

Stochastic and Deterministic Algorithms for Set-Based Data-Driven Computational Mechanics

Von der Fakultät für Bauingenieurwesen
der Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades eines Doktors der Ingenieurwissenschaften
genehmigte Dissertation

vorgelegt von

Erik Christian Prume, M. Sc.

Berichter*innen: Universitätsprofessorin Dr.-Ing. Stefanie Reese
Frank and Ora-Lee Marble Professor Emeritus Michael Ortiz

Tag der mündlichen Prüfung: 22. Juni 2026

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.

Dedicated to Enno & Carl

Acknowledgements

This dissertation consolidates the most important results from my time as a research assistant at the Institute of Applied Mechanics at RWTH Aachen University. During my time at the institute, I had the honor of meeting many wonderful colleagues, mentors, and friends. I would like to thank everyone who was part of this journey.

First of all, I would like to thank my doctoral advisor, Prof. Dr.-Ing. Stefanie Reese. She supported me as early as my bachelor's studies by providing me with a position at the institute, which was always extremely important to me. I am also deeply grateful for her trust in me, and for her scientific and professional guidance throughout my time as a research assistant. Without her support, the present thesis would not have been possible.

I am very grateful to my second supervisor, Prof. Michael Ortiz. It was a huge honor for me to work with him, and I am grateful for the time and patience he invested in me. Working under his visionary guidance has opened my scientific horizons in many ways and taught me that anything is possible if one doesn't give up.

I would also like to thank Prof. Laurent Stainier for making the French-German co-operation project on data-driven mechanics possible and for supporting me scientifically in his friendly way.

While I cannot name everyone, the whole IFAM team was a huge support for me. They gave me the motivation, help, feedback, and laughter to keep me going —thank you!

At this point, I want to thank all the wonderful friends I met during my studies. They made my time in Aachen joyful and very special, and I am truly grateful for all the moments we shared together.

My entire family was always there for me in every possible way, for which I am deeply thankful. I am especially grateful to my parents, my sisters, my grandmother, aunts and uncles, as well as my mother-in-law. Without their unconditional support, none of this would have been possible.

Finally, to my dear wife: thank you for everything—for your laughter and happiness, your patience, and for showing me new perspectives. You have brightened my life and helped me become the open-minded and happy person I am today.

Summary

Data-driven methods have transformed the field of computational mechanics in recent years. Inspired by remarkable advances in data science and artificial intelligence across various disciplines, researchers in computational mechanics have rethought the way to tackle existing challenges. In the field of solid mechanics, a central challenge is the realistic description of the material behavior due to the enormous complexity and variety of both natural and engineered materials. While the required field theories of equilibrium and kinematic compatibility are widely valid and mathematically very well characterized, the classical method of constitutive, phenomenological material modeling is a laborious task with many modeling uncertainties. The models have to be specifically designed, calibrated, verified and validated based on the current knowledge and data about the material. This process has to be repeated as new information and data are obtained, turning it into an open-ended process.

A significant portion of recently developed data-driven methods target the automation of this process by implementing universal function approximators such as artificial neural networks. Such methods are extremely flexible and are designed to handle large amounts of data efficiently. Furthermore, basic principles from materials theory can be implemented in order to regularize the functions and reduce the required amount of data.

In 2016, Kirchdoerfer and Ortiz [82] proposed an alternative approach, herein called the *set-based* data-driven method. Instead of working with functional approximations for the material behavior, the boundary value problem is formulated explicitly in the material data set. The principal goals of the method are 1) reduction of modeling uncertainties 2) no loss of information for noisy data 3) correct description of materials where no single-valued functional relationship exists and 4) simple and automated process-cycle of data acquisition and inference.

This cumulative dissertation aims to contribute to the set-based data-driven method with a set of novel algorithms addressing several open research questions. A central theme throughout the dissertation is the identification of whether either a stochastic or a deterministic algorithm is the best choice for a given problem. The findings are often surprising; for instance, computing the global minimum of the discrete optimization problem with combinatorial complexity is best achieved using a stochastic algorithm, whereas discovering all local minima from a non-convex energy landscape in high

dimensions was only tractable via a deterministic algorithm.

In the first article, a stochastic algorithm to compute the global minimum of the set-based minimization problem is developed, introducing a randomized data search which effectively helps to escape local minima emerging from the discrete nature of the problem. This solver further enables a new goal-oriented data acquisition scheme in the context of multiscale mechanics driven by a weighted k-means clustering. It was shown that the error can be consistently reduced by adding a minimal amount of optimally selected data, saving a substantial amount of costly microstructure simulations.

The second article introduces the foundation for a stochastic description of the set-based method by considering the whole posterior distribution of possible system outcomes as compared to a single, deterministic solution. A stochastic, simulated annealing algorithm explores the distribution by randomly propagating a Markov Chain. While the algorithm is capable of sampling multi-modal distributions by maintaining a set of solutions, it is limited to rather small system sizes.

This limitation is overcome by the third article. The proposed algorithm is based on a basis for the space of equilibrium stress fields, where sparsity of the basis is crucial for numerical efficiency. Existing methods to compute such a sparse null basis for general systems often rely heavily on expensive numerical factorizations or sophisticated graph algorithms, occasionally leading to conditioning problems. To overcome these problems, an efficient algorithm to compute the sparse null basis for the important special case of linear tetrahedral elements is proposed. Equipped with the basis for equilibrium stress fields, a deterministic annealing algorithm based on [144] is developed. While slowly resolving details of the multi-modal distribution, a partial eigenvalue decomposition in the space of equilibrium stress and compatible strain fields detects the formation of local minima and provides the directions where to initialize two separate deterministic minimization runs. By this process, a hierarchical solution set is constructed which provides the full posterior distribution, serving the purpose of uncertainty quantification. The method is demonstrated with a brittle fracture problem featuring a local material with stochastic tensile strength, where every solution represents one possible crack pattern. As the data set cannot be represented by a single-valued functional relationship, this result may be seen as a prime example for the set-based data-driven method.

Zusammenfassung

Datengestützte Methoden haben das Feld der Computational Mechanics in den letzten Jahren grundlegend verändert. Inspiriert von Fortschritten in Data Science und künstlicher Intelligenz über verschiedene Disziplinen hinweg, haben Forscher in der Mechanik die Herangehensweise an bestehende Herausforderungen neu überdacht. In der Festkörpermechanik bleibt die realistische Beschreibung des Materialverhaltens aufgrund der Komplexität natürlicher und technischer Werkstoffe eine zentrale Hürde. Während die Feldtheorien von Gleichgewicht und kinematischer Kompatibilität weitgehend gültig und mathematisch exakt charakterisiert sind, ist die klassische phänomenologische Materialmodellierung eine mühsame Aufgabe verbunden mit vielen Modellierungsunsicherheiten. Modelle müssen auf Basis des aktuellen Wissens spezifisch entworfen, kalibriert und validiert werden – ein Prozess, der sich bei neuen Daten stetig wiederholt und so zu einer Daueraufgabe wird.

Viele neuere datengestützte Methoden zielen darauf ab, diesen Prozess durch universelle Funktionsapproximatoren wie künstliche neuronale Netze zu automatisieren. Diese sind hochflexibel und verarbeiten große Datenmengen effizient. Zudem können materialtheoretische Grundprinzipien implementiert werden, um die Funktionen zu regularisieren und den Datenbedarf zu senken.

Im Jahr 2016 schlugen Kirchdoerfer und Ortiz [82] einen alternativen Ansatz vor, die hier als Set-basierte datengestützte Methode bezeichnet wird. Statt mit funktionalen Approximationen zu arbeiten, wird das Randwertproblem explizit im Materialdatensatz formuliert. Die Hauptziele sind: 1) Reduktion von Modellierungsunsicherheiten, 2) Erhalt von Informationen bei verrauschten Daten, 3) korrekte Beschreibung von Materialien ohne eindeutige funktionale Beziehung und 4) ein automatisierter Zyklus aus Datenerfassung und Inferenz.

Diese kumulative Dissertation trägt zur Set-basierten Methode durch neuartige Algorithmen für verschiedene Forschungsfragen bei. Ein zentrales Thema ist die Identifikation, ob ein stochastischer oder ein deterministischer Algorithmus für ein gegebenes Problem die beste Wahl ist. Die Erkenntnisse sind oft überraschend: So gelingt die globale Optimierung kombinatorisch komplexer Probleme am besten stochastisch, während das Auffinden aller lokalen Minima hochdimensionaler Energielandschaften nur deterministisch handhabbar war.

Im ersten Artikel wird ein stochastischer Algorithmus zur Lösung des Set-basierten Minimierungsproblems entwickelt. Eine randomisierte Datensuche hilft dabei, lokalen Minima zu entkommen, die aus der diskreten Natur des Problems resultieren. Dies ermöglicht zudem ein zielorientiertes Datengewinnungsschema für die Multiskalenmechanik, gesteuert durch ein gewichtetes k-means Clustering. Es wurde gezeigt, dass der Fehler durch optimal ausgewählten Daten konsistent sinkt, was teure Mikrostruktursimulationen einspart.

Der zweite Artikel führt die Grundlage für eine stochastische Beschreibung ein, indem die gesamte Posterior-Verteilung möglicher Systemausgänge statt einer Einzel-Lösung betrachtet wird. Ein stochastischer Simulated Annealing Algorithmus exploriert die Verteilung mittels Markov-Ketten. Obwohl der Algorithmus multimodale Verteilungen durch das Vorhalten mehrerer Lösungen abbilden kann, ist er auf eher kleine Systeme begrenzt.

Diese Limitierung wird im dritten Artikel überwunden. Der vorgeschlagene Algorithmus nutzt eine Basis für den Raum der Gleichgewichtsspannungsfelder, wobei deren Dünnbesetztheit entscheidend für die Effizienz ist. Da bestehende Methoden zur Berechnung solcher Nullbasen oft auf teuren Zerlegungen oder komplexen Graphalgorithmen beruhen, wird hier ein effizienter Algorithmus für den Spezialfall linearer Tetraederelemente vorgestellt. Darauf aufbauend wird ein Deterministic Annealing Algorithmus nach [144] entwickelt. Während die Details der Verteilung aufgelöst werden, erkennt eine partielle Eigenwertzerlegung im Raum der Spannungs- und Verzerrungsfelder lokale Minima und liefert die Richtungen für die Initialisierung von separaten, deterministischen Minimierungen. So entsteht eine hierarchische Lösungsmenge für die Unsicherheitsquantifizierung. Die Methode wird an einem Sprödbruchproblem mit stochastischer Zugfestigkeit demonstriert, wobei jede Lösung ein mögliches Rissmuster darstellt. Da der Datensatz keine eindeutige funktionale Beziehung zulässt, dient dieses Ergebnis als ein besonderes Beispiel für die Set-basierte datengestützte Methode.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	State-of-the-art in data-driven computational mechanics	3
1.2.1	Function-based data-driven methods	3
1.2.2	Set-based data-driven methods	5
1.3	Outline of the dissertation	9
2	Article 1:	
	Direct Data-Driven Algorithms for Multiscale Mechanics	13
2.1	Abstract	14
2.2	Introduction	14
2.3	Preliminaries: direct data-driven mechanics	16
2.4	Randomized fixed-point solver: an alternative solver scheme	19
2.4.1	Motivation	19
2.4.2	Proposed method: randomized fixed-point solver	20
2.4.3	Numerical example: stamp	21
2.5	Adaptive data generation using FFT-based microstructure models	25
2.5.1	Motivation	25
2.5.2	Proposed method: adaptive data generation	26
2.5.3	Algorithm summary	29
2.5.4	Numerical example: beam	30
2.6	Summary and concluding remarks	34
2.7	Acknowledgements	35

3 Article 2:

Model-Free Data-Driven Inference in Computational Mechanics	37
3.1 Abstract	38
3.2 Introduction	38
3.3 The inference problem	41
3.3.1 Example: trusses	41
3.3.2 The constraint set of admissible states	42
3.3.3 Material characterization	43
3.3.4 Example: local material behavior	44
3.3.5 Classical inference	45
3.3.6 Example: random trusses	45
3.3.7 Approximation by empirical data	47
3.3.8 Annealing schedule	48
3.3.9 Factorization and localization	49
3.4 Numerical implementation	51
3.4.1 Problem statement	51
3.4.2 Population annealing	51
3.5 Examples of application	57
3.5.1 Three-bar truss with sliding-Gaussian material data set	57
3.5.2 Three-bar truss with Weibull tensile-strength distribution	61
3.5.3 Lightweight space structure	64
3.6 Acknowledgements	66

4 Article 3: A Sparse Basis for Equilibrium Stress Fields with Application for Direct Data-Driven Mechanics

	69
4.1 Abstract	70
4.2 Introduction	70
4.3 Part I: An explicit and sparse basis for equilibrium stress fields	73
4.3.1 Notation	73
4.3.2 Discrete strain and divergence operators	73
4.3.3 Dimensionality and components of the null basis	75
4.3.4 Stress fields induced by edge force fields	77
4.3.5 Facet-wise mechanism	79

4.3.6	Edge-wise mechanism	80
4.3.7	Remaining null vectors	84
4.3.8	Sparse null basis assembly algorithm	86
4.3.9	Efficient Cholesky factorization of the null basis	89
4.4	Part II: Application for direct data-driven mechanics	90
4.4.1	Background: direct data-driven mechanics	90
4.4.2	Reformulation using null basis	93
4.4.3	Hierarchical solution set: deterministic annealing	96
4.5	Conclusions	103
4.6	Acknowledgements	105
5	Conclusions and Outlook	107
	List of Figures	111
	Bibliography	115

1 | Introduction

1.1 Motivation

The fundamental motivation of civil engineering is, as the name suggests, to serve the collective needs of society and improve the quality of human life under consideration of all associated aspects. This service includes the provision of safety, health, environmental stability and the advancement of supply of basic goods, productivity, mobility, and infrastructure. These objectives are no longer merely optional technical aspirations but have become absolute necessities driven by the pressures of an increasing global population and the growth in per-capita resource demands. The engineering community currently faces the daunting reality of resource scarcity, unequal distribution of materials, and the intensifying environmental degradation and climate change. In response, numerical simulations have evolved into one of the most versatile and efficient tools available to engineers and scientists for the development and optimization of new technologies. These digital methodologies empower engineers to understand the behavior of complex technical systems and facilitate large-scale virtual prototyping in an economically efficient manner.

In the specialized fields of structural and mechanical engineering, the Finite Element Method (FEM) stands as the central tool for the simulation of mechanical boundary value problems. Since its modern formalization in the mid-twentieth century, FEM has achieved a level of dominance due to its relatively simple yet highly efficient structure, which allows for the discretization of continuous domains into standard geometric elements. A mechanical boundary value problem provides the standardized mathematical description of any physical system, necessitating the precise notation of force equilibrium, kinematics, material laws, and boundary conditions. The equations of force equilibrium, derived from Newtonian axioms, dictate that the sum of all forces must equal zero in static systems. Kinematics establish the geometric relationships between

global displacements and localized strain fields, ensuring that deformations remain compatible. However, the closure of this mathematical system requires a constitutive law: a rule that defines the relationship between stresses and strains for the given material.

While kinematics and equilibrium are governed by axiomatic principles that are largely independent of the specific material being used, the constitutive law remains the most complex and variable component of the system. Historically, a vast amount of literature has been dedicated to characterizing industrial materials such as steel, aluminum, concrete, and rubber, as well as biological materials like soft tissues and cardiac cells. The behavior of these materials is often non-linear and characterized by diverse phenomena including elasticity, plasticity, viscosity, and eventual failure or damage. Traditional constitutive modeling attempts to capture these effects through a set of carefully chosen mathematical functions and parameters designed to satisfy fundamental principles such as objectivity and thermodynamic consistency. These parameters are typically calibrated using experimental data from uniaxial or multiaxial tests. However, the inherent complexity of materials often leads to modeling bias and uncertainty, as the prescribed functions may fail to represent the actual material response accurately, particularly under complex or unseen loading conditions. The consequence in practice is often that existing material models have to be continuously adapted and re-calibrated as soon as new material data becomes available, leading to a never-ending circle of significant economic costs.

This limitation has motivated a significant paradigm shift toward data-driven methods in computational mechanics. As experimental techniques like Digital Image Correlation (DIC) and automated test setups produce increasingly more data, researchers are looking for ways to overcome the drawbacks of phenomenological material modeling. The principal guideline is to move from a model-centric view to a data-centric view, where the simulation's performance is directly tied to the richness and quality of the available data. By automating the process of data acquisition and material behavior prediction in a more unbiased but still physical manner, efficient and reliable simulations are enabled, in particular when dealing with real-time integration of sensor data in conjunction with digital twin implementations. Within this emerging field, a distinction can be drawn between *function*-based data-driven approaches, such as Artificial Neural Networks (ANNs), and *set*-based data-driven approaches. While the former seeks to replace classical equations with flexible surrogate models, the latter incorporates no functional

material model but the raw data set directly into the mechanical solver. In the following state-of-the-art subsections, we will use these two approaches as a basic categorization, even though both fields have accumulated a large, diverse amount of literature, often with overlapping methods and ideas.

1.2 State-of-the-art in data-driven computational mechanics

1.2.1 Function-based data-driven methods

Function-based data-driven approaches center on the use of high-capacity function approximators, most notably artificial neural networks, to serve as constitutive surrogates. This methodology recognizes that while the governing equations of equilibrium and kinematics are known, the material law is often the weakest link in the simulation chain. By replacing traditional phenomenological material models with ANNs, researchers can capture complex, high-dimensional relationships that would be otherwise impossible to formalize mathematically. Furthermore, the flexible and generic model architecture can capture theoretically a broad range of materials with a systematic way to calibrate the parameters based on the data.

Artificial neural networks. The application of ANNs to constitutive modeling was pioneered by Ghaboussi and coworkers in the early 1990s [50]. They demonstrated that a backward propagation neural network could be trained to predict the stress-strain response of concrete based on experimental data. However, these early models were essentially "black boxes" [102]; they provided no guarantee that the predicted material behavior would be physically reasonable outside of the specific training regime. This lack of transparency and physical interpretability hindered the adoption of ANNs. The paradigm shifted with the advent of scientific machine learning and the development of architectures that directly integrate physical laws into the training process.

The class of Constitutive Artificial Neural Networks (CANNs) [108, 107] approaches enhance a neural network for the material model by insights from material theory and microstructural information. Developed to address the limitations of black-box models, CANNs utilize architectures that inherently satisfy kinematical principles. By using

strain invariants as inputs, CANNs guarantee material frame indifference (objectivity) and can be structured to enforce properties like polyconvexity, which is essential for the stability of hyperelastic materials. Further assumptions of the material behavior such as anisotropy can be additionally included [84, 44].

Inelasticity. Modeling inelastic materials like structural steel presents a significant challenge because the current stress depends not only on the current strain but on the entire history of deformation. Thermodynamics-Based Artificial Neural Networks (TANNs) [113, 155] include thermodynamic quantities like the dissipation rate directly into the network architecture. Thus, thermodynamic consistency can be ensured by means of non-negative dissipation rates: a critical requirement for modeling inelastic and damage phenomena. A detailed treatment of the multiplicative decomposition of the deformation gradient into an inelastic and elastic part along with the formulation of inelastic potentials is given in [66]. In [61], the yield function is replaced by a more flexible machine learning algorithm accounting for a wider range of plastic mechanisms. It is also possible to learn the whole deformation path with specialized network architectures such as Recurrent Neural Networks (RNNs) [119, 55, 13], networks enhanced by a proper orthogonal decomposition [67] and graph-based networks [65, 114]. This allows for a robust representation of path-dependent materials without the need for explicit internal variable construction.

Automated model selection. As proposed by [39, 153, 41], another strategy is to define a library of material model components and select the smallest subset of components sufficient to characterize the material behavior. The calibration works on a global level using displacement (such as from DIC measurements) and global force data. Further works on this approach show a close relationship with sparse regression [116] and optimal automated model selection [128].

Multiscale modeling. A primary application of data-driven constitutive models is in the field of multiscale computational mechanics [12]. In traditional FE^2 schemes [49], the evaluation of a microscopic Representative Volume Element (RVE) at every integration point of a macroscopic finite element mesh is computationally staggering. By training an ANN to act as a surrogate for the RVE, the computational cost can be reduced by orders of magnitude.

Advanced techniques further involve microstructural data into hierarchical layers that reflect the physical interactions within the material. These models have been successfully

applied to composite materials, dual-phase steels and concrete, where they learn effective constitutive laws at the microscale and transfer them to macroscale structural analyses [160]. A particularly useful network architecture for this challenge are Convolutional Neural Networks (CNNs) as they can learn directly from microscopic image data [148, 1].

Solving boundary value problems. A different application of ANNs in computational mechanics is to solve the boundary value problem itself by a neural network as opposed to using a FEM solver, with early approaches from the 1990s in [31, 99]. This approach has been popularized by the name Physics-Informed Neural Networks (PINNs) introduced by [161, 138, 146]. While not directly acting as a material model surrogate, PINNs incorporate the governing partial differential equations (PDEs) directly into the neural network's loss function. In structural mechanics, this allows for the solving of boundary value problems where a neural network approximates the displacement field while simultaneously satisfying equilibrium and boundary conditions. Even though the highly efficient, sparse structure of FEM for mechanical problems is superior from a numerical point of view, PINNs are applicable for a large range of different PDEs and show interesting similarities with meshless methods. This motivated further improvements in the form of Neural Operators [106, 91] which construct maps between infinite dimensional function spaces independent of the particular mesh discretization.

1.2.2 Set-based data-driven methods

In 2016, Kirchdoerfer and Ortiz [82] introduced a novel computing paradigm that directly includes the data set of stress-strain pairs into the formulation of boundary value problems. This approach represents a fundamental shift from classical function based methods and is in the literature referred to the different names *constitutive model-free*, *distance-minimizing*, *set-based* or *direct* data-driven method. In particular, any use of constitutive material models is avoided and thus no functional relationship between stress and strain is established, motivating the name 'set-based' in order to distinguish from 'function-based' approaches. While it recognizes kinematic compatibility and equilibrium as universal axiomatic laws which have to be fulfilled exactly, it aims to minimize any modeling uncertainties and loss of information regarding the complex material behavior.

The set-based data-driven method computes two solutions simultaneously: 1) the

assignment of data points to material points (of the finite element mesh) which are closest in satisfying compatibility and equilibrium 2) the compatible strain and equilibrium stress field which are closest to the given data set in each material point. As a consequence of the limited size of data sets, the distance between these two solutions is in general greater than zero, meaning that the data assignment cannot fulfill the physical constraints exactly. While this mismatch seems like a drawback due to a seemingly restrictive design choice, important information is preserved which is generally lost in model-based approaches: the method quantifies how sufficient the data set is for the given boundary value problem and where data is missing. It is noteworthy that while the method does not make predictions for the local material behavior, it has predictive capabilities about the global system response by joining the local data sets under the constraints of compatibility and equilibrium.

Various aspects of the direct data-driven method have been addressed in the existing literature, including mathematical proofs in particular on convergence behavior [23, 141], dynamical problems [83], finite strain theories [124, 24, 131, 93], snap-through problems [92], non-local micro-polar material behavior [76]. The concept has been further applied to multi-physical field problems like diffusion [90, 94], magneto-statics [48] and electro-mechanical coupled [112]. To reduce the required amount of data, the strict instance based approach has been relaxed by implementing local interpolation techniques [69, 63, 74, 35, 19]. The possibility to enhance efficiency by using a quantum computer has been explored in [159]. Furthermore, the relationship of data-driven methods to the concept of Constitutive Relation Error [98] has been pointed out in recent publications, [97] among others. Therein, the same basic guideline to quantify and reduce the error introduced by a material model while keeping known physical principles satisfied exactly, is pursued.

Data acquisition. A major advancement in the acquisition of material data is the Data-Driven Identification (DDI) method developed by [104, 149]. DDI provides an inverse formulation of the set-based data-driven framework that computes stress-strain pairs directly from a sequence of non-homogeneous displacement fields (measured via Digital Image Correlation or Digital Volume Correlation [15]) and boundary forces. This allows for the construction of high-fidelity material databases of complex deformations without the need to presuppose any constitutive equations or perform numerous simple experiments. The method has been successfully applied to identify data sets of

heterogeneous materials [154], large strain elasticity [26] and inelasticity [100]. More recently, [103, 60] derived conditions for the uniqueness of DDI solutions and estimates for the maximum number of data points which can be obtained from a given experiment. Related, but following a different approach is the contribution of [152], which constructs a 3D data base by decompressing 1D stress measurements for isotropic nonlinear elastic materials based on an eigenvalue decomposition of the Cauchy–Green tensor.

Another popular strategy to acquire sufficient data sets is to perform virtual experiments using microstructural RVE models [158, 89, 77, 54, 140] and perform the data-driven analysis on the homogenized data. Hereby, the RVE evaluation is informed by the boundary value problem to be solved which coined the term goal-oriented sampling. In particular, the information about the distance between data and compatible states is useful for completing missing regions of the data set. Compared to classical multiscale simulations techniques, the two main advantages of this approach are, that no constitutive model of the homogenized data has to be developed and that the data can be re-used, thus avoiding duplicate RVE evaluations. This idea was also exploited in [151], where a sequence of similar microstructural boundary value problems is solved by iteratively adding missing data.

Robustness of the data-driven solver. It is important to note that the set-based data-driven problem is of combinatorial complexity, as each data point can be in principle assigned to each material point. To tackle this challenge, [82] proposed a fixed-point solver which yields however only an approximate solution which might not be the global minimizer. In particular, the robustness of the fixed-point solver is dependent on the sparsity and scattering of the data set, non-linearity of the material behavior, the initial solution guess and metric parameters. As a consequence, several approaches were proposed to tackle the problems. In [73], a reformulation of the problem as Mixed-Integer Programming (MIP) is proposed. This ensures that the solver finds the global optimal solution rather than becoming trapped in local minima, increases however the computation times significantly. Another strategy is to fine-tune the metric parameters for each material point individually [48] and during the solution process [125]. More recently, [140] proposed a perturbation method in order to escape local minima, performed by choosing the *second* nearest data point at randomly chosen material points.

Uncertainty Quantification. A particularly appealing property of the method is that

all the material data is included without any loss of information, which is especially important for noisy data sets and uncertainty quantification purposes. Since the original fixed-point solver is quite susceptible with respect to outliers in the data set, [83] has proposed the formulation of a free energy function based on the density of data points, allowing for a weighted contribution from all data points. The derivation of the free energy function is based on information-theoretic arguments, with many similarities to soft clustering formulations [144]. Similarly to the fixed-point solver, the associated 'maximum-entropy solver' [83] then computes the maximum likelihood solution reducing the influence of outliers.

In order to acknowledge the intrinsic uncertainty of materials and measurement errors, further approaches were developed to enhance the set-based method. In [88, 163], a comparison between the stochastic finite element method (SFEM) and the direct data-driven method is given. Therein, the deterministic data-driven fixed-point solver is still used, examining the influence of random initializations and variance within the neighboring data. Based on the MIP approach, [75] suggested the use of segmented least squares to compute confidence bounds of the structural response. More recently in [142], it was shown how the direct data-driven method naturally includes the full set of non-unique solutions, which gets neglected by model-based approaches. The study demonstrated this with the example of buckling with an uncertain buckling failure strength.

Inelasticity. One of the most important and challenging tasks is the treatment of inelastic material behavior. In [33], a general method to condition the data set based on the previous history of deformation was proposed, with a focus on visco-elastic and plastic materials. Further approaches were developed, including the usage of tangent information [19], building history surrogate models [9, 132], viscoelasticity in the frequency domain [145]. In the context of multiscale modeling of composite structures, [62] exploits the knowledge of plastic internal variables to parameterize the data set. A formulation for brittle fracture is presented in [16] where the crack length is used as an history variable. In [105], a failure criterion is proposed based on the heuristic that failure occurs once the distance is exceedingly large, demonstrated with application for composite materials. Due to the incremental nature of the inelastic problems, using a nearest neighbor graph data structure has favorable properties for the data search as investigated in [27, 135]. Therein, it was also shown how thermodynamic quantities

like the dissipation rate can be used to restrict the data set to physically admissible trajectories.

1.3 Outline of the dissertation

This cumulative dissertation targets to contribute to the set-based data-driven method by identifying relevant research questions based on the state of the art, and develop mathematically sound as well as numerically practical algorithms to tackle the addressed problems. The articles are organized by thematic logic rather than chronological publication date.

The first article [136] addresses the following two research questions:

- Is there an algorithm which can efficiently and robustly compute the global minimum of the set-based data-driven problem?
- How to optimally enrich a data set to solve a boundary value problem up to a certain error tolerance, in particular for path-dependent materials?

First, a randomized solver algorithm is developed. The algorithm keeps the efficient fixed-point structure of [82], but modifies the data search by randomly deciding if a material point should optimize for closest strain *or* stress data in a given iteration. By this split, the dependency on the metric coefficients is effectively removed. As a result, the algorithm finds robustly in almost all cases the global minimum with almost no computational overhead compared to the original, deterministic fixed-point solver. The proposed algorithm is thus another example that randomization is often a powerful tool for solving complex, deterministic optimization problems.

The availability of a robust and efficient solver enabled the development of a converging data set completion scheme for path-dependent materials in the context of multiscale modeling. By exploiting the tangent information of the microscale model, an accurate proposal strain field is computed. Subsequently, a k-means clustering weighted by the minimum distance error computes the most important missing strain modes. Demonstrated with a numerical example, it was shown that even for a single boundary value problem a reduction of evaluated microstructures can be achieved. The computational gain of this approach becomes even more pronounced if a sequence of similar BVPs is computed, as demonstrated in [151].

The second and third article [134, 137] deal with noisy data sets and how to make stochastic inferences about the system response. Of particular interest are brittle fracture problems with uncertain tensile strength, resulting in highly non-convex energy landscapes with many non-unique solutions. Thus, the concerned problems are prime examples for the setting described in [142].

The following research questions are identified:

- How should the material likelihood function be formulated to ensure convergence to the true behavior with increasing data set size?
- How to sample efficiently from the resulting, non-convex posterior distributions?
- What is the best way to satisfy the force equilibrium constraint?

In the second article [134], the material likelihood function is derived based on the formulation presented in [83]. The difference is, that we are not only interested in the maximum likelihood solution, but in the whole distribution of system outcomes. In other words, the goal is to compute every possible system state which has a non-zero probability to occur, providing much richer information about the boundary value problem. With special care, the bandwidth parameter is determined as it is crucial for comparing different solutions which vary in their degree of incompatibility with the physical constraints. For the sampling, a population annealing algorithm based on the Markov-Chain Monte-Carlo (MCMC) method is developed, which is capable to explore different local minima, demonstrated on a simple truss example with brittle failure. To sample in the space of equilibrium stress fields, a principal component analysis (PCA) of a set of equilibrium stress fields was performed, providing a dense sampling basis. The correct convergence behavior with increasing data set sizes is demonstrated by comparing the sampled distribution with a derived analytical solution. However, due to the dense sampling space of equilibrium stress fields and the random movements of the Monte-Carlo method, only small problems were tractable.

This motivated the third article [137]. Therein, an efficient algorithm to compute a sparse basis for equilibrium stress fields is developed. Based on this, the data-driven problem is reformulated as an unconstrained minimization problem. This enabled the replacement of the Monte-Carlo sampling by a deterministic annealing algorithm based on [144, 143]. The algorithm computes a hierarchical set of solutions by detecting

formations of local minima, and creating a new solution instance for each local minimum which gets optimized by the annealed fixed-point solver [83]. The resulting solution set can be used for the purpose of stochastic inference and uncertainty quantification. This novel approach made it possible to compute realistic, three-dimensional problems with non-unique solutions which were previously not tractable.

2 | Article 1:

Direct Data-Driven Algorithms for Multiscale Mechanics

This article was published as:

Prume, E., Gierden, C., Ortiz, M., and Reese, S. [2025], 'Direct data-driven algorithms for multiscale mechanics', *Computer Methods in Applied Mechanics and Engineering* **433**, 117525.

Disclosure of the individual authors' contributions to the article:

E. Prume developed the algorithms of the randomized solver and adaptive data generation scheme. E. Prume drafted the manuscript, implemented the algorithms and computed the numerical examples. C. Gierden contributed expert knowledge about the FFT solvers. S. Reese and M. Ortiz supervised the research process and gave important scientific guidance. C. Gierden, S. Reese and M. Ortiz reviewed the draft and contributed valuable comments and edits. All authors approved the publication of the final version of the manuscript.

2.1 Abstract

We propose a randomized data-driven solver for multiscale mechanics problems which improves accuracy by escaping local minima and reducing dependency on metric parameters, while requiring minimal changes relative to non-randomized solvers. We additionally develop an adaptive data-generation scheme to enrich data sets in an effective manner. This enrichment is achieved by utilizing material tangent information and an error-weighted k-means clustering algorithm. The proposed algorithms are assessed by means of three-dimensional test cases with data from a representative volume element model.

2.2 Introduction

In this work, we consider two-scale problems where the finer scale is characterized by a representative volume element (RVE) that captures the details of the material microstructure and serves as a local material description of a macroscopic, mechanical boundary value problem (BVP). This type of material modeling is referred to as *upscaling* [49, 38]. In the FE^2 method [37], the RVE is evaluated at every material point, time step and possibly multiple iterations, which can be exceedingly costly. To reduce the number of RVE evaluations in an inelastic setting, [10] developed a k-means clustering FE^2 approach where material points of similar deformation behavior are grouped and jointly evaluated.

Another popular approach is to perform a series of evaluations (or "virtual experiments") with an RVE model and calibrate a phenomenological material model to the resulting data set. More recently, the material modeling is often performed by specialized *supervised* neural network architectures in order to automatize and standardize the modeling procedure for elastic [108, 4] and inelastic [113, 66] material behavior, see for instance [109, 45] for a review. A number of methods have been developed in order to reduce the number of RVE evaluations and improve the accuracy of the trained material model [46, 72].

The advantage of the supervised approach is the fast model evaluation, while the disadvantage is that additional modeling uncertainties are introduced by biased modeling assumptions, including which and how many virtual experiments are performed, how

validation and verification are performed and what features and functional basis are chosen for the model. Thus, when performing a predictive computation, i. e., a previously unseen material state, with a calibrated model, it is often not clear whether the local response of the material model needs to be validated by an RVE evaluation or may be assumed to be reliable. In addition, the cost of training the neural network architectures, while often not accounted for, can be very high, especially with large data sets and high-dimensional data.

