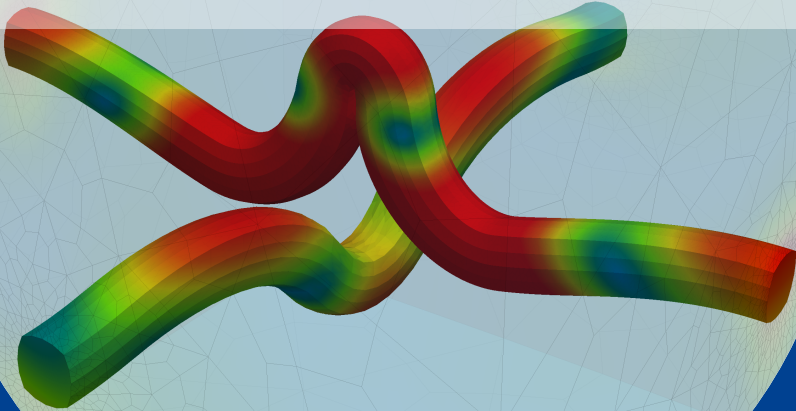




University of Stuttgart
Germany

**Oliver
Kunc**

Finite Strain
Hyperelastic Multiscale
Homogenization via Projection, Efficient
Sampling and Concentric Interpolation



9

Publication series of the
Institute of Applied Mechanics (IAM)



**Finite Strain Hyperelastic
Multiscale Homogenization
via Projection, Efficient Sampling
and Concentric Interpolation**

Vom Stuttgarter Zentrum für Simulationswissenschaften sowie
von der Fakultät Bau- und Umweltingenieurwissenschaften
der Universität Stuttgart zur Erlangung der Würde
eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Abhandlung

Vorgelegt von
Oliver Kunc, M.Sc.
aus Stuttgart

Hauptberichter: Prof. Dr.-Ing. Dipl.-Math. techn. Felix Fritzen
Mitberichter: Prof. Dr. rer. nat. Oliver Weeger

Tag der mündlichen Prüfung: 01. September 2021

Institut für Mechanik (Bauwesen) der Universität Stuttgart
Data Analytics in Engineering, Prof. Dr.-Ing. Felix Fritzen

2021

Publication series of the Institute of Applied Mechanics (IAM), Volume 9
Institute of Applied Mechanics
University of Stuttgart, Germany, 2021

Editors:

Prof. Dr.-Ing. Dipl.-Math. techn. F. Fritzen
Prof. Dr.-Ing. H. Steeb
Prof. Dr.-Ing. M.-A. Keip

© 2021 Oliver Kunc
Institut für Mechanik (Bauwesen)
Heisenberg Group for Data Analytics in Engineering (DAE)
Universität Stuttgart
Pfaffenwaldring 7
70569 Stuttgart, Deutschland

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, without the permission in writing of the author.

ISBN 978–3–937399–57–7
(D 93 – Dissertation, Universität Stuttgart)

Danksagungen

Die vorliegende Arbeit entstand während meiner Zeit in der Emmy-Noether-Gruppe EMMA (Effiziente Methoden zur Mechanischen Analyse). Ich möchte Prof. Fritzen meinen größten Dank für die konstruktive Zusammenarbeit und für die inspirierende und geduldige Betreuung aussprechen.

Die Nachwuchsgruppe war Teil des Instituts für Mechanik im Bauwesen an der Universität Stuttgart, wo ich die stets positive Atmosphäre und die angenehme Gesellschaft meines Kollegiums genoss. Der Institutsleitung durch Prof. Ehlers und Prof. Steeb sei an dieser Stelle ausdrücklich gedankt. Des Weiteren schätze ich die einzigartigen Möglichkeiten, die die Graduiertenschule des SC SimTech bot und die ich gerne nutzte. Ich danke Prof. Göttsche für die Co-Betreuung und die konstruktive Begleitung.

Zahlreiche Institutsmitarbeiter und Kontakte trugen durch erkenntnisreiche Diskussionen, als inspirierende Vorbilder oder durch Unterstützung sonstiger Art zum Ergebnis dieser Arbeit bei. Unter ihnen waren Mohammadreza Hassani, Dr. Matthias Leuschner, Dr. Mauricio Fernández, Julian Lißner, Dr. Shadi Alameddin, Barbara Teutsch und Dr. Arndt Wagner.

Größter Dank in fachfremder Hinsicht geht an die Tageseinrichtung für Kinder, die meine Kinder immer so gerne besuchten. Ohne diese seit meinem Mathematikstudium tragende Stütze für meine Familie wären weder meine berufliche noch meine private Entwicklung auf dieser Zeitskala möglich gewesen. Das Kitapersonal sowie das Konzept der Stadt Stuttgart als Trägerin dieser Einrichtung verdient allerhöchste Anerkennung.

An meine Lebensgefährtin, Kinder, Eltern sowie an meine ganze Familie richte ich einen besonders herzlichen Dank. Der Rückhalt, auf den ich mich immer verlassen konnte, und das Verständnis, das mir entgegengebracht wurde, war nicht selbstverständlich.

Stuttgart, im Oktober 2021

Oliver Kunc

Mechanics is the paradise of the mathematical sciences because by means of it one comes to the fruits of mathematics.

ursprünglich vermutlich von Leonardo da Vinci

Zusammenfassung

Realistische Simulationen mechanischer Bauteile erfordern Kompromisse zwischen dem für die Ergebnisse aufbrachten Aufwand und der Genauigkeit dieser Ergebnisse. Insbesondere wenn die Materialien mikroskopische Heterogenitäten aufweisen, erfordern klassische Lösungsansätze entweder exzessive Rechenkapazitäten oder scheitern aufgrund Übereinfachung an der Berechnung vertrauenswürdiger Ergebnisse. Dies gilt selbst im vergleichsweise einfachen Fall von Hyperelastizität, da enormer rechnerischer Aufwand allein schon durch feinkörnige Geometrien bedingt wird.

In der Literatur finden sich vergleichsweise effiziente Rechenschemata für Hyperelastizität, die auf dem Konzept der Skalenseparation beruhen. Diese beinhalten Yvonnet & He [2007], welches die Proper Orthogonal Decomposition (POD) auf mikroskopischen Verschiebungsvariablen anwendet, oder Temizer & Wriggers [2007], Yvonnet et al. [2009], wo homogenisierte Materialeigenschaften auf makroskopischen Dehnungs- oder Verformungsmaßen interpoliert werden. Die vorliegende Arbeit ist von diesen und anderen etablierten Verfahren inspiriert und schlägt eine neue, effizientere Methode vor. Dabei werden grundlegende Konzepte neu definiert, wie zum Beispiel welche Variablen durch POD-Projektion approximiert werden. Auch werden die völlig verschiedenen Methoden der Projektion und der Interpolation zu einem synergetischen Ansatz kombiniert, welcher die Vorteile beider Methoden vereint. Das Produkt ist ein Verfahren, welches auch für mehr als zwei Skalen hinreichend performant, robust und allgemein ist.

Nach bestem Wissen des Autors ist die vorliegende Arbeit die erste, die vierskalige, rechnergestützte hyperelastische Homogenisierung durchführt. Dabei werden keine Einschränkungen an die geometrische Nichtlinearität, die Anzahl der materiellen Komponenten, deren Form, Verteilung und hyperelastischen Eigenschaften sowie an das Vorhandensein von Poren getroffen. Die Tatsache, dass keine Höchstleistungs-Rechenkapazitäten benötigt werden und dass die Anzahl der handhabbaren Skalen beliebig ist, verdeutlicht wie in vielen anderen Fällen, dass Effizienz nicht nur der Leistungssteigerung dient, sondern auch neue Möglichkeiten eröffnet.

Ähnlichkeit der Skalen, wie zum Beispiel im Fall von fraktalen Mikrostrukturen, wird nicht vorausgesetzt, kann jedoch ohne Weiteres mit dieser Methode behandelt werden. Hochgradig anisotropes Effektivverhalten, geometrisch bedingte Versteifungen sowie skaleninduzierte Erweichungen sind darstellbar. Phasenkontraste von bis zu 1000 werden ohne größere Schwierigkeiten behandelt.

Der Vorschlag besteht auf jeder Skala aus zwei Stufen. Zuerst wird ein projektionsbasiertes, ordnungsreduziertes Modell konstruiert, welches kinematische Größen – jedoch nicht die Verschiebungen – durch eine Reduzierte Basis (RB) approximiert. Dadurch ist die Berechnung der effektiven Materialantwort deutlich effizienter im Vergleich zum unreduzierten Ausgangsmodell. In der zweiten Stufe werden sorgfältig ausgewählte Stichproben der homogenisierten Materialantwort mittels “Concentric Interpolation” (CI) konzentrisch interpoliert. Die Stützstellen hierfür werden in der ersten Stufe erstellt, wohingegen der Aufbau des RB-Modells auf Lösungsdaten beliebiger

hochgenauer Methoden basiert, wie zum Beispiel auf Finiten Elementen. Eine eigens entwickelte Art der Stichprobennahme nach konzentrischem Muster, "Concentric Sampling" (CS), mildert den Fluch der Dimensionalität und wird in jeder Stufe und auf jeder Skala auf den kinematischen Zustandsraum angewendet.

Die Leistungsfähigkeit des Gesamtschemas geht mit einem erhöhten Aufwand für den Modellaufbau einher. Dennoch sind Standardcomputer ausreichend. Der Aufbau hängt von je einem Satz Parameter für die RB- und die CI-Methode ab. Diese Parameter werden in dieser Fallstudie empirisch bestimmt.

Obgleich der fundamentalen Verschiedenheit der RB- und der CI-Methode wird CS in beiden Fällen für den Modellaufbau empfohlen. Dabei wird der Raum der (infinitesimalen oder Hencky-) Verzerrungstensenoren auf isotrope, konzentrische Weise abgetastet: Die Probestellen werden entlang homogen verteilter Richtungen an gewissen Radien platziert, wobei Letztere für alle Richtungen identisch sind. Hierdurch liegen die Probestellen auf konzentrischen Sphären um den Ursprung. Die Auswertung der homogenisierten Materialantwort zu Randbedingungen, die zu diesen kinematischen Zuständen der nächsthöheren Skala gehören, ermöglicht das Erfassen sowohl der effektiven Anisotropie als auch der Amplitudenabhängigkeit. Die Auflösung dieser Effekte ist mittels der Anzahl an Proberichtungen bzw. -radien kontrollierbar.

Tatsächlich empfiehlt sich CS für jegliches materielle Ersatzmodell, welches auf einer Diskretisierung des Raums der Hencky- oder der infinitesimalen Verzerrungstensenoren basiert. Dies gilt solange keine besonderen Anforderungen an die Probestellenverteilung gestellt werden, die über die allgemein wünschenswerten Eigenschaften der Grobheit (zwecks Leistungsfähigkeit) und der Dichte (zwecks vollständiger Erfassung) hinausgehen. Diese machen das klassische Dilemma der datenbasierten Mechanik aus.

Eine tiefgehende Analyse vorheriger Versuche, den Verzerrungsraum zu diskretisieren, wie zum Beispiel reguläre Gitter in Yvonnet & He [2007], führt zur Entwicklung von CS. Die Schlüsselidee für die konzentrische Anordnung der Probestellen ist, dass der Winkelabstand zwischen Richtungen im Raum der infinitesimalen oder Hencky-Verzerrungen ein Maß für die Verschiedenartigkeit der Belastungen ist. Deshalb trägt der Winkelabstand einen höheren Informationsgehalt als der Abstand in regulären Gittern. CS ist eine natürliche Folge der isotropen Erhaltung der Richtungsichte.

Die Verknüpfung von CS und CI sollte als Machbarkeitsstudie und ohne Anspruch auf Optimalität aufgefasst werden. Es wird vorgeführt, dass diese Kombination auch außerhalb der Mechanik anwendbar ist: Allgemeine Funktionen mit noch nicht näher bestimmten Glattheitseigenschaften können durch CI wie mit einer Blackbox interpoliert werden, selbst wenn die Bedeutung des Winkelabstands im Definitionsbereich unklar ist. CS ist auch in solchen Fällen vorteilhaft, da es dem Fluch der Dimensionalität nicht so stark unterliegt wie die Abtastung mittels regulärer Gitter.

Um sowohl die Zugänglichkeit zu den Methoden als auch die Reproduzierbarkeit der Ergebnisse zu erhöhen, ist der Quellcode der RB-Methode Kunc [2019], des allgemeinen CI-Schemas Fritzen & Kunc [2018a] und der Generierung homogen verteilter Richtungen in beliebigen Dimensionen Fritzen & Kunc [2018b] veröffentlicht.

Abstract

Realistic simulations of mechanical parts require compromises between resources spent on achieving the results and the accuracy of these results. Especially when significant microscopic material heterogeneities are present, classical solution methods either require excessive computational resources or fail to yield trustworthy results due to oversimplification. The comparatively simple case of hyperelasticity is no exception to this, as enormous computational complexity emerges from fine-grained geometries even in the absence of more sophisticated material models.

Building on the concept of separation of scales, rather efficient computational two-scale methods for hyperelasticity have been proposed in the literature. These include Yvonnet & He [2007] applying Proper Orthogonal Decomposition (POD) on the microscopic displacement variables, or Temizer & Wriggers [2007], Yvonnet et al. [2009] interpolating homogenized material properties on macroscopic measures of strain or deformation. The present work is inspired by these and other established techniques and proposes a novel method of significantly increased efficiency. Basic concepts of existing methods are redefined, e.g. which variables to approximate via POD-based projection. The unrelated methods of projection and of interpolation are combined into a synergetic approach inheriting the advantages of both paradigms. The resulting method is sufficiently performant, robust, and general in order to treat more than two spatial scales.

To the best of the author's knowledge, the present work is the first to conduct computational four-scale hyperelastic homogenization. The established method is general with respect to geometric nonlinearity, the number of material constituents, their shapes and distribution, their hyperelastic properties and the presence of pores. The facts that no high-performance computing environment is necessary and that the number of treatable scales is arbitrary highlight that – as in many cases – efficiency is not just an improvement but even an enablement.

Similarity among the scales, in which case one speaks of fractal micro-structures, is not required but can be treated without any modifications of the method. Highly anisotropic effective behaviour, geometric stiffening and scale softening is well reproducible. Solid phase contrasts of up to 1000 are applied without major hurdles.

The proposed method consists of two stages on each scale. First, a projection-based reduced order model is constructed via a Reduced Basis (RB) approximation of kinematic non-displacement quantities. The computation of the effective material behavior is conducted significantly more efficiently than with the unreduced, original model. Second, carefully selected samples of the homogenized material response are interpolated by means of Concentric Interpolation (CI). The support data for the CI is created at the first stage whereas the RB setup is based on solution data of any high-fidelity method, such as Finite Elements. An original Concentric Sampling strategy (CS) alleviates the curse of dimensionality and is applied to the kinematic state space at both stages and on every scale.

The performance comes at the cost of a notable setup time before the largest scale can be treated. Still, standard workstation computers suffice. The setup is dependent on a distinct set of parameters for each of the RB and the CI method. These parameters are found empirically in this case study.

Despite the fundamental dissimilarities between the RB and the CI method, CS is recommended for the setup of both. In this, the space of (infinitesimal or Hencky) strains is sampled in an isotropic and concentric manner: the samples are placed along homogeneously distributed directions at certain radii that are the same for all directions. Therefore, the samples lie on concentric spheres centered at the origin. When the homogenized material response is evaluated with respect to boundary conditions corresponding to these next-scale kinematic states, an efficient capture of both effective anisotropy and magnitude dependence is possible. The resolution of these effects is controllable via the number of sampling directions and radii, respectively.

In fact, CS is suggestive for use with any method for material model surrogation relying on discretizations of the space of Hencky strains or of infinitesimal strains. This holds at least as long as no special requirements are posed on the sample placement besides the generally desirable features of sparsity (for the purpose of performance) and density (for the purpose of representativeness), constituting the classical dilemma of data-based computational mechanics.

A thorough analysis of previous strain space discretization attempts, such as regular grids as employed in, e.g., Yvonnet et al. [2009], leads to the development of CS. The key idea for the concentric layout of the samples is that the angular distance between directions in the space of infinitesimal or Hencky strains is a measure of how much these kinematic loads differ in kind. Therefore, this angular distance is more meaningful than the spacing of regular grid discretizations. Conserving the directional sampling density in an isotropic way naturally leads to Concentric Sampling.

The combination of CS and CI should be seen as a proof of concept, optimality is not claimed. It is demonstrated that this combination is applicable beyond computational mechanics. General functions, with certain smoothness characteristics that are yet to be determined rigorously, may be interpolated in a black-box manner via CI even when the meaning of directional distance in the input space is unclear. CS is still beneficial in such cases as it significantly alleviates the curse of dimensionality when compared to regular grid discretization.

In order to increase the accessibility of the methods and of the reproducibility of the claims, source code is published for the RB method [Kunc, 2019], the general CI scheme [Fritzen & Kunc, 2018a] and for the generation of homogeneously distributed directions in arbitrary dimensions [Fritzen & Kunc, 2018b].

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Goal	3
1.3	Structure of this document	4
2	Preliminaries	5
2.1	Basic kinematics	5
2.2	Measures of strain	6
2.2.1	Geometric nonlinearity	6
2.2.2	Geometric linearity	8
2.2.3	The special role of the Hencky strain	8
2.3	Compatibility conditions	9
2.4	Balance equations	10
2.4.1	Stresses	10
2.4.2	Balance of momentum and conservation of mass	10
2.5	Hyperelasticity	11
3	Introduction to homogenization	15
3.1	Many-scale modeling	15
3.2	Representative Volume Elements and pores	18
3.3	The effective tangent modulus	19
3.4	Short summary of homogenization methods	20
3.4.1	Classical analytical methods	20
3.4.2	Modern computational and hybrid methods	21
3.4.3	Data-based methods	23
4	Outline	27
4.1	First publication – Two-stage data-driven homogenization of nonlinear solids using a Reduced Basis model	27
4.1.1	Projection based Reduced Order model	28
4.1.2	Sampling scheme	29
4.1.3	Interpolation scheme	31
4.2	Second publication – Finite Strain Homogenization Using a Reduced Basis and Efficient Sampling	32
4.2.1	Suitability of nonlinear kinematic measures	32

4.2.2	Projection based Reduced Basis model	32
4.2.3	Sampling scheme	33
4.3	Third publication – Generation of energy-minimizing point sets on spheres and their application in mesh-free interpolation and differentiation	36
4.3.1	General interpolation scheme	36
4.3.2	Point distributions and asymptotic uniformity	37
4.3.3	Kernel functions and energies of point sets	37
4.3.4	Energy-minimizing point sets as interpolation support	38
4.4	Fourth publication – Many-scale finite strain computational homogenization via Concentric Interpolation	41
4.4.1	Concentric Interpolation of the material response	41
4.4.2	Many-scale homogenization	42
4.5	List of publications	43
5	First publication – Two-stage data-driven homogenization of nonlinear solids using a Reduced Basis model	45
5.1	Introduction	46
5.1.1	Notation	48
5.2	Framework	49
5.2.1	Twoscale formulation	49
5.2.2	Hyperelastic material models	50
5.2.3	Effective hyperelastic material model	53
5.3	Hyperelastic Reduced Order Model	54
5.3.1	Galerkin Reduced Order Model	54
5.3.2	Generation of the reduced basis via snapshot POD	55
5.4	Data-assisted homogenization using reduced order models	56
5.4.1	Layout	56
5.4.2	Analysis of the NEXP of Yvonnet et al.	57
5.4.3	Sampling strategies: preliminary considerations	60
5.4.4	Discretization of load amplitudes and directions	61
5.5	Radial numerically explicit potentials (RNEXP)	62
5.5.1	Data sampling strategy	62
5.5.2	Interpolation scheme	63
5.5.3	Computation of stress and stiffness	65
5.5.4	Symmetry considerations	69
5.5.5	Choice of sampling points and adaption	69
5.6	Numerical examples	70
5.6.1	Layout	70
5.6.2	Calibration of the kernel width parameter	71

5.6.3	Validation for strain-driven RVE computations	73
5.6.4	Nonlinear hardening materials with anisotropic microstructure	74
5.6.5	RNEXP twoscale simulation with data-acceleration	75
5.7	Résumé	79
5.7.1	Summary	79
5.7.2	Discussion	80
5.7.3	Perspective	81
5.8	Appendix	83
5.8.1	A – Hyperelastic potentials and their gradients	83
5.8.2	B – Algorithm: Galerkin ROM for hyperelasticity	86
5.8.3	C – Algorithm: snapshot POD for hyperelasticity	87
5.8.4	D – Algorithm: RNEXP	87

6 Second publication – Finite Strain Homogenization Using a Reduced Basis and Efficient Sampling 89

6.1	Introduction	89
6.1.1	Purpose	89
6.1.2	State of the Art	90
6.1.3	Main Contributions and Outline	91
6.1.4	Notation	92
6.1.5	Material Models	94
6.1.6	Problem Setting of First Order Homogenization	95
6.2	Reduced Basis Homogenization for Hyperelasticity	97
6.2.1	Formulation	97
6.2.2	Identification of the Reduced Basis	99
6.2.3	Mathematical Motivation of the Reduced Basis model	100
6.2.4	Details on the Coefficient Optimization	101
6.3	Sampling	103
6.3.1	General Considerations	103
6.3.2	Large Strain Sampling Strategy	103
6.3.3	Application of the Stretch Tensor Trained Reduced Basis Model	108
6.4	Numerical Examples	109
6.4.1	Reduced Basis for a Fibrous Microstructure	109
6.4.2	Reduced Basis for a Stiffening Microstructure	114
6.5	Discussion	117
6.5.1	Discussion of the Reduced Basis Method	117
6.5.2	Discussion of the Sampling Strategy	120
6.6	Appendix	121
6.6.1	Material Objectivity	121

6.6.2	Effective Material Responses of the RB	122
6.6.3	Effective Stress	122
6.6.4	Effective Stiffness	122
6.6.5	Basis for Symmetric Traceless Second Order Tensors	123
7	Third publication – Generation of energy-minimizing point sets on spheres and their application in mesh-free interpolation and differentiation	125
7.1	Introduction	126
7.1.1	Part 1: Generation of LOG-energy minimizing point sets	126
7.1.2	Part 2: Mesh-free interpolation and differentiation	127
7.1.3	Outline	128
7.2	Minimum energy point sets	128
7.2.1	Kernels	128
7.2.2	Measures, potentials, energy, and asymptotic uniformity	129
7.2.3	Evaluation of the asymptotic LOG-energy	133
7.3	Proof of the asymptotic uniformity of LOG-optimal point sets (Theorem 1)	134
7.4	Generation of energy-minimizing point sets	138
7.4.1	Algorithmic implementation	138
7.4.2	Initial positions	139
7.5	Numerical examples I: point distributions	140
7.5.1	Mesh ratio	140
7.5.2	Empirical Distribution Functions (EDF)	140
7.5.3	Asymptotic behavior of the minimum energy	141
7.5.4	Universally optimal point sets	144
7.6	Concentric Interpolation	145
7.6.1	Introduction	145
7.6.2	Setup and Concentric Sampling	145
7.6.3	Interpolation	146
7.6.4	Differentiation	149
7.6.5	Symmetric Gaussian kernel function	150
7.7	Numerical examples II: Concentric Interpolation	150
7.7.1	Kernel parameter γ	150
7.7.2	Application to an eight-dimensional engineering model	152
7.8	Résumé	154
7.8.1	Summary	154
7.8.2	Discussion	156
7.9	Appendix	158
7.9.1	Generation of random points on \mathbb{S}^d	158
7.9.2	Algorithms	159

8	Fourth publication – Many-scale finite strain computational homogenization via Concentric Interpolation	161
8.1	Introduction	161
8.2	Notation, problem formulation and reduced order model	163
8.2.1	Notation	163
8.2.2	Multi-scale setting	166
8.2.3	Reduced Basis homogenization	169
8.3	Computational homogenization via Concentric Interpolation	171
8.3.1	Basic scheme	172
8.3.2	Efficient implementation	173
8.3.3	Concentric Interpolation of material laws	174
8.3.4	Concentric Sampling of material laws	175
8.3.5	Many-scale homogenization: overall algorithm	176
8.4	Numerical examples	182
8.4.1	Three-scale homogenization of fractal structures on standard workstations	182
8.4.2	Homogenized stress response of stiffening microstructure via CI	191
8.5	Summary and discussion	196
8.5.1	Summary	196
8.5.2	Interpolation given different strain measures	197
8.5.3	Number of sampling directions	198
8.5.4	Number and position of sampling magnitudes	198
8.5.5	Kernel parameter	198
8.5.6	Accuracy	199
8.5.7	Generalization to other material models	199
9	Outlook	201
9.1	Possible advancements of the current method	201
9.2	Extension to other non-dissipative material classes	202
9.3	Extension to dissipative material classes	202
A	Appendix	205
	Bibliography	206

Chapter 1:

Introduction

1.1 Motivation

The domain of simulation is indispensable to modern engineering. Simulations of mechanical behavior are fundamental to the design process of parts, components, and systems throughout many industries. Such computations can be regarded as virtual experiments with distinct advantages and disadvantages compared to the real-world equivalents.

A subset of the broad field of mechanical simulation is concerned with multiple spatial scales. Flow of information, both from smaller scales to larger scales and vice versa are possibly of interest, depending on the application. Examples include the design of composite materials for a specific purpose and the prediction of microscopic failure initialization, respectively. The former case is of primary interest in the context of the present work.

More specifically, it is assumed that all information about the complex microscopic structure of a mechanical part is given. In non-trivial applications however, this leads to an extremely fine-grained classical discretization as geometric features have to be resolved which are orders of magnitudes smaller than the macroscopic engineering part. In addition to the overly large size of the resulting model, material and geometric nonlinearities on all scales hinder a solution of the governing equations within a reasonable run time.

By consideration of Representative Volume Elements (RVE's), a concept that is known at least since Hill [1963], it is possible to treat different scales individually and thus more efficiently. The RVE contains a representative part of the microscopic structure and allows for homogenization of the mechanical properties. On the macroscopic scale, the micro-structure is then not resolved but incorporated only by the pre-computed effective material law. The complexity of the overall problem is greatly reduced as repetitive occurrences of sub-scale structures are avoided. In this way, flow of information from the smaller to the larger scale is made possible.

Iterative procedures take this a step further and switch back and forth among discretizations of the RVE and of the macroscopic problem. For instance, in Feyel [1999], Miehe et al. [1999b] this concept of separation and coupling of spatial scales is rigorously applied to complex problems. Millions of solutions of high-fidelity microscopic problems are necessary for just the closure of the equations on the macroscopic scale. Even with such methods that are fairly advanced in comparison to the naive approach (without sepa-

ration of scales) and that can be easily accelerated by means of parallel computing, the compute times for real-world problems are still unacceptable.

Certain order reduction methods provide a means to accelerate the solution of mathematical models of mechanics. Among these, projection via Proper Orthogonal Decomposition (POD), cf. e.g. Sirovich [1987], and Hyperreduction, cf. Ryckelynck [2005], are notable examples. The application of such methods to homogenization problems has proven especially fruitful, cf. Saeb et al. [2016], Geers & Yvonnet [2016]. It is this path of order reduction techniques that the present work intends to take, and to extend to even greater efficiency.

Such classical order reduction schemes have the feature of carrying micro-structural information with them, i.e. it is possible to recover the detailed microscopic state associated with a macroscopic state. This property may be important, especially if one is interested in the flow of information from the larger to the smaller scale. If one is interested in opposite direction only, as is the case in the current work, the burden of detailed information about micro-structural complexity is essentially unnecessary as only the effective material response is required. In this case, certain data-based methods promise to be game-changing in the way that seek to dispose of this ballast, taking computational economy to a different level while still providing satisfactory accuracy: if sufficient suitable data is available, this data may possibly replace the equations on which the homogenization is based. The quest for suitable data-based schemes and compatible data is highly non-trivial and depends on the particular context. Examples include interpolation methods, cf. e.g. Temizer & Wriggers [2007], and the very general set of Artificial Neural Networks (ANN's), cf. e.g. Le et al. [2015].

Synergies of model order reduction techniques and data-based methods do not seem to have been extensively investigated in the literature of homogenization prior to the beginning of this work. More specifically, it appears worthwhile to investigate whether advanced order reduction schemes may assist the setup of specialized data-based methods, combining the advantages of both.

Applications involving certain materials, such as polymers, foams, or soft biological tissues require a geometric description without the assumption of small strains or small rotations. In the context of efficient homogenization, this causes both modeling and numerical difficulties, ranging from a multitude of possible formulations to limited numerical stability. Even when assuming purely hyperelastic material models, the aforementioned difficulties remain obstructive for classical order reduction and data-based methods.

Although hyperelasticity is generally considered a relatively simple class of material models, sufficiently efficient homogenization methods allowing for finite strains on two or more spatial scales in three dimensions have not yet been established. A solution to this arguably elementary problem seems overdue.

1.2 Goal

The goal of the present work is to enable efficient realistic structural solid mechanical simulations under the following conditions.

- *Arbitrary number of spatial scales*: if desired, more than two scales may be considered in a simulation.
- *Finite strains*: a linearized geometrical description is optional but not necessary.
- *Hyperelasticity*: all constituents on the microscopic scale are assumed to be known and of general hyperelastic kind.
- *Limited computational resources*: all computations should be able to be carried out on current standard workstations. For instance, a typical hardware setup could consist of 12 modern CPU cores and 256 GB of RAM. Strongly depending on the actual complexity of the setting, rough estimations of the anticipated runtimes for realistic three-dimensional scenarios on two scales are
 - less than 24 hours for the *setup* of the macroscopic simulation.
 - few hours for the *execution* of the macroscopic simulation for a couple of interesting loading scenarios.

Applications which are covered by the just mentioned modeling conditions include, for instance, rubber foams with pores of different orders of magnitude. Furthermore, composites with possibly extreme phase contrast are within the scope of this work, such as metal fiber inclusions within a polymer matrix. Even some dissipative materials may possibly be covered by these conditions for proportional loadings, e.g. when speaking of pseudo-plasticity.

It is important to highlight the lack of certain additional conditions that are commonly imposed on micro-structured solids. Most notably, the effect of macroscopic anisotropy induced by microscopic (mis-)alignments is *not* assumed to be absent or restricted to a certain degree. This effect is especially significant when long fibers are involved.

Non-convex effective behavior is not specifically addressed by this work. A standard example of such a phenomenon is buckling which may occur within fibrous structures under loads that are compressive with respect to the fiber direction. Here, buckling scenarios are not excluded but uniqueness is assumed and stability is not investigated.

1.3 Structure of this document

The present dissertation is of cumulative nature. The essence of the work consists of four scientific publications that are wholly contained in Chapters 5–8. This document serves the purpose of providing additional context to the four articles, highlighting relationships among them, and summarizing the overall achievements and the resulting perspective.

In Chapter 2, the main notation and general modeling assumptions are stated. A broader view on the matter is established while select details are only covered in the respective publication chapters.

Chapter 3 introduces to the concept of many-scale homogenization. Theoretical foundations are addressed and the resulting models are stated. Classical and contemporary homogenization methods are summarized and discussed.

Chapter 4 aligns the four core works and states the overall line of thought. Conceptual connections among the individual parts and overarching decisions are explained.

Chapters 5–8 contain the published articles in their entirety and in chronological order. Owing to the format of a cumulative work, each of these chapters is self-contained with individual motivation, notation, summary etc.

In the closing Chapter 9, prospects on possible future developments are provided.

Readers of the digital version of this work are advised of the fact that both internal and external *references are hyperlinked and should be clickable* from standard document viewers, although no visual indication of this feature is given due to formal reasons.

Chapter 2:

Preliminaries

The foundation of solid mechanics, to which elasticity theory is a sub-domain, is the assumption of a *material continuum*: it is assumed that any physical body is composed of a continuum of material points, cf. [Wriggers, 2008, Chapter 3]. The continuity of the distribution of these points stands in contrast to what is assumed in particle mechanics and also to what is observed on sufficiently small scales of reality. All scales considered in this work are assumed to be large enough such that the continuum hypothesis holds and, therefore, the particulate or quantum nature of physical bodies does not need to be addressed.

Based on this, a *physical body* is mathematically modeled as a bounded, simply connected subset of the three-dimensional Euclidean space \mathbb{R}^3 . This subset describes the spatial extension of the body, by which it is said to be occupied. As the body is movable and deformable, i.e. not rigid, different sets may be occupied depending on the physical circumstances, i.e. for instance acting forces and mechanical properties of the material. This chapter is concerned with a brief introduction of a standard continuum mechanical formalism by means of which these physical circumstances and their influence on the state of the body may be described. For a more detailed view, the reader is referred to standard text books, e.g. Ogden [1984].

2.1 Basic kinematics

The *basic kinematic* quantities are sketched in Figure 2.1. Initially, the body occupies the set $\Omega_0 \subset \mathbb{R}^3$. In this state, each material point is described by its coordinate vector $\mathbf{X} \in \mathbb{R}^3$. The current configuration of the body, e.g. in a possibly deformed state, is denoted by $\Omega \subset \mathbb{R}^3$. Coordinate vectors of points in this set are denoted by $\mathbf{x} \in \mathbb{R}^3$ which relate to the initial coordinates by means of the displacement $\mathbf{u} \in \mathbb{R}^3$: $\mathbf{x} = \mathbf{X} + \mathbf{u}$. It is important to note that both the displacement and the current position are actually fields, i.e. $\mathbf{u} : \Omega_0 \rightarrow \mathbb{R}^3 : \mathbf{X} \mapsto \mathbf{u}(\mathbf{X})$ and $\mathbf{x} : \Omega_0 \rightarrow \Omega : \mathbf{X} \mapsto \mathbf{x}(\mathbf{X})$, although these functional relationships are not always explicitly stated.

In order to keep the necessary notation of this introductory chapter to a minimum, and by slight abuse of notation, the vectorial quantities \mathbf{X} , \mathbf{x} , and \mathbf{u} are also regarded as first order tensorial quantities, which does not lead to any ambiguity in this work.

The *Lagrangian description* is mostly used for fields, i.e. they are formulated as a function of the material coordinates, \mathbf{X} , if not stated otherwise. This is in contrast to the *Eulerian description*, which is with respect to spatial coordinates, \mathbf{x} . An

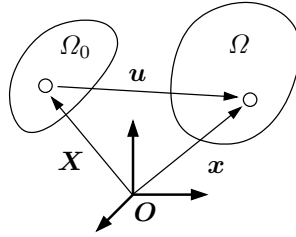


Figure 2.1: Configurations and basic kinematic quantities.

orthonormal, Euclidean basis $\mathbf{E}_I, \mathbf{E}_{II}, \mathbf{E}_{III} \in \mathbb{R}^3$ is assumed for Lagrangian quantities. The introduction of a spatial counterpart for Eulerian quantities is unnecessary for the present work and is therefore spared.

Correspondingly, the *displacement gradient* $\mathbf{H} = \partial \mathbf{u} / \partial \mathbf{X} = \mathbf{u} \otimes \nabla_{\mathbf{X}}$ and the *deformation gradient* $\mathbf{F} = \partial \mathbf{x} / \partial \mathbf{X} = \mathbf{x} \otimes \nabla_{\mathbf{X}}$ are second order tensorial quantities. These are related via $\mathbf{F} = \mathbf{H} + \mathbf{I}$, where \mathbf{I} is the corresponding second order identity tensor. The partial derivative notation of the deformation gradient follows from the relationship $d\mathbf{x} = \mathbf{F} d\mathbf{X}$ for infinitesimal line elements.

For physical reasons, the determinant of the deformation gradient, $J = \det(\mathbf{F})$, is strictly positive, $J > 0$. As it denotes the local change in volume corresponding to the deformation, a vanishing determinant would indicate a singular contraction of the body, while a negative determinant could be interpreted as self-penetration of the matter.

Hence, as the deformation gradient is invertible under any circumstance, its *polar decomposition* $\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R}$ is unique. It yields the *rotational part* \mathbf{R} , with the property $\mathbf{R}^T \mathbf{R} = \mathbf{R}\mathbf{R}^T = \mathbf{I}$ and the symmetric positive definite *right and left stretch tensors* \mathbf{U} and \mathbf{V} , respectively.

2.2 Measures of strain

2.2.1 Geometric nonlinearity

This subsection serves as a brief and standard introduction into the concept of strain tensors. Serving primarily the purpose of nomenclature definition, extended theoretical arguments are spared for the sake of brevity. The reader is referred to standard text books for more details, e.g. [Ogden, 1984, Chapter 3] or [Altenbach, 2018, Chapter 3.5]. All quantities introduced in this subsection are commonly referred to as *finite strains* or *large strains* for reasons that will be clear in the sequel.

A natural way of introducing the concept of strains in nonlinear mechanics is to start with infinitesimal line elements in the reference configuration Ω_0 and in the current configuration Ω , $d\mathbf{X}$ and $d\mathbf{x}$ respectively. Observing their coupled changes in length

due to deformation of the body, geometric considerations lead to the relationship $d\mathbf{x} \cdot d\mathbf{x} = d\mathbf{X} \cdot \mathbf{C} d\mathbf{X}$, with the *right Cauchy-Green tensor* $\mathbf{C} = \mathbf{F}^T \mathbf{F}$. This tensor hence acts as a metric tensor relating the square of the length of the line elements of the two configurations Ω_0 and Ω . Similar considerations lead to the *left Cauchy-Green tensor* $\mathbf{B} = \mathbf{F} \mathbf{F}^T$, the inverse of which has an analog geometric meaning. Just as the stretch tensors \mathbf{U} and \mathbf{V} , the right and left Cauchy-Green tensors \mathbf{C} and \mathbf{B} are rotationally invariant. This is important, as the subsequent definition of strains are required to show this invariance, too. One of the major reasons why the following line of thoughts is based on the Cauchy-Green tensors instead of the stretch tensors is that the computation of the former only requires simple arithmetics whereas that of the latter is more involved.

Measures of strain are required to yield zero not just for rotations but for rigid body motions in general. Therefore, one possibility to construct a strain tensor is by considering the difference of the squares of line elements, e.g. $d\mathbf{x} \cdot d\mathbf{x} - d\mathbf{X} \cdot d\mathbf{X}$. Substitutions and elementary modifications lead to *Green's strain tensor*,

$$\mathbf{E}_G = \frac{1}{2}(\mathbf{C} - \mathbf{I}) = \frac{1}{2}(\mathbf{H} + \mathbf{H}^T + \mathbf{H}^T \mathbf{H}), \quad (2.1)$$

as well as *Almansi's strain tensor*,

$$\mathbf{E}_a = \frac{1}{2}(\mathbf{I} - \mathbf{B}^{-1}). \quad (2.2)$$

The key characteristic by which these two quantities differ is the fact that Green's strain measures strains in the reference configuration, while Almansi's strain measures strains in the current configuration.

These two measures of strain are special cases of the more general one-parameter families

$$\begin{cases} \frac{1}{2r}(\mathbf{U}^{2r} - \mathbf{I}) & \text{if } r \in \mathbb{R} \setminus \{0\} \\ \log(\mathbf{U}) & \text{if } r = 0 \end{cases}, \quad (2.3)$$

$$\begin{cases} \frac{1}{2r}(\mathbf{I} - \mathbf{V}^{-2r}) & \text{if } r \in \mathbb{R} \setminus \{0\} \\ \log(\mathbf{V}) & \text{if } r = 0 \end{cases}, \quad (2.4)$$

parametrized by the real number r . The measures in (2.3) go back to the seminal works of Seth [1961] and Hill [1968] and are Lagrangian, while the measures in (2.4) are Eulerian quantities. The logarithmic quantities are continuous extensions to the case $r = 0$. For (2.3), this measure is called the *Hencky strain*,

$$\mathbf{E}_H = \log(\mathbf{U}), \quad (2.5)$$

which is crucial to the present work and which will be discussed after the next subsection.

2.2.2 Geometric linearity

The nonlinear nature of finite strain measures, e.g. the last term in (2.1), causes significant complexity in the formulation of the Finite Element Method (FEM). First and foremost, in case of the popular Green strain, the assembly of the element stiffness matrix is influenced by a second term, cf. [Wriggers, 2008, (4.67)]. This term does not depend on the material stiffness tensor but is a pure artifact of the nonlinear dependence of the kinematic quantity \mathbf{E}_G on the gradient \mathbf{H} of the displacement \mathbf{u} .

Furthermore, the FEM may be implemented with respect to either the Lagrangian or the Eulerian description, cf. [Wriggers, 2008, Chapter 4.2]. In combination with infinitely many possible strain measures, e.g. (2.3), (2.4), numerous considerations regarding physical modeling are to be made.

A standard simplification, cf. e.g. [Oliver & Agelet de Saracibar, 2017, Definition 2.5], [Altenbach, 2018, Chapter 10], is to assume *small displacement gradients*,

$$\|\mathbf{H}\|_{\text{Fro}} = \left\| \frac{\partial \mathbf{u}}{\partial \mathbf{X}} \right\|_{\text{Fro}} \ll 1. \quad (2.6)$$

Here, the Frobenius norm of the matrix representations of the respective tensors is used, $\|\bullet\|_{\text{Fro}} = \sqrt{\sum_{i,j=1}^3 \bullet_{ij}^2}$. In practice, this assumption is used to justify neglect of higher-order terms of the displacement gradient in strain measures.

For a displacement gradient with decreasing Frobenius norm the strain measures \mathbf{E}_G and \mathbf{E}_a asymptotically coincide. The resulting quantity emerging from either (2.1) or (2.2) is the popular *small* or *infinitesimal strain* measure,

$$\boldsymbol{\varepsilon} = \frac{1}{2}(\mathbf{H} + \mathbf{H}^\top). \quad (2.7)$$

Depending on the level of scrutiny, one may also argue that the configurations Ω_0 and Ω become indistinguishable. This is formally implied by (2.6) only for quantities depending on gradients of the displacement. An extension of this assertion to the position vectors themselves is possible, $\mathbf{X} \approx \mathbf{x}$, if rigid body movements are carefully excluded from the discussion. Here, the spatial coordinate of the infinitesimal strain field is taken to be identical to that of the displacement gradient $\mathbf{H} = \mathbf{H}(\mathbf{X})$, i.e. $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\mathbf{X})$.

Thus, the just mentioned drawbacks induced by nonlinearities of strain measures are circumvented. However, the assumption of small strains poses a severe limitation to the range of validity of the employed mathematical model. Most significant is the fact that rigid body rotations \mathbf{R} are no longer admissible as they would induce a non-zero infinitesimal strain $\boldsymbol{\varepsilon}$.

2.2.3 The special role of the Hencky strain

There is an intimate relationship between the infinitesimal strain $\boldsymbol{\varepsilon}$ and the Hencky strain \mathbf{E}_H . This has been comprehensively summarized in Neff et al. [2016] where geometric

motivations and thorough discussions regarding possible interpretations of \mathbf{E}_H are stated. For many reasons given therein, the Hencky strain should be regarded as the “natural” extension of the small strain tensor to the geometrically nonlinear description. Among the most intriguing arguments for this standpoint is the fact that the geodesic distance of the deformation gradient \mathbf{F} from the rotation group $SO(3)$ (which contains \mathbf{R}) can be expressed by the dilatational part of the Hencky strain, $\text{trace}(\mathbf{E}_H) = \log(J)$, the deviatoric part, $\mathbf{E}_H - \text{trace}(\mathbf{E}_H)/3 \mathbf{I}$ and two material dependent scalar parameters. This *physically meaningful decomposition* of the Hencky strain will be exploited multiple times in the present work.

Moreover, most of the benign properties possessed by the small strain tensor carry directly over to the Hencky strain. Among these is *linearity*: linear combinations of Hencky strains always result in Hencky strains. The analog is not true for the quantities \mathbf{F} , \mathbf{C} , \mathbf{E}_G , etc. This is due to multiple constraints that are violated by arbitrary linear combinations, the most obvious of which is the positivity of the determinant, $J > 0$. The fact that the space of Hencky strains is closed under linear combinations will prove essential in the sequel.

2.3 Compatibility conditions

The definitions of the deformation gradient, \mathbf{F} , and the small strain tensor, $\boldsymbol{\varepsilon}$, are essentially based upon the displacement, \mathbf{u} . Hence, there exists a unique field of deformation gradients and small strain tensors, respectively, for each field of displacements. Later in this work, it will be of importance to answer the inverse question: given a field of second order tensors, \mathbf{A} , does there exist a field of displacements such that \mathbf{A} is the corresponding field of deformation gradients or small strain tensors? The circumstances under which this question has an affirmative answer are known as the *compatibility conditions* and read

$$\nabla_{\mathbf{X}} \times \mathbf{F} = \mathbf{0}, \quad (2.8)$$

$$\nabla_{\mathbf{X}} \times (\nabla_{\mathbf{X}} \times \boldsymbol{\varepsilon}) = \mathbf{0} \quad (2.9)$$

for the deformation gradient, Acharya [1999], and the small strain tensor, [Lubliner, 2008, (1.2.4)], respectively. The employed notation is defined in Appendix A.

One may note that the assumption of small strains, while advantageous for the formulation of the mathematical model and the implementation of the FEM, comes along with increased complexity when considering the compatibility conditions. However, the left-hand sides of both (2.8) and (2.9) are linear, which is a key fact to the present theory.

These conditions rely on the non-trivial assumption made in Chapter 2 that the physical body be simply connected. For instance, RVE’s with continuous periodic voids are excluded by this assumption. This is not to say that such non-simply connected bodies could not be treated in general: compatibility conditions of the deformation gradient for such cases were found in Yavari [2013].

2.4 Balance equations

2.4.1 Stresses

Physical bodies exhibit reactions to changes of their kinematic state as well as to applications of forces. These reactions are forces exerted both on the spatial surrounding as well as internally among material points. Let the resulting force vector be $\mathbf{t}(\mathbf{x}) \in \mathbb{R}^3$ at each point of the current configuration.

Due to Cauchy's stress theorem, see e.g. [Wriggers, 2008, (3.64)], there exists a corresponding unique second-order tensor $\boldsymbol{\sigma}(\mathbf{x})$ that linearly relates the force vector $\mathbf{t}(\mathbf{x})$ to the normalized normal $\mathbf{n}(\mathbf{x}) \in \mathbb{R}^3$ of a (presumed) surface passing through the point $\mathbf{x} \in \Omega$,

$$\mathbf{t}(\mathbf{x}) = \boldsymbol{\sigma}(\mathbf{x})\mathbf{n}(\mathbf{x}). \quad (2.10)$$

The tensor $\boldsymbol{\sigma}$ is accordingly denoted *Cauchy stress* and entirely belongs to the Eulerian framework. The latter is to be understood in the sense that both the force vector and the surface normal are Eulerian quantities, i.e. they are defined with respect to the current configuration Ω .

Alternative stress tensors may be formulated with respect to other configurations. Most notably and most important to the present work are the following two examples. The *first Piola-Kirchhoff stress* $\mathbf{P}(\mathbf{X})$ relates surfaces normals $\mathbf{n}_0(\mathbf{X})$ of the reference configuration Ω_0 to forces $\mathbf{t}(\mathbf{x})$ in the current configuration Ω ,

$$\mathbf{t}(\mathbf{x}) = \mathbf{P}(\mathbf{X})\mathbf{n}_0(\mathbf{X}). \quad (2.11)$$

For the *second Piola-Kirchhoff stress* $\mathbf{S}(\mathbf{X})$, all quantities are assumed to be defined in the reference configuration Ω_0 ,

$$\mathbf{t}_0(\mathbf{X}) = \mathbf{S}(\mathbf{X})\mathbf{n}_0(\mathbf{X}). \quad (2.12)$$

The stress \mathbf{S} is, therefore, a Lagrangian quantity. A useful relation among the aforementioned two Piola-Kirchhoff stresses is the equation $\mathbf{P} = \mathbf{F}\mathbf{S}$.

2.4.2 Balance of momentum and conservation of mass

With these modeling capabilities at hand, the differential equations necessary for a complete description of the mechanical behavior of the considered physical body can be stated. Again, multiple different formulations are equally valid and may be chosen depending on the preferred framework. Here, the Lagrangian description with respect to the first Piola-Kirchhoff stress is sufficient:

$$\text{Div}_{\mathbf{X}}(\mathbf{P}) + \mathbf{b}_0(\mathbf{X}) = \mathbf{0} \quad + (\text{BC}) \quad (2.13)$$

The divergence operator is accordingly formulated with respect to the reference coordinates. The vector $\mathbf{b}(\mathbf{X}) \in \mathbb{R}^3$ denotes the volumetric force load. No inertia terms are

considered in this work. These equations, essentially based upon Newton's third law of motion, thus define the *quasi-static balance of linear momentum*. Alternative formulations as well as additional details regarding their derivations may be found in standard text books, e.g. [Wriggers, 2008, Chapter 3.2], [Altenbach, 2018, Chapter 5.2.2].

Sufficient boundary conditions (BC) have to be provided in order to render the problem (2.13) well-posed. Kinematic, static and mixed boundary conditions are suitable to this end, with some kinematic constraints eliminating at least the rigid body motion being necessary.

Considerations on the *balance of angular momentum* lead to the following condition on the stress tensors:

$$\mathbf{F}^{-1} \mathbf{P} = \mathbf{P}^T \mathbf{F}^{-T}. \quad (2.14)$$

Although this system of equations does not require attention in the formulation of standard solution schemes to (2.13), e.g. the FEM, attention must be paid with certain numerical methods that approximate the stress on a component-by-component basis.

Furthermore, *conservation of mass* is stated as

$$\rho_0 = J\rho, \quad (2.15)$$

where ρ_0 and ρ are the mass densities in the reference and in the current configuration, respectively. The physical interpretation of this equation is that related infinitesimal volume elements in the initial and the current configuration must contain the same mass. This highlights the meaning to the determinant J and its strict positivity requirement. Algorithmically, the condition (2.15) does not require dedicated attention in the numerical methods considered in the present work. However, the determinant J may become negative during intermediate iterations in which case the failure of the scheme is physically necessary. The circumstances under which this might occur are discussed later in this work.

In the case of geometric linearity the balance equations of linear momentum,

$$\text{Div}_{\mathbf{X}}(\boldsymbol{\sigma}) + \mathbf{b}_0(\mathbf{X}) = \mathbf{0} \quad + \text{(BC)}, \quad (2.16)$$

are formulated with respect to the unique measure of stress, $\boldsymbol{\sigma}$, and unique spatial coordinates, \mathbf{X} , as discussed in Chapter 2.2.2. Again, necessary kinematic boundary conditions must be prescribed. Angular momentum is balanced via

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^T. \quad (2.17)$$

2.5 Hyperelasticity

The partial differential equations (2.13) require closure, i.e. the relationship between the stress field \mathbf{P} and displacement field \mathbf{u} are yet to be defined. By the principle of translational invariance, the relationship is formulated in terms of gradient quantities

such as \mathbf{H} or \mathbf{F} . The most essential facts relevant to the present work are now stated and the reader is referred to standard literature for more details, e.g. [Ogden, 1984, Chapter 4], [Hackett, 2016, Chapters 1, 4].

In this work, all bodies are assumed to consist of *hyperelastic materials*. Hyperelasticity is characterized by the existence of a single scalar quantity depending only on the deformation gradient and completely describing the material response. This quantity is denoted the *stored-energy density* function $W_{\mathbf{F}}(\mathbf{F})$. The principle of local action is assumed, i.e. $W_{\mathbf{F}}(\mathbf{F}) = W_{\mathbf{F}}(\mathbf{F}(\mathbf{X}))$. Oftentimes, the stored-energy density is equivalently formulated with respect to other measures of deformation, e.g. $W_{\mathbf{C}}(\mathbf{C})$, or strain, e.g. $W_{\mathbf{E}_G}(\mathbf{E}_G)$. The physical dimension is $\frac{\text{energy}}{\text{volume}}$ and it carries the derived SI unit $\frac{\text{N}\cdot\text{m}}{\text{m}^3}$.

The stored-energy density function gives rise to the stress and the stiffness by means of the first and the second derivative, respectively, e.g.

$$\mathbf{P} = \frac{\partial W_{\mathbf{F}}}{\partial \mathbf{F}} \qquad \mathbb{C}_{\mathbf{F}} = \frac{\partial^2 W_{\mathbf{F}}}{\partial \mathbf{F}^2} \qquad (2.18)$$

$$\mathbf{S} = \frac{\partial W_{\mathbf{E}_G}}{\partial \mathbf{E}_G} \qquad \mathbb{C}_{\mathbf{E}_G} = \frac{\partial^2 W_{\mathbf{E}_G}}{\partial \mathbf{E}_G^2}. \qquad (2.19)$$

As the gradient kinematic quantities are dimensionless, the physical dimension associated with both the stress and the stiffness is $\frac{\text{force}}{\text{area}}$, employing the derived SI unit for pressure, Pa.

By the term *material response*, in this work, the stress and the stiffness are jointly addressed.

The determination of a stress which is energy-conjugate to a strain in the sense of (2.18) and (2.19) is highly non-trivial for some members of the strain families (2.3) and (2.4). Most notably, the stress conjugate to the important Hencky strain \mathbf{E}_H has widely only been utilized under additional simplifying assumptions, cf. [Ogden, 1984, Equation (3.5.23)]. The general \mathbf{E}_H -conjugate stress has been determined no earlier than in [Hoger, 1987, Equation (49)]. There, it was also proven that the Eulerian logarithmic strain $\log(\mathbf{V})$ – in general – does not have a conjugate stress.

In the context of small strains, i.e. for (2.16), uniqueness prevails and it is straightforward to define

$$\boldsymbol{\sigma} = \frac{\partial W_{\varepsilon}}{\partial \boldsymbol{\varepsilon}} \qquad \mathbb{C}_{\varepsilon} = \frac{\partial^2 W_{\varepsilon}}{\partial \boldsymbol{\varepsilon}^2}. \qquad (2.20)$$

More general kinds of elasticity theories exist, such as Cauchy elasticity, cf. [Ogden, 1984, Chapter 4.2], or Gradient elasticity, Mindlin & Eshel [1968]. While the latter enlarges the class of stored-energy density functions to also possibly depend on second gradients of the displacements, the former does not require a stored-energy density

function to exist at all but only assumes a functional relationship between the stresses and the strains. These theories are beyond the scope of the current work.

Unique solubility of the partial differential equations (2.13) respectively (2.16) is assumed throughout this work. That is, the topic of different kinds of convexity of the hyperelastic energy density function and related stability issues are not addressed here. The interested reader may find an introduction to this topic in [Ball, 1976].

Chapter 3:

Introduction to homogenization

Nevertheless, our result proves rigorously that even a perfect, complete description of the microscopic interactions between a material's particles is not always enough to deduce its macroscopic properties.

[Cubitt et al., 2018, p. 37]

3.1 Many-scale modeling

It is assumed that the physical body under consideration exhibits geometrical properties that are of $M + 1$ characteristic sizes on the order of

$$0 \text{ m} < \overset{(0)}{L} \ll \overset{(1)}{L} \ll \dots \ll \overset{(N)}{L} \ll \dots \ll \overset{(M)}{L} \quad (\text{unit m, } M \in \mathbb{N}). \quad (3.1)$$

The range of each of the sizes $\overset{(N)}{L}$ ($0 \leq N \leq M$) defines its corresponding length *scale*.

The scale with the smallest characteristic size $\overset{(0)}{L}$ is bounded from below only by the general assumptions of continuum mechanics. The number M is assumed small enough such that the theory of hyperelasticity can be considered reasonable.

Periodicity of the geometric features on scales $N < M$ is assumed throughout the whole domain. Details on local periodicity are postponed until later in this chapter. For theories avoiding the assumption of periodicity, the reader is referred to standard literature such as [Suquet, 1987, Chapter 2.1] and the references therein.

The many-scale model is now deduced successively in a scale-by-scale manner. Adapting the notation from Chapter 2 to the current setting, the original domain on the structural engineering scale M is denoted $\overset{(M)}{\Omega}_0 \ni \overset{(M)}{\mathbf{X}}$. At this scale, all geometrical features on the order of $\overset{(M-1)}{L}$ and below are neglected for the moment: $\overset{(M)}{\Omega}_0$ may be treated by the engineer as if no complicated micro-structure was present. In terms of classical FE analysis, this is to speak that the mesh of the domain $\overset{(M)}{\Omega}_0$ may be comparatively coarse and need not to resolve geometric sub-scale objects, topologies etc. At each material point $\overset{(M)}{\mathbf{X}}$, a homogeneous material is assumed to exist regardless of a possibly neglected sub-scale heterogeneity. Under this assumption, the balance of linear momentum (2.13) becomes

$$\text{Div}_{\overset{(M)}{\mathbf{X}}}(\overset{(M)}{\mathbf{P}}) + \overset{(M)}{\mathbf{b}}_0(\overset{(M)}{\mathbf{X}}) = \mathbf{0} \quad + \text{(BC)}. \quad (3.2)$$

Closure of these equations requires a constitutive law relating the stress $\mathbf{P}^{(M)}$ to the deformation gradient $\mathbf{F}^{(M)}(\mathbf{X}^{(M)})$. As in general no explicit form of this relationship is given, the sub-scale information that is neglected in (3.2) must be taken into account.

To this end, the concept of a *Representative Volume Element* (RVE) is utilized: let $\Omega_0^{(M-1)} \ni \mathbf{X}^{(M-1)}$ be a cuboid region within the original domain containing one period of the geometrical features of size $L^{(M-1)}$. This implies that no voids are present within the RVE. Porous domains will be treated later in this chapter. For efficiency reasons, RVE's are usually chosen to contain only a single period. There, the next smaller scale properties of sizes $L^{(N)}$, $0 \leq N < M - 1$, are neglected for the moment and only those on the order of $L^{(M-1)}$ are considered. The RVE $\Omega_0^{(M-1)}$ is representative of the scale $M - 1$ and thought to be attached to each material point $\mathbf{X}^{(M)}$ of the macroscopic scale. More details on RVE's will be discussed in the next section.

It is a key result of homogenization theory, cf. e.g. [Bensoussan et al., 1978, Chapter 1, Remark 8.5 and Section 16], Pruchnicki [1998], that the closure of (3.2) can be implicitly achieved via the following boundary value problem on the scale $M - 1$:

$$\text{Div}_{\mathbf{X}^{(M-1)}}(\mathbf{P}^{(M-1)}) = \mathbf{0} \quad + (\text{PFBC}; \mathbf{F}^{(M)}(\mathbf{X}^{(M)})). \quad (3.3)$$

These balance equations are coupled to (3.2) via kinematic *periodic fluctuation boundary conditions* (PFBC). These emanate from the kinematic state of the macro-structure at a material point, $\mathbf{F}^{(M)}(\mathbf{X}^{(M)})$, and are characterized by periodic fluctuations of the displacement $\mathbf{u}^{(M-1)}$ and anti-periodic fluctuations of the traction $\mathbf{t}_0^{(M-1)}$ on the boundary $\partial \Omega_0^{(M-1)}$ of the RVE. From here, elementary considerations lead to the scale coupling relationship

$$\langle \mathbf{F}^{(M-1)} \rangle_{\Omega_0^{(M-1)}} = \mathbf{F}^{(M)}(\mathbf{X}^{(M)}) \quad (3.4)$$

where the general *volume averaging operator*

$$\langle \bullet \rangle_{\Omega_0^{(N)}} = \frac{1}{\text{vol}(\Omega_0^{(N)})} \int_{\Omega_0^{(N)}} \bullet \, dV^{(N)} \quad (0 \leq N \leq M - 1) \quad (3.5)$$

is employed with $N = M - 1$.

In the two-scale setting, i.e. when $M = 1$, the equations (3.3) are closed as the material laws on the smallest scale, $N = 0$, are assumed to be known explicitly. In the *many-scale* case, i.e. when $M > 1$, the same procedure is “reiteratively” (coined in [Bensoussan et al., 1978, Chapter 1, Remark 8.5]) performed on scales of decreasing sizes: for each scale $N = M - 1, \dots, N = 0$, the boundary value problem

$$\text{Div}_{\mathbf{X}}^{(N)} \left(\mathbf{P} \right) = \mathbf{0} \quad + \text{(PFBC; } \mathbf{F}^{(N+1)} \left(\mathbf{X} \right) \text{)} \quad (3.6)$$

is stated with the boundary condition represented by the deformation gradient on the next larger scale, $\mathbf{F}^{(N+1)} \left(\mathbf{X} \right)$. Notably, there exists a cascade of domains $\Omega_0^{(N)}$ transitioning from the engineering domain on scale $N = M$ down to the RVE of the smallest considered scale $N = 0$. Accordingly, equation (3.4) becomes

$$\left\langle \mathbf{F} \right\rangle_{\Omega_0}^{(N)} = \mathbf{F}^{(N+1)} \left(\mathbf{X} \right) \quad (3.7)$$

The closure of the balance equations (3.6) is standard for $N = 0$ due to the assumption of explicitly known hyperelastic material laws. For $N > 0$, the closure is performed in ascending order, $N = 1, \dots, M$. To this end the fundamental *Hill-Mandel condition* is employed,

$$\left\langle \mathbf{P} \cdot \frac{d\mathbf{F}}{dt} \right\rangle_{\Omega_0}^{(N)} = \mathbf{P}^{(N+1)} \cdot \frac{d\mathbf{F}^{(N+1)}}{dt}, \quad (3.8)$$

cf. Hill [1965], Mandel [1972]. Here $t \in \mathbb{R}$ represents the temporal variable in non-static problems. This condition represents conservation of power across scales and allows for the formulation of the upscaling relations

$$\left\langle \mathbf{P} \right\rangle_{\Omega_0}^{(N)} = \mathbf{P}^{(N+1)} \left(\mathbf{X} \right), \quad (3.9)$$

which are applicable to the present quasi-static case. *The systems of equations (3.6) together with the characterizations (3.7) of the boundary conditions are posed in a top-down manner; $N = M - 1, \dots, 0$ and closed via (3.9) in reverse order, $N = 0, \dots, M - 1$.*

These systems of equations translate directly to the case of geometric linearity,

$$\text{Div}_{\mathbf{X}}^{(M)} \left(\boldsymbol{\sigma} \right) + \mathbf{b}_0^{(M)} \left(\mathbf{X} \right) = \mathbf{0} \quad + \text{(BC)}, \quad (3.10)$$

$$\text{Div}_{\mathbf{X}}^{(N)} \left(\boldsymbol{\sigma} \right) = \mathbf{0} \quad + \text{(PFBC; } \boldsymbol{\varepsilon}^{(N+1)} \left(\mathbf{X} \right) \text{)} \quad (3.11)$$

with $0 \leq N < M$, in correspondence to (2.16).

Lastly, as with the hyperelastic laws on the smallest scale $N = 0$ in Chapter 2.5, it shall be noted that stability issues are not subject to the current study. The interested reader is referred to [Braides, 1994] for classical and to [Schneider, 2016] for more recent results on this topic in the context of homogenization theory.

3.2 Representative Volume Elements and pores

The concept of a Representative Volume Element was introduced in the context of homogenization theory at least since Hill [1963], albeit rather informally. A modern and comprehensive disquisition on this topic can be found, e.g., in Ostoja-Starzewski [2006]. Multiple qualitative definitions are put into context in [Schröder, 2014, Section 2.4]. In essence, the idea behind the RVE is to choose a subdomain of a sub-scale structure, small enough to purge it of redundancy but large enough to maintain representativeness.

The definition of an RVE is straight-forward for the case of periodicity of microscopical features, as described in Chapter 3.1. However, non-periodicity is oftentimes present both in natural and in industrial composite materials. In such cases, more general concepts such as ergodicity or statistical homogeneity can be taken into account, cf. [Torquato, 2006, Section 2.2.2].

The many-scale case $M > 1$ is most meaningful when periodicity occurs only locally. Examples for this scenario are small inclusion particles surrounded by a heterogeneous matrix material or foams containing pressure-free voids with sizes of different orders of magnitudes. In such cases, one would choose a distinct RVE $\overset{(N-1)}{\Omega}_0(\overset{(N)}{\mathbf{X}})$ for a material point $\overset{(N)}{\mathbf{X}} \in \overset{(N)}{\Omega}_0$ ($N > 0$), depending on the micro-structure in the vicinity of $\overset{(N)}{\mathbf{X}}$. This includes the trivial cases of either an analytical material model if there is no heterogeneity near $\overset{(N)}{\mathbf{X}}$, or no material at all if $\overset{(N)}{\mathbf{X}}$ lies within a void. In the latter case, one must take the domain $\overset{(N)}{\Omega}_0$ to also contain the voids when evaluating the volume averaging operator. More precisely, for the operator defined by the right-hand side of (3.5) the RVE $\overset{(N)}{\Omega}_0$ is actually replaced by the smallest cuboid containing $\overset{(N)}{\Omega}_0$, and the integrated quantity is assumed to be zero-continued inside the voids.

The determination of an appropriate size for an RVE is a delicate topic. Even the term “size” may depend on the context since it may refer to either an absolute measure in the dimension of length or to a relative measure counting the number of features that are included in the RVE, cf. Fritzen et al. [2012]. Among others, Kanit et al. [2003] is a standard reference in this regard. Gitman et al. [2007] additionally addresses the issue of non-existence of RVE’s under certain conditions. In the present thesis, the existence of an RVE and the availability of its definition is assumed for $0 \leq N < M$.

3.3 The effective tangent modulus

For the widely used Finite Element Method, it is necessary to provide the tangent modulus to the stress. More precisely, for the formulation with respect to the initial configuration (3.6), it is usually required to have access to the stiffness tensor $\mathbb{C}_{\text{EG}}^{(N)} = \partial \mathbf{S} / (\partial \mathbf{C})$. The fully stated dependency $\mathbb{C}_{\text{EG}}^{(N)} = \mathbb{C}_{\text{EG}}^{(N)}(\mathbf{X}; \mathbf{S}(\mathbf{X}))$ is abbreviated for the sake of readability. For $N > 0$, there is no simple expression for the stiffness and multiple methods for evaluation or approximation were proposed in the past.

In [Miehe, 1996, Box 1], a finite difference approach was proposed for the approximation of general fourth-order stiffness tensors, i.e. not necessarily in the context of homogenization. Seven stress evaluations are required and the accuracy depends on the perturbation parameter which is not a straight-forward choice.

