrosion resistance at interior flow-facing surfaces, and stainless 316 is desirable for decent mechanical performance at much lower cost in the bulk of the structure. Figure 4.7a shows the initial 3D CAD model of the mechanical connector structure. Figure 4.7b shows the volumetrically-discretized structure using a voxelization with axial step sizes of $(\mathrm{d}x, \mathrm{d}y, \mathrm{d}z) = (1.5, 1.5, 1.5)$ in millimeters. This discretization, containing 4,550,941 voxels,

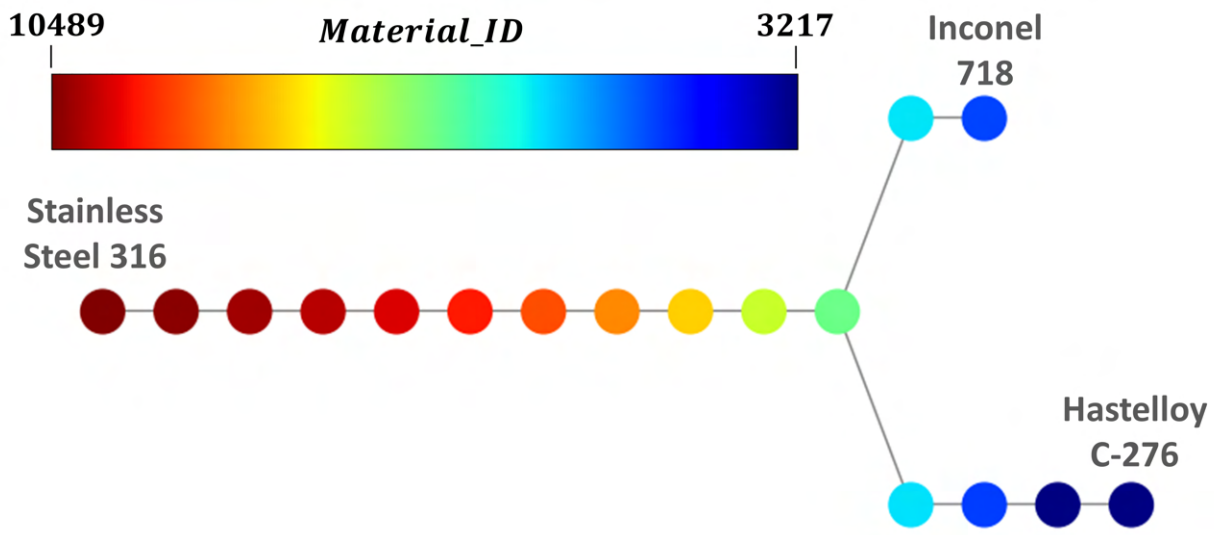

Figure 4.6: The optimized CGA tree, τ_{CGA}, with nodal $Alloy$ compositions and $JOINS\_TO$ edges between the three terminal alloys Hastelloy C-276, Inconel 718, and Stainless Steel 316, labeled and colored by the unique nodal $Material\_ID$ associated with each unique $Alloy$ node in G_{AM-F}.

was achieved using the *pyvista.voxelize* function from the PyVista [132] Python interface for the Visualization Toolkit (VTK) [133].

Figure 4.7c shows the initial label placement for Hastelloy C-276 and Inconel 718. In this case, Inconel 718 was placed on all surface voxels within an 80 mm radius of the central axis down the cylindrical cavity of the mechanical connector. Meanwhile, the Hastelloy C-276 was placed on the remainder of the surface for all voxels at least the minimum gradient distance away (5 steps based on τ_{CGA} for this case). This resulted in both materials covering almost the entire surface, with unlabeled compositions left around the Ring-Type Joint (RTJ) groove on the flange to place the gradient between these two terminal alloys. This structural discretization with the initial material labeling was stored as a spatial part graph G_P with a corresponding 4,550,941 cell nodes V_P and 57,557,168 pairwise cell adjacency relationships as edges in E_P.

4.4.5 Tree-based Material Adjacency Propagation

Given τ_{CGA}, ordered with Stainless 316 as the root node m_c and the initially labeled part graph G_P, the TreeMAP algorithm was applied to this problem, placing the gradient compositions in a

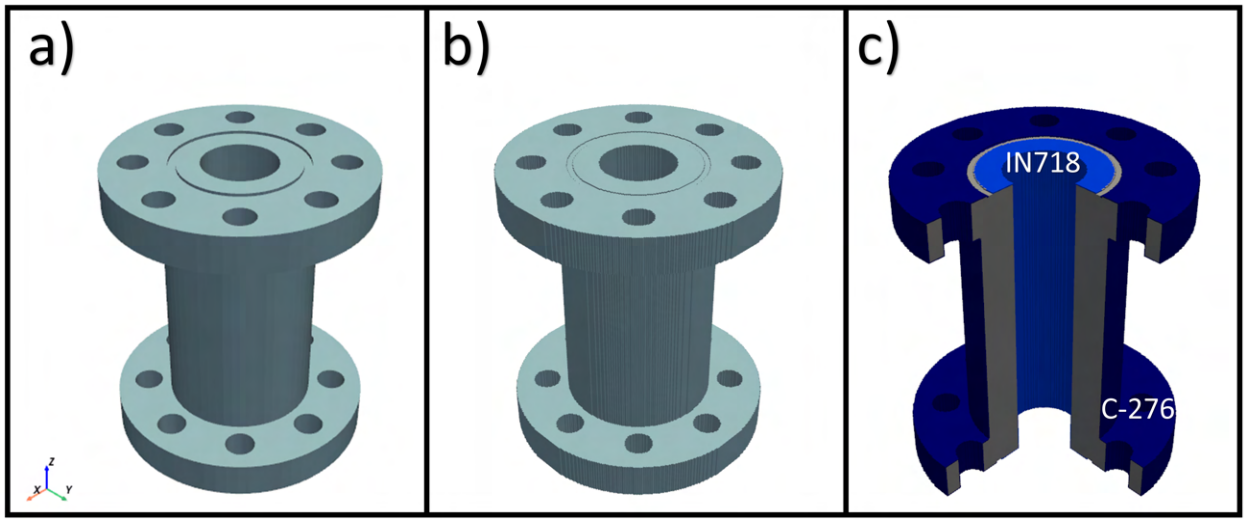

Figure 4.7: Illustration of the structural representations for this case study starting from **a)** the initial mechanical connector 3D CAD model represented as an STL, to **b)** volumetrically discretized 3D model, and, lastly, **c)** a cutaway showing the initial surface placement of the terminal alloy compositions Hastelloy C-276 and Inconel 718 since their boundary locations are known. Stainless 316 is left dependent on the CGA and placement of the other two terminal compositions and thus is used as m_c for TreeMAP.

CPU Time of 87.6 seconds using 32 Intel Xeon CPU cores at 2.20 GHz and 51.0 GB of system RAM. The resultant, fully labeled G_P is mapped back to the discretized 3D structure shown with a cutaway in Fig. 4.8 to illustrate the gradient placement throughout the structural volume. Notice that due to the convergence of the gradients with one another in the acyclic tree structure, and the close placement of IN718 and C-276 results in the compositions necessary for joining them being placed down on the surface, bypassing the gradient compositions necessary for SS316, which are added as the gradient progresses to the interior of the volume until the remainder of the structure can be filled with the significantly cheaper 316 stainless. The resultant structural design possesses local corrosion and wear resistance characteristics of C-276 and IN718 while being much cheaper due to the significant use of of cheaper stainless 316. Ultimately, the structural design demonstrates **a combination of properties not possible with any one of the terminal alloys**.

It is critical to understand how TreeMAP operates to properly set up a structural design problem and interpret the resulting solution in case adjustments need to be made. In this case, while we

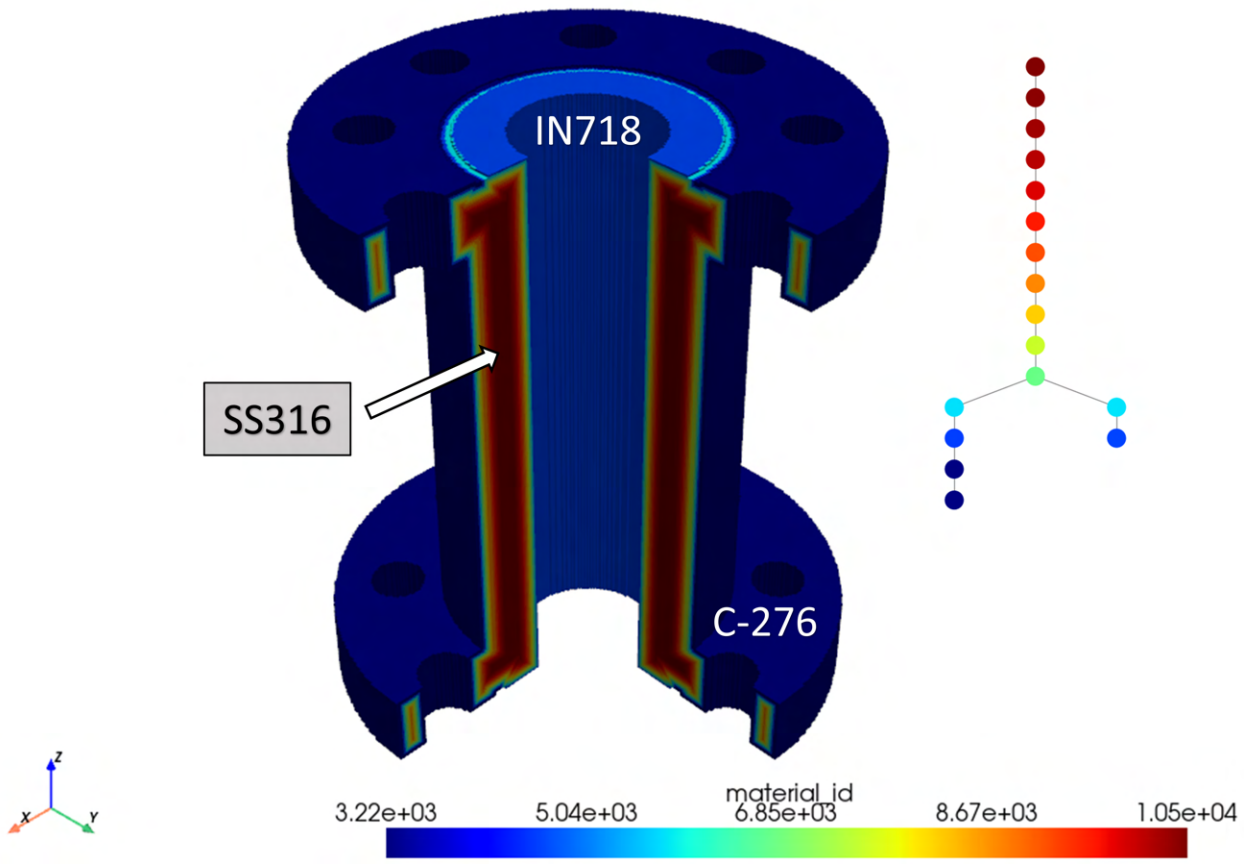

Figure 4.8: Illustration of the structural representations for this case study starting from **a)** the initial mechanical connector 3D CAD model represented as an STL, to **b)** volumetrically discretized 3D model, and, lastly, **c)** a cutaway showing the initial surface placement of the terminal alloy compositions Hastelloy C-276 and Inconel 718 since their boundary locations are known. Stainless 316 is left dependent on the CGA and placement of the other two terminal compositions and thus is used as m_c for TreeMAP.

see a sizable SS316 region placed in the center of the structural volume, there is not enough room in the flanges to complete the entire gradient based on the 3D model and discretization provided. If the additive technique used for this design can print finer features, then a finer discretization may address this, allowing the gradient to fully reach SS316 in the flanges. Alternatively, if the gradient compositions in the flanges still possess satisfactory structural properties at a lower cost than the surface material(s), then the design may be sufficient as-is. Otherwise, if the gradient is undesirable without the placement of SS316 in the flanges, then the design may require re-running of TreeMAP with a different 3D model, discretization, and/or initial placement of terminal alloys.

4.5 Scalability

Due to the vast size difference between G_P and τ_{CGA} for realistic problems such as the mechanical connector, the complexity of the TreeMAP algorithm is effectively dominated by the size of G_P. TreeMAP visits each node in G_P approximately twice, once when searching for nodes with a given material label, and a second time when finding unlabeled neighbors. The initially labeled nodes, and the unlabeled remaining nodes for placing m_c are only visited once. Thus, the algorithm can be considered to scale linearly with respect to the size of G_P, with total complexity on the order of $O(|V_P| + |E_P|)$. The CPU time of 87.6 seconds using 32 Intel Xeon CPU cores at 2.20 GHz and 51.0 GB demonstrated for the mechanical connector with spatial part graph G_P with 4,550,941 nodes V_P and 57,557,168 edges E_P can be used as reference and scaled to the G_P for an unseen problem based on the given linear time complexity relationship.

4.6 Chapter Summary

In summary, TreeMAP further builds upon ATLAS and the multi-terminal CGA design formulation, providing a means to map a CGA to a corresponding physical structure. This algorithm obeys the material adjacency required in the compositional gradient tree and yields a valid solution if valid inputs are given. The soundness and completeness of the TreeMAP algorithm were considered and discussed, as well as the time complexity. This proposed algorithm provides the last link from process design through to product performance, fully enabling the linkage for systematic

materials design. Although a full design demonstration was presented using real materials data, the design work focused on using existing alloys, such as Hastelloy C-276. However, for systematic materials design, the designer should not bias the problem by limiting the problem to conventional terminal alloys if possible, but instead, also consider the terminal alloys as a free design variable. In the following chapter, the three proposed components of 1) ATLAS, 2) the multi-terminal CGA design formulation, and 3) the TreeMAP algorithm, will be deployed in a unified framework for this purpose, fully enabling the systematic design of compositionally graded alloys and corresponding structures.

5. FRAMEWORK FOR SYSTEMATIC DESIGN OF COMPOSITIONALLY GRADED ALLOY STRUCTURES

With the introduction of the ATLAS, the alloy state space can be assessed using the constrained subgraphs for a given set of design constraints. These constrained subgraphs can be evaluated to identify where dissimilar alloys with desirable properties can coexist, since nodal $Alloy$ membership within the same constrained subgraph guarantees reachability. This enables property/performance-driven selection of compatible terminal alloys. Furthermore, the formulation of multi-terminal gradient design as a SPG alongside the proposed TreeMAP algorithm enables the design of multi-terminal CGAs and mapping these CGAs into corresponding discretized 3D structural geometries. This chapter presents a framework which integrates the capabilities of these methods for a unified approach to systematic design of CGA structures. Additional consideration is given to Scheil-Gulliver solidification simulations for predicting phases under the rapid cooling conditions of metal AM, and a relevant cracking susceptibility criterion and novel cost function are used to design a gradient that minimizes the worst-case cracking susceptibility. This framework is then applied to a real alloy system and engineering structure to show it can be successfully applied to design parts with localized performance that surpasses a traditional single alloy part.

5.1 Framework Overview

The framework used in this chapter unifies the capabilities of the ATLAS, multi-terminal CGA design, and TreeMAP algorithm into a single framework for the systematic design of compositionally graded alloy structures. This framework, illustrated in Fig 5.1, starts with collecting materials data in the alloy state space of interest (assuming data has not already been collected). This data is then used to train machine learning models for predicting missing or erroneous properties in the dataset due to computational or experimental errors. This completed dataset can be used to initialize or be added to an existing ATLAS graph or ATLAS materials graph. From this point, one can

filter the ATLAS materials graph based on the relevant design constraints to acquire a feasibility-constrained ATLAS materials graph $G_{\text{AM-F}}$ with resulting constrained subgraphs; these subgraphs can be evaluated, and the subgraph(s) with satisfactory property trade-offs can be selected. From these subgraphs, specific terminal alloys can be selected based on targeted property/performance criteria. These terminal alloys must be part of the same subgraph, but multiple sets can be selected to perform set-based design if multiple subgraphs have satisfactory property trade-offs. Next, an input a discretized 3D model of the corresponding structure can be used with the terminal alloys to create a labeled initial structure based with spatial placement of the terminal compositions. Meanwhile, property model(s) can be used to compute a relevant cost function for the $JOINS\_TO$ graph edges if necessary. This cost function and the identified terminal nodal $Alloy$ compositions can be used with an SPG solver to identify a multi-terminal CGA as a Steiner minimal tree $\tau_{\text{CGA}} = \tau_{\text{min}}$. Then, the compositional gradient τ_{CGA} and the part graph G_P of the labeled, discretized structural geometry are passed to the TreeMAP algorithm to place the CGA within the part, labeling the G_P cell nodes with the required gradient $Alloy$ compositions. Lastly, these compositions are mapped from the fully labeled G_P back to the discretized cells in the 3D structure to achieve a compositionally graded alloy part.

