



**BERGISCHE
UNIVERSITÄT
WUPPERTAL**

Constitutive modeling of sustainable materials enhanced by data-driven approaches

**Dissertation
zur Erlangung eines Doktorgrades (Dr.-Ing.)**

in der
Fakultät für Architektur und Bauingenieurwesen
der
Bergischen Universität Wuppertal

vorgelegt von
Birte Boes

Gutachtende: Prof. Dr.-Ing. habil. Jaan-Willem Simon
Prof. Dr.-Ing. habil. Ellen Kuhl

Tag der mündlichen Prüfung: 27.07.2026

Acknowledgements

This cumulative dissertation is the result of my scientific work over the past years. While it was initiated at the Institute of Applied Mechanics (IFAM) at RWTH Aachen University, the majority of the work was carried out at the Chair of Computational Applied Mechanics (CAM) at the University of Wuppertal. Looking back, I am very grateful for the many experiences I was able to gain and, in particular, for the people I had the privilege to meet and work with along the way. Their support, encouragement, and the shared exchange have been invaluable throughout this journey. I would therefore like to sincerely thank everyone who contributed to this work in one way or another.

First and foremost, I would like to thank my supervisor, Prof. Dr.-Ing. habil. Jaan-Willem Simon, for the invaluable time at the institute, his continuous support, trust, and guidance throughout my doctoral studies. I am especially grateful for the freedom and flexibility he provided, as well as for the many opportunities he created, including research stays, conference participation, and the chance to take on additional responsibilities besides research.

I would also like to thank Prof. Dr.-Ing. habil. Ellen Kuhl, my second examiner, for the opportunity to spend a research stay at her lab at Stanford University. This time was incredibly valuable, both academically and personally. I learned not only through the research itself, but also from her inspiring way of working—her enthusiasm for the field, her ability to grasp complex problems with remarkable speed, and her forward-thinking mindset left a lasting impression on me.

A very special thanks goes to Dr.-Ing. Hagen Holthusen. I am deeply grateful for his continuous support, guidance, and the time he dedicated to sharing his knowledge and ideas with me. Although never officially appointed as my supervisor, he played a central role throughout my entire doctoral journey and contributed significantly to its success. His patience, encouragement, and insights have shaped both my scientific development and my research direction in a fundamental way. Beyond academia, I am especially thankful to have gained not only a mentor but also a close friend, whose support extends far beyond this work. I am deeply grateful for my colleagues Jan Pfeifer, Greta Kloppenburg, and Nadir Kopic for the many shared experiences during my time at the institute. Your support, sense of humor, and companionship made this time truly enjoyable. I am very grateful for the great atmosphere, the shared challenges, the laughter, and the many memorable moments we experienced together. Further, I would like to thank Ylva van Luijt and Sofie Danowski, whom I had the pleasure of supervising during their bachelor's and master's theses as well as their time as student assistants. Working with you was truly rewarding, and I greatly appreciate the collaboration

and the insights I gained along the way. It was a pleasure to accompany your academic journey.

I would like to thank all my colleagues at the institute for the shared time, collaboration, and exchange. I am particularly grateful for the support in teaching and for the many valuable insights I gained through working together, especially in developing practical skills and approaches.

My thanks also extend to my former colleagues at IFAM and to all other colleagues from other institutes, whose contributions and interactions enriched my time significantly. I truly appreciate the supportive environment, the many shared coffee breaks and lunch times, and the sense of collegiality that made this time both enjoyable and meaningful.

My deepest gratitude goes to my family for their unconditional support. While my mother and sister guided me through tough times, I would like to thank my grandmother, who provided not only a place to work during some home office days, but also companionship and guidance in many situations throughout life.

Finally, I would like to thank my friends for their constant support, balance, and encouragement. They provided the necessary perspective beyond work and reminded me of the importance of enjoying the journey.

TLDR: I am deeply grateful for the people, who supported and inspired me, the challenges and opportunities, I was allowed to experience, and the mentorship, guidance and friendships throughout this journey. It is a pleasure, to be supported the way I was and still am, and I hope that the future brings the same privileges, joy and excitement as I was allowed to experience up to today.

Birte Boes
Wuppertal 2026

Summary

Sustainable material design and efficient resource usage are central challenges in modern engineering. Materials relevant in this context often exhibit complex mechanical behavior characterized by anisotropy, nonlinear and inelastic effects, as well as dependencies on manufacturing processes, microstructure, and material composition. Accurate constitutive models are therefore essential to enable the efficient use of such materials, improve material performance, reduce waste, and support reliable design processes. However, the inherent complexity of these materials makes the derivation of suitable and predictive material models challenging. While continuum mechanical frameworks provide a rigorous basis for modeling, recent advances in machine learning offer additional opportunities to capture complex material behavior and to identify constitutive relations directly from data. This dissertation builds on this synergy by combining physically motivated modeling with data-driven approaches to develop flexible, interpretable, and robust constitutive models. The proposed methodologies are exemplified particularly for paper and paperboard as anisotropic, inelastic materials and for tofu as a hydrated, composition-dependent soft material.

The first part of the work focuses on the development of a continuum mechanical framework for anisotropic materials, motivated by the behavior of paper and paperboard. In the first article, a novel formulation is introduced that enables a consistent decoupling of in-plane and out-of-plane material behavior through a decomposition of the deformation measures using structural tensors. This approach provides a physically sound and flexible basis for describing anisotropic responses at large deformations and overcomes limitations of existing models that rely on parameter-based decoupling. The framework is designed to be extensible and thereby allows straightforward incorporation of additional inelastic effects, demonstrated on visco-elasticity.

Building on this foundation, the second article extends the framework to account for various directional dependency in the inelastic material behavior. By introducing additional inelastic potentials and evolution equations, the model captures direction-dependent and -independent inelastic effects while maintaining the decoupled structure. This extension demonstrates the flexibility of the proposed formulation and highlights its capability to represent complex, anisotropic, and path-dependent material behavior within a unified three-dimensional setting. The second part of the dissertation addresses data-driven constitutive modeling. In the third article, the framework of inelastic constitutive artificial neural networks is extended to describe elasto-plastic behavior. By combining continuum mechanical principles with neural network representations of constitutive potentials, the approach enables automated model discovery

while ensuring thermodynamic consistency. The incorporation of recurrent structures allows the modeling of path-dependent processes, and the framework is validated using both synthetic and experimental data. This contribution demonstrates how data-driven approaches can enhance classical modeling by increasing flexibility and enabling the identification of underlying relations directly from data.

Finally, Article 4 applies the data-driven framework to the modeling of tofu as a representative hydrated soft material. Based on experimental data from confined compression tests, performed for this purpose and at varying strain rates and water contents, constitutive models are automatically identified that capture visco-elastic and feature-dependent behavior. By integrating feature networks, the approach accounts for variations in material composition and provides a unified model capable of describing both rate-dependent effects and composition-driven changes in mechanical response. This application highlights the potential of data-driven constitutive modeling not only for accurate prediction but also for gaining insight into structure-property relationships in complex materials.

Together, the contributions of this dissertation demonstrate that the combination of continuum mechanics and automated model discovery provides a powerful framework for modeling complex material behavior. The developed approaches enable both physically interpretable and flexible representations, offering new opportunities for the analysis, design, and optimization of sustainable materials.

Zusammenfassung

Nachhaltige Materialien und eine effiziente Nutzung von Ressourcen stellen zentrale Herausforderungen im modernen Ingenieurwesen dar. Materialien, die in diesem Kontext von Bedeutung sind, weisen häufig ein komplexes mechanisches Verhalten auf, das durch Anisotropie, nichtlineare und inelastische Effekte sowie Abhängigkeiten von Herstellungsprozessen, Mikrostruktur und Materialzusammensetzung gekennzeichnet ist. Zur effizienten Nutzung dieser Materialien, zur Verbesserung ihrer Leistungsfähigkeit, zur Reduzierung von Abfällen sowie zur Unterstützung zuverlässiger Auslegungsprozesse sind daher präzise konstitutive Modelle unerlässlich. Die inhärente Komplexität dieser Materialien erschwert jedoch die Herleitung geeigneter Materialmodelle erheblich.

Während kontinuumsmechanische Ansätze eine etablierte Grundlage für die Modellierung bieten, eröffnen jüngste Fortschritte im Bereich des maschinellen Lernens zusätzliche Möglichkeiten, komplexes Materialverhalten zu erfassen und konstitutive Beziehungen direkt aus Daten zu identifizieren. Die vorliegende Dissertation baut auf dieser Synergie auf, indem sie physikalisch motivierte Modellierung mit datengetriebenen Ansätzen kombiniert, um flexible, interpretierbare und robuste konstitutive Modelle zu entwickeln. Die vorgeschlagenen Methoden werden exemplarisch insbesondere an Papier und Pappkarton als anisotrope, inelastische Materialien sowie an Tofu als hydratisiertes, zusammensetzungsabhängiges weiches Material demonstriert.

Der erste Teil der Arbeit konzentriert sich auf die Entwicklung eines kontinuumsmechanischen Modells für anisotrope Materialien, motiviert durch das Verhalten von Papier und Karton. Im ersten Artikel wird eine neuartige Formulierung vorgestellt, die eine konsistente Entkopplung des in-plane- und out-of-plane-Verhaltens durch eine Zerlegung der Kinematik mithilfe Strukturtenoren ermöglicht. Dieser Ansatz bietet eine physikalisch fundierte und flexible Grundlage zur Beschreibung anisotroper Reaktionen bei großen Deformationen und überwindet die Einschränkungen bestehender Modelle, die auf parameterbasierter Entkopplung beruhen. Das Modell ist so konzipiert, dass es erweiterbar ist und somit eine einfache Integration zusätzlicher inelastischer Effekte erlaubt, was am Beispiel der Viskoelastizität demonstriert wird.

Aufbauend darauf wird der Ansatz im zweiten Artikel erweitert, um verschiedene richtungsabhängige Effekte im inelastischen Materialverhalten zu berücksichtigen. Durch die Einführung zusätzlicher inelastischer Potentiale und Evolutionsgleichungen kann das Modell sowohl richtungsabhängige als auch richtungsunabhängige inelastische Effekte abbilden, während die entkoppelte Struktur erhalten bleibt. Diese Erweiterung verdeutlicht die Flexibilität der Formulierung und zeigt ihre Fähigkeit, komplexes, anisotropes und pfadabhängiges Materialver-

halten innerhalb eines einheitlichen dreidimensionalen Rahmens zu beschreiben.

Der zweite Teil der Dissertation widmet sich der datengetriebenen konstitutiven Modellierung. Im dritten Artikel wird der Ansatz der inelastischen konstitutiven künstlichen neuronalen Netze erweitert, um elasto-plastisches Verhalten zu beschreiben. Durch die Kombination kontinuumsmechanischer Prinzipien mit neuronalen Netzwerken zur Darstellung konstitutiver Potentiale ermöglicht dieser Ansatz eine automatisierte Modellidentifikation bei gleichzeitiger Sicherstellung thermodynamischer Konsistenz. Die Einbindung in rekurrente Strukturen erlaubt die Abbildung pfadabhängiger Prozesse. Der Ansatz wird sowohl anhand synthetischer als auch experimenteller Daten validiert. Dieser Beitrag zeigt, wie datengetriebene Methoden klassische Modellierungsansätze erweitern können, indem sie die Flexibilität erhöhen und die Identifikation zugrunde liegender Zusammenhänge direkt aus Daten ermöglichen.

Abschließend wird im letzten Artikel der datengetriebene Ansatz auf die Modellierung von Tofu als repräsentativem hydratisierten weichen Material angewendet. Basierend auf experimentellen Daten aus hierfür durchgeführten Kompressionstests bei unterschiedlichen Dehnraten und Wassergehalten werden konstitutive Modelle automatisch identifiziert, die viskoelastisches und merkmalsabhängiges Verhalten erfassen. Durch die Integration von Feature-Netzwerken kann der Einfluss der Materialzusammensetzung berücksichtigt werden, sodass ein einheitliches Modell entsteht, das sowohl ratenabhängige Effekte als auch durch die Zusammensetzung bedingte Änderungen des mechanischen Verhaltens beschreibt. Diese Anwendung verdeutlicht das Potenzial datengetriebener konstitutiver Modellierung nicht nur für präzise Vorhersagen, sondern auch für das Verständnis von Struktur-Eigenschafts-Beziehungen in komplexen Materialien.

Insgesamt zeigen die Beiträge dieser Dissertation, dass die Kombination aus Kontinuumsmechanik und automatisierter Modellidentifikation einen leistungsfähigen Rahmen zur Beschreibung komplexen Materialverhaltens darstellt. Die entwickelten Ansätze ermöglichen sowohl physikalisch interpretierbare als auch flexible Modellierungen und eröffnen neue Möglichkeiten für die Analyse, das Design und die Optimierung nachhaltiger Materialien.

Contents

1. Introduction	1
1.1. Motivation	1
1.1.1. Modeling of sustainable materials	1
1.1.2. Advancements of machine learning in material modeling	3
1.2. State-of-the-art and research relevant questions	4
1.2.1. Anisotropic modeling of paper and paperboard	5
1.2.2. Modeling approaches for tofu	8
1.2.3. Machine learning in constitutive modeling	9
1.3. Results and discussion of the research contribution	14
1.4. Outline of the dissertation	16
2. Article 1:	
A novel continuum mechanical framework for decoupled material behavior in thickness and in-plane directions	19
2.1. Abstract	20
2.2. Introduction	20
2.3. Elastic material behavior	24
2.3.1. Classical isochoric-volumetric split	25
2.3.2. Decoupling of the in-plane and out-of-plane behaviors	25
2.4. Inelastic material formulation	31
2.4.1. Co-rotated intermediate configuration	32
2.4.2. Structural tensors in different configurations	33
2.4.3. Derivation based on the Clausius-Duhem inequality	34
2.4.4. Evolution equations	37
2.4.5. Numerical Implementation	40
2.5. Choice of Helmholtz free energy	40
2.5.1. Remarks on the formulation of Helmholtz free energy functions	40
2.5.2. Elastic material formulation - isoplanar-area-changing split	42
2.5.3. Elastic material formulation - isochoric-volumetric split	44

2.5.4. Inelastic material formulation	44
2.6. Numerical Examples	45
2.6.1. Single-element tests for the elastic material behavior	45
2.6.2. Cube under compression with elastic material behavior	46
2.6.3. Single-element tests for the inelastic material behavior	49
2.6.4. Pinched cylinder with inelastic material behavior	51
2.6.5. Symmetrically notched specimen with inelastic material behavior . .	56
2.7. Conclusion	57
2.8. Appendix	61
2.8.1. The modified right Cauchy-Green tensor in the isochoric-volumetric framework	61
2.8.2. Proof of decoupled formulation of second Piola-Kirchhoff stress for- mulation	61
2.8.3. Rates of tensors	61
2.8.4. Derivative of the Helmholtz free energy	62
2.8.5. Comparison of the Kirchhoff stress and its conjugate force	63
2.8.6. Calculation of invariants in different configurations	66
2.8.7. Partial Derivatives	67

3. Article 2:

A constitutive model for describing decoupled material behavior in thick- ness and in-plane directions	69
3.1. Abstract	70
3.2. Introduction	70
3.3. Decoupling of in-plane and out-of-plane material behavior	71
3.3.1. Structural tensors	71
3.3.2. Modified invariants	72
3.3.3. Isoplanar-area-changing split	72
3.4. Inelastic material formulation	73
3.4.1. Choice of potentials	75
3.4.2. Specific choice of the Helmholtz free energy	76
3.5. Numerical Examples	76
3.6. Conclusion	77

4. Article 3:

Accounting for plasticity: An extension of inelastic constitutive artificial neural networks	81
4.1. Abstract	82
4.2. Introduction	83
4.3. Constitutive Modeling	86
4.3.1. Kinematics	86
4.3.2. Derivation based on the Clausius-Planck inequality	87
4.4. Network Formulation and Architecture	89
4.4.1. Algorithmic treatment	90
4.4.2. Recurrent neural network architecture	92
4.4.3. Feed-forward neural network architecture	94
4.5. Automated model discovery	98
4.5.1. Specification of network architecture	99
4.5.2. Preprocessing	100
4.5.3. Discovering artificially generated data	101
4.5.4. Comparison to other neural network approaches	107
4.5.5. Experimental data	109
4.6. Discussion and limitations	112
4.7. Conclusion	114
4.8. Appendix	115
4.8.1. Impact of regularization techniques	115

5. Article 4:

The mechanics and physics of tofu: understanding hydrated soft solids through feature networks	119
5.1. Abstract	120
5.2. Introduction	120
5.3. Mechanical testing	123
5.4. Physics-based model	125
5.5. Neural network model	130
5.6. Results	137
5.7. Discussion	148
5.8. Appendix	154
5.8.1. Errors of discovered individual and unified models	154

5.8.2. Preliminary sensitivity analyses	156
6. Conclusions and Outlook	161
A. Appendix	165
List of Figures	167
List of Tables	175
Bibliography	177

1 | Introduction

1.1. Motivation

Efficient resource utilization has become a central objective in modern engineering, particularly in the development and application of sustainable materials. Many of these materials result from natural systems, which are inherently complex and highly optimized through evolutionary processes. Consequently, they exhibit sophisticated mechanical behavior often characterized by pronounced anisotropy, nonlinear responses, and inelastic effects.

To fully exploit the potential of such materials, a comprehensive understanding and accurate description of their mechanical behavior are essential. Constitutive modeling provides a systematic framework to relate observable quantities such as stress and strain and enables the prediction of material response under diverse loading conditions. Reliable material models thus form the foundation for improving material efficiency, optimizing structural performance, and minimizing unnecessary resource consumption. In this context, a better understanding of complex material behavior directly contributes to more efficient and sustainable resource usage.

1.1.1. Modeling of sustainable materials

Materials relevant to sustainable applications are frequently characterized by complex mechanical behavior that significantly influences their performance and usability. While such complexity is often associated with large-scale engineering systems, it also plays a crucial role in everyday applications. Although individual contributions may appear minor, their cumulative effect can be substantial on a global scale [Ganeshan, 2025]. This is particularly evident in sectors such as packaging and nutrition, where material selection and efficient utilization strongly impact overall sustainability. Improving the performance and usage of commonly employed materials is therefore essential for the transition toward more sustainable consumption patterns [Poore and Nemecek, 2018; Pauer et al., 2019].

A major challenge in this context arises from the intrinsic complexity of sustainable materials.

Their behavior is typically governed by anisotropy, nonlinearities, inelastic effects, as well as dependencies on environmental and compositional factors. Accurately capturing these characteristics is crucial for reliable predictions under realistic loading conditions. Constitutive models serve as a key tool for translating such complex behavior into a form suitable for analysis and simulation. Enhanced predictive capabilities enable improved material utilization, increased process reliability, and reduced resource consumption. These aspects can be investigated using several representative material classes: paper and paperboard as fibrous, anisotropic packaging materials, and tofu as a hydrated soft material relevant for sustainable nutrition. All materials play an important role in everyday life and exhibit fundamentally different, yet equally complex mechanical behavior. By addressing anisotropic and inelastic responses, this work aims to provide a deeper understanding of the underlying mechanisms and to demonstrate how advanced modeling approaches can contribute to the efficient and sustainable use of such materials.

Modeling of paper and paperboard. Paper and paperboard are widely used materials that constitute a cornerstone of the global packaging industry. As bio-based, recyclable, and versatile materials, they are central to the transition toward sustainable packaging solutions. In industrial processes such as creasing, folding, and forming, these materials are subjected to large deformations and complex loading conditions. The reliable simulation of such processes is essential for efficient product design and manufacturing and relies on accurate constitutive descriptions.

The development of such models is challenging due to the intrinsic complexity of paper-based materials. Their mechanical behavior is governed by a fibrous microstructure, which results in pronounced anisotropy. During the papermaking process, fibers predominantly align in the machine direction, leading to strongly direction-dependent properties [Alzweighi et al., 2022]. In addition, these materials exhibit nonlinear elasto-plastic deformation and rate-dependent effects [Tjahjanto et al., 2015].

Accurate material models that capture these effects can support the efficient resource usage and development of more sustainable packaging technologies.

Modeling of Tofu. For many soft materials relevant to sustainable applications, mechanical behavior is strongly influenced by both inelastic effects and material composition. Dependencies on factors such as compositional textures or processing conditions can significantly alter the response and therefore need to be incorporated into constitutive descriptions. Accounting for these influences is essential to achieve reliable and transferable material models.

Tofu provides a representative example in this context. As a plant-based protein source with

a low environmental footprint, it is an important component of sustainable diets that aim to reduce reliance on animal-based products. It consists of a hydrated protein network formed during the coagulation of soymilk [Wang et al., 2018], where variations in water content lead to a wide range of mechanical properties. This forms a bi-phasic material with a characteristic inelastic response that differs significantly depending on the composition.

In food science, these textures are commonly characterized using texture profile analysis, which has been established as a standard method for several decades. This approach typically relies on double-compression tests imitating the chewing process from which quantities such as stiffness, cohesiveness or springiness are derived [Bourne, 1968]. While these tests are widely applied, the interpretation of these measurements is usually framed in empirical or sensory terms rather than in a rigorous mechanical context.

Adopting a mechanics-based analysis enables a more fundamental understanding of tofu as a hydrated soft material. Embedding such perspectives within a constitutive modeling framework allows for a systematic and physically interpretable description of food materials. Moreover, due to its controllable composition and variable properties, tofu can serve as a model system and a suitable benchmark for hydrated soft materials.

Reliable material models capable of capturing nonlinear, anisotropic, and inelastic behavior are essential for improving resource efficiency across applications. A deeper understanding of materials such as paper and tofu might enable not only optimized product performance and reduced waste, but also supports sustainable product design strategies.

1.1.2. Advancements of machine learning in material modeling

In the context of material modeling, the central objective is to formulate constitutive relations that accurately describe the physical behavior observed in experimental data. This task becomes increasingly challenging with the emergence of new materials, which often exhibit complex, nonlinear, and coupled responses. While continuum mechanics provides a well-established theoretical foundation for elastic and inelastic behavior, identifying suitable functional forms that capture underlying mechanisms remains a nontrivial task.

In recent years, the field of machine learning has experienced rapid development across numerous scientific and engineering disciplines. Advances in computational power, data availability, and algorithmic design have led to efficiently represent complex, nonlinear relationships in data. Although many of these developments originate outside the field of mechanics, they provide valuable concepts and tools that can be transferred to constitutive modeling. Leveraging this progress is therefore a natural step, as it enables the use of established, well-developed methodologies to address challenges in material modeling.

Combining the developments in neural networks with established mechanical principles enables the discovery of enhanced constitutive models that retain physical laws while increasing modeling flexibility. In this context, data-driven approaches are not intended to replace classical theory but to complement it by facilitating the identification of suitable constitutive formulations that may be difficult to discover otherwise.

A key aspect of such models is the description of inelastic behavior under finite deformations, which governs irreversible processes such as plasticity and visco-elasticity. Potential-based formulations provide a flexible framework for this purpose, as they allow flexible formulation to account for different effects while satisfying thermodynamical constraints. Such formulations are well suited for integration with neural network-based approaches, as they enable to learn complex material responses from data while preserving a clear physical structure. The combination of nonlinear constitutive modeling and machine learning thus enables the development of models that are both learnable and physically sound.

Overall, the integration of modern machine learning methods with established concepts from continuum mechanics represents a promising direction for advancing material modeling. This synergy not only enhances existing models but also opens new pathways to discover novel constitutive formulations.

Despite the rapid growth of this research area and various data-driven approaches in constitutive modeling, significant challenges remain. These include the extension of existing modeling approaches toward more general and flexible frameworks that incorporate nonlinear effects such as elasto-plasticity, as well as feature-dependent behavior arising from variations in material composition. Addressing these aspects is essential to fully exploit the potential of machine learning in constitutive modeling and to develop reliable, efficient, and widely applicable modeling strategies.

1.2. State-of-the-art and research relevant questions

In order to further specify the scope of this work, the research-relevant questions are formulated in the following and placed in the context of existing literature. To this end, a review of relevant contributions is provided, linking the previously motivated challenges to the current state of research. Continuum mechanical modeling of materials, including anisotropy and inelasticity, is extensively studied. At the same time, recent developments in data-driven modeling have introduced new perspectives and methodologies for describing complex material behavior. The following literature review does not aim to provide a comprehensive overview of all existing work. Instead, it focuses on key developments relevant to this dissertation, highlighting both

classical approaches and recent advancements.

1.2.1. Anisotropic modeling of paper and paperboard

Paper and paperboard are widely used engineering materials in packaging and forming applications due to their low weight, sustainability, and favorable mechanical properties.

In many industrial processes—such as deep drawing, press forming, and tray forming—the material undergoes complex deformation states that involve both mechanisms in the sheet-plane and thickness direction. Accurately predicting this behavior is therefore essential for the reliable simulation and optimization of forming processes. Over the past decades, numerous continuum models have been developed to describe the mechanical behavior of paper and paperboard and to enable numerical simulation of forming operations such as deep drawing and press forming [Awais et al., 2017; Vishtal et al., 2014; Lindberg and Kulachenko, 2022] or creasing and folding [Beex and Peerlings, 2009; Huang et al., 2014]. Comprehensive reviews summarizing modeling approaches can be found by Fadji et al. [2018]; Östlund [2017] and Simon [2021].

From a microstructural perspective, paper is a fibrous material composed of a network of cellulose fibers. During the manufacturing process, a suspension of cellulose fibers in water is deposited onto a continuously moving forming fabric. The flow and deposition processes during manufacturing induce a preferred fiber orientation. This results in a pronounced mechanical anisotropy and three principal material directions: the machine direction (MD), the cross-machine direction (CD), and the thickness direction (ZD). Because most fibers align in the MD-CD plane, the stiffness in MD is typically two to three times higher than in CD, while the in-plane stiffness exceeds the out-of-plane stiffness by up to two orders of magnitude [Alzweighi et al., 2022; Stenberg et al., 2001b]. Although paper is a fibrous network at the microscale, the focus is restricted to the sheet-level modeling in the following treating it as a homogeneous anisotropic continuum. The strong anisotropy has led many modeling approaches to assume a mechanical decoupling between in-plane and out-of-plane responses, often accompanied by negligible or even vanishing Poisson's ratios between these directions [Stenberg and Fellers, 2002]. This assumption simplifies modeling, but it also highlights the need for formulations that can consistently represent direction-dependent behavior. While the microstructure strongly influences the mechanical response, many engineering simulations treat the material on the macroscale as a homogeneous, anisotropic continuum. This representation allows the effective mechanical behavior of the sheet to be described without explicitly resolving the underlying fiber network, while still capturing the key anisotropic features of the material.

In addition to anisotropy, paperboard exhibits pronounced inelastic behavior. The mechanical behavior depends strongly on environmental conditions such as moisture content and

on the applied deformation rate [Haslach Jr, 2000]. Among the various inelastic effects, plasticity is considered the dominant mechanism governing the irreversible deformation of paperboard [Nygårds, 2008; Alzweighi et al., 2022]. Time-dependent effects, such as creep, further contribute to the overall response and have been studied experimentally, highlighting the importance of visco-elastic effects in the mechanical response [Fadiji et al., 2019; Hussain et al., 2017; Köstner et al., 2017, 2018]. Moreover, damage and failure are critical aspects, particularly in forming processes, where material degradation and crack initiation may limit the performance [Linville and Östlund, 2016]. Together, these aspects significantly increase the complexity of constitutive modeling.

In-plane modeling. Modeling of the in-plane behavior has found wide attention, while both small-strain approaches [Mäkelä and Östlund, 2003; Huang and Nygårds, 2010; Tjahjanto et al., 2015; Wallmeier et al., 2015; Li et al., 2016; Bedzra et al., 2019; Pfeiffer and Kolling, 2019] and large-strain formulations that are typical for forming processes have been proposed [Xia et al., 2002; Harrysson and Ristinmaa, 2008; Borgqvist et al., 2014, 2015]. Inelasticity is typically modeled within anisotropic elasto-plastic frameworks. A widely used approach employs orthotropic yield functions based on Hill's criterion [Hill, 1952], which has been applied to paperboard in several studies [Hagman and Nygårds, 2017; Huang and Nygårds, 2010; Huang et al., 2014]. Alternative formulations aim to capture the anisotropic plastic response more explicitly. Multi-surface plasticity models provide a comparatively simple way to account for directional differences in yielding by defining separate yield conditions for different loading modes. Such models allow distinctions between tension and compression in both the machine and cross directions while maintaining a relatively small number of material parameters [Wallmeier et al., 2015; Bedzra et al., 2019]. A more elaborate approach was proposed by [Xia et al., 2002] who introduced a yield surface composed of six sub-surfaces describing the onset of plasticity under different in-plane loading conditions. This formulation has been applied to forming processes such as deep drawing [Nygårds et al., 2005, 2009]. Subsequent extensions incorporated coupled hardening evolution between the individual sub-surfaces [Borgqvist et al., 2014] and anisotropic kinematic hardening [Li et al., 2016; Pfeiffer and Kolling, 2019]. Other constitutive approaches have also been proposed for in-plane behavior. For example, orthotropic models based on plasticity with the Tsai-Wu failure criterion have been formulated to describe the response of fibrous materials [Phongphinitana and Jearanaisilawong, 2012]. The damage mechanism of in-plane paper are accounted for in continuum damage models [Isaksson et al., 2004; Isaksson and Hägglund, 2005; Isaksson et al., 2006; Alzweighi et al., 2023], and, more recently, phase-field formulations have been explored for modeling in-plane deformation mechanisms [Marengo et al., 2024], and additional continuum

approaches have been proposed to capture the mechanical behavior of paper sheets [Alajami et al., 2019].

Out-of-plane modeling. The out-of-plane response of paper and paperboard is essential to accurately describe forming processes. Experimental studies have provided detailed insight into the thickness-direction behavior and its characterization [Nygards et al., 2007; Nygård, 2008; Sanjon et al., 2024]. These investigations reveal that the mechanical response in the thickness direction is inherently inelastic, with pronounced elasto-plastic behavior [Stenberg et al., 2001*a,b*].

Early modeling efforts focused on small-deformation regimes, including the measurement of out-of-plane Poisson's ratios [Stenberg and Fellers, 2002], which resulted in a constitutive model of the thickness response [Stenberg, 2003]. The out-of-plane response is further characterized by anisotropy, particularly under compression and shear loading [Nygård, 2009]. More advanced approaches incorporate these features within elasto-plastic formulations [Li et al., 2018] and multi-surface plasticity models for forming applications such as deep drawing [Wallmeier et al., 2015]. Extensions to large deformation regimes have also been proposed, and recent combined experimental and modeling studies have further improved the understanding of the thickness-direction behavior [Robertsson et al., 2023*a*], and lead to a damage model for creasing and folding [Robertsson et al., 2023*b*].

Coupled modeling. While the works mentioned above demonstrate various and detailed modeling approaches, realistic simulations of forming processes require three-dimensional constitutive formulations capable of describing both in-plane and out-of-plane deformation simultaneously. However, only a limited number of models address the full three-dimensional behavior of paper and paperboard.

Existing approaches often combine in-plane elasto-plasticity with simplified out-of-plane behavior, which is frequently assumed purely elastic, and employ logarithmic strain measures and the yield function by Xia [Xia et al., 2002]. For industrial applications, computational efficiency is a key requirement. This has led to the development of models tailored for large-scale forming simulations, such as the approach in [Robertsson et al., 2018], which combines solid-shell finite elements with constitutive formulations based on the works by Borgqvist et al. [2015] and Borgqvist et al. [2016]. The latter ones describe a framework for large deformations with a decoupled model, where the elastic response is separated to allow distinct in-plane and out-of-plane properties, while plasticity is governed by a single distortional hardening yield surface. This model has been further extended to include rate dependence [Robertsson et al., 2021] and applied to compression and folding processes [Borgqvist et al., 2016].

Another line of research focuses on incorporating rate-dependent and time-dependent effects.

For instance, a viscoelastic-viscoplastic formulation for small deformations has been proposed by [Tjahjanto et al., 2015], including densification effects relevant under compressive loading. More recent developments aim to combine in-plane and out-of-plane behavior within unified formulations, for example by incorporating densification effects through structural tensors [Kopic-Osmanovic et al., 2023]. Despite these advances, many existing approaches either treat in-plane and out-of-plane responses separately or rely on complex coupling strategies. In some cases, the thickness direction is assumed to remain unchanged during deformation, which restricts the ability to fully represent the material response.

These limitations motivate the development of constitutive frameworks that consistently capture the complete three-dimensional deformation behavior. Rather than separating the constitutive description into independent components, an alternative approach is to decompose the kinematics of deformation that will be addressed in this work. This provides a flexible basis for incorporating various inelastic mechanisms while maintaining a coherent and extensible three-dimensional formulation.

1.2.2. Modeling approaches for tofu

The mechanical characterization of tofu and similar food materials has traditionally been approached from a food science perspective. A widely established method is texture profile analysis (TPA), which has been used for several decades to quantify the textural and sensory properties of food [Friedman et al., 1963]. In this approach, mechanical double-compression tests are employed to mimic the chewing process, and allow the extraction of parameters such as hardness, cohesiveness, springiness, and chewiness from the resulting force-time curves [Bourne, 1968; Dunne et al., 2025]. Due to its simplicity and direct relation to sensory perception, TPA remains a standard tool in food characterization [Paredes et al., 2022]. However, the lack of standardized testing protocols limits the comparability of results across different studies [Schreuders et al., 2021]. Furthermore, the analysis of texture profile analysis (TPA) curves in food science provides an opportunity to extract mechanically meaningful information, thereby enabling a consistent interpretation from a materials science perspective. To address these limitations, mechanics-based approaches have been introduced to analyze food materials such as tofu [Kuhl, 2025]. Mechanical testing methods, including tension, compression, and shear experiments, enable the extraction of fundamental material parameters such as stiffness, viscosity, and relaxation times [St. Pierre et al., 2024]. These quantities provide a direct link between the macroscopic mechanical response and the underlying microstructural architecture of the material [Budday et al., 2020]. Embedding the characterization of food materials into the framework of continuum mechanics therefore allows for a

more systematic and physically interpretable description of their behavior [Vervenne et al., 2025a].

As a hydrated protein gel formed by the coagulation of soymilk [Kohyama et al., 1995; Zheng et al., 2020], tofu exhibits a wide range of textures that are strongly influenced by its water content [Chen et al., 2023] and the specifics of the coagulation process [Wang et al., 2018]. This leads to pronounced nonlinear and often rate-dependent mechanical behavior, which requires expert knowledge in classical modeling approaches and is a time-consuming process. Moreover, modeling strategies based on a linear scaling of individual material contributions, such as volume-fraction-based approaches, may be appropriate for simple material behavior, but might fail to capture the complex, nonlinear dependencies that govern hydrated soft materials such as tofu.

These observations highlight that, despite the availability of experimental characterization techniques and initial mechanics-based interpretations, a consistent and physically grounded modeling framework for tofu remains limited. In particular, there is a need for approaches that not only reproduce experimental observations but also provide insight into the relationship between material composition, structure, and mechanical response. In this context, the present work aims to develop a continuum mechanical description of tofu as a hydrated protein gel that enables prediction of its mechanical behavior while providing physically interpretable insights into the underlying structure-property relationships.

1.2.3. Machine learning in constitutive modeling

Constitutive modeling aims to describe material behavior by establishing relationships between physical quantities such as stress and strain through constitutive laws. Traditionally, these relationships are formulated based on predefined functional forms derived from physical principles. In contrast, machine learning provides a complementary perspective by identifying such relationships directly from data, enabling the representation of complex and nonlinear material behavior without prescribing the functional structure a priori.

Machine learning and neural networks. Machine learning algorithms can generally be decomposed into four key components: a dataset, a parametrized model, a cost function, and an optimization procedure [Herrmann and Kollmannsberger, 2024]. During training, the model parameters are adjusted by minimizing a cost function that quantifies the discrepancy between predictions and observed data.

Neural networks constitute the most widely used class of machine learning models. Based on the *universal approximation theorem*, they are capable to approximate almost all functions

given a sufficient wide or deep network architecture [Hornik et al., 1990]. Throughout various neural network architectures, feed-forward neural networks propagate information from the input layer through a series of hidden layers to the output, with depth and width controlling their representational capacity. Networks with multiple hidden layers are commonly referred to as deep learning models [Goodfellow et al., 2016]. For applications involving path-dependent behavior, such as inelastic material responses, recurrent neural networks provide an extension by introducing internal states that evolve with the input sequence. This enables the modeling of path-dependent phenomena by propagating information across multiple steps. Training of neural networks is typically performed using backpropagation, which combines automatic differentiation with gradient-based optimization to efficiently compute parameter updates.

Depending on the available data and the learning objective, different types of learning can be distinguished that include supervised, unsupervised, semi-supervised, and reinforcement learning. The data is typically processed by clustering, classification, generative modeling, and regression. Among these, the latter one is of primary importance in the context of this thesis as it aims to approximate continuous relations of the underlying data.

To ensure reliable model performance, the dataset is typically divided into training, validation, and test sets to enable both prediction quality and generalization capability. A common choice for the cost function is the mean squared error between predicted and target values. Through modification of the model parameters, the optimization criterion is aimed to be minimized [Goodfellow et al., 2016]. A central challenge in this context is the balance between underfitting and overfitting, often addressed through appropriate model design and regularization techniques [McCulloch et al., 2024]. While underfitting can be handled by increased width and depth of the network architecture, regularization modifies the optimization problem to reduce the generalization error, also known as *bias-variance-trade-off* [Burkov, 2019].

Data-driven frameworks in constitutive modeling. The integration of machine learning into constitutive modeling has steadily increased over the past decades. While early contributions date back more than thirty years [Ghaboussi et al., 1991], advances in computational power and data availability have significantly accelerated developments in recent years [Le et al., 2015; Bessa et al., 2017]. This progress has led to a rapidly growing number of methodological frameworks aimed at learning material behavior directly from data, including purely data-driven or model-free approaches [Kirchdoerfer and Ortiz, 2016]. However, such approaches typically lack the incorporation of fundamental physical principles, which limits their robustness, generalization capability, and interpretability.

To address these limitations, recent research increasingly focuses on embedding physical knowl-

edge into machine learning models. One prominent approach is given by physics-informed neural networks (PINNs) [Raissi et al., 2019], which incorporate governing equations and constraints in a weak sense through additional loss function terms.

In contrast, methods that enforce these constraints strongly have increased widely. In this context, several conceptually distinct approaches have been proposed. Thermodynamics-based artificial neural networks (TANNs) explicitly incorporate thermodynamic principles into the model formulation [Masi et al., 2021; Masi and Stefanou, 2022]. Similarly, the EUCLID framework employs unsupervised learning strategies to identify constitutive relations directly from data [Flaschel et al., 2021]. Physics-augmented neural networks (PANNs) embed mechanical constraints such as polyconvexity into the network design [Klein et al., 2022a; Linden et al., 2023], whereas constitutive artificial neural networks (CANNs) focus on learning interpretable constitutive relations by construction [Linka et al., 2021; Linka and Kuhl, 2023].

Extensions based on Kolmogorov–Arnold networks emphasize the learning of explicit functional relationships rather than solely model parameters [Abdolazizi et al., 2025; Thakolkaran et al., 2025]. Comparative studies highlight the diversity of these fields [Taç et al., 2024; Jones and Fuhg, 2026].

Further developments include material fingerprinting [Flaschel et al., 2026], neural network-based finite element implementations [Weber et al., 2021], and input-specific architectures that combine different mapping properties within a single network [Jadoon et al., 2026]. A comprehensive overview of data-driven approaches for both path-independent and path-dependent behavior is provided by Fuhg et al. [2025].

Within these approaches, fundamental requirements such as objectivity, material symmetry, thermodynamic consistency, convexity, and growth conditions are embedded into the model design, ensuring their fulfillment *a priori*. Despite their shared objective of accurately representing material behavior, these frameworks differ fundamentally in the balance between model expressivity and interpretability. Dense neural network architectures, as commonly used in PANNs, offer high approximation capacity but often at the expense of transparency. In contrast, sparse and structured approaches, such as CANNs, prioritize interpretability by enabling the identification of physically meaningful relationships within the data. This distinction becomes particularly relevant in the presence of noisy or uncertain data, where interpretability plays a critical role in ensuring reliable and robust model predictions [Linka et al., 2025].

Inelasticity in neural network-based constitutive models. A broad spectrum of data-driven frameworks has been developed to describe inelastic material behavior, differing in how strongly physical principles are embedded. A central concept shared by many of these approaches is the introduction of internal variables to represent history-dependent effects.

Within this context, thermodynamic consistency can be enforced in a weak sense by augmenting the loss function with additional constraints, as proposed in thermodynamics-informed neural network formulations [Masi et al., 2021]. Similar strategies have been applied to elasto-plasticity [Vlassis and Sun, 2021; Malik et al., 2021], while multiscale extensions employ autoencoder-based architectures to extract internal variables from homogenization simulations [Masi and Stefanou, 2022]. More recently, attention-based neural ordinary differential equations have been introduced, enabling flexible representations of evolving internal states [Jones and Fuhg, 2025].

Thermodynamically consistent neural network models enforce physical constraints strongly and have been developed for elasto-plasticity models at small deformations [Meyer and Ekre, 2023; Fuhg et al., 2023b] and have been extended to finite deformation settings [Jadoon et al., 2025]. A unifying theoretical framework for many of these developments is provided by the theory of generalized standard materials (GSM) [Halphen and Nguyen, 1975; Hackl, 1997]. Within this framework, inelastic behavior is characterized through the use of dissipation potentials that are convex with respect to internal variables and satisfy specific normalization and stationarity conditions. Alternatively, dual dissipation potentials can be formulated in terms of thermodynamic forces, leading to an equivalent representation. This structure provides a natural basis for integrating thermodynamic principles into data-driven constitutive models.

The GSM concept has been employed in several data-driven frameworks, such as EUCLID [Flaschel et al., 2023]. This framework has been applied to various classes of inelastic behavior, including plasticity [Flaschel et al., 2022], non-associated pressure-sensitive plasticity [Xu et al., 2025], and visco-elasticity at small strains [Marino et al., 2023]. Related developments capture visco-elasticity at finite strains without relying on a multiplicative decomposition of the deformation gradient [As’ad and Farhat, 2023, 2026].

Similarly, physics-augmented neural networks incorporate dissipation-based formulations to represent inelastic behavior [Rosenkranz et al., 2024, 2023; Kalina et al., 2026], and have been extended to multi-physics settings to account for coupled phenomena [Klein et al., 2022b, 2024]. Variational approaches such as Onsager-based neural networks (VONNs) utilize potential-based formulations to ensure thermodynamic consistency [Huang et al., 2022].

In the context of GSM, stress-based potentials can be transformed into inelastic strain rate-based potentials via Legendre transformation if both are convex [Holthusen et al., 2025b]. This concept underlies the framework of inelastic constitutive artificial neural networks (iCANNs), introduced by Holthusen et al. [2024b] and has since been applied to a wide range of material behaviors, including viscoelasticity [Holthusen et al., 2024a], growth

and remodeling [Holthusen et al., 2025a], and plasticity [Boes et al., 2025; van der Velden et al., 2025]. Subsequent developments include generalized formulations within the GSM framework [Holthusen et al., 2025b], extensions to anisotropic material behavior [Holthusen and Kuhl, 2026], and the incorporation of multi-physics couplings [Holthusen et al., 2026]. In parallel, related developments such as viscoelastic CANNs (VCANNs) [Abdolazizi et al., 2024] and Kolmogorov–Arnold networks [Ji et al., 2026] have been presented that extend the initial framework to inelastic formulations.

Despite the wide range of existing data-driven methodologies for constitutive modeling, a key challenge remains the development of frameworks that are sufficiently general to capture a broad spectrum of material effects within a unified and physically consistent formulation. In particular, the ability to simultaneously represent nonlinear, path-dependent, and potentially coupled phenomena is essential for the realistic modeling of complex materials.

Within this landscape, inelastic constitutive artificial neural networks (iCANNs) provide a particularly promising foundation. Their formulation combines a rigorous thermodynamic structure with a flexible and interpretable representation, enabling the consistent integration of complex inelastic mechanisms. While the framework has already been successfully applied to several classes of material behavior, the systematic incorporation and analysis of elasto-plasticity remains a crucial step to further complete and consolidate its modeling capabilities. Plasticity plays a fundamental role in many materials due to its governing influence on irreversible deformation, and its inclusion not only enhances predictive performance but also enables the extraction of physically meaningful insights from the learned response.

Beyond classical inelastic effects, many materials exhibit behavior that is additionally influenced by a porous nature, compositional or environmental factors, such as hydration. Incorporating such dependencies into the framework represents an important extension toward more realistic material descriptions.

In combination, the systematic extension of iCANNs toward plasticity and composition-dependent behavior contributes to the development of more general and adaptable constitutive modeling frameworks. Such advancements provide an important step toward improving the predictive modeling of complex materials and, in a broader context, support the efficient and informed use of materials in applications where sustainability plays an increasing role.

1.3. Results and discussion of the research contribution

The contributions of this thesis span both the development of physically motivated constitutive models and their extension through data-driven approaches, with a particular focus on the anisotropy and inelasticity of sustainable materials.

Anisotropic model formulation for fibrous materials. For fibrous materials, a key result is the development of a continuum mechanical framework for anisotropic materials that enables a strict decoupling of in-plane and out-of-plane behavior. This is achieved through a decomposition of the right Cauchy-Green tensor using structural tensors and the introduction of a split into area-preserving and area-changing deformations. The formulation allows the definition of modified invariants that describe in-plane stretch and area change independently of out-of-plane deformation. This has successfully been extended to inelastic behavior, particularly visco-elasticity, where the anisotropy can be modified through different inelastic potentials. Numerical studies confirm that the model produces fully decoupled responses, i.e., in-plane loading does not induce out-of-plane stresses and vice versa, even under finite deformations and localized states.

In contrast to classical approaches, where decoupling is introduced through parameter reduction, the results show that anisotropy and decoupling can be embedded directly at the kinematic level, resulting in a flexible formulation that can be extended to other nonlinear effects, such as plasticity or damage, without altering its fundamental structure. In particular, the framework facilitates the incorporation of additional inelastic effects such as plasticity and damage, which are essential to accurately capture forming and folding processes. In this way, the developed approach contributes to the improved design and optimization of paper-based packaging materials.

Extension of a data-driven approach to plasticity. A further result is the extension of inelastic constitutive artificial neural networks to elasto-plastic material behavior. The developed framework enables the simultaneous identification of elastic and inelastic contributions, including yield behavior, plastic flow, and kinematic hardening, from data. By combining potential-based formulations with neural networks, the model learns both the Helmholtz free energy and the plastic dissipation mechanisms while it ensures thermodynamic consistency. The integration of a return-mapping algorithm allows the enforcement of plastic admissibility conditions during training, and enables stable and physically consistent predictions. Numerical results demonstrate that the framework can reproduce classical plasticity behavior, including nonlinear hardening and experimentally observed responses, while it also highlights limitations to capture highly complex yield surfaces and strongly nonlinear evolution laws.

In particular, the chosen model design allows for the systematic analysis of the influence of specific invariants on the mechanical response. This is of particular importance, as it provides direct insight into which physical quantities govern the observed behavior and thereby significantly enhances the interpretability of the learned models. Consequently, the extension of iCANNs to elasto-plasticity represents not only a methodological advancement but also a significant step toward completing the framework's capability to capture a broader class of inelastic phenomena. This contribution is particularly relevant, as plasticity constitutes a fundamental mechanism in many engineering materials. Its integration therefore enhances both the predictive power and the explanatory value of data-driven constitutive models, strengthening their applicability to complex and realistic material systems.

Discovery of hydrated soft material behavior. Finally, the application of automated model discovery to hydrated soft materials provides new insights into feature-dependent material behavior. Using experimental data, constitutive models are identified that capture strongly nonlinear, rate-dependent, and composition-dependent responses. The results show that the mechanical behavior can be described by a small number of dominant invariants, enabling a clear separation of elastic and inelastic effects. Furthermore, it is demonstrated that variations in water content significantly alter the material response in a highly nonlinear manner, which cannot be captured by classical mixture models. Instead, feature-dependent neural network formulations enable the identification of nontrivial dependencies between composition and mechanical behavior, allowing the formulation of unified models that generalize across different material states.

Taken together, these results demonstrate that the combination of structured constitutive modeling and data-driven approaches enables both accurate prediction and systematic analysis of complex material behavior. In particular, the findings highlight that embedding physical structure into the model formulation is essential to achieve interpretability and stability, while neural networks provide the flexibility required to capture nonlinear and coupled effects. This demonstrates that iCANNs are not limited to a specific class of materials or phenomena, but can be systematically extended to capture a wide range of complex behaviors. As such, the framework establishes a versatile foundation for linking methodological developments with application-oriented modeling, enabling the consistent description of both structural and compositional influences within a single constitutive approach.

From a broader perspective, the research contribution of this thesis lies in bridging the gap between methodological developments in data-driven constitutive modeling and their application to realistic material systems. The work shows that data-driven frameworks, when

combined with physically motivated architectures, can be used not only to reproduce experimental observations but also to identify governing mechanisms such as dominant invariants, directional dependencies, and feature-driven effects. This represents a shift from purely predictive modeling toward a more analytical use of machine learning in material modeling.

At the same time, the results reveal important challenges. The modeling of highly complex inelastic phenomena remains sensitive to the chosen architecture and training procedure, and the representation of complex yield surfaces and strongly nonlinear hardening behavior is still limited. Furthermore, the integration of additional dependencies, such as composition or environmental conditions, increases model complexity and requires careful design to ensure robustness and generalization.

Overall, this thesis establishes a unified and extensible framework for constitutive modeling that combines continuum mechanics and automated model discovery. By demonstrating its applicability to anisotropic fibrous materials, inelastic processes, and feature-dependent soft materials, it provides a foundation for the development of interpretable, flexible, and application-oriented material models.

1.4. Outline of the dissertation

The present cumulative dissertation consists of three peer-reviewed journal articles (Chapters 2, 4 and 5) as well as one contribution to conference proceedings (Chapter 3). The dissertation is structured as follows:

- **Article 1** Chapter 2 discusses a continuum mechanical constitutive model for fibrous materials motivated by the material behavior of paper and paperboard. A decoupled formulation for large strains is introduced that separates the in-plane and out-of-plane or thickness behavior. The separation is performed by modification of the kinematics in terms of the right Cauchy-Green tensor using structural tensors. Further, the Helmholtz free energy is divided into isoplanar and area-changing contributions, in analogy to the classical isochoric volumetric split. The framework is first introduced for elastic material behavior, while its extension to inelastic behavior follows. The latter one is derived using a co-rotated intermediate configuration for a consistent derivation and implementation using automated differentiation. Multiple inelastic potentials that define the evolution of inelastic deformations are introduced to describe different deformation mechanisms. Material point studies and structural simulations are used to examine the decoupling mechanisms of the elastic and various inelastic effects of the proposed formulation.

- **Article 2** Chapter 3 describes an extension of the model presented by Boes et al. [2023]. The derivation of the decoupled material model is presented for visco-elastic material behavior and large strains. It is extended by an additional inelastic potential that accounts anisotropic inelastic strain evolutions. By formulation of this potential, inelastic strains evolve only in the out-of-plane direction, while the in-plane remains unbothered. The model is validated through numerical simulations of a pinched cylinder and underlines the flexibility of the framework presented in Chapter 2.
- **Article 3** Chapter 4 focuses on the development of a neural network-based constitutive framework with the objective of enabling automated model discovery of elasto-plastic material behavior. The approach describes the extension of inelastic constitutive artificial neural networks Holthusen et al. [2024b] to the case of elasto-plastic behavior and combines continuum mechanical principles with the flexibility of neural networks. The latter are used to represent the functional forms of the constitutive relations. The proposed architecture is structured around a thermodynamically consistent formulation and includes neural network representations of the Helmholtz free energy, accounting for both linear and non-linear hardening contributions, as well as the yield surface and plastic potentials governing the evolution equations. To capture path-dependent behavior inherent to inelasticity, the framework is embedded within a recurrent neural network structure. In addition, a return-mapping algorithm is incorporated to ensure plastic admissibility, thereby maintaining consistency with classical computational plasticity concepts. Training on synthetic and experimental data of different materials is performed on varying sets of training data to study the model performance and impact of different amounts of training data.
- **Article 4** Chapter 5 is concerned with the discovery of the mechanics of tofu. Confined compression tests of different types of tofu are conducted at varying strain rates to extract the rate-dependent material behavior and impact of different water contents. The experimental results of each tofu type are used to discover three distinct visco-elastic material models using a generalized data-driven framework introduced by Holthusen et al. [2025a]. Elastic and inelastic terms are separately discovered and a thermodynamically consistent material model and fulfillment of physical principles are ensured by design of the network architectures. Combination of the model with feature networks extends the framework to account for water-content-dependency. Thereby, a unified material model that includes the rate-dependency and changes in the material behavior with varying water contents is discovered. Training and validation is performed on the

experimentally obtained stress-strain data to demonstrate the predictive capabilities and analyze the dependency of the obtained models on different invariants and impact of elastic and inelastic terms included in the model formulation.

Finally, an overall conclusion reviews the presented work and is provided together with an outlook on future topics to be discussed in Chapter 6.

2 | Article 1:

A novel continuum mechanical framework for decoupled material behavior in thickness and in-plane directions

This article was published as:

Boes, B., Simon, J. W., Reese, S., and Holthusen, H. (2023). A novel continuum mechanical framework for decoupled material behavior in thickness and in-plane directions. *Computer Methods in Applied Mechanics and Engineering*, **415**, 116192. <https://doi.org/10.1016/j.cma.2023.116192>.

Disclosure of the individual authors' contributions to the article:

B. Boes reviewed the relevant existing literature, worked together with H. Holthusen on the derivation of the constitutive framework and implemented the material model and the element routine into the finite element software *FEAP*. She set up and performed all simulations, interpreted the results and wrote the preliminary version of the article. All co-authors gave valuable conceptual advice, read the manuscript and gave suggestions for improvement. All authors approved the publication of the final version of the manuscript.

2.1. Abstract

A novel continuum mechanical framework for decoupling the material behavior in in-plane and thickness directions is presented that is motivated by the macroscopic behavior of paper and paperboard.

Due to the manufacturing process of these composite materials and the underlying microstructure resulting thereof, the material response is highly anisotropic and can be divided into in-plane and out-of-plane behavior that is independent of each other. This decoupling is mostly incorporated in existing material models by setting the corresponding material parameters to zero which limits the models to an extension to nonlinear problems and is restrictive from a continuum mechanical point of view.

Therefore, the aim of this paper is to derive a continuum mechanical framework that captures the in-plane and out-of-plane decoupling in a thermodynamically consistent manner for large deformations. Thus, the right Cauchy-Green deformation tensor was split into in-plane and out-of-plane deformations by use of structural tensors that were aligned with the preferential material directions. Therefore, the Helmholtz free energy was formulated in terms of modified invariants of the deformation tensor for in-plane and out-of-plane deformations separately, which lead to a decoupled framework that could easily be modified due to the flexible formulation of the Helmholtz free energy terms. The extension to inelasticity, in particular viscoelasticity here, was straightforward due to the split of the deformation tensor. Finally, the decoupling of the in-plane and out-of-plane response of the presented framework was demonstrated by structural examples for elastic and inelastic material behavior.

The presented model incorporates the split into in-plane and out-of-plane behavior in a continuum mechanical manner, and thus, is a flexible constitutive framework for decoupled material responses that can easily be adjusted to different material behavior.

2.2. Introduction

Fiber-reinforced composites are widely used today due to their high strength-to-weight ratio [Grzesik, 2017]. This holds particularly for paper and paperboard, which can be considered a special case of composite materials consisting of fiber networks on the micro-scale. Although paper and paperboard are mostly used in the packaging industry, a growing interest in other engineering applications emerges due to improved sustainability and recyclability when using renewable resources (Ojha et al. [2015]; Simon [2021]). This leads to a high interest in numerical models that capture the complex and anisotropic macroscopic material behavior in

Nomenclature	
a	Scalar quantity
$\mathbf{a}$	First order tensor
$\mathbf{A}$	Second order tensor
$\mathbb{A}$	Fourth order tensor
$(*)^\cdot$	Material time derivative
$(*)^T$	Transpose of a quantity
$(*)^{-1}$	Inverse of a quantity
$\text{tr}(*)$	Trace of a quantity
$\text{cof}(*)$	Cofactor of a quantity
$\det(*)$	Determinant of a quantity
$\ln(*)$	Logarithm of a quantity
$\text{sym}(*)$	Symmetric term of a quantity
$\mathbf{A} : \mathbf{B}$	Scalar product of two tensors $\mathbf{A}$ and $\mathbf{B}$ ($\mathbf{A} : \mathbf{B} = A_{ij}B_{ij}$)
$\mathbf{A} \otimes \mathbf{B}$	Dyadic product of two tensors $\mathbf{A}$ and $\mathbf{B}$
$(*)_{\text{e}}$	Elastic term of a quantity $(*)$
$(*)_{\text{i}}$	Inelastic term of a quantity $(*)$
$(*)_{\text{isoA}}$	Isoplanar (area-preserving) term of a quantity
$(*)_{\text{A}}$	Area-changing term of a quantity
$(*)_{\text{isov}}$	Isochoric (volume-preserving) term of a quantity
$(*)_{\text{V}}$	Volumetric (volume-changing) term of a quantity
$(*)_{\text{ip}}$	In-plane part of a quantity
$(*)_{\text{op}}$	Out-of-plane part of a quantity
cc	Current configuration
cic	Co-rotated intermediate configuration
ic	Intermediate configuration
rc	Reference configuration
$\tilde{(*)}$	Quantity defined in the ic
$\bar{(*)}$	Quantity defined in the cic
$\hat{(*)}$	Quantity defined in the cc

order to predict the material response under loading.

On the macroscale, paper is an orthotropic material and can be divided into in-plane (ip) and out-of-plane directions (op) as illustrated in Fig. 3.1. The op -direction - defined here in $\mathbf{n}_3$ direction - is orthogonal to the area spanned by the ip -directions which is introduced here by the director vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ and indicated by the grey area. The anisotropy results from the manufacturing process, where the fibers mostly align with the ip directions. As shown experimentally, e.g., by Stenberg and Fellers [2002], the ip and op behaviors are independent of each other, i.e., ip loadings do not lead to op stress states and vice versa. Thus, in contrast to most other materials, a decoupled material behavior for ip and op directions can be assumed.

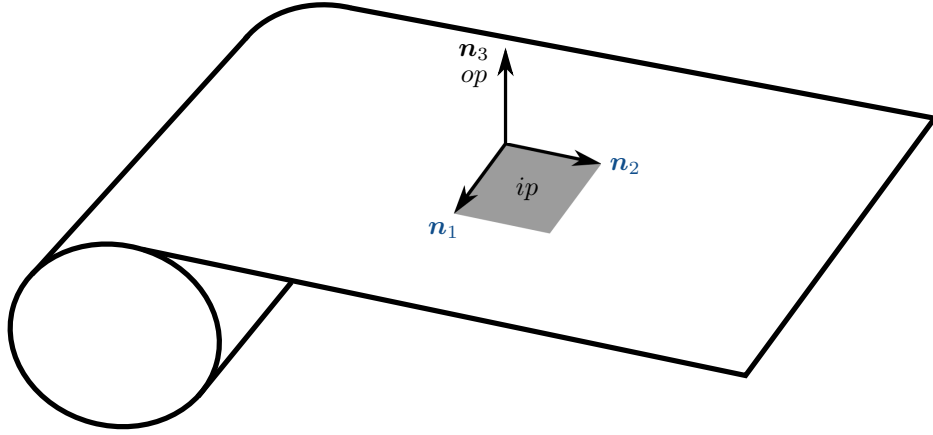


Figure 2.1.: Illustration of preferential material directions – in-plane area (*ip*) underlined in grey and out-of-plane direction (*op*).

This rather complex behavior is a challenge from the material modeling point of view. Up to now, several models exist that either capture the *ip* or the *op* behavior separately. The *ip* elasto-plastic model proposed by Xia et al. [2002] has found wide acceptance by several authors, for example, Nygård et al. [2005, 2009]; Borgqvist et al. [2014, 2015]; Tjahjanto et al. [2015]; Li et al. [2016]; Pfeiffer and Kolling [2019], to name a few. Other *ip* models capture elasto-plasticity by formulation of the yield criterion based on the Tsai-Wu criterion as employed by Harrysson and Ristinmaa [2007, 2008]. Huang and Nygård [2010] model the *ip* response by use of a quadratic Hill criterion and Mäkelä and Östlund [2003] use the concept of an isotropic plasticity equivalent material. On the contrary, the *op* material response is captured in models by, e.g., Stenberg [2003]; Nygård [2009]; Tjahjanto et al. [2015]; Li et al. [2018]. Only few models cover both, the *ip* and *op* behaviors, in a combined manner (cf. Nygård et al. [2009]; Huang and Nygård [2010]; Borgqvist et al. [2015]). In the first two models mentioned, the decoupling is incorporated by setting the coupling material parameters in the material tangent of the standard (coupled) model, for example, the *op* Poisson's ratio, to zero. While this approach works well in the small deformation regime, it cannot do justice to the complexity when material and geometric nonlinearities are considered. Then, a more general continuum mechanical framework is required so that large deformations can be taken into account in a thermodynamically consistent manner.

To the author's best knowledge, the only decoupled model for fibrous materials that is applicable also for large deformations has been proposed by Borgqvist et al. [2015] – extended to rate-dependent behavior by Robertsson et al. [2021] – and deals with elastic-plastic deformations. Anisotropy is incorporated here by use of three structural tensors that are aligned with the preferential material directions where the *op* direction is orthogonal to the *ip* ones.

Thus, the Helmholtz free energy is additively decomposed into one term for each preferential material direction. Furthermore, the decomposition is applied to the elastic energy formulation only. In contrast, our approach is based on a decomposition of the right Cauchy-Green deformation tensor by use of structural tensors that is incorporated into the Helmholtz free energy. With this, the energy formulation is directly divided into *ip* and *op* terms and can easily be extended to decoupled inelastic material behavior. In addition, a split of the deformation tensor into area-changing and area-preserving deformations is introduced in analogy to the classical isochoric-volumetric split for volumes.

One related modeling strategy is the modeling of generalized interfaces, as for example in Javili et al. [2014]; Esmaeili et al. [2017]; Chatzigeorgiou et al. [2017]. It is motivated from processes on the surfaces of solid bodies that result in surface properties that differ from the rest of the body due to, for example, oxidation, ageing and coating processes (c.f. Javili et al. [2013]). In the interface continuum theory, the interface is described with its own energetic structure as the thermo-mechanical properties might differ from the one of the bulk (c.f. Javili et al. [2014]). In addition, a link to surface elasticity can be made, see for example the works of Steigmann and Ogden [1999]; Steigmann [2009]. In the framework presented in this work, the basic continuum mechanical approach remains equal to the classical continuum mechanical approach and the decoupling of in-plane and out-of-plane behavior is introduced through a multiplicative split of the kinematics and an additional additive split of the Helmholtz free energy. With this, the weak form and discretization remains similar to a classical continuum mechanical model.

The novelty of this work, therefore, is the use of classical continuum mechanical approaches while the decoupling is included through the splits of the kinematics and the energy formulation. These two simple modifications directly lead to a decoupled material model in *ip* and *op* directions and make the model flexible for adjustments such that an extension to inelastic material behavior, such as visco-elasticity or plasticity, and adaption to different materials is straightforward. This flexibility is a major advantage of the presented framework as, for example, the extension to elasto-plasticity is an important aspect in modeling the material behavior of paper (c.f. Xia et al. [2002]; Nygård et al. [2005]; Borgqvist et al. [2014] and more). Thus, the modeling strategy of surface elasticity is restrictive from this point of material properties. In addition, as we include inelasticity through a multiplicative split of the deformation gradient, we favor to derive the model in a continuum mechanical approach rather than by use of membrane theory as used, for example, in Steigmann [2009]. Hence, the presented model is an enhancement of a continuum mechanical approach without loss of simple adaptations for different material behavior.

As described above, a continuum mechanical model that incorporates the decoupling of the *ip* and *op* material responses in a thermodynamically consistent manner is desired in a flexible approach such that it can easily be adapted to different material behaviors, for example, elasto-plasticity or visco-elasticity. For anisotropic, continuum mechanical models, a decomposition into fibre and matrix material response is common (see for example Spencer [2001]; Reese [2003]; Sansour and Bocko [2003]). We adapt this approach here for paper and paperboard and propose a material model for the matrix response as a first step towards a sophisticated material model for paper and paperboard. The extension to a fibre-matrix material to capture the in-plane anisotropy is then straightforward. However, this lies out of the scope of this work as our focus is on deriving a general framework for decoupling the *ip* and *op* behavior whereas the adaption to paper and paperboard will be captured in future works.

The starting point of the derivation is the Helmholtz free energy, that can be defined in terms of invariants of strain measures, for example, the right Cauchy-Green deformation tensor. It is decomposed into *ip* and *op* terms by introduction of structural tensors that are aligned with the preferential material directions. With this, a modified strain measure is formulated that captures *ip* and *op* strains separately and, thus, leads to a decomposition of the energy formulation into *ip* and *op* terms. As a result, the modification of the material model to different material behaviors and the extension to inelasticity, in particular viscoelasticity here, is straightforward. The concept of a co-rotated intermediate configuration is used for the derivation and implementation as it is - in contrast to the standard intermediate configuration - uniquely defined.

In the following, the framework for decoupling the *ip* and *op* behaviors is derived for elastic and inelastic material behaviors in Chapter 2.3 and Chapter 2.4, respectively. In Chapter 2.5, specific forms of the Helmholtz free energy are introduced whereas numerical examples for both material formulations are presented in Chapter 2.6. Finally, a conclusion and a short outlook are stated.

2.3. Elastic material behavior

Our idea for decoupling the material response in *ip* and *op* directions is based on a modification of the right Cauchy-Green deformation tensor, C , by means of structural tensors aligned with the *ip* and *op* directions. In analogy to the classical isochoric-volumetric split, a decomposition into isoplanar (area-preserving) and area-changing contributions is performed. This approach is general and can, therefore, be applied to numerous continuum mechanical material models. In the following, we focus on elastic material behavior before we move on to more complex models in the next chapter.

2.3.1. Classical isochoric-volumetric split

As mentioned above, the later on introduced framework is based on the idea of the classical isochoric-volumetric split, briefly described, for example, in Holzapfel [2002]. Instead of considering the total deformation gradient, $\mathbf{F}$, a split into volume-changing and volume-preserving terms can be performed by use of the volume ratio $J_{\mathbf{F}} := \det(\mathbf{F})$ which has originally been introduced by Flory [1961] and applied by, for example, Lubliner [1985]; Simo and Taylor [1991]; Ogden [1997]. Thus, the deformation gradient can be decomposed into $\mathbf{F} = \left(J_{\mathbf{F}}^{1/3} \mathbf{I}\right) \left(J_{\mathbf{F}}^{-1/3} \mathbf{F}\right)$. Due to reasons of objectivity, we will focus on the right Cauchy-Green tensor, $\mathbf{C} := \mathbf{F}^T \mathbf{F}$, rather than the deformation gradient, and thus, from now on focus on the formulation in terms of $\mathbf{C}$. In analogy to the deformation gradient, the right Cauchy-Green deformation tensor can be split multiplicatively into terms associated with volume preserving deformations, $\mathbf{C}_{\text{isov}}$, and volume changing deformations, $\mathbf{C}_V$. For the volume-preserving terms, it holds that $\left(\det \left(J_{\mathbf{F}}^{-1/3} \mathbf{F}\right)\right)^2 = \det \left(J^{-1/3} \mathbf{C}\right) = 1$. The decomposition is defined by

$$\mathbf{C} = \underbrace{J^{-1/3} \mathbf{C}}_{=: \mathbf{C}_{\text{isov}}} \underbrace{J^{1/3} \mathbf{I}}_{=: \mathbf{C}_V} \quad (2.1)$$

where $J = \det(\mathbf{C})$ is the Jacobian determinant or squared volume ratio and gives information about the volume change. It is also known as the third of three principal invariants that read

$$I_1 = \text{tr}(\mathbf{C}) \quad (2.2)$$

$$I_2 = \frac{1}{2} \left(\text{tr}(\mathbf{C})^2 - \text{tr}[\mathbf{C}^2] \right) \quad (2.3)$$

$$J := I_3 = \det(\mathbf{C}). \quad (2.4)$$

For a constitutive model that captures the *ip* behavior separated from the *op* response, an analogous formulation that deals with area changes instead of volume changes is searched for. Thus, a term that describes the area change is described in the following.

2.3.2. Decoupling of the in-plane and out-of-plane behaviors

To capture the deformations of an area only, the in-plane deformations must be separated from the out-of-plane ones. For this, director vectors, $\mathbf{n}_i$, of unit length are defined aligned with the preferential material directions in the reference configuration (*rc*) as applied in e.g. Boehler [1979]; Zhang and Rychlewski [1990]; Svendsen [1994]; Holthusen et al. [2020]. As shown in Fig. 2.1, the indices $i = 1, 2, 3$ refer to the preferential material directions, such that the

ip behavior is defined by the 1- and 2-directions, whereas the 3-direction refers to the *op*-behavior. Further, it is assumed that the *op* direction remains orthogonal to the *ip* area, such that $\mathbf{n}_3 = \mathbf{n}_1 \times \mathbf{n}_2$ and $\mathbf{n}_3 \cdot \mathbf{n}_1 = \mathbf{n}_3 \cdot \mathbf{n}_2 = 0$.

From this, a set of second-order structural tensors, $\mathbf{M}_i$, are defined by dyadic products, $\mathbf{M}_i = \mathbf{n}_i \otimes \mathbf{n}_i$. Moreover, the in-plane structural tensor, $\mathbf{\Omega}$, is introduced, which describes the entire in-plane behavior, such that only one single tensor is needed instead of $\mathbf{M}_1$ and $\mathbf{M}_2$. For this, the *op* structural tensor $\mathbf{M}_3$ is deduced from the identity tensor, $\mathbf{I}$, which includes all preferential directions. In consequence, the in-plane tensor in the reference configuration is expressed as

$$\mathbf{\Omega} = \mathbf{I} - \mathbf{M}_3. \quad (2.5)$$

It holds that $\mathbf{\Omega} = \mathbf{M}_1 + \mathbf{M}_2$.