In order to eschew these shortcomings of supervised models, in direct, or *unsupervised*, data-driven methods a pre-existing material data set is directly used in the solution of the boundary-value problem [82]. In this manner, only actual unbiased RVE data are used for purposes of material characterization. Evidently, the accuracy of the solution depends on the degree of coverage supplied by the material data set, which can be efficiently monitored using a distance function [82, 89, 54]. This quantitative metric can be used to identify regions of insufficient data coverage and to develop strategies to acquire new data in a systematic manner.

The direct data-driven method has been extended to large deformations [124, 131], inelastic material behavior [33, 16] and uncertainty quantification [58, 163, 134]. To improve the accuracy with sparse data sets, [35, 19] include material tangent information in the solving process. Furthermore, the method has been applied to other field equations than mechanics, such as magnetism [30] or electro-mechanical coupled problems [112]. In [104], the direct data-driven method has been reformulated in order to identify stress data from a sequence of strain field measurements. Similarly, [40] proposes an automated discovery scheme for plastic material laws using only strain data.

While this paper aims to enable efficient multiscale computations with synthetic microstructures, the direct data-driven method has been also used with real data measurements, based on the identification of two-dimensional stress fields [26, 150, 100] using Digital Image Correlation (DIC) measurements. Particularly challenging, but not impossible, is obtaining strain data in the three-dimensional case using Digital Volume Correlation (DVC) measurements [15], which in turn can be used to identify the corresponding stress data. Furthermore, domain specific measurement techniques are used to obtain real 3D data, see for example [145] which uses elastography data for a viscoelastic biomechanical problem.

This paper has two novel, significant improvements over previous work. First, it tackles

the long-standing problem of finding the global optimizer of the double minimization problem [82] which was also reported in [73, 48, 125]. This problem is in particular relevant for rather sparse data sets with three-dimensional, nonlinear material behavior. We propose a simple variation of the original fixed-point algorithm that effectively helps to escape local minima and reduces the dependence on the choice of metric coefficients. Compared to previous works, the particular advantages of the proposed approach are that it is easy to implement and does not increase the computational effort, yet is effective. Second, we contribute an incremental improvement for the systematic data acquisition problem in the multiscale context compared to previous work [89, 54]. In particular, we show how to exploit the material tangent information and the intrinsic error measure of the direct data-driven approach using a weighted k-means algorithm. The result is an effective and consistent data set enrichment scheme for nonlinear and path-dependent, force and displacement driven problems, thus minimizing the number of costly RVE evaluations in an unsupervised manner.

The paper is structured accordingly: in Section 2, we briefly review the theory of direct data-driven mechanics; in Section 3 the improved solver is formulated and tested on a numerical example; in Section 4 the adaptive data generation scheme is developed and evaluated with the aid of a numerical example. Finally, in Section 5 we summarize our findings and provide an outlook for further work.

2.3 Preliminaries: direct data-driven mechanics

In this section, we provide a compact review of the direct data-driven computing paradigm as proposed by [82] and the notation used in the following sections. We consider discretized mechanical systems of m material points, each characterized by a strain state $\varepsilon_e \in \mathbb{R}^d$, a stress state $\sigma_e \in \mathbb{R}^d$ and a volume $w_e \in \mathbb{R}$. Hereby, the index e indicates the specific material point and d is the dimension of stress and strain, that is $d = 6$ in the considered three-dimensional, geometrical linear case. In addition, we define the weighted Euclidean norms $\|\varepsilon_e\|^2 := \varepsilon_e \cdot \mathbb{C}_e \varepsilon_e$ and $\|\sigma_e\|^2 := \sigma_e \cdot \mathbb{C}_e^{-1} \sigma_e$ in order to account for different units of stresses and strains. The matrix $\mathbb{C}_e$ contains metric coefficients whose influence is studied in Section 2.4.

The idea of the direct data-driven approach is to formulate the solution of the boundary value problem explicitly in the given experimental stress-strain data, which is given in

the form of local data set(s)

$$D_e = \{(\varepsilon'_i, \sigma'_i)\}_{i=1}^{n_e} \quad (2.1)$$

with n_e material data points. The direct data-driven solution is then defined as that assignment of data points to each material point which fulfills the physical constraints of force equilibrium and kinematical compatibility simultaneously in an optimal sense.

We describe the kinematical compatibility by the set

$$E_\varepsilon := \{(\varepsilon_1, \dots, \varepsilon_m) \in \mathbb{R}^{md} : \varepsilon_e = B_e u, e = 1, \dots, m\} \quad (2.2)$$

of all strain fields $\varepsilon \in E_\varepsilon$ which are derived from displacement fields $u \in \mathbb{R}^n$ by the geometrical discretized gradient operator $B_e \in \mathbb{R}^{d \times n}$. Hereby, n is the number of degrees of freedom of the overall system including boundary conditions. Similarly, the force equilibrium is described by the set

$$E_\sigma := \left\{ (\sigma_1, \dots, \sigma_m) \in \mathbb{R}^{md} : \sum_{e=1}^m w_e B_e^T \sigma_e = f \right\} \quad (2.3)$$

of all stress fields $\sigma \in E_\sigma$ which fulfill the discretized force equilibrium given a force field $f \in \mathbb{R}^n$. Similarly to the local norms, we define $\|\varepsilon\|^2 := \sum_{e=1}^m w_e \|\varepsilon_e\|^2$ and $\|\sigma\|^2 := \sum_{e=1}^m w_e \|\sigma_e\|^2$ as the global norms accounting for the different material point volumes w_e .

Formally, we denote the set of all possible assignments of data points to material points by

$$D = D_1 \times D_2 \times \dots \times D_m \quad (2.4)$$

and one such assignment by $(\varepsilon', \sigma') = ((\varepsilon'_{e=1}, \dots, \varepsilon'_{e=m}), (\sigma'_{e=1}, \dots, \sigma'_{e=m})) \in D$. We note that Eq. (2.2) and (2.3) constrain the strain and stress fields independently of each other and the coupling between the two fields is only established by the material data set.

Using these notations, the direct data-driven problem can be formulated by

$$\min_{(\varepsilon', \sigma') \in D} \left(\min_{\varepsilon \in E_\varepsilon} \|\varepsilon - \varepsilon'\|^2 + \min_{\sigma \in E_\sigma} \|\sigma - \sigma'\|^2 \right) \quad (2.5)$$

with an optimal data assignment (ε', σ') and its corresponding, closest admissible strain and stress fields (ε, σ) as the solution.

For the computation of the closest compatible strain field and the closest equilibrium stress field we introduce the two projection operations $P_\varepsilon(\varepsilon') = \arg \min_{\varepsilon \in E_\varepsilon} \|\varepsilon - \varepsilon'\|^2$ and $P_\sigma(\sigma') = \arg \min_{\sigma \in E_\sigma} \|\sigma - \sigma'\|^2$ which consists in solving two linear, decoupled systems of equations

$$[B_{\text{red}}^T W \mathbb{C} B_{\text{red}}] u_{\text{red}} = B_{\text{red}}^T \mathbb{C} \varepsilon' \quad (2.6a)$$

$$[B_{\text{red}}^T W \mathbb{C} B_{\text{red}}] \eta_{\text{red}} = f_{\text{red}} - B_{\text{red}}^T W \sigma'. \quad (2.6b)$$

Here B_{red} is the assembled, discretized gradient operator, or B-operator after removing the degrees of freedom associated to Dirichlet boundary conditions, $W = \text{diag}(w_0, w_1, \dots, w_m)$ the material point volumes, $\mathbb{C} = \text{diag}(\mathbb{C}_1, \mathbb{C}_2, \dots, \mathbb{C}_m)$ the metric coefficients and f the applied external force vector.

From the two solutions of Eq. (2.6) the projection operations follow as

$$P_\varepsilon(\varepsilon') = Bu \quad (2.7)$$

$$P_\sigma(\sigma') = \sigma' + \mathbb{C}B\eta. \quad (2.8)$$

Note that B denotes the assembled B-operator including the degrees of freedom accounting for boundary conditions. Also, u contains u_{red} as well as the applied Dirichlet boundary conditions and η contains η_{red} as well as zeros at the constraint degrees of freedom.

For the two projection operations involved, i.e. the computation of the closest compatible strain field $P_\varepsilon(\varepsilon') = \arg \min_{\varepsilon \in E_\varepsilon} \|\varepsilon - \varepsilon'\|^2$ and the closest equilibrium stress field $P_\sigma(\sigma') = \arg \min_{\sigma \in E_\sigma} \|\sigma - \sigma'\|^2$ only two linear, decoupled system of equations have to be solved.

It is important to realize that the number of possible data assignments $|D| = \prod_{e=1}^m n_e$ is of combinatorial complexity and a solution to problem (2.5) is not directly computable. Therefore, several heuristic solver schemes has been proposed. The fixed-point solver proposed by [82] which we call in the following the "original fixed-point solver" is

summarized in Alg. (1).

Algorithm 1 Original Fixed-Point Solver

Require: Constraint set $E = E_\varepsilon \times E_\sigma$.

Require: Local data set(s) $D_e = \{(\varepsilon'_i, \sigma'_i)\}_{i=1}^{n_e}$.

Require: Convergence tolerance TOL

Initialize: $(\varepsilon', \sigma') \in D$ randomly

Initialize: $F_0 \leftarrow 0, F \leftarrow \infty$

while $|F - F_0| < \text{TOL}$ **do**

Set: $\varepsilon \leftarrow P_\varepsilon(\varepsilon')$

Set: $\sigma \leftarrow P_\sigma(\sigma')$

Set: $F_0 \leftarrow F$

Set: $F \leftarrow \frac{1}{m} \sum_{e=1}^m (\|\varepsilon_e - \varepsilon'_e\| + \|\sigma_e - \sigma'_e\|)$

for $e = 1 \dots m$ **do**

Set: $(\varepsilon'_e, \sigma'_e) \leftarrow \arg \min_{(\varepsilon'_i, \sigma'_i) \in D_e} (\|\varepsilon_e - \varepsilon'_i\|^2 + \|\sigma_e - \sigma'_i\|^2)$

end for

end while

return $(\varepsilon', \sigma'), (\varepsilon, \sigma)$

2.4 Randomized fixed-point solver: an alternative solver scheme

2.4.1 Motivation

The first contribution is an enhancement of the original solver scheme summarized in Alg. 1, which is beneficial for all kind of data-driven problems within the present framework. As also reported in [73, 125], the original solver fails to find the global minimizer of the optimization problem, in particular in the case of sparse data sets. In addition, the accuracy depends on the metric parameter $\mathbb{C}_e$, which has to be chosen in advance. To improve the accuracy, [73] proposes a mixed-integer global optimization technique, which becomes however computationally inefficient for larger systems. More closely to the original algorithm, [48, 125] proposes to adapt the metric coefficient online in an adaptive way in order to improve convergence properties, which requires however to assemble the stiffness matrix and build the data search structures multiple times during the online computation and additional local fitting procedures.

2.4.2 Proposed method: randomized fixed-point solver

We propose a variation of the original algorithm that overcomes the aforementioned issues while requiring only minimal changes. The basic idea of the algorithm is to perform the data search separated for stresses and strains in each material point in a randomized manner. In contrast to the original algorithm, which assigns to each material point the data point that minimizes simultaneously the distance of stresses and strains, our proposed algorithm searches only for the data which is closest with respect to *either* stresses *or* strains. If a given material point executes a strain search (and not a stress search), it is decided randomly with probability $p = 1/2$ at each iteration. Since we cannot expect that the error will decrease in every iteration, it is necessary to adapt the convergence criteria. The algorithm will be terminated if the optimal solution cannot be further improved after a certain number of iterations. The algorithm is summarized in Alg. 2.

Algorithm 2 Randomized Fixed-Point Solver

Require: Constraint set $E = E_\varepsilon \times E_\sigma$.
Require: Local data set(s) $D_e = \{(\varepsilon'_i, \sigma'_i)\}_{i=1}^{n_e}$.
Require: Maximum number of iterations with increasing error n_{trial}

Initialize: $(\varepsilon', \sigma') \in D$ randomly
Initialize: $F_{\min} \leftarrow \infty$
 $j_{\min} = 0, j = 0$
while $j < j_{\min} + n_{\text{trial}}$ **do**
 Set: $\varepsilon \leftarrow P_\varepsilon(\varepsilon')$
 Set: $\sigma \leftarrow P_\sigma(\sigma')$
 Set: $F \leftarrow \frac{1}{m} \sum_{e=1}^m (\|\varepsilon_e - \varepsilon'_e\| + \|\sigma_e - \sigma'_e\|)$
 if $F < F_{\min}$ **then**
 Set: $F_{\min} \leftarrow F$
 Set: $j_{\min} \leftarrow j$
 end if
 Set: $j \leftarrow j + 1$
 for $e = 1 \dots m$ **do**
 Set: $r \sim \text{Ber}(1/2)$ #Bernoulli random variable
 if $r = 0$ **then**
 Set: $(\varepsilon'_e, \sigma'_e) \leftarrow \arg \min_{(\varepsilon'_i, \sigma'_i) \in D_e} \|\varepsilon_e - \varepsilon'_i\|^2$
 else
 Set: $(\varepsilon'_e, \sigma'_e) \leftarrow \arg \min_{(\varepsilon'_i, \sigma'_i) \in D_e} \|\sigma_e - \sigma'_i\|^2$
 end if
 end for
end while
return $(\varepsilon', \sigma'), (\varepsilon, \sigma)$

2.4.3 Numerical example: stamp**2.4.3.1 Setup**

We compare the proposed randomized fixed-point solver with the original fixed-point solver by means of a numerical example. The considered boundary value problem is shown in Fig. (2.1). The data set in this section is set to the reference solution using a three-dimensional von-Mises plasticity material model with linear isotropic hardening containing 34606 data points.

The material parameters are chosen as follows: Young's modulus $E = 1000.0[\text{MPa}]$, Poisson's ratio $\nu = 0.3$, hardening modulus $H = 100.0[\text{MPa}]$ and initial yield strength

$\sigma_{y0} = 25.0[\text{MPa}]$. The shear loading $\bar{u}(t) = (10t/t_{\max}, 0, 0)$ is applied in 10 load steps. The finite element mesh consists of 3146 linear tetrahedral elements in three dimensional space.

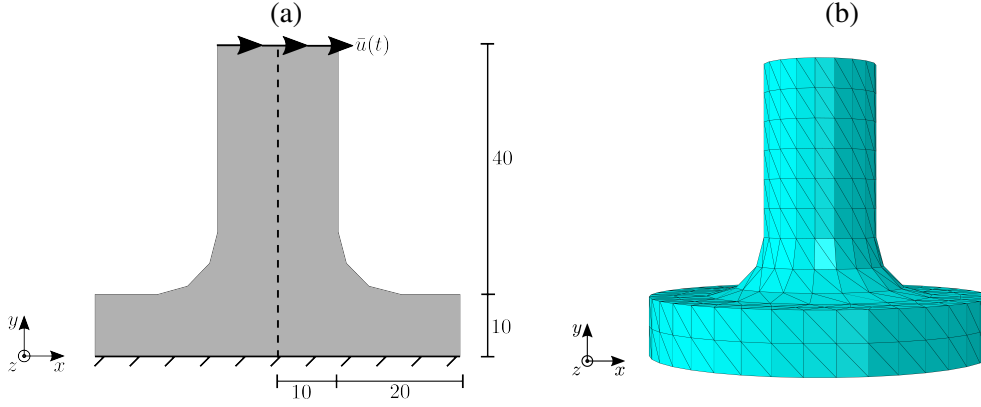


Figure 2.1: Stamp boundary value problem.

2.4.3.2 Results

First, we investigate the influence of the metric coefficients $\mathbb{C}_e = cI$, where we set the metric tensor proportional to the identity matrix. We consider a single time step at $t = 0.5t_{\max}$. The error is defined by the mean distance of the converged solution (ε', σ') to the reference solution $(\varepsilon_{\text{ref}}, \sigma_{\text{ref}})$ as $\mathcal{E}_{\text{ref}} = \frac{1}{m} \sum_{e=1}^m (\|\varepsilon_{e,\text{ref}} - \varepsilon'_e\| + \|\sigma_{e,\text{ref}} - \sigma'_e\|)$ using a fixed reference metric $\mathbb{C}_{e,\text{ref}} = 1000I$ so that the error measurement is independent of the metric parameter and thus comparable for different choices of c . Since the reference solution is included in the data set, we ideally expect any data-driven solver scheme to find a solution which satisfies equilibrium and compatibility up to numerical precision of the data. The results shown in Fig. 2.2 A) depicts the dependency of the original solver scheme on the metric coefficient c . Additionally, the original algorithm fails to find the global minimizer even for an optimal choice of the metric. In contrast, remarkably, the proposed randomized solver scheme computes independently of the chosen metric an almost optimal solution. Performing 30 independent runs, the original algorithm achieves in the best case for $c = 1000$ an error of 0.55 while the randomized solve achieves on average over all choices of c an error of $2.2 \cdot 10^{-5}$ with a standard deviation of $2.0 \cdot 10^{-4}$. These measures are heavily influenced by the outliers which

does not find the exact solution: 423 out of 450 (15 different metrics times 30 samples each) computations have an error below 10^{-16} meaning that the exact reference solution was found in these cases. Thus, the randomized solver finds independently of the metric coefficient c in most runs the optimal solution. In Fig 2.2 B) the evolution of the error during the iterations is shown for both algorithms for $c = 1000$. While the original algorithm takes in the first iterations large steps towards the minimum, it cannot effectively escape local minima. The randomized algorithm starts slower and does not necessarily decrease the error in each iteration, however, it has the capability to compute almost optimal solutions. Fig 2.2 C) visualizes the local data set and the computed solution of optimal data assignments of each material point.

The randomized solver comes with no computational overhead. The computation time per iteration for the original algorithm is on average 0.034[s] and 0.025[s] for the randomized algorithm. One explanation for this is that the data search in the randomized algorithm is distributed to two search structures of dimensionality 6 (dimension of stress/strain) instead of a single search structures of dimensionality 12, without adding any additional number of search operations. In addition, we emphasize that only minimal changes are required to implement the proposed randomized fixed-point solver with respect to the original fixed-point solver [82]. Nevertheless, we observe significant improvements in both accuracy and robustness with respect to the choice of the metric parameters.

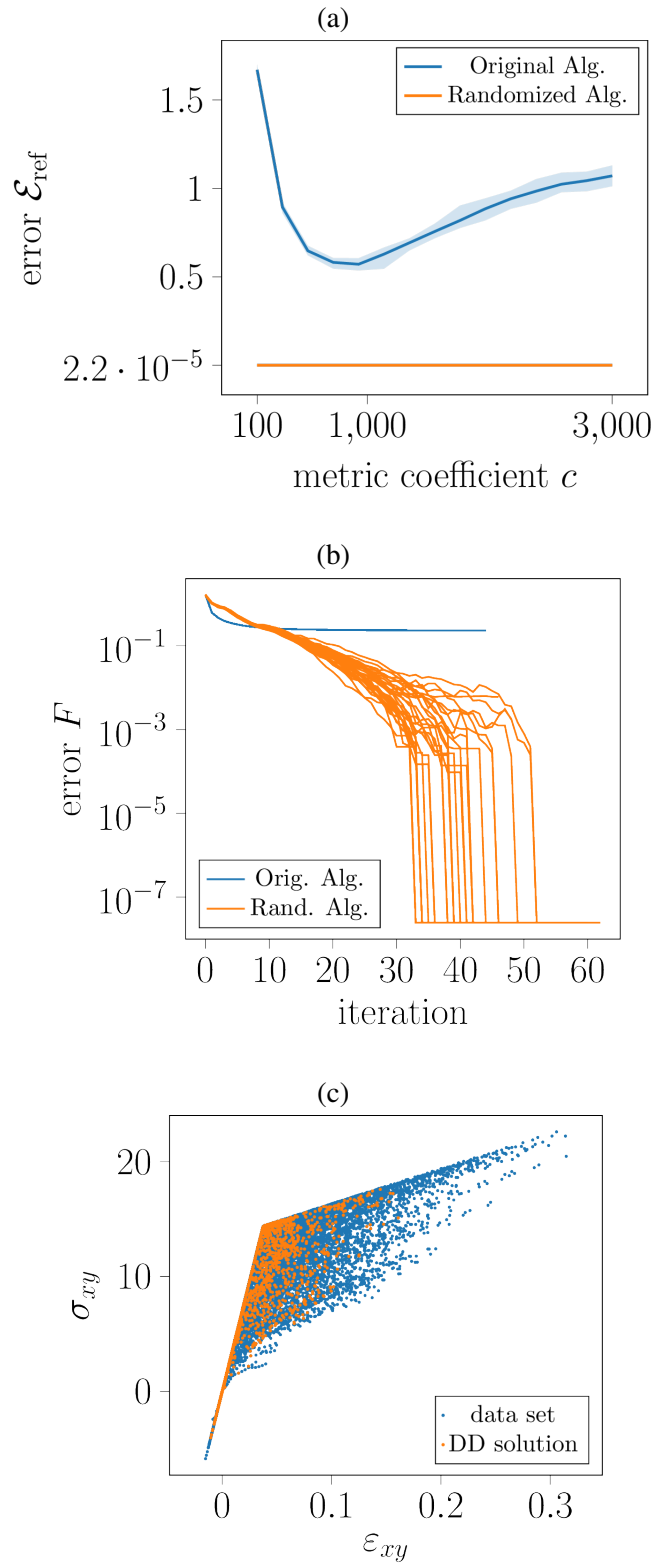


Figure 2.2: A) Influence of the metric parameter on the original and randomized algorithm. Shaded areas denote the maximum deviations from the mean of 30 independent runs. B) Error in constraints during the iterations for both algorithms for $c = 1000$. C) Blue: Local data set, orange: optimal data point assignment of each material point.

2.5 Adaptive data generation using FFT-based microstructure models

2.5.1 Motivation

In this section, we consider the more realistic scenario in which only an incomplete or even empty data set is available. We perform virtual experiments using a computational microstructure model (in the following also called representative volume element (RVE) model), using its homogenized stress and strain response as data points. Since we use the data from the RVE model directly in the macroscopic boundary value problem, we ensure that the material response is "ground-truth" and not affected by possibly wrong material model predictions. Of course, the "ground-truth" of the microstructure model is also assumption which has to be validated in real-world applications. However, computational RVE models offer a convenient way to systematically improve the accuracy of the complex material behavior by means of refinement of the RVE.

In particular, we develop a systematic and automated scheme for enriching a data set in order to solve a boundary value problem with a data-driven solver scheme with a desired accuracy, while minimizing the number of costly RVE evaluations. One advantage of the direct data-driven paradigm is that a data-driven solution, i.e., an optimal assignment of data points to material points, provides a direct measure of the data that is missing in order to fulfil equilibrium and kinematical compatibility for a given boundary value problem. Using this information leads to a so-called goal-oriented data generation scheme since it generates data for a specific BVP [54]. Nevertheless, this data may also be useful for further, different subsequent loading conditions or different BVPs with the same material.

This paper considers the behavior of a rate-independent plastic material. However, unloading is not considered, as the focus is on the performance of the data generation scheme, which is isolated from the effects of a history parametrization for the data-driven solver scheme. It is important to note that even simple loading is path-dependent, and thus the entire loading path must be considered when evaluating the RVE.

2.5.2 Proposed method: adaptive data generation

The general procedure of the proposed method is as follows: Step 1) Given the current data set compute a data-driven solution. Step 2) Compute new proposal strain paths using the given information of data, equilibrium, and compatibility. Step 3) Cluster these strain paths to avoid duplicate RVE evaluations using the error measurement. Step 4) Evaluate RVE with clustered strain paths and add the new stress-strain paths to the data set. Then repeat until desired tolerance is achieved. The following subsections provide the details for each step.

2.5.2.1 Step 1) data-driven solution

To compute the optimal assignment of data points to material points we use the randomized solver scheme proposed in section 2.4. Since a sequence of loadings is applied to the BVP, we label each solution by a time index $t \in [1, 2, \dots, T]$ as $(\varepsilon'_t, \sigma'_t)$ and $(\varepsilon_t, \sigma_t)$, thus obtaining *path data*. Recall that $(\varepsilon'_t, \sigma'_t) \in D$ is the optimal data field and $(\varepsilon_t, \sigma_t) \in E$ the closest admissible strain and stress field w.r.t. to that assignment of data. We indicate path data by an underline, for instance one local, compatible, strain path is denoted by $\underline{\varepsilon}_e = (\varepsilon_{e,1}, \varepsilon_{e,2}, \dots, \varepsilon_{e,T})$. We also store the local error $F_{e,t} = \|\varepsilon'_{e,t} - \varepsilon_{e,t}\|^2 + \|\sigma'_{e,t} - \sigma_{e,t}\|^2$ at every material point e in every time step t .

2.5.2.2 Step 2) strain paths proposals

We propose to store and use the homogenized material tangent information of the RVE, which may be either computed analytically or by means of strain perturbations. For each data point $(\varepsilon'_i, \sigma'_i) \in D_e$ the measured tangent is denoted by T'_i . Then, we denote the tangent associated to the local data point assignment $(\varepsilon'_{e,t}, \sigma'_{e,t}) \in D_e$ of material point e at time t by $T'_{e,t}$. The whole tangent field corresponding to the assignment $(\varepsilon'_t, \sigma'_t) \in D$ is then denoted by T'_t . Using the tangent information allows us to evaluate the RVE in a purely strain-driven manner, even in a force-driven BVP, which is easy to implement and always available in contrast to mixed-loading RVE evaluations.

We now compute proposal strain fields ε_t^* for each time step t independently, in such a way that each is kinematically compatible, and the associated tangent stress field is in equilibrium. Kinematical compatibility is ensured by expressing $\varepsilon_t^* = B_{\text{red}} u_{\text{red},t} + \varepsilon_t$. The associated tangential stress field around the computed data field $(\varepsilon'_t, \sigma'_t)$ is then given

by $\sigma(u_{\text{red},t}) = \sigma'_t + T'_t(B_{\text{red}}u_{\text{red},t} + \varepsilon_t - \varepsilon'_t)$. We enforce equilibrium by the discretized equilibrium equation $B_{\text{red}}^T W \sigma(u_{\text{red},t}) = f_{\text{red}}$. After inserting we obtain

$$(B_{\text{red}}^T W T'_t B_{\text{red}}) u_{\text{red},t}^* = f_{\text{red}} - B_{\text{red}}^T W (\sigma'_t + T'_t(\varepsilon_t - \varepsilon'_t)) \quad (2.9)$$

which can be solved for $u_{\text{red},t}^*$ and thus gives the proposal strain paths $\underline{\varepsilon}^* = \{\varepsilon_e^*\}_{e=1}^m = (\varepsilon_t^* : t \in [1, T], \varepsilon_t^* = B_{\text{red}} u_{\text{red},t}^* + \varepsilon_t)$.

The underlying assumption for this approach is, that the material behavior in the neighborhood of a data point $(\varepsilon'_i, \sigma'_i)$ is close to its material tangent T'_i . It is important to note that this procedure serves only to compute the proposal strain paths in an hopefully optimal manner for the RVE evaluations, i.e. no extrapolated stress-strain data is included in the data set, but only actually measured data.

2.5.2.3 Step 3) error-weighted clustering

For each local proposal strain path $\underline{\varepsilon}_e^*$ we define a weight which accounts for the error of the material point observed during the data-driven computation by

$$p_e = \exp\left(\frac{1}{\alpha} \sum_{t=1}^T (F_{e,t})\right) - 1, \quad \alpha = \frac{1}{m} \sum_{e=1}^m \sum_{t=1}^T (F_{e,t}). \quad (2.10)$$

Using this information, we employ a *weighted* k-means cluster algorithm [110] to cluster the strain paths using the weights p_e in order to give more importance to paths with larger error, with the result

$$\bar{\underline{\varepsilon}}_k^* = \frac{\sum_{e \in S_k} p_e \underline{\varepsilon}_e^*}{\sum_{e \in S_k} p_e}, \quad k = 1 \dots K \quad (2.11)$$

where $\bigcup_{k=1}^K S_k$ is the partition of all material points which minimizes the variance

$$\sum_{k=1}^K \sum_{e \in S_k} p_e \|\underline{\varepsilon}_e^* - \bar{\underline{\varepsilon}}_k^*\|^2. \quad (2.12)$$

For the sake of compactness we do not provide the details of the well-known (weighted) k-means clustering as mainstream libraries [127] provide implementations which can be

directly used without any modification.

The result of the k-means algorithm are the cluster centers $\{\bar{\varepsilon}_k^*\}_{k=1}^K \leftarrow \text{KMeans}(\{\varepsilon_e^*, p_e\}_{e=1}^m)$. Instead of using these cluster centers for the evaluation of the RVE we use the proposal strain path which are the nearest to each cluster center. This operation is performed by the nearest neighbor search

$$\underline{\varepsilon}_k^* = \arg \min_{\underline{\varepsilon}_e^* \in \bar{\varepsilon}^*} \|\underline{\varepsilon}_e^* - \bar{\varepsilon}_k^*\|^2, \quad \text{for } k = 1 \dots K. \quad (2.13)$$

The number of clusters K has to be set in advance. Given the number of material points m , we prescribe a certain cluster ratio $r = K/m$ from which K follows. We will study the influence of the cluster ratio in the numerical example 2.5.4.

2.5.2.4 Step 4) RVE evaluation

Finally, we compute the new stress data associated to the strain paths $\{\underline{\varepsilon}_k^*\}_{k=1}^K$. We perform the RVE evaluations using the Fast-Fourier-Transform (FFT) based fixed point algorithm proposed by [118]. More efficient solver schemes [71] are available for even faster evaluation times. For a review of the state of the art of these solvers and the computation of material tangents we refer to [147, 51, 53]. The FFT based solver schemes are efficient for the evaluation of periodic microstructures and can be conveniently evaluated with the strain paths without the need to compute an equivalent displacement loading. In the solver scheme, a reference stiffness $\mathbb{C}_0$ has to be defined, usually as the average elastic stiffness matrix of each constituent. For convenience, we use this reference stiffness also as the metric coefficients $\mathbb{C}_e$ of the data-driven solver scheme.

The availability and acquisition of realistic three-dimensional microstructures is still limited and subject to research. In this work, we obtained a sample from the public online library Microlib [81] which is based on a generative adversarial network (GAN) architecture that generates 3D microstructures based on real images of microstructure slices[80]. The workflow for generating the FFT model used in the following using this method is visualized in Fig. (2.3)

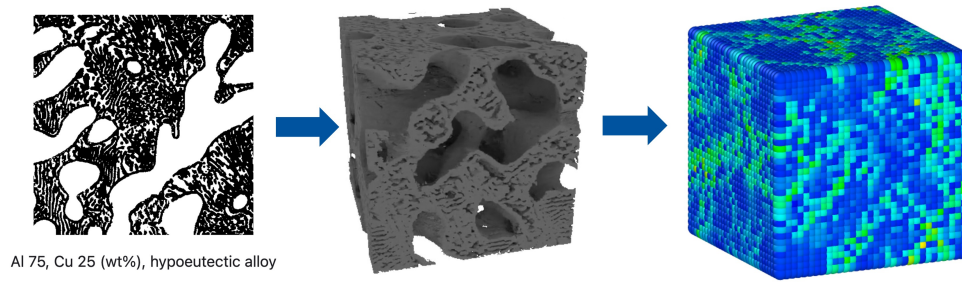


Figure 2.3: Workflow for the generation of the FFT model using SliceGAN (<https://microlib.io>) [80, 81]: 1) real image of a microstructure 2) generated 3D microstructure from a trained GAN 3) downsampled FFT-model with 32^3 pixels showing a strain field.

2.5.3 Algorithm summary

The proposed adaptive data generation algorithm is summarized in Algorithm 3.

Algorithm 3 Adaptive data generation**Require:** Constraint set(s) E_t .**Require:** Initial local data set(s) with tangents $D_e = \{(0, 0, T'_0)\}$.**Require:** Target error tolerance tol Initialize: $it \leftarrow 0$ **while** $F_{99\%} > \text{tol}$ **do** **for** $t = 1 \dots T$ **do** Set: $\varepsilon_t, \sigma_t, \varepsilon'_t, \sigma'_t, F_t \leftarrow \text{Min_Dist}(E_t, D_e)$ Solve Eq. (2.9) for $u_{\text{red},t}^*$ Set: $\varepsilon_t^* \leftarrow B_{\text{red}} u_{\text{red},t}^* + \varepsilon_t$ **end for**Set: $F_{99\%} \leftarrow \text{percentile}(F_{e,t}, 99)$ Set: p_e (Eq. (2.10))Set: $K = \text{int}(r(it) * m)$ # cluster ratio * number of material pointsSet: $\{\bar{\varepsilon}_k^*\}_{k=1}^K \leftarrow \text{KMeans}(\varepsilon^*, \text{sample_weights} = p_e)$ Set: $\underline{\varepsilon}_k^* \leftarrow \arg \min_{\varepsilon_e^* \in \varepsilon^*} \|\varepsilon_e^* - \bar{\varepsilon}_k^*\|^2$ **for** $k = 1 \dots K$ **do** Set: $\underline{\sigma}_k^*, \underline{T}_k^* \leftarrow \text{RVE}(\underline{\varepsilon}_k^*)$ Set: $D_e \leftarrow D_e \cup (\underline{\varepsilon}_k^*, \underline{\sigma}_k^*, \underline{T}_k^*)$ **end for**Set: $it \leftarrow it + 1$ **end while****return** D_e **2.5.4 Numerical example: beam****2.5.4.1 Setup**

As a numerical example, we consider the beam shown in Fig. (2.4). The monotonically increasing force $F(t)$ is distributed over the surface A as $F(t) = \int_A 5 \frac{t}{[s]} da$ [MN] for time $t \in (0, 1)[s]$ within 20 time steps, having the whole left edge of the beam clamped in all directions. The mesh consists of 13753 linear tetrahedral elements.