5.2 Case Study: Gas Turbine Blade

This case study demonstrates the capability of the proposed framework for systematic design of CGA structures, ultimately driven by properties and performance. Specifically, this demonstration shows the combined capabilities of the proposed ATLAS, multi-terminal gradient design formulation, and TreeMAP algorithm for computational design of a compositionally graded gas turbine blade with three separate spatial property optimization problems. For materials design, the Cr-Nb-V-W-Zr quinary system and its subsystems are sampled with a 5 mol % resolution grid sampling of compositions. CALPHAD equilibrium calculations and Scheil-Gulliver [134, 135] solidification simulations are used to identify a manufacturable set of alloy compositions expected to avoid deleterious phases and form a single BCC phase. Then, a variety of material property models are used to predict relevant properties for the turbine blade design. Using the ATLAS to construct a

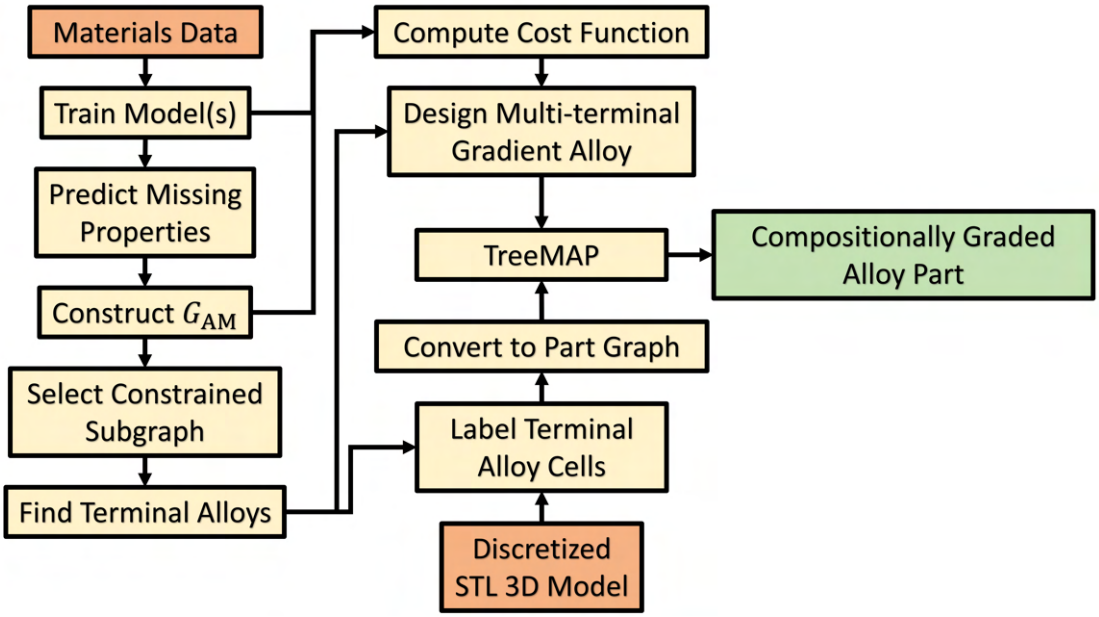

Figure 5.1: Unified systematic design framework for compositionally graded alloy structures based on the ATLAS, multi-terminal CGA design formulation, and the TreeMAP algorithm.

labeled property graph (LPG) from the materials data, the resultant materials graph is partitioned into MCCs to identify the constrained subgraphs based on both phase and property constraints. Then, these constrained subgraphs are evaluated based on the desired terminal alloy properties, resulting in the identification of a single satisfactory subgraph. Within this satisfactory subgraph, three terminal alloys for the turbine blade are selected by solving the initial property optimization problems over the discrete set of alloys in the subgraph. Next, a multi-terminal CGA is designed with the SPG formulation and the three identified alloys as terminals, using a cost function for minimizing the worst-case cracking susceptibility. With a CGA design in place, a discretized 3D gas turbine blade model is configured by defining the spatial placement of two of the three terminal alloys. Then, the TreeMAP algorithm is used to propagate the CGA throughout the discretized structure. Finally, the resultant gas turbine blade is shown to contain the manufacturable CGA, ultimately possessing a combination of properties and performance characteristics not otherwise possible using a single alloy composition in the same Cr-Nb-V-W-Zr quinary system.

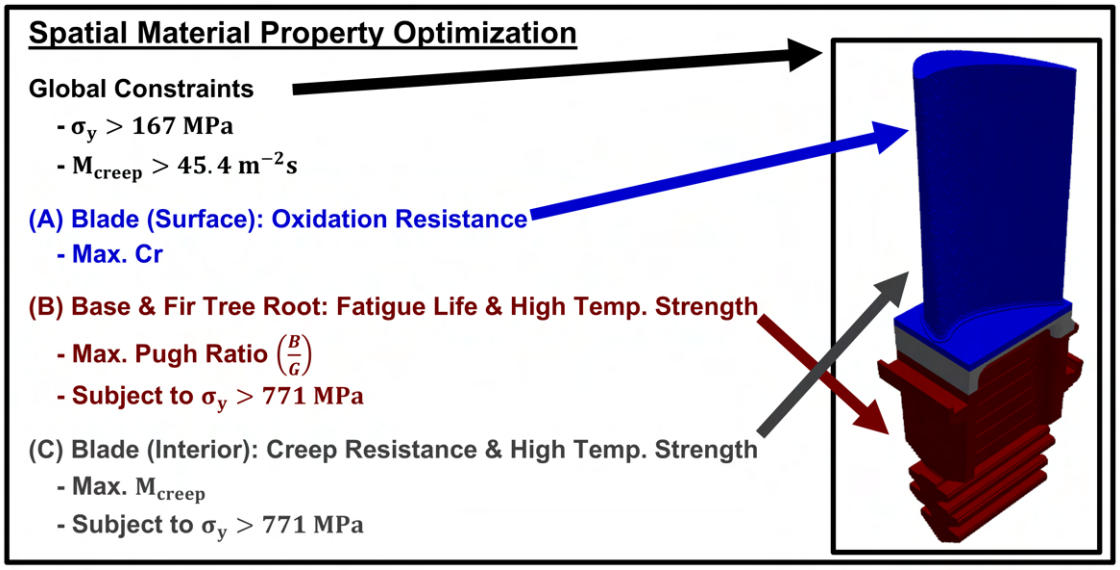

Figure 5.2: Localized spatial property optimization objectives & constraints for the gas turbine blade design demonstration.

5.2.1 Terminal Alloy Design Problems

The spatial property optimization problems for this turbine blade design demonstration are illustrated in Fig. 5.2. First, there is an overarching global requirement that the high-temperature yield strength, σ_y, and creep merit index [136], M_{creep}, should be above the 25th percentile value in order to avoid the use of extremely poor materials anywhere in the part. For σ_y, it is assumed that $1000°$C was a representative operating temperature to evaluate the yield strength for gas turbine blade alloys [137]. This results in respective lower bounds of 167 MPa for σ_y at $1000°$C and 45.4 m^{-2}s for M_{creep}. Note that all property percentile thresholds were calculated based on the properties of the feasible single-phase BCC compositions, not the entire dataset. The property/performance-driven optimization of terminal alloys for the compositional gradient can be expressed as separate optimization problems over the discrete nodal $Alloy$ compositions in the corresponding ATLAS materials graph which has been filtered for feasibility, $G_{\text{AM-F}} = (V_{\text{AM-F}}, E_{\text{AM-F}})$.

For the first terminal alloy, x_A, the structural design goal is to maximize the oxidation resistance on the high-temperature flow-facing surfaces of the blade. It is assumed that higher Cr content ap-

proximately indicates higher oxidation resistance, while recognizing that there are more rigorous metrics for characterizing oxidation resistance in alloys [138]. As a result, the materials design objective for this alloy is to maximize the chromium content. Since this is a demonstration, it is worth noting that implementing these metrics would not change the central framework demonstrated in this work. This optimization problem for the first terminal alloy can be expressed as

$$\begin{aligned} &\max_{v \in V_{\text{AM-F}}} \ \text{Cr}(v) \\ &\text{subject to: } \sigma_{\text{y}}(v) > 167\,\text{MPa}, \\ &\qquad M_{\text{creep}}(v) > 45.43\,\text{m}^{-2}\text{s} \end{aligned} \tag{5.1}$$

where $V_{\text{AM-F}}$ is the set of all $Feasible$-labeled $Alloy$ nodes in the ATLAS materials graph.

Next, for the second terminal alloy, x_{B}, the structural design goal is to maximize fatigue life and maintain exceptional high-temperature strength in the base/root of the blade [139]. Thus, for this alloy, the materials design objective is to maximize the fatigue life, indicated by the Pugh ratio ductility metric [140], with >95th percentile σ_{y}. This optimization problem can be expressed similarly as follows:

$$\begin{aligned} &\max_{v \in V_{\text{AM-F}}} \ \frac{\text{B}(v)}{\text{S}(v)} \\ &\text{subject to: } \sigma_{\text{y}}(v) > 771\,\text{MPa}, \\ &\qquad M_{\text{creep}}(v) > 45.43\,\text{m}^{-2}\text{s} \end{aligned} \tag{5.2}$$

where $\frac{\text{B}(v)}{\text{S}(v)}$ is the Pugh ratio (bulk modulus, B, divided by shear modulus, S) of a given nodal alloy composition, v.

Lastly, while not geometrically fixed to a specific bounded subregion within the structure, the last structural design goal is to maximize the high temperature creep resistance and maintain exceptional high temperature strength in the interior of the blade. Consequently, the last spatial material optimization was for a third terminal alloy, x_{C}, with high-temperature creep resistance and strength in the blade interior [137]. For this alloy, the objective is to maximize M_{creep} with >95th percentile σ_{y}. This optimization problem can be expressed as

$$\begin{aligned} &\max_{v \in V_{\text{AM-F}}} \; M_{\text{creep}}(v) \\ &\text{subject to: } \sigma_{\mathrm{y}}(v) > 771\,\text{MPa}, \\ &\qquad\qquad M_{\text{creep}}(v) > 45.43\,\text{m}^{-2}\text{s} \end{aligned} \tag{5.3}$$

Here, x_{C} is not fixed to a specific geometric region since its position is dependent on the number of compositions in the gradient to join the three terminal alloys in this design. This consideration is a direct and important consequence of integrating physical design, manufacturing considerations, and materials design into a unified methodology.

5.2.2 Methods

5.2.2.1 ATLAS Materials Graph Setup

The first step in assessing the alloy state space for an ATLAS graph is acquiring a sampling. For this work, a simplicial grid graph was efficiently generated for the Cr-Nb-V-W-Zr quinary system with a grid spacing of 5 mol % using the NIMPLEX library [97]. This sampling resulted in 10,626 unique compositions. Since this amount of compositions can be dealt with in local memory, and is part of a directed design effort involving limited graph filtering and queries, this sampling was used to initialize an ATLAS materials graph instead of a full ATLAS GDB.

5.2.2.2 Thermodynamic Calculations

To assess the phases in the sampling of the Cr-Nb-V-W-Zr quinary system, thermodynamic calculations were performed using Thermo-Calc [5], a commercial CALPHAD-based software package. In this work, equilibrium and Scheil-Gulliver solidification [134, 135] calculations were both used to assess the manufacturing feasibility of the thermodynamic space and solidification pathways. The process assumed by the Scheil-Gulliver solidification model results in an inhomogeneous solid, where different regions solidify at different temperatures, broadening the solidification range. This is more accurate to the rapid solidification conditions experienced by alloys processed under rapid cooling conditions, such as metal AM, which result in a non-equilibrium process. With the Scheil-Gulliver solidification model, there is complete mixing in the liquid phase, but there is no diffusion in the solid phase. Due to these conditions, the solidification process is progressive; the composition of the liquid remaining changes continuously according to the partition coefficient as solidification proceeds. The solute is rejected into the liquid as solid forms during this process, leading to segregation and compositional variation within the final solid material.

Equilibrium phase simulations were conducted for each composition using Thermo-Calc [5] version 2023b with the TCHEA6 database [6]. Temperatures from $1000°C$ to $2750°C$ were sampled at intervals of $250°C$ for each composition, collecting only the relevant (BCC) phases. Prior work [119] benchmarked the TCHEA4 database against experimental data, with results showing acceptable accuracy. Since the TCHEA6 database is more comprehensive, with updated thermodynamic assessments, it is expected to provide improved prediction accuracy compared to prior versions.

Thermo-Calc was also used to predict the Scheil-Gulliver curve. From this curve, the Scheil solidification range, the Kou hot-cracking criterion, and the phases present upon the end of rapid solidification were all obtained. The Scheil-Gulliver solidification range is more realistic to what actually occurs during rapid cooling compared to the equilibrium solidification range, which is often narrower. The former reflects the non-equilibrium conditions of rapid cooling more accurately, since incomplete diffusion in the solid phase results in a wider solidification range. Additionally,

there are a variety of quantities derived from the Scheil-Gulliver curve which are often used as indicator metrics for hot-cracking [59], such as the Kou cracking susceptibility criterion [141]. The Kou cracking susceptibility criterion is defined as the magnitude of the temperature derivative with respect to the square root of the solid fraction towards the end of solidification, $\mid dT/d\sqrt{f_{\mathrm{s}}} \mid$. Kou's criterion has been shown to be an accurate hot cracking indicator for the design of CGAs when compared to experimental results [59]. Lastly, the Scheil-Gulliver-predicted phase fractions at the end of the solidification curve were used as an estimate of what phases may form at the conclusion of rapid solidification during metal AM.

5.2.2.3 Material Property Modeling

A variety of models were used to integrate property/performance considerations into the design process. Predictions were specifically used to ultimately ensure that the designed structure satisfied constraints and achieved objectives for properties including high-temperature yield strength, ductility, and resistance to creep. In this work, higher ductility was used as an indicator of increased fatigue life performance.

In order to use ductility as a fatigue life indicator for the CGA design problem, the Pugh ratio ductility metric was used [140]. This ratio, given as the ratio of bulk modulus B to shear modulus S, is a common design metric for ductility in alloys. Higher values typically indicate more ductile behavior in concentrated multi-component alloys [142]. The threshold of B/S=1.75 is often used to indicate the transition from brittle to ductile behavior [143, 144, 145]. In the current work, a Rule-of-Mixtures average of the pure elemental properties was used to calculate these elastic constants for each composition.

For yield strength, a strengthening model for refractory BCC high entropy alloys originally proposed by Maresca and Curtin was used in this work [102]. Since the present work uses a single-phase BCC alloy design space, comprised largely of alloys with multiple principal elements, the model by Maresca and Curtin is an appropriate predictor of strength. For this model, an alloy's

shear strength is approximated across a wide range of temperatures as

$$\tau_{\text{y}}(T, \dot{\varepsilon}) = \tau_{\text{y0}} \exp\left[-\frac{1}{0.55}\left(\frac{kT}{\Delta E_b} \ln \frac{\dot{\varepsilon}_0}{\dot{\varepsilon}}\right)^{0.91}\right], \tag{5.4}$$

where k is the Boltzmann constant, T is the absolute temperature at which yield stress is predicted, τ_{y0} is the zero-temperature shear stress, and ΔE_b is the energy barrier for the motion of each dislocation segment. Also, $\dot{\epsilon}$ gives the applied strain rate, set as $\dot{\epsilon} = 10^{-3}s^{-1}$, and $\dot{\epsilon}_0$ gives a reference strain rate, estimated as $\dot{\epsilon}_0 = 10^{4}s^{-1}$. This equation has been shown to give a lower bound estimate [146, 147] for the tensile yield strength with

$$\sigma_{\text{y}}(T, \dot{\varepsilon}) = M\tau_{\text{y}}(T, \dot{\varepsilon}) \tag{5.5}$$

Here, $M = 3$ was used as the Taylor factor of a BCC metal. This model was used to predict yield strength at 1273 K to capture the yield strength of alloys across the alloy state space during high temperature operations of the turbine blade. A previously published description of the framework employed in this work for calculating the inputs to these equations can be found here [101]. This prior work by Acemi at al. also explains how Rule-of-Mixture averages of relevant properties of pure elements can be used to estimate the Maresca-Curtin yield strength.

For assessing high temperature creep performance, the creep merit index M_{creep} was used as a design indicator for creep resistance in the alloy state space. This index was originally proposed by Reed et al. [136] and is given as,

$$M_{\text{creep}} = \sum_i \frac{x_i}{\widetilde{D_i}}, \tag{5.6}$$

where x_i is the molar fraction of each alloy component and $\widetilde{D_i}$ is the interdiffusion coefficient for that respective component. This index is meant to capture an alloy's resistance to diffusion-based creep mechanisms such as Nabarro-Herring and Coble creep. Larger creep merit indices correlate with greater resistance to creep deformation. For this prediction, diffusion coefficients

were queried at 2273 K from Thermo-Calc's MOBHEA3 database, and the reference element was selected as the element with largest atomic radius in each alloy composition.