2.3.2.1. Isoplanar-area-changing split for the *ip* behavior

The introduced *ip* structural tensor, $\mathbf{\Omega}$, can be applied onto the right Cauchy-Green deformation tensor, $\mathbf{C}$. Consequently, *ip* stretches can be described separately from *op* stretches. For the special case of $\mathbf{C}$ and $\mathbf{\Omega}$ being coaxial, the following modified invariants for the *ip* response result that can be interpreted physically, i.e.

$$I'_1 = \text{tr}(\mathbf{C}\mathbf{\Omega}) = \lambda_1 + \lambda_2 \quad (2.6)$$

$$I'_2 = \frac{1}{2} (\text{tr}(\mathbf{C}\mathbf{\Omega})^2 - \text{tr}[(\mathbf{C}\mathbf{\Omega})^2]) = \lambda_1 \lambda_2 \quad (2.7)$$

$$I'_3 = \det(\mathbf{C}\mathbf{\Omega}) = 0 \quad (2.8)$$

where λ_i are the eigenvalues of $\mathbf{C}$. As the *op* terms are deducted, the third modified invariant I'_3 is zero and the first and second one give information about the *ip* length and area change, respectively. Proof, that I'_2 is the area change in general and not only for coaxiality of the right Cauchy-Green tensor and the *ip* structural tensor can be shown by Nansons formula. The latter describes the variation from an area element in the reference configuration, $d\mathbf{A} = dA \mathbf{n}_3$, to an area element in the current configuration, $d\mathbf{a} = da \hat{\mathbf{n}}_3$, i.e., $da \hat{\mathbf{n}}_3 = \det(\mathbf{F}) \mathbf{F}^{-T} \mathbf{n}_3 dA$ with $\mathbf{F}$ being the deformation gradient and $\hat{\mathbf{n}}_3$ the *op* director vector in the *cc*. It holds that $\|\mathbf{n}_3\| = \|\hat{\mathbf{n}}_3\| = 1$. By using $\hat{\mathbf{n}}_3 \cdot \hat{\mathbf{n}}_3 = 1$ and $\text{cof}(\mathbf{C}) := \det(\mathbf{C}) \mathbf{C}^{-1}$, the squared area change is then described by

$$da^2 = \mathbf{n}_3 \cdot \text{cof}(\mathbf{C}) \cdot \mathbf{n}_3 dA^2 \quad (2.9)$$

Rewriting Equation (2.7) by use of Equation (2.5), $\text{cof}(\mathbf{C}) = I_2(\mathbf{C})\mathbf{I} - I_1(\mathbf{C})\mathbf{C} + \mathbf{C}^2$, and $\text{tr}(\mathbf{M}_3) = 1$ yields

$$\begin{aligned}
 I'_2(\mathbf{C}\mathbf{\Omega}) &= \frac{1}{2} (\text{tr}(\mathbf{C}\mathbf{\Omega})^2 - \text{tr}[(\mathbf{C}\mathbf{\Omega})^2]) \\
 &= \frac{1}{2} (\text{tr}(\mathbf{C} - \mathbf{C}\mathbf{M}_3)^2 - \text{tr}[(\mathbf{C} - \mathbf{C}\mathbf{M}_3)^2]) \\
 &= I_2(\mathbf{C}) - I_1(\mathbf{C})\text{tr}(\mathbf{C}\mathbf{M}_3) + \text{tr}(\mathbf{C}^2\mathbf{M}_3) \\
 &= \mathbf{n}_3 \cdot \text{cof}(\mathbf{C}) \cdot \mathbf{n}_3.
 \end{aligned} \tag{2.10}$$

Thus, it is shown that the second modified invariant, I'_2 , reveals the area ratio during deformations. With this, a formulation for area changes analogously to the volume ratio J in Equation (2.4) has been found. As this area ratio will be used in a similar manner as the volume ratio J , we further denote it by J_A for readability and comparability. Therefore, we define

$$J_A := I'_2. \tag{2.11}$$

In analogy to the classical isochoric-volumetric split, a modified decomposition of the right Cauchy-Green deformation tensor, $\mathbf{C}$, is introduced. In Equation (2.1), the decomposition is defined by the squared volume change $\det(\mathbf{C}) = (dv/dV)^2$. We further decompose the volumetric term by using the relation $dv/dV = (d\mathbf{a}^T/d\mathbf{A}^T) \cdot (d\mathbf{l}/d\mathbf{L})$ such that we end up with the definition $\mathbf{C} = J^{-1/3} (d\mathbf{l}/d\mathbf{L})^{1/3} \mathbf{C} (J_A)^{1/3} \mathbf{I}$ which comes from Equation (2.1). By use of this relation, the definition of the volume ratio and adapting the exponent to $1/2$ as we consider area changes instead of volume changes, we motivate the introduction of the isoplanar-area-changing split by

$$\mathbf{C} = \underbrace{J_A^{-1/2} \mathbf{C}}_{=: \mathbf{C}_{\text{iso}_A}} \underbrace{J_A^{1/2} \mathbf{I}}_{=: \mathbf{C}_A}. \tag{2.12}$$

We introduce here a decomposition into isoplanar (area-preserving) terms, $\mathbf{C}_{\text{iso}_A}$, and area-changing terms, $\mathbf{C}_A$, motivated from the isochoric-volumetric split. To prove the characteristic of area-preserving deformations – which is similar to $\det(\mathbf{C}_{\text{iso}_V}) = 1$ for volumes – the isoplanar term of the right Cauchy-Green tensor, $\mathbf{C}_{\text{iso}_A}$, can be inserted into the second modified

invariant(c.f. Equation (2.7)), yielding

$$\begin{aligned}
 J_A(\mathbf{C}_{\text{iso}_A}) &= \frac{1}{2} \left(\text{tr}(\mathbf{C}_{\text{iso}_A} \boldsymbol{\Omega})^2 - \text{tr}[(\mathbf{C}_{\text{iso}_A} \boldsymbol{\Omega})^2] \right) \\
 &= \frac{1}{2} \left(J_A^{-1} \text{tr}(\mathbf{C} \boldsymbol{\Omega})^2 - J_A^{-1} \text{tr}[(\mathbf{C} \boldsymbol{\Omega})^2] \right) \\
 &= \frac{1}{J_A} J_A \equiv 1.
 \end{aligned} \tag{2.13}$$

Thus, the change in the area remains the same.

2.3.2.2. Clausius-Duhem inequality

In analogy to the classical isochoric-volumetric split and by using Equation (2.12), an *ip* Helmholtz free energy, ψ , is additively split into an isoplanar and area-changing term, ψ_{iso_A} and ψ_A , respectively, i.e.,

$$\psi(\mathbf{C}) = \psi_{\text{iso}_A}(\mathbf{C}_{\text{iso}_A}) + \psi_A(\mathbf{C}_A). \tag{2.14}$$

Starting with the Clausius-Duhem inequality for isothermal processes the material time derivative of ψ is calculated using the chain rule, i.e.

$$\dot{\psi} = \frac{\partial \psi}{\partial \mathbf{C}_{\text{iso}_A}} : \dot{\mathbf{C}}_{\text{iso}_A} + \frac{\partial \psi}{\partial \mathbf{C}_A} : \dot{\mathbf{C}}_A. \tag{2.15}$$

For this, the derivatives of the modified right Cauchy-Green tensors $\mathbf{C}_{\text{iso}_A}$ and $\mathbf{C}_A$ are derived by

$$\dot{\mathbf{C}}_{\text{iso}_A} = \frac{\partial \mathbf{C}_{\text{iso}_A}}{\partial \mathbf{C}} : \dot{\mathbf{C}} = J_A^{-\frac{1}{2}} \underbrace{\left(\mathbb{I} - \frac{1}{2} J_A^{-1} \mathbf{C} \otimes \frac{\partial J_A}{\partial \mathbf{C}} \right)}_{=:\mathbb{P}_{\text{iso}_A}^T} : \dot{\mathbf{C}} \tag{2.16}$$

$$\dot{\mathbf{C}}_A = \frac{\partial \mathbf{C}_A}{\partial \mathbf{C}} : \dot{\mathbf{C}} = \frac{1}{2} J_A^{-\frac{1}{2}} \underbrace{\left(\mathbf{I} \otimes \frac{\partial J_A}{\partial \mathbf{C}} \right)}_{=:\mathbb{P}_A^T} : \dot{\mathbf{C}}. \tag{2.17}$$

For comparability, the derivatives in the isochoric-volumetric framework are given in 2.8.1. $\mathbb{I}$ is the fourth order identity tensor. Similar to the classical isochoric-volumetric split, projection tensors in the reference configuration for the new split are defined here by $\mathbb{P}_{\text{iso}_A}$ and $\mathbb{P}_A$. The

derivative of the area ratio J_A with respect to the right Cauchy-Green tensor is

$$\frac{\partial J_A}{\partial \mathbf{C}} = \text{tr}(\mathbf{C}\boldsymbol{\Omega})\boldsymbol{\Omega} - \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega}. \quad (2.18)$$

By use of the Clausius-Duhem inequality, Equation (2.14) to (2.17) and the standard procedure by Coleman and Noll [1961], the second Piola-Kirchhoff stress tensor is defined by

$$\mathbf{S} := 2 \frac{\partial \psi}{\partial \mathbf{C}} = 2 \left(\frac{\partial \psi_{\text{isoA}}}{\partial \mathbf{C}} + \frac{\partial \psi_A}{\partial \mathbf{C}} \right) =: \mathbf{S}_{\text{isoA}} + \mathbf{S}_A. \quad (2.19)$$

It can be additively split into a purely isoplanar contribution, $\mathbf{S}_{\text{isoA}}$, and an area-changing term, $\mathbf{S}_A$, due to the split of the Helmholtz free energy (2.14). They are defined through the fourth-order projection tensors by

$$\mathbf{S}_{\text{isoA}} = 2 \frac{\partial \psi}{\partial \mathbf{C}_{\text{isoA}}} : \frac{\partial \mathbf{C}_{\text{isoA}}}{\partial \mathbf{C}} = \mathbb{P}_{\text{isoA}} : \mathbf{S}_{\text{isoA}}^* \quad (2.20)$$

$$\mathbf{S}_A = 2 \frac{\partial \psi}{\partial \mathbf{C}_A} : \frac{\partial \mathbf{C}_A}{\partial \mathbf{C}} = \mathbb{P}_A : \mathbf{S}_A^* \quad (2.21)$$

where $\mathbf{S}_{\text{isoA}}^* = 2 \frac{\partial \psi_{\text{isoA}}}{\partial \mathbf{C}_{\text{isoA}}}$ and $\mathbf{S}_A^* = 2 \frac{\partial \psi_A}{\partial \mathbf{C}_A}$.

The main equations of the isoplanar-area-changing split derived above are summarized and compared to the classical isochoric-volumetric split in Table 2.1.

Remark 1 The decoupled framework presented here is introduced such that the *ip* behavior can be described independently from the *op* one and vice versa. Thus, when only *ip* behavior is described no *op* response should be observed. It can be shown by defining the *ip* Helmholtz free energy, for example, in terms of the first two modified invariants I'_1 and $J_A := I'_2$. Recall that $I'_3 = 0$. Thus, we can write $\psi = \psi^*(I'_1, J_A)$. With this and Equation (2.19), the second Piola-Kirchhoff stress tensor is then defined by

$$\begin{aligned} \mathbf{S} &= 2 \left(\frac{\partial \psi}{\partial I'_1} \frac{\partial I'_1}{\partial \mathbf{C}} + \frac{\partial \psi}{\partial J_A} \frac{\partial J_A}{\partial \mathbf{C}} \right) \\ &= 2 \left(\frac{\partial \psi}{\partial I'_1} \left[\frac{\boldsymbol{\Omega}}{\sqrt{J_A}} - \frac{I'_1}{2J_A^{3/2}} (I'_1 \boldsymbol{\Omega} - \boldsymbol{\Omega} \mathbf{C} \boldsymbol{\Omega}) \right] + \frac{\partial \psi}{\partial J_A} (I'_1 \boldsymbol{\Omega} - \boldsymbol{\Omega} \mathbf{C} \boldsymbol{\Omega}) \right). \end{aligned} \quad (2.22)$$

This expression is directly derived from the definition of the Helmholtz free energy in the decoupled framework. As it is rather complex, its form can not intuitively be described. Hence, a simple introduction of the decoupled stress tensor is not possible. However, from the simple introduction of the *ip* structural tensor, $\boldsymbol{\Omega}$, and the isoplanar-area-changing split,

Table 2.1.: Overview of equations of the classical isochoric-volumetric split to the isoplanar-area-changing split.

Classical isochoric-volumetric split	Isoplanar-area-changing split
Kinematics:	
$\mathbf{C} = \underbrace{J^{-1/3}\mathbf{C}}_{\mathbf{C}_{\text{isov}}} \underbrace{J^{1/3}\mathbf{I}}_{\mathbf{C}_V} = J^{1/3}\mathbf{C}_{\text{isov}}$	$\mathbf{C} = \underbrace{J_A^{-1/2}\mathbf{C}}_{\mathbf{C}_{\text{isoA}}} \underbrace{J_A^{1/2}\mathbf{I}}_{\mathbf{C}_A} = J_A^{1/2}\mathbf{C}_{\text{isoA}}$
$J = \det(\mathbf{C})$	$J_A = \frac{1}{2} [\text{tr}(\mathbf{C}\mathbf{\Omega})^2 - \text{tr}([\mathbf{C}\mathbf{\Omega}]^2)]$ $= I_2(\mathbf{C}) - I_1(\mathbf{C}) \text{tr}(\mathbf{C}\mathbf{M}_3) + \text{tr}(\mathbf{C}^2\mathbf{M}_3)$
$J(\mathbf{C}_{\text{isov}}) = 1$	$J_A(\mathbf{C}_{\text{isoA}}) = 1$
Helmholtz free energy:	
$\psi(\mathbf{C}) = \psi_{\text{isov}}(\mathbf{C}_{\text{isov}}) + \psi_V(J)$	$\psi(\mathbf{C}) = \psi_{\text{isoA}}(\mathbf{C}_{\text{isoA}}) + \psi_A(J_A)$
$\dot{\psi} = \frac{\partial \psi}{\partial \mathbf{C}_{\text{isov}}} : \dot{\mathbf{C}}_{\text{isov}} + \frac{\partial \psi}{\partial J} \dot{J}$	$\dot{\psi} = \frac{\partial \psi}{\partial \mathbf{C}_{\text{isoA}}} : \dot{\mathbf{C}}_{\text{isoA}} + \frac{\partial \psi}{\partial J_A} \dot{J}_A$
$\dot{J} = J\mathbf{C}^{-1} : \dot{\mathbf{C}}$	$\dot{J}_A = (\text{tr}(\mathbf{C}\mathbf{\Omega})\mathbf{\Omega} - \mathbf{\Omega}\mathbf{C}\mathbf{\Omega}) : \dot{\mathbf{C}}$
$\dot{\mathbf{C}}_{\text{isov}} = J^{-1/3}\mathbb{P}_{\text{isov}} : \dot{\mathbf{C}}$	$\dot{\mathbf{C}}_{\text{isoA}} = J_A^{-1/2}\mathbb{P}_{\text{isoA}} : \dot{\mathbf{C}}$
Projection tensors:	
$\mathbb{P}_V = \frac{1}{3}J^{1/3}(\mathbf{C}^{-1} \otimes \mathbf{I})$	$\mathbb{P}_A = \frac{1}{2}J_A^{-1/2}\left(\frac{\partial J_A}{\partial \mathbf{C}} \otimes \mathbf{I}\right)$
$\mathbb{P}_{\text{isov}} = J^{-1/3}\left(\mathbb{I} - \frac{1}{3}\mathbf{C}^{-1} \otimes \mathbf{C}\right)$	$\mathbb{P}_{\text{isoA}} = J_A^{-1/2}\left(\mathbb{I} - \frac{1}{2}J_A^{-1}\frac{\partial J_A}{\partial \mathbf{C}} \otimes \mathbf{C}\right)$

its form comes rather handy. In 2.8.2, the decoupling of the second Piola-Kirchhoff stress is shown in more detail.

Remark 2 In the classical framework, the relation $\mathbb{P}_V^T : \mathbb{P}_{\text{isov}}^T : \mathbf{B} = 0$ holds true for any arbitrary tensor $\mathbf{B}$. Interestingly, a similar relation is obtained for the newly introduced projection tensors $\mathbb{P}_A$ and $\mathbb{P}_{\text{isoA}}$ such that

$$\mathbb{P}_A^T : \mathbb{P}_{\text{isoA}}^T : \mathbf{B} = \frac{1}{2}J_A^{-\frac{1}{2}}\left(\mathbf{I} \otimes \frac{\partial J_A}{\partial \mathbf{C}}\right) : J_A^{-\frac{1}{2}}\left(\mathbb{I} - \frac{1}{2}J_A^{-1}\mathbf{C} \otimes \frac{\partial J_A}{\partial \mathbf{C}}\right) : \mathbf{B} \quad (2.23)$$

$$= \frac{1}{2}J_A^{-1}\left[\text{tr}\left(\mathbf{B} \frac{\partial J_A}{\partial \mathbf{C}}\right)\mathbf{I} - \frac{1}{2}J_A^{-1}\text{tr}\left(\mathbf{B} \frac{\partial J_A}{\partial \mathbf{C}}\right)\text{tr}\left(\mathbf{C} \frac{\partial J_A}{\partial \mathbf{C}}\right)\mathbf{I}\right] = 0 \quad (2.24)$$

since $\text{tr}\left(\mathbf{C} \frac{\partial J_A}{\partial \mathbf{C}}\right) = 2J_A$.

Remark 3 Similarly, according to Holzapfel [2002], the relation $(\mathbb{P}_{\text{isov}} : \mathbf{B}) : \mathbf{C} = 0$ holds true in the classical framework for any arbitrary tensor $\mathbf{B}$. Again, a similar relation is

obtained for the newly introduced projection tensor $\mathbb{P}_{\text{isoA}}$, i.e.

$$(\mathbb{P}_{\text{isoA}} : \mathbf{B}) : \mathbf{C} = J_A^{-\frac{1}{2}} \left(\mathbf{B} - \frac{1}{2} J_A^{-1} \text{tr}(\mathbf{B}\mathbf{C}) \frac{\partial J_A}{\partial \mathbf{C}} \right) : \mathbf{C} = 0. \quad (2.25)$$

For completeness, the area-preserving projection tensor in the current configuration $\mathbb{P}_{\text{isoA}}^{\text{cc}}$ can be derived in analogy to the approach in Equation (2.25) and by use of the push forward definition between the second Piola-Kirchhoff stress tensor, $\mathbf{S}$, and the Kirchhoff stress tensor, $\boldsymbol{\tau}$, i.e. $\mathbf{S} = \mathbf{F}^{-1} \boldsymbol{\tau} \mathbf{F}^{-T}$, which leads to

$$\begin{aligned} \mathbb{P}_{\text{isoA}} : \mathbf{S} : \mathbf{C} &= J_A^{-\frac{1}{2}} \left(\mathbf{S} - \frac{1}{2} J_A^{-1} \text{tr}(\mathbf{S}\mathbf{C}) \frac{\partial J_A}{\partial \mathbf{C}} \right) : \mathbf{C} \\ &= J_A^{-\frac{1}{2}} \left(\mathbf{F}^{-1} \boldsymbol{\tau} \mathbf{F}^{-T} - \frac{1}{2} J_A^{-1} \text{tr}(\boldsymbol{\tau}) \frac{\partial J_A}{\partial \mathbf{C}} \right) : \mathbf{F}^T \mathbf{F} \\ &= J_A^{-\frac{1}{2}} \mathbf{F}^{-1} \left(\boldsymbol{\tau} - \frac{1}{2} J_A^{-1} \text{tr}(\boldsymbol{\tau}) \mathbf{F} \frac{\partial J_A}{\partial \mathbf{C}} \mathbf{F}^T \right) \mathbf{F}^{-T} : \mathbf{F}^T \mathbf{F} \\ &= J_A^{-\frac{1}{2}} \left(\boldsymbol{\tau} - \frac{1}{2} J_A^{-1} \text{tr}(\boldsymbol{\tau}) \mathbf{F} \frac{\partial J_A}{\partial \mathbf{C}} \mathbf{F}^T \right) : \mathbf{I} = 0. \end{aligned} \quad (2.26)$$

From this, the area-preserving projection tensor in the current configuration can be defined as

$$\mathbb{P}_{\text{iso}}^{\text{cc}} = \mathbb{I} - \frac{1}{2} J_A^{-1} \mathbf{F} \frac{\partial J_A}{\partial \mathbf{C}} \mathbf{F}^T \otimes \mathbf{I}. \quad (2.27)$$

The sum of the area-preserving and area-changing contributions gives the total tensor, for example for the Kirchhoff stress tensor, $\boldsymbol{\tau} = (\mathbb{P}_{\text{isoA}}^{\text{cc}} + \mathbb{P}_A^{\text{cc}}) : \boldsymbol{\tau}$. With this, the area-changing projection tensor in the current configuration $\mathbb{P}_A^{\text{cc}}$ is defined as the counterpart of the isoplanar term, yielding

$$\mathbb{P}_A^{\text{cc}} = \frac{1}{2} J_A^{-1} \mathbf{F} \frac{\partial J_A}{\partial \mathbf{C}} \mathbf{F}^T \otimes \mathbf{I}. \quad (2.28)$$

2.4. Inelastic material formulation

In order to model more complex material behavior than elasticity, a visco-elastic material formulation will be presented in the following. For that, we introduce the multiplicative split of the deformation gradient into an elastic part, $\mathbf{F}_e$, and an inelastic one, $\mathbf{F}_i$, i.e.,

$$\mathbf{F} = \mathbf{F}_e \mathbf{F}_i. \quad (2.29)$$

It is worth to mention that the inelastic part of the deformation gradient, and thus, the quantities in the intermediate configuration (*ic*), cannot be defined uniquely. Therefore, the material formulation is derived in a co-rotated intermediate configuration (*cic*) following the framework presented in Holthusen et al. [2023]. In this configuration, all quantities can directly be computed and implemented. Due to the increasing complexity when deriving the equations and calculating the derivatives needed for inelastic material models, we apply automated differentiation.

The configurations are distinguished by different notations in the following. Quantities in the *ic* are denoted by $(\tilde{\bullet})$, in the *cic* by $(\bar{\bullet})$ and in the *cc* by $(\hat{\bullet})$.

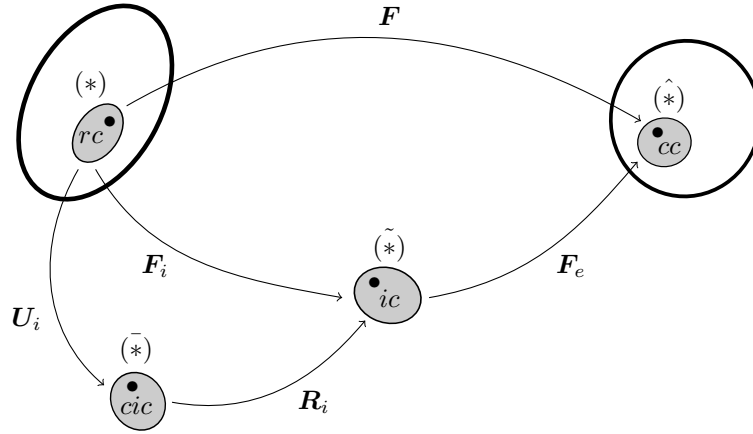


Figure 2.2.: Visualization of multiplicative decomposition of the deformation gradient F into F_e and F_i multiplicative decomposition of F_i into R_i und U_i and the corresponding configurations: *rc* - reference configuration, *cc* - current configuration, *ic* - intermediate configuration, *cic* - co-rotated intermediate configuration [cf. Holthusen et al., 2023].

2.4.1. Co-rotated intermediate configuration

The multiplicative decomposition of the deformation gradient leads to the introduction of an intermediate configuration (see Fig. 2.2). Most constitutive models dealing with inelasticity are derived in this very configuration. The Helmholtz free energy, ψ , is then assumed to be an isotropic function of the elastic right Cauchy-Green tensor, $C_e := F_e^T F_e$, i.e. $\psi = \psi^*(C_e)$. The arguments of ψ^* are defined in the intermediate configuration and cannot be computed directly due to the non-uniqueness, that results from the rotational non-uniqueness of R_i which itself is introduced by the polar decomposition of the inelastic deformation gradient, i.e. $F_i = R_i U_i$. For more details on the rotational non-uniqueness and a decomposition of the deformation gradient into uniquely defined terms, the reader is referred to Casey [2017].

As the Helmholtz free energy is assumed to be an isotropic function, it is not affected by the non-uniqueness and can be calculated. In contrast, the derivatives of the Helmholtz free energy with respect to the elastic right Cauchy-Green tensor are not unique. This can be seen in the definition of the derivative of the elastic right Cauchy-Green tensor which is desired for the total derivative of the energy, i.e. $\dot{\mathbf{C}}_e = \mathbf{F}_i^{-T} \dot{\mathbf{C}} \mathbf{F}_i^{-1} - \mathbf{C}_e \mathbf{L}_i - \mathbf{L}_i^T \mathbf{C}_e$, where $\mathbf{L}_i$ is the rate of inelastic strains and further introduced in Section 2.4.3. As can be seen, the inelastic part of the deformation gradient, $\mathbf{F}_i$, is included in the derivative of the elastic right Cauchy-Green deformation tensor, and can not be defined without any further assumptions. Thus, the derivative cannot be computed directly. Therefore, direct use of algorithmic differentiation is not possible but would enable a flexible adjustment of the material model as the energy formulation can be changed easily. Thus, it is a desired tool during the development process of a material model. In contrast to the tensor $\mathbf{R}_i$, the stretch tensor $\mathbf{U}_i$ can be determined. This leads to the introduction of the co-rotated intermediate configuration, *cic*, (cf. Fig. 2.2) as an additional configuration. By pull-back operations to this very configuration, the elastic right Cauchy-Green deformation tensor in the *cic* can be introduced by

$$\bar{\mathbf{C}}_e := \mathbf{R}_i^{-1} \mathbf{C}_e \mathbf{R}_i = \mathbf{U}_i^{-1} \mathbf{C} \mathbf{U}_i^{-1} \quad (2.30)$$

that can be determined since $\mathbf{U}_i$ and $\mathbf{C}$ are unique. The eigenvalues and symmetry properties between the *ic* and *cic* remain equal and the assumption of an unstressed *ic* (cf. Lee and Liu [1967]; Lee [1969]) can be transferred to the *cic*. As the Helmholtz free energy is an isotropic function of $\mathbf{C}_e$, it is invariant to superimposed rotations of $\mathbf{F}_i$ and, conclusively, can be represented by $\psi = \psi^*(\bar{\mathbf{C}}_e)$. With this, all quantities can directly be defined in the *cic* and implemented by automated differentiation without further pull-back operations. For more information, the interested reader is referred to Holthusen et al. [2023] and the literature cited therein.

2.4.2. Structural tensors in different configurations

The anisotropy in the model is captured by the structural tensors, $\mathbf{M}_i$. Due to the introduction of inelasticity, their counterparts in the different configurations need to be derived starting from the *rc*.

The *ip* director vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ are assumed to evolve with the deformation gradient $\mathbf{F}$ and are normalized to ensure unit length. Thus, the intermediate terms of the two *ip* vectors are

defined in analogy to Reese [2003] by

$$\tilde{\mathbf{n}}_1 = \frac{\mathbf{F}_i \mathbf{n}_1}{|\mathbf{F}_i \mathbf{n}_1|}, \quad \tilde{\mathbf{n}}_2 = \frac{\mathbf{F}_i \mathbf{n}_2}{|\mathbf{F}_i \mathbf{n}_2|}. \quad (2.31)$$

The mapping of the *op* structural tensor results directly from the previous definition and the relation $\tilde{\mathbf{n}}_3 = \tilde{\mathbf{n}}_1 \times \tilde{\mathbf{n}}_2$ such that

$$\begin{aligned} \tilde{\mathbf{M}}_3 &= \tilde{\mathbf{n}}_3 \otimes \tilde{\mathbf{n}}_3 \\ &= \left(\frac{\mathbf{F}_i \mathbf{n}_1}{|\mathbf{F}_i \mathbf{n}_1|} \times \frac{\mathbf{F}_i \mathbf{n}_2}{|\mathbf{F}_i \mathbf{n}_2|} \right) \otimes \left(\frac{\mathbf{F}_i \mathbf{n}_1}{|\mathbf{F}_i \mathbf{n}_1|} \times \frac{\mathbf{F}_i \mathbf{n}_2}{|\mathbf{F}_i \mathbf{n}_2|} \right) \end{aligned} \quad (2.32)$$

Using the relation $(\mathbf{F}_i \mathbf{n}_1) \times (\mathbf{F}_i \mathbf{n}_2) = \text{cof}(\mathbf{F}_i) \mathbf{n}_3$, the definition of the cofactor and normalization leads to the formulation of the *op* structural tensor in the *ic* of

$$\tilde{\mathbf{M}}_3 = \frac{1}{C_i^{-1} : \mathbf{M}_3} \mathbf{F}_i^{-T} \mathbf{M}_3 \mathbf{F}_i^{-1}. \quad (2.33)$$

with $C_i = U_i^2$. Thus, the mapping of the third structural tensor to the *ic* by $\mathbf{F}_i^{-T}$ and $\mathbf{F}_i^{-1}$ results from the classical mapping of the first and second director vector by $\mathbf{F}_i$. Its counterpart in the *cic* is pulled back analogously to Equation 2.30. This leads to

$$\bar{\mathbf{M}}_3 = \frac{1}{C_i^{-1} : \mathbf{M}_3} U_i^{-1} \mathbf{M}_3 U_i^{-1}. \quad (2.34)$$

Consequently, the *ip* structural tensor, $\bar{\boldsymbol{\Omega}}$, can be described in the *cic* by $\bar{\boldsymbol{\Omega}} = \mathbf{I} - \bar{\mathbf{M}}_3$. For later purposes, the structural tensor in the current configuration (*cc*) is given by

$$\hat{\mathbf{M}}_3 = \frac{1}{C_e^{-1} : \tilde{\mathbf{M}}_3} \mathbf{F}_e^{-T} \tilde{\mathbf{M}}_3 \mathbf{F}_e^{-1}. \quad (2.35)$$

An overview of the definitions of the third structural tensor in the different configurations is given in Table 2.2. The director vectors, $\mathbf{n}_j$ with $j = 1, 2, 3$, and the two *ip* structural tensors, $\mathbf{M}_1$ and $\mathbf{M}_2$, are neglected here, as they are not needed in the further derivation.

2.4.3. Derivation based on the Clausius-Duhem inequality

The Helmholtz free energy is defined in terms of the elastic right Cauchy-Green deformation tensor in the *cic*, $\bar{\mathbf{C}}_e$. Additionally, anisotropy is included in the formulation through the *op* structural tensor in the *cic*, $\bar{\mathbf{M}}_3$. Hence, we can define $\psi = \psi^*(\bar{\mathbf{C}}_e, \bar{\mathbf{M}}_3)$. Again, the Clausius-

Table 2.2.: Overview of the third structural tensor, $\mathbf{M}_3$, in the different configurations.

Configuration	Definition of $\mathbf{M}_3$ in the corresponding configuration
Reference configuration (rc)	$\mathbf{M}_3 = \mathbf{n}_3 \otimes \mathbf{n}_3$
Intermediate configuration (ic)	$\tilde{\mathbf{M}}_3 = \frac{1}{\mathbf{C}_i^{-1} : \mathbf{M}_3} \mathbf{F}_i^{-T} \mathbf{M}_3 \mathbf{F}_i^{-1}$
Co-rotated intermediate configuration (cic)	$\bar{\mathbf{M}}_3 = \frac{1}{\mathbf{C}_i^{-1} : \mathbf{M}_3} \mathbf{U}_i^{-1} \mathbf{M}_3 \mathbf{U}_i^{-1}$
Current configuration (cc)	$\hat{\mathbf{M}}_3 = \frac{1}{\mathbf{C}_e^{-1} : \tilde{\mathbf{M}}_3} \mathbf{F}_e^{-T} \tilde{\mathbf{M}}_3 \mathbf{F}_e^{-1}$

Duhem inequality for isothermal processes must be fulfilled, and thus, the time derivative of the Helmholtz free energy is derived, i.e.,

$$\dot{\psi} = \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} : \dot{\bar{\mathbf{C}}}_e + \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} : \dot{\bar{\mathbf{M}}}_3 \quad (2.36)$$

with the rates

$$\dot{\bar{\mathbf{M}}}_3 = 2 (\bar{\mathbf{L}}_i : \bar{\mathbf{M}}_3) \bar{\mathbf{M}}_3 - \bar{\mathbf{L}}_i \bar{\mathbf{M}}_3 - \bar{\mathbf{M}}_3 \bar{\mathbf{L}}_i^T \quad (2.37)$$

$$\dot{\bar{\mathbf{C}}}_e = \mathbf{U}_i^{-1} \dot{\mathbf{C}} \mathbf{U}_i^{-1} - \bar{\mathbf{C}}_e \bar{\mathbf{L}}_i - \bar{\mathbf{L}}_i^T \bar{\mathbf{C}}_e \quad (2.38)$$

where $\bar{\mathbf{L}}_i := \dot{\mathbf{U}}_i \mathbf{U}_i^{-1}$ is the rate of inelastic strains. For more details on the derivatives, the reader is referred to 2.8.3. Inserting Equations (2.36) to (2.38) into the Clausius-Duhem inequality and applying the procedure of Coleman and Noll [1961] leads to the definition of the second Piola-Kirchhoff stress tensor as

$$\mathbf{S} := 2 \mathbf{U}_i^{-1} \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} \mathbf{U}_i^{-1} \quad (2.39)$$

and the reduced dissipation inequality, i.e.

$$\mathcal{D}_{\text{red}} := \underbrace{\left(\underbrace{2 \bar{\mathbf{C}}_e \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e}}_{=: \bar{\Sigma}} + \underbrace{2 \bar{\mathbf{M}}_3 \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3}}_{=: \bar{\chi}} - \underbrace{2 \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} : (\bar{\mathbf{M}}_3 \otimes \bar{\mathbf{M}}_3)}_{=: \bar{\Pi}} \right)}_{=: \bar{\Gamma}} : \bar{\mathbf{L}}_i \geq 0. \quad (2.40)$$

A more detailed derivation can be found in 2.8.4. While the individual stress-like tensors $\bar{\Sigma}$, $\bar{\chi}$ and $\bar{\Pi}$ are generally not symmetric, their sum, $\bar{\Gamma}$, is symmetric (cf. Svendsen [1994]; Reese [2003]; Holthausen et al. [2022]). Hence, defining $\bar{D}_i := \text{sym}(\bar{L}_i)$, the reduced dissipation inequality (2.40) can be reformulated to

$$\bar{\Gamma} : \bar{D}_i \geq 0 \quad (2.41)$$

and must be fulfilled for arbitrary processes to ensure thermodynamical consistency. In analogy to Germain et al. [1983], we define inelastic potentials g being positive, convex and zero-valued functions of their thermodynamically consistent driving forces, and further, assume g to be a scalar-valued isotropic function, to ensure that Inequality (2.41) is satisfied. Thus, by introducing a potential $g = \bar{g}(\bar{\Gamma})$, the evolution equation is then defined by

$$\bar{D}_i = \frac{1}{\tau} \frac{\partial \bar{g}}{\partial \bar{\Gamma}} = \frac{1}{\tau} \mathbf{R}_i^{-1} \frac{\partial g}{\partial \bar{\Gamma}} \mathbf{R}_i \quad \longrightarrow \quad \dot{\bar{C}}_i = \frac{2}{\tau} \mathbf{U}_i \frac{\partial g}{\partial \bar{\Gamma}} \mathbf{U}_i \quad (2.42)$$

where τ is a relaxation time. The driving force $\bar{\Gamma}$ introduced above is the *cic*-counterpart of its intermediate expression, $\bar{\Gamma} = \mathbf{R}_i^{-1} \tilde{\Gamma} \mathbf{R}_i$. The potential is defined as a scalar-valued isotropic function of the driving force, and thus, the same physical interpretation can be assigned. With this, the set of constitutive equations has been defined in a general manner. Specific terms of the Helmholtz free energy, ψ , and the potential, g , must now be defined. An overview of the constitutive equations is given in Table 2.3.

Table 2.3.: Overview of constitutive equations of the elastic and inelastic material models.

Elastic material model	Inelastic material model
Structural tensors	Structural tensors
$\mathbf{n}_3 = \mathbf{n}_1 \times \mathbf{n}_2$	$\mathbf{n}_3 = \mathbf{n}_1 \times \mathbf{n}_2$
$\mathbf{M}_3 = \mathbf{n}_3 \otimes \mathbf{n}_3$	$\mathbf{M}_3 = \mathbf{n}_3 \otimes \mathbf{n}_3$
	$\bar{\mathbf{M}}_3 = \frac{1}{\mathbf{C}_i^{-1} : \mathbf{M}_3} \mathbf{U}_i^{-1} \mathbf{M}_3 \mathbf{U}_i^{-1}$
	Kinematics
	$\bar{\mathbf{C}}_e = \mathbf{U}_i^{-1} \mathbf{C} \mathbf{U}_i^{-1}$
Stress	Stresses
$\mathbf{S} = 2 \frac{\partial \psi}{\partial \mathbf{C}}$	$\mathbf{S} = 2 \mathbf{U}_i^{-1} \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} \mathbf{U}_i^{-1}$
	$\bar{\Gamma} = 2 \bar{\mathbf{C}}_e \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} + 2 \bar{\mathbf{M}}_3 \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} - 2 \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} : (\bar{\mathbf{M}}_3 \otimes \bar{\mathbf{M}}_3)$
	Evolution equation
	$\dot{\bar{\mathbf{C}}}_i = \frac{2}{\tau} \mathbf{U}_i \frac{\partial g}{\partial \bar{\Gamma}} \mathbf{U}_i$

2.4.4. Evolution equations

The potential g must be chosen appropriately. For the physical interpretation, the potentials are introduced in terms of the Kirchhoff stress $\boldsymbol{\tau}$. However, the Kirchhoff stress is defined in the cc , and thus, its counterpart in the cic has to be derived. As all potentials are described by the (mixed) invariants of $\boldsymbol{\tau}$ and $\hat{\mathbf{M}}_3$, we focus on the cic formulations of $\boldsymbol{\tau}$, $\boldsymbol{\tau}^2$ and $\boldsymbol{\tau}\hat{\mathbf{M}}_3$ specifically.

2.4.4.1. Relations between quantities in the current configuration and co-rotated intermediate configuration

As previously mentioned, the Helmholtz free energy is an isotropic function of $\bar{\mathbf{C}}_e$ and $\bar{\mathbf{M}}_3$ which both are defined in the cic . However, both are considered in the ic in the following for simplicity since the eigenvalues and symmetry properties remain the same between the ic and cic . Furthermore, pull-back operations from the ic to the cic are performed by applying $(\bar{\bullet}) = \mathbf{R}_i^{-1}(\tilde{\bullet})\mathbf{R}_i$. In the described invariants, the rotational tensors vanish, leading to a similar representation of the invariants in the ic and cic with their corresponding counterparts. Thus, the Helmholtz free energy can be represented as an isotropic function of $\mathbf{C}_e$ and $\tilde{\mathbf{M}}_3$ and, consequently, the tensors $\partial\psi/\partial\mathbf{C}_e$ and $\partial\psi/\partial\tilde{\mathbf{M}}_3$ can be described by the form

$$\frac{\partial\psi}{\partial\mathbf{C}_e} = \alpha_1\mathbf{I} + 2\alpha_2\mathbf{C}_e + 3\alpha_3\mathbf{C}_e^2 + \alpha_4\tilde{\mathbf{M}}_3 + \alpha_5(\mathbf{C}_e\tilde{\mathbf{M}}_3 + \tilde{\mathbf{M}}_3\mathbf{C}_e) \quad (2.43)$$

$$\frac{\partial\psi}{\partial\tilde{\mathbf{M}}_3} = \alpha_4\mathbf{C}_e + \alpha_5\mathbf{C}_e^2 \quad (2.44)$$

where $\alpha_i := \partial\psi/\partial I_i$ and I_i are the (mixed) invariants defined by $\mathbf{C}_e$ and $\tilde{\mathbf{M}}_3$ given in 2.8.5. Per definition, the Kirchhoff stress is $\boldsymbol{\tau} := \mathbf{F}\mathbf{S}\mathbf{F}^T$. Using the relations in Equations (2.29), (2.39), (2.43) and (2.44) leads to the description of the Kirchhoff stress $\boldsymbol{\tau}$ of

$$\begin{aligned} \boldsymbol{\tau} &= 2\mathbf{F}_e \frac{\partial\psi}{\partial\mathbf{C}_e} \mathbf{F}_e^T \\ &= 2 \left(\alpha_1 \mathbf{B}_e + 2\alpha_2 \mathbf{B}_e^2 + 3\alpha_3 \mathbf{B}_e^3 + \alpha_4 \mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e + \alpha_5 \left(\mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e + \mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^2 \right) \right) \end{aligned} \quad (2.45)$$

where $\mathbf{B}_e := \mathbf{F}_e \mathbf{F}_e^T$ is the elastic left Cauchy-Green deformation tensor and $\hat{\mathbf{M}}_3^* := \text{tr}(\mathbf{C}_e^{-1} \tilde{\mathbf{M}}_3) \hat{\mathbf{M}}_3$ is the third structural tensor in the cc without any scaling factor. More-

over, the stress-like quantities Γ , Σ , χ , and Π in the *ic*, introduced in Equation (2.40), read

$$\tilde{\Gamma} = \tilde{\Sigma} + \tilde{\chi} - \tilde{\Pi} \quad (2.46)$$

$$\tilde{\Sigma} = 2 \left(\alpha_1 C_e + 2 \alpha_2 C_e^2 + 3 \alpha_3 C_e^3 + \alpha_4 C_e \tilde{M}_3 + \alpha_5 \left(C_e^2 \tilde{M}_3 + C_e \tilde{M}_3 C_e \right) \right) \quad (2.47)$$

$$\tilde{\chi} = 2 \left(\alpha_4 \tilde{M}_3 C_e + \alpha_5 \tilde{M}_3 C_e^2 \right) \quad (2.48)$$

$$\tilde{\Pi} = 2 \left(\alpha_4 I_4 + \alpha_5 I_5 \right) \tilde{M}_3. \quad (2.49)$$

With these definitions, the relations between the *cc* and *ic* can be derived. Detailed calculations can be found in 2.8.5. It is shown that the relation $\text{tr}(\tau) = \text{tr}(\tilde{\Gamma})$ holds true. Moreover, the traces of the squared terms $\text{tr}(\tau^2)$ and $\text{tr}(\tilde{\Gamma}^2)$ are compared. However, they are only equal for $C_e \approx I$. In addition, the intermediate expression of $\tau : \hat{M}_3$ is required. The relation

$$\tau : \hat{M}_3 = \frac{1}{\beta} \tilde{\Gamma} : \text{sym}(C_e^{-1} \tilde{M}_3) \quad (2.50)$$

holds, where $\beta = \text{tr}(C_e^{-1} \tilde{M}_3)$.

The intermediate counterparts of the traces of τ , τ^2 , and $\tau \hat{M}_3$ are found. Consequently, potentials can be introduced in the physically sound *cc* and equivalently described and implemented in the *cic* in the following.

2.4.4.2. Choice of potentials

The described framework attempts to model the *ip op* decoupling and is derived in a general manner. To incorporate inelastic material behavior, three different potentials are presented in the following. Quadratic potentials are chosen such that the requirements mentioned in Section 2.4.3 are automatically fulfilled.

Potential 1

The first potential is defined by the trace of the Kirchhoff stress reduced by its *op* part given here by $\text{tr}(\tau \hat{M}_3)$, i.e.

$$g_1 = \frac{1}{2 K_2} \left(\text{tr}(\tau) - \text{tr}(\tau \hat{M}_3) \right)^2 \quad (2.51)$$

K_2 is a material parameter. Per definition of the potential, inelastic strains evolve only when *ip* loadings are applied. With the relations given in Equations (2.87) and (2.50), the potential can be described in the *ic*. Pull back operations to the *cic* are applied by $(\bar{\bullet}) = \mathbf{R}_i^{-1}(\bullet)\mathbf{R}_i$ where

the rotation tensors vanish in the trace. Thus, the potential in the *cic* can be written as

$$\bar{g}_1 = \frac{1}{2K_2} \left(\text{tr}(\bar{\Gamma}) - \frac{1}{\text{tr}(\bar{C}_e^{-1}\bar{M}_3)} \text{tr}(\bar{\Gamma}\bar{C}_e^{-1}\bar{M}_3) \right)^2. \quad (2.52)$$

It holds that $g_1 \equiv \bar{g}_1$.

Potential 2

A second potential g_2 is defined in analogy to the first potential. However, an additional term incorporating the evolution of inelastic strains in the *op* direction is included, yielding

$$g_2 = \frac{1}{2K_2} \left(\text{tr}(\boldsymbol{\tau}) - \text{tr}[\boldsymbol{\tau}\hat{M}_3] \right)^2 + \frac{1}{2K_3} \text{tr}(\boldsymbol{\tau}\hat{M}_3)^2. \quad (2.53)$$

Thus, the evolution of inelastic strains in *ip* and *op* directions is captured but remains decoupled. With this, *ip* and *op* stresses only lead to *ip* and *op* inelastic strain evolutions, respectively. The *cic* counterpart is determined similarly to Equation (2.52). Again, we can write $g_2 \equiv \bar{g}_2$ and

$$\bar{g}_2 = \frac{1}{2K_2} \left(\text{tr}(\bar{\Gamma}) - \frac{1}{\text{tr}(\bar{C}_e^{-1}\bar{M}_3)} \text{tr}(\bar{\Gamma}\bar{C}_e^{-1}\bar{M}_3) \right)^2 + \frac{1}{2K_3} \left(\frac{1}{\text{tr}(\bar{C}_e^{-1}\bar{M}_3)} \text{tr}(\bar{\Gamma}\bar{C}_e^{-1}\bar{M}_3) \right)^2. \quad (2.54)$$

Potential 3

With the potentials g_1 and g_2 , we will show that the desired decoupled behavior can be achieved. To show the flexibility of the framework, an additional coupling term is included in the third potential, g_3 . This will demonstrate that the general framework proposed herein is capable of dealing with both, coupled and decoupled behavior. The corresponding potential reads

$$g_3 = \frac{1}{4(K_1 + K_3)} \text{tr}(\text{dev}(\boldsymbol{\tau})^2) + \frac{1}{18(K_2 + K_3)} \text{tr}(\boldsymbol{\tau})^2 \quad (2.55)$$

$$= \frac{1}{4(K_1 + K_3)} \left(\text{tr}(\boldsymbol{\tau}^2) - \frac{\text{tr}(\boldsymbol{\tau})^2}{3} \right) + \frac{1}{18(K_2 + K_3)} \text{tr}(\boldsymbol{\tau})^2 \quad (2.56)$$

K_1 and K_3 are again material parameters. Here, the classical definition of the deviator, $\text{dev}(\bullet) = (\bullet) - \frac{1}{3} \text{tr}(\bullet) \mathbf{I}$, and its isochoric counterpart, $\text{tr}(\bullet)$, are used. This leads to consideration of volumetric and isochoric instead of *ip* and *op* deformations. Equation (2.56)

is represented in the *cic* by

$$\bar{g}_3 = \frac{1}{4(K_1 + K_3)} \left(\text{tr}(\bar{\mathbf{\Gamma}}^2) - \frac{\text{tr}(\bar{\mathbf{\Gamma}})^2}{3} \right) + \frac{1}{18(K_2 + K_3)} \text{tr}(\bar{\mathbf{\Gamma}})^2 \quad (2.57)$$

which is equal to the *cc* term for $\mathbf{C}_e \approx \mathbf{I}$ as derived in 2.8.5. Furthermore, we will show in Section 2.6.3 via numerical calculations that $\text{tr}(\boldsymbol{\tau}^2) \approx \text{tr}(\mathbf{\Gamma}^2)$. Thus, Equation (2.57) can be used as the *cic* counterpart for the third potential.

2.4.5. Numerical Implementation

An implicit time integration scheme has been used for the implementation of the evolution equations. For that, we chose the exponential map algorithm introduced by Dettmer and Reese [2004] and improved by Vladimirov et al. [2008] and Christ and Reese [2009] among others. It leads to the following residual format of the evolution equation (2.42):

$$\mathbf{r}_I = \mathbf{U}_{i_n}^2 - \mathbf{U}_i \exp \left(-2\Delta t \frac{1}{\tau} \frac{\partial \bar{g}}{\partial \bar{\mathbf{\Gamma}}} \right) \mathbf{U}_i = \mathbf{0} \quad (2.58)$$

The index n denotes values from the last converged time step. This equation naturally ensures the symmetry of $\mathbf{U}_i$. Due to this symmetry, a system of only six nonlinear equations is solved by the use of the Newton-Raphson scheme. For more information, the reader is referred to the literature mentioned above as well as Felder et al. [2020] or Holthusen et al. [2023].

The material models are implemented with the help of the algorithmic differentiation tool *AceGen* [Korelc, 2002; Korelc and Wriggers, 2016].

2.5. Choice of Helmholtz free energy

Up to now, the elastic and viscoelastic material models have been presented in a general manner which makes them flexible to adapt to different material behaviors. In the following, the specific choices of the Helmholtz free energy are introduced.

2.5.1. Remarks on the formulation of Helmholtz free energy functions

The choice of the Helmholtz free energy function is accompanied by some restrictions. One property we require is the fulfillment of the normalization condition that implies a zero-valued energy function at the reference configuration, $\mathbf{C} = \mathbf{I}$, such that $\psi = \psi(\mathbf{I}) = 0$. The energy function must be convex having its global minimum at the initial state. Thus, it is

non-negative for any deformation state, i.e. $\psi = \psi(\mathbf{C}) \geq 0$. These two requirements impose a global minimum at the initial state and ensure a stress-free reference configuration through vanishing gradients at the origin. For more information on the requirement of polyconvexity, the interested reader is referred to Hartmann and Neff [2003] and Schröder and Neff [2003]. In addition, the growth criterion must be fulfilled by the energy function which states that the Helmholtz free energy tends to $+\infty$ when $J = \det(\mathbf{C})$ approaches both, $+\infty$ or 0^+ , or when $\text{tr}(\mathbf{C})$ tends to $+\infty$. Physically, this implies that an infinite amount of energy is needed for an infinite expansion or compression of a continuum body. Thus, the Helmholtz free energy function must fulfill the criteria of convexity, non-negativeness, and the normalization and growth condition. For more information, the reader is referred to, for example, Ball [1976]; Ogden and Sternberg [1985]; Ciarlet [1988]; Holzapfel [2002].

In the classical volumetric framework, the Helmholtz free energy can be described in terms of the principal invariants of the right Cauchy-Green deformation tensor (cf. Equations (2.2) to (2.4)). Evaluation of the invariants at the initial configuration yields $I_1(\mathbf{I}) = I_2(\mathbf{I}) = 3$ and $I_3(\mathbf{I}) = 1$. With this and in compliance with the above criteria, the Helmholtz free energy can be represented as a series of

$$\psi(I_1, I_2, I_3) = \sum_{p,q,r=0}^{\infty} c_{pqr} (I_1 - 3)^p (I_2 - 3)^q (I_3 - 1)^{2r}, \quad c_{000} = 0 \quad (2.59)$$

c_{pqr} are material constants. Equation (2.59) can be separated into distortional and dilatational parts by setting $r = 0$ or $p = q = 0$, yielding

$$\psi(I_1, I_2, I_3) = \sum_{p,q=0}^{\infty} c_{pq0} (I_1 - 3)^p (I_2 - 3)^q + \sum_{r=0}^{\infty} c_{00r} (I_3 - 1)^{2r}. \quad (2.60)$$

Regarding the special case of $p = 1$, $q = 0$, and an incompressible material, i.e. $I_3 = 1$, the energy function reduces to the commonly known Neo-Hooke strain energy function, i.e.

$$\psi(I_1) = c_{10}(I_1 - 3). \quad (2.61)$$

It can be used for describing the isochoric deformations of a continuum body. This leads to the question of how to express the volumetric terms. Regarding the formulations presented in Ogden [1972]; Simo [1992]; Holzapfel [2002], a volumetric response function, ψ_{vol} , can be defined in terms of the volume ratio $J = \det(\mathbf{C})$ by

$$\psi_{\text{vol}}(J) = K \beta^{-2} (J^{-\beta} - 1 + \beta \ln[J]). \quad (2.62)$$

K and β are a material parameter and coefficient, respectively. The Helmholtz free energy functions defined above all fulfill the normalization and growth condition, are convex and non-negative.

2.5.2. Elastic material formulation - isoplanar-area-changing split

First, the definition of the elastic energy, ψ^e , using the isoplanar-area-changing split introduced in Section 2.3.2.1 is presented. It is additively divided into ip and op terms, ψ^{ip} and ψ^{op} , where the ip ones are further separated into isoplanar and area-changing terms, denoted by ψ_{iso}^{ip} and ψ_A^{ip} , respectively, i.e.

$$\psi^e = \psi^{ip} + \psi^{op} = \psi_{isoA}^{ip} + \psi_A^{ip} + \psi^{op} \quad (2.63)$$

The isoplanar term is based on Equation (2.61) and formulated in terms of the first modified invariant (cf. Equation (2.6)). Evaluation of this invariant in the reference configuration yields $I'_1(\mathbf{I}) = 2$. Due to the initial value of I'_1 , the area-preserving energy function is reformulated to

$$\psi_{isoA}^{ip} = \frac{K_1}{2} (\text{tr}(\mathbf{C}_{isoA} \mathbf{\Omega}) - 2). \quad (2.64)$$

Here, the isoplanar term of the deformation tensor, $\mathbf{C}_{isoA}$, is used rather than the total right Cauchy-Green tensor. K_i with $i = 1, 2, 3, 4$ are material parameters. The area-changing energy function is chosen in line with Equation (2.62) for $\beta = -1$, i.e.

$$\psi_A^{ip} = \frac{K_2}{4} (J_A - 1 - \ln(J_A)). \quad (2.65)$$

In contrast to the volume ratio J , the ip equivalent $J_A = I'_2$ has been used here. The op energy function is defined analogously to the volumetric formulation in Equation (2.62). However, rather than the total volume ratio, only the stretch in the op direction is incorporated. In addition, the second invariant modified by the op structural tensor is used. In detail, the op Helmholtz free energy reads

$$\psi^{op} = \frac{K_3}{4} (\text{tr}[\mathbf{C} \mathbf{M}_3] - 1 - \ln[\text{tr}(\mathbf{C} \mathbf{M}_3)]) + \frac{K_4}{2} (\text{tr}(\mathbf{C}^2 \mathbf{M}_3) - \text{tr}(\mathbf{C} \mathbf{M}_3)^2). \quad (2.66)$$

By definition of Equations (2.19) and (2.63) to (2.66), the second Piola-Kirchhoff stress tensor $\mathbf{S}$ is then defined by

$$\begin{aligned} \mathbf{S} = 2 \frac{\partial \psi}{\partial \mathbf{C}} = & K_1 \left(J_A^{-\frac{1}{2}} \boldsymbol{\Omega} - \frac{1}{2} J_A^{-\frac{3}{2}} (\text{tr}(\mathbf{C}\boldsymbol{\Omega})\boldsymbol{\Omega} - \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega}) \right) \\ & + \frac{K_2}{2} \left(\text{tr}(\mathbf{C}\boldsymbol{\Omega})\boldsymbol{\Omega} - J_A^{-1} \text{tr}(\mathbf{C}\boldsymbol{\Omega})\boldsymbol{\Omega} - \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega} + J_A^{-1} \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega} \right) \\ & + \frac{K_3}{2} \left(1 - \frac{1}{\text{tr}(\mathbf{C}\mathbf{M}_3)} \right) \mathbf{M}_3 \\ & + K_4 \left(\mathbf{C}\mathbf{M}_3 + \mathbf{M}_3\mathbf{C} - 2 \text{tr}(\mathbf{C}\mathbf{M}_3) \mathbf{M}_3 \right). \end{aligned} \quad (2.67)$$

Finally, the material tangent $\mathbb{C}$ can be obtained by

$$\begin{aligned} \mathbb{C} = 4 \frac{\partial^2 \psi}{\partial \mathbf{C} \partial \mathbf{C}} = & K_1 \left[- J_A^{-\frac{3}{2}} \left(\boldsymbol{\Omega} \otimes (\text{tr}(\mathbf{C}\boldsymbol{\Omega})\boldsymbol{\Omega} - \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega}) + (\text{tr}(\mathbf{C}\boldsymbol{\Omega})\boldsymbol{\Omega} - \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega}) \otimes \boldsymbol{\Omega} \right. \right. \\ & \left. \left. + \text{tr}(\mathbf{C}\boldsymbol{\Omega}) \left[\boldsymbol{\Omega} \otimes \boldsymbol{\Omega} - \frac{1}{2} \left((\boldsymbol{\Omega} \otimes \boldsymbol{\Omega}^T)^{\frac{23}{T}} + (\boldsymbol{\Omega} \otimes \boldsymbol{\Omega})^{\frac{24}{T}} \right) \right] \right) \right. \\ & \left. + \frac{3}{2} J_A^{-\frac{5}{2}} \text{tr}(\mathbf{C}\boldsymbol{\Omega}) (\text{tr}(\mathbf{C}\boldsymbol{\Omega})\boldsymbol{\Omega} - \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega}) \otimes (\text{tr}(\mathbf{C}\boldsymbol{\Omega})\boldsymbol{\Omega} - \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega}) \right] \\ & + K_2 \left[(1 - J_A^{-1}) \boldsymbol{\Omega} \otimes \boldsymbol{\Omega} + J_A^{-2} \text{tr}(\mathbf{C}\boldsymbol{\Omega}) \boldsymbol{\Omega} \otimes (\text{tr}(\mathbf{C}\boldsymbol{\Omega})\boldsymbol{\Omega} - \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega}) \right. \\ & \left. - \frac{1}{2} \left((\boldsymbol{\Omega} \otimes \boldsymbol{\Omega}^T)^{\frac{23}{T}} + (\boldsymbol{\Omega} \otimes \boldsymbol{\Omega})^{\frac{24}{T}} \right) - J_A^{-2} \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega} \otimes (\text{tr}(\mathbf{C}\boldsymbol{\Omega})\boldsymbol{\Omega} - \boldsymbol{\Omega}\mathbf{C}\boldsymbol{\Omega}) \right. \\ & \left. + \frac{1}{2} J_A^{-1} \left((\boldsymbol{\Omega} \otimes \boldsymbol{\Omega}^T)^{\frac{23}{T}} + (\boldsymbol{\Omega} \otimes \boldsymbol{\Omega})^{\frac{24}{T}} \right) \right] + K_3 \left[\text{tr}(\mathbf{C}\mathbf{M}_3)^{-2} \mathbf{M}_3 \otimes \mathbf{M}_3 \right] \\ & + K_4 \left[(\mathbf{I} \otimes \mathbf{M}_3^T)^{\frac{23}{T}} + (\mathbf{I} \otimes \mathbf{M}_3)^{\frac{24}{T}} + (\mathbf{M}_3 \otimes \mathbf{I})^{\frac{23}{T}} + (\mathbf{M}_3 \otimes \mathbf{I})^{\frac{24}{T}} \right. \\ & \left. - 4 \mathbf{M}_3 \otimes \mathbf{M}_3 \right]. \end{aligned} \quad (2.68)$$

For more detailed calculations, the reader is referred to 2.8.7. As a result, the components of the initial material tangent, i.e. $\mathbf{C} = \mathbf{I}$, for $\mathbf{n}_3 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ can be expressed via a 6×6

matrix in Voigt-notation as follows:

$$\mathbb{C}|_{C=I} = \begin{bmatrix} K_1 + K_2 & -K_1 + K_2 & 0 & 0 & 0 & 0 \\ -K_1 + K_2 & K_1 + K_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & K_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & K_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & K_4 & 0 \\ 0 & 0 & 0 & 0 & 0 & K_4 \end{bmatrix}. \quad (2.69)$$

The indices are ordered by the two *ip* directions first and then the *op* direction. It is obvious that the presented framework leads to no off-diagonal terms that cause coupling between the *ip* and *op* directions.

2.5.3. Elastic material formulation - isochoric-volumetric split

The introduced isoplanar-area-changing split will be compared to the commonly known isochoric-volumetric split presented in Section 2.3.1. Therefore, an elastic Helmholtz free energy, ψ_{class}^e – the subscript 'class' indicates the classical framework of the isochoric-volumetric split – is defined in analogy to the one in Section 2.5.2 by

$$\begin{aligned} \psi_{\text{class}}^e = & \frac{K_1}{2} (\text{tr}(\mathbf{C}_{\text{isov}} \mathbf{\Omega}) - 2) + \frac{K_2}{4} (J - 1 - \ln(J)) + \frac{K_3}{4} (\text{tr}(\mathbf{C} \mathbf{M}_3) - 1 - \ln[\text{tr}(\mathbf{C} \mathbf{M}_3)]) \\ & + \frac{K_4}{2} (\text{tr}(\mathbf{C}^2 \mathbf{M}_3) - \text{tr}(\mathbf{C} \mathbf{M}_3)^2). \end{aligned} \quad (2.70)$$

The right Cauchy-Green deformation tensor is divided into volume-changing and volume-preserving deformations as given in Equation (2.1). In order to observe the influence of the isoplanar-area-changing split in comparison to the isochoric-volumetric split, the energy formulation is chosen in analogy to the one in Section 2.5.2. Therefore, the energy is divided in *ip* and *op* directions through the use of the structural tensors.

In the following, the isoplanar-area-changing split presented in Chapter 2.3 and Section 2.5.2 will be denoted by Model I, whereas the isochoric-volumetric split given in this section is indicated by Model II.

2.5.4. Inelastic material formulation

A Helmholtz free energy for inelastic materials is defined in terms of the elastic right Cauchy-Green deformation tensor and the third structural tensor in the *cic*, $\bar{\mathbf{C}}_e$ and $\bar{\mathbf{M}}_3$. The *ip*

structural tensor in the *cic*, $\bar{\Omega}$, can be reformulated in terms of $\bar{M}_3$ and is used in the following for readability only. The definitions $\bar{C}_e = U_i^{-1} C U_i^{-1}$, $\bar{J}_A = 1/2 (\text{tr}(\bar{C}_e \bar{\Omega})^2 - \text{tr}[(\bar{C}_e \bar{\Omega})^2])$ and $\bar{C}_{\text{isoA}}^e = \bar{J}_A^{-1/2} \bar{C}_e$ are used in the following. Remark is made to the overbar which indicates a quantity defined in the *cic*.

The used Helmholtz free energy is defined as the sum of the elastic and inelastic energy formulations, i.e. $\psi = \psi^e + \psi^i$, where the elastic one is given in Equation (2.63) and the inelastic one by

$$\psi^i = \psi_{\text{isoA}}^i + \psi_A^i + \psi_{\text{op}}^i \quad (2.71)$$

$$\psi_{\text{isoA}}^i = \frac{K_1}{2} (\text{tr}(\bar{C}_{\text{isoA}}^e \bar{\Omega}) - 2) \quad (2.72)$$

$$\psi_A^i = \frac{K_2}{4} (\bar{J}_A - 1 - \ln(\bar{J}_A)) \quad (2.73)$$

$$\psi_{\text{op}}^i = \frac{K_3}{4} (\text{tr}(\bar{C}_e \bar{M}_3) - 1 - \ln[\text{tr}(\bar{C}_e \bar{M}_3)]) + \frac{K_4}{2} (\text{tr}(\bar{C}_e^2 \bar{M}_3) - \text{tr}(\bar{C}_e \bar{M}_3)^2). \quad (2.74)$$

2.6. Numerical Examples

In the following, the capabilities of the presented framework are demonstrated by numerical examples. First, the elastic material formulation described in Section 2.3 is evaluated by single-element tests in Section 2.6.1. Then, a comparison to the classical framework of the isochoric-volumetric split is presented for a structural example of a cube under compression in Section 2.6.2. The inelastic material formulation (cf. Section 2.4) is studied by single-element tests in Section 2.6.3 to examine the impact of the three different potentials. Afterwards, the structural example of a pinched cylinder is investigated with these three potentials in Section 2.6.4. Finally, the elastic and inelastic material formulations are compared by simulations of a symmetrically notched specimen (cf. Section 2.6.5). Throughout all simulations, a single Gaussian point concept for continuum elements (Q1STc) proposed by Barfusz et al. [2021] is used.

2.6.1. Single-element tests for the elastic material behavior

Uniaxial tension, uniaxial compression, and simple shear tests are performed in the *ip* and *op* directions on a single-element of unit length (1 mm). The decoupled material formulation presented in Section 2.3 with the energy formulation given in Section 2.5.2 is used here. The material parameters are $K_1 = 80 \left[\frac{\text{N}}{\text{mm}^2} \right]$, $K_2 = 400 \left[\frac{\text{N}}{\text{mm}^2} \right]$, $K_3 = 400 \left[\frac{\text{N}}{\text{mm}^2} \right]$ and

$K_4 = 0 \left[\frac{\text{N}}{\text{mm}^2} \right]$ for the single-element tests.

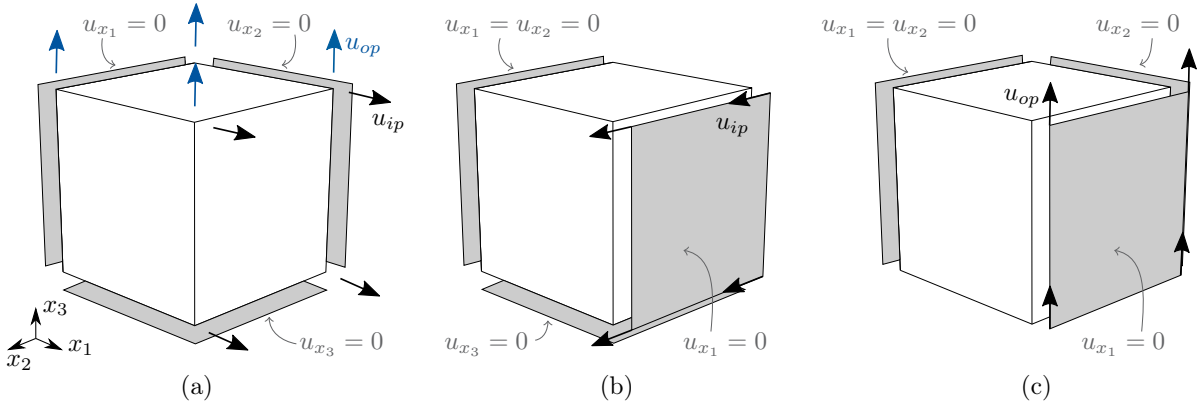


Figure 2.3.: Boundary conditions of the single-element tests: (a) uniaxial tension/compression (black - *ip*, blue - *op*), (b) in-plane simple shear, (c) out-of-plane simple shear.

A displacement of 0.2 mm is applied in all simulations. The results are shown in Figure 2.4. The upper line shows the *op* tension, compression, and shear tests, whereas the lower line gives the results of the *ip* tests. The components of the Green-Lagrange strain, $\mathbf{E}$, are plotted in the three preferential directions, where E_{11} and E_{22} refer to the two *ip* directions and E_{33} to the *op* direction.

The tension and compression tests show that loading in the *op* direction causes strain in the *op* direction only. Similarly, this holds true for *ip* loading, where strains evolve in the two *ip* directions, while the *op* direction remains unaffected. The simple shear tests both show an increase in the first *ip* direction. In the *op* shear test (cf. Fig. 2.3 (c)), a slight decrease in the opposite *ip* direction is observable, whereas the *ip* shear test (cf. Fig. 2.3 (b)) leads to no change in the *op* direction. Consequently, these single-element tests prove the capability of the framework that strains evolve independently in *ip* and *op* direction at *ip* and *op* loading. However, this behavior will additionally be shown in a structural example in the following.

2.6.2. Cube under compression with elastic material behavior

In order to investigate the models' performance under more complex deformation states, we examine a cube under compression. Because of symmetry, only one eighth is simulated. Each side has a length of 1 mm. A surface force of 100 N/mm^2 is applied in a linear fashion on one-quarter of the top surface indicated by the white top area in Fig. 2.5. The following material parameters are used: $K_1 = 100 \left[\frac{\text{N}}{\text{mm}^2} \right]$, $K_2 = 5000 \left[\frac{\text{N}}{\text{mm}^2} \right]$, $K_3 = 5000 \left[\frac{\text{N}}{\text{mm}^2} \right]$ and $K_4 = 1000 \left[\frac{\text{N}}{\text{mm}^2} \right]$. The *op* direction is defined in x_3 -direction.

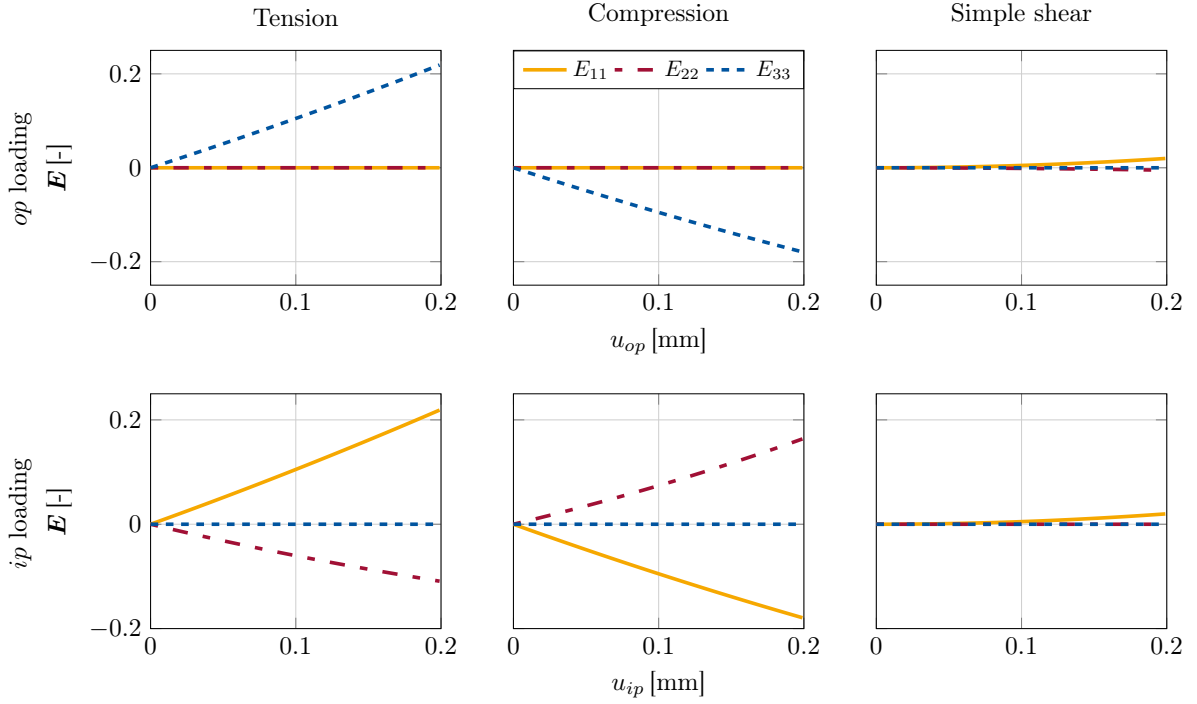


Figure 2.4.: Single-element study on the elastic material formulation: Components of the Green-Lagrange strain $\mathbf{E}$ at uniaxial tension (left), uniaxial compression (middle), and simple shear (right) in out-of-plane direction (upper line) and in-plane direction (lower line).

A mesh convergence study has been performed and is shown in Fig. 2.6. The final displacement in x_3 -direction at Node 1 (cf. Fig. 2.5) is plotted over the number of elements per edge. The results indicate that a mesh with $8 \times 8 \times 8$ elements can be considered sufficiently fine. Nevertheless, a mesh of 10 elements per edge is chosen.

To compare the new isoplanar-area-changing split (Model I) to the classical isochoric-volumetric one (Model II), the two elastic versions discussed in Sections 2.5.2 and 2.5.3, respectively, have been used in simulations of this cube under compression. In Fig. 2.7, the in-plane displacement u_{ip} is shown for both models after deformation. The classical split (Model II) leads to an expansion in both ip directions, specifically in the middle of the edges. In contrast, the new split (Model I) results in a deformation in the op direction, while the cube remains nearly undeformed in the ip directions. Thus, just as desired, the applied loading affects op deformations only, whereas the ip deformations remain nearly unaffected. This proves the capabilities of the presented decoupled framework even for complex deformation states.

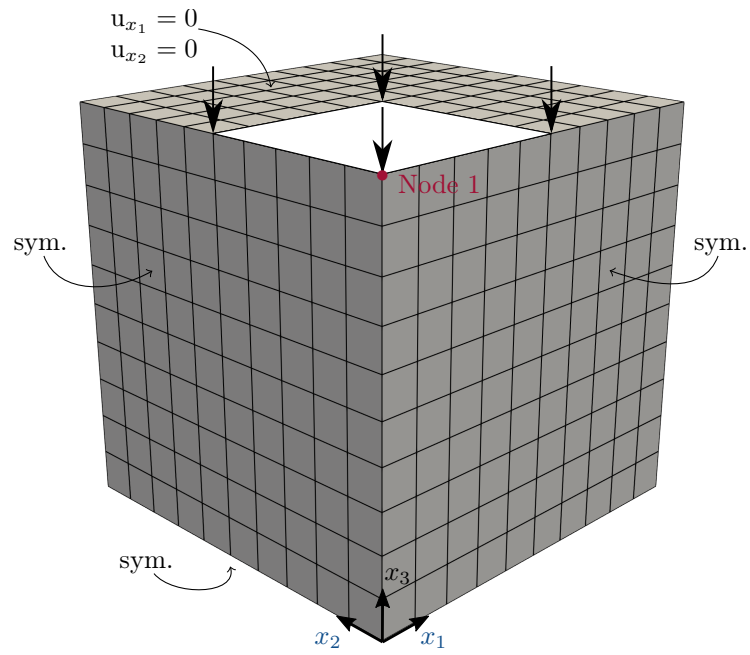


Figure 2.5.: Cube under compression: Geometry, discretization (10x10x10 elements), and boundary conditions - surface loading applied on the white top area.

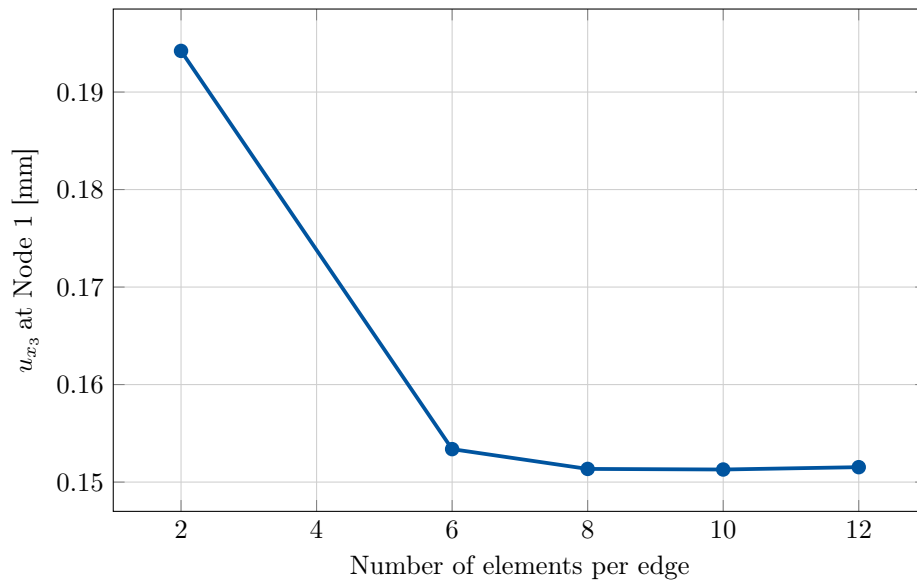


Figure 2.6.: Convergence study of cube under compression: Displacement in x_3 direction at node 1 over the number of elements per edge.