We assume homogeneous material behavior over the structure and no randomization of the microstructure, i.e. the material behavior of every material point is described by the same RVE. The RVE is resolved by 32^3 pixels as shown in Fig. (2.3) and contains two phases, Aluminium and Copper. We assume that the material behavior of each phase is less complex than the overall response of the RVE and can thus be described by

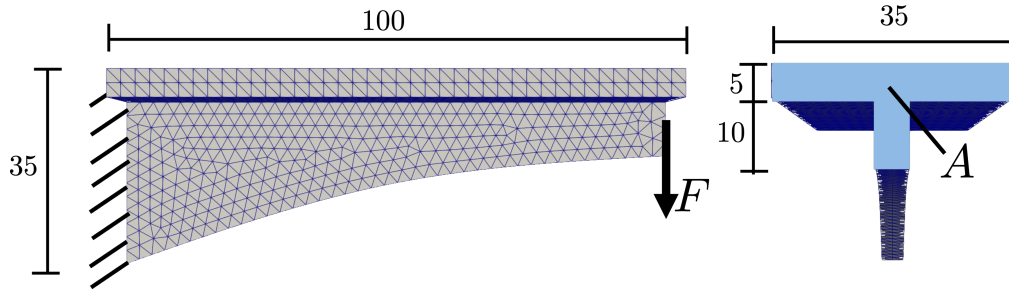


Figure 2.4: Beam boundary value problem.

a Von-Mises plasticity model with isotropic hardening. The material parameters for the two phases are chosen as follows: $E_{Al} = 68000[\text{MPa}]$, $\nu_{Al} = 0.36$, $\sigma_{y0,Al} = 240[\text{MPa}]$, $H_{Al} = E_{Al}/1000$ and $E_{Cu} = 110000[\text{MPa}]$, $\nu_{Cu} = 0.343$, $\sigma_{y0,Cu} = 33.3[\text{MPa}]$, $H_{Cu} = E_{Cu}/1000$.

2.5.4.2 Results

We evaluate the performance of the proposed adaptive data generation scheme for different cluster ratios. Every simulation starts with a data set containing only a zero data point and is iterated until the error tolerance 10^{-5} is achieved. For this, we measure the 99-percentile error $F_{99\%}$ of all local errors $F_{e,t}$. This error measure is much stricter than for instance the mean or median of all local errors, but excludes a small fraction of maximum errors which tend to fluctuate due to inaccuracies of the data-driven solver scheme.

First, we have run the adaptive data generation scheme for the cluster ratios $r = 1\%$ and $r = 10\%$. The error over the data set size is shown in Fig. 2.5). For a better comparison, we also provide the data set size per number of material points and time steps. We observe that it is much more efficient to add very few data (i.e. a cluster ratio of 1%) in the first iterations instead of adding larger amounts of data. The reason for this is that at this stage the strain paths are not yet converged and thus the generated data cannot be reused in later iterations. However, once the system has reached a certain level of convergence (the cutting point of the orange and blue line in Fig. (2.5) A)) a higher cluster ratio is more efficient, since a strong clustering smears out the accurately requested proposal paths too much.

This motivates our suggestion to start slowly by adding very little data, but then

successively increasing the cluster ratio. Based on this we have run the scheme using a successively increasing cluster ratio defined by the heuristic formula

$$r(it) = [0.1 + (1 - e^{(-it/20)})(10 - 0.1)]\% \quad (2.14)$$

where it is the iteration number and 0.1% and 10% chosen starting and final cluster ratios, respectively. As shown in Fig. (2.5) A) by the green curve this increasing cluster ratio scheme achieves the best results, i.e. achieves any error level with a minimum amount of data.

In order to assess what is a reasonable error level we evaluate the macroscopic response of the system as shown in Fig. (2.5) B). Here, we measure the average, vertical displacement of the right edge of the beam as function of the applied force density. From this, we conclude visually that an error range in the order of $F_{99\%} \sim 10^{-3} - 10^{-4}$ results in a converged macroscopic system response.

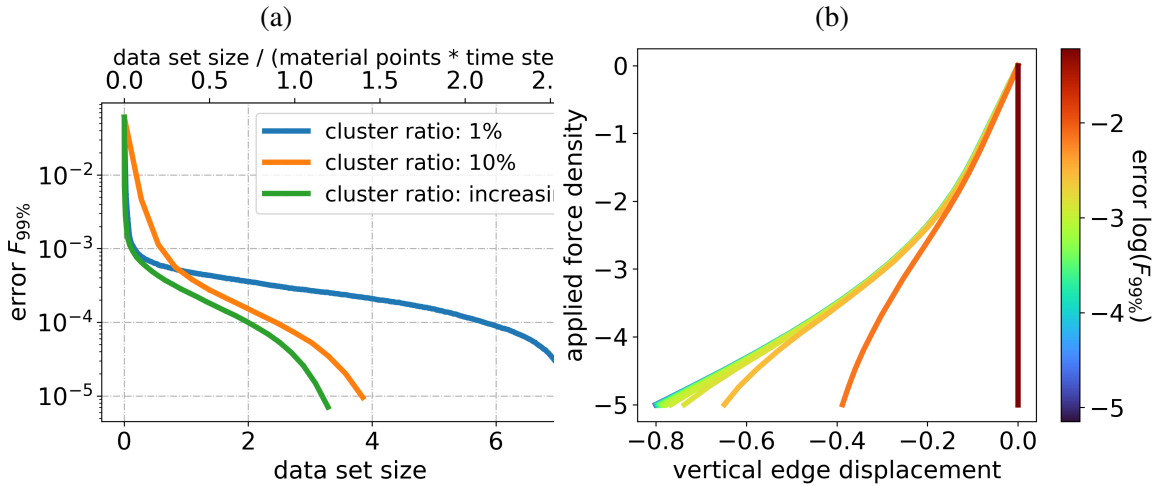


Figure 2.5: A) 99-percentile error for the iteratively enriched data sets using different cluster ratios. B) Force-Displacement curve for different iterations of data generation. Colored by the logarithm of the 99-percentile error.

The measured CPU times (as computed on a 24 cores shared cluster node) are reported in Fig. (2.6). We note that, for sufficiently large and realistic RVEs, the CPU time correlates directly to the amount of generated data, rather than to the number of iterations of data generation. The reason for this is, that the data-driven solving time compared to

the data generation time is much smaller, as shown in Fig. (2.6) B) for the increasing cluster ratio scheme. Note that the clustering is performed in the order of a few seconds, and it's computing time thus neglectable. Choosing a reasonable target tolerance of $F_{99\%} = 10^{-4}$ the overall computational cost is approximately 12 hours, thus enabling the solving of the three-dimensional multiscale problem in an acceptable time.

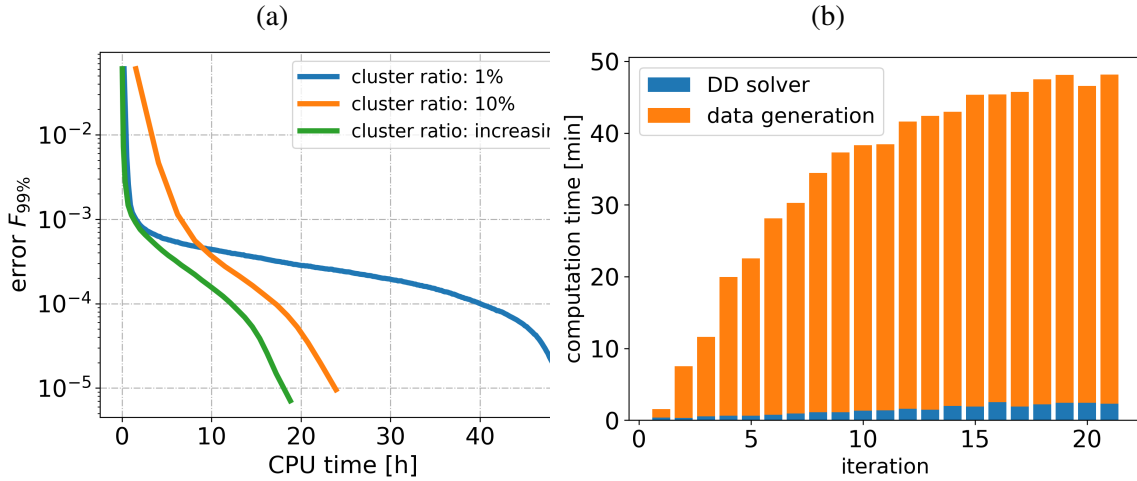


Figure 2.6: A) CPU times for the different cluster ratios. B) CPU times per iteration for the increasing cluster ratio scheme until a accuracy of $F_{99\%} < 10^{-4}$ is reached.

Finally, we compare the performance of the newly proposed randomized fixed-point solver scheme to the original solver as shown in Fig. (2.7). In addition to the 99-percentile error, we show the maximum error $F_{\max} = \max F_{e,t}$, i.e. the maximum local error of all material points and times steps. The randomized solver clearly outperforms the original solver by efficiently escaping local minima.

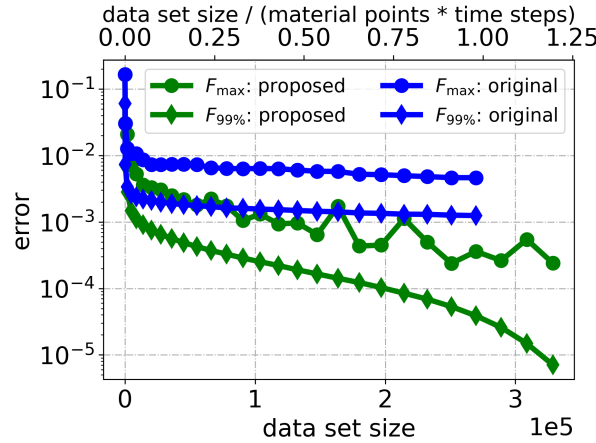


Figure 2.7: Comparison of the original and the newly proposed randomized fixed-point solver schemes for the increasing cluster ratio. In addition to the 99-percentile error, the maximum error is shown.

2.6 Summary and concluding remarks

We have presented two algorithms, an accurate solver algorithm and a strategy for adaptive data generation. Both have been shown to lead to significant improvements for relevant, three-dimensional multiscale problems.

The proposed solver algorithm is easy to implement and requires only minimal changes from the original solver algorithm. However, it can significantly improve the accuracy. This is achieved by searching separately for stresses and strains, thus effectively removing the impact of choice for metric coefficients. By randomization, local minima can be efficiently escaped. Since no additional nearest neighbor searches or other operations are performed, we observe no additional computational costs.

Based on this solver, we developed a new adaptive data generation algorithm using the material tangent information and an error-weighted k-means clustering. This automated and systematic approach to dataset enrichment has consistently minimized errors. For the numerical example under study, we have reached an acceptable error level at a dataset size below the number of material points times the number of time steps. In comparison to a classical inelastic finite element computation, where the RVE must be evaluated multiple times per time step, the proposed method demonstrates significant potential. We anticipate the benefits to be even more substantial when multiple boundary value

problems are computed, allowing for the reuse of previously computed data.

The error-weighted k-means clustering guidance for RVE evaluations can also be beneficial for model-based schemes. In this context, it can be used to validate the material model for predictions in regions without measured data by optimal selected, limited RVE evaluations. Using this for future works, a fair comparison between a model-based and direct data-driven scheme can be performed. While the clustering algorithm already incorporates strain path RVE evaluations, future works need to include history dependent data-driven solvers [33] in order to compute unloading scenarios. Finally, the proposed method can be extended to data from randomized microstructures and potentially from different pixel resolution levels for uncertainty quantification purposes.

2.7 Acknowledgements

EP thanks Kevin Linka for very helpful discussions. EP, CG and SR gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG) through the project "Multiscale modeling based on a novel combination of direct data-driven methods with Fourier transform-based microstructure simulation" (RE 1057/58-1, project number 532163998). MO gratefully acknowledges the support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) *via* project 211504053 - SFB 1060; project 441211072 - SPP 2256; and project 390685813 - GZ 2047/1 - HCM.

3 | **Article 2:**

Model-Free Data-Driven Inference in Computational Mechanics

This article was published as:

Prume, E., Reese, S., and Ortiz, M. [2023], 'Model-free data-driven inference in computational mechanics', *Computer Methods in Applied Mechanics and Engineering* **403**, 115704. ope

Disclosure of the individual authors' contributions to the article:

M. Ortiz drafted the initial formulation of the likelihood function. E. Prume developed the concepts into practicable forms and solver algorithms and subsequently drafted the manuscript. Further, E. Prume implemented the algorithms and computed the numerical examples. S. Reese and M. Ortiz reviewed the draft and made valuable comments and edits. All authors approved the publication of the final version of the manuscript.

3.1 Abstract

We extend the model-free Data-Driven computing paradigm to solids and structures that are stochastic due to intrinsic randomness in the material behavior. The behavior of such materials is characterized by a likelihood measure instead of a constitutive relation. We specifically assume that the material likelihood measure is known only through an empirical point-data set in material or phase space. The state of the solid or structure is additionally subject to compatibility and equilibrium constraints. The problem is then to infer the likelihood of a given structural outcome of interest. In this work, we present a Data-Driven method of inference that determines likelihoods of outcomes from the empirical material data and that requires no material or prior modeling. In particular, the computation of expectations is reduced to explicit sums over local material data sets and to quadratures over admissible states, i. e., states satisfying compatibility and equilibrium. The complexity of the material data-set sums is linear in the number of data points and in the number of members in the structure. Efficient population annealing procedures and fast search algorithms for accelerating the calculations are presented. The scope, cost and convergence properties of the method are assessed with the aid selected applications and benchmark tests.

3.2 Introduction

Some classes of materials exhibit behavior that is intrinsically random. A compelling example is set forth by brittle materials with random tensile strength [14]. Structural members made of such materials may fail—and lose their load-bearing capacity altogether—at tensile stresses that are sensitive to critical flaws in the material. Since such flaws have random geometry, the tensile strength of the material is likewise random (cf., e. g., refs. [139, 11] concerning the example of carbon fibers). In addition, empirical material data may be noisy due to experimental measurement error and scatter. This observational noise may be reduced by improving the precision of the experimental measurements. By contrast, the intrinsic randomness of the material is not eliminated by removing observational noise.

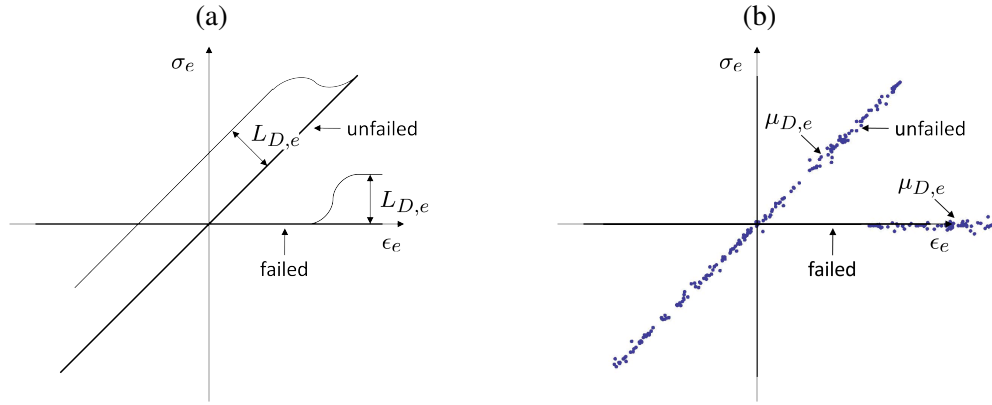


Figure 3.1: Brittle material with random tensile strength, member e of the structure. a) Local material likelihood function $L_{D,e}$ providing the likelihood that the local material state (ϵ_e, σ_e) be on the unfailed and failed branches of the local material set, respectively. b) Empirical local likelihood measure $\mu_{D,e}$ sampled from $L_{D,e}$.

The behavior of random materials is characterized by a likelihood measure instead of a constitutive relation. In principle, any local material state (ϵ_e, σ_e) is possible, albeit with varying likelihood. For instance, for a brittle material with random tensile strength undergoing monotonically increasing strains, the stress may take values in anyone of two branches, corresponding to a failed or un-failed material, with respective likelihoods determined by the distribution of tensile strength, Fig. 3.1a.

In addition, local stresses and strains in solids and structures are subject to compatibility and equilibrium constraints. These field constraints are known exactly, though the loading may also be random in some applications. The inference problem is then to determine the likelihood of structural outcomes of interest when the states of the structure are required to be admissible, in the sense of satisfying compatibility and equilibrium, and the material is characterized by a given material likelihood measure.

However, the requisite material likelihood measures are often unknown exactly and characterized only partially through empirical material data, Fig. 3.1b. A common response to that challenge is to model the empirical material data, e. g., by assuming the existence of an underlying reduced manifold in phase space masked by noise obeying a prespecified prior distribution. However, in general no such reduced manifold may exist and the form of the prior may be unknown and subject to surmise, which inevitably

introduces modeling bias and uncertainty and may prevent convergence with respect to the empirical data altogether [126].

For instance, for brittle materials with random tensile strength the material states may occupy one of two branches, corresponding to a failed or un-failed material, and no single-valued stress-strain relation exists, Fig. 3.1a. In addition, in approaches based on Bayesian statistics (cf., e. g., [28] for a review) it is common to presume a certain class of priors, most commonly Gaussian [85, 86], which may not contain the actual likelihood function of the material. For instance, if the tensile strength of a brittle material obeys a Weibull distribution (cf., e. g., [5, 122] for the example of carbon fibers), methods assuming a Gaussian prior inevitably converge to the wrong answer with increasing sample size.

In this work, we present an alternative method of inference that determines likelihoods of outcomes directly from the empirical material data and requires no material or prior-distribution *ansätze*. We begin by ‘thermalizing’ the data, which effectively replaces the empirical material likelihood by a sum of Gaussians of a certain width [21]. The thermalized likelihoods also follow variationally by minimizing a regularized Kullback-Leibler divergence [96, 95, 130] from the empirical material likelihood. Likewise, an optimal annealing schedule follows variationally by minimizing the regularized Kullback-Leibler divergence with respect to temperature. The sequence of thermalized likelihoods thus obtained is then used to approximate expectations of outcomes. In this manner, the computation of expectations is reduced to explicit sums over local material data sets and quadratures over admissible states, i. e., states satisfying compatibility and equilibrium. When restricted to the computation of maximum-likelihood outcomes, the present approach reduces to the minimum-distance (min-dist) Data-Driven method of [82] in the deterministic case with uniform data convergence, and to the maximum-entropy (max-ent) Data-Driven method of [83] in the case of noisy data with outliers.

It bears emphasis that, in the present Data-Driven approach, expectations are computed directly from the material data and that no material model or prior distribution need to be contrived at any point of the calculations. The complexity of the Data-Driven inference calculations is linear in the number of material data points and in the number of members in the structure or Gauss points in the solid. We formulate efficient population annealing procedures and fast-search algorithms for accelerating calculations further. Specifically, the population annealing algorithm reduces the complexity of the calculations and it

renders it linear in the size of the material data-point population. The calculations are further accelerated by recourse to a hierarchical k -means algorithm for the computation of the annealing energy with limits on the number backtracking operations.

The convergence properties of the method with respect to the empirical data are assessed with the aid of selected applications and benchmark tests, including a standard verification test based on Gaussian material data, brittle materials exhibiting random tensile strength, and a simple lightweight space structure. Robust convergence of posterior distributions is obtained in these tests, consistent with quantitative error estimates derived from analysis [21, 22]. The ability of the proposed Data-Driven method of inference to deal effectively with general material data sets and complex behavior without need for models, hypotheses or assumptions is quite remarkable. The numerical tests additionally attest to feasibility of the calculations and the effectiveness of acceleration methods such as population annealing and hierarchical k -means searches.

3.3 The inference problem

We consider finite-dimensional systems comprising m components whose state is characterized by two work-conjugate fields $\epsilon \equiv \{\epsilon_e \in \mathbb{R}^d, e = 1, \dots, m\}$ and $\sigma \equiv \{\sigma_e \in \mathbb{R}^d, e = 1, \dots, m\}$. We refer to the space of pairs $Z_e = \{z_e \equiv (\epsilon_e, \sigma_e) \in \mathbb{R}^d \times \mathbb{R}^d\}$ as the *local* phase space of component e , and $Z = Z_1 \times \dots \times Z_m = \mathbb{R}^{N \times N}$, $N = md$, as the *global* phase space of the system. We suppose that a suitable norm is defined in Z , e. g.,

$$\|z\| = \|(\epsilon, \sigma)\| = \left(\sum_{e=1}^m w_e \left(\mathbb{C}_e \epsilon_e \cdot \epsilon_e + \mathbb{C}_e^{-1} \sigma_e \cdot \sigma_e \right) \right)^{1/2}, \quad (3.1)$$

where $w_e > 0$ are weights and $\mathbb{C}_e \in \mathbb{R}_{sym,+}^{d \times d}$ are positive definite symmetric matrices, $e = 1, \dots, m$.

3.3.1 Example: trusses

We illustrate the essential structure of discrete field theories by means of the simple example of truss structures. Trusses are assemblies of bars that deform in uniaxial tension or compression. The bars are articulated at common joints, or nodes, that act as

hinges, i. e., cannot transmit moments. Trusses are examples of connected networks that obey conservation laws. Other examples in the same class include electrical circuits, pipeline networks, traffic networks, and others.

The material behavior of a bar e is characterized by a particularly simple relation between uniaxial strain ϵ_e and uniaxial stress σ_e . Thus, in this case the local phase spaces are $Z_e = \mathbb{R} \times \mathbb{R}$. These local states are subject to the following laws:

i) Compatibility: Suppose that bar e is connected to nodes a and b . Then, the strain in the bar is

$$\epsilon_e = \frac{u_b - u_a}{L_e} \cdot d_e, \quad (3.2)$$

where L_e is the length of the bar and d_e is the unit vector pointing from a to b .

ii) Equilibrium: Let S_a be the star of an unconstrained node a , i. e., the collection of members connected to a . Then, we must have

$$\sum_{e \in S_a} \sigma_e d_e A_e + f_a = 0, \quad (3.3)$$

where A_e is the cross-sectional area of bar e and f_a is the force applied to node a . $\square$

3.3.2 The constraint set of admissible states

The above example shows that the state of the system is subject to linear constraints of the general form

$$\sum_{e=1}^m w_e B_e^T \sigma_e = f, \quad (3.4a)$$

$$\epsilon_e = B_e u + g_e, \quad e = 1, \dots, m, \quad (3.4b)$$

where $u \in \mathbb{R}^n$ is the array of degrees of freedom of the system, w_e are positive weights, $B_e \in \mathbb{R}^{d \times n}$ is a discrete gradient operator, B_e^T is a discrete divergence operator, $f \in \mathbb{R}^n$ is a force array resulting from distributed sources and Neumann boundary conditions and the arrays $g_e \in \mathbb{R}^d$ follow from Dirichlet boundary conditions. The constraints (3.4) are material independent and define an affine subspace E of Z , the constraint set. The constraint set E encodes all the data of the problem, including geometry, loading and

boundary conditions. The constraints (3.4) can also be expressed in matrix form as

$$B^T \tau = f \quad (3.5a)$$

$$\epsilon = Bu + g, \quad (3.5b)$$

with $B = (B_1, \dots, B_m) \in \mathbb{R}^{N \times n}$, $\epsilon = (\epsilon_1, \dots, \epsilon_m) \in \mathbb{R}^N$, $\tau = (w_1 \sigma_1, \dots, w_m \sigma_m) \in \mathbb{R}^N$ and $g = (g_1, \dots, g_m) \in \mathbb{R}^N$.

We note that the affine space E defined by the constraints (3.5) is a translate of the linear space E_0 defined by the homogeneous constraints

$$B^T \tau = 0 \quad (3.6a)$$

$$\epsilon = Bu. \quad (3.6b)$$

Evidently, $E_0 = E_\epsilon \times E_\sigma$, where E_ϵ is the linear space defined by (3.6b) and E_σ is the linear space defined by (3.6a). Therefore, we have

$$\dim(E_0) = \dim(E_\epsilon) + \dim(E_\sigma) = \dim(\text{Im}(B)) + \dim(\text{Ker}(B^T)) = N. \quad (3.7)$$

Since $Z = \mathbb{R}^N \times \mathbb{R}^N$, it follows that the constraint set E is an affine subspace of Z of dimension N and co-dimension N .

This observation characterizes the structure and dimensionality of the phase space Z and of the subspace E of all admissible states in Z , or constraint set, i. e., the set of all the states that satisfy the conservation laws (3.5). To wit, the phase space has the conventional dual structure $Z = \mathbb{R}^N \times \mathbb{R}^N$ of work-conjugate pairs, whereas the constraint set E is a linear subspace of Z of dimension N and co-dimension N . This structure is particular to systems obeying conservation laws and sets them apart from other applications in data science where the data is heterogeneous and unstructured.

3.3.3 Material characterization

We assume that the material behavior is inherently random and characterized by a positive, continuous, likelihood function $L_D(y)$ representing the likelihood of observing a material state $y \in Z$, cf. Fig. 3.1a. Since material behavior is not localized to a bounded region of phase space in general, the material likelihood function L_D contains

infinite mass and cannot be normalized to define a probability density. The function $\Phi_D(y) = -\log L_D(y)$ is the corresponding *potential*.

3.3.4 Example: local material behavior

Material behavior is often *local* and can be characterized over each member phase space $Z_e = \mathbb{R}^d \times \mathbb{R}^d$ by a local material likelihood function $L_{D,e} \in \mathcal{M}(Z_e)$. Assuming that the behavior of the members is independent, then the *global* likelihood function is

$$L_D = L_{D,1} \times \cdots \times L_{D,m}, \quad (3.8)$$

and the *global* potential is

$$\Phi_D = \Phi_{D,1} + \cdots + \Phi_{D,m}. \quad (3.9)$$

Likelihood functions need not be integrable and do not define a probability density in general. For instance, suppose that the local behavior is characterized by a sliding Gaussian of the form, cf. Fig. 3.1a.,

$$L_{D,e}(y_e) = \exp\left(-\frac{w_e}{2s^2}\|\sigma_e - \mathbb{C}_e\epsilon_e\|^2\right) \quad (3.10)$$

where $y_e = (\epsilon_e, \sigma_e) \in Z_e$ and the parameter $s > 0$ measures the width of the distribution. The global likelihood function is, then,

$$L_D(y) = \exp\left(-\frac{1}{2s^2}\sum_{e=1}^m w_e\|\sigma_e - \mathbb{C}_e\epsilon_e\|^2\right), \quad \|\sigma_e\|^2 = \mathbb{C}_e^{-1}\sigma_e \cdot \sigma_e, \quad (3.11)$$

$y = (\epsilon, \sigma)$. Equivalently,

$$L_D(y) = \exp\left(-\frac{1}{2s^2}\left(\|\epsilon\|^2 + \|\sigma\|^2 - 2\sigma \cdot \epsilon\right)\right), \quad (3.12)$$

with

$$\|\epsilon\|^2 = \sum_{e=1}^m w_e \mathbb{C}_e \epsilon_e \cdot \epsilon_e, \quad \|\sigma\|^2 = \sum_{e=1}^m w_e \mathbb{C}_e^{-1} \sigma_e \cdot \sigma_e, \quad \sigma \cdot \epsilon = \sum_{e=1}^m w_e \sigma_e \cdot \epsilon_e, \quad (3.13)$$

or

$$L_D(y) = \exp \left(-\frac{1}{2s^2} \left(\|y\|^2 - 2\sigma \cdot \epsilon \right) \right) \quad (3.14)$$

in the norm (3.1). Evidently, $L_{D,e}$ decays away from the centerline $D_e = \{\sigma_e - \mathbb{C}_e \epsilon_e\}$ but is constant along the line. Hence,

$$\int_{Z_e} L_{D,e}(y_e) dy_e = +\infty, \quad (3.15)$$

and neither $L_{D,e}(y_e)$ nor $L_D(y)$ can be normalized. $\square$

3.3.5 Classical inference

If the constraint set E and the material likelihood function L_D are fully known, , cf. Fig. 3.1a, the expectation of any quantity of interest, represented by a bounded continuous function $f \in C_b(Z)$, is simply given by

$$\mathbb{E}[f] = \frac{\int_E f(z) L_D(z) dz}{\int_E L_D(z) dz}, \quad (3.16)$$

which fully characterizes the response of the system in the sense of probability.

3.3.6 Example: random trusses

Consider a truss such as defined in Example 3.3.1, with material behavior characterized by the likelihood function (3.11). In addition, we metrize phase space by means of the Euclidean norm (3.1). We begin by parameterizing the constraint space E . Recall that m is the number of members of the truss, $Z = \mathbb{R}^m \times \mathbb{R}^m$ is the phase space and $n < m$ the number of unconstrained degrees of freedom. Let $l = m - n$ and $A \in \mathbb{R}^{m \times l}$ the matrix whose columns define a basis of $\text{Ker}(B^T)$. Further, we define the matrix of Gauss point volumes as $\mathbb{W} = \text{diag}(w_1, \dots, w_m)$. Suppose non-degeneracy in the sense that σ satisfies the equilibrium condition $B^T \mathbb{W} \sigma = 0$ if and only if there is $v \in \mathbb{R}^l$ such that

$$\sigma = \mathbb{W}^{-1} A v. \quad (3.17)$$

We may thus regard v as a discrete Airy potential and A as a discrete Airy operator. In addition, suppose that ϵ satisfies the compatibility condition $A^T \epsilon = 0$ if and only if there

are displacements u such that $\epsilon = Bu$. The, the constraint set E_0 through the origin, corresponding to $f = 0$ and $g = 0$, admits the representation

$$E_0 = \{(\epsilon, \sigma) \in Z = \mathbb{R}^m \times \mathbb{R}^m : \epsilon = Bu, u \in \mathbb{R}^n; \sigma = \mathbb{W}^{-1}Av, v \in \mathbb{R}^l\}. \quad (3.18)$$

The general constraint set is then the translation

$$E = z_0 + E_0, \quad z_0 = (\epsilon_0, \sigma_0), \quad \epsilon_0 = g, \quad B^T \mathbb{W} \sigma_0 = f. \quad (3.19)$$

We verify that the dimension of Z is $2N = 2m$ and the dimension of E is $N = m = n + l$. We may regard $(u, v) \in \mathbb{R}^n \times \mathbb{R}^l$ as a set of coordinates parameterizing E . In this representation, the posterior likelihood function takes the form

$$L(u, v) = L_D(Bu + \epsilon_0, \mathbb{W}^{-1}Av + \sigma_0), \quad (3.20)$$

where we write $L_D(\epsilon, \sigma)$ for the material likelihood function. The corresponding posterior probability functions then follows by normalization of $L(u, v)$. For instance, if L_D is of the form (3.12), representation (3.20) gives

$$L(u, v) = \exp \left(-\frac{1}{2s^2} \left(\|Bu + \epsilon_0\|^2 + \|\mathbb{W}^{-1}Av + \sigma_0\|^2 - 2(\mathbb{W}^{-1}Av + \sigma_0) \cdot (Bu + \epsilon_0) \right) \right), \quad (3.21)$$

or, using the orthogonality properties of A and B ,

$$L(u, v) = \exp \left(-\frac{1}{2s^2} \left(\|Bu + \epsilon_0\|^2 + \|\mathbb{W}^{-1}Av + \sigma_0\|^2 - 2\mathbb{W}^{-1}Av \cdot \epsilon_0 - 2\sigma_0 \cdot Bu \right) \right). \quad (3.22)$$

From the properties of Gaussian integrals, the expected values $(\bar{u}, \bar{v})$ of the coordinates (u, v) are obtained by minimizing the potential $\Phi = -\log(L)$, with the result,

$$\bar{u} = (B^T \mathbb{W} C B)^{-1} B^T \mathbb{W} (\sigma_0 - C \epsilon_0), \quad (3.23a)$$

$$\bar{v} = (A^T \mathbb{W}^{-1} C^{-1} A)^{-1} A^T (\epsilon_0 - C^{-1} \sigma_0), \quad (3.23b)$$

which are computed by inverting the stiffness and compliance matrices of the truss, $B^T \mathbb{W} C B$ and $A^T \mathbb{W}^{-1} C^{-1} A$, respectively, with $C = \text{diag}(C_1, \dots, C_m)$. The posterior

probability density then follows as

$$L(u, v) = \sqrt{\det(B^T \mathbb{W} C B / 2\pi s^2)} \sqrt{\det(A^T \mathbb{W}^{-1} \mathbb{C}^{-1} A / 2\pi s^2)} \times \quad (3.24)$$

$$\exp \left(-\frac{1}{2s^2} \left((B^T \mathbb{W} C B)(u - \bar{u}) \cdot (u - \bar{u}) + (A^T \mathbb{W}^{-1} \mathbb{C}^{-1} A)(v - \bar{v}) \cdot (v - \bar{v}) \right) \right),$$

which provides a full account of the probability of outcomes. $\square$

3.3.7 Approximation by empirical data

Suppose, contrariwise, that, as in commonly the case in practice, the material likelihood function L_D is known only approximately through a sequence of empirical samples, e. g., drawn from experimental measurements, consisting of point-data sets $P_h = \{y_{h,i}, i = 1, \dots, M_h\} \subset Z$, $h = 0, 1, \dots$, of size M_h , cf. Fig. 3.1b. We additionally associate with each discrete data point a certain likelihood $c_{h,i} \in [0, 1]$ accounting, e. g., for experimental reliability and scatter [83]. Thus, $c_{h,i} = 1$ for a data point $y_{h,i}$ that is sure to be free of experimental error, whereas $c_{h,i} < 1$ if the experimental measurements show scatter or are of questionable provenance. For every material point-set P_h , we consider discrete expectations of the general form

$$\mathbb{E}_h[f] = \sum_{i=1}^{M_h} \left(\int_E p_{h,i}(z) f(z) dz \right), \quad (3.25)$$

where $f \in C_b(Z)$ is a continuous function representing a quantity of interest, and we require

$$p_{h,i}(z) \geq 0, \quad \sum_{i=1}^{M_h} \left(\int_E p_{h,i}(z) dz \right) = 1. \quad (3.26)$$

We envision an experimental campaign resulting in point-data samples P_h of increasing size, $M_h \uparrow \infty$ and increasing fidelity $c_{h,i} \uparrow 1$. The central question is, then, under what conditions on the data, if any, and what choices of probabilities $p_{h,i}(z)$ the approximate expectations $\mathbb{E}_h[f]$ converge to the exact ones $\mathbb{E}[f]$ in the limit as $h \rightarrow \infty$.

We specifically consider probabilities $p_h \equiv \{p_{h,i}(z), i = 1, \dots, M_h\}$ that minimize the regularized Kullback-Leibler (KL) divergence [96, 95]

$$G_\beta(p_h) = \sum_{i=1}^{M_h} \left(\int_E \left(\beta \|y_{h,i} - z\|^2 + \log \frac{p_{h,i}(z)}{c_{h,i}} - N \log \frac{\beta}{\beta_0} \right) p_{h,i}(z) dz \right), \quad (3.27)$$

for some reference $\beta_0 > 0$. We note that the form (3.25) assumed of the trial expectations ensures that the implied trial likelihoods are absolutely continuous with respect to the assumed material point-data set and deterministic loading, a requirement of the KL divergence. Evidently, the divergence function (3.27) penalizes distance to the constraint set E , i. e., it assigns an increasing discrepancy to material data points $y_{h,i}$ that are far from E , while simultaneously attempting to match the *a priori* likelihood $c_{h,i}$ of the point. The additional term $N \log(\beta/\beta_0)$ in (3.27) is introduced in order to ensure that the KL functions (G_β) are uniformly bounded with respect to β and it only plays a role in the variational characterization of β , cf. Section 3.3.8.