5.2.2.4 Deep Learning Regressor Model

A deep-learning property regression model [148] was employed to predict property values where needed. As input features, the model was trained the elemental compositions (i.e., Cr, Nb, V, W, Zr) with material properties as targets, including yield strength at various temperatures and the Kou criterion. The regressor captured the non-linear relationships between composition and a variety of properties. The overall performance of the property regressor is indicated by the model analysis shown in Fig. 5.3. In this work, the model was specifically used for 1) predicting missing/erroneous values of Yield Strength at $1000°$C and 2) to predict the Kou cracking susceptibility criterion across the $JOINS\_TO$ edges for a property-based cost function discussed in the following section.

5.2.2.5 Cost Function: Minimize Worst-case Cracking Susceptibility

Recent work on the manufacturability of CGAs indicates the benefit of using cracking susceptibility indicators for designing crack-free gradients [59]. However, as indicators, they suggest susceptibility and do not directly predict occurrence. Thus, this work proposes a novel cost function for minimizing the worst-case magnitude of the normalized Kou cracking susceptibility criterion [141], Kou$'$, with a secondary consideration for minimizing the path length. This cost function was defined as:

$$c(e) = \int_{x_a}^{x_b} \left(\text{Kou}'(x)\right)^P dx \tag{5.7}$$

where Kou$'(x)$ is the normalized Kou cracking susceptibility criterion [141] as a function of alloy composition x, and P is a user-defined power factor. This work used $P = 3$ to minimize the worst-case value of Kou$'$. It was found that $P > 3$ began to over-prioritize granular improvement in Kou$'$ for a drastically longer compositional gradient. This cost function minimizes the magnitude of Kou$'$ to improve the manufacturability of the gradient compositions by applying a power factor scaling. This power scaling ensures the planning process will prioritize a path with nominally

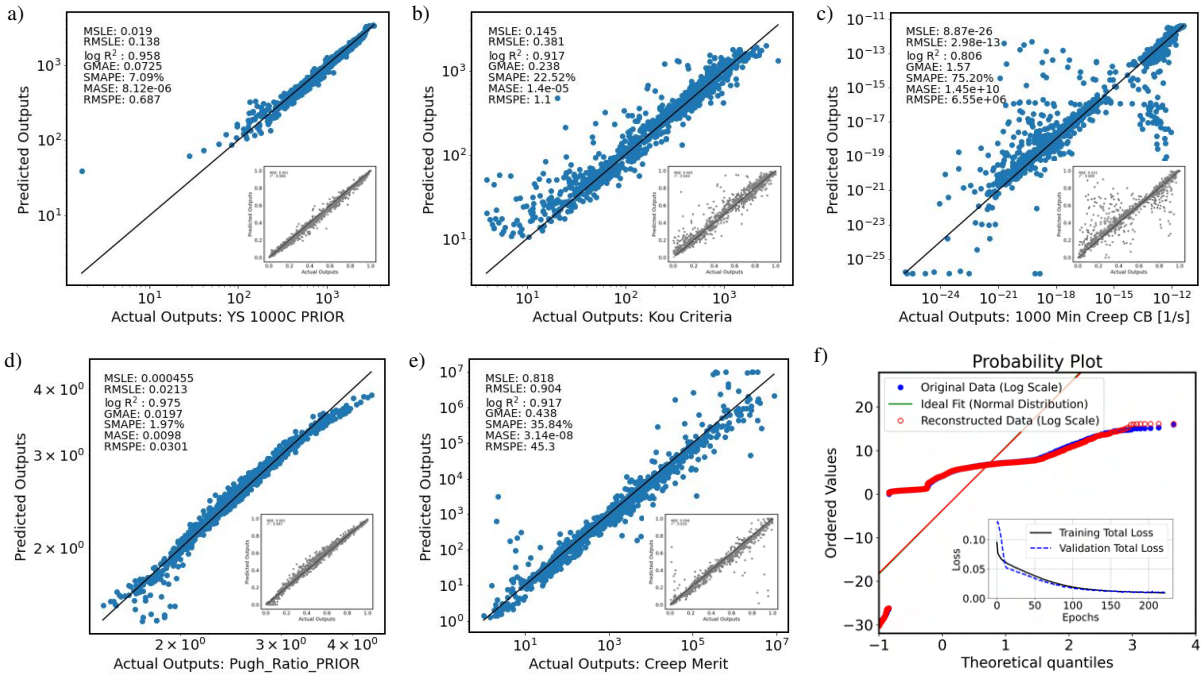

Figure 5.3: Prediction performance and probability analysis of the property regression model, shown for for various target material properties. Parity plots for **(a)** Yield Strength (YS) at 1000°C, **(b)** Kou Criteria, **(c)** Coble Creep at 1000°C, **(d)** Pugh Ratio, and **(e)** Creep Merit. These plots indicate strong agreement between predicted and actal points across different magnitudes. The smaller plots shown in each figure show the model performance for scaled data. Lastly, **(f)** shows a Quantile-Quantile plot comparing true and predicted values with an ideal fit for a vector of all features in log scale. This shows the model's effectiveness in reconstructing the true data. The inset shows the history of training and validation loss throughout the training process.

lower magnitudes of Kou′ regardless of compositional length. By integrating over composition, length is still considered so that for two paths with an identical sum of magnitude for Kou′, the gradient with a smaller change in composition will have a lower cost. The weight for each edge was computed by 1) interpolating the alloy compositions along each $JOINS\_TO$ edge with a step size of 0.1 mol %, 2) predicting Kou′ using the deep learning regression model with these interpolated compositions as input, 3) applying the power scaling to Kou′, and 3) computing the weight from the cost function in Eq. 5.7 using trapezoidal integration for the predicted & scaled Kou′ values over their corresponding interpolated compositions.

5.2.3 5-D Alloy State Space Topological Partitioning & Analysis

Using the 5 mol % grid sampling of the 5-component Cr-Nb-V-W-Zr BCC alloy composition space, the data from the equilibrium, Scheil-Gulliver solidification, and property predictions outlined in the prior sections was saved alongside each respective composition. In this work, the phase constraint for the alloy composition space was not only to avoid deleterious phases, but specifically to only use single phase BCC alloys. This constraint was applied by requiring > 99 mol% BCC phase across CALPHAD-predicted Scheil-Gulliver solidification [134, 135] simulations as well as equilibrium phase predictions. For equilibrium phase predictions, the worst case phase composition from the solidification temperature T_s of each alloy down to $1000°C$ was considered for the phase constraint. In addition to the single phase BCC constraint, property constraints were also applied based on the global property constraints for the gas turbine blade outlined in Fig. 5.2. This was done to ensure all alloys in the satisfactory alloy chemistry space and resulting CGA met these fundamental constraints. Applying all of these constraints to the 10,626 unique alloy compositions resulted in 169 feasible nodal $Alloy$ compositions in the filtered ATLAS materials graph $G_{\text{AM-F}}$. These feasible compositions are visualized in Fig. 5.4 using a UMAP dimensionally-reduced representation of the alloy state space[121, 122]. Evaluating the properties of these feasible compositions can indicate the possible material property trade-offs as well as the global distribution of properties across the alloy state space. Topological partitioning of the state space by identifying the MCCs of the $G_{\text{AM-F}}$ will enable evaluation of distinct constrained subgraphs for continuous compositional gradation between terminal alloys and their specific property trade-offs.

5.2.3.1 Feasible State Space

Analyzing the feasible nodal $Alloy$ compositions $V_{\text{AM-F}}$ for the feasible state space in $G_{\text{AM-F}}$ revealed information about the chemistry of these compositions and their properties for materials design. Analyzing the filtered data indicated that the feasible alloy chemistry space contains compositions in 4 binary systems: CrW, NbV, NbW, and VW; 4 ternary systems: CrNbV, CrNbW, CrVW, and NbVW; and 2 quaternary systems: CrNbVW and NbVWZr. From this information, it

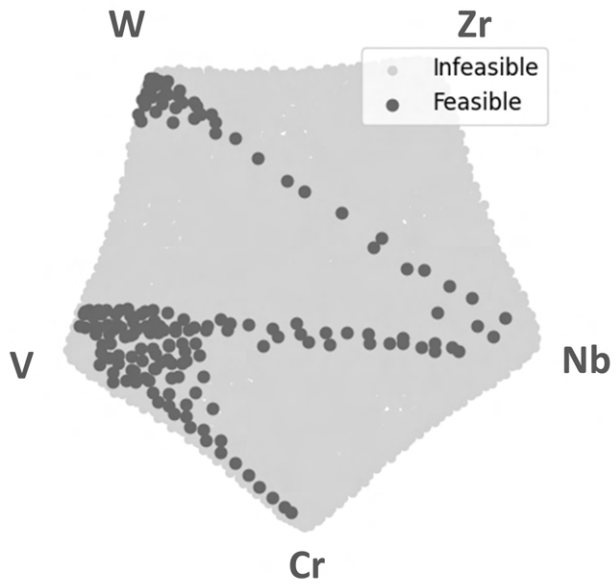

Figure 5.4: Uniform Manifold Approximation and Projection (UMAP) plot showing an embedded 2-dimensional representation of the infeasible alloy compositions (light grey) and feasible alloy compositions (dark gray) in the 5-component CrNbVWZr composition space. In this embedding, similar alloy compositions are qualitatively displayed closer together. Alloy compositions with lower configurational entropy are displayed closer to the exterior boundary of the hypocycloid shape made by the embedding, while alloys with higher configurational entropy are displayed closer to the center. Pure elements are projected at the vertices, which are labeled with their corresponding element.

was evident that Zr was only present in some feasible quaternary compositions and that no quinary compositions were feasible.

With this data, UMAPs with different color mappings, shown in Fig. 5.5a, were used to qualitatively indicate the variation of the properties of interest with respect to alloy chemistry. From these plots, several trends were apparent. First, σ_{y} appeared to decrease for compositions with higher V and Nb content, and increase for some W-rich compositions and compositions with relatively high configurational entropy near the V vertex. The creep merit index, M_{creep}, is shown on a log scale. $\log_{10} M_{\mathrm{creep}}$ appeared to increase as W content increases but does not show a significant trend with respect to the other elements. The Pugh ratio appears to be elevated for compositions containing higher amounts of Nb and V and lower for W and Cr. This trend is in agreement with the experimentally observed fact that Nb- and V-based alloys tend to be less brittle than W- and Cr-based ones. Lastly, the color map of Cr content indicates the presence of alloy compositions with moderate to high Cr content in the satisfactory alloy design space.

Meanwhile, the scatter plot matrix, shown in Fig. 5.5b, shows the distribution for each property in the histograms on the diagonal, as well as pair-wise correlation between these properties in the scatter plots off the diagonal. From these plots, we observed left-skewed distributions for σ_{y}, $\log_{10} M_{\mathrm{creep}}$, and Pugh ratio. The Cr content is strongly left-skewed due to the presence of compositions with minimal to no Cr content. From these plots, the property trade-offs between the available compositions become more apparent. For example, maximizing Cr content, the highest Cr content possible in the satisfactory alloy design space will be 0.85 mole fraction.

However, in order to fully evaluate the design space and select compatible terminal alloys for a CGA, the feasible ATLAS materials graph was used to identify the maximally connected components [98] of the graph and partition the design space into constrained subgraphs. The MCCs were identified using the breadth-first-search (BFS) approach implemented in NetworkX [7]. These subgraphs are significant, because every $Alloy$ node within a given subgraph is reachable from any other $Alloy$ within that subgraph using continuous compositional gradation along the $JOINS\_TO$ edges. This enables the selection of terminal alloy compositions for which there exists a gradient

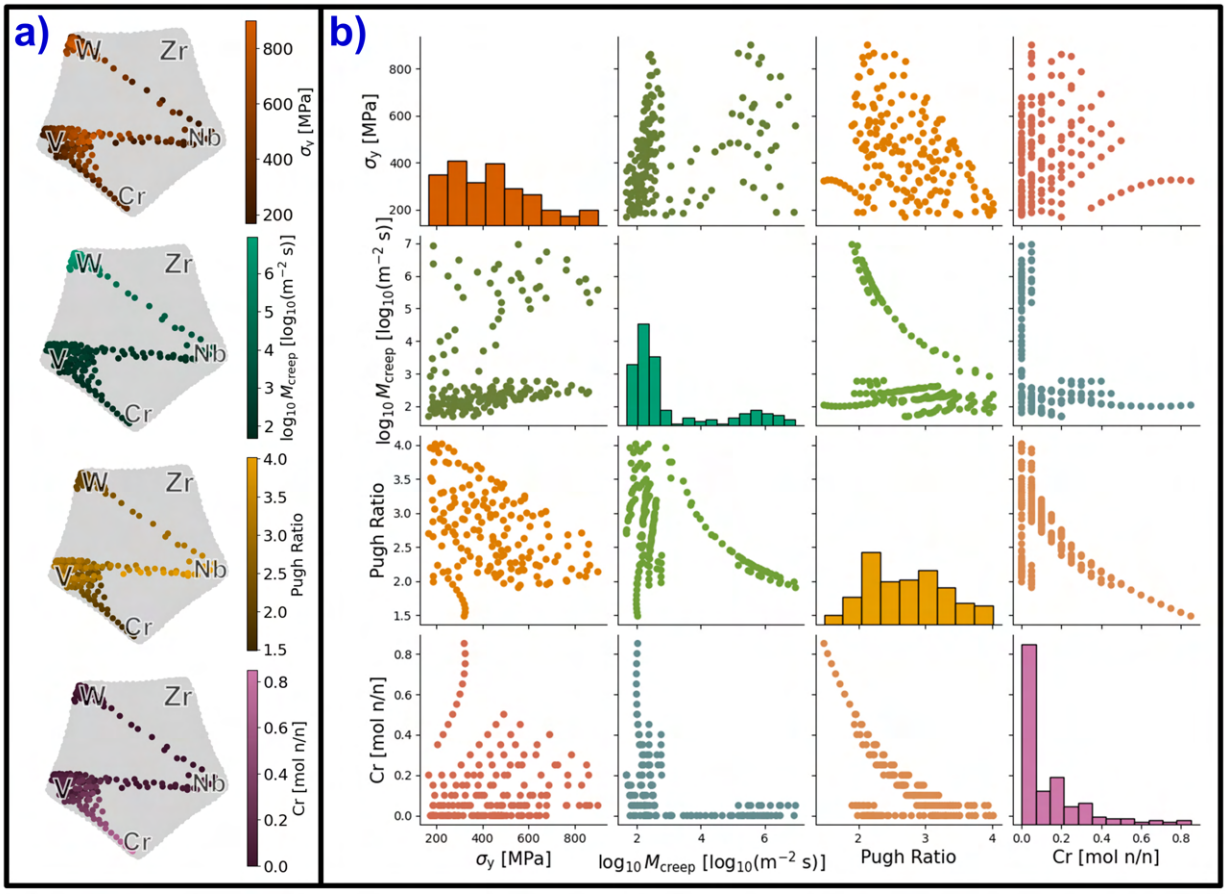

Figure 5.5: Relevant material properties for all feasible *Alloy* nodal compositions in $G_{\text{AM-F}}$. **a)** Uniform Manifold Approximation and Projection (UMAP) plots with color mapping illustrating the qualitative variation of the properties of interest with respect to alloy composition, and **b)** Scatterplot matrix of the properties of interest for this work, illustrating the pairwise correlation between properties (off-diagonal) and the distribution of each property (diagonal).

of satisfactory compositions prior to any effort to design a specific, optimized gradient path.

5.2.3.2 Analysis of Constrained Subgraphs

The four constrained subgraphs with the most *Alloy* nodes contained 80, 40, 37, and 5 nodes, respectively. Only the two largest subgraphs contained compositions with notable amounts of Cr. We did not consider the third largest subgraph due to our requirement for an oxidation-resistant (high-Cr) alloy, nor did we consider the fourth largest subgraph or those smaller due to their insignificant size.

As a result, the primary focus was directed towards the largest two subgraphs with 80 and 40

Alloy nodes, which we will refer to as $G_\alpha = (V_\alpha, E_\alpha)$ and $G_\beta = (V_\beta, E_\beta)$ respectively, where $V_\alpha, V_\beta \subset V_M,$ and $E_\alpha, E_\beta \subset E_M$. We performed the same analyses as the full satisfactory chemistry space. Observing the nodal alloy compositions for G_α, we found that it contains compositions in three ternary systems: CrNbV, CrVW and NbVW; and one quaternary system: CrNbVW. Meanwhile, for G_β, we found that it contains compositions in two ternary systems: CrVW and NbVW; and one quaternary system: CrNbVW. Thus, Zr was not present in any amount in either subgraph.

The same UMAP embedding with different color maps, shown in Fig. 5.6a, qualitatively indicated the variation of the properties of interest with alloy chemistry for the nodal alloy compositions in $G_\alpha \subset G_{\text{AM-F}}$. From these color maps, we saw that σ_{y} appears to decrease for compositions with higher V or Cr content and less configurational entropy. Meanwhile, the $\log_{10} M_{\text{creep}}$ increases for some compositions towards the V vertex. The Pugh ratio decreases for compositions containing higher amounts of Cr, and increases for higher Nb-content alloys. Lastly, the color map of Cr content indicates the presence of alloy compositions with moderate Cr content in this subset of the satisfactory design space, up to 0.45 mole fraction.