Additionally, the op displacements u_{op} are given in Fig. 2.8 for both models. It can be ob-

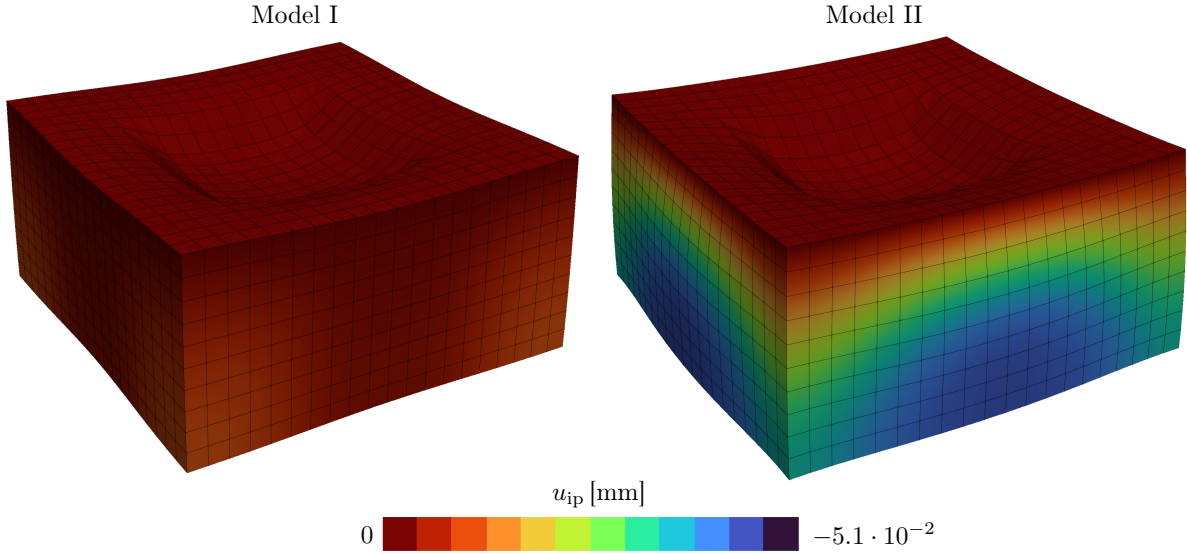


Figure 2.7.: In-plane displacement (magnitude of x_1 - and x_2 -displacement) for the new split (Model I, left) and classical split (Model II, right) (Scaling factor 2).

served that Model II shows a more significant deformation than Model I, which implies a softer behavior. This is underpinned by the load-displacement curves of Model I and Model II given in Fig. 2.9. The applied force F is plotted over the displacement in x_3 direction at Node 1. The difference in the final deformations between both models is around 14 %. The apparently stiffer material response of Model I can be explained through the decoupling of the loading directions. Since the applied force acts in op direction, the ip direction remains unaffected and nearly undeformed. Thus, the cube is compressed without expansion perpendicular to the loading direction. Therefore, as no expansions emerge in ip directions and the shearing terms are diminished in this model, it is harder to compress the cube. On the contrary, the classical split in Model II results in an expansion in the ip plane and, consequently, shows a softer material behavior than Model I.

2.6.3. Single-element tests for the inelastic material behavior

In the following, the inelastic material formulation presented in Chapter 2.4 is examined. For this, uniaxial tension tests are performed in ip and op directions for all three potentials defined in Section 2.4.4.2. The boundary conditions are shown in Fig 2.3 (a). A displacement of 0.2 mm is linearly applied over 20 seconds from t_0 to t_1 and reloaded over 20 seconds from t_1 to t_2 . The material parameters given in Section 2.6.1 and a relaxation time of $\tau = 0.5$ [1/s] are used.

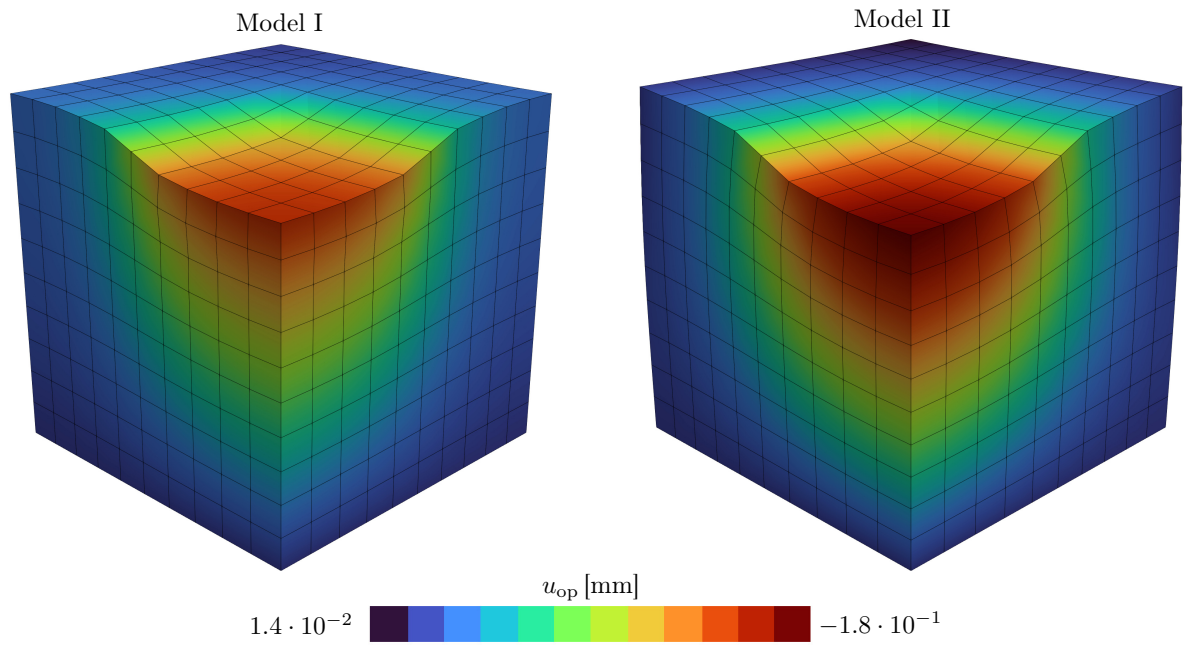


Figure 2.8.: Out-of-plane displacement for the new split (Model I, left) and classical split (Model II, right).

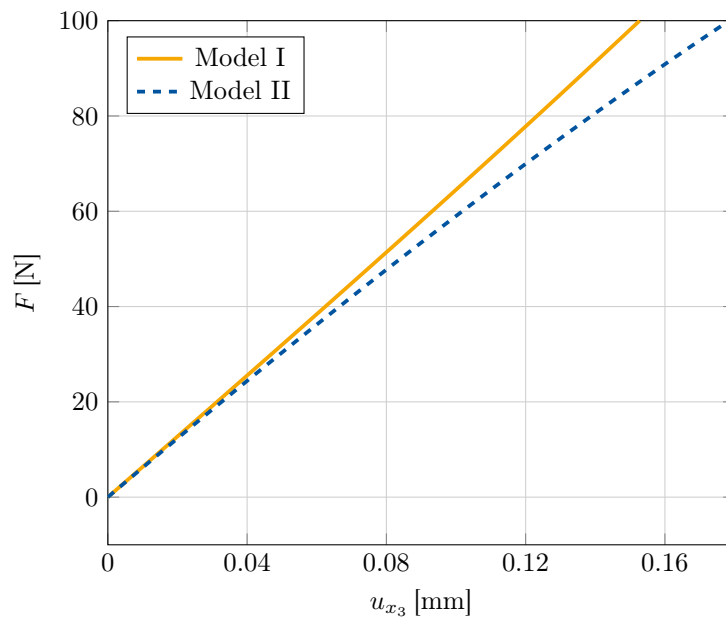


Figure 2.9.: Load-displacement-curve at node 1

The results are shown in Fig. 2.10. The upper line shows the *op* tension tests, whereas the lower line gives the results of the *ip* tests. The components of the inelastic right Cauchy-Green deformation tensor $\mathbf{C}_i$ are plotted in the three preferential directions, where C_{i11} and C_{i22} refer to the two *ip* directions and C_{i33} to the *op* one. Potential 1 describes an inelastic strain evolution only in the *ip* direction. This results in no inelastic strains in any direction when *op* loading is applied. For an *ip* loading, inelastic *ip* strains emerge identically as no distinction into the different *ip* directions is made. In the middle column, the results of the *ip* and *op* tension tests with the second potential are plotted. A difference in the *op* test compared to the first potential can be seen, as the inelastic strain in the *op* direction C_{i33} evolves with time. The *ip* tension test with potential 2 yields identical results compared to potential 1. This is reasonable as the only difference in both formulations is the additional *op* term in the second potential, which is not affected in an *ip* tension test. Consequently, no difference can be observed in the *ip* tests for the first and second potential.

In the third potential, the decoupling effect is removed since no structural tensors are involved in the formulation of the potential. Thus, an *op* displacement leads to an increase in inelastic strains in the *op* direction and a decrease in the *ip* directions during the loading process (cf. Fig. 2.10, upper right plot). During the reloading, a decrease in the *op* inelastic strain can be observed, whereas the *ip* directions expand. Additionally, the applied *ip* displacement leads to an increase of inelastic strains in the loading direction only, and a contraction in both other directions. For the deloading state, both directions orthogonal to the loading direction contract whereas the *ip* inelastic strain in loading direction decreases again. It can be concluded that the three different potential formulations are able to cover different material behavior and the decoupling in inelastic deformations can be diminished.

The deviation of the terms $\text{tr}(\boldsymbol{\tau}^2) = \text{tr}((\mathbf{C}\mathbf{S})^2)$ and $\text{tr}(\bar{\mathbf{I}}^2)$ (cf. Section 2.4.4.2) has been observed in the single-element tests. However, no graphical representation is shown here, as the difference is zero.

2.6.4. Pinched cylinder with inelastic material behavior

The characteristics of the inelastic material formulation are examined by use of structural examples in the following. Firstly, a pinched cylinder test, initially proposed by Simo and Kennedy [1992] and applied by, for example, Miehe [1997]; Sorić et al. [1997]; Eberlein and Wriggers [1999], is simulated. Its geometry and boundary conditions are shown in Fig. 2.11. Because of symmetry, only one-eighth of the cylinder is simulated as indicated in the graphical representation. The open edge of the cylinder (at $x_3 = 300$ mm) is fixed in x_1 and x_2 directions. A displacement of $w = 100$ mm is applied in x_2 direction over 100 seconds in the middle of the

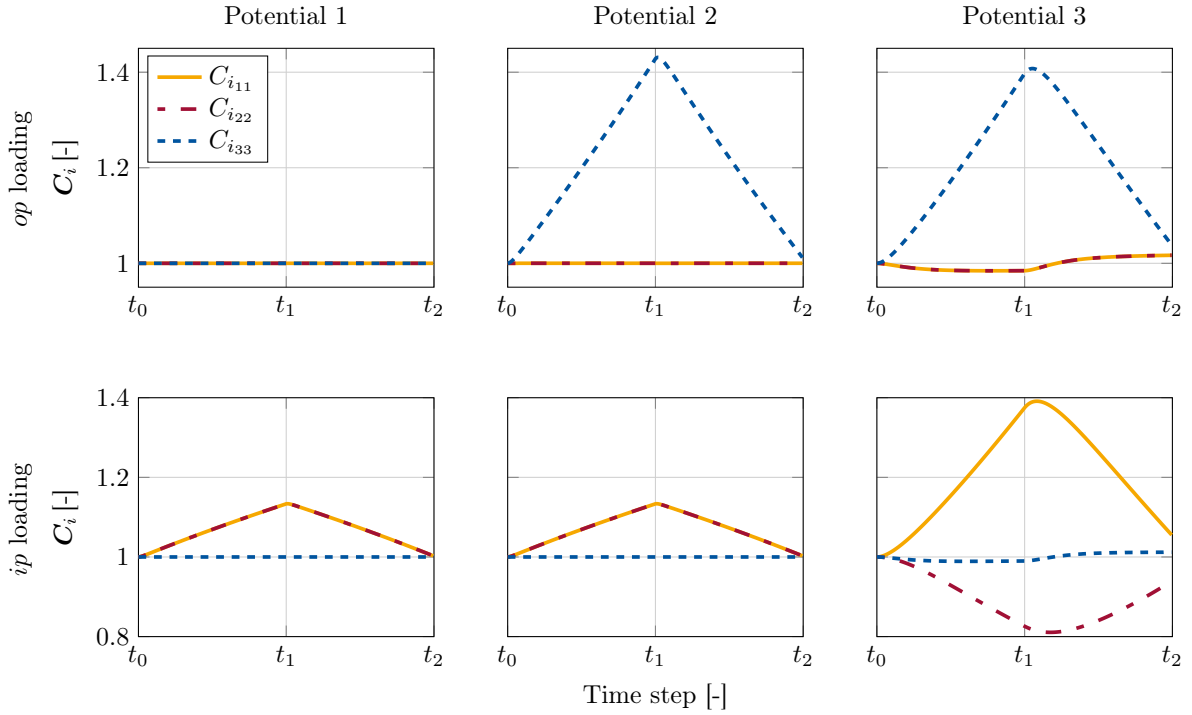


Figure 2.10.: Evolution of the components of the inelastic right Cauchy-Green strain for uniaxial tension in out-of-plane direction (x_3 -direction) (upper line) and in-plane direction (x_1 -direction) (lower line).

cylinder. The material parameters are $K_1 = 80 \left[\frac{\text{N}}{\text{mm}^2} \right]$, $K_2 = 400 \left[\frac{\text{N}}{\text{mm}^2} \right]$, $K_3 = 400 \left[\frac{\text{N}}{\text{mm}^2} \right]$, $K_4 = 10 \left[\frac{\text{N}}{\text{mm}^2} \right]$, and $\tau = 10 \left[\frac{1}{\text{s}} \right]$ are used. The *op* direction is defined as pointing outwards, normal to the cylinder surface, and is calculated for each element separately.

A mesh convergence study has been performed and is given in Fig. 2.12. The sum of all reaction forces F at the edge $x_3 = 0$ is plotted over the number of elements. It is given by the number of elements in thickness direction multiplied by the elements in x_3 -direction multiplied by the elements in radial direction. Convergence can be assumed for a mesh size of $2 \times 32 \times 64$ elements which is given in Fig. 2.12 on the right.

The calculations are performed with all three potentials defined in Section 2.4.4.2. As a result, Fig. 2.13 shows the *op* term of the inelastic right Cauchy-Green tensor, C'_{iop} . Since the displacement w is applied in the *op* direction, it is expected that inelastic strain evolution can only be observed for those potentials that explicitly depend on this *op* term, i.e. potentials 2 and 3.

The first potential defines an inelastic strain evolution only in the *ip* directions. Thus, the *op* component of the inelastic right Cauchy-Green tensor, C'_{iop} , does not evolve (cf. Fig. 2.13 (left)). Due to the additional *op* term in potential 2 in Equation (2.54), the corresponding plot (Fig. 2.13

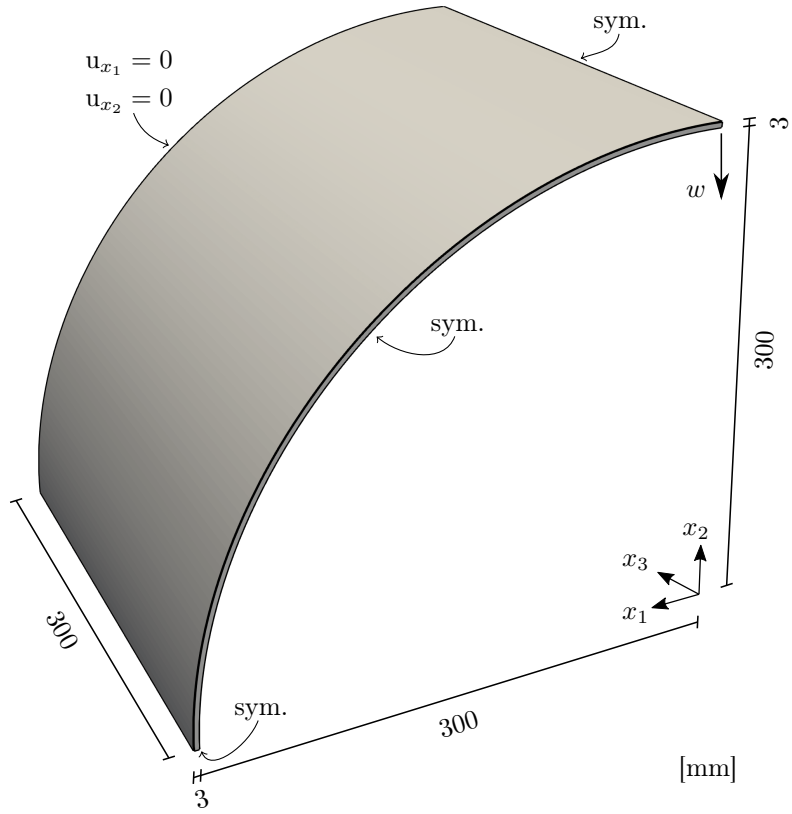


Figure 2.11.: Geometry and boundary conditions of the pinched cylinder.

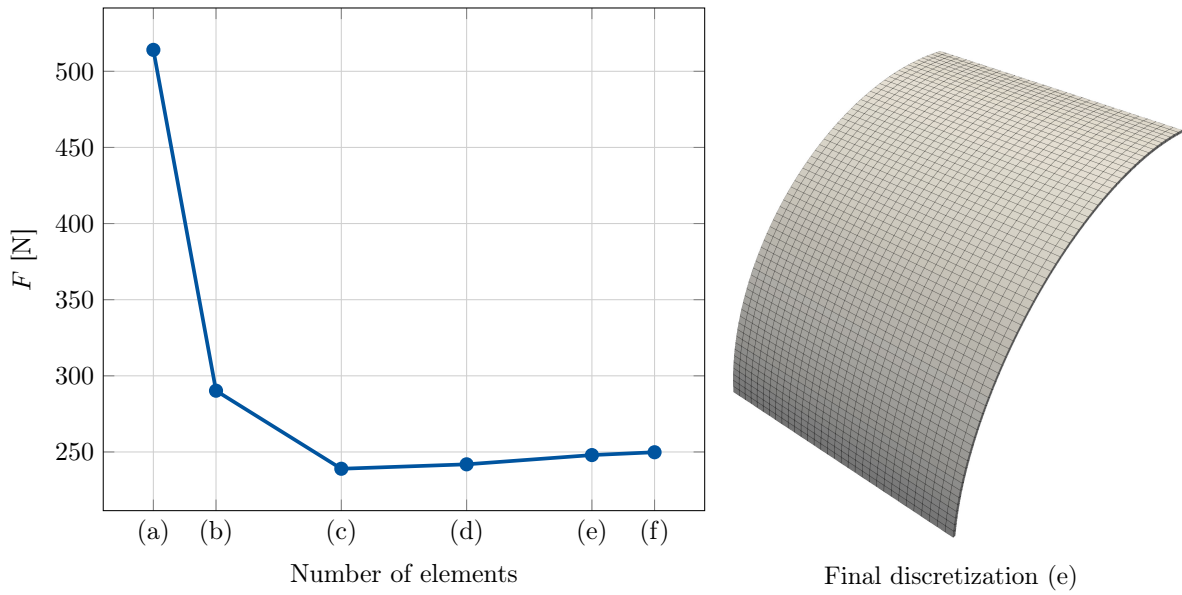


Figure 2.12.: Convergence study of pinched cylinder: Sum over all reaction forces F at edge $x_3 = 0$ over the number of elements: (a) $2 \times 4 \times 8$, (b) $2 \times 8 \times 16$ (c) $2 \times 16 \times 32$, (d) $2 \times 24 \times 48$, (e) $2 \times 32 \times 64$, (f) $2 \times 36 \times 72$ (left) and the converged geometry discretization (e) (right).

(middle)) reveals inelastic strains in the *op* direction. However, these inelastic *op* strains are relatively small and localized to a narrow region. On the contrary, when potential 3 is utilized, the resulting inelastic strains in *op* direction have higher magnitudes and their distribution is spread wider within the structure (cf. Fig. 2.13 (right)). This can be explained by the fact that the third potential prevents a decoupling for inelastic deformations, and consequently, the inelastic *op* strains are influenced by all deformation states.

In addition, the resulting inelastic *ip* strain distributions are displayed in Fig. 2.14. As expected, the potentials 1 and 2 yield the same results in the *ip* direction, because the two potentials only differ in the additional *op* term in potential 2, which does not influence the *ip* evolutions. When the third potential is used to compute the inelastic strains, the distribution within the structure is qualitatively similar compared to potentials 1 and 2. However, the values are quantitatively less significant.

In summary, the presented *ip-op*-decoupling yields promising and reasonable results even for inelastic deformations, as demonstrated by the pinched cylinder test with three different potentials. No inelastic *op* strains emerge when potential 1 is used. Potential 2 leads to a slight increase in inelastic *op* strains, and the third potential results in a significant increase in these strains.

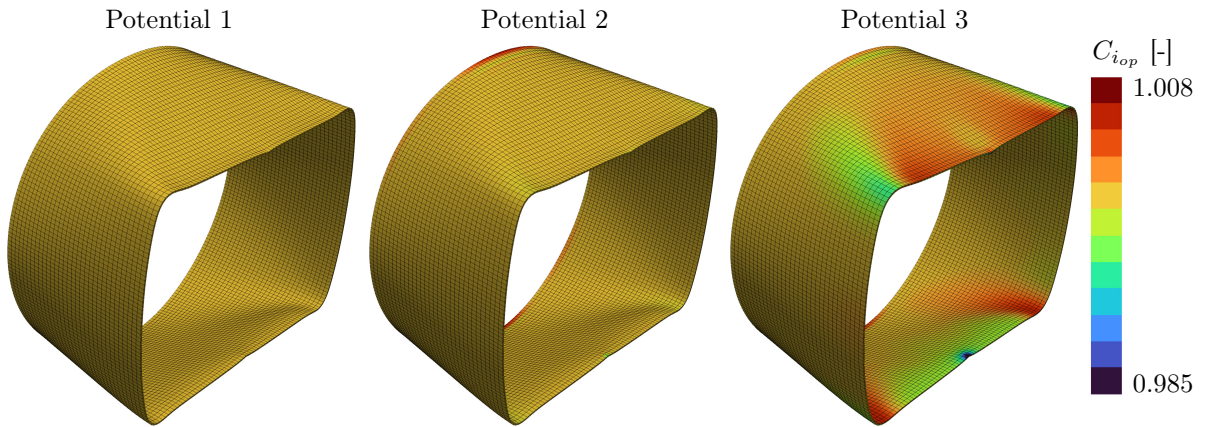


Figure 2.13.: Out-of-plane component, $C_{i_{op}}$, of the inelastic right Cauchy Green deformation tensor computed with potential 1 (left), potential 2 (middle) and potential 3 (right).

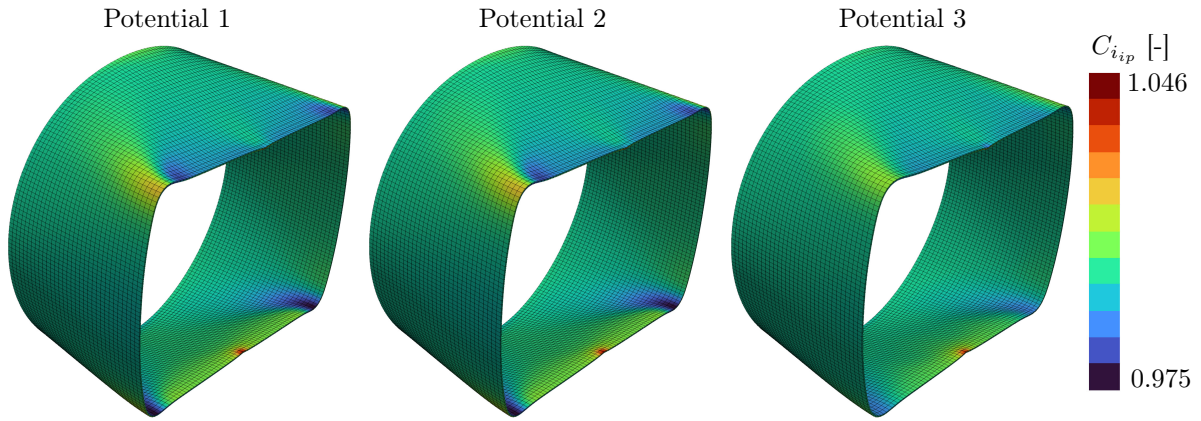


Figure 2.14.: In-plane component, C_{ip} , of the inelastic right Cauchy Green deformation tensor computed with potential 1 (left), potential 2 (middle) and potential 3 (right).

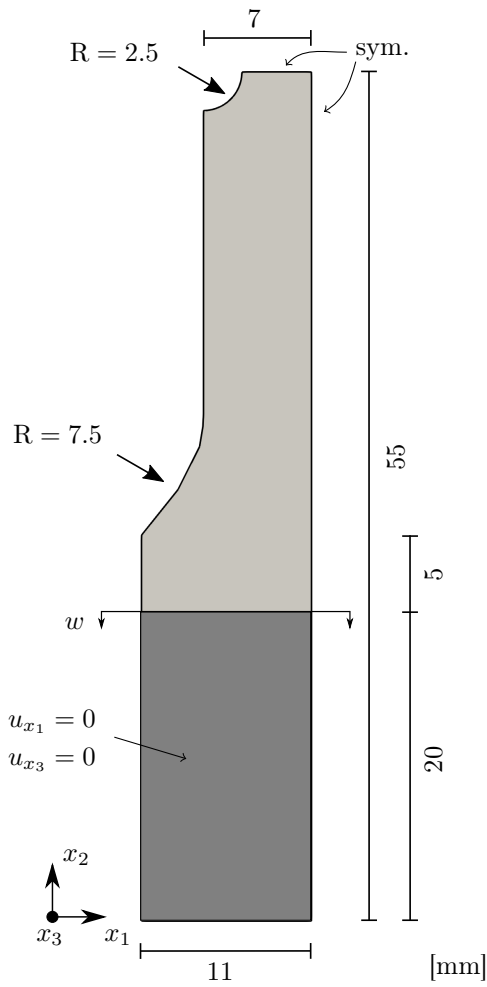


Figure 2.15.: Geometry and boundary conditions of the symmetrically notched specimen with a depth of 2.5 [mm].

2.6.5. Symmetrically notched specimen with inelastic material behavior

Finally, the elastic and inelastic formulations are compared in an additional structural example of a symmetrically notched specimen. Its geometry and boundary conditions are shown in Fig. 2.15 where the darker grey part is fixed in x_1 - and x_3 -directions and is pulled in negative x_2 -direction. Because of symmetry, only one-quarter of the specimen is simulated as indicated in the graphical representation. Symmetry effects in x_3 -direction, which is the *op* direction, are not exploited as the structure is already thin and only three elements are employed in this direction. The displacement $w = 110$ mm is applied linearly over 11 seconds at $x_2 = 20$ mm. The same material parameters as in the previous example are applied.

A mesh convergence study has been performed. For this, the mesh has mainly been refined close to the notch. In Fig. 2.16 on the left, the sum of the reaction forces in x_2 direction F_{x_2} at $x_2 = 55$ mm is plotted over the total number of elements. Convergence can be assumed for a number of elements $n_{el} = 4227$. The corresponding discretization is illustrated on the right in Fig. 2.16.

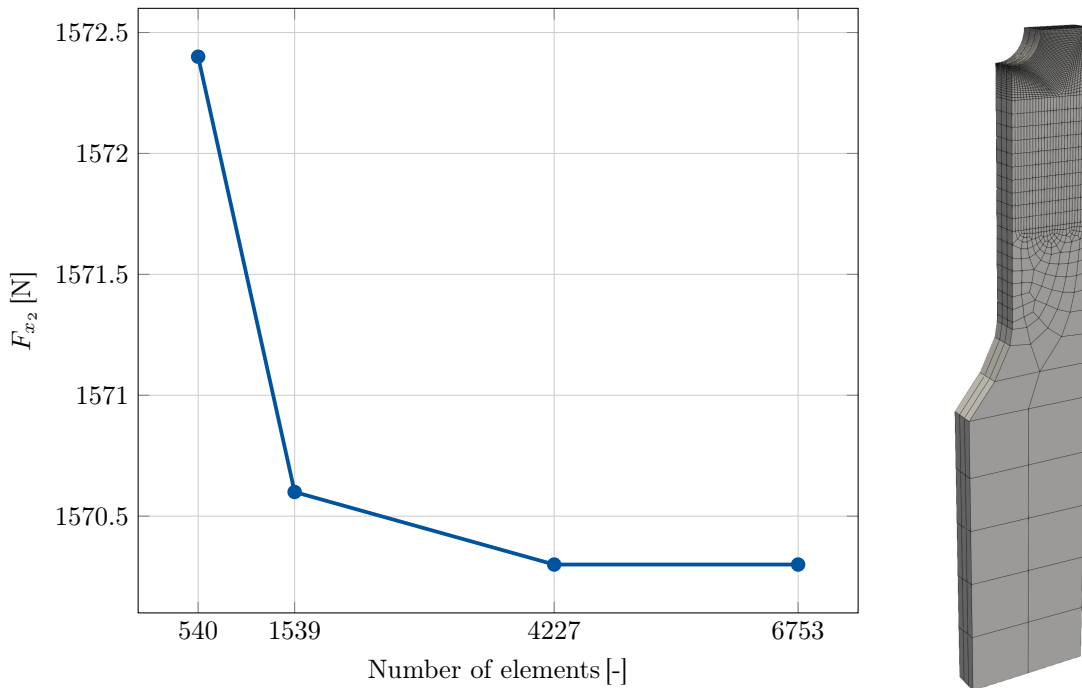


Figure 2.16.: Convergence study of symmetrically notched specimen: Sum of reaction forces in x_2 -direction F_{x_2} over the number of elements (left) and final discretization (right).

The example has been chosen since a stress concentration can be expected in the vicinity

of the notches, which normally leads to a contraction of the specimen in thickness direction particularly in this area. As a result, the distribution of the Cauchy stress in loading direction, σ_{22} , is given in the upper line in Fig. 2.17 for the elastic (a) and inelastic material behavior with all three potentials ((b) to (d)). Due to a rather constant distribution at the edges of the specimen, only the middle part is shown. As expected, a stress concentration near the notches can be observed. Without the decoupling of *ip* and *op*, a contraction in thickness direction can be anticipated. However, the presented framework prevents deformation in the *op* direction for the elastic material formulation and the inelastic one with the first and second potential (cf. Fig. 2.17, lower line (a) to (c)). In contrast, as previously described in Sections 2.4.4.2 and 2.6.3, the third potential removes the decoupling and inelastic strains emerge independently from *ip* or *op* directions. Thus, an evolution of the Green-Lagrange strain E_{33} can be observed for the calculation with the third potential as in the lower line in Fig. 2.17 (d). In addition to these stress and strain plots, the force-displacement curves for all four simulations are given in Fig. 2.18. In particular, the reaction forces in x_2 direction at $x_2 = 20$ mm, F_{x_2} , are given over the applied displacement w . A significant deviation of the elastic to the inelastic material formulation is observed. The curves of potentials 1 and 2 are the same, as the formulations only differ in an additional *op* term in the second potential which is not affected here. The use of the third potential removes the decoupling. Therefore, a softer material behavior can be observed, as discussed in Section 2.6.2 which leads to a less steep force-displacement curve (cf. Fig. 2.18).

2.7. Conclusion

The main objective of this work was to derive a constitutive model with focus on a decoupled material response for the in-plane (*ip*) and out-of-plane (*op*) behavior. This is motivated by the material behavior of paper and paperboard as its response in thickness direction is independent of the *ip* behavior perpendicular to it. Nonetheless, the formulation is not restricted to these materials.

The decoupling is, first, incorporated into an elastic material model. For this, structural tensors are used to capture the different material directions. An *ip* structural tensor is defined and applied on the right Cauchy-Green deformation tensor. This divides the deformation tensor into an *ip* and *op* part. The *ip* term is used for a split into area-changing and isoplanar (area-preserving) deformations which is derived in analogy to the classical isochoric-volumetric split for volumes. Using the split and the structural tensors, a Helmholtz free energy is formulated,

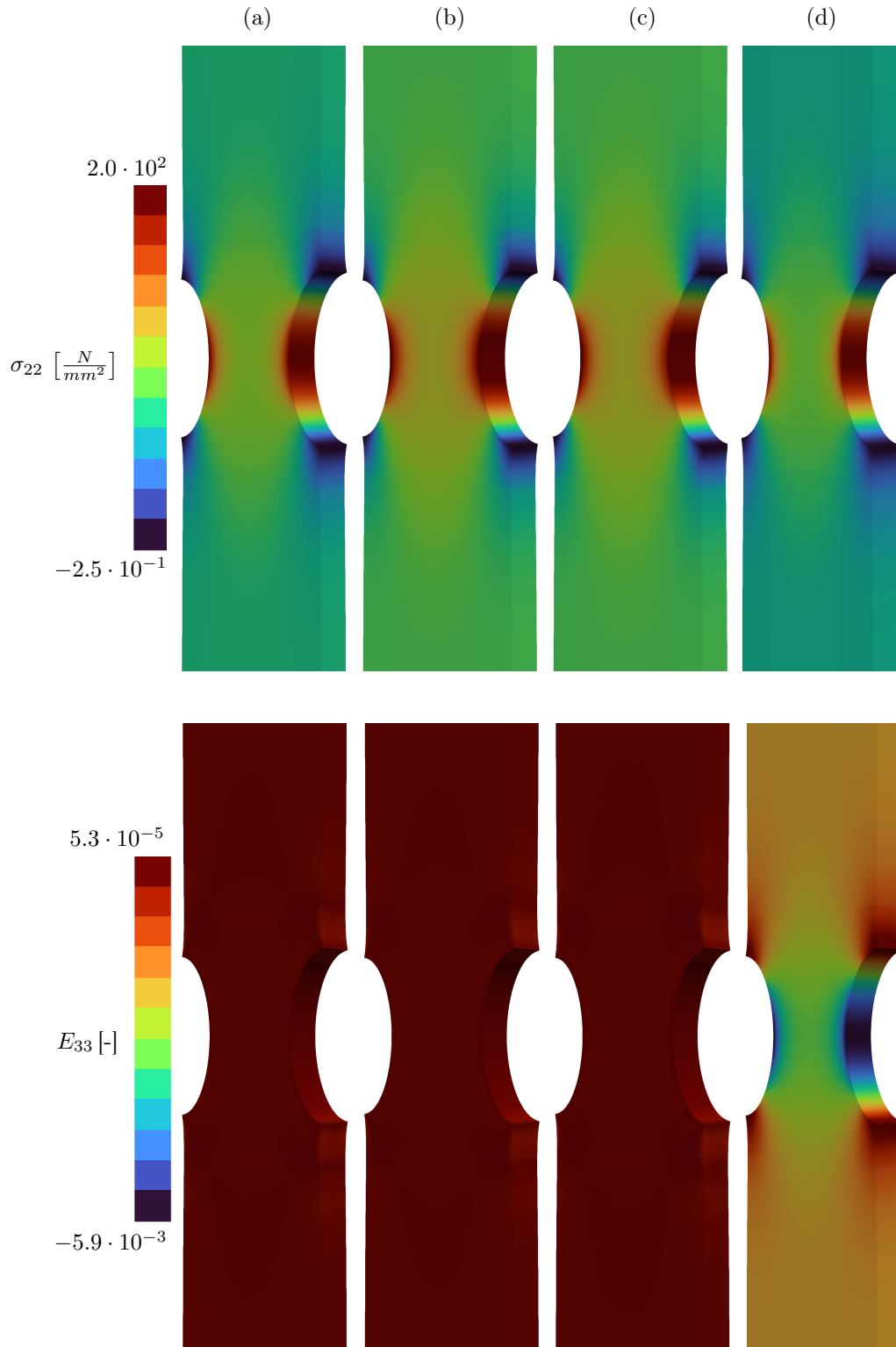


Figure 2.17.: Cauchy stress σ_{22} (upper line) and Green-Lagrange strain E_{33} (lower line) for elastic material (a), inelastic material with potential 1 (b), potential 2 (c), and potential 3 (d).

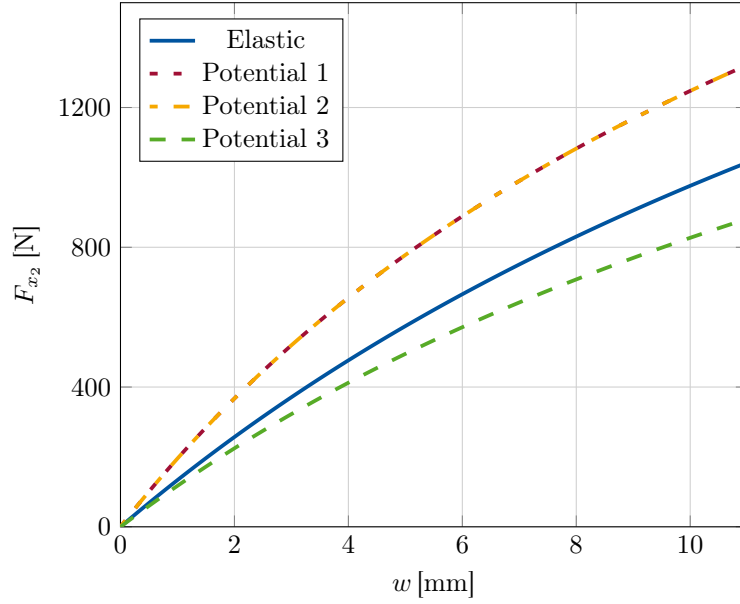


Figure 2.18.: Reaction forces in x_2 -direction, F_{x_2} , at $x_2 = 20$ mm over displacement w in x_2 -direction.

capturing *ip* and *op* terms independently and leading to a decoupled material response.

To prove the capabilities of the decoupled framework under more complex material behavior, an inelastic material formulation is introduced. To reduce the effort in the derivations and implementation, a co-rotated intermediate configuration is introduced that – in contrast to the standard intermediate configuration – is uniquely defined. Thus, all quantities can be directly derived and are implemented by automated differentiation. The mapping of the structural tensors into the different configurations has been shown. By use of the Clausius-Duhem inequality, the constitutive equations of the model are derived in the co-rotated intermediate configuration and evolution equations of the inelastic strains are defined through potentials. Three different potentials are introduced, each capturing different behaviors. The first one leads to inelastic in-plane evolution only when in-plane loadings are applied. Potential 2 differs from the first potential only by one additional out-of-plane term, whereas the decoupling of in-plane and out-of-plane deformations is removed in the third potential. All three potentials are introduced by the physically reasonable Kirchhoff stress in the current configuration. As the model is implemented in the co-rotated intermediate configuration, the corresponding counterparts in this very configuration are defined.

Finally, specific choices of the Helmholtz free energy are presented for the elastic and inelastic material formulations. To compare the presented decoupled framework to a coupled one, a free energy with the classical isochoric-volumetric split is given. With the introduction of a

specific energy formulation, the detailed forms of the second Piola-Kirchhoff and the material tangent are presented for the elastic case. In addition, the material tangent at the initial state is given to underline the successful decoupling.

The presented framework for elastic and inelastic material behavior is then examined by single-element tests that show the performance of the elastic material formulation at Gauss point level. In the single-element tests, the separation of the *ip* and *op* material response can be observed. In addition, the elastic material formulation with the new isoplanar-area-changing split is compared to the classical isochoric-volumetric split in the structural example of a cube under compression in *op* direction. The main difference lies in the expansion in *ip* directions in the classical model, whereas the decoupled framework remains nearly undeformed in *ip* directions, which underlines the capability of the decoupled framework. Furthermore, the inelastic material formulation is investigated by uniaxial tension tests at a single-element test, where the influence of the three different potentials is analyzed. The first potential leads to inelastic strain evolution only for applied *ip* loadings, whereas potential 2 captures *ip* and *op* elastic strain evolution separately but the decoupling of the material response remains unaffected. Furthermore, the use of potential 3 leads to a coupled material response for inelastic deformations. These investigations are underlined by the results of the structural example of a pinched cylinder. Finally, the inelastic material formulation with all three potentials is compared to the elastic material formulation at a symmetrically notched specimen where it can be observed that the decoupling remains even under localized stress states and only the formulation with the third potentials leads to *op* deformations when *ip* loadings are applied. Thus, the successful decoupling of *ip* and *op* behavior can be shown for inelastic material responses and complex deformation states.

Despite these promising results of the decoupled framework, an extension of the material formulation to, for example, plasticity is of interest. In addition, the model should be adapted specifically to the behavior of paper or paperboard, such that a comparison of numerical results to experimental data can be performed. For this, the anisotropy in the two in-plane directions should be added to the model in a first step as in the presented model the in-plane response is considered isotropic. To adequately model the material response of paper, this anisotropy is a major aspect for future works.

2.8. Appendix

2.8.1. The modified right Cauchy-Green tensor in the isochoric-volumetric framework

In the isochoric-volumetric framework, the right Cauchy-Green tensor is modified by the volume ratio J (cf. Equation (2.4)) as defined in Equation (2.1). Its material time derivative is then introduced by

$$\dot{\mathbf{C}}_{\text{isov}} = \frac{\partial \mathbf{C}_{\text{isov}}}{\partial \mathbf{C}} : \dot{\mathbf{C}} = J^{-1/3} \left(\mathbb{I} - \underbrace{\frac{1}{3} \mathbf{C} \otimes \mathbf{C}^{-1}}_{\mathbb{P}_{\text{isov}}^T} \right). \quad (2.75)$$

Here, the relation $J^{1/3} \frac{\partial J^{-1/3}}{\partial \mathbf{C}} = -\frac{1}{3} \mathbf{C}^{-1}$ has been used. $\mathbb{P}_{\text{isov}}^T$ is the transpose of the isochoric fourth-order projection tensor. For more information, the reader is referred to, for example, Holzapfel [2002].

2.8.2. Proof of decoupled formulation of second Piola-Kirchhoff stress formulation

To prove the plausibility, the relations

$$\mathbf{S} : (\mathbf{n}_3 \otimes \mathbf{n}_3) = \mathbf{S} : (\mathbf{n}_1 \otimes \mathbf{n}_3) = \mathbf{S} : (\mathbf{n}_3 \otimes \mathbf{n}_1) = \mathbf{S} : (\mathbf{n}_2 \otimes \mathbf{n}_3) = \mathbf{S} : (\mathbf{n}_3 \otimes \mathbf{n}_2) = 0 \quad (2.76)$$

can easily be shown by insertion of any arbitrary $\mathbf{C}$. Consequently, *ip* and *op* material behavior can be defined independently in the presented framework.

2.8.3. Rates of tensors

For the derivative of the Helmholtz free energy in Equation (2.36), the rate of the structural tensor $\bar{\mathbf{M}}_3$ has to be computed. Therefore, its form in the *cic* in Equation 2.34 is recalled. Its

rate then reads

$$\begin{aligned}\dot{\bar{\mathbf{M}}}_3 &= \frac{1}{\mathbf{C}_i^{-1} : \bar{\mathbf{M}}_3} \left(\dot{\mathbf{U}}_i^{-1} \bar{\mathbf{M}}_3 \mathbf{U}_i^{-1} + \mathbf{U}_i^{-1} \dot{\bar{\mathbf{M}}}_3 \mathbf{U}_i^{-1} + \mathbf{U}_i^{-1} \bar{\mathbf{M}}_3 \dot{\mathbf{U}}_i^{-1} \right) \\ &\quad - \frac{1}{(\mathbf{C}_i^{-1} : \bar{\mathbf{M}}_3)^2} \left(\dot{\mathbf{C}}_i^{-1} : \bar{\mathbf{M}}_3 + \mathbf{C}_i^{-1} : \dot{\bar{\mathbf{M}}}_3 \right) \mathbf{U}_i^{-1} \bar{\mathbf{M}}_3 \mathbf{U}_i^{-1} \\ &= 2 (\bar{\mathbf{L}}_i : \bar{\mathbf{M}}_3) \bar{\mathbf{M}}_3 - \bar{\mathbf{L}}_i \bar{\mathbf{M}}_3 - \bar{\mathbf{M}}_3 \bar{\mathbf{L}}_i^T.\end{aligned}\quad (2.77)$$

Additionally, the rate of $\bar{\mathbf{C}}_e$ reads

$$\begin{aligned}\dot{\bar{\mathbf{C}}}_e &= \dot{\mathbf{U}}_i^{-1} \mathbf{C} \mathbf{U}_i^{-1} + \mathbf{U}_i^{-1} \dot{\mathbf{C}} \mathbf{U}_i^{-1} + \mathbf{U}_i^{-1} \mathbf{C} \dot{\mathbf{U}}_i^{-1} \\ &= \mathbf{U}_i^{-1} \dot{\mathbf{C}} \mathbf{U}_i^{-1} - \bar{\mathbf{C}}_e \bar{\mathbf{L}}_i - \bar{\mathbf{L}}_i^T \bar{\mathbf{C}}_e.\end{aligned}\quad (2.78)$$

The relations $\bar{\mathbf{L}}_i = \dot{\mathbf{U}}_i \mathbf{U}^{-1}$ and $\dot{\mathbf{C}}_i = 2 \mathbf{U}_i \bar{\mathbf{D}}_i \mathbf{U}_i$ and $\bar{\mathbf{D}}_i := \text{sym}(\bar{\mathbf{L}}_i)$ have been used here.

2.8.4. Derivative of the Helmholtz free energy

Using the chain rule, the time derivative of the Helmholtz free energy, $\dot{\psi}$, as a scalar-valued function of $\bar{\mathbf{C}}_e$ and $\bar{\mathbf{M}}_3$ can be computed by

$$\begin{aligned}\frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} : \dot{\bar{\mathbf{M}}}_3 &= -\text{tr} \left(\frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} \bar{\mathbf{L}}_i^T \bar{\mathbf{M}}_3 \right) - \text{tr} \left(\frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} \bar{\mathbf{M}}_3 \bar{\mathbf{L}}_i \right) \\ &\quad + 2 \text{tr} \left(\frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} (\bar{\mathbf{M}}_3 \otimes \bar{\mathbf{M}}_3) \bar{\mathbf{L}}_i \right)\end{aligned}\quad (2.79)$$

$$= \left(-2 \bar{\mathbf{M}}_3 \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} + 2 \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} : (\bar{\mathbf{M}}_3 \otimes \bar{\mathbf{M}}_3) \right) : \bar{\mathbf{L}}_i \quad (2.80)$$

and

$$\frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} : \dot{\bar{\mathbf{C}}}_e = \mathbf{U}_i^{-1} \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} \mathbf{U}_i^{-1} : \dot{\mathbf{C}} - 2 \bar{\mathbf{C}}_e \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} : \bar{\mathbf{L}}_i. \quad (2.81)$$

This leads to

$$\begin{aligned}\dot{\psi} &= 2 \mathbf{U}_i^{-1} \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} \mathbf{U}_i^{-1} : \frac{1}{2} \dot{\mathbf{C}} + \left(-2 \bar{\mathbf{M}}_3 \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} + 2 \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} : (\bar{\mathbf{M}}_3 \otimes \bar{\mathbf{M}}_3) \right) : \bar{\mathbf{L}}_i \\ &\quad - 2 \bar{\mathbf{C}}_e \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} : \bar{\mathbf{L}}_i.\end{aligned}\quad (2.82)$$

2.8.5. Comparison of the Kirchhoff stress and its conjugate force

The Helmholtz free energy – depending on C_e and $\tilde{M}_3$ – can be formulated as an isotropic function of the following (mixed) invariants:

$$I_1 = \text{tr}(C_e), \quad I_2 = \text{tr}(C_e^2), \quad I_3 = \text{tr}(C_e^3), \quad I_4 = \text{tr}(C_e \tilde{M}_3), \quad I_5 = \text{tr}(C_e^2 \tilde{M}_3). \quad (2.83)$$

With this, the Kirchhoff stress can be defined as given in Equation (2.45) and the relation of the Kirchhoff stress τ to the stress-like quantity $\tilde{\Gamma}$ can be found to reformulate the potentials in the *ic* or *cic*. Thus, the traces of both are computed and compared, i.e.

$$\begin{aligned} \text{tr}(\tau) = & 2 \left[\alpha_1 \text{tr}(B_e) + 2 \alpha_2 \text{tr}(B_e^2) + 3 \alpha_3 \text{tr}(B_e^3) + \alpha_4 \text{tr}(B_e^2 \hat{M}_3^*) \right. \\ & \left. + \alpha_5 \left(\text{tr}(B_e^3 \hat{M}_3^*) + \text{tr}(B_e \hat{M}_3^*) \right) \right] \end{aligned} \quad (2.84)$$

$$\begin{aligned} = & 2 \left[\alpha_1 \text{tr}(C_e) + 2 \alpha_2 \text{tr}(C_e^2) + 3 \alpha_3 \text{tr}(C_e^3) + \alpha_4 \text{tr}(C_e \tilde{M}_3) \right. \\ & \left. + 2 \alpha_5 \text{tr}(C_e^2 \tilde{M}_3) \right] \end{aligned} \quad (2.85)$$

$$\begin{aligned} \text{tr}(\tilde{\Gamma}) = & 2 \left[\alpha_1 \text{tr}(C_e) + 2 \alpha_2 \text{tr}(C_e^2) + 3 \alpha_3 \text{tr}(C_e^3) + \alpha_4 \left(2 \text{tr}(C_e \tilde{M}_3) - I_4 \text{tr}(\tilde{M}_3) \right) \right. \\ & \left. + \alpha_5 \left(3 \text{tr}(C_e^2 \tilde{M}_3) - I_5 \text{tr}(\tilde{M}_3) \right) \right] \end{aligned} \quad (2.86)$$

$$= \text{tr}(\tau) \quad (2.87)$$

since $\text{tr}(\tilde{M}_3) = 1$. It can be seen that both formulations are equal and, thus, the potentials can equivalently be formulated in the *cic*.

The third potential requires the relation of $\text{tr}(\tau^2)$ to $\text{tr}(\tilde{\Gamma}^2)$ for the implementation. Thus, the

trace of the squared Kirchhoff stress reads

$$\begin{aligned}
 \text{tr}(\boldsymbol{\tau}^2) = & 4 \left[\alpha_1^2 \text{tr}(\mathbf{B}_e^2) + 2 \alpha_1 \alpha_2 \text{tr}(\mathbf{B}_e^3) + 3 \alpha_1 \alpha_3 \text{tr}(\mathbf{B}_e^4) + \alpha_1 \alpha_4 \text{tr}(\mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e) \right. \\
 & + \alpha_1 \alpha_5 \left[\text{tr}(\mathbf{B}_e^3 \hat{\mathbf{M}}_3^* \mathbf{B}_e) + \text{tr}(\mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e^2) \right] \\
 & + 2 \alpha_2 \alpha_1 \text{tr}(\mathbf{B}_e^3) + 4 \alpha_2^2 \text{tr}(\mathbf{B}_e^4) + 6 \alpha_2 \alpha_3 \text{tr}(\mathbf{B}_e^5) + 2 \alpha_2 \alpha_4 \text{tr}(\mathbf{B}_e^3 \hat{\mathbf{M}}_3^* \mathbf{B}_e) \\
 & + 2 \alpha_2 \alpha_5 \left[\text{tr}(\mathbf{B}_e^4 \hat{\mathbf{M}}_3^* \mathbf{B}_e) + \text{tr}(\mathbf{B}_e^3 \hat{\mathbf{M}}_3^* \mathbf{B}_e^2) \right] \\
 & + 3 \alpha_3 \alpha_1 \text{tr}(\mathbf{B}_e^4) + 6 \alpha_3 \alpha_2 \text{tr}(\mathbf{B}_e^5) + 9 \alpha_3^2 \text{tr}(\mathbf{B}_e^6) + 3 \alpha_3 \alpha_4 \text{tr}(\mathbf{B}_e^4 \hat{\mathbf{M}}_3^* \mathbf{B}_e) \\
 & + 3 \alpha_3 \alpha_5 \left[\text{tr}(\mathbf{B}_e^5 \hat{\mathbf{M}}_3^* \mathbf{B}_e) + \text{tr}(\mathbf{B}_e^4 \hat{\mathbf{M}}_3^* \mathbf{B}_e^2) \right] \\
 & + \alpha_4 \alpha_1 \text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^2) + 2 \alpha_4 \alpha_2 \text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^3) + 3 \alpha_4 \alpha_3 \text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^4) \\
 & + \alpha_4^2 \text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e) \\
 & + \alpha_4 \alpha_5 \left[\text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^3 \hat{\mathbf{M}}_3^* \mathbf{B}_e) + \text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e^2) \right] \\
 & + \alpha_5 \alpha_1 \left[\text{tr}(\mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e^2) + \text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^3) \right] \\
 & + 2 \alpha_5 \alpha_2 \left[\text{tr}(\mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e^3) + \text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^4) \right] \\
 & + 3 \alpha_5 \alpha_3 \left[\text{tr}(\mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e^4) + \text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^5) \right] \\
 & + \alpha_5 \alpha_4 \left[\text{tr}(\mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e) + \text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^3 \hat{\mathbf{M}}_3^* \mathbf{B}_e) \right] \\
 & + \alpha_5^2 \left[\text{tr}(\mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e^3 \hat{\mathbf{M}}_3^* \mathbf{B}_e) + 2 \text{tr}(\mathbf{B}_e \hat{\mathbf{M}}_3^* \mathbf{B}_e^4 \hat{\mathbf{M}}_3^* \mathbf{B}_e) \right. \\
 & \left. \left. + \text{tr}(\mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e^2 \hat{\mathbf{M}}_3^* \mathbf{B}_e^2) \right] \right] \quad (2.88)
 \end{aligned}$$

After reformulation, it can be described by

$$\begin{aligned}
 \text{tr}(\boldsymbol{\tau}^2) = & 4 \left[\alpha_1^2 \text{tr}(\mathbf{C}_e^2) + 4 \alpha_1 \alpha_2 \text{tr}(\mathbf{C}_e^3) + 6 \alpha_1 \alpha_3 \text{tr}(\mathbf{C}_e^4) + 2 \alpha_1 \alpha_4 \text{tr}(\mathbf{C}_e^2 \tilde{\mathbf{M}}_3) \right. \\
 & + 4 \alpha_1 \alpha_5 \text{tr}(\mathbf{C}_e^3 \tilde{\mathbf{M}}_3) + 4 \alpha_2^2 \text{tr}(\mathbf{C}_e^4) + 12 \alpha_2 \alpha_3 \text{tr}(\mathbf{C}_e^5) + 4 \alpha_2 \alpha_4 \text{tr}(\mathbf{C}_e^3 \tilde{\mathbf{M}}_3) \\
 & + 8 \alpha_2 \alpha_5 \text{tr}(\mathbf{C}_e^4 \tilde{\mathbf{M}}_3) + 9 \alpha_3^2 \text{tr}(\mathbf{C}_e^6) + 6 \alpha_3 \alpha_4 \text{tr}(\mathbf{C}_e^4 \tilde{\mathbf{M}}_3) + 12 \alpha_3 \alpha_5 \text{tr}(\mathbf{C}_e^5 \tilde{\mathbf{M}}_3) \\
 & + \alpha_4^2 \text{tr}(\mathbf{C}_e \tilde{\mathbf{M}}_3 \mathbf{C}_e \tilde{\mathbf{M}}_3) + 4 \alpha_4 \alpha_5 \text{tr}(\mathbf{C}_e^2 \tilde{\mathbf{M}}_3 \mathbf{C}_e \tilde{\mathbf{M}}_3) \\
 & \left. + \alpha_5^2 \left[\text{tr}(\mathbf{C}_e^2 \tilde{\mathbf{M}}_3 \mathbf{C}_e^2 \tilde{\mathbf{M}}_3) + 2 \text{tr}(\mathbf{C}_e \tilde{\mathbf{M}}_3 \mathbf{C}_e^3 \tilde{\mathbf{M}}_3) + \text{tr}(\mathbf{C}_e^3 \tilde{\mathbf{M}}_3 \mathbf{C}_e \tilde{\mathbf{M}}_3) \right] \right] \quad (2.89)
 \end{aligned}$$

The detailed formulation of $\text{tr}(\tilde{\Gamma}^2)$ gives

$$\begin{aligned}
\text{tr}(\tilde{\Gamma}^2) = & 4 \left[\alpha_1^2 \text{tr}(C_e^2) + 4 \alpha_1 \alpha_2 \text{tr}(C_e^3) + 6 \alpha_1 \alpha_3 \text{tr}(C_e^4) + 2 \alpha_1 \alpha_4 \left[2 \text{tr}(C_e^2 \tilde{M}_3) \right. \right. \\
& - \text{tr}(C_e \tilde{M}_3)^2 \left. \right] + 2 \alpha_1 \alpha_5 \left[3 \text{tr}(C_e^3 \tilde{M}_3) - \text{tr}(C_e^2 \tilde{M}_3) \text{tr}(C_e \tilde{M}_3) \right] \\
& + 4 \alpha_2^2 \text{tr}(C_e^4) + 12 \alpha_2 \alpha_3 \text{tr}(C_e^5) \\
& + 4 \alpha_2 \alpha_4 \left[2 \text{tr}(C_e^3 \tilde{M}_3) - \text{tr}(C_e^2 \tilde{M}_3) \text{tr}(C_e \tilde{M}_3) \right] \\
& + 4 \alpha_2 \alpha_5 \left(3 \text{tr}(C_e^4 \tilde{M}_3) - \text{tr}(C_e^2 \tilde{M}_3)^2 \right) \\
& + 9 \alpha_3^2 \text{tr}(C_e^6) + 6 \alpha_3 \alpha_4 \left[2 \text{tr}(C_e^4 \tilde{M}_3) - \text{tr}(C_e^3 \tilde{M}_3) \text{tr}(C_e \tilde{M}_3) \right] \\
& + 6 \alpha_3 \alpha_5 \left[3 \text{tr}(C_e^5 \tilde{M}_3) - \text{tr}(C_e^3 \tilde{M}_3) \text{tr}(C_e^2 \tilde{M}_3) \right] \\
& + \alpha_4^2 \left[\text{tr}\left((C_e \tilde{M}_3)^2\right) + 2 \text{tr}(C_e^2 \tilde{M}_3) - 3 \text{tr}(C_e \tilde{M}_3)^2 \right] \\
& + 2 \alpha_4 \alpha_5 \left[4 \text{tr}(C_e^2 \tilde{M}_3 C_e \tilde{M}_3) + 2 \text{tr}(C_e^3 \tilde{M}_3) - 3 \text{tr}(C_e^2 \tilde{M}_3) \text{tr}(C_e \tilde{M}_3) \right. \\
& - \left. \text{tr}(C_e \tilde{M}_3) \text{tr}\left((C_e \tilde{M}_3)^2\right) \right] + \alpha_5^2 \left[4 \text{tr}(C_e^3 \tilde{M}_3 C_e \tilde{M}_3) + 3 \text{tr}\left((C_e^2 \tilde{M}_3)^2\right) \right. \\
& + \left. 2 \text{tr}(C_e^4 \tilde{M}_3) - 3 \text{tr}(C_e^2 \tilde{M}_3)^2 - 2 \text{tr}(C_e^2 \tilde{M}_3) \text{tr}\left((C_e^2 \tilde{M}_3)^2\right) \right] \left. \right]. \quad (2.90)
\end{aligned}$$

These definitions show that $\text{tr}(\tau^2) \approx \text{tr}(\tilde{\Gamma}^2)$ if $C_e \approx I$. Therefore, the deviation is examined in Section 2.6.3.

Finally, the intermediate counterpart of $\tau : \hat{M}_3$ is sought. For this, the term in the current configuration is described by

$$\begin{aligned}
\tau : \hat{M}_3 = & 2 \left[\alpha_1 \text{tr}(B_e \hat{M}_3) + 2 \alpha_2 \text{tr}(B_e^2 \hat{M}_3) + 3 \alpha_3 \text{tr}(B_e^3 \hat{M}_3) + \alpha_4 \text{tr}(B_e \hat{M}_3 B_e \hat{M}_3) \right. \\
& + \left. \alpha_5 \left[\text{tr}(B_e^2 \hat{M}_3 B_e \hat{M}_3) + \text{tr}(B_e \hat{M}_3 B_e^2 \hat{M}_3) \right] \right] \\
= & 2 \left[\frac{\alpha_1}{\beta} \text{tr}(\tilde{M}_3) + 2 \frac{\alpha_2}{\beta} \text{tr}(C_e \tilde{M}_3) + 3 \frac{\alpha_3}{\beta} \text{tr}(C_e^2 \tilde{M}_3) + \frac{\alpha_4}{\beta} \text{tr}(\tilde{M}_3^2) \right. \\
& + \left. 2 \frac{\alpha_5}{\beta} \text{tr}(C_e \tilde{M}_3) \right]. \quad (2.91)
\end{aligned}$$

In the ic , one can write

$$\begin{aligned}
\tilde{\Gamma} : C_e^{-1} \tilde{M}_3 &= 2 \left[\alpha_1 \operatorname{tr} \left(C_e C_e^{-1} \tilde{M}_3 \right) + 2 \alpha_2 \operatorname{tr} \left(C_e^2 C_e^{-1} \tilde{M}_3 \right) + 3 \alpha_3 \operatorname{tr} \left(C_e^3 C_e^{-1} \tilde{M}_3 \right) \right. \\
&\quad + \alpha_4 \left[2 \operatorname{tr} \left(\tilde{M}_3 C_e C_e^{-1} \tilde{M}_3 \right) - I_4 \operatorname{tr} \left(\tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) \right] \\
&\quad + \alpha_5 \left[3 \operatorname{tr} \left(\tilde{M}_3 C_e^2 C_e^{-1} \tilde{M}_3 \right) + \operatorname{tr} \left(C_e \tilde{M}_3 C_e C_e^{-1} \tilde{M}_3 \right) \right. \\
&\quad \left. \left. + \operatorname{tr} \left(C_e^2 \tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) - I_5 \operatorname{tr} \left(\tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) \right] \right] \\
&= 2 \left[\alpha_1 \operatorname{tr} \left(\tilde{M}_3 \right) + 2 \alpha_2 \operatorname{tr} \left(C_e \tilde{M}_3 \right) + 3 \alpha_3 \operatorname{tr} \left(C_e^2 \tilde{M}_3 \right) \right. \\
&\quad + \alpha_4 \left[2 \operatorname{tr} \left(\tilde{M}_3 \right) + \operatorname{tr} \left(C_e \tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) - I_4 \operatorname{tr} \left(\tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) \right] \\
&\quad + \alpha_5 \left[3 \operatorname{tr} \left(C_e \tilde{M}_3 \right) + \operatorname{tr} \left(C_e \tilde{M}_3 \right) + \operatorname{tr} \left(C_e^2 \tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) \right. \\
&\quad \left. \left. - I_5 \operatorname{tr} \left(\tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) \right] \right]. \tag{2.92}
\end{aligned}$$

Thus, the relation $\tau : \mathbf{m}_3 = \frac{1}{\beta} \tilde{\Gamma} : \operatorname{sym}(C_e^{-1} \tilde{M}_3)$ is obtained and used for the definition of the potentials with $\beta = \operatorname{tr} \left(C_e^{-1} \tilde{M}_3 \right)$.

2.8.6. Calculation of invariants in different configurations

For the relation of invariants between the different configurations, the following relations hold

$$\operatorname{tr} \left(C_e \tilde{M}_3 \right) = \tilde{C}_{e_{ij}} (\tilde{\mathbf{m}}_i \otimes \tilde{\mathbf{m}}_j) : (\tilde{\mathbf{m}}_3 \otimes \tilde{\mathbf{m}}_3) = \tilde{C}_{e_{33}} \tag{2.93}$$

$$\operatorname{tr} \left(C_e^{-1} \tilde{M}_3 \right) = \tilde{C}_{e_{ij}}^{-1} (\tilde{\mathbf{m}}_i \otimes \tilde{\mathbf{m}}_j) : (\tilde{\mathbf{m}}_3 \otimes \tilde{\mathbf{m}}_3) = \tilde{C}_{e_{33}}^{-1} \tag{2.94}$$

$$\operatorname{tr} \left(C_e \tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) = \tilde{C}_{e_{ij}} (\tilde{\mathbf{m}}_i \otimes \tilde{\mathbf{m}}_j) : \tilde{C}_{e_{33}}^{-1} (\tilde{\mathbf{m}}_3 \otimes \tilde{\mathbf{m}}_3) = \tilde{C}_{e_{33}} \tilde{C}_{e_{33}}^{-1} \tag{2.95}$$

since $\tilde{M}_3 C_e^{-1} \tilde{M}_3 = \tilde{C}_{e_{33}}^{-1} (\tilde{\mathbf{m}}_3 \otimes \tilde{\mathbf{m}}_3)$. This leads to

$$\operatorname{tr} \left(C_e \tilde{M}_3 \right) \operatorname{tr} \left(\tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) = \operatorname{tr} \left(C_e \tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) \tag{2.96}$$

$$\operatorname{tr} \left(C_e^2 \tilde{M}_3 \right) \operatorname{tr} \left(\tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) = \operatorname{tr} \left(C_e^2 \tilde{M}_3 C_e^{-1} \tilde{M}_3 \right) \tag{2.97}$$

2.8.7. Partial Derivatives

The second Piola-Kirchhoff stress tensor and the material tangent in Equations (2.39) and (2.68) are derived by using the relations

$$\frac{\partial J_A}{\partial \mathbf{C}} = \text{tr}(\mathbf{C}\mathbf{\Omega})\mathbf{\Omega} - \mathbf{\Omega}\mathbf{C}\mathbf{\Omega} \quad (2.98)$$

$$\frac{\partial J_A^{-1/2}\mathbf{\Omega}}{\partial \mathbf{C}} = -\frac{1}{2}J_A^{-3/2}\mathbf{\Omega} \otimes (\text{tr}(\mathbf{C}\mathbf{\Omega})\mathbf{\Omega} - \mathbf{\Omega}\mathbf{C}\mathbf{\Omega}) \quad (2.99)$$

$$\frac{\partial \text{tr}(\mathbf{C}\mathbf{\Omega})\mathbf{\Omega}}{\partial \mathbf{C}} = \mathbf{\Omega} \otimes \mathbf{\Omega} \quad (2.100)$$

$$\frac{\partial J_A^{-1}\text{tr}(\mathbf{C}\mathbf{\Omega})\mathbf{\Omega}}{\partial \mathbf{C}} = -J_A^{-2}\text{tr}(\mathbf{C}\mathbf{\Omega})\mathbf{\Omega} \otimes (\text{tr}(\mathbf{C}\mathbf{\Omega})\mathbf{\Omega} - \mathbf{\Omega}\mathbf{C}\mathbf{\Omega}) + J_A^{-1}\mathbf{\Omega} \otimes \mathbf{\Omega} \quad (2.101)$$

$$\frac{\partial \mathbf{\Omega}\mathbf{C}\mathbf{\Omega}}{\partial \mathbf{C}} = \frac{1}{2} \left((\mathbf{\Omega} \otimes \mathbf{\Omega}^T)^{\frac{23}{T}} + (\mathbf{\Omega} \otimes \mathbf{\Omega})^{\frac{24}{T}} \right) \quad (2.102)$$

$$\begin{aligned} \frac{\partial J_A^{-1}\mathbf{\Omega}\mathbf{C}\mathbf{\Omega}}{\partial \mathbf{C}} &= -J_A^{-2}\mathbf{\Omega}\mathbf{C}\mathbf{\Omega} \otimes (\text{tr}(\mathbf{C}\mathbf{\Omega})\mathbf{\Omega} - \mathbf{\Omega}\mathbf{C}\mathbf{\Omega}) \\ &\quad + J_A^{-1}\frac{1}{2} \left((\mathbf{\Omega} \otimes \mathbf{\Omega}^T)^{\frac{23}{T}} + (\mathbf{\Omega} \otimes \mathbf{\Omega})^{\frac{24}{T}} \right) \end{aligned} \quad (2.103)$$

$$\frac{\partial \mathbf{C}\mathbf{M}_3}{\partial \mathbf{C}} = \frac{1}{2} \left((\mathbf{I} \otimes \mathbf{M}_3^T)^{\frac{23}{T}} + (\mathbf{I} \otimes \mathbf{M}_3^T)^{\frac{24}{T}} \right). \quad (2.104)$$

3 | Article 2:

A constitutive model for describing decoupled material behavior in thickness and in-plane directions

This article was published as:

Boes, B., Simon, J. W., and Holthusen, H. (2023). A constitutive model for describing decoupled material behavior in thickness and in-plane directions. *Proceedings in Applied Mathematics and Mechanics (PAMM 2023)*, **23(3)**. <https://doi.org/10.1002/pamm.202300057>.

Disclosure of the individual authors' contributions to the article:

B. Boes reviewed the relevant existing literature, worked together with H. Holthusen on the derivation of the constitutive framework and implemented the material model and the element routine into the finite element software *FEAP*. She set up and performed all simulations, interpreted the results and wrote the preliminary version of the article. It was her idea to extend the model presented in Boes et al. [2023] by an additional formulation to account for additional effects. All co-authors gave valuable conceptual advice, read the manuscript and gave suggestions for improvement. All authors approved the publication of the final version of the manuscript.

3.1. Abstract

An increasing demand for paper and paperboard in engineering applications has emerged throughout the years. This leads to a growing interest in simulation tools to predict the material behavior which is challenging due to the anisotropic structure of the material. It can be divided into in-plane and out-of-plane behavior which is mostly independent of each other resulting from the microstructure. Thus, a constitutive material model that captures in-plane and out-of-plane behavior in a consistent manner is desired to simulate complex deformation states. The decoupling will be incorporated by a decomposition of the kinematics and the energy formulation into in-plane and out-of-plane terms by use of structural tensors. This approach is flexible and can easily be extended to inelastic material behavior. In this work, the introduction of the decoupling into a constitutive material model for inelastic behavior is described. Visco-elasticity is included through inelastic potentials. Three different potential formulations are given where focus is on the third, novel formulation which defines inelastic strain evolution only in out-of-plane directions. The capabilities of this formulation are compared to the remaining potentials by simulations of a pinched cylinder that is loaded in the out-of-plane direction.

3.2. Introduction

Paper and paperboard are gaining importance in numerous technically relevant applications, in particular in the packaging industry, because they are extremely versatile, renewable materials and easily recyclable. This leads to a growing demand for simulation tools to predict the mechanical material behavior.

Due to the manufacturing process and the underlying microstructure resulting thereof, the material response is highly anisotropic. It can be divided into in-plane and out-of-plane behavior where the out-of-plane orientation is defined in thickness direction (c.f. Fig. 3.1). An independent material response between in-plane and out-of-plane directions can be observed. Thus, in-plane deformations lead to in-plane stresses only and similarly for the out-of-plane direction. There exist many models that capture either the in-plane behavior (cf. Xia et al. [2002]; Nygård et al. [2005, 2009]; Borgqvist et al. [2014, 2015]; Tjahjanto et al. [2015]; Li et al. [2016]; Pfeiffer and Kolling [2019]) or the out-of-plane one (c.f. Stenberg [2003]; Nygård [2009]; Li et al. [2018]). An overview on material models for paper and paperboard is given in Simon [2021]. Only few models combine the in-plane and out-of-plane behavior in a decoupled manner in one constitutive model. Some incorporate the decoupling by setting the

corresponding material parameters, for example the Poisson's ratio, to zero (c.f. Nygård et al. [2009]; Huang and Nygård [2010]) which is restrictive from a continuum mechanical point of view and limits the applicability to linear problems. In Borgqvist et al. [2015], the authors introduce the decoupling through structural tensors aligned with the preferential material directions in the elastic material formulation. Nevertheless, the demand for a constitutive material model that captures the decoupling effect in a flexible manner such that an extension to inelastic material behavior is straightforward is desired. Thus, a decoupled material model for visco-elastic behavior using structural tensors is presented. This model is in alignment with the one introduced in Boes et al. [2023] which is a flexible model that can easily be extended to inelastic material behavior. This extension is shown by including visco-elasticity that is incorporated by different inelastic potential formulations that present different inelastic strain evolution. In this work, an additional potential is introduced that defines inelastic strain evolution only for the out-of-plane direction. This is motivated by the densification effect and internal friction which are most dominant in the out-of-plane direction (c.f. Li et al. [2018]). Therefore, a model that captures inelastic strain evolution in the out-of-plane direction separately is straightforward.

3.3. Decoupling of in-plane and out-of-plane material behavior

The decoupling will be included by changing only the kinematics and the energy formulation. These two modifications are included easily in the model and lead to an independent material response in both directions. The split of the kinematics and the energy formulation is a general approach which makes the model flexible for an extension to inelastic material behavior or adaption to different materials.

3.3.1. Structural tensors

The anisotropy is incorporated by structural tensors aligned with the three preferential material directions in analogy as in, for example, Boehler [1979]; Zhang and Rychlewski [1990]; Svendsen [1994]; Holthusen et al. [2020]. For paper and paperboard, we define the machine direction (*MD*), cross machine direction (*CD*) and thickness direction (*ZD*). Their orientation with respect to a paper roll is given in Fig. 3.1.