It was briefly mentioned in [Smit et al., 1998, Section 5.2] that the effective tangent modulus may be computed explicitly from microscopical knowledge. This was much more comprehensively described in later works, e.g. [Miehe, 2003, Box 4]. There, a large system of equations for all boundary nodes of the RVE is assembled and intricately modified into the effective tangent modulus. This method utilizes the Schur complement of the global stiffness matrix of the RVE with respect to the boundary nodes.

Earlier, a closed form of the effective stiffness tensor $\mathbb{C}_{\text{F}}^{(N)}$ was given in [Miehe et al., 1999a, Equation (31)]. It was based on sensitivity analysis, cf. also [Saeb et al., 2016, Theorem 2.3.4], and has very descriptive physical interpretations: the naive and overly stiff volume average $\langle \mathbb{C}_{\text{F}}^{(N-1)} \rangle$ is softened by means of subtraction of contributions from fluctuations. However, the complexity of the FE discretization is not reduced, wherefore the computation is impractical for real-world applications.

The softening effect originating from microscopic fluctuations is comprehensively described in [Miehe, 2002, Section 3.3]. An analog phenomenon persists even if an order reduction technique is applied to the inelastic phases of the material, cf. [Fritzen & Hodapp, 2016, Section 4.4]. In the latter case, a compact method for the evaluation of the consistent tangent modulus is provided. It can be computed by means of six solutions of a comparatively low-dimensional linear system of equations.

A shortcut was taken in Yvonnet & He [2007]. There, the overall stiffness of a porous RVE was roughly approximated by assuming it to be identical with the stiffness of the matrix material. This is a harsh simplification of the real circumstances and may be severely detrimental to the accuracy of the stiffness, depending on the volume fractions of the voids or of different material phases. It is noteworthy that in the case of different phases, it is generally preferable to choose the highest of the present stiffnesses. This is due to the fact that an over-estimation of the tangent modulus is likely to just increase the number of necessary iterations while an under-estimation is detrimental to the stability of Newton-Raphson solution schemes.

An alternative formulation based on a penalty method was developed in Temizer & Wriggers [2008]. There, a single evaluation of the homogenized stiffness leads to compute time on the order of a minute. Even when accounting for the advances of available hardware ever since, this method is not suitable for realistic simulations of multiple scales.

Therefore, the important question of how the effective tangent modulus may be efficiently approximated is still open. Progress at this front will be presented in the sequel.

3.4 Short summary of homogenization methods

3.4.1 Classical analytical methods

The history of homogenization theory reaches back at least to the nineteenth century. At that time, pioneering German physicist Voigt expressed a certain discontent with the lack of agreement between the theoretical material models developed by Poisson on the one hand, and measurement data on the other hand, Voigt [1889]. By considering the multi-crystalline structure of metals and rocks, the question of the effective material stiffness was brought up. In this seminal work, the arithmetic mean of the local stiffness values was proposed as an approximation, [Voigt, 1889, eq. (7)]. Effectively, this method assumed a *homogeneous strain* distribution throughout the multi-crystalline microstructure.

Decades later, an alternative approach was introduced by Reuß who suggested to additionally consider the case of a *homogeneous stress* distribution on the microscopic scale, [Reuß, 1929, Section 5]. For the stiffness tensor, this assumption lead to taking the harmonic mean of the local stiffness values. Reuß explicitly stated that the assumption of either homogeneous stresses or homogeneous strains is necessary in order to perform any calculations at all. It was not within the scope of his work to discuss which of those exclusionary methods was more realistic. He did, however, compute approximations to the effective yield stress of heterogeneous crystalline materials, too.

The works by Taylor [1938] and Sachs [1929] are credited with additional key contributions in line with Voigt and Reuß, respectively, in the context of metal plasticity. These included the explicit consideration of the shear components of the strains and stresses, which were particularly meaningful for metals.

In Boas & Schmid [1934], the methods of Voigt and Reuß were compared to experimental data. It was found that Voigt's method yielded Young's moduli and torsion moduli above the measured data. In contrast to this, Reuß' method resulted in moduli lower than the experimental results. These observations were mathematically proven a general rule in Hill [1952], i.e. the effective material properties computed by means of these two different method constituted upper and lower bounds to the actual quantities.

Additional analytical, increasingly sophisticated methods were developed thereafter. For more information in this regard, the reader is referred to the references provided in the

above mentioned sources and exhaustive standard literature, such as Nemat-Nasser & Hori [1999], Yvonnet [2019]. Brief overviews are given by various review articles, such as Saeb et al. [2016], Geers & Yvonnet [2016].

3.4.2 Modern computational and hybrid methods

In Smit et al. [1998], the concept of “multi-level” FE simulations was introduced. The paradigm followed by this branch of multi-scale mechanics is a compromise between the analytical school, e.g. as in Chapter 3.4.1, and the naive, brute-force approach of fully resolving the microstructure on the engineering scale. The former suffers from the drawback of relying on restrictive or overly simplifying assumptions, the latter is impractical for realistic applications due to the extreme demands it makes on computational resources.

In essence, multi-level FE methods are specific realizations of the homogenization scheme of Chapter 3.1 for $M = 1$ with both the macroscopic problem (3.2) and the sub-scale problems (3.3) being solved by FE computations. It was the work Feyel [1999] that popularized this concept under the descriptive term “FE²” (read: “F E square”). Depending on the specific setup, possibly much less memory is required compared to the full resolution of the microstructured macroscopic simulation. Furthermore, multi-level schemes always allow for trivial parallelization during the assembly of the macroscopic quantities. However, the computational complexity is still extreme, rendering methods of this kind impractical for real-world applications even today. On the upside, it must be mentioned that, at least in theory, arbitrary accuracy can be realized by multi-level schemes. This is also due to the fact that no simplifying or phenomenological modeling assumptions are required, except for the separability of scales which is mathematically backed by convergence theories. Therefore, after refinement and generalization, e.g. by Miehe [2003], this class of schemes became the benchmark for future ideas and still serve this purpose today.

Earlier, another line of thought was pursued with the establishment of the “Transformation Field Analysis” (TFA), cf. Dvorak & Benveniste [1992], Dvorak et al. [1994]. Assuming the scale-coupled problem formulations (3.2), (3.3), and the existence of an RVE, the TFA is a way of reducing the number of unknowns in inelastic processes. Initially, on the microscopic scale, the internal variables are assumed to be constant within each material phase. Thus, the coefficients of the microscopic fields of internal variables become macroscopic quantities. However, this simplifying assumption in turn requires assumptions on the evolution equations of these macroscopic objects. From a modeling perspective, this reliance on additional closure equations is a significant disadvantage. Furthermore, it soon became obvious that the assumption of phasewise constant internal variables was an oversimplification of reality and that accuracy suffered from this. This assumption was therefore soon eased to piecewise uniform fields within the material phases. Also, the introduction of additional sub-domain within the phases could increase the accuracy, albeit at additional computational costs. An extension of the method

including damage models can be found in Kruch & Chaboche [2011] and the review article Kanouté et al. [2009] is suggested.

As it became clear that more resolution on the microscopic scale was necessary, a generalization of the TFA was proposed in Michel et al. [2002], Michel & Suquet [2003]. The main idea behind this “Nonuniform Transformation Field Analysis” (NTFA) was to allow the fields of internal variables to be linear combinations of previously determined, actually occurring, varying fields. However, the determination of these fields was proposed in an ad hoc manner only, lacking generality. It shall be noted that the concept of an offline stage was established within the homogenization community at least since then. While much more accurate than the TFA, this method still suffered from the inherent drawback of relying on additional modeling assumptions on the macroscopic scale.

Another significant step in the evolution of this family of methods was the “Potential Based Reduced Basis Model Order Reduction” (pRBMOR), Fritzen & Leuschner [2013]. By means of this hybrid computational method, many of the issues associated with the NTFA were solved. First and foremost, no modeling assumptions on the macroscopic scale were necessary any longer. In addition, the incorporated alternative formulation variational formulation was less restrictive in terms of material laws on the micro-scale. Concerning the offline phase, a rather systematic procedure was proposed based on classical best-approximation theory. It was demonstrated that the pRBMOR is applicable to real-world multi-scale problems by means of massively parallel implementations on GPU’s, Fritzen et al. [2014], Fritzen & Hodapp [2016]. However, this method was limited to the case of small strains. The generalization to finite strains is still an open question.

Although no longer competitive by itself, the TFA approach was picked up rather recently and combined with other methods. The “Self-Consistent Clustering Analysis” (SCCA), Liu et al. [2016], combined the original TFA idea with the Lippmann-Schwinger equation (which is comprehensibly derived in, e.g., [Schneider et al., 2017, Section 4.1]). This novel method was adapted to the regime of finite strains, Yu et al. [2019]. Notably, the online CPU time was shown to be reduced with respect to reference methods under certain circumstances. On the other hand, the CPU time of the SCCA grows with a higher order than that of its benchmark schemes as the resolution of the microstructure is increased. This is significant since only structured voxel discretizations, i.e. regular grids, are considered and, therefore, very fine-grained microstructures could not be treated as efficiently as with sophisticated FE meshing techniques.

In fact, this issue with voxel discretizations is shared by many homogenization schemes based on the “Fast Fourier Transformation” (FFT), which is a popular operation whenever convolutions need to be computed en masse. Prominent examples include the already mentioned family of Lippmann-Schwinger equation based methods, e.g. Moulinec & Suquet [1998], Schneider et al. [2017] in addition to the SCCA. A method equivalent to Schneider et al. [2017] but derived from FE-formulations and not explicitly dependent on the Lippmann-Schwinger equation is presented Leuschner & Fritzen [2018]. Another use case of the FFT is the projection of vectors onto curl-free or

divergence-free spaces, as has been done in Rambašek et al. [2019] for homogenization of magneto-active materials. Although distinct algorithmic differences are present, all of these schemes share the essence of evaluating convolutions in Fourier space. Another commonality is the reliance on regular discretization. This is a core feature as the employed formulations in Fourier space are based on the translational invariance of the spatial discretization. On one hand, this is possibly advantageous for automated image-based modeling. However, the required global spatial resolution is possibly very large and may result in much more degrees of freedom than an unstructured finite element mesh, e.g. when comparing to second-order tetrahedra.

Another successful branch of efficient mechanical simulations is based on projection. Belonging to the set of “Reduced Order Modeling” (ROM), these methods classically employ a projection of the primal variables, e.g. displacements, onto a low-dimensional subspace. The key characteristic of ROM is that the underlying mathematical problem, e.g. (2.13) or (2.13), is unchanged but only the complexity associated with its solution is reduced. Naturally, a certain sacrifice of accuracy has to be accepted along with the order reduction.

The reduction of the number of degrees of freedom is widely conducted by means of the “Karhunen-Loève expansion”, Karhunen [1946], Loève [1963], which also goes by the name “Proper Orthogonal Decomposition” (POD), Sirovich [1987]. Essentially, it extracts typical fields of the primal variables from a large set pre-computed, unreduced solution fields. Famously, the “a priori Hyperreduction”, Ryckelynck [2005], contributed to the field by reducing the large amounts of costly pre-computations and by coupling this with an adaptive procedure. In the field of mechanical homogenization of solids, the projection method was introduced with the work Yvonnet & He [2007] the latest. There, the displacement degrees of freedom of the micro-scale problem (3.3) are reduced. The above mentioned pRB MOR utilizes the POD on the micro-scale, too, although not on the primal but on the internal variables. This has even been extended to displacement discontinuities at microstructural cohesive interfaces, see Fritzen & Leuschner [2015].

In general, POD based schemes are supported by accessible mathematical theorems that admit physical interpretations. However, original implementations may come along with difficulties when strong nonlinearities are involved or when adaptivity is required. Countermeasures include, e.g., a larger set of pre-computed solution fields and a larger set of basis fields, i.e. less reduction of the number of degrees of freedom.

3.4.3 Data-based methods

The era of modern computing initiated another branch of efficient homogenization methods. Just as in many other parts of science, rapid advancements on both the hardware and the software side lead to increasingly available *in silico* data. As a consequence, it was worthwhile exploring whether the explicit mathematical modeling on the micro-scale (3.3) could be replaced by data. The basic idea was that this data would implicitly contain the underlying physical laws and serve as the setup data for a numerical surrogate

of the effective material law. In contrast to Reduced Order Modeling, where solely the order of the approximation was efficiently reduced, data-based methods compromise the original model for efficiency.

Perhaps the earliest occurrence of such a method in homogenization literature was the idea of a "material map", Temizer & Zohdi [2007]. Assuming effective isotropy of the microstructure, this method interpolated the coefficients of a simple constitutive equation on the macro-scale between values determined by means of pre-computations. It was set up in the context of geometric nonlinearity and quickly extended to the case of effective orthotropy, Temizer & Wriggers [2007]. Although only academic numerical experiments were considered, the limited available compute power was shown to be significantly more efficiently used with the proposed method than with the FEM. One of the most significant contributions was the extremely quick computation of the macroscopic stiffness, which is the most expensive part of the unreduced FE².

A more general way of approximating the effective material response was proposed in Yvonnet et al. [2009] with so-called "Numerically Explicit Potentials" (NEXP). There, the effective hyperelastic energy function was approximated by means of interpolation techniques in the space of small strains. Among the candidate methods, the ansatz based on separation of variables proved most efficient due to the fact that the response could be approximated by means of multiple one-dimensional interpolations. The interpolation was supported by pre-computed data which had been obtained by means of unreduced FE simulations on the microscale. Therefore, the setup phase was very demanding in terms of computational efforts and crucially depended on the number support points, i.e. on the density of the sampling of the strain space. More difficulties were encountered with the adaption to geometric nonlinearity, Yvonnet et al. [2013], where the microscopic problem, similar to (3.3) could not be solved for all of the sampled macroscopic right Cauchy-Green tensors $\overset{(M)}{\mathbf{C}}$ ($M = 1$). Hence, an extrapolation based on additional assumptions was necessary. Overall, this interpolation method delivered highly accurate results for simple test cases and clearly outperformed the classical FE² method. However, it obviously suffered from the curse of dimensionality during the setup phase, not reducing the computational efforts for realistic three-dimensional examples to manageable levels.

A very general method of function approximation are Artificial Neural Networks (ANN's). An outline of recent developments in this active field of research can be found in Abiodun et al. [2018], as well as a short introduction. Probably the first application of this popular method in the field of computational mechanics was in the context of material modeling, Wu et al. [1990], Ghaboussi et al. [1991]. Homogenization was subject to treatment by ANN's, e.g., in Le et al. [2015]. There, it was attempted to approximate the effective hyperelastic energy density of an RVE as a function of the (small) strain. The well-known drawbacks of classical ANN's, especially the very large amounts of necessary setup data, proved a limitation to the ansatz: as the training data was obtained by means of unreduced FE simulations, significant compromises had to be made either on the density of the sampling, on the complexity of the RVE or on the compute time. Nonetheless, many others have employed ANN's in the context of

homogenization, including but not limited to Barkaoui et al. [2013], Lißner & Fritzen [2019], Göküzüm et al. [2019], Fritzen et al. [2019], Nguyen-Thanh et al. [2020], Rao & Liu [2020]. There, the details of the specific ansatzes vary widely, from the direct approximation of the target function and iterative ANN applications to hybrid methods that adaptively switch between ANN's and more classical methods.

Hyperelastic materials and any other path-independent types of materials may appear obvious candidates for modeling via ANN's. This is due to the fact that path dependency, such as with dissipative material laws, leads to a significant increase of the dimension of the input space. As this input space requires comprehensive sampling, an overly large dimension may render the (naive) setup of ANN's practically impossible. However, there are delicate issues that require special attention in the case of hyperelasticity, too. One example is the fundamental requirement of zero energy corresponding to zero strain. Also, smoothness and monotonicity of the energy density and the stress are strongly desirable both from a modeling perspective and from the viewpoint of numerical stability. One can conclude that this class of methods faces significant challenges but exhibits a very promising development with increasing success.

The "Matrial Knowledge System" (MKS), Kalidindi et al. [2010], is a mathematical framework for efficient homogenization of microstructure properties based on regular voxel discretizations, the FFT and influence coefficients. These influence factors are pre-computed in the setup phase of the method by means of a regression analysis based on FE results. After this, the homogenized response function can be quickly evaluated. The key contribution is that the opposite direction of homogenization, i.e. the determination of the micro-structural fields from macroscopic loading conditions, can also be computed very efficiently. This is a major advantage over more methods based on function approximation, such as ANN's and interpolation schemes, in which the micro-structural information is lost during the homogenization. An associated software project, Brough et al. [2017], has become very popular and contributed to the accessibility of the method.

Chapter 4:

Outline

The present work proposes a multi-staged computational method for the efficient approximation of hyperelastic material responses that are homogenized successively on multiple spatial scales. This generalizes and advances the field of *in silico data-driven methods*, cf. Chapter 3.4.3. Certain aspects of this work also contribute to more general areas of research, such as efficient interpolation of general functions in moderate-dimensional spaces. Moreover, insight into the purely mathematical field of potential theory is gained while investigating the setup of such interpolation schemes.

Along the way to these achievements, issues both on the computational and on the conceptual side have to be solved. The computational issues mostly originate from the degree of detail with which RVE's are resolved – not only on one but on *numerous hierarchical scales*. From the theoretical side, the treatment of both geometric and material nonlinearities poses a significant challenge. Sophisticated methods are required for both the setup and the application phase of the data-driven model: nonlinearities of both kinds necessitate specialized techniques for the *discretization of the input space*. The outputs, too, are obtained by means of dedicated techniques formulated with respect to carefully chosen *function spaces*. The overall achievement is reached with four journal articles that are outlined in the following sections.

Only the main concepts are stated and few key equations are presented in this outline. In particular no emphasis is put on completeness at this point. Comprehensive details are given in the chapters containing the papers.

4.1 First publication – Two-stage data-driven homogenization of nonlinear solids using a Reduced Basis model

To begin with, the comparatively simple case of geometric linearity and only a single scale transition is addressed, $M = 1$. By means of this simplifying choice, the main ideas can be conceptually studied without the hurdle of nonlinear kinematic quantities and without cumbersome arbitrary-scale case differentiation. However, it is important to point out that pronounced nonlinearities are considered with respect to the material laws: pseudo-plasticity models with and without hardening imitate dissipative material behavior for proportional loading conditions.

The *two-stage reduction concept* is introduced in this initial phase of research. Building on data generated by means of high-fidelity simulations, e.g. FE analyses, the first and the second reduction stage each introduce a significant decrease of computational complexity as compared to the original high-fidelity method. Naturally, trade-offs have to be accepted, which in this case consist of additional approximation errors and increased setup times.

4.1.1 Projection based Reduced Order model

Roughly speaking, during the *first stage*, data from the FE simulations is processed in order to setup a projection based reduced order model. This model serves the purpose of approximating solution strain fields, i.e. the strain fields $\overset{(0)}{\boldsymbol{\varepsilon}}(\mathbf{X}; \overset{(1)}{\boldsymbol{\varepsilon}})$ associated with the solutions of (3.11). This efficient strain approximation on the RVE in turn yields an efficient approximation of the homogenized material response to the strain on the macroscopic scale, $\overset{(1)}{\boldsymbol{\varepsilon}}$.

More precisely, for the setup of the first stage, the strain fields corresponding to FE solutions of (3.11) for different boundary conditions are processed into fluctuation fields, $\tilde{\boldsymbol{\varepsilon}}(\mathbf{X}; \overset{(1)}{\boldsymbol{\varepsilon}}) = \overset{(0)}{\boldsymbol{\varepsilon}}(\mathbf{X}; \overset{(1)}{\boldsymbol{\varepsilon}}) - \overset{(1)}{\boldsymbol{\varepsilon}}$. The notion fluctuation is motivated by the zero mean property, $\langle \tilde{\boldsymbol{\varepsilon}} \rangle_{\Omega_0} = \mathbf{0}$. Furthermore, a classical, truncated analysis of correlations within this set of solution fluctuation fields is carried out. The result is a set of fields carrying most of this linear relation information,

$$\mathcal{B}_\varepsilon = \{\tilde{\boldsymbol{\varepsilon}}^{(1)}(\mathbf{X}), \dots, \tilde{\boldsymbol{\varepsilon}}^{(N_{\text{RB}})}(\mathbf{X})\}, \quad (4.1)$$

denoted *Reduced Basis for the strain ε* , or *ε -RB*. The *size* or *dimension* of this basis, N_{RB} , is by many orders of magnitude smaller than the number of Degrees of Freedom (DOF) of the FE model.

On a side note, it is interesting to note that this ε -RB is orthogonal with respect to the standard $L^2(\Omega_0)$ inner product. It might be interpreted as a *\mathbf{u} -RB*, orthogonal with respect to a Sobolev $H_0^1(\Omega_0)$ -like inner product. This theoretical consideration is not directly relevant for the sequel but helps justifying the approach and makes the relationships between ε -RB's and \mathbf{u} -RB's more obvious.

After this setup of the second stage, general solution strain fields are approximated via the RB ansatz

$$\overset{(0)}{\boldsymbol{\varepsilon}}(\mathbf{X}; \overset{(1)}{\boldsymbol{\varepsilon}}) \approx \overset{(0)}{\boldsymbol{\varepsilon}}^{\text{RB}}(\mathbf{X}; \overset{(1)}{\boldsymbol{\varepsilon}}) = \overset{(1)}{\boldsymbol{\varepsilon}} + \sum_{i=1}^{N_{\text{RB}}} \xi_i \tilde{\boldsymbol{\varepsilon}}^{(i)}(\mathbf{X}). \quad (4.2)$$

The vector of coefficients $\underline{\xi} = \underline{\xi}^{(1)}(\underline{\varepsilon}) \in \mathbb{R}^{N_{\text{RB}}}$ is determined by numerical approximation of the solution to the unconstrained minimization problem

$$\underline{\xi}^{(1)}(\underline{\varepsilon}) = \arg \min_{\underline{\xi} \in \mathbb{R}^{N_{\text{RB}}}} \left\langle \overset{(0)}{W}_{\varepsilon} \left(\overset{(0)}{\varepsilon}{}^{\text{RB}} \left(\overset{(0)}{\mathbf{X}}; \overset{(1)}{\varepsilon} \right) \right) \right\rangle_{\Omega_0}. \quad (4.3)$$

On one hand, this formulation is based on the fundamental variational principle of mechanics, cf. Hill [1956]. Related formulations have been used for decades in works on fundamental theorems of homogenization, e.g. [Ponte-Castañeda, 1991, (2.14)]. On the other hand, this method is notably different from established *efficient* homogenization methods based on a projection of the solution displacement field, e.g. Ryckelynck [2005], Yvonnet & He [2007]. In contrast to these methods, the presented RB scheme directly operates on the strains and completely avoids computations of strains from displacements, which leads to a significant decrease of the overall computational effort.

One can verify the consistency of the RB approximation process (4.3) with the solution of the original problem (3.11) in multiple ways. Here, it shall only be noted explicitly that the RB ansatz (4.2) unconditionally yields a compatible displacement field. This is due to the linearity of the small strain compatibility conditions (2.9).

After the energy minimization process (4.3), the effective stress is readily obtained via

$$\overset{(1)}{\sigma}(\underline{\varepsilon}) = \left\langle \sigma \left(\overset{(0)}{\varepsilon}{}^{\text{RB}} \left(\overset{(0)}{\mathbf{X}}; \overset{(1)}{\varepsilon} \right) \right) \right\rangle_{\Omega_0}. \quad (4.4)$$

Even more importantly, the effective stiffness is computable via

$$\overset{(1)}{\mathbb{C}}_{\varepsilon}(\underline{\varepsilon}) = \left\langle \overset{(0)}{\mathbb{C}}_{\varepsilon} \right\rangle_{\Omega_0} - \sum_{i,j=1}^{N_{\text{RB}}} D_{ij}^{-1} \left\langle \overset{(0)}{\mathbb{C}}_{\varepsilon} \left(\overset{(0)}{\varepsilon}{}^{\text{RB}} \right) [\tilde{\varepsilon}^{(i)}] \right\rangle_{\Omega_0} \otimes \left\langle \tilde{\varepsilon}^{(j)} : \overset{(0)}{\mathbb{C}}_{\varepsilon} \left(\overset{(0)}{\varepsilon}{}^{\text{RB}} \right) \right\rangle_{\Omega_0}, \quad (4.5)$$

$$D_{ij} = \left\langle \tilde{\varepsilon}^{(i)} : \overset{(0)}{\mathbb{C}}_{\varepsilon} \left(\overset{(0)}{\varepsilon}{}^{\text{RB}} \right) [\tilde{\varepsilon}^{(j)}] \right\rangle_{\Omega_0} = D_{ji}, \quad (4.6)$$

where the colon and the bracket operators are defined as the standard double contraction of a fourth-order tensor with a second-order tensor from the left and from the right, respectively.

The recipe (4.5) – although still computationally demanding – is significantly more efficient than the methods described in Chapter 3.3 and is one of the major contributions.

4.1.2 Sampling scheme

The sampling procedure by means of which the setup data of the first stage is gathered, i.e. the way in which the space of the macroscopic strain $\overset{(1)}{\varepsilon}$ is sampled, is a novelty in its own right. Previously available methods, as described in Sections 3.4.2 and 3.4.3, are

driven by variations of regular Cartesian discretization of the components of the strain tensor or by somewhat controlled randomness. In contrast to these, the method proposed here seeks to discretize the space of macroscopic strains in a way that is considerate of the possibly present anisotropy of the effective material law. It is desired that one can explicitly control the degree with which the phenomenon of effective anisotropy is resolved. Anisotropy denotes the dependence of the material response to the *kind* of the applied kinematic loading condition – as opposed to its *magnitude*. This distinction is respected by the novel sampling scheme, which is now briefly outlined.

By introduction of the multiplicative split of the macroscopic strain tensor $\overset{(1)}{\boldsymbol{\varepsilon}}$ into its *magnitude* ε and *direction* \mathbf{N} , i.e.

$$\overset{(1)}{\boldsymbol{\varepsilon}} = \varepsilon \mathbf{N}, \quad \varepsilon = \|\overset{(1)}{\boldsymbol{\varepsilon}}\|_{\text{Fro}}, \quad \mathbf{N} = \frac{\overset{(1)}{\boldsymbol{\varepsilon}}}{\varepsilon}, \quad (4.7)$$

the macroscopic strain space may be sampled concentrically:

$$\mathcal{S}_\varepsilon = \{\varepsilon_i \mathbf{N}^{(j)}\}_{i=1, j=1}^{i=N_{\text{mag}}, j=N_{\text{dir}}}. \quad (4.8)$$

Figure 4.1 illustrates the way in which the set of sampling points \mathcal{S}_ε contains sites at identical N_{mag} magnitudes (or radii) ε_i along each of N_{dir} directions $\mathbf{N}^{(j)}$. The particulate choices for the sampling magnitudes ε_i are made on an empirically derived rule. For the sampling directions $\mathbf{N}^{(j)}$, a discretely uniform distribution in the space of small strain tensors is promoted. This way, no assumption on the anisotropy of the generally unknown effective material law is made. The larger the number N_{dir} , the more information on anisotropy can be captured. This parameter is independent of the number N_{mag} which controls the resolution of the effective parameter behavior along each direction, i.e. the evolution of the response for a constant kind of loading.

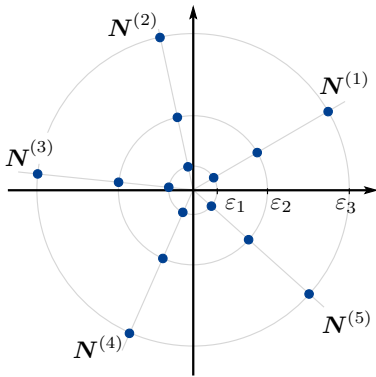


Figure 4.1: Schematic example of placement of samples in the space of infinitesimal strains.

This sampling method is utilized twice in the overall homogenization scheme. For the preparation of the first stage, the set of FE boundary conditions, i.e. $\overset{(1)}{\boldsymbol{\varepsilon}}$ in (3.11), is constructed in this manner. For the setup of the second stage, too, a comparable set of

boundary conditions is created, i.e. $\underline{\varepsilon}^{(1)}$ in (4.3). However, since the RB method is more efficient than the FEM, the number of boundary conditions $N_{\text{mag}}N_{\text{dir}}$ is chosen larger for the RB simulations.

4.1.3 Interpolation scheme

The *second reduction stage* extremely accelerates the evaluation of the effective material law. It does so by interpolating the values of the effective energy $\widetilde{W}_\varepsilon^{(1)}$ provided by the RB model, cf. left-hand side of (4.3). Thus, the interpolation method is constructed in a way specialized to this geometric feature of the sampling method (4.8). It is defined as

$$\widetilde{W}_\varepsilon^{(1)}(\underline{\varepsilon}^{(1)}) \approx \widetilde{W}(\varepsilon, \mathbf{N}) = \underline{p}(\varepsilon)^\top \underline{K}^{-1} \underline{k}(\mathbf{N}), \quad (4.9)$$

where $\underline{p}(\varepsilon), \underline{k}(\mathbf{N}) \in \mathbb{R}^{N_{\text{dir}}}$ and $\underline{K} \in \mathbb{R}^{N_{\text{dir}} \times N_{\text{dir}}}$. The vector $\underline{p}(\varepsilon)$ consists of N_{dir} one-dimensional interpolants of the effective energy as a function of ε . The components of the vector $\underline{k}(\mathbf{N})$ are spherical Gaussian kernel functions centered at the support directions $\mathbf{N}^{(j)}$, cf. (4.8), and evaluated at the query direction \mathbf{N} ,

$$k_j(\mathbf{N}) = \exp\left(-\gamma \text{acos}^2(\mathbf{N}^{(j)} \cdot \mathbf{N})\right) \quad (j = 1, \dots, N_{\text{dir}}, \gamma > 0). \quad (4.10)$$

These functions utilize an angular measure of distance and, hence, in combination with the corresponding kernel matrix \underline{K} , form a spherical interpolation scheme. The coefficients of the latter are defined via $K_{ij} = K_{ji} = k_j(\mathbf{N}^{(i)})$.

The effective stress $\underline{\sigma}^{(1)}$ and stiffness $\mathbb{C}_\varepsilon^{(1)}$ are computed by means of differentiation of \widetilde{W} . This is performed in an algorithmically efficient manner resulting in unconditional compatibility of \widetilde{W} , $\underline{\sigma}^{(1)}$ and $\mathbb{C}_\varepsilon^{(1)}$. Notably, the latter is also unconditionally symmetric.

Overall, this two-staged method falls into the class of data-based methods, cf. Chapter 3.4.3, and stands out from it for several reasons. The two most significant improvements over concurring methods are

- Speed-up of the setup phase: the sampling data set (4.8) is determined in a manner based upon physical reasoning (consideration of anisotropy) instead of indiscriminate brute force (e.g. Cartesian grid sampling) or chance (e.g. random sampling).
- Speed-up of the application phase: the interpolation method (4.9) is demonstrably applicable to realistic two-scale simulations in three dimensions on standard computers.

4.2 Second publication – Finite Strain Homogenization Using a Reduced Basis and Efficient Sampling

4.2.1 Suitability of nonlinear kinematic measures

Building on the successful demonstration of the ε -RB approach (4.2), and still settling with two scales, $M = 1$, the concept of an RB for kinematic non-displacement quantities is transferred to the case of large strain hyperelasticity. That is, the first stage of the two-stage procedure of Chapter 4.1 are now modified in order to be applicable to geometric nonlinearity. However, the lack of the small strain assumption (2.6) comes along with additional difficulties. Among these is the question of which kinematic non-displacement quantity should be approximated by means of an RB. For instance, the displacement gradient \mathbf{H} , the deformation gradient \mathbf{F} , the right Cauchy-Green tensor \mathbf{C} and any of the infinitely many measures of strain (2.3), (2.4) could be candidates for an approximation similar to (4.2).

There are several reasons for and against each of these possibilities. Among them, the dimension of the involved quantity directly affects the computational efforts necessary to determine a particular RB approximation. For example, symmetric tensors such as the right Cauchy-Green tensor are readily represented as six-dimensional vectors. In contrast to this, the deformation gradient is generally asymmetric and therefore only representable by nine-dimensional quantities. This is a significant factor as it means that 50% more memory is required. The number of algebraic operations involved in the evaluation of the homogenized material response also grows at least linearly with the dimension of the involved kinematic quantity. There is even some quadratic dependence on this number, cf. [Kunc & Fritzen, 2019a, Table 1].

Another factor is the effort necessary to ensure compatibility with the displacement field \mathbf{u} . While the compatibility condition for the deformation gradient (2.8) is linear, the one for the right Cauchy-Green tensor involves sophisticated differential geometrical considerations, cf. Acharya [1999] and the references given therein. In the context of this work, preliminary studies of an efficient approximation of the latter have not come to fruition. For this reason, an \mathbf{F} -RB approach is chosen here. This is not to say that other kinematic quantities may not possibly be successfully approximated by means of reduced bases and additional computational steps. However, this alternative line of thought is not pursued in this work.

4.2.2 Projection based Reduced Basis model

Hence, by choosing an \mathbf{F} -RB ansatz of the form

$$\mathbf{F}^{(0)}(\mathbf{X}; \mathbf{F}^{(0)}) \approx \mathbf{F}^{\text{RB}}(\mathbf{X}; \mathbf{F}^{(0)}) = \mathbf{F}^{(1)} + \sum_{i=1}^{N_{\text{RB}}} \xi_i \tilde{\mathbf{F}}^{(i)}(\mathbf{X}), \quad (4.11)$$

compatible fields of deformation gradients are produced without any assumptions on the vector of coefficients $\underline{\xi} \in \mathbb{R}^{N_{\text{RB}}}$. Similar to (4.1), the $L^2(\Omega_0)$ -orthogonal RB

$$\mathcal{B}_F = \{\tilde{\mathbf{F}}^{(1)}(\mathbf{X}), \dots, \tilde{\mathbf{F}}^{N_{\text{RB}}}(\mathbf{X})\} \quad (4.12)$$

is obtained by means of a standard, truncated analysis of correlations among a set of fluctuation snapshots. These have the form $\tilde{\mathbf{F}}(\mathbf{X}; \mathbf{F}) = \mathbf{F}^{(0)}(\mathbf{X}; \mathbf{F}) - \mathbf{F}^{(1)}$, where the fields $\mathbf{F}^{(0)}(\mathbf{X}; \mathbf{F})$ correspond to solutions of (3.3) obtained during the first stage.

In analogy to (4.3), the RB coefficients $\xi_i = \xi_i^{(1)}(\mathbf{F})$ are determined by means of numerical minimization of the approximation of the effective hyperelastic energy,

$$\xi_i^{(1)}(\mathbf{F}) = \arg \min_{\underline{\xi} \in \mathbb{R}^{N_{\text{RB}}}} \left\langle W_F(\mathbf{F}^{\text{RB}}(\mathbf{X}; \mathbf{F})) \right\rangle_{\Omega_0}^{(0)}. \quad (4.13)$$

Likewise, the expressions for $\mathbf{P}^{(1)}(\mathbf{F})$ and $\mathbf{C}_F^{(1)}(\mathbf{F})$ are obtained in direct correspondence to their small strain equivalent, cf. (4.4)-(4.6). As with the ε -RB, a major advantage of this \mathbf{F} -RB over the methods mentioned in Chapter 3.4.2 is the efficient representation of the homogenized stiffness.

4.2.3 Sampling scheme

The sampling procedure for the first stage, i.e. the choice of the boundary conditions of (3.3), comes along with increased complexity compared with the method developed in the geometrically linear case, cf. (4.8). The considerations there were based on the split of the loading condition $\tilde{\boldsymbol{\varepsilon}}^{(1)}$ into its magnitude ε and its direction \mathbf{N} . The correspondence between this loading direction and the *kind* of the loading $\tilde{\boldsymbol{\varepsilon}}^{(1)}$ was exploited. However, this does not immediately carry over to the case of the deformation gradient. For example, consider the case of pure planar shear which can be represented by

$$\mathbf{F}^{(1)}(\alpha) \sim \begin{bmatrix} \sqrt{1 + \alpha^2} & \alpha & 0 \\ \alpha & \sqrt{1 + \alpha^2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4.14)$$

where $\alpha \geq 0$ parametrizes the loading from the relaxed state ($\alpha = 0$) to the final condition. The same kind of loading is representable as a small strain tensor,

$$\begin{aligned} \stackrel{(1)}{\boldsymbol{\varepsilon}}(\alpha) &\sim \begin{bmatrix} \sqrt{1+\alpha^2}-1 & \alpha & 0 \\ \alpha & \sqrt{1+\alpha^2}-1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &\xrightarrow{\alpha \rightarrow 0} \begin{bmatrix} 0 & \alpha & 0 \\ \alpha & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \underbrace{\sqrt{2\alpha}}_{\varepsilon} \underbrace{\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\mathbf{N}}. \end{aligned} \quad (4.15)$$

It is the small strain linearization (2.6) that simplifies the isochoric condition $\det(\stackrel{(1)}{\mathbf{F}}(\alpha)) = \stackrel{(1)}{J}(\alpha) = 1$ to $\text{trace}(\stackrel{(1)}{\boldsymbol{\varepsilon}}(\alpha)) = 0$, respectively shaping (4.14) and the last expression in (4.15).

A representation of this example for large strains that is more similar to the small strain formulation (4.15) can be obtained via the Hencky strain

$$\stackrel{(1)}{\mathbf{E}}_{\text{H}}(\alpha) = \log\left(\stackrel{(1)}{\mathbf{F}}(\alpha)\right) \sim \underbrace{\sqrt{2} \sinh^{-1}(\alpha)}_t \underbrace{\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\mathbf{N}}, \quad (4.16)$$

where t and \mathbf{N} are defined in analogy to (4.7),

$$\stackrel{(1)}{\mathbf{E}}_{\text{H}} = t\mathbf{N}, \quad t = \|\stackrel{(1)}{\mathbf{E}}_{\text{H}}\|_{\text{FrO}}, \quad \mathbf{N} = \frac{\stackrel{(1)}{\mathbf{E}}_{\text{H}}}{t}. \quad (4.17)$$

The resemblance of the last terms of (4.16) and (4.15) is striking.

The zero-trace condition for isochoric deformations holds for the Hencky strain, too. However, in contrast to the infinitesimal strain tensor, not a direct linearization in the sense of a negligence of higher-order terms is the cause. Instead, the well-known general relationship $\det(\exp(t\mathbf{M})) = \exp(\text{trace}(t\mathbf{M}))$ ($\mathbf{M} \in \mathbb{R}^{n \times n}$, $n \in \mathbb{N}$, $t \in \mathbb{R}$), cf.

[Kühnel, 2011, Satz 6.2], implies $\text{trace}(\stackrel{(1)}{\mathbf{E}}_{\text{H}}) = 0$ as the isochoric condition of the Hencky strain space. The Hencky strain space can be considered a linearization in a differential geometric context: it is the tangent space to the space of isochoric stretch tensors, i.e. the set of all \mathbf{U} with $\det(\mathbf{U}) = 1$. Readers interested in this fundamental relationship are referred to standard textbooks on Lie groups, such as [Faraud, 2008, Kühnel, 2011] and to [Neff et al., 2016] for a mechanical viewpoint.

Another important consideration for the setup of the sampling procedure is made possible by the well-known principle of material frame-indifference: $W_{\text{F}}(\mathbf{Q}\mathbf{F}) =$

$W_F(\mathbf{F})$ for all \mathbf{Q} with $\mathbf{Q}\mathbf{Q}^\top = \mathbf{Q}^\top\mathbf{Q} = \mathbf{I}$, representing rigid body rotations of the deformed configuration. Substituting the inverse \mathbf{R}^\top of the rotational part of the polar decomposition $\mathbf{F} = \mathbf{R}\mathbf{U}$ proves that such rotations do not affect the energy. In fact, the rotation \mathbf{R} needs not to be considered when sampling the effective material response $W_F(\mathbf{F})$, $\mathbf{P}(\mathbf{F})$, $\mathbb{C}_F(\mathbf{F})$: the effect of the rotation \mathbf{R} on $W_F(\mathbf{U})$ is trivial and can be applied to $\mathbf{P}(\mathbf{U})$ and $\mathbb{C}_F(\mathbf{U})$ obtained in the canonical frame, i.e. one may assume $\mathbf{R} = \mathbf{I}$ during the sampling phase without loss of generality.

Therefore, the sampling of the boundary condition $\mathbf{F} = \mathbf{U}$ in the balance of linear momentum (3.3) may be conducted in the six-dimensional space of Hencky strains. The formulation

$$S_F = \left\{ \left(J^{(m)} \right)^{1/3} \exp \left(t^{(p)} \mathbf{N}^{(n)} \right) \right\}_{m,n,p=1}^{m=N_{\text{det}}, n=N_{\text{dir}}, p=N_{\text{mag}}} \quad (4.18)$$

is suggested. Here N_{det} is the number of sampled determinants, N_{dir} denotes the number of sampled directions of isochoric Hencky strains and N_{mag} states the number of sampled isochoric Hencky strain magnitudes. Moreover, the scalars $J^{(m)} \in [J_{\text{min}}^{(1)}, J_{\text{max}}^{(1)}]$ and $t^{(p)} \in [0, t_{\text{max}}]$ denote certain measures of the magnitudes of the dilatational and of the deviatoric strain, respectively. Their limits $J_{\text{min}}^{(1)}$, $J_{\text{max}}^{(1)}$ and t_{max} may be chosen reasonably and independently of each other. The tensors $\mathbf{N}^{(n)}$ are directions in the five-dimensional space of isochoric Hencky strains, i.e. they are symmetric and traceless. Just as for the small strain case, an approximately uniform distribution of these directions is advised in order to sample the possibly anisotropic behavior in an unbiased manner.

The Hencky strain \mathbf{E}_H is identified as the key ingredient to this sampling procedure: directions in the space of Hencky strains carry the same information as those in the space of small strains $\boldsymbol{\varepsilon}$. The close relationship between these two quantities is discussed in great detail in Neff et al. [2016]. *The proposal to conduct the sampling of the kinematic state space via the Hencky strain tensors is a major novelty.* A large part of the data-driven methods described in Chapter 3.4.3 could directly benefit from the utilization of Concentric Sampling in the Hencky strain space.

Summarizing the finite strain RB method, the most important advancements compared to the methods mentioned in Chapter 3.4.2 are the following:

- Speed-up of the setup phase: the selective sampling data set (4.18) is considerate of both the anisotropy and the determinant J . Therefore, it places the sampling sites within a physically meaningful range of the Hencky strain space. Simpler methods (e.g. Cartesian grid or random sampling of the components of \mathbf{U}) suffer much more from the six-dimensionality of the sampling space and fail either to contain the samples

within a reasonable range of the determinant (e.g. $J \in [1 - \epsilon, 1 + \epsilon]$ ⁽¹⁾), to explore this range thoroughly or to keep the number of samples within practical bounds.

- Speed-up of the application phase: the F -RB method yields good results for the homogenization of realistic, complex three-dimensional microstructures for arbitrary loading conditions, given that the latter lie within the sampling range of the setup phase. At the same time, computational complexity is significantly reduced. In many cases, the algorithmic complexity of the F -RB is advantageous to that of competing displacement-based methods, cf. Kunc & Fritzen [2019a].

In analogy to the case of infinitesimal strains, it is intended to establish an interpolating third stage, and to create its supporting points using the efficient F -RB model from the second stage. Before elaborating on the adaption of the interpolation method from Chapter 4.1 to the case of finite strains, more studies of the method in general are conducted.

4.3 Third publication – Generation of energy-minimizing point sets on spheres and their application in mesh-free interpolation and differentiation

4.3.1 General interpolation scheme

The performance of the interpolation ansatz (4.9) is influenced by the geometrical distribution of the supporting directions. Here, it is investigated which influence this distribution has on the interpolation quality. Moreover, special emphasis is put on the way such distributions can be found.

As the context is – for the moment – more general and purely mathematical, a few changes are made to the underlying conceptions. For one, the dimension of the interpolation domain is generalized to \mathbb{R}^{d+1} , $d \in \mathbb{N}$. Consequently, the interpolant is generalized to be a scalar valued approximant,

$$f(\mathbf{x}) \approx \tilde{f}(x, \mathbf{n}) = \underline{p}(x)^\top \underline{K}^{-1} \underline{k}(\mathbf{n}), \quad (4.19)$$

where $\mathbf{x} \in \mathbb{R}^{d+1}$, $x = \|\mathbf{x}\|$, $\mathbf{n} = \frac{\mathbf{x}}{x} \in \mathbb{S}^d \subset \mathbb{R}^{d+1}$ and \underline{p} , \underline{K} , \underline{k} are defined in analog to Chapter 4.1 with $\mathbf{n}^{(j)} \in \mathbb{S}^d$. Furthermore, instead of *directions* in \mathbb{R}^{d+1} only *points* on \mathbb{S}^d are spoken of. The number of support points (formerly the number of support directions N_{dir}) is renamed N in this publication, i.e. $j = 1, \dots, N$.

The term *Concentric Interpolation (CI)* denotes the ansatz (4.19) due to the straightforward interpretation that for each query point $\mathbf{x} \in \mathbb{R}^{d+1}$ the one-dimensional interpolants $p_j(x)$ take values on the surface of the origin-centered sphere of radius x , i.e. on $x\mathbb{S}^d$.

4.3.2 Point distributions and asymptotic uniformity

In order to formalize the concept of *distributions of points*, a measure theoretic approach is chosen. For the sake of this outline, two kinds of measures on \mathbb{S}^d are sufficient. First, the discrete measure corresponding to a point set $X_N = \{\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(N)}\} \subset \mathbb{S}^d$ is

$$\mu_{X_N}(A) = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{n}^{(i)}}(A) \quad (A \subset \mathbb{S}^d, \text{measurable}), \quad (4.20)$$

where the single point measure $\delta_{\mathbf{n}^{(i)}}$ is the characteristic function of the point $\mathbf{n}^{(i)}$,

$$\delta_{\mathbf{n}^{(i)}}(A) = \begin{cases} 1 & \text{if } \mathbf{n}^{(i)} \in A \\ 0 & \text{else} \end{cases}. \quad (4.21)$$

The main features of μ_{X_N} are the equal weight associated with each point and the normalization property, $\mu_{X_N}(\mathbb{S}^d) = 1$. Second, the standard uniform measure on the sphere is required in the sequel:

$$\sigma(A) = \frac{\text{surface area of } A}{\text{surface area of } \mathbb{S}^d} \quad (A \subset \mathbb{S}^d, \text{measurable}). \quad (4.22)$$

At this point, a *distribution* of a spherical point set X_N is characterized by means of the properties of its corresponding measure μ_{X_N} . A very special feature of certain kinds of point sets is asymptotically uniform distribution:

Definition 1 A sequence of point sets $(X_N)_{N \in \mathbb{N}}$ is said to be *asymptotically uniformly distributed* if the corresponding sequence of measures $(\mu_{X_N})_{N \in \mathbb{N}}$ converges to the standard uniform measure σ as $N \rightarrow \infty$. A suitable definition of convergence is chosen to this end. \square

Such convergence behavior is relevant to many kinds of applications. Roughly speaking, whenever it is desired to have point sets that are “evenly spread out” across the sphere \mathbb{S}^d , asymptotic uniformity is usually the first choice formal definition. A detailed list of research projects from various domains utilizing such point sets is given in the full paper.

4.3.3 Kernel functions and energies of point sets

One possible realization of such appealing convergence behavior is by means of *energy-minimizing* point sets. This concept is now briefly formalized.

It is standard to base the theory of energy-minimizing spherical points on the family of Riesz s -kernels,

$$k_s(\mathbf{x}, \mathbf{y}) = \frac{1}{\|\mathbf{x} - \mathbf{y}\|^s} \quad (\mathbf{x}, \mathbf{y} \in \mathbb{S}^d, s \in \mathbb{R}_{>0}). \quad (4.23)$$

This family of functions was originally motivated by physical considerations. For instance, the repulsive behavior of equally charged electric particles confined to a

conductive spherical surface ($d = 1, 2$) is modeled by the case $s = 2$. This is due to the resemblance of the 2-kernel with the electrostatic force according to Coulomb's law.

The associated potential-theoretic s -energy of a point set $X_N \subset \mathbb{S}^d$ is defined as

$$I_s(X_N) = \frac{1}{N(N-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^N k_s(\mathbf{x}_i, \mathbf{x}_j). \quad (4.24)$$

It is this energy of which a minimization is sought-after. Again, the minimization has a physical background as mentioned above. Mathematically speaking, the following well-known theorem states a characteristic of energy-minimizing point sets that is crucial to the present work.

Theorem 1 *Sequences of s -energy-minimizing points $(X_N)_{N \in \mathbb{N}}$ are asymptotically uniformly distributed, cf. Definition 1. \square*

4.3.4 Energy-minimizing point sets as interpolation support

In this sense, minimum energy points, i.e. energy-minimizing points, are “evenly distributed”. It is this descriptive property upon which the intuitive choice of energy-minimizing points as support for the interpolation (4.19) is based. Notably, there exist alternatives which base the choice of the supporting points on direct, formal association with the interpolant. For instance, in Chen et al. [2011] the distribution of the support points is optimized with respect to the condition number of the interpolation matrix. However, it is known that the resulting point sets of certain alternative methods are also asymptotically uniformly distributed, cf. Marzo & Ortega-Cerdà [2009].

Although the choice of minimum energy points as interpolation support is somewhat arbitrary, there are good reasons to take this path. For one, one may take advantage of the long-grown respective theory and of existing databases, e.g. Hardin et al. [1997], Womersley [2003]. Additionally, some of the most fundamental open problems in Mathematics are concerned with energy-minimizing points, see Smale [1998]. Possibly, there may be interdisciplinary synergies when investigating the suitability of minimum energy points for different kinds of applications.

It is known that notable differences of the resulting point distributions occur when the value of s is changed. In particular, larger values lead to more locally repulsive behavior, while smaller values result in a more global influence of each point.

Here, it is suspected that the interpolation (4.19) should generally perform more accurately for lower values of s . Thus, an exemplary numerical study is conducted with the aim of quantifying this effect for a constant target function $f(\mathbf{x}) = 1$.

Moreover, the effect is anticipated to be even more significant for the log-kernel, $k_{\log}(\mathbf{x}, \mathbf{y}) = \log(2/\|\mathbf{x} - \mathbf{y}\|)$. As a novelty, the LOG-kernel

$$k_{\text{LOG}}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \left(\log \frac{\|\mathbf{x} - \mathbf{y}\|}{2} - 1 \right) + 2 \quad (4.25)$$

is introduced. In a certain sense, the functions k_{\log} and k_{LOG} are the first and the second primitive of k_1 , respectively. Hence, with the new kernel function k_{LOG} , the interaction of the points within the set X_N is even more dominated by global effects than with the previously available kernels.

From a physical perspective, it is interesting to observe the behavior of the energy and its derivative associated with this kernel. On one hand, k_{LOG} still leads to infinitely strong repulsive forces for touching particles, in alignment with many real-world observations. On the other hand, the energy I_{LOG} is bounded even for coinciding particle locations, $\mathbf{x} = \mathbf{y}$. This fact very significantly simplifies the analysis of the energy of point sequences.

Numerical experiments show that support points generated for lower values of s indeed lead to more accurate CI approximations of the constant function. Also, the log-kernel and the LOG-kernel outperform the s -kernels in this sense. The novel LOG-kernel becomes superior to the log-kernel for larger dimensions, i.e. $d \geq 5$.

The main result of this work consists of the proof of Theorem 1 for the novel LOG-case. This is an essential contribution to the rich history of potential theory, where numerous proofs for the log-kernel and the Riesz s -kernels have been long-known. Furthermore, the following contributions are made

- to the field of energy-minimizing spherical point sets:
 - open source MATLAB code for the generation of energy-minimizing point sets, with emphasis on an easy-to-use GUI and automated point quality evaluation, cf. Figure 4.2
 - introduction of symmetrized variants of the kernel functions for highly efficient generation of symmetric, large, energy-minimizing point sets, cf. Figure 4.3
- to the field of interpolation on spheres \mathbb{S}^d and their surrounding spaces \mathbb{R}^{d+1} :
 - definition of the general Concentric Interpolation method
 - open source C++ code with emphasis on efficiency, e.g. allowing for tens of thousands of evaluations per second on a single CPU core, depending on the dimension d and the number of support points N
 - demonstration of highly accurate black-box applicability of the CI method for an arbitrary engineering function in \mathbb{R}^8
 - exemplary demonstration that, for increasingly high dimension d , support points generated by the LOG-kernel lead to slightly more accurate interpolation results than points associated with classical log- or s -kernels.

With this strong confidence in the CI method at hand, finally, the ambitious task of designing a second stage on top of the RB method of Chapter 4.2 is addressed.

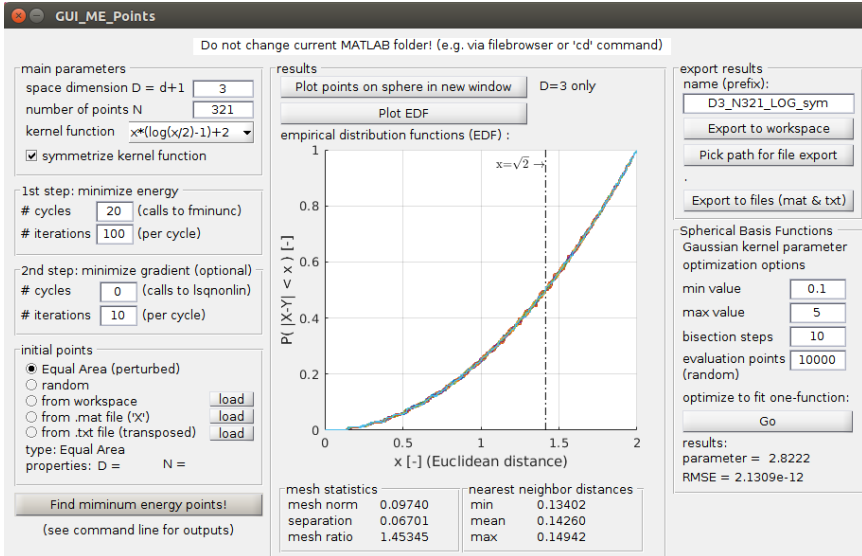


Figure 4.2: Graphical User Interface (GUI) of the provided open source MATLAB code. In this example, 321 points are arranged on the sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ such that they minimize the energy corresponding to the symmetrized LOG-kernel. The result is a set of $N=642$ symmetric points which is visualized in Figure 4.3.

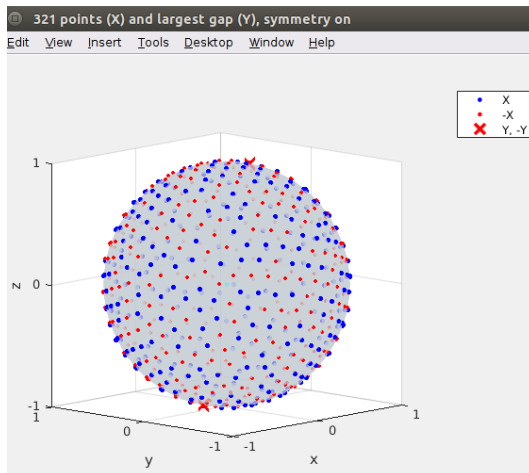


Figure 4.3: Plot of the 321 points (blue dots) and their antipodes (red dots) resulting from the program run depicted in Figure 4.2. The red crosses mark estimates of the centers of the largest point gaps.

4.4 Fourth publication – Many-scale finite strain computational homogenization via Concentric Interpolation

4.4.1 Concentric Interpolation of the material response

The two-staged scheme of Chapter 4.1 is now adapted to the case of large strains. To this end, the RB method of Chapter 4.2 is employed in the first stage in order to sample the material response. Subsequently, the CI method of Chapter 4.3 is utilized in the second stage for the efficient interpolation of these samples. Here, in contrast to the case of small strains, multiple quantities are interpolated. This is now addressed in brevity.

As argued in Chapter 4.2, the deformation gradient fluctuation snapshots required in the setup phase of the RB method are obtained via Concentric Sampling of the space of Hencky strain tensors. For the same reasons, the supporting points of the Concentric Interpolation method are placed in the Hencky strain space. However, the choice of which quantity to interpolate has significant impact on the overall many-scale procedure.

Since it is desired to conduct realistic FE simulations on the macroscopic scale, the CI implementation ideally returns quantities that do not require post-processing by the routines from which it is called. This is crucial for efficiency as the evaluation of the material response takes place at the innermost loop, i.e. at integration point level. Efficiency is thus increased if the CI scheme directly interpolates the stress and the stiffness in the formulation used by the FE method.

A common choice is to formulate the FE method with respect to the initial configuration Ω_0 . In this case, the second Piola-Kirchhoff stress tensor \underline{S} and its associated stiffness tensor \underline{C} are standard choices as they lead to efficient assembly of the residuum and the element stiffness. For this reason, the CI scheme is here set up to interpolate 27 scalar values simultaneously:

$$\begin{bmatrix} \underline{S} \\ \underline{C} \end{bmatrix} \approx \tilde{\underline{f}}(t, \underline{N}) = \underline{\underline{p}}(t)^\top \underline{\underline{K}}^{-1} \underline{\underline{k}}(\underline{N}) \quad (4.26)$$

Here, $\tilde{\underline{f}} \in \mathbb{R}^{27}$, $\underline{\underline{p}} \in \mathbb{R}^{N \times 27}$, $\underline{\underline{K}} \in \mathbb{R}^{N \times N}$ and $\underline{\underline{k}} \in \mathbb{R}^N$. The Hencky strain \underline{E}_H is split into t and \underline{N} in analog to (4.17). Therefore, each component of an appropriate vectorization $\underline{S} \in \mathbb{R}^6$ of \underline{S} and $\underline{C} \in \mathbb{R}^{21}$ of \underline{C} is interpolated separately on the space of Hencky strains.

Although the ansatz (4.26) does not guarantee consistency among the stress and the stiffness, it turns out to lead to highly accurate results in cases in which a reference solution is available. Moreover, the convergence of CI-based FEM on the next-larger scale is rather robust as long as the load is confined within certain (reasonable) bounds.

These bounds can be controlled by the user during the setup phase of the scheme, i.e. during the Concentric Sampling steps for the setup and the application of the first stage.

4.4.2 Many-scale homogenization

If the domain on this larger scale is an RVE in its own right, one could wonder whether the material behavior could be upscaled once more, i.e. to obtain the effective response on a presumed third scale. To this end, certain gentle assumptions on the physical model must be made, which do not represent any practical limitation of the present method. This topic has been thoroughly addressed in one of the first comprehensive summaries of homogenization theory in general, [Bensoussan et al., 1978, Remark 8.5], where it was coined “reiterated homogenization”.

The suggested homogenization scheme iterates through scales and is a numerical solution approximation to the many-scale scheme introduced in Chapter 3.1. It is visually described in Figure 4.4.

It is demonstrated that *bridging of four scales on a standard workstation* is viable from a computational point of view. Scale-softening for porous micro-structures is observable, however high-fidelity reference solutions were not findable to the author for such sophisticated homogenization examples. Physically speaking, four is probably close to the upper limit of the number of scales that can be realistically homogenized under the assumption of hyperelasticity although no such claim is made at this point. Essentially, the entire range of hyperelastic scales is treatable by means of this truly many-scale homogenization method.

A multitude of parameters is involved in the formal statement of the proposed many-scale homogenization scheme. This fact calls for a thorough investigation of the interdependencies and of how each choice affects the robustness of the overall method. The present work settles with a *successful proof of principle that evaluations of a fourth-scale homogenized material law can be quickly evaluated on laptop computers.*

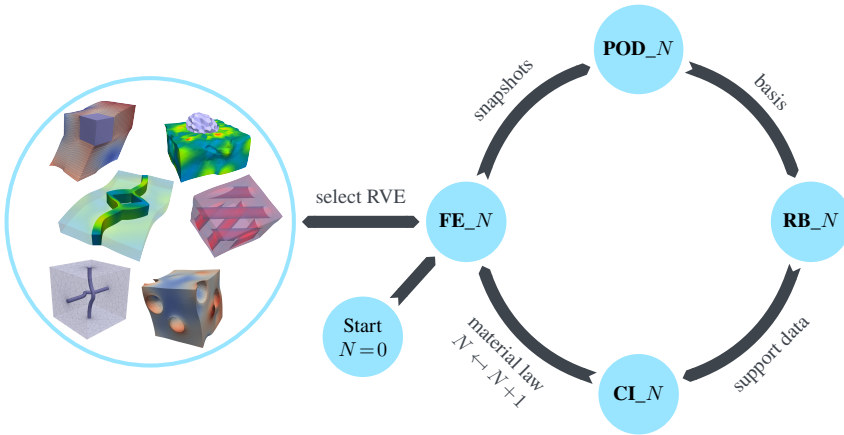


Figure 4.4: The proposed many-scale homogenization scheme begins with the Finite Element Method on the smallest modeled scale, $N = 0$. Deformation gradient fluctuation snapshots are generated for the subsequent, classical Proper Orthogonal Decomposition. This yields the Reduced Basis, which is utilized for efficient generation of the supporting points for the Concentric Interpolation method. Its computational efficiency allows for FE simulations on the next larger scale, $N = 1$, with this CI scheme substituting for the homogenized material response emerging from the previous scale. Subsequently, many upscaling processes are conducted until the FEM is set up on the engineering scale, $N = M$.

4.5 List of publications

This cumulative dissertation contains the following published articles.

- Fritzen, F. & Kunc, O.: Two-stage data-driven homogenization for nonlinear solids using a reduced order model. *European Journal of Mechanics - A/Solids* **69** (2018c), 201–220, Elsevier BV, doi:10.1016/j.euromechsol.2017.11.007
- Kunc, O. & Fritzen, F.: Finite strain homogenization using a reduced basis and efficient sampling. *Mathematical and Computational Applications* **24** (2019b), MDPI, doi:10.3390/mca24020056
- Kunc, O. & Fritzen, F.: Generation of energy-minimizing point sets on spheres and their application in mesh-free interpolation and differentiation. *Advances in Computational Mathematics* **45** (2019c), 3021–3056, Springer Science and Business Media LLC, doi:10.1007/s10444-019-09726-5
- Kunc, O. & Fritzen, F.: Many-scale finite strain computational homogenization via Concentric Interpolation. *International Journal for Numerical Methods in Engineering* (2020), Wiley, doi:10.1002/nme.6454

Chapter 5:

First publication – Two-stage data-driven homogenization of nonlinear solids using a Reduced Basis model

Original publication:

Fritzen, F. & Kunc, O.: Two-stage data-driven homogenization for nonlinear solids using a reduced order model. *European Journal of Mechanics - A/Solids* **69** (2018c), 201–220, Elsevier BV, doi:10.1016/j.euromechsol.2017.11.007

Abstract. The nonlinear behavior of materials with three-dimensional microstructure is investigated using a data-driven approach. The key innovation is the combination of two hierarchies of precomputations with sensibly chosen sampling sites and adapted interpolation functions: First, finite element (FE) simulations are performed on the microstructural level. A sophisticated sampling strategy is developed in order to keep the number of costly FE computations low. Second, the generated simulation data serves as input for a reduced order model (ROM). The ROM allows for considerable speed-ups on the order of 10-100. Still, its performance is below the demands for actual twoscale simulations. In order to attain the needed speed-ups, in a third step, the use of radial numerically explicit potentials (RNEXP) is proposed. The latter combine uni-directional cubic interpolation functions with radial basis functions operating on geodesic distances. The evaluation of the RNEXP approximation is realized almost in real-time. It benefits from the computational efficiency of the ROM since a higher number of sampling points can be realized than if direct FE simulations were used. By virtue of the dedicated sampling strategy less samples and, thus, precomputations (both FE and ROM) are needed than in competing techniques from literature. These measures render the offline cost of the RNEXP manageable on workstation computers. Additionally, the chosen sampling directions show favorable for the employed kernel interpolation. Numerical examples for highly nonlinear hyperelastic (pseudo-plastic) composite materials with isotropic and anisotropic microstructure are investigated. Twoscale simulations involving more than 10^6 DOF on the structural level are solved using the RNEXP and the influence of the microstructure on the structural behavior is quantified.

Keywords: reduced order model; data-driven computational homogenization; RNEXP

5.1 Introduction

The aim of homogenization methods is the prediction of the effective properties of microstructured solids. Examples for such materials are particle or fiber reinforced thermoplastics, porous metals, foams and polycrystalline solids. Many promising methods have been suggested which extend the early first order estimates proposed by Voigt [1910] and Reuß [1929]. For linear materials estimates and bounds of various type have been proposed, e.g. the well-known upper and lower Hashin-Shtrikman bounds [Hashin & Shtrikman, 1963]. Various monographs on the topic have been published (e.g. Nemat-Nasser & Hori [1999], Torquato [2006]). For nonlinear materials the estimation is more difficult, especially when path-dependent problems are considered (see, e.g., Klusemann et al. [2012], Matouš et al. [2017] for review articles). The simple first order Taylor and Sachs estimates lead to overly stiff or soft predictions. Variational estimates can achieve better accuracy (e.g. Ponte-Castañeda [1992, 2002], Agoras & Castañeda [2011]), but are more involved. For specific applications the general homogenization problem is often specialized, e.g. in the Gurson model [Gurson, 1977] for porous metals and the many extensions thereof (e.g. Bilger et al. [2007], Shen et al. [2012]).

While analytical methods are often useful in the presence of isotropic microstructures or simple topologies and morphologies (e.g. spheroids with specific orientation distribution), computational methods can help to predict the material behavior in the presence of more complicated microstructures. Simulations can provide reliable information on the hardening and damaging behavior and they can predict viscous effects. For solid mechanical problems the finite element method (FEM) can be considered as the method of choice for general problems although schemes based on the Fast Fourier Transform (FFT) are seeing increasing attention for homogenization problems, e.g. [Mishra et al., 2016].