The scatter plot matrix for the nodal alloy compositions in G_α, is shown in Fig. 5.6b. From the plots in this matrix, we observed an approximately normal distribution for yield strength, with a range from approximately 150 to 875 MPa. In this subgraph, σ_{y}, $\log_{10} M_{\text{creep}}$, and Pugh ratio are approximately uniform in distribution. However, the Cr content is strongly left-skewed due to the presence of a compositions in the NbVW ternary with no Cr content. By partitioning the graph into maximally connected component subgraphs, we now had *the guarantee when evaluating these properties that all nodal alloy compositions were reachable from all other nodal alloy compositions* by traversing satisfactory alloy compositions. Since we leveraged the graph structure to achieve this reachability guarantee, we then observed the attainable property trade-offs for selecting terminal alloys for the compositionally graded part. From the off-diagonal scatter plots in Fig. 5.6b, it was evident that we could select a composition with a moderate Cr content of 0.45 mole fraction, which is over twice the Cr content of many Ni-based superalloys. As a bonus, this can be achieved while maintaining a yield strength value above the average of 439 MPa among sat-

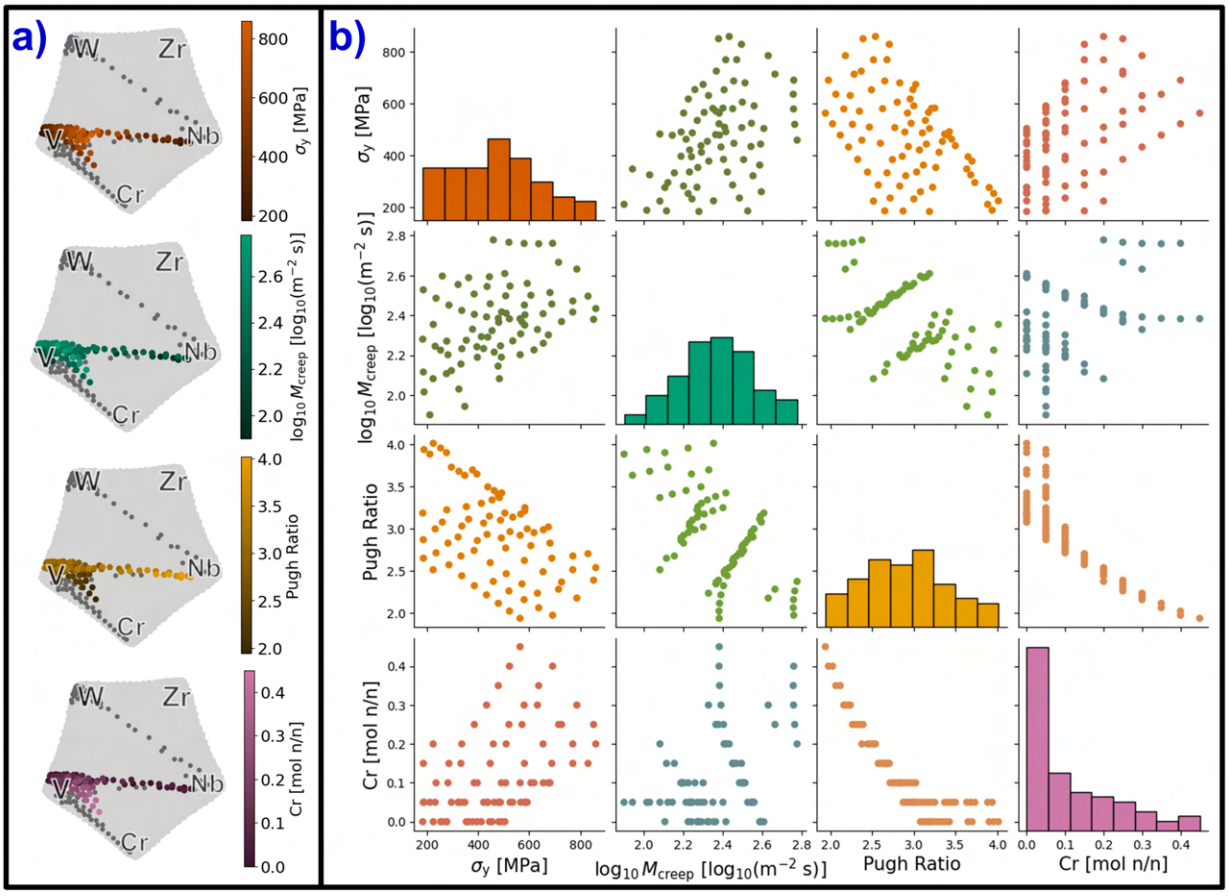

Figure 5.6: Relevant material properties for the *Alloy* nodal compositions in the largest constrained subgraph G_α. **a)** color maps showing the qualitative variation of the properties of interest with respect to alloy composition using Uniform Manifold Approximation and Projection (UMAP) embedded coordinates, and **b)** Scatterplot matrix of the properties of interest for this work showing the pairwise correlation between properties (off-diagonal) and the distribution of each property (diagonal) for the reachable nodal compositions.

isfactory compositions. Furthermore, some of the compositions with the highest value of the creep merit index in this subgraph also has an above-average value for σ_y. Lastly, we also observed that we can achieve a Pugh ratio above the average well above the lower magnitude threshold of 1.75 typically used to indicate ductility [140], while achieving a yield strength above the 95th percentile threshold of 771 MPa or more.

For the second largest subgraph, $G_\beta \subset G_{\text{AM-F}}$, color maps for the UMAP embedded coordinates, shown in Fig. 5.7a, were used to qualitatively illustrate the properties of interest in relation

to alloy composition. Similarly to G_α, we saw that σ_y and M_creep appear to decrease for compositions with higher V or Cr content and less configurational entropy. Meanwhile, the Pugh ratio seems to have a negative correlation with Cr content. Lastly, the color map of Cr content indicates the presence of the alloy compositions with up to the highest Cr content in the overall satisfactory alloy design space, 0.85 mole fraction. The scatter plot matrix for the nodal alloy compositions in G_β, shown in Fig. 5.7b, indicates that almost all of the compositions have below average values for σ_y, with none exceeding the 95th percentile threshold of 771 MPa required for the turbine blade interior and base materials. Thus, despite the availability of alloys with higher Cr content and M_creep, we did not select terminal alloy compositions from G_β due to the low σ_y values of its alloy compositions. This ultimately left G_α as the only viable subgraph for selecting terminal alloy compositions based on our property requirements.

5.2.4 Results

5.2.4.1 Terminal Alloy Design

First, three terminal alloy compositions were identified from the selected subgraph G_α based on the spatial property optimization problems. For alloy x_A, an oxidation-resistant alloy with maximum Cr content, the optimal composition was identified as $Cr_{45}V_{35}W_{20}$. This alloy has more than twice the Cr content of many Ni-based superalloys. For alloy x_B, an alloy with maximum fatigue-resistance (indicated by Pugh ratio) and $\sigma_\mathrm{y} \geq 771$ MPa, the optimal composition was identified as $Cr_{15}Nb_{25}V_{40}W_{20}$. This alloy has a predicted Pugh ratio of 2.71, above the typical 1.75 ductility threshold, and predicted $\sigma_\mathrm{y} = 830$ MPa. Lastly, for alloy x_C, an alloy with maximum high temperature creep-resistance and $\sigma_\mathrm{y} \geq 771$ MPa, the optimal composition was identified as $Cr_{30}Nb_5V_{40}W_{25}$. This alloy has a predicted $M_\mathrm{creep} = 429$ $m^{-2}s$, and $\sigma_\mathrm{y} = 786$ MPa. With these terminal alloy compositions selected, the next step is to proceed with identifying a multi-terminal CGA.

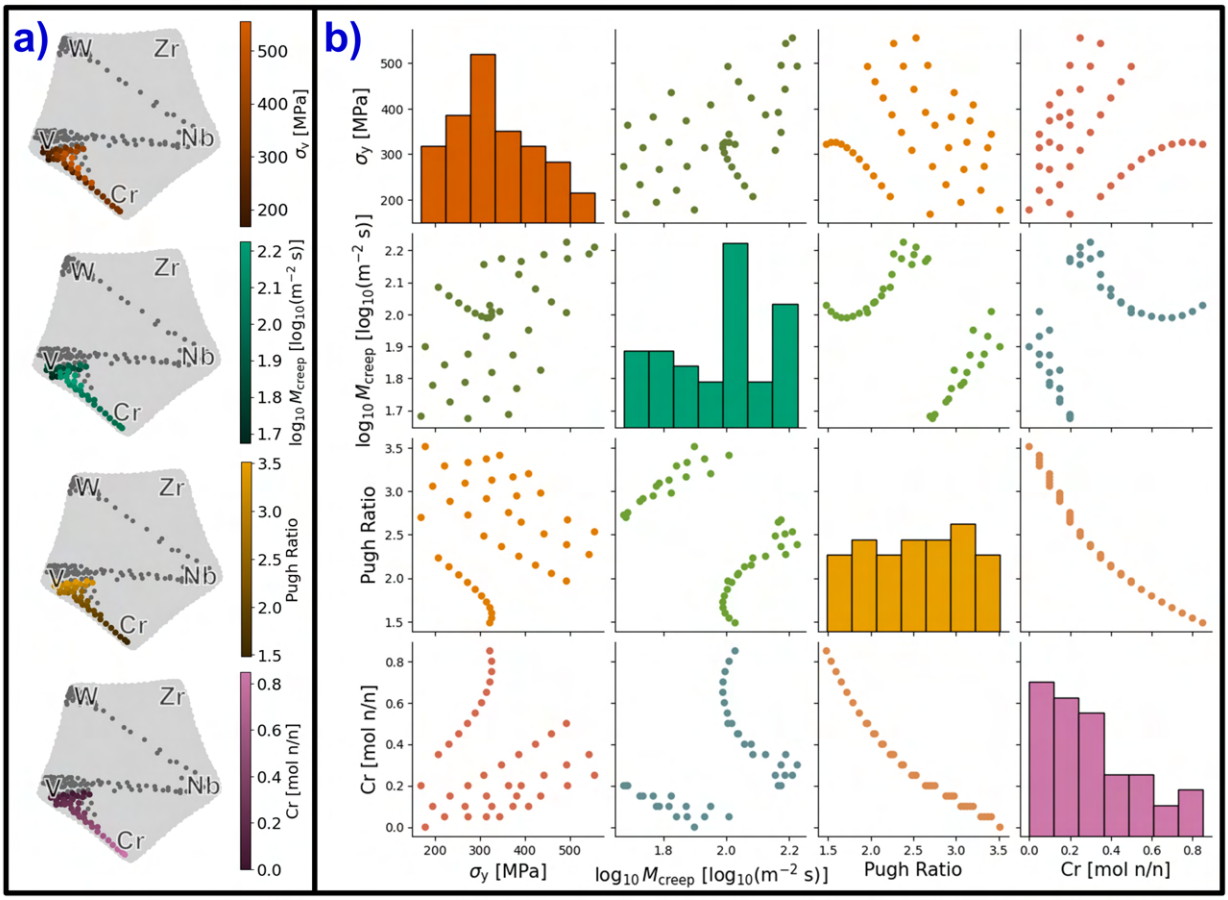

Figure 5.7: Relevant material properties for the $Alloy$ nodal compositions in the largest constrained subgraph G_β. **a)** color maps showing the qualitative variation of the properties of interest with respect to alloy composition using Uniform Manifold Approximation and Projection (UMAP) embedded coordinates, and **b)** Scatterplot matrix of the properties of interest for this work showing the pairwise correlation between properties (off-diagonal) and the distribution of each property (diagonal) for the reachable nodal compositions.

5.2.4.2 *Multi-terminal Gradient Alloy Design*

With these terminal alloys, we then used a cost function to minimize the worst-case cracking susceptibility and the existing heuristic solver by Mehlhorn [78] for the SPG. As mentioned prior, the solution given by this solver is guaranteed to have a total weight no more than $2\left(1-\frac{1}{l_{\text{min}}}\right)$ times that of a Steiner minimal tree, where l_{min} is the minimum number of leaves in the optimal tree. Since there are three terminal nodes, the worst case is that there are three leaves, making the weight of the approximated Steiner minimal tree, at most, $\frac{4}{3}$ times that of a true Steiner minimal tree. From this input, we designed a multi-terminal gradient tree, τ_{CGA}, that connects the three terminal alloys. The topological structure of the tree is visualized in Fig. 5.8 alongside a unique nodal $Alloy$ composition identifier, '*Material_ID*'. These nodal alloy compositions and their properties from τ_{CGA} are enumerated in Table 5.1 ordered from left to right based on Fig. 5.8, and specifically identified by $Material\_ID$. As a result of the graph construction, this gradient is composed of single-phase BCC alloys with M_{creep} and σ_{y} entirely above the bottom quartile. Furthermore, as a result of the cost function, this gradient has a low worst-case cracking susceptibility $\text{Kou}' = 0.066$ for a non-terminal alloy in the τ_{CGA}, with total Euclidean compositional edge length of 1.20 mole fraction [n/n]. Meanwhile, simply using Euclidean composition distance as the cost function results in a higher worst-case value of $\text{Kou}' = 0.079$, but a shorter total Euclidean compositional edge length of 0.57 mole fraction [n/n]. This indicates that the cost function works as intended to achieve the desired trade-off between length and worst-case cracking susceptibility.

5.2.4.3 *Conformal Compositional Gradient Mapping & Final Design*

Finally, with a gradient tree of alloy compositions identified, the multi-terminal gradient was ready for placement within the part geometry. In order to conformally place this gradient within a part geometry, the TreeMAP algorithm was used to maximize the use of the creep-resistant x_B in the interior of the blade after the required gradient layers are placed. With this approach, a spatial part graph, $G_P = (V_P, E_P)$, corresponding to a discretized 3D model of our part is necessary for efficient traversal. This graph has nodes, V_P, corresponding to each unique discrete geometric cell,

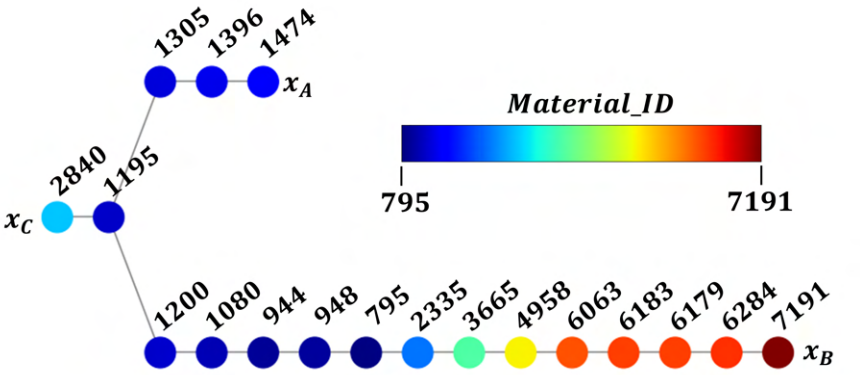

Figure 5.8: The optimized CGA tree, τ_{CGA}, with nodal $Alloy$ compositions and $JOINS\_TO$ edges between the three terminal alloys x_A, x_B, & x_C, labeled and colored by $Material\_ID$

Table 5.1: Gradient alloy tree nodal compositions & relevant properties

Material_ID	**Composition [mol %]**	σ_{y} **[MPa]**	M_{creep} **[m^{-2} s]**	$\frac{B}{S}$	Kou′
2840	$Cr_{30}Nb_5V_{40}W_{25}$	786	429	2.18	0.077
1195	$Cr_{30}V_{45}W_{25}$	580	576	2.16	0.027
1200	$Cr_{30}V_{50}W_{20}$	433	247	2.21	0.024
1305	$Cr_{35}V_{45}W_{20}$	479	243	2.22	0.022
1080	$Cr_{25}V_{55}W_{20}$	386	254	2.12	0.026
1396	$Cr_{40}V_{40}W_{20}$	523	241	2.32	0.021
944	$Cr_{20}V_{60}W_{20}$	337	266	2.44	0.028
1474	$Cr_{45}V_{35}W_{20}$	564	242	1.94	0.020
948	$Cr_{20}V_{65}W_{15}$	225	121	2.52	0.022
795	$Cr_{15}V_{70}W_{15}$	187	131	2.65	0.024
2335	$Cr_{10}Nb_5V_{70}W_{15}$	282	159	2.83	0.043
3665	$Cr_5Nb_{10}V_{70}W_{15}$	337	171	3.03	0.043
4958	$Cr_5Nb_{15}V_{65}W_{15}$	422	171	3.07	0.042
6063	$Cr_5Nb_{20}V_{50}W_{15}$	488	177	3.10	0.041
6183	$Cr_{10}Nb_{20}V_{55}W_{15}$	565	159	2.92	0.047
6179	$Cr_{10}Nb_{20}V_{50}W_{20}$	681	327	2.83	0.060
6284	$Cr_{15}Nb_{20}V_{45}W_{20}$	770	297	2.68	0.066
7191	$Cr_{15}Nb_{25}V_{40}W_{20}$	830	314	2.71	0.067

and edges, E_P, between geometrically adjacent cells. A $Material\_ID$ property was applied to the nodes V_P to indicate the geometric placement of materials based on their general node index in the ATLAS materials graph A_{AM}.