The optimal regularized discrete probability p_h then follows by minimization of G_β , with the explicit result

$$p_{h,i}(z) = \frac{c_{h,i} e^{-\beta \|y_{h,i} - z\|^2}}{\sum_{i=1}^{M_h} \int_E c_{h,i} e^{-\beta \|y_{h,i} - z\|^2} dz}, \quad z \in E. \quad (3.28)$$

We see from this expression that material data points $y_{h,i}$ that are far from the constraint set E contribute less than points are close, as expected. The parameter β sets the scale of the distance as $1/\sqrt{\beta}$.

3.3.8 Annealing schedule

We note that, for $\beta \rightarrow +\infty$, the posterior probabilities (3.28) degenerate in general. This degeneracy simply reflects the fact that the point sets P_h are unlikely to have a consequential intersection with the admissible set E , i. e., the likelihood that a point $y_{h,i}$ be admissible with respect to a given deterministic loading is zero. As $\beta \rightarrow 0$, the posterior probabilities (3.28) also degenerate. There must therefore be an optimal value of β where the probabilities are least degenerate. Conveniently, this optimal value β_h of β again follows by minimization of the KL divergence (3.27). Thus, evaluating $G_\beta(p_h)$ at the minimizer (3.28), we obtain

$$\min G_\beta = -\log \left(\sum_{i=1}^{M_h} \int_E c_{h,i} e^{-\beta \|y_{h,i} - z\|^2} dz \right) - N \log \frac{\beta}{\beta_0}. \quad (3.29)$$

Minimizing with respect to β gives the optimality condition

$$\frac{1}{\beta_h} = \frac{1}{N} \frac{\sum_{i=1}^{M_h} \int_E c_{h,i} \|y_{h,i} - z\|^2 e^{-\beta_h \|y_{h,i} - z\|^2} dz}{\sum_{i=1}^{M_h} \int_E c_{h,i} e^{-\beta_h \|y_{h,i} - z\|^2} dz}, \quad (3.30)$$

which defines β_h implicitly. Finally, inserting (3.30) into (3.25) we obtain

$$\mathbb{E}_h[f] = \frac{\sum_{i=1}^{M_h} \left(\int_E c_{h,i} e^{-\beta_h \|y_{h,i} - z\|^2} f(z) dz \right)}{\sum_{i=1}^{M_h} \left(\int_E c_{h,i} e^{-\beta_h \|y_{h,i} - z\|^2} dz \right)}, \quad (3.31)$$

which supplies an expression for the approximate expectations that is explicit up to quadratures.

3.3.9 Factorization and localization

By exchanging the integrals over E and the sums over data in (3.31), the approximate expectations have the equivalent form

$$\mathbb{E}_h[f] = \frac{\int_E \sum_{i=1}^{M_h} f(z) e^{-\beta_h \|y_{h,i} - z\|^2} c_{h,i} dz}{\int_E \sum_{i=1}^{M_h} e^{-\beta_h \|y_{h,i} - z\|^2} c_{h,i} dz}. \quad (3.32)$$

The great advantage of this alternative form is that it factorizes with respect to local data, which eliminates the combinatorial complexity of the first form. Thus, suppose that

$$D_h = \prod_{e=1}^m D_{e,h}, \quad (3.33)$$

where $D_{e,h} = \{y_{e,h,i}, i = 1, \dots, M_{e,h}\}$ are local data files. Then,

$$\int_E \sum_{i=1}^{M_h} f(z) e^{-\beta_h \|y_{h,i} - z\|^2} c_{h,i} dz = \int_E f(z) L_h(z) dz, \quad (3.34)$$

where

$$L_h(z) = \sum_{i=1}^{M_h} e^{-\beta_h \|y_{h,i} - z\|^2} c_{h,i} \quad (3.35)$$

is an inferred likelihood density. For local data, the sum over data factorizes, with the result

$$L_h(z) = \prod_{e=1}^m L_{e,h}(z_e), \quad (3.36)$$

where

$$L_{e,h}(z_e) = \sum_{i=1}^{M_{e,h}} e^{-\beta_h \|y_{e,h,i} - z_e\|^2} c_{e,h,i} \quad (3.37)$$

are local likelihood density functions. We note that the computational complexity of the expectation is $\sum_{e=1}^m M_{e,h}$, instead of $\prod_{e=1}^m M_{e,h}$, as before, which renders the evaluation of approximate expectations computationally feasible.

We further note that, for local data, we can generalize (3.37) to

$$L_{e,h}(z_e) = \sum_{i=1}^{M_{e,h}} e^{-\beta_{e,h} \|y_{e,h,i} - z_e\|^2} c_{e,h,i}, \quad (3.38)$$

where we localize temperature. We can then estimate $1/\beta_{e,h}$ as the average minimum distance squared between points in $D_{e,h}$, i. e.,

$$\frac{1}{\beta_{e,h}} \sim \frac{1}{M_{e,h}} \sum_{i=1}^{M_{e,h}} \min_{j \neq i} \|y_{e,h,i} - y_{e,h,j}\|^2. \quad (3.39)$$

This local approximation of Eq. (3.30) has the advantage that $\beta_{e,h}$ can be explicitly pre-computed for each local data set and is therefore used in the following numerical examples.

Finally, the integrals over E can be conveniently computed by stochastic quadrature. These and other matters of implementation are discussed next.

3.4 Numerical implementation

3.4.1 Problem statement

In evaluating (3.32), two main operations need to be performed: the integration over the constraint set E and the summation over the global material data set for the evaluation of the material likelihood. The main computational challenges pertaining to these operations are:

1. The evaluation of the local material likelihood involves a summation over local data sets for each material point that can be computationally demanding. However, most data points have negligible influence on the total sum, which suggests the use of efficient, carefully designed *a propos* data structures.
2. The Monte Carlo integration over the constraint set E requires the generation of random points on E , which requires the repeated solution of global systems of equations if the sampling is done by projection. It is therefore advantageous to identify sampling strategies that reduce the number of projections.
3. The thermalized likelihood function implied by (3.31) is exceedingly multi-modal, featuring peaks at every data point (cf. [73] for a discussion in the context of *min-dist* DD). This intricate likelihood landscape arises even if the underlying likelihood function is smooth and compounds the evaluation of the integrals over E in (3.31).

These challenges are addressed next in turn.

3.4.2 Population annealing

The complexity of the thermalized likelihood landscape renders the direct application of standard Markov chain importance sampling strategies, such as the classical Metropolis–Hastings algorithm [117], difficult. Therefore, a more tailored sampling strategy is required, as discussed next.

3.4.2.1 Population annealing

Population Annealing (PA) [68, 111, 157] is a sequential Monte Carlo method designed for complex likelihood functions and scalable parallel computations. A set of states, the population, is maintained in order to explore different regions of phase space that are possibly separated by energy barriers resulting from local minima. Each population member performs a relatively short Markov chain importance sampling. In addition, the population is resampled by a combination of pruning and enriching, in analogy to *Genetic Algorithms*. The process is driven by a *Simulated Annealing* schedule, which slowly increases the inverse local temperature β_e . For every new temperature, the population is re-equilibrated by resampling and a specified number of Markov chain moves. In this manner, at low β_e the population can escape readily from local minima and perform a more exhaustive exploration. At high β_e , the population is increasingly driven towards the constraint set E .

3.4.2.2 Initialization and resampling step

The principal idea of the resampling step is to remove and duplicate the population samples according to their energy while increasing β_e in order to accelerate the convergence to low-energy states. Resampling is done in such a way that the population size fluctuates around a target population size N_P^* and equilibrium of the distribution is maintained for every new temperature.

Initially, at $\beta_e = 0$, a population $\mathcal{P} = \{z_p \in E : p = 1 \dots N_P\}$ of N_P states is generated either randomly by the closest-point projection of random global data points onto E or using a *min-dist* DD solver [82]. At initialization, the population size N_P may be chosen equal to the target size N_P^* , or set to a different value, in which case the size gets adapted to the target size during resampling. The energy of the population is initially set to infinity. For $\beta_e > 0$, the energy is defined as

$$e_p = -\frac{1}{\beta_e} \sum_{e=1}^m \log \left(\frac{1}{M_e} \sum_{i=1}^{M_e} e^{-\beta_e \|y_{e,i} - z_{e,p}\|^2} \right). \quad (3.40)$$

where we assume local data and $c_i = 1$ for all data points. The efficient evaluation of (3.40) is discussed in Section 3.4.2.4.

At each iteration, or *quench*, β_e is increased according to

$$\beta_e = \beta_e + \Delta\beta_e \quad (3.41)$$

for a specified number of quenches N_Q . For simplicity, in calculations we choose a constant increment

$$\Delta\beta_e = \beta_{e,f}/N_Q \quad (3.42)$$

for a pre-computed target $\beta_{e,f}$. The simple choice of $\beta_{e,f}$ given by the inverse average nearest neighbor distance (3.39) is used in the numerical examples presented in Sections 3.5.1 to 3.5.3.

For resampling, the relative weights

$$\hat{\tau}_p = N_P^* \frac{e^{-\Delta\beta_e e_p}}{\sum_{i=1}^{N_P} e^{-\Delta\beta_e e_i}} \quad (3.43)$$

are computed for each population member. The required number of copies of population member p is then conveniently estimated as [156, 7]

$$l_p = \begin{cases} \lfloor \hat{\tau}_p \rfloor, & \text{if } u > \hat{\tau}_p - \lfloor \hat{\tau}_p \rfloor, \\ \lfloor \hat{\tau}_p \rfloor + 1, & \text{otherwise,} \end{cases} \quad (3.44)$$

where u is a random number drawn uniformly from $[0, 1)$ and the floor function $\lfloor x \rfloor$ is the largest integer less or equal x . If $l_p = 0$, member p is removed from the population. The resampling step incurs a modest computational overhead that is offset by the overall performance gain resulting from reconfiguring the population to the relevant regions of the phase space.

3.4.2.3 Markov chain importance sampling step

Resampling introduces a bias since population members may be duplicated. In order to eliminate this bias, a Markov chain importance sampling step is performed. In this step, every population member executes a random walk, thereby exploring new states and decorrelating the population. The random walk consists of repeated random moves in the constraint set E . We recall that the constraint set E is a linear subspace of

dimension $N = md$ (number of material points times dimension of stress/strain) in an $2N$ -dimensional global phase space Z . We can conveniently exploit this affine structure and represent E in terms of an N -dimensional basis.

We compute the requisite basis by means of a principal component analysis (PCA). To this end, we generate $K > N$ random global data points $z_i \in Z$ and project them onto E . For consistency with the norm (3.1), we further weight the strains and stresses by $\sqrt{w_e} \mathbb{C}_e^{1/2}$ and $\sqrt{w_e} \mathbb{C}_e^{-1/2}$, respectively. The weighted states z_i^w (here and subsequently the superscript w indicates weighting) are collected in a matrix $T \in \mathbb{R}^{K \times 2N}$ in order to compute its covariance matrix $C_E \in \mathbb{R}^{2N \times 2N}$. Computing the eigenvectors of C_E associated to the N largest eigenvalues then gives the basis vectors of the constraint set as a matrix $A_E \in \mathbb{R}^{2N \times N}$.

To effect a move, N random numbers are drawn from a standard normal distribution and multiplied by a step size s_p , transformed back by A_E^T and added to the weighted state z_p^w

$$z_{p,trial}^w = z_p^w + s_p g A_E^T, \quad g \sim \mathcal{N}(0, I_N). \quad (3.45)$$

The proposed move to the trial state $z_{p,trial}$ (obtained by reverse weighting of $z_{p,trial}^w$) is then either accepted or rejected according to the Metropolis-Hastings acceptance probability

$$P_{\beta_e} = \min(1, \exp(-\beta_e(e_{p,trial} - e_p))), \quad (3.46)$$

where $e_{p,trial}$ is the energy of $z_{p,trial}$. This acceptance criterion allows moves that increase the energy slightly, in particular at the initial stages of annealing characterized by low β_e . In addition, the step size s_p influences the acceptance rate and, therefore, the efficiency of the algorithm. In calculations, we chose the simple adaption scheme

$$s_p = s_p + (r_p - r^*)s_p, \quad (3.47)$$

with a target acceptance rate $r^* \in [0, 1]$. The acceptance rate r_p is measured as the ratio of accepted moves to the total number of trials N_T . The adaption scheme has the effect of reducing the step size if the acceptance rate is smaller than r^* , which signals an overly-aggressive exploration strategy, and increasing it if the acceptance rate is too small, which contrariwise signals an overly-timid exploration strategy.

3.4.2.4 Energy computation

The complexity of the evaluation of the energy (3.40) is $\sum_e M_e$. However, for sufficiently large β_e a vast majority of the data points contribute negligibly to (3.40), which suggests restricting the sum with the aid of supporting data structures. In [34], a number of data structures for local data sets were investigated in the framework of min-dist DD solvers and the hierarchical k -means tree [47] was found to be particularly well-suited for executing approximate nearest neighbor searches.

In this work, we likewise use the hierarchical k -means implementation of the FLANN library [120]. The approximate nearest neighbor search provided by the library is performed using a priority queue of maximum length N_{checks} in order to reduce backtracking operations. We employ a radius search technique, which returns all data points $y_{e,i} \in D_e$ within a specified radius

$$r_{\text{TOL}}^2 = -\frac{1}{\beta_e} \log(\text{TOL}). \quad (3.48)$$

around the query point $z_{p,e}$. This technique has the advantage over a k -nearest-neighbor (k NN) search that the number of nearest neighbors k need not be specified. Instead, the sum is restricted to points whose contribution exceeds the tolerance TOL .

Our numerical tests reveal that using a radius search increases performance markedly and that, remarkably, the performance gains are insensitive to the choice of tolerance. Thus, in tests values of $\text{TOL} = 10^{-8}, 10^{-16}, 10^{-32}$ delivered nearly identical performance. In view of this tradeoff, in all calculations presented subsequently we use $\text{TOL} = 10^{-16}$. By contrast, the performance of the algorithm is strongly dependent on the degree of backtracking as set by N_{checks} , cf. Section 3.5.3.

It bears emphasis that only one material data structure needs to be set up for each specific material in the structure, since all members made of the same material access the same local data set. Likewise, only one inverse local temperature $\beta_{e,f}$ and one quenching scheme is required for each specific material. In particular, if all structural members are made of the same material, only one material data structure and quenching scheme need to be maintained and run.

3.4.2.5 Summary of the algorithm

The described algorithm is summarized in Algorithm 4. It returns the final population $\mathcal{P}$ of size N_P . The expectation of a quantity of interest f is evaluated as

$$\mathbb{E}[f] = \frac{1}{N_P} \sum_{p=1}^{N_P} f(z_p), \quad z_p \in \mathcal{P}. \quad (3.49)$$

One advantage of the algorithm is that does not require extensive fine-tuning of the input variables. Thus, the target population size N_P^* , the number of quenches N_Q and the number of Markov chain moves N_T are parameters that are chosen as large as permitted by computational resources. In all calculations, the target acceptance rate is set to $r^* = 0.25$. The initial step size s_0 influences only the first few iterations as it is adapted during the quenching schedule. The step size tends to decrease with increasing dimensionality [57]. Lastly, we note that the population member perform their moves independently of each other, which renders the algorithm *embarrassingly parallel*.

Algorithm 4 Population annealing algorithm

Require: Constraint set $E \subset Z$, basis vectors A_E .
Require: Local data set(s) D_e
Require: $N_Q, N_T, N_P^*, s_0, r^*, \beta_{e,f}$
Initialize: $\mathcal{P} \leftarrow \{z_p \in E\}$ randomly or heuristically
Initialize: step size $s_p \leftarrow s_0$
Initialize: energy $e_p \leftarrow \infty$
Initialize: $\beta_e \leftarrow 0$
Set: $\Delta\beta_e \leftarrow \beta_{e,f}/N_Q$
for $q = 0, \dots, N_Q - 1$ **do**
 Set: $\beta_e \leftarrow \beta_e + \Delta\beta_e$
 Set: $\mathcal{P} \leftarrow \text{Resampling}(\mathcal{P}, \Delta\beta_e)$
 for each $z_p \in \mathcal{P}$ **do**
 for $t = 0, \dots, N_T - 1$ **do**
 Set: $z_{p,\text{trial}} \leftarrow \text{RandomMove}(z_p, s_p, A_E)$
 Set: $e_{p,\text{trial}} \leftarrow \text{Energy}(D_e, \beta_e, z_{p,\text{trial}})$
 Set: $u \leftarrow \text{UniformRand}(0, 1)$
 Set: $w \leftarrow \exp(-\beta_e(e_{p,\text{trial}} - e_p))$
 if $u \leq w$ **then**
 Set: $z_p \leftarrow z_{p,\text{trial}}$
 Set: $e_p \leftarrow e_{p,\text{trial}}$
 end if
 end for
 end for
 Set: $r_p \leftarrow \text{AcceptanceRate}$
 Set: $s_p \leftarrow s_p + (r_p - r^*)s_p$
end for
return $\mathcal{P}$

3.5 Examples of application

We conclude with selected examples of application that illustrate simply the range, scope and convergence properties of the approach.

3.5.1 Three-bar truss with sliding-Gaussian material data set

We begin with a simple verification example concerned with the three-bar truss shown in Figure 3.2a loaded by a force P and undergoing a corresponding displacement Δ .

In calculations, we set $P = 100$, $L = 1$ and all cross-sectional areas $A = 1$ (here and subsequently units are omitted). The outcome quantity of interest is chosen to be the displacement Δ .

In order to have an analytical reference solution, we assume that the material data set obeys a sliding-Gaussian distribution, in which case the exact solution of the inference problem is given in Section 3.3.6. The local material data set is generated according to Section 3.3.4 with $\mathbb{C}_e = 10000$ and standard deviation $s = 5 \times 10^{-4}$. A typical data set is shown in Fig. 3.2b. We emphasize that such material data sets are the only input to the calculations and, once generated, the underlying distribution whence they are sampled is jettisoned altogether and presumed unknown.

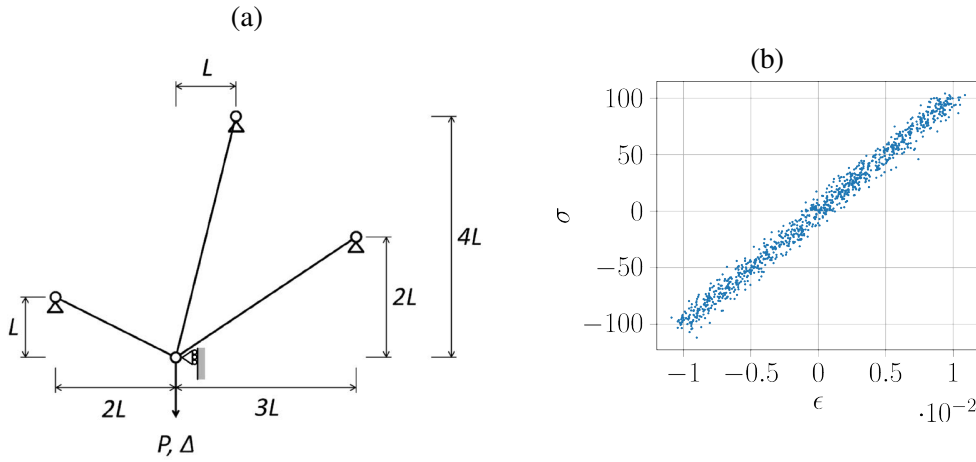


Figure 3.2: a) Truss with prescribed applied force P undergoing a displacement Δ .
b) Local material data set with 1000 data points sampled from a sliding-Gaussian likelihood function.

Calculations are carried out using Algorithm 4 with number of quenches $N_Q = 100$, number of trials $N_T = 20$ and initial step size $s_0 = 1.0$. Nearest-neighbor searches with radius $\text{TOL} = 10^{-16}$ are performed exactly, i. e., with an unlimited number of backtracking checks N_{checks} . As a measure of the error we choose the Kolmogorov–Smirnov statistic e_{KS} , which measures the maximum difference between the exact and computed cumulative distributions.

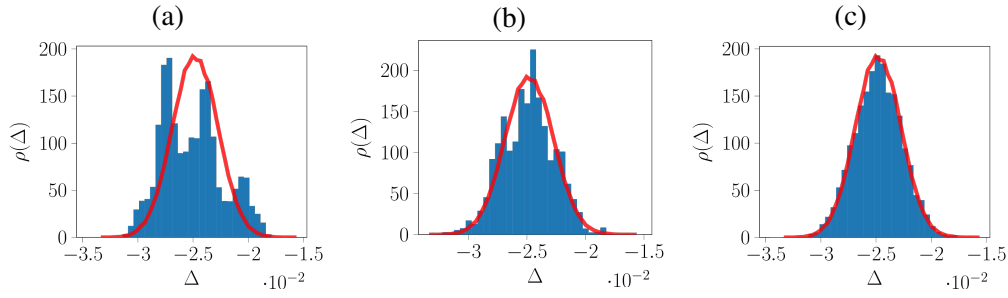


Figure 3.3: Three-bar truss with sliding-Gaussian material data. Computed histograms vs. exact distribution (red) of displacement Δ for material data-sets of sizes: a) $M_e = 10^3$; b) $M_e = 10^4$; c) $M_e = 10^5$.

Fig. 3.3 collects three computed histograms of displacement Δ for three different material data-set sizes $\{10^3, 10^4, 10^5\}$. The exact Gaussian distribution is shown in red for comparison. A clear trend towards convergence of the computed histogram to the exact distribution is apparent from the figure. It bears emphasis that, as expected, both the exact distribution and the computed histograms of displacement are strictly *unimodal*. Unimodality is indicative of a lack of complexity of the behavior of a system. In particular, the maximum-likelihood state of the system—equivalently, the minimum free-energy solution—is uniquely defined [83] and the outcomes of the system cluster around the maximum-likelihood state. For the given problem, the maximum-likelihood solution is computed to $\Delta_{\text{ML}} = -2.48 \cdot 10^{-2}$.

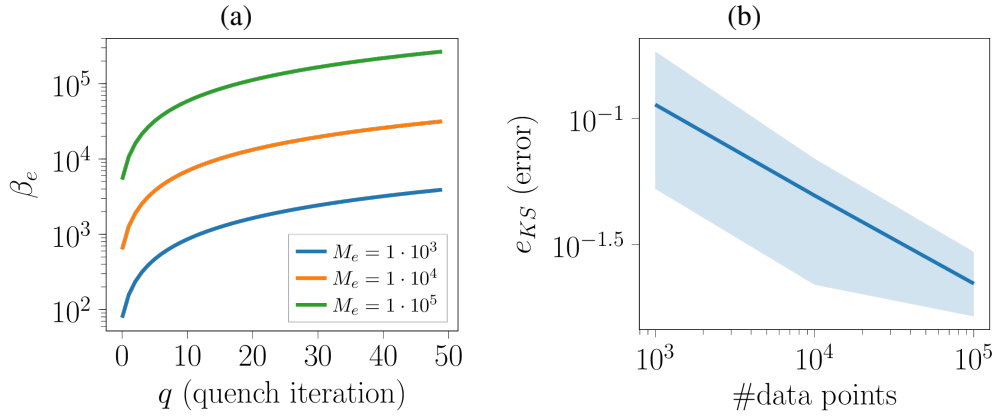


Figure 3.4: Three-bar truss with sliding-Gaussian material data. Convergence of temperatures and computed histograms of displacement Δ . a) Evolution of β_e during annealing process. b) Kolmogorov–Smirnov (KS) error vs. material data-set size, $N_P = 10^4$.

Fig. 3.4a shows the evolution of β_e during the annealing process. As may be seen from the figure, β_e increases monotonically with the number of quenches and tends towards a final value. This final value itself increases with the material data-set size as the average minimum distance between material data points decreases. Fig. 3.4b shows the dependence of the Kolmogorov–Smirnov error e_{KS} on the number of material data points, which again exhibits a clear trend towards convergence. An analysis of the convergence plot reveals that, asymptotically for large material data sets, $e_{KS} \sim M_e^{-\alpha}$, where $\alpha \sim 0.36$ is the rate of convergence. These results are consistent with quantitative error estimates derived from analysis [22]. For completeness, Fig. 3.5a shows the dependence of e_{KS} on the population size N_P . Fig. 3.5b shows the dependence on the number of quenches N_Q for different number of trials N_T using a population size of $N_P = 5000$.

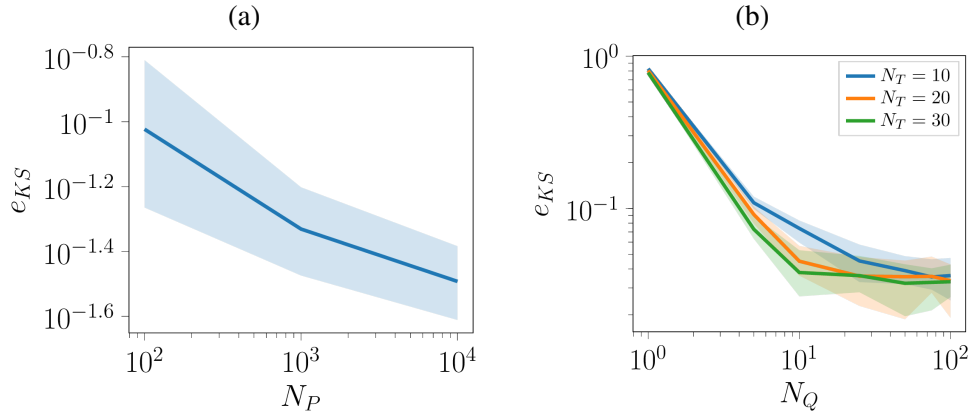


Figure 3.5: Three-bar truss with sliding-Gaussian material data, $M_e = 10^5$. a) KS error vs. population size, $N_T = 20$ and $N_Q = 100$. b) KS error vs. number of quenches N_Q for different number of trials N_T using $N_P = 5000$. The lines show the mean of 20 repeated computations, the shaded area indicate the maximum deviation from mean.

3.5.2 Three-bar truss with Weibull tensile-strength distribution

Next, we present a simple example that is designed to illustrate the ability of the method to deal effectively with complex data and behavior. The example concerns the same structure as in Section 3.5.1, Fig. 3.6a, but the material is now assumed to be brittle with random tensile strength.

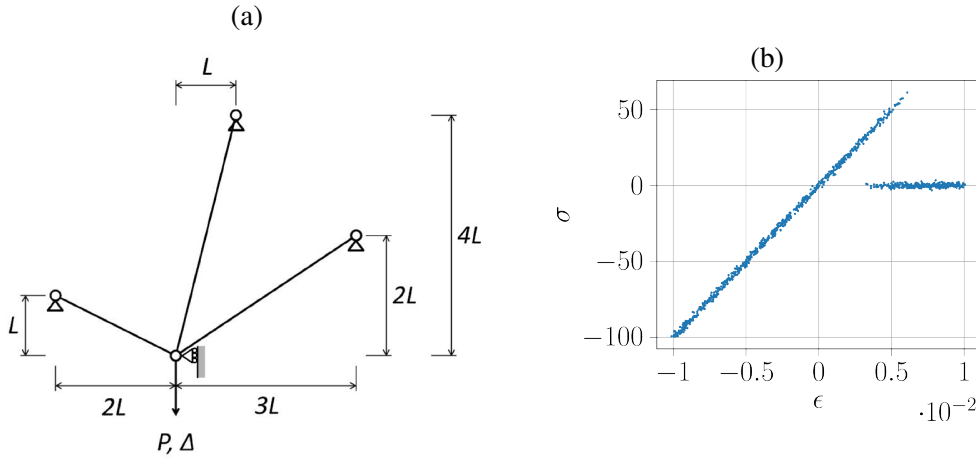


Figure 3.6: a) Truss with prescribed displacement Δ and reaction force P . b) Local material data set with 1000 data points sampled from a bimodal likelihood function with failed and unfailed branches selected according to Weibull strength statistics, plus superimposed Gaussian noise.

The control parameter is now the displacement Δ , which is increased monotonically. The outcome of interest is the corresponding reaction force P . By virtue of the monotonicity of Δ , the strains in all bars also increase monotonically. The behavior of the bars in compression is linear elastic. By contrast, in bars in tension the stress increases linearly up to the tensile strength $\sigma_t > 0$ of the material and subsequently drops to zero at larger strains. In addition, the tensile strength of the material is assumed to be random and to obey Weibull statistics, with distribution

$$W(\sigma_t) = 1 - e^{-(\sigma_t/\sigma_0)^p}, \quad (3.50)$$

where σ_0 and p are material-specific parameters. Under these conditions, the likelihood of observing a local tensile state (ϵ_e, σ_e) is *bimodal* and consists of: a branch $\sigma_e = \mathbb{C}_e \epsilon_e$ with likelihood $1 - W(\mathbb{C}_e \epsilon_e)$, corresponding to an unfailed bar; and a branch $\sigma_e = 0$ with likelihood $W(\mathbb{C}_e \epsilon_e)$, corresponding to a failed bar.

A typical material data set sampled from this likelihood function, with superimposed Gaussian noise of standard deviation $s = 10^{-4}$, is shown in Fig. 3.6b. The data set is intended to mimic empirical data sets such as might be generated experimentally, including experimental scatter. Again, we emphasize that such material data sets are the

only input to the calculations and, once generated, the underlying distribution whence they are sampled is jettisoned altogether and presumed unknown.

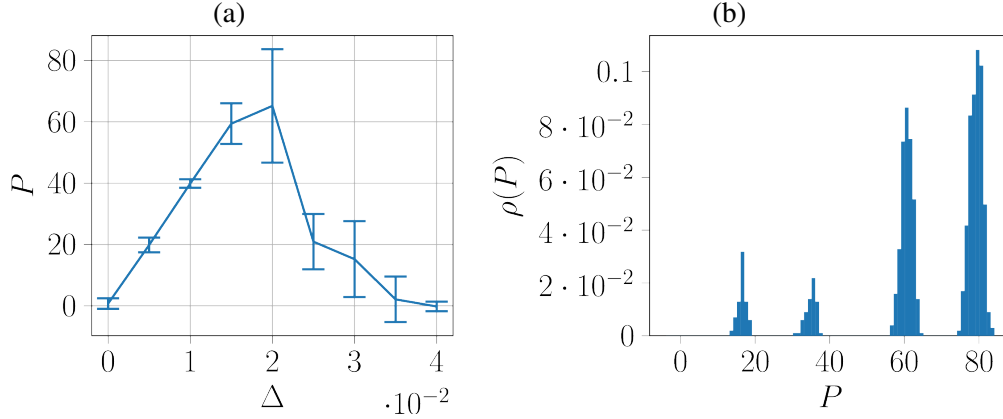


Figure 3.7: Three-bar truss with Weibull tensile strength data. a) Computed mean and standard deviation of the reaction force P for prescribed displacement Δ . b) Computed distribution of P at $\Delta = 2 \times 10^{-2}$.

For the simulation parameters, we choose $N_Q = 50$, $N_T = 10$ and $N_P = 2 \times 10^4$ with data-set size of $M_e = 10^3$. Fig. 3.7a shows the computed mean and standard deviation of the reaction force P over a range of prescribed displacements Δ . Interestingly, the figure reveals three regimes: i) an initial linear elastic response with low uncertainty, as measured by the standard deviation; ii) an intermediate regime combining failed and unfailed bars characterized by a decreasing reaction and high uncertainty; and iii) full failure, or *collapse*, with low uncertainty.

Fig. 3.7b in turn shows the computed distribution of P at $\Delta = 2 \times 10^{-2}$. Remarkably, the distribution is strongly *multimodal*, a telltale hallmark of complexity. The various modes correspond to several different possible combinations of failed and unfailed bars and their attendant likelihoods. For the data used in calculations, the right bar is computed to remain unfailed with probability 1 or, equivalently, the probability of the right bar failing is computed to be 0. By contrast, the remaining two bars can be failed or unfailed with varying likelihoods: no failure (right peak), double failure (left peak), single failures (two middle peaks). From a global perspective, the various modes in the posterior likelihood function of P correspond to a *clustering*, or *classification*, of global material data points close to the constraint set. The ability of the method to classify

outcomes automatically and effectively bears remark.

3.5.3 Lightweight space structure

By way of demonstration, we consider the lightweight structure shown in Fig. 3.8 designed to be part of a modular space telescope [101]. We specifically aim to assess the feasibility of the proposed approach for higher-dimensional problems (constraint set dimension $N = 39$) and its computational performance in terms of resource requirements and execution times.

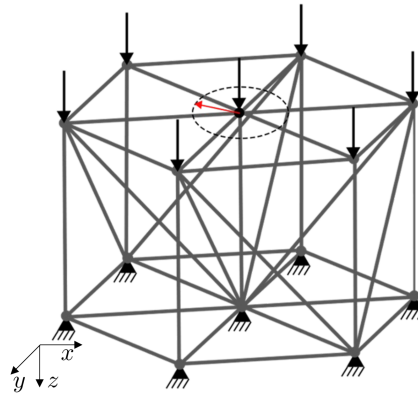


Figure 3.8: Lightweight space structure [101]. The red arrow indicates the quantity of interest (joint eccentricity u_e) to be inferred from the analysis.

In calculations, we set the cross-sectional area of the bars to $A = 0.1$ and the length of the non-diagonal bars to 1. The structure is constrained throughout its base and is loaded by longitudinal joint forces $P = 5$ as shown in Fig. 3.8. As a quantity of interest, we choose the top central joint eccentricity $u_e = \sqrt{u_x^2 + u_y^2}$ whereby u_x and u_y are the in-plane displacements of the joint. We specifically consider a material data set with $M_e = 10^4$ data points sampled from a sliding Gaussian as in Section 3.5.1, cf. Fig. 3.2. Conveniently, this assumption affords a closed-form analytical solution for the distribution of u_e that can be used by way of reference, cf. Section 3.3.6.