A short-coming of the FEM (and more general of simulation methods) is their high computational cost: each simulation requires billions of algebraic operations and produces a lot of information that often needs to be stored (e.g. history variables). For instance the numerical studies presented for porous materials in [Fritzen et al., 2012] dealt with rather simple materials but required massive computational resources. In order to alleviate the computational burden, reduced order models have become rather fashionable, especially in the context of nonlinear homogenization. A major development in this direction was the Nonuniform Transformation Field Analysis (NTFA, [Michel & Suquet, 2003, 2004]). Later the NTFA has been extended in terms of the potential-based reduced basis model order reduction (pRB MOR) of the authors [Fritzen & Leuschner, 2013]. The pRB MOR is based on a mixed incremental variational principle that provides the evolution law for the reduced degrees of freedom. In combination with GPU acceleration [Fritzen et al., 2014, Fritzen & Hodapp, 2016] this variational method has proven to provide reliable predictions at computational complexity allowing for three-dimensional twoscale simulations with consideration of the microstructure of the material and of the nonlinearities therein. However, the computational demands remain

significant. A recent proposal for tackling some disadvantages of the NTFA is the self-consistent clustering analysis (SCCA, [Liu et al., 2016]). The SCCA consists of two steps: First, a partitioning of the spatial domain according to a cluster analysis of the linear elastic strain localization operator is performed. Second, a TFA-like (e.g. [Dvorak & Benveniste, 1992]) procedure is applied to the subdomains. Other state of the art techniques from the model reduction community include the (a priori) hyper-reduction [Ryckelynck, 2005] or the proper generalized decomposition (PGD, e.g. [Néron & Ladevèze, 2010, Relun et al., 2013]). Despite the recent progress, such reduced order models are computationally still too involved. Notably, the NTFA, SCCA and pRB MOR all incorporate micromechanical considerations into the reduced model which is a key difference to general (Galerkin) projection based reduced order models such as the R3M proposed by Yvonnet & He [2007]. The reduced basis is obtained via snapshot proper orthogonal decomposition (see also [Radermacher & Reese, 2015, Soldner et al., 2017]) for applications). The derived reduced problems are low-dimensional and nonlinear.

A different branch has developed in terms of data-driven techniques (or generalized look-up-tables) during the past decade. Temizer & Zohdi [2007] have developed material maps. They calibrated a linear interpolation of the parameters of an asserted effective constitutive model by means of numerical computations (see also [Temizer & Wriggers, 2007]). Another approach has been pursued by Yvonnet et al. [2009] in terms of numerically explicit potentials (NEXP; see also [Clément et al., 2012, Yvonnet et al., 2013]). They approximated the effective hyperelastic material model of a microstructured solid based on a substantial amount of sampling points in strain space. Each sample consists of the effective strain energy of the microstructure subjected to the given loading (defined by the position of the sample). The NEXP are differentiable and, hence, the symmetry of the tangent stiffness operator is assured. Unfortunately, the number of required offline computations is overly large for general three-dimensional problems. An interesting application of the NEXP in multiscale optimization is due to Xia & Breitkopf [2015].

Le et al. [2015] used neural networks in order to determine black-box function approximations in the spirit of the NEXP for multi-parametric problems. The resolution of the sampling (in strain space) is rather limited, especially at very small strain amplitudes. This could lead to some deficiencies in the prediction of the homogenized material response. Further, the number of precomputations (hundreds of thousands are reported in [Le et al., 2015]) exceed nowadays computing capabilities by far.

Addressing the topic from a different perspective, Bhattacharjee & Matouš [2016] proposed a homogenization method for (moderately) nonlinear hyperelastic solids at finite strains in terms of a manifold-based reduced order model (MNR OM). A machine learning approach is used that exploits large data sets obtained from several thousand FE simulations. The computational cost of the offline phase of this approach remains pronounced due to the large number of required high-fidelity solutions (e.g., $\approx 16,000$ FE solutions in [Bhattacharjee & Matouš, 2016]). It must be emphasized that the MNR OM works purely data-driven, i.e. independent of (reduced) equilibrium conditions, which is

a notable difference to reduced basis methods for hyperelastic solids (e.g. [Yvonnet & He, 2007, Soldner et al., 2017]).

In the current investigation we focus on hyperelasticity (here confined to the small strain regime). In this context we propose a highly structured overall strategy which combines different computational techniques with a particular layout of the sampling positions. The latter plays a crucial role and contributes to both accuracy and performance of the resulting scheme. As a starting point the hyperelastic problem is rephrased as a minimization problem (Section 5.2.3). Next, a snapshot proper orthogonal decomposition (similar to, e.g., [Yvonnet & He, 2007]) is employed to construct a Galerkin reduced order model (ROM) in Section 5.3. The ROM is used to approximate the high-fidelity model at significantly reduced computational cost, thereby allowing for thousands of samples for the subsequent data-driven approach. The general layout of the homogenization procedure is developed in Section 5.4. Following this layout the radial numerically explicit potentials (coined RNEXP analogous to the NEXP of [Yvonnet et al., 2009]) are constructed through the two stages developed in Section 5.5. Detailed algorithms of POD, ROM and RNEXP are provided in the Appendix. A selection of numerical examples is provided in Section 5.6. These include a comparison of RNEXP twoscale simulations to FE^{2R}-type computations (FE Square Reduced, cf. [Fritzen & Hodapp, 2016]) that are computationally significantly more expensive. The capability to treat realistic three-dimensional geometries with underlying microstructure is exemplified via realistic twoscale simulations involving more than 10^6 structural degrees of freedom in manageable computing time.

5.1.1 Notation

In the following an index free notation of tensors is used. Bold face letters denote first order (e.g. \mathbf{x} , \mathbf{u} , \mathbf{n}), second order (e.g. $\boldsymbol{\sigma}$, $\boldsymbol{\varepsilon}$) and fourth order (e.g. \mathbb{C}) tensor, respectively. Differentiation of a field \bullet with respect to \mathbf{y} is written as $\partial_{\mathbf{y}}\bullet$. Alternatively, the differentiation of functions of multiple arguments \mathbf{y}_1, \dots with respect to \mathbf{y}_i is denoted by $\partial_{\mathbf{y}_i}\bullet$. In the case of the spatial variables \mathbf{x} and $\bar{\mathbf{x}}$ the common gradient operators $\nabla_{\mathbf{x}}$ and $\nabla_{\bar{\mathbf{x}}}$ are employed. For the sake of simplicity we confine attention to an orthonormal basis and introduce vector representations of symmetric second and fourth order tensors. Vectors and matrices are denoted by underlined quantities and, if not mentioned otherwise, canonical identities such as $\boldsymbol{\sigma} \leftrightarrow \underline{\boldsymbol{\sigma}}$ hold. The selection of an orthonormal basis is crucial (see also [Yvonnet & He, 2007]), e.g. in terms of

$$\mathcal{B}_1 = \mathbf{e}_x \otimes \mathbf{e}_x, \quad \mathcal{B}_2 = \mathbf{e}_y \otimes \mathbf{e}_y, \quad \mathcal{B}_3 = \mathbf{e}_z \otimes \mathbf{e}_z, \quad (5.1)$$

$$\mathcal{B}_4 = \sqrt{2} \text{sym}(\mathbf{e}_x \otimes \mathbf{e}_y), \quad \mathcal{B}_5 = \sqrt{2} \text{sym}(\mathbf{e}_x \otimes \mathbf{e}_z), \quad \mathcal{B}_6 = \sqrt{2} \text{sym}(\mathbf{e}_y \otimes \mathbf{e}_z). \quad (5.2)$$

For the computation of effective properties the averaging operator $\langle \bullet \rangle$ is defined as

$$\langle \bullet \rangle = \frac{1}{|\Omega|} \int_{\Omega} \bullet \, dV. \quad (5.3)$$

5.2 Framework

5.2.1 Twoscale formulation

In the following we assume a standard twoscale framework at small deformations as considered in many works related to linear and nonlinear homogenization (see, e.g., [Feyel, 1999, Fritzen et al., 2012, Fritzen & Leuschner, 2013]). A crucial assumption in this context is the asserted separation of length scales: the gradients of the fluctuations of the macroscopic (or effective) quantities must be *small* relative to the physical length of the reference volume element (RVE). The latter is assumed to exist, to be well-selected (cf., e.g., [Galli et al., 2012]) and, further, the material properties at the small length scale are considered to be available.

On the macroscopic (or structural) domain $\bar{\Omega}$ we consider the displacement field $\bar{\mathbf{u}}$, the infinitesimal strain tensor $\bar{\boldsymbol{\varepsilon}}$ and the stress tensor $\bar{\boldsymbol{\sigma}}$. On the microscopic domain, i.e. inside of the RVE Ω , the respective quantities are \mathbf{u} , $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$. On either scale the strain tensor is the symmetric gradient of the displacements

$$\bar{\boldsymbol{\varepsilon}} = \frac{1}{2} \left(\nabla_{\bar{\mathbf{x}}} \bar{\mathbf{u}} + (\nabla_{\bar{\mathbf{x}}} \bar{\mathbf{u}})^{\top} \right), \quad \boldsymbol{\varepsilon} = \frac{1}{2} \left(\nabla_{\mathbf{x}} \mathbf{u} + (\nabla_{\mathbf{x}} \mathbf{u})^{\top} \right), \quad (5.4)$$

where $\nabla_{\bar{\mathbf{x}}}$ denotes the gradient with respect to the macroscopic coordinate $\bar{\mathbf{x}}$ and $\nabla_{\mathbf{x}}$ with respect to the microscopic coordinate \mathbf{x} . Balance of linear and angular momentum on both scales needs to be satisfied in the sense that

$$\nabla_{\bar{\mathbf{x}}} \cdot \bar{\boldsymbol{\sigma}} + \bar{\mathbf{b}} = \mathbf{0} \quad \bar{\boldsymbol{\sigma}} = \bar{\boldsymbol{\sigma}}^{\top} \quad + \overline{\text{BC}}, \quad (5.5)$$

$$\nabla_{\mathbf{x}} \cdot \boldsymbol{\sigma} = \mathbf{0} \quad \boldsymbol{\sigma} = \boldsymbol{\sigma}^{\top} \quad + \text{BC}, \quad (5.6)$$

where $\overline{\text{BC}}$ denotes the usual (Dirichlet and Neumann type) boundary conditions on the structural level and $\bar{\mathbf{b}}$ are volumetric forces. On the microscopic scale the boundary conditions BC are obtained from the scale coupling relations

$$\begin{aligned} \bar{\boldsymbol{\sigma}} &= \frac{1}{|\Omega|} \int_{\partial\Omega} \mathbf{t} \otimes \mathbf{x} \, dA = \frac{1}{|\Omega|} \int_{\Omega} \boldsymbol{\sigma} \, dV, \\ \bar{\boldsymbol{\varepsilon}} &= \frac{1}{|\Omega|} \int_{\partial\Omega} \text{sym}(\mathbf{u} \otimes \mathbf{n}) \, dA = \frac{1}{|\Omega|} \int_{\Omega} \boldsymbol{\varepsilon} \, dV, \end{aligned} \quad (5.7)$$

with the surface traction \mathbf{t} and the unit surface normal \mathbf{n} . Pores call for special attention, for instance in terms of the consideration of cavity stresses (pore pressure, whose

diffusion can actually be interpreted as a viscous contribution on the macro-scale, e.g. as in [Jänicke et al., 2015]) and cavity strains (see, e.g., [Fritzen et al., 2012]). The Hill-Mandel condition restricts the microscopic fields to satisfy an integral identity (e.g. [Bishop & Hill, 1951])

$$\frac{1}{|\Omega|} \int_{\Omega} \boldsymbol{\sigma} \cdot \dot{\boldsymbol{\varepsilon}} \, dV = \frac{1}{|\Omega|} \int_{\partial\Omega} (\boldsymbol{\sigma} \mathbf{n}) \cdot \dot{\mathbf{u}} \, dA. \quad (5.8)$$

The condition (5.8) is satisfied by linear displacement boundary conditions or by periodic displacement fluctuations. Due to reasons discussed in detail by, e.g., Miehe [2002] we focus on periodic fluctuation fields which have favorable properties. We refer to the large body of literature on the topic with respect to other types of boundary conditions and their algorithmic realization (see also, e.g., [Ostoja-Starzewski, 2006, Fritzen & Böhlke, 2010]).

5.2.2 Hyperelastic material models

In the following attention is confined to hyperelastic materials with free energy density $\psi(\boldsymbol{\varepsilon})$ on the microscopic scale. For these materials the stress tensor $\boldsymbol{\sigma}$ and the stiffness \mathbb{C} are given by

$$\boldsymbol{\sigma} = \frac{\partial \psi(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}}, \quad \mathbb{C} = \frac{\partial^2 \psi(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon} \partial \boldsymbol{\varepsilon}}. \quad (5.9)$$

Of course the consideration of dissipative materials is a key objective which cannot be achieved by hyperelasticity. Still, hyperelastic materials allow for many meaningful investigations regarding for example porous materials [Bilger et al., 2005] or the investigation of materials undergoing (mostly) proportional loading: hyperelastic models can mimic the stress-strain response observed in elasto-plastic materials.

In this work two different hyperelastic models mimicking plasticity are studied: first, a deformation type plasticity model without and with linear hardening is considered. It is closely related to the one used by Bilger et al. [2005] and others. Its free energy is given by ($\boldsymbol{\varepsilon}' = \boldsymbol{\varepsilon} - \text{tr}(\boldsymbol{\varepsilon})/3 \mathbf{I}$)

$$\begin{aligned} \psi'_e(\boldsymbol{\varepsilon}) &= G \|\boldsymbol{\varepsilon}'\|^2, \\ \psi'_{1,p}(\boldsymbol{\varepsilon}) &= \sqrt{\frac{2}{3}} \sigma_c \left(\|\boldsymbol{\varepsilon}'\| - \frac{\varepsilon_c}{2} \right) + \frac{1}{2} \frac{2Gh}{3G+h} \left(\|\boldsymbol{\varepsilon}'\| - \varepsilon_c \right)^2, \\ \psi_1(\boldsymbol{\varepsilon}) &= \frac{K}{2} (\mathbf{I} \cdot \boldsymbol{\varepsilon})^2 + \begin{cases} \psi'_e(\boldsymbol{\varepsilon}) & \|\boldsymbol{\varepsilon}'\| \leq \varepsilon_c, \\ \psi'_{1,p}(\boldsymbol{\varepsilon}) & \|\boldsymbol{\varepsilon}'\| > \varepsilon_c, \end{cases} \text{ with } \varepsilon_c = \sqrt{\frac{2}{3}} \frac{\sigma_c}{2G}, \end{aligned} \quad (5.11)$$

where K, G, h denote bulk, shear and hardening modulus, respectively. More details can be found in Appendix 5.8.1. The linear elastic material parameters K and G correspond to classic isotropic linear elasticity and σ_c, h correspond to the yield stress and hardening modulus in von Mises elasto-plasticity, respectively. The free energy of

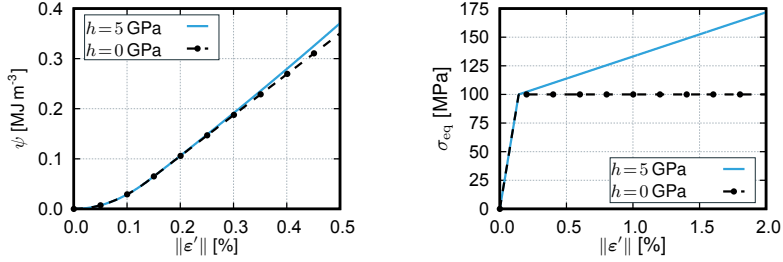


Figure 5.1: Deformation plasticity model as, e.g., used in the limit analysis of porous materials (e.g. [Bilger et al., 2005]) without hardening (**black, dashed**) and with linear hardening (**blue, solid line**) (left: free energy; right: corresponding (deviatoric) stress-strain curve)

the model with and without hardening and the relation between the deviatoric stress and the applied deviatoric strain are illustrated in Figure 5.1. Note that unlike elasto-plastic models, the presented model has no dissipation and the stress-strain relation becomes unphysical upon unloading. Still, the constitutive response can mimic elasto-plasticity for proportional strain paths.

Further, a nonlinear material similar to the one employed in [Yvonnet et al., 2009] is investigated:

$$\epsilon_c = \epsilon_0 \left(\frac{\sigma_0}{3G\epsilon_0} \right)^{\frac{1}{1-p}}, \quad \epsilon_{\text{eq}} = \sqrt{\frac{2}{3}} \|\epsilon'\|, \quad (5.12)$$

$$\psi'_{2,p}(\epsilon) = \frac{\sigma_0 \epsilon_0}{p+1} \left[\left(\frac{\epsilon_{\text{eq}}}{\epsilon_0} \right)^{1+p} - \left(\frac{\epsilon_c}{\epsilon_0} \right)^{1+p} \right] + \frac{3G}{2} \epsilon_c^2, \quad (5.13)$$

$$\psi_2(\epsilon) = \frac{K}{2} (\mathbf{I} \cdot \epsilon)^2 + \begin{cases} \psi'_e(\epsilon) & \epsilon_{\text{eq}} \leq \epsilon_c, \\ \psi'_{2,p}(\epsilon) & \epsilon_{\text{eq}} > \epsilon_c. \end{cases} \quad (5.14)$$

In addition to the elastic constants K, G , this model has the hardening exponent $0 < p < 1$, the reference stress σ_0 and the reference strain ϵ_0 as additional parameters (see 5.8.1 for details). In contrast to the model in [Yvonnet et al., 2009] our modified free energy ψ_2 shows linear elastic behavior with prescribed shear modulus G until the (pseudo-)yield stress is reached. Thereby the overly stiff behavior shown by the model proposed and used in [Yvonnet et al., 2009] (i.e. stiffer than the linear elastic material) at low strains due to the singularity of the gradient of the free energy is prevented. For higher strains the nonlinear stress-strain behavior of the original model is preserved. Graphically, the difference between the original and modified free energy and the related stress-strain relations are shown in Figure 5.2.

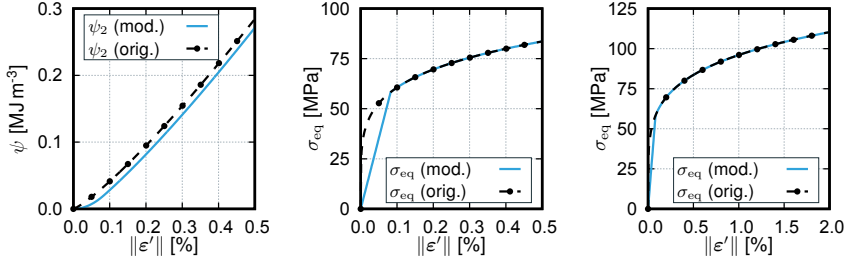


Figure 5.2: Nonlinear material model (**black, dashed**; see, e.g., [Yvonnet et al., 2009]) and a modified version thereof (**blue, solid line**) with linear elastic regime at small strains (*left*: free energy; *middle, right*: deviatoric stress-strain curve for different ranges of $\|\varepsilon'\|$)

Remark 1 Differentiability of the free energy ψ The potentials that we propose are continuous up to the first derivative with respect to ε' and smooth with respect to the volumetric strain. The second derivative (i.e. the stiffness tensor) is, however, discontinuous for $\|\varepsilon'\| = \varepsilon_c$. While the continuity of the first gradient is important since it guarantees continuous stresses, the continuity of the second derivative is not mandatory. Further, we would like to emphasize that both models possess a strictly convex free energy which is desirable from the viewpoint of solution theory (i.e. existence and uniqueness are a priori guaranteed) if $h > 0$ is considered for ψ_1 . The original model from [Yvonnet et al., 2009] is not differentiable at $\varepsilon = \mathbf{0}$ which is delicate from a computational perspective.

It shall be noted that, assuming a Galerkin formulation, the weak form of balance of linear momentum of the hyperelastic RVE problem can be expressed by the minimum problem (similar to [Moulinec & Suquet, 2003])

$$W(\bar{\varepsilon}) = \min_{\varepsilon \in \mathcal{K}^\varepsilon} \bar{\Pi}(\varepsilon), \quad \text{with} \quad \bar{\Pi}(\varepsilon) = \frac{1}{|\Omega|} \int_{\Omega} \psi(\varepsilon) \, dV. \quad (5.15)$$

Here \mathcal{K}^ε denotes the strain fields derived from the set of all admissible displacement fields \mathcal{K}^u that are in accordance with the prescribed deformation $\bar{\varepsilon}$ and with the boundary conditions. In the following \mathcal{K}^ε and \mathcal{K}^u are used alternately since they are isomorphic $\mathcal{K}^\varepsilon \sim \mathcal{K}^u$. Analogously we identify $\bar{\Pi}(u) \equiv \bar{\Pi}(\varepsilon)$, where $\varepsilon = \varepsilon(u)$. The sets $\mathcal{K}^u, \mathcal{K}^\varepsilon$ can be written in terms of a displacement fluctuation field w and the macroscopic strain $\bar{\varepsilon}$ via

$$\begin{aligned} \mathcal{K}^u &= \{u = u_0 + w \text{ with } u_0 = \bar{\varepsilon}x, w \in \mathcal{H}_\#^1\}, \\ \mathcal{K}^\varepsilon &= \{\varepsilon = \bar{\varepsilon} + \bar{\varepsilon}(w) \text{ with } w \in \mathcal{H}_\#^1\}. \end{aligned} \quad (5.16)$$

Here $\mathcal{H}_\#^1$ is the space of all periodic fields in the usual Sobolev space of continuous functions with square integrable gradient. Since the macroscopic strain $\bar{\varepsilon}$ is prescribed in

strain-driven homogenization, the minimum problem (5.15) can be written as: Find the displacement fluctuations $\tilde{\mathbf{u}} \in \mathcal{H}_{\#}^1$ such that

$$\delta \bar{\Pi}(\bar{\boldsymbol{\varepsilon}} + \tilde{\boldsymbol{\varepsilon}}(\tilde{\mathbf{u}})) = \frac{1}{|\Omega|} \int_{\Omega} \tilde{\boldsymbol{\varepsilon}}(\mathbf{w}) \cdot \boldsymbol{\sigma}(\bar{\boldsymbol{\varepsilon}} + \tilde{\boldsymbol{\varepsilon}}(\tilde{\mathbf{u}})) \, dV \stackrel{!}{=} 0 \quad \forall \mathbf{w} \in \mathcal{H}_{\#}^1. \quad (5.17)$$

5.2.3 Effective hyperelastic material model

The effective energy $W(\bar{\boldsymbol{\varepsilon}})$ defines the effective stress and stiffness via

$$\bar{\boldsymbol{\sigma}} = \frac{\partial W(\bar{\boldsymbol{\varepsilon}})}{\partial \bar{\boldsymbol{\varepsilon}}} = \left\langle \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}} \right\rangle, \quad \bar{\mathbb{C}} = \frac{\partial^2 W(\bar{\boldsymbol{\varepsilon}})}{\partial \bar{\boldsymbol{\varepsilon}} \partial \bar{\boldsymbol{\varepsilon}}} = \langle \mathbb{C}(\boldsymbol{\varepsilon}) \rangle + \left\langle \mathbb{C}(\boldsymbol{\varepsilon}) \cdot \frac{\partial \tilde{\boldsymbol{\varepsilon}}}{\partial \bar{\boldsymbol{\varepsilon}}} \right\rangle. \quad (5.18)$$

The tensor $\bar{\mathbb{C}}$ has the two minor and the major symmetry and is positive definite for convex microscopic strain energy functions. By inspection of (5.18) the solution of the microscale variational problem (5.17) can be omitted, if the function W and its gradients are known. Hence, the hyperelastic homogenization problem is reformulated as:

Find $W \in C^{2,1}(\text{Sym}(\mathbb{R}^3), \mathbb{R})$ such that $W(\bar{\boldsymbol{\varepsilon}}) = \inf_{\tilde{\mathbf{u}} \in \mathcal{H}_{\#}^1} \bar{\Pi}(\bar{\boldsymbol{\varepsilon}} + \tilde{\boldsymbol{\varepsilon}}(\tilde{\mathbf{u}})).$ **[H]**

Here $C^{2,1}$ denotes all piecewise twice differentiable functions with continuous first derivative. The structure of **[H]** immediately confirms that the actual field \mathbf{w} realizing the infimum of $\bar{\Pi}$ is not necessarily needed, if W is available. In the following the homogenization problem is re-phrased in terms of a function approximation problem: a surrogate $\tilde{W} \approx W$ is sought-after.

The idea to approximate W has a rich history. The simplest approximation is given by the Taylor approximation (which goes back to the early work of Voigt [1910]). It defines a strict upper bound for the effective strain energy by asserting $\tilde{\mathbf{u}} = \mathbf{0}$, i.e. asserting constant strain in the entire microstructure. Other approaches include the Hashin-Shtrikman estimate [Hashin & Shtrikman, 1963] or the various variational methods by Ponte-Castañeda, Suquet and others [Ponte-Castañeda, 1991, Lahellec & Suquet, 2007a,b, Agoras et al., 2016]. The (semi-) analytical techniques have their up- and downsides: often they are computationally rather inexpensive. However, they cannot account for arbitrary microstructural geometries. The presence of pronounced geometrical features attracts computational methods. Most often these methods are realized in terms of finite element simulations on the microstructural level or, more recently, the techniques based on the Fast Fourier Transform (FFT; e.g. [Zeman et al., 2010, Kabel et al., 2014, Mishra et al., 2016]). In the following an approach that is realized using a FE discretization is used. However, FFT or Finite Difference based schemes could be used, as well.

5.3 Hyperelastic Reduced Order Model

5.3.1 Galerkin Reduced Order Model

The reduced order model (ROM) for the hyperelastic problem is constructed based on the minimum problem (5.15). Therefore it is assumed that a reduced basis (RB) \mathcal{E} of the strain fluctuations $\tilde{\varepsilon}$ is given (see Section 5.3.2 regarding the construction of the RB). The RB thereby defines a subset of kinematically admissible strain fields $\tilde{\mathcal{K}}^\varepsilon$ via

$$\tilde{\mathcal{K}}^\varepsilon(\bar{\varepsilon}) = \left\{ \varepsilon = \bar{\varepsilon} + \tilde{\varepsilon} \quad \text{with} \quad \tilde{\varepsilon} = \sum_{i=1}^N \xi_i \mathcal{E}_i \right\} \subset \mathcal{K}^\varepsilon, \quad (5.19)$$

where \mathcal{E}_i denotes the i th basis function of the N -dimensional basis \mathcal{E} (derived from a displacement fluctuation $\tilde{\mathbf{u}}_i$) and ξ_i is the related activity coefficient. For brevity the fluctuation field in (5.19) is denoted by $\tilde{\varepsilon}(\underline{\xi})$. The reduced problem stemming from (5.17) reads

$$\tilde{W}(\bar{\varepsilon}) = \min_{\varepsilon \in \tilde{\mathcal{K}}^\varepsilon} \frac{1}{|\Omega|} \int_{\Omega} \psi(\varepsilon) \, dV = \min_{\underline{\xi} \in \mathbb{R}^N} \frac{1}{|\Omega|} \int_{\Omega} \psi(\bar{\varepsilon} + \tilde{\varepsilon}(\underline{\xi})) \, dV. \quad (5.20)$$

Allowing for arbitrary variations of $\underline{\xi}$ leads to a set of N nonlinear equations representing the necessary conditions for a minimum of (5.20):

$$f_i(\underline{\xi}) = \frac{1}{|\Omega|} \int_{\Omega} \mathcal{E}_i \cdot \frac{\partial \psi(\bar{\varepsilon} + \tilde{\varepsilon}(\underline{\xi}))}{\partial \varepsilon} \, dV \stackrel{!}{=} 0 \quad (i = 1, \dots, N). \quad (5.21)$$

For convenience a vector notation for the strains is used (see Section 5.1.1), i.e.

$$\varepsilon \leftrightarrow \underline{\varepsilon}, \quad \bar{\varepsilon} \leftrightarrow \underline{\bar{\varepsilon}}, \quad \sigma \leftrightarrow \underline{\sigma}, \quad \dots \quad (5.22)$$

Then the RB for the strain fluctuation $\tilde{\varepsilon}$ can be expressed by a matrix-valued field

$$\underline{E}(\mathbf{x}) = \left[\underline{\mathcal{E}}_1(\mathbf{x}) \quad \underline{\mathcal{E}}_2(\mathbf{x}) \quad \dots \quad \underline{\mathcal{E}}_N(\mathbf{x}) \right] \in \Omega \mapsto \mathbb{R}^{6 \times N}. \quad (5.23)$$

The reduced problem (5.21) is then rewritten in the compact form

$$\underline{f}(\underline{\xi}) = \frac{1}{|\Omega|} \int_{\Omega} \underline{E}^\top \sigma \, dV \stackrel{!}{=} \underline{0}, \quad \underline{\sigma} \leftrightarrow \sigma = \frac{\partial \psi(\bar{\varepsilon} + \tilde{\varepsilon}(\underline{\xi}))}{\partial \varepsilon}. \quad (5.24)$$

Due to $\tilde{\mathcal{K}}^\varepsilon \subset \mathcal{K}^\varepsilon$ the effective free energy in the reduced setting \tilde{W} denotes an upper bound to the effective strain energy density W for arbitrary macroscopic strains $\bar{\varepsilon}$

$$\tilde{W}(\bar{\varepsilon}) \geq W(\bar{\varepsilon}). \quad (5.25)$$

The numerical solution of the nonlinear problem (5.24) can be obtained via a Newton-Raphson iteration requiring the Jacobian

$$\underline{\underline{J}}(\underline{\xi}) = \frac{1}{|\Omega|} \int_{\Omega} \underline{\underline{E}}^{\top} \underline{\underline{C}} \underline{\underline{E}} \, dV \quad \underline{\underline{C}} \leftrightarrow \frac{\partial^2 \psi}{\partial \underline{\varepsilon} \partial \underline{\varepsilon}}. \quad (5.26)$$

The Jacobian is also needed for the computation of the algorithmic stiffness operator. In order to derive the latter, the differential of (5.24) is investigated:

$$d\underline{f} = \underline{\underline{J}} \, d\underline{\xi} + \underline{\underline{\Sigma}} \, d\underline{\bar{\varepsilon}} = \underline{0} \quad \Rightarrow \quad d\underline{\xi} = -\underline{\underline{J}}^{-1} \underline{\underline{\Sigma}} \, d\underline{\bar{\varepsilon}}, \quad \underline{\underline{\Sigma}} = \left\langle \underline{\underline{E}}^{\top} \underline{\underline{C}} \right\rangle \quad (5.27)$$

The condition $d\underline{f} = \underline{0}$ enforces the stationary condition $\underline{f} = \underline{0}$ under changes of the strain $d\underline{\bar{\varepsilon}}$. Substituting $d\underline{\xi}$ from (5.27) into the differential of the microscopic strain $d\underline{\varepsilon} = d\underline{\bar{\varepsilon}} + \underline{\underline{E}} \, d\underline{\xi}$ implies

$$d\underline{\sigma} = \langle \underline{\underline{C}} \, d\underline{\varepsilon} \rangle = \langle \underline{\underline{C}} \rangle \, d\underline{\bar{\varepsilon}} + \langle \underline{\underline{C}} \underline{\underline{E}} \rangle \, d\underline{\xi} = \left[\langle \underline{\underline{C}} \rangle - \underline{\underline{\Sigma}}^{\top} \underline{\underline{J}}^{-1} \underline{\underline{\Sigma}} \right] \, d\underline{\bar{\varepsilon}}. \quad (5.28)$$

The symmetry of $\underline{\underline{C}}$ (since $\underline{\underline{C}}$ is the Hessian of ψ) is exploited in (5.28). The tangent stiffness is thus

$$\underline{\underline{C}}_{\text{A}} = \frac{d\underline{\sigma}}{d\underline{\bar{\varepsilon}}} = \langle \underline{\underline{C}} \rangle - \underline{\underline{\Sigma}}^{\top} \underline{\underline{J}}^{-1} \underline{\underline{\Sigma}}. \quad (5.29)$$

The algebraic structure of (5.29) confirms that the effective stiffness operator is unconditionally symmetric. Further it is bounded from above by the Taylor/Voigt approximation $\underline{\underline{C}} = \langle \underline{\underline{C}} \rangle$. In order to show this an arbitrary strain $\underline{e} \leftrightarrow \underline{\varepsilon}$ is introduced. Then

$$\underline{e}^{\top} \underline{\underline{C}}_{\text{A}} \underline{e} = \underline{e}^{\top} \underline{\underline{C}} \underline{e} - \underline{e}^{\top} \underline{\underline{\Sigma}}^{\top} \underline{\underline{J}}^{-1} \underline{\underline{\Sigma}} \underline{e} \quad (5.30)$$

holds. The local stiffness $\underline{\underline{C}}$ is positive (semi-) definite and, thus, the Jacobian $\underline{\underline{J}}$ is symmetric positive (semi-) definite. By defining $\underline{y} = \underline{\underline{\Sigma}} \underline{e}$ and accounting for the positivity of $\underline{\underline{J}}$, equation (5.30) induces the inequality

$$\underline{e}^{\top} \underline{\underline{C}}_{\text{A}} \underline{e} = \underline{e}^{\top} \underline{\underline{C}} \underline{e} - \underline{y}^{\top} \underline{\underline{J}}^{-1} \underline{y} \leq \underline{e}^{\top} \underline{\underline{C}} \underline{e} \quad \Leftrightarrow \quad \rho(\underline{\underline{C}}_{\text{A}}) \leq \rho(\underline{\underline{C}}), \quad (5.31)$$

where $\rho(\bullet)$ denotes the spectral radius.

A comprehensive algorithm for the Galerkin ROM for hyperelastic solids is presented in Appendix 5.8.2.

5.3.2 Generation of the reduced basis via snapshot POD

The reduced basis \mathcal{E} (RB) (or \underline{E} , respectively) was assumed to be given in Section 5.3.1. In the current work the RB is identified using a snapshot proper orthogonal decomposition (POD). The POD is performed on the strain fluctuation fields $\tilde{\varepsilon}$ gathered during precomputations performed using the high-fidelity discretization (here: using the FEM) on the fully resolved microscale problem. The standard $L^2(\Omega)$ inner product was cho-

sen for simplicity (operating on the strain fluctuations), but different inner products for the computation of the snapshot correlation matrix could be chosen at will. The authors would like to emphasize that taking the $L^2(\Omega)$ inner product is equivalent to a weighted inner product on the periodic displacement fluctuations w but is different from taking the $L^2(\Omega)$ product of the displacements which is a crucial observation.

In contrast to standard POD Galerkin procedures the method proposed in Section 5.3.1 does not operate directly on primal fields (i.e. on the displacements). Instead it operates directly on the strain fluctuations which is favorable from an algorithmic view-point. In contrast to assembling a displacement field and deducing therefrom the strains, the direct computation of the strains requires less floating point operations and allows for a straight-forward implementation at moderate programming effort. Still, the displacement field leading to the strain is stored and, therefore, the deformations can be computed for postprocessing purposes. An algorithm for the implementation of the snapshot POD can be found in Appendix 5.8.3.

5.4 Data-assisted homogenization using reduced order models

5.4.1 Layout

In order to find the function \tilde{W} approximating the solution of the hyperelastic homogenization problem **[H]** (see p. 53), a hierarchical procedure based on three approximation levels is proposed:

- **Level 1:** A (small) set of computationally involved finite element simulations are effected and $N_{\text{comp}}^{\text{FE}}$ field solutions are computed.
- **Level 2:** The FE results are processed by a snapshot POD in order to identify a reduced order model (ROM; see Section 5.3). The ROM is solved $N_{\text{comp}}^{\text{ROM}}$ times in order to generate snapshots of W and $\bar{\sigma}$.
- **Level 3:** Based on the data generated from the mechanically and mathematically supported ROM, a data-assisted interpolation algorithm is trained (see Section 5.5).

The procedure is summarized in Figure 5.3. The scores at the bottom of each box rate some of the properties, i.e. the reproduction capability (recover solution used for the model generation), the prediction ability (prediction of solutions for new loadings), the amount of recovered field information (e.g. local stress/strain information) and the overall computational efficiency (for computation of W , $\bar{\sigma}$, \mathbb{C}).

The RNEXP (Level 3) is constructed in two stages:

- Stage 1: Construction of the ROM [L.2] using FE solution from [L.1].
- Stage 2: Construction of the data-driven RNEXP [L.3] based on data from Stage 1.

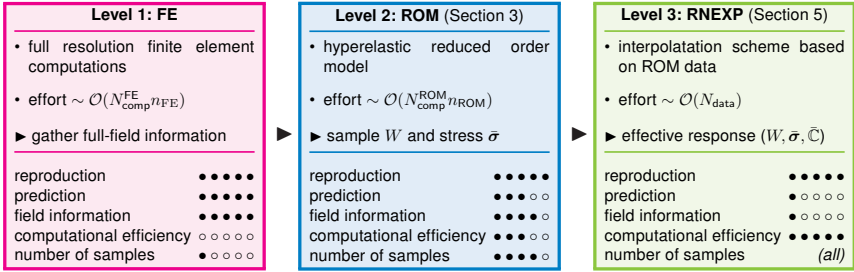


Figure 5.3: Sampling strategy for the staggered reduced order/data-driven homogenization procedure

Key ingredients in this construction are:

- Clever design of experiments to reduce the number of sampling points (see Sections 5.4.3 and 5.4.4). The definition of the sampling points has implications for all levels and is essential for the proposed homogenization scheme.
- The use of a (ROM, c.f. Section 5.3) for fast data-provisioning.
- A data-driven function approximation related to the NEXP [Yvonnet et al., 2009, Clément et al., 2012, Le et al., 2015] is used: the RNEXP.

5.4.2 Analysis of the NEXP of Yvonnet et al.

Yvonnet and co-workers [Yvonnet et al., 2009, Clément et al., 2012] have suggested the use of Numerically EXplicit Potentials (NEXP) in order to approximate the effective free energy W of hyperelastic, microstructured materials. The NEXP approximation is obtained by first computing via microscale FE simulations the effective strain energy on a regular grid¹ in strain space followed by a subsequent spline interpolation of the gathered data. The identification of the spline was much eased by using a tensor decomposition (referred to as NEXP2 in [Yvonnet et al., 2009] but called NEXP in the remainder of this paper). Similar tensor decomposition techniques include but are not limited to the recent TT-Cross [Drozdo et al., 2017] and hierarchical Tucker tensors (e.g. [Grasedyck, 2010]).

The main outputs of the NEXP are the effective stress $\bar{\sigma}_{\text{NEXP}}$ and the stiffness \bar{C}_{NEXP} which are determined by the first and second gradient of the data-driven approximation of the effective free energy \tilde{W}_{NEXP} , respectively:

$$\bar{\sigma}_{\text{NEXP}} = \frac{\partial \tilde{W}_{\text{NEXP}}}{\partial \bar{\epsilon}}, \quad \bar{C}_{\text{NEXP}} = \frac{\partial^2 \tilde{W}_{\text{NEXP}}}{\partial \bar{\epsilon} \partial \bar{\epsilon}}. \quad (5.32)$$

¹Irregular or nonuniform strain grids are briefly mentioned in Yvonnet et al. [2009], but have not been investigated to the best of our knowledge (also not in the follow-ups [Clément et al., 2012, Yvonnet et al., 2013]).

While the NEXP approach is generally appealing, the construction of the surrogate \tilde{W}_{NEXP} requires a substantial amount of training simulations. In the following let D be the resolution of the grid spanning the strain space along one direction (with three directions being needed for plane strain and six for general three dimensional problems). Then the grid required for building the approximation \tilde{W}_{NEXP} consists of D^6 sampling points for general 3D problems. At each sampling point a finite element simulation on the RVE level is needed. The number of D^6 simulations is prohibitive, even if D is rather small/moderate: a coarse grid consisting of $D = 10$ points in each direction in strain space would require a total of 1 Mio. finite element solutions of the nonlinear RVE problem. This is clearly beyond nowadays computing capabilities and (most likely) also beyond the ones that are expected in the near future, especially, if additional parameter variations accounting for changing microscopic characteristics are considered.

After closer inspection it is also found that the regular grid in strain space assumed for the original NEXP has an additional short-coming: At low strain amplitudes (i.e. for $\|\bar{\boldsymbol{\varepsilon}}\| \ll 1$), the regular grid has a limited accuracy with respect to resolving possible anisotropic material response of the microstructured solid. On the other hand the uniform grid spacing leads to a sort of oversampling at larger strain amplitudes. In order to illustrate this more clearly, the six-dimensional effective strain $\bar{\boldsymbol{\varepsilon}}$ is decomposed into its amplitude $\bar{\varepsilon}$ and a six-dimensional direction \mathbf{N}^ε (with $\|\mathbf{N}^\varepsilon\| = 1$) according to

$$\bar{\boldsymbol{\varepsilon}} = \bar{\varepsilon} \mathbf{N}^\varepsilon, \quad \bar{\varepsilon} = \|\bar{\boldsymbol{\varepsilon}}\|_2, \quad \mathbf{N}^\varepsilon = \frac{\bar{\boldsymbol{\varepsilon}}}{\bar{\varepsilon}}. \quad (5.33)$$

Making use of an orthonormal basis $\{\mathbf{B}^{(i)}\}_{i=1}^6$ spanning $Sym(\mathbb{R}^3)$, the grid points are parameterized via

$$\bar{\boldsymbol{\varepsilon}} = \sum_{i=1}^6 n_i h \mathbf{B}^{(i)}, \quad n_i \in \{-n_h, -(n_h - 1), \dots, n_h - 1, n_h\} \quad (5.34)$$

where $n_h = (D - 1)/2$, n_i is the discrete grid position and h is the grid spacing². At the lowest non-zero amplitudes (i.e. for $\bar{\boldsymbol{\varepsilon}} = h \mathbf{B}^{(i)}$ for arbitrary $i \in \{1, \dots, 6\}$) the minimum angle to a neighboring³ (not co-linear) data point is

$$\phi_{\min}(h \mathbf{B}^{(i)}) = \text{acos} \left(\frac{h \mathbf{B}^{(i)} \cdot h (\mathbf{B}^{(i)} + \mathbf{B}^{(j)})}{\sqrt{2} h^2} \right) = \text{acos} \left(\frac{\sqrt{2}}{2} \right) = \frac{\pi}{4} \equiv 45^\circ \quad (5.35)$$

and the largest angle to a neighboring grid point is

$$\phi_{\max}(h \mathbf{B}^{(i)}) = \text{acos} \left(\frac{1}{\sqrt{6}} \right) = \text{acos} \left(\frac{\sqrt{6}}{6} \right) \approx 0.366\pi \equiv 65.9^\circ. \quad (5.36)$$

A graphical representation of ϕ_{\min} and ϕ_{\max} is shown in Figure 5.4.

²(D should be odd in order to include zero strain as a sample point)

³Neighboring points denote points that can be reached by moving (at most) one grid width along each coordinate axis.

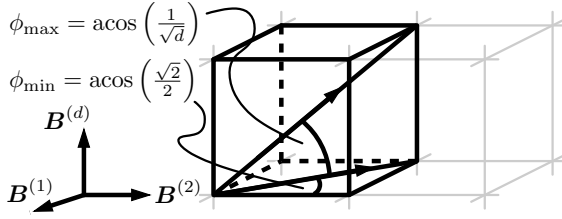


Figure 5.4: Hyper-cube for illustration of $\phi_{\min} = \pi/4$, $\phi_{\max} = \text{acos}(1/\sqrt{d})$ (where d is the spacial dimension, i.e. $d = 6$ for strains in $Sym(\mathbb{R}^3)$; here $d = 3$ is used in order to allow for a reasonable graphical representation)

Thus, directional information can only be provided in a rather scarce way which implies that the anisotropy of the effective behavior cannot be captured accurately for low strain amplitudes. However, at higher strain amplitudes we have

$$\phi_{\min}(n_h h \mathbf{B}^{(i)}) = \text{acos}\left(\frac{n_h^2}{n_h \sqrt{n_h^2 + 1}}\right) = \text{acos}\left(\frac{n_h}{\sqrt{n_h^2 + 1}}\right) \approx \text{acos}(1) = 0. \quad (5.37)$$

More precisely, data is provided at states for which the directions are nearly parallel and the amplitudes are comparable, too. Even for moderate values like $D = 11$, the angle gets as low as $0.0628\pi \equiv 11.31^\circ$ and for $D = 21$ it is $0.0317\pi \equiv 5.71^\circ$. Note also that along the diagonal of the hyper-cube in strain space parameterized via

$$\mathbf{D}_k = kh \sum_{i=1}^6 \mathbf{B}^{(i)}, \quad (5.38)$$

the minimum nearest neighbor angle is much lower (numerical values are for $D = 11$):

$$\Theta_{\min}(\mathbf{D}_1) = \text{acos}\left(\frac{7}{3\sqrt{6}}\right) \stackrel{n_h=5}{\approx} 0.0984\pi \equiv 17.72^\circ, \quad (5.39)$$

$$\Theta_{\min}(\mathbf{D}_{n_h}) = \text{acos}\left(\frac{6n_h^2 + n_h}{\sqrt{6} n_h \sqrt{5n_h^2 + (n_h + 1)^2}}\right) \stackrel{n_h=5}{\approx} 0.0229\pi \equiv 4.13^\circ. \quad (5.40)$$

In summary for $p = 11$ the nearest angle to neighboring data points varies between approximately 4.13° and 45° for a uniform grid. For $D = 21$ (i.e. for $n_h = 10$) the number of data points is increased by a factor $\approx 48 = (21/11)^6$. Still, around the origin the minimum nearest neighbor angle is $\pi/4 \equiv 45^\circ$ and the smallest possible nearest neighbor angle is $0.0117\pi \equiv 2.10^\circ$. This means that the resolution of the data sampling with respect to the direction (i) depends on the strain amplitude and (ii) is anisotropic.

Another limitation of the regular grid in strain space is the coarse resolution in radial direction (i.e. with respect to the strain amplitude): at low strains the nonlinearity of the

materials presented in Section 5.2.2 is pronounced and it decays quickly at higher strain amplitudes.

The above observations show that – although it can lead to good results of the interpolation function in parts of the strain space – the representation of the data is heterogeneous within the grid and, hence, a lot of near redundant computations are performed at the outer boundary. Additionally it is found that the behavior for low amplitudes of the strain tensor is approximated in a rather coarse way although this domain is likely to be important for practical applications.

The authors of the NEXP have more recently accounted for some of these short-comings in [Le et al., 2015] where they combine the idea of numerically explicit potentials with neural networks trained by the high-dimensional model representation method (HDMR; see, e.g., [Manzhos & Carrington Jr, 2006]). The data sampling is performed using Monte Carlo type sampling, i.e. unstructured random samples are drawn.

5.4.3 Sampling strategies: preliminary considerations

The levels 1 and 2 provide information samples for the next level of the model which is eventually used for the prediction of the homogenized response in the purely data-driven level 3. In order to perform the data sampling on either of the levels 1 and 2 (in the following [L.1] and [L.2], respectively) efficiently, the sought-after outputs and the algorithmic properties of the underlying numerical procedure (i.e. FE for [L.1] vs. ROM for [L.2]) must be considered:

- **Stage 1 ([L.1] ► data for [L.2]): Finite Element simulations for building the ROM**

In order to provide reasonable inputs for [L.2] the full parameter space should be explored. This implies that (i) strain amplitudes from 0 to the expected maximum strain should be considered, (ii) a sufficient accuracy in the straining direction (e.g. in terms of angle to neighboring directions) should be considered and (iii) at low strain amplitudes the load increments should be small while they may increase gradually⁴.

Input during [L.1] are the microstructural geometry (in terms of a FE mesh), the material properties and the loading directions and amplitudes. The FE simulations benefit from constant loading direction which lead to good performance during Newton-Raphson iterations. The outputs of [L.1] used for [L.2] comprise the displacement field $\mathbf{u}(\mathbf{x})$ and the strain field $\boldsymbol{\varepsilon}(\mathbf{x})$.

- **Stage 2 ([L.2] ► data for [L.3]): Data generation using the ROM for training the RNEXP**

The reduced order model (see Section 5.3) allows for rapid computations of the effective free energy $W(\bar{\boldsymbol{\varepsilon}})$ and of the effective stress $\bar{\boldsymbol{\sigma}}$. The performance of the ROM computations benefits from the proportional loading for the same reasons stated

⁴The optimal strain amplitudes depend strongly on the microscopic materials and partially on the considered microstructure.

in Stage 1 for the FE simulations. The data samples $(W, \bar{\sigma})$ gathered in [L.2] provide the inputs for the data-assisted predictions of the RNEXP [L.3].

5.4.4 Discretization of load amplitudes and directions

In order to easily account for the directional dependency of the constitutive response as well as for the higher strain amplitude sensitivity at low amplitudes, a coordinate separation into direction and amplitude is suggested according to (5.33). The proposed split can be seen as a special case of *stratified sampling* (cf. [Mckay et al., 2000]). It is noteworthy that for the purpose of general twoscale simulations a large part of the strain space (i.e. arbitrary directions up to a considerable load amplitude) need to be sampled. This is due to the fact that the loading within the macroscopic structure is usually a priori unknown and can be strongly heterogeneously distributed with unexpected peaks. Therefore, no lower dimensional manifold can be identified for the sampling which could allow for further compression of the sampling set.

In this section, we explain the proposed discretization of the strain space in general. The question of exactly how many points to choose and where to locate them for a particular problem will be addressed in Section 5.5.5, as well as notes on refinement.

Along a constant direction N_i^ε different load amplitudes $\bar{\varepsilon}_j$ are considered. In order to capture the often pronounced dependency on the load amplitude at low strains, a simple geometric series is proposed with parameters $\bar{\varepsilon}_1 > 0, R \geq 0$:

$$\Delta\bar{\varepsilon}_1 = \bar{\varepsilon}_1 > 0, \quad \Delta\bar{\varepsilon}_{k+1} = (1 + R)\Delta\bar{\varepsilon}_k, \quad \bar{\varepsilon}_{k+1} = \bar{\varepsilon}_k + \Delta\bar{\varepsilon}_{k+1} \quad (k = 1, \dots, q). \quad (5.41)$$

For given number of radial steps q and final strain amplitude $\bar{\varepsilon}_{\max}$, the initial load step $\bar{\varepsilon}_1$ can be computed as a function of R through

$$\bar{\varepsilon}_1(R, q, \bar{\varepsilon}_{\max}) = \frac{\bar{\varepsilon}_{\max} R}{(1 + R)^q - 1}. \quad (5.42)$$

Alternatively, R can be identified for given $\bar{\varepsilon}_1, \bar{\varepsilon}_{\max}$ numerically. An example for $R = 0.3, q = 10$ and $\bar{\varepsilon}_{\max} = 2.5\%$ is given in Table 5.1. These numerical values show that a high sampling point density at small loadings (e.g. close to the yield point) can be recovered.

Table 5.1: Example values for the nonuniform sampling in radial direction ($R = 0.3, q = 10, \bar{\varepsilon}_{\max} = 2.5\%$)

i	1	2	3	4	5
$\bar{\varepsilon}_i$ [%]	0.0587	0.1349	0.2341	0.3629	0.5305
i	6	7	8	9	10
$\bar{\varepsilon}_i$ [%]	0.7483	1.031	1.400	1.878	2.500

In order to efficiently generate training directions the approach presented by Fritzen & Leuschner [2015] (see Section 3.6 therein) is adopted. In brief, a procedure mimicking electrostatic repulsion is used to attain a near-homogeneous distribution of points on the hyper-sphere in \mathbb{R}^6 . The symmetry of the material behavior is accounted for by using a symmetrized version of the repulsive potential (in accordance with Section 5.5.4; see below). The method takes as parameters the number of directions to be generated and one fixed direction ($\mathbf{N}_1^\varepsilon = \mathbf{e}_x \otimes \mathbf{e}_x$ in our examples). A refined version of the algorithm is currently under development and will be the topic of another study.

Examples of the outputs of the algorithm are shown in Figure 5.5: First, preliminary results for two-dimensional point sets are presented which highlight that a perfectly homogeneous neighbor distance distribution function can be achieved. For the general

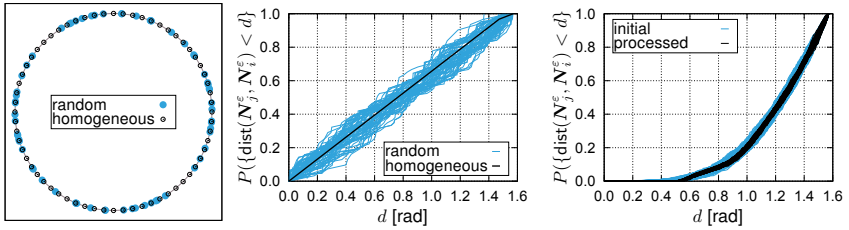


Figure 5.5: *Left:* Illustration of point sets on a two-dimensional hyper-sphere (with point symmetry); initial random points blue and after repulsion algorithm (black); distribution of the geodesic distance function for random points in \mathbb{R}^2 (*middle*) and \mathbb{R}^6 (*right*); initial data: blue lines; processed data: black lines; each line represents the probability distribution for one direction \mathbf{N}_i^ε , $i = 1, \dots, 512$ (symmetry cf. (5.72) applies).

six dimensional space the probability distribution of the neighbor distance is significantly improved after the processing: The distribution functions of the k -nearest neighbor for different \mathbf{N}_i^ε are almost identical and they have an almost uniform nearest neighbor distance which avoids unwanted oscillations and overshooting in the kernel based interpolation of Section 5.5.

5.5 Radial numerically explicit potentials (RNEXP)

5.5.1 Data sampling strategy

For the reasons given in Section 5.4.4, the data sampling is done by using an amplitude–direction split. Therefore, normalized loading directions $\mathbf{N}_i^\varepsilon \in \text{Sym}(\mathbb{R}^3)$ ($i = 1, \dots, n$) and nonuniformly spaced amplitudes $r_j \in (0, \bar{\varepsilon}_{\max})$ ($j = 1, \dots, q$) are considered. This split results into a total of $N_{\text{comp}} = nq$ sampling positions, where N_{comp} corresponds to $N_{\text{comp}}^{\text{FE}}$ (in [L.1], Fig. 5.3) and $N_{\text{comp}}^{\text{ROM}}$ (in [L.2], Fig. 5.3) and q could depend likewise on the hierarchical approximation level. It is chosen identical in this study. The proposed

discretization helps to retain the important physical information while minimizing the number of sampling positions.

For each of these sampling positions the free energy and the stress in radial direction are stored:

$$D_{ij} = W(r_j \mathbf{N}_i^\varepsilon), \quad \underline{\underline{D}} \in \mathbb{R}^{n \times q}, \quad i \in \{1, \dots, n\}, \quad j \in \{1, \dots, q\}, \quad (5.43)$$

$$T_{ij} = \mathbf{N}_i^\varepsilon \cdot \bar{\boldsymbol{\sigma}}(r_j \mathbf{N}_i^\varepsilon), \quad \underline{\underline{T}} \in \mathbb{R}^{n \times q}, \quad i \in \{1, \dots, n\}, \quad j \in \{1, \dots, q\}. \quad (5.44)$$

The coefficients T_{ij} correspond to the uniaxial stress in the loading direction. At the same time, T_{ij} is the derivative of W with respect to the load amplitude:

$$\frac{\partial W(\bar{\boldsymbol{\varepsilon}})}{\partial r} = \frac{\partial W(\bar{\boldsymbol{\varepsilon}})}{\partial \bar{\boldsymbol{\varepsilon}}} \cdot \frac{\partial \bar{\boldsymbol{\varepsilon}}}{\partial r} = \bar{\boldsymbol{\sigma}} \cdot \mathbf{N}^\varepsilon \stackrel{(5.44)}{\leftrightarrow} T_{ij}. \quad (5.45)$$

The storage of both, D_{ij} and T_{ij} is a noteworthy difference to existing interpolation strategies where only W is sampled. Note that the amplitudes r_j have nonuniform spacing cf. Section 5.4.4, Table 5.1.

5.5.2 Interpolation scheme

In order to attain the sought-after outputs (i.e. the effective free energy, the effective stress and the stiffness operator), accurate approximations of the free energy and its first and second gradient are required based on the sampling of Section 5.5.1.

Along each training direction \mathbf{N}_i^ε a C^1 -continuous piecewise cubic polynomial $\mathcal{S}_i(r)$ is fitted. The function $\mathcal{S}_i \in \mathcal{P}^3(\mathbb{R}) : [r_j, r_{j+1}] \mapsto \mathbb{R}$ needs to satisfy the conditions

$$\mathcal{S}_i(r_j) = D_{ij}, \quad \mathcal{S}'_i(r_j) = T_{ij}, \quad \mathcal{S}_i(r_{j+1}) = D_{i(j+1)}, \quad \mathcal{S}'_i(r_{j+1}) = T_{i(j+1)}, \quad (5.46)$$

which uniquely define cubic polynomials on all intervals. The additional conditions $W(\mathbf{0}) = 0$ and $\bar{\boldsymbol{\sigma}}(\mathbf{0}) = \mathbf{0}$ are used, i.e. the undeformed configuration is assumed to be stress-free and, hence, with zero free energy⁵. In Figure 5.6 a representative example is shown in which the interpolation (blue curve) and the provided data are shown together. At each sampling point the slope (represented by T_{ij}) is illustrated via black line segments. By the metric of vision the piecewise cubic interpolation seems to provide accurate approximations of the data which was also confirmed by comparison to additional computations with different load increments. The functions \mathcal{S}_i are straight-forward to adjust to the data and they are independent of the nonuniform radial discretization. In contrast to cubic spline interpolations no additional algebraic conditions are needed at the boundary.

⁵Chemical, thermal and other contributions to the free energy can be disregarded in the present, purely mechanical, setting.

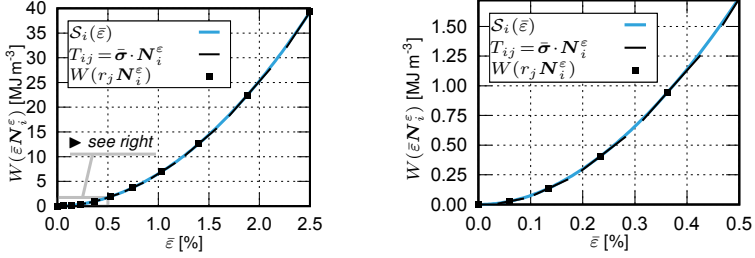


Figure 5.6: Comparison of the data (energies D_{ij} and slopes T_{ij} ; **black symbols/line segments**) to the piecewise cubic interpolation (**blue**) defined by (5.46); the data corresponds to a training loadings for a microstructure including 10% particles); full load range (*left*) and first load steps (*right*)

Next, a coordinate separation is proposed according to

$$W(\bar{\epsilon} N^\epsilon) \approx \tilde{W}(\bar{\epsilon}, N^\epsilon) = \sum_{i=1}^n w_i(N^\epsilon) S_i(\bar{\epsilon}). \quad (5.47)$$

Here $w_i(N^\epsilon)$ denotes the interpolation weight for the i th radial interpolation function $S_i(\bar{\epsilon})$. More precisely, first the energy along the training directions N_i^ϵ is interpolated via $S_i(\bar{\epsilon})$. Then the weights $w_i(N^\epsilon)$ perform the tangential interpolation on the hyper-sphere in \mathbb{R}^6 at constant strain amplitude $\bar{\epsilon}$. They depend on the distance of N^ϵ to the training directions N_i^ϵ only, and they are constrained by

$$\sum_{i=1}^n w_i(N^\epsilon) = 1. \quad (5.48)$$

In order to allow for accurate reproduction of the data D_{ij} and T_{ij} the weights need to satisfy

$$w_i(N_j^\epsilon) = \delta_{ij} = \begin{cases} 0 & i \neq j, \\ 1 & i = j. \end{cases} \quad (5.49)$$

In the following the weight w_i is computed by a radial basis functions (RBF) operating on the geodesic distance ξ_i of the direction N^ϵ to the training direction N_i^ϵ . The distance ξ_i is computed through

$$\theta_i(N^\epsilon) = N^\epsilon \cdot N_i^\epsilon = \cos(\xi_i) \in [-1, 1] \quad \Rightarrow \quad \xi_i(N^\epsilon) = \text{acos}(\theta_i). \quad (5.50)$$

Next the Gaussian kernel function

$$\zeta_i(N^\epsilon) = \exp(-\gamma \xi_i^2(N^\epsilon)) \quad (5.51)$$

with positive parameter $\gamma > 0$ is considered as RBF. For the training directions \mathbf{N}_i^ε the symmetric kernel matrix

$$K_{ij} = K_{ji} = \exp(-\gamma \xi_i^2(\mathbf{N}_j^\varepsilon)) \quad (5.52)$$

is defined. Note that for unique (i.e. non-repeated) training directions \underline{K} is positive definite, i.e. invertible. Then the weights $\underline{w}(\mathbf{N}^\varepsilon) \in \mathbb{R}^n$ can be computed via

$$\underline{w}(\mathbf{N}^\varepsilon) = \underline{K}^{-1} \underline{\zeta}(\mathbf{N}^\varepsilon), \quad (5.53)$$

where $\underline{\zeta} \in \mathbb{R}^n$ is a vector with components defined via (5.51). Through (5.53) the Kronecker delta property (5.49) is a priori satisfied. The normalization condition (5.48) is more intricate to investigate. The authors have found that for the training directions provided through the algorithm of Section 5.4.4, the weights (very) closely satisfied the partition of unity condition for all γ that lead to reasonable interpolations, i.e. in the considered examples the deviations were on the order of 10^{-5} and below which can be considered negligible (for $n = 512$ training directions). Note that the condition (5.48) could be enforced by uniform a posteriori correction of the weights.

Finally, the approximation \tilde{W} is expressed through

$$\tilde{W}(\bar{\varepsilon}, \mathbf{N}^\varepsilon) = \underline{\mathcal{D}}^\top(\bar{\varepsilon}) \underline{K}^{-1} \underline{\zeta}(\mathbf{N}^\varepsilon) = \underline{\mathcal{D}}^\top(\bar{\varepsilon}) \underline{w}(\mathbf{N}^\varepsilon), \quad w_i = \sum_{k=1}^n \left(\underline{K}^{-1} \right)_{ik} \left(\underline{\zeta}(\mathbf{N}^\varepsilon) \right)_k. \quad (5.54)$$

The representation (5.54) emphasizes the radial–tangential (or amplitude–direction) split: first the data in *radial* direction is interpolated along all training directions \mathbf{N}_i^ε at the amplitude $\bar{\varepsilon}$ of the query point. This data is then interpolated in *tangential* direction, i.e. on the hyper-sphere in \mathbb{R}^6 , through the weights \underline{w} emerging from the Gaussian kernel.

Remark 2 Oscillations and overshooting Kernel methods with the Kronecker property (5.49) can produce strongly oscillatory function interpolations. Here, the sensible selection of the training directions via the procedure described in Section 5.4.4 pays off: the geodesic distance between the training directions is (almost) identically distributed. This leads to a well-conditioned kernel matrix for near arbitrary values of γ and, importantly, the accuracy of the RNEXP is not too sensitive with respect to γ .

5.5.3 Computation of stress and stiffness

The stress and the stiffness will now be computed via the gradients of (5.54). For this purpose one could, principally speaking, store only the energy (5.43) resulting from [L.2]. Instead, the stress in radial direction (5.44) is also stored and used as input to the radial interpolation, cf. (5.46). Taking this gradient information into account increases the accuracy of the piecewise cubic polynomial interpolation considerably, see also Figure 5.6. Although generally possible within the proposed interpolation scheme, a direct interpolation of the stresses is not followed in this work by purpose. This is motivated by the fact that the related tangent operator would lose symmetry which

contradicts the existence of a hyperelastic energy function. Further, the energy density reconstructed from a direct interpolation of the stresses could lead to numerically induced path-dependency which should be avoided.

In order to evaluate (5.54) the constant kernel matrix \underline{K} is pre-computed and its inverse \underline{K}^{-1} is stored. Then at each evaluation point $\bar{\varepsilon} = \bar{\varepsilon} \mathbf{N}^\varepsilon$ the cubic interpolations along the training directions $\underline{S}(\bar{\varepsilon})$ and the vector $\underline{\zeta}(\mathbf{N}^\varepsilon)$ composed of the Gaussian kernel functions is computed. The effort for one evaluation is $\mathcal{O}(n^2)$ due to the (dense) matrix-vector product involving the inverse of the kernel matrix \underline{K} . However, all operations are simple explicit function evaluations operating on independent data in an algorithmically favorable form (see also Appendix 5.8.4).