For the 3D model, a publicly available STL file of a gas turbine blade from GrabCAD was used [149]. CAD software (ANSYS® SpaceClaim® Direct Modeler™, version 2023) was then used to process the STL. To develop a discretized representation, the STL surface mesh was then converted into a voxelization with axial step sizes of $(\mathrm{d}x, \mathrm{d}y, \mathrm{d}z) = (0.4, 0.4, 1.0)$ in millimeters using the *pyvista.voxelize* function from PyVista [132] Python interface for the Visualization Toolkit (VTK) [133]. The voxel step sizes were based on the reported hatch spacing and the compositional resolution in the build direction for multi-material AM with steels using LP-DED by Salas et al. [20]. Here, the x- and y-axis resolutions were based on the reported hatch spacing, and the z-resolution was based on the reported number of layers in the build direction to achieve a distinct compositional band. The resultant voxelization, stored as an unstructured PyVista grid, contained 2.49×10^6 cells. This was then converted into a spatial part graph G_P for input to the TreeMAP algorithm. First, the graph G_P was defined, with corresponding nodes, V_P, for each cell and edges, E_P, between cells that share a vertex in the unstructured grid. Constructing G_P for the gas turbine blade resulted in a graph with 2.49×10^6 nodes and 3.13×10^7 edges.

The terminal alloy for the surface of the turbine blade, x_A was located within the part geometry by using PyVista to import a mesh file encompassing the blade, which translates to $z > -3$ in the STL mesh file coordinates. This mesh was used to select the voxels in the entire blade $z > -3$, which were then down-selected to only the exterior surface voxels. These surface voxels on the blade were labeled with $Material\_ID = 1474$ for x_A.

The terminal alloy for the base of the turbine blade, x_B, was located based on the location of x_A and the number of edges (16) in the tree between x_A & x_B. All voxels more than 16 edges away from x_A in the part graph G_P, and with $z < -3$ (to avoid encroaching on the blade core), were assigned x_B with $Material\_ID = 7191$.

The placement of the terminal alloy x_C could be determined similarly to x_B. However, in

this work, we defined the TreeMAP parameter $m_c = x_C$, $Material\_ID = 2840$, allowing the algorithm to place the required conformal gradient layers and fill the remaining unlabeled voxels with x_C as the last step to completing the gradient. Lastly, all unselected nodes were left unlabeled for TreeMAP to complete the conformal multi-terminal CGA mapping.

With G_P and the gradient given as τ_{CGA}, the TreeMAP algorithm was used to automatically map the multi-terminal CGA onto the discretized part geometry. In this case, since an overarching goal was to maximize the usageof the high-temperature creep-resistant material in the interior blade volume, x_C was used as the coalescent material algorithm parameter, m_c. Once the necessary gradient layers are placed, TreeMAP filled any remaining unlabeled voxels by assigning $m_c = x_C$ as the material. In essence, this amounts to maximizing the usage of the high-performing alloy x_C in the interior of the part where its high-temperature creep and strength properties are particularly desirable. While using $m_c = x_C$ may result in x_C being used elsewhere in the part, x_C possesses notable mechanical properties compared to the other compositions in the gradient, so it is desirable to use it rather than distribute the gradient compositions with subpar material properties over a larger physical area.

As a result, G_P was labeled for a compositionally graded alloy between the three terminal compositions in a CPU time of 38.05 seconds. Mapping the resultant, labeled voxel nodes from the G_P back to the PyVista voxelization resulted in the final CGA part. For the materials in the resultant gas turbine blade design, shown externally in Fig. 5.10d and internally in Fig. 5.10a by $Material\_ID$, the terminal alloy compositions can be seen in their prescribed locations, as well as the conformal layers of alloy compositions from τ_{CGA}. Figure 5.9a-d illustrates the evolution of the part design from the initial STL 3D model representation to the final compositionally graded gas turbine blade structure.

5.2.5 Discussion

Recalling the original goals for this design outlined in Fig. 5.2, Fig. 5.10 shows the resulting design from the systematic design approach taken in this work. First and foremost, Fig. 5.10a shows the unique materials by their $Material\_ID$ and indicates that the materials in the part

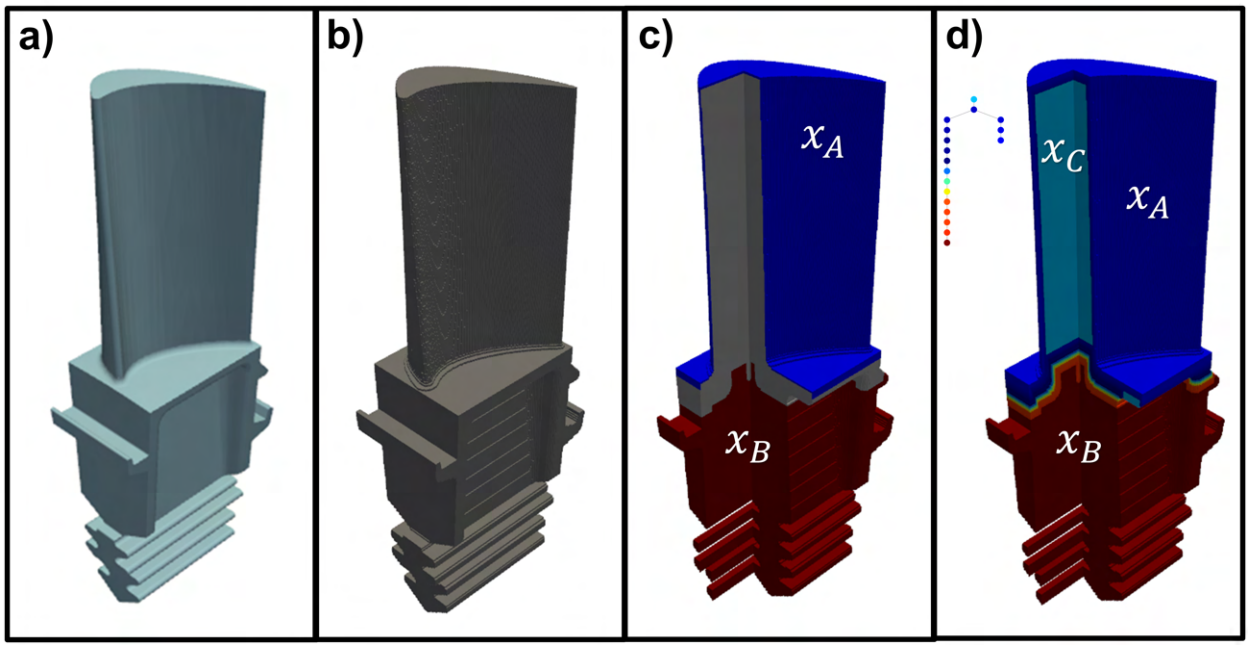

Figure 5.9: Illustration of the gas turbine blade throughout the design process, namely a) the initial STL 3D surface mesh, b) the volumetrically-discretized model from PyVista, c) the discretized model labeled with the spatial placement of terminal alloy compositions x_A & x_B, and d) the final discretized model with conformal CGA placement, colored based on the corresponding $Material\_ID$ from the gradient tree, T_{CGA}, of alloy compositions.

follow the adjacency required by the Steiner tree to ensure single BCC phases, avoiding both deleterious phases and any other undesirable (non-BCC) phases. In Fig. 5.10b, it is evident that σ_{y} does not fall below the 25th percentile value of 167 MPa. Here we also see that x_C in the blade core, and x_B in the base are both possess an extremely high yield strength, above the 95th percentile value of 771 MPa. For oxidation resistance, Fig. 5.10c highlights the placement of x_A at the blade surface with 45 mol % Cr, which is dramatically higher than the other compositions in the gradient. Next, Fig. 5.10d indicates the high creep merit index of x_C. While not as high as some other alloys in the gradient, this value is noteworthy due to the significant yield strength of this alloy. Lastly, Fig. 5.10e indicates a similar situation for the x_B concerning the Pugh ratio. Its magnitude is not the highest, but it is noteworthy when considering the exceptional yield strength of the alloy. Overall, the **resultant design has a combination of properties that would not be possible if limited to one alloy composition for conventional manufacturing**. These results

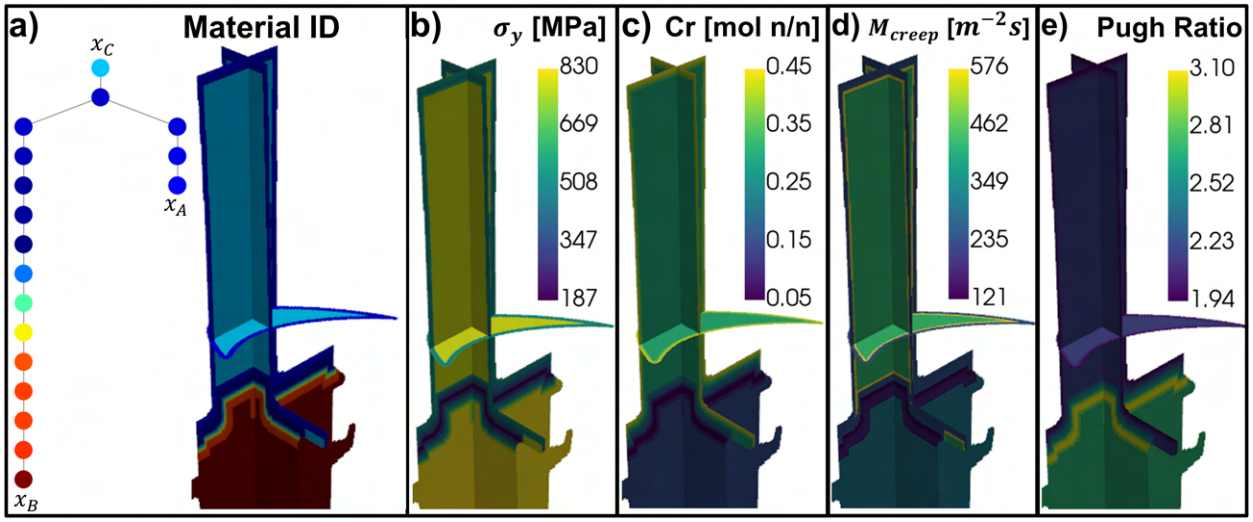

Figure 5.10: Orthogonal section views showing the spatial distribution of the alloys by their a) $Material\_ID$, b) yield strength, c) chromium content, d) creep merit index, and e) Pugh ratio for the voxels in the final turbine blade design

indicate the utility of the presented workflow for systematic materials design of compositionally graded alloy parts with enhanced performance compared to parts comprised of a singular alloy composition, and extends the consideration of AM for simplifying assemblies even when they are comprised of parts with different alloy compositions.

In addition to observing the final design in light of the design intent, we can also observe some new considerations for CGAs from Fig. 5.10 concerning the design of physical parts. It is evident from the cross-sections shown in Fig. 5.10 that the number of materials in a gradient translates to more physical space required, illustrated by the small gradient region between x_A & x_C compared to x_B & x_C. This indicates the practical value of the simple Euclidean composition distance cost function used for several early works on computational CGA design [12, 64]. If the gradient between x_A & x_B had required many more gradient material steps, x_B may have occupied an inconsequential amount of space on the interior of the blade, if there was any space at all. Since the discretization was based on the spatial, compositional resolution of a real metal AM machine [20], this also points to the benefit of using AM technologies with finer printing resolution. These technologies could deposit compositional gradients with a given number of discrete steps over a

smaller physical volume.

It is also apparent that using $m_c = x_C$ did result in the use of x_C elsewhere besides the blade core, which indicates additional space in this portion of the structure beyond the required space necessary for placing the gradient compositions. Figure 5.10a and Figure 5.9d both show that the gradient also proceeds to x_C elsewhere in the structure besides the blade core. This is apparent by observing the presence of the light blue color corresponding to x_C both in the blade core as intended but also in the central overhanging region at the base of the blade. If the use of x_C is undesirable in this region, the TreeMAP inputs can be revised (such as using a different initial material labeling) and used to re-run TreeMAP to alter the resultant structural design.

Overall, the results show that the combined capabilities of recent methods can be used for the property-driven design of complex CGA parts with enhanced performance capabilities. Based on the materials data collected, these capabilities would be unattainable using any other single-phase BCC alloy composition in this alloy system. If current efforts [150] to develop complementary capabilities in manufacturing technology are successful, these structures may soon become a physical reality.

5.3 Chapter Summary

Ultimately, the integration of the novel components of this work, namely 1) ATLAS, 2) multi-terminal CGA design, and 3) TreeMAP, results in a framework capable of systematic design of CGAs and corresponding structures. Identifying the feasible constrained regions with ATLAS helped narrow the relevant, feasible alloy processing space to a specifically constrained subgraph that passed all the relevant structural performance criteria. From this subgraph, the optimal terminal alloys were identified based on given design constraints and objectives. Then, a CGA was designed using these terminal nodal $Alloy$ compositions and a novel cost function to improve manufacturability by minimizing the worst-case magnitude of a relevant cracking susceptibility criterion. This approach, supported by state-of-the-art property models and an ML regressor to aid in calculating the cost function and missing/erroneous property values, produced a compositional gradient given as a graph tree. This CGA was then mapped into a discretized 3D gas turbine blade

structure using TreeMAP, yielding a structural design with performance characteristics that are not attainable using a single material composition from this same alloy system.

Additionally, the spatial design of the structure indicated new considerations. Specifically, considering the number of required material steps in the CGA and the spatial placement and discretization became important considerations to ensure a) the gradient can be placed and b) the terminal alloy compositions can occupy adequate portions of the final design. While a cost function for minimizing the Euclidean compositional distance has been used prior, this is the first known illustration of its relevance in structural design. While the designed structures in this work extend beyond current manufacturing capabilities, they serve as motivating examples that could be realized if current, related research efforts in manufacturing software and technology are successful [150, 19].

6. CONCLUSION

6.1 Summary of Research Contributions

Compositionally graded alloys show significant promise for developing structures with spatially-varying properties for elevated performance compared to traditional single-material structures. Metal additive manufacturing is particularly well-suited for realizing these materials due to the capability for spatial composition control during fabrication. However, there are significant fundamental gaps in literature for realizing compositionally graded alloys and structures, particularly, 1) identifying multi-terminal CGAs to join arbitrarily many compositions in a monolithic structure, and 2) mapping multi-terminal CGAs into corresponding physical structures.

While methods to address the primary materials design challenges of deleterious phase formation and solidification cracking in CGAs, these methods each have notable shortcomings that limit their effectiveness. Furthermore, the complex material requirements for CGAs has limited the use of these materials in the design of engineering structures, particularly those with multi-terminal material requirements beyond the capabilities of current two-terminal design paradigms. This work presents a novel computational framework that leverages the benefits of both materials design approaches, and establishes a novel link between the multi-terminal CGAs and performance-driven structural design, enabling holistic, systematic materials design of CGAs.

The first contribution of this work is a novel information model, the Alloy Topology-Linked informAtion Schema (ATLAS), which was developed to provide the benefits of prior CGA design methods while alleviating their respective shortcomings. Notably, graphs instantiated according to the ATLAS ('ATLAS graphs') satisfy the five schema requirements identified in the analysis originally summarized in Table 2.1, namely: 1) storing and reusing data with database write and read queries, 2&3) capturing the topology of the alloy state space and relating alloy compositions for automated path planning using $Alloy$ nodes and $JOINS\_TO$ edges, 4) scaling to arbitrary data dimensionality using node/edge properties, and 5) updating with new knowledge using write

queries and the flexible nature of graph databases.

In particular, the significant capabilities of ATLAS graphs for topological partitioning and analysis of the alloy state space were demonstrated on several materials datasets. These datasets were constrained based on realistic manufacturability criteria, particularly the avoidance of deleterious brittle phases. This capability of ATLAS graphs provided significant insight, particularly in high dimensions, and was paired with dimensionality reduction techniques for 2D visualizations such as UMAP. Once calculated for a given set of constraints, the constrained subgraph partitioning can be used to determine if nodal $Alloy$ compositions can be joined with a feasible compositional gradient, before moving forward with the design of a specific compositional gradient. This capability eliminates trial and error in the computational gradient design process and was not possible beyond ternary alloy systems using prior methods.