For algorithm parameters, we choose $N_P = 2 \times 10^3$, $N_T = 10$, $N_Q = 100$ and $s_0 = 0.01$. For faster convergence to low-energy states, we initialize the population by means of 1000 random min-dist solutions [82] for noise-free data. As expected, the posterior distribution derived from the initial population is far from the exact distribution,

Fig. 3.9a. By contrast, a close agreement is achieved following the application of the population annealing algorithm, Fig. 3.9b, which is remarkable in view of the relatively small population and material-data sizes used in the calculations. For comparison, we note that the maximum-likelihood solution [15] is computed to be $u_{e,ML} = 0$. This comparison underscores that, while the maximum-likelihood solution can deal effectively with observational noise, it only identifies the most likely state of the system. By contrast, for intrinsically random systems it is often desirable to infer the entire posterior distribution of states, as in the present framework.

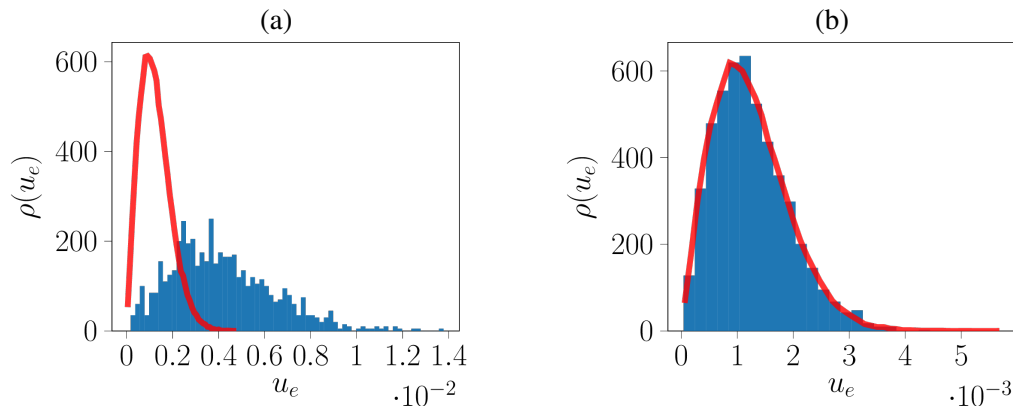


Figure 3.9: Lightweight space structure [101]. Computed posterior distributions of the joint eccentricity u_e and exact reference distribution (red curves). a) Posterior distribution obtained from an initial population of 1000 random min-dist solutions. b) Posterior distribution following application of the population annealing algorithm.

Finally, we assess the impact of the approximate nearest neighbor search on accuracy and CPU times for the population size $N_P = 2 \times 10^3$. The absolute CPU times are, of course, implementation and machine dependent. In this work, the computations were performed on a 12-Core AMD Ryzen 9 3900X machine with Python bindings [70] to a C++ implementation using Standard Template Library Parallel Algorithms.

The CPU times for a direct evaluation of Eq. (3.40) without any data structuring or acceleration is ~ 662 seconds. The radius search with full backtracking terminates in ~ 92 seconds, or a seven-fold acceleration. Fig. 3.10 shows the effect of the degree of backtracking measured in terms of N_{checks} . We see from the figure that, whereas the computation time increases rapidly with the degree of backtracking, accuracy is not

substantially improved above $N_{\text{checks}} = 2^3$, at which limit the CPU time is reduced to 8 seconds. This tradeoff further suggests that the performance of the algorithm can benefit greatly from data structuring techniques such as approximate nearest-neighbor searches [34], especially for large systems and material data sets. Furthermore, we recall that the algorithm is embarrassingly parallel, which opens the way for efficient parallel implementations, e. g., on GPU machines [7].

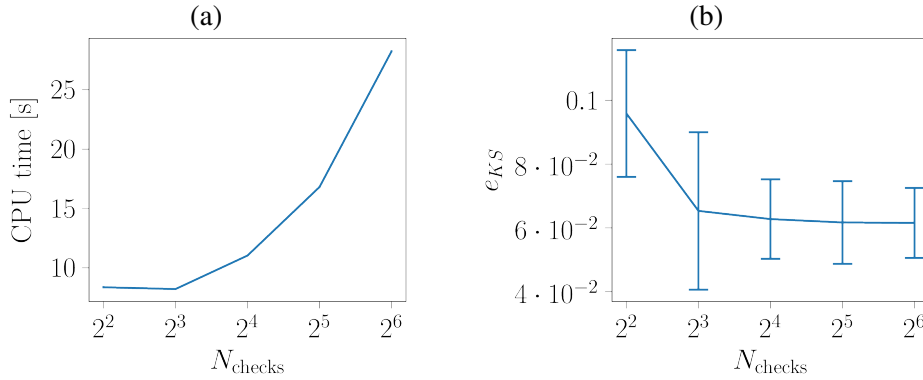


Figure 3.10: Lightweight space structure [101]. Influence of the number of checks for backtracking approximate in the nearest-neighbor search. a) CPU times in seconds on a 12-Core AMD Ryzen 9 3900X machine. b) Error e_{KS} of node eccentricity u_e with error bars derived from a sample of 25 calculations.

3.6 Acknowledgements

All authors gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG) and French Agence Nationale de la Recherche (ANR) through the project "Direct Data-Driven Computational Mechanics for Anelastic Material Behaviours" (ANR-19-CE46-0012-01, RE 1057/47-1, project number 431386925) within the French-German Collaboration for Joint Projects in Natural, Life and Engineering (NLE) Sciences. SR gratefully acknowledges the funding of the DFG projects "Model order reduction in space and parameter dimension - towards damage-based modeling of polymorphic uncertainty in the context of robustness and reliability" (SPP 1886, project number 312911604) and "Coupling of intrusive and non-intrusive locally decomposed model order reduction techniques for rapid simulations of road systems" (TP B05/TRR 339, project number 453596084). MO is grateful for support from the

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) *via* project 211504053 - SFB 1060; project 441211072 - SPP 2256; and project 390685813 - GZ 2047/1 - HCM.

4 | Article 3: A Sparse Basis for Equilibrium Stress Fields with Application for Direct Data-Driven Mechanics

This article was published as: Prume, E., Ji, C., Reese, S., Ortiz, M. [2026], ‘A Sparse Basis for Equilibrium Stress Fields with Application for Direct Data-Driven Mechanics’, *Computer Methods in Applied Mechanics and Engineering* **453**, 118803.

Disclosure of the individual authors’ contributions to the article:

E. Prume developed the concept and algorithms for computing the sparse null basis and the novel data-driven solver algorithms. E. Prume drafted the manuscript, implemented the algorithms and computed the numerical examples. C. Ji contributed by testing the solvers on different numerical examples and giving important feedback. S. Reese and M. Ortiz provided scientific guidance, reviewed the draft and contributed valuable comments and edits. All authors approved the publication of the final version of the manuscript.

4.1 Abstract

We present a new class of solvers for direct data-driven mechanical problems based on a sparse basis representation of equilibrium stress fields. Our first contribution is an efficient algorithm for computing the required sparse null-space basis on tetrahedral meshes. Only a single QR decomposition is needed to compute a small remaining set of dense basis vectors associated with boundary conditions and topological holes which can be handled efficiently via a partitioned Cholesky factorization. Building on this, we demonstrate how standard iterative solvers—such as the Newton–Raphson method—can be applied to direct data-driven formulations. The proposed approach is particularly valuable for challenging problems with complex data distributions requiring systematic exploration of the space of equilibrium stress fields. To this end, we introduce an algorithm that constructs a hierarchical solution set through an eigenvalue decomposition in the joint space of equilibrium stress and compatible strain fields. We demonstrate the proposed methodology with a numerical example involving brittle fracture with probabilistic tensile strength. The resulting family of failure patterns offers valuable insights for uncertainty quantification and design decision-making.

4.2 Introduction

In this paper, we develop a new class of solvers for direct data-driven problems [82, 83] by explicitly constructing a basis for equilibrium stress fields on tetrahedral meshes. The basic principle of the data-driven method is to satisfy the equations of kinematic compatibility and equilibrium exactly while minimizing the modeling uncertainties associated with a phenomenological material model. This is achieved by incorporating the data set of stress-strain measurements directly into the formulation of the boundary value problem. Suitable solutions are then defined as the compatible strain fields and equilibrium stress fields that minimize a distance to the data set.

Compatibility can be explicitly enforced by parameterizing the strain fields ε as

$$\varepsilon = \mathbf{B}\mathbf{u} + \varepsilon_0 \quad (4.1)$$

where $\mathbf{B}$ is the discrete strain operator, $\mathbf{u}$ is the displacement field, and ε_0 is an offset

accounting for Dirichlet boundary conditions.

In contrast, the equilibrium condition for a stress field σ is given by the linear constraint

$$(B^T V)\sigma = E\sigma = p. \quad (4.2)$$

Here, E is the discrete divergence operator, V is a diagonal matrix of material point volumes, and p is the external load vector.

Thus, to derive a basis Z for all equilibrium stress fields we need to compute a basis for the null space, or kernel, of E satisfying

$$EZ = 0. \quad (4.3)$$

Every column of the null basis Z is called a null vector and can be interpreted as a linear independent stress field which is in equilibrium without external forces; it is therefore also called a self-stress field [129]. By taking linear combinations of the null vectors, all possible equilibrium stress fields can be parameterized as

$$\sigma = Zv + \sigma_0 \quad (4.4)$$

where v represents the degrees of freedom (sometimes referred as redundant forces) and σ_0 is a particular solution satisfying

$$E\sigma_0 = p. \quad (4.5)$$

In principle, computing the null basis Z can be simply done by a QR decomposition or singular value decomposition of E . However, despite the usual sparsity of E , this approach leads to a dense null basis which quickly becomes computationally inefficient for large 3D problems. This difficulty motivated the development of greedy algorithms in the 1980s [64, 52, 133, 78] to compute sparse null bases as approximations for the computationally NP-hard problem of finding the sparsest possible null basis [20]. In [56], the relationship to the numerically related problem of matrix sparsification was established. More recently, [79] has proposed an graph-based algorithm for tetrahedral meshes using numerical factorizations on substructures of the mesh to achieve sparsity of the null basis.

In the first part of this paper we develop an algorithm that efficiently computes a sparse null basis for tetrahedral elements with a very small fraction of dense columns. In particular, we provide an exact count for different types of self-stress fields for simply connected domains and show how they are related in order to avoid linear dependencies. Then, based on the seminal work of [42] for triangular meshes, we derive the null vectors from self-equilibrium force fields defined on the set of edges. Finally, we require only a single QR decomposition in order to compute the remaining, typically very small fraction of null vectors corresponding to indeterminate supports and topological holes. As these null vectors cannot be expected to be very sparse in general, we argue that incorporating a small fraction of dense columns is the most practical approach and show how a partitioned Cholesky factorization can be used for efficient application.

In the second part of the paper, we show how the direct data-driven method can be reformulated as an unconstrained minimization problem using the null basis. In particular, we develop algorithms for two kind of procedures: energy minimization (or exploitation) and exploration of energy landscapes. For energy minimization, we numerically demonstrate the identical convergence behavior of the original fixed-point solver and a modified Newton-Raphson solver for the proposed, unconstrained problem formulation. However, computational gains can only be expected for systems with a very low degree of static indeterminacy. The parameterization of equilibrium stress fields becomes particularly useful for the exploration of multi-modal energy landscapes resulting from non-trivial, noisy data sets where no unique stress-strain relationship can be established, such as in the case of brittle damage [134]. To tackle this challenge, we finally develop a new algorithm based on the Deterministic Annealing framework [144, 143] which computes a hierarchical solution set that can be useful for uncertainty quantification purposes. Finally, we conclude the paper in Sec. 4.5 and give an outlook for further applications and developments.

4.3 Part I: An explicit and sparse basis for equilibrium stress fields

4.3.1 Notation

Part I involves the topology and geometry of tetrahedral meshes. We use the following notations. Let N, E, F, T denote the sets of nodes, edges, facets and tetrahedrons, respectively. Let $|\bullet|$ denote the size of a set, for instance $|N|$ denotes the number of vertices. The elements of the sets are labeled by the lower case letters $n \in N, e \in E, f \in F$ and $t \in T$. Furthermore, a geometrical quantity is a set of lower order quantities, for instance the nodes of a tetrahedron $t \in T$ are denoted by $n \in t$. Using the set difference notation, we can identify a facet of t opposite to node n by $(t \setminus n)$. We further use the shorthand notation for conditional sums like $\sum_{n:n \in t} := \sum_{n \in N: n \in t}$ to indicate the sum over all nodes of tetrahedron t and likewise $\sum_{t:n \in t}$ for the sum over all tetrahedrons which share a node n . The volumes of various elements are denoted by edge length l_e , facet area A_f and tetrahedral volume V_t .

4.3.2 Discrete strain and divergence operators

We consider discrete mechanical problems with linearized kinematics defined on tetrahedral meshes for the three-dimensional case. The basic field quantities involved are the infinitesimal strain field

$$\varepsilon : T \rightarrow \text{sym}(\mathbb{R}^3 \times \mathbb{R}^3) \quad (4.6)$$

and the stress field

$$\sigma : T \rightarrow \text{sym}(\mathbb{R}^3 \times \mathbb{R}^3) \quad (4.7)$$

which are symmetric tensor fields that are constant over each tetrahedron. We denote the strain and stress in a given tetrahedron $t \in T$ by ε_t and σ_t .

Both fields are subject to linear constraints: kinematical compatibility and force equilibrium. We enforce compatibility by deriving the strain field as the symmetric gradient of a displacement field

$$\mathbf{u} : N \rightarrow \mathbb{R}^3 \quad (4.8)$$

as

$$\boldsymbol{\varepsilon}_t := \frac{1}{2} \left(\text{Grad}_t(\mathbf{u}) + \text{Grad}_t^T(\mathbf{u}) \right). \quad (4.9)$$

Using a standard barycentric interpolation [162], the displacement gradient reads

$$\text{Grad}_t(\mathbf{u}) := - \sum_{n:n \in t} \frac{A_{t \setminus n}}{3V_t} \mathbf{u}_n \otimes \mathbf{n}_{t \setminus n} \quad (4.10)$$

where $A_{t \setminus n}$ is the area of the facet of tetrahedron t opposite to node n and $\mathbf{n}_{t \setminus n}$ denotes the unit normal of that facet pointing *outwards* of t , see Fig. 4.1 for a visualization. We note that the barycentric interpolation yields the same form as if derived from an isoparametric concept with linear shape functions, but this form proves useful to derive and show further geometric relationships in the following.

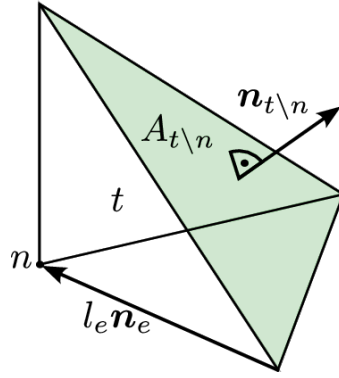


Figure 4.1: Tetrahedron t . The set notation $(t \setminus n)$ denotes the facet of t opposite to node n with area size $A_{t \setminus n}$ and outward unit normal $\mathbf{n}_{t \setminus n}$.

Using a suitable Voigt notation

$$\boldsymbol{\varepsilon}_t = (\varepsilon_{t,11}, \varepsilon_{t,22}, \varepsilon_{t,33}, 2\varepsilon_{t,12}, 2\varepsilon_{t,13}, 2\varepsilon_{t,23}) \quad (4.11)$$

$$\boldsymbol{\sigma}_t = (\sigma_{t,11}, \sigma_{t,22}, \sigma_{t,33}, \sigma_{t,12}, \sigma_{t,13}, \sigma_{t,23}), \quad (4.12)$$

the strain field can be expressed by the matrix vector product

$$\boldsymbol{\varepsilon} = \mathbf{B}^{\text{full}} \mathbf{u}^{\text{full}} = \mathbf{B} \mathbf{u} + \boldsymbol{\varepsilon}_0 \quad (4.13)$$

where the superscript 'full' indicates that the constraint degrees of freedom are included.

For a cleaner notation, we do not use a different notation for Voigt and tensor forms but give remarks if required. Rewriting Eq. (4.13) in Voigt notation, the discrete strain operator consists of the block matrices

$$\mathbf{B}_{t,n}^{\text{full}} = - \sum_{n:n \in t} \frac{A_{t \setminus n}}{3V_t} \begin{bmatrix} n_{t \setminus n,1} & 0 & 0 & n_{t \setminus n,2} & n_{t \setminus n,3} & 0 \\ 0 & n_{t \setminus n,2} & 0 & n_{t \setminus n,1} & 0 & n_{t \setminus n,3} \\ 0 & 0 & n_{t \setminus n,3} & 0 & n_{t \setminus n,1} & n_{t \setminus n,2} \end{bmatrix}^T. \quad (4.14)$$

We say a stress field $\boldsymbol{\sigma}$ is in equilibrium if its nodal divergence balances the external forces $\mathbf{p} : N \rightarrow \mathbb{R}^3$ as

$$\text{Div}_n(\boldsymbol{\sigma}) + \mathbf{p}_n = \mathbf{0} \quad \forall n \in N. \quad (4.15)$$

By Cauchy's stress theorem, we define the divergence

$$\text{Div}_n(\boldsymbol{\sigma}) := \sum_{t:n \in t} \frac{A_{t \setminus n}}{3} \boldsymbol{\sigma}_t \cdot \mathbf{n}_{t \setminus n} \quad (4.16)$$

as the net outward force flux of node n .

We call the matrix form of the (negative) divergence operator in Voigt notation the equilibrium matrix $\mathbf{E}^{\text{full}}$ consisting of block matrices

$$\mathbf{E}_{n,t}^{\text{full}} = V_t (\mathbf{B}_{t,n}^{\text{full}})^T \quad (4.17)$$

noting the adjoint relationship between the strain and divergence operator. Thus, the assembled equilibrium matrix after removing the constraint degrees of freedom can be written as

$$\mathbf{E} = \mathbf{B}^T \mathbf{V} \quad (4.18)$$

where $\mathbf{V}$ is the diagonal matrix of tetrahedral volumes.

4.3.3 Dimensionality and components of the null basis

We start by identifying the the number of linear independent null vectors for simply connected domains, i.e. domains without holes.

By the rank-nullity theorem, we have

$$\dim(\text{Im}(\mathbf{E})) + \dim(\text{Ker}(\mathbf{E})) = 6|T| \quad (4.19)$$

where 'dim' denotes the dimension, 'Im' the image space and 'Ker' the kernel or null space.

Assuming that the supports does not allow any rigid body movements, the dimension of the image space of the reduced equilibrium matrix is

$$\dim(\text{Im}(\mathbf{E})) = 3|N| - |D| \quad (4.20)$$

with $|D|$ being the number of degrees of freedom constraint by Dirichlet boundary conditions. Thus, the dimension of the null space $\mathbf{Z}$ of $\mathbf{E}$ is

$$\dim(\mathbf{Z}) = \dim(\text{Ker}(\mathbf{E})) = 6|T| - (3|N| - |D|). \quad (4.21)$$

We can get further insights by differentiating between internal quantities within the domain and external quantities on the boundary of the domain, indicated by subscripts 'int' and 'ext', respectively. From the counting rules

$$|N| = |N_{\text{int}}| + |N_{\text{ext}}| \quad (4.22)$$

$$4|T| = 2|F_{\text{int}}| + |F_{\text{ext}}| \quad (4.23)$$

$$2|E_{\text{ext}}| = 3|F_{\text{ext}}| \quad (4.24)$$

and Euler's characteristic for closed surfaces

$$|N_{\text{ext}}| - |E_{\text{ext}}| + |F_{\text{ext}}| = 2 \quad (4.25)$$

we obtain

$$\dim(\mathbf{Z}) = 3|F_{\text{int}}| - 3|N_{\text{int}}| + |D| - 6. \quad (4.26)$$

To derive independent mechanisms, we propose the split

$$\dim(\mathbf{Z}) = (3|F_{\text{int}}| - |E_{\text{int}}|) + (|E_{\text{int}}| - 3|N_{\text{int}}|) + (|D| - 6). \quad (4.27)$$

The motivation for the split is that a group of internal facets is linked by one internal edge, creating a dependency which needs to be resolved, see Fig. 4.2 a) for a visualization. The consequence is that the following mechanisms are required to produce linear independent null vectors:

- Facet-wise mechanism: $(3|F_{\text{int}}| - |E_{\text{int}}|)$ null vectors. Each of the 3 edges of every internal facet creates one null vector which is supported on the pair of tetrahedrons adjacent to the internal facet. Every internal edge has to be left out *once*.
- Edge-wise mechanism: $(|E_{\text{int}}| - 3|N_{\text{int}}|)$ null vectors. An internal edge left out from the facet-wise mechanism can create one, more complex, null vector supported by the group of tetrahedrons adjacent to the internal edge. For every internal node, three internal edge-wise mechanism has to be left out to ensure linear independent null vectors.
- Dirichlet mechanism: $(|D| - 6)$ null vectors. Every Dirichlet boundary condition creates one null vectors except for 6 supports preventing rigid body motions.
- Holes in the domain induce further null vectors. A general treatment of homologies in the sense of algebraic topology [121] is, however, outside the scope of this paper.

Before we derive the algorithms for each mechanism we establish a useful geometric relation between force fields supported on edges and stress fields as shown next.

4.3.4 Stress fields induced by edge force fields

The preceding analysis suggests that the degrees of freedom are scalars supported on edges. We follow an intuitive picture and think of force fields supported on the edges of each tetrahedron from which the stresses can be computed consistently as shown in this subsection. This approach is a generalization to the three dimensional case of the concepts introduced in [42].

We consider scalar fields

$$P : E \times T \rightarrow \mathbb{R} \quad (4.28)$$

subject to the constraint

$$\sum_{e:n \in e} \left(\sum_{t:e \in t} P_{e,t} \right) (\mathbf{x}_{e \setminus n} - \mathbf{x}_n) / l_e = \mathbf{0} \quad \forall n \in N \quad (4.29)$$

where $\mathbf{x}_n \in \mathbb{R}^3$ and $\mathbf{x}_{e \setminus n} \in \mathbb{R}^3$ are the nodal coordinates of an edge e adjacent to node n .

Interpreting $P_e = \sum_{t:e \in t} P_{e,t}$ as the resultant force magnitude of edge e , Eq. (4.29) resembles the classical equilibrium of nodal forces. Then, $P_{e,t}$ can be also seen as the part of P_e distributed to tetrahedron t .

We define the associated stress of a tetrahedron $t \in T$ as

$$\boldsymbol{\sigma}_t = \sum_{e:e \in t} \frac{l_e}{V_t} P_{e,t} (\mathbf{n}_e \otimes \mathbf{n}_e) \quad (4.30)$$

where $\mathbf{n}_e$ is a unit vector along the edge e with an arbitrary but fixed orientation.

By inserting Eq. (4.30) into the divergence Eq. (4.16)

$$\text{Div}_n(\boldsymbol{\sigma}) = \sum_{t:n \in t} \frac{A_{t \setminus n}}{3} \left(\sum_{e:e \in t} \frac{l_e}{V_t} P_{e,t} \mathbf{n}_e \otimes \mathbf{n}_e \right) \cdot \mathbf{n}_{t \setminus n} \quad (4.31)$$

$$= \sum_{t:n \in t} \sum_{e:e \in t} P_{e,t} \frac{(l_e \mathbf{n}_e) \cdot (A_{t \setminus n} \mathbf{n}_{t \setminus n})}{3V_t} \mathbf{n}_e \quad (4.32)$$

$$= \sum_{e:n \in e} \left(\sum_{t:e \in t} P_{e,t} \right) \text{sign}(\mathbf{n}_e \cdot \mathbf{n}_{t \setminus n}) \mathbf{n}_e = \mathbf{0} \quad (4.33)$$

we can show that if a force field P satisfies constraint Eq. (4.29), then the associated stress field is divergence free. Note in particular that

$$\frac{(l_e \mathbf{n}_e) \cdot (A_{t \setminus n} \mathbf{n}_{t \setminus n})}{3V_t} = \begin{cases} +1, & \text{if } \mathbf{n}_e \text{ points away from } n \\ -1, & \text{if } \mathbf{n}_e \text{ points towards } n \\ 0, & \text{if } e \in (t \setminus n) \end{cases} \quad (4.34)$$

holds (see again Fig. 4.1).

4.3.5 Facet-wise mechanism

In this section, we give the algorithm how to compute one null vector $\mathbf{z}$ based on the facet-wise mechanism. Each null vector corresponds to one internal facet $f \in F_{\text{int}}$ and one of its edges $e \in f$. Along the given edge e , a pair of opposite unit forces is applied on the two adjacent tetrahedrons t_0 and t_1 as

$$P_{e,t_0} = +1 \quad (4.35)$$

$$P_{e,t_1} = -1 \quad (4.36)$$

noting that the required constraint (4.29) is fulfilled trivially since $P_e = P_{e,t_0} + P_{e,t_1} = 0$ and $P_{e,t} = 0$ everywhere else, see Fig. 4.2 b) for a visualization.

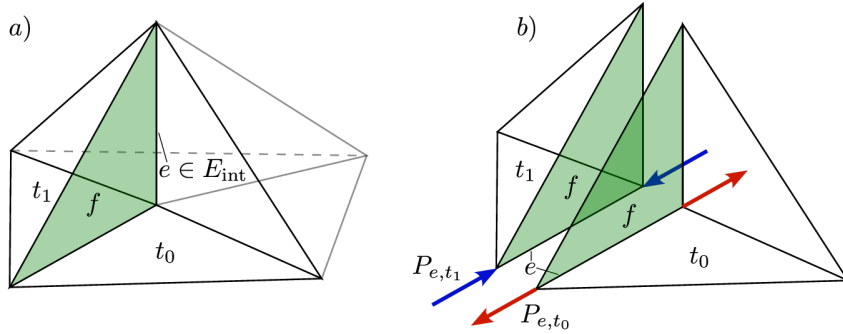


Figure 4.2: a) Facets adjacent to an internal edge. b) Pair of equal but opposite directing forces along a shared edge of two adjacent tetrahedrons.

The resulting null vector $\mathbf{z}$ is thus supported on t_0 and t_1 is given by evaluating the stress function (4.30) to

$$\mathbf{z}_{t_0} = +(l_e/V_{t_0})(\mathbf{n}_e \otimes \mathbf{n}_e) \quad (4.37)$$

$$\mathbf{z}_{t_1} = -(l_e/V_{t_1})(\mathbf{n}_e \otimes \mathbf{n}_e). \quad (4.38)$$

We remark that this approach is a direct generalization of the two-dimensional case proposed in [42]. The procedure is summarized in Algorithm (5).

Algorithm 5 Facet-wise mechanism.**Require:** Internal facet $f \in F_{int}$ **Require:** Edge $e \in f$ $\{t_0, t_1\} \leftarrow \{t \in T : f \in t\}$ $\mathbf{z} \leftarrow \mathbf{0}$ $\mathbf{z}_{t_0} \leftarrow +(l_e/V_{t_0})(\mathbf{n}_e \otimes \mathbf{n}_e)$ $\mathbf{z}_{t_1} \leftarrow -(l_e/V_{t_1})(\mathbf{n}_e \otimes \mathbf{n}_e)$ **return** TensorToVoigt($\mathbf{z}$)**4.3.6 Edge-wise mechanism**

In this section, we derive an algorithm to compute one null vector $\mathbf{z}$ based on the edge-wise mechanism.

Every null vector corresponds to one internal edge e . Let $T_e = \{t \in T : e \in t\}$ be set of tetrahedrons adjacent to e , which we call a basic module. Further, let $N_e^{\text{cycle}} = \{n \in N : n \in T_e, n \notin e\}$ be the set of nodes forming a cycle around e . We note that in a typical mesh, N_e^{cycle} contains at least 3 nodes (which is a very well studied case in classical literature [115]) but usually $|N_e^{\text{cycle}}|$ is larger than 3. However, independent of the size of the basic module, the number of edges of a basic module is $3|N_e^{\text{cycle}}| + 1$ which implies that there is exactly one self-stress field for each basic module. In order to derive an edge-tetrahedron force field $P_{e,t}$ satisfying Eq. (4.29), We implement the following procedure:

1. Compute an equilibrium force field P_e which is only non-zero on the edges of the basic module.
2. Compute $P_{e,t}$ by distributing the edge forces P_e to the tetrahedrons of the basic module.
3. Compute $\mathbf{z}$ from $P_{e,t}$ using the stress formula Eq. (4.30).

In order to obtain a self-equilibrium edge force field, we generalize an algorithm of [2] for basic modules with $|N_e^{\text{cycle}}| \geq 3$. The algorithm exploits the fact that every node in the cycle is shared by exactly four edges with a non-zero force, which is the minimum number of forces in three dimensions to form equilibrium, see Fig. 4.3 for a

visualization. The edge forces can then be determined by a purely geometrical procedure using a recursive application of signed volume ratios as described in the following.

Every node $n \in N_e^{\text{cycle}}$ is joined by the four edges $\mathbf{v}_j := (\mathbf{x}_{e_j \setminus n} - \mathbf{x}_n)$, $j \in (1, 2, 3, 4)$ of the basic module with force magnitudes P_j . To identify the ratio between two of the edge forces, P_1 and P_2 for example, the equation of force equilibrium

$$P_1 \mathbf{v}_1 / l_1 + P_2 \mathbf{v}_2 / l_2 + P_3 \mathbf{v}_3 / l_3 + P_4 \mathbf{v}_4 / l_4 = \mathbf{0} \quad (4.39)$$

is multiplied by a cross product and contraction of the other two edges as

$$P_1 (\mathbf{v}_1 \times \mathbf{v}_3) \cdot \mathbf{v}_4 / l_1 + P_2 (\mathbf{v}_2 \times \mathbf{v}_3) \cdot \mathbf{v}_4 / l_2 = 0 \quad (4.40)$$

yielding

$$\frac{P_2}{P_1} = -\frac{l_2 (\mathbf{v}_1 \times \mathbf{v}_3) \cdot \mathbf{v}_4}{l_1 (\mathbf{v}_2 \times \mathbf{v}_3) \cdot \mathbf{v}_4}. \quad (4.41)$$

Choosing $P_1 = 1$ as an initial unit force, all other edge forces can be recursively computed by evaluating ratio P_j / P_1 , $j \in (2, 3, 4)$. It is important that edge 1 is an edge of the cycle in order to ensure non-zero denominators as the denominators are then always signed volumes of single tetrahedrons. After repeating this process for all nodes of the cycle in an sorted manner, the force of the central edge e can be computed simply by completing force equilibrium

$$P_e = -\frac{1}{l_e} (\mathbf{x}_{e \setminus n} - \mathbf{x}_n) \cdot \sum_{e': n \in e', e' \neq e} \frac{P_{e'}}{l_{e'}} (\mathbf{x}_{e' \setminus n} - \mathbf{x}_n). \quad (4.42)$$

at one of its node $n \in e$. For the second step, the edge forces are equally distributed to the tetrahedrons of the basic module as

$$P_{e,t} = \frac{\mathbb{1}(t \in T_e)}{\sum_{t \in T_e} \mathbb{1}(e \in t)} P_e \quad (4.43)$$

where $\mathbb{1}(\bullet)$ is the indicator function evaluating to 1 if the argument is true and 0 otherwise. We note that the choice (4.43) is not unique but ensures condition (4.29). Having computed the force field $P_{e,t}$, the associated null vector follows again from the stress formula (4.30). A basic module is visualized in Fig. 4.3 and the procedure is summarized in Algorithm (6).

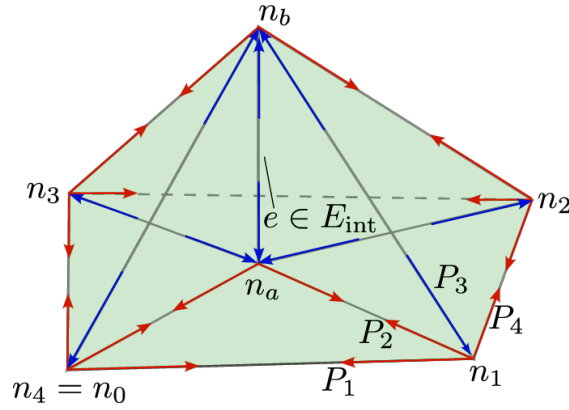


Figure 4.3: Edge forces of an basic module around an internal edge e . The outer nodes form a cycle $(n_1, n_2, n_3, n_4, n_1)$ of length $K = 4$.

Algorithm 6 Edge-wise mechanism.

Require: Internal edge $e \in E_{\text{int}}$

```

 $z \leftarrow 0$ 
 $T_e \leftarrow \{t \in T : e \in t\}$ 
 $\{n_a, n_b\} \leftarrow \{n \in N : n \in e\}$ 
 $N_e^{\text{cycle}} \leftarrow \{n \in N : n \in T_e\} \setminus \{n_a, n_b\}$ 
 $(n_1, n_2, \dots, n_K = n_0, n_1) \leftarrow \text{Sort}(N_e^{\text{cycle}})$  //Sort as an closed cycle using node-edge
connectivity
 $P_{(n_0, n_1)} = 1$ 
for  $k = 1 \dots K$  do
     $P_{(n_k, n_a)} \leftarrow \alpha P_{(n_{k-1}, n_k)}$  // $\alpha$  :signed volume ratio using Eq. (4.41)
     $P_{(n_k, n_b)} \leftarrow \alpha P_{(n_{k-1}, n_k)}$ 
     $P_{(n_k, n_{k+1})} \leftarrow \alpha P_{(n_{k-1}, n_k)}$ 
end for
 $P_{(n_a, n_b)} \leftarrow \text{Eq. (4.42)}$  // force of central edge  $e$ 
for all  $t \in T_e$  do
    for all  $e' \in t$  do
         $D = |\{t \in T_e : e' \in t\}| \in \{1, 2, K\}$  //Distribution factor of Eq.(4.43)
         $P_{e', t} \leftarrow P_{e'} / D$ 
         $z_t \leftarrow z_t + (l_{e'} / V_t)(P_{e', t})(\mathbf{n}_{e'} \otimes \mathbf{n}_{e'})$ 
    end for
end for
return TensorToVoigt( $z$ )

```

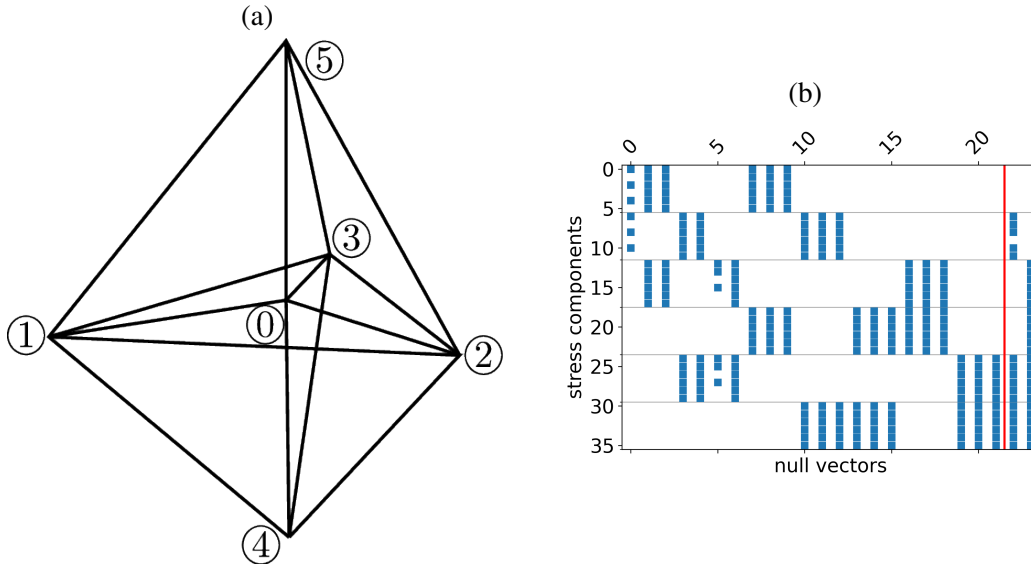
As an alternative, one can also consider the related lumped stress method [43], where

self-equilibrium force field supported on the edges of triangular meshes is derived from jumps in the gradient of an Airy stress function. However, a generalization to the three-dimensional case is not straightforward [87, 59], which is why the present approach is preferred.