In order to evaluate the gradients of the approximation (5.54) some intermediate steps are useful. The following relations hold for the derivatives of θ_i and ξ_i with respect to the total strain:

$$\frac{\partial \theta_i}{\partial \bar{\varepsilon}} = \frac{1}{\bar{\varepsilon}} \mathbb{P}^\varepsilon \mathbf{N}_i^\varepsilon = \frac{1}{\bar{\varepsilon}} (\mathbf{N}_i^\varepsilon - \theta_i \mathbf{N}^\varepsilon) = \frac{\sin(\xi_i)}{\bar{\varepsilon}} \mathbf{Q}_i^\varepsilon, \quad (5.55)$$

$$\begin{aligned} \frac{\partial^2 \theta_i}{\partial \bar{\varepsilon} \partial \bar{\varepsilon}} &= \frac{1}{\bar{\varepsilon}^2} (3\theta_i \mathbf{N}^\varepsilon \otimes \mathbf{N}^\varepsilon - \theta_i \mathbb{I} - 2 \text{sym}(\mathbf{N}_i^\varepsilon \otimes \mathbf{N}^\varepsilon)) \\ &= -\frac{1}{\bar{\varepsilon}^2} (\theta_i \mathbb{P}^\varepsilon + 2 \sin(\xi_i) \text{sym}(\mathbf{Q}_i^\varepsilon \otimes \mathbf{N}^\varepsilon)), \end{aligned} \quad (5.56)$$

where $\mathbb{P}^\varepsilon = \mathbb{I} - \mathbf{N}^\varepsilon \otimes \mathbf{N}^\varepsilon$. Geometric considerations motivate the definition of the direction \mathbf{Q}_i^ε pointing from \mathbf{N}^ε towards \mathbf{N}_i^ε within the tangent plane defined by the normal \mathbf{N}^ε (see also Figure 5.7):

$$\mathbf{Q}_i^\varepsilon = \frac{1}{\sin(\xi_i)} \mathbb{P}^\varepsilon \mathbf{N}_i^\varepsilon = \frac{1}{\sin(\xi_i)} (\mathbf{N}_i^\varepsilon - \theta_i \mathbf{N}^\varepsilon), \quad \|\mathbf{Q}_i^\varepsilon\| = 1, \quad \mathbf{N}^\varepsilon \cdot \mathbf{Q}_i^\varepsilon = 0. \quad (5.57)$$

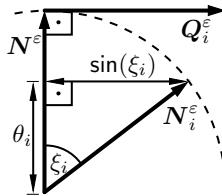


Figure 5.7: Geometric interpretation of \mathbf{Q}_i^ε

Using the presented abbreviations the derivative of ξ_i with respect to $\bar{\varepsilon}$ can be expressed via

$$\frac{\partial \xi_i}{\partial \bar{\varepsilon}} = -\frac{1}{\sin(\xi_i)} \frac{\partial \theta_i}{\partial \bar{\varepsilon}} = -\frac{1}{\bar{\varepsilon}} \mathbf{Q}_i^\varepsilon, \quad (5.58)$$

$$\frac{\partial^2 \xi_i}{\partial \bar{\varepsilon} \partial \bar{\varepsilon}} = \frac{1}{\bar{\varepsilon}^2} \left(\frac{\theta_i}{\sin(\xi_i)} (\mathbb{P}^\varepsilon - \mathbf{Q}_i^\varepsilon \otimes \mathbf{Q}_i^\varepsilon) + 2 \text{sym}(\mathbf{Q}_i^\varepsilon \otimes \mathbf{N}^\varepsilon) \right) \stackrel{\text{def.}}{=} \frac{1}{\bar{\varepsilon}^2} \mathbf{Q}_i. \quad (5.59)$$

Combining the derivatives of ζ_i

$$\zeta_i' = \frac{\partial \zeta_i}{\partial \xi_i} = -2\gamma \xi_i \zeta_i, \quad \zeta_i'' = \frac{\partial^2 \zeta_i}{\partial \xi_i \partial \xi_i} = -2\gamma (\zeta_i + \xi_i \zeta_i') = -2\gamma \zeta_i (1 - 2\gamma \xi_i^2), \quad (5.60)$$

with (5.58) and (5.59), the derivatives of the kernel function ζ_i with respect to the macroscopic strain $\bar{\varepsilon}$ are

$$\frac{\partial \zeta_i}{\partial \bar{\varepsilon}} = \zeta_i' \frac{\partial \xi_i}{\partial \bar{\varepsilon}} = -\frac{\zeta_i'}{\bar{\varepsilon}} \mathbf{Q}_i^\varepsilon = \frac{2\gamma \xi_i \zeta_i}{\bar{\varepsilon}} \mathbf{Q}_i^\varepsilon, \quad (5.61)$$

$$\frac{\partial^2 \zeta_i}{\partial \bar{\varepsilon} \partial \bar{\varepsilon}} = \zeta_i' \frac{\partial^2 \xi_i}{\partial \bar{\varepsilon} \partial \bar{\varepsilon}} + \zeta_i'' \frac{\partial \xi_i}{\partial \bar{\varepsilon}} \otimes \frac{\partial \xi_i}{\partial \bar{\varepsilon}} = \frac{\zeta_i'}{\bar{\varepsilon}^2} \mathbf{Q}_i + \frac{\zeta_i''}{\bar{\varepsilon}^2} \mathbf{Q}_i^\varepsilon \otimes \mathbf{Q}_i^\varepsilon. \quad (5.62)$$

The derivatives of the radial approximation \mathcal{S}_i are

$$\mathcal{S}_i' = \frac{\partial \mathcal{S}_i}{\partial \bar{\varepsilon}}, \quad \mathcal{S}_i'' = \frac{\partial^2 \mathcal{S}_i}{\partial \bar{\varepsilon} \partial \bar{\varepsilon}}, \quad \frac{\partial \mathcal{S}_i}{\partial \bar{\varepsilon}} = \mathcal{S}_i' \mathbf{N}^\varepsilon, \quad \frac{\partial^2 \mathcal{S}_i}{\partial \bar{\varepsilon} \partial \bar{\varepsilon}} = \mathcal{S}_i' \frac{1}{\bar{\varepsilon}} \mathbb{P}^\varepsilon + \mathcal{S}_i'' \mathbf{N}^\varepsilon \otimes \mathbf{N}^\varepsilon. \quad (5.63)$$

Last, the following partial derivatives of \tilde{W} are provided in convenient matrix-vector notation:

$$\frac{\partial W}{\partial \underline{\mathcal{S}}} = \underline{\underline{K}}^{-1} \underline{\zeta}, \quad \frac{\partial W}{\partial \underline{\zeta}} = \underline{\underline{K}}^{-1} \underline{\mathcal{S}}, \quad \frac{\partial^2 W}{\partial \underline{\mathcal{S}} \partial \underline{\zeta}} = \frac{\partial^2 W}{\partial \underline{\zeta} \partial \underline{\mathcal{S}}} = \underline{\underline{K}}^{-1} \quad (5.64)$$

With these technical, yet required, abbreviations at hand the effective stress can be derived:

$$\bar{\sigma} = \frac{\partial W}{\partial \bar{\varepsilon}} = \frac{\partial W}{\partial \underline{\mathcal{S}}} \frac{\partial \underline{\mathcal{S}}}{\partial \bar{\varepsilon}} \mathbf{N}^\varepsilon + \frac{\partial W}{\partial \underline{\zeta}} \frac{\partial \underline{\zeta}}{\partial \bar{\varepsilon}} = \mathbf{N}^\varepsilon \left(\mathcal{S}'^T \underline{\underline{K}}^{-1} \underline{\zeta} \right) - \frac{1}{\bar{\varepsilon}} \sum_{i=1}^n \left(\underline{\underline{K}}^{-1} \underline{\mathcal{S}} \right)_i \zeta_i' \mathbf{Q}_i^\varepsilon. \quad (5.65)$$

The stress can be rewritten in a convenient format via

$$\bar{\sigma} = \sigma_0 \mathbf{N}^\varepsilon + \sum_{i=1}^n \tau_i \mathbf{Q}_i^\varepsilon, \quad \sigma_0 = \underline{\mathcal{S}}'^T \underline{\underline{K}}^{-1} \underline{\zeta}, \quad \tau_i = -\frac{1}{\bar{\varepsilon}} \left(\underline{\underline{K}}^{-1} \underline{\mathcal{S}} \right)_i \zeta_i', \quad (5.66)$$

where σ_0 is the stress acting in direction of the current strain and the components τ_i denote stress contributions in the normal plane defined by \mathbf{N}^ε pointing towards the

i th training direction, see Figure 5.8 for a schematic representation. Note that the computational effort for the stress prediction is proportional to the number of training directions, i.e. $\mathcal{O}(n)$, in addition to the computation of the matrix-vector products in (5.65) which are in $\mathcal{O}(n^2)$.

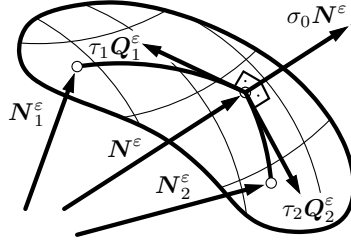


Figure 5.8: Schematic representation of the stress cf. eq. (5.65); σ_0 is the stress in loading direction N^ε and τ_i denotes components within the tangent plane defined by its normal N^ε

After some straight-forward but nevertheless rather technical steps the tangent stiffness is obtained as

$$\bar{\mathbb{C}} = \frac{\partial^2 W}{\partial \underline{\varepsilon} \partial \underline{\varepsilon}} = \frac{\partial W}{\partial \underline{\mathcal{S}}} \frac{\partial^2 \underline{\mathcal{S}}}{\partial \underline{\varepsilon} \partial \underline{\varepsilon}} + \frac{\partial W}{\partial \underline{\zeta}} \frac{\partial \underline{\zeta}}{\partial \underline{\varepsilon}} \frac{\partial \underline{\zeta}}{\partial \underline{\varepsilon}} + \frac{\partial^2 W}{\partial \underline{\mathcal{S}} \partial \underline{\zeta}} \frac{\partial \underline{\mathcal{S}}}{\partial \underline{\varepsilon}} \otimes \frac{\partial \underline{\zeta}}{\partial \underline{\varepsilon}} + \frac{\partial^2 W}{\partial \underline{\zeta} \partial \underline{\mathcal{S}}} \frac{\partial \underline{\zeta}}{\partial \underline{\varepsilon}} \otimes \frac{\partial \underline{\mathcal{S}}}{\partial \underline{\varepsilon}} \quad (5.67)$$

$$= \bar{\alpha} \mathbb{I}^s + \sum_{i=1}^n \beta_i \underline{Q}_i^\varepsilon \otimes \underline{Q}_i^\varepsilon + 2 \text{sym}(\bar{\underline{Q}}^\varepsilon \otimes N^\varepsilon) + \bar{\mu} N^\varepsilon \otimes N^\varepsilon. \quad (5.68)$$

The coefficients $\bar{\alpha}$, $\bar{\mu}$, β_i , ω_i and the tensor $\bar{\underline{Q}}^\varepsilon$ are given by

$$\alpha_i = \frac{1}{\bar{\varepsilon}^2} \frac{\theta_i}{\sin(\xi_i)} \left(\underline{\mathcal{S}}^\top \underline{K}^{-1} \right)_i \zeta_i', \quad \bar{\alpha} = \frac{1}{\bar{\varepsilon}} \left(\underline{\mathcal{S}}'^\top \underline{K}^{-1} \underline{\zeta} \right) + \sum_{i=1}^n \alpha_i, \quad (5.69)$$

$$\bar{\mu} = \underline{\mathcal{S}}''^\top \underline{K}^{-1} \underline{\zeta} - \bar{\alpha}, \quad \beta_i = \left(\underline{K}^{-1} \underline{\mathcal{S}} \right)_i \frac{\zeta_i''}{\bar{\varepsilon}^2} - \alpha_i, \quad (5.70)$$

$$\bar{\underline{Q}}^\varepsilon = \frac{1}{\bar{\varepsilon}^2} \sum_{i=1}^n \left(\underline{K}^{-1} (\underline{\mathcal{S}} - \bar{\varepsilon} \underline{\mathcal{S}}') \right)_i \zeta_i' \underline{Q}_i^\varepsilon. \quad (5.71)$$

Obviously, $\bar{\mathbb{C}}$ is unconditionally symmetric. Further algorithmic advantage could be drawn from substituting $\underline{Q}_i^\varepsilon$ through N^ε and N_i^ε via (5.57) which is omitted here for brevity.

Similar to the representation of the effective stress cf. (5.66), the tangent stiffness (5.68) has a particular structure. The assembly of the tangent stiffness can be realized with good computational efficiency:

- one matrix-vector product with $\mathcal{O}(n^2)$ in order to get $\underline{\underline{K}}^{-1} \underline{\underline{\zeta}}'$ (needed for $\tau_i, \overline{\mathbf{Q}}^\varepsilon$),
- n rank one updates (due to $\mathbf{Q}_i^\varepsilon \otimes \mathbf{Q}_i^\varepsilon$),
- one tensor averaging (for $\overline{\mathbf{Q}}^\varepsilon$),
- few algebraic operations (e.g., averaging of scalars; final rank one updates containing $\overline{\mathbf{Q}}^\varepsilon$ etc.).

5.5.4 Symmetry considerations

In order to improve the accuracy of the interpolation scheme without the need of additional training data, the consideration of symmetries in the material response is of interest. Assuming symmetric material behavior in the sense of $W(\bar{\varepsilon}) = W(-\bar{\varepsilon})$, the modified kernel function

$$\tilde{\zeta}_i = \zeta_i + \exp(-\gamma(\pi - \xi_i)^2). \quad (5.72)$$

can be defined. The derivatives of $\tilde{\zeta}_i$ are given by

$$\tilde{\zeta}'_i = \zeta'_i + 2\gamma(\pi - \xi_i) \exp(-\gamma(\pi - \xi_i)^2), \quad (5.73)$$

$$\tilde{\zeta}''_i = \zeta''_i - 2\gamma \exp(-\gamma(\pi - \xi_i)^2) + 4\gamma^2(\pi - \xi_i)^2 \exp(-\gamma(\pi - \xi_i)^2). \quad (5.74)$$

Replacing ζ in eq. (5.51) by the symmetrized kernel function defined through (5.72) and substituting ζ'_i and ζ''_i by $\tilde{\zeta}'_i$ and $\tilde{\zeta}''_i$ respectively, the interpolation scheme automatically considers the material symmetry:

$$\begin{aligned} \tilde{\zeta}_i(\mathbf{N}^\varepsilon) &= \exp(-\gamma\xi_i^2) + \exp(-\gamma(\pi - \xi_i)^2) \\ &= \exp(-\gamma(\pi - \xi_i)^2) + \exp(-\gamma\xi_i^2) = \tilde{\zeta}_i(-\mathbf{N}^\varepsilon). \end{aligned} \quad (5.75)$$

The symmetry consideration allows to double the number of effective sampling points without additional training simulations.

Further, the implementation via the symmetrized kernel function $\tilde{\zeta}$ affects the computing time only marginally. This is a key advantage over a modified data-set which includes the mirrored loadings: Then the number of points in the interpolation would be doubled and the computing time would increase by a factor ranging from two to four, since the effort for the multiplication with the (inverse) kernel matrix is $\mathcal{O}(n^2)$ and all other terms are $\mathcal{O}(n)$.

5.5.5 Choice of sampling points and adaption

The question of how to choose the numbers q and n of different amplitudes $\bar{\varepsilon}_i$ and of the directions \mathbf{N}_j^ε , respectively, is important for the overall performance of the method and depends on the microstructure under consideration. The coordinate separation isolates directional information, i.e. information about the anisotropy. This means that

the constitutive response along different N_j^ε can vary considerably. In case that the anisotropy is pronounced, n should be large in order for the RBFs to yield accurate geodesic information, i.e. the sampling directions must be sufficiently dense. If the distribution of the sampling directions is chosen according to Section 5.4.4, the proposed interpolation can be applied as black-box technique (mostly) independent of the anisotropy. Further, the calibration of the kernel parameter γ benefits from the homogeneous sampling directions (see Section 5.6.2). For example, in the numerical investigations of the next section, the number $n = 512$ is chosen as a result of pre-calculations with different values, i.e. the decision-making is experience-driven. Due to the sign symmetry of W discussed in Section 5.5.4, 1024 directions are considered effectively.

The number q of amplitudes $\bar{\varepsilon}_i$ is motivated by the stress–strain curves of the considered constitutive laws (see Section 5.2.2). Since the nonlinearities of the matrix materials occur at low strains, the sampling of the macroscopic strain amplitudes is chosen denser at low values (e.g., $q = 10$ is taken in the examples). Note that the sampling strategy is limited to hyperelasticity, meaning that a thorough adaption will be required if dissipative effects are taken into account. This is postponed to future investigations.

A refinement or coarsening of the sampling directions in the current proposal is not possible in a straight-forward way: Adding directions would imply a loss of the homogeneity of the distribution of the sampling directions which is unfavorable for the accuracy of the underlying kernel method (e.g. in the sense of the partition of unity property). However, it shall be noted that the radial discretization can be chosen individually for each training direction without further modifications.

5.6 Numerical examples

5.6.1 Layout

First composites with spherical inclusions are investigated. The matrix is assumed to be “plastic” without hardening (model 1; $\sigma_c = 100$ MPa; $h = 0$ MPa) and the elastic parameters of aluminium are taken ($E_m = 75$ GPa, $\nu_m = 0.3$). The inclusions are modeled as linear elastic ($E_i = 400$ GPa, $\nu_i = 0.2$). Three different particle volume fractions f are considered: 1, 5 and 10%. For each volume fraction two different geometric realizations are considered: first, a simple RVE with a single centered particle is investigated. Then a near isotropic distribution of $n_P = 20$ particles is considered. All microstructures are periodic. The single particle composites lead to a cubic response and the (near) isotropic microstructures show near isotropic overall behavior. Four of the six different discretizations are shown in Fig. 5.9. The DOFs for the different models are on the order of 200k for the single inclusion models and 1000k, 540k and 300k for the 20 particles and 1, 5 and 10% vol. of the inclusions, respectively.

The same training strategy was pursued for all of the microstructures

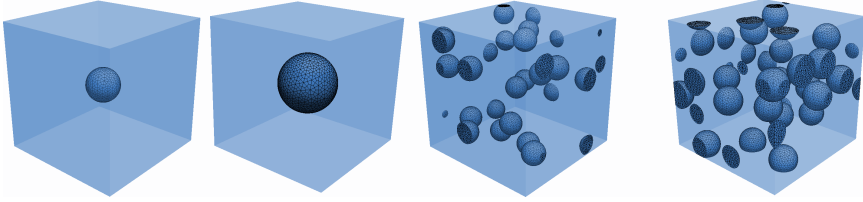


Figure 5.9: RVEs for particle volume fractions 1% and 5% (*leftmost two*; $n_P = 1$), 5% and 10% (*rightmost two*; $n_P = 20$)

- In [L.1] FE simulations along 32 distinct directions (generated cf. Section 5.4.4) with 10 load increments for each leading (cf. Table 5.1) with maximum strain amplitudes of 2.5% are performed. More precisely 320 snapshots are collected for each of the six RVEs.
- A POD (see Appendix 5.8.3) of the snapshots provides a reduced bases with 80 strain fluctuation modes for each RVE.
- In [L.2] the ROM (Section 5.3) is evaluated along 512 directions (same load increments), leading to 5120 samples of \bar{W}_{ROM} and $\bar{\sigma}_{\text{ROM}}$ for each RVE.⁶
- The RNEXP is trained from the 5120 samples of [L.2]. The symmetric kernel function of Section 5.5.4 is used.

The accuracy and performance of the model are assessed for examples of increasing complexity.

5.6.2 Calibration of the kernel width parameter

The only parameter of the RNEXP is $\gamma \in \mathbb{R}_+$ which controls the width of the Gaussian kernel function. It is identified in a two-step procedure: first, a “leave one out” (LOO; e.g. [Rippa, 1999, Hickernell & Hon, 1999]) optimization of γ is performed in order to get a good initial value without additional computations. Then some additional ROM computations are performed in order to find optimal values for directions that could be seen as furthest away from the training data. Therefore, the geodesic distance map of all training directions is computed first. Then k pairs $(\mathbf{N}_{i,1}^\varepsilon, \mathbf{N}_{i,2}^\varepsilon)$ ($i = 1, \dots, k$) with the largest nearest neighbor distances are selected. For each pair the intermediate direction is computed with account for the symmetry of the directions:

$$a = \text{sign}(\mathbf{N}_{i,1}^\varepsilon \cdot \mathbf{N}_{i,2}^\varepsilon), \quad M_i^\varepsilon = \frac{\mathbf{N}_{i,1}^\varepsilon + a \mathbf{N}_{i,2}^\varepsilon}{\|\mathbf{N}_{i,1}^\varepsilon + a \mathbf{N}_{i,2}^\varepsilon\|_2}. \quad (5.76)$$

⁶The ROM computations require few hours (per microstructure; single CPU core) in contrast to almost one full day for the 320 FE snapshots. Since the direct comparison of the ROM to the FE problem is not the focus of the current paper, no detailed benchmarks were done.

The geodesic distance of M_i to its generating directions is exactly half of the distance between these points. In practice M_i can be seen as a tough validation case due its maximum distance to the available data. Along the k new directions M_i additional ROM computations are performed. Then the value of γ is adjusted such that the new data is captured accurately.

The results for the interpolation of W are verified along geodesic lines connecting arbitrary point $(N_\alpha^\varepsilon, N_\beta^\varepsilon)$ pairs. They are constructed for $\theta \in [0; \angle(N_\alpha^\varepsilon, N_\beta^\varepsilon)]$ via

$$N^\varepsilon(\theta; \alpha, \beta) = N_\alpha^\varepsilon \cos(\theta) + Q_{\alpha, \beta} \sin(\theta). \quad (5.77)$$

The construction of the bi-sector M_i^ε and the interpolation $N^\varepsilon(\theta; \alpha, \beta)$ are illustrated in Figure 5.10.

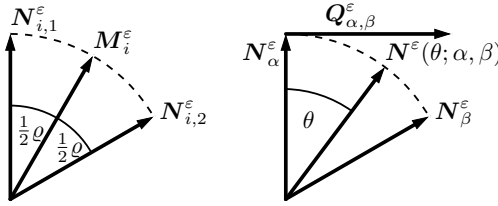


Figure 5.10: Construction of bi-sector M_i^ε ($\rho = \text{acos}(N_{i,1}^\varepsilon \cdot N_{i,2}^\varepsilon)$) and interpolation $N^\varepsilon(\theta; \alpha, \beta)$ of directions $N_\alpha^\varepsilon, N_\beta^\varepsilon$ cf. (5.77)

Along this geodesic line the ROM is used to sample verification data ($W, \bar{\sigma}, \bar{\mathbb{C}}$) that can directly be compared to the prediction of the interpolation scheme. Comparisons of the RNEXP predictions for one of the RVEs presented in Section 5.6.1 ($f = 10\%$, $n_P = 20$ inclusions) are plotted in Fig. 5.11. The energy is predicted to a relative accuracy $\approx 10^{-4}$ and better. Regarding the stress predictions, the relative errors are well below 1% (for the full tensor) and on the order of few percent for the deviatoric stress. The comparison of the deviatoric stress could lack from the bad proportion of hydrostatic vs. deviatoric stress.

In order to confirm that for low hydrostatic stresses the error is not increasing unproportionally, four additional deviatoric load cases that are chosen randomly and differ from the training data are considered. The results are shown in Figure 5.12 for the relative error in the energy and in the effective stress. The findings reported in Figure 5.11 and 5.12 are confirmed for other loadings and for the other microstructures. During these comparisons it was also found that the kernel parameter, which was fixed to $\gamma = 2.0$, has rather little influence on the overall accuracy, which is related to the homogeneous nearest neighbor distance of the sampling sites. The latter emphasizes the importance of the layout of the sampling directions following Section 5.4.4.

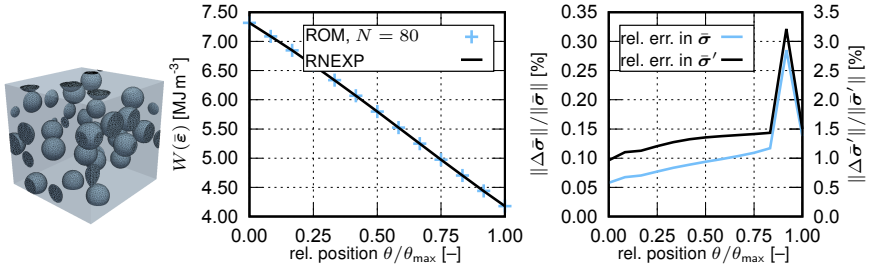


Figure 5.11: Verification of the interpolation along a geodesic line connecting a pair of training points for the inclusion microstructure from Section 5.6.1 with $f = 10\%$, $n_P = 20$; *left*: energy W of RNEXP interpolation versus ROM (for $\bar{\varepsilon} = 2.5\%$); *right*: relative error in $\bar{\sigma}$ and in the deviator of $\bar{\sigma}$ (kernel parameter $\gamma = 2.0$; 512 sampling directions)

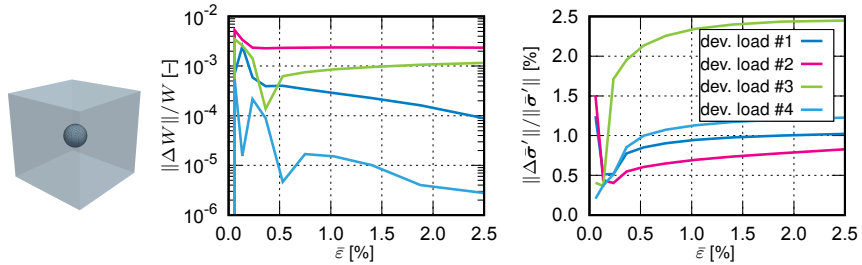


Figure 5.12: Verification of the interpolation for four additional purely deviatoric strain loadings (different from training) ($f = 10\%$, $n_P = 20$); *left*: relative error of the energy W (RNEXP versus ROM, $N = 80$); *right*: relative error in the deviatoric stress

5.6.3 Validation for strain-driven RVE computations

Next, the error of the interpolation scheme with respect to ROM computations ($N = 80$ modes are used throughout this section) performed over a set of 512 validation directions (10 load steps cf. Table 5.1) is evaluated in order to judge on the accuracy of the RNEXP and, thereby, on its suitability for large twoscale simulations. The validation loadings are not randomly chosen, but they are determined using the same procedure as for the training directions albeit with different initialization leading to a set of complementary directions.

The results for microstructures consisting of a single inclusion at 1% inclusion volume fraction are shown in Figure 5.13 as well as the ones for 10% and 20 particles in Figure 5.14. The results state good accuracy and the 99% percentiles indicate less than 1% of all validation cases showed an error in the effective stress prediction of more than $\approx 1\%$ for either microstructure.

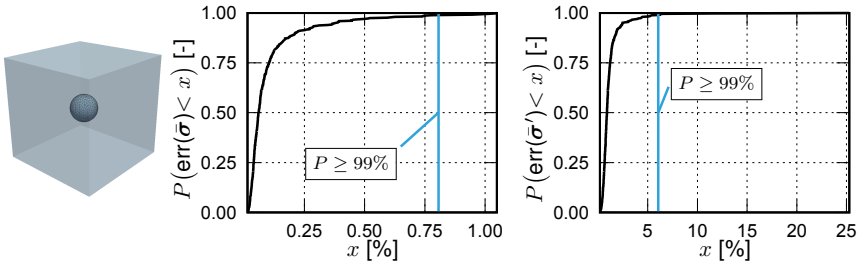


Figure 5.13: Distribution of the error of the RNEXP in the total stress (*left*) and deviatoric stress (*right*); the 99% percentile is also shown (microstructure: $f = 1\%$, $n_P = 1$)

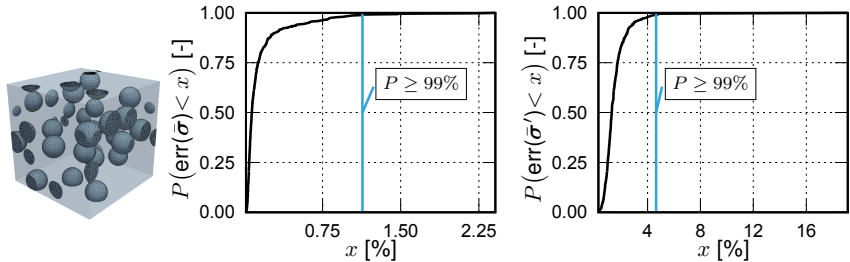


Figure 5.14: Distribution of the error of the RNEXP in the total stress (*left*) and deviatoric stress (*right*); the 99% percentile is also shown (microstructure: $f = 10\%$, $n_P = 20$)

5.6.4 Nonlinear hardening materials with anisotropic microstructure

Besides the previous investigations without hardening, the nonlinear hardening material (cf. model 2 in Section 5.2.2) can also be processed using the RNEXP. The same procedure discussed before was applied to a material characterized by an anisotropic microstructure with high particle concentration within a planar layer (10% particles overall, locally $\approx 40\%$). The FE mesh (169,349 nodes/508,047 DOF; 122,918 TET10 elements) and the FE results of one training simulation are shown in Figure 5.15.

The same FE and ROM computations as in the previous examples were performed. Importantly, the anisotropy was not explicitly considered during training, but the RNEXP was applied as black box procedure. The statistic distribution of the relative stress errors over the 512 validation cases (comparison RNEXP vs. ROM, $N = 80$) is shown in Figure 5.16. The results confirm good accuracy of the RNEXP for the anisotropic material in the presence of nonlinear hardening.

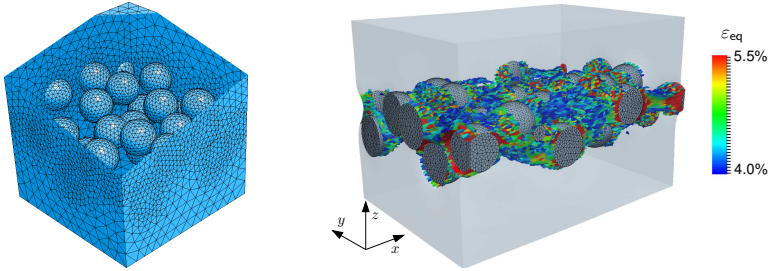


Figure 5.15: Anisotropic microstructure with spherical inclusions concentrated into a layer; left: mesh; right: result of the first FE training simulation (matrix material for $\varepsilon_{\text{eq}} \geq 4\%$ is shown)

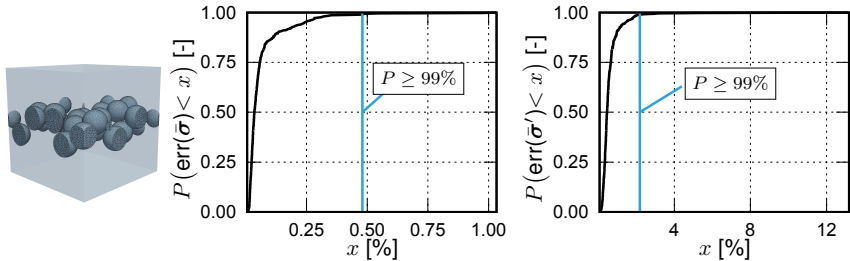


Figure 5.16: Distribution of the error of the RNEXP in the total stress (*left*) and deviatoric stress (*right*); the 99% percentile is also shown (anisotropic particle microstructure)

5.6.5 RNEXP twoscale simulation with data-acceleration

Comparison of multilevel FE using RNEXP and ROM

The previous tests validated the accuracy of the RNEXP predictions for selected loading conditions. Thereby confidence is gained for the application of the RNEXP in actual twoscale computations. In this Section twoscale problems of gradually increasing complexity are explored.

First, a simple macroscopic domain consisting of cuboid with centered spherical pore (10% porosity; see Fig. 5.17, left) is used. At the integration point level the homogenized material behavior of the RVE with single inclusion and 5% inclusion volume fraction was considered (RVE of Figure 5.9 second from left; material model 1 without hardening).

Different approximations of the actual homogenized material response were used:

- reduced order models cf. Section 5.3.1 using different mode sets ($N \in \{16, 32, 80\}$; examples of the modes are shown in Figure 5.17);

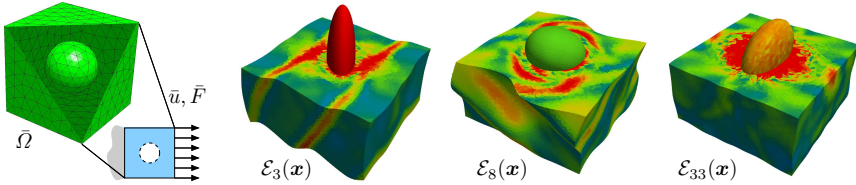
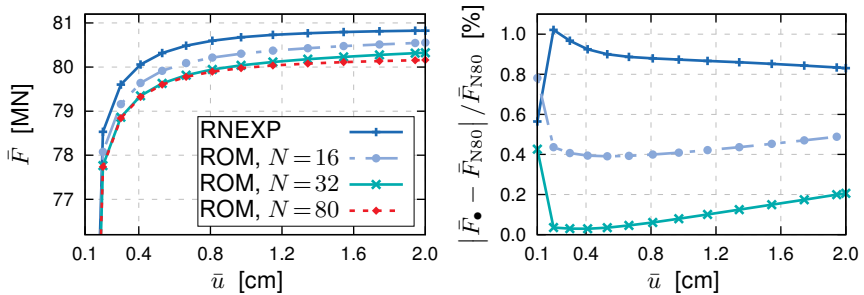


Figure 5.17: Porous macroscopic FE mesh (dimension: $1 \times 1 \times 1 \text{ m}^3$; *left*) and three strain fluctuation fields of the reduced basis for the single-inclusion microstructure (coloring indicates $\|\varepsilon'\|$)



(a) Effective 0.8 force-displacement curves of the different models

(b) Relative error against the ROM with $N = 80$ modes (key as in (a))

Figure 5.18: Results of the uniaxial tension test for the porous microstructure

- RNEXP model obtained from 512 training directions with 10 load steps each and trained using the ROM with $N = 80$ modes (c.f. Section 5.5).

The porous cuboid made of the heterogeneous material was subjected to uniaxial tension with 2% final strain applied along one of the structures' principle axes. A rather coarse macroscopic FE discretization (see Fig. 5.17, left) was chosen (6567 DOF; 1082 quadratic tetrahedra) in order to keep computing times for the largest ROM in a reasonable range. Due to the excessive computational cost, we did not consider FE^2 simulations (i.e. substituting the microscopic material model by FE simulations on RVE level).

The accuracy is investigated through comparison of the effective force-displacement curves (\bar{F} vs. \bar{u}). The results obtained using the ROM with $N = 80$ modes is considered as reference solution. The relative deviations with respect to this reference are shown in Figure 5.18, right. As stated in Section 5.3.1 the increasing dimension of the RB used within the ROM leads to slightly softer overall response. It shall however be noted that even when using only 16 modes, the deviation to the fine scale ROM was well below 1%

during the entire macroscopic simulation. The RNEXP overestimates the force by only 1% and, hence, good accuracy is found.

The slopes of the force-displacement curves are approximated via finite differences in Figure 5.19 using a logarithmic scale. It is found that all models predict the same effective stiffness of the nonlinear problem up to 1% of tensile deformation (corresponding to \bar{u} of 1 cm). Then saturation is observed and all slopes tend towards 0. Here it must be emphasized that the logarithmic scale suggests larger deviations than actually observed.

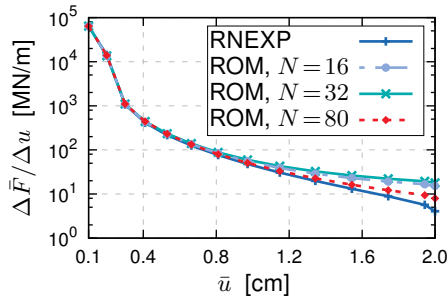


Figure 5.19: Secant stiffness of porous macrostructure with heterogeneous micromaterial

The number of Newton-Raphson iterations was 45 (RNEXP) and 42 or 43 (ROM, $N \in \{16, 32, 80\}$), respectively. This states that the accuracy of the tangent stiffness operator provided by the RNEXP works reliably. Concerning runtimes an estimate of the performance is provided in this work in terms of the relative runtime in non-laboratory conditions. All models are fairly efficiently implemented (i.e. the implementations are ruled out as limitations of the performance). The simulations were run on the same

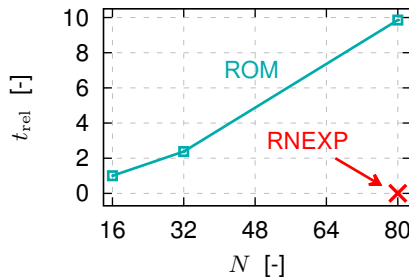


Figure 5.20: Relative overall runtime t_{rel} of the twoscale simulation of the voided macrostructure relative to ROM with $N = 16$ (approximately 3.3 h; for comparison: RNEXP 117 s)

workstation without using parallelization (in order to not deteriorate the benchmark) and

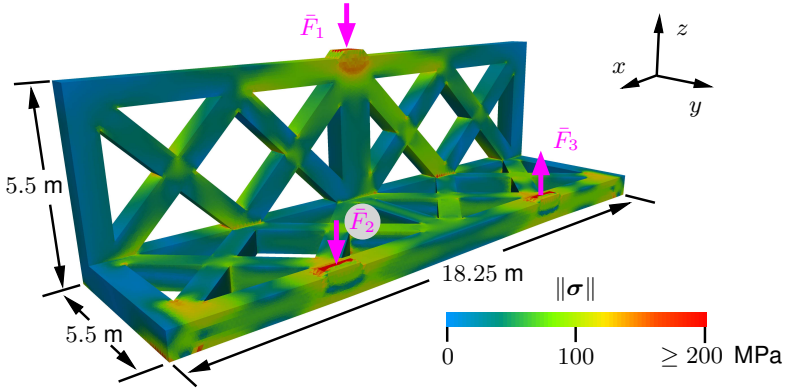


Figure 5.21: Structure for investigation of the influence of the microstructure’s inclusion volume fraction on the reaction forces $\bar{F}_1, \bar{F}_2, \bar{F}_3$ (arrows, pink) recorded at the inhomogeneous Dirichlet boundaries

no computationally intense tasks were executed in parallel⁷. The recorded computation times are normalized with respect to the time of the ROM with $N = 16$ (≈ 3.3 h). The nondimensional solution time denoted by t_{rel} is shown in Figure 5.20. The findings indicate a superlinear scaling of the runtime with respect to the number of modes N used in the ROM ($t_{rel} \sim N^{1.4}$), which is due to the efficient implementation including quasi-Newton methods⁸. This is important as the observed scaling is significantly more “gentle” than the quadratic dependence expected for standard Newton-Raphson procedures where assembly of the Jacobian dominates the overall cost. In summary, the RNEXP computation (computation time: 117 s) could outperform the ROM by factors ranging from 100 ($N = 16$) to 1000 ($N = 80$).

Fine scale twoscale simulation

The second macroscopic structure consists of a complex three-dimensional geometry with a truss-like structure, see Figure 5.21. A fine scale FE discretization involving 376,060 nodes ($\approx 1.1 \cdot 10^6$ DOF) and 234,802 quadratic tetrahedra is used. Homogeneous Dirichlet boundary conditions are prescribed to the bear surfaces extremal w.r.t. the x -direction on the rear side. Inhomogeneous Dirichlet conditions are imposed on the tops of the three transmission spots. The reaction forces $\bar{F}_1, \bar{F}_2, \bar{F}_3$ indicate the direction of the prescribed displacements. The load amplitudes are 3.3 cm, 4.4 cm and 3.3 cm respectively.

⁷Although not relevant, some hardware specifications are provided: 2x Intel Xeon(R) E5-2643 v3; 256 GB RAM

⁸As the solution of the ROM is not the focus of this work details may be presented in later works.

Three different simulations of the structure were performed:

- the material model 1 without hardening defined in Section 5.6.1 (mimicking von Mises plasticity);
- the RNEXP model for the single-inclusion RVE ($f = 5\%$, $n_P = 1$; see Figure 5.9 second from left);
- the RNEXP model for the multi-inclusion RVE ($f = 10\%$, $n_P = 20$; see Figure 5.9 right).

Details on the computations are provided in Table 5.2. Interestingly, the number of Newton-Raphson iterations of the RNEXP are basically identical to the case without inclusions, where the exact material stiffness was provided. This confirms the quality of the tangent operator of the RNEXP.

Table 5.2: Number of required Newton iterations for each increment for the truss structure.

material	# iterations in . . .				
	1 st inc.	2 nd inc.	3 rd inc.	4 th inc.	5 th inc.
homogeneous matrix	4	5	6	6	4
RNEXP, $f = 5\%$	4	5	8	6	4
RNEXP, $f = 10\%$	4	5	7	6	4

A comparison of the forces recorded during the three simulations can help to quantify the effect of particle volume fraction on the nonlinear structural behavior. The results are shown in Figure 5.22. Both the nonlinearity due to the pseudo-plasticity in the microscopic matrix material, and the stiffening due to increasing inclusion volume fraction are captured.

5.7 Résumé

5.7.1 Summary

A two-stage computational homogenization procedure starting from high-fidelity finite element simulations [L.1], a therefrom derived reduced order model [L.2] and the data-driven RNEXP [L.3] is presented. Each computational level has advantages and drawbacks. More specifically high-accuracy is generally associated with increased computing times and vice versa. The different computational schemes on each level are combined by a particular layout of the directions and load amplitudes in order to achieve accurate data (devoid of over- and undersampling) and, hence, predictions of the constitutive response with rather few sampling points. The consideration of symmetries (see Section 5.5.4) states that at (almost) constant computing time, the accuracy of the

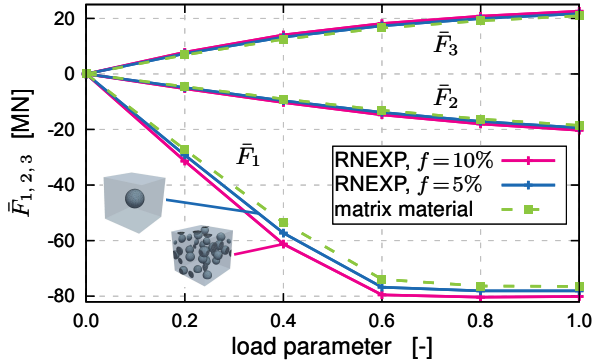


Figure 5.22: Forces versus nondimensional load parameter for the truss structure as a function of the microstructure: no inclusion (green, dashed), $f = 5\%$ and $n_p = 1$ (blue, solid line), $f = 10\%$ and $n_p = 20$ (magenta, solid)

proposal can be improved considerably. In the presented examples 320 FE field solutions (along 32 directions with 10 amplitudes each) yield the ROM for [L.2]. Using the latter, 5120 samples per microstructure are computed. The entire preprocessing was realized on a single workstation within a single day per microstructure without the use of parallelization. The accuracy and the performance of the RNEXP are validated in several stages: First the calibration of the kernel parameter (cf. Section 5.6.2) makes use of loads along geodesic lines connecting different training directions. Next, a comparison to rather extensive sets of validation loadings (cf. Section 5.6.3; here: of the same size as the sampling data) is performed. Third, twoscale simulations are effected using different ROMs and the RNEXP and the outputs are compared. Finally, the RNEXP is used in large scale structural simulations in order to assess the effect of the microstructure in Section 5.6.5 for which twoscale computations using the ROM are no longer practicable.

5.7.2 Discussion

The key innovation of the RNEXP is the combination of different (established) techniques with a particular layout of the sampling sites. The latter is a key ingredient of the overall procedure: It allows for an adaptive grid density and provides samples with nearly homogeneous distribution of the loading direction. In our view this is a major improvement over Monte Carlo sampling (e.g. [Le et al., 2015]) and regular grids (e.g. [Yvonnet et al., 2009] and follow-ups) which do not provide optimal information content per sample.

Further, two different discretizations and solvers (FE and ROM) are used. Each technique is applied as sparsely as possible in order to build reliable foundations for the next level/hierarchy: The accuracy of the ROM in [L.2] depends on the richness of the

training data (gathered in [L.1]). Since the ROM contains information of the geometry and of the microscopic materials on the microstructural level, a moderate amount of FE simulations suffices to feed the ROM. Again, the homogeneous distribution of the training directions is beneficial and leads to few training computations. Analogously, the RNEXP predictions in [L.3] require samples of the energy and radial stress provided by means of the ROM [L.2]. Since the RNEXP has no physics built-in, a larger number of scalar inputs is required than in order to set up the ROM. Note that the use of the radial slope of the stored energy helps to gain accuracy at no additional computational cost and without sacrificing the symmetry of the tangent stiffness. On a side note, the adjustment of the radial cubic interpolation functions is straight-forward and the radial interpolation is unique without the need for further (algebraic) constraints.

Lastly, the kernel approximation used for the directional interpolation of the RNEXP benefits from the homogeneous distribution of the directions: Overshooting and oscillations are prevented, the solutions are not too sensible with respect to the kernel parameter γ , and the kernel interpolation satisfies the partition of unity condition (cf. (5.48) on p. 64) almost exactly. The applied radial-tangential (or amplitude–direction) split used in the interpolation appears somehow intuitive given the origin of the training data, although the choice is of course not unique and could be altered in the future.

The authors would like to point out that a machine learning related method using reduced approximations was recently realized in terms of the MNROM of Bhattacharjee & Matouš [2016]. Some ideas are similar to the present technique: a dedicated sampling strategy in strain space employing the polar decomposition is used to set up training simulations for moderately nonlinear hyperelastic solids in a finite strain framework. The strain parametrization used in [Bhattacharjee & Matouš, 2016] can be considered a variant of the method used herein with a preceding partitioning of the parameter/strain space into four “modes”. The MNROM uses neural networks and manifold learning, i.e. the approach is strongly data-oriented. In direct comparison to the MNROM the RNEXP has advantages due to the lower algorithmic complexity (CPU time, memory) and the availability of a symmetric algorithmic tangent stiffness. Ideas of combining concepts of the MNROM with the RNEXP could potentially be investigated, e.g. during the sampling [L.2].

5.7.3 Perspective

The present study focuses on small strain (pseudo-)hyperelasticity. We would like to emphasize that the methodology can be generalized in several ways. Within the field of small strain mechanics, any rate-independent material can be mimicked pseudo-hyperelastically as long as there is next to proportional loading or if the path-dependency is limited. This holds true for many structures undergoing simple load cases that are relevant for practical use. The adaption to large strain hyperelasticity is subject to ongoing research and has already yielded promising results. There, the choice of a measure of deformation or strain has crucial influence on the particular form and the

performance of computational levels [L.2], [L.3], and of the sampling strategy of section 5.4.4.

As more complex materials often show path-dependency, it is desirable to account for this property in the present framework. However, the sampling strategy essentially relies on the fact that the state space is rather low-dimensional (here: 6-dimensional). The space of all possible macroscopic load trajectories has infinite dimensions and can thus not be sampled directly. Any sampling of loading histories would require restriction to some finite-dimensional subspace. Even with a rather coarse resolution of the trajectories, the dimensionality of the remaining quantities would increase the computational demands, quickly exceeding today's compute capabilities. It is due to this obstacle that we cannot provide a solution to the problem of path-dependency as of today.

Moreover, there remain open questions concerning the sampling strategy even in the path- and rate-independent case. One of which is about the assessment of the quality of a chosen point set, for both the training of the ROM in [L.2] and for the setup of the interpolation method in [L.3]. For both of these stages, error estimators could be one way to gain information about the suitability of the sampling points. For [L.2], a posteriori error estimators are subject to current research. There, it can be exploited that both the ROM and the FEM are Galerkin methods, which does not hold for [L.3]. Since the latter is a purely mathematical model without any directly incorporated physics (e.g. variational principle), error estimation requires a different approach for [L.3]. Although some simplistic a posteriori tests were carried out in Section 5.6.2, this topic cannot be regarded as sufficiently covered yet.

The RNEXP method is not restricted to mechanics but can be adapted to other kinds of elliptic problems in a straight-forward way. To give just one example, nonlinear thermal conduction problems could be treated analogously in [L.2] by identifying a RB for the space of temperature gradients and in [L.3] by interpolating between samples of this three-dimensional space. The proposed sampling strategy is general with respect to the spatial dimension, and the radial resolution can be adapted to suit any constitutive law. This emphasizes the generality of the chosen approach.

Last, the radial tangential split could as well be used for data of arbitrary dimension that is provided along radial sampling paths. Notably the selection of the directions conforming to section 5.4.4 is essential as it contributes to the efficiency and accuracy of the radial basis function interpolation.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

The use of the kernel method was inspired by discussions with Prof. Bernard Haasdonk in the scope of the scientific network CoSiMOR funded by the German Research Foundation (Deutsche Forschungsgemeinschaft) under grant FR2702/4, /7. Further, the authors gratefully acknowledge the financial support of this research through grant DFG-FR2702/6 of the German Research Foundation (Deutsche Forschungsgemeinschaft) within the scope of the Emmy-Noether program of DFG. The authors acknowledge the input gained in valuable discussions within the scope of the DFG Cluster of Excellence Simulation Technology (DFG EXC 310 SimTech).

5.8 Appendix

5.8.1 A – Hyperelastic potentials and their gradients

Deformation plasticity with linear hardening

We recall the energy of the model similar to the one used, e.g., by Bilger et al. [2005] that mimics plasticity with linear hardening

$$\begin{aligned} \psi'_e(\boldsymbol{\varepsilon}) &= G \|\boldsymbol{\varepsilon}'\|^2, \\ \psi'_{1,p}(\boldsymbol{\varepsilon}) &= \sqrt{\frac{2}{3}} \sigma_c \left(\|\boldsymbol{\varepsilon}'\| - \frac{\varepsilon_c}{2} \right) + \frac{Gh}{3G+h} \left(\|\boldsymbol{\varepsilon}'\| - \varepsilon_c \right)^2, \end{aligned} \quad (5.78)$$

$$\psi_1(\boldsymbol{\varepsilon}) = \frac{K}{2} (\mathbf{I} \cdot \boldsymbol{\varepsilon})^2 + \begin{cases} \psi'_e(\boldsymbol{\varepsilon}) & \|\boldsymbol{\varepsilon}'\| \leq \varepsilon_c, \\ \psi'_{1,p}(\boldsymbol{\varepsilon}) & \|\boldsymbol{\varepsilon}'\| > \varepsilon_c, \end{cases}, \quad \text{with } \varepsilon_c = \sqrt{\frac{2}{3}} \frac{\sigma_c}{2G}. \quad (5.79)$$

The parameters of the model are the bulk modulus K , the shear modulus G , the linear hardening modulus h and the initial yield stress σ_c . In the case of $\|\boldsymbol{\varepsilon}'\| \geq \varepsilon_c$ the apparent (not physical) plastic strain in the model is (with $\langle \bullet \rangle_+ = \max\{0, \bullet\}$)

$$\boldsymbol{\varepsilon}^p = E_* \langle \|\boldsymbol{\varepsilon}'\| - \varepsilon_c \rangle_+ \mathbf{N}', \quad E_* = \frac{3G}{3G+h}, \quad \mathbf{N}' = \frac{\boldsymbol{\varepsilon}'}{\|\boldsymbol{\varepsilon}'\|}. \quad (5.80)$$

The abbreviation E_* comes in handy and allows to represent the hardening stress σ_h and the stress tensor in a rather straight-forward notation

$$\sigma_h = \sigma_y - \sigma_c = \sqrt{\frac{2}{3}} h \|\boldsymbol{\varepsilon}^p\| = \sqrt{\frac{2}{3}} h E_* \left(\|\boldsymbol{\varepsilon}'\| - \varepsilon_c \right) \quad (5.81)$$

$$\boldsymbol{\sigma} = \frac{\partial \psi_1}{\partial \boldsymbol{\varepsilon}} = K (\mathbf{I} \cdot \boldsymbol{\varepsilon}) \mathbf{I} + \begin{cases} 2G \boldsymbol{\varepsilon}' & \|\boldsymbol{\varepsilon}'\| \leq \varepsilon_c \\ \left[\sqrt{\frac{2}{3}} \sigma_c + \frac{2}{3} E_* h \left(\|\boldsymbol{\varepsilon}'\| - \varepsilon_c \right) \right] \frac{\boldsymbol{\varepsilon}'}{\|\boldsymbol{\varepsilon}'\|} & \|\boldsymbol{\varepsilon}'\| > \varepsilon_c. \end{cases} \quad (5.82)$$

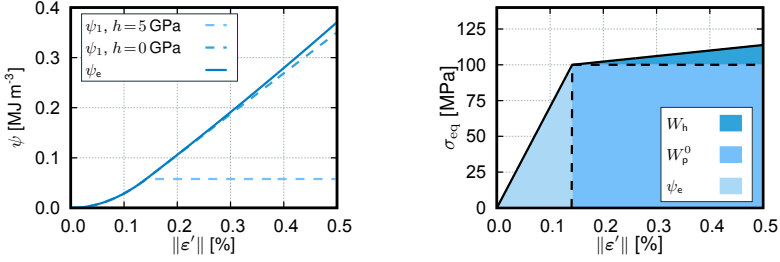


Figure 5.23: von Mises stress vs. deviatoric strain and the corresponding energy split into three parts: elastic energy (up to σ_c) ψ_e , plastic work without hardening W_p^0 and work due to $h = 5$ GPa (W_h)

The second gradient of the free energy provides the stiffness of the material. At sub-critical strains (i.e. for $\|\epsilon'\| \leq \epsilon_c$) the isotropic elastic stiffness tensor

$$\mathbb{C}_0 = K \mathbf{I} \otimes \mathbf{I} + 2G \left(\mathbb{I}^s - \frac{1}{3} \mathbf{I} \otimes \mathbf{I} \right) = 3K \mathbb{P}_1 + 2G \mathbb{P}_2 \quad (5.83)$$

is found. At strains leading to states by-passing the (pseudo-)yield stress σ_c , the stiffness is

$$\mathbb{C}_c = 3K \mathbb{P}_1 + \frac{2}{3} h E_* \mathbb{P}_2 + \sqrt{\frac{2}{3}} \frac{E_* \sigma_0}{\|\epsilon'\|} (\mathbb{P}_2 - \mathbf{N}' \otimes \mathbf{N}'), \quad (5.84)$$

where the abbreviation E_* and the relation between σ_c and ϵ_c are incorporated. Note that \mathbb{C}_c is not positive definite for non-hardening materials, i.e. for $h = 0$. The first part contributes the volumetric stiffness, the second is due to the hardening while the last part can be interpreted as a secant stiffness for all strain changes that are orthogonal to the current strain. It states that any change of ϵ' perpendicular to the current direction will only affect the direction of the stress but not its amplitude which is in full accordance with the proposed model.

In summary, the stiffness of ψ_1 is

$$\mathbb{C} = \frac{\partial^2 \psi_1}{\partial \epsilon \partial \epsilon} = \begin{cases} \mathbb{C}_0 & \|\epsilon'\| \leq \epsilon_c, \\ \mathbb{C}_c & \|\epsilon'\| > \epsilon_c. \end{cases} \quad (5.85)$$

A graphical representation of the energy contributions is given in Figure 5.23: By ψ_e the energy due to the initial elastic deformation is expressed, i.e. before reaching the yield stress. Then the energy W_p^0 is the plastic work due to constant yield stress σ_c and W_h is the energy due to the linear hardening.

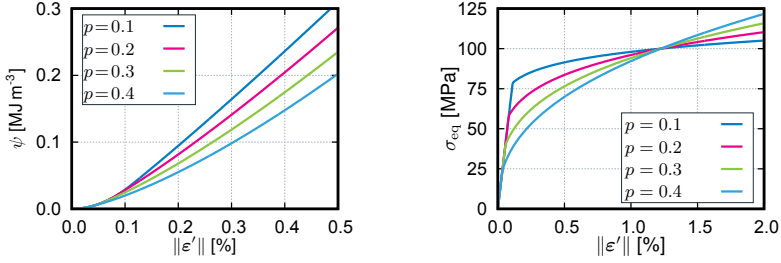


Figure 5.24: Comparison of the free energy $\psi(\boldsymbol{\varepsilon})$ for $\sigma_0 = 100$ MPa, $\varepsilon_0 = 0.01$ and $p \in \{0.1, 0.2, 0.3, 0.4\}$

Nonlinear hardening model with linear elastic domain

The second model under investigation is similar to the nonlinear model employed by Yvonnet et al. [2009]. It has a pronounced nonlinear (hardening-like) behavior. In order to prevent unphysical behavior at small strains and to preclude non-differentiability at $\boldsymbol{\varepsilon} = \mathbf{0}$, the model has been modified in the sense that a linear elastic domain has been added. The free energy of the model is

$$\varepsilon_c = \varepsilon_0 \left(\frac{\sigma_0}{3G\varepsilon_0} \right)^{\frac{1}{1-p}}, \quad \varepsilon_{\text{eq}} = \sqrt{\frac{2}{3}} \|\boldsymbol{\varepsilon}'\|, \quad (5.86)$$

$$\psi'_{2,p}(\boldsymbol{\varepsilon}) = \frac{\sigma_0 \varepsilon_0}{p+1} \left[\left(\frac{\varepsilon_{\text{eq}}}{\varepsilon_0} \right)^{1+p} - \left(\frac{\varepsilon_c}{\varepsilon_0} \right)^{1+p} \right] + \frac{3G}{2} \varepsilon_c^2, \quad (5.87)$$

$$\psi_2(\boldsymbol{\varepsilon}) = \frac{K}{2} (\mathbf{I} \cdot \boldsymbol{\varepsilon})^2 + \begin{cases} \psi'_e(\boldsymbol{\varepsilon}) & \varepsilon_{\text{eq}} \leq \varepsilon_c, \\ \psi'_{2,p}(\boldsymbol{\varepsilon}) & \varepsilon_{\text{eq}} > \varepsilon_c. \end{cases} \quad (5.88)$$

Parameters of the model are the isotropic elastic moduli K and G , the critical stress σ_0 and exponent $p \in (0, 1]$. The stress in the elastic domain (i.e. for strains satisfying $\varepsilon_{\text{eq}} \leq \varepsilon_c$) is linear in the strain and coincides with the previous model. In the nonlinear regime the stress and the stiffness are

$$\boldsymbol{\sigma} = K (\mathbf{I} \cdot \boldsymbol{\varepsilon}) \mathbf{I} + \sqrt{\frac{2}{3}} \sigma_0 \left(\frac{\varepsilon_{\text{eq}}}{\varepsilon_0} \right)^p \mathbf{N}', \quad (5.89)$$

$$\mathbb{C} = 3K \mathbb{P}_1 + \frac{2}{3} \frac{\sigma_0}{\varepsilon_{\text{eq}}} \left(\frac{\varepsilon_{\text{eq}}}{\varepsilon_0} \right)^p [\mathbb{P}_2 + (p-1) \mathbf{N}' \otimes \mathbf{N}']. \quad (5.90)$$

Example free energy and deviatoric stress-strain curves are shown in Figure 5.24 for $\sigma_0 = 100$ MPa, $\varepsilon_0 = 0.01$ and different values of the exponent p (elastic parameters: $E = 75$ GPa, $\nu = 0.3$).

5.8.2 B – Algorithm: Galerkin ROM for hyperelasticity

Algorithm 5.1: Galerkin Reduced Order Model for hyperelasticity

Input : macroscopic strain $\bar{\underline{\underline{\epsilon}}} \leftrightarrow \bar{\underline{\underline{\xi}}}$;

Reduced Basis at integration points $\underline{\underline{\mathbf{x}}}_j: \underline{\underline{E}}^{(j)}$;

normalized weights $\tilde{w}_j = w_j/V$ ($V = \sum_{j=1}^{n_{\text{gp}}} w_j$);

tolerance TOL;

[OPTIONAL] initial guess $\underline{\underline{\xi}}_0$

Output: effective energy \tilde{W} ;

effective stress $\bar{\underline{\underline{\sigma}}}$;

[OPTIONAL] algorithmic stiffness $\bar{\underline{\underline{C}}}_A$

1 initialize $\underline{\underline{\xi}} = \underline{\underline{0}}$ or $\underline{\underline{\xi}} = \underline{\underline{\xi}}_0$; set convergence flag to FALSE

2 **while** convergence flag = FALSE **do**

3 set $\underline{\underline{f}} = \underline{\underline{0}} \in \mathbb{R}^N$, $\underline{\underline{J}} = \underline{\underline{0}} \in \mathbb{R}^{N \times N}$

4 **for** $j = 1, \dots, n_{\text{gp}}$ **do**

 // compute local strain

5 $\underline{\underline{\epsilon}}^{(j)} \leftarrow \bar{\underline{\underline{\epsilon}}} + \underline{\underline{E}}^{(j)} \underline{\underline{\xi}}$

 // compute material response

6 $\psi^{(j)} \leftarrow \psi(\underline{\underline{\mathbf{x}}}_j, \underline{\underline{\epsilon}}^{(j)})$; $\underline{\underline{\sigma}}^{(j)} \leftarrow \partial_{\underline{\underline{\epsilon}}} \psi(\underline{\underline{\mathbf{x}}}_j, \underline{\underline{\epsilon}}^{(j)})$; $\underline{\underline{C}}^{(j)} \leftarrow \partial_{\underline{\underline{\epsilon}} \underline{\underline{\epsilon}}}^2 \psi(\underline{\underline{\mathbf{x}}}_j, \underline{\underline{\epsilon}}^{(j)})$

 // compute residual and Jacobian

7 $\underline{\underline{f}} \leftarrow \underline{\underline{f}} + \tilde{w}_j \underline{\underline{E}}^{(j)T} \underline{\underline{\sigma}}^{(j)}$; $\underline{\underline{J}} \leftarrow \underline{\underline{J}} + \tilde{w}_j \underline{\underline{E}}^{(j)T} \underline{\underline{C}}^{(j)} \underline{\underline{E}}^{(j)}$

8 **end**

9 **if** $\|\underline{\underline{f}}\| < \text{TOL}$ **then**

10 set convergence flag to TRUE

11 **else**

12 update $\underline{\underline{\xi}} \leftarrow \underline{\underline{\xi}} - \underline{\underline{J}}^{-1} \underline{\underline{f}}$

13 **end**

14 **end**

15 set $\tilde{W} = 0$, $\bar{\underline{\underline{\sigma}}} = \underline{\underline{0}} \in \mathbb{R}^6$, $\bar{\underline{\underline{C}}}_A = \underline{\underline{0}} \in \mathbb{R}^{6 \times 6}$, $\underline{\underline{\Sigma}} = \underline{\underline{0}} \in \mathbb{R}^{N \times 6}$

16 **for** $j = 1, \dots, n_{\text{gp}}$ **do**

17 $\tilde{W} \leftarrow \tilde{W} + \tilde{w}_j \psi^{(j)}$; $\bar{\underline{\underline{\sigma}}} \leftarrow \bar{\underline{\underline{\sigma}}} + \tilde{w}_j \underline{\underline{\sigma}}^{(j)}$; $\bar{\underline{\underline{C}}}_A \leftarrow \bar{\underline{\underline{C}}}_A + \tilde{w}_j \underline{\underline{C}}^{(j)}$;

$\underline{\underline{\Sigma}} \leftarrow \underline{\underline{\Sigma}} + \tilde{w}_j \underline{\underline{E}}^{(j)T} \underline{\underline{C}}^{(j)}$

18 **end**

19 $\bar{\underline{\underline{C}}}_A \leftarrow \bar{\underline{\underline{C}}}_A - \underline{\underline{\Sigma}}^T \underline{\underline{J}}^{-1} \underline{\underline{\Sigma}}$

5.8.3 C – Algorithm: snapshot POD for hyperelasticity

Algorithm 5.2: Snapshot POD for the generation of an RB for the ROM of Appendix 5.8.2

Input : snapshot data:
 $\underline{\underline{\varepsilon}}^{(k)}$ and strain field $\underline{\underline{\varepsilon}}^{(k)}(\mathbf{x}_j)$ ($k \in \{1, \dots, n_s\}$: snapshot index);
normalized weights $\tilde{w}_j = w_j/V$ ($V = \sum_{j=1}^{n_{gp}} w_j$);
POD threshold δ or number of modes N ;

Output: reduced basis $\underline{\underline{E}}(\mathbf{x}_j)$

```
// compute correlation matrix
1 for  $i, j = 1, \dots, n_s$  do
2 |  $S_{ij} \leftarrow S_{ij} + \sum_{k=1}^{n_{gp}} \tilde{w}_k \underline{\underline{\varepsilon}}^{(i)}(\mathbf{x}_k) \cdot \underline{\underline{\varepsilon}}^{(j)}(\mathbf{x}_k)$ 
3 end
4 compute eigen-decomposition  $\underline{\underline{S}} = \underline{\underline{U}} \underline{\underline{\Lambda}} \underline{\underline{U}}^T$  (diagonal  $\underline{\underline{\Lambda}}$  with decreasing entries)
5 if threshold  $\delta > 0$  is given then
6 | find  $N$  such that  $\sum_{i=1}^N \Lambda_{ii} > (1 - \delta^2) \sum_{i=1}^{n_s} \Lambda_{ii}$ 
7 end
8 truncate  $\underline{\underline{U}}$  to the first  $N$  columns
9 set  $\underline{\underline{E}}^{(j)} = \underline{\underline{0}} \in \mathbb{R}^{6 \times N}$  ( $j = 1, \dots, n_{gp}$ )
10 for  $i = 1, \dots, N$  and  $j = 1, \dots, n_{gp}$  do
11 |  $\underline{\underline{E}}_{\bullet i}^{(j)} = \sum_{k=1}^{n_s} U_{ki} \underline{\underline{\varepsilon}}^{(k)}(\mathbf{x}_j)$ 
12 end
```

5.8.4 D – Algorithm: RNEXP

Initialization of the kernel method

- Given n directions \mathbf{N}_i^ε ($i = 1, \dots, n$) compute the kernel matrix $\underline{\underline{K}}$ and its inverse $\underline{\underline{K}}^{-1}$ via

$$K_{ij} = K_{ji} = \zeta \left(\text{acos}(\mathbf{N}_i^\varepsilon \cdot \mathbf{N}_j^\varepsilon) \right). \quad (5.91)$$

- For given data D_{ij} , T_{ij} , cf. equations (5.43), (5.44), define the piecewise cubic interpolation functions $\mathcal{S}_i(\bar{\varepsilon})$ through conditions defined via (5.46).