Secondly, this work contributed a novel formulation of the multi-terminal CGA materials design problem as a minimum Steiner tree problem in graphs. This enables the optimization of a CGA for an arbitrary number of terminal nodal $Alloy$ compositions simultaneously, and a resultant CGA represented as a graph tree. This concept can also be extended to designing gradients as minimal spanning trees for experimental exploration and characterization of the alloy state space.

Third, based on the tree representation of the CGA, the TreeMAP algorithm was proposed to map a multi-terminal CGA into a corresponding physical structure based on initial disparate labeled alloy regions, particularly for the terminal compositions. The soundness and completeness of this algorithm were discussed, and the complete algorithm was shown to run quickly even on larger-scale real-world design problems.

Lastly, the novel proposed methods of ATLAS, multi-terminal CGA design, and the TreeMAP algorithm were unified into a singular framework and used to demonstrate systematic design of a compositionally graded alloy and corresponding structure for specific structural performance criteria. This ICME framework demonstrated novel capabilities for both material and structural design, and notably extends beyond the capabilities of current methods, enabling systematic design of multi-terminal CGAs and structures for the first time.

6.2 Limitations & Future Work

Although this framework enables the novel design of 3D compositionally graded parts, the computational design capabilities in this field are outpacing the manufacturing technology, leaving a gap in the experimental realization of these novel structures. Furthermore, while the proposed framework enables a host of novel capabilities and advantages compared to prior techniques, it has several noteworthy limitations regarding its use in practice. The following sections will outline these limitations and suggest potential future research to address them.

6.2.1 Manufacturing Limitations

6.2.1.1 Compositional Control

Current AM technology is insufficient in several key areas for realizing the multi-terminal compositionally graded structures designed in this work. Perhaps the most significant hindrance is the lack of multi-material slicing and process planning software. This need has been clearly illustrated by the recent June 2024 Small Business Innovation Research (SBIR) solicitation by the US Army with the objective of developing software to process CAD models with user-defined gradients in composition and automatically generate a AM tool path for grading in at least 3 directions [150].

In addition to software, the primary bottleneck with respect to hardware is the limited compositional degrees of freedom in many metal AM systems. Among commercial LP-DED machines, which are the most mature for manufacturing CGAs, the typical limit on independent compositional feeds is four. While there are some specialized systems for alloy development which increase this number, such as the Alloy Development Feeder by FormAlloy [151], these systems have not been broadly adopted. Another obstacle to manufacturing CGAs is the limited availability of custom metal powders. Due to the unique design objectives of compositionally grading, CGAs frequently use eccentric compositions which are not typically available off-the-shelf. These compositions can be made by blending alloy feedstocks in-situ, and in the case of powder-based AM they can be mixed in a feeder or hopper prior to fabrication. However, these approaches typically result in greater compositional deviation from the planned path compared to using pre-alloyed

feedstock. Thus, increased adoption of lab-scale powder production methods would aid in the feasibility and quality of many CGAs.

6.2.1.2 Digital Thread Integration

One particular important topic in AM is the so-called digital thread [152], defined in the 2013 Global Horizons Report as "the use of digital tools and representations for design, evaluation, and life cycle management" [153]. The digital thread for AM includes solutions like STL for representing geometry, and G-code for setting production tool paths [152]. Additional standard geometric representations have been proposed, such as STL 2.0, which can handle multiple and graded materials compared to previous representations [154], though its use in practice has been limited [152]. Standards that integrate the entire digital chain, comprised of computer aided design (CAD), computer aided process planning (CAPP), computer aided manufacturing (CAM), computer numerical control (CNC), and coordinate-measuring machine (CMM), have also been proposed, such as the Standard for the Exchange of Product model data compliant Numerical Control (STEP-NC) [155, 156, 152]. Ultimately, data continuity across this digital thread including materials, design geometry in CAD, manufacturing, post-processing, qualification, and implementation enables more standardized workflows for metal AM across a wider variety of relevant use cases. As the tools and representations in the AM digital thread continue to mature, they should include considerations for multi-material design and manufacturing in order to span the breadth of metal AM use cases.

6.2.2 Information Schema

6.2.2.1 Processing Parameters

In the proposed ATLAS, a single processing condition is assumed for $Alloy$ nodes since the current practice for manufacturing CGAs generally targets a single set of AM processing parameters to simplify the manufacturing process [63]. Due to the flexibility of GDBs, the proposed schema can be readily adapted in the future as process control for relevant methods like LP-DED becomes better understood and predicted for unseen compositions. While experimental printability

assessment methods have been proposed for both powder bed methods [157] and powder blown methods [158], rigorous processing prediction for unseen compositions has only been recently proposed for laser powder bed fusion (LPBF) [159, 160, 161], which is considered the most mature metal AM technique [162]. As process prediction improves for LP-DED in future work, or LPBF becomes more popular for CGA manufacturing, this consideration could be integrated into the schema by adding processing parameter dimensions to the alloy state space, where $Alloy$ nodes are redefined as unique compositions subject to *specific processing conditions*. The definition of $JOINS\_TO$ edges could likewise be extended to also connect $Alloy$ nodes of similar composition with compatible processing conditions so that a CGA design can be performed in all alloy chemistry and processing degrees of freedom for further control over the resultant structure and properties. This is particularly relevant for integrating gradient design with structural design, since process parameters may need to be varied throughout a structure to mitigate undesirable effects, such as the buildup of residual thermal stresses.

6.2.2.2 Blended Alloy Feedstocks

The current implementation of this work focuses on a global perspective of the alloy state space, particularly with respect to chemistry, by using the pure elemental components for constructing simplicial grid graphs. With this approach, the entire range of combinations for a given set of alloy components is represented in the ATLAS materials graph. However, in many instances it is more cost-effective in manufacturing to use one or more alloyed feedstocks with a fixed composition ratio. Thus, a graph representation with stepwise changes according to these feedstock ratios is beneficial for designing CGAs. While an ATLAS materials graph could be extended to use these ratios, ratio-based grid points would need to either be stored in addition to a fixed elemental grid, which would increase the complexity, or one would need to replace portions of the elemental grid which may reduce the breadth of the alloy state space represented in the ATLAS graph. If either of these issues is not prohibitive, then these could be taken as potential solutions. As digital storage capabilities locally and on the cloud continue to improve, the storage of redundant elemental and alloyed grid samplings may not be prohibitive, but this extension may require strategic filtering

prior to performing graph queries that are sensitive to the number of nodes and edges.

Since these blended alloys are primarily needed from the standpoint of cost and manufacturing practicality, there are several simple solutions which could be added to the existing methods for two-terminal or multi-terminal gradient design to find blended alloy gradients while maintaining an elemental ATLAS graph database. To find a CGA with blended alloy feedstocks in mind, one could generate a simplicial grid using potential blended feedstock vertices to generate a separate materials graph. The nodal $Alloy$ compositions and $JOINS\_TO$ edges in the blended alloy ATLAS materials graph could then be assessed for the necessary labels and properties using methods such as CALPHAD (expensive), or using surrogate models trained on the existing elemental ATLAS graph data (relatively inexpensive). Then, a blended alloy CGA could be identified in one of several ways.

One approach would be to identify the gradient path or tree on the elemental ATLAS materials graph and then map it to blended feedstocks by finding the nearest feasible nodal $Alloy$ composition in the blended alloy ATLAS materials graph, repeating for each $Alloy$ node in the elemental gradient path/tree. This does not consider the path or tree structure though, so it could result in the elemental path or tree becoming invalid. Nevertheless, it does benefit from using the existing data collected for the ATLAS graph for the initial CGA design.

Another approach would be to find the nearest feasible neighbors of the terminal alloy compositions in the blended feedstock graph, and then compute the gradient tree path or tree in the blended alloy materials graph. This method guarantees a gradient if the mapped terminals are still in the same subgraph. However, it requires the computation of the cost function over the new edges in the blended alloy materials graph. Though, it is worth noting that the computational expense of this calculation can be limited by restricting it to only the blended alloy constrained subgraph which contains the terminal alloys.

6.2.3 Structural Design

6.2.3.1 Volumetric Discretization

While the voxelization method used in this dissertation was successful in discretizing 3D parts in preparation for the TreeMAP algorithm, this approach significantly reduced the representative quality of the 3D model compared to the original 3D STL representation. While this is arguably a realistic loss of detail if the voxel step sizes are tied to the AM print resolution, it does not account for the conformal tool path planning capabilities of some AM methods. It also does not account for the initial spatial placement of the terminal alloys, which could be simplified into larger volumetric cells to reduce the complexity of the discretized model and corresponding spatial part graph G_P.

Thus, improved discretization strategies could refine the gradient placement process. As computational materials design becomes better integrated into the digital chain, discretization strategies which leverage information such as tool path planning, process-specific print resolution, spatial placement of terminal alloys, and necessary gradient material steps as prior information could be used to adaptively tailor the part discretization in cooperation with both the materials design and manufacturing considerations.

6.2.3.2 Terminal Alloy Labeling

Placing terminal alloys in a discretized part is difficult without the use of a CAD interface due to the need for tailoring the volumetric partitioning and selecting volumetric cells to label with terminal alloy compositions. Simple geometric filters can be implemented relatively easily using spatial filters in Python for the PyVista [132] interface for the Visualization Toolkit (VTK) [133], but complex interior surfaces and structural geometries are more challenging to address. The proposed framework would greatly benefit from a refined interface/approach to labeling the initial material placement in a discretized structure before proceeding with the TreeMAP algorithm.

6.2.4 Structural Performance Metrics

Since this work links the design of multi-terminal gradients with structural design, there are a multitude of metrics which can now be considered during the design of engineering structures.

These metrics include, but are not limited to, safety, functionality, lifespan, risk and cost, which are dependent on the structural geometry and properties. Due to the core importance of these metrics for assessing the holistic merit of a structural design, these performance criteria should be further integrated in the design process.

The first step to integrating these types of criteria is by computing them. Thus, the most notable addition would be the use of finite element analysis (FEA), which can be used to perform simulated assessment of of performance metrics such as factor of safety, compliance, heat transfer coefficient, and maximum temperature. These metrics would help assess the merit of a given structural design under its operating conditions.

Metrics such as raw material criticality [21, 22], and material cost are also significant in the design of engineering structures. These metrics in particular have a significant relationship with the spatial placement of a CGA when other cost-related variables such as the manufacturing method are fixed. Since the elemental composition is known volumetrically with the gradient assigned to a discretized structure, the total cost or criticality of a designed component can be computed and used to assess the merit of various structural designs. A structure using a gradient that uses a relatively low proportion of an expensive material, but requires many discrete compositional layers for manufacturing, may be significantly more expensive in practice than a gradient which uses a large proportion of expensive material with much fewer discrete compositional layers. For example, there is recent gradient design work which uses Ag as an intermediary for joining Al and Cu [13]. While at face value this could be considered a high-cost gradient due to the inclusion of Ag, the economical viability of this CGA can only be truly assessed by evaluating its merits when used in a physical structure.

6.2.5 Topology Optimization

Lastly, this work did not consider the physical geometry or topology of the structures as a variable in the design process, and instead focused exclusively on the materials design of a corresponding gradient path and its placement within fixed 3D structural geometries. It was noted in Chapter 1 that the primary limitation of existing TO methods is their incapability to accurately

represent the complexity of the alloy processing landscape with respect to feasibility (e.g. avoiding deleterious phases), and properties. The proposed ATLAS enables the development of GDBs as living repositories of materials data, with significant capabilities for targeted experimental and computational assessment for efficiently developing a tailored data representation of the alloy processing space. This rich data representation could be used to train ML models to predict alloy feasibility and properties for materials-informed TO with CGAs.

REFERENCES

[1] M. Allen, R. Arróyave, and R. Malak, "A Graph Algorithm for the Design of Functionally Graded Metal Components," *Volume 3A: 50th Design Automation Conference (DAC)*, Aug. 2024.

[2] M. Allen, R. Arróyave, and R. Malak, "A Graph Database Schema for Metal Additive Informatics," *Volume 2B: 44th Computers and Information in Engineering Conference (CIE)*, Aug. 2024.

[3] G. B. Olson, "Computational design of hierarchically structured materials," *Science*, vol. 277, no. 5330, pp. 1237–1242, 1997.

[4] W. Zuo and K. Saitou, "Multi-material topology optimization using ordered SIMP interpolation," *Structural and Multidisciplinary Optimization*, vol. 55, pp. 477–491, Feb. 2017.

[5] J.-O. Andersson, T. Helander, L. Höglund, P. Shi, and B. Sundman, "Thermo-calc & dictra, computational tools for materials science," *Calphad*, vol. 26, no. 2, pp. 273–312, 2002.

[6] Thermo-Calc Software, "Thermodynamic database for high entropy alloys (tchea) version 6," 2024.

[7] A. A. Hagberg, D. A. Schult, and P. J. Swart, "Exploring network structure, dynamics, and function using networkx," *Proceedings of the 7th Python in Science Conference*, pp. 11 – 15, 2008.

[8] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proceedings of the American Mathematical Society*, vol. 7, no. 1, pp. 48–50, 1956.

[9] D. C. Hofmann, J. Kolodziejska, S. Roberts, R. Otis, R. P. Dillon, J.-O. Suh, Z.-K. Liu, and J.-P. Borgonia, "Compositionally graded metals: A new frontier of additive manufacturing," *Journal of Materials Research*, vol. 29, pp. 1899–1910, Sept. 2014.

[10] A. Reichardt, A. A. Shapiro, R. Otis, R. P. Dillon, J. P. Borgonia, B. W. McEnerney, P. Hosemann, and A. M. Beese, "Advances in additive manufacturing of metal-based functionally graded materials," *International Materials Reviews*, vol. 66, pp. 1–29, Jan. 2021.

[11] T. M. Pollock, J. E. Allison, D. G. Backman, M. C. Boyce, M. Gersh, E. A. Holm, R. Lesar, M. Long, A. C. Powell IV, J. Schirra, D. Whitis, and C. Woodward, *Integrated Computational Materials Engineering: A Transformational Discipline for Improved Competitiveness and National Security*. Washington, D.C., United States: National Academies Press, 2008.

[12] T. Kirk, E. Galvan, R. Malak, and R. Arróyave, "Computational Design of Gradient Paths in Additively Manufactured Functionally Graded Materials," *Journal of Mechanical Design*, vol. 140, Sept. 2018.

[13] H. Sun, B. Pan, Z. Yang, A. M. Krajewski, B. Bocklund, S.-L. Shang, J. Li, A. M. Beese, and Z.-K. Liu, "Materialsmap: A calphad-based tool to design composition pathways through feasibility map for desired dissimilar materials, demonstrated with resistance spot welding joining of ag-al-cu," *Materialia*, vol. 36, p. 102153, 2024.

[14] M. Allen, T. Kirk, R. Malak, and R. Arróyave, "A Subspace-Inclusive Sampling Method for the Computational Design of Compositionally Graded Alloys," *Journal of Mechanical Design*, vol. 144, p. 041704, Feb. 2022.

[15] S. G. Price, Z. Cao, and I. McCue, "A novel path planning framework for constructing compositional gradients in many-element systems," in *Proceedings of the SMST 2024: Extended Abstracts from the International Conference on Shape Memory and Superelastic Technologies*, Shape Memory Proceedings, (Cascais, Portugal), pp. 90–91, ASM, May 2024.

[16] D. C. Hofmann, S. Roberts, R. Otis, J. Kolodziejska, R. P. Dillon, J.-o. Suh, A. A. Shapiro, Z.-K. Liu, and J.-P. Borgonia, "Developing Gradient Metal Alloys through Radial Deposition Additive Manufacturing," *Scientific Reports*, vol. 4, p. 5357, June 2014.

[17] J. Walker, J. R. Middendorf, C. C. Lesko, and J. Gockel, "Multi-material laser powder bed fusion additive manufacturing in 3-dimensions," *Manufacturing Letters*, vol. 31, pp. 74–77,

Jan. 2022.

[18] B. Vela, C. Acemi, P. Singh, T. Kirk, W. Trehern, E. Norris, D. D. Johnson, I. Karaman, and R. Arróyave, "High-throughput exploration of the wmovtanbal refractory multi-principal-element alloys under multiple-property constraints," *Acta Materialia*, vol. 248, p. 118784, 2023.

[19] Defense Advanced Research Projects Agency (DARPA), "Multiobjective engineering and testing of alloy structures (metals)," 2023. Broad Agency Announcement HR001123S0029.

[20] D. Salas, D. Ebeperi, M. Elverud, R. Arróyave, R. J. Malak, and I. Karaman, "Embedding hidden information in additively manufactured metals via magnetic property grading for traceability," *Additive Manufacturing*, vol. 60, p. 103261, Dec. 2022.

[21] S. Glöser, L. Tercero Espinoza, C. Gandenberger, and M. Faulstich, "Raw material criticality in the context of classical risk assessment," *Resources Policy*, vol. 44, pp. 35–46, June 2015.