4.3.6.1 Minimal example

Before continuing, we provide a minimal example featuring the facet- and edge-wise mechanisms as shown in Fig. 4.4 A). The facet-wise mechanism contributes $3|F_{\text{int}}| - |E_{\text{int}}| = 3 * 9 - 5 = 22$ null vectors, which are shown as the first 22 columns in the sparsity pattern in Fig. 4.4 B). Each null vector is supported on 2 adjacent tetrahedrons with 6 stress components each. The edge-wise mechanism contributes $|E_{\text{int}}| - 3|N_{\text{int}}| = 5 - 3 = 2$ null vectors, shown as the last 2 columns in Fig. 4.4 B). The first null vector of this type cycles around edge $(0, 4)$ supported on 3 tetrahedrons, while the second null vector cycles around edge $(0, 3)$ supported on 4 tetrahedrons. The rank of equilibrium matrix is $\text{rank}(\mathbf{E}) = 3|N| - 6 = 12$ since the rigid body motions are not prevented by Dirichlet boundary conditions. Thus, we fulfill the rank-nullity condition (4.19) by $\text{rank}(\mathbf{E}) + \text{rank}(\mathbf{Z}) = 12 + 24 = 6|T|$.

Figure 4.4: A) Minimal example with 6 tetrahedrons, 1 internal node (labeled 0), 5 internal edges and 9 internal facets. B) Sparsity pattern of the null base $\mathbf{Z}$.



4.3.7 Remaining null vectors

Every constrained displacement degree of freedom creates one null vector z of which 6 are linear dependent assuming the boundary conditions do not admit any rigid body motions of the whole body. Furthermore, every hole in the domain introduces additional null vectors. For these remaining null vectors, no closed form solution like the facet-wise or edge-wise mechanism can be generally stated, as they depend on the distribution of the boundary conditions and geometry of holes. In addition, we cannot expect in general highly sparse null vectors as the holes can be quite large or boundary conditions may be sparsely distributed. However, it is important to note that this part of the null space is significantly smaller than the previous parts as the number of boundary conditions scales at most quadratic of the surface and not cubic. Furthermore, the number of holes is not dependent on the discretization and is thus limited. In principle, sparsification strategies can be employed for this part. We have investigated this approach using methods such as least angle regression (LARS) [32] or iteratively reweighted least squares (IRLS) [17]. However, these tests proved uncompetitive due to a significantly larger computational overhead and numerical instabilities such as an ill-conditioned null basis. For our purpose we found that it is sufficient, numerically more robust and more practicable to compute the remaining, dense part of the null base Z_D as the null vectors of equilibrium matrix E which are orthogonal to the existing, sparse null vectors Z_S .

In general, an orthogonal (but dense) null basis Z of matrix E can be computed by a sparse QR decomposition [29] of E^T as

$$E^T P = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \quad (4.44)$$

with column permutation P . Then, Q_2 is a null basis satisfying $EQ_2 = 0$. The reason for this is that any null basis Z also satisfies $Z^T E^T = 0$, i.e. the null space of E is orthogonal to the row space of E . Then, since Q_1 is a basis for the row space of E , Q_2 spans the null space as it is orthogonal to Q_1 .

We are now interested in computing only the remaining null vectors of E which are orthogonal to the known part of the null basis Z_S . This can be achieved by a QR

decomposition of the augmented matrix

$$[\mathbf{E}^T, \mathbf{Z}_S] \mathbf{P} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} \quad (4.45)$$

from which one obtains $\mathbf{Z}_D = \mathbf{Q}_2$. The complete null basis is then given by

$$\mathbf{Z} = [\mathbf{Z}_S, \mathbf{Z}_D] \quad (4.46)$$

where the sparse part $\mathbf{Z}_S$ contains the null vectors from the facet-wise and edge-wise mechanisms and $\mathbf{Z}_D$ contains all remaining, missing null vectors. It is interesting to note that this mechanism is an instance of the Fault Tolerant Computing [3] theory. If, due to numerical instabilities, $\mathbf{Z}_S$ contains some linear dependent null vectors up to a certain precision, the orthogonalization procedure will add all missing null vectors, thus automatically correcting previous errors. In practice, however, we did not encounter this case.

As an alternative, if one is only interested in computing the null vectors corresponding to the Dirichlet boundary conditions, it is numerically more efficient to decompose

$$[\mathbf{E}^{\text{full}} \mathbf{E}^T] \mathbf{P} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} \quad (4.47)$$

and compute $\mathbf{Z}_D = (\mathbf{E}^{\text{full}})^T \mathbf{Q}_2$. This approach exploits again the orthogonality property of the row space and the null space: since we consider only null vectors for $\mathbf{E}$ which are in the row space of $\mathbf{E}^{\text{full}}$, we ensure that those are orthogonal to its null basis $\mathbf{Z}_S$.

The algorithms are summarized in Algorithm (7) and Algorithm (8). While Algorithm (7) computes *all* remaining null vectors, Algorithm (8) computes only null vectors from Dirichlet boundary conditions in a more efficient way.

Algorithm 7 Remaining null vectors

Require: Equilibrium matrix $\mathbf{E}$

Require: Known part of the null base $\mathbf{Z}_S$

QR decomposition with column ordering: $[\mathbf{E}^T, \mathbf{Z}_S] \mathbf{P} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix}$

$\mathbf{Z}_D \leftarrow \mathbf{Q}_2$

return $\mathbf{Z}_D$

Algorithm 8 Dirichlet null vectors (alternative to Alg. (7))

Require: Full and reduced equilibrium matrices $\mathbf{E}^{\text{full}}, \mathbf{E}$

QR decomposition with column ordering: $[\mathbf{E}^{\text{full}} \mathbf{E}^T] \mathbf{P} = [\mathbf{Q}_1 \quad \mathbf{Q}_2] \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix}$

$\mathbf{Z}_D \leftarrow (\mathbf{E}^{\text{full}})^T \mathbf{Q}_2$

return $\mathbf{Z}_D$

4.3.8 Sparse null basis assembly algorithm

In this section, we evaluate the previously developed mechanisms while correctly handling the conditions for linear dependencies as derived in section 4.3.3. While it is in principle possible to compute the edge-wise and facet-wise part of the null basis separately, the combined evaluation has the advantage that no column sorting is required as the sparse matrix is automatically also banded. In addition, since all null vectors are linear independent by construction, we do not required an additional factorization in order to remove redundant null vectors. The final algorithm is given in Alg. 9.

Algorithm 9 Assembly algorithm.

```

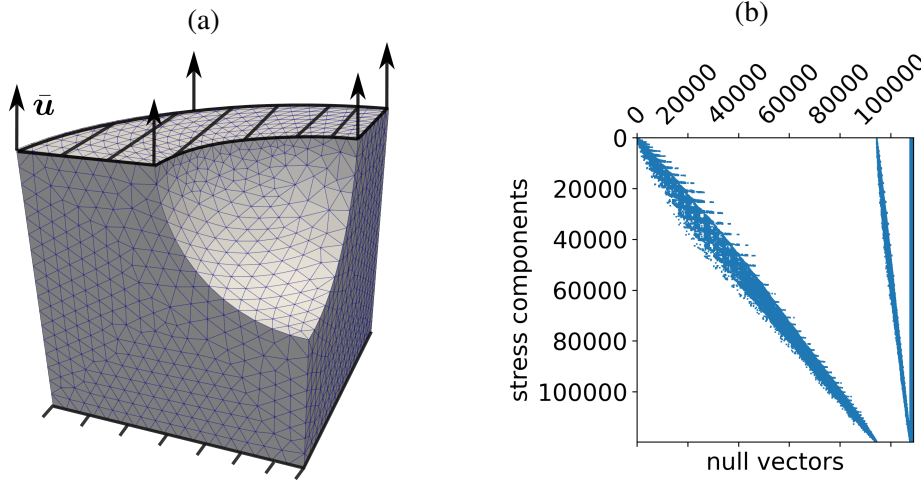
 $Z_S \leftarrow 0$ 
 $I_e \leftarrow \text{false} \forall e \in E_{\text{int}}$ 
 $I_e \leftarrow \text{true} \forall e \in E_{\text{ext}}$ 
 $I_n \leftarrow 0 \forall n \in N_{\text{int}}$ 
 $I_n \leftarrow 3 \forall n \in N_{\text{ext}}$ 
 $i \leftarrow 1$ 
for all  $f \in F_{\text{int}}$  do
  for all  $e \in f$  do
    if  $I_e = \text{false}$  then
       $I_e \leftarrow \text{true}$  //skip every internal edge once
       $\{n_a, n_b\} \leftarrow \{n \in N : n \in e\}$ 
      if  $I_{n_a} < 3$  then
         $I_{n_a} \leftarrow I_{n_a} + 1$  //skip every internal node three times
      else if  $I_{n_b} < 3$  then
         $I_{n_b} \leftarrow I_{n_b} + 1$  //skip every internal node three times
      else
         $Z_{S,i} \leftarrow \text{Edge-wise-mechanism}(e)$  (Alg. 2)
         $i \leftarrow i + 1$ 
      end if
    else
      Facet-wise mechanism
       $Z_{S,i} \leftarrow \text{Facet-wise-mechanism}(f, e)$  (Alg. 1)
       $i \leftarrow i + 1$ 
    end if
  end for
end for
 $Z_D \leftarrow \text{RemainingColumns}(E^{\text{full}}, E, Z_S)$  (Alg. 3 or Alg. 4)
return  $Z_S, Z_D$ 

```

4.3.8.1 Benchmark example

In order to demonstrate how the proposed algorithm performs with a larger problem, we consider the mesh shown in Fig. 4.6 A). The mesh consists of 19954 elements and 4030 nodes. The displacements of the nodes at the top and bottom sides are prescribed in all directions, removing in total 1299 displacement degrees of freedoms. The required topological connectivities are obtained from the open-source finite element library dolfinx [8].

Figure 4.6: A) Benchmark example boundary value problem. B) Sparsity pattern of the null base. For visualization, the edge-wise and facet-wise parts are plotted separately.



The sparsity pattern of the null base shown in Fig. 4.6 B) visualizes that the facet-wise mechanism (first block of columns) is by far the largest contribution with 94373 columns, followed by the edge-wise mechanism (middle block) with 13267 columns. Lastly, the Dirichlet boundary conditions contribute 1299 dense columns (last block) which are approximately 1.2% of all columns. The sparsity, as measured by the number of nonzero elements divided by the matrix dimensions, of the sparse part (facet-wise and edge-wise mechanism) is with less than 0.012% very low. The computation times are promising: a Python implementation running on an AMD Ryzen 9 7950X 16-Core Processor CPU requires 4 seconds for the sparse part and 3 seconds for the dense part using Alg. (8). The error of the null basis as measured by the norm $\|\mathbf{E}\mathbf{Z}\|_2 \approx 10^{-12}$ is at machine precision.

From a practical point of view, forming large dense matrices has to be avoided as much as possible. For instance, $\mathbf{Q}_2$ should be computed by an application of Householder transformations (as provided by the libraries) on the last 1299 columns of an identity matrix instead of forming the whole matrix $\mathbf{Q}$ explicitly.

4.3.9 Efficient Cholesky factorization of the null basis

As elaborated in the previous sections, we consider a null basis $\mathbf{Z} = [\mathbf{Z}_S, \mathbf{Z}_D]$ consisting of a sparse part $\mathbf{Z}_S$ and a dense part $\mathbf{Z}_D$ which has much less columns than the sparse part. For the later application we require factorizations of $\mathbf{H} = \mathbf{Z}^T \mathbf{D} \mathbf{Z}$ where $\mathbf{D}$ is a symmetric block diagonal matrix. While $\mathbf{Z}^T \mathbf{Z}$ is indeed a sparse and banded matrix due to the orthogonality of $\mathbf{Z}_D$, a matrix $\mathbf{D}$ destroys this property in general leading to an unfavorable sparsity pattern of $\mathbf{H}$.

In this subsection, we will briefly present a *partitioned* Cholesky factorization which stores $\mathbf{Z}_D$ separately as a dense matrix leading to much more efficient factorization and solving operations. As visualized in Fig. 4.8, matrix $\mathbf{H}$ consists of the large, sparse block $\mathbf{H}_{SS} = \mathbf{Z}_S^T \mathbf{D} \mathbf{Z}_S$ and the small, dense blocks $\mathbf{H}_{DS} = \mathbf{Z}_D^T \mathbf{D} \mathbf{Z}_S$, $\mathbf{H}_{DS}^T$ and $\mathbf{H}_{DD} = \mathbf{Z}_D^T \mathbf{D} \mathbf{Z}_D$.

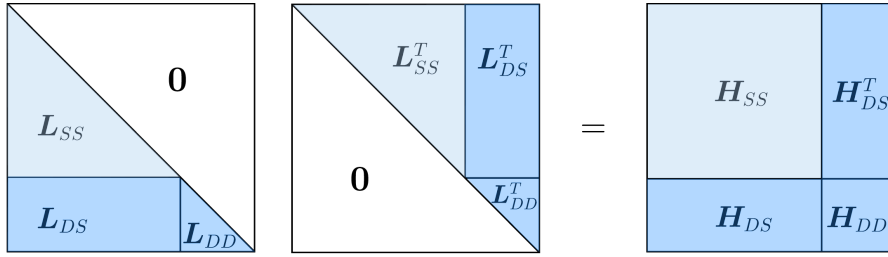


Figure 4.8: Partitioned Cholesky factorization $\mathbf{L}\mathbf{L}^T = \mathbf{H}$. Light blue and dark blue blocks are sparse and dense matrices, respectively.

The main block $\mathbf{L}_{SS}$ of the lower triangular matrix $\mathbf{L}$ can then be efficiently computed by a sparse, supernodal Cholesky decomposition [18] as $\mathbf{L}_{SS}\mathbf{L}_{SS}^T = \mathbf{H}_{SS}$. Using this factorization, we can solve $\mathbf{L}_{SS}\mathbf{L}_{DS}^T = \mathbf{H}_{DS}^T$ for $\mathbf{L}_{DS}^T$. Finally, $\mathbf{L}_{DD}$ is obtained by a dense but small Cholesky factorization $\mathbf{L}_{DD}\mathbf{L}_{DD}^T = \mathbf{H}_{DD} - \mathbf{L}_{DS}\mathbf{L}_{DS}^T$.

Then, for solving $\mathbf{H}\mathbf{x} = \mathbf{b}$, the forward and backward substitution of $\mathbf{L}$ can also be carried out separately for the sparse and dense parts. The necessary steps are summarized in Algorithm (10).

Algorithm 10 Partitioned Cholesky factorization and solving**Require:** Sparse columns $\mathbf{Z}_S$, dense columns $\mathbf{Z}_D$ **Require:** Symmetric block diagonal matrix $\mathbf{D}$ **Require:** Right-hand side $\mathbf{b}$

#Factorization

$$\mathbf{H}_{SS} \leftarrow \mathbf{Z}_S^T \mathbf{D} \mathbf{Z}_S$$

$$\mathbf{H}_{DS} \leftarrow \mathbf{Z}_D^T \mathbf{D} \mathbf{Z}_S$$

$$\mathbf{H}_{DD} \leftarrow \mathbf{Z}_D^T \mathbf{D} \mathbf{Z}_D$$

$$\mathbf{L}_{SS} \leftarrow \text{SparseCholesky}(\mathbf{H}_{SS})$$

$$\mathbf{L}_{DS}^T \leftarrow \mathbf{L}_{SS}^{-1} \mathbf{H}_{DS}^T$$

$$\mathbf{L}_{DD} \leftarrow \text{DenseCholesky}(\mathbf{H}_{DD} - \mathbf{L}_{DS} \mathbf{L}_{DS}^T)$$

#Solving

$$[\mathbf{b}_S, \mathbf{b}_D] \leftarrow \mathbf{b}$$

$$\mathbf{y}_S = \mathbf{L}_{SS}^{-1}(\mathbf{b}_S)$$

$$\mathbf{y}_D = \mathbf{L}_{DD}^{-1}(\mathbf{b}_D - \mathbf{L}_{DS} \mathbf{y}_S)$$

$$\mathbf{x}_D \leftarrow \mathbf{L}_{DD}^{-T} \mathbf{y}_D$$

$$\mathbf{x}_S \leftarrow \mathbf{L}_{SS}^{-T}(\mathbf{y}_S - \mathbf{L}_{DS}^T \mathbf{x}_D)$$

return $[\mathbf{x}_S, \mathbf{x}_D]$

4.4 Part II: Application for direct data-driven mechanics

4.4.1 Background: direct data-driven mechanics

We consider compatible and equilibrium states

$$(\boldsymbol{\varepsilon}, \boldsymbol{\sigma}) \in \mathbb{R}^{6|T|} \times \mathbb{R}^{6|T|} \quad \text{s.t.} \quad \boldsymbol{\varepsilon} = \mathbf{B}\mathbf{u} + \boldsymbol{\varepsilon}_0 \quad \text{and} \quad \mathbf{B}^T \mathbf{V} \boldsymbol{\sigma} = \mathbf{p} \quad (4.48)$$

of mechanical system with $|T|$ material points. Furthermore, we assume to have a data set

$$\mathbf{D} = \{(\boldsymbol{\varepsilon}'_i, \boldsymbol{\sigma}'_i)\}_{i=1}^{n_D}, \quad (\boldsymbol{\varepsilon}'_i, \boldsymbol{\sigma}'_i) \in \mathbb{R}^6 \times \mathbb{R}^6 \quad (4.49)$$

of n_D stress-strain pair measurements. We assume a homogeneous material behavior, i.e. every material point shares the same material data set.

As proposed by [82], the feasibility of an admissible state can be directly quantified

based on the available data by the sum of nearest neighbor distances

$$F(\boldsymbol{\varepsilon}, \boldsymbol{\sigma}) = \sum_{t=1}^{|T|} V_t \min_{(\boldsymbol{\varepsilon}'_i, \boldsymbol{\sigma}'_i) \in D} \text{dist}_t^2((\boldsymbol{\varepsilon}_t, \boldsymbol{\sigma}_t), (\boldsymbol{\varepsilon}'_i, \boldsymbol{\sigma}'_i)) \quad (4.50)$$

in each material point. The local distances are defined as

$$\text{dist}_t^2((\boldsymbol{\varepsilon}_t, \boldsymbol{\sigma}_t), (\boldsymbol{\varepsilon}'_i, \boldsymbol{\sigma}'_i)) = (\boldsymbol{\varepsilon}'_i - \boldsymbol{\varepsilon}_t) \cdot \mathbb{C}_t(\boldsymbol{\varepsilon}'_i - \boldsymbol{\varepsilon}_t) + (\boldsymbol{\sigma}'_i - \boldsymbol{\sigma}_t) \cdot \mathbb{C}_t^{-1}(\boldsymbol{\sigma}'_i - \boldsymbol{\sigma}_t) \quad (4.51)$$

with metric coefficients $\mathbb{C}_t$ and material point volumes V_t . As proposed in [83, 134], this rather strict measure can be relaxed but by the weighted distance to all data points as given by the energy

$$\mathcal{E}_\beta(\boldsymbol{\varepsilon}, \boldsymbol{\sigma}) = -\frac{1}{\beta} \sum_{t=1}^{|T|} \log \left(\sum_{i=1}^{n_D} \exp(-\beta V_t \text{dist}_t^2((\boldsymbol{\varepsilon}_t, \boldsymbol{\sigma}_t), (\boldsymbol{\varepsilon}'_i, \boldsymbol{\sigma}'_i))) \right) \quad (4.52)$$

with a regularization parameter β . Notably, the original distance measure (4.50) can be recovered by the limit

$$\lim_{\beta \rightarrow \infty} \mathcal{E}_\beta(\boldsymbol{\varepsilon}, \boldsymbol{\sigma}) = F(\boldsymbol{\varepsilon}, \boldsymbol{\sigma}) \quad (4.53)$$

reducing the sum over all data points to only the closest data point, which we will refer in the following as the 'min-dist' limit.

With this at hand, the direct data-driven problem can be written as the constraint double minimization problem

$$\min_{\mathbf{u}, \boldsymbol{\sigma}} \mathcal{E}_\beta(\mathbf{B}\mathbf{u} + \boldsymbol{\varepsilon}_0, \boldsymbol{\sigma}) \quad \text{s.t.} \quad \mathbf{B}^T \mathbf{V} \boldsymbol{\sigma} = \mathbf{p}. \quad (4.54)$$

The solution of this problem is thus the compatible strain and equilibrium stress field which minimizes the weighted distance to measured stress-strain pairs in all material points. In [82], a fixed-point iteration was proposed to solve the 'min-dist' limit case which was extended in [83] to solve the more general, 'min-energy' case. Therein, an annealing scheme is used which slowly increases $\beta \rightarrow \infty$ in order to converge to the global minima. For each smoothness level β , a fixed-point iteration is applied.

The fixed point iteration consists of the two operations of data search and projection

to admissible states and is repeated until convergence is achieved. The data search is a local operation. Given an admissible, local state $(\varepsilon_t, \sigma_t)$, we define the generalized data search as

$$\mathbb{E}_D[\varepsilon'_t] = \sum_{i=1}^{n_D} p_i(\varepsilon_t, \sigma_t) \varepsilon'_i \quad (4.55)$$

$$\mathbb{E}_D[\sigma'_t] = \sum_{i=1}^{n_D} p_i(\varepsilon_t, \sigma_t) \sigma'_i \quad (4.56)$$

with the softmax function

$$p_i(\varepsilon_t, \sigma_t) = \frac{\exp(-\beta V_t \text{dist}_t^2((\varepsilon_t, \sigma_t), (\varepsilon'_i, \sigma'_i)))}{\sum_{i=1}^{n_D} \exp(-\beta V_t \text{dist}_t^2((\varepsilon_t, \sigma_t), (\varepsilon'_i, \sigma'_i)))}. \quad (4.57)$$

The name data search comes from the fact that this operation reduces in the 'min-dist' limit to a nearest neighbor search which can be very efficiently performed, see [34] for a comparison in this context.

The projection step computes the closest compatible strain field to $\varepsilon' := \{\mathbb{E}_D[\varepsilon'_t]\}_{t=1}^{|T|}$ and the closest equilibrium stress field to $\sigma' := \{\mathbb{E}_D[\sigma'_t]\}_{t=1}^{|T|}$. This can be efficiently performed by solving two decoupled system of equations

$$(B^T V \mathbb{C} B) u = B^T V \mathbb{C} \varepsilon' \quad (4.58)$$

$$(B^T V \mathbb{C} B) \eta = p - B^T V \sigma' \quad (4.59)$$

$$\varepsilon = B u + \varepsilon_0 \quad (4.60)$$

$$\sigma = \sigma' + \mathbb{C} B \eta \quad (4.61)$$

with $\mathbb{C} = \text{diag}(\mathbb{C}_t)$. Hereby, η are Lagrange multipliers enforcing the equilibrium constraint, see [82, 83] for further details. Finally, we note that several solver variations has been proposed in the literature. For better accuracy, [136] proposed a randomized variant of the data search step, while [125] considered an adaptive scheme for the metric coefficients. An alternative approach is also the reformulation as a mixed-integer problem [73]. Particularly noteworthy is also the inverse problem formulation [104] which can be used to obtain data sets from digital image correlation experiments. Further strategies for adaptive, goal-oriented data acquisition were presented in [54].

4.4.2 Reformulation using null basis

The availability of a sparse basis for equilibrium stress fields allows us to formulate new solver schemes for the class of direct data-driven problems using the parameterization

$$\boldsymbol{\varepsilon} = \mathbf{B}\mathbf{u} + \boldsymbol{\varepsilon}_0 \quad (4.62)$$

$$\boldsymbol{\sigma} = \mathbf{Z}\mathbf{v} + \boldsymbol{\sigma}_0. \quad (4.63)$$

With this at hand, we can rewrite problem (4.54) as an unconstrained minimization problem

$$\min_{\mathbf{u}, \mathbf{v}} \mathcal{E}_\beta(\mathbf{B}\mathbf{u} + \boldsymbol{\varepsilon}_0, \mathbf{Z}\mathbf{v} + \boldsymbol{\sigma}_0) \quad (4.64)$$

which can be minimized by standard algorithms like the Newton-Raphson method. In particular, we can derive explicit expressions for the Jacobian and Hessian of $\mathcal{E}_\beta$ with respect to $\mathbf{u}$ and $\mathbf{v}$ as shown in the following. The Jacobian reads

$$\mathbf{J}_u(\mathbf{u}, \mathbf{v}) = \frac{\partial \mathcal{E}_\beta}{\partial \mathbf{u}} = \mathbf{B}^T \frac{\partial \mathcal{E}_\beta}{\partial \boldsymbol{\varepsilon}} = \mathbf{B}^T \text{diag} \left(\frac{\partial \mathcal{E}_\beta}{\partial \boldsymbol{\varepsilon}_t} \right) \quad (4.65)$$

$$\mathbf{J}_v(\mathbf{u}, \mathbf{v}) = \frac{\partial \mathcal{E}_\beta}{\partial \mathbf{v}} = \mathbf{Z}^T \frac{\partial \mathcal{E}_\beta}{\partial \boldsymbol{\sigma}} = \mathbf{Z}^T \text{diag} \left(\frac{\partial \mathcal{E}_\beta}{\partial \boldsymbol{\sigma}_t} \right) \quad (4.66)$$

with

$$\frac{\partial \mathcal{E}_\beta(\boldsymbol{\varepsilon}, \boldsymbol{\sigma})}{\partial \boldsymbol{\varepsilon}_t} = -2V_t \mathbb{C}_t (\mathbb{E}_D[\boldsymbol{\varepsilon}'_t] - \boldsymbol{\varepsilon}_t) \quad (4.67)$$

$$\frac{\partial \mathcal{E}_\beta(\boldsymbol{\varepsilon}, \boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}_t} = -2V_t \mathbb{C}_t^{-1} (\mathbb{E}_D[\boldsymbol{\sigma}'_t] - \boldsymbol{\sigma}_t). \quad (4.68)$$

Similarly, the Hessian is given by

$$\mathbf{H}_u(\mathbf{u}, \mathbf{v}) = \mathbf{B}^T \text{diag} \left(\frac{\partial^2 \mathcal{E}_\beta}{\partial \boldsymbol{\varepsilon}_t^2} \right) \mathbf{B} \quad (4.69)$$

$$\mathbf{H}_v(\mathbf{u}, \mathbf{v}) = \mathbf{Z}^T \text{diag} \left(\frac{\partial^2 \mathcal{E}_\beta}{\partial \boldsymbol{\sigma}_t^2} \right) \mathbf{Z} \quad (4.70)$$

with

$$\frac{\partial^2 \mathcal{E}_\beta}{\partial \boldsymbol{\varepsilon}_t^2} = 2V_t \mathbb{C}_t - 4\beta V_t^2 \mathbb{C}_t \left[\sum_{i=1}^{n_D} p_i(\boldsymbol{\varepsilon}_t, \boldsymbol{\sigma}_t) (\boldsymbol{\varepsilon}'_i \otimes \boldsymbol{\varepsilon}'_i) - \mathbb{E}_D[\boldsymbol{\varepsilon}'_t] \otimes \mathbb{E}_D[\boldsymbol{\varepsilon}'_t] \right] \mathbb{C}_t \quad (4.71)$$

$$\frac{\partial^2 \mathcal{E}_\beta}{\partial \boldsymbol{\sigma}_t^2} = 2V_t \mathbb{C}_t^{-1} - 4\beta V_t^2 \mathbb{C}_t^{-1} \left[\sum_{i=1}^{n_D} p_i(\boldsymbol{\varepsilon}_t, \boldsymbol{\sigma}_t) (\boldsymbol{\sigma}'_i \otimes \boldsymbol{\sigma}'_i) - \mathbb{E}_D[\boldsymbol{\sigma}'_t] \otimes \mathbb{E}_D[\boldsymbol{\sigma}'_t] \right] \mathbb{C}_t^{-1}. \quad (4.72)$$

Thus, the Hessian matrices consist of the constant terms

$$\bar{\mathbf{H}}_u = 2\mathbf{B}^T \mathbf{V} \mathbb{C} \mathbf{B} \quad (4.73)$$

$$\bar{\mathbf{H}}_v = 2\mathbf{Z}^T \mathbf{V} \mathbb{C}^{-1} \mathbf{Z} \quad (4.74)$$

and a remaining term capturing the covariance around an equilibrium state $(\mathbf{u}, \mathbf{v})$. Thus, we can write

$$\mathbf{H}_u(\mathbf{u}, \mathbf{v}) = \bar{\mathbf{H}}_u - \beta \text{Cov}_u(\mathbf{u}, \mathbf{v}) \quad (4.75)$$

$$\mathbf{H}_v(\mathbf{u}, \mathbf{v}) = \bar{\mathbf{H}}_v - \beta \text{Cov}_v(\mathbf{u}, \mathbf{v}). \quad (4.76)$$

The availability of the Jacobian and Hessian with respect to the degrees of freedom $\mathbf{u}$ and $\mathbf{v}$ allows us to use standard solver schemes like the Newton-Raphson scheme to solve Problem (4.64) as shown in Alg. (11). However, for sufficiently large β , the Hessian is not positive semidefinite in general. Thus, the problem is non-convex and the solution of a minimization scheme depends on the initial values. Therefore, one should also use a *modified* Newton-Raphson scheme which involves only the constant parts $\bar{\mathbf{H}}_u$ and $\bar{\mathbf{H}}_v$ in order to guarantee positive definiteness. Interestingly, using only the constant part of the Hessian, the Newton-Raphson update step $-\bar{\mathbf{H}}^{-1} \mathbf{J}(\mathbf{u}, \mathbf{v})$ resembles exactly one fixed point iteration of data search and projection as shown numerically in the next example.

Algorithm 11 (Modified) Newton-Raphson Solver

Require: Geometry B, Z
Require: Load vectors ε_0, σ_0
Require: Regularization parameter β
Require: Initial solution (u, v)
 Factorize $\bar{H}_u = L_u L_u^T, \bar{H}_v = L_v L_v^T$
while not converged do
 $u \leftarrow u - \bar{H}_u^{-1} J_u(u, v)$
 $v \leftarrow v - \bar{H}_v^{-1} J_v(u, v)$
end while
 $(\varepsilon, \sigma) \leftarrow (Bu + \varepsilon_0, Zv + \sigma_0)$
return $(u, v), (\varepsilon, \sigma)$

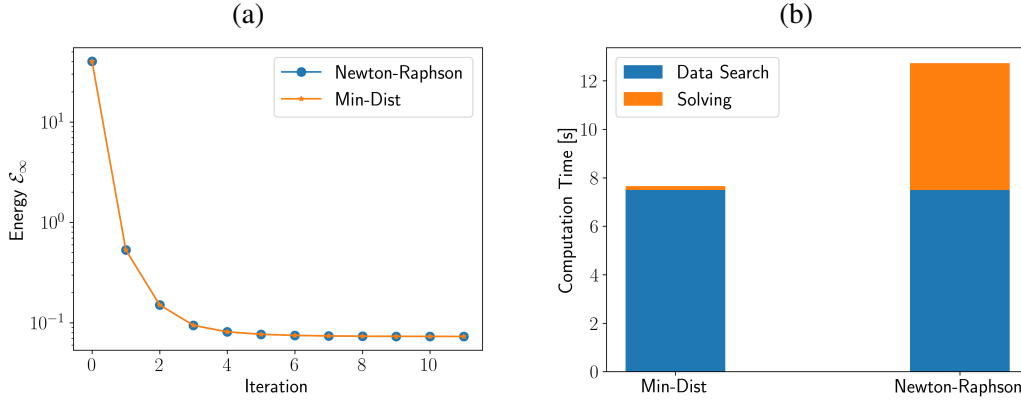
4.4.2.1 Numerical example: Newton-Raphson and fixed-point iteration

We compare the proposed Newton-Raphson solver with the original fixed-point iteration solver numerically with the boundary value defined in section (4.3.8.1). We consider the 'min-dist' limit with $\beta \rightarrow \infty$. In this case, the covariance terms in the Hessian vanishes, thus, the constant approximation of the Hessian becomes exact. At the top nodes, a (dimensionless) displacement of $\bar{u} = (0, 1, 0)$ is prescribed. We assume a linear elastic material behavior with elastic stiffness matrix C and sample 99770 data points normally distributed around the reference solution. Furthermore, the metric coefficients are chosen to $\mathbb{C}_t = K$.

In Fig. 4.9 the convergence of the energy and computation times for both solver schemes are reported. Remarkably, both solvers show exactly the same convergence after starting at the same, randomly chosen, equilibrium state. The solving time of approximately 12 seconds is still satisfactory given the size of the boundary value problem and data set.

However, for this example, the solving time and the memory requirements of the Newton-Raphson solver are significantly higher than the original fixed-point solver [82] since a bigger system of equation is solved. Thus, the fixed-point iteration is generally preferable if available. However, the Newton-Raphson solver might be a better choice for systems with a very low degree of static indeterminacy, or in situations when a general purpose, or 'black-box' solver is required for specific implementations.

Figure 4.9: Numerical comparison of the Newton-Raphson algorithm and the original fixed-point iteration 'min-dist' solver [82]. A) Convergence of the energy B) Computational times.