Computation of stress and stiffness

Algorithm 5.3: Interpolation of the stress $\bar{\sigma}$ and of the tangent stiffness $\underline{\underline{C}}_{\underline{\underline{A}}}$

Input : strain $\bar{\varepsilon}$;

kernel matrix $\underline{\underline{K}}$;

$\underline{\underline{N}} \in \mathbb{R}^{6 \times n}$ (columns correspond to the n training directions);

Output: effective stress $\bar{\sigma}$;

[OPTIONAL] effective stiffness $\underline{\underline{C}}_{\underline{\underline{A}}}$

- 1 decompose $\bar{\varepsilon}$ into amplitude $\bar{\varepsilon}$ and direction $\underline{\underline{N}}^\varepsilon \leftrightarrow \underline{\underline{N}}^\varepsilon$
 - 2 evaluate $\underline{\theta} = \underline{\underline{N}}^\top \underline{\underline{N}}^\varepsilon$; set $\xi_i = \text{acos}(\theta_i)$ // compute geodesic distance
 - 3 compute $\underline{\underline{S}} \leftarrow \underline{\underline{S}}(\bar{\varepsilon})$, $\underline{\underline{S}}' \leftarrow \underline{\underline{S}}'(\bar{\varepsilon})$, $\zeta_i \leftarrow \zeta(\xi_i)$, $\zeta'_i \leftarrow \zeta'(\xi_i)$ // helpers
 - 4 compute $\underline{w} = \underline{\underline{K}}^{-1} \underline{\zeta}$, $\underline{y} = \underline{\underline{K}}^{-1} \underline{\underline{S}}$ // weights and helper vector
 - 5 set $\sigma_0 = \underline{w}^\top \underline{\underline{S}}'$, $\underline{\tau} = -\frac{1}{\bar{\varepsilon}} \underline{y} \odot \underline{\zeta}'$ // normal and tangential stress components
 - 6 evaluate $\tilde{\sigma}_0 = \sigma_0 - \sum_{i=1}^n \frac{\theta_i}{\sin(\xi_i)} \tau_i$ and $\tilde{\tau}_i = \frac{\tau_i}{\sin(\xi_i)}$ // substitute $\underline{\underline{Q}}_i^\varepsilon$
 - 7 compute stress via $\bar{\sigma} = \tilde{\sigma}_0 \underline{\underline{N}}^\varepsilon + \underline{\underline{N}} \tilde{\tau}$ // (5.66)
// optional: computation of effective stiffness
 - 8 compute $\zeta''_i \leftarrow \zeta''(\xi_i)$, $\underline{\underline{S}}'' \leftarrow \underline{\underline{S}}''(\bar{\varepsilon})$, $\underline{z} = \underline{\underline{K}}^{-1} \underline{\underline{S}}'$ // helpers
 - 9 set $\alpha_i = \frac{\theta_i}{\sin(\xi_i)} \frac{y_i \zeta'_i}{\bar{\varepsilon}^2}$, $\bar{\alpha} = \frac{1}{\bar{\varepsilon}} \underline{w}^\top \underline{\underline{S}}' + \sum_{i=1}^n \alpha_i$, $\bar{\mu} = \underline{w}^\top \underline{\underline{S}}'' - \bar{\alpha}$,
 $\beta_i = \frac{1}{\bar{\varepsilon}^2} y_i \zeta''_i - \alpha_i$
 - 10 determine direction $\bar{\underline{\underline{Q}}}^\varepsilon = \frac{1}{\bar{\varepsilon}^2} \sum_{i=1}^n (y_i - \bar{\varepsilon} z_i) \zeta'_i \underline{\underline{Q}}_i^\varepsilon$ // (5.71)
 - 11 compute $\underline{\underline{C}}_{\underline{\underline{A}}} = \bar{\alpha} \underline{\underline{I}} + \bar{\mu} \underline{\underline{N}}^\varepsilon \underline{\underline{N}}^{\varepsilon\top} + 2 \text{sym}(\bar{\underline{\underline{Q}}}^\varepsilon \underline{\underline{N}}^{\varepsilon\top}) + \sum_{i=1}^n \beta_i \underline{\underline{Q}}_i^\varepsilon \underline{\underline{Q}}_i^{\varepsilon\top}$ // (5.68)
-

Chapter 6:

Second publication – Finite Strain Homogenization Using a Reduced Basis and Efficient Sampling

Original publication:

Kunc, O. & Fritzen, F.: Finite strain homogenization using a reduced basis and efficient sampling. *Mathematical and Computational Applications* **24** (2019b), MDPI, doi:10.3390/mca24020056

Abstract. The computational homogenization of hyperelastic solids in the geometrically nonlinear context has yet to be treated with sufficient efficiency in order to allow for real-world applications in true multiscale settings. This problem is addressed by a problem-specific surrogate model founded on a reduced basis approximation of the deformation gradient on the microscale. The setup phase is based upon a snapshot POD on deformation gradient fluctuations, in contrast to the widespread displacement-based approach. In order to reduce the computational offline costs, the space of relevant macroscopic stretch tensors is sampled efficiently by employing the Hencky strain. Numerical results show speed-up factors in the order of 5–100 and significantly improved robustness while retaining good accuracy. An open-source demonstrator tool with 50 lines of code emphasizes the simplicity and efficiency of the method.

Keywords: computational homogenization; large strain; finite deformation; geometric nonlinearity; reduced basis; reduced-order model; sampling; Hencky strain

MSC: 74Q05, 74B20, 74S30

6.1 Introduction

6.1.1 Purpose

The description of solid mechanics under finite strains is of particular interest in both academia and industry. It allows for accurate descriptions of rotations and stretches under mild assumptions. Thus, many geometric effects can be captured. For instance,

alignments and rearrangements of the respective structures may trigger pronounced stiffening or softening effects.

In such cases where rotations and deformations are not suitable for linearization, dissipative effects also play a notable role for many materials. Regardless of the kind of dissipation involved in a certain process, hyperelasticity usually persists to a certain extent. Therefore, it is worthwhile investigating this comparatively simple case at first, before introducing history dependence into the description. Prominent examples of materials that require a hyperelastic description at finite strains include carbon black-filled rubber [Rendek & Lion, 2010] and amorphous glassy polymers [Nguyen et al., 2016], to name just two.

The main purpose of this work is the *computationally efficient quasi-static homogenization of hyperelastic solids with full account for geometric nonlinearities*. The employed methodology is twofold. First, a Reduced Basis (RB) model for the microscopic problem is established. The term Reduced Basis, used in this work, is not to be confused with the homonymous method introduced by Barrault, Maday, Nguyen, and Patera Barrault et al. [2004]. Once set up, it enables more efficient evaluations of the homogenized material response as compared to the Finite Element Method (FEM). Second, an efficient strategy for sampling of the space of macroscopic kinematic states is proposed. This renders the setup phase of the RB model more rational.

6.1.2 State of the Art

Efficiently determining the overall solid–mechanical properties of microstructures has been investigated for decades, and a large body of literature is available. Comprehensive review articles, such as Geers & Yvonnet [2016] and Saeb et al. [2016], summarize the progress. Here, attention is restrained to few methods most similar or relevant to the present work.

The FE^2 method Feyel [1999] is theoretically capable of performing realistic two-scale simulations with arbitrary accuracy. Therefore, it serves as a reference method in the context of first-order homogenization based on the assumption of separated length scales. In the FE^2 , the evaluations of the unknown macroscopic constitutive law are approximated by microscopic FE simulations. However, this comes along with computational costs that quickly exceed the capabilities of common workstations, both at present and in the foreseeable future. Roughly speaking, the computational effort required on the microscale multiplies with that of the macroscale, hence the method’s name. It is thus worthwhile to develop order reduction methods for the microscopic problem.

A common approach within the field of computational homogenization (and well beyond) is to extract essential information from provided in silico data. To this end, schemes based on the Proper Orthogonal Decomposition (POD) compute correlations within snapshot data, [Sirovich, 1987]. Such methods include the R3M Yvonnet & He [2007] and can be further enhanced by the use of, e.g., the EIM, as in Radermacher & Reese [2015].

Numerical comparisons of various schemes were conducted in Radermacher & Reese [2013], Soldner et al. [2017]. To the best of the authors' knowledge, all published POD-based methods addressing the finite strain hyperelastic problem choose to reduce the number of degrees of freedom (DOF) of the displacement field. This results in sometimes significant speed-ups. Another important feature is that they allow for reconstruction of the microscopic displacement fields. The application of the snapshot POD to *gradients* of the primal variables has been studied—e.g., for infinitesimal strain hyperelasticity Fritzen & Kunc [2018c] and fluid mechanics Akkari et al. [2019]—but does not appear to have been investigated for finite strain hyperelasticity yet.

Still, the solution of the reduced equations remains a complex task. It requires evaluations of material laws and numerical integration over the microstructure. Promising progress has been made in the field of efficient integration schemes, see for instance An et al. [2008], Hernández et al. [2017]. A main reason for the speed-up of these methods is the reduced number of function evaluations.

The highest speed-ups are achievable if the computational effort of the determination of effective microstructural responses can be fully decoupled from underlying microstructural discretizations. Such homogenization methods directly approximate the effective material law by means of a dedicated numerical scheme. Technically, this can be seen as the direct surrogation of unknown functions, e.g., of the effective free energy or stress. For instance, the Material Map Temizer & Zohdi [2007] interpolates the coefficients of an assumed macroscopic material model. Another example is the NEXP method Yvonnet et al. [2013], where the effective stored energy density is approximated using a tensor product of one-dimensional splines. The authors treated the case of small strains by introducing the RNEXP method Fritzen & Kunc [2018c], where the effective stored energy is interpolated by a dedicated kernel scheme.

However, interpolatory and regression methods suffer the inherent drawback of not providing any explicit information on the microscale. For instance, microscopic displacement or stress fields cannot be reconstructed from the solutions of macroscopic interpolation. Another important open question is how to provide the supporting data points for the interpolation in an efficient manner. The data at these points is usually provided by the solution of a full-order model (FOM) and come along with the corresponding numerical costs. Hence, the positions of data points in the parameter space should be chosen carefully, as unnecessary or redundant solutions of the FOM should be avoided. On the other hand, too sparsely seeded points might not capture the homogenized properties of the microstructure appropriately.

6.1.3 Main Contributions and Outline

The present work generalizes parts of the previous paper Fritzen & Kunc [2018c] to the finite strain regime. It aims at reducing the computational complexity for the determination of the homogenized microstructural response, which is parametrized by the macroscopic deformation gradient acting as a boundary condition. This is achieved

by means of a *Reduced Basis approximation of the microscopic deformation gradient*. The basis is obtained with the aid of a *POD of snapshots of fluctuation fields of the deformation gradient*. Thus, the application of the RB model *does not necessitate the computation of gradients of displacement fields*, and even does not require the displacements to be available at all. In other words, microscopic displacement fields are completely avoided. However, they *can be reconstructed* from the RB approximation of the deformation gradient, uniquely up to rigid body motion.

Another key advantage is the *sleek implementation* of the method. A *demonstration* containing a minimum working example of the RB model with 50 lines of MATLAB/Octave code is provided, [Kunc, 2019].

As for the setup phase, the snapshot data is created by means of an *efficient sampling procedure* for the microscopic boundary condition. To this end, the set of macroscopic *Hencky strains* is identified as a suitable *linear parameter space*, within which the sampling sites are placed based upon *physical interpretation*. This allows for *control* of the resolution of certain key characteristics of the effective material response while keeping the total number of samples within bounds.

The Reduced Basis method is presented in Section 6.2. The basis identification is based on the sampling strategy developed in Section 6.3. Numerical examples are presented in Section 6.4. Both the numerical and the theoretical findings are summarized and discussed in Section 6.5.

6.1.4 Notation

The set of real numbers and the subset of positive numbers greater than zero are denoted by \mathbb{R} and \mathbb{R}_+ , respectively. Matrices are marked by two underlines and vectors by one underline, e.g., $\underline{\underline{A}}$, \underline{a} . Vectors are assumed to be columns, and the dot product of two vectors of the same size is understood as the Euclidean scalar product, $\underline{x} \cdot \underline{y} = \underline{x}^T \underline{y}$. First order and second order tensors in coordinate-free description are denoted by bold letters, e.g., \mathbf{A} , \mathbf{a} . No conclusion can be drawn on the order of a tensor based on its capitalization. Here, the underlying space is always the Euclidean space \mathbb{R}^3 with its standard basis. First order and second order tensors can be represented as vectors and matrices, e.g., $\mathbf{A} \leftrightarrow \underline{\underline{A}} \in \mathbb{R}^3$ and $\mathbf{B} \leftrightarrow \underline{\underline{B}} \in \mathbb{R}^{3 \times 3}$, respectively. Norms of vectors and matrices respectively denote the Euclidean and the Frobenius norm. The norm of a tensor of second order equals the norm of its matrix representation for the chosen basis. Fourth order tensors are denoted by blackboard bold symbols other than \mathbb{R} , e.g., \mathbb{C} and \mathbb{I} . Components of tensors of order M are with respect to the Euclidean tensorial basis

$\mathbf{e}^{(1)} \otimes \mathbf{e}^{(2)} \otimes \dots \otimes \mathbf{e}^{(M)}$, e.g., A_{ij} , B_{ij} for second order tensors \mathbf{A} , \mathbf{B} and C_{ijkl} , C'_{ijkl} for \mathbb{C} , \mathbb{C}' . The following contractions are defined:

$$\mathbf{A} \cdot \mathbf{B} = \sum_{i,j=1}^3 A_{ij} B_{ij},$$

$$\mathbb{C} \cdot \mathbf{B} = \sum_{i,j,k,l=1}^3 C_{ijkl} B_{kl} \mathbf{e}^{(i)} \otimes \mathbf{e}^{(j)},$$

$$\mathbf{A} \cdot \mathbb{C} = \sum_{i,j,k,l=1}^3 A_{ij} C_{ijkl} \mathbf{e}^{(k)} \otimes \mathbf{e}^{(l)},$$

$$\mathbb{C} \cdot \mathbb{C}' = \sum_{i,j,k,l,m,n=1}^3 C_{ijmn} C'_{mnkl} \mathbf{e}^{(i)} \otimes \mathbf{e}^{(j)} \otimes \mathbf{e}^{(k)} \otimes \mathbf{e}^{(l)}.$$

Let $\Omega \subset \mathbb{R}^3$ be the domain occupied by a physical body undergoing elastic deformations, and let Ω_0 be its initial configuration. Then, \mathbf{x} and \mathbf{X} describe the coordinates of material points within the current configuration Ω and within the reference state Ω_0 , respectively. Their difference is the displacement $\mathbf{u} = \mathbf{x} - \mathbf{X}$, see Figure 6.1. The gradient of a vector field $\mathbf{v} = \mathbf{v}(\mathbf{X})$ is defined as a *right gradient* and denoted by $\frac{\partial \mathbf{v}}{\partial \mathbf{X}} = \mathbf{v} \otimes \nabla_{\mathbf{X}}$. The divergence of a second order tensor field is the vector field resulting from *row-wise* divergence. The boundaries of the respective configurations are denoted by $\partial\Omega$ and $\partial\Omega_0$. The set of square-integrable Lebesgue functions on the reference domain is tagged $L^2(\Omega_0)$.

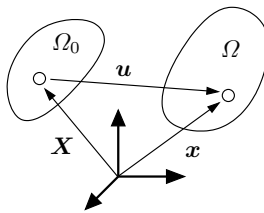


Figure 6.1: Initial (Ω_0) and current (Ω) configuration; elementary kinematic quantities.

The displacement gradient $\mathbf{H} = \mathbf{u} \otimes \nabla_{\mathbf{X}}$ and the deformation gradient $\mathbf{F} = \mathbf{x} \otimes \nabla_{\mathbf{X}}$ are related through $\mathbf{F} = \mathbf{H} + \mathbf{I}$, where \mathbf{I} is the second order identity tensor in three dimensions. The determinant $J = \det(\mathbf{F})$ measures the relative volumetric change due to the present deformation.

Unimodular quantities, i.e., second order tensors with determinant ones, may be emphasized by a hat, e.g., $\widehat{\mathbf{F}} = J^{-1/3} \mathbf{F}$. This multiplicative decomposition is

sometimes attributed to Flory Flory [1961] and also goes by the name Dilatational-Deviatoric Multiplicative Split (DDMS).

In the two-scale context, overlined symbols represent quantities on the macroscopic scale, e.g., $\overline{\mathbf{A}}$, $\overline{\mathbf{a}}$, while symbols without overline correspond to their microscopic counterpart, e.g., \mathbf{A} , \mathbf{a} . Equivalently, macroscopic quantities are called *global* and microscopic ones are called *local*. The volume average of a general local field φ

$$\langle \varphi \rangle = \langle \varphi(\bullet) \rangle = \frac{1}{|\Omega_0|} \int_{\Omega_0} \varphi(\bullet) dV \quad (6.1)$$

is essential to the theory. The dependence of a microscopic quantity \mathbf{A} on both the microscopic coordinates \mathbf{X} and a macroscopic quantity $\overline{\mathbf{B}}$ is denoted by $\mathbf{A} = \mathbf{A}(\mathbf{X}; \overline{\mathbf{B}})$. In such a case, the components of the macroscopic quantity $\overline{\mathbf{B}}$ are called *parameters* of the microscopic function $\mathbf{A}(\bullet; \overline{\mathbf{B}})$. The application of the volume averaging operator is abbreviated by $\langle \mathbf{A} \rangle = \langle \mathbf{A}(\bullet; \overline{\mathbf{B}}) \rangle$. The case of a concatenated function $f(\mathbf{A}) = f(\mathbf{A}(\mathbf{X}; \overline{\mathbf{B}}))$ is analogous, i.e. $\langle f \rangle = \langle f(\mathbf{A}) \rangle = \langle f(\mathbf{A}(\bullet; \overline{\mathbf{B}})) \rangle$, regardless of the tensorial order of the image of the function f .

6.1.5 Material Models

In this work, hyperelastic materials are investigated. They are characterized by *stored energy density functions* $W = W(\mathbf{F})$. The first Piola–Kirchhoff stress

$$\mathbf{P}(\mathbf{F}) = \frac{\partial W}{\partial \mathbf{F}}(\mathbf{F}) \quad (6.2)$$

and the corresponding fourth-order stiffness tensor

$$\mathbb{C}(\mathbf{F}) = \frac{\partial^2 W}{\partial \mathbf{F}^2}(\mathbf{F}) \quad (6.3)$$

characterize the material response.

Henceforth, for reasons of readability, the stored energy density function W will be spoken of as an energy, and the terms *stored* and *density* will not always be mentioned explicitly. In the infinitesimal strain framework, hyperelastic energies have been formulated to model deformation plasticity [e.g., Fritzen & Kunc, 2018c, Yvonnet et al., 2013, Bilger et al., 2005]. Although these models are only valid for purely proportional loading conditions, they provide means to simulate highly nonlinear material behavior in certain scenarios comparably easily within the context of hyperelasticity. Note that genuine dissipative processes require additional state describing variables with corresponding evolution laws.

The proposed method is *suitable for any type of hyperelastic constitutive law*. As the modeling of complex material behavior is not the main focus of this study, the Neo-Hookean law

$$W(\mathbf{F}) = W_{\text{DDMS}}(J, \widehat{\mathbf{F}}) = \frac{K}{4} [(J-1)^2 + (\ln J)^2] + \frac{G}{2} (\text{tr}(\widehat{\mathbf{F}}^T \widehat{\mathbf{F}}) - 3) \quad (6.4)$$

is used, with K the bulk modulus and G the shear modulus. The volumetric part of the energy is taken from Doll & Schweizerhof [1999]. Using the DDMS, a decoupled dependence on the volumetric and isochoric part of the deformation is assumed, which is a common way to model the distinct material behavior with respect to these two contributions, see e.g., [Simo, 1988].

6.1.6 Problem Setting of First Order Homogenization

Assuming stationarity and separability of scales, the following coupled and deformation-driven problems can be derived by means of asymptotic expansion of the displacement \mathbf{u} and subsequent first order approximation. This procedure is carried out in Pruchnicki [1998] with much detail. Here, the technical process is omitted and only the resulting equations are stated.

Macroscopic Problem

Balance of linear momentum

$$\text{Div}_{\overline{\mathbf{X}}}(\overline{\mathbf{P}}) + \overline{\mathbf{b}} = \mathbf{0}, \quad (6.5)$$

where $\overline{\mathbf{b}}$ denote bulk forces, and balance of angular momentum

$$\overline{\mathbf{F}}^{-1} \overline{\mathbf{P}} = \overline{\mathbf{P}}^T \overline{\mathbf{F}}^{-T}, \quad (6.6)$$

along with well-posed boundary conditions constitute the *macroscopic boundary value problem*. This system of equations is closed by means of the hyperelasticity law, cf. (6.2),

$$\overline{\mathbf{P}}(\overline{\mathbf{F}}) = \frac{\partial \overline{W}}{\partial \overline{\mathbf{F}}}. \quad (6.7)$$

Note that, in general, \overline{W} is a priori not available and (6.7) is thus a purely formal relation. For reasons of readability, the dependence of any quantity on the macroscopic material coordinate $\overline{\mathbf{X}}$ is usually spared, e.g., $\overline{\mathbf{F}} = \overline{\mathbf{F}}(\overline{\mathbf{X}})$, $\overline{\mathbf{H}} = \overline{\mathbf{H}}(\overline{\mathbf{X}})$, or $\overline{\mathbf{u}} = \overline{\mathbf{u}}(\overline{\mathbf{X}})$.

Microscopic Problem

The *microscopic boundary value problem* is given by the balance equations

$$\text{Div}_{\mathbf{X}}(\mathbf{P}) = \mathbf{0} \quad (6.8)$$

$$\mathbf{F}^{-1} \mathbf{P} = \mathbf{P}^{\top} \mathbf{F}^{-\top} \quad (6.9)$$

and suitable boundary conditions, e.g., as discussed in Miehe [2003]. In this work, *periodic displacement fluctuation boundary conditions* are employed. The microscopic displacements take the form

$$\mathbf{u}(\mathbf{X}; \bar{\mathbf{F}}) = \bar{\mathbf{u}} + \bar{\mathbf{H}}\mathbf{X} + \mathbf{w}(\mathbf{X}; \bar{\mathbf{F}}). \quad (6.10)$$

Therein, the macroscopic displacement is independent of microscopic quantities. The second term, $\bar{\mathbf{H}}\mathbf{X}$, corresponds to a homogeneous deformation of the microstructure. The third term, $\mathbf{w}(\mathbf{X})$, is a *displacement fluctuation* with the *zero mean property* $\langle \mathbf{w} \rangle = \mathbf{0}$. Hence, the deformation gradient reads

$$\mathbf{F}(\mathbf{X}; \bar{\mathbf{F}}) = \bar{\mathbf{F}} + \widetilde{\mathbf{H}}(\mathbf{X}; \bar{\mathbf{F}}) = \bar{\mathbf{F}} + \widetilde{\mathbf{F}}(\mathbf{X}; \bar{\mathbf{F}}), \quad (6.11)$$

where the fluctuation part $\widetilde{\mathbf{H}} = \mathbf{w} \otimes \nabla_{\mathbf{X}} = \widetilde{\mathbf{F}}$, has the zero mean property

$$\langle \widetilde{\mathbf{F}} \rangle = \mathbf{0}. \quad (6.12)$$

Thus, the volume average of the local deformation gradient equals its macroscopic counterpart,

$$\bar{\mathbf{F}} = \langle \mathbf{F} \rangle. \quad (6.13)$$

This motivates the identification of $\bar{\mathbf{F}}$ as *the boundary condition* to the microscopic problem (6.8). As for the material response, the following relationships can be deduced:

$$\bar{\mathbf{W}} = \langle \mathbf{W} \rangle, \quad (6.14)$$

$$\bar{\mathbf{P}} = \langle \mathbf{P} \rangle, \quad (6.15)$$

$$\bar{\mathbf{C}} \neq \langle \mathbf{C} \rangle. \quad (6.16)$$

Equations (6.13) and (6.15) are called kinematic and static *coupling relations*, respectively. The inequality (6.16) generally holds for heterogeneous microstructures, even in the most simple case of infinitesimal strains and linear elasticity. More precisely, the volume average overestimates the effective stiffness in the spectral sense,

$$\mathbf{x} \cdot \langle \mathbf{C} \rangle \cdot \mathbf{x} \geq \mathbf{x} \cdot \bar{\mathbf{C}} \cdot \mathbf{x} \quad \forall \text{ second order tensors } \mathbf{x}. \quad (6.17)$$

6.2 Reduced Basis Homogenization for Hyperelasticity

6.2.1 Formulation

The *Reduced Basis (RB)* scheme is based on a direct approximation of the microscopic deformation gradient \mathbf{F} from Equation (6.11) without the need to explicitly have the corresponding displacement at hand. The initial approximation is given by

$$\mathbf{F}_\xi(\mathbf{X}; \bar{\mathbf{F}}, \underline{\xi}) = \bar{\mathbf{F}} + \sum_{i=1}^N \mathbf{B}^{(i)}(\mathbf{X}) \xi_i. \quad (6.18)$$

The arguments $\bar{\mathbf{F}}$ and $\underline{\xi} \in \mathbb{R}^N$ are the boundary condition to (6.8) and the reduced coefficient vector, respectively. Note that the macroscopic coordinate $\bar{\mathbf{X}}$ is not assumed to influence the RB approximation, i.e., the same approximation is made throughout the macrostructure. The set $\{\mathbf{B}^{(i)}\}_{i=1}^N$ is linearly independent within the space $L^2(\Omega_0)$ and is called *RB of \mathbf{F}* . In a later context, it will also be referred to as the set of *ansatz functions*. In order to enforce the relationship

$$\langle \mathbf{F}_\xi \rangle = \bar{\mathbf{F}} \quad (6.19)$$

regardless of the reduced coefficients $\underline{\xi}$, the basis functions are asserted to have the *fluctuation property*, i.e., for $i = 1, \dots, N$

$$\langle \mathbf{B}^{(i)} \rangle = \mathbf{0}. \quad (6.20)$$

For now, the RB is assumed to be given.

The ansatz (6.18) allows for evaluation of the energy at the macroscale as a function of the macroscopic kinematic variable $\bar{\mathbf{F}}$ and of the reduced coefficients $\underline{\xi}$,

$$\bar{W}_\xi(\bar{\mathbf{F}}, \underline{\xi}) = \langle W(\mathbf{F}_\xi) \rangle. \quad (6.21)$$

By the principle of minimum energy, the optimal coefficients

$$\underline{\xi}^*(\bar{\mathbf{F}}) = \arg \min_{\underline{\xi} \in \mathbb{R}^N} \bar{W}_\xi(\bar{\mathbf{F}}, \underline{\xi}) \quad (6.22)$$

are sought after. The unconstrained and unique solvability of this task is assumed for the moment and will be addressed in Section 6.2.4. The solution of (6.22) defines the *RB approximation* of the deformation gradient

$$\mathbf{F}^{\text{RB}}(\mathbf{X}; \bar{\mathbf{F}}) = \bar{\mathbf{F}} + \sum_{i=1}^N \mathbf{B}^{(i)}(\mathbf{X}) \xi_i^*(\bar{\mathbf{F}}). \quad (6.23)$$

The microscopic energy, stress, and stiffness within the microstructure are then approximated by

$$W^{\text{RB}}(\mathbf{X}; \bar{\mathbf{F}}) = W(\mathbf{F}^{\text{RB}}(\mathbf{X}; \bar{\mathbf{F}})), \quad (6.24)$$

$$\mathbf{P}^{\text{RB}}(\mathbf{X}; \bar{\mathbf{F}}) = \frac{\partial W}{\partial \mathbf{F}}(\mathbf{F}^{\text{RB}}(\mathbf{X}; \bar{\mathbf{F}})), \quad (6.25)$$

$$\mathbb{C}^{\text{RB}}(\mathbf{X}; \bar{\mathbf{F}}) = \frac{\partial^2 W}{\partial \mathbf{F}^2}(\mathbf{F}^{\text{RB}}(\mathbf{X}; \bar{\mathbf{F}})), \quad (6.26)$$

respectively. The resulting effective energy is readily given by

$$\bar{W}^{\text{RB}}(\bar{\mathbf{F}}) = \langle W^{\text{RB}} \rangle. \quad (6.27)$$

Also, the effective responses $\bar{\mathbf{P}}^{\text{RB}}(\bar{\mathbf{F}})$ and $\bar{\mathbb{C}}^{\text{RB}}(\bar{\mathbf{F}})$ may now be calculated. However, before going into detail on that, it is advantageous to first elaborate on the solution of the minimization problem (6.22). This short survey will reveal essential properties of some occurring quantities that are important for the determination of the effective material response.

The necessary, first order optimality conditions to (6.22) define the components of the residual vector $\underline{r} \in \mathbb{R}^N$,

$$r_i(\bar{\mathbf{F}}, \underline{\xi}) = \frac{\partial \bar{W}_\xi}{\partial \xi_i}(\bar{\mathbf{F}}, \underline{\xi}) = \left\langle \frac{\partial W}{\partial \mathbf{F}}(\mathbf{F}_\xi) \cdot \frac{\partial \mathbf{F}_\xi}{\partial \xi_i} \right\rangle = \left\langle \mathbf{P}(\mathbf{F}_\xi) \cdot \mathbf{B}^{(i)} \right\rangle = 0. \quad (6.28)$$

Note 1 The solution stress field \mathbf{P}^{RB} is $L^2(\Omega_0)$ -orthogonal to the RB ansatz $\{\mathbf{B}^{(i)}\}_{i=1}^N$.

A viable choice for the solution of the minimization problem (6.22) is the Newton–Raphson scheme, which necessitates the Jacobian $\underline{\underline{D}} \in \text{Sym}(\mathbb{R}^{N \times N})$ with the components

$$D_{ij} = \left\langle \mathbf{B}^{(i)} \cdot \frac{\partial \mathbf{P}}{\partial \xi_j}(\mathbf{F}_\xi) \right\rangle = \left\langle \mathbf{B}^{(i)} \cdot \mathbb{C}(\mathbf{F}_\xi) \cdot \mathbf{B}^{(j)} \right\rangle = D_{ji} \quad (i, j = 1, \dots, N). \quad (6.29)$$

Then, the k^{th} iteration to the solution $\underline{\xi}^*(\bar{\mathbf{F}})$ reads

$$\underline{\xi}^{(k)} = \underline{\xi}^{(k-1)} - \left(\underline{\underline{D}}^{(k-1)} \right)^{-1} \underline{r}^{(k-1)} \quad (k \geq 1). \quad (6.30)$$

The initial guess $\underline{\xi}^{(0)}$ can be zero or a more sophisticated alternative.

The deduction of the effective material response by means of Note 1 and the Jacobian \underline{D} is given in Appendix 6.6.2. The following list summarizes the homogenized quantities arising from the \mathbf{F} -RB:

$$\bar{\mathbf{F}} = \langle \mathbf{F}^{\text{RB}} \rangle \quad \text{see (6.19)}$$

$$\bar{W}^{\text{RB}} = \langle W^{\text{RB}} \rangle \quad \text{see (6.27)}$$

$$\bar{\mathbf{P}}^{\text{RB}} = \langle \mathbf{P}^{\text{RB}} \rangle \quad \text{see Appendix 6.6.3}$$

$$\bar{\mathbb{C}}^{\text{RB}} = \langle \mathbb{C}^{\text{RB}} \rangle - \sum_{i,j=1}^N D_{ij}^{-1} \langle \mathbb{C}^{\text{RB}} \cdot \mathbf{B}^{(i)} \rangle \otimes \langle \mathbf{B}^{(j)} \cdot \mathbb{C}^{\text{RB}} \rangle \quad \text{see Appendix 6.6.4}$$

In total, the problem of determining both the local field \mathbf{F} and the homogenized material properties depends only on N degrees of freedom, namely on the N coefficients ξ_i . Usually, the number of DOF N is in the range of 10 to 150, which compares to the full order model's complexity that can easily reach more than 10^5 DOF.

Remark 3 Despite this impressive reduction of the number of DOF, the computational costs associated with the assembly of the residual \underline{r} and of the Jacobian \underline{D} still relate to the number of quadrature points of the microstructural discretization.

This method extends corresponding methods for the geometrically linear case, where the infinitesimal strain tensor $\boldsymbol{\varepsilon} = \text{sym}(\mathbf{H})$ is considered. For more information on these topics, the reader is referred to the authors' previous work Fritzen & Kunc [2018c] or standard literature, such as [Castañeda & Suquet, 1997, part II.C].

6.2.2 Identification of the Reduced Basis

The basis $\{\mathbf{B}^{(i)}\}_{i=1}^N$ is computed by means of a classical snapshot POD. In contrast to many other POD based reduction methods, it is important to point out that here, the primal variable is *not* taken to be the displacement field, \mathbf{u} . Instead, the POD is performed on deformation gradient fluctuations, $\tilde{\mathbf{F}}$.

During the snapshot POD, *snapshots* are first created by means of high-fidelity solutions to the nonlinear microscopic problem (6.8) with different *snapshot boundary conditions* $\bar{\mathbf{F}}^{(j)}$, $j = 1, \dots, N_s$, which are also called *training boundary conditions*. Each of these boundary conditions leads to a solution field $\mathbf{F}^{(j)}(\mathbf{X}; \bar{\mathbf{F}})$. Typically, the Finite Element Method (FEM) or solvers making use of the Fast Fourier Transform [e.g., Kabel et al., 2014] are used to this end. The RB method presented here is independent of the discretization method utilized to obtain full field solutions. It is merely necessary to know the quadrature weights and the related discrete values of the solutions $\mathbf{F}^{(j)}(\mathbf{X}; \bar{\mathbf{F}}^{(j)})$. For better readability, the continuous notation is maintained for the moment. The

corresponding fluctuation fields are computed by means of local subtraction of the macroscopic deformation gradient

$$\tilde{\mathbf{F}}^{(j)}(\mathbf{X}; \overline{\mathbf{F}}^{(j)}) = \mathbf{F}^{(j)}(\mathbf{X}; \overline{\mathbf{F}}^{(j)}) - \overline{\mathbf{F}}^{(j)} \quad (j = 1, \dots, N_s). \quad (6.31)$$

Each of these N_s fluctuation fields $\tilde{\mathbf{F}}^{(j)}$ represent one snapshot.

Second, the most dominant features of the snapshots are extracted. This is done by means of the eigendecomposition of the correlation matrix. It consists of the mutual $L^2(\Omega_0)$ scalar products of the snapshots, $\langle \tilde{\mathbf{F}}^{(i)} \cdot \tilde{\mathbf{F}}^{(j)} \rangle$ ($i, j = 1, \dots, N_s$), cf. (6.1). The remaining procedure is standard, see for instance [Sirovich, 1987] or [Quarteroni et al., 2016]: the N_s eigenvalues λ_j , corresponding to the eigenvectors $\underline{\mathbf{E}}^{(j)} \in \mathbb{R}^{N_s}$, are sorted in descending order and truncated by only considering the first N values, $\lambda_1 \geq \dots \geq \lambda_N$. The decision on a particular threshold index N is based on consideration of the Schmidt–Eckhard–Young–Mirsky theorem. Finally, the RB is constructed as

$$\mathbf{B}^{(i)}(\mathbf{X}) = \sum_{j=1}^{N_s} \frac{1}{\sqrt{\lambda_i}} \left(\underline{\mathbf{E}}^{(i)} \right)_j \tilde{\mathbf{F}}^{(j)}(\mathbf{X}) \quad (i = 1, \dots, N) \quad (6.32)$$

where the factor $1/\sqrt{\lambda_i}$ accounts for $L^2(\Omega_0)$ normalization of the base elements. We conclude that the RB is a collection of $L^2(\Omega_0)$ orthonormal $\tilde{\mathbf{H}}$ -like quantities.

6.2.3 Mathematical Motivation of the Reduced Basis model

Next, the obtained deformation gradient field \mathbf{F}^{RB} and the associated stress field \mathbf{P}^{RB} are shown to weakly solve the original problem (6.8), $\text{Div}_{\mathbf{X}}(\mathbf{P}) = \mathbf{0}$.

Let $\delta \mathbf{w} \in H_0^1(\Omega_0)$ be an admissible test function, i.e., a once weakly differentiable periodic displacement fluctuation field, and let $\delta \mathbf{H} = \delta \mathbf{w} \otimes \nabla_{\mathbf{X}}$ denote its gradient. Then, the well-known weak form of (6.8) is equivalent to the principle of virtual work,

$$\delta \overline{W} = \int_{\Omega_0} \mathbf{P} \cdot \delta \mathbf{H} \, dV = 0. \quad (6.33)$$

The residual \underline{r} from (6.28) coincides with the integral of the weak form, if the test function $\delta \mathbf{w}$ is chosen suitably. As the basis elements $\mathbf{B}^{(i)}$ are linear combinations of deformation gradient fluctuations $\mathbf{w}^{(j)} \otimes \nabla_{\mathbf{X}}$, cf. (6.32), the equivalence of (6.28) and (6.33) is obvious.

It follows that the Reduced Basis scheme is a *Galerkin* procedure, taking the displacement fields corresponding to the RB elements $\mathbf{B}^{(i)}$ as both ansatz and test functions. Hence, the *RB model is equivalent to the FEM*, but with basis functions that are globally supported in $\Omega_0 \setminus \partial \Omega_0$. In other words, the basis functions of the RB method span a subspace of dimension N within the high-dimensional space of FE basis functions. Although this property is shared with RB schemes based on POD of displacement snap-

shots, a notable difference is that this novel approach directly operates on fields entering the constitutive equations.

6.2.4 Details on the Coefficient Optimization

The coefficient optimization task (6.22) leads to a weak solution of the microscopic boundary value problem, as was just shown. Hence, the well-established theories on which the FEM is built, e.g., the calculus of variations, are applicable to the presented method just as much. This implies that the well-known issues with suitable convexity conditions and with existence and uniqueness of minimizers apply to the RB method, too. We focus on ad hoc numerical treatments of these issues. For more details on the theoretical part, the reader is referred to standard literature, such as Ball [1976], and recent developments in this matter, e.g., Schneider [2016].

A constraint to the optimization problem is the physical condition

$$J = \det(\mathbf{F}_\xi(\mathbf{X})) > 0 \quad \forall \mathbf{X} \in \Omega_0, \quad (6.34)$$

which means that no singular ($J = 0$) or self-penetrating ($J < 0$) deformations shall occur. This reduces the set of admissible coefficients ξ to a subset of \mathbb{R}^N that is nontrivial to access. The positiveness of the microscopic determinant of the deformation gradient introduces a constraint to the, thus far, unconstrained minimum problem (6.22), representing the weak form of (6.8) in the RB setting.

In case of a violation of the inequality (6.34), the implementation of the RB method is prone to failure as soon as the constitutive law (6.4) is evaluated. This occurs only in the context of volume averaging, i.e., for the assembly of the residual, the Jacobian, or the effective energy, stress, or stiffness. The numerical quadrature approximating the volume averaging operation is

$$\langle \bullet \rangle = \frac{1}{|\Omega_0|} \int_{\Omega_0} \bullet(\mathbf{X}) \, dV \approx \frac{1}{|\Omega_0|} \sum_{p=1}^{N_{\text{qp}}} \bullet(\mathbf{X}_p) w_p. \quad (6.35)$$

Here, N_{qp} , \mathbf{X}_p , and w_p respectively denote the number of quadrature points, their positions, and the corresponding positive weights. Moreover, even if the inequality (6.34) is satisfied everywhere, the local field \mathbf{F}_ξ might possibly have some positive but overly small values of the determinant, $0 < \det(\mathbf{F}_\xi(\mathbf{X})) \ll 1$, that are unphysical. In such a case, the energy optimization, cf. (6.22), would be dominated by these nearly singular points. Instead of allowing the optimization to focus almost exclusively on these exceptional points, we interpret unphysically small values of the determinant as limitations to the reliability of the RB method. On the other hand, large values $\det(\mathbf{F}_\xi(\mathbf{X})) \gg 1$ are not too detrimental to the functionality of the scheme, although such values are just as questionable.

Thus, the following *weighted numerical quadrature rule* is introduced:

$$\langle \bullet \rangle \approx \left(\sum_{p=1}^{N_{\text{qp}}} \bullet(\mathbf{X}_p) \phi(J_p) w_p \right) / \left(\sum_{p=1}^{N_{\text{qp}}} \phi(J_p) w_p \right). \quad (6.36)$$

Therein, the almost smooth *cutoff function* $\phi : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ is empirically defined by

$$\phi(J) = \begin{cases} 1 & \text{if } J > 0.6 \\ 0.5 \operatorname{erf}(30J - 15) + 0.5 & \text{if } 0.4 < J \leq 0.6 \\ 0 & \text{if } J \leq 0.4 \end{cases}. \quad (6.37)$$

which makes use of the well-known error function. The cutoff function is depicted in Figure 6.2. This reliability indicator could, in principle, be modified, e.g., the steepness, the smoothness, and its center could be adjusted. Thus, it should be regarded as an example only.

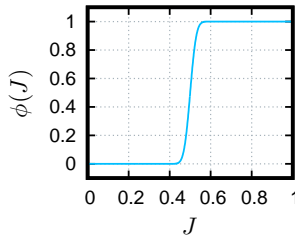


Figure 6.2: Cutoff function ϕ . Its value is used as a reliability factor in the numerical quadrature.

This weighted numerical quadrature rule is used henceforward for all *numerical* volume averaging operations. Its application will not be noted explicitly. However, the *theoretical* derivation of the RB method, as described in Section 6.2.1, is not affected by this, i.e., volume averaging operations remain exact as far as the theory is concerned. The two most important consequences of this numerical tweak are:

- The RB method is robust with respect to outlier values of the determinant. The modified quadrature rule extends the set of coefficient vectors $\underline{\xi}$ for which effective quantities can be computed, albeit approximately, to the whole space \mathbb{R}^N .
- The significance of local fields varies with the value of the cutoff function. When ϕ attains values less than one, information is considered accordingly less reliable. In this sense, microscopic information is filtered based on a trust region for J defined by ϕ can be seen as a reliability indicator.

6.3 Sampling

6.3.1 General Considerations

The proposed sampling strategy builds on the previous work Fritzen & Kunc [2018c], in which the authors proposed an analogous sampling procedure for the small strain setting. However, substantial modifications are required in order to account for the finite rotational part, $\overline{\mathbf{R}}$, of the macroscopic deformation gradient, $\overline{\mathbf{F}}$, and the nonlinearity of the volumetric part of the deformation, J , with respect to the local displacements, \mathbf{u} . For the setup of the Reduced Basis model as described in Section 6.2.2, the space of macroscopic deformation gradients,

$$\overline{\mathcal{F}} = \{\text{second order tensors } \overline{\mathbf{F}} \mid \det(\overline{\mathbf{F}}) > 0\}, \quad (6.38)$$

needs to be sampled, i.e., the discrete sampling set $\overline{\mathcal{F}}_s = \{\overline{\mathbf{F}}^{(m)}\}_{m=1}^{N_s} \subset \overline{\mathcal{F}}$ is to be defined. Two contradictory requirements need to be satisfied when constructing $\overline{\mathcal{F}}_s$:

1. The samples should be *densely and homogeneously distributed* within the space of all admissible macroscopic kinematic configurations. This is owing to the desire that the POD may extract correlation information from a holistic and unbiased set. In other words, the samples should be as *uniformly random* as possible *within the anticipated query domain of the surrogate*.
2. The sample number N_s should not exceed a certain limit. Only with this property may the RB be identified within the bounds of available computational resources (e.g., memory and CPU time).

6.3.2 Large Strain Sampling Strategy

The set of admissible macroscopic deformation gradients $\overline{\mathcal{F}}$ is a subset of a nine-dimensional space ($\overline{\mathbf{F}} \in \mathbb{R}^{3 \times 3} \sim \mathbb{R}^9$) restricted by the inequality

$$\det(\overline{\mathbf{F}}) > 0. \quad (6.39)$$

Therefore, a regular grid in the components of $\overline{\mathbf{F}}$ might lead to a prohibitively large amount of samples, and even to a violation of (6.39). For instance, such a grid with a rather moderate resolution of just 10 samples of each component would require 1 billion solutions of the FOM. Also, the subsequent POD would involve a snapshot matrix and/or correlation matrix of accordingly huge dimensionality.

In order to decrease the dimension of the sampling space, recall the polar decomposition of the deformation gradient, $\overline{\mathbf{F}} = \overline{\mathbf{R}}\overline{\mathbf{U}}$, where $\overline{\mathbf{R}}$ is a rotation and $\overline{\mathbf{U}}$ is the symmetric positive definite (s.p.d.) stretch tensor. Material objectivity implies the energy function to be independent of the frame of reference,

$$\overline{W}(\overline{\mathbf{R}}\overline{\mathbf{U}}) = \overline{W}(\overline{\mathbf{U}}), \quad (6.40)$$

and the transformation behavior

$$\overline{\mathbf{P}}(\overline{\mathbf{R}}\overline{\mathbf{U}}) = \overline{\mathbf{R}}\overline{\mathbf{P}}(\overline{\mathbf{U}}) \quad (6.41)$$

for the first Piola–Kirchhoff stress and

$$\overline{\mathbb{C}}_{ijkl}(\overline{\mathbf{R}}\overline{\mathbf{U}}) = \sum_{m,n=1}^3 \overline{R}_{im}\overline{\mathbb{C}}_{mjnl}(\overline{\mathbf{U}})\overline{R}_{kn} \quad (i, j, k, l = 1, 2, 3) \quad (6.42)$$

for the components of the corresponding stiffness tensor \mathbb{C} , see Appendix 6.6.1. These known facts lead to

Note 2 *In order to collect representative samples of the hyperelastic response functions $\overline{\mathbf{W}}$, $\overline{\mathbf{P}}$, and $\overline{\mathbb{C}}$, it suffices to evaluate them on samples of the stretch tensor $\overline{\mathbf{U}} \sim \mathbb{R}^6$ instead of evaluating them on samples of the deformation gradient $\overline{\mathbf{F}} \sim \mathbb{R}^9$.*

This effectively reduces the number of dimensions of $\overline{\mathbf{F}}$ from nine to six. The same dimensionality is attained when considering the response functions with respect to a symmetric measure of strain, e.g., as is done in Miehe [1996] where the tangent stiffness is efficiently computed using a perturbation technique. However, such measures are unsuitable for reduction by means of the proposed RB method.

The remaining six-dimensional space of s.p.d. tensors is not linear but a convex cone (which does not include the zero element). Moreover, linearly combining elements within this space quickly leads to values of $\overline{J} = \det(\overline{\mathbf{U}})$ describing unphysically large changes in volume. For instance, $\overline{\mathbf{U}} = 1.3\mathbf{I}$ equates to more than 100% volumetric extension, which is well beyond the regime of usual hyperelastic materials that are often nearly incompressible. On the other hand, 100% deviatoric strain is within the range of many standard materials, such as rubber. Hence, *in order to describe the space of practically relevant stretch tensors*, we propose to apply the DDMS to the macroscopic stretch tensor,

$$\overline{\mathbf{U}} = \overline{J}^{1/3} \widehat{\mathbf{U}}. \quad (6.43)$$

Let $\widehat{\mathcal{U}}$ denote the manifold of unimodular macroscopic stretch tensors $\widehat{\mathbf{U}} = (\overline{J})^{-1/3}\overline{\mathbf{U}}$. The multiplicative split (6.43) is the basis for:

Proposition 1 *The set of practically relevant macroscopic stretch tensors $\overline{\mathbf{U}}$ can be sampled via sampling of both the macroscopic determinants,*

$$\left\{ \overline{J}^{(m)} \right\}_{m=1}^{N_{\text{dil}}} \subset \mathbb{R}_+,$$

and the macroscopic unimodular stretch tensors,

$$\left\{ \widehat{\mathbf{U}}^{(j)} \right\}_{j=1}^{N_{\text{dev}}} \subset \widehat{\mathcal{U}},$$

where N_{dil} and N_{dev} are the numbers of the samples. The sampling set is determined by the product set

$$\left\{ \left(\bar{J}^{(m)} \right)^{1/3} \widehat{\underline{U}}^{(j)} \right\}_{m,j=1}^{m=N_{\text{dil}}, j=N_{\text{dev}}} \subset \bar{\underline{U}}. \quad (6.44)$$

The choice of the dilatational samples is relatively straightforward. For many common materials, the expected range of \bar{J} is rather narrow, so a few equisized or adaptive sub-intervals around $\bar{J} = 1$ deliver sufficient resolution.

For the space of unimodular s.p.d. matrices representing $\widehat{\underline{U}} \in \widehat{\underline{U}}$, basic results of Lie group theory can be exploited. We restrict to stating well-known facts that are necessary at this point. For more details, the interested reader is referred to standard text books, such as Faraut [2008].

Corollary 1 *Let*

$$\text{symsl} = \left\{ \underline{Y} \in \mathbb{R}^{3 \times 3} \mid \underline{Y} = \underline{Y}^T, \text{tr}(\underline{Y}) = 0 \right\}$$

be the tangent space and

$$\text{SymSL}_+ = \left\{ \underline{U} \in \mathbb{R}^{3 \times 3} \mid \underline{U} = \underline{U}^T, \det(\underline{U}) = 1, \underline{x}^T \underline{U} \underline{x} > 0 \forall \underline{x} \in \mathbb{R}^3 \right\}$$

be the manifold of unimodular s.p.d. matrices. The matrix exponential maps the tangent space bijectively onto the manifold,

$$\exp : \text{symsl} \rightarrow \text{SymSL}_+.$$

The proof of this statement is standard, e.g., by means of the eigenvalue decomposition, and does not necessitate the reference to the abstract setting of Lie groups. In fact, all of the following arguments could be given without the notion of tangent spaces and manifolds. However, this notion is a fundamental concept in nonlinear mechanics. For a very descriptive and comprehensive work on this topic, the reader is referred to Neff et al. [2016]. We choose to build upon this concept, as it comes along with vivid interpretations of the function spaces $\bar{\underline{U}}$ and $\widehat{\underline{U}}$.

The set SymSL_+ is the set of matrix representations of the stretch tensors $\widehat{\underline{U}} \in \widehat{\underline{U}}$. The tangent space symsl is the set of *Hencky strains*, which is linear. Hence, by virtue of the matrix exponential, the sampling of the nonlinear manifold $\widehat{\underline{U}}$ can be reduced to the sampling of a linear space. Moreover, the restrictions of symmetry and zero trace render the tangent space five-dimensional. This property is, by definition, shared with the manifold SymSL_+ .

The two-dimensional case is now addressed for the sake of visualization. In this setting, the nonlinearity of the manifold and the structure of the sampling set $\bar{\underline{U}}$ can be illustrated

figuratively. With the subscript (2) denoting two-dimensional quantities, a basis of the tangent space is given by

$$\underline{\underline{Y}}_{(2)}^{(1)} = \sqrt{\frac{1}{2}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad \underline{\underline{Y}}_{(2)}^{(2)} = \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The stretch tensors are obtained through

$$\overline{\underline{\underline{U}}}_{(2)} = \overline{J}^{\frac{1}{2}} \exp \left(t \left(\alpha \underline{\underline{Y}}_{(2)}^{(1)} + \beta \underline{\underline{Y}}_{(2)}^{(2)} \right) \right) = \begin{bmatrix} a & b \\ b & d \end{bmatrix}.$$

A visualization of such samples is depicted in Figure 6.3. There, for the sake of visual clarity, the determinant \overline{J} is sampled by four equidistant (and rather unrealistic) values between 0.1 and 4. The value $t \in [-2, 2]$ is called *deviatoric amplitude*. A densely uniform sampling $\varphi \in [0, 2\pi)$ yields the coefficients $\alpha = \cos \varphi$ and $\beta = \sin \varphi$.

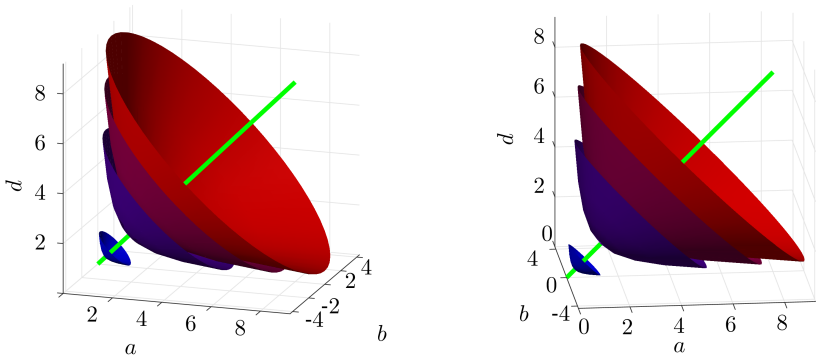


Figure 6.3: Visualizations of the family of $\overline{\underline{\underline{U}}}$ -manifolds with constant determinant $\overline{J} \in \{0.1, \dots, 4.0\}$ for the two-dimensional case from two different perspectives. The green line represents the set $\{\lambda \mathbf{I} \mid \lambda > 0\}$.

This emphasizes the important role of the DDMS in the context of sampling, as utilized in (6.44) – it allows for the identification of a physically meaningful sampling domain that is much smaller than the surrounding space of all admissible stretch tensors. On a side note, the isodet surfaces are perpendicular to the line representing the dilatational stretch tensors.

The proposed sampling procedure for $\overline{\underline{\underline{U}}}$ in three dimensions is given in Algorithm 6.1. For this purpose, an orthonormal basis $\underline{\underline{Y}}^{(1)}, \dots, \underline{\underline{Y}}^{(5)}$ of the tangent space symsl is fixed, cf. Appendix 6.6.5. The numbers of different kind of samples N_{det} , N_{dir} , and

N_{amp} relate to the quantities N_{dil} and N_{dev} introduced in (6.44) by $N_{\text{det}} = N_{\text{dil}}$ and $N_{\text{dir}} N_{\text{amp}} = N_{\text{dev}}$.

Algorithm 6.1: Sampling of the macroscopic stretch tensor.

Input : $\bar{J}_{\min}, \bar{J}_{\max}$ minimum and maximum determinant with
 $\bar{J}_{\min} < 1 < \bar{J}_{\max}$
 $t_{\max} > 0$ maximum deviatoric amplitude
 N_{det} number of macroscopic determinants
 N_{dir} number of deviatoric directions
 N_{amp} number of deviatoric amplitudes

Output: $N_{\text{det}} N_{\text{dir}} N_{\text{amp}}$ samples of \bar{U}

- 1 Place N_{det} determinants regularly between the extremal values,

$$\bar{J}_{\min} = \bar{J}^{(1)} < \dots < 1 < \dots < \bar{J}^{(N_{\text{det}})} = \bar{J}_{\max} .$$

- 2 Generate any approximately uniform distribution of N_{dir} directions in \mathbb{R}^5 , e.g., as in Fritzen & Kunc [2018c],

$$\left\{ \underline{N}^{(n)} \right\}_{n=1}^{N_{\text{dir}}} \subset \{ \underline{N} \in \mathbb{R}^5 : \|\underline{N}\| = 1 \} .$$

- 3 Place N_{amp} deviatoric amplitudes regularly between 0 and the expected maximum value,

$$0 \leq t^{(1)} < \dots < t^{(N_{\text{amp}})} = t_{\max} .$$

- 4 Return the set of samples of \bar{U} :

$$\left\{ \left(\bar{J}^{(m)} \right)^{1/3} \exp \left(t^{(p)} \left[\sum_{k=1}^5 \left(\underline{N}^{(n)} \right)_k \underline{Y}^{(k)} \right] \right) \right\}_{m,n,p=1}^{m=N_{\text{det}}, n=N_{\text{dir}}, p=N_{\text{amp}}} \subset \bar{U} \quad (6.45)$$

The order of Steps 1 to 3 is interchangeable. Details on these parts are now presented:

- Step 1. Uniform seeding of the determinants is actually not required, but any pattern implying the sampling determinants $\{ \bar{J}^{(k)} \}_{k=1}^{N_{\text{det}}}$ to be dense in $[\bar{J}_{\min}, \bar{J}_{\max}]$ as $N_{\text{det}} \rightarrow \infty$ works without loss of generality. In this way, the dilatational response may be resolved adaptively.
- Step 2. The generation of uniform point distributions on spheres is a research topic on its own, see [Brauchart & Grabner, 2015] for an overview. The method

described in Fritzen & Kunc [2018c] is based on energy minimization, which is also used in the present work. Some point sets of various sizes are included in the example program [Kunc, 2019]. More detailed investigations on this topic and an open-source code of a point generation program are part of another work, [Kunc & Fritzen, 2019c]. Alternatively, Equal Area Points [Leopardi, 2006] may be used as a rough but quickly computable approximation of such point sets.

Step 3. As in Step 1, the uniform placement of the deviatoric amplitudes, $t^{(p)}$, may be substituted by adaptive alternatives. In [Fritzen & Kunc, 2018c], we have suggested to use an exponential distance function.

The result of Steps 2 and 3, i.e., the sampling of the tangent space, is exemplified in Figure 6.4 for the two-dimensional case ($\mathbf{u} \in \mathbb{R}^2$) and for $N_{\text{dir}} = 5$ and $N_{\text{amp}} = 3$. There, an adaptive spacing of the deviatoric amplitudes is applied. This might be beneficial for capturing strongly changing material behavior near the relaxed state.

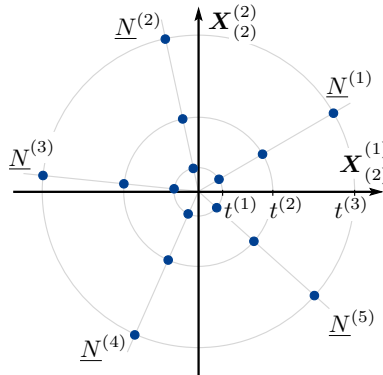


Figure 6.4: Example of a discretization of the two-dimensional tangent space. The samples are placed along the equidistant (in higher dimensions—uniformly distributed) directions with nonuniformly increasing amplitude.

In general, the vector $\underline{N} \in \mathbb{R}^5$, $\|\underline{N}\| = 1$, corresponding to a macroscopic Hencky strain characterizes the *direction of the applied stretch* $\overline{\mathbf{U}}$, which we also coin the *load case*.

6.3.3 Application of the Stretch Tensor Trained Reduced Basis Model

Since the RB model is trained on only the rotationally invariant part of $\overline{\mathbf{F}}$ but should be applied to general deformation gradients, the transformation rules (6.41) and (6.42) are incorporated during the evaluation of the surrogate model. Details on the procedure are given in Algorithm 6.2.

Algorithm 6.2: Online phase of the stretch tensor trained Reduced Basis method

-
- Input :** $\bar{\mathbf{F}}$ macroscopic deformation gradient
Output: $\bar{\mathbf{P}}^{\text{RB}}(\bar{\mathbf{F}})$, $\bar{\mathbb{C}}^{\text{RB}}(\bar{\mathbf{F}})$ effective material response
- 1 Compute polar decomposition $\bar{\mathbf{F}} = \bar{\mathbf{R}}\bar{\mathbf{U}}$.
 - 2 Compute approximations of effective stress $\bar{\mathbf{P}}^{\text{RB}}(\bar{\mathbf{U}})$ and effective stiffness $\bar{\mathbb{C}}^{\text{RB}}(\bar{\mathbf{U}})$.
 - 3 Transform effective responses to correct frame $\bar{\mathbf{P}}^{\text{RB}}(\bar{\mathbf{F}})$, $\bar{\mathbb{C}}^{\text{RB}}(\bar{\mathbf{F}})$, using $\bar{\mathbf{R}}$, cf. (6.41), (6.42).
-

6.4 Numerical Examples

6.4.1 Reduced Basis for a Fibrous Microstructure

The applicability of the proposed RB method in combination with the sampling scheme is now numerically studied for a fibrous microstructure roughly resembling polymers with woven reinforcements. The goal is to prove the efficiency of the \mathbf{F} -RB scheme in principle and under “worst-case” conditions, the latter meaning that the phase contrast is chosen to be rather large. Yet, at this stage, it is explicitly not aspired to provide fully realistic examples. These would require investigations on the proper size of the microstructure and should employ dissipative material laws, both of which are not within the scope of this work.

A cubical microstructure with two fibrous inclusions is considered, see Figure 6.5a and cf. [Kim & Swan, 2003] for a related example. The inclusions are periodic and occupy approximately 0.7% of the volume. The mesh contains 35, 516 nodes in 25, 633 quadratic tetrahedron elements (C3D10).

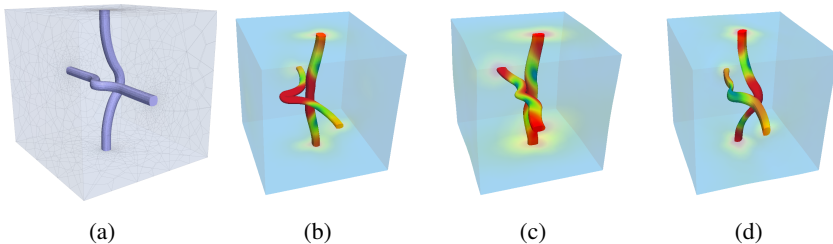


Figure 6.5: (a) fibrous microstructure. (b)–(d) random \mathbf{F} -reduced basis (RB) elements.

For the matrix, the material constants are chosen to be $K_m = 400$ MPa and $G_m = 0.4$ MPa, resembling rubber-like material properties. For the fibers, the values are $K_f = 800$ MPa

and $G_f = 240$ MPa. The latter parameters approximate the behavior of stiffer polymers, such as polyethylene. The phase contrast is 600 with respect to the shear moduli, and the Poisson ratios are $\nu_m = 0.4995$ and $\nu_f = 0.3636$.

The training boundary conditions are defined with $N_{\text{dir}} = 128$ deviatoric directions, $\underline{N}^{(n)}$, and $N_{\text{amp}} = 10$ deviatoric amplitudes, $t^{(p)}$, which are regularly distributed in the interval $[0.05, 0.5]$. In order to also consider response to volumetric extension in the training data, an additional set of $N_{\text{det}}=10$ boundary conditions of the form $(\bar{J}^{(m)})^{1/3} \mathbf{I}$ is included in the training set, with the determinant $\bar{J}^{(m)}$ being linearly increased from 1 to 1.02.

The validation load cases are 640 mixed dilatational-deviatoric boundary conditions. Along $N_{\text{dir}} = 64$ new deviatoric directions, both the deviatoric amplitudes ($t^{(p)} = 0.05, \dots, 0.5$) and the dilatational amplitudes ($\bar{J}^{(m)} = 0.9995, \dots, 0.9950$) are applied in 10 equidistant increments.

The results for various values N of the RB-size are compared with the results of FE simulations with the same boundary conditions. To this end, the error measures

$$\text{err}_W = \frac{\|\bar{W}^{\text{RB}} - \bar{W}^{\text{FEM}}\|}{\|\bar{W}^{\text{FEM}}\|} \quad \text{and} \quad \text{err}_P = \frac{\|\bar{P}^{\text{RB}} - \bar{P}^{\text{FEM}}\|}{\|\bar{P}^{\text{FEM}}\|} \quad (6.46)$$

are employed. Figures 6.6 and 6.7 visualize the results.

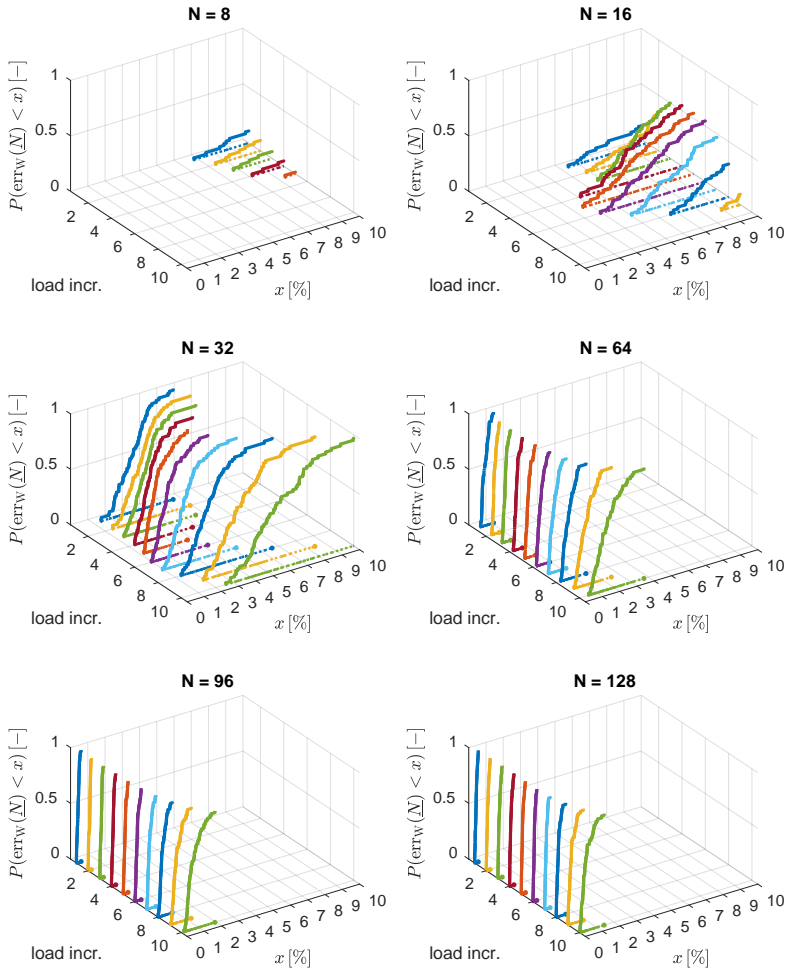


Figure 6.6: Cumulative energy error distribution per direction for the *RB* of the fibrous microstructure under validation boundary conditions.

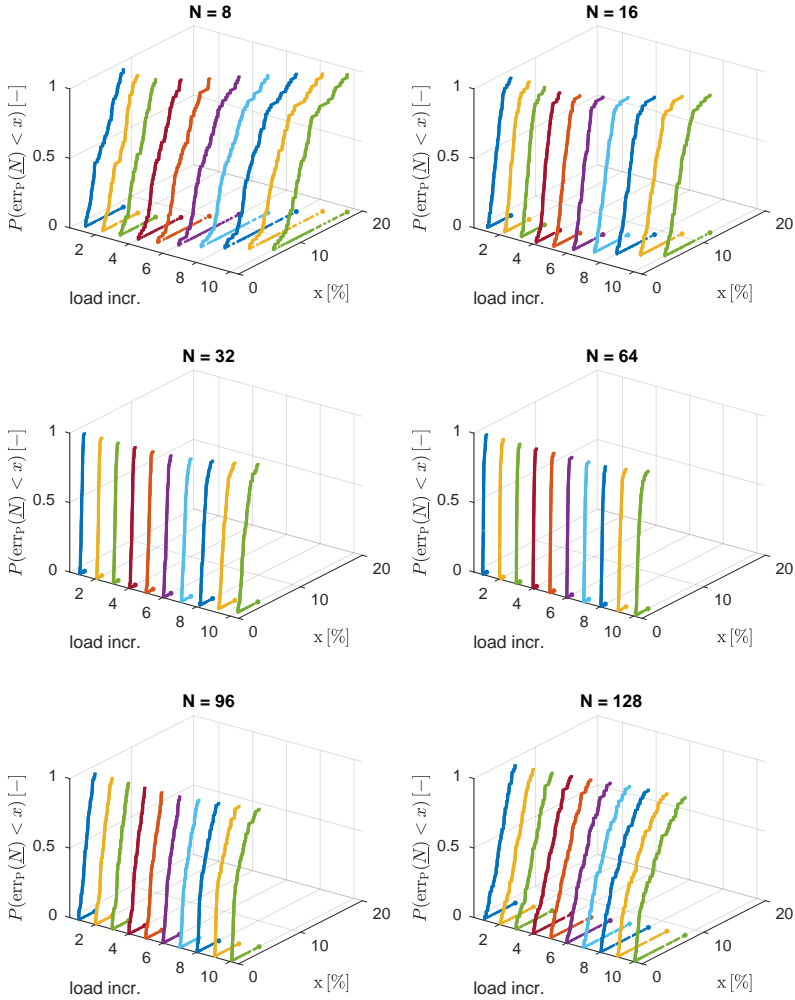


Figure 6.7: Cumulative stress error distribution per direction for the *RB* of the fibrous microstructure under validation boundary conditions.