[22] D. Schrijvers, A. Hool, G. A. Blengini, W.-Q. Chen, J. Dewulf, R. Eggert, L. van Ellen, R. Gauss, J. Goddin, K. Habib, C. Hagelüken, A. Hirohata, M. Hofmann-Amtenbrink, J. Kosmol, M. Le Gleuher, M. Grohol, A. Ku, M.-H. Lee, G. Liu, K. Nansai, P. Nuss, D. Peck, A. Reller, G. Sonnemann, L. Tercero, A. Thorenz, and P. A. Wäger, "A review of methods and data to determine raw material criticality," *Resources, Conservation and Recycling*, vol. 155, p. 104617, 2020.

[23] A. Reichardt, R. P. Dillon, J. P. Borgonia, A. A. Shapiro, B. W. McEnerney, T. Momose, and P. Hosemann, "Development and characterization of Ti-6Al-4V to 304L stainless steel gradient components fabricated with laser deposition additive manufacturing," *Materials & Design*, vol. 104, pp. 404–413, Aug. 2016.

[24] B. E. Carroll, R. A. Otis, J. P. Borgonia, J.-o. Suh, R. P. Dillon, A. A. Shapiro, D. C. Hofmann, Z.-K. Liu, and A. M. Beese, "Functionally graded material of 304L stainless steel and inconel 625 fabricated by directed energy deposition: Characterization and thermodynamic modeling," *Acta Materialia*, vol. 108, pp. 46–54, Apr. 2016.

[25] B. Chen, Y. Su, Z. Xie, C. Tan, and J. Feng, "Development and characterization of 316L/Inconel625 functionally graded material fabricated by laser direct metal deposition," *Optics & Laser Technology*, vol. 123, p. 105916, Mar. 2020.

[26] W. Meng, Y. Xiaohui, W. Zhang, F. Junfei, G. Lijie, M. Qunshuang, and C. Bing, "Additive manufacturing of a functionally graded material from Inconel625 to Ti6Al4V by laser synchronous preheating," *Journal of Materials Processing Technology*, vol. 275, Jan. 2020.

[27] M. Ansari, E. Jabari, and E. Toyserkani, "Opportunities and challenges in additive manufacturing of functionally graded metallic materials via powder-fed laser directed energy deposition: A review," *Journal of Materials Processing Technology*, vol. 294, p. 117117, Aug. 2021.

[28] Z. Sun and J. C. Ion, "Laser welding of dissimilar metal combinations," *Journal of Materials Science*, vol. 30, pp. 4205–4214, Sept. 1995.

[29] Y. Fang, X. Jiang, D. Mo, D. Zhu, and Z. Luo, "A review on dissimilar metals' welding methods and mechanisms with interlayer," *The International Journal of Advanced Manufacturing Technology*, vol. 102, pp. 2845–2863, June 2019.

[30] B. Blakey-Milner, P. Gradl, G. Snedden, M. Brooks, J. Pitot, E. Lopez, M. Leary, F. Berto, and A. du Plessis, "Metal additive manufacturing in aerospace: A review," *Materials & Design*, vol. 209, p. 110008, 2021.

[31] W. E. Frazier, "Metal Additive Manufacturing: A Review," *Journal of Materials Engineering and Performance*, vol. 23, pp. 1917–1928, June 2014.

[32] L. Li, W. Syed, and A. Pinkerton, "Rapid additive manufacturing of functionally graded structures using simultaneous wire and powder laser deposition," *Virtual and Physical Prototyping*, vol. 1, pp. 217–225, Dec. 2006.

[33] W. U. H. Syed, A. J. Pinkerton, Z. Liu, and L. Li, "Coincident wire and powder deposition by laser to form compositionally graded material," *Surface and Coatings Technology*, vol. 201, pp. 7083–7091, May 2007.

[34] F. Wang, J. Mei, and X. Wu, "Compositionally graded Ti6Al4V+TiC made by direct laser fabrication using powder and wire," *Materials & Design*, vol. 28, pp. 2040–2046, Jan. 2007.

[35] M. Simonelli, N. Aboulkhair, M. Rasa, M. East, C. Tuck, R. Wildman, O. Salomons, and R. Hague, "Towards digital metal additive manufacturing via high-temperature drop-on-demand jetting," *Additive Manufacturing*, vol. 30, p. 100930, 2019.

[36] N. Gilani, A. Foerster, and N. T. Aboulkhair, "Material jetting," in *Springer Handbook of Additive Manufacturing* (E. Pei, A. Bernard, D. Gu, C. Klahn, M. Monzón, M. Petersen, and T. Sun, eds.), pp. 371–388, Springer, 2023.

[37] K. I. Schwendner, R. Banerjee, P. C. Collins, C. A. Brice, and H. L. Fraser, "Direct laser deposition of alloys from elemental powder blends," *Scripta Materialia*, vol. 45, pp. 1123–1129, Nov. 2001.

[38] R. Banerjee, P. C. Collins, D. Bhattacharyya, S. Banerjee, and H. L. Fraser, "Microstructural evolution in laser deposited compositionally graded α/β titanium-vanadium alloys," *Acta Materialia*, vol. 51, pp. 3277–3292, June 2003.

[39] P. C. Collins, R. Banerjee, S. Banerjee, and H. L. Fraser, "Laser deposition of compositionally graded titanium–vanadium and titanium–molybdenum alloys," *Materials Science and Engineering: A*, vol. 352, pp. 118–128, July 2003.

[40] L. D. Bobbio, B. Bocklund, R. Otis, J. P. Borgonia, R. P. Dillon, A. A. Shapiro, B. McEnerney, Z.-K. Liu, and A. M. Beese, "Characterization of a functionally graded material of Ti-6Al-4V to 304L stainless steel with an intermediate V section," *Journal of Alloys and Compounds*, vol. 742, pp. 1031–1036, Apr. 2018.

[41] L. D. Bobbio, R. A. Otis, J. P. Borgonia, R. P. Dillon, A. A. Shapiro, Z.-K. Liu, and A. M. Beese, "Additive manufacturing of a functionally graded material from Ti-6Al-4V to Invar: Experimental characterization and thermodynamic calculations," *Acta Materialia*, vol. 127, pp. 133–142, Apr. 2017.

[42] F. Wei, S. Wei, K. B. Lau, W. H. Teh, J. J. Lee, H. L. Seng, C. C. Tan, P. Wang, and U. Ramamurty, "Compositionally graded AlxCoCrFeNi high-entropy alloy manufactured by laser powder bed fusion," *Materialia*, vol. 21, p. 101308, Mar. 2022.

[43] Z. Wu, A. E. Wilson-Heid, R. J. Griffiths, and E. S. Elton, "A review on experimentally observed mechanical and microstructural characteristics of interfaces in multi-material laser powder bed fusion," *Frontiers in Mechanical Engineering*, vol. 9, p. 1087021, July 2023.

[44] J. Wang, Z. Pan, Y. Ma, Y. Lu, C. Shen, D. Cuiuri, and H. Li, "Characterization of wire arc additively manufactured titanium aluminide functionally graded material: Microstructure, mechanical properties and oxidation behaviour," *Materials Science and Engineering: A*, vol. 734, pp. 110–119, Sept. 2018.

[45] T. A. Rodrigues, N. Bairrão, F. W. C. Farias, A. Shamsolhodaei, J. Shen, N. Zhou, E. Maawad, N. Schell, T. G. Santos, and J. P. Oliveira, "Steel-copper functionally graded material produced by twin-wire and arc additive manufacturing (T-WAAM)," *Materials & Design*, vol. 213, p. 110270, Jan. 2022.

[46] D. J. John, K. R. Sajja Rama, L. K. Adhimoolam, and J. Amalraj, "Studies on In Situ Alloy Formation Using Mild Steel–Inconel 625 Twin Filler Wire Gas Tungsten Arc Weld Deposition," *Engineering Proceedings*, vol. 61, no. 1, p. 28, 2024.

[47] H. Dobbelstein, E. L. Gurevich, E. P. George, A. Ostendorf, and G. Laplanche, "Laser metal deposition of compositionally graded TiZrNbTa refractory high-entropy alloys using elemental powder blends," *Additive Manufacturing*, vol. 25, pp. 252–262, Jan. 2019.

[48] Y. Zhao, K. B. Lau, W. H. Teh, J. J. Lee, F. Wei, M. Lin, P. Wang, C. C. Tan, and U. Ramamurty, "Compositionally graded CoCrFeNiTi high-entropy alloys manufactured by laser powder bed fusion: A combinatorial assessment," *Journal of Alloys and Compounds*, vol. 883, p. 160825, Nov. 2021.

[49] E. Stach, B. DeCost, A. G. Kusne, J. Hattrick-Simpers, K. A. Brown, K. G. Reyes, J. Schrier, S. Billinge, T. Buonassisi, I. Foster, C. P. Gomes, J. M. Gregoire, A. Mehta, J. Montoya,

E. Olivetti, C. Park, E. Rotenberg, S. K. Saikin, S. Smullin, V. Stanev, and B. Maruyama, "Autonomous experimentation systems for materials development: A community perspective," *Matter*, vol. 4, pp. 2702–2726, Sept. 2021.

[50] S. C. Tjong and Y.-W. Mai, "Processing-structure-property aspects of particulate- and whisker-reinforced titanium matrix composites," *Composites Science and Technology*, vol. 68, pp. 583–601, Mar. 2008.

[51] D. de Faoite, D. J. Browne, F. R. Chang-Díaz, and K. T. Stanton, "A review of the processing, composition, and temperature-dependent mechanical and thermal properties of dielectric technical ceramics," *Journal of Materials Science*, vol. 47, pp. 4211–4235, May 2012.

[52] J. Smith, W. Xiong, W. Yan, S. Lin, P. Cheng, O. L. Kafka, G. J. Wagner, J. Cao, and W. K. Liu, "Linking process, structure, property, and performance for metal-based additive manufacturing: Computational approaches with experimental support," *Computational Mechanics*, vol. 57, pp. 583–610, Apr. 2016.

[53] S. N. Leung, "Thermally conductive polymer composites and nanocomposites: Processing-structure-property relationships," *Composites Part B: Engineering*, vol. 150, pp. 78–92, Oct. 2018.

[54] G. D. Goh, Y. L. Yap, H. K. J. Tan, S. L. Sing, G. L. Goh, and W. Y. Yeong, "Process–Structure–Properties in Polymer Additive Manufacturing via Material Extrusion: A Review," *Critical Reviews in Solid State and Materials Sciences*, vol. 45, pp. 113–133, Mar. 2020.

[55] A. R. Moustafa, A. Durga, G. Lindwall, and Z. C. Cordero, "Scheil ternary projection (STeP) diagrams for designing additively manufactured functionally graded metals," *Additive Manufacturing*, vol. 32, Mar. 2020.

[56] B. Bocklund, L. D. Bobbio, R. A. Otis, A. M. Beese, and Z.-K. Liu, "Experimental validation of scheil–gulliver simulations for gradient path planning in additively manufactured

functionally graded materials," *Materialia*, vol. 11, p. 100689, 2020.

[57] B. Bocklund, *Computational design of additively manufactured functionally graded materials by thermodynamic modeling with uncertainty quantification*. PhD thesis, The Pennsylvania State University, 2021.

[58] L. D. Bobbio, B. Bocklund, E. Simsek, R. T. Ott, M. J. Kramer, Z.-K. Liu, and A. M. Beese, "Design of an additively manufactured functionally graded material of 316 stainless steel and Ti-6Al-4V with Ni-20Cr, Cr, and V intermediate compositions," *Additive Manufacturing*, vol. 51, p. 102649, Mar. 2022.

[59] Z. Yang, H. Sun, Z.-K. Liu, and A. M. Beese, "Design methodology for functionally graded materials: Framework for considering cracking," *Additive Manufacturing*, vol. 73, p. 103672, 2023.

[60] T. Kirk, R. Malak, and R. Arróyave, "Computational Design of Compositionally Graded Alloys for Property Monotonicity," *Journal of Mechanical Design*, vol. 143, Nov. 2020.

[61] O. Adiyatov and H. A. Varol, "Rapidly-exploring random tree based memory efficient motion planning," *2013 IEEE International Conference on Mechatronics and Automation*, pp. 354–359, 2013.

[62] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," *The International Journal of Robotics Research*, vol. 30, pp. 846–894, June 2011.

[63] O. V. Eliseeva, T. Kirk, P. Samimi, R. Malak, R. Arróyave, A. Elwany, and I. Karaman, "Functionally Graded Materials through robotics-inspired path planning," *Materials & Design*, vol. 182, p. 107975, Nov. 2019.

[64] T. Q. Kirk, *Computational Design of Compositionally Graded Alloys*. PhD thesis, Texas A&M University, Department of Mechanical Engineering, Dec. 2020.

[65] R. Tavakoli and S. M. Mohseni, "Alternating active-phase algorithm for multimaterial topology optimization problems: A 115-line MATLAB implementation," *Structural and Multidisciplinary Optimization*, vol. 49, pp. 621–642, Apr. 2014.

[66] C. Woischwill and I. Y. Kim, "Multimaterial multijoint topology optimization," *International Journal for Numerical Methods in Engineering*, vol. 115, no. 13, pp. 1552–1579, 2018.

[67] L. Crispo, S. W. K. Roper, R. Bohrer, R. Morin, and I. Y. Kim, "Multi-Material and Multi-Joint Topology Optimization for Lightweight and Cost-Effective Design," *Volume 3B: 47th Design Automation Conference (DAC)*, Aug. 2021.

[68] Q. Xia and M. Y. Wang, "Simultaneous optimization of the material properties and the topology of functionally graded structures," *Computer-Aided Design*, vol. 40, pp. 660–675, June 2008.

[69] A. Radman, X. Huang, and Y. M. Xie, "Topology optimization of functionally graded cellular materials," *Journal of Materials Science*, vol. 48, pp. 1503–1510, Feb. 2013.

[70] T. Liu, S. Guessasma, J. Zhu, W. Zhang, and S. Belhabib, "Functionally graded materials from topology optimisation and stereolithography," *European Polymer Journal*, vol. 108, pp. 199–211, Nov. 2018.

[71] Y.-C. Chan, D. Da, L. Wang, and W. Chen, "Remixing functionally graded structures: Data-driven topology optimization with multiclass shape blending," *Structural and Multidisciplinary Optimization*, vol. 65, p. 135, Apr. 2022.

[72] R. F. Silva, P. G. Coelho, F. M. Conde, B. R. Santos, and J. P. Oliveira, "Minimizing the maximum von mises stress of elastic continuum structures using topology optimization and additively manufactured functionally graded materials," *Computers & Structures*, vol. 301, p. 107469, 2024.

[73] Z. Hashin and S. Shtrikman, "A variational approach to the theory of the elastic behaviour of multiphase materials," *Journal of the Mechanics and Physics of Solids*, vol. 11, no. 2, pp. 127–140, 1963.

[74] S. Paul, A. Mitra, and C. Koner, "A Review on Graph Database and its representation," in *2019 International Conference on Recent Advances in Energy-efficient Computing and Communication (ICRAECC)*, (Nagercoil, India), pp. 1–5, IEEE, Mar. 2019.

[75] M. S. Rahman, *Basic Graph Theory*. Undergraduate Topics in Computer Science, Springer Cham, 2017.

[76] S. E. Dreyfus and R. A. Wagner, "The steiner problem in graphs," *Networks*, vol. 1, no. 3, pp. 195–207, 1971.

[77] L. Kou, G. Markowsky, and L. Berman, "A fast algorithm for Steiner trees," *Acta Informatica*, vol. 15, pp. 141–145, June 1981.

[78] K. Mehlhorn, "A faster approximation algorithm for the Steiner problem in graphs," *Information Processing Letters*, vol. 27, pp. 125–128, Mar. 1988.

[79] L. Kaufman and H. Bernstein, *Computer Calculation of Phase Diagrams with Special Reference to Refractory Metals*. Refractory Materials, Academic Press, 1970.

[80] B. Reitz, C. Lotz, N. Gerdes, S. Linke, E. Olsen, K. Pflieger, S. Sohrt, M. Ernst, P. Taschner, J. Neumann, E. Stoll, and L. Overmeyer, "Additive Manufacturing Under Lunar Gravity and Microgravity," *Microgravity Science and Technology*, vol. 33, p. 25, Mar. 2021.

[81] A. Zocca, J. Wilbig, A. Waske, J. Günster, M. P. Widjaja, C. Neumann, M. Clozel, A. Meyer, J. Ding, Z. Zhou, and X. Tian, "Challenges in the technology development for additive manufacturing in space," *Chinese Journal of Mechanical Engineering: Additive Manufacturing Frontiers*, vol. 1, no. 1, p. 100018, 2022.

[82] C. Körner, "Additive manufacturing of metallic components by selective electron beam melting — a review," *International Materials Reviews*, vol. 61, no. 5, pp. 361–377, 2016.

[83] M. Allen, R. Arroyave, and R. Malak, "Deep Ensembles for Modeling Uncertain Phase Constraints In Compositionally Graded Alloy Design," *Volume 3A: 48th Design Automation Conference (DAC)*, Aug. 2022.