4.4.3 Hierarchical solution set: deterministic annealing

As previously studied in [83, 134], the energy landscape $\mathcal{E}_\beta(\mathbf{u}, \mathbf{v})$ becomes usually very rough for increasing values of β , i.e. forming multiple local minima. Moreover, these minima can occur at different scales of β . At low temperatures (large β), individual data points start to resolve, causing many local minima scattered around each data point. However, local minima can also occur at higher temperatures due to structural effects and well separated clusters within local data sets. This motivates us to develop an algorithm which can systematically identify all, or at least multiple, local minima at each temperature.

Following the framework of Deterministic Annealing [144, 143], the key insight is that local minima occur if an eigenvalue of the Hessian matrix becomes zero. Furthermore, the associated eigenvector carries the information in which direction the local minima are separated. In physical analogy, this process is also called a phase transformation at a critical (inverse) temperature β_{crit} .

Since the Hessian matrix of $\mathcal{E}_\beta$ consists of two separated matrices $\mathbf{H}_u$ and $\mathbf{H}_v$ (Eqs. 4.75 and 4.76), we need to solve two generalized eigenvalue problems in the form of

$$\det(\bar{\mathbf{H}}_k - \beta \text{Cov}_k) = 0 \quad (4.77)$$

with $k \in \{u, v\}$. Using the Cholesky factorization of the constant term $\bar{\mathbf{H}}_k = \mathbf{L}_k \mathbf{L}_k^T$, this is equivalent to the standard eigenvalue problem

$$\det(\mathbf{I} - \beta \mathbf{L}_k^{-1} \text{Cov}_k \mathbf{L}_k^{-T}) = 0. \quad (4.78)$$

This determines the lowest value of β at which a phase transformation occurs as

$$\beta_{\text{crit}} = \min \left(\frac{1}{\lambda_{u,\max}}, \frac{1}{\lambda_{v,\max}} \right) \quad (4.79)$$

where $\lambda_{k,\max}$ is the largest eigenvalue of $\mathbf{L}_k^{-1} \text{Cov}_k(\mathbf{u}, \mathbf{v}) \mathbf{L}_k^{-T}$.

Furthermore, the direction of the split is given by

$$\mathbf{r}_u = \begin{cases} \mathbf{0}, & \lambda_{u,\max} < \lambda_{v,\max} \\ \mathbf{L}_u^{-T} \mathbf{w}_{u,\max}, & \text{else} \end{cases} \quad (4.80)$$

$$\mathbf{r}_v = \begin{cases} \mathbf{0}, & \lambda_{v,\max} < \lambda_{u,\max} \\ \mathbf{L}_v^{-T} \mathbf{w}_{v,\max}, & \text{else} \end{cases} \quad (4.81)$$

where $\mathbf{w}_{k,\max}$ is the eigenvector associated to the largest eigenvalue $\lambda_{k,\max}$. We remark that the eigenvalue problem is defined on the Hessian with respect to $\mathbf{u}$ and $\mathbf{v}$, i.e. directly in the space of kinematical compatible strains and equilibrium stress fields. In particular, this step cannot be replaced by a projection type operation, making the availability of the basis $\mathbf{Z}$ crucial.

From a practical point of view, one can note that solving for the largest eigenvalue of a symmetric matrix can be efficiently performed without forming the full matrix, for instance by Lanczos algorithm [25]. Nevertheless, the sparsity of $\mathbf{B}$ and $\mathbf{Z}$ and the partitioned Cholesky factorization (Sec. 4.3.9) is important for a good performance.

Since the covariance matrix itself depends on β , Eq. (4.79) is an implicit equation. Thus, the critical temperatures cannot be directly computed, but has to be estimated by an annealing schedule which slowly increases β . In each iteration, or quench, every solution state $(\mathbf{u}, \mathbf{v}) \in \mathcal{S}$ of the current solution set $\mathcal{S}$ is updated by a deterministic minimization scheme at the current temperature. For this step, either the fixed-point iteration [83] or the Newton-Raphson algorithm may be used. Then, the maximum eigenvalue problem is solved and checked if the actual value of β exceeds β_{crit} . If this is the case, the solution

is duplicated and perturbed along the principal axis as $(\mathbf{u} \pm \gamma_u \mathbf{r}_u, \mathbf{v} \pm \gamma_v \mathbf{r}_v)$ by a certain perturbation amplification γ_u and γ_v . We note that the γ parameters should be chosen sufficiently large to ensure that the new solutions do not converge again into the same local minima. The final algorithm is summarized in Alg. (12).

Algorithm 12 Deterministic Annealing

Require: Geometry $\mathbf{B}, \mathbf{Z}$

Require: Load vectors $\boldsymbol{\varepsilon}_0, \boldsymbol{\sigma}_0$

Require: Number of quenches N_Q , maximum number of solutions $K_{\max}$

Require: Initial β_0 and final $\beta_{\max}$

Require: Perturbation parameters γ_u, γ_v

$\mathcal{S} \leftarrow \{(\mathbf{u}_0, \mathbf{v}_0)\}$ // random initialization of solution set

$\alpha \leftarrow (\beta_{\max}/\beta_0)^{1/N_Q}$

$\beta \leftarrow \beta_0$

$K \leftarrow 1$

for $q = 1 \dots N_Q$ **do**

for $i = 1 \dots |\mathcal{S}|$ **do**

$\mathcal{S}_i \leftarrow \arg \min \mathcal{E}_\beta(\mathcal{S}_i)$ // Newton-Raphson or fixed-point iteration

$\mathcal{S}' \leftarrow \{\}$

if $K < K_{\max}$ **then**

$(\mathbf{u}, \mathbf{v}) \leftarrow \mathcal{S}_i$

$\beta_{\text{crit}}, \mathbf{r}_u, \mathbf{r}_v \leftarrow \text{MaxEvProblem}(\mathbf{u}, \mathbf{v}, \beta)$ // Eqs. (4.79-4.81)

if $\beta \geq \beta_{\text{crit}}$ **then**

$\mathcal{S}_i \leftarrow (\mathbf{u} - \gamma_u \mathbf{r}_u, \mathbf{v} - \gamma_v \mathbf{r}_v)$

$\mathcal{S}' \leftarrow \mathcal{S}' \cup (\mathbf{u} + \gamma_u \mathbf{r}_u, \mathbf{v} + \gamma_v \mathbf{r}_v)$

$K \leftarrow K + 1$

end if

end if

end for

$\mathcal{S} \leftarrow \mathcal{S} \cup \mathcal{S}'$

$\beta \leftarrow \alpha \beta$

end for

$\beta \leftarrow \beta_{\max}$

for $i = 1 \dots K$ **do**

$\mathcal{S}_i \leftarrow \arg \min \mathcal{E}_\beta(\mathcal{S}_i)$

end for

return $\mathcal{S}$

The algorithm is, as the name suggests, purely deterministic, which constitutes a key

difference from the algorithm presented in the previous work [134]. In that work, a probabilistic Monte Carlo sampling scheme based on the related concept of simulated annealing was employed. Therein, random moves are proposed, and the inverse temperature β controls the probability of accepting moves with increasing energy. Notably, the proposed parametrization of equilibrium stress fields allows for the implementation of more efficient, gradient-based Monte Carlo sampling schemes, such as Hamiltonian Monte Carlo [123]. However, our tests showed that the method of deterministic annealing proved more efficient in finding low-energy states within complex, multi-modal energy landscapes.

4.4.3.1 Numerical example: probabilistic failure of a plate with holes

We revisit the simple, 1-dimensional probabilistic failure example studied in [134] in a more complex, 3-dimensional setting. We consider the boundary value problem shown in Fig. 4.11 A) and assume a linear elastic material behavior with a Weibull-distributed Drucker-Prager stress tensile strength as

$$\sigma_{\text{DP}} = \frac{1}{2}(1 - \sigma_t/\sigma_c) \text{tr}(\boldsymbol{\sigma}) + \frac{1}{2}(1 + \sigma_t/\sigma_c) \sqrt{\frac{1}{2} \boldsymbol{\sigma}_D : \boldsymbol{\sigma}_D} \quad (4.82)$$

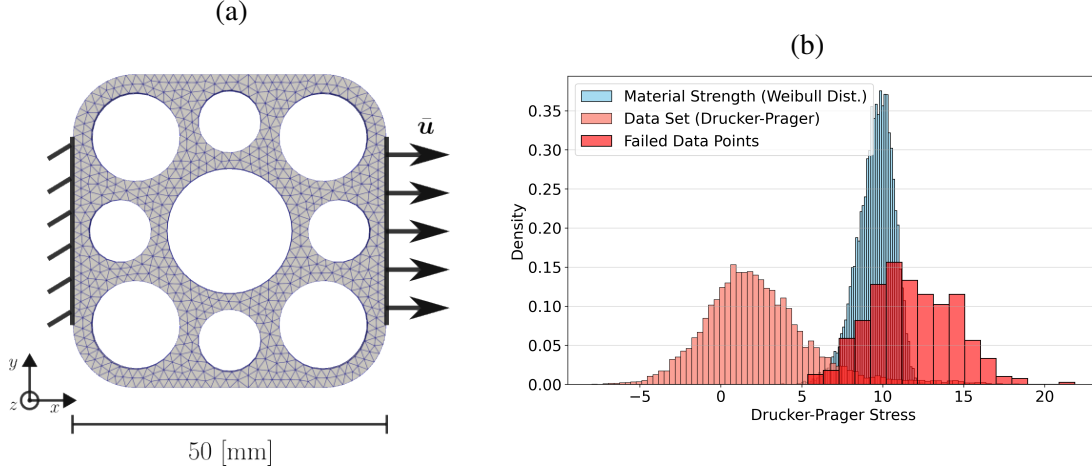
$$P_{\text{fail}}(\sigma_{\text{DP}}) = 1 - e^{-(\sigma_{\text{DP}}/\sigma_t)^k} \quad (4.83)$$

where $\boldsymbol{\sigma}_D$ is the stress deviator $\boldsymbol{\sigma} - \frac{1}{3} \text{tr}(\boldsymbol{\sigma}) \mathbf{I}$ and 'tr' denotes the trace of a matrix.

Specifically, the artificial data set was created using the following procedure: 1) Compute the linear elastic solution of the boundary value problem, giving $|T|$ data points. 2) Add a normally distributed noise term to each strain state. 3) Compute the stress components by linear elasticity. 4) Draw $|T|$ Weibull-distributed tensile strengths. 5) For each data point, set the stress components to zero if its Drucker-Prager stress exceeds the tensile strength. 4) Optional: repeat the process.

The following parameters were chosen: elastic parameters $E = 100$ and $\nu = 0.3$; tensile and compression strength $\sigma_t = 10$ and $\sigma_c = 100$ (positive per definition); Weibull shape parameter $k = 10$; noise term with standard deviation $s = 0.01$. In total, 3 iterations of data generations were performed, resulting in $3|T| = 10563$ data points. The resulting distribution of Drucker-Prager stresses of the data set and the sampled tensile strengths are visualized in Fig. 4.11 B).

Figure 4.11: A) Plate with holes boundary value problem. B) Drucker-Prager stress distribution of the local data set. The failure probability is Weibull distributed.



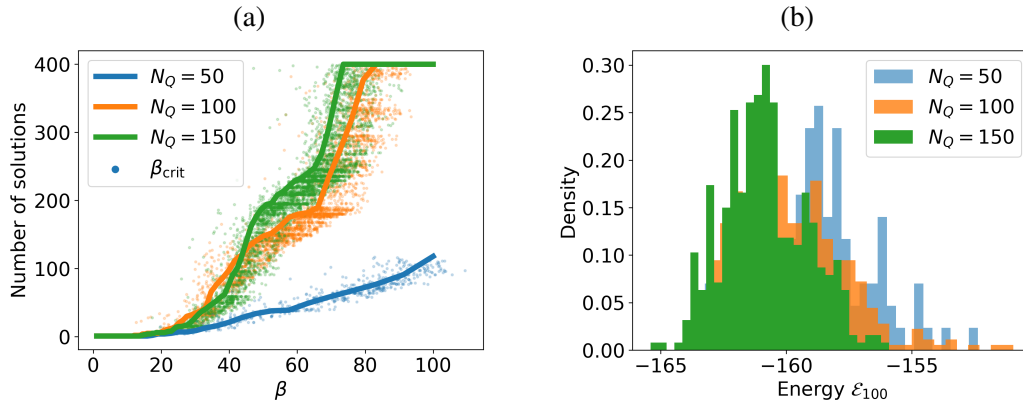
The sparse basis for equilibrium stress fields has been computed using the algorithms from Part 1 of this paper with the general Algorithm (7) since the domain contains holes. The total dimension of $\mathbf{Z}$ is 16896, containing 14619 facet-wise and 1977 edge-wise null vectors as well as 252 null vectors from Dirichlet boundary conditions and 48 null vectors from the holes.

First, we study the influence of the annealing schedule. Following [134], we estimate the order of magnitude of $\beta_{\max}$ with $\beta_{\max}^{-1}$ proportional to the average nearest neighbor distance of each data point in the data set times the average material point volume size. Based on this, we choose conservatively $\beta_{\max} = 100$. We further choose the perturbation parameters $\gamma_u = 10|\mathbf{u}|/|\mathbf{r}_u|$ and $\gamma_v = 10|\mathbf{v}|/|\mathbf{r}_v|$.

In Fig. 4.13 A), we show the evolution of the number of solutions for different number of quenches $N_Q \in \{50, 100, 150\}$ with a maximum number of allowed solution $K_{\max} = 400$. In addition, we show for each solution its critical inverse temperature β_{crit} in each iteration. Thus, a new solution is generated whenever a point β_{crit} is below the actual β of the iteration (that is, on the left side of the line). If the annealing speed is too fast ($N_Q = 50$, blue line), many possible phase transitions are not detected and the solution space is not sufficiently explored. For slower annealing speeds ($N_Q \in \{100, 150\}$), we observe a more exhaustive development of solutions. Furthermore, we can observe two distinct regions of phase transformations: the first one for approximately

$\beta \in [20, 60]$ where phase transitions occur at a relatively high temperature and the second one for lower temperatures $\beta > 70$ with an exponential increase in solutions resolving finer details of the data set.

Figure 4.13: A) Evolution of the number of solutions during the annealing procedure with critical temperatures for different annealing rates. The maximum number of solution is set to 400. B) Energy distribution of the final solution set for different annealing rates.



The following computation times for the different number of quenches were recorded in a Python implementation: 32 min ($N_Q = 50$), 103 min ($N_Q = 100$), 134 min ($N_Q = 150$). The critical operations are the energy, gradient and Hessian computations, since the summation is over all data points in all material points. We have used an NVIDIA GeForce RTX 4090 to perform the embarrassingly parallel summation in a brute force manner. For larger data sets, it may be beneficial to use data search algorithms to include only relevant data points in the summation, as previously studied in [34, 134]. Furthermore, the algorithm can benefit from a parallel implementation as each solution evolves independently from each other.

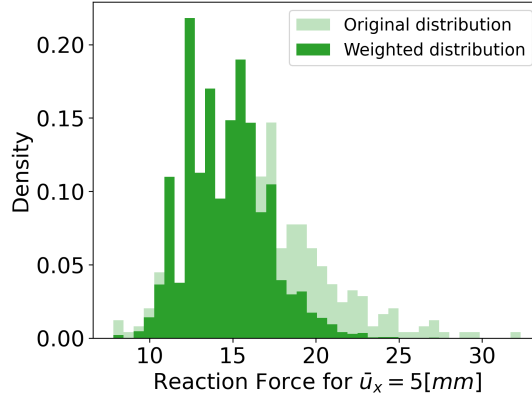
In Fig. 4.13 B), the energy distributions of the solution sets at the end of the annealing schedule for the three different annealing speeds are shown. The important conclusion is that all solutions have an energy in the same order of magnitude, thus there is not a single solution as an unique minimizer. Nevertheless, the energy of every solution can be directly used to quantify how close the equilibrium solution state $(\mathbf{u}, \mathbf{v})$ is to the given data set D . Using the histogram re-weighting method [36], the information about the relative energy differences can also be used to evaluate the results in an energy-weighted

manner by computing the weights

$$w_k = e^{-\Delta\beta(\mathcal{E}_{\beta,k} - \mathcal{E}_{\beta,\min})} \quad (4.84)$$

where $\Delta\beta$ is a factor which can be used to control how strictly energy differences should be handled relative to the solution with minimum energy $\mathcal{E}_{\beta,\min}$. In Fig. 4.15, the reaction force distribution of the solution set after $N_Q = 150$ quenches is shown unweighted (light green) as well weighted by the energy differences according to Eq. (4.84) with $\Delta\beta = 1$. The reaction forces are computed by summing $(\mathbf{B}^{\text{full}})^T \boldsymbol{\sigma}$ at the degrees of freedom in x -direction at the clamped node set, see Fig. 4.11 A).

Figure 4.15: Original and energy-weighted reaction force distribution of 400 solutions at a given displacement loading step after $N_Q = 150$ quenches.

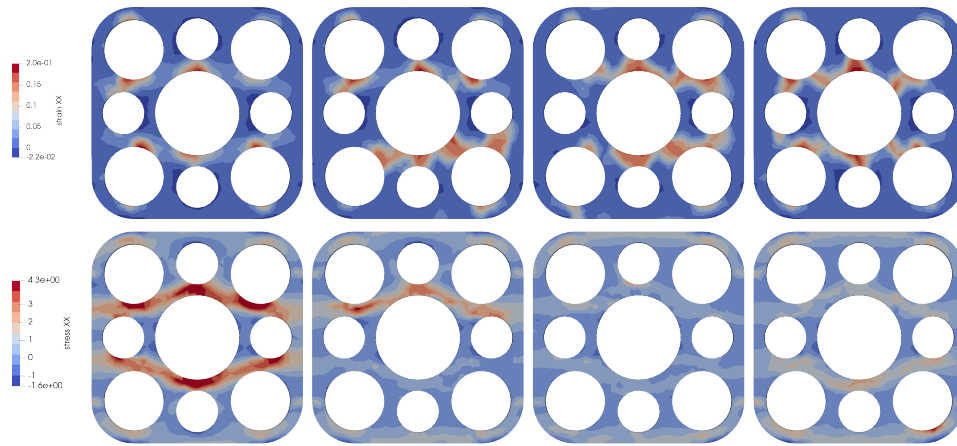


Finally, we show in Fig. 4.16 the strain and stress fields in x -direction for a selection of solutions after $N_Q = 150$ iterations. We emphasize that the boundary value problem is for every solution the same, but due to the probabilistic tensile strength and thus noisy data set, no unique stress-strain relationship exists resulting in a large variety of possible outcomes.

The visualization shows that every solution represents one possible crack pattern, giving important information for uncertainty quantification and design decisions. Even though the local data set is relatively simple, the possible outcomes can become, due to the global equilibrium and compatibility constraints, quite complex. We argue that in such cases with non-trivial variability in the stresses, the consistent parametrization of the space of equilibrium stress fields is particularly useful since the eigenvalue analysis, and

thus the solution space exploration, can include all information about data, compatibility and equilibrium.

Figure 4.16: A selection of solutions for the same boundary value problem highlighting the different failure possibilities. Top row: strain ε_{xx} fields, bottom row: stress σ_{xx} fields. From left to right: Large reaction force (≈ 25 [N]), moderate reaction force (≈ 12 [N]), low reaction force (≈ 8 [N]), minimum energy solution.



4.5 Conclusions

In this paper, we have developed a novel class of solvers for direct data-driven mechanics using a sparse basis for equilibrium stress fields. First, we developed an algorithm capable of computing a sparse null basis of the equilibrium matrix for large three-dimensional problems within seconds. A detailed analysis of the different components of the null basis enabled us to derive linearly independent generator mechanisms based on self-equilibrium force fields. Furthermore, a partitioned Cholesky factorization was presented to efficiently process the small fraction of dense columns related to indeterminate supports and holes.

Based on this foundation, we have demonstrated how to use standard solvers, such as a Newton-Raphson algorithm, to solve the data-driven problem. However, except

for system with a low degree of static indeterminacy, the approach is computationally less efficient than the original fixed point solver. Nevertheless, it might be useful for situations where the implementation requires the use of an unconstrained solver. Another potential application might be situations when no system of equations should be solved, since any gradient descent scheme can be directly used.

We conclude that the true strength of this approach lies in problems with complex and non-trivial data sets requiring an consistent exploration of the space of equilibrium stress fields. The characteristic of such problems are rough, multi-modal energy landscapes which are difficult to explore. As an example, we have presented a brittle fracture problem with probabilistic failure strength. An eigenvalue analysis in the space of equilibrium stress fields and compatible strain fields enabled an effective exploration of the complex solution space. As a result, different possible failure modes were discovered giving important information for uncertainty quantification and design processes.

This paper introduces several research questions and possibilities for future work. First, an interesting question from the field of algebraic topology [121] is how to analyze the topological homologies to precisely determine the number of null vectors associated with holes. Furthermore, as an alternative to the proposed practical approach of using dense columns for the supports and holes, developing techniques to find sparse representations could be considered.

A further important open research question concerns the locking behavior of the proposed method. Standard tetrahedron element formulations that utilize linear displacement interpolation often exhibit artificial stiffness when approaching the incompressible limit. However, as noted by [124], the data-driven formulation bears a close relation to mixed finite element formulations, which are generally known to demonstrate much more favorable locking behaviors. Consequently, further investigations are required to fully characterize this aspect. Additionally, it would be interesting to investigate whether the presented concepts could be adapted to other element types, such as hexahedral elements.

The approach of computing a solution set using deterministic annealing can be applied to different data distributions and boundary value problems. To achieve this, developing robust choices for parameters like annealing speed, maximum solution set size, and perturbation sizes is crucial. Further research is required to develop efficient data sampling strategies that cover all relevant states of deformation. This is particularly

important for path-dependent materials, for which the response depends on the loading trajectory [136]. We believe that many applications, particularly safety-critical ones, benefit from the much deeper insights provided by the proposed approach.

4.6 Acknowledgements

EP and SR gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG) through the project "Multiscale modeling based on a novel combination of direct data-driven methods with Fourier transform-based microstructure simulation" (RE 1057/58-1, project number 532163998). MO gratefully acknowledges the financial support of the Centre Internacional de Mètodes Numèrics a l'Enginyeria (CIMNE) of the *Universitat Politecnica de Catalunya* (UPC), Spain, through the *UNESCO* Chair in Numerical Methods in Engineering.

5 | Conclusions and Outlook

The landscape of computational mechanics is undergoing one of its most significant transformations. By transitioning from a paradigm of phenomenological material modeling to data-driven approaches, the simulations are driven by the quality and amount of available material data, rather than by specific modeling choices. This point of view becomes particularly clear through the set-based data-driven method. This cumulative dissertation has contributed several new methods to this branch of research:

First, an efficient randomized fixed-point algorithm was developed which can compute the global minimum of the distance-minimizing problem with almost no computational overhead compared to the original fixed-point solver. This enabled the development of a convergent goal-oriented data acquisition scheme for path-dependent materials in the context of multiscale mechanics. It was shown that even for a single boundary value problem a significant number of RVE evaluations could be avoided. The efficiency of both methods was also demonstrated in follow-up work [151] for a sequence of boundary value problems.

At the core of this work, a mathematical framework for stochastic inference was developed. A rigorous treatment of intrinsically random materials and noisy experimental datasets, made it possible for the first time to consistently compute all non-unique solutions of complex, stochastic mechanical problems. This aspect is one of the biggest strengths of the set-based data-driven method, as it is not limited to single-valued functions. While this may seem like a minor detail at the local material level, it can have huge consequences on the overall global system response as the material points are coupled with each other by compatibility and equilibrium.

Finally, efficient solver algorithms have been developed to make the stochastic framework applicable to realistic three-dimensional fracture problems. Special care has been taken for the equilibrium constraint by computing a sparse basis for equilibrium stress fields for tetrahedral elements. By an eigenvalue decomposition in the space of equi-

librium stress fields, the multitude of solutions can be consistently discovered. The resulting solution set of the data-driven analysis provides deep insights into the problem. Importantly, it is directly possible to quantify how close each solution is to the measured data and thus, how trustworthy the result is.

The new contributions of this dissertation open various future research possibilities. While the goal-oriented data acquisition scheme is already designed for path-dependent materials, further research is required to robustly parameterize the deformation history, in particular for complex, mixed plastic hardening behaviors. One can expect that the computational gain of this approach is maximal for problems with an expensive material model evaluation (such as in inelastic multiscale simulations) and for multiple similar boundary value problems (such as in microstructure optimization). In addition, the data-driven identification for inelastic materials is still challenging and requires further research.

The application of a basis for equilibrium stress fields in the field of material property optimization [64] is well known. The proposed, efficient algorithm to compute the sparse basis for three dimensional problems is an alternative to classical turnback-style algorithms and can support new methods in this field. In addition, it may be used to compute the number of data points obtainable from data-driven identification [60], and more broadly in the field of topology optimization [6].

The proposed algorithm to compute the full set of all possible system outcomes can be applied in various other fields, for instance transport problems, buckling or non-linear elasticity. Further research is in particular required for the set-based data-driven method for large deformations, as many works including this dissertation are limited to small deformations. However, many concepts such as the basis for equilibrium stress fields are transferable, especially if a nominal formulation in terms of the first Piola-Kirchhoff stress tensor is chosen.

Finally, it remains to emphasize that one particular strength of the set-based data-driven method is the simplicity of the data collection and inference workflow cycle. Since no intermediate modeling step, which might require further verification steps, is required, the process can be easily automated. Furthermore, the method is extremely robust with noisy data sets, notably because no material tangent is required for solving. These properties are in particular beneficial in data-rich scenarios, for instance with data streams from multiple, different sources and continuous inference tasks. Future works

can explore such setups for shared data sets and automated predictions.

List of Figures

2.1	Stamp boundary value problem.	22
2.2	A) Influence of the metric parameter on the original and randomized algorithm. Shaded areas denote the maximum deviations from the mean of 30 independent runs. B) Error in constraints during the iterations for both algorithms for $c = 1000$. C) Blue: Local data set, orange: optimal data point assignment of each material point.	24
2.3	Workflow for the generation of the FFT model using SliceGAN (https://microlib.io) [80, 81]: 1) real image of a microstructure 2) generated 3D microstructure from a trained GAN 3) downsampled FFT-model with 32^3 pixels showing a strain field.	29
2.4	Beam boundary value problem.	31
2.5	A) 99-percentile error for the iteratively enriched data sets using different cluster ratios. B) Force-Displacement curve for different iterations of data generation. Colored by the logarithm of the 99-percentile error. . .	32
2.6	A) CPU times for the different cluster ratios. B) CPU times per iteration for the increasing cluster ratio scheme until a accuracy of $F_{99\%} < 10^{-4}$ is reached.	33
2.7	Comparison of the original and the newly proposed randomized fixed-point solver schemes for the increasing cluster ratio. In addition to the 99-percentile error, the maximum error is shown.	34
3.1	Brittle material with random tensile strength, member e of the structure. a) Local material likelihood function $L_{D,e}$ providing the likelihood that the local material state (ϵ_e, σ_e) be on the unfailed and failed branches of the local material set, respectively. b) Empirical local likelihood measure $\mu_{D,e}$ sampled from $L_{D,e}$	39

3.2	a) Truss with prescribed applied force P undergoing a displacement Δ . b) Local material data set with 1000 data points sampled from a sliding-Gaussian likelihood function.	58
3.3	Three-bar truss with sliding-Gaussian material data. Computed histograms vs. exact distribution (red) of displacement Δ for material data-sets of sizes: a) $M_e = 10^3$; b) $M_e = 10^4$; c) $M_e = 10^5$	59
3.4	Three-bar truss with sliding-Gaussian material data. Convergence of temperatures and computed histograms of displacement Δ . a) Evolution of β_e during annealing process. b) Kolmogorov–Smirnov (KS) error vs. material data-set size, $N_P = 10^4$	60
3.5	Three-bar truss with sliding-Gaussian material data, $M_e = 10^5$. a) KS error vs. population size, $N_T = 20$ and $N_Q = 100$. b) KS error vs. number of quenches N_Q for different number of trials N_T using $N_P = 5000$. The lines show the mean of 20 repeated computations, the shaded area indicate the maximum deviation from mean.	61
3.6	a) Truss with prescribed displacement Δ and reaction force P . b) Local material data set with 1000 data points sampled from a bimodal likelihood function with failed and unfailed branches selected according to Weibull strength statistics, plus superimposed Gaussian noise.	62
3.7	Three-bar truss with Weibull tensile strength data. a) Computed mean and standard deviation of the reaction force P for prescribed displacement Δ . b) Computed distribution of P at $\Delta = 2 \times 10^{-2}$	63
3.8	Lightweight space structure [101]. The red arrow indicates the quantity of interest (joint eccentricity u_e) to be inferred from the analysis.	64
3.9	Lightweight space structure [101]. Computed posterior distributions of the joint eccentricity u_e and exact reference distribution (red curves). a) Posterior distribution obtained from an initial population of 1000 random min-dist solutions. b) Posterior distribution following application of the population annealing algorithm.	65

3.10	Lightweight space structure [101]. Influence of the number of checks for backtracking approximate in the nearest-neighbor search. a) CPU times in seconds on a 12-Core AMD Ryzen 9 3900X machine. b) Error e_{KS} of node eccentricity u_e with error bars derived from a sample of 25 calculations.	66
4.1	Tetrahedron t . The set notation $(t \setminus n)$ denotes the facet of t opposite to node n with area size $A_{t \setminus n}$ and outward unit normal $\mathbf{n}_{t \setminus n}$	74
4.2	a) Facets adjacent to an internal edge. b) Pair of equal but opposite directing forces along a shared edge of two adjacent tetrahedrons.	79
4.3	Edge forces of an basic module around an internal edge e . The outer nodes form a cycle $(n_1, n_2, n_3, n_4, n_1)$ of length $K = 4$	82
4.4	A) Minimal example with 6 tetrahedrons, 1 internal node (labeled 0), 5 internal edges and 9 internal facets. B) Sparsity pattern of the null base $\mathbf{Z}$	83
4.6	A) Benchmark example boundary value problem. B) Sparsity pattern of the null base. For visualization, the edge-wise and facet-wise parts are plotted separately.	88
4.8	Partitioned Cholesky factorization $\mathbf{LL}^T = \mathbf{H}$. Light blue and dark blue blocks are sparse and dense matrices, respectively.	89
4.9	Numerical comparison of the Newton-Raphson algorithm and the original fixed-point iteration 'min-dist' solver [82]. A) Convergence of the energy B) Computational times.	96
4.11	A) Plate with holes boundary value problem. B) Drucker-Prager stress distribution of the local data set. The failure probability is Weibull distributed.	100
4.13	A) Evolution of the number of solutions during the annealing procedure with critical temperatures for different annealing rates. The maximum number of solution is set to 400. B) Energy distribution of the final solution set for different annealing rates.	101
4.15	Original and energy-weighted reaction force distribution of 400 solutions at a given displacement loading step after $N_Q = 150$ quenches.	102

- 4.16 A selection of solutions for the same boundary value problem highlighting the different failure possibilities. Top row: strain ε_{xx} fields, bottom row: stress σ_{xx} fields. From left to right: Large reaction force (≈ 25 [N]), moderate reaction force (≈ 12 [N]), low reaction force (≈ 8 [N]), minimum energy solution. 103

Bibliography

- [1] Fadi Aldakheel, Elsayed S Elsayed, Tarek I Zohdi, and Peter Wriggers. Efficient multiscale modeling of heterogeneous materials using deep neural networks. *Computational Mechanics*, 72(1):155–171, 2023.
- [2] Omar Aloui, Jessica Flores, David Orden, and Landolf Rhode-Barbarigos. Cellular morphogenesis of three-dimensional tensegrity structures. *Computer methods in applied mechanics and engineering*, 346:85–108, 2019.
- [3] Anish Arora and Mohamed Gouda. Closure and convergence: A foundation of fault-tolerant computing. *IEEE Transactions on Software Engineering*, 19(11):1015–1027, 1993.
- [4] Faisal As’ad, Philip Avery, and Charbel Farhat. A mechanics-informed artificial neural network approach in data-driven constitutive modeling. *International Journal for Numerical Methods in Engineering*, 123(12):2738–2759, 2022.
- [5] M. G. Bader, K. L. Pickering, A. Buxton, A. Rezaifard, and P. A. Smith. Failure micromechanisms in continuous carbon-fibre/epoxy-resin composites. *Composites Science and Technology*, 48:135–142, 1993.
- [6] William F Baker, Lauren L Beghini, Arkadiusz Mazurek, Juan Carrion, and Alessandro Beghini. Maxwell’s reciprocal diagrams and discrete michell frames. *Structural and multidisciplinary optimization*, 48(2):267–277, 2013.
- [7] Lev Yu Barash, Martin Weigel, Michal Borovský, Wolfhard Janke, and Lev N Shchur. Gpu accelerated population annealing algorithm. *Computer Physics Communications*, 220:341–350, 2017.

-
- [8] Igor A Baratta, Joseph P Dean, Jørgen S Dokken, Michal Habera, Jack HALE, Chris N Richardson, Marie E Rognes, Matthew W Scroggs, Nathan Sime, and Garth N Wells. Dolfinx: the next generation fenics problem solving environment. 2023.
 - [9] Thorsten Bartel, Marius Harnisch, Ben Schweizer, and Andreas Menzel. A data-driven approach for plasticity using history surrogates: Theory and application in the context of truss structures. *Computer Methods in Applied Mechanics and Engineering*, 414:116138, 2023.
 - [10] Mohamed Amine Benaïmeche, Julien Yvonnet, Benoit Bary, and Qi-Chang He. A k-means clustering machine learning-based multiscale method for anelastic heterogeneous structures with internal variables. *International Journal for Numerical Methods in Engineering*, 123(9):2012–2041, 2022.
 - [11] S. C. Bennett, D. J. Johnson, and W. Johnson. Strength-structure relationships in pan-based carbon fibres. *Journal of Material Science*, 18:3337–3347, 1983.
 - [12] Dana Bishara, Yuxi Xie, Wing Kam Liu, and Shaofan Li. A state-of-the-art review on machine learning-based multiscale modeling, simulation, homogenization and design of materials: D. bishara et al. *Archives of computational methods in engineering*, 30(1):191–222, 2023.
 - [13] Colin Bonatti and Dirk Mohr. On the importance of self-consistency in recurrent neural network models representing elasto-plastic solids. *Journal of the Mechanics and Physics of Solids*, 158:104697, 2022.
 - [14] A. F. Bower. *Applied Mechanics of Solids*. CRC Press, Boca Raton, Fla., 2010.
 - [15] Ante Buljac, Clément Jailin, Arturo Mendoza, Jan Neggers, Thibault Taillandier-Thomas, Amine Bouterf, Benjamin Smaniotto, François Hild, and Stéphane Roux. Digital volume correlation: review of progress and challenges. *Experimental Mechanics*, 58:661–708, 2018.
 - [16] Pietro Carrara, Laura De Lorenzis, Laurent Stainier, and Michael Ortiz. Data-driven fracture mechanics. *Computer Methods in Applied Mechanics and Engineering*, 372:113390, 2020.