We introduce director vectors of unit length $\mathbf{n}_i$ where the index $i = 1, 2, 3$ is related to the three material directions *MD*, *CD* and *ZD*, respectively. The out-of-plane vector, $\mathbf{n}_3$, is defined

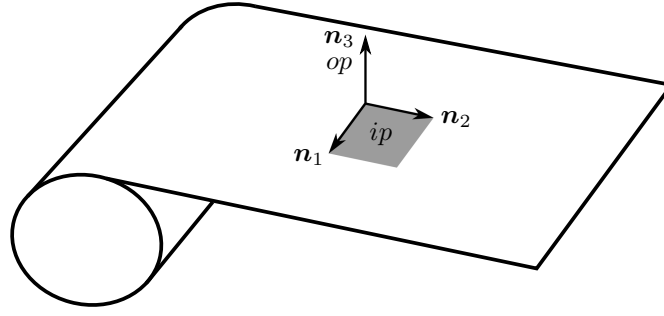


Figure 3.1.: Illustration of preferential material directions – in-plane area (*ip*) underlined in grey and out-of-plane direction (*op*).

orthogonal to the in-plane ones yielding to the relations $\mathbf{n}_3 = \mathbf{n}_1 \times \mathbf{n}_2$ and $\mathbf{n}_3 \cdot \mathbf{n}_2 = \mathbf{n}_3 \cdot \mathbf{n}_1 = 0$. From this, we define structural tensors by dyadic products of the corresponding director vectors, i.e. $\mathbf{M}_i = \mathbf{n}_i \otimes \mathbf{n}_i$. As we want to decompose the material answer into out-of-plane and in-plane behavior, we further introduce an in-plane structural tensor by $\mathbf{\Omega} = \mathbf{I} - \mathbf{M}_3$ which can be assumed to include all material directions except the out-of-plane one. Thus, we have the two structural tensors $\mathbf{\Omega}$ and $\mathbf{M}_3$ to define the material behavior separately.

3.3.2. Modified invariants

The Helmholtz free energy is commonly described in terms of invariants of strain measures, more precisely, in terms of the right Cauchy-Green deformation tensor $\mathbf{C}$. This motivates the introduction of modified invariants to decompose the energy formulation into in-plane and out-of-plane terms in a simple manner. Thus, we apply the in-plane structural tensor, $\mathbf{\Omega}$, onto the right Cauchy-Green tensor, $\mathbf{C}$, and formulate modified invariants by

$$I'_1 = \text{tr}(\mathbf{C}\mathbf{\Omega}) = \lambda_1 + \lambda_2, \quad I'_2 = \frac{1}{2} (\text{tr}(\mathbf{C}\mathbf{\Omega})^2 - \text{tr}[(\mathbf{C}\mathbf{\Omega})^2]) = \lambda_1 \lambda_2, \quad I'_3 = \det(\mathbf{C}\mathbf{\Omega}) = 0 \quad (3.1)$$

λ_i are the eigenvalues of $\mathbf{C}$. With this, the out-of-plane part is reduced in the formulation and only in-plane terms are described. The in-plane energy can be defined analogously to classical material formulations with these modified invariants which makes the physical interpretation easy to follow.

3.3.3. Isoplanar-area-changing split

An additional approach that we adapt is the classical isochoric-volumetric split. We introduce an isoplanar-area-changing split where we further decompose the deformations into area-

preserving and area-changing terms.

In the classical approach, the third invariant, $J := I_3 = \det(\mathbf{C})$, is used for the decomposition into volume-preserving and volume-changing deformations as it defines the squared volume ratio. Analogously, the second modified invariant I'_2 is used in our approach as it is the squared area ratio which can be shown by Nansons formula. It is further denoted by $J_A := I'_2$ for comparability to the squared volume ratio J used in the isochoric-volumetric split. The decomposition into isoplanar and area-changing deformations is defined by

$$\mathbf{C} = \underbrace{J_A^{-1/2} \mathbf{C}}_{=: \mathbf{C}_{\text{iso}_A}} \underbrace{J_A^{1/2} \mathbf{I}}_{=: \mathbf{C}_A}. \quad (3.2)$$

where $\mathbf{C}_{\text{iso}_A}$ refers to the isoplanar term and $\mathbf{C}_A$ to the area-changing one. The area-preserving characteristic can be shown by inserting $\mathbf{C}_{\text{iso}_A}$ into the second modified invariant J_A instead of the total right Cauchy-Green tensor which yields to

$$J_A(\mathbf{C}_{\text{iso}_A}) = \frac{1}{2} (\text{tr}(\mathbf{C}_{\text{iso}_A} \boldsymbol{\Omega})^2 - \text{tr}[(\mathbf{C}_{\text{iso}_A} \boldsymbol{\Omega})^2]) = \frac{1}{2} J_A^{-1} (\text{tr}(\mathbf{C} \boldsymbol{\Omega})^2 - \text{tr}[(\mathbf{C} \boldsymbol{\Omega})^2]) \equiv 1 \quad (3.3)$$

Thus, we have decomposed the kinematics into in-plane and out-of-plane deformations by use of structural tensors and further split into isoplanar and area-changing terms.

3.4. Inelastic material formulation

A multiplicative decomposition of the deformation gradient of $\mathbf{F} = \mathbf{F}_e \mathbf{F}_i$ into inelastic terms, $\mathbf{F}_i$, and elastic ones, $\mathbf{F}_e$, is introduced. This results in an additional configuration which is named the intermediate configuration. Inelastic material models are commonly derived in this configuration but pull-back operations are required for the implementation of these models as the intermediate configuration is not uniquely defined. This non-uniqueness motivates the introduction of a co-rotated intermediate configuration which results from the multiplicative decomposition of $\mathbf{F}_i = \mathbf{R}_i \mathbf{U}_i$ into its rotational part $\mathbf{R}_i$ and the stretch tensor $\mathbf{U}_i$. The co-rotated intermediate configuration is defined by a pull-back from the intermediate configuration by $\mathbf{R}_i$. With this, we can define co-rotated elastic right Cauchy-Green tensor

$$\bar{\mathbf{C}}_e := \mathbf{R}_i^{-1} \mathbf{C}_e \mathbf{R}_i = \mathbf{U}_i^{-1} \mathbf{C} \mathbf{U}_i^{-1}. \quad (3.4)$$

The overbar ($\bar{\bullet}$) indicates a quantity defined in the co-rotated intermediate configuration. The eigenvalues and symmetry properties between both intermediate configurations remain equal

but the co-rotated one is not affected by the rotational non-uniqueness. Thus, all quantities in the co-rotated intermediate configuration are uniquely defined and can be implemented directly using automated differentiation tools such as *Ace Gen* [Korelc, 2002; Korelc and Wriggers, 2016]. For a more detailed discussion on this topic, the reader is referred to Holthausen et al. [2023].

The two in-plane director vectors are assumed to evolve with the deformation gradient, and thus, are defined in the intermediate configuration by $\tilde{\mathbf{n}}_1 = (\mathbf{F}_i \mathbf{n}_1) / (|\mathbf{F}_i \mathbf{n}_1|)$ and $\tilde{\mathbf{n}}_2 = (\mathbf{F}_i \mathbf{n}_2) / (|\mathbf{F}_i \mathbf{n}_2|)$ (c.f. Reese [2003]). These definitions and a pull-back operation to the co-rotated intermediate configuration lead to the description of the out-of-plane structural tensor in this configuration as

$$\bar{\mathbf{M}}_3 = \frac{1}{\mathbf{C}_i^{-1} : \mathbf{M}_3} \mathbf{U}_i^{-1} \mathbf{M}_3 \mathbf{U}_i^{-1}. \quad (3.5)$$

We define a Helmholtz free energy in terms of the elastic right Cauchy-Green tensor and the out-of-plane structural tensor in the co-rotated intermediate configuration, i.e. $\psi = \psi^*(\bar{\mathbf{C}}_e, \bar{\mathbf{M}}_3)$, and insert its time derivative into the Clausius-Duhem inequality, $-\dot{\psi} + \mathbf{S} : 1/2 \dot{\bar{\mathbf{C}}} \geq 0$. Applying the procedure of Coleman and Noll (cf. Coleman and Noll [1961]) leads to the definition of the second Piola-Kirchhoff stress tensor, $\mathbf{S}$, and the reduced dissipation inequality, $\mathcal{D}_{\text{red}}$, i.e.

$$\mathbf{S} := 2 \mathbf{U}_i^{-1} \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} \mathbf{U}_i^{-1}, \quad \mathcal{D}_{\text{red}} := \underbrace{\left(2 \bar{\mathbf{C}}_e \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} + 2 \bar{\mathbf{M}}_3 \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} - 2 \frac{\partial \psi}{\partial \bar{\mathbf{M}}_3} : (\bar{\mathbf{M}}_3 \otimes \bar{\mathbf{M}}_3) \right)}_{=: \bar{\mathbf{\Gamma}}} : \bar{\mathbf{L}}_i \geq 0 \quad (3.6)$$

where $\bar{\mathbf{L}}_i := \dot{\mathbf{U}}_i \mathbf{U}_i^{-1}$ holds and $\bar{\mathbf{\Gamma}}$ is a stress-like quantity which is symmetric (c.f. Svendsen [1994]; Reese [2003]; Holthausen et al. [2022]). Hence, by defining $\bar{\mathbf{D}}_i := \text{sym}(\bar{\mathbf{L}}_i)$, the reduced dissipation inequality can be reformulated to $\mathcal{D}_{\text{red}} = \bar{\mathbf{\Gamma}} : \bar{\mathbf{D}}_i \geq 0$ which must be fulfilled for thermodynamical consistency. Therefore, we introduce a potential g as a positive, convex, zero-valued and scalar-valued isotropic function of the stress-like quantity $\bar{\mathbf{\Gamma}}$ and define the evolution equation of the inelastic strain by

$$\bar{\mathbf{D}}_i = \frac{1}{\tau} \frac{\partial \bar{g}}{\partial \bar{\mathbf{\Gamma}}} \quad \longrightarrow \quad \dot{\bar{\mathbf{C}}}_i = 2 \mathbf{U}_i \bar{\mathbf{D}}_i \mathbf{U}_i = \frac{2}{\tau} \mathbf{U}_i \frac{\partial g}{\partial \bar{\mathbf{\Gamma}}} \mathbf{U}_i \quad (3.7)$$

where τ is a relaxation time.

3.4.1. Choice of potentials

We introduce several potentials g to capture different effects and introduce them in the current configuration for physical interpretation. Thus, they are described in terms of the Kirchhoff stress tensor $\boldsymbol{\tau}$ and the out-of-plane structural tensor in the current configuration $\hat{\mathbf{M}}_3$. In Boes et al. [2023], it is shown that $\text{tr}(\boldsymbol{\tau}) = \text{tr}(\bar{\boldsymbol{\Gamma}})$ and $\boldsymbol{\tau} : \hat{\mathbf{M}}_3 = (1/\beta) \bar{\boldsymbol{\Gamma}} : \text{sym}(\mathbf{C}_e^{-1} \bar{\mathbf{M}}_3)$ holds. By use of these relations, the potentials introduced in the current configuration, g_i , are then reformulated in terms of the co-rotated intermediate configuration, $\bar{g}_i$.

The first two potentials are in alignment with Boes et al. [2023] where the first one describes inelastic strain evolution in in-plane directions when in-plane loadings are applied. It reads

$$g_1 = \frac{1}{2 K_2} \left(\text{tr}(\boldsymbol{\tau}) - \text{tr}(\boldsymbol{\tau} \hat{\mathbf{M}}_3) \right)^2, \quad (3.8)$$

$$\bar{g}_1 = \frac{1}{2 K_2} \left(\text{tr}(\bar{\boldsymbol{\Gamma}}) - \frac{1}{\text{tr}(\bar{\mathbf{C}}_e^{-1} \bar{\mathbf{M}}_3)} \text{tr}(\bar{\boldsymbol{\Gamma}} \bar{\mathbf{C}}_e^{-1} \bar{\mathbf{M}}_3) \right)^2. \quad (3.9)$$

The second potential is defined analogously but with an additional separate out-of-plane term such that inelastic strains might also evolve in the out-of-plane direction but remain decoupled from the in-plane behavior. Thus, we define

$$g_2 = \frac{1}{2 K_2} \left(\text{tr}(\boldsymbol{\tau}) - \text{tr}[\boldsymbol{\tau} \hat{\mathbf{M}}_3] \right)^2 + \frac{1}{2 K_3} \text{tr}(\boldsymbol{\tau} \hat{\mathbf{M}}_3)^2, \quad (3.10)$$

$$\begin{aligned} \bar{g}_2 = & \frac{1}{2 K_2} \left(\text{tr}(\bar{\boldsymbol{\Gamma}}) - \frac{1}{\text{tr}(\bar{\mathbf{C}}_e^{-1} \bar{\mathbf{M}}_3)} \text{tr}(\bar{\boldsymbol{\Gamma}} \bar{\mathbf{C}}_e^{-1} \bar{\mathbf{M}}_3) \right)^2 \\ & + \frac{1}{2 K_3} \left(\frac{1}{\text{tr}(\bar{\mathbf{C}}_e^{-1} \bar{\mathbf{M}}_3)} \text{tr}(\bar{\boldsymbol{\Gamma}} \bar{\mathbf{C}}_e^{-1} \bar{\mathbf{M}}_3) \right)^2. \end{aligned} \quad (3.11)$$

In this work, we investigate the effect of a third potential which is only defined for out-of-plane inelastic strain evolution. The densification and internal friction effect that are most dominant in the out-of-plane direction (c.f. Stenberg et al. [2001b,a]; Nygards et al. [2007]; Nygård [2008]) motivate the introduction of an inelastic potential defined only for an out-of-plane evolution. Thus, it is defined by

$$g_3 = \frac{1}{2 K_3} \text{tr}(\boldsymbol{\tau} \hat{\mathbf{M}}_3)^2, \quad \bar{g}_3 = \frac{1}{2 K_3} \left(\frac{1}{\text{tr}(\bar{\mathbf{C}}_e^{-1} \bar{\mathbf{M}}_3)} \text{tr}(\bar{\boldsymbol{\Gamma}} \bar{\mathbf{C}}_e^{-1} \bar{\mathbf{M}}_3) \right)^2. \quad (3.12)$$

3.4.2. Specific choice of the Helmholtz free energy

The Helmholtz free energy is additively split into an elastic and inelastic term, ψ^e and ψ^i , which both are further decomposed into isoplanar in-plane, area-preserving in-plane and out-of-plane parts denoted by ψ_{isoA} , ψ_A and ψ_{op} , respectively. This yields to

$$\psi = \psi^e + \psi^i, \quad \psi^e = \psi_{\text{isoA}}^e + \psi_A^e + \psi_{\text{op}}^e, \quad \psi^i = \psi_{\text{isoA}}^i + \psi_A^i + \psi_{\text{op}}^i. \quad (3.13)$$

The elastic energy is defined by

$$\psi_{\text{isoA}}^e = \frac{K_1}{2} (\text{tr}(\mathbf{C}_{\text{isoA}} \mathbf{\Omega}) - 2), \quad \psi_A^e = \frac{K_2}{4} (J_A - 1 - \ln(J_A)) \quad (3.14)$$

$$\psi_{\text{op}}^e = \frac{K_3}{4} (\text{tr}[\mathbf{C} \mathbf{M}_3] - 1 - \ln[\text{tr}(\mathbf{C} \mathbf{M}_3)]) + \frac{K_4}{2} (\text{tr}(\mathbf{C}^2 \mathbf{M}_3) - \text{tr}(\mathbf{C} \mathbf{M}_3)^2) \quad (3.15)$$

whereas the inelastic one reads

$$\psi_{\text{isoA}}^i = \frac{K_1}{2} (\text{tr}(\bar{\mathbf{C}}_{\text{isoA}}^e \bar{\mathbf{\Omega}}) - 2), \quad \psi_A^i = \frac{K_2}{4} (\bar{J}_A - 1 - \ln(\bar{J}_A)) \quad (3.16)$$

$$\psi_{\text{op}}^i = \frac{K_3}{4} (\text{tr}(\bar{\mathbf{C}}_e \bar{\mathbf{M}}_3) - 1 - \ln[\text{tr}(\bar{\mathbf{C}}_e \bar{\mathbf{M}}_3)]) + \frac{K_4}{2} (\text{tr}(\bar{\mathbf{C}}_e^2 \bar{\mathbf{M}}_3) - \text{tr}(\bar{\mathbf{C}}_e \bar{\mathbf{M}}_3)^2). \quad (3.17)$$

The relations $\bar{\mathbf{C}}_e = \mathbf{U}_i^{-1} \mathbf{C} \mathbf{U}_i^{-1}$, $\bar{J}_A = 1/2 (\text{tr}(\bar{\mathbf{C}}_e \bar{\mathbf{\Omega}})^2 - \text{tr}[(\bar{\mathbf{C}}_e \bar{\mathbf{\Omega}})^2])$, $\bar{\mathbf{C}}_{\text{isoA}}^e = \bar{J}_A^{-1/2} \bar{\mathbf{C}}_e$, and $\bar{\mathbf{\Omega}} = \mathbf{I} - \bar{\mathbf{M}}_3$ have been used here.

3.5. Numerical Examples

In the following, the influence of the third potential is compared to the first and second one. For this, the numerical example of a pinched cylinder was simulated as already given in Boes et al. [2023] for the first two potentials.

The geometry and boundary conditions are given in Fig. 3.2. Only one eighth of the cylinder was simulated due to symmetry effects. The out-of-plane direction was assumed pointing outwards – normal to the cylinder surface – and the specimen was fixed in x_1 and x_2 direction at $x_3 = 300$ [mm]. A displacement $w = 100$ [mm] was applied over 100 seconds at the node at $x_2 = 300$ [mm] and $x_1 = x_3 = 0$ [mm]. After a convergence study has been performed, a discretization into $2 \times 32 \times 64$ elements was chosen. A single Gaussian point element formulation for continuum elements (Q1STc) proposed by Barfusz et al. [2021] was used. The material parameters were $K_1 = 80 \left[\frac{\text{N}}{\text{mm}^2} \right]$, $K_2 = 400 \left[\frac{\text{N}}{\text{mm}^2} \right]$, $K_3 = 400 \left[\frac{\text{N}}{\text{mm}^2} \right]$, $K_4 = 10 \left[\frac{\text{N}}{\text{mm}^2} \right]$, and $\tau = 10 \left[\frac{1}{\text{s}} \right]$.

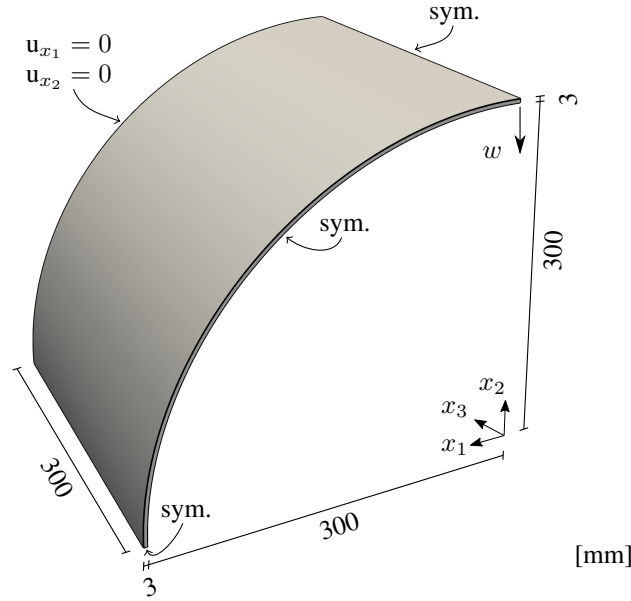


Figure 3.2.: Geometry and boundary conditions of the pinched cylinder.

The aim of these simulations was to investigate the inelastic strain evolution with the three different potentials. Figure 3.3 shows the in-plane component of the inelastic right Cauchy-Green deformation tensor, C_{ip} , computed with potential 1 to 3 from left to right. It can be seen that the first two potentials lead to an inelastic strain evolution due to the loading. No inelastic strains in the in-plane directions can be observed for the third potential which has been newly introduced here and is defined for out-of-plane inelastic strain evolution only.

Figure 3.4 illustrates the out-of-plane inelastic right Cauchy-Green deformation C_{iop} . The results on the left that show calculations with the first potential show no evolution as the potential only defines in-plane inelastic strain evolution. The out-of-plane part of the second and third potential (Fig. 3.4 middle and right) are defined in a similar manner (c.f. Eq. (3.11) and (3.12)), leading to similar localized inelastic strain evolutions. Thus, the introduction of the third potential shows the possibility of defining out-of-plane inelastic strain evolution separately from the in-plane one.

3.6. Conclusion

A model for decoupled material behavior in thickness (out-of-plane) and in-plane directions has been presented that is motivated by the macroscopic material behavior of paper and paperboard. The decoupling has been introduced by a decomposition of the kinematics and the Helmholtz free energy into in-plane and out-of-plane parts. The kinematics, precisely the

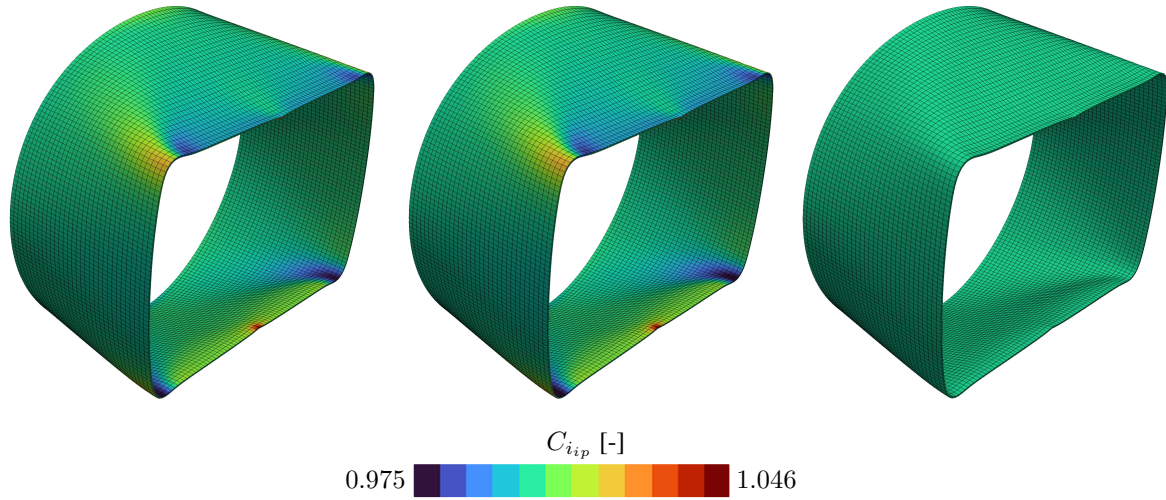


Figure 3.3.: In-plane component of the inelastic right Cauchy-Green deformation tensor, C_{ip} , computed with Potential 1 (left), Potential 2 (middle) and Potential 3 (right).

right Cauchy-Green deformation tensor, is further decomposed into area-preserving and area-changing deformations leading to the isoplanar-area-changing split which is in alignment with the classical isochoric-volumetric split. Including the decoupling only by decomposing the right Cauchy-Green deformation tensor and the Helmholtz free energy leads to the flexibility of this model for adjustments to different material responses, and thus, to extending it to inelastic material behavior which is a major advantage. To prove this flexibility, a visco-elastic material formulation has been introduced in terms of a co-rotated intermediate configuration. The latter one has been used as it is uniquely defined in contrast to the intermediate configuration, and therefore, quantities defined within this configuration can be implemented directly using automated differentiation tools.

The model presented is in alignment with the one depicted in Boes et al. [2023]. In this work, we present an additional inelastic potential that defines inelastic strain evolution only in the out-of-plane direction when out-of-plane loading is applied. The influence of the potential is shown by simulations of a pinched cylinder that have been presented in Boes et al. [2023] for the first two potentials and are compared to the newly introduced potential. It is shown that by the specific choice of the potential different material behaviors can be obtained. The introduction of the third potential given here indicates inelastic material behavior for the out-of-plane direction only which is an advantage for adjusting the material model more precisely to the material behavior of paper and paperboard in the following since the out-of-plane response shows inelastic behavior and is independent of the in-plane one.

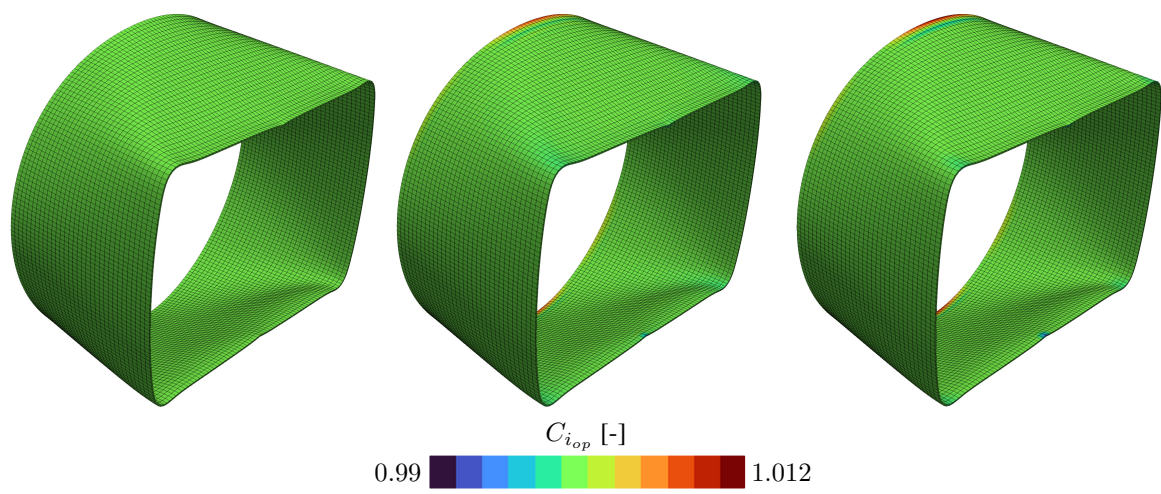


Figure 3.4.: Out-of-plane component of the inelastic right Cauchy-Green deformation tensor, C_{iop} , computed with Potential 1 (left), Potential 2 (middle) and Potential 3 (right).

4 | **Article 3:**

Accounting for plasticity: An extension of inelastic constitutive artificial neural networks

This article was published as:

Boes, B., Simon, J.-W., and Holthusen, H. (2026). Accounting for plasticity: An extension of inelastic constitutive artificial neural networks. *European Journal of Mechanics - A/Solids*, **117**, 105998. <https://doi.org/10.1016/j.euromechsol.2025.105998>.

Disclosure of the individual authors' contributions to the article:

B. Boes reviewed the relevant existing literature, collected and created the data, implemented advancements to the model presented by Holthusen et al. [2024b] and trained the model. She interpreted the results and wrote the preliminary version of the article. H. Holthusen gave valuable conceptual advice regarding implementation and theoretical aspects. All co-authors gave valuable conceptual advice, read the manuscript and gave suggestions for improvement. All authors approved the publication of the final version of the manuscript.

4.1. Abstract

In this work, we extend the framework of inelastic constitutive artificial neural networks (iCANNs) by incorporating plasticity. Thereby, we increase their applicability to capture more complex material behavior by solving the underlying inequality constraints associated with plasticity, aiming to automatically discover the formulation that matches the experimental data most accurately. Learning on stress-strain data, our approach enables the automatic discovery of the constitutive equations for elasto-plastic materials – including the identification of the material parameters and the underlying functional forms of the Helmholtz free energy, yield function and evolution equations for the inelastic deformations including linear and nonlinear kinematic hardening. These are captured using feed-forward neural networks. The proposed methodology ensures objectivity, material symmetry, and thermodynamic consistency, providing a robust and interpretable basis for automatic model discovery at finite strains. Investigation of our model's prediction showed that the extended iCANNs successfully predict both linear and nonlinear kinematic hardening behavior based on experimental and artificially generated datasets, showcasing promising capabilities of this framework. Nonetheless, challenges remain in fully capturing complex yield criteria – such as those exhibiting tension-compression asymmetry – and predicting complex nonlinear hardening.

Notations

a	Scalar quantity
$\mathbf{a}$	First order tensor
$\mathbf{A}$	Second order tensor
$\dot{\mathbf{A}}$	Total derivative of $\mathbf{A}$ with respect to time
$\mathbf{A}^T$	Transpose of $\mathbf{A}$
$\mathbf{A}^{-1}$	Inverse of $\mathbf{A}$
$\text{cof}(\mathbf{A})$	Cofactor of $\mathbf{A}$
$\det(\mathbf{A})$	Determinant of $\mathbf{A}$
$\text{tr}(\mathbf{A})$	Trace of $\mathbf{A}$
$\ln(\mathbf{A})$	Logarithm of $\mathbf{A}$
$\text{sym}(\mathbf{A})$	Symmetric term of $\mathbf{A}$
$\mathbf{A} : \mathbf{B}$	Scalar product of two tensors $\mathbf{A}$ and $\mathbf{B}$
$I_1^{\mathbf{A}}$	$\text{tr}(\mathbf{A})$
$I_2^{\mathbf{A}}$	$\text{tr}(\text{cof}(\mathbf{A}))$
$I_3^{\mathbf{A}}$	$\det(\mathbf{A})$
$J_2^{\mathbf{A}}$	$\frac{1}{2}\text{tr}(\text{dev}(\mathbf{A})^2)$
$J_3^{\mathbf{A}}$	$\frac{1}{3}\text{tr}(\text{dev}(\mathbf{A})^3)$

4.2. Introduction

In the context of constitutive modeling, a vast variety of material models have been derived throughout the years. Identifying the optimal constitutive model and its associated parameters is essential for accurately describing material behavior and performing finite element simulations, making the selection of an appropriate model a challenging task. Traditionally, this process has relied on expert knowledge and reasoning, which inherently introduces user bias. Recent advances in machine learning, particularly neural network-based methods, have enhanced the process of model selection by discovering both model parameters and constitutive equations automatically from experimental data, revealing the formulation that most accurately matches the underlying data. This shift towards data-driven methods has the potential to significantly improve the efficiency and accuracy of material modeling, particularly in capturing complex behaviors. While the concept of neural networks in constitutive modeling dates back decades [Ghaboussi et al., 1991], the trend to automatically discover these equations has emerged significantly throughout the last years. A variety of different approaches have been proposed, spanning from data-driven methods in a model-free approach [Kirchdoerfer and Ortiz, 2016; Eggersmann et al., 2021], over physically driven approaches [Flaschel et al., 2021; Yoshida et al., 2002] to physics-informed [Haghighat et al., 2021; Rezaei et al., 2024; Yang et al., 2025] and physics-augmented [Linden et al., 2023, 2025; Kalina et al., 2024] neural networks, to name a few only. All of these approaches aim to accurately capture the complex behavior of materials under various loading conditions, including elastic and inelastic responses. For a comprehensive overview of these advancements, the reader is referred to the reviews by Fuhg et al. [2024a] and Watson et al. [2024].

Although the *universal approximation theorem* states that neural networks can approximate any function given a sufficiently wide or deep architecture [Hornik et al., 1990], increased model complexity often comes at the cost of interpretability. Therefore, automated model discovery involves balancing between accuracy and transparency. Among all approaches, the framework of constitutive artificial neural networks (CANNs), initially presented by Linka et al. [2021] and Linka and Kuhl [2023], has gained significant attention due to its interpretability. Learning from experimental stress-strain data, it is capable of automatically discovering the constitutive equations of elastic materials, including the functional form of the Helmholtz free energy, which is represented in terms of feed-forward neural networks (FFNs). A key advantage of this framework is its custom-designed architecture, which inherently satisfies the principles of objectivity, material symmetry, and thermodynamic consistency, thereby ensuring that the discovered models are physically sound. Although it is based on a sparse architecture, it conforms

to the definition of a neural network, where it is essential to balance between interpretability and accuracy. While one might argue that sparse architectures could lack sufficient approximation capabilities, applications in diverse fields such as Alzheimer’s disease research [Zhang et al., 2024; Stockman et al., 2024], human tissues, such as brain cortices [Linka et al., 2023a; Hou et al., 2024], cardiac tissues [Martonová et al., 2024; Vervenne et al., 2025a], human skin [Linka et al., 2023b], and atria [Peirlinck et al., 2025] of have demonstrated that CANNs achieve highly accurate approximations despite their reduced complexity. Additionally, its applications extend to the mechanics of artificial meat [Pierre et al., 2023; St. Pierre et al., 2024] and textile structures [McCulloch and Kuhl, 2024], and it has recently served as a foundation for the extension to Kolmogorov-Arnold networks by Abdolazizi et al. [2025]. The variety of these examples highlights its versatility and strength. Notably, recent work by Vervenne et al. [2025a] has demonstrated that automatically discovered models can surpass even established, sophisticated material models, highlighting that traditional models do not always yield the most accurate results.

While automated model discovery has been shown to be a powerful tool in the field of elasticity, extending these approaches to inelasticity is crucial to capture more complex material behavior, which is essential for engineering applications. Inelastic behavior, such as plasticity and visco-elasticity, significantly influences the performance and durability of materials under various loading conditions. Consequently, expanding in particular the CANNs framework to address inelasticity is of high interest. CANNs have recently been extended to visco-elasticity using a Prony series Abdolazizi et al. [2024]. Further, a general formulation for inelasticity through inelastic constitutive artificial neural networks (iCANNs) has been introduced Holthusen et al. [2024b], which allows for the automatic discovery of inelastic constitutive equations. By design of the architecture, the framework ensures not only objectivity and material symmetry, but also obtains thermodynamically consistent formulations by custom-designed architectures of the feed-forward networks, providing a robust basis for automated model discovery at finite strains. Further, the framework provides insights into the underlying mechanical behavior of the material. Recently, a comprehensive formulation capable of capturing a wider range of material behaviors through a generalized formulation of the potential was presented by Holthusen et al. [2025b]. The extension to iCANNs broadens the framework’s applicability to a wide range of material behaviors, already applied to, for example, visco-elastic materials [Holthusen et al., 2024b], and growth modeling [Holthusen et al., 2024a, 2025a], while the application to plasticity remains to be addressed.

Since plasticity significantly influences the durability and consequently the applicability of materials, expanding the use of neural networks in the context of predicting yield criteria

and evolution laws has been of high interest recently. For instance, neural networks have been employed to predict the yield criteria [Settgast et al., 2020; van der Velden et al., 2025], while a hybrid strategy to represent complex yield surfaces and non-associated flow directions via neural networks was proposed by Malik et al. [2021]. Data-driven plasticity formulations have been presented by augmenting phenomenological yield functions with a correction term [Fuhg et al., 2023a]. Ensuring thermodynamic consistency in modeling inelasticity has been approached in various ways. While one strategy is through modification of the loss function [Rosenkranz et al., 2023; Masi et al., 2021; Masi and Stefanou, 2022], convexity of the predicted networks is enforced, dealing with input convex neural networks [Benady et al., 2024; Flaschel et al., 2025b; Fuhg et al., 2023b] or constrained neural networks [Weber et al., 2023]. Further, elasto-plasticity has been included in recurrent neural networks [Mozaffar et al., 2019; Bonatti and Mohr, 2022; Heidenreich et al., 2024; Heidenreich and Mohr, 2024], though these often require substantial training data and can suffer from reduced interpretability. Anisotropic hardening behavior has been modelled using two ANNs to predict the hardening behavior and the anisotropic parameters by [Ebrahim et al., 2024]. Further, a generalized plasticity model for small strain theory embedded into ANNs was introduced Wang et al. [2021], and recent efforts have advanced the prediction of complex hardening behavior [Zhang and Mohr, 2020; Li et al., 2022; Flaschel et al., 2022] as well as non-associative plasticity [Xu et al., 2025]. Recently, work on benchmarks to evaluate neural networks capturing elasto-plastic behavior has been described [Lesueur et al., 2025]. Given the rapid growth of automated model discovery and the advancing emphasis on plasticity, extending iCANNs to incorporate plasticity is vital for their applicability and will be discussed in this work. A similar approach to the one presented here was recently introduced Jadoon et al. [2025], highlighting the relevance of this research direction. While their method leverages fully-connected neural networks to achieve high accuracy in the yield prediction, it lacks interpretability, which we address in this work using iCANNs.

In contrast to the general discussion on inelasticity of iCANNs, we include predictions of the yield surface and evolution equations of inelastic stretches in terms of learning formulations of plastic potentials, and discovering elastic and inelastic formulations of the Helmholtz free energy. This enables the simultaneous discovery of both elastic and inelastic contributions accounting for hardening effects. We further employ a return-mapping algorithm to address the underlying inequality constraints associated with plasticity.

Outline

We provide the theoretical derivation of the constitutive modeling of plasticity, including linear and nonlinear kinematic hardening in Section 4.3. It is derived in a co-rotated intermediate configuration, simplifying the implementation. The network formulation and architecture are discussed in Section 4.4, giving insights into the requirements of the architecture. We propose a formulation of the Helmholtz free energies to account for elastic material behavior and hardening effects, as well as plastic potentials describing the yield surface and evolution equations. The application of the neural network framework is presented in Section 4.5, which details the algorithmic implementation and learning strategy. This section demonstrates the framework's capabilities by predicting results for artificially generated data, including the von Mises and Tschoegel yield criteria, a comparison to results obtained by Liao and Luo [2025], and predicting experimental data for steel [Cao et al., 2009]. The outcomes, along with current limitations, are discussed together in Section 4.6, and the main conclusions are summarized in Section 4.7.

4.3. Constitutive Modeling

In the following, the constitutive framework is presented, which is the basis for our network architecture. The characteristic equations are derived following the principles of thermodynamics such that a thermodynamically consistent formulation is found.

4.3.1. Kinematics

We consider finite strains and apply the multiplicative decomposition of the deformation gradient, $\mathbf{F}$, into an elastic and plastic component according to $\mathbf{F} = \mathbf{F}_e \mathbf{F}_p$ [Lee and Liu, 1967; Lee, 1969]. To incorporate nonlinear kinematic hardening into the model formulation, we perform an additional multiplicative split of the plastic deformation gradient, i.e. $\mathbf{F}_p = \mathbf{F}_{p_e} \mathbf{F}_{p_i}$, where $\mathbf{F}_{p_e}$ can be referred to elastic deformations due to dislocations and $\mathbf{F}_{p_i}$ to local plastic deformations through slip systems. This additional multiplicative split introduces a physically motivated additional intermediate configuration of kinematic hardening [Lion, 2000].

To derive a thermodynamically consistent model formulation for inelastic materials, we formulate a general Helmholtz free energy, ψ , that must satisfy the requirements of objectivity, material symmetry, and rotational non-uniqueness. Therefore, we define it in terms of the elastic right Cauchy-Green tensor, $\mathbf{C}_e := \mathbf{F}_e^T \mathbf{F}_e$, and the inelastic left Cauchy-Green tensors, $\mathbf{B}_p := \mathbf{F}_p \mathbf{F}_p^T$ and $\mathbf{B}_{p_e} := \mathbf{F}_{p_e} \mathbf{F}_{p_e}^T$, which are defined in analogy to the total right and left

Cauchy-Green tensors, $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ and $\mathbf{B} = \mathbf{F} \mathbf{F}^T$, respectively. Thus, the dependency reads $\psi = \hat{\psi}(\mathbf{C}_e, \mathbf{B}_p, \mathbf{B}_{p_e})$. The multiplicative split is associated with a fictitious, unstressed intermediate configuration that suffers from a rotational non-uniqueness. Thus, neither stress- nor strain-like quantities in the intermediate configuration can be computed. Following the concept presented by Holthausen et al. [2023] and applied by Boes et al. [2023], a co-rotated intermediate configuration is introduced as a co-rotated pull-back from the intermediate configuration by $(\bar{\bullet}) = \mathbf{R}_p^{-1}(\bullet)\mathbf{R}_p$, where the rotation tensor $\mathbf{R}_p$ together with the stretch tensor $\mathbf{U}_p$ result from the polar decomposition of the plastic term of the deformation gradient, $\mathbf{F}_p = \mathbf{R}_p \mathbf{U}_p$. The eigenvalues and symmetry properties between the intermediate configuration and the co-rotated one remain the same, and the assumption of an unstressed intermediate configuration holds true here. Consequently, we can introduce the co-rotated Cauchy-Green tensors

$$\bar{\mathbf{C}}_e := \mathbf{R}_p^{-1} \mathbf{C}_e \mathbf{R}_p = \mathbf{U}_p^{-1} \mathbf{C} \mathbf{U}_p^{-1} \quad (4.1)$$

$$\bar{\mathbf{B}}_p := \mathbf{R}_p^{-1} \mathbf{B}_p \mathbf{R}_p \equiv \mathbf{C}_p \quad (4.2)$$

$$\bar{\mathbf{B}}_{p_e} := \mathbf{R}_p^{-1} \mathbf{B}_{p_e} \mathbf{R}_p = \mathbf{U}_p \mathbf{C}_{p_i}^{-1} \mathbf{U}_p = \mathbf{U}_p \mathbf{U}_{p_i}^{-2} \mathbf{U}_p, \quad (4.3)$$

where $\mathbf{C}_{p_i} = \mathbf{F}_{p_i}^T \mathbf{F}_{p_i}$ is the right Cauchy-Green tensor resulting from the additional multiplicative split of the plastic deformation gradient and $\mathbf{U}_{p_i}$ is the corresponding stretch tensor. The co-rotated tensors $\bar{\mathbf{C}}_e$, $\bar{\mathbf{B}}_p$, and $\bar{\mathbf{B}}_{p_e}$ are invariant under rigid body motions and can be used to describe the Helmholtz free energy as an isotropic function of these quantities, i.e. $\psi = \hat{\psi}(\bar{\mathbf{C}}_e, \mathbf{C}_p, \bar{\mathbf{B}}_{p_e})$. Thus, these quantities and the derivatives with respect to them can be defined and computed using automated differentiation.

4.3.2. Derivation based on the Clausius-Planck inequality

The constitutive equations of our model are derived by evaluating the Clausius-Planck inequality that reads $-\dot{\psi} + \mathbf{S} : \frac{1}{2} \dot{\mathbf{C}} \geq 0$, where $\mathbf{S}$ is the second Piola-Kirchhoff stress and $\dot{\mathbf{C}}$ the time derivative of the right Cauchy-Green tensor. Therefore, we insert the time derivative of the Helmholtz free energy

$$\dot{\psi} = \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} : \dot{\bar{\mathbf{C}}}_e + \frac{\partial \psi}{\partial \mathbf{C}_p} : \dot{\mathbf{C}}_p + \frac{\partial \psi}{\partial \bar{\mathbf{B}}_{p_e}} : \dot{\bar{\mathbf{B}}}_{p_e}, \quad (4.4)$$

with the rates

$$\dot{\bar{\mathbf{C}}}_e = \mathbf{U}_p^{-1} \dot{\mathbf{C}} \mathbf{U}_p^{-1} - \bar{\mathbf{C}}_e \bar{\mathbf{L}}_p - \bar{\mathbf{L}}_p^T \bar{\mathbf{C}}_e \quad (4.5)$$

$$\dot{\mathbf{C}}_p = \bar{\mathbf{L}}_p \mathbf{C}_p + \mathbf{C}_p \bar{\mathbf{L}}_p^T \quad (4.6)$$

$$\dot{\bar{\mathbf{B}}}_{p_e} = -\mathbf{U}_p \mathbf{C}_{p_i}^{-1} \dot{\mathbf{C}}_{p_i} \mathbf{C}_{p_i}^{-1} \mathbf{U}_p + \bar{\mathbf{L}}_p \bar{\mathbf{B}}_{p_e} + \bar{\mathbf{B}}_{p_e} \bar{\mathbf{L}}_p^T, \quad (4.7)$$

where $\bar{\mathbf{L}}_p := \dot{\mathbf{U}}_p \mathbf{U}_p^{-1}$ holds. Further, we introduce the relation $\bar{\mathbf{L}}_{p_i} := \dot{\mathbf{U}}_{p_i} \mathbf{U}_{p_i}^{-1}$ leading to the dissipation inequality of the form

$$\begin{aligned} \left(\mathbf{S} - 2 \mathbf{U}_p^{-1} \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} \mathbf{U}_p^{-1} \right) : \frac{1}{2} \dot{\mathbf{C}} + \underbrace{\left(\underbrace{2 \bar{\mathbf{C}}_e \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e}}_{=: \bar{\Sigma}} - 2 \underbrace{\frac{\partial \psi}{\partial \mathbf{C}_p} \mathbf{C}_p}_{=: \bar{\chi}} - 2 \underbrace{\frac{\partial \psi}{\partial \bar{\mathbf{B}}_{p_e}} \bar{\mathbf{B}}_{p_e}}_{=: \bar{\Xi}} \right)}_{=: \bar{\Gamma}} : \bar{\mathbf{L}}_p \\ + \underbrace{2 \mathbf{U}_{p_i}^{-1} \mathbf{U}_p \frac{\partial \psi}{\partial \bar{\mathbf{B}}_{p_e}} \mathbf{U}_p \mathbf{U}_{p_i}^{-1}}_{=: \bar{\Theta}} : \bar{\mathbf{L}}_{p_i} \geq 0. \end{aligned} \quad (4.8)$$

Applying the procedure by Coleman and Noll [1961], we obtain the second Piola-Kirchhoff stress, $\mathbf{S}$, and define the state relations of the relative stress, $\bar{\Gamma} := \bar{\Sigma} - \bar{\chi} - \bar{\Xi}$, defined by the Mandel stress, $\bar{\Sigma}$, the backstresses, $\bar{\chi}$ and $\bar{\Xi}$, and the Mandel stress corresponding to kinematic hardening, $\bar{\Theta}$. The backstresses are related to kinematic hardening and describe the translation of the origin of the yield surface during plastic deformations. While neither the backstresses, $\bar{\chi}$ and $\bar{\Xi}$, nor the Mandel stress related to linear kinematic hardening, $\bar{\Sigma}$, are generally symmetric, the stress quantities $\bar{\Gamma}$ and $\bar{\Theta}$ are symmetric [Svendsen, 2001]. Consequently, we introduce the relations $\bar{\mathbf{D}}_p := \text{sym}(\bar{\mathbf{L}}_p)$ and $\bar{\mathbf{D}}_{p_i} := \text{sym}(\bar{\mathbf{L}}_{p_i})$, and obtain the reduced dissipation inequality, i.e.

$$\mathcal{D}_{\text{red}} := \bar{\Gamma} : \bar{\mathbf{D}}_p + \bar{\Theta} : \bar{\mathbf{D}}_{p_i} \geq 0. \quad (4.9)$$

For more detailed derivations, the interested reader is referred to the works of Vladimirov et al. [2008]; Brepols et al. [2020]; Felder et al. [2020]; Holthausen et al. [2023]. The reduced dissipation inequality must be fulfilled for arbitrary thermodynamical processes. Therefore, we introduce evolution equations for the strain-like quantities in terms of general plastic potentials, g_1 and g_2 . We describe them by scalar-valued isotropic functions of the thermodynamic driving

forces, i.e. $g_1 = \hat{g}_1(\bar{\Gamma})$ and $g_2 = \hat{g}_2(\bar{\Theta})$, such that

$$\bar{D}_p = \lambda \frac{\partial g_1}{\partial \bar{\Gamma}} \rightarrow \dot{C}_p = 2\lambda U_p \frac{\partial g_1}{\partial \bar{\Gamma}} U_p \quad (4.10)$$

$$\bar{D}_{p_i} = \lambda \frac{\partial g_2}{\partial \bar{\Theta}} \rightarrow \dot{C}_{p_i} = 2\lambda U_{p_i} \frac{\partial g_2}{\partial \bar{\Theta}} U_{p_i}. \quad (4.11)$$

The latter one can be used to account for nonlinear kinematic hardening of Frederick-Armstrong type [Armstrong et al., 1966]. Further, we restrict the potentials to be convex, zero-valued, and non-negative, thereby automatically fulfilling Inequality (4.9) (see Holthusen et al. [2025a]). Regarding the wide class of inelasticity, λ can be interpreted, for example, as a relaxation time in visco-elasticity or a growth multiplier for growth modeling (c.f. Lamm et al. [2021]; Lamm and Holthusen [2022]). We interpret it as a plastic multiplier, considering inelastic effects of plasticity. The plastic multiplier is (by definition) non-negative and is defined through the Karush-Kuhn-Tucker conditions, which read

$$\lambda \geq 0, \quad \Phi \leq 0, \quad \lambda \Phi = 0. \quad (4.12)$$

They complete the set of constitutive equations and introduce the yield function Φ , which encloses the set of all admissible stress states. It is zero on the yield surface denoting the plastic regime and negative in the elastic regime. All quantities are listed in Table 4.1 to give a brief overview of the constitutive model equations. A rheological representation of this model is given in Figure 4.1 for visualization purposes, where the Helmholtz free energy, ψ , is represented by three springs – one accounting for the elastic material response ψ_e , and the remaining two for the plastic terms, ψ_p , and ψ_{p_e} . The latter ones describe linear kinematic hardening, detailed by a linear spring, and nonlinear kinematic hardening, indicated by a serial connection of a dashpot element governed by the plastic potential, g_2 , and a spring. The plastic flow is represented through a slider being inactive for $\Phi < 0$ and evolving through g_1 . It is connected in parallel to the elements accounting for linear and nonlinear kinematic hardening.

4.4. Network Formulation and Architecture

So far, the constitutive equations have been defined in a general manner. It remains to find specific formulations for the Helmholtz free energy, ψ , the yield function, Φ , and the potentials, g_1 and g_2 . The total Helmholtz free energy is defined by $\psi := \psi_e + \psi_p + \psi_{p_e}$. In classical constitutive material modeling, these are often predefined in a specific manner – e.g. in terms

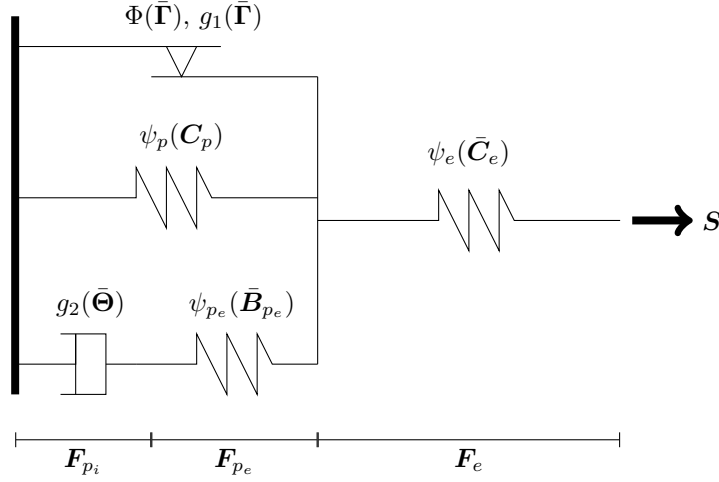


Figure 4.1.: Rheological representation of the iCANN capturing plasticity consisting out of three springs, ψ_e , ψ_p , and ψ_{pe} . A slider, Φ and g_1 , and a dashpot element, g_2 , with the states $\bar{U}_p$ and $\bar{U}_{pi}$ that are captured through the strain tensors $\bar{C}_p$ and $\bar{B}_{pe}$, are connected in parallel.

of a Neo-Hookean type energy or von Mises yield surface. In contrast, our approach replaces the predefined formulations by feed-forward neural networks that are trained on experimental data to discover the best matching model. However, the formulations of the FFNs are designed to follow physical requirements such as objectivity, material symmetry, and thermodynamic consistency, and must follow specific requirements to ensure that the learned constitutive relations remain physically sound and interpretable. The overall framework is built in analogy to the work presented by Holthusen et al. [2024b]. In the following, we describe the design of the FFNs, outline the key aspects of our architecture and algorithmic treatment, and highlight the novel extensions related to plasticity.

4.4.1. Algorithmic treatment

An explicit time integration scheme is applied to numerically evaluate the evolution equations (4.10) and (4.11). They are discretized within a time interval $t \in [t_n, t_{n+1}]$, where t_n denotes the last and t_{n+1} the current time step. Notably, the index $n+1$ is omitted in the following for readability. An exponential integration scheme is applied, such that Equations (4.10)

Table 4.1.: Summary of all constitutive relations.

Constitutive equations	
Stresses	
Second Piola-Kirchhoff stress	$\mathbf{S} = 2 \mathbf{U}_p^{-1} \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e} \mathbf{U}_p^{-1}$
Mandel stress	$\bar{\mathbf{\Sigma}} = 2 \bar{\mathbf{C}}_e \frac{\partial \psi}{\partial \bar{\mathbf{C}}_e}$
Linear backstress	$\bar{\mathbf{\chi}} = 2 \frac{\partial \psi}{\partial \bar{\mathbf{C}}_p} \bar{\mathbf{C}}_p$
Nonlinear backstress	$\bar{\mathbf{\Xi}} = 2 \frac{\partial \psi}{\partial \bar{\mathbf{B}}_{p_e}} \bar{\mathbf{B}}_{p_e}$
Relative stress	$\bar{\mathbf{\Gamma}} = \bar{\mathbf{\Sigma}} - \bar{\mathbf{\chi}} - \bar{\mathbf{\Xi}}$
Mandel stress related to kinematic hardening	$\bar{\mathbf{\Theta}} = 2 \mathbf{U}_{p_i}^{-1} \mathbf{U}_p \frac{\partial \psi}{\partial \bar{\mathbf{B}}_{p_e}} \mathbf{U}_p \mathbf{U}_{p_i}^{-1}$
Evolution Equations	
	$\dot{\mathbf{C}}_p = 2 \lambda \mathbf{U}_p \frac{\partial g_1}{\partial \bar{\mathbf{\Gamma}}} \mathbf{U}_p$
	$\dot{\mathbf{C}}_{p_i} = 2 \lambda \mathbf{U}_{p_i} \frac{\partial g_2}{\partial \bar{\mathbf{\Theta}}} \mathbf{U}_{p_i}$
Yield criterion	
	$\Phi(\bar{\mathbf{\Gamma}}) \leq 0$

and (4.11) yield

$$\mathbf{C}_p = \mathbf{U}_{p_n} \exp(2 \Delta \lambda \bar{\mathbf{D}}_p) \mathbf{U}_{p_n}, \quad \mathbf{U}_p = +\sqrt{\mathbf{C}_p} \quad (4.13)$$

$$\mathbf{C}_{p_i} = \mathbf{U}_{p_{i_n}} \exp(2 \Delta \lambda \bar{\mathbf{D}}_{p_i}) \mathbf{U}_{p_{i_n}}, \quad \mathbf{U}_{p_i} = +\sqrt{\mathbf{C}_{p_i}}, \quad (4.14)$$

with $\Delta \lambda := \Delta t \lambda$ and $\Delta t = t - t_n$. This formulation avoids artificial interaction between shear and volumetric stresses.

Unlike CANNs or iCANNs used for visco-elasticity, we must fulfill the Karush-Kuhn-Tucker conditions in Equation (4.12). Therefore, we apply an elastic predictor-plastic corrector scheme by using a Newton-Raphson iteration. To this end, we introduce a residual $r = \Phi^{\text{trial}} \stackrel{!}{=} 0$, which is solved if a trial step yields $\Phi^{\text{trial}} > 0$ and then minimized iteratively until it converges to zero by updating the plastic multiplier of the current iteration step, k , to the next step, $k + 1$, by

$$\lambda^{(k+1)} = \lambda^{(k)} - \left(\frac{\partial r^{(k)}}{\partial \lambda^{(k)}} \right)^{-1} r^{(k)}, \quad (4.15)$$

until convergence is achieved, ensuring that the current stress state lies within the set of admissible stresses that are defined by the yield surface.

4.4.2. Recurrent neural network architecture

In total, the architecture comprises the following six feed-forward neural networks:

- ψ_e – elastic Helmholtz free energy
- ψ_p – Helmholtz free energy to account for linear kinematic hardening
- ψ_{pe} – Helmholtz free energy to account for nonlinear kinematic hardening
- Φ – yield surface
- g_1 – plastic potential defining the evolution of C_p
- g_2 – plastic potential defining the evolution of $\bar{B}_{pe}$

All feed-forward neural networks are embedded in a recurrent neural network, which is evaluated at each time step. The structure of the recurrent neural network is visualized as a flowchart in Figure 4.2. We aim to learn the constitutive equations of the material from pairs of inputs, C , and outputs, σ , where the Cauchy stress, σ , is computed from the second Piola-Kirchhoff stress by

$$\sigma = \det(F)^{-1} F S F^T. \quad (4.16)$$

The latter ones are used as loss measure where we compute the mean squared error between the experimentally observed and predicted stresses. All history variables from the last time step, U_{p_n} , $U_{pp_{t_n}}$, and C_n , are passed through the hidden states. As visualized in Figure 4.2, first, we compute all kinematic, stress, and stress-like quantities from the last time step t_n (upper gray box) by evaluating the Helmholtz free energies. These are described by FFNs, which comprise the same architecture, visualized by orange boxes. Next, we evaluate the plastic potentials, $g_{1_n} := g_1(t = t_n)$ and $g_{2_n} := g_2(t = t_n)$, to update the explicit evolution equations. These are again formulated by FFNs and colored by blue boxes to distinguish between the different architectures of the FFNs for the Helmholtz free energy and the potentials. We then pass the current right Cauchy-Green tensor, compute all kinematic quantities from the current time step (lower gray box), t_{n+1} , and determine the current stress and stress-like quantities by evaluating the Helmholtz free energy (orange boxes). As we consider plasticity, we enforce the Karush-Kuhn-Tucker conditions in Equation (4.12) to be fulfilled. Thus, we employ an elastic trial step and, if necessary, update the plastic multiplier iteratively until convergence is achieved. Therefore, the FFN of the yield function (blue box) is evaluated. Finally, the current stress state is computed and returned as an output (blue trapezoid).

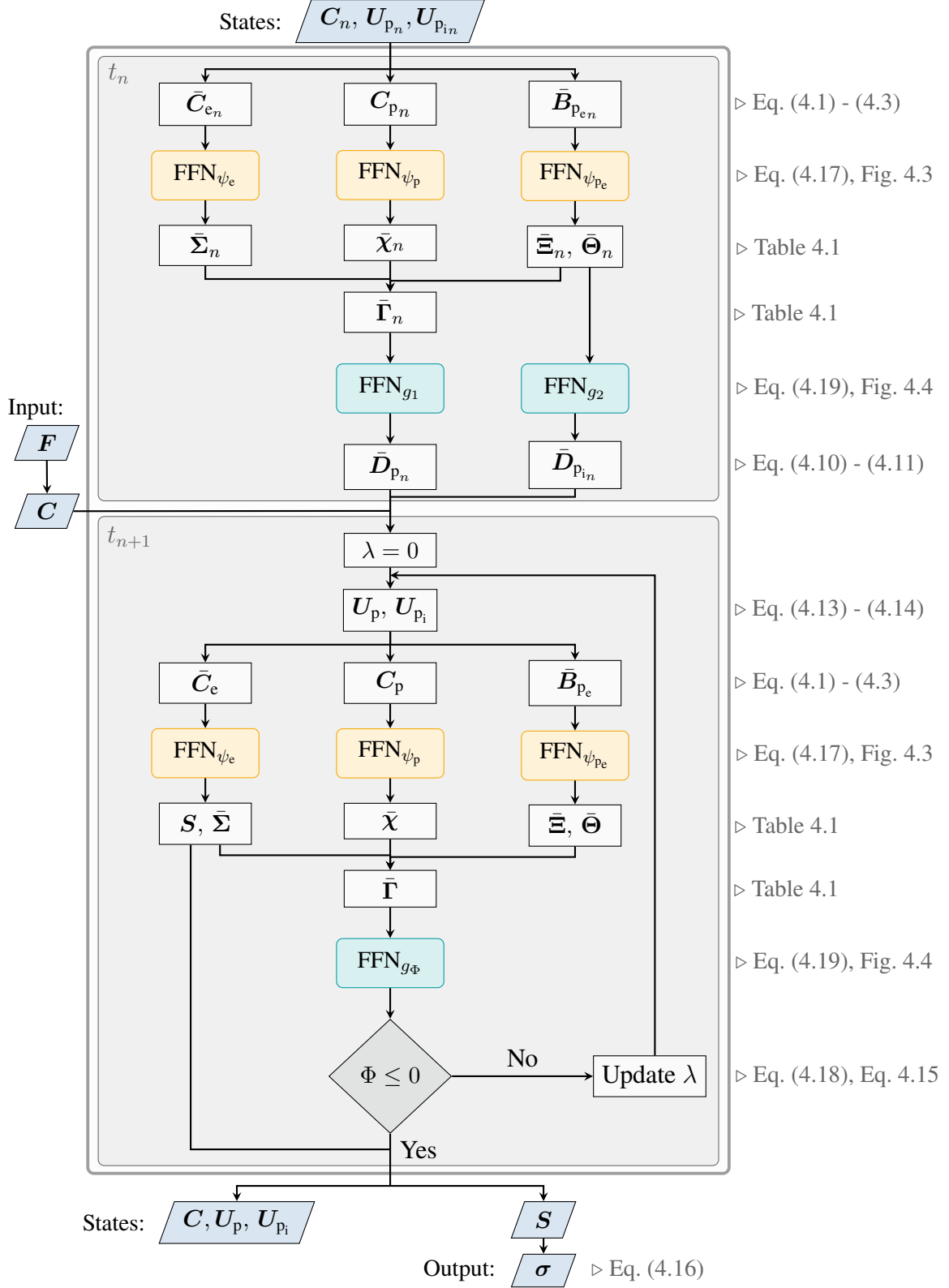


Figure 4.2.: Flowchart of the recurrent neural network (outer gray box) called at each time step. First, all constitutive equations from the last time step t_n (upper gray box) are computed. Internal variables are explicitly updated to compute all constitutive equations of the current time step t_{n+1} (lower gray box). The index $n + 1$ is omitted for readability. Feed-forward networks of the Helmholtz free energies are indicated by orange boxes, while the ones of the plastic potentials are shown by blue boxes. Inputs, outputs and states are visualized by blue trapezoids.

4.4.3. Feed-forward neural network architecture

We obtain two different types of FFNs, one for the Helmholtz free energy and one for the plastic potentials, highlighted by orange and turquoise boxes in Figure 4.2, respectively. For associative plasticity, we obtain $\partial\Phi/\partial\bar{\Gamma} = \partial g_1/\partial\bar{\Gamma}$, reducing the number of FFNs by one. Excluding hardening effects further reduces the amount of trainable neural networks. The structure of these neural networks and their requirements are discussed in the following.

4.4.3.1. Helmholtz free energy

Prerequisites for the formulation. To define a generic formulation of the Helmholtz free energy, we define ψ to be an isotropic function of its arguments and express it in terms of its invariants. Furthermore, we employ the classical isochoric-volumetric split originally introduced by Flory [1961] to satisfy the growth criterion and ensure a stress-free initial configuration at $\bar{C}_e = \mathbf{I}$. In addition, we exclude all mixed invariants of $\bar{C}_e$, C_p , and $\bar{B}_{pe}$ such that we decompose the total Helmholtz free energy to $\psi = \psi_e(\bar{C}_e) + \psi_p(C_p) + \psi_{pe}(\bar{B}_{pe})$, where each term depends on one strain solely such that we can easily prove polyconvexity. Resulting from this decomposition, we employ three separate networks to discover the total Helmholtz free energy. Together with the requirements of fulfilling the growth criterion and ensuring polyconvexity, each formulation must fulfill the normalization condition, i.e. being zero-valued at the origin, and ensuring a stress-free reference configuration. Thus, we can employ the same structure for all of them.

Generic formulation. While using a feed-forward network for the elastic Helmholtz free energy has successfully been applied in CANNs and iCANNs already [Linka et al., 2021; Holthusen et al., 2024b], we follow the same FFN architecture but extend the model by two additional FFNs to discover relations for kinematic hardening. The structure of the FFNs is visualized in Figure 4.3. Here, $\mathbf{A}$ denotes the corresponding strain measure, that is either $\bar{C}_e$, C_p , or $\bar{B}_{pe}$, while $a(\mathbf{A})$ denotes the corresponding output, that is either $\psi_e(\bar{C}_e)$, $\psi_p(C_p)$, or $\psi_{pe}(\bar{B}_{pe})$, respectively.

The invariants of the related strain measure serve as input quantities (see Notations) and are passed through a two-layer network. As described above, we employ the isochoric-volumetric invariants: $\tilde{I}_1 = I_1/I_3^{1/3}$, $\tilde{I}_2 = I_2/I_3^{2/3}$, and I_3 . The former ones are evaluated at the initial state, $\mathbf{A} = \mathbf{I}$, to fulfill the normalization condition and ensure convexity. Here, the second invariant has been evaluated to the power of $3/2$ to ensure polyconvexity [Hartmann and Neff, 2003; Schröder and Neff, 2003]. Previous works have already demonstrated that a sparse architecture is capable of accurately approximating both elastic and inelastic material

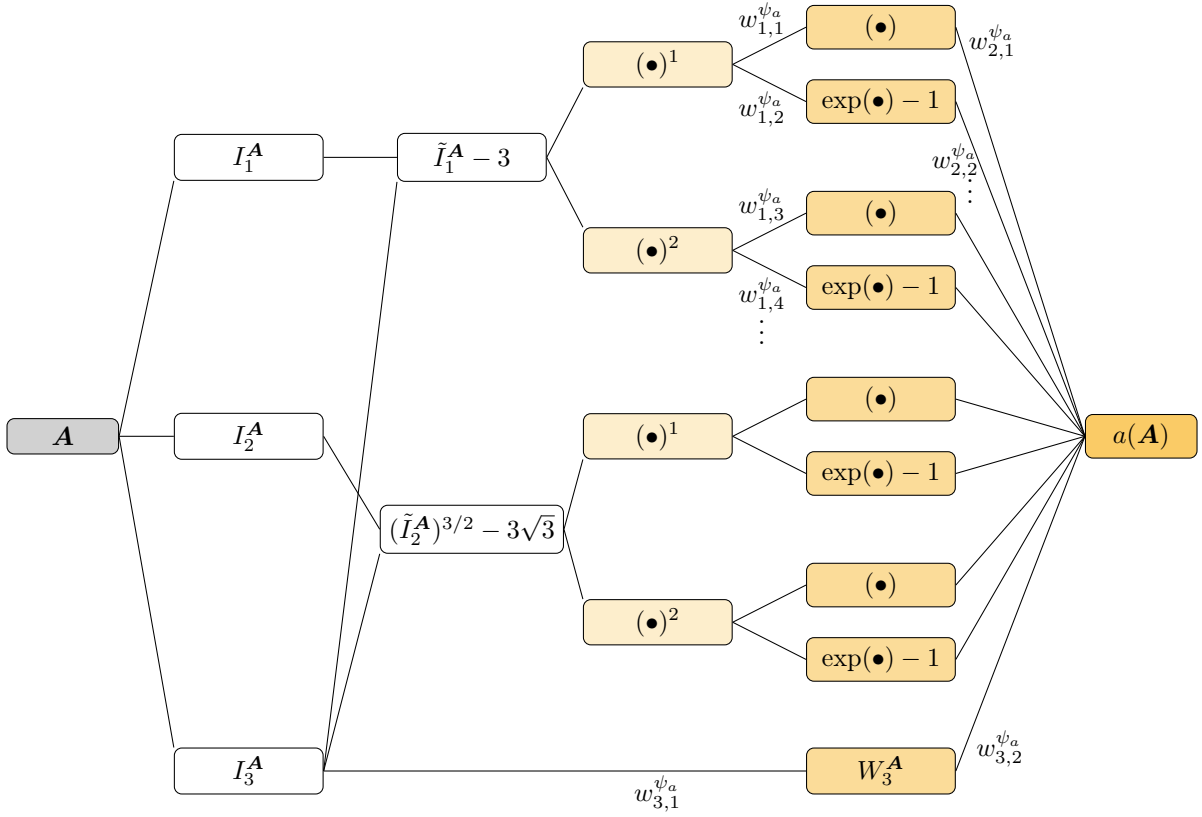


Figure 4.3.: Schematic illustration of the feed-forward network of the elastic and plastic Helmholtz free energies. Elastic formulation: $\mathbf{A} = \bar{\mathbf{C}}_e$ and $a = \psi_e(\bar{\mathbf{C}}_e)$, plastic formulations: $\mathbf{A} = \bar{\mathbf{C}}_p$ and $a = \psi_p(\bar{\mathbf{C}}_p)$, as well as $\mathbf{A} = \bar{\mathbf{B}}_{p_e}$ and $a = \psi_{p_e}(\bar{\mathbf{B}}_{p_e})$.

responses [Linka and Kuhl, 2023; Pierre et al., 2025; Holthusen et al., 2024b,a, 2025b]. Despite its sparsity, the proposed architecture offers a physically interpretable structure, which is particularly beneficial in material modeling, as it facilitates a deeper understanding of the predicted behavior. Given that the present work focuses on modeling plastic effects, and considering the proven effectiveness of the energy formulation in prior applications, we adopt the same architecture herein. Specifically, where linear and quadratic activation functions are applied in the first layer, while the second layer consists of linear and exponential functions. These are well suited for the Helmholtz free energy being monotonic, continuous, and both continuous differentiable and zero-valued at the origin. In addition, they are unbounded such that the growth criterion is fulfilled. The volumetric term is designed following Ogden [1972]. The weights of the first layer define the shape of the activation functions, while the weights of the second layer scale their contribution and can be interpreted as material parameters. All weights are trainable parameters and are optimized during the training process. Fulfilling the requirements given above, we restrict the weights to be non-negative to obtain physically

reasonable results, as negative weights would lead to negative Helmholtz free energies that are not admissible. The FFN of the Helmholtz free energy can be expressed equation-wise as

$$\begin{aligned}
\psi(\mathbf{A}) = & w_{2,1}^\psi (\tilde{I}_1^{\mathbf{A}} - 3) + w_{2,2}^\psi \left(\exp(w_{1,2}^\psi [\tilde{I}_1^{\mathbf{A}} - 3]) - 1 \right) \\
& + w_{2,3}^\psi (\tilde{I}_1^{\mathbf{A}} - 3)^2 + w_{2,4}^\psi \left(\exp \left(w_{1,4}^\psi [\tilde{I}_1^{\mathbf{A}} - 3]^2 \right) - 1 \right) \\
& + w_{2,5}^\psi \left((\tilde{I}_2^{\mathbf{A}})^{3/2} - 3\sqrt{3} \right) + w_{2,6}^\psi \left(\exp(w_{1,6}^\psi [(\tilde{I}_2^{\mathbf{A}})^{3/2} - 3\sqrt{3}]) - 1 \right) \\
& + w_{2,7}^\psi \left((\tilde{I}_2^{\mathbf{A}})^{3/2} - 3\sqrt{3} \right)^2 + w_{2,8}^\psi \left(\exp \left(w_{1,8}^\psi [(\tilde{I}_2^{\mathbf{A}})^{3/2} - 3\sqrt{3}]^2 \right) - 1 \right) \\
& + w_{3,2}^\psi \underbrace{\left[(I_3^{\mathbf{A}})^{w_{3,1}^\psi} - 1 - w_{3,1}^\psi \ln(I_3^{\mathbf{A}}) \right]}_{W_3^{\mathbf{a}}}. \tag{4.17}
\end{aligned}$$

All redundant weights of the first layer corresponding to the linear activation functions have been omitted here.

4.4.3.2. Plastic potential

In the context of plasticity, we must formulate functions for both the yield surface and the evolution equations. Focussing on the yield surface, we aim to establish a general formulation that can be expressed through a neural network-based approach. In general, the yield surface can be described in terms of a stress-dependent function, $g_\Phi(\bar{\Gamma})$, such that the yield function reads

$$\Phi(\bar{\Gamma}) = g_\Phi(\bar{\Gamma}) - 1 \leq 0. \tag{4.18}$$

Yielding occurs when $g_\Phi(\bar{\Gamma}) = 1$. The shape of the yield surface is defined through the function g_Φ , which is denoted as a plastic potential. In the context of associative plasticity, it defines the direction of plastic flow, i.e. $g_1 = g_\Phi$. To flexibly discover the underlying yield behavior, we model it in terms of an FFN and employ the same network structure that will be introduced for the potentials in Equations (4.10) and (4.11).

Requirements on the potentials. The potentials are designed to be scalar-valued isotropic functions of their arguments such that we can define them in terms of stress invariants. While the choice of these invariants is not unique, we choose a formulation in terms of the integrity basis of their inputs, ensuring rotational non-uniqueness. Specifically, we use the invariants I_1 and J_2 (see Notations) to account for hydrostatic pressures and shear stresses. We omit the third deviatoric invariant, J_3 , here. Although often used in plasticity, it does not fulfill

convexity in general [Holthusen et al., 2025b]. Further, the additionally gained information about the stress behavior is often negligible [Hill, 1952].