The distribution of the energy error, err_w , improves monotonically as the RB is enriched from $N = 8$ to $N = 128$ elements. This enrichment corresponds to the inclusion of additional subtrahends in the computation of $\overline{\mathbb{C}}^{\text{RB}}$, improving the spectral overestimation by the volume average of the stiffness, cf. (6.17). It is also noteworthy that

the error tends to be higher for larger magnitudes of the applied kinematic boundary condition, although that is not always the case.

In contrast to this, the stress error err_P distribution monotonically improves only up to a certain threshold value of the number of basis elements, which is $N = 64$ in this example. For the greater bases with $N = 96$ and $N = 128$, the quality of the results deteriorates as far as the stress error is concerned. This is most likely due to an excessively oscillatory nature of the higher order modes—at some critical level $1 \ll i = N_{\text{crit}} < N$, the POD constructs eigenvectors $\underline{E}^{(i)}$ with the $L^2(\Omega_0)$ -norm $\sqrt{\lambda_i} \ll 1$. Therefore, the POD would construct basis vectors out of *numerical fluctuations*, which would be unphysical. While the enrichment of the optimization space with unphysical information cannot increase the minimum energy error err_W , it might lead to fluctuations in the displacement field that have significant impact on the overall stress response. This is especially the case for numerical fluctuations within the stiff inclusion phase where low overall displacement errors still could lead to notable impact on the effective stress.

Nonetheless, all observed errors are less than 20% and stay below 3% for the optimal sampled size $N = 64$. For half the basis size, $N = 32$, the errors max at approximately 5%, which is still acceptable considering the uncertainties involved in realistic two-scale simulations. Note that the maximum errors strongly depend on the maximum load amplitude, which is chosen to be very large in this example (50% deviatoric strain and 0.5% compression).

The runtimes of the RB model for different sizes N are depicted in Figure 6.8. A nearly linear growth of the runtime with respect to the basis size can be asserted. It is noteworthy that the online time of the RB method is strongly dominated by the assembly of the Jacobian \underline{D} . Therefore, a *Quasi-Newton implementation* was chosen, resulting in only two assemblies per load increment.

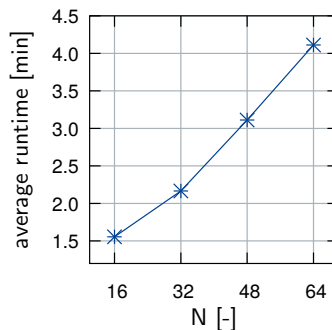


Figure 6.8: Laptop computer runtimes of the RB model for the fibrous microstructure for various sizes N of the basis. Each data point represents the time needed for all 10 load increments. The spread of the individual times of the 64 validation cases around these average values is negligible.

Speed-ups become impressive when very large load increments are considered. In all examples observed thus far, *the RB converges to the final load amplitude in a single increment, requiring 7–13 Quasi-Newton iterations, with only 2–4 assemblies of the Jacobi matrix \underline{D} and a runtime of 10–50 seconds.* This is in strong contrast to the FEM which is very sensitive to large load increments as they come along with a high probability of a violation of the condition $\det(\mathbf{F}) > 0$. By means of standard implementation, such occurrences are usually treated with cutbacks of the load increment, which is detrimental to the runtime of the FEM.

No rigorous speed-up analysis is intended at this point. Both the codes of the FEM and of the RB method are fairly efficient in-house C/C++ developments and perform exact line searches. While the FEM has not yet been equipped with a Quasi-Newton procedure, the linear solver makes use of parallelization. This is in contrary to the current implementation of the RB method. Depending on the microstructure (especially the geometry, material nonlinearities, and phase contrast), the loading conditions, and the size N of the RB, *speed-up factors are in the order of 5–100.*

6.4.2 Reduced Basis for a Stiffening Microstructure

The second example takes the “worst-case” approach further by considering a noncubical microstructure with even higher phase contrast and significant topological nonlinearity. To this end, a cuboid microstructure occupying the domain $[-0.5, 0.5] \times [-0.3, 0.3] \times [-0.05, 0.05] \subset \mathbb{R}^3$ and containing a hash-like inclusion is investigated, see Figure 6.9a. The mesh is periodic and contains 33,923 nodes in 21,726 quadratic tetrahedron elements (C3D10). The reinforcement makes up approximately 13.3% of the volume. Due to this large volume fraction, a pronounced geometry-induced nonlinearity of the effective response is expected under uniaxial loading conditions along the x -axis. As it is elongated, the hashlike part is straightened and thus increasingly aligned with the external loading, see Figure 6.9b. Such effects might be desirable when designing microstructures for functional materials.

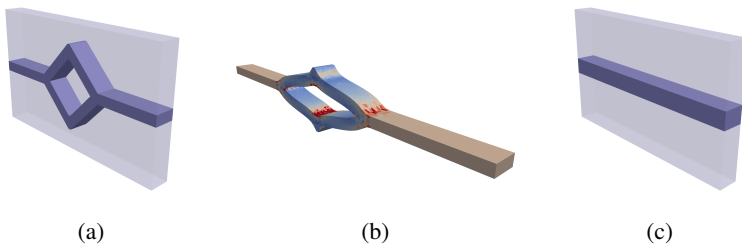


Figure 6.9: (a) Cuboidal microstructure with hashlike inclusion phase. (b) Deformed state under uniaxial tension loading. Only inclusion is shown, coloring indicates \bar{P}_{xx} . (c) Straight inclusion substitute microstructure, leading to a comparable effective stress.

The material parameters are $K_m = 19.867$ MPa, $G_m = 0.4$ MPa, $K_f = 19,867$ MPa, and $G_f = 400$ MPa, implying a Poisson ratio of 0.49 in both materials and a phase contrast of 1000.

The training boundary conditions are the deviatoric ones of the set considered in Section 6.4.1, i.e., $N_{\text{dir}} = 128$ deviatoric directions and $N_{\text{amp}} = 10$ regularly spaced deviatoric amplitudes from the interval $[0.05, 0.5]$. No dilatational training cases are considered, i.e., only points from a five-dimensional submanifold of the space $\bar{\mathcal{U}}$ are sampled.

Uniaxial tension boundary conditions are applied for the validation. More precisely, the stretch component \bar{U}_{xx} is increased from 1.0 to 1.5 in 10 increments of equal size. The other components are chosen such that all but the xx -component of the effective stress $\bar{\mathbf{P}}$ vanish.

Figure 6.10 depicts the results for different sizes N of the RB. The influence of the stiffening effect on the stress curve is emphasized by the black dashed line corresponding to a similar microstructure with a straight, cuboid inclusion that leads to the same final stress value under these boundary conditions, see Figure 6.9c.

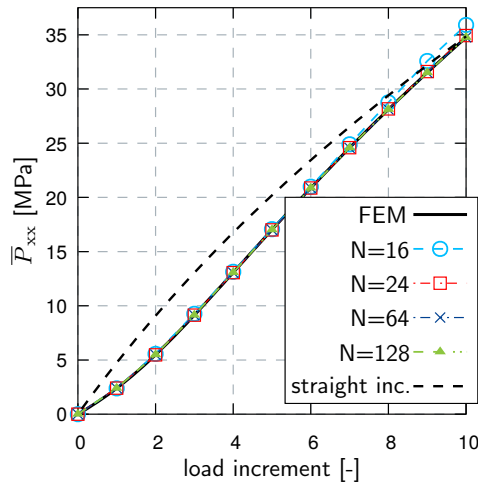


Figure 6.10: Stress curves for a microstructure with geometric stiffening (cf. Figure 6.9a), comparing the FEM result to the RB for various number of basis elements N . A similar microstructure without geometric stiffening but with the same final stress value (cf. Figure 6.9c) leads to the black dashed curve.

In this example, the geometric stiffening effect is captured by the RB with high accuracy, with as few as $N = 24$ basis elements. For moderate stretches, even an RB size of $N = 16$ suffices. These results are somewhat more impressive when noticing that the

applied boundary condition contains more than 1.2% volumetric compression, i.e. the validation loading actually lies outside the submanifold covered during training.

In order to examine the action of the cutoff function ϕ , the following two indicators are introduced:

$$c_{\text{qp}} = \# \text{ quadrature points with } (\phi(J) < 1), \quad (6.47)$$

$$V_{\text{excl}} = \left(|\Omega_0| - \sum_{p=1}^{N_{\text{qp}}} \phi(J_p) w_p \right) / |\Omega_0|. \quad (6.48)$$

The first quantity, c_{qp} , counts the number of quadrature points at which the cutoff function has an influence. The second one, V_{excl} , measures the relative excluded volume, interpreting the value of ϕ as a scaling of the corresponding quadrature weight. The values of these indicators are depicted in Figure 6.11.

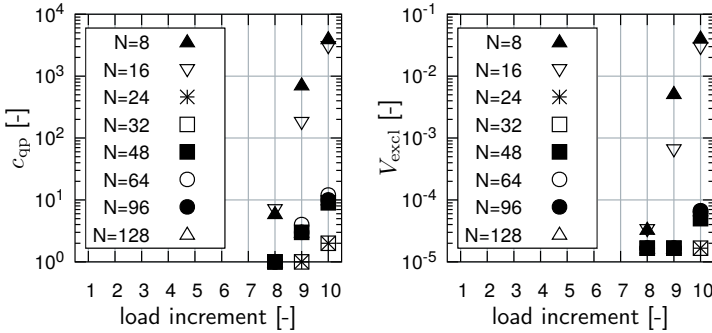


Figure 6.11: (Left) Number of quadrature points at which the cutoff function ϕ attains a value less than one. (Right) Relative excluded volume.

Most notably, the cutoff function does not have any effect before the eighth load increment in this example. Only for large load amplitudes does this numerical stability tweak become necessary. Even then, the number of points at which it has an influence is small, considering the total number of quadrature points, $N_{\text{qp}} = 86,904$. This example is representative for all conducted numerical experiments.

6.5 Discussion

6.5.1 Discussion of the Reduced Basis Method

Relation of the RB Homogenization to Analytical Estimates

Zero coefficients, $\underline{\xi} = \underline{0}$, correspond to the Taylor homogenization, i.e., to the nonlinear counterpart of the Voigt estimate [Voigt, 1910], which provides an upper bound for the material response, cf. (6.17). Starting with the initial guess $\underline{\xi}^{(0)} = \underline{0}$, the evolution of the coefficients corresponds to a (possibly not monotonous) relaxation of this overly stiff response into a more natural state. In view of improved computational efficiency, a nonzero initial guess $\underline{\xi}^{(0)}$ combined with an exact line search has proven reasonable and easy to realize. For instance, such a guess might stem from previous load steps or an interpolation/extrapolation of available coefficient data.

Reconstruction of Displacement Fields

Given an RB approximation of the deformation gradient, \mathbf{F}^{RB} , one can reconstruct the corresponding displacement field uniquely up to rigid body motion. This is possible due to the linear dependence of the deformation gradient fluctuations on the displacement fluctuations. Recall the definition of the RB in (6.32),

$$\mathbf{B}^{(i)} = \sum_{j=1}^{N_s} \frac{1}{\sqrt{\lambda_i}} \left(\underline{E}^{(i)} \right)_j \tilde{\mathbf{F}}^{(j)} \quad (i = 1, \dots, N).$$

The corresponding displacement fluctuations are

$$\tilde{\mathbf{u}}_{\text{B}}^{(i)} = \sum_{j=1}^{N_s} \frac{1}{\sqrt{\lambda_i}} \left(\underline{E}^{(i)} \right)_j \tilde{\mathbf{u}}^{(j)} \quad (i = 1, \dots, N). \quad (6.49)$$

The displacement fluctuation fields $\tilde{\mathbf{u}}^{(j)}$ are defined by $\tilde{\mathbf{u}}^{(j)}(\mathbf{X}) = \mathbf{u}^{(j)}(\mathbf{X}) - \overline{\mathbf{H}}^{(j)} \mathbf{X}$, where the displacement fields $\mathbf{u}^{(j)}$ are the solutions computed during training, and $\overline{\mathbf{H}}^{(j)} = \overline{\mathbf{U}}^{(j)} - \mathbf{I}$. Thus, a displacement field compatible to the RB result $\mathbf{F}^{\text{RB}}(\mathbf{X}; \overline{\mathbf{F}})$ is given by

$$\mathbf{u}^{\text{RB}}(\mathbf{X}; \overline{\mathbf{F}}) = \overline{\mathbf{H}} \mathbf{X} + \sum_{i=1}^N \xi_i^*(\overline{\mathbf{F}}) \tilde{\mathbf{u}}_{\text{B}}^{(i)}(\mathbf{X}). \quad (6.50)$$

The missing term $\overline{\mathbf{u}}(\overline{\mathbf{X}})$, cf. (6.10), cannot be reconstructed.

Relation to Classical Displacement-Based POD Methods

In a certain sense, the entries of the correlation matrix used in the offline phase, cf. Section 6.2.2, are “*weighted*” scalar products of the displacement fluctuation fields $\tilde{\mathbf{w}}^{(i)}$ within the Sobolev space $H_0^1(\Omega_0)$. “*Weighted*” is to be understood in that the zeroth order derivative is multiplied by zero. Classical displacement-based POD methods compute correlations of the fluctuations $\tilde{\mathbf{w}}^{(i)}$ within the Lebesgue space $L^2(\Omega_0)$. The change to $H_0^1(\Omega_0)$ -like scalar products is physically motivated by the fact that the local energy $W = W(\mathbf{F})$ explicitly depends on the gradient of the displacement, $\mathbf{F} = \mathbf{u} \otimes \nabla_{\mathbf{x}} + \mathbf{I}$, but *not* on the displacement, \mathbf{u} .

Advantages Compared to General Displacement-Based Schemes

The proposed method is advantageous compared to both displacement-based POD methods and the classical FEM for the following reasons:

- *No gradients need to be computed from displacement fields*, which displacement-based schemes always require prior to the evaluation of the material law.
- The residual \underline{r} and the Jacobian \underline{D} are *algorithmically sleek and trivial to implement*.
- The *absence of element formulations* in the assembly of the reduced residual \underline{r} and of the Jacobi matrix \underline{D} contributes to both the simplicity and the efficiency of the method—no incidence matrices occur, allowing for *linear memory access*. Moreover, the algebraic operations associated with reference element formulations are bypassed. This is also in favor of *parallel computations*. Such an implementation is still outstanding for the problem at hand, but has been conducted for related problems in the small strain setting in [Fritzen & Hodapp, 2016].

Although the storage of the basis $\{\mathbf{B}^{(i)}\}_{i=1}^N$ requires $9N_{\text{qp}}N$ double precision values, the basis is compact enough to be completely loaded into random access memory of standard computers. For instance, the bases considered in Section 6.4 occupy only ~200 Mb of memory for $N = 32$.

We now briefly address the *algorithmic complexity* associated with the proposed \mathbf{F} -RB method and with the \mathbf{u} -RB method that was employed in previous works, such as [Radermacher & Reese, 2015] and [Yvonnet & He, 2007]. To this end, the fully discretized versions of the residual \underline{r} and of the Jacobian \underline{D} as well as discrete quantities associated with the \mathbf{u} -RB method are introduced in the following listing.

- $\underline{P}(\mathbf{X}_p) \in \mathbb{R}^9$: Nine values of the stress $\mathbf{P}(\mathbf{F}_\xi(\mathbf{X}_p))$ at the quadrature point $\mathbf{X}_p \in \Omega_0$
- $\underline{C}(\mathbf{X}_p) \in \mathbb{R}^{9 \times 9}$: Symmetric stiffness tensor
- $\underline{B}(\mathbf{X}_p) \in \mathbb{R}^{9 \times N}$: \mathbf{F} -RB matrix, nine values of each basis elements $\mathbf{B}^{(i)}$ as columns
- w_p : The quadrature weight at \mathbf{X}_p

- N_{DOF} : Three times the number of nodes, $\sim N_{\text{qp}}$
- $\underline{r}^{\text{FE}} \in \mathbb{R}^{N_{\text{DOF}}}$: Global FE residual vector
- $\underline{\underline{B}}^{\text{u}} \in \mathbb{R}^{N_{\text{DOF}} \times N}$: \mathbf{u} -RB matrix, columns contain the nodal displacement values
- $\underline{\underline{K}}^{\text{FE}} \in \mathbb{R}^{N_{\text{DOF}} \times N_{\text{DOF}}}$: Global FE stiffness matrix

Table 6.1 compares the *algorithmic complexity* of the presented \mathbf{F} -RB method with that of standard \mathbf{u} -RB schemes. First of all, both methods share a *quadratic* dependence of their Jacobi matrices on the number of modes, N . Therefore, the assembly of the Jacobian is usually the most costly operation. Secondly, both approaches' complexities suffer a linear dependency on the number of quadrature points, N_{qp} . In the displacement-based approach, this is included in the assembly of the residual and of the stiffness, which relate to the factor 9 and 9^2 , respectively. Thirdly, the novel \mathbf{F} -RB scheme spares the computational overhead associated with FE formulations $\underline{r}^{\text{FE}}$ and $\underline{\underline{K}}^{\text{FE}}$. More details on this matter are currently being investigated.

Table 6.1: Algorithmic complexities of the well-established \mathbf{u} -based RB method and the novel \mathbf{F} -based RB method. In any case, N denotes the size of the RB. The assembly of the FE residual and of the FE stiffness depend on N_{qp} .

RB Method	Quantity	Complexity
\mathbf{F} -based	$\underline{r} = \sum_{p=1}^{N_{\text{qp}}} \underline{\underline{B}}(\mathbf{X}_p)^\top \underline{P}(\mathbf{X}_p) w_p$	$\mathcal{O}(9NN_{\text{qp}})$
	$\underline{\underline{D}} = \sum_{p=1}^{N_{\text{qp}}} \underline{\underline{B}}(\mathbf{X}_p)^\top \underline{\underline{C}}(\mathbf{X}_p) \underline{\underline{B}}(\mathbf{X}_p) w_p$	$\mathcal{O}([9^2N + 9N^2]N_{\text{qp}})$
\mathbf{u} -based	$\underline{r} = \left(\underline{\underline{B}}^{\text{u}}\right)^\top \underline{r}^{\text{FE}}$	$\mathcal{O}(NN_{\text{DOF}})$ + assembly of $\underline{r}^{\text{FE}}$
	$\underline{\underline{D}} = \left(\underline{\underline{B}}^{\text{u}}\right)^\top \underline{\underline{K}}^{\text{FE}} \underline{\underline{B}}^{\text{u}}$	$\mathcal{O}(NN_{\text{DOF}}^2 + N^2N_{\text{DOF}})$ + assembly of $\underline{\underline{K}}^{\text{FE}}$

Outlook

Future research should aim at an application of the introduced Reduced Basis method within realistic two-scale simulations, in analog to [Fritzen & Kunc, 2018c, Fritzen & Hodapp, 2016, Rambauser et al., 2019, Kochmann et al., 2017]. Hyperreduction methods, cf. Ryckelynck [2005], might give rise to additional speed-ups in the online phase. Further, modifications of the cutoff function, ϕ , should be investigated—a function with compact support might be more appropriate. The construction of the RB from large sets of snapshots is computationally intense, as much data needs to

be processed. In the above examples, the POD consumed multiple hours of time. Hierarchical approximations, such as Himpe et al. [2018], might mitigate the effects by enabling parallel computations. Overall, the long-term perspective is to extend this RB framework efficiently to the context of dissipative materials.

6.5.2 Discussion of the Sampling Strategy

The proposed sampling strategy is meant to serve as a template. As exemplified in Section 6.4.1, the samplings can be modified and still lead to a coverage of the set of macroscopic boundary conditions that is sufficient for the problem at hand. The example of Section 6.4.2 took this idea further and showed that it might not even be necessary to sample the macroscopic determinant. Hence, the sampling can sometimes be reduced to the five-dimensional subspace of isochoric macroscopic stretch tensors.

In any case, the exact choice of both the inputs to Algorithm 6.1 and the distributions of the deviatoric amplitudes and the macroscopic determinants remains to be based on knowledge and sophisticated guesses, at least at the current state of the art. Further research on this matter might lead to a refined alternative to Algorithm 6.1, possibly involving the evaluation of error estimators.

Author Contributions

Conceptualization, Oliver Kunc and Felix Fritzen; Funding acquisition, Felix Fritzen; Methodology, Oliver Kunc and Felix Fritzen; Project administration, Felix Fritzen; Software, Oliver Kunc; Supervision, Felix Fritzen; Writing—original draft, Oliver Kunc and Felix Fritzen.

Funding

This research was funded by Deutsche Forschungsgemeinschaft (DFG) within the Emmy–Noether programm under grant DFG-FR2702/6.

Conflicts of Interest

The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

RB	Reduced Basis
FE(M)	Finite Element (Method)
POD	Proper Orthogonal Decomposition
DOF	degree(s) of freedom
FOM	full-order model
s.p.d.	symmetric positive definite
DDMS	Dilatational-Deviatoric Multiplicative Split

6.6 Appendix

6.6.1 Material Objectivity

The transformation behavior of the components of the stiffness tensor \mathbb{C} is now deduced. To this end, Green's strain tensor $\bar{\mathbf{E}} = \frac{1}{2}(\bar{\mathbf{F}}^\top \bar{\mathbf{F}} - \mathbf{I})$, the corresponding stored energy density function $\bar{W}^E(\bar{\mathbf{E}}) = \bar{W}(\bar{\mathbf{F}})$, the second Piola–Kirchhoff stress $\bar{\mathbf{S}} = \partial \bar{W}^E / \partial \bar{\mathbf{E}}|_{\bar{\mathbf{E}}}$, and the corresponding stiffness tensor $\bar{\mathbb{C}}^E = \partial^2 \bar{W}^E / (\partial \bar{\mathbf{E}})^2|_{\bar{\mathbf{E}}}$ are introduced. Starting from the well-known relationship $\bar{\mathbf{P}} = \bar{\mathbf{F}} \bar{\mathbf{S}}$ between $\bar{\mathbf{S}}$ and the first Piola–Kirchhoff stress $\bar{\mathbf{P}} = \partial \bar{W} / \partial \bar{\mathbf{F}}|_{\bar{\mathbf{F}}}$ (see for instance Bertram [2008]), we express the components of $\bar{\mathbb{C}}$ in terms of those of $\bar{\mathbf{S}}$ and of $\bar{\mathbb{C}}^E$:

$$\bar{C}_{ijkl} = \frac{\partial^2 \bar{W}}{\partial \bar{\mathbf{F}}_{ij} \partial \bar{\mathbf{F}}_{kl}} = \frac{\partial \bar{P}_{ij}}{\partial \bar{\mathbf{F}}_{kl}} = \sum_{m=1}^3 \frac{\partial \bar{F}_{im} \bar{S}_{mj}}{\partial \bar{\mathbf{F}}_{kl}} \quad (6.51)$$

$$= \sum_{m=1}^3 \left(\delta_{ik} \delta_{lm} \bar{S}_{mj} + \bar{F}_{im} \frac{\partial \bar{S}_{mj}}{\partial \bar{\mathbf{F}}_{kl}} \right) \quad (6.52)$$

$$= \delta_{ik} \bar{S}_{lj} + \sum_{m,n,o=1}^3 \bar{F}_{im} \frac{\partial \bar{S}_{mj}}{\partial \bar{E}_{no}} \frac{\partial \bar{E}_{no}}{\partial \bar{\mathbf{F}}_{kl}} \quad (6.53)$$

$$= \delta_{ik} \bar{S}_{lj} + \sum_{m,n,o=1}^3 \bar{F}_{im} \bar{C}_{mjno}^E \frac{\partial \bar{E}_{no}}{\partial \bar{\mathbf{F}}_{kl}} \quad (6.54)$$

$$= \delta_{ik} \bar{S}_{lj} + \sum_{m,p=1}^3 \bar{F}_{im} \bar{C}_{mjpl}^E \bar{F}_{kp} .$$

In the last step, the minor symmetry $\bar{C}_{mjno}^E = \bar{C}_{mjon}^E$ has been exploited, and $i, j, k, l = 1, 2, 3$ above and throughout. From this, the inverse relation

$$\bar{C}_{ijkl}^E = - \left(\bar{U}^{-2} \right)_{ik} \bar{S}_{lj} + \sum_{m,n=1}^3 \left(\bar{\mathbf{F}}^{-1} \right)_{im} \bar{C}_{mjnl} \left(\bar{\mathbf{F}}^{-\top} \right)_{nk} \quad (6.55)$$

can be derived. The fact that Green's strain tensor is frame invariant, i.e. $\overline{\mathbf{E}}(\overline{\mathbf{R}}\overline{\mathbf{U}}) = \overline{\mathbf{E}}(\overline{\mathbf{U}})$, implies that both the left hand side $\overline{\mathbf{C}}_{ijkl}^{\mathbf{E}} = \overline{\mathbf{C}}_{ijkl}^{\mathbf{E}}(\overline{\mathbf{E}})$ and the second Piola–Kirchhoff stress $\overline{\mathbf{S}}_{ij} = \overline{\mathbf{S}}_{ij}(\overline{\mathbf{E}})$ are independent of $\overline{\mathbf{R}}$. This is in contrast to $\overline{\mathbf{C}}_{mjnl} = \overline{\mathbf{C}}_{mjnl}(\overline{\mathbf{R}}\overline{\mathbf{U}})$ from which follows that

$$\sum_{m,n=1}^3 \left(\overline{\mathbf{F}}^{-1}\right)_{im} \overline{\mathbf{C}}_{mjnl}(\overline{\mathbf{R}}\overline{\mathbf{U}}) \left(\overline{\mathbf{F}}^{-\top}\right)_{nk} = \sum_{m,n=1}^3 \left(\overline{\mathbf{U}}^{-1}\right)_{im} \overline{\mathbf{C}}_{mjnl}(\overline{\mathbf{U}}) \left(\overline{\mathbf{U}}^{-\top}\right)_{nk}. \quad (6.56)$$

By contraction of the indices i and k with the second index of $\overline{\mathbf{F}}$ and the first index of $\overline{\mathbf{F}}^{\top}$, respectively, equation (6.42) follows.

6.6.2 Effective Material Responses of the RB

Let \mathbb{I} denote the fourth order identity tensor and let the arguments of the \mathbf{F} -RB approximation (6.23) be omitted, i.e. here $\mathbf{F}^{\text{RB}} = \mathbf{F}^{\text{RB}}(\mathbf{X}; \overline{\mathbf{F}})$. Its derivative with respect to the boundary condition $\overline{\mathbf{F}}$ is

$$\frac{\partial \mathbf{F}^{\text{RB}}}{\partial \overline{\mathbf{F}}} = \mathbb{I} + \sum_{i=1}^N \mathbf{B}^{(i)} \otimes \frac{\partial \xi_i^*}{\partial \overline{\mathbf{F}}}(\overline{\mathbf{F}}). \quad (6.57)$$

6.6.3 Effective Stress

$$\begin{aligned} \overline{\mathbf{P}}^{\text{RB}} &= \frac{\partial \overline{W}^{\text{RB}}}{\partial \overline{\mathbf{F}}} = \frac{\partial}{\partial \overline{\mathbf{F}}} \langle W^{\text{RB}} \rangle = \left\langle \frac{\partial W^{\text{RB}}}{\partial \mathbf{F}} \cdot \frac{\partial \mathbf{F}^{\text{RB}}}{\partial \overline{\mathbf{F}}} \right\rangle \\ &\stackrel{(6.57)}{=} \langle \mathbf{P}^{\text{RB}} \rangle + \sum_{i=1}^N \left\langle \mathbf{P}^{\text{RB}} \cdot \left(\mathbf{B}^{(i)} \otimes \frac{\partial \xi_i^*}{\partial \overline{\mathbf{F}}}(\overline{\mathbf{F}}) \right) \right\rangle \\ &= \langle \mathbf{P}^{\text{RB}} \rangle + \sum_{i=1}^N \left\langle \mathbf{P}^{\text{RB}} \cdot \mathbf{B}^{(i)} \right\rangle \otimes \frac{\partial \xi_i^*}{\partial \overline{\mathbf{F}}}(\overline{\mathbf{F}}) \quad \stackrel{(6.28)}{=} \langle \mathbf{P}^{\text{RB}} \rangle \end{aligned} \quad (6.58)$$

6.6.4 Effective Stiffness

$$\overline{\mathbb{C}}^{\text{RB}} = \frac{\partial \overline{\mathbf{P}}^{\text{RB}}}{\partial \overline{\mathbf{F}}} \stackrel{(6.58)}{=} \frac{\partial}{\partial \overline{\mathbf{F}}} \langle \mathbf{P}^{\text{RB}} \rangle = \left\langle \frac{\partial^2 W^{\text{RB}}}{\partial \overline{\mathbf{F}} \partial \mathbf{F}} \right\rangle = \left\langle \frac{\partial^2 W^{\text{RB}}}{\partial \mathbf{F}^2} \cdot \frac{\partial \mathbf{F}^{\text{RB}}}{\partial \overline{\mathbf{F}}} \right\rangle \quad (6.59)$$

$$\begin{aligned}
&\stackrel{(6.57)}{=} \langle \mathbb{C}^{\text{RB}} \rangle + \sum_{i=1}^N \left\langle \mathbb{C}^{\text{RB}} \cdot \left(\mathbf{B}^{(i)} \otimes \frac{\partial \xi_i^*}{\partial \underline{\mathbf{F}}}(\underline{\mathbf{F}}) \right) \right\rangle \\
&= \langle \mathbb{C}^{\text{RB}} \rangle + \sum_{i=1}^N \left\langle \mathbb{C}^{\text{RB}} \cdot \mathbf{B}^{(i)} \right\rangle \otimes \frac{\partial \xi_i^*}{\partial \underline{\mathbf{F}}}(\underline{\mathbf{F}})
\end{aligned} \tag{6.59 continued}$$

For $\frac{\partial \xi_i^*}{\partial \underline{\mathbf{F}}}(\underline{\mathbf{F}})$, we demand that the residual $r_i(\underline{\mathbf{F}}, \underline{\xi})$ from (6.28) is stable with respect to the boundary condition $\underline{\mathbf{F}}$ when converged to $r_i(\underline{\mathbf{F}}, \underline{\xi}^*(\underline{\mathbf{F}})) = 0$,

$$\begin{aligned}
\frac{\partial r_i}{\partial \underline{\mathbf{F}}}(\underline{\mathbf{F}}, \underline{\xi}^*(\underline{\mathbf{F}})) &= \left\langle \mathbf{B}^{(i)} \cdot \frac{\partial \mathbf{P}^{\text{RB}}}{\partial \underline{\mathbf{F}}} \right\rangle \\
&= \left\langle \mathbf{B}^{(i)} \cdot \left(\frac{\partial \mathbf{P}^{\text{RB}}}{\partial \underline{\mathbf{F}}} \frac{\partial \underline{\mathbf{F}}^{\text{RB}}}{\partial \underline{\mathbf{F}}} \right) \right\rangle \\
&= \left\langle \mathbf{B}^{(i)} \cdot \mathbb{C}^{\text{RB}} \right\rangle + \sum_{j=1}^N \left\langle \mathbf{B}^{(i)} \cdot \mathbb{C}^{\text{RB}} \cdot \frac{\partial \underline{\mathbf{F}}^{\text{RB}}}{\partial \xi_j^*} \frac{\partial \xi_j^*}{\partial \underline{\mathbf{F}}} \right\rangle \\
&= \left\langle \mathbf{B}^{(i)} \cdot \mathbb{C}^{\text{RB}} \right\rangle + \sum_{j=1}^N \underbrace{\left\langle \mathbf{B}^{(i)} \cdot \mathbb{C}^{\text{RB}} \cdot \mathbf{B}^{(j)} \right\rangle}_{D_{ij}} \frac{\partial \xi_j^*}{\partial \underline{\mathbf{F}}} = \mathbf{0}.
\end{aligned} \tag{6.60}$$

It follows that

$$\frac{\partial \xi_j^*}{\partial \underline{\mathbf{F}}}(\underline{\mathbf{F}}) = - \sum_{i=1}^N \left(\underline{\underline{D}}^{-1} \right)_{ij} \left\langle \mathbf{B}^{(i)} \cdot \mathbb{C}^{\text{RB}} \right\rangle. \tag{6.61}$$

6.6.5 Basis for Symmetric Traceless Second Order Tensors

$$\begin{aligned}
\underline{\underline{Y}}^{(1)} &= \sqrt{\frac{1}{6}} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} & \underline{\underline{Y}}^{(2)} &= \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \\
\underline{\underline{Y}}^{(3)} &= \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \underline{\underline{Y}}^{(4)} &= \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\
\underline{\underline{Y}}^{(5)} &= \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}
\end{aligned} \tag{6.62}$$

Chapter 7:

Third publication – Generation of energy-minimizing point sets on spheres and their application in mesh-free interpolation and differentiation

Original publication:

Kunc, O. & Fritzen, F.: Generation of energy-minimizing point sets on spheres and their application in mesh-free interpolation and differentiation. *Advances in Computational Mathematics* **45** (2019c), 3021–3056, Springer Science and Business Media LLC, doi:10.1007/s10444-019-09726-5

Abstract. It is known that discrete sets of uniformly distributed points on the hypersphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ can be obtained from minimizing the energy functional corresponding to Riesz s -kernels $k_s(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^{-s}$ ($s > 0$) or the logarithmic kernel $k_{\log}(\mathbf{x}, \mathbf{y}) = -\log\|\mathbf{x} - \mathbf{y}\| + \log 2$. We prove the same for the kernel $k_{\text{LOG}}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|(\log \frac{\|\mathbf{x} - \mathbf{y}\|}{2} - 1) + 2$ which is a front-extension of the sequence of derivatives $k_{\log}, k_1, k_2, k_3, \dots$, up to sign and constants. The boundedness of the kernel simplifies the classical potential-theoretical proof of the asymptotic uniformity of the point distributions. Still, the property of a singular derivative for $\mathbf{x} \rightarrow \mathbf{y}$ is preserved, with the physical interpretation of infinite repulsive forces for touching particles. The quality of the resulting point distributions is exemplary compared to that of Riesz- and classical logarithmic point sets, and found to be competitive.

Originally motivated by problems of high-dimensional data, the applicability of LOG-optimal point sets with a novel concentric interpolation and differentiation scheme is demonstrated. The method is significantly optimized by the introduction of symmetrized kernels for both the generation of the minimum energy points and the spherical basis functions.

Both the point generation and the Concentric Interpolation software are available as Open Source software under the GNU GPL License v3 and selected point sets are provided.

Keywords: points on spheres, uniform distribution, minimum energy points, Riesz kernel, logarithmic kernel, mesh-free interpolation, kernel interpolation, numerical differentiation, high-dimensional data

Mathematics Subject Classification (2010): 31B05, 31B10, 31B15, 31E05, 31C20, 65D05, 65D25

7.1 Introduction

In the first part of the present work, the existence, the generation and the properties of energy minimizing point sets on spheres $\mathbb{S}^d = \{\mathbf{x} \in \mathbb{R}^{d+1} : \|\mathbf{x}\| = 1\}$, where $\|\mathbf{x}\|$ denotes the Euclidean norm of \mathbf{x} , in arbitrary dimension $d \geq 1$ is studied.⁹

The application of these point sets as support sites for data-assisted, mesh-free methods represents the second part of this work. An easy to use interpolation scheme including the automatic and computationally efficient differentiation of the data is provided.

7.1.1 Part 1: Generation of LOG-energy minimizing point sets

Directional data and sampling over the sphere are required in earth sciences, medicine, psychology and physics, see e.g. Mardia & Jupp [1999]. Further, sampling the high-dimensional state space of physical problems in order to identify pronounced material properties requires the non-trivial generation of sample directions for $d \geq 4$, see e.g. Fritzen & Kunc [2018c]. Other fields of research applying certain distributions of directions include microstructure characterization: Dietrich et al. [2012]; anisotropy evaluation: Pérez-Ramírez et al. [2013]; crystallography: Morawiec [2004]; flow visualization: Ma et al. [2013]; electromagnetism: Martini et al. [2010]; astronomy: Roşca [2010]; meteorology: Staniforth & Thuburn [2011]; distributed computing: Andelfinger et al. [2014] and image compression: Lovisolo & Da Silva [2001]. It is therefore of notable importance to develop schemes for the generation of directions with specific properties. In many cases, uniformity of the distribution of directions is important and the main objective is to reduce the number of samples while meeting accuracy requirements of the subsequent computational schemes.

Directions in \mathbb{R}^{d+1} can be represented as points on the unit sphere \mathbb{S}^d . The generation of point sets on \mathbb{S}^d with particular distribution has been an active field of research for many years, cf., e.g., Saff & Kuijlaars [1997] for a brief introduction.

Among the characteristics of point sets that might be desired, asymptotic uniformity is a prominent example. The term was introduced more than a century ago, probably by Weyl [1916], and initiated a branch of research that is still seeing notable attention today. A physical approach for the approximation of asymptotically uniform point sets was the

⁹In the present work the name sphere is used for arbitrary dimension d , i.e. it is used as a synonym for hypersphere. Also, we call \mathbb{S}^d the d -sphere.

minimization of an assumed electric potential. As a generalization of this special case, the Riesz kernel $k_s(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^{-s}$ ($s > 0$) is central to the theory.

By standard arguments of potential theory, cf. Landkof [1972], it is well-known that point sets minimizing the energy corresponding to the Riesz s -kernel k_s with $0 < s < d$ are asymptotically uniformly distributed. With some modifications, cf. Landkof [1972], the arguments prove the same for the log-kernel $k_{\log}(\mathbf{x}, \mathbf{y}) = -\log\|\mathbf{x} - \mathbf{y}\| + \log 2$, where the non-standard additive constant $\log 2$ does not affect the validity of the proof. These cases are part of the *potential-theoretic regime*, as the double integral of the respective kernels is finite, cf. Brauchart & Grabner [2015]. The case $s \geq d$ comes along with integrability issues. In Hardin & Saff [2005] it is shown that in this *hyper-singular regime*, too, *energy-optimality is sufficient for asymptotic uniformity of the distribution*.

It is interesting to note that the Riesz-kernels for integral values of s and the log-kernel can be connected in the following way: let the Euclidean distance of two points be denoted by $z(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$, and let the kernel functions with respect to this scalar argument be $\tilde{k}_{\log}(z)$, $\tilde{k}_s(z)$. Then each element of the series

$$\left((-1)^n a_n \frac{\partial^n \tilde{k}_{\log}(z)}{(\partial z)^n} \right)_{n \in \mathbb{N}_{\geq 0}} = \tilde{k}_{\log}(z), \tilde{k}_1(z), \tilde{k}_2(z), \tilde{k}_3(z), \dots, \quad (7.1)$$

with the factor $a_0 = 1$ and $a_n = \frac{1}{(n-1)!}$ for $n \in \mathbb{N}$, is known to lead to asymptotically uniform point distributions. **The main contribution of this work** is to show that the same holds for a front-extension of the above series by the kernel $\tilde{k}_{\text{LOG}}(z) = z(\log \frac{z}{2} - 1) + 2$, which satisfies $\partial \tilde{k}_{\text{LOG}}(z) / \partial z = -\tilde{k}_{\log}(z)$. Thus, to the best knowledge of the authors, the set of potential-theoretic cases is extended.

7.1.2 Part 2: Mesh-free interpolation and differentiation

Energy minimizing point sets are not only appealing due to their asymptotic uniformity, which could also be realized by a proper sequence of random points. Rather, their beauty lies in their inherent “evenness” or “regularity”, two terms that intuitively describe multiple formal definitions. The goal of the second part of the present work is to exemplarily prove the applicability of LOG-optimal points to a novel concentric interpolation and differentiation scheme. Based on spherical basis functions, a sense of “regularity” is a property one would expect sets of supporting points to have in order to be well-suited for this application.

The Concentric Interpolation method addresses the problem of high-dimensional data. In science and engineering, modern experimental methods produce ever increasing amounts of data. Novel in silico methods can efficiently produce data at exactly predefined positions within the admissible parameter domain, which may be of large dimension. However, the processing of the data sets often remains challenging, e.g., if differential operators such as the divergence or gradient should be applied to discrete data, cf., e.g., Esmailbeigi et al. [2017]. The higher the dimension and the greater the size

of the data set, the more desirable it is to have a robust meshless interpolation and differentiation method at hand. Generality of the method, also known as “black-box” character, contributes to the ease of use and to avoiding operational errors.

The Concentric Interpolation method provides an application of the point sets from the first part in terms of a mesh-free interpolation and differentiation scheme on the sphere \mathbb{S}^d with an extension to the surrounding space \mathbb{R}^{d+1} .

7.1.3 Outline

The first part begins with Section 7.2 where the theorem of asymptotic uniformity of LOG-optimal point sets is set up. The theorem is proved in Section 7.3. In Section 7.4, the algorithm behind the provided Matlab code for the generation of sets of energy-minimizing points is described. The first part concludes with Section 7.5, where the quality of LOG-optimal point sets and the evolution of their energy is numerically studied.

The second part begins with Section 7.6. There, the Concentric Interpolation scheme is introduced. The suitability of various kinds of energy-optimal point sets is numerically exemplified in Section 7.7. The paper closes with a summary and discussion in Section 7.8.

7.2 Minimum energy point sets

7.2.1 Kernels

By kernels, we denote a family of symmetric functions $k_s : \mathbb{S}^d \times \mathbb{S}^d \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ which are parametrized by $s \in \mathbb{R}_{\geq 0} \cup \{\log, \text{LOG}\}$ and which are defined by

$$k_s(\mathbf{x}, \mathbf{y}) = \begin{cases} \frac{1}{\|\mathbf{x} - \mathbf{y}\|^s} & \text{if } s \geq 0, \\ \log \frac{2}{\|\mathbf{x} - \mathbf{y}\|} & \text{if } s = \log, \\ \|\mathbf{x} - \mathbf{y}\|(\log \frac{\|\mathbf{x} - \mathbf{y}\|}{2} - 1) + 2 & \text{if } s = \text{LOG}. \end{cases} \quad (7.2)$$

If $s > 0$ in the first case, the kernel k_s is usually denoted a Riesz-kernel. In the second case, k_{\log} is usually referred to as the logarithmic kernel, where the natural logarithm is employed. Note the constant $\log 2$ by which this definition deviates from the standard literature. In order to avoid possible confusions, we will denote this the “log-kernel” and speak of the “log-case”, in contrast to the capitalized “LOG-kernel” of the “LOG-case” for $s = \text{LOG}$. In the latter, the definition is continuously extended for equal arguments, $k_{\text{LOG}}(\mathbf{x}, \mathbf{x}) = 2$. In contrast to this, the Riesz- and the log-kernel are only lower semi-continuous and exhibit a singularity, $\lim_{\mathbf{y} \rightarrow \mathbf{x}} k_s(\mathbf{x}, \mathbf{y}) = \infty$. The case $s = 0$ is only

of formal interest and does not deserve special nomenclature. As the case $s = \text{LOG}$ is introduced by the current work, the other cases ($s \neq \text{LOG}$) are referred to as *classic*.

The family of kernels k_s depends only on the Euclidean distance $z = \|\mathbf{x} - \mathbf{y}\|$. Therefore, it is useful to introduce the univariate function $\tilde{k}_s : [0, 2] \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ and its derivative \tilde{k}'_s , defined by

$$\tilde{k}_s(z) = k_s(\mathbf{x}, \mathbf{y}) \quad \tilde{k}'_s(z) = \frac{d\tilde{k}_s}{dz}(z). \quad (7.3)$$

Some of the introduced kernels are visualized in Figure 7.1.

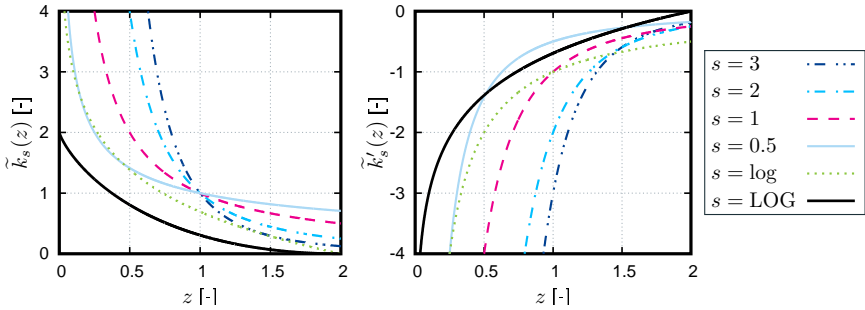


Figure 7.1: *Left:* various Riesz-kernels, the log-kernel and the LOG-kernel. The latter (solid black thick line) has infinite slope at $z = 0$ but is bounded. *Right:* respective derivatives.

The univariate kernel function may be symmetrized with respect to $z = \sqrt{2}$, corresponding to a mirror symmetry at the origin, independently of s , via

$$\tilde{k}_s^{\text{sym}}(z) = \tilde{k}_s(z) + \tilde{k}_s(\sqrt{4 - z^2}). \quad (7.4)$$

The expression $\sqrt{4 - z^2}$ arises from the curvature of the manifold \mathbb{S}^d and measures the Euclidean distance to an antipode, e.g. if $z = \|\mathbf{x} - \mathbf{y}\|$, then $\sqrt{4 - z^2} = \|\mathbf{x} + \mathbf{y}\|$. This function \tilde{k}_s^{sym} is the effective kernel resulting from two antipodal centers. Such symmetrized kernels are exemplarily visualized in Figure 7.2.

7.2.2 Measures, potentials, energy, and asymptotic uniformity

Let \mathcal{M} denote the set of non-negative Borel-measures on \mathbb{S}^d with respect to the topology that is induced by the Euclidean metric $z(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{d+1}$. Additionally, the geodesic metric $z_g(\mathbf{x}, \mathbf{y}) = \text{acos}(\mathbf{x}^\top \mathbf{y})$ is introduced, $\mathbf{x}, \mathbf{y} \in \mathbb{S}^d$, which relates to the former metric via $z = 2 \sin(z_g/2)$. Let $\mathcal{B}(\mathbb{S}^d)$ be the Borel sigma algebra, and \mathcal{H}^d

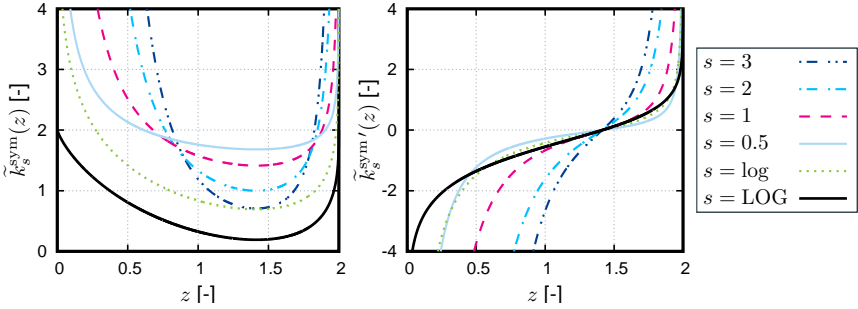


Figure 7.2: *Left:* symmetrized kernels of Riesz, log, and LOG type. The minimum value of all kernels is attained at $z = \sqrt{2}$. *Right:* respective derivatives with respect to z .

denote the d -dimensional Hausdorff measure on the metric space (\mathbb{S}^d, z_g) . Then the *uniform measure* σ , defined by

$$\sigma(A) = \frac{\mathcal{H}^d(A)}{\mathcal{H}^d(\mathbb{S}^d)}, \quad (7.5)$$

where $A \in \mathcal{B}(\mathbb{S}^d)$, is a common example for an element of \mathcal{M} . Another element of \mathcal{M} is the Dirac measure corresponding to a single point $\mathbf{x}^* \in \mathbb{S}^d$,

$$\delta_{\mathbf{x}^*}(A) = \begin{cases} 1 & \text{if } A \ni \mathbf{x}^* \\ 0 & \text{else} \end{cases}, \quad (7.6)$$

with $A \in \mathcal{B}(\mathbb{S}^d)$. These two examples both belong to the subset of normalized measures $\overline{\mathcal{M}} = \{\mu \in \mathcal{M} \mid \mu(\mathbb{S}^d) = 1\}$. Henceforth, we shall restrict our attention to $\overline{\mathcal{M}}$, which contains all measures that are of interest for our purposes, although many definitions equally apply to the larger set \mathcal{M} .

Convex combinations of countably many Dirac measures shall be called *discrete* measures, while all others are referred to as *continuum* measures. The term *continuum* is to be understood in the sense of uncountably large sets, not addressing continuity. The *support* of a measure, $\text{supp}(\mu)$, is defined as the complement in \mathbb{S}^d of the largest open set $A \in \mathcal{B}(\mathbb{S}^d)$ with $\mu(A) = 0$. A sequence of measures $(\mu_n)_{n \in \mathbb{N}} \subset \overline{\mathcal{M}}$ *weakly converges* to $\mu \in \overline{\mathcal{M}}$, denoted by $\mu_n \xrightarrow{w} \mu$, if for all *test functions* $f \in C_K = \{\text{compactly supported, continuous functions} : \mathbb{S}^d \rightarrow \mathbb{R}\}$

$$\lim_{n \rightarrow \infty} \int_{\mathbb{S}^d} f(\mathbf{x}) d\mu_n(\mathbf{x}) = \int_{\mathbb{S}^d} f(\mathbf{x}) d\mu(\mathbf{x}). \quad (7.7)$$

Finite sets of arbitrary but fixed points on the hypersphere will be denoted by $X_N = (\mathbf{x}_1, \dots, \mathbf{x}_N) \subset \mathbb{S}^d$. Although the order within point sets is irrelevant for our purposes, the N -tuple notation $(\mathbf{x}_1, \dots, \mathbf{x}_N)$ is used (in contrast to the notation $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ of

an unordered set) in order to unambiguously allow for identical elements, i.e. possibly $\mathbf{x}_i = \mathbf{x}_j$ for some $i \neq j$. For convenience, only the term (*point set*) is used hereafter. Further, the notation of a sequence of point sets $(X_N)_{N \in \mathbb{N}}$ allows for non-hierarchical sets, i.e. possibly $X_N \not\subset X_{N+1}$. For every point set X_N , there is a corresponding discrete normalized measure $\mu_{X_N} \in \overline{\mathcal{M}}$ with equal weights,

$$\mu_{X_N} = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{x}_i}. \quad (7.8)$$

Discrete measures of this form are *weakly dense* in $\overline{\mathcal{M}}$, meaning that for all $\mu \in \overline{\mathcal{M}}$ there exists a sequence of point sets $(X_N)_{N \in \mathbb{N}}$ such that $\mu_{X_N} \rightarrow \mu$.

For fixed kernel k_s and measure $\mu \in \overline{\mathcal{M}}$, the function $U_s^\mu : \mathbb{S}^d \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ is defined by

$$U_s^\mu(\mathbf{x}) = \int_{\mathbb{S}^d} k_s(\mathbf{x}, \mathbf{y}) \, d\mu(\mathbf{y}) \quad (7.9)$$

and is denoted the *s-potential of μ at $\mathbf{x} \in \mathbb{S}^d$* . The functional $I_s : \overline{\mathcal{M}} \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$, called the *s-energy of μ* , is associated with the *s-potential* by

$$I_s(\mu) = \int_{\mathbb{S}^d} U_s^\mu(\mathbf{x}) \, d\mu(\mathbf{x}) = \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} k_s(\mathbf{x}, \mathbf{y}) \, d\mu(\mathbf{x}) \, d\mu(\mathbf{y}), \quad (7.10)$$

which by the Fubini theorem is finite only if $U_s^\mu(\mathbf{x})$ is finite μ -almost everywhere (μ -a.e.). Analogous conditions hold for the *mutual s-energy* of the measures $\mu, \nu \in \overline{\mathcal{M}}$, $I_s : \overline{\mathcal{M}} \times \overline{\mathcal{M}} \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$, defined by

$$I_s(\mu, \nu) = \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} k_s(\mathbf{x}, \mathbf{y}) \, d\mu(\mathbf{x}) \, d\nu(\mathbf{y}) = \int_{\mathbb{S}^d} U_s^\nu(\mathbf{x}) \, d\mu(\mathbf{x}) = \int_{\mathbb{S}^d} U_s^\mu(\mathbf{y}) \, d\nu(\mathbf{y}). \quad (7.11)$$

This is notationally distinct from the *s-energy* of a single measure by the number of arguments, and relates to it by $I_s(\mu) = I_s(\mu, \mu)$.

In the discrete setting, the finiteness of the *s-energy*

$$I_s(\mu_{X_N}) = \frac{1}{N^2} \sum_{i,j=1}^N k_s(\mathbf{x}_i, \mathbf{x}_j) \quad (7.12)$$

depends on the parameter s . It is finite if $s \in \{0, \text{LOG}\}$. In fact, $I_0 \equiv 1$ for any discrete measure μ_{X_N} . For the other classical cases, $s \in \mathbb{R}_{>0} \cup \{\log\}$, one usually considers the *modified s-energy*

$$\tilde{I}_s(\mu_{X_N}) = \frac{1}{N(N-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^N k_s(\mathbf{x}_i, \mathbf{x}_j) = \frac{2}{N(N-1)} \sum_{1 \leq i < j \leq N} k_s(\mathbf{x}_i, \mathbf{x}_j) \quad (7.13)$$

which is only meaningful for discrete measures $\mu_{X_N} \in \overline{\mathcal{M}}$. In the context of Riesz-kernels it is naturally assumed that the point set X_N contains no repeated elements. As we mostly focus on the newly introduced LOG-case, this assumption is not made at this point. Its minimum value,

$$\tilde{\mathcal{I}}_s(N) = \min_{X_N \in \mathbb{S}^d} \tilde{I}_s(\mu_{X_N}), \quad (7.14)$$

exists due to the boundedness of the kernels, $0 \leq k_s$, and the compactness of the domain \mathbb{S}^d . It is well-known, cf. Brauchart & Grabner [2015], that the classical kernels are related via the s -expansion of the modified energy at $s = 0$,

$$\tilde{I}_s(\mu_{X_N}) = \tilde{I}_0 + s(\tilde{I}_{\log}(\mu_{X_N}) - \log(2)) + o(s) \quad (s \rightarrow 0). \quad (7.15)$$

Note that the minimum LOG-energy with respect to point sets of size N exists,

$$\mathcal{I}_{\text{LOG}}(N) = \min_{X_N \subset \mathbb{S}^d} I_{\text{LOG}}(\mu_{X_N}). \quad (7.16)$$

and denote a point set that realizes this minimum N -point LOG-energy by

$$\Xi_N = \arg \min_{X_N \subset \mathbb{S}^d} I_{\text{LOG}}(\mu_{X_N}) = (\xi_1^{(N)}, \dots, \xi_N^{(N)}). \quad (7.17)$$

Such point sets are coined *energy optimal* or *energy minimizing*. Note that energy minimizing point sets generally need not be unique, not even up to orthogonal transformations.

A sequence of point sets $(X_N)_{N \in \mathbb{N}}$ is said to be *asymptotically uniformly distributed* (a.u.d.), cf. Brauchart & Grabner [2015], if it weakly converges to the uniform measure, i.e. $\mu_{X_N} \rightharpoonup \sigma$. It is well-known, cf. Brauchart & Grabner [2015], that energy minimizing point sets for the cases $s \in \mathbb{R}_{>0} \cup \{\log\}$ are a.u.d. The main contribution of the present work is the following

Theorem 1 *Sequences of LOG-energy optimal point sets $(\Xi_N)_{N \in \mathbb{N}}$ are asymptotically uniformly distributed, i.e. $\mu_{\Xi_N} \rightharpoonup \sigma$. \square*

The proof is given in Section 7.3. There, it is also shown that the optimal N -point energy $\mathcal{I}_{\text{LOG}}(N)$ converges to the infimum value $\mathcal{I}_{\text{LOG}}^c$ of the energy $I_{\text{LOG}}(\mu)$, i.e.

$$\lim_{N \rightarrow \infty} \mathcal{I}_{\text{LOG}}(N) = \mathcal{I}_{\text{LOG}}^c := \inf_{\mu \in \overline{\mathcal{M}}} I_{\text{LOG}}(\mu). \quad (7.18)$$

The upper index c emphasizes that continuum measures are also considered in the infimum operator. This is distinct from the definition (7.16) of $\mathcal{I}_{\text{LOG}}(N)$, where the minimum is taken from a subset of $\overline{\mathcal{M}}$ – albeit a dense subset.

7.2.3 Evaluation of the asymptotic LOG-energy

The energy of optimal point sets and its limit behavior has been subject of investigation for years, cf. e.g. Brauchart et al. [2012], Hardin & Saff [2005] and the literature mentioned therein. As shown in the proof of Theorem 1 in Section 7.3, for a.u.d. point sequences the relation

$$\lim_{N \rightarrow \infty} \mathcal{I}_{\text{LOG}}(N) = \mathcal{I}_{\text{LOG}}^c = \int_{\mathbb{S}^d \times \mathbb{S}^d} k_{\text{LOG}}(\mathbf{x}, \mathbf{y}) \, d\sigma(\mathbf{x}) \, d\sigma(\mathbf{y}) \quad (7.19)$$

holds. For the potential-theoretic Riesz-case, $0 < s < d$, is known that the minimum energy $\mathcal{I}_s^c = \inf_{\mu \in \overline{\mathcal{M}}} I_s(\mu)$ has the explicit expression

$$\mathcal{I}_s^c = 2^{d-1-s} \frac{\Gamma((d+1)/2) \Gamma((d-s)/2)}{\sqrt{\pi} \Gamma(d-s/2)}, \quad (7.20)$$

cf. [Brauchart & Grabner, 2015, (3.12)]. The hyper-singular case, $s \geq d$, is characterized by the lack of integrability, $\mathcal{I}_s^c = \infty$. A straight-forward, general, integral representation of \mathcal{I}_s^c for all potential-theoretic cases can be constructed by means of the following modification of (7.19).

By Lemma 3 of Section 7.3, the potential $U_s^\sigma(\mathbf{x})$ is σ -a.e. constant. Therefore, a single integral is sufficient,

$$\mathcal{I}_s^c = \int_{\mathbb{S}^d} \tilde{k}_s(z(\mathbf{x}, \mathbf{y})) \, d\sigma(\mathbf{y}) \quad \forall \mathbf{x} \in \mathbb{S}^d. \quad (7.21)$$

For further simplification and for later reference, the notion of a *spherical cap* is introduced,

$$C^d(\mathbf{c}, r) = \{\mathbf{x} \in \mathbb{S}^d : z_g(\mathbf{c}, \mathbf{x}) \leq r\}, \quad (7.22)$$

where $\mathbf{c} \in \mathbb{S}^d$ and $r \in [0, \pi]$ shall be called the *center* and the *radius* of the cap, respectively. Note that in contrast to the kernels, spherical caps are defined with respect to the geodesic metric z_g . Then the (geodesic) *distribution function* $F^d : [0, \pi] \rightarrow [0, 1]$ associated with the measure σ is given by

$$F^d(r) = \sigma(C(\mathbf{c}, r)), \quad (7.23)$$

which is independent of the cap center $\mathbf{c} \in \mathbb{S}^d$. For example, the distribution functions for $d = 1, 2, 3$ are

$$F^1(r) = \frac{r}{\pi}, \quad F^2(r) = \frac{1 - \cos(r)}{2} \quad \text{and} \quad F^3(r) = \frac{r - \sin(r) \cos(r)}{\pi}. \quad (7.24)$$

The respective *density function* $f^d : [0, \pi] \rightarrow \mathbb{R}_{\geq 0}$ is given by

$$f^d(r) = \frac{\partial F^d(r)}{\partial r}. \quad (7.25)$$

With this at hand, (7.21) is equivalent to the simple equation

$$\mathcal{I}_s^c = \int_0^\pi \tilde{k}_s(2 \sin(z_g/2)) f^d(z_g) dz_g. \quad (7.26)$$

Most importantly, the integration reduces to a one-dimensional integral, allowing for arbitrarily exact evaluation using numerical quadrature. This holds for all potential-theoretic cases, e.g. for all s satisfying $0 < s < d$, $s = \log$ or $s = \text{LOG}$.

7.3 Proof of the asymptotic uniformity of LOG-optimal point sets (Theorem 1)

The proof is largely based on the frequently cited standard textbook on potential theory [Landkof, 1972, pp. 131–137 and 160–163], where it was conducted for the case $0 < s < d$. The mentioned source also contains an adaption of the proof to the case $s = \log$. The boundedness of k_{LOG} implies that the case $s = \text{LOG}$ should be included in the potential-theoretic regime, too, and hints at that the proof can be adapted. As it turns out, the proof can actually be significantly simplified, such that the essential ideas can be provided here comparatively briefly. For instance, the classic potential-theoretic proof is based on the concepts of capacity and of the Fourier transform (of kernels or potentials). Both of these are unnecessary in the LOG-case. Furthermore, the finiteness of potentials and of energies is trivial to show. Also, the kernel requires neither “truncation” nor additional limit considerations related to semi-continuity. Where technical details are lengthy and require no modifications, we omit copying them and refer the interested reader to the original source Landkof [1972]. We begin by collecting helpful facts in

Corollary 2 *Let $(\mu_n)_{n \in \mathbb{N}} \subset \overline{\mathcal{M}}$ be a sequence of measures that weakly converges to $\mu \in \overline{\mathcal{M}}$.*

(a) *The sequence of LOG-potentials converges to the LOG-potential of the weak limit, i.e.*

$$\lim_{n \rightarrow \infty} U_{\text{LOG}}^{\mu_n}(\mathbf{x}) = U_{\text{LOG}}^\mu(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{S}^d. \quad (7.27)$$

(b) *The LOG-potential of any measure $\mu \in \overline{\mathcal{M}}$ is a test function, i.e. $U_{\text{LOG}}^\mu(\mathbf{x}) \in C_K$.*

(c) *The sequence of LOG-energies converges to the LOG-energy of the weak limit, i.e.*

$$\lim_{n \rightarrow \infty} I_{\text{LOG}}(\mu_n) = I_{\text{LOG}}(\mu). \quad (7.28)$$

PROOF (a) In the definition of weak convergence (7.7), choose $f(\mathbf{x}) = k_{\text{LOG}}(\mathbf{x}, \mathbf{y}^*)$ which is clearly in C_K for fixed $\mathbf{y}^* \in \mathbb{S}^d$.

(b) This follows from the continuity of k_{LOG} and from the linearity of the integral operator. In other words, the LOG-potential is a test function because the LOG-kernel $k_{\text{LOG}}(\mathbf{x}, \mathbf{y}^*)$ is a test function for fixed $\mathbf{y}^* \in \mathbb{S}^d$, cf. (a).

7.3 Proof of the asymptotic uniformity of LOG-optimal point sets (Theorem 1) 35

(c) Rewrite the energy of $\tilde{\mu}_n$ as $I_{\text{LOG}}(\mu_n) = \int_{\mathbb{S}^d} U_{\text{LOG}}^{\mu_n}(\mathbf{x}) \, d\mu_n(\mathbf{x})$. By (a), the integrand converges to $U_{\text{LOG}}^{\mu}(\mathbf{x})$. By (b), both the sequence of potentials $(U_{\text{LOG}}^{\mu_n}(\mathbf{x}))_{n \in \mathbb{N}}$ and its limit are test functions. Hence, the consecutive limit $\lim_{n \rightarrow \infty} I_{\text{LOG}}(\mu_n)$ equals $I_{\text{LOG}}(\mu)$. ■

Next, the existence and the uniqueness of a LOG-energy minimizing measure is deduced in Lemmas 1-3.

Lemma 1 *There exists $\lambda \in \overline{\mathcal{M}}$ such that for all $\mu \in \overline{\mathcal{M}}$: $I_{\text{LOG}}(\lambda) \leq I_{\text{LOG}}(\mu)$.* □

PROOF Due to the bounds $0 \leq k_{\text{LOG}} \leq 2$, the compactness of \mathbb{S}^d , the global support of k_{LOG} and the fact that $k_{\text{LOG}}(z)$ does not take its minimum value at $z = 0$, the infimum value of the energy $\mathcal{I}_{\text{LOG}}^c$ exists and lies within the bounds

$$0 < \mathcal{I}_{\text{LOG}}^c < 2. \quad (7.29)$$

Therefore, there exists a sequence of energy minimizing measures,

$$(\mu_n)_{n \in \mathbb{N}} \subset \overline{\mathcal{M}} : \quad \lim_{n \rightarrow \infty} I_{\text{LOG}}(\mu_n) = \mathcal{I}_{\text{LOG}}^c. \quad (7.30)$$

As this sequence is a subset of $\overline{\mathcal{M}}$, it contains a weakly convergent subsequence,

$$(\mu_{n_k})_{k \in \mathbb{N}} : \quad \mu_{n_k} \xrightarrow{*} \lambda^*, \quad (7.31)$$

cf. [Landkof, 1972, Theorem 0.6]. It is clear that the weak limit measure λ^* is contained in $\overline{\mathcal{M}}$, too. Due to (7.28), $\lim_{k \rightarrow \infty} I(\mu_{n_k}) = I(\lambda^*)$. By (7.30) and (7.18), we can set $\lambda = \lambda^*$ and the condition of the Lemma is met. ■

By definition, the value $\mathcal{I}_{\text{LOG}}^c$ is unique. However, the uniqueness of the minimizing measure λ requires additional considerations in the form of Lemmas 2 and 3. In the former, it is shown that the energy minimizing potential is constant almost everywhere on its support. In the latter, the support is shown to be global on \mathbb{S}^d . Both lemmas share the similar argument that, if λ does not satisfy the respective property, a contradiction can be created by moving a part of the “mass” of λ .