[84] H. Ko, Y. Lu, Z. Yang, N. Y. Ndiaye, and P. Witherell, "A framework driven by physics-guided machine learning for process-structure-property causal analytics in additive manufacturing," *Journal of Manufacturing Systems*, vol. 67, pp. 213–228, Apr. 2023.

[85] J. Miller, "Graph Database Applications and Concepts with Neo4j," *SAIS 2013 Proceedings*, May 2013.

[86] B. G. Tudorica and C. Bucur, "A comparison between several NoSQL databases with comments and notes," in *2011 RoEduNet International Conference 10th Edition: Networking in Education and Research*, (Iasi, Romania), pp. 1–5, IEEE, June 2011.

[87] S. Medhi and H. Baruah, "Relational database and graph database: A comparative analysis," *Journal of Process Management. New Technologies*, vol. 5, no. 2, pp. 1–9, 2017.

[88] S. Almabdy, "Comparative analysis of relational and graph databases for social networks," *2018 1st International Conference on Computer Applications & Information Security (ICCAIS)*, pp. 1–4, 2018.

[89] N. Jatana, S. Puri, M. Ahuja, I. Kathuria, and D. Gosain, "A Survey and Comparison of Relational and Non-Relational Database," *International Journal of Engineering Research*, vol. 1, no. 6, 2012.

[90] C. Vicknair, M. Macias, Z. Zhao, X. Nan, Y. Chen, and D. Wilkins, "A comparison of a graph database and a relational database: A data provenance perspective," in *Proceedings of the 48th Annual Southeast Regional Conference*, (Oxford Mississippi), pp. 1–6, ACM, Apr. 2010.

[91] B. L. Boyce and M. D. Uchic, "Progress toward autonomous experimental systems for alloy development," *MRS Bulletin*, vol. 44, pp. 273–280, Apr. 2019.

[92] M. Besta, R. Gerstenberger, E. Peter, M. Fischer, M. Podstawski, C. Barthels, G. Alonso, and T. Hoefler, "Demystifying Graph Databases: Analysis and Taxonomy of Data Organization, System Designs, and Graph Queries," *ACM Computing Surveys*, vol. 56, pp. 1–40, Feb. 2024.

[93] Z. Chen, Y. Wan, Y. Liu, and A. Valera-Medina, "A knowledge graph-supported information fusion approach for multi-faceted conceptual modelling," *Information Fusion*, vol. 101, p. 101985, Jan. 2024.

[94] A. Donkers, D. Yang, and N. Baken, "Linked data for smart homes: Comparing rdf and labeled property graphs," *Proceedings of the 15th International Conference on Web Information Systems and Technologies (WEBIST)*, 06 2020.

[95] Neo4j, Inc., "Neo4j: The graph database platform," 2024. Graph Database Management System.

[96] C. Sharma and R. Sinha, "A Schema-First Formalism for Labeled Property Graph Databases: Enabling Structured Data Loading and Analytics," in *Proceedings of the 6th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies*, (Auckland New Zealand), pp. 71–80, ACM, Dec. 2019.

[97] A. M. Krajewski, A. M. Beese, W. F. Reinhart, and Z.-K. Liu, "Efficient generation of grids and traversal graphs in compositional spaces towards exploration and path planning," *npj Unconventional Computing*, vol. 1, p. 12, Nov. 2024.

[98] A. Lovrencic, "Maximally connected component." Dictionary of Algorithms and Data Structures [online], Nov. 2020.

[99] G. van Rossum, "Python tutorial," Tech. Rep. CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, May 1995.

[100] B. Yin and W. A. Curtin, "First-principles-based prediction of yield strength in the RhIrPdPtNiCu high-entropy alloy," *npj Computational Materials*, vol. 5, p. 14, Feb. 2019.

[101] C. Acemi, B. Vela, E. Norris, W. Trehern, K. C. Atli, C. Cleek, R. Arróyave, and I. Karaman, "Multi-objective, multi-constraint high-throughput design, synthesis, and characterization of tungsten-containing refractory multi-principal element alloys," *Acta Materialia*, vol. 281, p. 120379, 2024.

[102] F. Maresca and W. A. Curtin, "Mechanistic origin of high strength in refractory bcc high entropy alloys up to 1900k," *Acta Materialia*, vol. 182, pp. 235–249, 2020.

[103] J. A. Slotwinski, E. J. Garboczi, P. E. Stutzman, C. F. Ferraris, S. S. Watson, and M. A. Peltz, "Characterization of Metal Powders Used for Additive Manufacturing," *Journal of Research of the National Institute of Standards and Technology*, vol. 119, p. 460, Sept. 2014.

[104] H. M. Rietveld, "Line profiles of neutron powder-diffraction peaks for structure refinement," *Acta Crystallographica*, vol. 22, pp. 151–152, Jan 1967.

[105] H. M. Rietveld, "A profile refinement method for nuclear and magnetic structures," *Journal of Applied Crystallography*, vol. 2, pp. 65–71, Jun 1969.

[106] L. B. McCusker, R. B. Von Dreele, D. E. Cox, D. Louër, and P. Scardi, "Rietveld refinement guidelines," *Journal of Applied Crystallography*, vol. 32, pp. 36–50, Feb 1999.

[107] X. Wang and W. Xiong, "Uncertainty quantification and composition optimization for alloy additive manufacturing through a CALPHAD-based ICME framework," *npj Computational Materials*, vol. 6, pp. 1–11, Dec. 2020.

[108] R. Arroyave, "Ultimate: Birdshot (final technical report)," tech. rep., Texas A & M University, Nov. 2023.

[109] K. S. Vecchio, O. F. Dippo, K. R. Kaufmann, and X. Liu, "High-throughput rapid experimental alloy development (ht-read)," *Acta Materialia*, vol. 221, p. 117352, 2021.

[110] M. Ciglan, A. Averbuch, and L. Hluchy, "Benchmarking Traversal Operations over Graph Databases," *2012 IEEE 28th International Conference on Data Engineering Workshops*, pp. 186–189, Apr. 2012.

[111] R. Otis and Z.-K. Liu, "Pycalphad: CALPHAD-based Computational Thermodynamics in Python," *Journal of Open Research Software*, vol. 5, Jan. 2017.

[112] B. Sundman, U. R. Kattner, M. Palumbo, and S. G. Fries, "Opencalphad-a free thermodynamic software," *Integrating Materials and Manufacturing Innovation*, vol. 4, pp. 1–15, 2015.

[113] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.

[114] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.

[115] B. Haeupler, R. Hladík, V. Rozhoň, R. Tarjan, and J. Tětek, "Universal optimality of dijkstra via beyond-worst-case heaps," *arXiv*, 2024.

[116] Thermo-Calc Software, "Thermo-calc software high entropy alloys database, version 4," 2021. Accessed: February 2021.

[117] Y. Chen, B. Xie, B. Liu, Y. Cao, J. Li, Q. Fang, and P. K. Liaw, "A focused review on engineering application of multi-principal element alloy," *Frontiers in Materials*, vol. 8, p. 816309, 2022.

[118] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne.," *Journal of machine learning research*, vol. 9, no. 11, 2008.

[119] T. Kirk, B. Vela, S. Mehalic, K. Youssef, and R. Arróyave, "Entropy-driven melting point depression in fcc heas," *Scripta Materialia*, vol. 208, p. 114336, 2022.

[120] D. Khatamsaz, B. Vela, P. Singh, D. D. Johnson, D. Allaire, and R. Arróyave, "Multi-objective materials bayesian optimization with active learning of design constraints: Design of ductile refractory multi-principal-element alloys," *Acta Materialia*, vol. 236, p. 118133, 2022.

[121] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv*, 2020.

[122] B. Vela, T. Hastings, M. Allen, and R. Arróyave, "Visualizing high entropy alloy spaces: methods and best practices," *Digital Discovery*, 2025.

[123] R. M. Karp, *Reducibility among Combinatorial Problems*, pp. 85–103. Boston, MA: Springer US, 1972.

[124] D. Rehfeldt and T. Koch, "Implications, conflicts, and reductions for Steiner trees," *Mathematical Programming*, vol. 197, pp. 903–966, Feb. 2023.

[125] T. Polzin, *Algorithms for the Steiner Problem in Networks*. PhD thesis, Universität des Saarlandes, Saarbrücken, Germany, 2003. Advisor: Prof. Dr. Kurt Mehlhorn, MPI für Informatik.

[126] T. Polzin and S. Vahdati-Daneshmand, *Approaches to the Steiner Problem in Networks*, pp. 81–103. Springer Berlin Heidelberg, 2009.

[127] M. P. de Aragão and R. F. Werneck, "On the Implementation of MST-Based Heuristics for the Steiner Problem in Graphs," in *Algorithm Engineering and Experiments* (D. M. Mount and C. Stein, eds.), Lecture Notes in Computer Science, pp. 1–15, Springer Berlin Heidelberg, 2002.

[128] V. Jarník, "O jistém problému minimálním," *Moravská Přírodovědecká Společnost v Brně (Acta Societatis Scientiarum Naturalium Moraviae)*, vol. 6, pp. 57–63, 1930.

[129] R. C. Prim, "Shortest connection networks and some generalizations," *The Bell System Technical Journal*, vol. 36, no. 6, pp. 1389–1401, 1957.

[130] U. Meyer and P. Sanders, "Δ-stepping: A parallelizable shortest path algorithm," *Journal of Algorithms*, vol. 49, pp. 114–152, Oct. 2003.

[131] H. Mao, H.-L. Chen, and Q. Chen, "TCHEA1: A Thermodynamic Database Not Limited for "High Entropy" Alloys," *Journal of Phase Equilibria and Diffusion*, vol. 38, no. 4, pp. 353–368, 2017.

[132] C. B. Sullivan and A. A. Kaszynski, "PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK)," *Journal of Open Source Software*, vol. 4, p. 1450, May 2019.

[133] W. Schroeder, K. Martin, and B. Lorensen, *The Visualization Toolkit (4th ed.)*. Clifton Park, NY, USA: Kitware, 2006.

[134] E. Scheil, "Bemerkungen zur schichtkristallbildung," *International Journal of Materials Research*, vol. 34, no. 3, pp. 70–72, 1942.

[135] G. Gulliver, "The quantitative effect of rapid cooling upon the constitution of binary alloys," *Journal of the Institute of Metals*, vol. 9, pp. 120–157, 1913.

[136] R. Reed, T. Tao, and N. Warnken, "Alloys-by-design: Application to nickel-based single crystal superalloys," *Acta Materialia*, vol. 57, no. 19, pp. 5898–5913, 2009.

[137] M. Rezazadeh Reyhani, M. Alizadeh, A. Fathi, and H. Khaledi, "Turbine blade temperature calculation and life estimation - a sensitivity analysis," *Propulsion and Power Research*, vol. 2, no. 2, pp. 148–161, 2013.

[138] D. Sauceda, P. Singh, G. Ouyang, O. Palasyuk, M. J. Kramer, and R. Arróyave, "High throughput exploration of the oxidation landscape in high entropy alloys," *Mater. Horiz.*, vol. 9, pp. 2644–2663, 2022.

[139] J. Hou, B. J. Wicks, and R. A. Antoniou, "An investigation of fatigue failures of turbine blades in a gas turbine engine by mechanical analysis," *Engineering Failure Analysis*, vol. 9, no. 2, pp. 201–211, 2002.

[140] S. Pugh, "XCII. Relations between the elastic moduli and the plastic properties of polycrystalline pure metals," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 45, pp. 823–843, Aug. 1954.

[141] S. Kou, "A simple index for predicting the susceptibility to solidification cracking," *Weld. J*, vol. 94, no. 12, pp. 374–388, 2015.

[142] P. Singh, B. Vela, G. Ouyang, N. Argibay, J. Cui, R. Arróyave, and D. D. Johnson, "A ductility metric for refractory-based multi-principal-element alloys," *Acta Materialia*, vol. 257, p. 119104, 2023.

[143] L. Liu, R. Paudel, Y. Liu, X.-L. Zhao, and J.-C. Zhu, "Theoretical and Experimental Studies of the Structural, Phase Stability and Elastic Properties of AlCrTiFeNi Multi-Principle Element Alloy," *Materials*, vol. 13, no. 19, p. 4353, 2020.

[144] Y. Liu, K. Wang, H. Xiao, G. Chen, Z. Wang, T. Hu, T. Fan, and L. Ma, "Theoretical study of the mechanical properties of CrFeCoNiMox ($0.1 \leq x \leq 0.3$) alloys," *RSC advances*, vol. 10, no. 24, pp. 14080–14088, 2020.

[145] H. Zhang, X. Sun, S. Lu, Z. Dong, X. Ding, Y. Wang, and L. Vitos, "Elastic properties of AlxCrMnFeCoNi ($0 \leq x \leq 5$) high-entropy alloys from ab initio theory," *Acta Materialia*, vol. 155, pp. 12–22, 2018.

[146] C. Baruffi, F. Maresca, and W. Curtin, "Screw vs. edge dislocation strengthening in body-centered-cubic high entropy alloys and implications for guided alloy design," *Mrs Communications*, vol. 12, no. 6, pp. 1111–1118, 2022.

[147] B. Vela, D. Khatamsaz, C. Acemi, I. Karaman, and R. Arróyave, "Data-augmented modeling for yield strength of refractory high entropy alloys: A bayesian approach," *Acta Materialia*, vol. 261, p. 119351, 2023.

[148] V. Attari and R. Arroyave, "Decoding non-linearity and complexity: Deep tabular learning approaches for materials science," *arXiv*, 2024.

[149] GrabCAD Community, "Turbine blade." `https://grabcad.com/library/gas-turbine-blade-4`, n.d. Model no longer available. Originally retrieved on July 28, 2024.

[150] Small Business Innovation Research (SBIR), "Automated functional grading of materials for directed energy deposition additive manufacturing," 2024. SBIR Program Topic A244-043, U.S. Army.

[151] Formalloy, "Manufacturing solutions," 2023. https://www.formalloy.com/manufacturing-solutions.

[152] R. Bonnard, J.-Y. Hascoët, P. Mognol, and I. Stroud, "STEP-NC digital thread for additive manufacturing: Data model, implementation and validation," *International Journal of Computer Integrated Manufacturing*, vol. 31, pp. 1141–1160, Nov. 2018.

[153] Task Force, USAF Global Science and Technology Vision, "Global Horizons: Final Report," June 2013.

[154] J. Hiller and H. Lipson, "STL 2.0: A proposal for a universal multi-material Additive Manufacturing File format," *20th Annual International Solid Freeform Fabrication Symposium, SFF 2009*, pp. 266–278, Jan. 2009.

[155] International Organization for Standardization, "ISO 14649-17:2020 Industrial automation systems and integration – Physical device control – Data model for computerized numerical controllers – Part 17: Process data for additive manufacturing," Standard ISO 14649-17:2020, ISO, Geneva, Switzerland, Aug. 2020.

[156] M. S. Ryou, H. S. Jee, W. H. Kwon, and Y. B. Bang, "Development of a data interface for rapid prototyping in STEP-NC," *International Journal of Computer Integrated Manufacturing*, vol. 19, pp. 614–626, Sept. 2006.

[157] B. Zhang, R. Seede, L. Xue, K. C. Atli, C. Zhang, A. Whitt, I. Karaman, R. Arróyave, and A. Elwany, "An efficient framework for printability assessment in laser powder bed fusion metal additive manufacturing," *Additive Manufacturing*, vol. 46, p. 102018, 2021.

[158] M. Vaughan, M. Elverud, J. Ye, R. Seede, S. Gibbons, P. Flater, B. Gaskey, R. Arróyave, A. Elwany, and I. Karaman, "Development of a process optimization framework for fabricating fully dense advanced high strength steels using laser directed energy deposition," *Additive Manufacturing*, vol. 67, p. 103489, 2023.

[159] S. Sheikh, B. Vela, P. Honarmandi, P. Morcos, D. Shoukr, I. Karaman, A. Elwany, and R. Arróyave, "An automated computational framework to construct printability maps for

additively manufactured metal alloys," *npj Computational Materials*, vol. 10, p. 252, Nov. 2024.

[160] S. Sheikh, B. Vela, P. Honarmandi, P. Morcos, D. Shoukr, A. M. Kotb, R. Arróyave, I. Karaman, and A. Elwany, "High-throughput alloy and process design for metal additive manufacturing," *arXiv*, 2023.

[161] S. Sheikh, B. Vela, V. Attari, X. Huang, P. Morcos, J. Hanagan, C. Acemi, I. Karaman, A. Elwany, and R. Arróyave, "Exploring chemistry and additive manufacturing design spaces: a perspective on computationally-guided design of printable alloys," *Materials Research Letters*, vol. 12, no. 4, pp. 235–263, 2024.

[162] I. Gibson, D. Rosen, B. Stucker, and A. Khorasani, *Additive Manufacturing Technologies*. Springer, Nov. 2020.