In analogy to the energy, we additively decompose each potential into sub-potentials that depend separately on the invariants, i.e. $g(\mathbf{A}) = g_I(I_1^{\mathbf{A}}) + g_{II}(J_2^{\mathbf{A}})$ ¹. Thus, the neurons are not connected to all neurons in the subsequent layer, leading to a network architecture that is not fully connected. However, following this architecture, convexity can be ensured easily.²

Generic formulation. We employ a two-layered network architecture. Again, the choice of the activation functions is not unique, but we choose linear and quadratic terms for the first layer (light blue layer in Figure 4.4) to ensure convexity and non-negativity. The second layer (darker blue layer in Figure 4.4) consists of the activation functions that are the absolute value function $\text{abs}(\bullet)$, the logarithm of the hyperbolic cosine function $\ln(\cosh(\bullet))$, and the hyperbolic cosine function minus one $\cosh(\bullet) - 1$. These activation functions are chosen to ensure convexity, non-negativity, and to be zero-valued at the origin. The FFN architecture is visualized in Figure 4.4, while the resulting equation for the feed-forward network reads

$$\begin{aligned} g(\mathbf{A}) = & w_{2,1}^g \text{abs}(I_1^{\mathbf{A}}) + w_{2,2}^g \ln[\cosh(w_{1,1}^g I_1^{\mathbf{A}})] + w_{2,3}^g [\cosh(w_{1,2}^g I_1^{\mathbf{A}}) - 1] \\ & + w_{2,4}^g \text{abs}\left((I_1^{\mathbf{A}})^2\right) + w_{2,5}^g \ln[\cosh(w_{1,3}^g (I_1^{\mathbf{A}})^2)] + w_{2,6}^g [\cosh(w_{1,4}^g (I_1^{\mathbf{A}})^2) - 1] \\ & + w_{2,7}^g \text{abs}(J_2^{\mathbf{A}}) + w_{2,8}^g \ln[\cosh(w_{1,5}^g J_2^{\mathbf{A}})] + w_{2,9}^g [\cosh(w_{1,6}^g J_2^{\mathbf{A}}) - 1] \\ & + w_{2,10}^g \text{abs}\left((J_2^{\mathbf{A}})^2\right) + w_{2,11}^g \ln[\cosh(w_{1,7}^g (J_2^{\mathbf{A}})^2)] + w_{2,12}^g [\cosh(w_{1,8}^g (J_2^{\mathbf{A}})^2) - 1]. \end{aligned} \quad (4.19)$$

Here, $\mathbf{A}$ should be replaced by either $\bar{\Gamma}$ for the potential associated with the yield surface, g_Φ

¹The additive decomposition into sub-functions can lead to a non-smooth potential, which might cause problems of non-unique flow directions, non-differentiability at the intersections of the sub-functions and numerical instability. To avoid these problems, a smoothing technique as described by Gesto Beiroa et al. [2011], such as the p-norm regularization Mollica and Srinivasa [2002], can be applied. The potential can then be modified to $g(\bar{\Gamma}) = (\sum_{i=1}^n |g_i(\bar{\Gamma})|^p)^{1/p}$, where p is a positive parameter and n is the number of sub-potentials. In contrast to the approach by Mollica and Srinivasa [2002], it is redundant to apply the absolute function due to our choice of the sub-functions g_i to be convex, non-negative, and zero-valued.

²Anisotropic material behavior is of interest in many engineering applications. A commonly used method is the use of structural tensors [Spencer, 1971, 1984; Boehler, 1979; Zhang and Rychlewski, 1990; Schröder and Neff, 2003]. Although the current framework does not yet account for elastic or plastic anisotropy, their incorporation is possible in principle. By introducing structural tensors, the set of invariants used would be expanded to include the additional mixed invariants, which extend the input layer of the FFNs, as demonstrated by Vervenne et al. [2025a]. Other approaches have been shown, for example, by Gärtner et al. [2021], or Klein et al. [2022a]. Anisotropic plasticity, in particular, can be modeled following the framework proposed by Vladimirov et al. [2010]. However, the mapping of the structural tensors into the intermediate configuration is required but not unique Holthusen et al. [2023]. Consequently, inclusion of elastic and plastic anisotropy into our current framework remains applicable but requires extensive theoretical derivation and is therefore out of the scope of the current work.

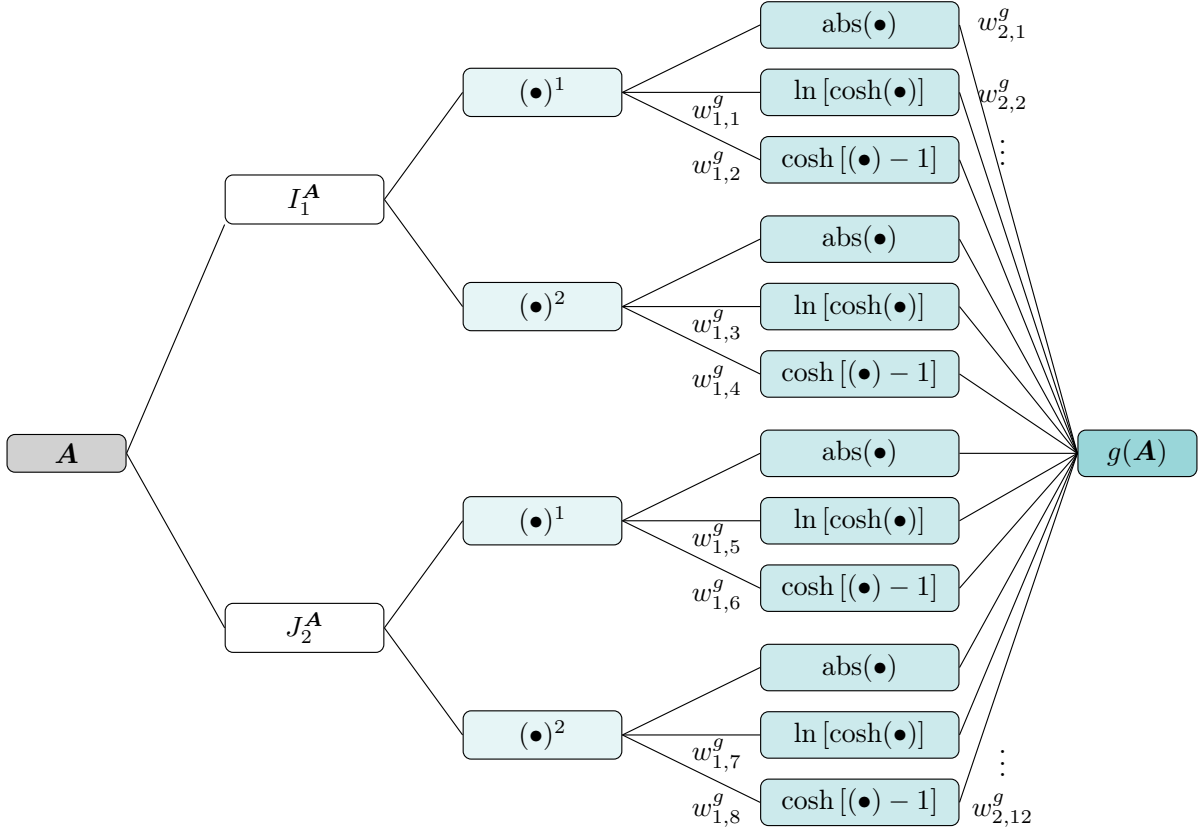


Figure 4.4.: Schematic illustration of the feed-forward network of the plastic potential of Equation (4.19) with $A = \bar{\Gamma}$ for g_1 and g_Φ , and $A = \bar{\Theta}$ for g_2 .

(Equation (4.18)), and the evolution of plastic strains, g_1 (Equation (4.10)), or by $\bar{\Theta}$ for the potential g_2 (Equation (4.11)).

4.5. Automated model discovery

In the following, numerical results are demonstrated to validate the presented framework of neural networks to learn elasto-plastic behavior. This is achieved using various training datasets, including artificially generated data based on the von Mises yield function, data exhibiting tension-compression asymmetry, comparisons with other neural network approaches, and training on experimental data of steel. The framework is implemented into the library *TensorFlow* [Abadi et al., 2016]. We restrict this work to incompressible material behavior. Therefore, we add a Lagrangian term to the energy that reads $p(I_3^{C_e} I_3^{C_p} - 1)$. Here, p can be interpreted as the hydrostatic pressure that is computed from the boundary conditions. Con-

sidering coaxial loads only, we can define p such that $S_{33} = 0$ holds. We compute the loss, $\mathcal{L}$, by taking the mean squared error (MSE) of the first entries of the experimentally observed and predicted Cauchy stress of all data points, i , summarized in σ_{11} and $\hat{\sigma}_{11}$, respectively, which then reads

$$\mathcal{L}(\sigma_{11}, \mathbf{w}) = \frac{1}{n_{\text{exp}}} \sum_{n=1}^{n_{\text{exp}}} \left(\frac{1}{n_{\text{data}}} \sum_{i=1}^{n_{\text{data}}} (\sigma_{11_i} - \hat{\sigma}_{11_i})^2 \right) + R(\mathbf{w}). \quad (4.20)$$

The number of experiments used for training are denoted by n_{exp} , while n_{data} are the number of data points per experiment, and n_w the number of trainable weights regularized, which are summarized in $\mathbf{w}$. Throughout the training process, different regularization schemes have been investigated. These are captured in the loss function through the regularization term $R(\mathbf{w})$, which is defined as

$$R(\mathbf{w}) = L_2 \sum_{j=1}^{n_w} w_j^2. \quad (4.21)$$

Here, L_2 is the ridge regularization factor, and w_j are the trainable weights of the network. The regularization has been applied with $L_2 = 0.001$ for all FFNs of the Helmholtz free energies and $L_2 = 0.0001$ for all potentials if not mentioned otherwise. In addition to the L_2 -regularization, a comparison of the impact of different regularization techniques on our model is given in 4.8.1, where Lasso-regularization (L_1) is compared to ridge regularization (L_2). Throughout all trainings, the ADAM optimization [Kingma and Ba, 2014] has been used. In addition, gradient-clipping has been applied to prevent training against nonphysical values by $G_{\text{clip}} = G \cdot \min(1, c/|G|)$. Here, G is the initially computed gradient and $c = 0.01$ the clip-norm threshold.

4.5.1. Specification of network architecture

The formulation of the Helmholtz free energy FFNs is given in Equation (4.17). To point out the main idea of our framework and reduce the numerical effort, we neglect the squared terms of the first layer of the energies for linear and non-linear kinematic hardening, such that we

obtain

$$\begin{aligned} \psi(\mathbf{C}_p) = & w_{2,1}^{\psi_p} \left(\tilde{I}_1^{\mathbf{C}_p} - 3 \right) + w_{2,2}^{\psi_p} \left(\exp \left(w_{1,1}^{\psi_p} [\tilde{I}_1^{\mathbf{C}_p} - 3] \right) - 1 \right) \\ & + w_{2,3}^{\psi_p} \left(\left(\tilde{I}_2^{\mathbf{C}_p} \right)^{3/2} - 3\sqrt{3} \right) + w_{2,4}^{\psi_p} \left(\exp \left(w_{1,2}^{\psi_p} \left[\left(\tilde{I}_2^{\mathbf{C}_p} \right)^{3/2} - 3\sqrt{3} \right] \right) - 1 \right) \\ & + w_{3,2}^{\psi_p} \left[\left(I_3^{\mathbf{C}_p} \right)^{w_{3,1}^{\psi_p}} - 1 - w_{3,1}^{\psi_p} \ln \left(I_3^{\mathbf{C}_p} \right) \right] \end{aligned} \quad (4.22)$$

$$\begin{aligned} \psi(\bar{\mathbf{B}}_{pe}) = & w_{2,1}^{\psi_{pe}} \left(\tilde{I}_1^{\bar{\mathbf{B}}_{pe}} - 3 \right) + w_{2,2}^{\psi_{pe}} \left(\exp \left(w_{1,1}^{\psi_{pe}} [\tilde{I}_1^{\bar{\mathbf{B}}_{pe}} - 3] \right) - 1 \right) \\ & + w_{2,3}^{\psi_{pe}} \left(\left(\tilde{I}_2^{\bar{\mathbf{B}}_{pe}} \right)^{3/2} - 3\sqrt{3} \right) + w_{2,4}^{\psi_{pe}} \left(\exp \left(w_{1,2}^{\psi_{pe}} \left[\left(\tilde{I}_2^{\bar{\mathbf{B}}_{pe}} \right)^{3/2} - 3\sqrt{3} \right] \right) - 1 \right) \\ & + w_{3,2}^{\psi_{pe}} \left[\left(I_3^{\bar{\mathbf{B}}_{pe}} \right)^{w_{3,1}^{\psi_{pe}}} - 1 - w_{3,1}^{\psi_{pe}} \ln \left(I_3^{\bar{\mathbf{B}}_{pe}} \right) \right]. \end{aligned} \quad (4.23)$$

It should be noted that the weights have been renumbered accordingly to maintain a consecutive numbering of all remaining weights. In addition, we choose an associative evolution law for the evolution of $\mathbf{C}_p$, such that the potential of the yield criterion and of the plastic stretches are the same, i.e. $g_\Phi = g_1$. Thus, the two potentials g_1 and g_2 remain to be learned. We describe both by the following reduced formulations, which are

$$\begin{aligned} g_1(\bar{\Gamma}) = & w_{2,1}^{g_1} \text{abs} \left(I_1^{\bar{\Gamma}} \right) + w_{2,2}^{g_1} \ln \left[\cosh \left(w_{1,1}^{g_1} I_1^{\bar{\Gamma}} \right) \right] \\ & + w_{2,3}^{g_1} \text{abs} \left((I_1^{\bar{\Gamma}})^2 \right) + w_{2,4}^{g_1} \ln \left[\cosh \left(w_{1,2}^{g_1} (I_1^{\bar{\Gamma}})^2 \right) \right] \\ & + \tilde{w}_{2,5}^{g_1} \text{abs} \left(\tilde{J}_2^{\bar{\Gamma}} \right) + w_{2,6}^{g_1} \ln \left[\cosh \left(\tilde{w}_{1,3}^{g_1} \tilde{J}_2^{\bar{\Gamma}} \right) \right] \end{aligned} \quad (4.24)$$

$$\begin{aligned} g_2(\bar{\Theta}) = & w_{2,1}^{g_2} \text{abs} \left(I_1^{\bar{\Theta}} \right) + w_{2,2}^{g_2} \ln \left[\cosh \left(w_{1,1}^{g_2} I_1^{\bar{\Theta}} \right) \right] \\ & + w_{2,3}^{g_2} \text{abs} \left((I_1^{\bar{\Theta}})^2 \right) + w_{2,4}^{g_2} \ln \left[\cosh \left(w_{1,2}^{g_2} (I_1^{\bar{\Theta}})^2 \right) \right] \\ & + \tilde{w}_{2,5}^{g_2} \text{abs} \left(\tilde{J}_2^{\bar{\Theta}} \right) + w_{2,6}^{g_2} \ln \left[\cosh \left(\tilde{w}_{1,3}^{g_2} \tilde{J}_2^{\bar{\Theta}} \right) \right]. \end{aligned} \quad (4.25)$$

Notably, we have scaled the second deviatoric invariant by $\tilde{J}_2 = 3 J_2$. Thus, we obtain for the corresponding weights the relation $w_{2,5}^g = 3 \tilde{w}_{2,5}^g$ and $w_{1,3}^g = 3 \tilde{w}_{1,3}^g$, which is in line with for example Taç et al. [2023].

4.5.2. Preprocessing

All stresses were normalized before being used for training to avoid excessive sensitivity to randomly initialized weights. To our experience, and in contrast to visco-elasticity, the

satisfaction of the Karush-Kuhn-Tucker conditions causes a more unstable training and often does not converge or abort. Therefore, we performed a pretraining of the initial weights by setting all weights to zero and non-trainable, and began with training the weights of the elastic Helmholtz free energy. We then sequentially pretrained the weights of the yield function, the linear hardening energy, the nonlinear hardening energy, and the corresponding plastic potential. This strategy ensured that the initial weights were already close to their final values, leading to less error-prone and more stable training processes.

In particular, a key aspect of this approach was to enforce plastic behavior within the training domain to ensure that the loss was dependent on the weights of the plastic potential. Without this enforcement, a purely elastic iteration step would cause these weights to be automatically set to zero in the next iteration. Thus, any plastic dependency would be eliminated, causing the model to predict purely elastic material behavior.

4.5.3. Discovering artificially generated data

In the following, we investigate the prediction of artificially generated data. For this, we first use data of a model that is included in our network architecture. Subsequently, we aim to find a model that can not be captured by our architecture, as it includes tension-compression asymmetry.

4.5.3.1. Artificially generated data

We investigated the performance first by training on artificially generated data of an elastoplastic material with nonlinear kinematic hardening. The underlying data was generated using a Neo-Hookean model type, where the Helmholtz free energy was defined by

$$\begin{aligned} \psi(\bar{\mathbf{C}}_e) = & \frac{\mu}{2} \left(\frac{\text{tr}(\bar{\mathbf{C}}_e)}{\det(\bar{\mathbf{C}}_e)^{1/3}} - 3 \right) + \frac{K}{4} (\det(\bar{\mathbf{C}}_e) - 1 - \ln[\det(\bar{\mathbf{C}}_e)]) \\ & + \frac{c}{2} \left(\frac{\text{tr}(\mathbf{C}_p)}{\det(\mathbf{C}_p)^{1/3}} - 3 \right) - p(I_3^C - 1). \end{aligned} \quad (4.26)$$

The last term was added as a Lagrange term where p represented the hydrostatic pressure, derived from the boundary conditions. The yield function was chosen as von Mises type, i.e. given by

$$\Phi(\bar{\mathbf{\Gamma}}) = \frac{3}{2} \text{tr}(\text{dev}(\bar{\mathbf{\Gamma}})^2) - \sigma_{y0}^2, \quad (4.27)$$

with the shear modulus $\mu = 12.5$ kPa, the bulk modulus $K = 25$ kPa and the yield stress $\sigma_{y0} = 2$ kPa. The hardening parameter is $c = 8.5$ kPa, while the evolution of hardening is described by a plastic potential of $g_2 = b/2 \text{tr}(\text{dev}(\bar{\Theta})^2)$ with $b = 3$ kPa.

To evaluate the model's predictive capability, we trained it on different sets of loading conditions. Given that a large amount of data from experiments in reality is often limited due to costs and time, we restricted our data to a realistic set of load sets:

- Uniaxial tension (UT)
- Uniaxial compression (UC)
- Equibiaxial tension (EB)
- Uniaxial tension with unloading (UT-unl)
- Uniaxial cyclic loading (Cyc)

These data sets were used in five different combinations to examine the model prediction with different training data. Particularly, the increase in data was investigated by increasing the uniform loads: (UT) $\rightarrow$ (UT, EB) $\rightarrow$ (UT, EB, UC). Further, we studied the influence of more complex loading paths: (UT-unl) and (Cyc). Each training set was run for a maximum of 5000 epochs, while early stopping of the training process was triggered if the loss did not decrease for 50 consecutive epochs. The stress-strain and corresponding loss curves of the different combinations of uniform loads are shown in Figure 4.5, where training on (UT) is shown in (a), training on (UT, EB) in (b), and (UT, EB, UC) in (c). Similarly, the results for training on (UT-unl) and (Cyc) are compared in Figure 4.6 (b) and (c), respectively. The training on (UT) is given in (a) as a reference. The predicted weights of the yield surface and the Helmholtz free energies accounting for kinematic hardening are summarized in Tables 4.2 and 4.3.

Results. The stress-strain curves indicate a qualitative improvement in the model prediction with increasing training data. While training on uniaxial tension alone provided already satisfactory results for this load case, the load cases used for testing revealed offsets. Expanding the dataset with additional uniform load sets improved the model's prediction as can be seen in Figure 4.5 (b) and (c). The best prediction was obtained for complex loading paths, in particular for cyclic loading, see Figure 4.6 (b) and (c).

These observations align with the predicted weights. Ideally, the prediction should yield $w_{2,5}^{g\Phi} = 0.25$, $w_{2,5}^{g_2} = 1/3$, and $w_{1,1}^{\psi_{pe}} = 8.5$. From Table 4.2, we see that training on (UT) produces weights close to the expected value of $w_{2,5}^{g\Phi}$ but with some noise, as there remain some non-zero weights. The minimal number of non-zero weights was observed in the training on

Table 4.2.: Artificially generated data: Learned weights of the yield function (Equation (4.24)) for the different load cases used for training (UT – Uniaxial tension, EB – Equibiaxial tension, UC – Uniaxial compression, UT-unl – Uniaxial tensions with unloading).

Loadcase	UT	UT, EB	UT, EB, UC	UT-unl	Cyclic
$w_{1,1}^{g\Phi}$	0	0	0	0	0
$w_{1,2}^{g\Phi}$	0	0	0	0	0
$w_{1,3}^{g\Phi}$	0	0	0	0	0
$w_{2,1}^{g\Phi}$	1.878e-6	6.342e-4	1.129e-4	3.444e-5	0
$w_{2,2}^{g\Phi}$	0	0	0	0	0
$w_{2,3}^{g\Phi}$	7.044e-11	5.773e-9	1.068e-9	4.261e-10	0
$w_{2,4}^{g\Phi}$	0	0	0	0	0
$w_{2,5}^{g\Phi}$	2.536e-1	2.603e-1	2.481e-1	2.427e-1	2.349e-1
$w_{2,6}^{g\Phi}$	0	0	0	0	0

cyclic loading. For this load case, only the weight $w_{2,5}^{g\Phi}$ was non-zero. Hence, the ideally predicted weight was obtained. Additionally, analyzing the weights for linear and nonlinear kinematic hardening in Table 4.3 reveals that the training on cyclic loading led to elimination of all linear hardening terms, ensuring that all hardening effects originated from the nonlinear hardening terms, which was in line with the analytical solution.

Interestingly, when comparing the results of the load sets with the greatest variety – (UT, EB, UC) and (Cyc) – we observed similar predictions but faster convergence during training when multiple load cases were included (UT, EB, UC) compared to the cyclic training data, highlighting the advantage of incorporating diverse load sets in the training data. Thus, we can conclude from both qualitatively (stress-strain data) and quantitatively (predicted weights and convergence) that increasing the variety in training data enhances the predictive accuracy of our model, which aligns with the information entropy principle [Shannon, 1948]. Particularly, including loading and unloading paths is crucial for capturing history-dependent material behavior and ensuring realistic hardening predictions. Cyclic loading data is especially important to distinguish between different hardening effects. While we did not consider isotropic hardening here, this would even further contribute to the desire for cyclic loading.

4.5.3.2. Artificially generated data with tension-compression asymmetry

In the previous section, we analyzed the model performance for data that can be captured by our network architecture. To consider cases that could not be captured by the iCANN, we investigated an elasto-plastic material with perfect plasticity and a yield criterion that captured tension-compression asymmetry. This model was based on a compressible Neo-Hookean type

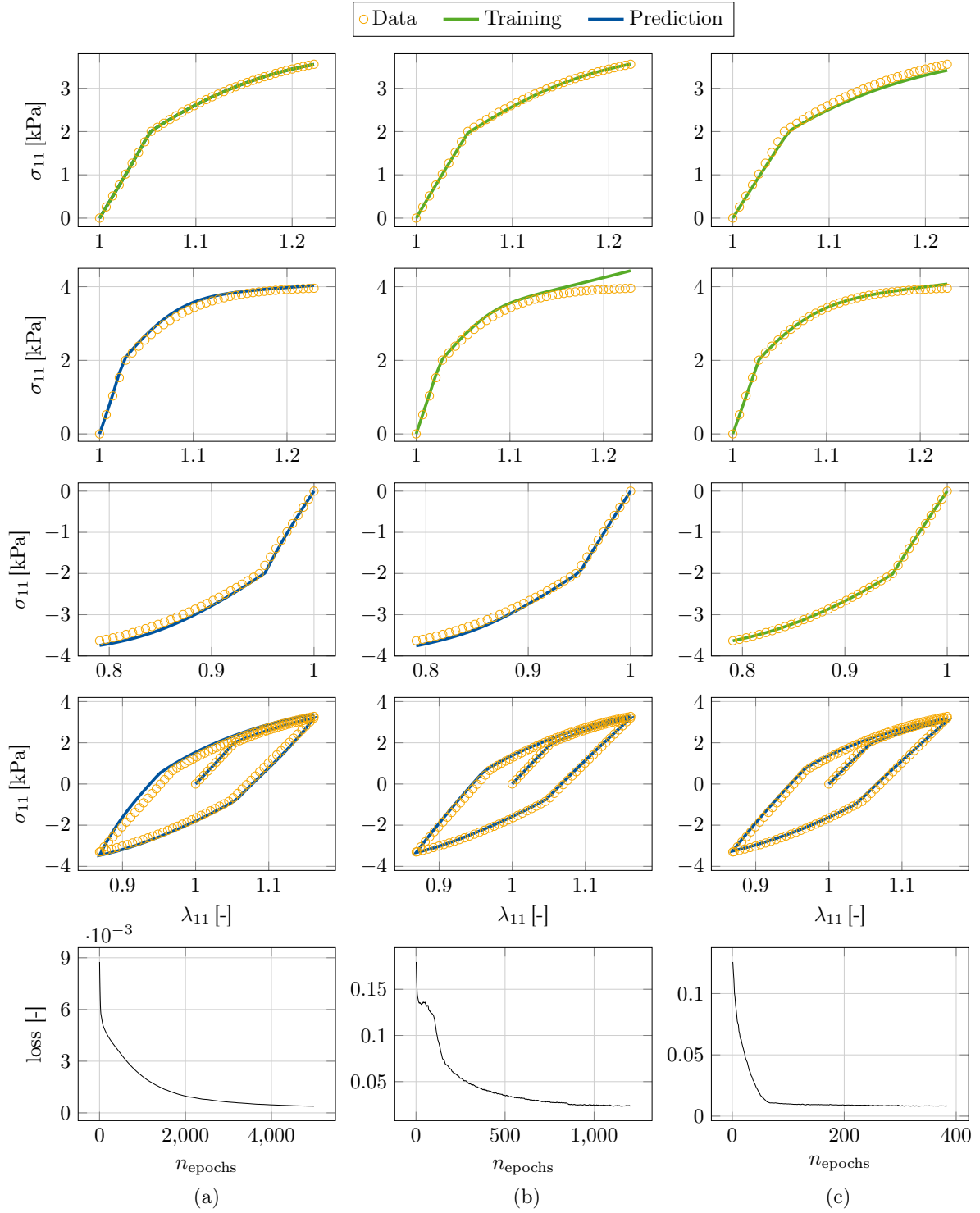


Figure 4.5.: Artificially generated data: Stress strain curves of uniaxial tension (1st row), equibiaxial tension (2nd row), uniaxial compression (3rd row), cyclic loading (4th row), and corresponding loss curves (5th row). Trained on (a) uniaxial tension, (b) uniaxial tension and equibiaxial tension, (c) uniaxial tension, equibiaxial tension, uniaxial compression. Early stopping was enforced after 50 consecutive epochs of non-decreasing loss values.

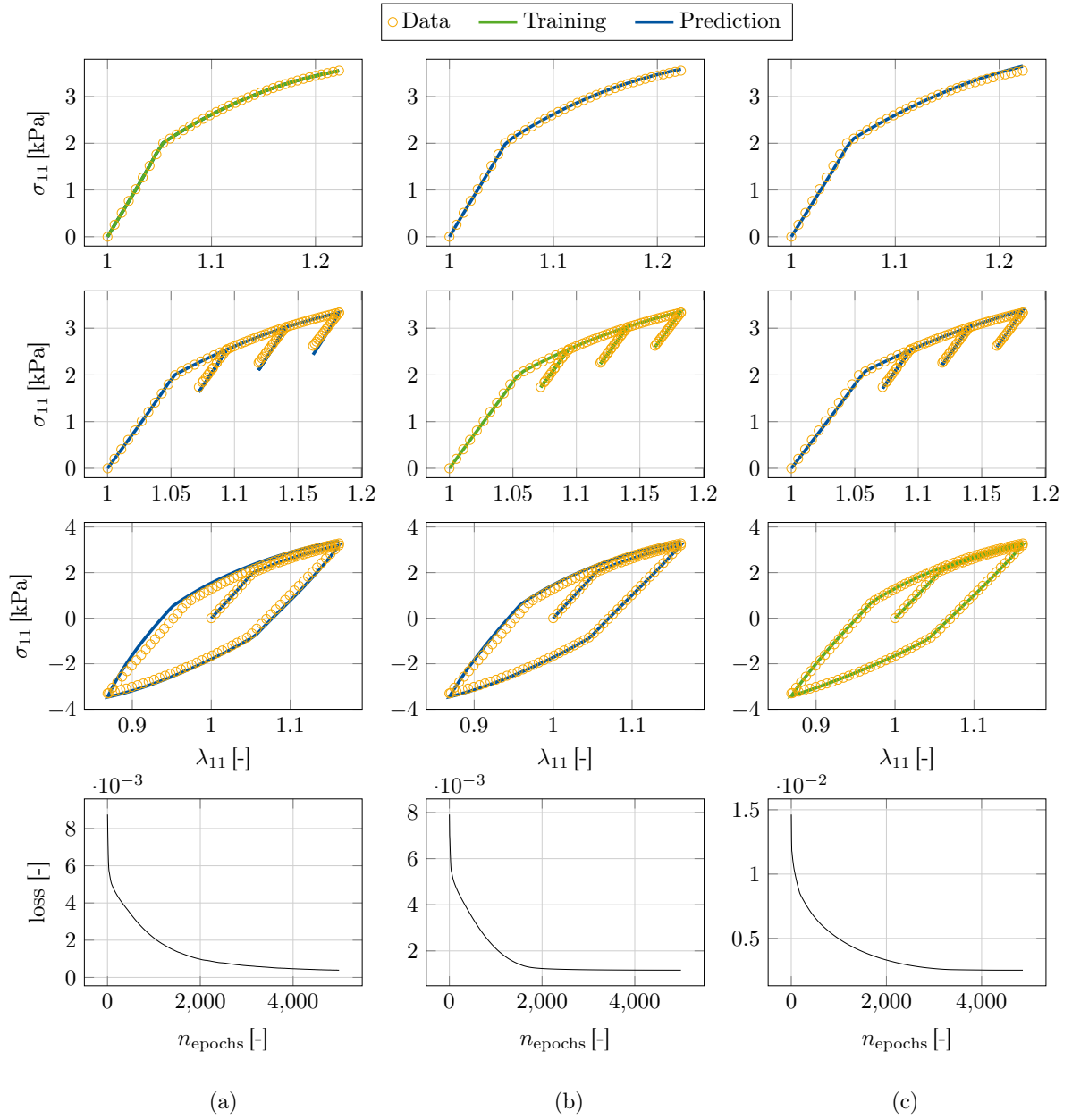


Figure 4.6.: Artificially generated data: Stress strain curves of uniaxial tension (1st row), uniaxial tension with unloading (2nd row), cyclic loading (3rd row), and corresponding loss curves (5th row). Trained on (a) uniaxial tension, (b) uniaxial tension with unloading, (c) cyclic loading. Early stopping was enforced after 50 consecutive epochs of non-decreasing loss values.

Table 4.3.: Artificially generated data: Learned weights of the Helmholtz energy accounting for linear kinematic hardening (ψ_p) and nonlinear kinematic hardening (ψ_{pe}).

	UT	UT, EB	UT, EB, UC	UT-unl	Cyc
$w_{1,1}^{\psi_p}$	0	0	0	0	0
$w_{1,2}^{\psi_p}$	0	0	0	0	0
$w_{2,1}^{\psi_p}$	1.510e-2	2.853e-1	6.461e-2	3.933e-2	0
$w_{2,2}^{\psi_p}$	0	0	0	0	0
$w_{2,3}^{\psi_p}$	4.081e-2	1.234e-1	5.536e-2	9.599e-2	0
$w_{2,4}^{\psi_p}$	0	0	0	0	0
$w_{1,1}^{\psi_{pe}}$	6.063e+0	4.644e+0	4.648e+0	4.848e+0	5.197e+0
$w_{1,2}^{\psi_{pe}}$	4.959e+0	3.311e+0	3.392e+0	3.747e+0	4.835e+0
$w_{2,1}^{\psi_{pe}}$	1.856e-2	1.111e-1	5.815e-2	2.694e-2	1.598e-2
$w_{2,2}^{\psi_{pe}}$	1.121e-1	2.869e-1	1.863e-1	1.524e-1	9.283e-2
$w_{2,3}^{\psi_{pe}}$	4.922e-2	1.434e-1	1.479e-1	6.838e-2	5.210e-2
$w_{2,4}^{\psi_{pe}}$	2.321e-1	2.215e-1	2.729e-1	3.018e-1	2.536e-1

energy with a paraboloid yield function following Tschoegl [1971], i.e.

$$\Phi = 3 J_2 + (\sigma_C - \sigma_T) I_1 - \sigma_c \sigma_T. \quad (4.28)$$

Here, the initial yield stress was set to $\sigma_T = 2$ kPa in tension and to $\sigma_C = 4$ kPa in compression. Tension-compression asymmetry was captured through the sign of the first invariant I_1 in Equation (4.28). Notably, the activation functions used in our model are non-negative and convex, thereby excluding any dependency of the sign of the inputs and preventing the network from learning the tension-compression asymmetry. Again, different combinations of load cases were investigated: (UT), (UT, UC), and (UT, EB). As the main interest lied in investigating the predicted yield criterion, we focused solely on the yield surface prediction by fixing the elastic energy to ensure correct elastic behavior and setting all other weights to zero. The discovered yield surfaces are visualized in Figure 4.7 in 2D in the σ_{11} - σ_{22} -plane in the upper line and in 3D in the lower line. The different load cases used for training are shown in (a) for uniaxial tension, (b) uniaxial and equibiaxial tension, and (c) uniaxial tension and compression. The initial yield stresses corresponding to each load set are represented by black markers, providing a visual reference of the training data. The yield surface of Equation (4.28) is visualized in orange and the predicted one in blue.

Results. It can be observed that for the training using the UT data only, the predicted yield stress matches the reference value accurately. However, the overall shape of the yield surface

Table 4.4.: Artificially generated data with tension-compression asymmetry: Learned weights of the yield function (Equation (4.24)) for the different load cases used for training (UT – Uniaxial tension, EB – Equibiaxial tension, UC – Uniaxial compression).

Loadcase	UT	UT, EB	UT, UC
$w_{1,1}^{g\Phi}$	0	0	0
$w_{1,2}^{g\Phi}$	0	0	0
$w_{1,3}^{g\Phi}$	1.286e-1	0	0
$w_{2,1}^{g\Phi}$	4.024e-7	8.562e-2	3.709e-2
$w_{2,2}^{g\Phi}$	0	0	0
$w_{2,3}^{g\Phi}$	2.553e-12	9.457e-10	6.491e-9
$w_{2,4}^{g\Phi}$	0	0	0
$w_{2,5}^{g\Phi}$	2.485e-1	2.606e-1	1.187e-1
$w_{2,6}^{g\Phi}$	4.614e-2	0	4.303e-3

is not in line with the underlying model (see Figure 4.7 (a)), where the prediction resembles the von Mises type yield surface. Investigation of the predicted weights given in Table 4.4 reveals that the invariant J_2 most dominantly affected the result. When the training data set has been increased to include UT and EB, the form of the yield function had changed due to an increased influence of the first invariant and aligned more closely with the training data in the tensile octant (c.f. Figure 4.7 (b)). Despite this change in the form, the model had failed to capture the asymmetry of the Tschoegel yield function, as the discovered and reference solution still differed significantly in the compressive regime. Training on UT and UC resulted in a compromise between tension and compression, but still failed to represent the full asymmetry reflected in the yield stresses. These findings confirm that without explicit mechanisms to capture sign-dependent behavior, the model struggles to learn tension-compression asymmetry. This highlights the need for alternative architectures or additional constraints to improve yield surface predictions for asymmetric materials.

4.5.4. Comparison to other neural network approaches

There are numerous other neural network approaches to discover material models including effects of plasticity. One of them is an efficient kernel learning-based constitutive model (EKLC) presented by Liao and Luo [2025]. This model is significantly faster than other neural network approaches while effectively capturing elasto-plastic material behavior, including both linear and nonlinear hardening effects. A key distinction between EKLC and our approach is that our model fulfills thermodynamic consistency by design, and it is formulated within the finite strain theory, making it more flexible to be applied to various materials exhibiting large

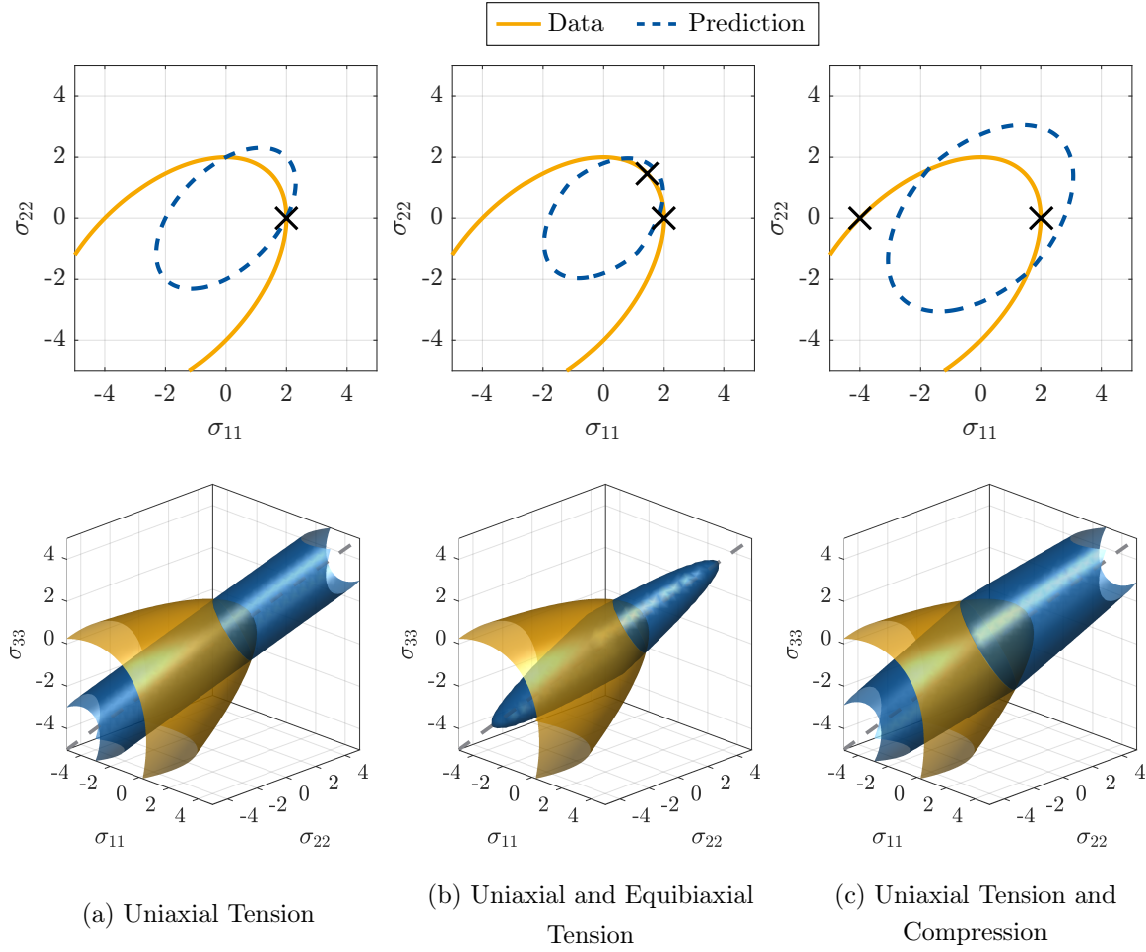


Figure 4.7.: Artificially generated data with tension-compression asymmetry: Underlying yield surface (blue) and predicted yield surfaces (orange) in 2D (upper row) and 3D (lower row) in the stress space (in kPa) for training on (a) uniaxial tension, (b) uni- and equibiaxial tension, and (c) uniaxial tension and compression (hydrostatic axis is given for reference in gray).

deformations. Additionally, we do not only learn the yield function and hardening law but also discover the complete constitutive relations, further enhancing the model's adaptability. As described in Section 4.4, the weights of our feed-forward networks can be interpreted as material parameters due to the structure of the networks.

The EKLC model was trained on four cycles and tested on cyclic loading with varying step sizes and showed good agreement with the training and testing data. Taking the same data set into account, we performed training on a reduced amount of data and investigated our model prediction. The regularization factor of the energy FFNs was chosen to $L_2 = 0.0001$, regularizing the FFNs of the energies and the potentials similarly. Three different training sets were used: tensile loading, tensile and compressive loading, and the first cyclic loading. The

corresponding stress-stretch curves are shown in Figure 4.8, where training data is visualized in green, testing data in orange, and the model predictions in solid blue lines. **Results.** It can be observed that the training on the first tensile loading already yielded a good fit to the cyclic data. However, a slight offset in the elastic behavior, the yield strength, and the hardening behavior was observed at larger stretches, particularly, in the compressive regime, where the transition into the plastic regime was slightly overshoot. Overall, the cyclic tension-compression behavior is met quite accurately when training only on a limited amount of data.

Expanding the training dataset to include tension-compression and tension-compression-tension loading paths enhanced the model's ability to predict data of unseen stretch ratios beyond the training regime. While including tension-compression data into the training set led to a better alignment of experimental data and model predictions, further increasing the load cycles did not result in a significant improvement. Although it generally revealed that more data points improve the numerical accuracy, the trade-off between computational costs and predictive performances suggests that training with tension-compression data alone already yielded sufficiently accurate and reliable model predictions, which aligns with previous observations.

4.5.5. Experimental data

To validate our model using experimental data, we examined the cyclic tension-compression tests on DP600 steel conducted by Cao et al. [2009]. Throughout the years, multiple material models have been derived to predict the material response of steels. We aim to use the proposed framework to automatically discover the constitutive equations from experimental data, thereby demonstrating its advantages.

The energy FFNs have been regularized using a ridge regularization factor of $L_2 = 0.0001$. We have trained on three training sets. The first one used only the tensile part of the first cyclic loading path, while the second one was expanded to include both tensile and compressive states from the first cycle, and the last one was extended to tension-compression-tension data. The results are visualized in Figure 4.9, where the data points used for training are increasing from the upper to the lower line. The data used for training is marked in green, data used for testing by orange markers, and the model prediction by solid blue lines.

Results. The experimental stress curves demonstrated pronounced nonlinear behavior. Training the model solely on tensile loadings resulted in a good fit within the training data range. In the predictions beyond the training data, the model exhibited a linear hardening behavior, which led to large discrepancies between experimental data and prediction, most notably in the compressive regime, failing to fully capture the nonlinearity observed in the

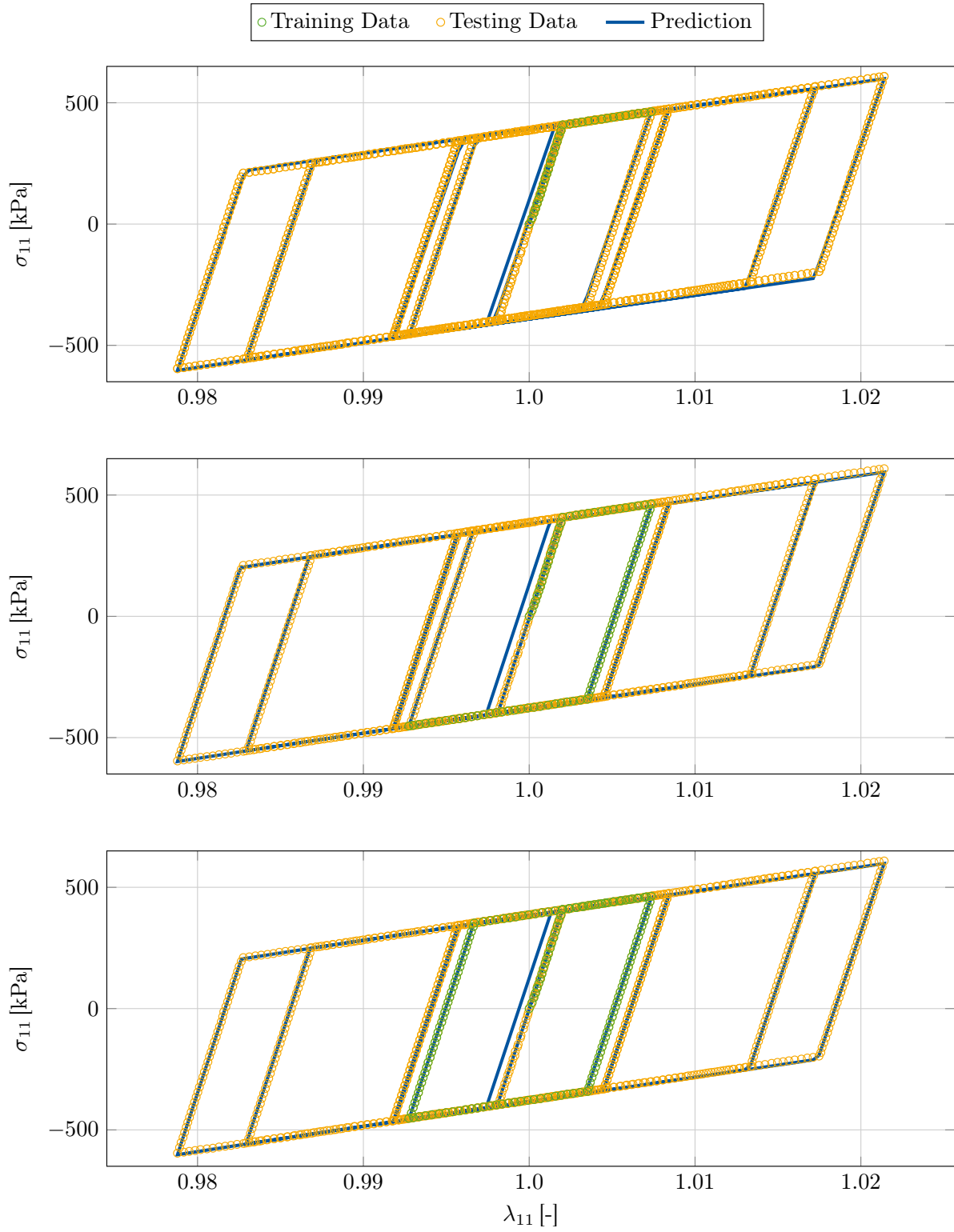


Figure 4.8.: Training on data taken from Liao and Luo [2025]. Stress-stretch curves for training on the first tensile loading (1st row), tensile and compressive loading (2nd row), and on the first cycle (3rd row). Data points used for training are shown as green markers and data points used for testing as orange markers. Model predictions are given as solid blue lines.

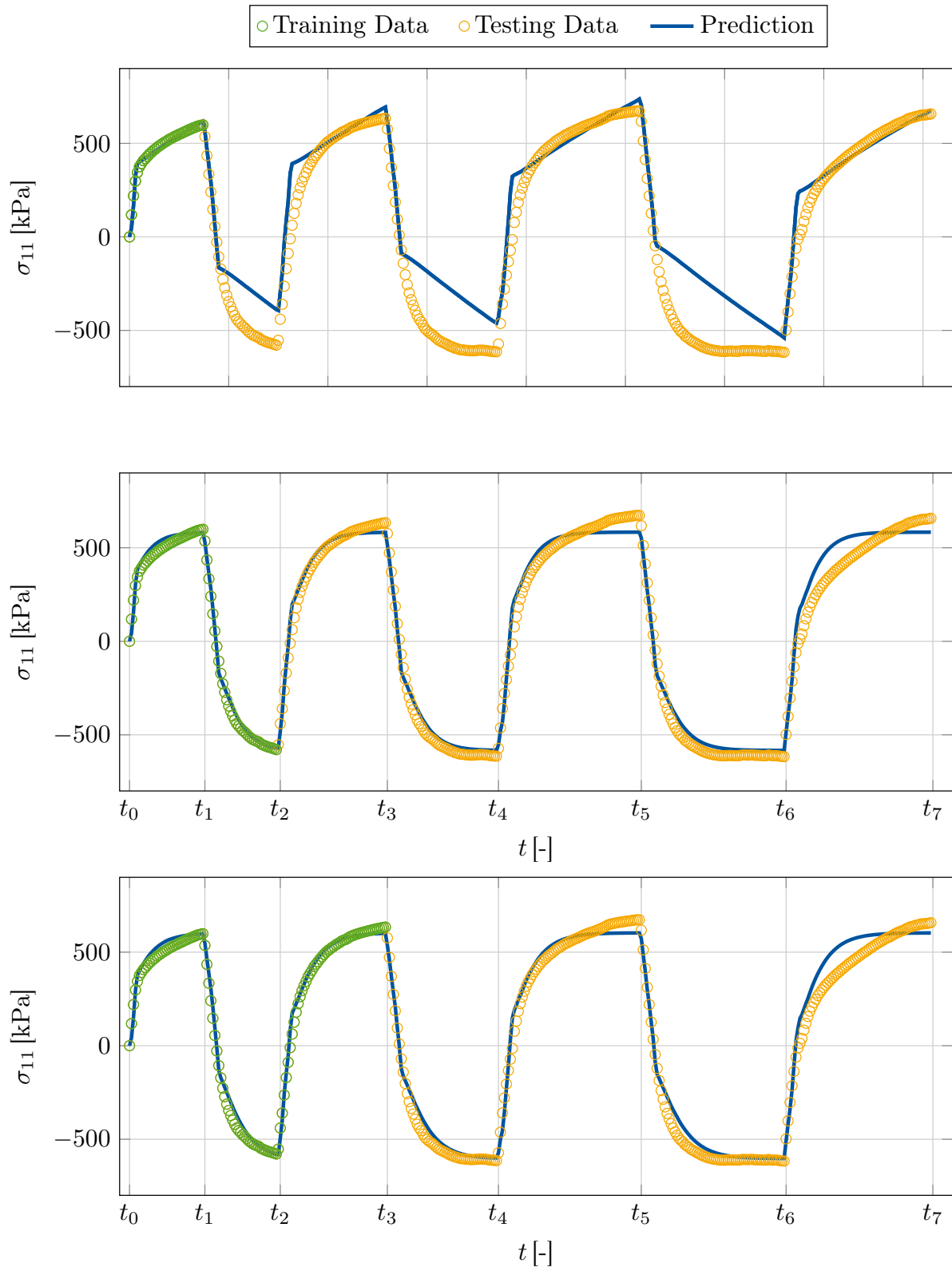


Figure 4.9.: Experimental data of DP600 steel taken from Cao et al. [2009]: Stress curves over time under cyclic tension-compression loading for training on the first tension loading (upper line), the first tension and compression loading (middle line), and tension-compression-tension loading (lower line). Data points used for training are shown as green markers and data points used for testing as orange markers. Model predictions are given as solid blue lines.

experimental data. This limitation revealed that the training data lacked sufficient information to fully capture the observed nonlinearity. Expanding the training dataset to include both tension and compression improved the model's performance, yielding a smoother transition from elastic to plastic behavior and enhancing alignment with the experimental data, as shown in the middle row in Figure 4.9. The fit to compressive behavior, in particular, became quite accurate. Despite these improvements, further loading revealed that offsets in predicted and experimental stresses persisted, particularly under tensile loading, indicating that the model still did not fully capture the hardening behavior. Extending the training data further to incorporate an additional tensile loading cycle – similar to the findings in Section 4.5.4 – resulted only in marginal improvements. Although slight improvements in the compressive regime are visible, the overall model performance did not improve qualitatively. Further advancements could potentially be achieved by incorporating isotropic hardening into the formulation. However, this was out of the scope of this work, since – in contrast to kinematic hardening – it introduces a strongly coupled but otherwise arbitrary hardening variable.

We can conclude that one cycle of tension and compression loads already provides sufficiently accurate results within our current modeling framework. Nonetheless, persistent discrepancies in the hardening regime during subsequent load cycles indicate that the current architecture might be too limited to capture the full complexity of nonlinear hardening, implying that a wider network architecture should be investigated in future works.

4.6. Discussion and limitations

Architecture. While our model successfully captured elasto-plastic material behavior with both linear and nonlinear kinematic hardening effects, several limitations have been observed. A nonlinear kinematic hardening model for steel was successfully identified. However, the prediction exhibited offsets in increased loading, likely due to isotropic hardening mechanisms not yet included in our framework. Future works should incorporate these to enhance accuracy and a wider set of mechanisms to be captured. Further, capturing more complex material behavior, such as described by asymmetric yield functions, was not possible due to the current architecture of the potentials. Although the presented network architecture promotes interpretability, it limits the accurate identification of more complex material behavior. Experimental data and multiscale simulations often reveal behaviors that require denser architectures capable of discovering formulations beyond those embedded in the present framework. The difficulty in representing asymmetric yield surfaces, therefore, underscores the need for a more flexible and expressive network design. Automated model discovery inherently involves a trade-off

between interpretability and predictive accuracy. The present framework can be considered near the interpretability-oriented end of this spectrum, coming with the risk of reduced accuracy in the model predictions. Nevertheless, the results obtained here still demonstrate high accuracy. For future works, a hybrid approach – positioned closer to the balance point between interpretability and accuracy – could yield results that maintain interpretable material models while achieving even greater predictive performance.

Moreover, the analysis of the Tschoegel yield criterion shows that diverse loading conditions are essential in the training data to capture complex yielding behavior. Classical uniaxial tension tests alone do not provide sufficient information to characterize the full yield surface or the evolution of inelastic stretches.

Numerical stabilization. During pre-studies prior to this work, we observed that training on the Cauchy stress rather than the second Piola-Kirchhoff stress was more stable, in particular for compression data. This effect was particularly pronounced when the initial model predictions were too stiff, as the absolute value of the second Piola-Kirchhoff stress further increases under compression, even for perfect plasticity, whereas the Cauchy stress remains constant. Consequently, we have adopted the Cauchy stress as the primary loss measure.

To enhance the training stability, we performed pretraining for the initial weights, leading to more efficient training. The initial weights for the pretraining were set randomly. Although one might argue that this procedure introduces a supervised training process, the overall model discovery remains automated, with a preprocessing step in advance. In future works, optimization of this preprocessing should be investigated. Without the pretraining step, we frequently encountered trainings where the weight optimization led to divergence (NaN values). This was circumvented by applying pre-processing, normalizing the stresses, and using gradient clipping. Including plasticity into the framework resulted in additional numerical complexity due to the Newton-Raphson iteration that required evaluation of the gradients of the residual to be minimized to solve for the Karush-Kuhn-Tucker conditions. This additional evaluation might have led to problems in the computation of the global gradients. Thus, a derivative-free iteration procedure should be investigated in future work.

Effects of regularization. Using the L_2 -regularization, we observed that using a smaller or equally large regularization factor for the potentials compared to the energies improved accuracy. This choice ensured that the weights of the energies would be penalized more severely, such that they were more likely to be changed to fit the experimental data instead of reducing the weights of the potentials. In contrast, larger regularization factors of the potentials would lead to reduced weights of the potentials. Given our network architecture, where the yield function is normalized (see Equation (4.18)), small initial weights could shift

the initial onset of plasticity outside the training data, resulting in an entirely elastic model prediction that attempts to approximate the nonlinear stress-strain curve.

4.7. Conclusion

In conclusion, we presented a thermodynamically consistent framework for elasto-plasticity in constitutive artificial neural networks. Dealing with finite strains, mechanisms of linear and nonlinear kinematic hardening were included using a multiplicative decomposition of both the total and the plastic part of the deformation gradient. The constitutive equations of the elastic, linear and nonlinear hardening behavior, as well as the yield surface, and the evolution equations of the inelastic stretches were described by feed-forward networks to be automatically discovered by obtaining physically reasonable models due to the design of the networks.

The performance of our model was thoroughly investigated using artificially generated data (incorporating both a symmetric yield function and one capturing tension-compression asymmetry), a comparison to other network approaches, and experimental data for steels. While the model generally predicted most material behaviors accurately, discrepancies were observed when evaluating the tension-compression yield surface and experimental data. These differences stem from limitations in the current model's design, specifically the potential architecture, its sparsity, and the absence of isotropic hardening. The sparse design limits the training process to identifying the best parameters for pre-defined custom models already included in the architecture. Consequently, this approach does not enable the pure discovery of entirely new material models, unlike methods that use dense network structures. Enhancing the network architecture is essential to capture a wider spectrum of material behaviors. Despite these shortcomings, the model preserves the interpretability characteristic of CANNs, automatically discovering transparent material models during training.

Adapting the architecture of the networks to a denser structure, as well as including isotropic hardening, will be subject to future work. Denser network structures then might require different regularization techniques that should be treated in a new study.

While numerical instabilities had been circumvented by a pretraining of initial weights, normalization of the stresses prior to the training, gradient clipping, and regularization, future works should investigate the implementation of gradient-free iteration processes to solve the local iteration scheme. However, this remained out of the scope of our present work, dealing with the basis of including plasticity into the general design of iCANNs.

4.8. Appendix

4.8.1. Impact of regularization techniques

Optimization and regularization of the weights has proven to be an important tool, as shown by McCulloch et al. [2024], Fuhg et al. [2024b], and Flaschel et al. [2025a]. Therefore, we explored the impact of different regularization techniques on the model performance. In addition to L_2 , we investigated L_1 - or Lasso-regularization, which promotes sparsity of the model parameters by enforcing them to zero and enhances interpretability further. Thus, the regularization term given in Equation (4.21) was extended to include the L_1 penalty by

$$\hat{\mathcal{R}} = \mathcal{R} + L_1 \sum_{j=1}^{n_w} |w_j|, \quad (4.29)$$

where L_1 is the Lasso-regularization factor. While our model was already designed following a sparse architecture for interpretability, using L_1 -regularization to further enforce sparsity was of interest as we wanted to investigate whether this might suppress plastic behavior, thereby prioritizing simplicity over accuracy. We applied L_1 -regularization of $L_1 = 0.0001$ and $L_1 = 0.01$ to the last layers of the FFNs to penalize redundant weights. Figure 4.10 compares these results (dashed purple lines for $L_1 = 0.0001$ and dotted gray lines for $L_1 = 0.01$) with experimental data (dots) and the predictions using L_2 -regularization, as described in Section 4.5.4 (solid blue lines). Comparing L_1 - and L_2 -regularization with a similar regularization factor, a notable difference only appeared at large stretch ratios when the model was trained exclusively on tensile data, where L_2 regularization provided a closer fit to the underlying data compared to L_1 . Aside from this region, only negligible effects were obtained. Increasing the regularization factor highlighted the promotion of sparsity, as all hardening effects were essentially removed by the regularization (see Table 4.5). While both elastic and plastic energies were subjected to regularization, only the plastic terms were eliminated. Applying strong regularization to all elastic energy weights would have resulted in stresses approaching zero, leading to a substantially increased mean squared error, and thus a higher overall loss, $\mathcal{L}$, even as the regularization term, $\hat{\mathcal{R}}$, itself diminished. This observation highlights how identical regularization strategies can affect elastic and plastic components differently within the loss function. The results presented in Figure 4.10 underscore the importance of considering a hyperparameter study to determine the optimal regularization parameters for any specific model, ensuring a balance between sparsity, accuracy, and interpretability.

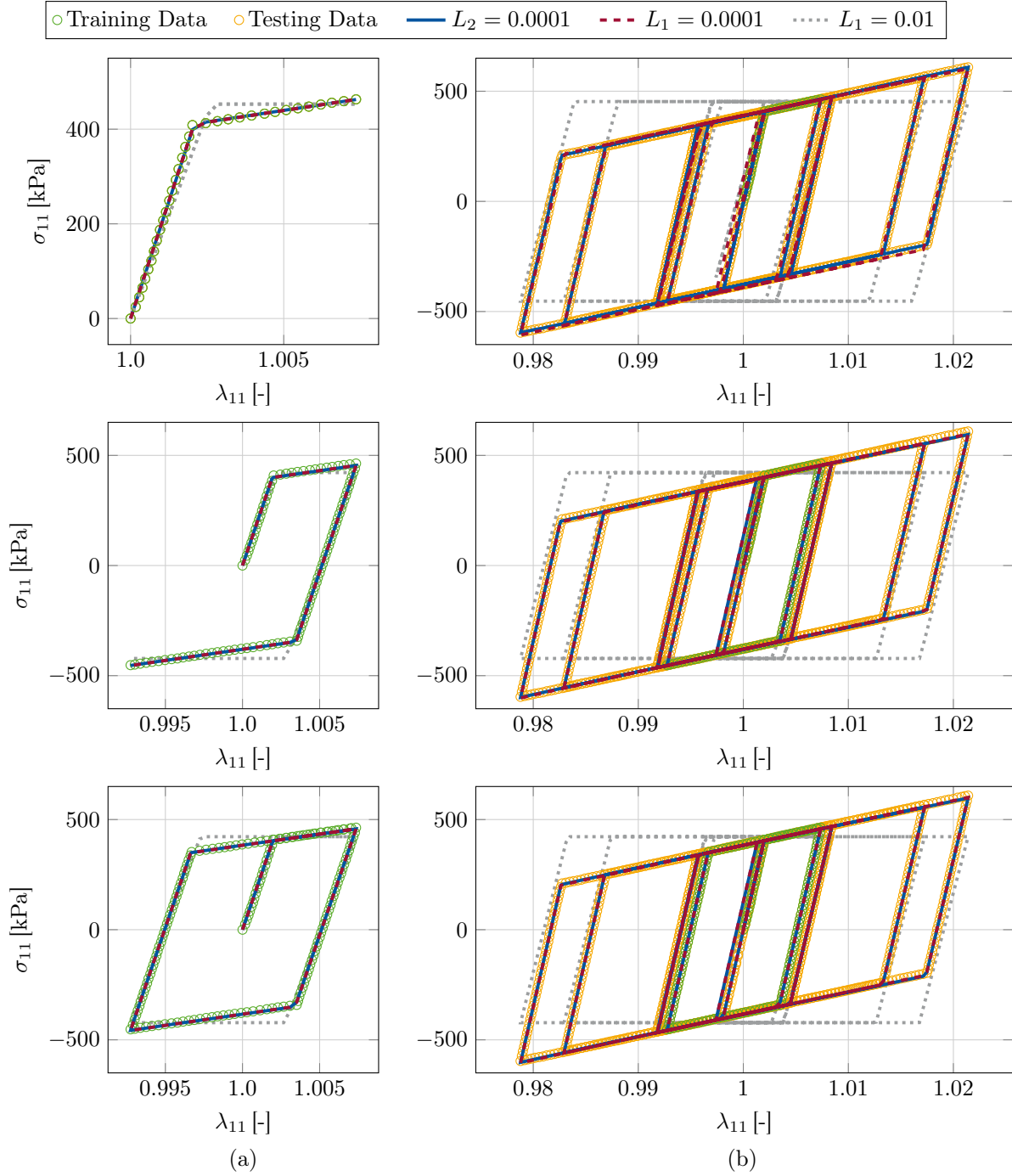


Figure 4.10.: Comparison of regularization schemes – Training on data taken from Liao and Luo [2025]. Stress-stretch curves for training on the first tensile loading (1st row), tensile and compressive loading (2nd row), and on the first cycle (3rd row). Training data (green marker) and model prediction for regularization with regularization of $L_2 = 0.0001$ (blue, solid), $L_1 = 0.0001$ (red, dashed), and $L_1 = 0.01$ (gray, dotted). Testing data (orange marker) is given on the right.

Table 4.5.: Comparison of the learned weights of the Helmholtz energy accounting for linear kinematic hardening (ψ_p) and nonlinear kinematic hardening (ψ_{pe}) for regularization with L_1 -regularization.

Trained on Regularization	UT		UT, UC		Cyc	
	L_2	L_1	L_2	L_1	L_2	L_1
$w_{1,1}^{\psi_p}$	1.412e+1	6.502e+0	1.125e+1	5.719e+0	1.074e+1	5.904e+0
$w_{1,2}^{\psi_p}$	1.937e+1	2.204e+1	1.356e+1	9.697e-1	1.408e+1	9.958e-1
$w_{2,1}^{\psi_p}$	2.738e-3	0	3.960e-3	0	3.528e-3	0
$w_{2,2}^{\psi_p}$	3.837e-2	0	4.392e-2	1.004e-5	3.846e-2	1.403e-5
$w_{2,3}^{\psi_p}$	7.008e-3	0	1.056e-2	0	9.851e-3	0
$w_{2,4}^{\psi_p}$	1.357e-1	1.322e-1	1.419e-1	2.111e-5	1.358e-1	0
$w_{1,1}^{\psi_{pe}}$	0	0	0	0	0	0
$w_{1,2}^{\psi_{pe}}$	0	0	0	0	0	0
$w_{2,1}^{\psi_{pe}}$	2.738e-3	0	3.960e-3	1.056e-2	3.528e-3	0
$w_{2,2}^{\psi_{pe}}$	3.837e-2	0	0	0	0	0
$w_{2,3}^{\psi_{pe}}$	7.008e-3	0	1.056e-2	0	9.851e-3	3.234e-6
$w_{2,4}^{\psi_{pe}}$	1.357e-1	0	0	0	0	0

5 | Article 4:

The mechanics and physics of tofu: understanding hydrated soft solids through feature networks

This article was published as:

Boes, B., Simon, J.-W., Holthusen, H., and Kuhl, E. (2026). The mechanics and physics of tofu: understanding hydrated soft solids through feature networks. *Journal of the Mechanics and Physics of Solids*, **213**, 106632. <https://doi.org/10.1016/j.jmps.2026.106632>.

Disclosure of the individual authors' contributions to the article:

B. Boes reviewed the relevant existing literature, conducted all experiments and trained the model presented by Holthusen et al. [2025b] supported by H. Holthusen. She interpreted the results and wrote the preliminary version of the article. All co-authors gave valuable conceptual advice, read the manuscript and gave suggestions for improvement. All authors approved the publication of the final version of the manuscript.

5.1. Abstract

Tofu remains one of the world’s most important plant-based foods, famous for its cultural legacy, nutritional benefits, and environmental sustainability. Made up of only two ingredient—soy beans and water—it has nourished people for centuries; yet, its rheology remains incompletely understood. Here, we use tofu as a minimal model system to study how composition governs the mechanics and physics of hydrated soft solids. We perform over a hundred controlled compression tests on silken, firm, and extrafirm tofu, and observe a strong nonlinearity, a pronounced viscoelasticity, and a more than ten-fold increase in stress with only six percent decrease in water content. We demonstrate that automated model discovery can identify inelastic constitutive laws that accurately capture these unique characteristics across multiple loading magnitudes and rates. The discovered models autonomously separate elastic and inelastic mechanisms and reveal that tofu elasticity depends primarily on the second isochoric invariant, while tofu inelasticity depends on a combination of the second and third deviatoric invariants. This functional form is universal across all three tofu types, but scales in magnitude with water content. Our new water-content-based feature network reveals that this correlation is not linear, as assumed by traditional bi-phasic theories, but rather highly non-linear and sensitive to the individual model terms. These results position tofu as a quantitative benchmark for nonlinear poroviscoelastic solids and demonstrate how physics-informed machine learning can uncover the constitutive structure of hydrated soft materials.

“Where soy meets water, tofu becomes matter.”

5.2. Introduction

Tofu feeds more than a billion people worldwide and has done so for over two thousand years [Shurtleff and Aoyagi, 1975]. It is one of the simplest and most influential plant-based foods in human history [Watanabe, 1997]. Beyond its cultural role, tofu promotes human health [Qin et al., 2022], contributes to planetary sustainability [Poore and Nemecek, 2018], and remains the cornerstone of a modern plant-based diet [Rizzo and Baroni, 2018]. Yet, the material itself is remarkably simple [Kohyama et al., 1995]: Soymilk and a mineral coagulant form a hydrated protein gel with only two essential ingredients, soybeans and water [Zheng et al., 2020]. Whether we make our own tofu at home or buy it in supermarkets, Figure 5.1 illustrates that its texture can range from silken to extrafirm, determined almost entirely by

its water content [Chen et al., 2023] and by the specifics of its coagulation process [Wang et al., 2018]. This simplicity makes tofu an ideal prototype material for testing, validating, and revealing the power of automated model-discovery for soft matter systems.

For more than half a century, texture profile analysis has defined the gold standard in food science to quantify the textural and sensory attributes of food [Friedman et al., 1963]. In texture profile analysis, mechanical double-compression tests simulate the behavior of chewing [Bourne, 1968], we can easily extract its parameters such as stiffness, hardness, cohesiveness, springiness, resilience, and chewiness from the resulting force–time curves [Dunne et al., 2025]. Although texture profile analysis continues to be widely used [Paredes et al., 2022], the lack of standardized testing protocols limits comparability across the different studies and labs [Schreuders et al., 2021]. Moreover, the interpretation of texture profile analysis curves remains largely rooted in food science rather than in materials science, even though these tests inherently capture valuable mechanical information.

Mechanical testing allows us to interpret texture through the lens of mechanics and to extract fundamental material properties directly from experimental data [St. Pierre et al., 2024]. We can quantify material parameters such as stiffness, viscosity, and relaxation time—commonly used in mechanics—from tension, compression, and shear tests to reveal the underlying macrostructural behavior and correlate it to the microstructural architecture [Budday et al., 2020]. Grounding texture evaluation in the principles of material modeling enables a more rigorous and interpretable mechanistic analysis [Vervenne et al., 2025a]. This raises the question: Can we characterize the behavior of food through robust constitutive models that reliably reproduce experimental data and confidently predict the effects of different loading conditions?

Food materials form an emerging class of complex, heterogeneous systems that challenge conventional modeling techniques [Datta et al., 2025]. Their nonlinear, often rate-dependent response calls for new approaches that are capable of identifying the governing equations directly from experimental data [Kuhl, 2025]. Recent advances in data-driven mechanics have introduced frameworks that automatically discover constitutive relationships [Linka and Kuhl, 2023] without pre-selecting the specific functional form of the model [Holthusen et al., 2024b]. Instead of repeatedly fitting and refining candidate models, automated model discovery infers the governing equations in a self-consistent and physics-informed manner [Flaschel et al., 2021]. This motivates us to ask: Can we adopt the concept of automated model discovery to discover the models and parameters that best characterize tofu as a minimal benchmark system for soft, structurally engineered foods?

A wide range of data-driven frameworks has emerged in recent years, including unsuper-

vised learning [Flaschel et al., 2021] as well as mechanics-informed [As’ad and Farhat, 2023], physics-constrained [Kalina et al., 2023b], physics-augmented [Kalina et al., 2023a; Klein et al., 2023], thermodynamics-based [Masi et al., 2021], and input convex neural networks [Amos et al., 2017]. Among these, the framework of constitutive artificial neural networks has proven particularly effective because it hardwires continuum mechanical principles directly into the network architecture [Linka et al., 2021]. This design ensures that the resulting models remain both physically sound and interpretable [Linka and Kuhl, 2023]. The sparse network structure balances accuracy and interpretability and allows the simultaneous discovery of concise yet descriptive constitutive models and their model parameters.

Automated model discovery has successfully identified constitutive models across diverse applications, including soft biological tissues such as skin [Linka et al., 2023b], arteries [Ver-venne et al., 2025a], the heart [Martonová et al., 2024], and the brain [Hou et al., 2024; Linka et al., 2023a], and, most recently, plant-based and animal meats [Pierre et al., 2023]. Most studies have focused on discovering models for the elastic behavior; yet, an accurate representation of realistic material responses—especially at larger loading magnitudes and rates—often requires the inclusion of inelastic effects. Recent extensions of the framework have incorporated viscoelasticity through Prony-series representations [Abdolazizi et al., 2024; Wang et al., 2023]. Building on this idea, inelastic constitutive artificial neural networks, extend the formulation by discovering both, the Helmholtz free energy for the elastic behavior and a dual potential for the inelastic behavior [Holthusen et al., 2024b]. This approach preserves thermodynamic consistency and interpretability, while expanding the range of representable material responses. The general framework of inelastic constitutive neural networks has since been applied to viscoelasticity [Holthusen et al., 2024a], growth modeling [Holthusen et al., 2025a], and plasticity [Boes et al., 2025; van der Velden et al., 2025]. Importantly, inelastic constitutive neural networks autonomously discover whether inelastic effects are necessary, and reduce unnecessary model complexity when the data do not justify it [Holthusen et al., 2025b].