Lemma 2 *The potential of the minimizing measure λ of Lemma 1 equals the minimum energy λ -almost everywhere, i.e.*

$$U_{\text{LOG}}^{\lambda}(\mathbf{x}) = \mathcal{I}_{\text{LOG}}^c \quad \text{for } \lambda\text{-almost all } \mathbf{x} \in \mathbb{S}^d. \quad (7.32)$$

PROOF First, we note that λ -a.e.

$$U_{\text{LOG}}^{\lambda}(\mathbf{x}) \geq \mathcal{I}_{\text{LOG}}^c. \quad (7.33)$$

To prove this by contradiction, assume there was a compact set $S \in \mathcal{B}(\mathbb{S}^d)$ with $\lambda(S) > 0$ such that $U_{\text{LOG}}^{\lambda}(\mathbf{x}) < \mathcal{I}_{\text{LOG}}^c$ for all $\mathbf{x} \in S$. Take $\nu \in \overline{\mathcal{M}}$ with support in S , for instance, $\nu = \delta_{\mathbf{x}^*}$ with $\mathbf{x}^* \in S$. Then

$$I_{\text{LOG}}(\lambda, \nu) < \mathcal{I}_{\text{LOG}}^c. \quad (7.34)$$

But, considering a convex combination $\mu = a\nu + (1-a)\lambda \in \overline{\mathcal{M}}$ with any $0 \leq a \leq 1$, the opposite can be shown, too: the energy of the convex combination,

$$I_{\text{LOG}}(\mu) = a^2 I_{\text{LOG}}(\nu) + 2a(1-a)I_{\text{LOG}}(\nu, \lambda) + (1-a)^2 I_{\text{LOG}}(\lambda), \quad (7.35)$$

is not less than that of λ , i.e. $I_{\text{LOG}}(\mu) \geq I_{\text{LOG}}(\lambda)$. Hence, the inequality

$$I_{\text{LOG}}(\lambda, \nu) - I_{\text{LOG}}(\lambda) + \frac{a}{2} (I_{\text{LOG}}(\nu) - I_{\text{LOG}}(\nu, \lambda) + I_{\text{LOG}}(\lambda)) \geq 0 \quad (7.36)$$

holds for all $0 < a \leq 1$. Therefore, the relation

$$I_{\text{LOG}}(\lambda, \nu) \geq \mathcal{I}_{\text{LOG}}^c \quad (7.37)$$

is true, contradicting (7.34).

Second, the inequality

$$U_{\text{LOG}}^\lambda(\mathbf{x}) \leq \mathcal{I}_{\text{LOG}}^c \quad (7.38)$$

holds everywhere in $\text{supp}(\lambda)$. If not, there would exist $\mathbf{x}^* \in \text{supp}(\lambda)$ with $U_{\text{LOG}}^\lambda(\mathbf{x}^*) > \mathcal{I}_{\text{LOG}}^c$. Corollary 2(b) implies that there is a neighborhood $S \in \mathcal{B}(\mathbb{S}^d)$ of \mathbf{x}^* with $\lambda(S) > 0$ on which $U_{\text{LOG}}^\lambda(\mathbf{x}) > \mathcal{I}_{\text{LOG}}^c$ holds for all $\mathbf{x} \in S$. Hence,

$$\begin{aligned} \mathcal{I}_{\text{LOG}}^c &= \int_S U_{\text{LOG}}^\lambda(\mathbf{x}) d\lambda(\mathbf{x}) + \int_{\mathbb{S}^d \setminus S} U_{\text{LOG}}^\lambda(\mathbf{x}) d\lambda(\mathbf{x}) \\ &\stackrel{(7.33)}{>} \mathcal{I}_{\text{LOG}}^c \lambda(S) + \mathcal{I}_{\text{LOG}}^c (1 - \lambda(S)) = \mathcal{I}_{\text{LOG}}^c, \end{aligned} \quad (7.39)$$

which is impossible. Combining (7.33) and (7.38), the Lemma is proven. \blacksquare

Lemma 3 *The potential $U_{\text{LOG}}^\lambda(\mathbf{x})$ corresponding to the minimizing measure λ of Lemma 1 is supported σ -almost everywhere, i.e. $\sigma(\text{supp}(U_{\text{LOG}}^\lambda)) = 1$. \square*

PROOF If $\sigma(\text{supp}(U_{\text{LOG}}^\lambda)) \neq 1$, then there exists a compact set $S \in \mathcal{B}(\mathbb{S}^d)$ with $\sigma(S) > 0$ on which $U_{\text{LOG}}^\lambda(\mathbf{x}) = 0$ everywhere. Define the convex combination $\mu = (1-a)\lambda + a\delta_{\mathbf{x}^*}$ with $\mathbf{x}^* \in S$ and $0 < a \leq 1$. Then

$$I_{\text{LOG}}(\mu) = a^2 \underbrace{k_{\text{LOG}}(\mathbf{x}^*, \mathbf{x}^*)}_{=2} + 2(1-a) \underbrace{U_{\text{LOG}}^\lambda(\mathbf{x}^*)}_{=0} + (1-a)^2 I_{\text{LOG}}(\lambda) < I_{\text{LOG}}(\lambda) \quad (7.40)$$

is equivalent to

$$a < \frac{2I_{\text{LOG}}(\lambda)}{2 + I_{\text{LOG}}(\lambda)}. \quad (7.41)$$

This is satisfied by, e.g., $a = \mathcal{I}_{\text{LOG}}^c / (2 + \mathcal{I}_{\text{LOG}}^c) < 1$. This does not contradict $a > 0$ due to $\mathcal{I}_{\text{LOG}}^c > 0$, cf. (7.29). Hence, (7.40) is true, which contradicts the definition of λ . Therefore, the initial assumption $\sigma(\text{supp}(U_{\text{LOG}}^\lambda)) \neq 1$ is wrong. \blacksquare

7.3 Proof of the asymptotic uniformity of LOG-optimal point sets (Theorem 1) 37

Lemmas 2 and 3 imply that the minimizing measure λ must be rotationally invariant. At least since Christensen [1970], it is known that the normalized measure with this property is unique, implying $\lambda = \sigma$.

In order to prove Theorem 1, it remains to be shown that the discrete measure associated with LOG-optimal points Ξ_N converges to the rotationally invariant measure σ , i.e. $\mu_{\Xi_N} \xrightarrow{\cdot} \sigma$. By (7.16) and (7.17),

$$\mathcal{I}_{\text{LOG}}(N) \leq I_{\text{LOG}}(\mu_{X_N}) \quad (7.42)$$

holds for an arbitrary point sequence $(X_N)_{N \in \mathbb{N}}$ and its associated discrete measure μ_{X_N} . Again employing the theorem about convergence of sequences in $\overline{\mathcal{M}}$ [Landkof, 1972, Theorem 0.6], both the minimum energy sequence $(\mu_{\Xi_N})_{N \in \mathbb{N}}$ and the arbitrary sequence $(\mu_{X_N})_{N \in \mathbb{N}}$ contain a weakly convergent subsequence,

$$\mu_{\Xi_{N_k}} \xrightarrow{\cdot} \gamma \quad (k \rightarrow \infty) \qquad \mu_{X_{N_k}} \xrightarrow{\cdot} \nu \quad (k \rightarrow \infty). \quad (7.43)$$

By Corollary 2(c), the inequality (7.42) is preserved in the limit $k \rightarrow \infty$,

$$I_{\text{LOG}}(\gamma) \leq I_{\text{LOG}}(\nu). \quad (7.44)$$

As the sequence $(X_N)_{N \in \mathbb{N}}$ is arbitrary, the density of discrete measures in $\overline{\mathcal{M}}$ implies that (7.44) holds for arbitrary $\nu \in \overline{\mathcal{M}}$. Hence, the measure γ is energy minimizing, i.e. $\gamma = \sigma$. The same argument can be applied to any subsequence of $(\Xi_N)_{N \in \mathbb{N}}$, so the whole sequence $(\mu_{\Xi_N})_{N \in \mathbb{N}}$ weakly converges to σ .

Alternatively, avoiding the density argument, one can study the evolution of the minimum N -point energy and conclude asymptotic uniformity from

Lemma 4 *The sequence of minimal N -point energies $(\mathcal{I}_{\text{LOG}}(N))_{N \in \mathbb{N}}$ is convergent. \square*

PROOF The optimal N -point energy $\mathcal{I}_{\text{LOG}}(N)$ is decomposed into the “self-energy” of the points and the mutual terms, i.e.

$$\mathcal{I}_{\text{LOG}}(N) = \frac{1}{N^2} \left(\sum_{i=1}^N \underbrace{k_{\text{LOG}}(\xi_i^{(N)}, \xi_i^{(N)})}_{=2} + 2 \sum_{1 \leq i < j \leq N} k_{\text{LOG}}(\xi_i^{(N)}, \xi_j^{(N)}) \right) \quad (7.45)$$

$$= \frac{2}{N} + \frac{2}{N^2} \sum_{1 \leq i < j \leq N} k_{\text{LOG}}(\xi_i^{(N)}, \xi_j^{(N)}). \quad (7.46)$$

Obviously, the self-energy $2/N$ vanishes in the limit $N \rightarrow \infty$. The mutual part asymptotically behaves the same as $\tilde{\mathcal{I}}_{\text{LOG}}(N)$. The latter is monotonically increasing, which can be shown using the identity

$$\sum_{1 \leq i < j \leq N} k_{ij} = \frac{1}{N-2} \sum_{l=1}^N \sum_{\substack{1 \leq i < j \leq N \\ i, j \neq l}} k_{ij}, \quad (7.47)$$

for arbitrary $k_{ij} \in \mathbb{R}$, and exploiting the fact that the optimal point sets are in general not hierarchical, i.e. $\Xi_{N-1} \not\subset \Xi_N$. The argument is similarly carried out in [Landkof, 1972, p. 160]. Once more, the boundedness of the kernel k_{LOG} is invoked, this time to show the existence of a limit of $\tilde{\mathcal{I}}_{\text{LOG}}(N)$, and hence of $\mathcal{I}_{\text{LOG}}(N)$. ■

Since the limit measures of arbitrary subsequences of $(\mu_{\Xi_N})_{N \in \mathbb{N}}$ all attain the energy $\mathcal{I}_{\text{LOG}}^c$, these limits are energy minimizing. Thus, they equal the uniform measure σ , wherefore again $\mu_{\Xi_N} \xrightarrow{\cdot} \sigma$.

7.4 Generation of energy-minimizing point sets

7.4.1 Algorithmic implementation

The computation of energy-minimizing point sets requires the solution of a nonlinear constraint optimization problem. The number of unknowns is $(d+1)N$ (i.e. N vectors with $d+1$ components each) and the number of nonlinear constraints is N . In order to work around the nonlinear constraints, an algorithm based on *unconstrained* optimization has been developed and implemented in Matlab. The key ingredients are:

- initialization of the point positions, see Section 7.4.2
- the gradient of the energy with respect to the point positions, $g \in \mathbb{R}^{(d+1)N}$, is computed in closed form
- the unconstrained gradient descent algorithm is interrupted every $n_{\text{it}}^{\text{min}}$ iterations in order to perform a projection of the points onto \mathbb{S}^d
- [OPTIONAL] refine via the application of a minimum least squares solver to the gradient g

Algorithm 7.1 is provided in Appendix 7.9.2. Despite its simplicity and the lack of incorporating the closed form expression of the Hessian of I_s , reasonable performance and robust results are achieved. However, large amounts of points N and high dimensions d lead to substantial runtimes.

The runtimes may be notably reduced by employing symmetrized kernel functions, \tilde{k}_s^{sym} . A minor drawback of this approach is that N is required to be an even number. As an advantage, the number of unknowns is reduced to $(d+1)N/2$. Only after the optimization of the point set $X_{N/2}$ is converged, it may be symmetrized, i.e. $(X_{N/2}, -X_{N/2})$. As the computational complexity of the algorithm grows superlinearly in N , symmetrized kernels speed up the computation by a factor of more than two.

In order to ensure an empirical robustness of the attained energy minimum, the converged set is randomly perturbed by a small amount, and then further optimized. Only rarely, this leads to a better local energy minimum, thus it is not included in the description of Algorithm 7.1.

The algorithm is implemented in Matlab and the software is made available as Open Source under terms of the GNU GPL License v3, see the repository provided in Fritzen & Kunc [2018b]. For already available databases of point sets of minimum energy type and others, the reader is referred to Womersley [2003], Hardin et al. [1997].

7.4.2 Initial positions

In principle, any point set can be chosen to initialize the optimization algorithm, as long as all points are separated (i.e. $z(\mathbf{x}_i, \mathbf{x}_j) > 0 \forall i \neq j$). Thus, *random points*, e.g. as in Marsaglia [1972], are a natural choice. Such points can be generated efficiently by means of the standard normal distribution. A random number generator with such distribution is used to define each component of the coordinate vector of each point separately. This is followed by subsequent normalization in order to ensure the initial points being located on \mathbb{S}^d . A detailed description is given in Appendix 7.9.1.

From a computational point of view the selection of the initial positions for the nonlinear problem can heavily influence the convergence behavior. The closer an initial point set is to its equilibrium state, the smaller the number of iterations and, hence, the computing time. Further, the employed gradient-based optimization method is theoretically prone to local minima. However, no significant outliers have been observed yet.

In the present work the easy to handle Matlab software provided by Leopardi's Recursive Zonal Equal Area Sphere Partitioning Toolbox Leopardi [2006] was used. It constructs point sets on \mathbb{S}^d such that each of the points is associated with an area of equal size and bounded diameter. Such sets are denoted *EQ point sets* in the sequel. The algorithm works in a recursive, analytical manner and thousands of points can be generated in fractions of a second on a laptop computer, even for large dimensions. Although EQ sets are not energy-minimizing for any of the investigated kernels, it is obvious that they are a.u.d. In our program, such kinds of sets are included as one possible option for the initial positions. However, small random perturbations are applied to initial EQ sets in order to mitigate possible artifacts that might come along with this specific initialization.

Furthermore, EQ sets $(\mathbf{x}_1, \dots, \mathbf{x}_N)$ are a practical choice for numerical integration on \mathbb{S}^d , i.e.

$$\int_{\mathbb{S}^d} f(\mathbf{x}) d\sigma(\mathbf{x}) \approx \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i). \quad (7.48)$$

Any a.u.d. point set has this integral approximation property (cf. Brauchart & Grabner [2015]), but EQ sets captivate with the numerical performance of their generating algorithm. Thus, for spherical integration in the manner of (7.48), EQ sets with $N = 10000$ are used in what follows.

7.5 Numerical examples I: point distributions

In this section, numerical results with respect to minimum energy points are presented. Various properties are compared for the cases $s = \text{LOG}, \log, 0.5, 1, 2, 3$, and for equal area (EQ) point sets, cf. Section 7.4.2. The sets are sized $N = 32, 41, 64, 83, 128, 193, 256, 383, 512, 1024, 1319, 1889, 2048$. The prime numbers are chosen randomly in order to avoid possible non-generic regularities that might be implied by the powers of two. For instance, EQ point sets of size $N = 2^k$, $k \in \mathbb{N}$, are symmetric. All of the studied point sets are included as examples along with the source code Fritzen & Kunc [2018b].

A rigorous study of the influence of the parameters of Algorithm 7.1 (stopping criteria, initial positions, etc.) is not intended at this point. Also, since the number of local minimum values of the energy $I_{\text{LOG}}(\mu_{X_N})$ might grow rapidly as N increases, a comprehensive numerical investigation would have to address this issue, too. For instance, in Ballinger et al. [2009] many starting positions were used in order to gain confidence in the found energy minimum. Here, the aim is to exemplarily compare the properties of LOG-optimal points with the classical cases. Few, random examples of different point sets with equal initial conditions have been compared and found to yield only negligible deviations.

7.5.1 Mesh ratio

A common quality indicator for point sets and meshes is the *mesh ratio* ρ , see e.g. Shewchuk [1996], Si [2015], Hesse et al. [2010], Narcowich et al. [2006a]. It is defined by

$$\rho(X_N) = \frac{h(X_N)}{q(X_N)}, \quad \text{where} \quad h(X_N) = \max_{\mathbf{y} \in \mathbb{S}^d} \min_{\mathbf{x} \in X_N} z(\mathbf{x}, \mathbf{y})$$

$$\text{and} \quad q(X_N) = \frac{1}{2} \min_{\substack{\mathbf{x}, \mathbf{y} \in X_N \\ \mathbf{x} \neq \mathbf{y}}} z(\mathbf{x}, \mathbf{y}). \quad (7.49)$$

The mesh ratio ρ is plotted for various point sets on \mathbb{S}^2 in Figure 7.3. The plot suggests that the mesh ratio of minimum energy point sets is almost always bounded from above by the EQ set's mesh ratio. In other words, most of the times it is worthwhile to conduct the energy optimization when starting with EQ points. No clear tendency between the different energy minimum point sets is observed.

7.5.2 Empirical Distribution Functions (EDF)

Intuitively, one could describe the “regularity” of an energy optimal point set Ξ_N as a measure of how much the neighbor relations vary among the points. To put this more

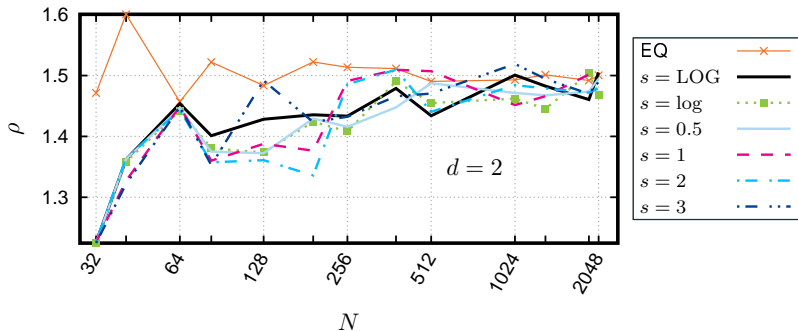


Figure 7.3: Mesh ratio ρ as a function of the size N of point sets on \mathbb{S}^2 . Data is linearly interpolated for better visibility. The equal area set (EQ, orange solid line with points) is used (after slight perturbation) as initial configuration for the generation of the minimum energy point sets.

formal, we employ the definition of the Empirical Distribution Function of the geodesic point distance towards a fixed point $\xi_i^{(N)} \in \Xi_N$,

$$P(z_g(\xi_i^{(N)}, \xi_j^{(N)}) < r) \quad \text{for } r \in [0, \pi] \text{ and } j = 1, \dots, N. \quad (7.50)$$

There, P denotes the usual probability measure on discrete sets. The set Ξ_N may be regarded as more regular the better the individual EDF's (7.50) approximate the distribution function $F^d(r)$ of the rotationally invariant measure σ , cf. (7.23).

The EDF's of LOG-optimal and of EQ point sets on the 2-sphere are compared in Figure 7.4. The very significant step-like curve of the EQ set for $N = 128$ originates from the two poles, of which the neighbors are distributed along circles of latitudes. This indicates comparatively strong non-uniformity. Nonetheless, convergence towards the limit function $F^2(r)$ can be observed for either kind of set.

EDF's of energy minimizing point sets on the 2-sphere for the cases $s = \text{LOG}$ and $s = 2$ are depicted in Figure 7.5. It is noteworthy that the curves behave less like step functions as N increases. By the metric of vision, i.e. qualitatively, no significant difference in the uniformity of the different kinds of energy optimal point sets can be observed.

7.5.3 Asymptotic behavior of the minimum energy

The asymptotic behavior of the modified energy $\tilde{\mathcal{I}}_s(N)$ for the classical cases is subject of a large body of literature. Especially in the hypersingular case, $d \leq s$, where the limit of $\tilde{\mathcal{I}}_s(N)$ for $N \rightarrow \infty$ is not finite, the growth rate of $\tilde{\mathcal{I}}_s(N)$ in orders of N is of much interest, cf. Brauchart & Grabner [2015]. For instance, in Kuijlaars & Saff [1998] the

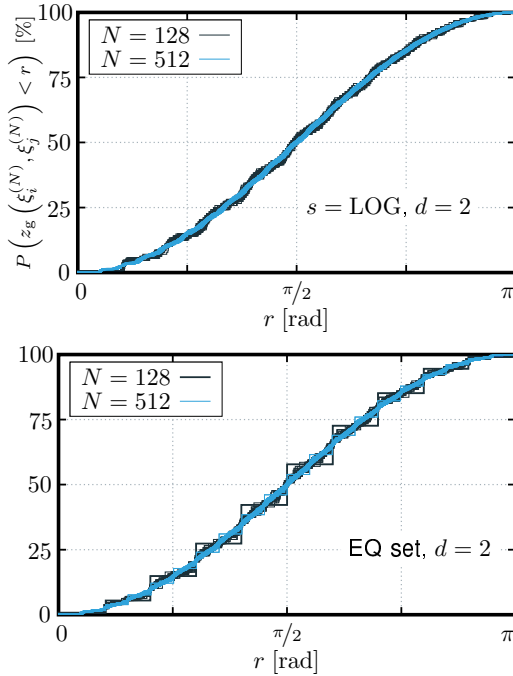


Figure 7.4: All EDF's of minimum LOG-energy point sets (*above*) and of EQ point sets (*below*) on \mathbb{S}^2 .

case $d = s = 2$ is theoretically analyzed. Therein, the first order term of the energy growth is proven to be $\log N$, see Figure 7.6.

The limit value of the modified energy $\tilde{\mathcal{I}}_s(N)$ is finite and can be computed in the potential-theoretic case. A noteworthy property of the bounded kernel k_{LOG} is that even the energy $\mathcal{I}_{\text{LOG}}(N) = I_{\text{LOG}}(\mu_{\Xi_N})$ is finite, which has certain advantages over the modified energy $\tilde{\mathcal{I}}_{\text{LOG}}(N)$.

First, additional insight into the nature of energy minimizing point sets can be gained. Apart from the slightly different normalization factors, the difference between $\tilde{\mathcal{I}}_{\text{LOG}}(N)$ and $\mathcal{I}_{\text{LOG}}(N)$ is mainly the self-energy $N^{-2} \sum_{i=1}^N k_{\text{LOG}}(\xi_i, \xi_i) = 2/N$, which was introduced in Lemma 4. It is due to the finiteness of the self-energy contribution that an *evaluation of the energy functional* $I_{\text{LOG}} : \mathcal{M} \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ *at discrete measures is finite*. Consequently, the concept of capacity (used in Landkof [1972], Brauchart & Grabner [2015]) is unnecessary in the proof of Theorem 1.

Secondly and consequently, $\mathcal{I}_{\text{LOG}}(N)$ is an approximation of the limit value $\mathcal{I}_{\text{LOG}}^c$ from *above*. In Figure 7.7 (*left*), this is visualized. Moreover, $\tilde{\mathcal{I}}_{\text{LOG}}(N)$ is an approximation from *below*. The difference between the two lines is essentially related to the self-energy.

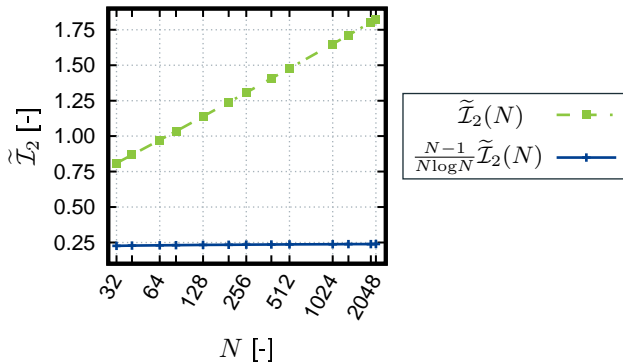
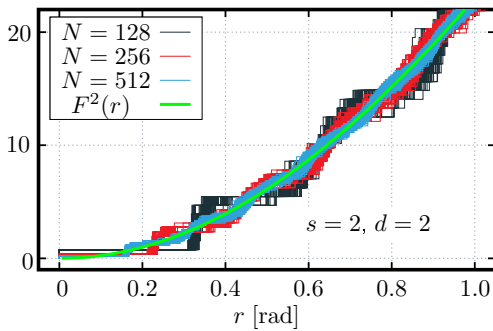
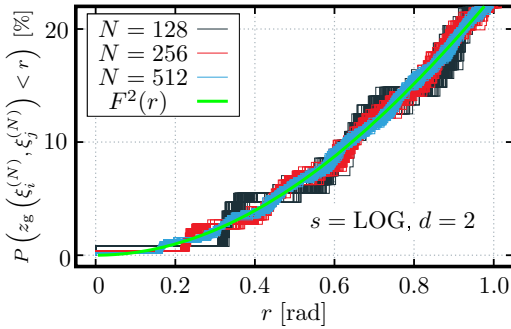


Figure 7.5: All EDF's corresponding to LOG-optimal point sets (above) and to Riesz 2-optimal point sets (below) on \mathbb{S}^2 . The limit distribution function $F^2(r)$ is included (green).

Figure 7.6: Modified energy of minimizing point sets on \mathbb{S}^2 for the Riesz case $s = 2$ (green, dashed line) and a scaled variant (blue, solid line). The latter appears to be monotonically converging from below to 0.25, supporting the theoretical result Kuijlaars & Saff [1998, Theorem 3].

Thirdly, Figure 7.7 (left) suggests that the quantity $\mathcal{I}_{\text{LOG}}(N)$ is a *more accurate* approximation of $\mathcal{I}_{\text{LOG}}^c$ than $\tilde{\mathcal{I}}_{\text{LOG}}(N)$. Additionally, Figure 7.7 (right) is numerical evidence for a *faster convergence* of $\mathcal{I}_{\text{LOG}}(N)$.

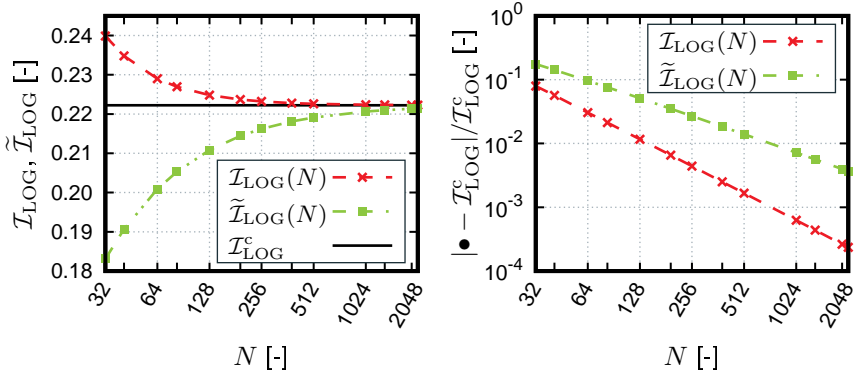


Figure 7.7: Left: evolution of the energy $\mathcal{I}_{\text{LOG}}(N)$ and of the modified energy $\tilde{\mathcal{I}}_{\text{LOG}}(N)$ for point sets on \mathbb{S}^2 . Right: relative deviations of both quantities with respect to $\mathcal{I}_{\text{LOG}}^c$.

7.5.4 Universally optimal point sets

A point set is called *universally optimal*, cf. Cohn & Kumar [2007], if it is energy optimal with respect to *any* kernel $\tilde{k}(z)$ for which $(-1)^m \tilde{k}^{(m)}(z) > 0$ holds for all z in the open interval $(0, 2)$ and for all $m \in \mathbb{N}_{\geq 0}$. The kernel \tilde{k}_{LOG} obviously has this property, and is thus called *completely monotonic*, cf. Cohn & Kumar [2007].

It is known, cf. Cohn & Kumar [2007, Table 1], that there are exactly six universally optimal point configurations on \mathbb{S}^2 : A single point, two antipodal points, an equilateral triangle, a regular tetrahedron, an octahedron, and an icosahedron. The provided energy minimizing program was tested for its ability to find these point configurations. All of the pre-implemented kernel functions ($s = \text{LOG}, \log, 0.5, 1, 2, \dots, 10$) were tried multiple times. As initial configuration, both perturbed EQ point sets and random point sets were tested.

In all runs, and by the metric of vision, the program returned the respective universally optimal point set. For the cases $N = 3$ and $N = 6$, the symmetrized kernel functions yielded half of the point sets of the octahedron and of the icosahedron, respectively. A rigorous analysis with a high number of varying parameters (initial positions, convergence criteria, optimization routines, etc.) and with an automated procedure to identify the points up to rotation has not been conducted.

It is interesting to observe that the EDF's of all points within a universally optimal configuration on \mathbb{S}^2 coincide. Random samples in higher dimensions yielded the same

result. For an extensive numerical investigation on universal optimality, the reader is referred to Ballinger et al. [2009].

7.6 Concentric Interpolation

7.6.1 Introduction

The Concentric Interpolation is a generalized version of the RNEXP of the authors Fritzen & Kunc [2018c] which was suggested for a specific physical application. The generalization lies in the arbitrariness of the dimension d and of the radial interpolants as explained below. Further, the source code is now available via GitHub Fritzen & Kunc [2018a].

In many scientific applications an approximation of a continuous function, $f : \mathbb{R}^{d+1} \rightarrow \mathbb{R}$, by a surrogate model, \tilde{f} , is sought-after. If the function f is sufficiently smooth so that the gradient $\nabla_{\mathbf{x}} f$ or the Hessian $\nabla_{\mathbf{x}}^2 f$ exists, i.e.

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}), \quad \nabla_{\mathbf{x}}^2 f(\mathbf{x}) = \frac{\partial^2 f}{\partial \mathbf{x}^2}(\mathbf{x}), \quad (7.51)$$

then it is desirable to also approximate these derivatives by means of the same surrogate \tilde{f} . Thus, under these circumstances, it is worthwhile if the surrogate relationship reads

$$\mathbf{x} \mapsto \tilde{f}(\mathbf{x}) \approx f(\mathbf{x}), \quad \nabla_{\mathbf{x}} \tilde{f}(\mathbf{x}) \approx \nabla_{\mathbf{x}} f(\mathbf{x}), \quad \nabla_{\mathbf{x}}^2 \tilde{f}(\mathbf{x}) \approx \nabla_{\mathbf{x}}^2 f(\mathbf{x}). \quad (7.52)$$

The Concentric Interpolation method provides means to realize such an approximation of f and, if existent, of its first two derivatives.

7.6.2 Setup and Concentric Sampling

The starting point for the Concentric Interpolation (CI) scheme is the split of the coordinate \mathbf{x} into amplitude (or radius) $x = \|\mathbf{x}\|$ and direction \mathbf{n} :

$$\mathbf{x} = x \mathbf{n}. \quad (7.53)$$

The support points of the CI scheme are required to be arranged concentrically around the origin of the domain \mathbb{R}^{d+1} of f in the following sense: assume that $P \in \mathbb{N}$ support coordinates are provided in the form

$$\{\mathbf{x}_p\}_{p=1}^P = \{l_r \mathbf{n}_i\}_{r=1, i=1}^{r=R, i=N}, \quad (7.54)$$

utilizing the split notation (7.53). For simplicity, the R radii l_r are assumed to be identical along each of the N directions $\mathbf{n}_i \in X_N \subset \mathbb{S}^d$. Further, the directions X_N are assumed to be an s -energy minimizing set of points, e.g. $X_N = \Xi_N$ for the case $s = \text{LOG}$. Therefore, $P = RN$ and this way of placing sampling sites is descriptively coined *Concentric Sampling*, see Figure 7.8 (left).

Note that this sampling is considerably different from sampling using a tensor product of predefined point positions along different Cartesian coordinates (see Figure 7.8, right). Concentric Sampling yields a “uniform” point density in all directions and the radial refinement/coarsening is controlled by the choice of the sampling amplitudes $\{l_r\}_{r=1}^R$.

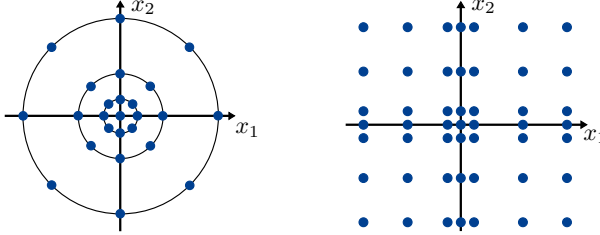


Figure 7.8: Illustration of sampling strategies in \mathbb{R}^2 . *Left:* Concentric Sampling including visualized sampling radii. *Right:* nonuniform Cartesian grid. In both graphs, the density of the nine points closest to the origin is similar.

The specific structure of the coordinate split allows for both efficient sampling and interpolation/differentiation of concentric data. The idea can be outlined as follows: first, the concentrically sampled data is used to build one-dimensional interpolants along the individual training directions, e.g. by using piecewise defined polynomials with continuous first derivative, in order to allow for *Radial Interpolation* (RI). This is followed by a *Tangential Interpolation* (TI) via Radial Basis Functions (RBF), more precisely employing Gaussian kernel functions operating on the geodesic distance z_g . The latter are well-known from the context of Spherical Basis Functions (SBF, see e.g. Sommariva & Womersley [2005]). In other words, TI interpolates data on the sphere \mathbb{S}^d , while RI interpolates data between scaled spheres $l_r \mathbb{S}^d = \{\mathbf{x} \in \mathbb{R}^{d+1} : \|\mathbf{x}\| = l_r\}$, $r = 1, \dots, R$. The general procedure is represented by Figure 7.9.

7.6.3 Interpolation

Radial Interpolation (RI)

Along all training directions $\mathbf{n}_1, \dots, \mathbf{n}_N$ scalar interpolation functions $\mathcal{R}_i(x)$ are defined. The interpolation conditions

$$\tilde{f}(l_r \mathbf{n}_i) = \mathcal{R}_i(l_r) \quad \forall i \in \{1, \dots, N\}, \forall r \in \{1, \dots, R\} \quad (7.55)$$

hold. The radial interpolants \mathcal{R}_i must be smooth enough in order to allow for the desired derivatives of \tilde{f} . Apart from this, no further conditions are imposed.

For example, in the recently proposed RNEXP \mathcal{R}_i was taken to be a piecewise cubic polynomial with continuous first derivative. In the present work, piecewise quadratic polynomials are utilized. In the following the univariate vector-valued function $\underline{\mathcal{R}}(x) \in \mathbb{R}^N$ represents a general interpolation along all N directions at radius x .

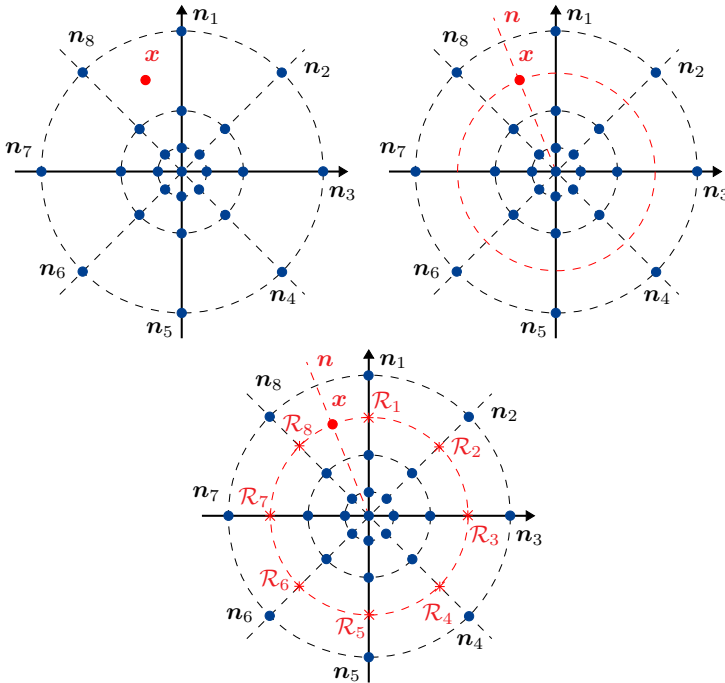


Figure 7.9: Schematic representation of the Concentric Interpolation. *Top left:* query point \mathbf{x} , sampling points \mathbf{x}_p (blue dots) and sampling directions \mathbf{n}_i . *Top right:* query radius x (red circle) for Radial Interpolation and query direction \mathbf{n} for Tangential Interpolation. *Right:* radial interpolants \mathcal{R}_i evaluated along the directions \mathbf{n}_i .

Tangential Interpolation (TI) using radial basis functions

Interpolation on spheres has already been widely investigated, e.g. Fasshauer & Schumaker [1998] addressing scattered data problems, Sloan & Womersley [2000] and Womersley & Sloan [2001] considering spherical harmonics for this purpose, and Wang et al. [2017] using needlet approximation. For the tangential part of the interpolation, we choose the well-established concept of Spherical Basis Functions (see below).

The classic kernel approximation of a general function $g : \mathbb{R}^{d+1} \rightarrow \mathbb{R}$ can be written as (cf., e.g., Fasshauer & McCourt [2015])

$$g(\mathbf{x}) \approx \tilde{g}(\mathbf{x}) = \sum_{i=1}^N w_i k_i(\mathbf{x}) = \underline{w}^\top \underline{k}(\mathbf{x}), \quad (7.56)$$

where the weight vector $\underline{w} = [w_1, \dots, w_N]^\top$ is constant and the components of $\underline{k}(\mathbf{x}) = [k_1(\mathbf{x}), \dots, k_N(\mathbf{x})]^\top$ are symmetric positive definite kernel functions $k_i : \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \rightarrow \mathbb{R}$ centered at (general) sampling sites \mathbf{x}_i , i.e. $k_i(\mathbf{x}) = k(\mathbf{x}_i, \mathbf{x})$.

The TI method is obtained by placing the kernel centers on the sphere \mathbb{S}^d at the training directions \mathbf{n}_i , and by employing kernels acting on the geodesic distance between the center \mathbf{n}_i and the query direction \mathbf{n} , i.e. $k(\mathbf{n}_i, \mathbf{n}) = \tilde{k}(z_g(\mathbf{n}_i, \mathbf{n})) = k_i(\mathbf{n})$. Gaussian kernel functions with kernel width parameter $\gamma > 0$ have proven a viable choice,

$$k_i(\mathbf{n}) = \exp(-\gamma(z_g(\mathbf{n}_i, \mathbf{n}))^2). \quad (7.57)$$

However, the TI scheme is general with respect to the employed kernel function, as long as it is positive definite and operates on the geodesic distance.

The same kernel parameter γ is used for the kernel functions k_i at all supporting points \mathbf{n}_i . This is a simplifying assumption which may be dropped if desired, i.e. the TI method does not rely on the kernel parameters being equal. An algorithm for the optimization of the parameter γ is described in Appendix 7.9.2, Algorithm 7.2. It makes use of a bisection algorithm after a coarse-grained initial trial-and-error search.

In order to compute the weights, the Radial Interpolation must be accounted for, i.e. the weight vector depends on the amplitude:

$$\tilde{f}(\mathbf{x}) = \underline{w}(x)^\top \underline{k}(\mathbf{n}). \quad (7.58)$$

The weight vector $\underline{w}(x)$ is specified such that the function values at the provided inputs from Concentric Sampling are reproduced,

$$\tilde{f}(l_r \mathbf{n}_i) = f(l_r \mathbf{n}_i) = \mathcal{R}_i(l_r). \quad (7.59)$$

In order to obtain the weight $\underline{w}(x)$ the kernel vector \underline{k} is evaluated at the directions X_N in order to build the symmetric and positive definite (albeit poorly conditioned) kernel matrix

$$\left(\underline{K}\right)_{ij} = \left(\underline{K}\right)_{ji} = k_i(\mathbf{n}_j) \quad (i, j \in \{1, \dots, N\}). \quad (7.60)$$

Equations (7.58), (7.59) and (7.60) yield

$$\underline{\mathcal{R}}(l_r) = \underline{K} \underline{w}(l_r) \quad \Rightarrow \quad \underline{w}(x) = \underline{K}^{-1} \underline{\mathcal{R}}(x) \quad (7.61)$$

and the overall surrogate reads

$$\tilde{f}(\mathbf{x}) = \underline{\mathcal{R}}(x)^\top \underline{K}^{-1} \underline{k}(\mathbf{n}). \quad (\text{CI})$$

Thus, the radial dependence of the weight vector $\underline{w}(x)$ yields an extension of the interpolation on the sphere (TI) to the surrounding space via radial interpolation (RI).

7.6.4 Differentiation

The differentiation of the surrogate \tilde{f} is effected step by step exploiting the chain rule. First, the gradients of the amplitude x , of the direction \mathbf{n} and of the geodesic distance $z_i = z_g(\mathbf{n}_i, \mathbf{n})$ are computed and transferred into matrix-vector notation (underlines):

$$\nabla_{\mathbf{x}} x = \frac{\mathbf{x}}{\|\mathbf{x}\|} = \mathbf{n} \quad \leftrightarrow \underline{\mathbf{n}} \in \mathbb{R}^{d+1}, \quad (7.62)$$

$$\nabla_{\mathbf{x}}^2 x = \nabla \mathbf{n} = \frac{1}{x} (\mathbf{I} - \mathbf{n} \otimes \mathbf{n}) = \mathbf{P} \quad \leftrightarrow \underline{\underline{\mathbf{P}}} \in \mathbb{R}^{(d+1) \times (d+1)}, \quad (7.63)$$

$$\nabla_{\mathbf{x}} z_i = -\frac{1}{\sin(z_i)} \mathbf{n}_i \quad \equiv -\frac{1}{\sin(z_i)} \underline{\mathbf{n}}_i \in \mathbb{R}^{d+1}. \quad (7.64)$$

The second gradient of \mathbf{n} can be replaced by combination of the previous results. Next, the gradient of \tilde{f} is expressed in vector matrix notation (for convenience the arguments x and \mathbf{n} are omitted in the following):

$$\nabla_{\mathbf{x}} \tilde{f} = \frac{\partial \tilde{f}}{\partial x} \underline{\mathbf{n}} + \underline{\underline{\mathbf{P}}} \frac{\partial \tilde{f}}{\partial \mathbf{n}} = \frac{\partial \mathcal{R}}{\partial x} \underline{\underline{\mathbf{K}}}^{-1} \underline{\mathbf{k}} + \underline{\underline{\mathbf{P}}} \frac{\partial \underline{\underline{\mathbf{K}}}}{\partial \underline{\mathbf{n}}} \underline{\underline{\mathbf{K}}}^{-1} \underline{\mathcal{R}}, \quad \frac{\partial k_i}{\partial \underline{\mathbf{n}}} = -\frac{\tilde{k}'(z_i)}{\sin(z_i)} \underline{\underline{\mathbf{P}}} \underline{\mathbf{n}}_i. \quad (7.65)$$

For algorithmic convenience the abbreviation $\underline{\mathcal{R}}' = \partial_x \underline{\mathcal{R}}$, the matrix of input directions $\underline{\underline{\mathbf{N}}}$ and the diagonal matrix $\underline{\underline{\kappa}}$ defined by

$$\underline{\underline{\mathbf{N}}} = \begin{bmatrix} \underline{\mathbf{n}}_1^\top \\ \underline{\mathbf{n}}_2^\top \\ \vdots \\ \underline{\mathbf{n}}_N^\top \end{bmatrix} \in \mathbb{R}^{N \times d+1}, \quad \underline{\underline{\kappa}} = -\text{diag} \left(\frac{\tilde{k}'(z_1)}{\sin(z_1)}, \dots, \frac{\tilde{k}'(z_N)}{\sin(z_N)} \right) \quad (7.66)$$

are introduced. Thereby the gradient can be rewritten as

$$\nabla_{\mathbf{x}} \tilde{f} = \underline{\mathbf{n}} \underline{\mathcal{R}}'^\top \underline{\underline{\mathbf{K}}}^{-1} \underline{\mathbf{k}} + \underline{\underline{\mathbf{P}}} \underline{\underline{\mathbf{N}}}^\top \underline{\underline{\kappa}} \underline{\underline{\mathbf{K}}}^{-1} \underline{\mathcal{R}}. \quad (7.67)$$

The computation is efficiently carried out as follows:

- compute $\underline{\mathbf{v}} = \underline{\underline{\mathbf{K}}}^{-1} \underline{\mathcal{R}}$ and $\underline{\mathbf{v}}' = \underline{\underline{\mathbf{K}}}^{-1} \underline{\mathcal{R}}'$
- compute $(\underline{\mathbf{y}})_i = -\frac{v_i k'_i}{\sin(z_i)}$
- set

$$\begin{aligned} \nabla_{\mathbf{x}} \tilde{f} &= \underline{\mathbf{n}} \left(\underline{\mathbf{v}}'^\top \underline{\mathbf{k}} \right) + \underline{\underline{\mathbf{P}}} \left(\underline{\underline{\mathbf{N}}}^\top \underline{\mathbf{y}} \right) \\ &= \underline{\mathbf{n}} \left(\underline{\mathcal{R}}'^\top(x) \underline{\tilde{\mathbf{w}}}(\underline{\mathbf{n}}) \right) + \frac{1}{x} \left(\underline{\underline{\mathbf{I}}} - \underline{\mathbf{n}} \underline{\mathbf{n}}^\top \right) \underline{\underline{\mathbf{N}}}^\top \underline{\underline{\kappa}}(\underline{\mathbf{n}}) \underline{\mathbf{w}}(x) \end{aligned} \quad (7.68)$$

By construction the gradient is constituted by a part in radial direction represented by the first term in (7.68) and a part within the tangential plane represented by the second term.

The second gradient is computed in a straight-forward way following the same concept [see Fritzen & Kunc, 2018c, for details of the closely related RNEXP].

7.6.5 Symmetric Gaussian kernel function

In the spirit of the symmetric kernels used for the generation of energy minimizing point sets, the use of symmetric kernel functions for TI is computationally appealing for symmetric functions $f(\boldsymbol{x}) = f(-\boldsymbol{x})$. Therefore, the kernel function may be replaced with

$$\tilde{k}(z) \rightarrow \tilde{k}^{\text{sym}}(z) = \tilde{k}(z) + \tilde{k}(\pi - z). \quad (7.69)$$

Then only half the number of points must be considered in comparison to the scheme using the standard kernel function. Additionally, the symmetry of f is reproduced exactly in the surrogate \tilde{f} . Note that the number of points enters quadratically into the computational effort due to the multiplication with the kernel matrix, i.e. the computational savings due to the use of a symmetric kernel function are above a factor of two.

7.7 Numerical examples II: Concentric Interpolation

In this section, we present numerical examples of the CI method introduced in Section 7.6. All of the direction sets used in this section are available in the Matlab software package's subfolder `examples/exports`, see Fritzen & Kunc [2018b], and in the C++ software packages subfolder `data/directions`, see Fritzen & Kunc [2018a].

7.7.1 Kernel parameter γ

Algorithm 7.2 is employed for the optimization of the global kernel parameter γ . First, it is applied to energy minimizing point sets for the cases $s = \text{LOG}, \log, 0.5, 1, 2, 3$ and to EQ point sets of size $N = 512$ on the spheres $\mathbb{S}^3, \mathbb{S}^4, \mathbb{S}^5$. As a target function, the constant one-function, $f(\boldsymbol{x}) \equiv 1$, is chosen. The evolution of the absolute error $\|\tilde{f} - 1\|_{L^2(\mathbb{S}^d)}$ as a function of γ , in the vicinity of the optimal parameter γ_{best} , is depicted in Figure 7.10.

Most notably, all energy optimal point sets exhibit similar overall behavior. Generally, the approximation error is in the order of 10^{-10} to 10^{-7} , with the minimum error seemingly increasing with the dimension d . The apparent differences among the various cases of s are rather subtle. If any, then the kernels $\tilde{k}_s(z)$ that are closer to the beginning of

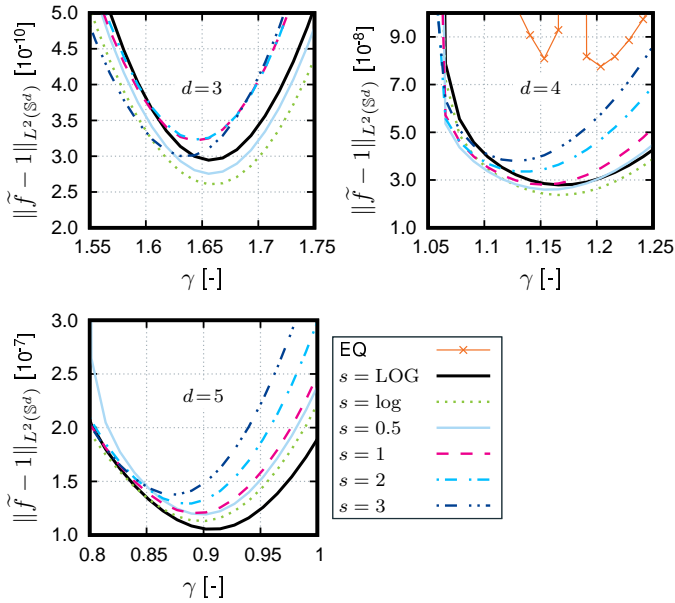


Figure 7.10: Partition of unity property of the interpolant, as a function of the kernel width parameter γ in different dimensions. The $N = 512$ support points are varied among different kinds of minimum energy point sets and the EQ point set. For the latter, the minimum value is $1.5 \cdot 10^{-9}$ on S^3 (top left, out of range) and $2.1 \cdot 10^{-6}$ on S^5 (bottom, out of range).

the sequence of derivatives (7.1), front-extended by $\tilde{k}_{\text{LOG}}(z)$, are somewhat superior to those that appear later in the sequence. This might be due to the fact that the former kernels have more “global” behavior than the latter, in the sense that the singularity of the kernel as $z \rightarrow 0$ is of increasing order as the kernel sequence index n increases. In fact, the singularity is not present for $s = \text{LOG}$ and only appears for $s = \text{log}$ and for the Riesz kernels. However, not enough numerical evidence is present to thoroughly judge this conjecture, as multiple point sets of the same kind would have to be taken into consideration. This is necessary if representativeness has to be guaranteed, i.e. when the effect of possibly existent local minima in the energy optimization process should be mitigated. However, it is very clear that minimum energy points are superior to their initial EQ point positions in this context, as the errors of the EQ sets are mostly out of range in Figure 7.10.

Further point sets are investigated in the dimensions $d = 2$ and $d = 7$. To this end, LOG-optimal point sets of sizes $N = 83, 256, 1319, 2048$ are created. This time, the symmetrized LOG-kernel is employed in the case $d = 7$. Consequently, the symmetric

interpolant \tilde{f}^{sym} corresponding to the symmetrized Gaussian kernel (7.69) is applied in this case. The results are shown in Figure 7.11

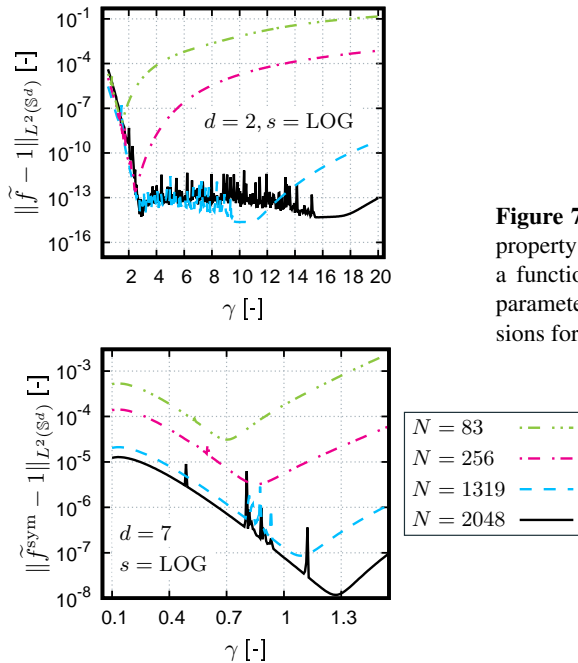


Figure 7.11: Partition of unity property of the interpolant, as a function of the kernel width parameter γ in different dimensions for the LOG-case.

It is noteworthy that for the low-dimensional case, $d = 2$, larger sets of supporting points do not necessarily increase the accuracy, as the level of machine precision is reached even for moderate values of N , in this case for $N = 1319$. The symmetrization procedure, combined in both the generation of the point set and in the spherical interpolation method, enables accurate approximations of constant functions.

7.7.2 Application to an eight-dimensional engineering model

This section is concluded by a demonstration of the applicability of the Concentric Interpolation method to a realistic example. In Morris et al. [1993], the eight-dimensional scalar function $f : \mathbf{x} = [x_1, \dots, x_8] \rightarrow \mathbb{R}$,

$$\hat{f}(\mathbf{x}) = 2\pi x_3 (x_5 - x_6) \left(\log \left(\frac{x_2}{x_1} \right) \left(1 + \frac{2x_7 x_3}{\log \left(\frac{x_2}{x_1} \right) x_1^2 x_8 + \frac{x_3}{x_4}} \right) \right)^{-1}, \quad (7.70)$$

is stated as a model of the water flow rate through a borehole. The ranges of the eight variables are given in Table 7.1.

Table 7.1: Domain of the function \hat{f} from (7.70).

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
min	0.05	100	63070	63.1	990	700	1120	9855
max	0.15	50000	115600	116	1110	820	1680	12045

The CI scheme requires the support points to be centered around the origin, and the present implementation also necessitates the function value to be zero at the origin. Therefore, the argument is shifted by the center coordinate of the domain, $\mathbf{x}_{\text{center}}$, and the corresponding function value, $\hat{f}(\mathbf{x}_{\text{center}})$, is subtracted. Furthermore, the coordinates are scaled in order to ensure that the split in amplitude and direction (7.53) respects each of the eight dimensions equally. Thus, the equivalent function

$$f(x\mathbf{n}) = \hat{f} \left(x \begin{bmatrix} \Delta_{x_1} & & 0 \\ & \ddots & \\ 0 & & \Delta_{x_8} \end{bmatrix} \mathbf{n} + \mathbf{x}_{\text{center}} \right) - \hat{f}(\mathbf{x}_{\text{center}}) \quad (7.71)$$

is approximated, where $\Delta_{x_i} = x_{i,\text{max}} - x_{i,\text{min}}$ is the range of the i -th coordinate.

For the supporting points on \mathbb{S}^7 , minimum LOG-energy points are created in an efficient manner by symmetrizing the point sets of Section 7.7.1 for the case $d = 7$. Therefore, the supporting point sets are of sizes $N = 166, 512, 2638, 4096$, although the minimum energy program was run with only $N/2$ points. This emphasizes the benefit of symmetrized kernel functions.

The optimization of the kernel parameter γ , however, was not reused from Section 7.7.1, as it was performed there with respect to the constant one-function and with symmetrized spherical basis functions. The function f of (7.71) is significantly anisotropic, i.e. dependent on the direction \mathbf{n} , and not symmetric. Hence, the optimal kernel parameter is determined for the usual, i.e. unsymmetrized kernel function (7.57) such that the absolute error $\|\hat{f} - f\|_{(\mathbb{S}^7)}$ is minimized for each point set. The resulting values are $\gamma = 0.6355, 0.3923, 0.308, 0.290$ respectively for the point sets of sizes $N = 166, 512, 2638, 4096$.

As for the setup of the CI scheme, function values at the respective point positions are provided at the radii $x = 0, 0.25, 0.5, 0.75, 1$. This means the total number of samples provided is $4N + 1$. Considering the range of N , this can be regarded as a quite sparse sampling in \mathbb{R}^8 . These samples are used to set up two piecewise quadratic polynomials as radial interpolants along each direction. Differentiability is not guaranteed in this case, but could easily be achieved if required.

Results are depicted in Figure 7.12 with respect to the error $\|\tilde{f} - f\|_{L^2(r\mathbb{S}^7)} / \|f\|_{L^2(r\mathbb{S}^7)}$, i.e. on the spheres of radius $r > 0$. The initial decline of the error as r increases is expected, as it is a relative error measure and the function f is zero for $r = 0$. It is observable that the error attains local minimum values in the vicinity of the intermediate supporting radii $r = 0.25, 0.5, 0.75$, at least for larger values of N . Most notably, the error appears to be mostly monotonically decreasing as N increases.

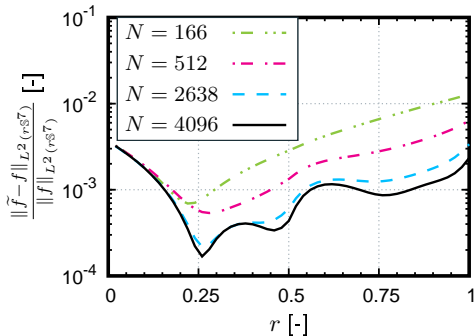


Figure 7.12: Relative error of the CI approximation \tilde{f} of the eight-dimensional function f from (7.71) on the spheres $r\mathbb{S}^7$ for 50 equidistant radii r . The points in the graph are linearly interpolated. The supporting points are placed at the radii $r = 0, 0.25, 0.5, 0.75$, and 1 along the symmetrizations $(X_{N/2}, -X_{N/2})$ of the corresponding point sets $X_{N/2}$ of Section 7.7.1, cf. Figure 7.11 (right).

7.8 Résumé

7.8.1 Summary

It is shown that LOG-optimal point sets are asymptotically uniformly distributed. This implies that all kernels of the sequence

$$\left((-1)^n a_n \frac{\partial^n \tilde{k}_{\text{LOG}}}{(\partial z)^n}(z) \right)_{n \in \mathbb{N}_{\geq 0}} = \tilde{k}_{\text{LOG}}(z), \tilde{k}_{\log}(z), \tilde{k}_1(z), \tilde{k}_2(z), \tilde{k}_3(z), \dots, \quad (7.72)$$

where the factor sequence $(a_n)_{n \in \mathbb{N}_{\geq 0}}$ is defined as

$$a_n = \begin{cases} 1 & \text{if } n = 0, 1 \\ \frac{1}{(n-2)!} & \text{else,} \end{cases} \quad (7.73)$$

lead to a.u.d. point sets. In this sense, the regime of potential-theoretic cases is “naturally” extended. These kernels are further compared in Table 7.2, where non-integral values of $s \in \mathbb{R}_{\geq 0}$ are included. The property $\tilde{k}'_s(2) = 0$ is investigated for numerical reasons, as it might be computationally favorable if the gradient of the energy of two antipodal points vanishes, and not just its projection onto the tangential plane of the sphere. This may possibly be an issue for high accuracy computations where round-off errors are not negligible. Furthermore, superharmonicity, as defined

in [Landkof, 1972, p. 52], is included in this comparison. This property is interesting if the a.u.d. characteristic is not only investigated on the sphere \mathbb{S}^d but also on the ball $\mathbb{B}^{d+1} = \{\mathbf{x} \in \mathbb{R}^{d+1} : \|\mathbf{x}\| \leq 1\}$. For superharmonic kernels, the corresponding energy optimal point sets lie on the boundary of \mathbb{B}^{d+1} , i.e. on \mathbb{S}^d , cf. [Landkof, 1972, p. 163]. As the Laplacian of the LOG-kernel is $\Delta_{\mathbf{x}} \tilde{k}_{\text{LOG}}(\|\mathbf{x}\|) = (1 + d \log(\|\mathbf{x}\|/2)) / \|\mathbf{x}\|$, it is superharmonic on \mathbb{B}^{d+1} for $d > 1$. On a side note, it shall be mentioned that the LOG-kernel is strictly majorized by some Riesz kernels, i.e. $\tilde{k}_{\text{LOG}}(z) < \tilde{k}_s(z)$ for all $z \in [0, 2]$ and for $1.0624 \leq s \leq 1.2292$, which are numerically determined bounds.

Table 7.2: Comparison of various properties of the different cases for $s \in \mathbb{R}_{\geq 0} \cup \{\text{LOG}, \log\}$. If $s = \text{LOG}$ or $s = \log$, the formal index “ $s + 1$ ” means $\text{LOG} + 1 = \log$ or $\log + 1 = 1$, respectively.

	$s = \text{LOG}$	$s = \log$	$0 < s < d$	$d \leq s$
$k_s(\mathbf{x}, \mathbf{x}) = \tilde{k}_s(0)$	2	∞	∞	∞
$\tilde{k}'_s(2) = \tilde{k}_{s+1}(2)$	0	> 0	> 0	> 0
$k_s(\mathbf{x}, -\mathbf{x}) = \tilde{k}_s(2)$	0	0	> 0	> 0
superharmonicity of k_s	if $d > 1$	if $d > 1$	if $s < d - 1$	never
\mathcal{I}_s^c	$< \infty$	$< \infty$	$< \infty$	∞

Conclusively, one can say that \tilde{k}_{LOG} extends the series of kernel derivatives more “regularly”, as it inherits all gentle properties of \tilde{k}_{\log} and additionally lacks a singularity. Generally speaking, even if this intimate relationship to the classical cases were not present, there would be some motivation to study new, yet unknown kernels. For instance, in Tumanov [2013], biquadratic potentials were investigated due to the conjecture that “known solutions for several potentials may yield a solution for another potential”.

Energy optimal point sets for any case of s lead to comparable mesh ratio and similar empirical distribution functions. Also, the suitability as supporting points for spherical basis functions in the context of meshless interpolation is comparable for all cases of s . Kernels that are closer to the beginning of the sequence (7.72) yield point sets that tend to be – slightly – more accurate when interpolating constant functions, compared to kernels that appear later in the sequence. However, the provided numerical evidence for this observation is not sufficient to claim this as a fact.

By all of the investigated quality indicators, the energy optimization procedure improves the properties of Leopardi’s equal area point sets. They prove, however, to be an excellent choice for the initial positions.

The Concentric Interpolation method extends the well-known interpolation on \mathbb{S}^d by means of spherical basis functions to the surrounding space \mathbb{R}^{d+1} . It allows for accurate approximation of continuous scalar functions and, if existent, their gradient and their Hessian. This claim is proved by means of a real-world example in \mathbb{R}^8 , where an

engineering function with pronounced nonlinearities is interpolated by the CI method in a “black-box” manner, i.e. without considering any specific properties of the function.

The efficiency of both the energy minimizing routine and the CI scheme may be significantly enhanced if the problem under consideration permits the utilization of symmetric kernel functions.

The developed software, both for the generation of the point sets and for the interpolation, is published as open source [Fritzen & Kunc, 2018b,a]. Special emphasis is put on self-contained, modular, and accessible code.

7.8.2 Discussion

A notable feature of the CI method is that the radial accuracy is principally a mere matter of offline cost. This is due to the fact that the evaluation of the radial interpolants $\mathcal{R}_1, \dots, \mathcal{R}_N(x)$ does not influence the algorithmic complexity of evaluating the surrogate \tilde{f} and its derivatives. In practice, a piecewise definition of the radial interpolants had only negligibly negative effects on the runtime. By the results depicted in Figure 7.12, it is expected that the insertion of additional supporting points along the same training directions would notably improve the accuracy. As the radial interpolation is independent of the tangential interpolation, such a radial refinement of an already set up CI scheme can be done at negligible computational costs, if the corresponding data is provided.

The currently employed LDL solver for the application of the inverse kernel matrix \underline{K}^{-1} is not unconditionally stable as the kernel parameter γ approaches zero. However, it greatly improves the stability of the method as compared to pre-computation of \underline{K}^{-1} . Various suggestions on an improvement of the accuracy for small values of the kernel parameter have been discussed in the literature. To name just two out of many, Rashidinia et al. [2016] proposed a reformulation of the interpolant by means of Mercer’s theorem, while Wright & Fornberg [2017] utilized vector-valued rational approximations. Further, the efficiency of the method could possibly be improved by employing the idea of localized bases, e.g. as in Fuselier et al. [2013]. These issues should be addressed in future work on the software. A thorough error estimation within appropriate function spaces, e.g. in the sense of Narcowich et al. [2006b], is yet outstanding.

One can also consider minimum energy points with respect to the kernels applied to the geodesic distance, $\tilde{k}_s(z_g)$, cf. e.g. Leopardi [2013]. In this case, the kernels and their symmetrized variants have the appearance visualized in Figure 7.13. In this context, the LOG-kernel is substituted with $z_g(\log \frac{z_g}{\pi} - 1) + \pi$.

The energy minimization algorithm may easily be adapted to this geodesic case. Random samples of such point sets did not exhibit notable differences to the corresponding point sets of the Euclidean case. However, it shall be noted that geodesic Riesz kernels come along with an additional drawback: if just two points are considered, then the gradient of the respective energy, $I_s(N = 2)$, with respect to the point positions

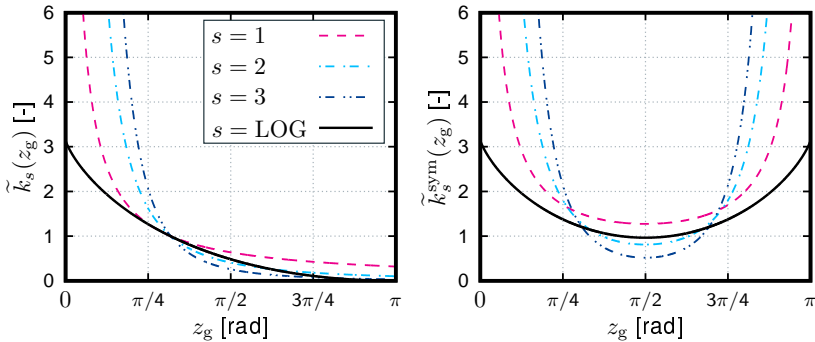


Figure 7.13: Geodesic kernels $\tilde{k}_s(z_g)$ and their symmetric versions $\tilde{k}_s^{\text{sym}}(z_g)$ for the cases $s = 1, 2, 3, \text{LOG}$. See also Figures 7.1 and 7.2.

does not continuously approach zero as the points reach antipodal configuration. This discontinuity in the gradient may be relevant in combination with certain stopping criteria, but has not been investigated yet. The log- and the LOG-kernel do not have this possibly disadvantageous property.

Finally, it might be worth speculating about the implications the newly introduced LOG-kernel could have on Smale's 7th problem, cf. Smale [1998]. This problem calls for an algorithm that yields log-optimal N -point configurations on spheres \mathbb{S}^d within a time that is polynomial in N . Progress in this respect has been made, e.g. in Nerattini et al. [2014] and by many others. The regularity of the LOG-case, cf. Table 7.2, might possibly give rise to new approaches. For instance, the much studied theory of Reproducing Kernel Hilbert Spaces, e.g. as in Damelin et al. [2010], could be the starting point for new progress in this regard. We postpone further discussion of this highly speculative claim to future works.

Conflict of interest

The authors declare that they have no conflict of interest.

Data availability statement

All of the point sets used in the numerical examples are available in the Matlab software package's subfolder `examples/exports`, see [Fritzen & Kunc, 2018b], and in the C++ software packages subfolder `data/directions`, see [Fritzen & Kunc, 2018a].

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – FR2702/6 and FR2702/8 – in the scope of the Emmy-Noether and Heisenberg funding lines. Oliver Kunc also received funding by the DFG within the Cluster of Excellence in Simulation Technology (Grant EXC 310/2, Project PN1-23) at the University of Stuttgart which is highly appreciated. The authors thank the anonymous reviewers for valuable comments that lead to significant improvements of the manuscript.