- [17] Rick Chartrand and Wotao Yin. Iteratively reweighted algorithms for compressive sensing. In *2008 IEEE international conference on acoustics, speech and signal processing*, pages 3869–3872. IEEE, 2008.
- [18] Yanqing Chen, Timothy A Davis, William W Hager, and Sivasankaran Rajamanickam. Algorithm 887: Cholmod, supernodal sparse cholesky factorization and update/downdate. *ACM Transactions on Mathematical Software (TOMS)*, 35(3): 1–14, 2008.
- [19] Kerem Ciftci and Klaus Hackl. Model-free data-driven simulation of inelastic materials using structured data sets, tangent space information and transition rules. *Computational Mechanics*, 70(2):425–435, 2022.
- [20] Thomas F Coleman and Alex Pothén. The null space problem i. complexity. *SIAM Journal on Algebraic Discrete Methods*, 7(4):527–537, 1986.
- [21] S. Conti, F. Hoffmann, and M. Ortiz. Model-free data-driven inference. *arXiv.org*, math:arXiv:2106.02728, 2021.
- [22] S. Conti, F. Hoffmann, and M. Ortiz. Ansatz-free data-driven inference in physically constrained problems. *arXiv.org*, math:to be submitted, 2022.
- [23] Sergio Conti, Stefan Müller, and Michael Ortiz. Data-driven problems in elasticity. *Archive for Rational Mechanics and Analysis*, 229(1):79–123, 2018.
- [24] Sergio Conti, Stefan Müller, and Michael Ortiz. Data-driven finite elasticity. *Archive for Rational Mechanics and Analysis*, 237(1):1–33, 2020.
- [25] Jane K Cullum and Ralph A Willoughby. *Lanczos algorithms for large symmetric eigenvalue computations: Vol. I: Theory*. SIAM, 2002.
- [26] Marie Dalémat, Michel Coret, Adrien Leygue, and Erwan Verron. Measuring stress field without constitutive equation. *Mechanics of Materials*, 136:103087, 2019.
- [27] Héloïse Dandin, Adrien Leygue, and Laurent Stainier. Graph-based representation of history-dependent material response in the data-driven computational mechan-

- ics framework. *Computer Methods in Applied Mechanics and Engineering*, 419: 116694, 2024.
- [28] Masoumeh Dashti and Andrew M. Stuart. *The Bayesian Approach to Inverse Problems*, pages 311–428. Springer International Publishing, Cham, 2017.
- [29] Timothy A Davis. Algorithm 915, suitesparseqr: Multifrontal multithreaded rank-revealing sparse qr factorization. *ACM Transactions on Mathematical Software (TOMS)*, 38(1):1–22, 2011.
- [30] Herbert De Gersem, Armin Galetzka, Ion Gabriel Ion, Dimitrios Loukrezis, and Ulrich Römer. Magnetic field simulation with data-driven material modeling. *IEEE Transactions on Magnetics*, 56(8):1–6, 2020.
- [31] MWM Gamini Dissanayake and Nhan Phan-Thien. Neural-network-based approximations for solving partial differential equations. *communications in Numerical Methods in Engineering*, 10(3):195–201, 1994.
- [32] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. Least angle regression. 2004.
- [33] Robert Eggersmann, Trenton Kirchdoerfer, Stefanie Reese, Laurent Stainier, and Michael Ortiz. Model-free data-driven inelasticity. *Computer Methods in Applied Mechanics and Engineering*, 350:81–99, 2019.
- [34] Robert Eggersmann, Laurent Stainier, Michael Ortiz, and Stefanie Reese. Efficient data structures for model-free data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 382:113855, 2021.
- [35] Robert Eggersmann, Laurent Stainier, Michael Ortiz, and Stefanie Reese. Model-free data-driven computational mechanics enhanced by tensor voting. *Computer Methods in Applied Mechanics and Engineering*, 373:113499, 2021.
- [36] Alan M Ferrenberg, David P Landau, and Robert H Swendsen. Statistical errors in histogram reweighting. *Physical Review E*, 51(5):5092, 1995.

- [37] Frédéric Feyel and Jean-Louis Chaboche. Fe2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre sic/ti composite materials. *Computer methods in applied mechanics and engineering*, 183(3-4):309–330, 2000.
- [38] Jacob Fish, Gregory J Wagner, and Sinan Keten. Mesoscopic and multiscale modelling in materials. *Nature materials*, 20(6):774–786, 2021.
- [39] Moritz Flaschel, Siddhant Kumar, and Laura De Lorenzis. Unsupervised discovery of interpretable hyperelastic constitutive laws. *Computer Methods in Applied Mechanics and Engineering*, 381:113852, 2021.
- [40] Moritz Flaschel, Siddhant Kumar, and Laura De Lorenzis. Discovering plasticity models without stress data. *npj Computational Materials*, 8(1):91, 2022.
- [41] Moritz Flaschel, Siddhant Kumar, and Laura De Lorenzis. Automated discovery of generalized standard material models with euclid. *Computer Methods in Applied Mechanics and Engineering*, 405:115867, 2023.
- [42] Baudouin Fraeijs de Veubeke. Upper and lower bounds in matrix structural analysis. *Rapport du LTAS*, 1963.
- [43] Fernando Fraternali, Maurizio Angelillo, and Antonio Fortunato. A lumped stress method for plane elastic problems and the discrete-continuum approximation. *International journal of solids and structures*, 39(25):6211–6240, 2002.
- [44] Jan N Fuhg, Nikolaos Bouklas, and Reese E Jones. Learning hyperelastic anisotropy from data via a tensor basis neural network. *Journal of the Mechanics and Physics of Solids*, 168:105022, 2022.
- [45] Jan N Fuhg, Govinda Anantha Padmanabha, Nikolaos Bouklas, Bahador Bahmani, WaiChing Sun, Nikolaos N Vlassis, Moritz Flaschel, Pietro Carrara, and Laura De Lorenzis. A review on data-driven constitutive laws for solids: In fuhg et al. *Archives of Computational Methods in Engineering*, 32(3):1841–1883, 2025.
- [46] Jan Niklas Fuhg, Christoph Böhm, Nikolaos Bouklas, Amelie Fau, Peter Wriggers, and Michele Marino. Model-data-driven constitutive responses: Application to

- a multiscale computational framework. *International Journal of Engineering Science*, 167:103522, 2021.
- [47] Keinosuke Fukunaga and Patrenahalli M. Narendra. A branch and bound algorithm for computing k-nearest neighbors. *IEEE transactions on computers*, 100(7):750–753, 1975.
- [48] Armin Galetzka, Dimitrios Loukrezis, and Herbert De Gerssem. Data-driven solvers for strongly nonlinear material response. *International Journal for Numerical Methods in Engineering*, 122(6):1538–1562, 2021.
- [49] Marc GD Geers, Varvara G Kouznetsova, and W Brekelmans. Multi-scale computational homogenization: Trends and challenges. *Journal of computational and applied mathematics*, 234(7):2175–2182, 2010.
- [50] Jamshid Ghaboussi, JH Garrett Jr, and Xiping Wu. Knowledge-based modeling of material behavior with neural networks. *Journal of engineering mechanics*, 117(1):132–153, 1991.
- [51] Christian Gierden, Julian Kochmann, Johanna Waimann, Bob Svendsen, and Stefanie Reese. A review of fe-fft-based two-scale methods for computational modeling of microstructure evolution and macroscopic material behavior. *Archives of Computational Methods in Engineering*, 29(6):4115–4135, 2022.
- [52] John R Gilbert and Michael T Heath. Computing a sparse basis for the null space. *SIAM Journal on Algebraic Discrete Methods*, 8(3):446–459, 1987.
- [53] Felix Selim Göküzüm and Marc-André Keip. An algorithmically consistent macroscopic tangent operator for fft-based computational homogenization. *International Journal for Numerical Methods in Engineering*, 113(4):581–600, 2018.
- [54] Anna Gorgogianni, Konstantinos Karapiperis, Laurent Stainier, Michael Ortiz, and José E Andrade. Adaptive goal-oriented data sampling in data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 409:115949, 2023.

- [55] Maysam B Gorji, Mojtaba Mozaffar, Julian N Heidenreich, Jian Cao, and Dirk Mohr. On the potential of recurrent neural networks for modeling path dependent plasticity. *Journal of the Mechanics and Physics of Solids*, 143:103972, 2020.
- [56] Lee-Ad Gottlieb and Tyler Neylon. Matrix sparsification and the sparse null space problem. In *International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 205–218. Springer, 2010.
- [57] Todd L Graves. Automatic step size selection in random walk metropolis algorithms. *arXiv preprint arXiv:1103.5986*, 2011.
- [58] Xu Guo, Zongliang Du, Chang Liu, and Shan Tang. A new uncertainty analysis-based framework for data-driven computational mechanics. *Journal of Applied Mechanics*, 88(11):111003, 2021.
- [59] Márton Hablicsek, Masoud Akbarzadeh, and Yi Guo. Algebraic 3d graphic statics: Reciprocal constructions. *Computer-Aided Design*, 108:30–41, 2019.
- [60] Nour Hachem, Adrien Leygue, and Laurent Stainier. Mathematical and numerical assessment of data-driven identification method applied to nonlinear elasticity. *Computer Methods in Applied Mechanics and Engineering*, 446:118273, 2025.
- [61] Alexander Hartmaier. Data-oriented constitutive modeling of plasticity in metals. *Materials*, 13(7):1600, 2020.
- [62] Ji He, Shengda Jiang, and Cong Guo. A data-driven method for constitutive-model-free multiscale concurrent analysis: theoretical modeling and numerical simulation. *Composite Structures*, page 119830, 2025.
- [63] Qizhi He and Jiun-Shyan Chen. A physics-constrained data-driven approach based on locally convex reconstruction for noisy database. *Computer Methods in Applied Mechanics and Engineering*, 363:112791, 2020.
- [64] Michael T Heath, Robert J Plemmons, and Robert C Ward. Sparse orthogonal schemes for structural optimization using the force method. *SIAM Journal on Scientific and Statistical Computing*, 5(3):514–532, 1984.

- [65] Yousef Heider, Kun Wang, and WaiChing Sun. So (3)-invariance of informed-graph-based deep neural network for anisotropic elastoplastic materials. *Computer Methods in Applied Mechanics and Engineering*, 363:112875, 2020.
- [66] Hagen Holthusen, Lukas Lamm, Tim Brepols, Stefanie Reese, and Ellen Kuhl. Theory and implementation of inelastic constitutive artificial neural networks. *Computer Methods in Applied Mechanics and Engineering*, 428:117063, 2024.
- [67] Dengpeng Huang, Jan Niklas Fuhg, Christian Weißenfels, and Peter Wriggers. A machine learning based plasticity model using proper orthogonal decomposition. *Computer Methods in Applied Mechanics and Engineering*, 365:113008, 2020.
- [68] Yukito Iba. Population monte carlo algorithms. *Transactions of the Japanese Society for Artificial Intelligence*, 16(2):279–286, 2001.
- [69] Ruben Ibañez, Domenico Borzacchiello, Jose Vicente Aguado, Emmanuelle Abisset-Chavanne, Elias Cueto, Pierre Ladeveze, and Francisco Chinesta. Data-driven non-linear elasticity: constitutive manifold construction and problem discretization. *Computational Mechanics*, 60(5):813–826, 2017.
- [70] Wenzel Jakob, Jason Rhinelander, and Dean Moldovan. pybind11—seamless operability between c++ 11 and python, 2017. URL <https://github.com/pybind/pybind11>, 2019.
- [71] Matthias Kabel, Thomas Böhlke, and Matti Schneider. Efficient fixed point and newton–krylov solvers for fft-based homogenization of elasticity at large deformations. *Computational Mechanics*, 54(6):1497–1514, 2014.
- [72] Karl A Kalina, Lennart Linden, Jörg Brummund, and Markus Kästner. Fe ann: an efficient data-driven multiscale approach based on physics-constrained neural networks and automated data mining. *Computational Mechanics*, 71(5):827–851, 2023.
- [73] Yoshihiro Kanno. Mixed-integer programming formulation of a data-driven solver in computational elasticity. *Optimization Letters*, 13(7):1505–1514, 2019.

- [74] Yoshihiro Kanno. Alternating minimization for data-driven computational elasticity from experimental data: kernel method for learning constitutive manifold. *Theoretical and Applied Mechanics Letters*, 11(5):100289, 2021.
- [75] Yoshihiro Kanno. Data-driven confidence bound for structural response using segmented least squares: a mixed-integer programming approach. *Japan Journal of Industrial and Applied Mathematics*, 41(3):1501–1534, 2024.
- [76] Konstantinos Karapiperis, Michael Ortiz, and José E Andrade. Data-driven non-local mechanics: Discovering the internal length scales of materials. *Computer Methods in Applied Mechanics and Engineering*, 386:114039, 2021.
- [77] Konstantinos Karapiperis, Laurent Stainier, Michael Ortiz, and José E Andrade. Data-driven multiscale modeling in mechanics. *Journal of the Mechanics and Physics of Solids*, 147:104239, 2021.
- [78] A Kaveh. Suboptimal cycle bases of graphs for the flexibility analysis of skeletal structures. *Computer methods in applied mechanics and engineering*, 71(3): 259–271, 1988.
- [79] A Kaveh and M S Massoudi. Efficient finite element analysis using graph-theoretical force method; tetrahedron elements. *International Journal of Civil Engineering*, 12(2), 2014.
- [80] Steve Kench and Samuel J Cooper. Generating three-dimensional structures from a two-dimensional slice with generative adversarial network-based dimensionality expansion. *Nature Machine Intelligence*, 3(4):299–305, 2021.
- [81] Steve Kench, Isaac Squires, Amir Dahari, and Samuel J Cooper. Microlib: A library of 3d microstructures generated from 2d micrographs using sliceGAN. *Scientific Data*, 9(1):645, 2022.
- [82] Trenton Kirchdoerfer and Michael Ortiz. Data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 304:81–101, 2016.
- [83] Trenton Kirchdoerfer and Michael Ortiz. Data driven computing with noisy material data sets. *Computer Methods in Applied Mechanics and Engineering*, 326:622–641, 2017.

-
- [84] Dominik K Klein, Mauricio Fernández, Robert J Martin, Patrizio Neff, and Oliver Weeger. Polyconvex anisotropic hyperelasticity with neural networks. *Journal of the Mechanics and Physics of Solids*, 159:104703, 2022.
- [85] B. T. Knapik, A. W. van der Vaart, and J. H. van Zanten. Bayesian inverse problems with gaussian priors. *Annals of Statistics*, 39:2626–2657, 2011.
- [86] B. T. Knapik, B. T. Szabó, A. W. van der Vaart, and J. H. van Zanten. Bayes procedures for adaptive inference in inverse problems for the white noise model. *Probab. Theory Relat. Fields*, 164(3–4):771–813, 2016.
- [87] Marina Konstantatou, Pierluigi D’Acunto, and Allan McRobie. Polarities in structural analysis and design: n-dimensional graphic statics and structural transformations. *International Journal of Solids and Structures*, 152:272–293, 2018.
- [88] Tim Fabian Korzeniowski and Kerstin Weinberg. A comparison of stochastic and data-driven fem approaches to problems with insufficient material data. *Computer Methods in Applied Mechanics and Engineering*, 350:554–570, 2019.
- [89] Tim Fabian Korzeniowski and Kerstin Weinberg. A multi-level method for data-driven finite element computations. *Computer Methods in Applied Mechanics and Engineering*, 379:113740, 2021.
- [90] Tim Fabian Korzeniowski and Kerstin Weinberg. Data-driven diffusion with uncertainties. *PAMM*, 20(1):e202000325, 2021.
- [91] Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to pdes. *Journal of Machine Learning Research*, 24(89):1–97, 2023.
- [92] Zengtao Kuang, Xiaowei Bai, Qun Huang, Jie Yang, Wei Huang, Salim Belouettar, and Heng Hu. Data-driven computational framework for snap-through problems. *International Journal of Solids and Structures*, 269:112226, 2023.
- [93] Zengtao Kuang, Wei Yan, Kun Yu, Rui Xu, Liang Li, Qun Huang, Jie Yang, Gaetano Giunta, and Salim Belouettar. Data-driven computing for nonlinear

- problems of composite structures based on sub-domain search technique. *Computers & Structures*, 279:106982, 2023.
- [94] Adriana Kuliková, Andrei G Shvarts, Łukasz Kaczmarczyk, and Chris J Pearce. Conservative data-driven finite element framework with adaptive hp-refinement for diffusion problems with material uncertainty. *Computer Methods in Applied Mechanics and Engineering*, 452:118703, 2026.
- [95] S Kullback. *Information Theory and Statistics*. John Wiley & Sons, New-York, 1959.
- [96] S. Kullback and R. A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86, 1951.
- [97] Pierre Ladevèze and Ludovic Chamoin. Data-driven material modeling based on the constitutive relation error. *Advanced Modeling and Simulation in Engineering Sciences*, 11(1):23, 2024.
- [98] Pierre Ladeveze and Dominique Leguillon. Error estimate procedure in the finite element method and applications. *SIAM journal on numerical analysis*, 20(3):485–509, 1983.
- [99] Isaac E Lagaris, Aristidis Likas, and Dimitrios I Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *IEEE transactions on neural networks*, 9(5):987–1000, 1998.
- [100] Raphaël Langlois, Michel Coret, and Julien Réthoré. Non-parametric stress field estimation for history-dependent materials: Application to ductile material exhibiting piobert–lüders localization bands. *Strain*, 58(4):e12410, 2022.
- [101] Nicolas N Lee, Joel W Burdick, Paul Backes, Sergio Pellegrino, Kristina Hogstrom, Christine Fuller, Brett Kennedy, Junggon Kim, Rudranarayan Mukherjee, Carl Seubert, et al. Architecture for in-space robotic assembly of a modular space telescope. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2(4):041207, 2016.

- [102] Marek Lefik and Bernhard A Schrefler. Artificial neural network as an incremental non-linear constitutive model for a finite element code. *Computer methods in applied mechanics and engineering*, 192(28-30):3265–3283, 2003.
- [103] Adrien Leygue. On the formulation and convergence of data-driven identification. *Comptes Rendus. Mécanique*, 353(G1):761–773, 2025.
- [104] Adrien Leygue, Michel Coret, Julien Réthoré, Laurent Stainier, and Erwan Veron. Data-based derivation of material response. *Computer Methods in Applied Mechanics and Engineering*, 331:184–196, 2018.
- [105] Liang Li, Jie Yang, Xiaowei Bai, Qun Huang, Gaetano Giunta, Salim Belouettar, Huicui Li, and Heng Hu. Criterion-free and model-free data-driven framework for failure prediction of composites. *Composite Structures*, page 119900, 2025.
- [106] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.
- [107] Lennart Linden, Dominik K Klein, Karl A Kalina, Jörg Brummund, Oliver Weeger, and Markus Kästner. Neural networks meet hyperelasticity: A guide to enforcing physics. *Journal of the Mechanics and Physics of Solids*, 179: 105363, 2023.
- [108] Kevin Linka, Markus Hillgärtner, Kian P Abdolazizi, Roland C Aydin, Mikhail Itskov, and Christian J Cyron. Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning. *Journal of Computational Physics*, 429:110010, 2021.
- [109] Xin Liu, Su Tian, Fei Tao, and Wenbin Yu. A review of artificial neural networks in the constitutive modeling of composite materials. *Composites Part B: Engineering*, 224:109152, 2021.
- [110] Stuart Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137, 1982.

- [111] Jonathan Machta. Population annealing with weighted averages: A monte carlo method for rough free-energy landscapes. *Physical Review E*, 82(2):026704, 2010.
- [112] Eduard Marenić, Guillem Seychal, and Jean-Charles Passieux. Data driven approach in multiphysics framework: Application to coupled electro-mechanical problems. *Computer Methods in Applied Mechanics and Engineering*, 395: 114959, 2022.
- [113] Filippo Masi, Ioannis Stefanou, Paolo Vannucci, and Victor Maffi-Berthier. Thermodynamics-based artificial neural networks for constitutive modeling. *Journal of the Mechanics and Physics of Solids*, 147:104277, 2021.
- [114] Marco Maurizi, Chao Gao, and Filippo Berto. Predicting stress, strain and deformation fields in materials and structures with graph neural networks. *Scientific reports*, 12(1):21834, 2022.
- [115] J Clerk Maxwell. Xlv. on reciprocal figures and diagrams of forces. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 27(182): 250–261, 1864.
- [116] Jeremy A McCulloch, Skyler R St. Pierre, Kevin Linka, and Ellen Kuhl. On sparse regression, lp-regularization, and automated model discovery. *International Journal for Numerical Methods in Engineering*, 125(14):e7481, 2024.
- [117] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
- [118] Hervé Moulinec and Pierre Suquet. A numerical method for computing the overall response of nonlinear composites with complex microstructure. *Computer methods in applied mechanics and engineering*, 157(1-2):69–94, 1998.
- [119] Mojtaba Mozaffar, Ramin Bostanabad, W Chen, K Ehmann, Jian Cao, and MA Bessa. Deep learning predicts path-dependent plasticity. *Proceedings of the National Academy of Sciences*, 116(52):26414–26420, 2019.

- [120] Marius Muja and David G Lowe. Scalable nearest neighbor algorithms for high dimensional data. *IEEE transactions on pattern analysis and machine intelligence*, 36(11):2227–2240, 2014.
- [121] James R Munkres. *Elements of algebraic topology*. CRC press, 2018.
- [122] K. Naito, Y. Tanaka, J.-M. Yang, and Y. Kagawa. Tensile properties of ultrahigh strength pan-based, ultrahigh modulus pitch-based and high ductility pitch-based carbon fibers. *Carbon*, 46:189–195, 2008.
- [123] Radford M Neal. Mcmc using hamiltonian dynamics. *Handbook of markov chain monte carlo*, pages 47–95, 2011.
- [124] Lu Trong Khiem Nguyen and Marc-André Keip. A data-driven approach to nonlinear elasticity. *Computers & Structures*, 194:97–115, 2018.
- [125] Lu Trong Khiem Nguyen, Roland Can Aydin, and Christian Johannes Cyron. Accelerating the distance-minimizing method for data-driven elasticity with adaptive hyperparameters. *Computational Mechanics*, 70(3):621–638, 2022.
- [126] H. Owhadi, C. Scovel, and T. Sullivan. On the brittleness of bayesian inference. *SIAM REVIEW*, 57(4):566–582, 2015.
- [127] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [128] Mathias Peirlinck, Kevin Linka, Juan A Hurtado, and Ellen Kuhl. On automated model discovery and a universal material subroutine for hyperelastic materials. *Computer Methods in Applied Mechanics and Engineering*, 418:116534, 2024.
- [129] Sergio Pellegrino and Christopher Reuben Calladine. Matrix analysis of statically and kinematically indeterminate frameworks. *International Journal of Solids and Structures*, 22(4):409–428, 1986.

- [130] F. J. Pinski, G. Simpson, A. M. Stuart, and H. Weber. Kullback-leibler approximation for probability measures on infinite dimensional spaces. *SIAM Journal on Mathematical Analysis*, 47(6):4091–4122, 2015.
- [131] Auriane Platzter, Adrien Leygue, Laurent Stainier, and Michael Ortiz. Finite element solver for data-driven finite strain elasticity. *Computer Methods in Applied Mechanics and Engineering*, 379:113756, 2021.
- [132] Klaas Poelstra, Thorsten Bartel, and Ben Schweizer. A data-driven framework for evolutionary problems in solid mechanics. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 103(3):e202100538, 2023.
- [133] Alex Pothén. Sparse null basis computations in structural optimization. *Numerische Mathematik*, 55(5):501–519, 1989.
- [134] Erik Prume, Stefanie Reese, and Michael Ortiz. Model-free data-driven inference in computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 403:115704, 2023.
- [135] Erik Prume, Laurent Stainier, Michael Ortiz, and Stefanie Reese. A data-driven solver scheme for inelastic problems. *PAMM*, 23(1):e202200153, 2023.
- [136] Erik Prume, Christian Gierden, Michael Ortiz, and Stefanie Reese. Direct data-driven algorithms for multiscale mechanics. *Computer Methods in Applied Mechanics and Engineering*, 433:117525, 2025.
- [137] Erik Prume, Chenyi Ji, Stefanie Reese, and Michael Ortiz. A sparse basis for equilibrium stress fields with application for direct data-driven mechanics. *Computer Methods in Applied Mechanics and Engineering*, 453:118803, 2026.
- [138] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.
- [139] W. N. Reynolds and J. V. Sharp. Crystal shear limit to carbon fibre strength. *Carbon*, 12:103—110, 1974.

- [140] Felipe Rocha, Auriane Platzer, Adrien Leygue, and Laurent Stainier. On-the-fly adaptive sampling strategy for data-driven computational mechanics: Applications to computational homogenisation. *Mechanics of Materials*, 207:105382, 2025.
- [141] Matthias Röger and Ben Schweizer. Relaxation analysis in a data driven problem with a single outlier. *Calculus of Variations and Partial Differential Equations*, 59(4):119, 2020.
- [142] Ignacio Romero and Michael Ortiz. A note on data-driven methods for mechanical problems with non-unique solutions. *Meccanica*, 61(1):26, 2026.
- [143] Kenneth Rose. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proceedings of the IEEE*, 86(11):2210–2239, 2002.
- [144] Kenneth Rose, Eitan Gurewitz, and Geoffrey C Fox. Statistical mechanics and phase transitions in clustering. *Physical review letters*, 65(8):945, 1990.
- [145] Hossein Salahshoor and Michael Ortiz. Model-free data-driven viscoelasticity in the frequency domain. *Computer Methods in Applied Mechanics and Engineering*, 403:115657, 2023.
- [146] Esteban Samaniego, Cosmin Anitescu, Somdatta Goswami, Vien Minh Nguyen-Thanh, Hongwei Guo, Khader Hamdia, Xiaoying Zhuang, and Timon Rabczuk. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. *Computer Methods in Applied Mechanics and Engineering*, 362:112790, 2020.
- [147] Matti Schneider. A review of nonlinear fft-based computational homogenization methods. *Acta Mechanica*, 232(6):2051–2100, 2021.
- [148] Alexander Sorini, Evan J Pineda, Joshua Stuckner, and Peter A Gustafson. A convolutional neural network for multiscale modeling of composite materials. In *AIAA Scitech 2021 Forum*, page 0310, 2021.

- [149] Laurent Stainier, Adrien Leygue, and Michael Ortiz. Model-free data-driven methods in mechanics: material data identification and solvers. *Computational Mechanics*, 64(2):381–393, 2019.
- [150] Tung-Huan Su, Nian-Hu Lu, Chih-Hsuan Chen, and Chuin-Shan Chen. Full-field stress and strain measurements revealing energy dissipation characteristics in martensitic band of cu-al-mn shape memory alloy. *Materials Today Communications*, 24:101321, 2020.
- [151] Misato Suzuki, Mayu Muramatsu, Stefanie Reese, and Erik Prume. Efficient evaluation of mechanical properties for two-phase materials using a direct data-driven approach. *Materials & Design*, page 114946, 2025.
- [152] Shan Tang, Gang Zhang, Hang Yang, Ying Li, Wing Kam Liu, and Xu Guo. Map123: A data-driven approach to use 1d data for 3d nonlinear elastic materials modeling. *Computer Methods in Applied Mechanics and Engineering*, 357: 112587, 2019.
- [153] Prakash Thakolkaran, Akshay Joshi, Yiwen Zheng, Moritz Flaschel, Laura De Lorenzis, and Siddhant Kumar. Nn-euclid: Deep-learning hyperelasticity without stress data. *Journal of the Mechanics and Physics of Solids*, 169:105076, 2022.
- [154] Gabriel Valdés-Alonzo, Christophe Binetruy, Benedikt Eck, Alberto Garcia-Gonzalez, and Adrien Leygue. Phase distribution and properties identification of heterogeneous materials: A data-driven approach. *Computer Methods in Applied Mechanics and Engineering*, 390:114354, 2022.
- [155] Nikolaos N Vlassis and WaiChing Sun. Sobolev training of thermodynamic-informed neural networks for interpretable elasto-plasticity models with level set hardening. *Computer Methods in Applied Mechanics and Engineering*, 377: 113695, 2021.
- [156] Wenlong Wang, Jonathan Machta, and Helmut G Katzgraber. Comparing monte carlo methods for finding ground states of ising spin glasses: Population annealing, simulated annealing, and parallel tempering. *Physical Review E*, 92(1):013303, 2015.

- [157] Martin Weigel, Lev Barash, Lev Shchur, and Wolfram Janke. Understanding population annealing monte carlo simulations. *Physical Review E*, 103(5):053301, 2021.
- [158] Rui Xu, Jie Yang, Wei Yan, Qun Huang, Gaetano Giunta, Salim Belouettar, Hamid Zahrouni, Tarak Ben Zineb, and Heng Hu. Data-driven multiscale finite element method: From concurrence to separation. *Computer Methods in Applied Mechanics and Engineering*, 363:112893, 2020.
- [159] Yongchun Xu, Jie Yang, Zengtao Kuang, Qun Huang, Wei Huang, and Heng Hu. Quantum computing enhanced distance-minimizing data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 419: 116675, 2024.
- [160] Shibo Yan, Xi Zou, Mohammad Ilkhani, and Arthur Jones. An efficient multiscale surrogate modelling framework for composite materials considering progressive damage based on artificial neural networks. *Composites Part B: Engineering*, 194:108014, 2020.
- [161] Bing Yu et al. The deep ritz method: a deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics*, 6(1):1–12, 2018.
- [162] Olgierd Cecil Zienkiewicz and Robert Leroy Taylor. *The finite element method for solid and structural mechanics*. Elsevier, 2005.
- [163] Selina Zschocke, Ferenc Leichsenring, Wolfgang Graf, and Michael Kaliske. A concept for data-driven computational mechanics in the presence of polymorphic uncertain properties. *Engineering Structures*, 267:114672, 2022.

Publications

2026

- Prume, E., Ji, C., Reese, S., & Ortiz, M. (2026). A sparse basis for equilibrium stress fields with application for direct data-driven mechanics. *Computer Methods in Applied Mechanics and Engineering*, 453, 118803.

2025

- Prume, E., Gierden, C., Ortiz, M., & Reese, S. (2025). Direct data-driven algorithms for multiscale mechanics. *Computer Methods in Applied Mechanics and Engineering*, 433, 117525.
- Suzuki, M., Muramatsu, M., Reese, S., & Prume, E. (2025). Efficient evaluation of mechanical properties for two-phase materials using a direct data-driven approach. *Materials & Design*, 114946.

2023

- Prume, E., Reese, S., & Ortiz, M. (2023). Model-free data-driven inference in computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 403, 115704.
- Prume, E., Stainier, L., Ortiz, M., & Reese, S. (2023). A data-driven solver scheme for inelastic problems. *PAMM*, 23(1), e202200153.
- Kopic-Osmanovic, N., Prume, E., Felder, S., Kloppenburg, G., & Simon, J. W. (2023). Modeling the anisotropic elasto-plastic material behavior of paper and paperboard at finite deformations. *PAMM*, 22(1), e202200293.

Conferences

2024

- Prume, E., Gierden, C., Ortiz, M., Reese, S. (2024). Adaptive Material Data Set Enrichment for Multi-Scale Analysis. In 26th International Congress of Theoretical and Applied Mechanics (ICTAM).
- Reese, S., Gierden, C., Macek, D., Ortiz, M., Prume, E., Ritzert, S., Svendsen, B. (2024). Multiscale modeling by combining data-driven approaches and Fourier transforms. In 9th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS).

2023

- Prume, E., Gierden, C., Reese, S. (2023). A Data-Driven Framework for Multiscale Material Analysis. In XVII International Conference on Computational Plasticity (COMPLAS 2023).

2022

- Prume, E., Eggersmann, R., Reese, S., Ortiz, M. (2022). Model-Free Data-Driven Inference in Computational Mechanics. In 92nd Annual Meeting of the International Association of Applied Mathematics and Mechanics (GAMM).
- Reese, S., Prume, E., Gierden, C., Svendsen, B., Ortiz, M. (2022). Data-driven mechanics for elastic and inelastic problems. In XXV Convegno AIMETA.
- Prume, E., Reese, S., Ortiz, M. (2022). Model-Free Data-Driven Inference in Computational Mechanics. In 9th GACM Colloquium on Computational Mechanics (GACM).

2021

- Prume, E., Eggersmann, R., Stainier, L., Ortiz, M., Reese, S. (2021). Solver strategies for data-driven computations with inelastic materials. In XVI International Conference on Computational Plasticity (COMPLAS 2021).
- Prume, E., Eggersmann, R., Stainier, L., Ortiz, M., Reese, S. (2021). Data-driven computations with inelastic materials using commercial solvers. In Mechanistic Machine Learning and Digital Twins for Computational Science, Engineering & Technology.
- Prume, E., Eggersmann, R., Reese, S., Ortiz, M., (2021). A probabilistic model-free data-driven solver scheme. In 16th U.S. National Congress on Computational Mechanics
- Reese, S., Eggersmann, R., Prume, E., Stainier, L., Ortiz, M. (2021). Efficient data structures for data-driven mechanics, In Presentations and videos to 10th International Conference on Adaptive Modeling and Simulation (ADMOS).
- Reese, S., Prume, E., Eggersmann, R., Stainier, L., Ortiz, M. (2021). Data-driven mechanics – interesting aspects in the modelling of inelastic behaviour. In XVI International Conference on Computational Plasticity (COMPLAS 2021).
- Reese, S., Gierden, C., Schmidt, A., Prume, E., Weimann, J., Svendsen, B., Ortiz, M. (2021) Efficient multi-scale modelling by means of FE-FFT and model order reduction techniques - enhanced by data-driven mechanics. In MCIC 2021 – Materials Discovery and Processing for Energy.
- Reese, S., Prume, E., Eggersmann, R., Gierden, C., Svendsen, B., Stainier, L., Ortiz, M. (2021). Data-driven Mechanics – theoretical and numerical aspects in the context of inelastic material behavior. In 25th International Congress of Theoretical and Applied Mechanics (ICTAM).