Motivated by the growing interest in meat alternatives, this study investigates the mechanics and physics of tofu by replacing the traditional texture profile analysis with classical mechanical testing. We perform compression experiments on multiple tofu types with varying water content and analyze the resulting force–time curves to explore the correlation between structure and function. First, we adopt an inelastic constitutive neural network to discover individual constitutive models for each tofu type and identify the dominant mechanisms that govern the material behavior. Then, we postulate that the water content is the single distinguishing feature

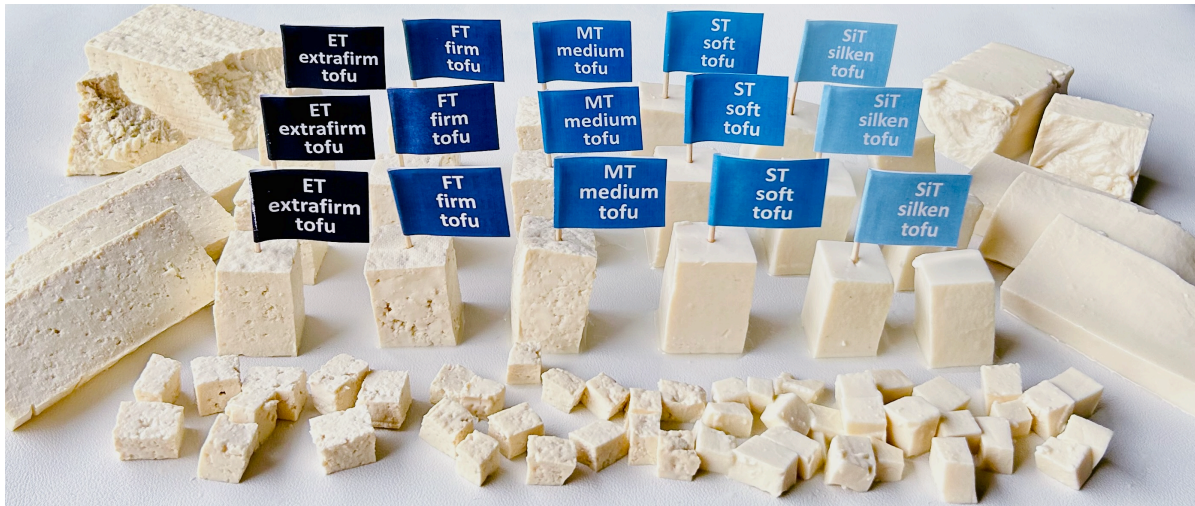


Figure 5.1.: **Five different types of tofu.** Extra-firm, firm, medium, soft, and silken tofu illustrate how small changes in composition and processing create large differences in mechanical behavior. Tofu combines soy beans with water to form a biphasic protein gel whose stiffness, porosity, and fracture characteristics arise directly from its water content and its coagulation pathway. Whether we make our own tofu at home or obtain it through commercial retail distribution, we encounter the same continuum of textures, from a dense cohesive protein matrix to a nearly fluid-like gel.

across all individual models. We thus design a second neural network that takes the water content as input and outputs a feature vector that modulates the inelastic constitutive neural network. This architecture enables a flexible and unbiased mapping between composition and mechanical behavior. It results in a single unified model that provides predictive insight into how changes in formulation—such as water or protein content—affect the texture of tofu, or more generally, of any alternative protein product.

Section 5.3 describes the mechanical testing. Sections 5.4 and 5.5 introduce the physics-based model and the neural network model. Section 5.6 summarizes our results and Section 5.7 discusses our main findings and outlines future directions.

5.3. Mechanical testing

We test a total of $n = 108$ samples with three different water contents across three different loading magnitudes at five different rates. Specifically, we test three types of tofu, silken, firm, and extrafirm (Azumaya, Nasoya Foods, Ayer, Massachusetts). Table 5.1 summarizes their serving sizes, calories, and macronutrient compositions per serving according to their

individual nutrition fact labels. We estimate the water mass as the difference between the serving size and the listed masses of fat, carbs, protein, and ash. We assume that tofu fat mainly

Table 5.1.: **Nutrition facts of silken, firm, and extrafirm tofu.** Serving size, calories per serving, macronutrient composition per serving, and fluid mass and volume fractions for the three different tofu types.

tofu type	serving size	calories per serving	composition per serving					fluid mass fraction	fluid volume fraction
			fat	carbs	protein	ash	water		
silken	85g	40cal	2g	2g	4g	1g	76g	0.89	0.92
firm	85g	70cal	4g	2g	7g	1g	71g	0.84	0.87
extrafirm	85g	80cal	4g	2g	8g	1g	70g	0.82	0.86

presents as lipid droplets that do not move freely under mechanical loading, and conclude that the fluid phase of tofu is made up exclusively of the freely moving water. To calculate the mass fraction of the fluid w_f , we divide the water mass, $m_f = \{76\text{g}, 71\text{g}, 70\text{g}\}$, by the total mass per serving of 85g, resulting in $w_f = \{0.89, 0.84, 0.82\}$ for the three tofu types. To calculate the volume fraction of the fluid $\nu_f = V_f/[V_f + V_s]$, we divide the volume of the fluid, $V_f = w_f/\rho_f$, by the total volume of fluid and solid, $V_f + V_s$, with $V_s = [1 - w_f]/\rho_s$, assuming fluid and solid densities of $\rho_f = 1.0\text{ g/ml}$ and $\rho_s = 1.3\text{ g/ml}$, resulting in $\nu_f = \{0.92, 0.87, 0.86\}$ for the three tofu types. For each tofu type, we perform four sets of tests, three triple-compression tests at different rates, and one double-compression test. For each set, we test at least $n = 8$ samples.

Sample preparation. To prepare our samples, we first cut the tofu into 10 mm thick blocks and then extract cylindrical specimens of 8 mm diameter using biopsy punches. Figure 5.2, left, illustrates our sample preparation. To prevent the samples from dehydration, we store them between wetted tissues and test them immediately after preparation.

Sample testing. We perform all mechanical tests using an HR20 Discovery Hybrid Rheometer (TA Instruments, New Castle, Delaware). Figure 5.2, right, illustrates a representative sample inside a cylindrical chamber, mounted in the rheometer, to perform compression tests with controlled axial displacements and constrained lateral deformation. Before testing, we calibrate the rheometer, tare the force sensor, and adjust the initial gap between the loading nod and the specimen to ensure contact without exceeding a force of 0.05 N. All experiments follow a displacement-controlled protocol with four distinct loading programs: three triple-compression tests that serve as training data and one double-compression test that serves as test data. During triple compression in Figure 5.2 middle left, we compress each sample three times to a stretch of $\lambda = 0.7$, hold for 0.5 s, unload to $\lambda = 0.9$, and hold for another 0.5 s, at slow, medium, and fast stretch rates of $\dot{\lambda} = \{0.2/\text{s}, 0.4/\text{s}, 0.6/\text{s}\}$. During double compression in Figure 5.2

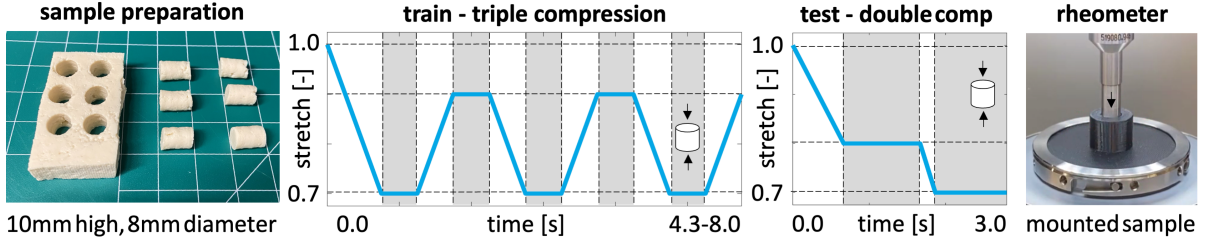


Figure 5.2.: **Mechanical testing.** Preparation of cylindrical samples; triple compression test for training; double compression test for testing; and rheometer with mounted sample. During triple compression, we load and unload between $\lambda = 0.7$ and $\lambda = 0.9$ at slow, medium, and fast stretch rates of $\dot{\lambda} = \{0.2/\text{s}, 0.4/\text{s}, 0.6/\text{s}\}$ and hold for 0.5s in between. During double compression, we load to $\lambda = 0.8$ at $\dot{\lambda} = 0.3/\text{s}$, hold for 1.0s, load to $\lambda = 0.7$ at $\dot{\lambda} = 0.5/\text{s}$, and hold for 1.0s. White regions represent loading and unloading, grey regions represent holding.

middle right, we compress each specimen first to a stretch of $\lambda = 0.8$ at a rate of $\dot{\lambda} = 0.3/\text{s}$, hold for 1.0 s, further compress to $\lambda = 0.7$ at a rate of $\dot{\lambda} = 0.5/\text{s}$, and hold for another 1.0 s.

Data processing. We record the gap l and the force F throughout the loading history and smoothen both with a Gaussian filter using the *imgaussfilt* function in Matlab R2022b [The MathWorks, Inc., 2022] with a filter width of $\sigma_{\text{gauss}} = \{8, 3, 3\}$ for silken, firm, and extrafirm tofu. We calculate and report the axial stretch $\lambda = l/L$ as the ratio between the current gap l and the initial gap L and the axial second Piola Kirchhoff stress $S = F/(A\lambda)$ as the recorded force F divided by the initial cross section area $A = \pi \cdot (4 \text{ mm})^2$ and the stretch λ .

5.4. Physics-based model

Discovering a physically sound constitutive model that explains the experimental data requires a thermodynamically consistent framework. In this section, we outline the continuum equations that motivate the design of our neural network architecture.

Kinematics. To characterize the experimentally observed rate dependence and nonlinear relaxation, we adopt a multiplicative decomposition of the deformation gradient $\mathbf{F}$ into elastic and inelastic parts, $\mathbf{F}_e$ and $\mathbf{F}_i$,

$$\mathbf{F} = \mathbf{F}_e \mathbf{F}_i. \quad (5.1)$$

As deformation measures, we introduce the right Cauchy Green deformation tensor $\mathbf{C}$ and its elastic and inelastic counterparts, $\mathbf{C}_e$ and $\mathbf{C}_i$,

$$\mathbf{C} = \mathbf{F}^T \mathbf{F} \quad \text{and} \quad \mathbf{C}_e = \mathbf{F}_e^T \mathbf{F}_e = \mathbf{F}_i^T \mathbf{C} \mathbf{F}_i^{-1} \quad \text{and} \quad \mathbf{C}_i = \mathbf{F}_i^T \mathbf{F}_i. \quad (5.2)$$

To characterize the isotropic volume-preserving and volume-changing behavior of tofu, we introduce the isochoric invariants $\tilde{I}_1^C = I_1^C/(I_3^C)^{1/3}$ and $\tilde{I}_2^C = I_2^C/(I_3^C)^{2/3}$ in terms of the classical invariants $I_1^C = \text{tr}(\mathbf{C})$ and $I_2^C = \text{tr}(\text{cof}(\mathbf{C})) = \frac{1}{2}[(\text{tr}(\mathbf{C}))^2 - \text{tr}(\mathbf{C}^2)]$, and the volumetric invariant $I_3^C = \det(\mathbf{C})$ of the right Cauchy Green tensor $\mathbf{C}$,

$$\tilde{I}_1^C = \text{tr}(\mathbf{C})/(I_3^C)^{1/3} \quad \text{and} \quad \tilde{I}_2^C = \text{tr}(\text{cof}(\mathbf{C}))/(I_3^C)^{2/3} \quad \text{and} \quad I_3^C = \det(\mathbf{C}), \quad (5.3)$$

where $\text{tr}(\circ) = (\circ) : \mathbf{I}$ denotes the trace, $\text{cof}(\circ) = I_3^{(\circ)}(\circ)^{-\text{T}}$ the cofactor, $\text{tr}(\text{cof}(\circ)) = \frac{1}{2}[(\text{tr}(\mathbf{C}_e))^2 - \text{tr}(\mathbf{C}_e^2)]$ the trace of the cofactor, and $\det(\circ)$ the determinant of the second-order tensor $(\circ)$. The first invariant $\tilde{I}_1^C$ measures the isochoric stretch, the second invariant $\tilde{I}_2^C$ measures the isochoric area change, and the third invariant I_3^C measures the volume change. To characterize the elastic behavior only, we introduce the elastic isochoric invariants $\tilde{I}_1^{C_e} = I_1^{C_e}/(I_3^{C_e})^{1/3}$ and $\tilde{I}_2^{C_e} = I_2^{C_e}/(I_3^{C_e})^{2/3}$ in terms of the classical elastic invariants $I_1^{C_e} = \text{tr}(\mathbf{C}_e)$ and $I_2^{C_e} = \text{tr}(\text{cof}(\mathbf{C}_e))$, and the volumetric invariant $I_3^{C_e} = \det(\mathbf{C}_e)$, of the elastic right Cauchy Green tensor $\mathbf{C}_e$,

$$\tilde{I}_1^{C_e} = \text{tr}(\mathbf{C}_e)/(I_3^{C_e})^{1/3} \quad \text{and} \quad \tilde{I}_2^{C_e} = \text{tr}(\text{cof}(\mathbf{C}_e))/(I_3^{C_e})^{2/3} \quad \text{and} \quad I_3^{C_e} = \det(\mathbf{C}_e). \quad (5.4)$$

Elastic potential. We model tofu as an isotropic, inelastic solid and express its total Helmholtz free energy ψ as the sum of an equilibrium contribution ψ^{eq} for the elastic part that depends on the right Cauchy Green tensor $\mathbf{C}$, or, alternatively, on its invariants $\tilde{I}_1^C, \tilde{I}_2^C, I_3^C$, and a non-equilibrium contribution ψ^{neq} for the inelastic part that depends on the elastic right Cauchy Green tensor $\mathbf{C}_e$, or on its invariants $\tilde{I}_1^{C_e}, \tilde{I}_2^{C_e}, I_3^{C_e}$,

$$\psi = \psi^{\text{eq}}(\mathbf{C}) + \psi^{\text{neq}}(\mathbf{C}_e) = \psi^{\text{eq}}(\tilde{I}_1^C, \tilde{I}_2^C, I_3^C) + \psi^{\text{neq}}(\tilde{I}_1^{C_e}, \tilde{I}_2^{C_e}, I_3^{C_e}). \quad (5.5)$$

To ensure thermodynamic consistency, we evaluate the Clausius–Planck inequality, $\mathcal{D} := \mathbf{S} : \frac{1}{2}\dot{\mathbf{C}} - \dot{\psi} \geq 0$, where $\mathbf{S}$ is the second Piola Kirchhoff stress and the dot denotes the material time derivative,

$$\mathcal{D} := \left[\mathbf{S} - 2 \frac{\partial \psi^{\text{eq}}}{\partial \mathbf{C}} - 2 \mathbf{F}_i^{-1} \frac{\partial \psi^{\text{neq}}}{\partial \mathbf{C}_e} \mathbf{F}_i^{-\text{T}} \right] : \frac{1}{2}\dot{\mathbf{C}} + 2 \mathbf{C}_e \frac{\partial \psi^{\text{neq}}}{\partial \mathbf{C}_e} : \mathbf{D}_i \geq 0, \quad (5.6)$$

where $\mathbf{D}_i = \text{sym}(\dot{\mathbf{F}}_i \mathbf{F}_i^{-1})$ is the inelastic rate of deformation tensor. From the Clausius–Planck inequality (5.6), we obtain the definitions of the second Piola Kirchhoff stress $\mathbf{S} := 2 \partial \psi / \partial \mathbf{C}$,

$$\mathbf{S} = 2 \frac{\partial \psi^{\text{eq}}}{\partial \mathbf{C}} + 2 \mathbf{F}_i^{-1} \frac{\partial \psi^{\text{neq}}}{\partial \mathbf{C}_e} \mathbf{F}_i^{-\text{T}} = \mathbf{S}^{\text{eq}} + \mathbf{S}^{\text{neq}}, \quad (5.7)$$

its equilibrium part $\mathbf{S}^{\text{eq}} := 2 \partial \psi^{\text{eq}} / \partial \mathbf{C}$,

$$\begin{aligned} \mathbf{S}^{\text{eq}} = & 2 \frac{\partial \psi^{\text{eq}}}{\partial \tilde{I}_1^{\mathbf{C}}} [\mathbf{I} - \frac{1}{3} I_1^{\mathbf{C}} \mathbf{C}^{-1}] / (I_3^{\mathbf{C}})^{1/3} + 2 \frac{\partial \psi^{\text{eq}}}{\partial \tilde{I}_2^{\mathbf{C}}} [I_1^{\mathbf{C}} \mathbf{I} - \mathbf{C} - \frac{2}{3} I_2^{\mathbf{C}} \mathbf{C}^{-1}] / (I_3^{\mathbf{C}})^{2/3} \\ & + 2 \frac{\partial \psi^{\text{eq}}}{\partial \tilde{I}_3^{\mathbf{C}}} I_3^{\mathbf{C}} \mathbf{C}^{-1}, \end{aligned} \quad (5.8)$$

and its non-equilibrium part $\mathbf{S}^{\text{neq}} := 2 \mathbf{F}_i^{-1} \partial \psi^{\text{neq}} / \partial \mathbf{C}_e \mathbf{F}_i^{-\text{T}}$,

$$\begin{aligned} \mathbf{S}^{\text{neq}} = & 2 \frac{\partial \psi^{\text{neq}}}{\partial \tilde{I}_1^{\mathbf{C}_e}} [\mathbf{C}_i^{-1} - \frac{1}{3} I_1^{\mathbf{C}_e} \mathbf{C}^{-1}] / (I_3^{\mathbf{C}_e})^{1/3} + 2 \frac{\partial \psi^{\text{neq}}}{\partial \tilde{I}_2^{\mathbf{C}_e}} [I_1^{\mathbf{C}_e} \mathbf{C}_i^{-1} - \mathbf{C}_i^{-1} \mathbf{C} \mathbf{C}_i^{-1} \\ & - \frac{2}{3} I_2^{\mathbf{C}_e} \mathbf{C}^{-1}] / (I_3^{\mathbf{C}_e})^{2/3} + 2 \frac{\partial \psi^{\text{neq}}}{\partial \tilde{I}_3^{\mathbf{C}_e}} I_3^{\mathbf{C}_e} \mathbf{C}^{-1}. \end{aligned} \quad (5.9)$$

In Section 5.5, we reverse-engineer two constitutive neural networks, one for the equilibrium potential ψ^{eq} and one for the non-equilibrium potential ψ^{neq} , from which we derive the equilibrium and non-equilibrium stresses, $\mathbf{S}^{\text{eq}}$ and $\mathbf{S}^{\text{neq}}$, according to equations (5.8) and (5.9).

Dissipative potential. From the Clausius-Planck inequality (5.6), we motivate the Mandel stress $\boldsymbol{\Sigma}$ as the driving force for dissipative processes, with $\boldsymbol{\Sigma} := 2 \mathbf{C}_e \partial \psi^{\text{neq}} / \partial \mathbf{C}_e$,

$$\begin{aligned} \boldsymbol{\Sigma} = & 2 \frac{\partial \psi^{\text{neq}}}{\partial \tilde{I}_1^{\mathbf{C}_e}} [\mathbf{C}_e - \frac{1}{3} I_1^{\mathbf{C}_e} \mathbf{I}] / (I_3^{\mathbf{C}_e})^{1/3} + 2 \frac{\partial \psi^{\text{neq}}}{\partial \tilde{I}_2^{\mathbf{C}_e}} [I_1^{\mathbf{C}_e} \mathbf{C}_e - \mathbf{C}_e^2 - \frac{2}{3} I_2^{\mathbf{C}_e} \mathbf{I}] / (I_3^{\mathbf{C}_e})^{2/3} \\ & + 2 \frac{\partial \psi^{\text{neq}}}{\partial \tilde{I}_3^{\mathbf{C}_e}} I_3^{\mathbf{C}_e} \mathbf{I}. \end{aligned} \quad (5.10)$$

To characterize the dissipative volume-changing and volume-preserving behavior of tofu, we introduce the volumetric invariant $I_1^{\boldsymbol{\Sigma}}$ and the deviatoric invariants $J_2^{\boldsymbol{\Sigma}}$ and $J_3^{\boldsymbol{\Sigma}}$ of the Mandel stress $\boldsymbol{\Sigma}$,

$$I_1^{\boldsymbol{\Sigma}} = \text{tr}(\boldsymbol{\Sigma}) \quad \text{and} \quad J_2^{\boldsymbol{\Sigma}} = \frac{1}{2} \text{tr}(\text{dev}(\boldsymbol{\Sigma})^2) \quad \text{and} \quad J_3^{\boldsymbol{\Sigma}} = \frac{1}{3} \text{tr}(\text{dev}(\boldsymbol{\Sigma})^3), \quad (5.11)$$

with derivatives $\partial I_1^{\boldsymbol{\Sigma}} / \partial \boldsymbol{\Sigma} = \mathbf{I}$ and $\partial J_2^{\boldsymbol{\Sigma}} / \partial \boldsymbol{\Sigma} = \text{dev}(\boldsymbol{\Sigma})$ and $\partial J_3^{\boldsymbol{\Sigma}} / \partial \boldsymbol{\Sigma} = \text{dev}(\text{dev}(\boldsymbol{\Sigma})^2)$, where $\text{tr}(\circ) = (\circ) : \mathbf{I}$ denotes the trace and $\text{dev}(\circ) = (\circ) - \frac{1}{3} I_1^{\circ}(\circ)$ denotes the deviator of a second-order tensor $(\circ)$. The first invariant $I_1^{\boldsymbol{\Sigma}}$ measures the hydrostatic stress, the second invariant $J_2^{\boldsymbol{\Sigma}}$ measures the magnitude of shear, and the third invariant $J_3^{\boldsymbol{\Sigma}}$ measures the magnitude and mode of shear. To characterize the evolution of the inelastic deformation, we introduce the dissipative potential ϕ that depends on the Mandel stress $\boldsymbol{\Sigma}$ or, alternatively, on its invariants $I_1^{\boldsymbol{\Sigma}}, J_2^{\boldsymbol{\Sigma}}, J_3^{\boldsymbol{\Sigma}}$,

$$\phi = \phi(\boldsymbol{\Sigma}) = \phi(I_1^{\boldsymbol{\Sigma}}, J_2^{\boldsymbol{\Sigma}}, J_3^{\boldsymbol{\Sigma}}). \quad (5.12)$$

To a priori satisfy the reduced Clausius–Planck inequality,

$$\mathcal{D}_{\text{red}} = \boldsymbol{\Sigma} : \mathbf{D}_i \geq 0 \quad (5.13)$$

we postulate that the dissipative potential ϕ is convex, non-negative, and zero at the origin and define the inelastic rate of deformation tensor as $\mathbf{D}_i := \partial\phi/\partial\boldsymbol{\Sigma}$,

$$\mathbf{D}_i = \frac{\partial\phi}{\partial I_1^\Sigma} \mathbf{I} + \frac{\partial\phi}{\partial J_2^\Sigma} \text{dev}(\boldsymbol{\Sigma}) + \frac{\partial\phi}{\partial J_3^\Sigma} \text{dev}(\text{dev}(\boldsymbol{\Sigma})^2). \quad (5.14)$$

The inelastic rate of deformation tensor $\mathbf{D}_i$ introduces the evolution of the inelastic right Cauchy Green tensor, $\dot{\mathbf{C}}_i = 2 \mathbf{F}_i^T \mathbf{D}_i \mathbf{F}_i$, which we now discretize in time to calculate the inelastic right Cauchy Green tensor $\mathbf{C}_i$ and its elastic counterpart $\mathbf{C}_e$.

Time discretization. To advance the governing equations in time, from the previous time point t^n with known deformation gradient $\mathbf{F}^n$, inelastic tensor $\mathbf{F}_i^n$, right Cauchy Green tensor $\mathbf{C}_i^n$, and rate of deformation tensor $\mathbf{D}_i^n$, to the new time point $t^{n+1} = t^n + \Delta t$ with time increment Δt , we adopt an explicit time integration scheme using the exponential update of the inelastic right Cauchy Green tensor,

$$\mathbf{C}_i^{n+1} = (\mathbf{F}_i^n)^T \exp(2 \Delta t \mathbf{D}_i^n) \mathbf{F}_i^n \quad \text{with} \quad \mathbf{F}_i^n = +(\mathbf{C}_i^n)^{1/2}. \quad (5.15)$$

In Section 5.5, we design a neural network for the dissipative potential ϕ from which we derive the inelastic rate of deformation tensor $\mathbf{D}_i$, the inelastic right Cauchy Green tensor $\mathbf{C}_i$, and the inelastic part of the deformation gradient $\mathbf{F}_i$ at the new time point t^{n+1} .

Features. In many engineering applications, additional state variables or fields such as the density, temperature, or water content may affect the material behavior. To account for these effects, we extend the constitutive framework by introducing an input feature vector $\boldsymbol{\alpha}$ that can modify the elastic and dissipative potentials,

$$\psi = \psi^{\text{eq}}(\mathbf{C}, \boldsymbol{\alpha}) + \psi^{\text{neq}}(\mathbf{C}_e, \boldsymbol{\alpha}) \quad \text{and} \quad \phi = \phi(\boldsymbol{\Sigma}, \boldsymbol{\alpha}). \quad (5.16)$$

Here, we acknowledge the biphasic nature of tofu and propose a single scalar-valued input feature, the water content $\alpha = n^F$. In Section 5.5, we design a feature network that we couple to the networks for the equilibrium potential ψ^{eq} , the non-equilibrium potential ψ^{neq} , and the dissipative potential ϕ to explore the dependencies of tofu on the water content.

Kinematics of finite inelastic confined cylindrical compression. Our experiment considers

a cylindrical sample under confined compression in the z -direction. Its principal stretches in cylindrical coordinates, $\lambda_r = 1$, $\lambda_\theta = 1$, $\lambda_z = \lambda$, define the total deformation gradient $\mathbf{F}$ that we multiplicatively decompose into elastic and inelastic parts, $\mathbf{F}_e$ and $\mathbf{F}_i$,

$$\mathbf{F} = \text{diag} \{ 1, 1, \lambda \} = \mathbf{F}_e \mathbf{F}_i, \quad (5.17)$$

with the Jacobian, $J = \det(\mathbf{F}) = \lambda_r \lambda_\theta \lambda_z = \lambda$, defining the volumetric change of the sample. We assume that the inelastic deformation is isotropic and volume preserving, $J_i = \det(\mathbf{F}_i) = 1$, and that all volumetric changes are associated with the elastic deformation only, $J_e = \det(\mathbf{F}_e) = \lambda$ [Holzapfel, 2002; Reese and Govindjee, 1998; Simo, 1987], such that elastic and inelastic parts of the deformation gradient are

$$\mathbf{F}_e = \text{diag} \{ \lambda_i^{1/2}, \lambda_i^{1/2}, \lambda/\lambda_i \} \quad \text{and} \quad \mathbf{F}_i = \text{diag} \{ \lambda_i^{-1/2}, \lambda_i^{-1/2}, \lambda_i \}. \quad (5.18)$$

These explicit definitions introduce the right Cauchy Green tensor $\mathbf{C}$ and its elastic counterpart $\mathbf{C}_e$,

$$\mathbf{C} = \mathbf{F}^t \mathbf{F} = \text{diag} \{ 1, 1, \lambda^2 \} \quad \text{and} \quad \mathbf{C}_e = \mathbf{F}_e^t \mathbf{F}_e = \text{diag} \{ \lambda_i, \lambda_i, \lambda^2/\lambda_i^2 \} \quad (5.19)$$

with invariants

$$\begin{aligned} I_1^{\mathbf{C}} &= 2 + \lambda^2 & I_2^{\mathbf{C}} &= 1 + 2\lambda^2 & I_3^{\mathbf{C}} &= \lambda^2 \\ \tilde{I}_1^{\mathbf{C}} &= 2/\lambda^{2/3} + \lambda^{4/3} & \tilde{I}_2^{\mathbf{C}} &= 1/\lambda^{4/3} + 2\lambda^{2/3} & I_3^{\mathbf{C}} &= \lambda^2 \\ \tilde{I}_1^{\mathbf{C}_e} &= 2\lambda_i/\lambda^{2/3} + \lambda^{4/3}/\lambda_i^2 & \tilde{I}_2^{\mathbf{C}_e} &= 2\lambda^{2/3}/\lambda_i + \lambda_i^2/\lambda^{4/3} & I_3^{\mathbf{C}_e} &= \lambda^2. \end{aligned} \quad (5.20)$$

We note that, consistent with axisymmetric confined compression, all tensors are diagonal in the principal directions $\{r, \theta, z\}$, i.e., there are no shear components, and the radial and circumferential components are identical since the deformation is axisymmetric. For this special case of inelastic confined cylindrical compression, we can explicitly calculate the equilibrium stress in the axial direction,

$$S^{\text{eq}} = \frac{4}{3} \left[1 - \frac{1}{\lambda^2} \right] \left[\frac{1}{\lambda^{2/3}} \right] \frac{\partial \psi^{\text{eq}}}{\partial \tilde{I}_1^{\mathbf{C}}} + \frac{4}{3} \left[1 - \frac{1}{\lambda^2} \right] \left[\frac{1}{\lambda^{4/3}} \right] \frac{\partial \psi^{\text{eq}}}{\partial \tilde{I}_2^{\mathbf{C}}} + 2 \frac{\partial \psi^{\text{eq}}}{\partial I_3^{\mathbf{C}}}, \quad (5.21)$$

and the non-equilibrium stress in axial direction,

$$S^{\text{neq}} = \frac{4}{3} \left[\frac{1}{\lambda_i^2} - \frac{\lambda_i}{\lambda^2} \right] \left[\frac{1}{\lambda^{2/3}} \right] \frac{\partial \psi^{\text{neq}}}{\partial \tilde{I}_1^{\mathbf{C}_e}} + \frac{4}{3} \left[\frac{1}{\lambda_i} - \frac{\lambda_i^2}{\lambda^2} \right] \left[\frac{1}{\lambda^{4/3}} \right] \frac{\partial \psi^{\text{neq}}}{\partial \tilde{I}_2^{\mathbf{C}_e}} + 2 \frac{\partial \psi^{\text{neq}}}{\partial I_3^{\mathbf{C}_e}}, \quad (5.22)$$

solely in terms of total and inelastic stretches, λ and λ_i , and the derivatives of the elastic potentials, ψ^{eq} and ψ^{neq} , that we will now discover using constitutive neural networks.

5.5. Neural network model

To discover the model and parameters that best describe our experimental data, we introduce a combination of four constitutive neural networks for the equilibrium potential ψ^{eq} [Linka and Kuhl, 2023], the non-equilibrium potential ψ^{neq} [Holthusen et al., 2025a], the dissipative potential ϕ [Holthusen et al., 2025b], and the weight vector F^{max} . This architecture combines two feed forward neural networks for the equilibrium and non-equilibrium potentials with a recurrent neural network for the dissipative potential within the framework of inelastic constitutive neural networks, and, if desired, combines the resulting overall architecture with a parallel feed forward network for the weight vector. It takes the known deformation state F^n and the internal variables F_i^n from the previous time point t^n as input, integrates them in time using the exponential update (5.15), stores the internal variables F_i^{n+1} , and outputs the second Piola Kirchhoff stress S^{n+1} at the new time point t^{n+1} .

Elastic networks. To model the equilibrium potential ψ^{eq} [Linka and Kuhl, 2023] and the non-equilibrium potential ψ^{neq} [Holthusen et al., 2025b], we introduce two structurally similar but input-independent feed forward networks. The equilibrium network uses the invariants of the total right Cauchy Green tensor $\tilde{I}_1^C, \tilde{I}_2^C, I_3^C$ from equation (5.3) [Linka et al., 2023a; Martonová et al., 2024; Peirlinck et al., 2025] and the non-equilibrium network uses the invariants of the elastic right Cauchy Green tensor $\tilde{I}_1^{C_e}, \tilde{I}_2^{C_e}, I_3^{C_e}$ from equation (5.4) [Holthusen et al., 2024b,a, 2025a] as inputs. For both networks, we ensure expressiveness and interpretability by adopting a sparse network architecture with two hidden layers, where the activation functions f^I of the first layer are the first and second powers and the activation functions f^{II} of the second layer are the identity, the exponential function, and the natural logarithm,

$$f^I \in \left\{ \begin{pmatrix} (\circ)^1 \\ (\circ)^2 \end{pmatrix} \right\} \quad \text{and} \quad f^{II} \in \left\{ \begin{pmatrix} (\circ) \\ \exp(\circ) - 1 \\ (\circ) - 1 - \ln(\circ) \end{pmatrix} \right\}.$$

Figure 5.3 visualizes the architecture of both elastic networks that a priori satisfy the thermodynamic requirements of polyconvexity, growth, and normalization [Boes et al., 2025; Holthusen et al., 2025b; Linka and Kuhl, 2023]. We constrain all weights to be non-negative and add Lasso regularization to the weights of the final layer. Importantly, although the two feed for-

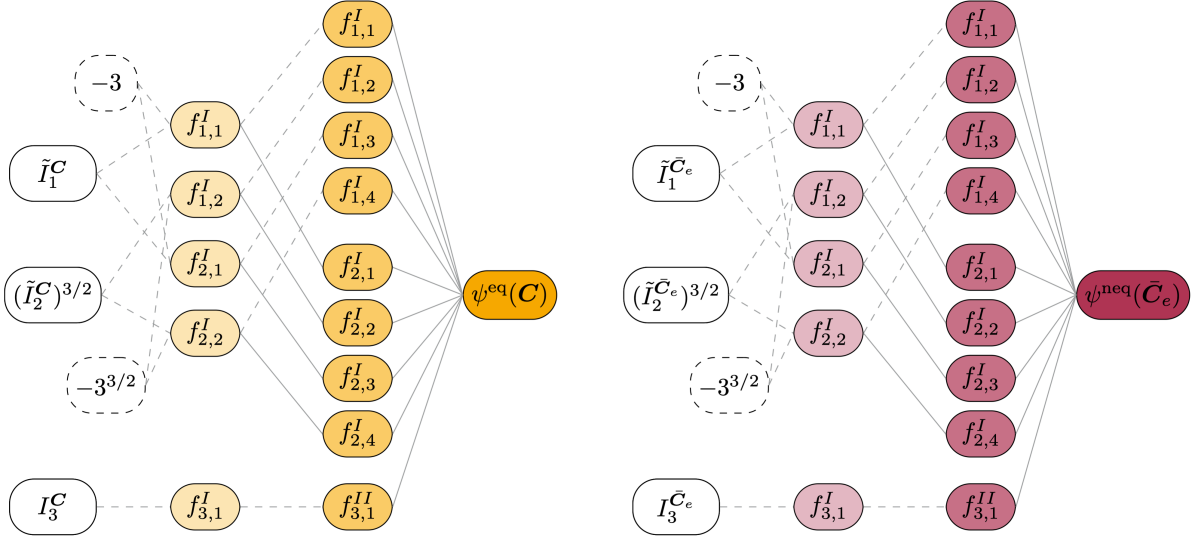


Figure 5.3.: **Elastic networks.** The elastic networks have two hidden layers with 5 and 9 nodes per layer. The feed forward networks take the invariants of the total right Cauchy Green tensor $\tilde{I}_1^C$, $\tilde{I}_2^C$, I_3^C , left, or of the elastic right Cauchy Green tensor $\tilde{I}_1^{C_e}$, $\tilde{I}_2^{C_e}$, $I_3^{C_e}$, right, as input; the first layer applies first or second powers, the second layer applies the identity, exponential function, or natural logarithm; the network learns the equilibrium or non-equilibrium potentials ψ^{eq} , left, or ψ^{neq} , right, as output. Activation functions $f_{\beta,\gamma}^\delta$ are labeled by activation type β , node index γ , and layer index δ ; dashed ellipses denote constant, non-trainable biases; dashed lines indicate fixed weights equal to one.

ward networks for the equilibrium and non-equilibrium potentials are input-independent, they are coupled through their derivatives.

Dissipative network. To model the dissipative potential ϕ [Holthusen et al., 2025b], we introduce a recurrent neural network that uses the invariants of the Mandel stress I_1^Σ , J_2^Σ , J_3^Σ from equation (5.11) as input. Specifically, to maintain consistent physical units, we take their first, second, and third roots as input. We adopt a fully connected network architecture with three hidden layers, where the activation functions f^I of the first layer are trainable powers p of the absolute input values $|(\circ)|$, and the activation functions f^{II} and f^{III} of the second and third layers are the maximum or exponential functions,

$$f^I \in \left\{ |(\circ)|^{p_1+1}, |(\circ)|^{p_2+1}, |(\circ)|^{p_3+1} \right\} \quad \text{and} \quad f^{II} \in \left\{ \max((\circ), 0), \exp(\circ) - 1 \right\} \quad \text{and} \quad f^{III} \in \left\{ \max((\circ), 0), \exp(\circ) - 1 \right\}.$$

Figure 5.4 visualizes the architecture of the dissipative network that a priori satisfies thermo-

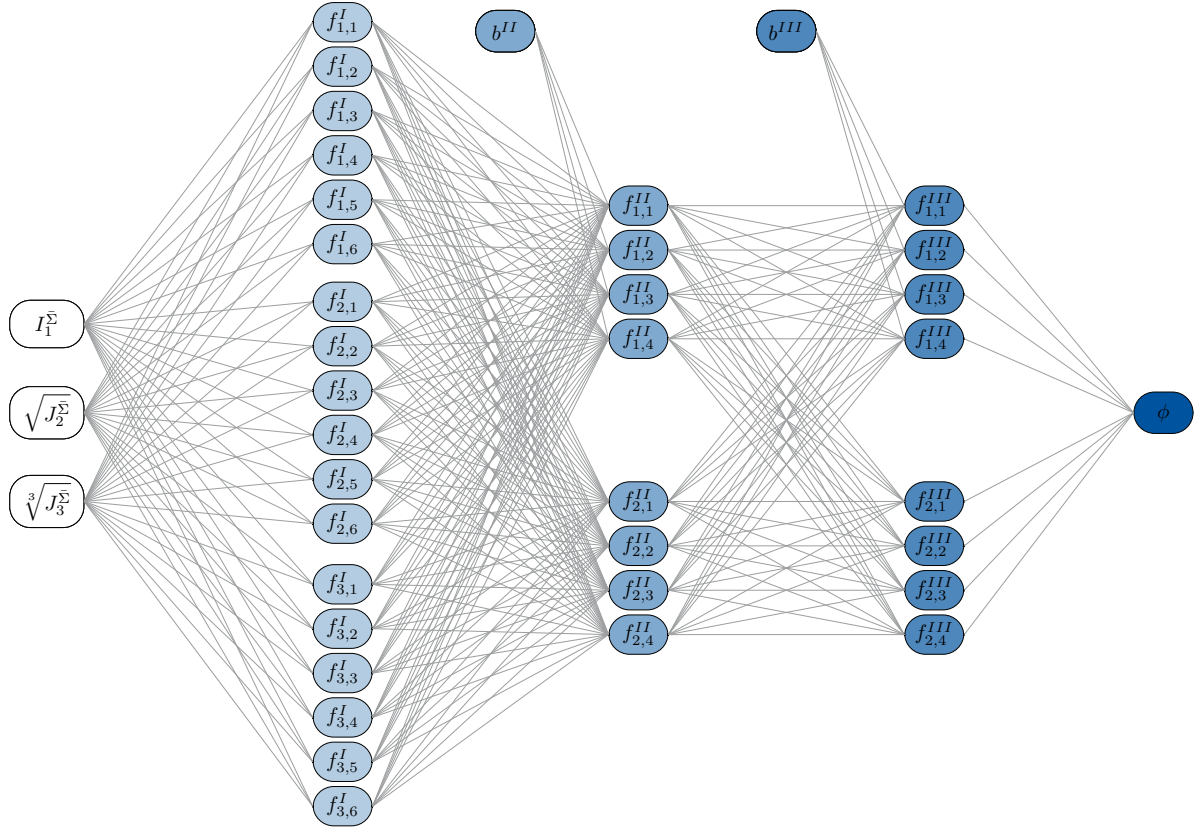


Figure 5.4.: **Dissipative network.** The dissipative network has three hidden layers with 18, 8, and 8 nodes per layer. It takes the invariants of the Mandel stress I_1^Σ , J_2^Σ , J_3^Σ as input; the first layer applies trainable powers p to the absolute input values, the second and third layers apply the maximum or exponential functions; the network learns the dissipative potential ϕ as output. Activation functions $f_{\beta,\gamma}^\delta$ are labeled by activation type β , node index γ , and layer index δ ; biases b^δ are applied to the maximum functions in the second and third hidden layers.

dynamic consistency by remaining non-negative, convex, and zero-valued at equilibrium to ensure that the reduced dissipation inequality $\mathcal{D}_{\text{red}} \geq 0$ (5.13) is satisfied. In the first layer, we smoothen the absolute-value functions $|(\circ)|$ near the origin for differentiability, constrain the exponents p_1, p_2, p_3 to be non-negative, and allow the network weights to take positive or negative values [Holthusen et al., 2025b]. In the second and third layers, we apply additional non-positive biases to the maximum functions and constrain the weights to be non-negative. To guarantee non-negativity, we introduce a final layer that mirrors the third layer, but uses the maximum function without additional weights. After evaluating multiple network configurations with varying numbers of neurons per layer, we select the architecture that achieves the best performance in a preliminary study: a three-layered network with $\{18, 8, 8\}$ nodes per layer. The resulting dissipative network has a total number of 281 trainable parameters.

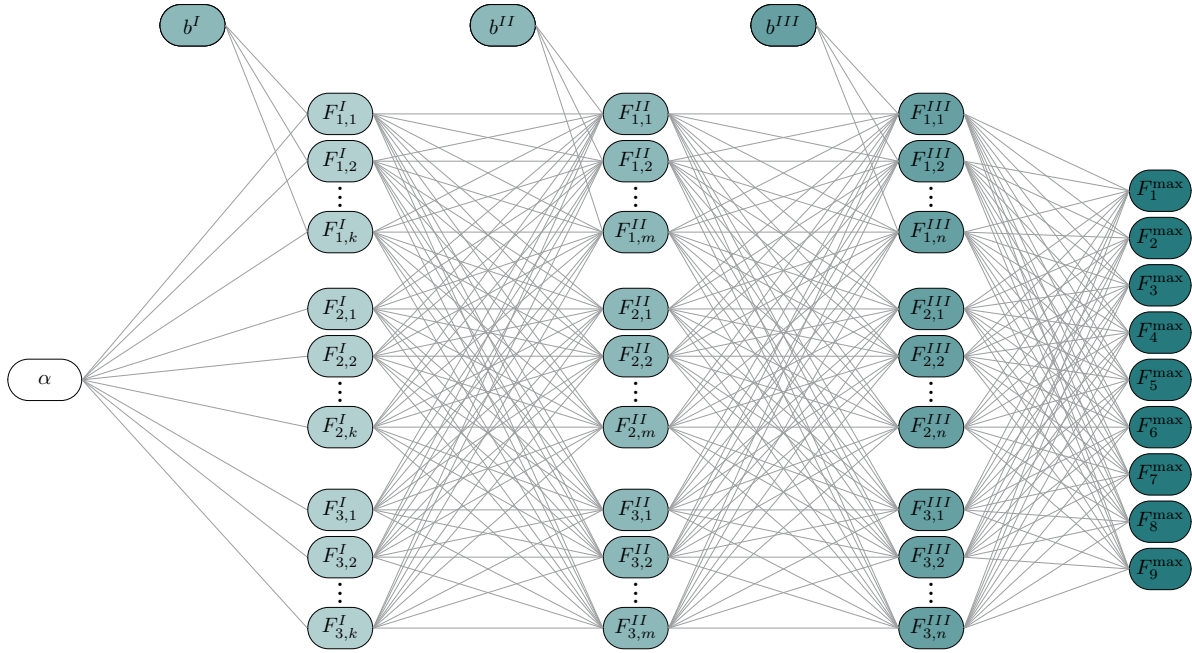


Figure 5.5.: **Feature network.** The feature network has three hidden layers with $k \times 4$ nodes per layer. It takes the feature α , in our case the scalar-valued water content, as input; the three hidden layers apply the maximum, identity, hyperbolic tangent, and softplus functions; the output of the final layer applies the maximum function; the network learns the feature weight vector $F^{\max}$ as output. Activation functions $f_{\beta,\gamma}^\delta$ are labeled by activation type β , node index γ , and layer index δ ; biases b^δ are applied to all maximum functions.

5.8.2 provides additional sensitivity analyses for varying network architectures and varying activation functions.

Feature network. To model the impact of the feature vector on the mechanical response, we introduce a feed forward network that uses the water content as a single scalar-valued input feature α . We adopt a fully connected network architecture with three hidden layers, where the activation functions F^I , F^{II} , F^{III} are the maximum, identity, hyperbolic tangent, and softplus functions and the output of the final layer $F^{\max}$ applies the maximum function,

$$F^I, F^{II}, F^{III} \in \left\{ \begin{array}{l} \max((\circ), 0) \\ (\circ) \\ \tanh(\circ) \\ \tanh(\circ)^2 \end{array} \right\} \quad \text{and} \quad F^{\max} \in \left\{ \max((\circ), 0) \right\}.$$

Figure 5.5 schematically visualizes the architecture of the feature network. To a priori ensure non-negative, non-zero-valued outputs $F^{\max}$, we add biases to all maximum functions.

After evaluating different network architectures, we select feature networks with $\{14,15,11\}$ nodes per activation function per layer for the equilibrium potential, $\{13,17,15,9\}$ for the non-equilibrium potential, and $\{20,6,6\}$ for the dissipative potential. The resulting feature network adds a total number of 19,494 trainable parameters. The number of nodes in the final layer $n^{\max}$ matches those of the mechanical network to which it is coupled, with $n^{\max} = 9$ nodes for the elastic networks in Figure 5.3 and $n^{\max} = 8$ nodes for the dissipative network in Figure 5.4.

Figure 5.6 illustrates the coupling between one of the three mechanical networks and the feature

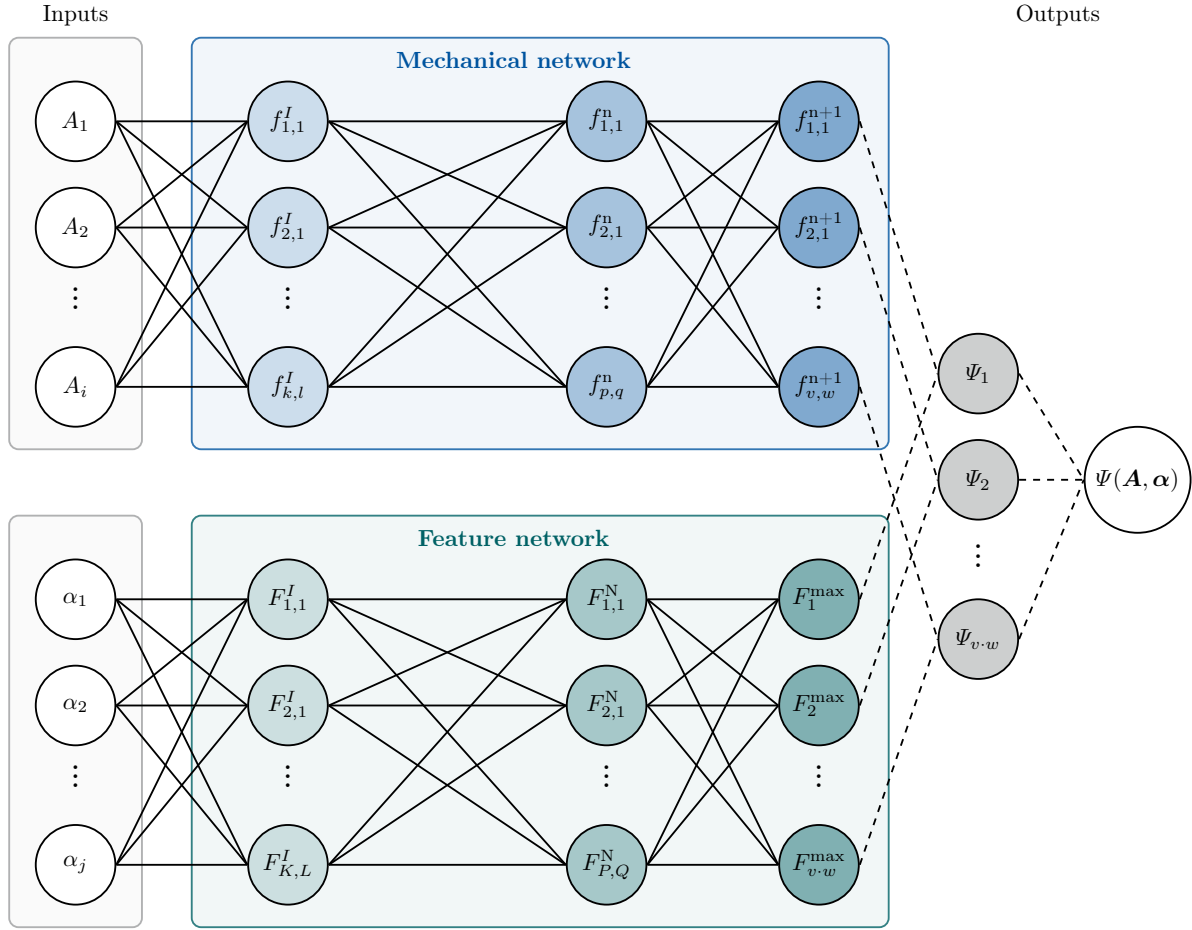


Figure 5.6.: Coupling of mechanical and feature networks. The mechanical network from Figure 5.3 or 5.4 receives a strain- or stress-like input $\mathbf{A}$ and the feature network from Figure 5.5 receives a feature-like input α . The networks are multiplicatively coupled in the nodes of their final layers. The sum of the final layer defines the network output, the feature-weighted mechanical potential $\Psi(\mathbf{A}, \alpha)$. Activation functions f_{γ}^{δ} and F_{γ}^{δ} of the mechanical and feature networks are labeled by node index γ and layer index δ .

network. The mechanical networks from Figure 5.3 or 5.4 receive a strain- or stress-like input $\mathbf{A}$ and the feature network from Figure 5.5 receives a feature-like input α . The networks are multiplicatively coupled by summing up the product of the nodes in their final layers to obtain feature-weighted potential Ψ ,

$$\Psi(\mathbf{A}, \alpha) = \sum_{i=1}^{n_{\max}} f_i^{\max}(\mathbf{A}) F_i^{\max}(\alpha) \quad \text{with} \quad \Psi = \{ \psi^{\text{eq}}, \psi^{\text{neq}}, \phi \} \quad \text{and} \quad \mathbf{A} = \{ \mathbf{C}, \mathbf{C}^e, \boldsymbol{\Sigma} \},$$

where the potential Ψ is either the equilibrium or non-equilibrium potential ψ^{eq} or ψ^{neq} , or the dissipative potential ϕ , and the input $\mathbf{A}$ is either the total or the elastic right Cauchy Green tensor $\mathbf{C}$ or $\mathbf{C}^e$ or the Mandel stress $\boldsymbol{\Sigma}$. This multiplicative coupling ensures that the combined output is non-negative and remains zero for zero-valued mechanical outputs. Multiplicative feature weighting inherently preserves polyconvexity and other thermodynamic requirements.

Network training. To discover the best model and parameters to characterize tofu, we train and test our networks using the experimental data from Section 5.3. We introduce the loss function as the mean squared error between the model stress S and the experimental stress $\hat{S}$ in the direction of loading, averaged across all n_{exp} experiments and all n_{data} data points. To induce sparsity and improve generalization, we supplement the loss with L_1 -regularization or Lasso and L_2 -regularization or ridge regression across all n_w weights, where α_1 and α_2 are the regularization coefficients. The total loss function reads

$$\begin{aligned} L(S, \mathbf{w}) = & \frac{1}{n_{\text{exp}}} \sum_{k=1}^{n_{\text{exp}}} \frac{1}{n_{\text{data}}} \sum_{m=1}^{n_{\text{data}}} \left\| S_m(\mathbf{C}_m, \mathbf{F}_{i,m}, \mathbf{w}) - \hat{S}_m \right\|_k^2 + \sum_{n=1}^{n_w} \alpha_1 |w_n| \\ & + \sum_{n=1}^{n_w} \alpha_2 (w_n)^2. \end{aligned} \quad (5.23)$$

We perform two types of model discovery: First, we analyze *each tofu type independently* by using only the elastic and dissipative networks to discover *three individual tofu models* for each different tofu type. Then, we analyze *all three tofu types combined* by integrating the elastic and dissipative networks with the feature network to discover a *single unified tofu model* with the water content as input feature.

Discovering three individual tofu models. First, we train the elastic and dissipative networks to discover individual tofu models using the stress–strain data from Section 5.3. We neglect water-content dependency and discover three independent material models by simultaneously learning the equilibrium and non-equilibrium potentials ψ^{eq} and ψ^{neq} and the dissipative potential ϕ . We use the triple-compression tests at three different strain rates as training data and the

double-compression tests at intermediate strain rates as test data. To capture history-dependent effects, we adopt a staggered training strategy that progressively increases the batch size and gradually introduces additional training data along the load path. The training proceeds in two phases: In the first phase, we only include the first $n = 50$ data points of the first cycle of the triple-compression test. In the second phase, we include all cycles and apply a regularization with $\alpha_1 = 0.0001$ and $\alpha_2 = 0.0001$. The weights and biases from the first phase initialize the second phase to ensure continuity and convergence throughout the process.

Discovering a single unified tofu model. Then, we train the feature networks to discover a unified tofu model for the stress–strain data from Section 5.3. We incorporate the water-content as a scalar-valued input feature and discover a single material model by learning feature-weighted terms of the equilibrium, non-equilibrium, and dissipative potentials, ψ^{eq} and ψ^{neq} and ϕ . Yet, unlike the elastic and dissipative networks, we do not train the feature network directly on the experimental data. Instead, we train on the *outputs of the discovered models*. We generate synthetic training data by evaluating our three discovered models for various sets of invariants $I_i^A = \{I_i^C, I_i^{C_e}, I_i^\Sigma\}$ of our network input $\mathbf{A} = \{\mathbf{C}, \mathbf{C}_e, \boldsymbol{\Sigma}\}$, the total and elastic right Cauchy Green tensors $\mathbf{C}$ and $\mathbf{C}_e$ and the Mandel stress $\boldsymbol{\Sigma}$. According to equations (5.8), (5.9), and (5.14), the derivatives of the potentials $\Psi = \{\psi^{\text{eq}}, \psi^{\text{neq}}, \phi\}$ with respect to these invariants, $\partial\Psi/\partial I_i^A$, determine both the stress response and the inelastic flow direction $\{\mathbf{S}^{\text{eq}}, \mathbf{S}^{\text{neq}}, \mathbf{D}_i\}$. We train the combined feature–mechanical network on these derivatives, which contain the essential physical information.

Figure 5.7 illustrates how we generate our synthetic training data $\partial\hat{\Psi}_\alpha/\partial I_i^A$: We evaluate each of the three discovered network potentials $\hat{\Psi}_\alpha$ for a series of invariants I_i^A , where $\alpha = \{n_{\text{silken}}, n_{\text{firm}}, n_{\text{extra}}\}$ is the water content of silken, firm, and extrafirm tofu and $i = \{1, 2, 3\}$ are the three isotropic invariants. The derivatives of the potentials with respect to the invariants, $\partial\hat{\Psi}_\alpha/\partial I_i^A$, represent the training data that we seek to approximate by the unified model $\Psi(\mathbf{A}, \alpha, \mathbf{w})$ of the feature-mechanical network that takes both the invariants of $\mathbf{A}$ and the feature α as input. During training, the networks learn the weights $\mathbf{w}$ by minimizing the loss function that measures the error between the derivatives of the unified model $\partial\Psi_{\alpha,m}/\partial I_i^A$ and the derivatives of the synthetic training data $\partial\hat{\Psi}_{\alpha,m}/\partial I_i^A$ at $m = 1, \dots, n_{\text{data}}$ data points,

$$L(\partial\Psi/\partial I_i^A, \mathbf{w}) = \sum_{\alpha=1}^{n_{\text{tofu}}} \sum_{m=1}^{n_{\text{data}}} \sum_{i=1}^3 \left\| \frac{\partial\Psi_m(\mathbf{A}_m, \alpha_m, \mathbf{w})}{\partial I_i^A} - \frac{\partial\hat{\Psi}_m}{\partial I_i^A} \right\|. \quad (5.24)$$

We adopt the Adam optimizer and train the equilibrium network for 5,000 epochs using 500 training samples with a learning rate of 10^{-3} and a clip norm of 0.001, and the non-

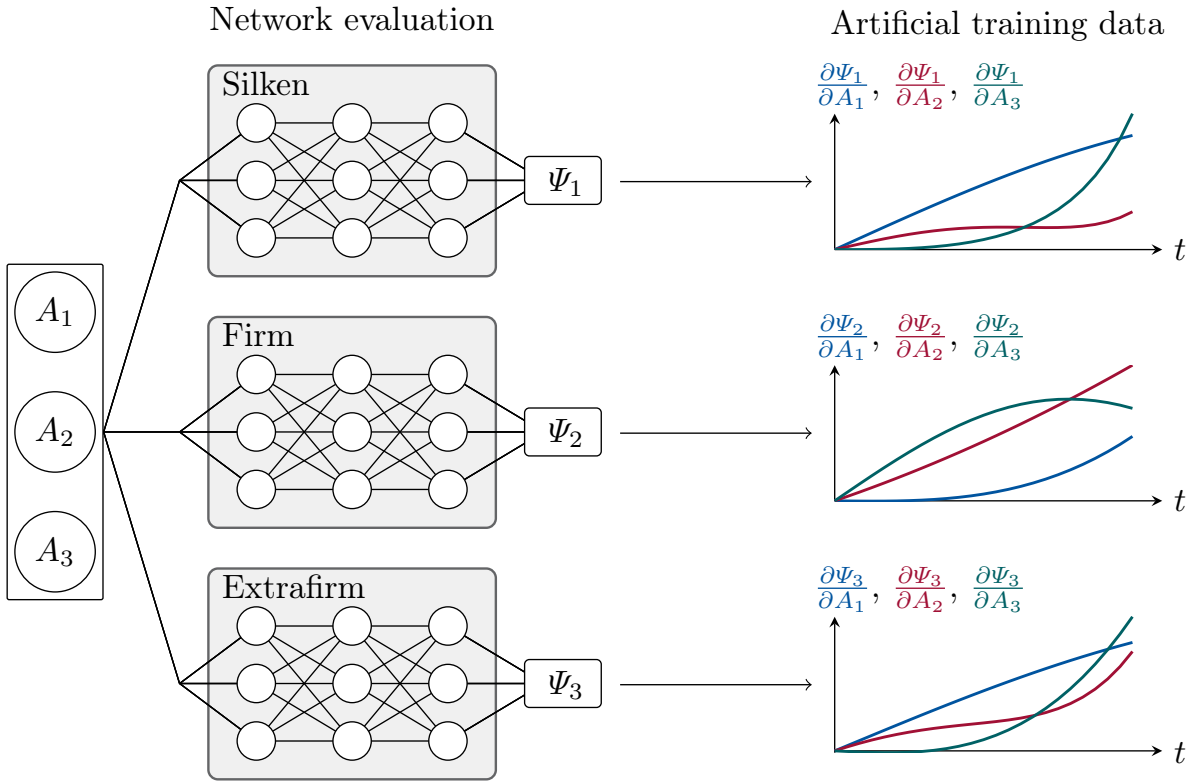


Figure 5.7.: **Generating synthetic training data.** After discovering the best models and parameters for for silken, firm, and extrafirm tofu by training the equilibrium and non-equilibrium networks in Figures 5.3 and the dissipative network in Figure 5.4 on the experimental data, we generate synthetic data to train the coupled mechanical and feature networks in Figure 5.6. We evaluate the three tofu models for sets of invariants I_i^A , and their derivatives $\hat{\Psi}/\partial I_i^A$ provide the training data for the unified feature–mechanical model.

equilibrium and dissipative networks for 15,000 and 10,000 epochs with a learning rate of 10^{-4} . This approach eliminates the need to retrain the entire recurrent network on experimental stress–strain pairs and significantly accelerates the learning process. Importantly, it produces a *single tofu model* that captures tofu types of different water content within a unified constitutive framework.

5.6. Results

Mechanical testing. Figures 5.8 and 5.9 summarize the applied stretch and stress histories $\lambda(t)$ and $S(t)$ during the triple-compression and double-compression tests. Each row corresponds to one tofu type—silken, firm, and extrafirm—and each column represents a different

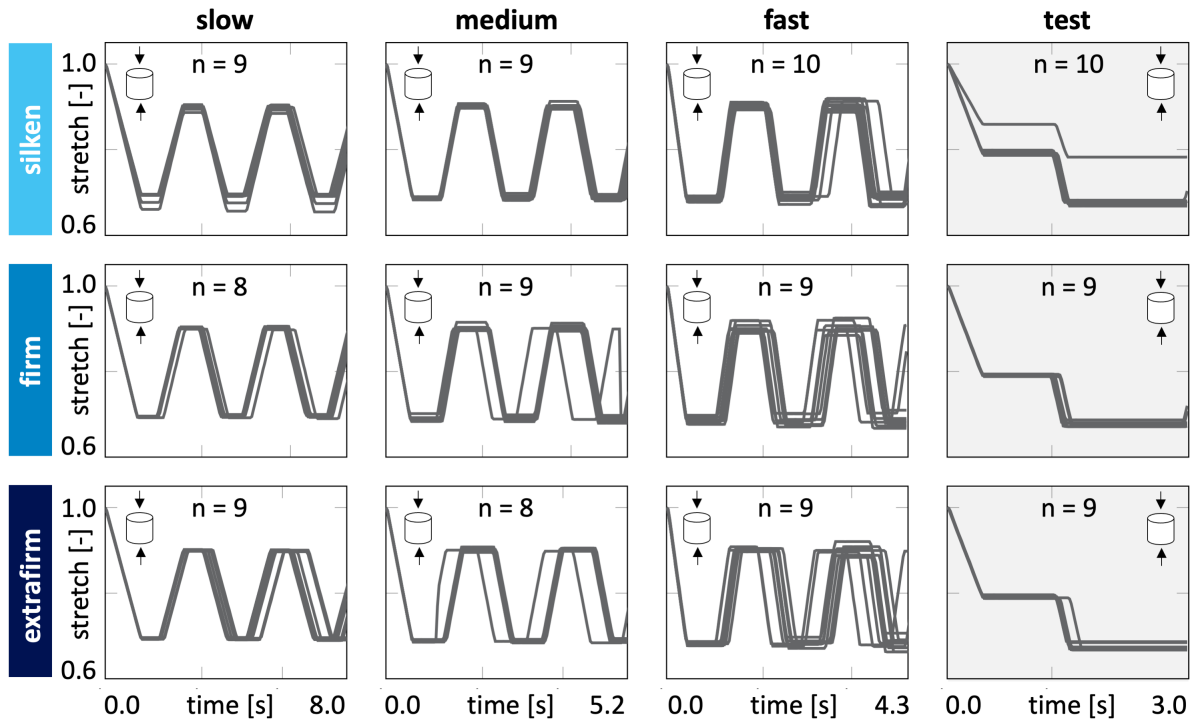


Figure 5.8.: **Mechanical testing - Experimentally applied stretches.** Stretch histories $\lambda(t)$ during triple-compression and double-compression tests for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests at slow, medium, and fast strain rates of 20%/s, 40%/s, and 60%/s (from left to right), where each specimen is loaded to a stretch of 0.7 and unloaded to 0.9. The last column shows the double-compression tests at a strain rates of 30%/s to a stretch of 0.8 followed by a strain rate of 50%/s to 0.7.

loading condition. The first three columns show the triple-compression tests at different strain rates and the final column shows the double-compression tests. The curves display high reproducibility across all tofu types and loading conditions and confirm that we can robustly record stretch and stress histories at different stretch magnitudes and rates. However, especially at higher strain rates, we observe slight variations in the applied stretch profiles resulting in slight inconsistencies in the stress profiles, which complicate direct averaging of the curves. We thus report each curve as an independent measurement. As a general trend, we observe a non-linear stress response that is sensitive to both strain rate and tofu type: extrafirm tofu is the stiffest, closely followed by firm tofu, and silken tofu is the softest. The stiffness decreases with increasing water content, consistent with the transition from solid-like to fluid-like behavior.

Discovering three individual tofu models. Figure 5.10 compares the computationally predicted stresses in blue and the experimentally recorded stresses in gray. The rows correspond to silken, firm, and extrafirm tofu, and the columns represent distinct loading cases with vary-

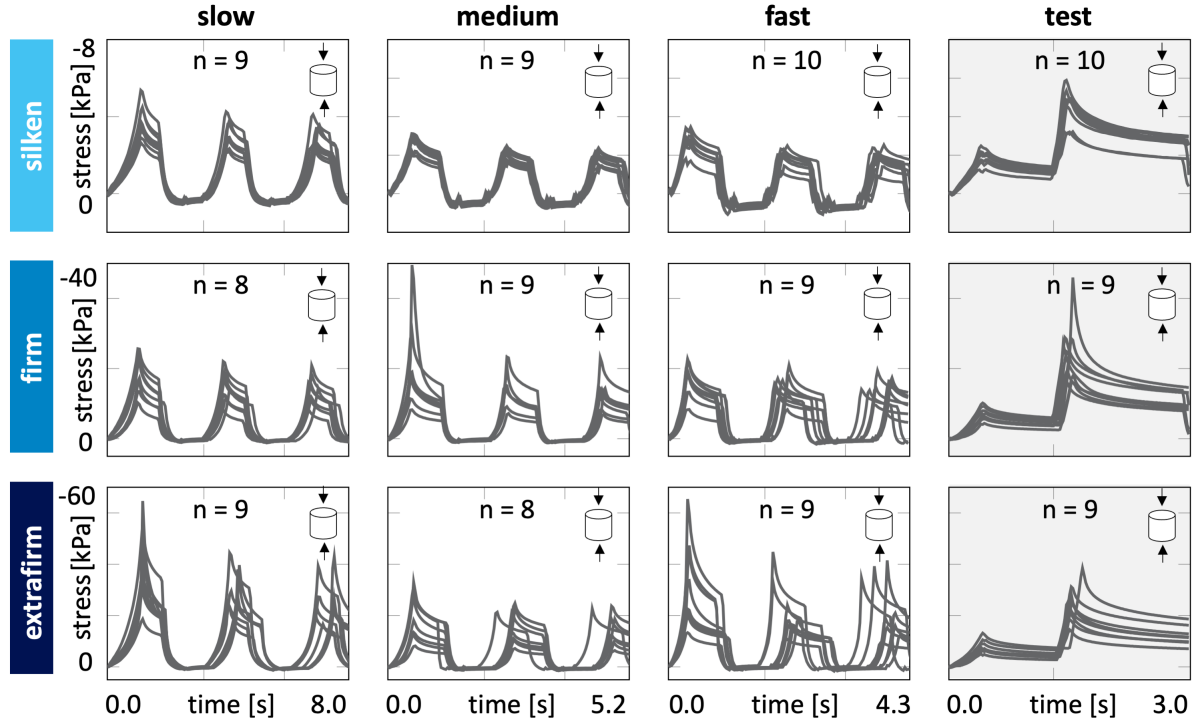


Figure 5.9.: **Mechanical testing - Experimentally recorded stresses.** Stress histories $S(t)$ during triple-compression and double-compression tests for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests at slow, medium, and fast strain rates of 20%/s, 40%/s, and 60%/s (from left to right), where each specimen is loaded to a stretch of 0.7 and unloaded to 0.9. The last column shows the double-compression tests at a strain rates of 30%/s to a stretch of 0.8 followed by a strain rate of 50%/s to 0.7.

ing stretch magnitudes and rates. The three discovered models and parameter sets for each tofu type accurately reproduce the experimental trends: For the training data in the first three columns, the modeled stresses generally fall within the measured stress ranges and accurately capture the mean behavior during loading, holding, and unloading. For the test data in the last column, the silken tofu model accurately reproduces the first peak and the relaxation behavior, but slightly underestimates the second peak; the firm tofu model displays a similar trend; and the extrafirm model aligns closely with all experimental curves. Across all tofu types, the discovered models capture a rate dependent behavior: the higher the strain rate, the stiffer the response and the larger the peak stress. In a few cases, the slowest strain rate generates higher peak stresses than the intermediate rate, but overall, the discovered models correctly reproduce strain-rate stiffening.

For silken tofu, the discovered model features a quadratic second-invariant term and a loga-

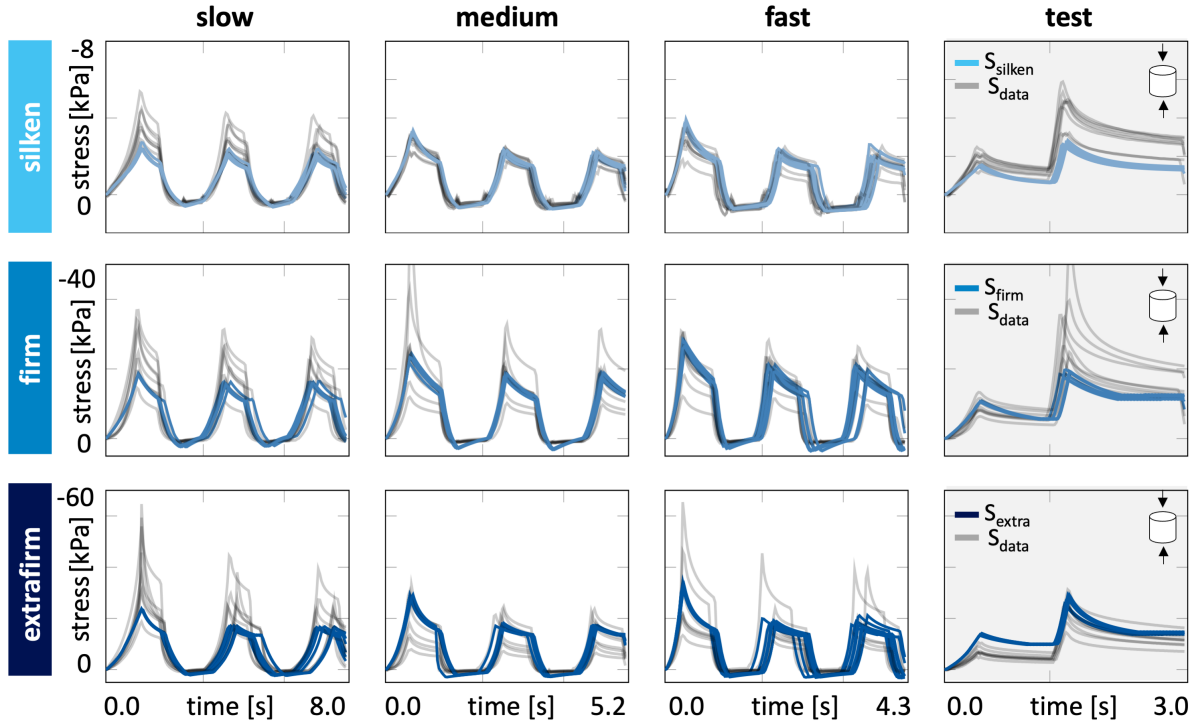


Figure 5.10.: **Model discovery - Experimentally recorded and computationally predicted stresses of the three individual tofu models.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests used as training data. The last column shows double-compression tests used as test data. Experimental data are plotted in gray, model predictions for silken, firm, and extrafirm tofu in light to dark blue.

rithmic third-invariant term for the equilibrium potential and a linear second-invariant term for the non-equilibrium potential,

$$\begin{aligned}\psi_{\text{silken}}^{\text{eq}}(\mathbf{C}) &= \frac{1}{2} \mu_7 [(\tilde{I}_2^{\mathbf{C}})^{3/2} - 3^{3/2}]^2 + \alpha_9 [(I_3^{\mathbf{C}})^{\beta_9} - 1 - \beta_9 \ln(I_3^{\mathbf{C}})] \\ \psi_{\text{silken}}^{\text{neq}}(\mathbf{C}_e) &= \frac{1}{2} \mu_5 [(\tilde{I}_2^{\mathbf{C}_e})^{3/2} - 3^{3/2}],\end{aligned}$$

with network weights $\mu_7 = 0.158 \text{ kPa}$, $\alpha_9 = 0.052 \text{ kPa}$, $\beta_9 = 5.507$ for the equilibrium part and $\mu_5 = 1.179 \text{ kPa}$ for the non-equilibrium part that have the physical unit of stress and a clear physical interpretation as stiffness-like parameters. For firm tofu, the discovered model features a quadratic first-invariant term and a linear second-invariant term for the equilibrium potential and a linear second-invariant term and a logarithmic third-invariant term for the

non-equilibrium potential,

$$\begin{aligned}\psi_{\text{firm}}^{\text{eq}}(\mathbf{C}) &= \frac{1}{2} \mu_3 [\tilde{I}_1^{\mathbf{C}} - 3]^2 + \frac{1}{2} \mu_5 [(\tilde{I}_2^{\mathbf{C}})^{3/2} - 3^{3/2}] \\ \psi_{\text{firm}}^{\text{neq}}(\mathbf{C}_e) &= \frac{1}{2} \mu_5 [(\tilde{I}_2^{\mathbf{C}_e})^{3/2} - 3^{3/2}] + \alpha_9 [(I_3^{\mathbf{C}_e})^{\beta_9} - 1 - \beta_9 \ln(I_3^{\mathbf{C}_e})],\end{aligned}$$

with network weights $\mu_3 = 5.250$ kPa, $\mu_5 = 0.136$ kPa for the equilibrium part and $\mu_5 = 4.952$ kPa, $\alpha_9 = 0.867$ kPa, $\beta_9 = 1.439$ for the non-equilibrium part, again, all with a clear physical interpretation. For extrafirm tofu, the discovered model features a quadratic second-invariant term and a logarithmic third-invariant term for the equilibrium potential and a linear and a quadratic second-invariant term for the non-equilibrium potential,

$$\begin{aligned}\psi_{\text{extra}}^{\text{eq}}(\mathbf{C}) &= \frac{1}{2} \mu_7 [(\tilde{I}_2^{\mathbf{C}})^{3/2} - 3^{3/2}]^2 + \alpha_9 [(I_3^{\mathbf{C}})^{\beta_9} - 1 - \beta_9 \ln(I_3^{\mathbf{C}})] \\ \psi_{\text{extra}}^{\text{neq}}(\mathbf{C}_e) &= \frac{1}{2} \mu_5 [(\tilde{I}_2^{\mathbf{C}_e})^{3/2} - 3^{3/2}] + \frac{1}{2} \mu_7 [(\tilde{I}_2^{\mathbf{C}_e})^{3/2} - 3^{3/2}]^2\end{aligned}$$

with $\mu_7 = 0.151$ kPa, $\alpha_9 = 0.103$ kPa, $\beta_9 = 4.920$ for the equilibrium part and $\mu_5 = 8.160$ kPa, $\mu_7 = 2.735$ kPa for the non-equilibrium part. Notably, the most prevalent invariant across all six potentials is the second invariant $\tilde{I}_2^{\mathbf{C}_e}$ featured in seven terms, followed by the third invariant $I_3^{\mathbf{C}_e}$ in three terms, and the first invariant $I_3^{\mathbf{C}_e}$ in only one term. None of the models features any of the four exponential terms, neither linear nor quadratic, neither in the first nor in the second invariant, and none of the models features a quadratic first-invariant term.

The dissipative part of our model discovery uses the dissipative network in Figure 5.4 to discover the dissipative potential ϕ . Figures 5.11-5.13 display the three discovered potentials for silken, firm, and extrafirm tofu. The colors represent the magnitude of the final weights and biases and enable direct interpretation of active pathways. For silken tofu in Figure 5.11, the dissipative potential is driven by both the second and third deviatoric invariants J_2^{Σ} and J_3^{Σ} ,

$$\begin{aligned}\phi_{\text{silken}} &= \max (1.653 [\exp(1.770 \cdot \exp(1.837 \cdot | - 0.058 (J_2^{\Sigma})^{1/2} \\ &\quad + 0.148 (J_3^{\Sigma})^{1/3} |^{0.14}) - 1) - 1], 0).\end{aligned}$$

For firm tofu in Figure 5.12, the dissipative potential is driven exclusively by the second deviatoric invariant J_2^{Σ} ,

$$\phi_{\text{firm}} = \max (0.780 [\exp(0.814 \cdot \exp(0.857 \cdot | 0.457 (J_2^{\Sigma})^{1/2} |^{0.036}) - 1) - 1], 0).$$

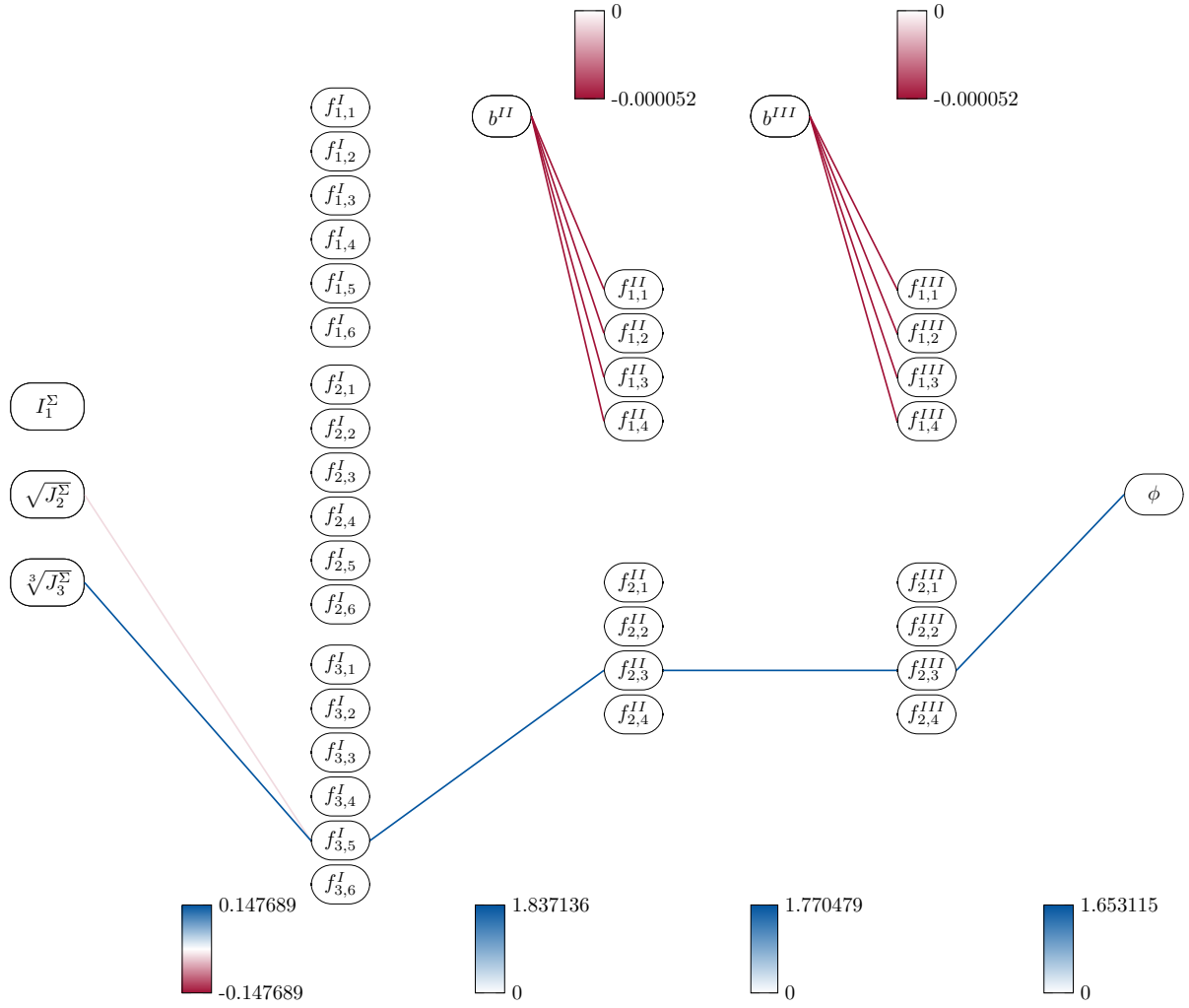


Figure 5.11.: **Model discovery - Trained dissipative network for silken tofu.** The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a single-term potential for silken tofu that only depends on the third deviatoric invariant J_3^Σ of the Mandel stress Σ .

For extrafirm tofu in Figure 5.13, the dissipative potential is driven exclusively by the third deviatoric invariant J_3^Σ ,

$$\phi_{\text{extra}} = \max \left(1.698 \left[\exp \left(1.710 \cdot \exp \left(1.717 \cdot |0.156 (J_3^\Sigma)^{1/3}|^{1.42} \right) - 1 \right) - 1 \right], 0 \right).$$

Strikingly, for all three tofu types, the networks are sparse and discover the same functional form: a single term that contains weighted powers of the second and third deviatoric invariants J_2^Σ and J_3^Σ embedded in two nested exponential functions. None of the discovered potentials

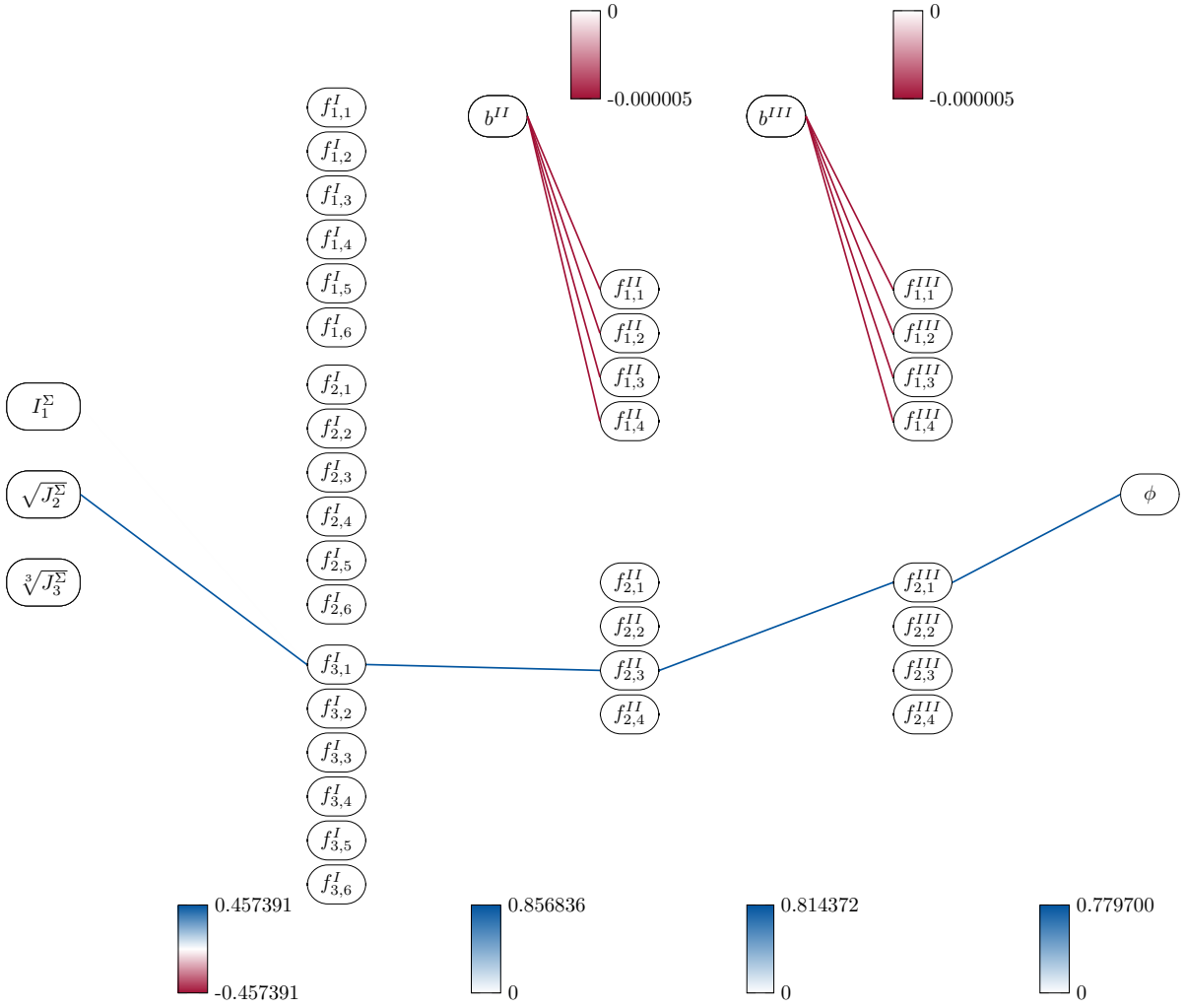


Figure 5.12.: **Model discovery - Trained dissipative network for firm tofu.** The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a single-term potential for firm tofu that depends on the second and third deviatoric invariants J_2^Σ and J_3^Σ of the Mandel stress Σ .

depends on the volumetric invariant I_1^Σ , and none of them features any biases that would result in additive contributions.

Figure 5.14 illustrates the stresses and the inelastic deformation rates of the discovered models for silken, firm, and extrafirm tofu for a strain rate of 60%/s. The first row shows the applied stretches λ , followed by the total stress S , its equilibrium and non-equilibrium parts S^{eq} and S^{neq} , and the inelastic deformation rate $\dot{C}_i$. The color shades denote contributions from the three invariants $\{\tilde{I}_1^C, \tilde{I}_2^C, I_3^C\}$ and $\{\tilde{I}_1^{C_e}, \tilde{I}_2^{C_e}, I_3^{C_e}\}$ and $\{I_1^\Sigma, J_2^\Sigma, J_3^\Sigma\}$. Strikingly, the predicted peak stress S in the second row increases notably with decreasing water content, from

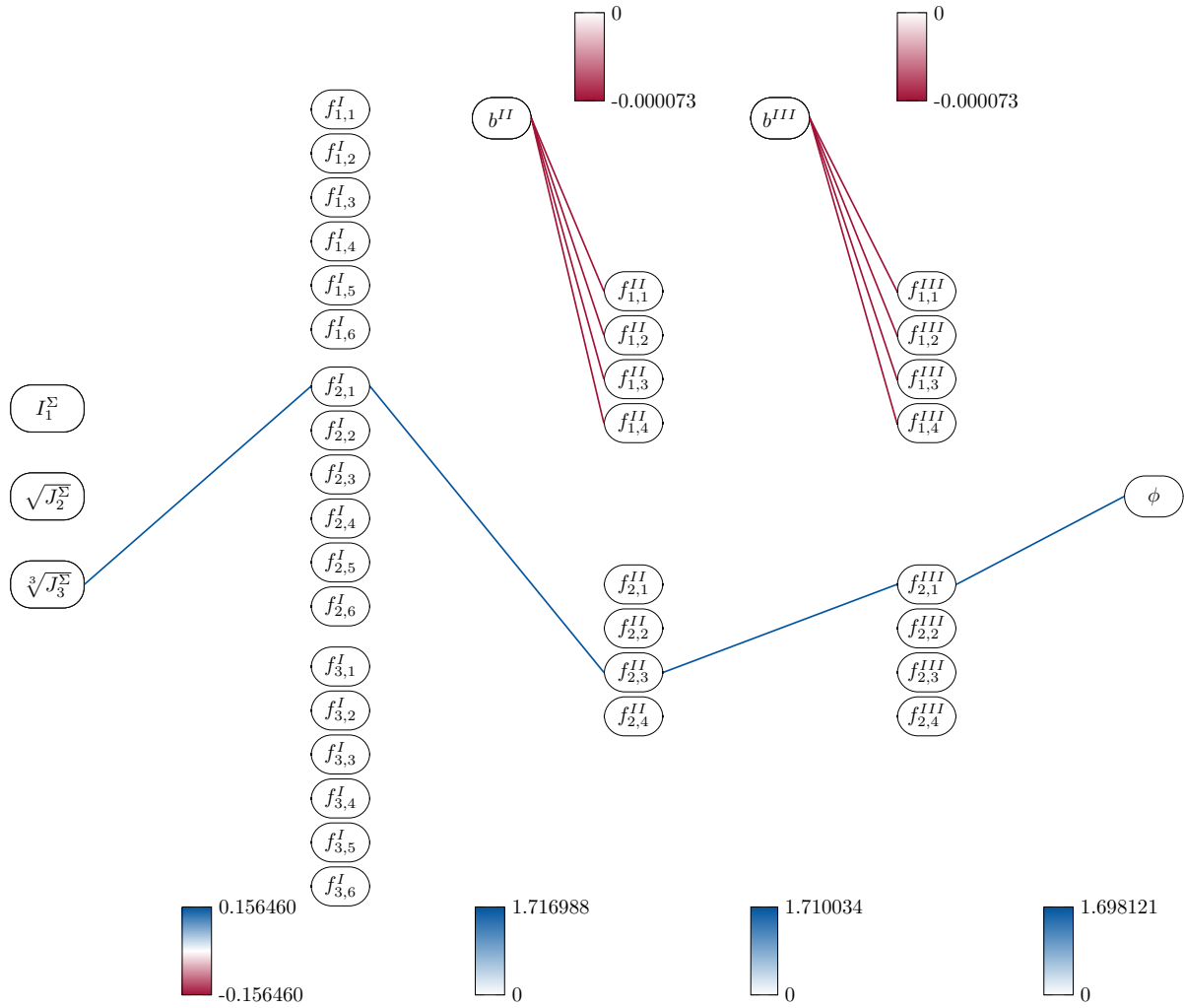


Figure 5.13.: **Model discovery - Trained dissipative network for extrafirm tofu.** The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a four-term potential for extrafirm tofu that depends on complex combinations of the second and third deviatoric invariants J_2^Σ and J_3^Σ of the Mandel stress Σ .

-3.90 kPa for silken to -19.62 kPa for firm and -33.85 kPa for extrafirm tofu, from left to right. The peak equilibrium stress S^{eq} in the third row increases also, from -1.26 kPa for silken to -2.44 kPa for firm and -2.76 kPa for extrafirm tofu, but less drastically than the total stress S . Notably, the peak non-equilibrium stress S^{neq} in the fourth row is almost entirely made up of second-invariant terms in $\tilde{I}_2^C$ across all three tofu types. Similarly, the evolution of the inelastic right Cauchy tensor $\dot{C}_i$ in the bottom row is made up only of second- and third-invariant terms in J_2^Σ and J_3^Σ , suggesting that the inelastic deformation remains entirely deviatoric.

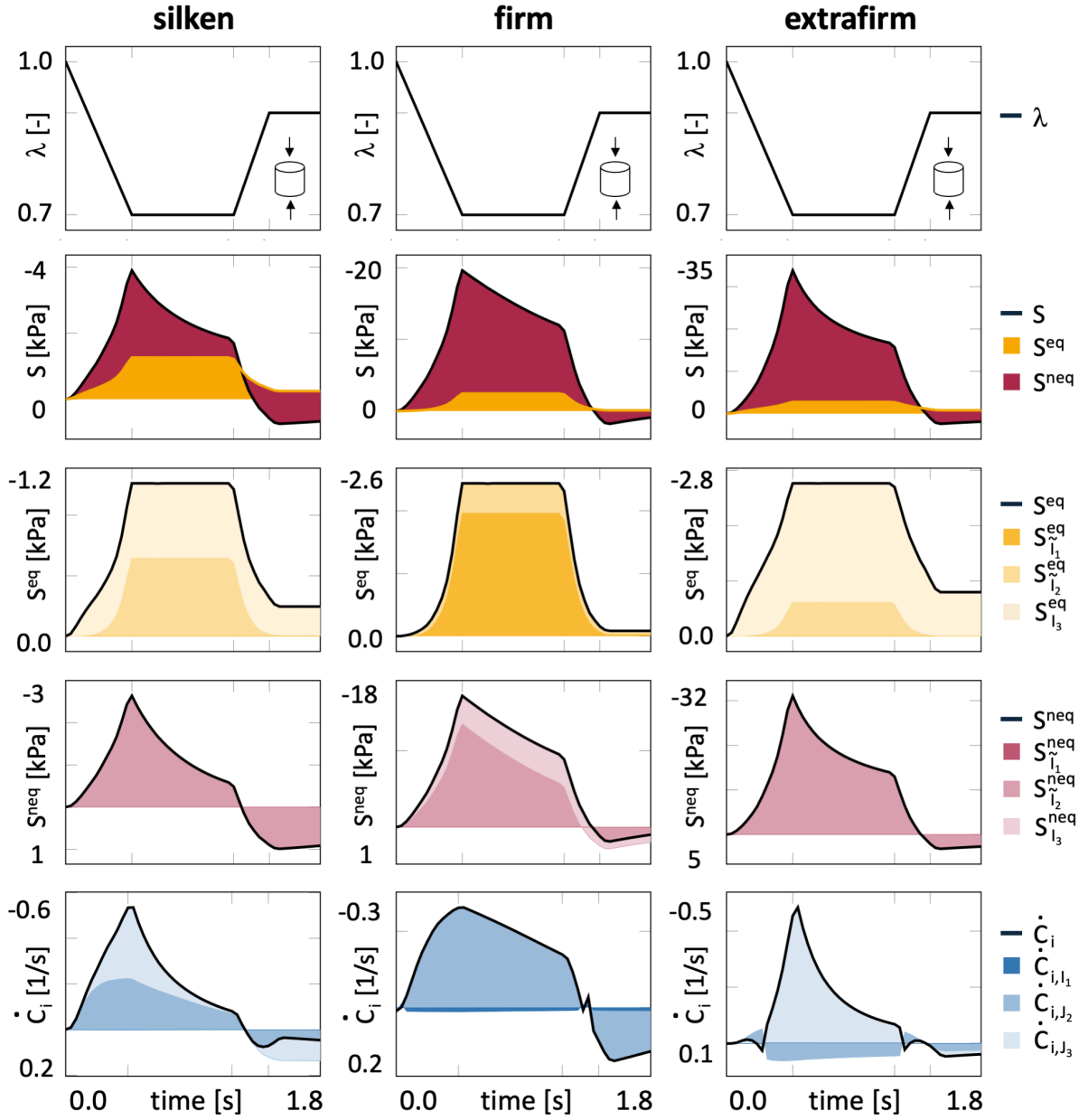


Figure 5.14.: **Model discovery - Predicted stresses and inelastic deformation rates of the three individual tofu models.** Stress histories and evolution of the inelastic deformation rate for silken (left), firm (middle), and extrafirm (right) tofu for a strain rate of 60%/s. Applied stretches λ (top); total stress S , equilibrium stress S^{eq} , and non-equilibrium stress S^{neq} (middle), and inelastic deformation rate $\dot{C}_i$ (bottom). Distinct patterns denote contributions from the three invariants $\{\tilde{I}_1^C, \tilde{I}_2^C, I_3^C\}$ and $\{\tilde{I}_1^{C_e}, \tilde{I}_2^{C_e}, I_3^{C_e}\}$ and $\{I_1^\Sigma, J_2^\Sigma, J_3^\Sigma\}$.

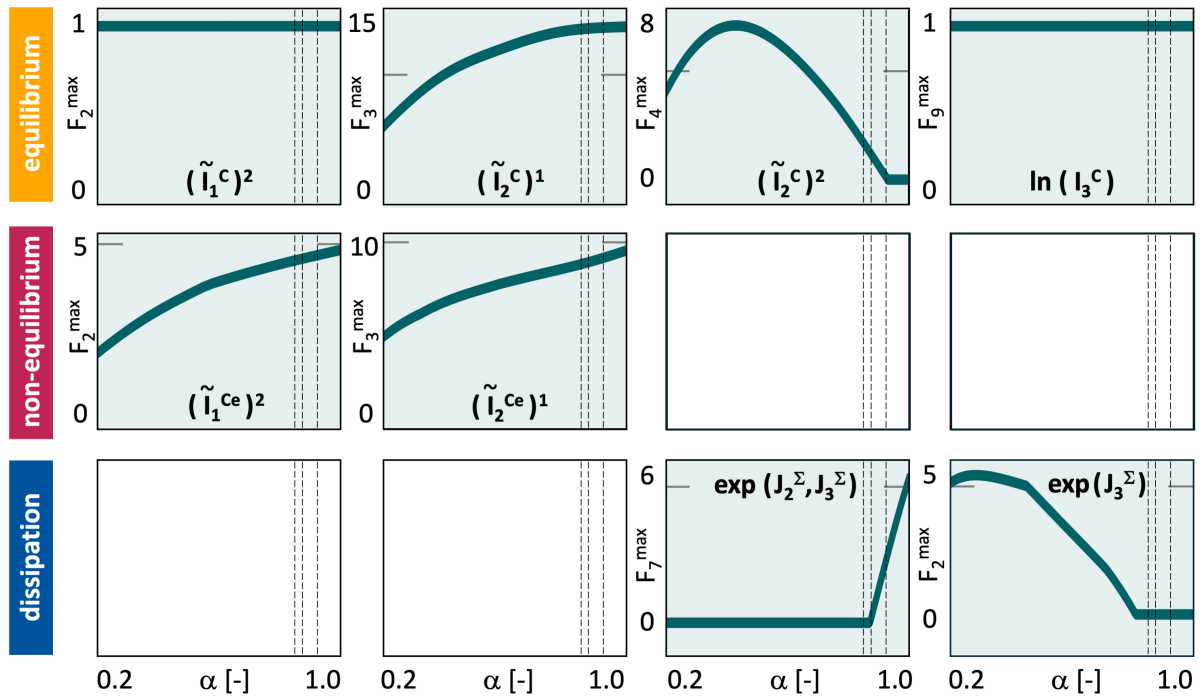


Figure 5.15.: **Model discovery - Water-content dependency of the single unified tofu model.** Functional dependency of the equilibrium potential ψ^{eq} (top), non-equilibrium potential ψ^{neq} (middle), and dissipative potential ϕ (bottom) on the water content α . Columns show the discovered dependencies on the individual terms. Dashed lines highlight the water contents of extrafirm, firm, and silken tofu.

Importantly, as the water content decreases, from left to right, the equilibrium contribution S^{eq} to the total stress S decreases: It accounts for 1/3 of the total stress in silken, but for only 1/8 in firm and 1/12 in extrafirm tofu.

Discovering a single unified tofu model. Figure 5.15 summarizes the water-content dependency of the single unified tofu model. The rows correspond to the equilibrium, non-equilibrium, and dissipative potentials, the columns show the discovered dependencies of the individual terms, and the dashed lines indicate the water contents of our extrafirm, firm, and silken tofu. Interestingly, the majority of terms, six of eight, depend on the water content of tofu, while two terms do not: The equilibrium potential ψ^{eq} features water-content-independent quadratic first-invariant and logarithmic third-invariant terms, combined with water-content-dependent linear and quadratic second-invariant terms, where the first increases and the second decreases with water content. The non-equilibrium potential ψ^{neq} features water-content-dependent quadratic first-invariant and linear second-invariant terms, where both increase with water content. The dissipative potential ϕ features water-content-dependent exponential second- and third-invariant terms, where the first increases and the second decreases with water

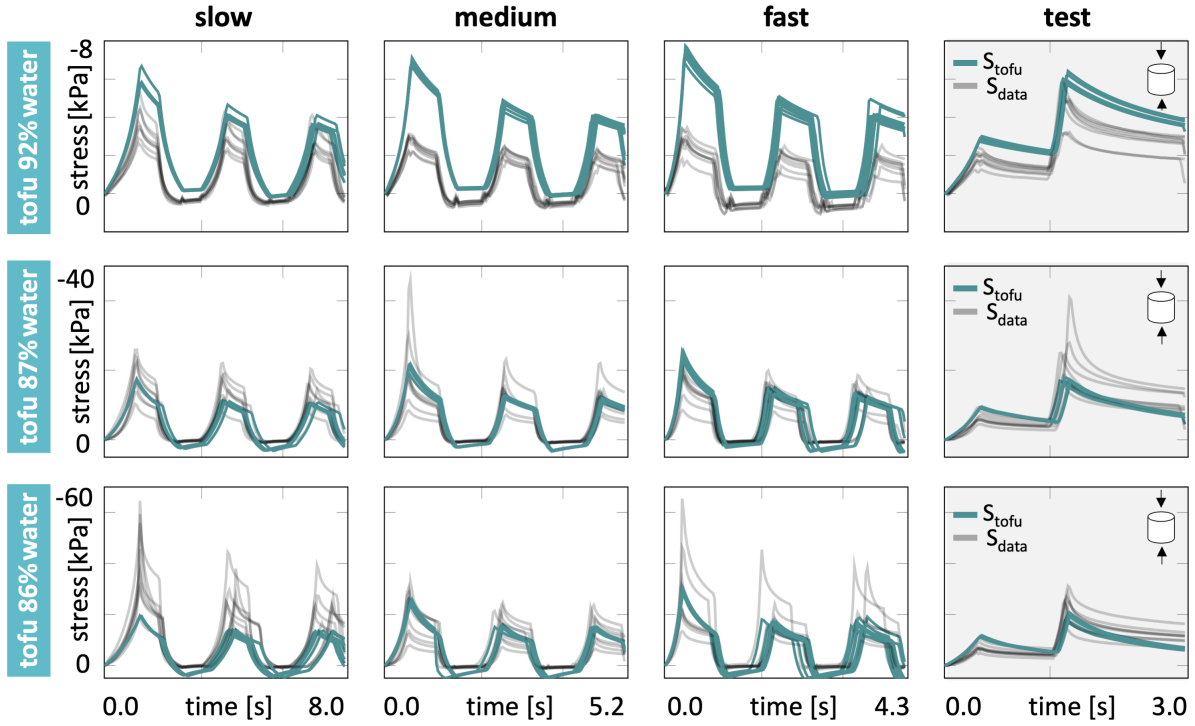


Figure 5.16.: **Model discovery - Experimentally recorded and computationally predicted stresses of the single unified tofu model.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests used as training data. The last column shows double-compression tests used as test data. Experimental data are plotted in gray, model predictions in green.

content.

Figure 5.16 compares the computationally predicted stresses in green and the experimentally recorded stresses in gray. The rows correspond to silken, firm, and extrafirm tofu, and the columns represent distinct loading cases with varying stretch magnitudes and rates. In contrast to Figure 5.10, the stress predictions now result from a single model that includes the water content as an explicit input feature. This unified model results in a smaller coefficient of determination with $R^2 = 0.6900$ compared to the single models with $R^2 = 0.7449$, but still captures the underlying trends effectively. Section 5.8.1 provides additional details about the errors. Although the model is trained only on synthetic data, it reproduces the major experimental trends: For firm and extrafirm tofu, in the second and third rows, the computationally predicted stresses align closely with the experimentally recorded stresses across all strain rates, especially during loading and unloading. During holding, the predicted stresses follow the experimental averages, but display slightly less nonlinearity than the experimental

measurements. For silken tofu, in the first row, the model overestimates the compressive peak stresses by a factor two, which is reflected in its negative R^2 -values. Despite this discrepancy, the silken tofu model captures the overall trends of nonlinear relaxation and decreasing stress magnitude with successive load cycles.

5.7. Discussion

Tofu is a nonlinear viscoelastic material. Tofu is an exceptional material made up of only soy beans and water. This simple yet tunable model system offers unique conditions to study composition–mechanics relationships. Our mechanical testing of three different tofu types across different water contents, strain magnitudes, and rates in Figures 5.8 and 5.9 reveals several interesting rheological characteristics: First, during the entirely loading path—loading, holding, and unloading—the response of tofu is markedly *nonlinear*. Second, during holding, stresses relax at constant stretch, and the response of tofu is notably *time-dependent*. During consecutive load cycles with similar peak stretches, the recorded peak stresses decrease, which supports the existence of stress relaxation. We conclude that the material behavior of tofu is made up of a history-independent *elastic* part and a history-dependent *inelastic* part. Our observations agree well with rheological tests of different foods throughout the past four decades [Lee et al., 1983; Rao and Steffe, 1992]. Specifically, our confined compression setup imposes non-permeable boundaries that prevent water from escaping the sample. This suggests that the observed stress relaxation arises primarily from the viscoelastic behavior of the solid [Budday et al., 2017a], rather than from the fluid-filled matrix [Franceschini et al., 2006]. We therefore first adopt the *theory of finite viscoelasticity* [Reese and Govindjee, 1998; Simo, 1987] to explain the observed nonlinear time-dependent response.

Tofu is loading-rate and water-content dependent. We further observe two general trends: stresses increase with increasing loading rate and with decreasing the water content. Figure 5.9 of the experimental data and Figure 5.10 of the discovered model highlight the loading-rate dependence from left to right and the water-content dependence from top to bottom. The *loading-rate dependence* agrees with our previous findings of fungi-based steak, where the stiffness increased by a factor five when varying the loading rate across four orders of magnitude [Vervenne et al., 2025b]. This trend is visible in Figures 5.9 and 5.10, although less pronounced, since our loading rates only vary by a factor three. The *water-content dependence* is more obvious, for example, in the first column, where mean peak stresses are lowest in silken tofu with -3.74 ± 0.86 kPa, followed by firm tofu with -18.91 ± 5.50 kPa, and extrafirm tofu with -42.84 ± 16.22 kPa. These stress ranges agree well with our previous tension, com-

pression, and shear tests of firm and extrafirm tofu with mean stiffnesses of 26.4 ± 4.0 kPa and 27.5 ± 5.6 kPa [St. Pierre et al., 2024] and with our rheologically measured storage moduli of 5.7 ± 0.5 kPa and 11.0 ± 1.0 kPa [Dunne et al., 2025]. Our observations also agree with previous viscoelastic tests on lab-made tofu with different water contents [Cheng et al., 2002], where the elastic, viscous, and relaxation parameters increased notably with decreasing water content [Cheng et al., 2005]. We can potentially explain these observations with the *theory of finite poroelasticity* [De Boer, 2012; Mow et al., 1984], and propose to supplement the theory of finite viscoelasticity with volume fractions of the solid and fluid phases.

Inelastic constitutive neural networks differentiate elastic and inelastic effects. Our experiments suggest that the constitutive behavior of tofu consists of elastic and inelastic parts. Yet, their individual contributions to the overall stresses remain poorly understood. To discover which fraction of the experimentally recorded stress is reversibly elastic and which is not, we utilize an *inelastic constitutive neural network* [Abdolazizi et al., 2024; Holthusen et al., 2024b] made up of feed forward networks for the equilibrium and non-equilibrium potentials combined with a recurrent neural network for the dissipative potential [Holthusen et al., 2024a]. The consistent agreement between our computationally predicted and experimentally measured stretch–strain curves, across all three tofu types, for both training and test data in Figure 5.10, confirms that our network accurately captures the essential mechanical features of tofu. From the visualization of the equilibrium and non-equilibrium contributions to the overall stress in Figure 5.14, we conclude that the network autonomously differentiates elastic and inelastic effects. This differentiation provides an important physical insight: The equilibrium stress increases, from -1.26 kPa for silken to -2.44 kPa for firm and -2.76 kPa for extra firm tofu, as the water content decreases. This trend mirrors the classical constitutive theory for fluids, where a perfect fluid at rest cannot support shear stresses. In other words, its rheological model lacks a Hookean spring and only consists of a Maxwell element. Several rheological models with a single Maxwell model, a generalized Maxwell model, and a four-element Maxwell model have been proposed to explain the stress relaxation of tofu with different water contents [Cheng et al., 2002]. Our automated discovery process captures these Maxwell-type characteristics [Holthusen et al., 2024a], and naturally reproduces the gradual transition from a viscoelastic solid to a nearly fluid-like response.

The elastic behavior of tofu is dominated by the second isochoric invariant. To characterize the isotropic elastic behavior of tofu, we have parameterized its equilibrium and non-equilibrium potentials in terms of *two isochoric invariants* and *one volumetric invariant* of the total and elastic right Cauchy Green tensors [Holzapfel, 2002], $\{ \tilde{I}_1^C, \tilde{I}_2^C, I_3^C \}$ and

$\{\tilde{I}_1^{C_e}, \tilde{I}_2^{C_e}, I_3^{C_e}\}$. Yet, their individual contributions to the two elastic potentials are unknown. To discover which invariants drive the equilibrium and non-equilibrium responses, we consult our elastic networks in Figure 5.3 to identify the most relevant contributions out of $2^9 = 512$ possible combinations of terms [Holthusen et al., 2024b,a]. Despite quantitative differences, the discovered models for all three tofu types, silken, firm, and extrafirm, share several interesting trends: Across all tofu types, both elastic potentials, equilibrium and non-equilibrium, are dominated by the *second isochoric invariants*, $\tilde{I}_2^C$ or $\tilde{I}_2^{C_e}$, which measure the isochoric area change. The equilibrium potential ψ^{eq} features a quadratic second-invariant term for silken and extrafirm tofu, and a linear second-invariant term for firm tofu. The non-equilibrium potential ψ^{neq} features a linear second-invariant term for all three tofu types, and an additional quadratic second-invariant term for extrafirm tofu. The complete *absence of exponential terms*—in both potentials and across all three tofu types—is in line with our general understanding that exponential terms represent the strain-stiffening of elastin and collagen fibers in soft biological tissues [Holzapfel, 2002], which are not present in tofu. The *dominance of the second invariant* [Kuhl and Goriely, 2024] is in line with previously discovered models for human brain [Linka et al., 2023a] and plant-based meat [Linka and Kuhl, 2024], and suggests that these hydrated soft materials share common rheological trends. Our results parallel classical constitutive modeling, where a single constitutive model can describe a family of materials simply by adjusting its parameter values. For example, in previous studies on lab-made tofu, a four-element Burgers model, a rheological model with two parallel dashpot-springs, was able to explain the rheological behavior of tofu with varying soymilk concentrations of $\{5\%, 6\%, 7\%, 8\%, 9\%\}$, simply by varying the elastic and viscous parameter values and the relaxation time [Cheng et al., 2005]. Our discovered models naturally reflect this principle: Tofu with varying solid volume fractions of $\{8\%, 13\%, 14\%\}$ results in varying parameterizations of the model, but not in fundamentally different functional forms.

The dissipative behavior of tofu is driven by the second and third deviatoric invariants. To characterize the dissipative behavior of tofu, we have parameterized its dissipative potentials in terms of *one volumetric invariant* and *two deviatoric invariants* of the Mandel stress [Holthusen et al., 2024b], $\{I_1^\Sigma, J_2^\Sigma, J_3^\Sigma\}$. Yet, their individual contributions to the dissipative potential are unknown. To discover which invariants drive the dissipative responses, we consult our dissipative network in Figure 5.4. Strikingly, across all three tofu types, the discovered dissipative potentials ϕ feature a single term that contains weighted powers of the second and third deviatoric invariants, J_2^Σ and J_3^Σ , embedded in two nested exponential functions. None of the discovered potentials depends on the volumetric invariant I_1^Σ , and none of them features any biases that would result in additive contributions. These observations confirm that our dissipa-

tive network autonomously eliminates irrelevant terms, and discovers sparse and interpretable models [McCulloch et al., 2024]: The first stress invariant I_1^Σ measures the hydrostatic stress, and its absence in the dissipative potential indicates a *pressure insensitivity*, unlike in the classical Drucker-Prager model [Drucker and Prager, 1952]. The second stress invariant J_2^Σ measures the magnitude of shear, and its presence in the dissipative potential indicates a *shear sensitivity*, like in the classical von Mises model [Mises, 1913]. The third stress invariant J_3^Σ measures the mode of shear via the ratio between the principal stresses, and its presence in the dissipative potential indicates a *stress ratio sensitivity*, like in the classical Tresca [Tresca, 1869] and Mohr-Coulomb models [Coulomb, 1776; Mohr, 1900]. While simple gels are often modeled as von-Mises type viscoelastic materials that depend exclusively on the second invariant, our dissipative network discovers a Lode-angle sensitive behavior [Lode, 1925] that depends on both the second and third invariants and distinguishes between tension-shear, pure shear, and compression-shear of the same shear magnitude [Bruhns, 2024]. These considerations underscore that our discovered models not only accurately reproduce the experimental behavior of tofu, but also offer interpretable, data-driven insight into its nonlinear viscoelastic behavior.

The mechanics of tofu depend nonlinearly on its water content. Tofu is one of the simplest man-made foods; it consists of only two ingredients: soy beans and water. We can easily tune its consistency by adjusting its processing parameters: the higher the pressure and the longer its duration, the lower the water content [Shurtleff and Aoyagi, 1975]. Making and testing tofu with varying water content provides valuable insight into how we can modulate the rheology of food simply by adding or removing water [Cheng et al., 2005]. From our three individually discovered tofu models, we conclude that the nonlinear and time-dependent behavior of tofu is highly sensitive to its water content. Yet, the explicit impact of water on the individual tofu models is unknown. A naive ad hoc explanation would be to consult the biphasic theory [Mow et al., 1984], adopt the concept of volume fractions [De Boer and Ehlers, 1990], associate the equilibrium behavior with the solid and the non-equilibrium behavior with the fluid, and weigh both with their individual volume fractions [Ehlers, 1993], $S = \nu_s \cdot S^{\text{eq}} + \nu_f \cdot S^{\text{neq}}$. In Table 5.1, the volume fractions of the fluid and solid are $\nu_f = \{0.92, 0.87, 0.86\}$ and $\nu_s = 1 - \nu_f = \{0.08, 0.13, 0.14\}$, for silken, firm, and extrafirm tofu. We can easily see that the *biphasic theory alone fails* to explain the observed stiffness variations: For example, the peak stresses in the first column of Figure 5.9, $S_{\text{peak}} = \{-3.74 \pm 0.86 \text{ MPa}, -18.91 \pm 5.50 \text{ MPa}, -42.84 \pm 16.22 \text{ MPa}\}$ vary by more than an order of magnitude across the three tofu types, while the volume fractions vary by only six percent. This agrees with our previous observations of firm and extrafirm tofu for which

the stiffness, storage, and loss moduli varied by a factor two, while the water contents differed by only one percent [Dunne et al., 2025]. We conclude that volume-fraction-based scaling might be appropriate for linear mixtures with low-to-moderate water contents [De Boer, 2012], but a more complex nonlinear water-content dependence is critical to accurately model highly hydrated soft materials like tofu.

Feature network modeling reveals correlations between water content and mechanics.

Feature networks learn nonlinear mappings from raw input features to weights that parameterize or modulate terms of another network model. To explore how water content modulates the mechanics of tofu, we introduce a feature network that takes the scalar-valued water content α as input, and outputs the weights for each node of the three mechanical networks. Our results in Figure 5.15 confirm that the individual terms of the equilibrium, non-equilibrium, and dissipative potentials display different types of water-content dependency: Two terms are entirely independent of the water content, while four increase and two decrease with increasing water content. We conclude that incorporating the *water content as a feature* can successfully introduce material-specific variability into a unified constitutive model. The close agreement between the computationally predicted and experimental recorded stresses in Figure 5.16, especially for firm and extrafirm tofu, confirms that the feature network can indeed effectively modulate the mechanical response. Comparisons with other hydrated soft biological tissues suggest that neither a purely viscoelastic [Budday et al., 2017b], nor a purely poroelastic model [Franceschini et al., 2006] can accurately capture this characteristic behavior, but that a poroviscoelastic model is needed instead [Comellas et al., 2020]. Indeed, hydrated soft solids including the cornea, meniscus, intervertebral disk, articular cartilage, and hydrogels have long been known to display a complex nonlinear poroviscoelastic behavior [Holmes and Mow, 1990]. Our feature network provides the inherent flexibility to model these complex rheologies. It naturally combines microstructural information and macrostructural experiments into a data driven, scalable approach that represents families of hydrated soft materials within a single feature-weighted model. More generally, our results suggest that feature network modeling is a flexible approach to integrate multimodal data and discover how various multiphysics inputs—water content, density, temperature, or fiber angles—modulate the mechanics and physics of soft matter systems.

Limitations. We note a few limitations of our study: First, we have assumed that the *inelastic deformation* is isotropic and *volume preserving*, and that volumetric changes are associated with the elastic deformation only [Holzapfel, 2002; Reese and Govindjee, 1998; Simo, 1987]. In a follow up study, we are currently performing both confined and unconfined compression tests

on tofu to explore to which extent this assumption is justified. Second, our currently findings are restricted the *compressive loading* only. Although compression is the dominant loading mode during chewing and most relevant to our sensory perception of texture [Vervenne et al., 2025a], we have previously tested firm and extrafirm tofu in compression and shear [Dunne et al., 2025], and in tension, compression, and shear [St. Pierre et al., 2024]; yet, testing silken tofu in tension remains challenging. Third, we have only tested over a *small range of loading rates*. Testing food at loading rates across several orders of magnitude and observing fluid flow can help delineate poroelastic and viscoelastic phenomena [Vervenne et al., 2025a]. Fourth, we have assumed that water in tofu exists only in a single state [De Boer, 2012], freely moving. In reality, tofu is a hydrated soy protein gel that contains water in three distinct states, *bound*, *immobilized*, and *free* [Yuan et al., 2025], where bound and immobilized water modulate the elastic response, while free water modulates the inelastic response [Lucey, 2002]. A next logical step would be to characterize the volume fractions of bound, immobilized, and free water [Kainz et al., 2023, 2025], then integrate all three volume fractions as distinct features of the network, and discover how each contribution modulates the behavior of tofu, and of hydrated soft solids in general. Finally—and this is probably more a philosophical question than a true limitation—in line with the literature [Brunton et al., 2016; Flaschel et al., 2021; Linka and Kuhl, 2023], we have used the terminology *discover* in the context of discovering sparse, interpretable functional representations of potentials within a pre-specified constitutive ansatz. Following the common approach, we reverse-engineer this pre-defined ansatz from popular constitutive terms and functions and generalize it to a library of activation functions that satisfy thermodynamics and physical constraints by design [Kalina et al., 2023b; Klein et al., 2023; Masi et al., 2021].

Conclusion

Tofu is a hydrated protein gel composed only of soybeans and water. It provides a minimal model system to explore how material composition governs the mechanics and physics of food. A series of compression tests of three tofu types reveals its strong nonlinearity, its pronounced stress relaxation, and its more than ten-fold increase in peak stress with only six percent decrease in water content. Using automated model discovery, we discover inelastic constitutive models and parameters that accurately explain both our training and test data of silken, firm, and extrafirm tofu across three loading magnitudes and five loading rates. The discovered models autonomously separate elastic and inelastic behavior and reveal that the elastic response is dominated by the second isochoric invariant, while the inelastic response is governed by a

combination of the second and third deviatoric invariants. Water content modulates the magnitude of these mechanisms, but not their functional form. This motivates the integration of all results into a single water-content weighted model for different tofu types. Our results position tofu as a simple quantitative benchmark for nonlinear poroviscoelastic solids and demonstrate how physics-informed machine learning can expose the underlying constitutive structure of hydrated soft materials. Knowing the precise rheology of tofu is essential to reliably reproduce, engineer, and design the texture of alternative protein products in the benefit of human and planetary health. More broadly, our results suggest that combined feature-mechanics networks can uncover complex multiphysics behavior in hydrated soft materials, including biopolymer hydrogels, engineered protein networks, and hydrated porous composites.

5.8. Appendix

5.8.1. Errors of discovered individual and unified models

To quantify the performance of our discovered models, we calculate the error between the stress of the model S and the data $\hat{S}$ across all data points n_{data} . In addition to the coefficients of determination, $R^2 = 1 - \sum_{i=1}^m [S_i - \hat{S}_i]^2 / \sum_{i=1}^m [S_i - \bar{S}]^2$ in Figures 5.10 and 5.16, we report the mean absolute error, $\text{MAE} = \sum_{m=1}^{n_{\text{data}}} |S_m - \hat{S}_m| / n_{\text{data}}$, the root mean squared error, $\text{RMSE} = [\sum_{m=1}^{n_{\text{data}}} (S_m - \hat{S}_m)^2 / n_{\text{data}}]^{1/2}$, normalized root mean squared error, $\text{NRMSE} = \text{RMSE} / \max\{\hat{S}\}$, where $\max\{\hat{S}\}$ denotes the maximum experimental stress, and the relative L^2 error, $L^2 = [\sum_{i=1}^m [S_i - \hat{S}_i]^2]^{1/2} / [\sum_{i=1}^m S_i^2]^{1/2}$. Table 5.2 summarizes the errors for the discovered individual tofu models in Figure 5.10 and for the discovered unified tofu model in Figure 5.16. To interpret the sources of error, Fig. 5.17 highlights the training performance on individual curves extracted from the datasets. We incrementally increase the amount of training data, from slow loading, to slow and medium, to slow, medium and fast, from top to bottom. Fig 5.18 shows the model prediction vs. ground truth data for these three training scenarios. For training on a single curve, top row, the fit is nearly perfect, and it is evident that the remaining error originates primarily from the data themselves. The performance on the training data with $R^2 = 0.9543$ is notably better than the performance on the test data with $R^2 = 0.7287$ and $R^2 = 0.8416$. For training on two curves, middle row, the fit of the second curve improves significantly, from $R^2 = 0.7287$ to $R^2 = 0.9495$, while the first shows a minor reduction in fit quality. For training on three curves, bottom row, the fit of the third curve improves, from $R^2 = 0.8416$ to $R^2 = 0.9457$, while the first and second curve fits remain largely unchanged. The plain testing performance, last column, remains weaker

Table 5.2.: **Errors of discovered models.** Mean absolute error MAE, root mean squared error RMSE, normalized root mean squared error NRMSE, and relative L^2 error for the discovered individual tofu models in Figure 5.10, bottom, and for the discovered unified tofu model in Figure 5.16, top.

individual	error	slow	medium	fast	test
silken	RMSE	0.633	0.298	0.460	1.206
	MAE	0.402	0.223	0.322	0.970
	NRMSE	0.118	0.095	0.133	0.205
	L^2	0.381	0.228	0.344	0.479
firm	RMSE	2.911	2.891	2.476	4.203
	MAE	1.871	1.684	1.625	2.707
	NRMSE	0.112	0.063	0.113	0.103
	L^2	0.361	0.352	0.836	0.496
extrafirm	RMSE	5.746	4.877	5.200	4.644
	MAE	3.400	3.386	3.659	3.922
	NRMSE	0.089	0.147	0.080	0.148
	L^2	0.402	0.521	0.419	0.411

unified	error	slow	medium	fast	test
silken	RMSE	1.473	2.318	2.643	1.633
	MAE	1.182	1.893	2.150	1.435
	NRMSE	0.274	0.742	0.762	0.277
	L^2	0.8855	1.780	1.973	0.6481
firm	RMSE	2.867	3.140	3.056	3.880
	MAE	2.131	2.231	2.321	2.728
	NRMSE	0.111	0.068	0.141	0.095
	L^2	0.3557	0.3821	0.3729	0.3554
extrafirm	RMSE	7.632	4.030	5.405	4.080
	MAE	4.982	3.290	3.642	3.297
	NRMSE	0.118	0.122	0.083	0.131
	L^2	0.5342	0.4307	0.4357	0.3606

than the slow, medium, and fast columns, since the model is evaluated on previously unseen loading magnitudes and rates. Interestingly, the predictions remain accurate for slower loading rates, while larger deviations occur at higher loading rates. This low predictive performance indicates that the current training data do not sufficiently capture the full range of the material behavior under different rates. Finally, the reported R^2 values in Figures 5.10 and Figures 5.16 and in Table 5.2 additionally reflect the fact that here, a single model represents the collective response of $n = 8$ to $n = 10$ individual curves, with a large variance between individual curves. Taken together, from Figures 5.17 and 5.18, we conclude that the R^2 value

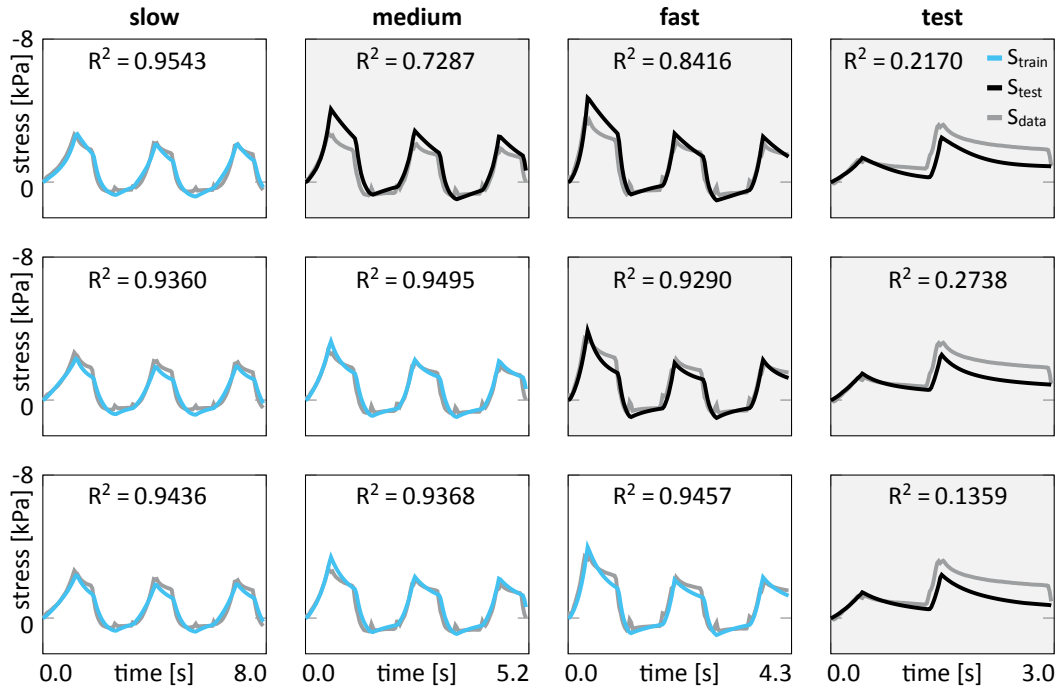


Figure 5.17.: **Sensitivity of error measures for varying training sets.** Single stress histories $S(t)$ for silken tofu trained on the slow strain rate (top), slow and medium rates (middle), and slow, medium, and fast rates (bottom). Predictions during training shown in blue, testing in black.

for training on single curves is on the order of 0.95%, and that the remaining error of 5% reflects measurement inconsistencies in the training data. The R^2 values for testing on single curves vary between 0.73% and 0.93%, and the remaining error reflects the collective effects of measurement inconsistencies and prediction errors to unseen loading rates. This error increases further when making predictions into unseen loading magnitudes and rates. The seemingly low R^2 values in Figures 5.10 and Figures 5.16 and in Table 5.2 reflect a *collective result of measurement inconsistencies in single curves, a high variance across different curves, and prediction errors into previously unseen regimes.*

5.8.2. Preliminary sensitivity analyses

Prior to generating the main results of this study, we performed a systematic series of sensitivity analyses. Here we showcase three exemplary studies that, amongst many others, have guided our selection of network architectures, activation functions, and hyperparameters.

Network architectures. Figure 5.19 summarizes our preliminary sensitivity analysis for varying network architectures. Specifically, we study a finer network with $\{30, 16, 12\}$ neurons

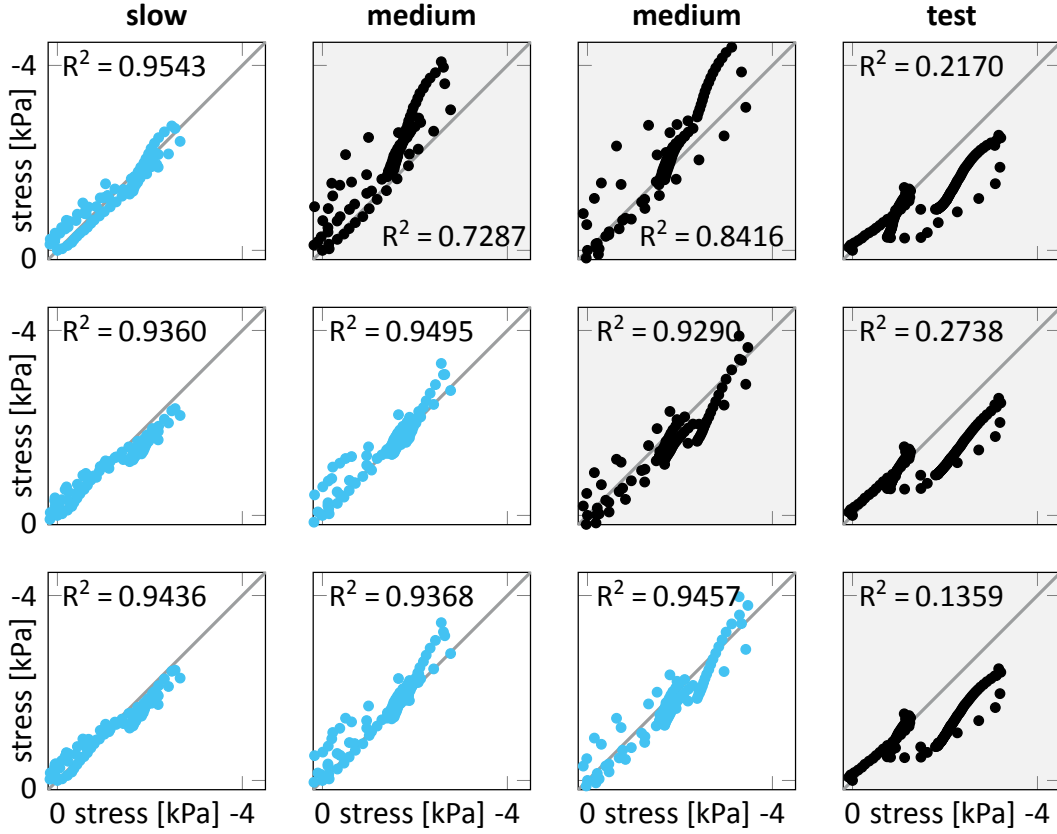


Figure 5.18.: **Sensitivity of error measures on varying training sets.** Ground truth vs. model predictions for silken tofu trained on the slow strain rate (top), slow and medium rates (middle), and slow, medium, and fast rates (bottom), corresponding to Fig. 5.17. Predictions during training shown in blue, testing in black.

per layer and a coarser network with $\{4, 4, 2\}$ neurons per layer, compared to our reference network with $\{18, 8, 8\}$ neurons per layer that achieved a goodness of fit of $R^2 = 0.7449$. The finer $\{30, 16, 12\}$ network results in a reduced goodness of fit of $R^2 = 0.7085$, associated with the stress predictions in red. The coarser $\{4, 4, 2\}$ network results in a similarly reduced goodness of fit of $R^2 = 0.7091$, associated with the stress predictions in orange. Interestingly, both network architectures lead to numerical instabilities, most visible in the results for firm and extrafirm tofu. These instabilities could be a result of time-step sensitivity related to the explicit time integration and might require a reduced time step size or more steps in the staggered training approach.

Activation functions. Figure 5.20 summarizes our preliminary sensitivity analysis for varying activation functions. Specifically, we vary the activation functions in the first layer of the dissipative network by fixing the powers to $p = 1$ and $p = 2$, instead of leaving p as a trainable

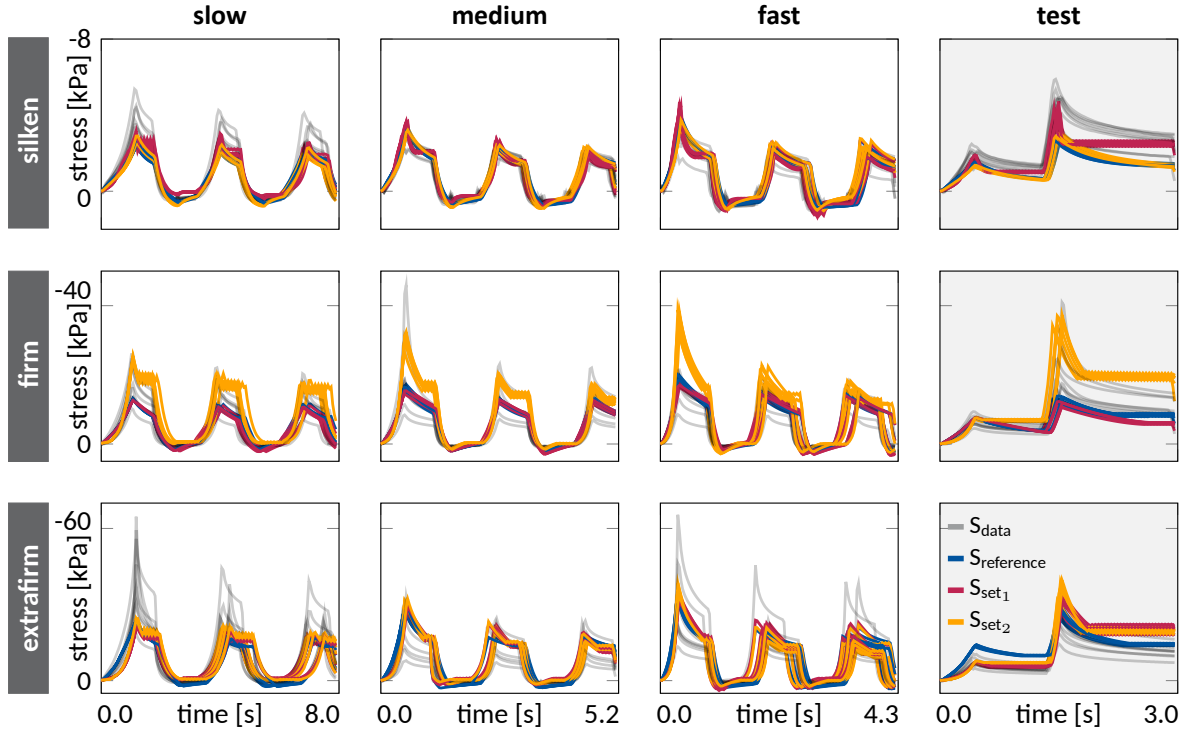


Figure 5.19.: **Preliminary sensitivity analysis for varying network architectures.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 5.10 with $\{18, 8, 8\}$ neurons per layer are plotted in blue; predictions with $\{30, 16, 12\}$ and $\{4, 4, 2\}$ neurons per layer are shown in red as set 1 and in orange as set 2.

parameter which achieved a goodness of fit of $R^2 = 0.7449$. Fixing the powers to $p = 1$ and $p = 2$ reveals similar trends and a notable reduction of the goodness of fit to $R^2 = 0.7212$ for $p = 1$, associated with the red curves, and even further to $R^2 = 0.6586$ for $p = 2$, associated with the orange curves. For the latter, non-linear effects almost vanish entirely for firm and extrafirm tofu. Overall, our selection of activation functions is guided by thermodynamical considerations, and is in line with previous elastic [Linka and Kuhl, 2023; Linka et al., 2023a,b] and dissipative [Holthausen et al., 2024b, 2025a,b] networks similar to Figures 5.3 and 5.4 that have demonstrated high robustness and reproducibility.

Regularization parameters. Figure 5.21 summarizes our preliminary sensitivity analysis for varying regularization coefficients α_1 and α_2 that fine-tune the level of L_1 and L_2 regularization. Specifically, we set both hyperparameters a magnitude higher and lower than our selected reference values of $\alpha_1 = \alpha_2 = 0.0001$ that achieved a goodness of fit of $R^2 = 0.7449$. Higher regularization coefficients of $\alpha_1 = \alpha_2 = 0.001$ result in a reduced goodness of fit of

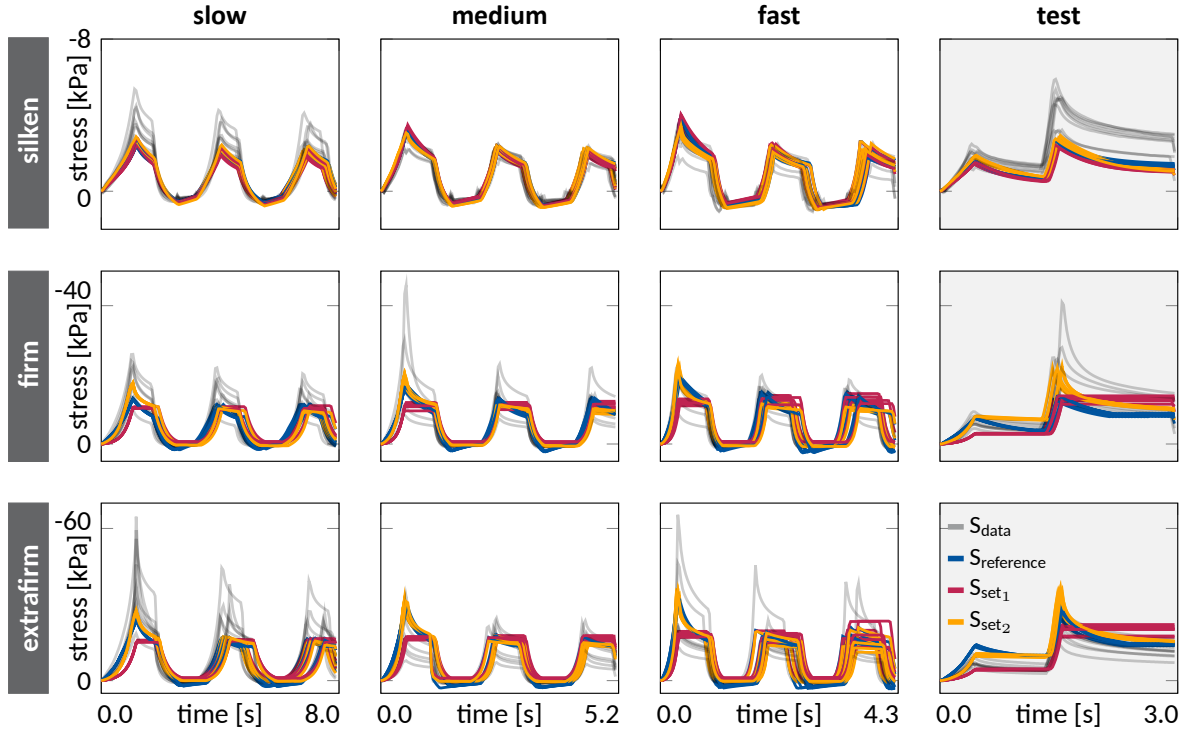


Figure 5.20.: **Preliminary sensitivity analysis for varying activation functions.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 5.10 with trainable parameter p are plotted in blue; predictions with fixed parameters $p = 1$ and $p = 2$ are shown in red as set 1 and in orange as set 2.

$R^2 = 0.6556$ with stress predictions that display almost no nonlinearity for firm and extrafirm tofu, shown in red. Lower regularization coefficients of $\alpha_1 = \alpha_2 = 0.00001$ result in a similar goodness of fit of $R^2 = 0.7510$ as the reference model, with stress predictions, shown in orange, that predict similar features as the reference model, shown in blue.

Taken together, these results illustrate the sensitivity of network training with respect to the network architecture, activation functions, and regularization parameters. While we only show a few selected cases here, we have performed a wide variety of preliminary studies and recommend systematic sensitivity analyses before training the final model.

Data availability

Our source code and examples are available at <https://doi.org/10.5281/zenodo.16993236>.

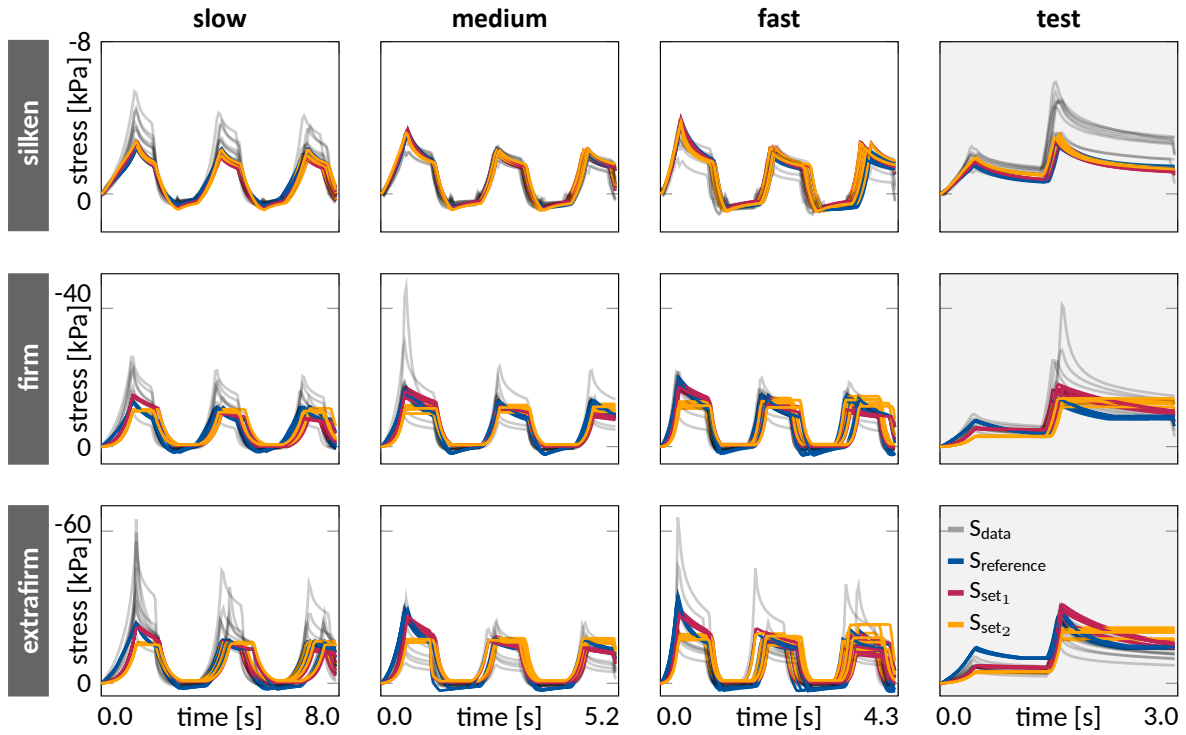


Figure 5.21.: **Preliminary sensitivity analysis for varying regularization parameters.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 5.10 with regularization parameters $\alpha_1 = \alpha_2 = 0.0001$ are plotted in blue; predictions with $\alpha_1 = \alpha_2 = 0.001$ and $\alpha_1 = \alpha_2 = 0.00001$ are shown in red as set 1 and in orange as set 2.

6 | Conclusions and Outlook

The present dissertation was concerned with the modeling of sustainable materials in the sense of constitutive modeling aided by data-driven frameworks. Particular attention was paid to the phenomena of plasticity, visco-elasticity, anisotropic material behavior, and compositional properties, particularly hydration. While various materials have been considered, focus was oriented on fibrous materials such as paper and paperboard and tofu as a meat-replacement product.

Decoupled anisotropic constitutive modeling for fibrous materials. A novel continuum mechanical framework was developed that enables a thermodynamically consistent decoupling of in-plane and out-of-plane material behavior at finite strains. Based on the use of structural tensors and invariant-based formulations, the approach overcomes limitations of classical parameter-driven decoupling and provides a general and flexible modeling concept for anisotropic materials. The framework was systematically extended to inelasticity, demonstrating that direction-specific dissipation mechanisms can be introduced and controlled through distinct potential formulations. In particular, the introduction of a dedicated out-of-plane inelastic potential completes the framework by enabling the independent representation of dominant thickness-direction dissipation mechanisms, without artificial coupling to in-plane behavior. This establishes a modular constitutive framework that allows the targeted modeling of anisotropic and direction-dependent inelastic effects, which are characteristic of fibrous materials such as paper and paperboard.

An important aspect for future developments lies in extending the framework toward fully coupled anisotropic inelastic models that incorporate additional mechanisms such as plasticity and in-plane anisotropy. In particular, fibrous materials exhibit pronounced anisotropy not only in the thickness direction but also within the in-plane directions, which requires more comprehensive three-dimensional formulations. Furthermore, the integration of multi-physics effects, such as moisture and temperature, is essential for realistic process simulations, as these factors significantly influence the mechanical response in industrial applications. Another promising direction is the development of fiber–matrix-based approaches to establish direct links between microstructure and macroscopic material behavior.

Extension of data-driven approaches to plasticity. To address the challenge of identifying suitable constitutive formulations for complex material behavior, the framework of inelastic constitutive artificial neural networks (iCANNs) was extended to elasto-plasticity. This extension enables the simultaneous identification of elastic and plastic contributions, including yield behavior, flow rules, and hardening mechanisms, directly from data. By combining potential-based formulations with neural networks, the approach preserves thermodynamic consistency and fundamental continuum mechanical principles while allowing for automated model discovery. The integration of a return-mapping algorithm ensures the enforcement of plastic admissibility conditions during training and leads to stable and physically consistent predictions.

Beyond predictive capabilities, the framework provides the possibility to systematically analyze the learned material behavior. In particular, the identification of relevant invariants governing the mechanical response enables the extraction of physically meaningful insights from the training process. The extension of iCANNs to plasticity therefore represents a crucial step toward completing the framework's capability to capture a broader class of inelastic phenomena, as plasticity constitutes a fundamental mechanism in many engineering materials.

Future work should focus on extending the framework toward more general plasticity formulations, including anisotropic yield criteria, combined isotropic and kinematic hardening, and more complex evolution laws. In addition, improving robustness and data efficiency for experimental datasets remains an important challenge. The integration of such approaches into finite element simulations and the development of standardized workflows for automated model discovery are further key steps toward practical applicability.

Hydrated materials modeled using data-driven approaches. Building on the developed data-driven framework, an experimental and computational study of tofu as a representative hydrated soft material was conducted. Using automated model discovery, constitutive models were identified that capture strongly nonlinear, rate-dependent, and composition-dependent behavior directly from experimental data. The results demonstrate that the mechanical response can be described by a small number of dominant invariants, enabling a clear separation of elastic and inelastic contributions.

A key finding is that the material behavior is strongly governed by the water content, which influences the mechanical response in a highly nonlinear manner. This dependency cannot be adequately described by classical mixture models. Instead, the introduction of feature-dependent neural network formulations allows the identification of nontrivial relationships between composition and mechanical behavior. This enables the formulation of unified constitutive models that generalize across different material states and compositions.

Future developments should focus on extending this approach to incorporate multiple interacting features, such as temperature, microstructure, or additional compositional variables. Furthermore, the combination with multiphase and poromechanical modeling approaches offers significant potential for improving the physical interpretation of hydrated materials. Beyond tofu, the methodology provides a promising pathway for analyzing and designing a wide range of bio-based and soft materials, contributing to a more systematic understanding of structure-property relationships in complex material systems.

While the theoretical methodologies are profound, future problems lie in the transition from theoretical model development toward applicability of these approaches. Current data-driven methods often remain user-dependent and require expert knowledge, which limits their applicability. Therefore, research should aim to develop frameworks that enable a seamless workflow from experimental data to predictive constitutive models.

In this context, the use of full-field experimental data offers significant potential, especially for enabling inverse material design and improving the quality of learned models. Furthermore, the extension of these approaches to a broad range of materials and applications is essential, particularly in fields where accurate material modeling is critical. Another key objective is the development of efficient and real-time capable models, which would allow data-driven constitutive approaches to be integrated into industrial processes and large-scale simulations. Overall, future work should focus on leveraging the ongoing advances in machine learning to facilitate automated model discovery, while shifting attention toward the interpretation and utilization of the extracted information. This will enable not only improved predictive capabilities, but also a deeper understanding of material behavior, ultimately contributing to more efficient resource usage and the optimization of sustainable material systems.

A | Appendix

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work AI tools were used in order to improve readability and for language refinement. After using these tools, the content was reviewed and edited as needed and full responsibility for the content of this thesis is taken by the author.

Declaration of adapted notations

The notation of the original publication in Chapter 5 has been slightly adapted for consistency throughout this dissertation. These changes are purely notational and include the transpose of a tensor and the tensor product of two second order tensors. This does not affect the content or results of the original works.

List of Figures

2.1. Illustration of preferential material directions – in-plane area (<i>ip</i>) underlined in grey and out-of-plane direction (<i>op</i>).	22
2.2. Visualization of multiplicative decomposition of the deformation gradient $\mathbf{F}$ into $\mathbf{F}_e$ and $\mathbf{F}_i$ multiplicative decomposition of $\mathbf{F}_i$ into $\mathbf{R}_i$ und $\mathbf{U}_i$ and the corresponding configurations: <i>rc</i> - reference configuration, <i>cc</i> - current configuration, <i>ic</i> - intermediate configuration, <i>cic</i> - co-rotated intermediate configuration [cf. Holthusen et al., 2023].	32
2.3. Boundary conditions of the single-element tests: (a) uniaxial tension/compression (black - <i>ip</i> , blue - <i>op</i>), (b) in-plane simple shear, (c) out-of-plane simple shear.	46
2.4. Single-element study on the elastic material formulation: Components of the Green-Lagrange strain $\mathbf{E}$ at uniaxial tension (left), uniaxial compression (middle), and simple shear (right) in out-of-plane direction (upper line) and in-plane direction (lower line).	47
2.5. Cube under compression: Geometry, discretization (10x10x10 elements), and boundary conditions - surface loading applied on the white top area.	48
2.6. Convergence study of cube under compression: Displacement in x_3 direction at node 1 over the number of elements per edge.	48
2.7. In-plane displacement (magnitude of x_1 - and x_2 -displacement) for the new split (Model I, left) and classical split (Model II, right) (Scaling factor 2). . .	49
2.8. Out-of-plane displacement for the new split (Model I, left) and classical split (Model II, right).	50
2.9. Load-displacement-curve at node 1	50
2.10. Evolution of the components of the inelastic right Cauchy-Green strain for uniaxial tension in out-of-plane direction (x_3 -direction) (upper line) and in-plane direction (x_1 -direction) (lower line).	52
2.11. Geometry and boundary conditions of the pinched cylinder.	53

2.12. Convergence study of pinched cylinder: Sum over all reaction forces F at edge $x_3 = 0$ over the number of elements: (a) $2 \times 4 \times 8$, (b) $2 \times 8 \times 16$ (c) $2 \times 16 \times 32$, (d) $2 \times 24 \times 48$, (e) $2 \times 32 \times 64$, (f) $2 \times 36 \times 72$ (left) and the converged geometry discretization (e) (right).	53
2.13. Out-of-plane component, $C'_{i_{op}}$, of the inelastic right Cauchy Green deformation tensor computed with potential 1 (left), potential 2 (middle) and potential 3 (right).	54
2.14. In-plane component, $C'_{i_{ip}}$, of the inelastic right Cauchy Green deformation tensor computed with potential 1 (left), potential 2 (middle) and potential 3 (right).	55
2.15. Geometry and boundary conditions of the symmetrically notched specimen with a depth of 2.5 [mm].	55
2.16. Convergence study of symmetrically notched specimen: Sum of reaction forces in x_2 -direction F_{x_2} over the number of elements (left) and final discretization (right).	56
2.17. Cauchy stress σ_{22} (upper line) and Green-Lagrange strain E_{33} (lower line) for elastic material (a), inelastic material with potential 1 (b), potential 2 (c), and potential 3 (d).	58
2.18. Reaction forces in x_2 -direction, F_{x_2} , at $x_2 = 20$ mm over displacement w in x_2 -direction.	59
3.1. Illustration of preferential material directions – in-plane area (<i>ip</i>) underlined in grey and out-of-plane direction (<i>op</i>).	72
3.2. Geometry and boundary conditions of the pinched cylinder.	77
3.3. In-plane component of the inelastic right Cauchy-Green deformation tensor, $C'_{i_{ip}}$, computed with Potential 1 (left), Potential 2 (middle) and Potential 3 (right).	78
3.4. Out-of-plane component of the inelastic right Cauchy-Green deformation tensor, $C'_{i_{op}}$, computed with Potential 1 (left), Potential 2 (middle) and Potential 3 (right).	79
4.1. Rheological representation of the iCANN capturing plasticity consisting out of three springs, ψ_e , ψ_p , and ψ_{p_e} . A slider, Φ and g_1 , and a dashpot element, g_2 , with the states U_p and U_{p_i} that are captured through the strain tensors C_p and $\bar{B}_{p_e}$, are connected in parallel.	90

- 4.2. Flowchart of the recurrent neural network (outer gray box) called at each time step. First, all constitutive equations from the last time step t_n (upper gray box) are computed. Internal variables are explicitly updated to compute all constitutive equations of the current time step t_{n+1} (lower gray box). The index $n + 1$ is omitted for readability. Feed-forward networks of the Helmholtz free energies are indicated by orange boxes, while the ones of the plastic potentials are shown by blue boxes. Inputs, outputs and states are visualized by blue trapezoids. 93
- 4.3. Schematic illustration of the feed-forward network of the elastic and plastic Helmholtz free energies. Elastic formulation: $\mathbf{A} = \bar{\mathbf{C}}_e$ and $a = \psi_e(\bar{\mathbf{C}}_e)$, plastic formulations: $\mathbf{A} = \mathbf{C}_p$ and $a = \psi_p(\mathbf{C}_p)$, as well as $\mathbf{A} = \bar{\mathbf{B}}_{p_e}$ and $a = \psi_{p_e}(\bar{\mathbf{B}}_{p_e})$ 95
- 4.4. short 98
- 4.5. Artificially generated data: Stress strain curves of uniaxial tension (1st row), equibiaxial tension (2nd row), uniaxial compression (3rd row), cyclic loading (4th row), and corresponding loss curves (5th row). Trained on (a) uniaxial tension, (b) uniaxial tension and equibiaxial tension, (c) uniaxial tension, equibiaxial tension, uniaxial compression. Early stopping was enforced after 50 consecutive epochs of non-decreasing loss values. 104
- 4.6. Artificially generated data: Stress strain curves of uniaxial tension (1st row), uniaxial tension with unloading (2nd row), cyclic loading (3rd row), and corresponding loss curves (5th row). Trained on (a) uniaxial tension, (b) uniaxial tension with unloading, (c) cyclic loading. Early stopping was enforced after 50 consecutive epochs of non-decreasing loss values. 105
- 4.7. Artificially generated data with tension-compression asymmetry: Underlying yield surface (blue) and predicted yield surfaces (orange) in 2D (upper row) and 3D (lower row) in the stress space (in kPa) for training on (a) uniaxial tension, (b) uni- and equibiaxial tension, and (c) uniaxial tension and compression (hydrostatic axis is given for reference in gray). 108
- 4.8. Training on data taken from Liao and Luo [2025]. Stress-stretch curves for training on the first tensile loading (1st row), tensile and compressive loading (2nd row), and on the first cycle (3rd row). Data points used for training are shown as green markers and data points used for testing as orange markers. Model predictions are given as solid blue lines. 110

- 4.9. Experimental data of DP600 steel taken from Cao et al. [2009]: Stress curves over time under cyclic tension-compression loading for training on the first tension loading (upper line), the first tension and compression loading (middle line), and tension-compression-tension loading (lower line). Data points used for training are shown as green markers and data points used for testing as orange markers. Model predictions are given as solid blue lines. 111
- 4.10. Comparison of regularization schemes – Training on data taken from Liao and Luo [2025]. Stress-stretch curves for training on the first tensile loading (1st row), tensile and compressive loading (2nd row), and on the first cycle (3rd row). Training data (green marker) and model prediction for regularization with regularization of $L_2 = 0.0001$ (blue, solid), $L_1 = 0.0001$ (red, dashed), and $L_1 = 0.01$ (gray, dotted). Testing data (orange marker) is given on the right. 116
- 5.1. **Five different types of tofu.** Extra-firm, firm, medium, soft, and silken tofu illustrate how small changes in composition and processing create large differences in mechanical behavior. Tofu combines soy beans with water to form a biphasic protein gel whose stiffness, porosity, and fracture characteristics arise directly from its water content and its coagulation pathway. Whether we make our own tofu at home or obtain it through commercial retail distribution, we encounter the same continuum of textures, from a dense cohesive protein matrix to a nearly fluid-like gel. 123
- 5.2. **Mechanical testing.** Preparation of cylindrical samples; triple compression test for training; double compression test for testing; and rheometer with mounted sample. During triple compression, we load and unload between $\lambda = 0.7$ and $\lambda = 0.9$ at slow, medium, and fast stretch rates of $\dot{\lambda} = \{0.2/\text{s}, 0.4/\text{s}, 0.6/\text{s}\}$ and hold for 0.5s in between. During double compression, we load to $\lambda = 0.8$ at $\dot{\lambda} = 0.3/\text{s}$, hold for 1.0s, load to $\lambda = 0.7$ at $\dot{\lambda} = 0.5/\text{s}$, and hold for 1.0s. White regions represent loading and unloading, grey regions represent holding. 125

- 5.3. **Elastic networks.** The elastic networks have two hidden layers with 5 and 9 nodes per layer. The feed forward networks take the invariants of the total right Cauchy Green tensor $\tilde{I}_1^C, \tilde{I}_2^C, I_3^C$, left, or of the elastic right Cauchy Green tensor $\tilde{I}_1^{C^e}, \tilde{I}_2^{C^e}, I_3^{C^e}$, right, as input; the first layer applies first or second powers, the second layer applies the identity, exponential function, or natural logarithm; the network learns the equilibrium or non-equilibrium potentials ψ^{eq} , left, or ψ^{neq} , right, as output. Activation functions $f_{\beta,\gamma}^\delta$ are labeled by activation type β , node index γ , and layer index δ ; dashed ellipses denote constant, non-trainable biases; dashed lines indicate fixed weights equal to one. 131
- 5.4. **Dissipative network.** The dissipative network has three hidden layers with 18, 8, and 8 nodes per layer. It take the invariants of the Mandel stress $I_1^\Sigma, J_2^\Sigma, J_3^\Sigma$ as input; the first layer applies trainable powers p to the absolute input values, the second and third layers apply the maximum or exponential functions; the network learns the dissipative potential ϕ as output. Activation functions $f_{\beta,\gamma}^\delta$ are labeled by activation type β , node index γ , and layer index δ ; biases b^δ are applied to the maximum functions in the second and third hidden layers. . . . 132
- 5.5. **Feature network.** The feature network has three hidden layers with $k \times 4$ nodes per layer. It take the feature α , in our case the scalar-valued water content, as input; the three hidden layer apply the maximum, identity, hyperbolic tangent, and softplus functions; the output of the final layer applies the maximum function; the network learns the feature weight vector F^{max} as output. Activation functions $f_{\beta,\gamma}^\delta$ are labeled by activation type β , node index γ , and layer index δ ; biases b^δ are applied to all maximum functions. 133
- 5.6. **Coupling of mechanical and feature networks.** The mechanical network from Figure 5.3 or 5.4 receives a strain- or stress-like input $\mathbf{A}$ and the feature network from Figure 5.5 receives a feature-like input α . The networks are multiplicatively coupled in the nodes of their final layers. The sum of the final layer defines the network output, the feature-weighted mechanical potential $\Psi(\mathbf{A}, \alpha)$. Activation functions f_γ^δ and F_γ^δ of the mechanical and feature networks are labeled by node index γ and layer index δ 134

- 5.7. **Generating synthetic training data.** After discovering the best models and parameters for for silken, firm, and extrafirm tofu by training the equilibrium and non-equilibrium networks in Figures 5.3 and the dissipative network in Figure 5.4 on the experimental data, we generate synthetic data to train the coupled mechanical and feature networks in Figure 5.6. We evaluate the three tofu models for sets of invariants I_i^A , and their derivatives $\hat{\Psi}/\partial I_i^A$ provide the training data for the unified feature–mechanical model. 137
- 5.8. **Mechanical testing - Experimentally applied stretches.** Stretch histories $\lambda(t)$ during triple-compression and double-compression tests for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests at slow, medium, and fast strain rates of 20%/s, 40%/s, and 60%/s (from left to right), where each specimen is loaded to a stretch of 0.7 and unloaded to 0.9. The last column shows the double-compression tests at a strain rates of 30%/s to a stretch of 0.8 followed by a strain rate of 50%/s to 0.7. 138
- 5.9. **Mechanical testing - Experimentally recorded stresses.** Stress histories $S(t)$ during triple-compression and double-compression tests for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests at slow, medium, and fast strain rates of 20%/s, 40%/s, and 60%/s (from left to right), where each specimen is loaded to a stretch of 0.7 and unloaded to 0.9. The last column shows the double-compression tests at a strain rates of 30%/s to a stretch of 0.8 followed by a strain rate of 50%/s to 0.7. 139
- 5.10. **Model discovery - Experimentally recorded and computationally predicted stresses of the three individual tofu models.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests used as training data. The last column shows double-compression tests used as test data. Experimental data are plotted in gray, model predictions for silken, firm, and extrafirm tofu in light to dark blue. 140
- 5.11. **Model discovery - Trained dissipative network for silken tofu.** The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a single-term potential for silken tofu that only depends on the third deviatoric invariant J_3^Σ of the Mandel stress Σ 142

- 5.12. **Model discovery - Trained dissipative network for firm tofu.** The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a single-term potential for firm tofu that depends on the second and third deviatoric invariants J_2^Σ and J_3^Σ of the Mandel stress Σ 143
- 5.13. **Model discovery - Trained dissipative network for extrafirm tofu.** The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a four-term potential for extrafirm tofu that depends on complex combinations of the second and third deviatoric invariants J_2^Σ and J_3^Σ of the Mandel stress Σ 144
- 5.14. **Model discovery - Predicted stresses and inelastic deformation rates of the three individual tofu models.** Stress histories and evolution of the inelastic deformation rate for silken (left), firm (middle), and extrafirm (right) tofu for a strain rate of 60%/s. Applied stretches λ (top); total stress S , equilibrium stress S^{eq} , and non-equilibrium stress S^{neq} (middle), and inelastic deformation rate $\dot{C}_i$ (bottom). Distinct patterns denote contributions from the three invariants $\{\tilde{I}_1^C, \tilde{I}_2^C, I_3^C\}$ and $\{\tilde{I}_1^{C_e}, \tilde{I}_2^{C_e}, I_3^{C_e}\}$ and $\{I_1^\Sigma, J_2^\Sigma, J_3^\Sigma\}$ 145
- 5.15. **Model discovery - Water-content dependency of the single unified tofu model.** Functional dependency of the equilibrium potential ψ^{eq} (top), non-equilibrium potential ψ^{neq} (middle), and dissipative potential ϕ (bottom) on the water content α . Columns show the discovered dependencies on the individual terms. Dashed lines highlight the water contents of extrafirm, firm, and silken tofu. 146
- 5.16. **Model discovery - Experimentally recorded and computationally predicted stresses of the single unified tofu model.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests used as training data. The last column shows double-compression tests used as test data. Experimental data are plotted in gray, model predictions in green. 147
- 5.17. **Sensitivity of error measures for varying training sets.** Single stress histories $S(t)$ for silken tofu trained on the slow strain rate (top), slow and medium rates (middle), and slow, medium, and fast rates (bottom). Predictions during training shown in blue, testing in black. 156

- 5.18. **Sensitivity of error measures on varying training sets.** Ground truth vs. model predictions for silken tofu trained on the slow strain rate (top), slow and medium rates (middle), and slow, medium, and fast rates (bottom), corresponding to Fig. 5.17. Predictions during training shown in blue, testing in black. 157
- 5.19. **Preliminary sensitivity analysis for varying network architectures.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 5.10 with $\{18, 8, 8\}$ neurons per layer are plotted in blue; predictions with $\{30, 16, 12\}$ and $\{4, 4, 2\}$ neurons per layer are shown in red as set 1 and in orange as set 2. 158
- 5.20. **Preliminary sensitivity analysis for varying activation functions.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 5.10 with trainable parameter p are plotted in blue; predictions with fixed parameters $p = 1$ and $p = 2$ are shown in red as set 1 and in orange as set 2. 159
- 5.21. **Preliminary sensitivity analysis for varying regularization parameters.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 5.10 with regularization parameters $\alpha_1 = \alpha_2 = 0.0001$ are plotted in blue; predictions with $\alpha_1 = \alpha_2 = 0.001$ and $\alpha_1 = \alpha_2 = 0.00001$ are shown in red as set 1 and in orange as set 2. 160

List of Tables

2.1. Overview of equations of the classical isochoric-volumetric split to the isoplanar-area-changing split.	30
2.2. Overview of the third structural tensor, M_3 , in the different configurations. . .	35
2.3. Overview of constitutive equations of the elastic and inelastic material models.	36
4.1. Summary of all constitutive relations.	91
4.2. Artificially generated data: Learned weights of the yield function (Equation (4.24)) for the different load cases used for training (UT – Uniaxial tension, EB – Equibiaxial tension, UC – Uniaxial compression, UT-unl – Uniaxial tensions with unloading).	103
4.3. Artificially generated data: Learned weights of the Helmholtz energy accounting for linear kinematic hardening (ψ_p) and nonlinear kinematic hardening (ψ_{pe}).106	
4.4. Artificially generated data with tension-compression asymmetry: Learned weights of the yield function (Equation (4.24)) for the different load cases used for training (UT – Uniaxial tension, EB – Equibiaxial tension, UC – Uniaxial compression).	107
4.5. Comparison of the learned weights of the Helmholtz energy accounting for linear kinematic hardening (ψ_p) and nonlinear kinematic hardening (ψ_{pe}) for regularization with L_1 -regularization.	117
5.1. Nutrition facts of silken, firm, and extrafirm tofu. Serving size, calories per serving, macronutrient composition per serving, and fluid mass and volume fractions for the three different tofu types.	124
5.2. Errors of discovered models. Mean absolute error MAE, root mean squared error RMSE, normalized root mean squared error NRMSE, and relative L^2 error for the discovered individual tofu models in Figure 5.10, bottom, and for the discovered unified tofu model in Figure 5.16, top.	155

Bibliography

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M. et al. [2016], ‘Tensorflow: Large-scale machine learning on heterogeneous distributed systems’, *arXiv preprint arXiv:1603.04467* .
- Abdolazizi, K. P., Aydin, R. C., Cyron, C. J. and Linka, K. [2025], ‘Constitutive kolmogorov–arnold networks (ckans): Combining accuracy and interpretability in data-driven material modeling’, *Journal of the Mechanics and Physics of Solids* p. 106212.
- Abdolazizi, K. P., Linka, K. and Cyron, C. J. [2024], ‘Viscoelastic constitutive artificial neural networks (vcanns)—a framework for data-driven anisotropic nonlinear finite viscoelasticity’, *Journal of Computational Physics* **499**, 112704.
- Alajami, A., Li, Y. and Simon, J.-W. [2019], Continuum modelling of the anisotropic elastic-plastic in-plane behavior of paper in small and large strains regimes, *in* ‘Plasticity, Damage and Fracture in Advanced Materials’, Springer, pp. 1–18.
- Alzweighi, M., Mansour, R., Tryding, J. and Kulachenko, A. [2022], ‘Evaluation of hoffman and xia plasticity models against bi-axial tension experiments of planar fiber network materials’, *International Journal of Solids and Structures* **238**, 111358.
- Alzweighi, M., Tryding, J., Mansour, R., Borgqvist, E. and Kulachenko, A. [2023], ‘Anisotropic damage behavior in fiber-based materials: Modeling and experimental validation’, *Journal of the Mechanics and Physics of Solids* **181**, 105430.
- Amos, B., Xu, L. and Kolter, J. Z. [2017], Input convex neural networks, *in* D. Precup and Y. W. Teh, eds, ‘Proceedings of the 34th International Conference on Machine Learning’, Vol. 70 of *Proceedings of Machine Learning Research*, PMLR, pp. 146–155.
- Armstrong, P. J., Frederick, C. O. et al. [1966], *A mathematical representation of the multiaxial Bauschinger effect*, Vol. 731, Berkeley Nuclear Laboratories Berkeley, CA.

- As'ad, F. and Farhat, C. [2023], A mechanics-informed neural network framework for data-driven nonlinear viscoelasticity, in 'AIAA SCITECH 2023 Forum'.
- As'ad, F. and Farhat, C. [2026], 'A staggered training framework for mechanics-informed neural networks in tractable multiscale homogenization with application to woven fabrics', *Computer Methods in Applied Mechanics and Engineering* **452**, 118666.
- Awais, M., Sorvari, J., Tanninen, P. and Leppänen, T. [2017], 'Finite element analysis of the press forming process', *International Journal of Mechanical Sciences* **131**, 767–775.
- Ball, J. M. [1976], 'Convexity conditions and existence theorems in nonlinear elasticity', *Archive for rational mechanics and Analysis* **63**, 337–403.
- Barfusz, O., Brepols, T., van der Velden, T., Frischkorn, J. and Reese, S. [2021], 'A single gauss point continuum finite element formulation for gradient-extended damage at large deformations', *Computer Methods in Applied Mechanics and Engineering* **373**, 113440.
- Bedzra, R., Li, Y., Reese, S. and Simon, J.-W. [2019], 'A comparative study of a multi-surface and a non-quadratic plasticity model with application to the in-plane anisotropic elastoplastic modelling of paper and paperboard', *Journal of Composite Materials* **53**(6), 753–767.
- Beex, L. and Peerlings, R. [2009], 'An experimental and computational study of laminated paperboard creasing and folding', *International Journal of Solids and Structures* **46**(24), 4192–4207.
- Benady, A., Baranger, E. and Chamoïn, L. [2024], 'Unsupervised learning of history-dependent constitutive material laws with thermodynamically-consistent neural networks in the modified constitutive relation error framework', *Computer Methods in Applied Mechanics and Engineering* **425**, 116967.
- Bessa, M. A., Bostanabad, R., Liu, Z., Hu, A., Apley, D. W., Brinson, C., Chen, W. and Liu, W. K. [2017], 'A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality', *Computer Methods in Applied Mechanics and Engineering* **320**, 633–667.
- Boehler, J.-P. [1979], 'A simple derivation of representations for non-polynomial constitutive equations in some cases of anisotropy', *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* **59**(4), 157–167.

- Boes, B., Simon, J.-W. and Holthusen, H. [2025], 'Accounting for plasticity: An extension of inelastic constitutive artificial neural networks', *European Journal of Mechanics-A/Solids* p. 105998.
- Boes, B., Simon, J.-W., Reese, S. and Holthusen, H. [2023], 'A novel continuum mechanical framework for decoupled material behavior in thickness and in-plane directions', *Computer Methods in Applied Mechanics and Engineering* **415**, 116192.
- Bonatti, C. and Mohr, D. [2022], '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.
- Borgqvist, E., Lindström, T., Tryding, J., Wallin, M. and Ristinmaa, M. [2014], 'Distortional hardening plasticity model for paperboard', *International journal of solids and structures* **51**(13), 2411–2423.
- Borgqvist, E., Wallin, M., Ristinmaa, M. and Tryding, J. [2015], 'An anisotropic in-plane and out-of-plane elasto-plastic continuum model for paperboard', *Composite structures* **126**, 184–195.
- Borgqvist, E., Wallin, M., Tryding, J., Ristinmaa, M. and Tudisco, E. [2016], 'Localized deformation in compression and folding of paperboard', *Packaging Technology and Science* **29**(7), 397–414.
- Bourne, M. C. [1968], 'Texture profile of ripening pears', *Journal of Food Science* **33**(2), 223–226.
- Brepols, T., Wulfinghoff, S. and Reese, S. [2020], 'A gradient-extended two-surface damage-plasticity model for large deformations', *International Journal of Plasticity* **129**, 102635.
- Bruhns, O. T. [2024], 'Some remarks on lode angle and lode parameter', *PAMM* **24**(2), e202400004.
- Brunton, S. L., Proctor, J. L. and Kutz, J. N. [2016], 'Discovering governing equations from data by sparse identification of nonlinear dynamical systems', *Proceedings of the national academy of sciences* **113**(15), 3932–3937.
- Budday, S., Ovaert, T. C., Holzapfel, G. A., Steinmann, P. and Kuhl, E. [2020], 'Fifty shades of brain: A review on the mechanical testing and modeling of brain tissue: S. budday et al.', *Archives of Computational Methods in Engineering* **27**(4), 1187–1230.

- Budday, S., Sommer, G., Haybaeck, J., Steinmann, P., Holzapfel, G. A. and Kuhl, E. [2017a], 'Rheological characterization of human brain tissue', *Acta biomaterialia* **60**, 315–329.
- Budday, S., Sommer, G., Holzapfel, G. A., Steinmann, P. and Kuhl, E. [2017b], 'Viscoelastic parameter identification of human brain tissue', *Journal of the mechanical behavior of biomedical materials* **74**, 463–476.
- Burkov, A. [2019], *The hundred-page machine learning book*, Vol. 1, Andriy Burkov Quebec City, QC, Canada.
- Cao, J., Lee, W., Cheng, H. S., Seniw, M., Wang, H.-P. and Chung, K. [2009], 'Experimental and numerical investigation of combined isotropic-kinematic hardening behavior of sheet metals', *International Journal of Plasticity* **25**(5), 942–972.
- Casey, J. [2017], 'A convenient form of the multiplicative decomposition of the deformation gradient', *Mathematics and Mechanics of Solids* **22**(3), 528–537.
- Chatzigeorgiou, G., Meraghni, F. and Javili, A. [2017], 'Generalized interfacial energy and size effects in composites', *Journal of the Mechanics and Physics of Solids* **106**, 257–282.
- Chen, C.-C., Hsieh, J.-F. and Kuo, M.-I. [2023], 'Insight into the processing, gelation and functional components of tofu: A review', *Processes* **11**(1), 202.
- Cheng, Y., Shimizu, N. and Kimura, T. [2002], 'Texture evaluation of tofu (soybean curd) by viscoelastic tests', *Journal of the Japanese Society of Agricultural Machinery* **64**(6), 137–144.
- Cheng, Y., Shimizu, N. and Kimura, T. [2005], 'The viscoelastic properties of soybean curd (tofu) as affected by soymilk concentration and type of coagulant', *International Journal of Food Science and Technology* **40**(4), 385–390.
- Christ, D. and Reese, S. [2009], 'A finite element model for shape memory alloys considering thermomechanical couplings at large strains', *International Journal of Solids and Structures* **46**(20), 3694–3709.
- Ciarlet, P. G. [1988], *Mathematical elasticity: Three-dimensional elasticity*, Elsevier.
- Coleman, B. D. and Noll, W. [1961], 'Foundations of linear viscoelasticity', *Reviews of modern physics* **33**(2), 239.

- Comellas, E., Budday, S., Pelteret, J.-P., Holzapfel, G. A. and Steinmann, P. [2020], 'Modeling the porous and viscous responses of human brain tissue behavior', *Computer Methods in Applied Mechanics and Engineering* **369**, 113128.
- Coulomb, M. [1776], 'Essai sur une application des règles de maximis & minimis à quelques problèmes de statique, relatifs à l'architecture', *Mem. Acad. Roy. Div. Sav.* **7**, 343–387.
- Datta, B., Buehler, M. J., Chow, Y., Gligoric, K., Jurafsky, D., Kaplan, D. L., Ledesma-Amaro, R., Del Missier, G., Neidhardt, L., Pichara, K. et al. [2025], 'Ai for sustainable future foods', *arXiv preprint arXiv:2509.21556*.
- De Boer, R. [2012], *Theory of porous media: highlights in historical development and current state*, Springer Science & Business Media.
- De Boer, R. and Ehlers, W. [1990], 'The development of the concept of effective stresses', *Acta Mechanica* **83**, 77–92.
- Dettmer, W. and Reese, S. [2004], 'On the theoretical and numerical modelling of armstrong–frederick kinematic hardening in the finite strain regime', *Computer Methods in Applied Mechanics and Engineering* **193**(1-2), 87–116.
- Drucker, D. C. and Prager, W. [1952], 'Soil mechanics and plastic analysis or limit design', *Quarterly of applied mathematics* **10**(2), 157–165.
- Dunne, R. A., Darwin, E. C., Medina, V. A. P., Levenston, M. E., Pierre, S. R. S. and Kuhl, E. [2025], 'Texture profile analysis and rheology of plant-based and animal meat', *Food Research International* **205**, 115876.
- Eberlein, R. and Wriggers, P. [1999], 'Finite element concepts for finite elastoplastic strains and isotropic stress response in shells: theoretical and computational analysis', *Computer Methods in Applied Mechanics and Engineering* **171**(3-4), 243–279.
- Ebrahim, A. S., Zhang, Q. and Ha, J. [2024], 'Artificial neural network enhanced plasticity modelling and ductile fracture characterization of grade-1 commercial pure titanium', *International Journal of Plasticity* p. 104044.
- Eggersmann, R., Stainier, L., Ortiz, M. and Reese, S. [2021], 'Model-free data-driven computational mechanics enhanced by tensor voting', *Computer Methods in Applied Mechanics and Engineering* **373**, 113499.

- Ehlers, W. [1993], Constitutive equations for granular materials in geomechanical context, in ‘Continuum mechanics in environmental sciences and geophysics’, Springer, pp. 313–402.
- Esmaeili, A., Javili, A. and Steinmann, P. [2017], ‘Highly-conductive energetic coherent interfaces subject to in-plane degradation’, *Mathematics and Mechanics of Solids* **22**(8), 1696–1716.
- Fadiji, T., Coetzee, C. J., Berry, T. M., Ambaw, A. and Opara, U. L. [2018], ‘The efficacy of finite element analysis (fea) as a design tool for food packaging: A review’, *Biosystems Engineering* **174**, 20–40.
- Fadiji, T., Coetzee, C. J. and Opara, U. L. [2019], ‘Analysis of the creep behaviour of ventilated corrugated paperboard packaging for handling fresh produce—an experimental study’, *Food and Bioproducts Processing* **117**, 126–137.
- Felder, S., Holthusen, H., Hesseler, S., Pohlkemper, F., Gries, T., Simon, J.-W. and Reese, S. [2020], ‘Incorporating crystallinity distributions into a thermo-mechanically coupled constitutive model for semi-crystalline polymers’, *International Journal of Plasticity* **135**, 102751.
- Flaschel, M., Hastie, T. and Kuhl, E. [2025a], ‘Non-smooth optimization meets automated material model discovery’, *arXiv preprint arXiv:2507.10196*.
- Flaschel, M., Kumar, S. and De Lorenzis, L. [2021], ‘Unsupervised discovery of interpretable hyperelastic constitutive laws’, *Computer Methods in Applied Mechanics and Engineering* **381**, 113852.
- Flaschel, M., Kumar, S. and De Lorenzis, L. [2022], ‘Discovering plasticity models without stress data’, *npj Computational Materials* **8**(1), 91.
- Flaschel, M., Kumar, S. and De Lorenzis, L. [2023], ‘Automated discovery of generalized standard material models with euclid’, *Computer Methods in Applied Mechanics and Engineering* **405**, 115867.
- Flaschel, M., Martonová, D., Veil, C. and Kuhl, E. [2026], ‘Material fingerprinting: A shortcut to material model discovery without solving optimization problems’, *Computer Methods in Applied Mechanics and Engineering* **450**, 118573.
- Flaschel, M., Steinmann, P., De Lorenzis, L. and Kuhl, E. [2025b], ‘Convex neural networks learn generalized standard material models’, *Journal of the Mechanics and Physics of Solids* **200**, 106103.

- Flory, P. [1961], ‘Thermodynamic relations for high elastic materials’, *Transactions of the Faraday Society* **57**, 829–838.
- Franceschini, G., Bigoni, D., Regitnig, P. and Holzapfel, G. A. [2006], ‘Brain tissue deforms similarly to filled elastomers and follows consolidation theory’, *Journal of the Mechanics and Physics of Solids* **54**(12), 2592–2620.
- Friedman, H. H., Whitney, J. E. and Szczesniak, A. S. [1963], ‘The texturometer—a new instrument for objective texture measurement’, *Journal of Food Science* **28**(4), 390–396.
- Fuhg, J. N., Anantha Padmanabha, G., Bouklas, N., Bahmani, B., Sun, W., Vlassis, N. N., Flaschel, M., Carrara, P. and De Lorenzis, L. [2024a], ‘A review on data-driven constitutive laws for solids’, *Archives of Computational Methods in Engineering* pp. 1–43.
- Fuhg, J. N., Anantha Padmanabha, G., Bouklas, N., Bahmani, B., Sun, W., Vlassis, N. N., Flaschel, M., Carrara, P. and De Lorenzis, L. [2025], ‘A review on data-driven constitutive laws for solids: In fuhg et al.’, *Archives of Computational Methods in Engineering* **32**(3), 1841–1883.
- Fuhg, J. N., Fau, A., Bouklas, N. and Marino, M. [2023a], ‘Enhancing phenomenological yield functions with data: challenges and opportunities’, *European Journal of Mechanics-A/Solids* **99**, 104925.
- Fuhg, J. N., Hamel, C. M., Johnson, K., Jones, R. and Bouklas, N. [2023b], ‘Modular machine learning-based elastoplasticity: generalization in the context of limited data’, *Computer Methods in Applied Mechanics and Engineering* **407**, 115930.
- Fuhg, J. N., Jones, R. E. and Bouklas, N. [2024b], ‘Extreme sparsification of physics-augmented neural networks for interpretable model discovery in mechanics’, *Computer Methods in Applied Mechanics and Engineering* **426**, 116973.
- Ganeshan, K. [2025], Everyday choices, global impact: The power of individual sustainability actions, in ‘Proceedings of the Smart Sustainable Development Conference 2025 (SSD 2025)’, Springer Nature, p. 183.
- Gärtner, T., Fernández, M. and Weeger, O. [2021], ‘Nonlinear multiscale simulation of elastic beam lattices with anisotropic homogenized constitutive models based on artificial neural networks’, *Computational Mechanics* **68**(5), 1111–1130.

- Germain, P., Suquet, P. and Nguyen, Q. S. [1983], 'Continuum thermodynamics', *ASME Journal of Applied Mechanics* **50**, 1010–1020.
- Gesto Beiroa, J. M., Gens Solé, A. and Vaunat, J. [2011], Smoothing of yield surfaces and a reformulation of multi-surface plasticity, in 'COMPLAS XI: proceedings of the XI International Conference on Computational Plasticity: fundamentals and applications', CIMNE, pp. 921–931.
- Ghaboussi, J., Garrett, J. H. and Wu, X. [1991], 'Knowledge-based modeling of material behavior with neural networks', *Journal of Engineering Mechanics* **117**, 132–153.
- Goodfellow, I., Bengio, Y., Courville, A. and Bengio, Y. [2016], *Deep learning*, Vol. 1, MIT press Cambridge.
- Grzesik, W. [2017], 'Machinability of engineering materials', *Advanced Machining Processes of Metallic Materials; Elsevier: Amsterdam, The Netherlands* pp. 241–264.
- Hackl, K. [1997], 'Generalized standard media and variational principles in classical and finite strain elastoplasticity', *Journal of the Mechanics and Physics of Solids* **45**(5), 667–688.
- Haghighat, E., Raissi, M., Moure, A., Gomez, H. and Juanes, R. [2021], 'A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics', *Computer Methods in Applied Mechanics and Engineering* **379**, 113741.
- Hagman, A. and Nygåards, M. [2017], 'Thermographical analysis of paper during tensile testing and comparison to digital image correlation', *Experimental mechanics* **57**(2), 325–339.
- Halphen, B. and Nguyen, Q. S. [1975], 'Sur les matériaux standard généralisés', *Journal de mécanique* **14**(1), 39–63.
- Harrysson, A. and Ristinmaa, M. [2008], 'Large strain elasto-plastic model of paper and corrugated board', *International Journal of Solids and Structures* **45**(11-12), 3334–3352.
- Harrysson, M. and Ristinmaa, M. [2007], 'Description of evolving anisotropy at large strains', *Mechanics of materials* **39**(3), 267–282.
- Hartmann, S. and Neff, P. [2003], 'Polyconvexity of generalized polynomial-type hyperelastic strain energy functions for near-incompressibility', *International journal of solids and structures* **40**(11), 2767–2791.

- Haslach Jr, H. W. [2000], ‘The moisture and rate-dependent mechanical properties of paper: a review’, *Mechanics of time-dependent materials* **4**(3), 169–210.
- Heidenreich, J. N., Bonatti, C. and Mohr, D. [2024], ‘Transfer learning of recurrent neural network-based plasticity models’, *International Journal for Numerical Methods in Engineering* **125**(1), e7357.
- Heidenreich, J. N. and Mohr, D. [2024], ‘Recurrent neural network plasticity models: Unveiling their common core through multi-task learning’, *Computer Methods in Applied Mechanics and Engineering* **426**, 116991.
- Herrmann, L. and Kollmannsberger, S. [2024], ‘Deep learning in computational mechanics: a review’, *Computational Mechanics* **74**(2), 281–331.
- Hill, R. [1952], ‘The mathematical theory of plasticity’, *Bull. Am. Math. Soc* **58**, 507–512.
- Holmes, M. and Mow, V. C. [1990], ‘The nonlinear characteristics of soft gels and hydrated connective tissues in ultrafiltration’, *Journal of biomechanics* **23**(11), 1145–1156.
- Holthusen, H., Brepols, T., Linka, K. and Kuhl, E. [2025a], ‘Automated model discovery for tensional homeostasis: Constitutive machine learning in growth and remodeling’, *Computers in Biology and Medicine* **186**, 109691.
- Holthusen, H., Brepols, T., Reese, S. and Simon, J.-W. [2020], ‘An anisotropic constitutive model for fiber-reinforced materials including gradient-extended damage and plasticity at finite strains’, *Theoretical and Applied Fracture Mechanics* **108**, 102642.
- Holthusen, H., Brepols, T., Reese, S. and Simon, J.-W. [2022], ‘A two-surface gradient-extended anisotropic damage model using a second order damage tensor coupled to additive plasticity in the logarithmic strain space’, *Journal of the Mechanics and Physics of Solids* **163**, 104833.
- Holthusen, H. and Kuhl, E. [2026], ‘A complement to neural networks for anisotropic inelasticity at finite strains’, *Computer Methods in Applied Mechanics and Engineering* **450**, 118612.
- Holthusen, H., Lamm, L., Brepols, T., Reese, S. and Kuhl, E. [2024a], ‘Polyconvex inelastic constitutive artificial neural networks’, *PAMM* **24**(3), e202400032.
- Holthusen, H., Lamm, L., Brepols, T., Reese, S. and Kuhl, E. [2024b], ‘Theory and implementation of inelastic constitutive artificial neural networks’, *Computer Methods in Applied Mechanics and Engineering* **428**, 117063.

- Holthusen, H., Linka, K., Kuhl, E. and Brepols, T. [2025b], ‘A generalized dual potential for inelastic constitutive artificial neural networks: A jax implementation at finite strains’, *Journal of the Mechanics and Physics of Solids* **206**, 106337.
- Holthusen, H., Rothkranz, C., Lamm, L., Brepols, T. and Reese, S. [2023], ‘Inelastic material formulations based on a co-rotated intermediate configuration – application to bioengineered tissues’, *Journal of the Mechanics and Physics of Solids* **172**, 105174.
- Holthusen, H., Steinmann, P. and Kuhl, E. [2026], ‘A convex route to thermomechanics: Learning internal energy and dissipation’, *arXiv preprint arXiv:2603.28707*.
- Holzapfel, G. A. [2002], ‘Nonlinear solid mechanics: a continuum approach for engineering science’.
- Hornik, K., Stinchcombe, M. and White, H. [1990], ‘Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks’, *Neural networks* **3**(5), 551–560.
- Hou, J., Chen, X., Wu, T., Kuhl, E. and Wang, X. [2024], ‘Automated data-driven discovery of material models based on symbolic regression: A case study on the human brain cortex’, *Acta Biomaterialia* **188**, 276–296.
- Huang, H., Hagman, A. and Nygård, M. [2014], ‘Quasi static analysis of creasing and folding for three paperboards’, *Mechanics of Materials* **69**(1), 11–34.
- Huang, H. and Nygård, M. [2010], ‘A simplified material model for finite element analysis of paperboard creasing’, *Nordic pulp & paper research journal* **25**(4), 502–509.
- Huang, S., He, Z., Chem, B. and Reina, C. [2022], ‘Variational onsager neural networks (vonns): a thermodynamics-based variational learning strategy for non-equilibrium pdes’, *Journal of the Mechanics and Physics of Solids* **163**, 104856.
- Hussain, S., Coffin, D. W. and Todoroki, C. [2017], ‘Investigating creep in corrugated packaging’, *Packaging Technology and Science* **30**(12), 757–770.
- Isaksson, P., Gradin, P. and Kulachenko, A. [2006], ‘The onset and progression of damage in isotropic paper sheets’, *International Journal of Solids and Structures* **43**(3-4), 713–726.
- Isaksson, P. and Häggglund, R. [2005], ‘A mechanical model of damage and delamination in corrugated board during folding’, *Engineering Fracture Mechanics* **72**(15), 2299–2315.

- Isaksson, P., Hägglund, R. and Gradin, P. [2004], ‘Continuum damage mechanics applied to paper’, *International journal of solids and structures* **41**(16-17), 4731–4755.
- Jadoon, A. A., Meyer, K. A. and Fuhg, J. N. [2025], ‘Automated model discovery of finite strain elastoplasticity from uniaxial experiments’, *Computer Methods in Applied Mechanics and Engineering* **435**, 117653.
- Jadoon, A. A., Seidl, D. T., Jones, R. E. and Fuhg, J. N. [2026], ‘Input specific neural networks’, *Computational Mechanics* pp. 1–23.
- Javili, A., Kaessmair, S. and Steinmann, P. [2014], ‘General imperfect interfaces’, *Computer Methods in Applied Mechanics and Engineering* **275**, 76–97.
- Javili, A., McBride, A. and Steinmann, P. [2013], ‘Thermomechanics of solids with lower-dimensional energetics: on the importance of surface, interface, and curve structures at the nanoscale. a unifying review’, *Applied Mechanics Reviews* **65**(1).
- Ji, C., Abdolazizi, K. P., Holthusen, H., Cyron, C. J. and Linka, K. [2026], ‘Inelastic constitutive kolmogorov-arnold networks: A generalized framework for automated discovery of interpretable inelastic material models’, *arXiv preprint arXiv:2602.17750* .
- Jones, R. E. and Fuhg, J. N. [2025], ‘An attention-based neural ordinary differential equation framework for modeling inelastic processes’, *Computer Methods in Applied Mechanics and Engineering* **445**, 118186.
- Jones, R. E. and Fuhg, J. N. [2026], ‘A hierarchy of thermodynamics learning frameworks for inelastic constitutive modeling’, *arXiv preprint arXiv:2603.02645* .
- Kainz, M. P., Greiner, A., Hinrichsen, J., Kolb, D., Comellas, E., Steinmann, P., Budday, S., Terzano, M. and Holzapfel, G. A. [2023], ‘Poro-viscoelastic material parameter identification of brain tissue-mimicking hydrogels’, *Frontiers in bioengineering and biotechnology* **11**, 1143304.
- Kainz, M. P., Polz, M., Ziesel, D., Nowakowska, M., Üçal, M., Kienesberger, S., Hasibapappas, S., Winter, R., Tabrizi-Wizsy, N. G., Kager, S. et al. [2025], ‘Biointegration of soft tissue-inspired hydrogels on the chorioallantoic membrane: An experimental characterization’, *Proceedings of the 12th European Solid Mechanics Conference ESMC 2025* .

- Kalina, K. A., Brummund, J. and Kästner, M. [2026], ‘A physics-augmented neural network framework for finite strain incompressible viscoelasticity’, *Computer Methods in Applied Mechanics and Engineering* **455**, 118892.
- Kalina, K. A., Gebhart, P., Brummund, J., Linden, L., Sun, W. and Kästner, M. [2024], ‘Neural network-based multiscale modeling of finite strain magneto-elasticity with relaxed convexity criteria’, *Computer Methods in Applied Mechanics and Engineering* **421**, 116739.
- Kalina, K. A., Gebhart, P., Brummund, J., Linden, L., Sun, W. and Kästner, M. [2023a], ‘Neural network-based multiscale modeling of finite strain magneto-elasticity with relaxed convexity criteria’, *Computer Methods in Applied Mechanics and Engineering* .
- Kalina, K. A., Linden, L., Brummund, J. and Kästner, M. [2023b], ‘Fe ann: an efficient data-driven multiscale approach based on physics-constrained neural networks and automated data mining’, *Computational Mechanics* **71**(5), 827–851.
- Kingma, D. P. and Ba, J. [2014], ‘Adam: A method for stochastic optimization’, *arXiv preprint arXiv:1412.6980* .
- Kirchdoerfer, T. and Ortiz, M. [2016], ‘Data-driven computational mechanics’, *Computer Methods in Applied Mechanics and Engineering* **304**, 81–101.
- Klein, D. K., Fernández, M., Martin, R. J., Neff, P. and Weeger, O. [2022a], ‘Polyconvex anisotropic hyperelasticity with neural networks’, *Journal of the Mechanics and Physics of Solids* **159**, 104703.
- Klein, D. K., Ortigosa, R., Martínez-Frutos, J. and Weeger, O. [2022b], ‘Finite electro-elasticity with physics-augmented neural networks’, *Computer Methods in Applied Mechanics and Engineering* **400**, 115501.
- Klein, D. K., Ortigosa, R., Martínez-Frutos, J. and Weeger, O. [2024], ‘Nonlinear electro-elastic finite element analysis with neural network constitutive models’, *Computer Methods in Applied Mechanics and Engineering* **425**, 116910.
- Klein, D. K., Roth, F. J., Valizadeh, I. and Weeger, O. [2023], ‘Parametrized polyconvex hyperelasticity with physics-augmented neural networks’, *Data-Centric Engineering* **4**.
- Kohyama, K., Sano, Y. and Doi, E. [1995], ‘Rheological characteristics and gelation mechanism of tofu (soybean curd)’, *Journal of Agricultural and Food Chemistry* **43**(7), 1808–1812.

- Kopic-Osmanovic, N., Prume, E., Felder, S., Kloppenburg, G. and Simon, J.-W. [2023], ‘Modeling the anisotropic elasto-plastic material behavior of paper and paperboard at finite deformations’, *PAMM* **22**(1), e202200293.
- Korelc, J. [2002], ‘Multi-language and multi-environment generation of nonlinear finite element codes’, *Engineering with computers* **18**(4), 312–327.
- Korelc, J. and Wriggers, P. [2016], *Automation of Finite Element Methods*, Springer.
- Köstner, V., Ressel, J. B., Sadlowsky, B. and Böröcz, P. [2018], ‘Measuring the creep behaviour of corrugated board by cascade and individual test rig’, *Journal of Applied Packaging Research* **10**(1), 4.
- Köstner, V., Ressel, J., Sadlowsky, B. and Böröcz, P. [2017], ‘Individual test rig for measuring the creep behaviour of corrugated board for packaging’, *Acta Technica Jaurinensis* **10**(2), 148–156.
- Kuhl, E. [2025], ‘Ai for food: accelerating and democratizing discovery and innovation’, *npj Science of Food* **9**(1), 82.
- Kuhl, E. and Goriely, A. [2024], ‘I too i2: A new class of hyperelastic isotropic incompressible models based solely on the second invariant’, *Journal of the Mechanics and Physics of Solids* **188**, 105670.
- Lamm, L. and Holthusen, H. [2022], ‘A macroscopic approach for stress-driven anisotropic growth in bioengineered soft tissues’, *Biomechanics and Modeling in Mechanobiology* **21**(2), 627–645.
- Lamm, L., Holthusen, H., Böhm, C., Jockenhövel, S. and Reese, S. [2021], ‘Macroscopic modelling of stress driven anisotropic growth in bioengineered tissues’, *PAMM* **21**(1), e202100046.
- Le, B., Yvonnet, J. and He, Q.-C. [2015], ‘Computational homogenization of nonlinear elastic materials using neural networks’, *International Journal for Numerical Methods in Engineering* **104**(12), 1061–1084.
- Lee, E. H. [1969], ‘Elastic-plastic deformation at finite strains’, *Journal of Applied Mechanics* **36**(1), 1–6.
- Lee, E. and Liu, D. [1967], ‘Finite-strain elastic-plastic theory with application to plane-wave analysis’, *Journal of applied physics* **38**(1), 19–27.

- Lee, Y., Rosenau, J. and Peleg, M. [1983], 'Rheological characterization of tofu 2', *Journal of Texture Studies* **14**(2), 143–154.
- Lesueur, L., Thorin, A. and Weisz-Patrault, D. [2025], 'A benchmark for elasto-plasticity in finite strain'.
- Li, X., Roth, C. C., Bonatti, C. and Mohr, D. [2022], 'Counterexample-trained neural network model of rate and temperature dependent hardening with dynamic strain aging', *International Journal of Plasticity* **151**, 103218.
- Li, Y., Stapleton, S. E., Reese, S. and Simon, J.-W. [2016], 'Anisotropic elastic-plastic deformation of paper: In-plane model', *International Journal of Solids and Structures* **100**, 286–296.
- Li, Y., Stapleton, S. E., Reese, S. and Simon, J.-W. [2018], 'Anisotropic elastic-plastic deformation of paper: Out-of-plane model', *International Journal of Solids and Structures* **130**, 172–182.
- Liao, Y. and Luo, H. [2025], 'An efficient kernel learning-based constitutive model for cyclic plasticity in nonlinear finite element analysis', *Computer Methods in Applied Mechanics and Engineering* **436**, 117700.
- Lindberg, G. and Kulachenko, A. [2022], 'Tray forming operation of paperboard: A case study using implicit finite element analysis', *Packaging Technology and Science* **35**(2), 183–198.
- Linden, L., Kalina, K. A., Brummund, J., Riemer, B. and Kästner, M. [2025], 'A dual-stage constitutive modeling framework based on finite strain data-driven identification and physics-augmented neural networks', *Computer Methods in Applied Mechanics and Engineering* **447**, 118289.
- Linden, L., Klein, D. K., Kalina, K. A., Brummund, J., Weeger, O. and Kästner, M. [2023], 'Neural networks meet hyperelasticity: A guide to enforcing physics', *Journal of the Mechanics and Physics of Solids* **179**, 105363.
- Linka, K., Hillgärtner, M., Abdolazizi, K. P., Aydin, R. C., Itskov, M. and Cyron, C. J. [2021], 'Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning', *Journal of Computational Physics* **429**, 110010.
- Linka, K., Holzapfel, G. A. and Kuhl, E. [2025], 'Discovering uncertainty: Bayesian constitutive artificial neural networks', *Computer Methods in Applied Mechanics and Engineering* **433**, 117517.

- Linka, K. and Kuhl, E. [2023], ‘A new family of constitutive artificial neural networks towards automated model discovery’, *Computer Methods in Applied Mechanics and Engineering* **403**, 115731.
- Linka, K. and Kuhl, E. [2024], ‘Best-in-class modeling: A novel strategy to discover constitutive models for soft matter systems’, *Extreme Mechanics Letters* p. 102181.
- Linka, K., Pierre, S. R. S. and Kuhl, E. [2023a], ‘Automated model discovery for human brain using constitutive artificial neural networks’, *Acta Biomaterialia* **160**, 134–151.
- Linka, K., Tepole, A. B., Holzapfel, G. A. and Kuhl, E. [2023b], ‘Automated model discovery for skin: Discovering the best model, data, and experiment’, *Computer Methods in Applied Mechanics and Engineering* **410**, 116007.
- Linville, E. and Östlund, S. [2016], ‘Biaxial in-plane yield and failure of paperboard’, *Nordic Pulp & Paper Research Journal* **31**(4), 659–667.
- Lion, A. [2000], ‘Constitutive modelling in finite thermoviscoplasticity: a physical approach based on nonlinear rheological models’, *International Journal of Plasticity* **16**(5), 469–494.
- Lode, W. [1925], ‘Versuche über den einfluß der mittleren hauptspannung auf die fließgrenze’, *Zeitschrift Angewandte Mathematik und Mechanik* **5**, 142–144.
- Lubliner, J. [1985], ‘A model of rubber viscoelasticity’, *Mechanics Research Communications* **12**(2), 93–99.
- Lucey, J. [2002], ‘Formation and physical properties of milk protein gels’, *Journal of dairy science* **85**(2), 281–294.
- Mäkelä, P. and Östlund, S. [2003], ‘Orthotropic elastic–plastic material model for paper materials’, *International Journal of Solids and Structures* **40**(21), 5599–5620.
- Malik, A., Abendroth, M., Hütter, G. and Kiefer, B. [2021], ‘A hybrid approach employing neural networks to simulate the elasto-plastic deformation behavior of 3d-foam structures’, *Advanced Engineering Materials* **24**, 2100641.
- Marengo, A., Perego, U., Borgqvist, E., Tryding, J. and Ristinmaa, M. [2024], ‘An in-plane phase-field ductile fracture model for orthotropic paperboard material’, *International Journal of Solids and Structures* **294**, 112763.

- Marino, E., Flaschel, M., Kumar, S. and De Lorenzis, L. [2023], ‘Automated identification of linear viscoelastic constitutive laws with euclid’, *Mechanics of Materials* **181**, 104643.
- Martonová, D., Peirlinck, M., Linka, K., Holzapfel, G. A., Leyendecker, S. and Kuhl, E. [2024], ‘Automated model discovery for human cardiac tissue: Discovering the best model and parameters’, *Computer Methods in Applied Mechanics and Engineering* **428**, 117078.
- Masi, F. and Stefanou, I. [2022], ‘Multiscale modeling of inelastic materials with thermodynamics-based artificial neural networks (tann)’, *Computer Methods in Applied Mechanics and Engineering* **398**, 115190.
- Masi, F., Stefanou, I., Vannucci, P. and Maffi-Berthier, V. [2021], ‘Thermodynamics-based artificial neural networks for constitutive modeling’, *Journal of the Mechanics and Physics of Solids* **147**, 104277.
- McCulloch, J. A. and Kuhl, E. [2024], ‘Automated model discovery for textile structures: The unique mechanical signature of warp knitted fabrics’, *Acta Biomaterialia* **189**, 461–477.
- McCulloch, J. A., St. Pierre, S. R., Linka, K. and Kuhl, E. [2024], ‘On sparse regression, lp-regularization, and automated model discovery’, *International Journal for Numerical Methods in Engineering* p. e7481.
- Meyer, K. A. and Ekre, F. [2023], ‘Thermodynamically consistent neural network plasticity modeling and discovery of evolution laws’, *Journal of the Mechanics and Physics of Solids* **180**, 105416.
- Miehe, C. [1997], ‘A formulation of finite elastoplasticity in shells based on dual co-and contra-variant eigenvectors normalized with respect to a plastic metric’, *CIMNE, Barcelona* **1929**.
- Mises, R. v. [1913], ‘Mechanik der festen körper im plastisch-deformablen zustand’, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* pp. 582–592.
- Mohr, C. O. [1900], ‘über die darstellung des spannungszustandes und des bruchverhaltens fester körper’, *Zeitschrift des Vereines Deutscher Ingenieure* **44**, 1524–1527, 1572–1577.
- Mollica, F. and Srinivasa, A. [2002], ‘A general framework for generating convex yield surfaces for anisotropic metals’, *Acta mechanica* **154**, 61–84.

- Mow, V. C., Holmes, M. H. and Lai, W. M. [1984], 'Fluid transport and mechanical properties of articular cartilage: a review', *Journal of biomechanics* **17**(5), 377–394.
- Mozaffar, M., Bostanabad, R., Chen, W., Ehmann, K., Cao, J. and Bessa, M. [2019], 'Deep learning predicts path-dependent plasticity', *Proceedings of the National Academy of Sciences* **116**(52), 26414–26420.
- Nygårds, M. [2008], 'Experimental techniques for characterization of elasticplastic material properties in paperboard', *Nordic pulp & paper research journal* **23**(4), 432–437.
- Nygårds, M. [2009], 'Modelling the out-of-plane behaviour of paperboard', *Nordic Pulp & Paper Research Journal* **24**(1), 72–76.
- Nygårds, M., Fellers, C. and Ostlund, S. [2007], 'Measuring out-of-plane shear properties of paperboard', *Journal of pulp and paper science* **33**(2), 105–109.
- Nygårds, M., Hallbäck, N., Just, M. and Tryding, J. [2005], A finite element model for simulations of creasing and folding of paperboard, in 'Abaqus users' conference'.
- Nygårds, M., Just, M. and Tryding, J. [2009], 'Experimental and numerical studies of creasing of paperboard', *International journal of solids and structures* **46**(11–12), 2493–2505.
- Ogden, R. and Sternberg, E. [1985], 'Nonlinear elastic deformations'.
- Ogden, R. W. [1972], 'Large deformation isotropic elasticity—on the correlation of theory and experiment for incompressible rubberlike solids', *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **326**(1567), 565–584.
- Ogden, R. W. [1997], *Non-linear elastic deformations*, Courier Corporation.
- Ojha, A., Sharma, A., Sihag, M. and Ojha, S. [2015], 'Food packaging—materials and sustainability—a review', *Agricultural Reviews* **36**(3), 241–245.
- Östlund, S. [2017], 'Three-dimensional deformation and damage mechanisms in forming of advanced structures in paper', *Advances in Pulp and Paper Research, Oxford* pp. 489–594.
- Paredes, J., Cortizo-Lacalle, D., Imaz, A. M., Aldazabal, J. and Vila, M. [2022], 'Application of texture analysis methods for the characterization of cultured meat', *Scientific Reports* **12**(1).

- Pauer, E., Wohner, B., Heinrich, V. and Tacker, M. [2019], 'Assessing the environmental sustainability of food packaging: An extended life cycle assessment including packaging-related food losses and waste and circularity assessment', *Sustainability* **11**(3), 925.
- Peirlinck, M., Linka, K. and Kuhl, E. [2025], Atrial constitutive neural networks, in 'International Conference on Functional Imaging and Modeling of the Heart', Springer, pp. 249–259.
- Pfeiffer, M. and Kolling, S. [2019], 'A non-associative orthotropic plasticity model for paperboard under in-plane loading', *International Journal of Solids and Structures* **166**, 112–123.
- Phongphinitana, E. and Jearanaisilawong, P. [2012], 'A microstructurally-based orthotropic elasto-plastic model for paper and paperboard', *Applied Science and Engineering Progress* **5**(1), 19–25.
- Pierre, S. R. S., Rajasekharan, D., Darwin, E. C., Linka, K., Levenston, M. E. and Kuhl, E. [2023], 'Discovering the mechanics of artificial and real meat', *Computer Methods in Applied Mechanics and Engineering* **415**, 116236.
- Pierre, S. R. S., Sibley, L. S., Tran, S., Tran, V., Darwin, E. C. and Kuhl, E. [2025], 'Biaxial testing and sensory texture evaluation of plant-based and animal deli meat', *Current Research in Food Science* p. 101080.
- Poore, J. and Nemecek, T. [2018], 'Reducing food's environmental impacts through producers and consumers', *Science* **360**(6392), 987–992.
- Qin, P., Wang, T. and Luo, Y. [2022], 'A review on plant-based proteins from soybean: Health benefits and soy product development', *Journal of Agriculture and Food Research* **7**, 100265.
- Raissi, M., Perdikaris, P. and Karniadakis, G. E. [2019], '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.
- Rao, M. A. and Steffe, J. F. [1992], *Viscoelastic properties of foods*, Elsevier Applied Science New York.
- Reese, S. [2003], 'Meso-macro modelling of fibre-reinforced rubber-like composites exhibiting large elastoplastic deformation.', *International Journal of Solids and Structures* **980**(40(4):951).
- Reese, S. and Govindjee, S. [1998], 'A theory of finite viscoelasticity and numerical aspects', *International Journal of Solids and Structures* **35**(26–27), 3455–3482.

- Rezaei, S., Moeineddin, A. and Harandi, A. [2024], ‘Learning solutions of thermodynamics-based nonlinear constitutive material models using physics-informed neural networks’, *Computational Mechanics* **74**(2), 333–366.
- Rizzo, G. and Baroni, L. [2018], ‘Soy, soy foods and their role in vegetarian diets’, *Nutrients* **10**(1), 43.
- Robertsson, K., Borgqvist, E., Wallin, M., Ristinmaa, M., Tryding, J., Giampieri, A. and Perego, U. [2018], ‘Efficient and accurate simulation of the packaging forming process’, *Packaging Technology and Science* **31**(8), 557–566.
- Robertsson, K., Engqvist, J., Wallin, M., Ristinmaa, M., Tryding, J. and Borgqvist, E. [2023a], ‘Out-of-plane uniaxial loading of paperboard: experimental procedure and evaluation’, *Nordic Pulp & Paper Research Journal* **38**(3), 389–398.
- Robertsson, K., Jacobsson, E., Wallin, M., Borgqvist, E., Ristinmaa, M. and Tryding, J. [2023b], ‘A continuum damage model for creasing and folding of paperboard’, *Packaging Technology and Science* **36**(12), 1037–1050.
- Robertsson, K., Wallin, M., Borgqvist, E., Ristinmaa, M. and Tryding, J. [2021], ‘A rate-dependent continuum model for rapid converting of paperboard’, *Applied Mathematical Modelling* **99**, 497–513.
- Rosenkranz, M., Kalina, K. A., Brummund, J. and Kästner, M. [2023], ‘A comparative study on different neural network architectures to model inelasticity’, *International Journal for Numerical Methods in Engineering* **124**(21), 4802–4840.
- Rosenkranz, M., Kalina, K. A., Brummund, J., Sun, W. and Kästner, M. [2024], ‘Viscoelasticity with physics-augmented neural networks: model formulation and training methods without prescribed internal variables’, *Computational Mechanics* **74**(6), 1279–1301.
- Sanjon, C. W., Leng, Y., Yan, Y., Groche, P., Hauptmann, M., Ludat, N. and Majschak, J.-P. [2024], ‘Characterization and modeling of out-of-plane behavior of fiber-based materials: Numerical illustration of wrinkle in deep drawing’, *Materials* **17**(17), 4177.
- Sansour, C. and Bocko, J. [2003], ‘On the numerical implications of multiplicative inelasticity with an anisotropic elastic constitutive law’, *International journal for numerical methods in engineering* **58**(14), 2131–2160.

- Schreuders, F. K., Schlangen, M., Kyriakopoulou, K., Boom, R. M. and van der Goot, A. J. [2021], 'Texture methods for evaluating meat and meat analogue structures: A review', *Food Control* **127**, 108103.
- Schröder, J. and Neff, P. [2003], 'Invariant formulation of hyperelastic transverse isotropy based on polyconvex free energy functions', *International journal of solids and structures* **40**(2), 401–445.
- Settgast, C., Hütter, G., Kuna, M. and Abendroth, M. [2020], 'A hybrid approach to simulate the homogenized irreversible elastic-plastic deformations and damage of foams by neural networks', *International Journal of Plasticity* **126**, 102624.
- Shannon, C. E. [1948], 'A mathematical theory of communication', *The Bell system technical journal* **27**(3), 379–423.
- Shurtleff, W. and Aoyagi, A. [1975], *The Book of Tofu: Protein Source of the Future*, Soyinfo Center, Lafayette.
- Simo, J. [1987], 'On a fully three-dimensional finite-strain viscoelastic damage model: Formulation and computational aspects', *Computer Methods in Applied Mechanics and Engineering* **60**(2), 153–173.
- Simo, J. [1992], 'Associative coupled thermoplasticity at finite strains: Formulation, numerical analysis and implementation', *Comput. Meth. Appl. Mech. Engrg.* **98**, 41–194.
- Simo, J. C. and Kennedy, J. [1992], 'On a stress resultant geometrically exact shell model. part v. nonlinear plasticity: formulation and integration algorithms', *Computer Methods in Applied Mechanics and Engineering* **96**(2), 133–171.
- Simo, J. C. and Taylor, R. L. [1991], 'Quasi-incompressible finite elasticity in principal stretches. continuum basis and numerical algorithms', *Computer methods in applied mechanics and engineering* **85**(3), 273–310.
- Simon, J.-W. [2021], 'A review of recent trends and challenges in computational modeling of paper and paperboard at different scales', *Archives of Computational Methods in Engineering* **28**(4), 2409–2428.
- Sorić, J., Montag, U. and Krätzig, W. [1997], 'An efficient formulation of integration algorithms for elastoplastic shell analysis based on layered finite element approach', *Computer methods in applied mechanics and engineering* **148**(3-4), 315–328.

- Spencer, A. [2001], 'A theory of viscoplasticity for fabric-reinforced composites', *Journal of the Mechanics and Physics of Solids* **49**(11), 2667–2687.
- Spencer, A. J. [1971], 'Part iii. theory of invariants', *Continuum physics* **1**, 239–353.
- Spencer, A. J. M. [1984], Constitutive theory for strongly anisotropic solids, in 'Continuum theory of the mechanics of fibre-reinforced composites', Springer, pp. 1–32.
- St. Pierre, S. R., Darwin, E. C., Adil, D., Aviles, M. C., Date, A., Dunne, R. A., Lall, Y., Parra Vallecillo, M., Perez Medina, V. A., Linka, K. et al. [2024], 'The mechanical and sensory signature of plant-based and animal meat', *npj Science of Food* **8**(1), 94.
- Steigmann, D. J. [2009], 'A concise derivation of membrane theory from three-dimensional nonlinear elasticity', *Journal of Elasticity* **97**, 97–101.
- Steigmann, D. J. and Ogden, R. [1999], 'Elastic surface—substrate interactions', *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **455**(1982), 437–474.
- Stenberg, N. [2003], 'A model for the through-thickness elastic–plastic behaviour of paper', *International Journal of Solids and Structures* **40**(26), 7483–7498.
- Stenberg, N. and Fellers, C. [2002], 'Out-of-plane poisson's ratios of paper and paperboard', *Nordic pulp & paper research journal* **17**(4), 387–394.
- Stenberg, N., Fellers, C. and Östlund, S. [2001a], 'Plasticity in the thickness direction of paperboard under combined shear and normal loading', *J. Eng. Mater. Technol.* **123**(2), 184–190.
- Stenberg, N., Fellers, C. and Östlund, S. [2001b], 'Measuring the stress-strain properties of paperboard in the thickness direction', *Journal of Pulp and Paper Science (JPPS)* **27**(6), 213–221.
- Stockman, C. A., Goriely, A., Kuhl, E., Initiative, A. D. N. et al. [2024], 'Two for tau: Automated model discovery reveals two-stage tau aggregation dynamics in alzheimer's disease', *Brain Multiphysics* **7**, 100103.
- Svendsen, B. [1994], 'On the representation of constitutive relations using structure tensors', *International journal of engineering science* **32**(12), 1889–1892.

- Svendsen, B. [2001], ‘On the modelling of anisotropic elastic and inelastic material behaviour at large deformation’, *International Journal of Solids and Structures* **38**(52), 9579–9599.
- Taç, V., Linka, K., Sahli-Costabal, F., Kuhl, E. and Tepole, A. B. [2024], ‘Benchmarking physics-informed frameworks for data-driven hyperelasticity’, *Computational mechanics* **73**(1), 49–65.
- Taç, V., Rausch, M. K., Costabal, F. S. and Tepole, A. B. [2023], ‘Data-driven anisotropic finite viscoelasticity using neural ordinary differential equations’, *Computer Methods in Applied Mechanics and Engineering* **411**, 116046.
- Thakolkaran, P., Guo, Y., Saini, S., Peirlinck, M., Alheit, B. and Kumar, S. [2025], ‘Can kan cans? input-convex kolmogorov-arnold networks (kans) as hyperelastic constitutive artificial neural networks (cans)’, *Computer Methods in Applied Mechanics and Engineering* **443**, 118089.
- The MathWorks, Inc. [2022], ‘MATLAB version 9.13.0 (R2022b)’, <https://www.mathworks.com>. MATLAB.
- Tjahjanto, D. D., Girlanda, O. and Östlund, S. [2015], ‘Anisotropic viscoelastic–viscoplastic continuum model for high-density cellulose-based materials’, *Journal of the Mechanics and Physics of Solids* **84**, 1–20.
- Tresca, H. E. [1869], *Mémoire sur l’écoulement des corps solides*, Imprimerie impériale.
- Tschoegl, N. [1971], ‘Failure surfaces in principal stress space’, **32**(1), 239–267.
- van der Velden, T., Boes, B., Brepols, T., Kuhl, E. and Holthusen, H. [2025], ‘A note on constitutive artificial neural networks for finite strain plasticity’, *Available at SSRN 5588650* .
- Vervenne, T., Peirlinck, M., Famaey, N. and Kuhl, E. [2025a], ‘Constitutive neural networks for main pulmonary arteries: discovering the undiscovered’, *Biomechanics and Modeling in Mechanobiology* pp. 1–20.
- Vervenne, T., Pierre, S. R. S., Famaey, N. and Kuhl, E. [2025b], ‘Probing mycelium mechanics and taste: The moist and fibrous signature of fungi steak’, *Acta Biomaterialia* .
- Vishtal, A., Hauptmann, M., Zelm, R., Majschak, J.-P. and Retulainen, E. [2014], ‘3d forming of paperboard: The influence of paperboard properties on formability’, *Packaging technology and science* **27**(9), 677–691.

- Vladimirov, I. N., Pietryga, M. P. and Reese, S. [2008], ‘On the modelling of non-linear kinematic hardening at finite strains with application to springback – comparison of time integration algorithms’, *International Journal for Numerical Methods in Engineering* **75**(1), 1–28.
- Vladimirov, I. N., Pietryga, M. P. and Reese, S. [2010], ‘Anisotropic finite elastoplasticity with nonlinear kinematic and isotropic hardening and application to sheet metal forming’, *International Journal of Plasticity* **26**(5), 659–687.
- Vlassis, N. and Sun, W. [2021], ‘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.
- Wallmeier, M., Linvill, E., Hauptmann, M., Majschak, J.-P. and Östlund, S. [2015], ‘Explicit fem analysis of the deep drawing of paperboard’, *Mechanics of materials* **89**, 202–215.
- Wang, L. M., Linka, K. and Kuhl, E. [2023], ‘Automated model discovery for muscle using constitutive recurrent neural networks’, *Journal of the Mechanical Behavior of Biomedical Materials* **145**, 106021.
- Wang, X., Luo, K., Liu, S., Zeng, M., Adhikari, B., He, Z. and Chen, J. [2018], ‘Textural and rheological properties of soy protein isolate tofu-type emulsion gels: Influence of soybean variety and coagulant type’, *Food Biophysics* **13**(3), 324–332.
- Wang, Z., Estrada, J. B., Arruda, E. M. and Garikipati, K. [2021], ‘Inference of deformation mechanisms and constitutive response of soft material surrogates of biological tissue by full-field characterization and data-driven variational system identification’, *Journal of the Mechanics and Physics of Solids* **153**, 104474.
- Watanabe, T. [1997], ‘The science of tofu’, *Food Journal Co.* .
- Watson, J., Song, C., Weeger, O., Gruner, T., Le, A. T., Hansel, K., Hendawy, A., Arenz, O., Trojak, W., Cranmer, M. et al. [2024], ‘Machine learning with physics knowledge for prediction: A survey’, *arXiv preprint arXiv:2408.09840* .
- Weber, P., Geiger, J. and Wagner, W. [2021], ‘Constrained neural network training and its application to hyperelastic material modeling’, *Computational Mechanics* **68**(5), 1179–1204.
- Weber, P., Wagner, W. and Freitag, S. [2023], ‘Physically enhanced training for modeling rate-independent plasticity with feedforward neural networks’, *Computational Mechanics* .

- Xia, Q. S., Boyce, M. C. and Parks, D. M. [2002], 'A constitutive model for the anisotropic elastic–plastic deformation of paper and paperboard', *International journal of solids and structures* **39**(15), 4053–4071.
- Xu, H., Flaschel, M. and De Lorenzis, L. [2025], 'Discovering non-associated pressure-sensitive plasticity models with euclid', *Advanced Modeling and Simulation in Engineering Sciences* **12**(1), 1.
- Yang, S., Levin, M., Padmanabha, G. A., Borshevsky, M., Cohen, O., Seidl, D. T., Jones, R. E., Bouklas, N. and Cohen, N. [2025], 'Physics augmented machine learning discovery of composition-dependent constitutive laws for 3d printed digital materials', *International Journal of Engineering Science* **217**, 104381.
- Yoshida, F., Uemori, T. and Fujiwara, K. [2002], 'Elastic-plastic behavior of steel sheets under in-plane cyclic tension-compression at large strain', *International journal of plasticity* **18**(5-6), 633–659.
- Yuan, Y., Zhang, Q., Lin, S. and Li, J. [2025], 'Water: The soul of hydrogels', *Progress in Materials Science* **148**, 101378.
- Zhang, A. and Mohr, D. [2020], 'Using neural networks to represent von mises plasticity with isotropic hardening', *International Journal of Plasticity* **132**, 102732.
- Zhang, J. and Rychlewski, J. [1990], 'Structural tensors for anisotropic solids', *Archives of Mechanics* **42**(3), 267–277.
- Zhang, Z., Zou, Z., Kuhl, E. and Karniadakis, G. E. [2024], 'Discovering a reaction–diffusion model for alzheimer's disease by combining pinns with symbolic regression', *Computer Methods in Applied Mechanics and Engineering* **419**, 116647.
- Zheng, L., Regenstein, J. M., Teng, F. and Li, Y. [2020], 'Tofu products: A review of their raw materials, processing conditions, and packaging', *Comprehensive Reviews in Food Science and Food Safety* **19**(6), 3683–3714.