7.9 Appendix

7.9.1 Generation of random points on \mathbb{S}^d

The generation of *uniformly random points* on \mathbb{S}^d , can be realized in several ways. Two prominent and convenient examples include:

- **Generation using the normal distribution \mathcal{N}** [cf. Marsaglia, 1972]

The easiest way of generating the sought-after points is based on random variables n_i following a normal distribution \mathcal{N} . A point $\mathbf{x} \in \mathbb{S}^d$ is obtained by setting the components of its coordinate vector $x_i \sim \mathcal{N}$ ($i = 1, \dots, d + 1$), followed by normalization of the whole vector. A purely technical improvement of the algorithm is to abandon random samples which have a vector norm close to machine precision before the normalization in order to prevent numerical truncation.

- **Generation using the uniform distribution \mathcal{U} on $[-1, 1]$**

Another option is to seed all components of candidate points according to a uniform distribution \mathcal{U} on the interval $[-1, 1]$. Next, points having a norm greater than 1 and smaller than a tolerance determined by machine precision are rejected, i.e. only points contained in a spherical shell are accepted. The remaining points are then projected onto \mathbb{S}^d .

The advantage of the method relying on random variables following a normal distribution is that there is virtually no rejection while the second algorithm will lead to a substantial amount of rejected points: for $d = 1$ the chance for rejection is 21.46%, for $d = 2$ it is 47.64% and for arbitrary spherical dimension d it is defined by

$$P(\text{candidate is rejected}) = 1 - \frac{L^{d+1}(\mathbb{B}^{d+1})}{2^{d+1}}, \quad (7.74)$$

where L^{d+1} denotes the Lebesgue measure in \mathbb{R}^{d+1} and $\mathbb{B}^{d+1} = \{\mathbf{x} \in \mathbb{R}^{d+1} : \|\mathbf{x}\| \leq 1\}$. Since Gaussian random variables are available at little numerical expense in many software libraries, the first algorithm is available as an option in the graphical user interface of our Matlab software. If it is selected, the resulting point set is passed to Algorithm 7.1 as initial configuration.

7.9.2 Algorithms

Algorithm 7.1: Minimum energy algorithm without nonlinear constraints.

Input : number of points N ; case parameter s ;
 symmetrization flag (if true, then symmetrized kernels are employed);
 number of cycles n_C^{\min} and iterations per cycle n_{it}^{\min} for the
 minimization of the energy;

[OPTIONAL] number of cycles n_C^F and iterations per cycle n_{it}^F for the
 subsequent minimization of the gradient;

[OPTIONAL] initial configuration $\underline{X} \in \mathbb{R}^{(d+1) \times N}$

Output: locally s -energy optimal point set $\underline{X} \in \mathbb{R}^{(d+1) \times N}$

```

1 if no initial  $\underline{X}$  provided then
2   |  $\underline{X} \leftarrow$  equal area points, cf. Leopardi Leopardi [2006]
3   |  $\underline{X} \leftarrow$  RenormalizeColumns(  $\underline{X} + 0.01$  RandomDirections(d+1,N) )
   | // see Appendix 7.9.1
4 end
5 for  $i_C^{\min} = 1, \dots, n_C^{\min}$  do // part 1: energy minimization, unconstrained
6   |  $\underline{X} \leftarrow$  fminunc( $\underline{X}$ ) // max.  $n_{it}^{\min}$  iterations
7   |  $\underline{X} \leftarrow$  RenormalizeColumns(  $\underline{X}$  )
8   | check function and gradient: accept and exit? (else: continue)
9 end
10 for  $i_C^F = 1, \dots, n_C^F$  do // part 2: gradient minimization, unconstrained
11   |  $\underline{X} \leftarrow$  lsqnonlin( $\underline{X}$ ) // max.  $n_{it}^F$  iterations
12   |  $\underline{X} \leftarrow$  RenormalizeColumns(  $\underline{X}$  )
13   | check function and gradient: accept and exit? (else: continue)
14 end
15 return  $\underline{X}$ 

```

For Algorithm 7.2, the quantity \tilde{f}_γ denotes the surrogate model \tilde{f} , cf. (CI), with respect to the kernel function $\tilde{k}(z_g) = \exp(-\gamma z_g^2)$, approximating the original function f .

Algorithm 7.2: Optimization of the kernel parameter γ .

Input : $\gamma_{\min}, \gamma_{\max}$: limits for the gamma search, $0 < \gamma_{\min} < \gamma_{\max}$;

$N_{\text{regular}}, N_{\text{bisection}}$: maximum numbers of iterations for regular and bisectional optimization (≥ 0);

β : bisection factor, $0 < \beta < 2$, defaults to $\beta = 1$

Output: γ_{best} minimizing $\tilde{\Phi}$

// initialization

1 set $\Delta\gamma = \frac{\gamma_{\max} - \gamma_{\min}}{N_{\text{regular}} + 1}$ and $\gamma_{\text{best}} = \gamma_{\min}$

// part 1: global optimization on regular γ grid

2 for $\gamma = \gamma_{\min}, \gamma_{\min} + \Delta\gamma, \dots, \gamma_{\max}$ // $N_{\text{regular}} + 2$ values of γ

3 do

4 | if $\|\tilde{f}_\gamma - f\|_{L^2(\mathbb{S}^d)} < \|\tilde{f}_{\gamma_{\text{best}}} - f\|_{L^2(\mathbb{S}^d)}$ then $\gamma_{\text{best}} = \gamma$

5 end

6 if $\gamma_{\text{best}} = \gamma_{\min}$ or $\gamma_{\text{best}} = \gamma_{\max}$ then exit with error message “Please restart with different boundaries”

// part 2: local optimization via bisection

7 $\gamma_2 = \gamma_{\text{best}} \pm \Delta\gamma$, i.e. neighbor of γ_{best} with the lower value of $\|\tilde{f}_\gamma - f\|_{L^2(\mathbb{S}^d)}$

8 for $n_{\text{bisection}} = 0, \dots, N_{\text{bisection}} - 1$ do

9 | $\gamma_{\text{mid}} = \gamma_{\text{best}} + \frac{\beta}{2} (\gamma_2 - \gamma_{\text{best}})$

10 | if $|\gamma_{\text{mid}} - \gamma_{\text{best}}| < \text{TOL}_{L\gamma}$ then return γ_{best}

11 | if $\|\tilde{f}_{\gamma_{\text{mid}}} - f\|_{L^2(\mathbb{S}^d)} < \|\tilde{f}_{\gamma_{\text{best}}} - f\|_{L^2(\mathbb{S}^d)}$ then

12 | | set $\gamma_2 = \gamma_{\text{best}}$ and $\gamma_{\text{best}} = \gamma_{\text{mid}}$

13 | else

14 | | $\gamma_2 = \gamma_{\text{mid}}$

15 | end

16 | if $\left| \frac{\|\tilde{f}_{\gamma_2} - f\|_{L^2(\mathbb{S}^d)} - \|\tilde{f}_{\gamma_{\text{best}}} - f\|_{L^2(\mathbb{S}^d)}}{\|\tilde{f}_{\gamma_{\text{best}}} - f\|_{L^2(\mathbb{S}^d)}} \right| < \text{TOL}_{\text{func}}$ then return γ_{best}

17 end

18 return γ_{best}

Chapter 8:

Fourth publication – Many-scale finite strain computational homogenization via Concentric Interpolation

Original publication:

Kunc, O. & Fritzen, F.: Many-scale finite strain computational homogenization via Concentric Interpolation. *International Journal for Numerical Methods in Engineering* (2020), Wiley, doi:10.1002/nme.6454

Abstract. A method for efficient computational homogenization of hyperelastic materials under finite strains is proposed. Multiple spatial scales are homogenized in a recursive procedure: starting on the smallest scale, few high fidelity FE computations are performed. The resulting fields of deformation gradient fluctuations are processed by a snapshot POD resulting in a Reduced Basis model (RB). By means of the computationally efficient RB model, a large set of samples of the homogenized material response is created. This data set serves as the support for the Concentric Interpolation scheme (CI), interpolating the effective stress and stiffness. Then, the same procedure is invoked on the next larger scale with this CI surrogating the homogenized material law. A three-scale homogenization process is completed within few hours on a standard workstation. The resulting model is evaluated within minutes on a laptop computer in order to generate fourth-scale results. Open source code is provided.

Keywords: computational homogenization, multi-scale, hyperelasticity, Hencky strain, reduced basis, concentric interpolation

8.1 Introduction

Homogenization of mechanical properties of solid bodies has caught the attention of researchers for at least 110 years, [Voigt, 1910]. The ever-increasing interest in efficient two-scale homogenization methods is fueled by the sustainable trend to digitize and automate engineering tasks. This is relevant for both academia and industry.

The sub-field of finite strain homogenization comes along with additional difficulties when compared to the infinitesimal strain context: the infinite amount of possible

measures of strain and stress, dependencies on finite rotations, and the physical constraint of material self-impenetrability. These are just few examples of increased conceptual and computational complexity as compared to the infinitesimal strain context.

Nonetheless, notable progress in the field of two-scale finite strain homogenization was made recently. An example is the production, experimental investigation and analytical modeling of fibre reinforced rubber-like materials citepCiarletta2011. In other cases, glass fibre reinforced polymers were efficiently homogenized using a Fast Fourier Transformation (FFT) scheme, [Kabel et al., 2014]. A related method was developed for magneto-active materials, [Rambausek et al., 2019].

Another interesting movement within the homogenization community is the quest for multi-scale schemes that are capable of bridging more than one scale. For instance, upscaling techniques for three-scale problems involving carbon fibers and epoxy matrices, [Yuan & Fish, 2009], layered solids, [Ramírez-Torres et al., 2018], or models of porous materials with cracks, [Takano & Okuno, 2004], were developed. Impressively, the mechanical properties of a biological tissue modeled on seven scales were reported, [Nikolov et al., 2010]. These and other works exemplify the applicability of homogenization techniques based on multiple, hierarchical separations of scales.

Multiple spatial scales are also studied within scenarios of fractal or self-similar microstructures. Recent examples for this include the analytical homogenization of the elastic properties in the presence of fractal pores, [Wu et al., 2018] or (without assumed separation of scales) fractal interfaces, [Heida et al., 2020, Jarroudi & Er-Riani, 2018]. Such considerations are motivated, e.g., by geology or fracture mechanics.

The present work investigates a computational method for the mechanical homogenization of a large amount of spatial scales – hence the term “many-scale homogenization”. It relies on the presence of exclusively hyperelastic constitutive laws on the smallest scale. There, no assumptions are made on the degree of anisotropy or nonlinearity of the material laws. Also, it allows for an arbitrary amount of material phases – including voids – in any geometric layout, given periodic Representative Volume Elements (RVE). The latter may or may not differ on each scale. The case of similar RVE’s on each scale relates this method to fractal schemes.

The proposed scheme is based on the idea of transitioning from computationally demanding, high fidelity methods to methods compromising some of the accuracy for efficiency. This transition is conducted in a staged manner on each scale, and many scales are processed successively. The scheme is illustrated in Figure 8.1. Starting on Scale 0, high fidelity Finite Element (FE) solutions are produced. These are then processed by means of a classical Proper Orthogonal Decomposition (POD) which yields a Reduced Basis (RB). The latter is in turn used to evaluate the homogenized material response at numerous specifically chosen sampling sites within the strain space. This set of data serves as the support of the Concentric Interpolation (CI) scheme, which rapidly approximates the effective constitutive response. With this fast, purely numerical surrogate at hand, the next scale can be homogenized in a similar manner. Eventually,

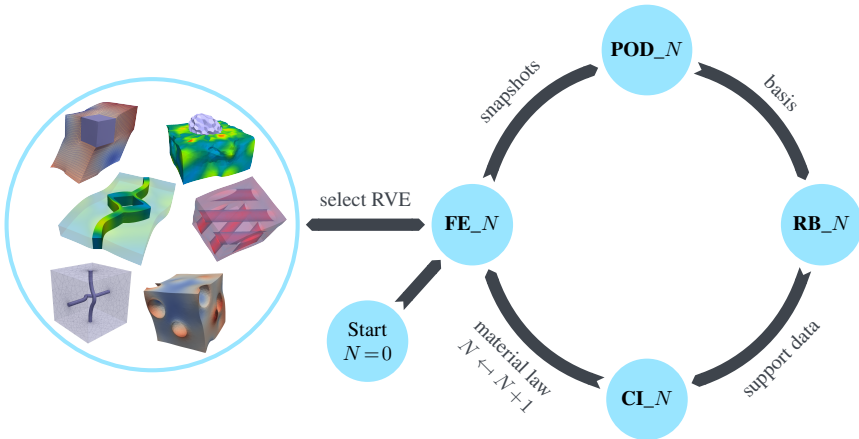


Figure 8.1: $N = 0$ Proposed many-scale homogenization scheme for finite strain hyperelasticity. Each complete cycle corresponds to the homogenization of one scale. The procedure is finished after the homogenization of Scales $0, 1, \dots, M - 1$. At this point, the material law $\mathbf{E} \mapsto \mathbf{S}^{(M)}, \mathbb{C}^{(M)}$ on Scale M is efficiently approximated by Concentric Interpolation. Note that the FE method may be replaced by any high-fidelity method, e.g. FFT.

the resulting CI method approximates the material law that is relevant on the engineering Scale M .

The outline is as follows. In Section 8.2, essential notation is introduced and the multi-scale problem as well as the RB scheme are recalled. Section 8.3 contains the many-scale homogenization algorithm with detailed comments on each step. Numerical examples are presented in Section 8.4. A summary and additional discussions are given in Section 8.5.

8.2 Notation, problem formulation and reduced order model

8.2.1 Notation

Coordinate-free descriptions of first and second order tensors are denoted by bold letters, e.g. \mathbf{X}, \mathbf{F} . The order of a tensor is not related to the capitalization of its representative symbol. Fourth order tensors are written in blackboard bold symbols, e.g. \mathbb{C} . Only the set of real numbers \mathbb{R} , its subset $\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$, and the hypersphere $\mathbb{S}^5 = \{\underline{x} \in \mathbb{R}^6 : \|\underline{x}\| = 1\}$ are exceptions to this rule. Real vectors and matrices are denoted by a single and a double underline, e.g. \underline{X} and $\underline{\underline{F}}$, respectively. Vectors are assumed to be columns, unless transposed, e.g. \underline{X}^T .

In this work, all simulations are conducted in three spatial dimensions. Let $\Omega_0 \subset \mathbb{R}^3$ be the domain of the simply connected physical body in its undeformed state, and $\Omega \subset \mathbb{R}^3$ be the domain occupied by the deformed body. Points within these two sets are referred to as $\mathbf{X} \in \Omega_0$ and $\mathbf{x} \in \Omega$, respectively.

In order to measure the deformation, the deformation gradient $\mathbf{F} = \partial \mathbf{x} / \partial \mathbf{X}$ and the right Cauchy-Green tensor $\mathbf{C} = \mathbf{F}^\top \mathbf{F}$ are introduced. The physical impossibility of material self-penetration implies the positiveness of the determinant $J = \det(\mathbf{F}) > 0$, which in turn guarantees the unique existence of the polar decomposition $\mathbf{F} = \mathbf{R}\mathbf{U}$. This yields the rotation tensor \mathbf{R} , with the properties $\det(\mathbf{R}) = 1$ and $\mathbf{R}^\top \mathbf{R} = \mathbf{I}$, where \mathbf{I} is the identity tensor, as well as the symmetric positive definite right stretch tensor \mathbf{U} . The Hencky strain $\mathbf{E} = \log(\mathbf{U})$ is of utmost importance to the present work.

Hyperelastic material laws are characterized by elastic energy density functions $W_F(\mathbf{F})$ or $W_C(\mathbf{C})$, respectively, which are equal if their arguments correspond to each other. The \mathbf{F} -energy W_F density gives rise to the first Piola-Kirchhoff stress tensor via differentiation, $\mathbf{P} = \partial W(\mathbf{F}) / \partial \mathbf{F}$, while the \mathbf{C} -energy density W_C defines the second Piola-Kirchhoff stress tensor, $\mathbf{S} = \partial W_C(\mathbf{C}) / (2\partial \mathbf{C})$. These two stress measures are convertible by the relation $\mathbf{P} = \mathbf{F}\mathbf{S}$. The tangent moduli of the first and of the second Piola-Kirchhoff stresses read $\mathbb{C}_F = \partial^2 W_F(\mathbf{F}) / \partial \mathbf{F}^2$ and $\mathbb{C} = \partial^2 W_C(\mathbf{C}) / (2\partial \mathbf{C})^2$, respectively.

For spatial descriptions, the standard Cartesian basis $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \mathbf{e}^{(3)}$ is employed. It is orthonormal with respect to the Euclidean inner/dot product and the norm $\|\bullet\|$ induced thereby, i.e. $\underline{\mathbf{e}}^{(i)} \cdot \underline{\mathbf{e}}^{(j)} = \underline{\mathbf{e}}^{(i)\top} \underline{\mathbf{e}}^{(j)} = \delta_{ij}$ ($i, j = 1, \dots, 3$) holds for the vectorized counterpart of the basis, and δ_{ij} denotes the Kronecker symbol. Components of tensorial quantities are with respect to the tensorial basis $\{\mathbf{e}^{(i_1)} \otimes \dots \otimes \mathbf{e}^{(i_A)}\}_{i_1, \dots, i_A=1,2,3}$, where $A \in \{1, 2, 4\}$ is the respective tensorial order. Thus, there are natural representations of tensors of first and second order as vectors and matrices respectively, e.g. $\mathbf{X} \leftrightarrow \underline{\mathbf{X}} \in \mathbb{R}^3$ or $\mathbf{E} \leftrightarrow \underline{\mathbf{E}} \in \mathbb{R}^{3 \times 3}$. For vectors and matrices of arbitrary dimensions, the Euclidean norm and, respectively, the Frobenius norm are employed.

Two kinds of special tensorial bases are employed. Using a Voigt-like notation, vectorizations of the symmetric second order stress tensor $\mathbf{S} \leftrightarrow \underline{\mathbf{S}} \in \mathbb{R}^6$ are formulated with respect to the basis

$$\begin{aligned} \mathbf{b}^{(1)} &= \mathbf{e}^{(1)} \otimes \mathbf{e}^{(1)}, & \mathbf{b}^{(2)} &= \mathbf{e}^{(2)} \otimes \mathbf{e}^{(2)}, & \mathbf{b}^{(3)} &= \mathbf{e}^{(3)} \otimes \mathbf{e}^{(3)}, \\ \mathbf{b}^{(4)} &= \frac{1}{\sqrt{2}} \mathbf{e}^{(1)} \otimes \mathbf{e}^{(2)}, & \mathbf{b}^{(5)} &= \frac{1}{\sqrt{2}} \mathbf{e}^{(1)} \otimes \mathbf{e}^{(3)}, & \mathbf{b}^{(6)} &= \frac{1}{\sqrt{2}} \mathbf{e}^{(2)} \otimes \mathbf{e}^{(3)}. \end{aligned} \quad (8.1)$$

Likewise, the matrix variant of the fourth-order stiffness tensor $\mathbb{C} \leftrightarrow \underline{\underline{\mathbf{C}}} \in \mathbb{R}^{6 \times 6}$ with both minor and major symmetry employs the same basis. This basis is compatible with the Euclidean norm, i.e. $\|\mathbf{S}\| = \|\underline{\mathbf{S}}\|$. In general, ambiguity of the components of a second order tensor and of its vectorized counterpart is not possible since the former are indexed by two variables, whereas the latter has only one index variable.

Another basis is used with regard to the vectorization of the symmetric Hencky strain $\mathbf{E} \leftrightarrow \underline{\mathbf{E}} \in \mathbb{R}^6$. For this purpose, a family of special bases $\mathcal{Y}_{J^*} = \{\underline{\mathbf{Y}}^{(1)}, \dots, \underline{\mathbf{Y}}^{(5)}, \underline{\mathbf{Y}}^{(6)}(J^*)\}$ is introduced:

$$\begin{aligned} \underline{\mathbf{Y}}^{(1)} &= \sqrt{\frac{1}{6}} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, & \underline{\mathbf{Y}}^{(2)} &= \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \\ \underline{\mathbf{Y}}^{(3)} &= \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & \underline{\mathbf{Y}}^{(4)} &= \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, & (8.2) \\ \underline{\mathbf{Y}}^{(5)} &= \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, & \underline{\mathbf{Y}}^{(6)}(J^*) &= \frac{\log(J^*)}{3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

These bases are parametrized by the constant dilatational scaling factor,

$$1 < J^* = \text{const}, \quad (8.3)$$

the meaning of which will be given towards the end of this subsection. The components E_1, \dots, E_6 of $\underline{\mathbf{E}}$ are the coefficients of the \mathcal{Y}_{J^*} -representation of $\underline{\mathbf{E}}$,

$$\underline{\mathbf{E}} = \sum_{i=1}^6 E_i \underline{\mathbf{Y}}^{(i)} \leftrightarrow \mathbf{E}, \quad (8.4)$$

where the explicit dependence of the sixth basis element on the scaling factor, $\underline{\mathbf{Y}}^{(6)} = \underline{\mathbf{Y}}^{(6)}(J^*)$, is dropped for notational convenience. The equivalence (8.4) occasionally leads to a slight abuse of notation implying $\underline{\mathbf{E}} = \mathbf{E}$, which can always be uniquely resolved from the context. Further, the vector $\underline{\mathbf{E}}$ is decomposed as

$$\underline{\mathbf{E}} = t\underline{\mathbf{N}}, \quad t = \|\underline{\mathbf{E}}\|, \quad \underline{\mathbf{N}} = \frac{1}{t}\underline{\mathbf{E}}, \quad (8.5)$$

where $t \in \mathbb{R}$ and $\underline{\mathbf{N}} \in \mathbb{S}^5 \subset \mathbb{R}^6$ are called the magnitude and the direction in Hencky strain space, respectively. While the magnitude, t , is one possible scalar measure of strain, the direction in Hencky strain space, $\underline{\mathbf{N}}$, characterizes the kind of the strain. For instance, $\underline{\mathbf{N}} = [1, 0, 0, 0, 0, 0]^T \leftrightarrow \underline{\mathbf{Y}}^{(1)}$ corresponds to isochoric uniaxial extension and $\underline{\mathbf{N}} = [0, 0, 0, 0, 0, 1]^T \leftrightarrow \underline{\mathbf{Y}}^{(6)}(J^*)$ corresponds to pure dilatation.

This vectorized notation requires a particular choice of \mathcal{Y}_{J^*} . This choice will always be clear from the context. For simplicity, we only consider Hencky strains with the restriction

$$t \in [0, T_{\max}] \quad (0 < T_{\max} \leq 1). \quad (8.6)$$

Therefore, the constant dilatational scaling factor, J^* , determines the maximum and the minimum possible determinant of the associated stretch tensor,

$$J = \det(\exp(t(\pm 1)\underline{Y}^{(6)}(J^*))) \in \left[\frac{1}{J^*}, J^*\right]. \quad (8.7)$$

In general, adaptations to T_{\max} as well as to the normalization of the basis elements $\underline{Y}^{(1)}, \dots, \underline{Y}^{(5)}$ are admissible. For incompressible materials, it is reasonable to consider only the five-dimensional Hencky strain space spanned by the first five elements of (8.2) to which the following theory can be easily adapted, [Kunc & Fritzen, 2019b].

It is important to notice that the magnitude, which is the modulus of the vector representation, is generally not the same as the norm of the matrix representation, i.e. $t = \|\underline{E}\| \neq \|\underline{\underline{E}}\|$ whenever $J^* \neq \exp(\sqrt{3})$ and $[0, 0, 0, 0, 0, 1] \underline{N} \neq 0$. While the quantity $\|\underline{\underline{E}}\|$ might be of significance in other contexts, we only ever use t instead.

8.2.2 Multi-scale setting

Recap of two-scale homogenization

The starting point is the (quasi-)static formulation of the balance equations of linear momentum with respect to the reference configuration Ω_0 , $\text{Div}_{\mathbf{X}}(\mathbf{P}) = \mathbf{B}_{\text{ext}}$, where \mathbf{B}_{ext} denotes the body forces. Complementary, sufficient boundary conditions are assumed to be provided and the balance equations of angular momentum hold, $\mathbf{F}^{-1}\mathbf{P} = \mathbf{P}^T\mathbf{F}^{-T}$. The latter will not be addressed in the sequel for the sake of brevity.

In many cases, the domain Ω_0 consists of a very detailed and heterogeneous micro-structure that must be resolved for a sufficiently accurate approximation of the solution of the above problem. On one hand, the microscopic details considered here are of a much smaller characteristic length than that of the overall domain Ω_0 . On the other hand, these micro-features are still large enough to justify the assumptions of general continuum mechanics.

A well-known method to avoid an excessively fine discretization of the domain Ω_0 (e.g. with a FE mesh containing an overly large number of elements) is the method of asymptotic homogenization with assumed separation of scales, [Pruchnicki, 1998]. For more details on this well-established procedure, the reader is referred to standard literature, [Nemat-Nasser & Hori, 1999].

Multi-scale problem

As mentioned in the introduction, some realistic scenarios make multiple, hierarchical separations of scales desirable. Formally, this is achieved by iteratively stating the assumptions of the classical two-scale procedure. A rigorous derivation and mathematical analysis of this procedure can be found in the seminal works by Bensoussan et al., [Ben-

soissan et al., 1978, Chapter 1, Section 8]. An informal motivating description of this process is now stated briefly.

Consider a setting in which, after a single scale separation, the emanating Representative Volume Element (RVE) of the smaller scale in turn exhibits geometric features that are of a much smaller characteristic length than the RVE. Then, another scale separation leads to an even smaller scale with its own RVE and so forth. The total number of scales that is considered within this context is $M + 1$. We count the scales zero-based and in ascending order, i.e. Scale 0 denotes the smallest scale and Scale M denotes the engineering scale, e.g. the one on which we would like to resolve the original domain. The index $0 \leq N \leq M$ is reserved for reference to an arbitrary scale.

Quantities on Scale N are denoted by N overlines, e.g. $\overline{\mathbf{F}}(\mathbf{X})$ and $\overline{\overline{\mathbf{F}}}(\overline{\overline{\mathbf{X}}})$ are the deformation gradients on Scales 0 and 1, respectively. However, if $N > 3$ or if N is not specified, the notation is switched to the corresponding number in parentheses as a superset, e.g. $\overline{\overline{\mathbf{P}}} = \overline{\overline{\overline{\mathbf{P}}}}^{(2)}$ or $\overline{\overline{\mathbf{C}}} = \overline{\overline{\overline{\overline{\mathbf{C}}}}}^{(N)}$. For instance, the characteristic lengths of all scales are

$$0 < \underbrace{L}_{\text{Scale 0}} \ll \underbrace{\overline{L}}_{\text{Scale 1}} \ll \underbrace{\overline{\overline{L}}}_{\text{Scale 2}} \ll \dots \ll \underbrace{\overline{\overline{\overline{L}}}}_{\text{Scale } N} \ll \dots \ll \underbrace{\overline{\overline{\overline{\overline{L}}}}}_{\text{Scale } M}. \quad (8.8)$$

Furthermore, for each separated Scale N ($0 \leq N < M$) there is a RVE which is denoted by $\overline{\overline{\Omega}}_0^{(N)}$ in its undeformed configuration. Note that from now on Ω_0 represents the RVE of Scale 0 and $\overline{\overline{\Omega}}_0^{(M)}$ denotes the whole, original domain without geometric features that are of characteristic length $\leq \overline{\overline{\overline{L}}}^{(M-1)}$.

The scale separation approach requires compatible kinematic boundary conditions. Multiple such boundary conditions are available and they are characterized by the fact that they fulfill the Hill-Mandel condition. For more details and a discussion of some possible choices, the reader is referred to standard literature, [Miehe, 2003]. Here, periodic fluctuation boundary conditions for the displacements, $\overline{\overline{\mathbf{u}}} = \overline{\overline{\mathbf{x}}} - \overline{\overline{\mathbf{X}}}$, are chosen. These imply the kinematic scale-coupling relations

$$\langle \overline{\overline{\mathbf{F}}} \rangle = \overline{\overline{\mathbf{F}}}^{(N+1)}(\overline{\overline{\mathbf{X}}})^{(N+1)}, \quad (8.9)$$

employing the volume averaging operator

$$\langle \overline{\overline{\bullet}} \rangle = \frac{1}{|\overline{\overline{\Omega}}_0^{(N)}|} \int_{\overline{\overline{\Omega}}_0^{(N)}} \overline{\overline{\bullet}}^{(N)}(\overline{\overline{\mathbf{X}}})^{(N)} dV. \quad (8.10)$$

Moreover, the right-hand side of (8.9) representing the scale coupling induces a kinematic boundary condition for the problem on Scale N . The complete set of coupled balance equations is stated in the following box:

Scale M :	$\text{Div}_{\mathbf{X}}^{(M)}(\mathbf{P}) = \mathbf{B}_{\text{ext}}^{(M)}$	(BC)	(8.11)
Scale $M - 1$:	$\text{Div}_{\mathbf{X}}^{(M-1)}(\mathbf{P}) = \mathbf{0}$	$\langle \mathbf{F} \rangle = \mathbf{F}^{(M)}(\mathbf{X})$	(8.12)
\vdots	\vdots	\vdots	
Scale N :	$\text{Div}_{\mathbf{X}}^{(N)}(\mathbf{P}) = \mathbf{0}$	$\langle \mathbf{F} \rangle = \mathbf{F}^{(N+1)}(\mathbf{X})$	(8.13)
\vdots	\vdots	\vdots	
Scale 1:	$\text{Div}_{\mathbf{X}}(\overline{\mathbf{P}}) = \mathbf{0}$	$\langle \overline{\mathbf{F}} \rangle = \overline{\overline{\mathbf{F}}}(\overline{\mathbf{X}})$	(8.14)
Scale 0:	$\text{Div}_{\mathbf{X}}(\mathbf{P}) = \mathbf{0}$	$\langle \mathbf{F} \rangle = \overline{\mathbf{F}}(\overline{\mathbf{X}})$	(8.15)

The boundary conditions (BC) to Scale M may be of any generally admissible type, e.g. kinematic, static or mixed. For each material point $\mathbf{X}^{(N)}$ on Scale N ($0 < N \leq M$), the deformation gradient $\mathbf{F}^{(N)}(\mathbf{X})$ acts as a parameter on Scale $N - 1$, [Bensoussan et al., 1978, Chapter 1, Remark 8.5] representing the corresponding boundary condition.. Hence, the problems (8.12)–(8.15) need to be solved in a nested manner. On Scale 0, the material laws of all constituents are assumed to be given explicitly. The ridiculously large computational effort associated with the solution of (8.11) becomes even more obvious when considering that multiple iterations (e.g. of the Newton method) are required – on each scale. This exponential growth of the algorithmic complexity is long-known and was even literally indicated in the denomination of the FE² method, [Feyel, 1999].

For notational clarity, the argument of the boundary condition will be omitted in the sequel, $\mathbf{F}^{(N)}(\mathbf{X}) = \mathbf{F}^{(N)}$. It is important to note that the body forces on all but the top scale must equal zero, $\mathbf{B}_{\text{ext}}^{(M-1)} = \dots = \overline{\mathbf{B}}_{\text{ext}} = \mathbf{B}_{\text{ext}} = \mathbf{0}$, in order for the asymptotic homogenization ansatz to remain valid.

As with the classical two-scale results, the hyperelasticity property is conserved by upscaling, i.e. the characterizing energy density function on the next scale is the volume average of the respective function on the current scale. The same holds for all additive quantities, but especially *not* for those that are non-linear with respect to the displacement $\mathbf{u}^{(N)} = \mathbf{x}^{(N)} - \mathbf{X}^{(N)}$. The most relevant statements in this regard are given in the following box:

kinematics:	$\langle \mathbf{F} \rangle = \mathbf{F}^{(N+1)}$	(8.16)
	$\langle \mathbf{C} \rangle \neq \mathbf{C}^{(N+1)} = \mathbf{F}^{(N+1)\top} \mathbf{F}^{(N+1)}$	(8.17)
energies:	$\langle W_F \rangle = W_F^{(N+1)}$	(8.18)
	$\langle W_C \rangle = W_C^{(N+1)}$	(8.19)
stresses:	$\langle \mathbf{P} \rangle = \mathbf{P}^{(N+1)} = \partial W_F^{(N+1)} / \partial \mathbf{F}^{(N+1)}$	(8.20)
	$\langle \mathbf{S} \rangle \neq \mathbf{S}^{(N+1)} = \partial W_C^{(N+1)} / (2\partial \mathbf{C}^{(N+1)})$	(8.21)
tangent moduli:	$\langle \mathbb{C}_F \rangle \neq \mathbb{C}_F^{(N+1)} = \partial^2 W_F^{(N+1)} / (\partial \mathbf{F}^{(N+1)})^2$	(8.22)
	$\langle \mathbb{C} \rangle \neq \mathbb{C}^{(N+1)} = \partial^2 W_C^{(N+1)} / (2\partial \mathbf{C}^{(N+1)})^2$	(8.23)

The inequality (8.17)₁ stems from the quadratic nature of the right Cauchy-Green tensor. From this, the inequality (8.21)₁ follows, which can be shown by the corresponding form of the Hill-Mandel condition. The non-additivity of stiffness tensors (8.22)₁ and (8.23)₁ is adding to the complexity of the computational multi-scale homogenization. Having the stiffness tensor for the subsequent scale accurately represented is required in order to setup reliable multi-scale FE simulations with low numbers of nonlinear Newton-Raphson iterations. The classical computational evaluation of $\mathbb{C}_F^{(N+1)}$ requires the usage of numerical perturbation of the corresponding stress, [Miehe, 1996], or of a computationally very demanding Schur complement technique, [Kouznetsova et al., 2001]. The authors have previously introduced a Reduced Basis method, [Kunc & Fritzen, 2019b], that can efficiently compute the effective stiffness tensor in a two-scale setting. This method will now be briefly introduced and formally generalized to multiple scales. It is crucial to the proposed computational many-scale homogenization method.

8.2.3 Reduced Basis homogenization

The essentials of the Reduced Basis (RB) homogenization method, [Kunc & Fritzen, 2019b], are now summarized in brevity. For more details, the reader is referred to the original open access source. An algorithmic comparison with competing alternative methods can be found in [Kunc & Fritzen, 2019a]. In the following, the transition from Scale N to Scale $N + 1$ is considered, i.e. $0 \leq N < M$ and $M \geq 1$. In terms of Figure 8.1, the **RB** _{N} method is now described.

At the core of the RB ansatz is the approximation of the deformation gradient field $\mathbf{F}^{(N)}(\mathbf{X})$. It is additively split into a constant part and a fluctuating part. The constant part, $\mathbf{F}^{(N+1)}$, represents the boundary condition of the respective boundary value problem on Scale N . The fluctuating part vanishes when volume averaged and is a linear combination of the $L^2(\Omega_0^{(N)})$ -orthonormal RB ansatz functions, $\{\mathbf{B}^{(N)(i)}(\mathbf{X})\}_{i=1,\dots,N_{\text{RB}}} = \mathcal{B}$. For the notation of both the set of basis functions, $\mathcal{B} = \mathcal{B}^{(N)}$, and its size, $N_{\text{RB}} = N_{\text{RB}}^{(N)}$, we omit the explicit N -dependence. The basis functions are assumed to be given at the moment, their construction is algorithmically described in the next section of this paper. Thus, the boundary condition and the vector of RB coefficients, $\underline{\xi} = \underline{\xi}^{(N)} \in \mathbb{R}^{N_{\text{RB}}}$, are included as parameters in the RB approximation

$$\mathbf{F}^{(N)}(\mathbf{X}) \approx \mathbf{F}_{\text{RB}}^{(N)}(\mathbf{X}; \mathbf{F}^{(N+1)}, \underline{\xi}) = \mathbf{F}^{(N+1)} + \sum_{i=1}^{N_{\text{RB}}} \xi_i \mathbf{B}^{(N)(i)}(\mathbf{X}). \quad (8.24)$$

As a fundamental principle of mechanics, the energy integral is sought to be minimized,

$$\langle \mathcal{W}_{\text{F}}^{(N)}(\mathbf{F}_{\text{RB}}^{(N)}(\mathbf{X}; \mathbf{F}^{(N+1)}, \underline{\xi})) \rangle \rightarrow \min_{\underline{\xi} \in \mathbb{R}^{N_{\text{RB}}}} ! \quad (8.25)$$

From this and from the fact that the ansatz (8.24) is a low-dimensional approximation within the space of all deformation gradients, it follows from [Kunc & Fritzen, 2019b, Sec. 2.3] that

$$\langle \mathcal{W}_{\text{F}}^{(N)}(\mathbf{F}_{\text{RB}}^{(N)}(\mathbf{X}; \mathbf{F}^{(N+1)}, \underline{\xi})) \rangle \geq \mathcal{W}_{\text{F}}^{(N+1)}(\mathbf{F}^{(N+1)}), \quad (8.26)$$

meaning that the solution to (8.25) realizes the best approximation of the true effective energy density.

Once sufficient convergence, cf. [Kunc & Fritzen, 2019b, (28)], of (8.25) is achieved, the RB approximations of the effective first Piola-Kirchhoff stress tensor $\mathbf{P}_{\text{RB}}^{(N+1)}$ and of its corresponding tangent modulus $\mathbb{C}_{\text{F};\text{RB}}^{(N+1)}$ are readily available, cf. [Kunc & Fritzen, 2019b, Appendix B]. Next, these are converted to their equivalents with respect to the right Cauchy-Green tensor,

$$\mathbf{S}_{\text{RB}}^{(N+1)} = \mathbf{F}^{(N+1)-1} \mathbf{P}_{\text{RB}}^{(N+1)} \quad (8.27)$$

$$\begin{aligned} \left(\mathbb{C}_{\text{RB}}^{(N+1)} \right)_{ijkl} = & \left(\sum_{a,b=1}^3 \left(\mathbf{F}^{(N+1)-1} \right)_{ia} \left(\mathbb{C}_{\text{F;RB}}^{(N+1)} \right)_{ajbl} \left(\mathbf{F}^{(N+1)-\text{T}} \right)_{bk} \right) \\ & - \left(\mathbf{C}^{(N+1)-1} \right)_{ik} \left(\mathbf{S}_{\text{RB}}^{(N+1)} \right)_{lj} \quad (i, j, k, l = 1, 2, 3). \end{aligned} \quad (8.28)$$

The formulation with respect to the second Piola-Kirchhoff stress and its tangent modulus is widely used for the implementation of material routines, e.g. in the context of the FEM. Thus, this RB method may be easily wrapped for the employment in such simulation software. This holds true even for commercial products as long as a user-defined material interface is provided, which usually is the case. The computational effort for a single evaluation of the stress in (8.27) is significantly reduced compared to the FEM as the solution of the minimization task (8.25) depends on just N_{RB} DOF. Usually, N_{RB} is on the order of 20–100. Apart from this speed-up, a good approximation of the tangent modulus is given by (8.28).

However, the volume averaging operator requires integration over the RVE $\Omega_0^{(N)}$, cf. (8.10). This operator is invoked multiple times per iteration of the minimization task (8.26). Thus, although the number of DOF is drastically reduced in comparison to the FEM, the original spatial complexity is still present by the number of integration points. For this reason, the unmodified RB method as described above is not suitable for online application in many-scale simulations. Several opportunities for further speed-up are discussed in the original paper, [Kunc & Fritzen, 2019b]. For now, we settle with the moderate speed-up factor of 10–100 with respect to the FEM (not accounting for the setup phase of the RB), as the RB method is part of the offline phase of the proposed many-scale scheme.

8.3 Computational homogenization via Concentric Interpolation

The Concentric Interpolation (CI) method was recently introduced by the authors, [Kunc & Fritzen, 2019c], in its general form. Previously, it was employed in the RNEXP method, [Fritzen & Kunc, 2018c], in an ad-hoc manner. It serves as another major building block for the proposed many-scale computational homogenization scheme. In terms of Figure 8.1, the RB method efficiently provides some evaluations of the homogenized material law. This data is used as supporting data for the subsequent CI method. While the RB method as treated in this context is based on the fundamental physical principle of energy minimization, the CI method is a purely mathematical means of function approximation. However, its particular design is specialized to the

characteristics of material laws and enables very efficient numerical treatment of material data, accounting for the usually rather sparse availability of the latter.

The C/C++ implementation of the CI scheme employed in this study is provided as open source code, [Fritzen & Kunc, 2018a]. The repository includes an example in which the hyperelastic material law ETM from Section 8.4.1 is interpolated. For this, compile `demo_largestrain` in the subfolder `examples`, and then execute the corresponding binary in the `bin` folder.

8.3.1 Basic scheme

The CI method is a specialization of the established general kernel interpolation method. It is standard, [Fasshauer & McCourt, 2015], that the approximation \tilde{f} of a scalar function $f : \mathbb{R} \rightarrow \mathbb{R}$ by N^{supp} symmetric positive definite kernel functions $k_i : \mathbb{R} \rightarrow \mathbb{R}_+$ ($i = 1, \dots, N^{\text{supp}}$) takes the form

$$f(x) \approx \tilde{f}(x) = \underline{p}^T \underline{\underline{K}}^{-1} \underline{k}(x). \quad (8.29)$$

The kernel functions $k_i(x) = k(x_i, x)$ are identical up to the locations of the support points, $x_1, \dots, x_{N^{\text{supp}}}$. These functions constitute the components of the kernel vector $\underline{k}(x)$, whereas the kernel matrix $\underline{\underline{K}} \in \mathbb{R}^{N^{\text{supp}} \times N^{\text{supp}}}$ is composed of the values $K_{ij} = k_i(x_j) = K_{ji}$. The function values at the kernel centers define the components of the vector \underline{p} , i.e. $p_i = f(x_i)$.

In the multivariate case, $f : \mathbb{R}^D \rightarrow \mathbb{R}$, $D \geq 1$, radial basis functions of the kind $k(\underline{x}, \underline{y}) = \Phi(\|\underline{x} - \underline{y}\|)$, with $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, are common for use as kernel functions. The restriction to the hypersphere, $f : \mathbb{S}^{D-1} \rightarrow \mathbb{R}$, $\mathbb{S}^{D-1} = \{\underline{x} \in \mathbb{R}^D : \|\underline{x}\| = 1\}$, by substituting the geodesic (i.e. angular) distance function, $\text{acos}(\underline{x} \cdot \underline{y})$, for the argument in Φ is known as spherical basis function ansatz, [Fasshauer & Schumaker, 1998, §6]. A popular choice for kernel functions is the Gaussian kernel, which then takes the form

$$k(\underline{x}, \underline{y}) = \exp(-\gamma \text{acos}^2(\underline{x} \cdot \underline{y})). \quad (8.30)$$

In the present work, attention is restricted to this kernel function and investigations of possible alternatives is not within the current scope. Since the kernel centers are points on a unit hypersphere, they are from now on denoted support directions, and accordingly their number is $N_{\text{dir}}^{\text{supp}}$.

The CI method is an extension of this spherical interpolation on the hypersphere \mathbb{S}^{D-1} to the surrounding Euclidean space \mathbb{R}^D . This is achieved by first splitting the argument $\underline{x} \in \mathbb{R}^D$ into its magnitude t and its direction \underline{N} ,

$$\underline{x} = t\underline{N} \qquad t = \|\underline{x}\| \qquad \underline{N} = \frac{1}{t}\underline{x}. \quad (8.31)$$

Then, the spherical interpolation ansatz from above, classically acting exclusively on $\underline{N} \in \mathbb{S}^{D-1}$, is complemented by the introduction of a dependence of the vector of function values on the magnitude $t \in \mathbb{R}_+$,

$$f(\underline{x}) \approx \tilde{f}(t, \underline{N}) = \begin{cases} \underline{p}(t)^\top \underline{K}^{-1} \underline{k}(\underline{N}) & \text{if } t \neq 0 \\ 0 & \text{if } t = 0 \end{cases}. \quad (8.32)$$

Here, it is assumed without loss of generality that $f(\underline{0}) = 0$, which can always be ensured by shifting the function values. Accordingly, the components of the function value vector $\underline{p}(t)$, $p_i(t) : \mathbb{R}_+ \rightarrow \mathbb{R}$, are chosen with the property $p_i(0) = 0$. More precisely, these functions are defined as piecewise linear polynomials and the same set of support magnitudes $0 < t_1 < \dots < t_{N_{\text{mag}}^{\text{supp}}} = 1$ is chosen for all $i = 1, \dots, N_{\text{dir}}^{\text{supp}}$. Hence, there is a distinct one-dimensional approximation of the function values along each of the $N_{\text{dir}}^{\text{supp}}$ supporting directions.

The total number of function values necessary for the setup of the CI scheme is $N_{\text{dir}}^{\text{supp}} N_{\text{mag}}^{\text{supp}}$ and can be adjusted according to the properties of the target function f . Also, the placement of the magnitude support points $t_1, \dots, t_{N_{\text{mag}}^{\text{supp}}}$ may be adapted in a problem specific manner. However, the choice of the support directions $\underline{N}^{(1)}, \dots, \underline{N}^{N_{\text{dir}}^{\text{supp}}}$ is strongly constrained by the choice of identical kernel functions of the kind (8.30) for each of the components of the kernel vector $\underline{k}(\underline{N})$, i.e. all kernel functions share the same kernel parameter γ . Thus, the directions are sought to be asymptotically uniformly distributed, [Kunc & Fritzen, 2019c].

8.3.2 Efficient implementation

The approximant \tilde{f} from (8.32) can be highly efficiently implemented: by means of dedicated matrix-vector representations of the involved quantities, the inverse kernel matrix, \underline{K}^{-1} , can be multiplied with static data during the setup phase. Then, each evaluation of $\tilde{f}(t, \underline{N})$ reduces to the evaluation of the radial interpolation vector, $\underline{p}(t)$, the kernel vector, $\underline{k}(\underline{N})$, and some matrix-vector multiplications.

For the sake of simplicity, this is now exemplified assuming the radial interpolation vector, $\underline{p}(t) \in \mathbb{R}^{N_{\text{dir}}^{\text{supp}}}$, is linear on the whole domain \mathbb{R}_+ , i.e. not just piecewise. Then, it has the discrete representation

$$\underline{p}(t) = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_{N_{\text{dir}}^{\text{supp}}} & b_{N_{\text{dir}}^{\text{supp}}} \end{bmatrix} \begin{bmatrix} t \\ 1 \end{bmatrix}. \quad (8.33)$$

Here, the components a_i and b_i are the polynomial coefficients of the component $p_i = a_i t + b_i$, $i = 1, \dots, N_{\text{dir}}^{\text{supp}}$. Substitution into (8.32) yields

$$\tilde{f}(t, \underline{N}) = \left[t \quad 1 \right] \underbrace{\begin{bmatrix} a_1 & a_2 & \dots & a_{N_{\text{dir}}^{\text{supp}}} \\ b_1 & b_2 & \dots & b_{N_{\text{dir}}^{\text{supp}}} \end{bmatrix}}_{\text{constant}} \underline{K}^{-1} \underline{k}(\underline{N}). \quad (8.34)$$

In the actual implementation with piecewise linearity of p , the polynomial coefficients are stored in $N_{\text{mag}}^{\text{supp}}$ sets, each of which is multiplied with the inverse kernel matrix in analog to (8.34) during the setup phase.

8.3.3 Concentric Interpolation of material laws

Finite Element software oftentimes provides an interface to user-defined material routines. In finite strain hyperelasticity, material laws are usually defined with respect to rotationally invariant kinematic descriptors, such as \underline{U} or $\underline{C} = \underline{U}^2$. However, as the authors previously emphasized, [Kunc & Fritzen, 2019b], it is advantageous for a number of reasons to sample mechanical response functions on the space of Hencky strains, $\underline{E} = \log(\underline{U})$. Additionally, the space of Hencky strains is isomorphic to \mathbb{R}^6 , cf. (8.4), allowing for the application of CI, $\tilde{f} : \mathbb{R}^6 \rightarrow \mathbb{R}$. In principle, the form of the interpolant, cf. (8.32), is suitable for approximating the hyperelastic energy density function, $W_C = W_C(\exp(2\underline{E}))$. Further, if sufficiently smooth one-dimensional radial interpolation functions, p_i , were used instead of piecewise linear polynomials, the stress, \underline{S} , and its tangent modulus, \underline{C} , could be computed by means of differentiation. However, differentiation comes along with significant increases of the number of linear operations, [Fritzen & Kunc, 2018c, Sec. 5.3]. Moreover, the involved derivatives of the matrix exponential and of the matrix logarithm are highly non-trivial to implement. A different approach is pursued here for these reasons.

Instead of interpolating the energy density function, the second Piola-Kirchhoff stress, $\underline{S} \leftrightarrow \underline{S} \in \mathbb{R}^6$, and its tangent modulus, $\underline{C} \leftrightarrow \underline{C} \in \mathbb{R}^{21}$ are directly interpolated over the space of Hencky strains, $\underline{E} \leftrightarrow \underline{E} \in \mathbb{R}^6$. It is important to recall that the basis (8.1) is employed for the stress and the stiffness, while the Hencky strain is formulated with respect to the basis (8.2). The CI ansatz is modified into

$$\begin{bmatrix} \underline{S}(\underline{E}) \\ \underline{C}(\underline{E}) \end{bmatrix} \approx \tilde{f}(t, \underline{N}) = \left[\underline{p}^{(1)}(t), \dots, \underline{p}^{(27)}(t) \right]^T \underline{K}^{-1} \underline{k}(\underline{N}) \in \mathbb{R}^{6+21}, \quad (8.35)$$

meaning that all 27 components (6 for the stress and 21 for the stiffness) are interpolated separately but simultaneously: the respective functional data is independently interpolated by the $N_{\text{dir}}^{\text{supp}}$ distinct components of the vector $\underline{p}^{(i)}(t)$ ($i = 1, \dots, 27$). The magnitude, t , and the direction, \underline{N} , of the query Hencky strain, \underline{E} , are computed as defined in (8.5).

By interpolating the components of the stress and of the stiffness separately, their consistency is generally lost. Thus, a possibly negative effect on the convergence behavior of Newton-Raphson is to be expected and will be investigated in the numeric section of this paper. One way to principally attenuate this drawback could be to incorporate FFT-based homogenization schemes, [Spahn et al., 2014, Leuschner & Fritzen, 2018], on the lower scales as such schemes are generally less sensitive to the tangent modulus. Alternatively, Automatic Differentiation (AD), [Rall, 1981], of the energy density function W_C could possibly be employed for the evaluation of derivatives, similar to previous works, [Fritzen & Kunc, 2018c].

8.3.4 Concentric Sampling of material laws

The CI method crucially depends on the Concentric Sampling (CS) method. The CS algorithm employed in this work is a slight modification of the original, [Kunc & Fritzen, 2019b, Algorithm 1], and is essentially described by the following construction of Hencky strains, cf. (8.4), (8.5):

$$\mathbf{E}^{(m,n)} \leftrightarrow \underline{\underline{E}}^{(m,n)} = t_m \sum_{i=1}^6 N_i^{(n)} \underline{\underline{Y}}^{(i)} \quad (m = 1, \dots, N_{\text{mag}}, n = 1, \dots, N_{\text{dir}}) \quad (8.36)$$

Here, a Hencky strain basis $\underline{\underline{Y}}^{(i)} \in \mathcal{Y}_{J^*}$ ($i = 1, \dots, 6$) with fixed scaling factor J^* is considered. Further, a set of asymptotically uniformly distributed points on the five-dimensional unit hypersphere, $\underline{\underline{N}}^{(n)} \in \mathbb{S}^5 \subset \mathbb{R}^6 \sim \text{span}(\mathcal{Y}(J^*))$, $n = 1, \dots, N_{\text{dir}}$, is assumed to be provided. Such point sets may be constructed, e.g., by means of the `MinimumEnergyPoints` open source Matlab/Octave code provided by the authors, [Fritzen & Kunc, 2018b]. These hyperspherical points are interpreted as directions within the Hencky strain space along which samples are placed. The distribution of the samples along each direction is given by the set of sampling magnitudes, $0 < t_1, \dots, t_{N_{\text{mag}}} = 1$. A visualization in two dimensions is provided in Figure 8.2, where $N_{\text{dir}} = 5$ and $N_{\text{mag}} = 3$.

When the Hencky strains obtained through CS are utilized as support points for the CI method, the nomenclature $N_{\text{dir}} = N_{\text{dir}}^{\text{supp}}$ and $N_{\text{mag}} = N_{\text{mag}}^{\text{supp}}$ is adopted. On the other hand, if the CS points are evaluation points of the CI scheme, this is denoted by $N_{\text{dir}} = N_{\text{dir}}^{\text{eval}}$ and $N_{\text{mag}} = N_{\text{mag}}^{\text{eval}}$. Accordingly, when the resulting stretch tensors, $\exp(\mathbf{E}^{(m,n)})$, are applied as boundary conditions to the boundary value problems (8.12)–(8.15), this is again denoted by $N_{\text{dir}} = N_{\text{dir}}^{\text{eval}}$ and $N_{\text{mag}} = N_{\text{mag}}^{\text{eval}}$, regardless of the employed solution method, FE or the RB.

By means of this sampling approach, *representative* and *unbiased* sampling of the Hencky strain space and, thus, after exponentiation, of the *physically meaningful* part of the space of stretch tensors is realized:

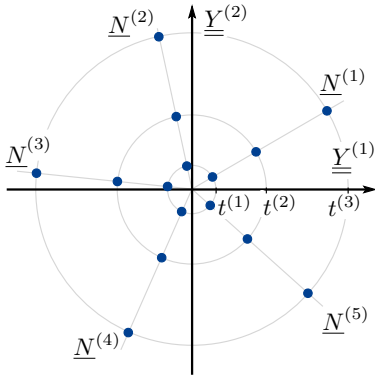


Figure 8.2: $\underline{N}^{(1)}$ Sketch of Concentric Sampling of the Hencky strain space. For visualization purposes, only two of the six coordinate axes of $\mathcal{Y}(J^*)$ are considered. N_{mag} samples are placed along N_{dir} approximately uniformly distributed directions. The former characterize the load magnitudes, the latter correspond to the type of the load.

- *Representative* means that the sampled space is covered in a manner that is dense with respect to a certain metric, given a limited number of discrete sampling points. In the case of material laws, the metric is the angular distance between different load directions, $\arccos(\underline{N}^{(i)} \cdot \underline{N}^{(j)})$. Sampling the directions in an approximately uniform manner ensures that the resolution of the anisotropy of the present material law can be adjusted.
- *Unbiased* means that the resolution by which different regions of the domain are covered is either constant or deliberately adapted. The radial density of the samples in the CS method may be adjusted according to expected features of the present material law. For instance, if a significant change of the material response is expected at certain load magnitude, e.g. a transition from linear to non-linear behavior, then a neighborhood of this critical point may be resolved more thoroughly.
- *Physically meaningful*, in the context of sampling of kinematic quantities, such as the stretch tensor $\underline{U} = \exp(\underline{E})$, means that the sampling domain is confined in such a way that all contained points correspond to realistic deformations that might actually occur in real-world problems. This is one of the main advantages of the CS method in contrast to classical approaches that directly sample the space of stretch tensors, right Cauchy-Green tensors, or deformation gradients with uniform Cartesian grids, [Yvonnet et al., 2013, Brands et al., 2019]. Such grid-based sampling methods are unable to keep the determinant $J = \det(\underline{U})$ within realistic bounds (e.g. $1 \pm \varepsilon$ in the quasi-incompressible case) while also exploring isochoric regions of the respective kinematic state spaces to their whole physically meaningful extent, e.g. more than 100% strain for rubber-like materials.

8.3.5 Many-scale homogenization: overall algorithm

The proposed method is a many-scale solution to the multi-scale problem: it is emphasized that not just more than one scale but actually numerous scales are transitioned. As indicated by Figure 8.1, the lowest M scales of the general multi-scale problem, cf.

(8.12)–(8.15), are homogenized in a recursive manner. Then, the resulting numerical surrogate of the homogenized material law can be applied on the engineering scale. The complete procedure is formally stated in Algorithm 8.1. A thorough step-by-step description is given in the following. We omit the scale index N on top of some quantities for the sake of enhanced readability.

Algorithm 8.1: Many-scale homogenization for finite strain hyperelasticity

Input : number M of scales to be homogenized;

set of RVE's $\{\Omega_0^{(N)}\}_{N=0,\dots,M-1}$;

scalar parameters for the sub-algorithms 8.2–8.5

```

// loop Scales
1 for  $N = 0, \dots, M - 1$  do
  // generate snapshots
2  call Algorithm 8.2:  $\mathbf{FE}_N(\Omega_0, N_{\text{dir}}^{\text{eval}}, N_{\text{mag}}^{\text{eval}}, J^*, T_{\text{max}}) \rightarrow \mathcal{S}$ 
  // determine RB
3  call Algorithm 8.3:  $\mathbf{POD}_N(\mathcal{S}, N_{\text{RB}}) \rightarrow \mathcal{B}$ 
  // generate CI support data
4  call Algorithm 8.4:  $\mathbf{RB}_N(\mathcal{B}, N_{\text{dir}}^{\text{eval}}, N_{\text{mag}}^{\text{eval}}, J^*, T_{\text{max}}) \rightarrow \mathcal{D}$ 
  // homogenized material law
5  setup  $\mathbf{CI}_N(\mathcal{D}, N_{\text{dir}}^{\text{supp}}, N_{\text{mag}}^{\text{supp}}, \gamma)$ 
6  if  $N < M - 1$  then
  | // use Algorithm 8.5
  | set  $\mathbf{CI}_N$  as next scale material law in (one phase of) the RVE  $\Omega_0^{(N+1)}$ 
8  else
  | // use Algorithm 8.5
  | set  $\mathbf{CI}_N$  as material law in (one phase of) the engineering Scale  $M$ 
9

```

For the sake of brevity, the scaling parameter J^* is fixed throughout the whole algorithm, i.e. across all scales, which conforms with the numerical experiments of the next section. If the query Hencky strain was ever out of the sampling range defined by T_{max} and J^* , the CI routine will still return a result by means of extrapolation.

Finite Element Simulations

On Scale N , the FE method is invoked by a call to Algorithm 8.2 which is denoted \mathbf{FE}_N . Kinematic periodic fluctuation boundary conditions, defined in terms of stretch tensors corresponding to Hencky strains (8.36) on Scale $N + 1$, are prescribed to the RVE $\Omega_0^{(N)}$. These boundary conditions are chosen by means of the Concentric Sampling method of Section 8.3.4. In view of limited computational resources the parameters

$N_{\text{dir}}^{\text{eval}}$ and $N_{\text{mag}}^{\text{eval}}$ are chosen in order to balance data availability (and thereby accuracy) and computational cost. A discussion of their physical interpretation is included in Section 8.3.4. Moreover, the choice of these parameters depends on the complexity of FE simulations, which is strongly influenced by the mesh of $\Omega_0^{(N)}$, the local material laws, the stability of the element formulation, and the efficiency of the FE implementation. Empirically speaking, $N_{\text{dir}}^{\text{eval}} = 100$ and $N_{\text{mag}}^{\text{eval}} = 10$ are practical starting points for the present study.

Algorithm 8.2: Finite Element sub-routine **FE_N**

Input : RVE $\Omega_0^{(N)}$;
 number of load directions $N_{\text{dir}}^{\text{eval}}$;
 number of load magnitudes $N_{\text{mag}}^{\text{eval}}$;
 scaling factor J^* of Hencky strain basis;
 maximum load magnitude T_{max} ;
 load directions $\{\underline{N}^{(n)}\}_{n=1, \dots, N_{\text{dir}}^{\text{eval}}} \subset \mathbb{S}^6$;
 load magnitudes $\{t_m\}_{m=1, \dots, N_{\text{mag}}^{\text{eval}}} \subset [0, T_{\text{max}}]$

Output: deformation gradient fluctuation snapshots \mathcal{S} with $|\mathcal{S}| = N_{\text{snap}}$
 ($N_{\text{snap}} \leq N_{\text{dir}}^{\text{eval}} N_{\text{mag}}^{\text{eval}}$)

- 1 initialize snapshot list $\mathcal{S} = \{\}$
 // loop directions
- 2 **for** $n = 1, \dots, N_{\text{dir}}^{\text{eval}}$ **do**
 // loop magnitudes
- 3 **for** $m = 1, \dots, N_{\text{mag}}^{\text{eval}}$ **do**
 // set boundary condition
- 4 $\underline{U}^{(N+1)} = \exp\left(t_m \sum_{i=1}^6 N_i^{(n)} \underline{Y}^{(i)}\right)$ ($\underline{Y}^{(i)} \in \mathcal{Y}_{J^*}$)
- 5 try to solve boundary value problem (8.13) on $\Omega_0^{(N)}$ with boundary
 condition $\underline{U}^{(N+1)}$ using FEM
- 6 **if convergence then**
 // set fluctuation snapshot
- 7 subtract homogeneous deformation from local field,
 $\underline{\tilde{F}}^{(N)}(\underline{X}) = \underline{F}^{(N)}(\underline{X}) - \underline{U}^{(N+1)}$
- 8 include fluctuation snapshot in list, $\mathcal{S} \leftarrow \mathcal{S} \cup \left\{ \underline{\tilde{F}}^{(N)}(\underline{X}) \right\}$

It is important to note the order in which the boundary conditions are applied: the FE method generally converges quicker if only the load magnitude, t , is incremented between load steps and the kind of the load, i.e. its direction \underline{N} with respect to the

Hencky strain representation, is kept constant. Furthermore, very simple parallelization of the overall workload is possible by running multiple instances of the same program with different sets of loading directions. For instance, if the number of available compute threads is 10 and $N_{\text{dir}}^{\text{eval}} = 100$, one might run 10 instances in parallel, each applying 10 different load directions successively.

The Newton iterations of the FEM are stopped when the maximum nodal residual is below $10^{-5} \text{MPa} \cdot L^{(N)2}$ (where $L^{(N)}$ is the side length of the cubic RVE) or if its ratio to the maximum reaction force is below 10^{-3} . In the present study, the same stopping FE criterion is employed on all scales. This is in contrast to other works, [Özdemir et al., 2008], that eased the stopping criterion on larger scales.

Usual convergence issues are expected at some evaluation directions above few critical load magnitudes, [Yvonnet et al., 2013, Section 3.2.5]. In practice, certain states of (large) deformations simply cannot be reached by a given FE model. This effect depends on the maximum magnitude T_{max} , the properties of the mesh, the element stabilization method, the material laws employed at the current Scale N , the scaling parameter J^* , among other factors. Therefore, the number of available solutions is usually lower than the number of boundary conditions provided by CS. In order to conservatively assess the practicality of the proposed method, no FE stabilization method is employed in the present study.

The converged resulting fields of deformation gradients are post-processed by pointwise subtraction of the prescribed homogeneous deformation. Thus, the set \mathcal{S} of deformation gradient fluctuation fields is computed, denoted snapshots, and returned to the main routine.

Proper Orthogonal Decomposition

The set of deformation gradient fluctuation snapshots, \mathcal{S} , is processed by the POD Algorithm 8.3, which is dubbed **POD** $_N$. Snapshot POD methods of this kind have been successfully applied before, [Sirovich, 1987].

A popular way of deciding when to truncate the emanating set of basis functions, i.e. how large to choose the number $N_{\text{RB}} = |\mathcal{B}|$ of basis functions, is to employ the Eckart–Young–Mirsky theorem. By this, the resulting projection error can be controlled based on the eigenvalues of the correlation matrix. However, the authors have experienced a deteriorating effect in terms of the stress approximation error with N_{RB} values greater than a certain optimum, [Kunc & Fritzen, 2019b]. This effect is interpreted as an artifact stemming from spurious displacement patterns (e.g. strongly oscillating fields) mainly located at the interface of high contrast materials. The severity of this effect has yet to be investigated systematically.

For the moment, we settle with the empirically determined, [Kunc & Fritzen, 2019b], number of $N_{\text{RB}} = 30$ basis elements, which appears to give reasonably good

Algorithm 8.3: Proper Orthogonal Decomposition sub-routine **POD_{-N}**

Input : snapshot list \mathcal{S} with $|\mathcal{S}| = N_{\text{snap}}$;
 Reduced Basis size N_{RB} with $1 \leq N_{\text{RB}} \leq N_{\text{snap}}$;
 RVE domain Ω_0 (implicit)

Output: $\mathcal{B} = \{ \overset{(N)}{\mathbf{B}}^{(i)}(\overset{(N)}{\mathbf{X}}) \}_{i=1, \dots, N_{\text{RB}}}$
 // correlation matrix $\underline{\Gamma}$

- 1 **for** $i = 1, \dots, N_{\text{snap}}$ **do**
- 2 **for** $j = i, \dots, N_{\text{snap}}$ **do**
- 3 $\Gamma_{ij} = \langle \overset{(N)}{\mathbf{F}}^{(i)}, \overset{(N)}{\mathbf{F}}^{(j)} \rangle = \Gamma_{ji}$

// eigendecomposition

- 4 solve eigenvalue problem of $\underline{\Gamma} \in \mathbb{R}^{N_{\text{snap}} \times N_{\text{snap}}}$
 → eigenvalues $\kappa_1 > \dots > \kappa_{N_{\text{snap}}} > 0$,
 → eigenvectors $\underline{V}^{(1)}, \dots, \underline{V}^{(N_{\text{snap}})} \in \mathbb{R}^{N_{\text{snap}}}$ corresponding to $\kappa_1, \dots, \kappa_{N_{\text{snap}}}$

// assembly and truncation

- for** $i = 1, \dots, N_{\text{RB}}$ **do**
- 5 $\overset{(N)}{\mathbf{B}}^{(i)}(\overset{(N)}{\mathbf{X}}) = \sum_{j=1}^{N_{\text{snap}}} \frac{1}{\sqrt{\kappa_i}} V_j^{(i)} \overset{(N)}{\mathbf{F}}^{(j)}(\overset{(N)}{\mathbf{X}})$ // normalized Reduced Basis

approximation accuracy of both the energy density and the stresses. Also, the resulting speed-up compared with the FE method remains significant. The truncated basis \mathcal{B} is returned to the main algorithm.

Reduced Basis Simulations

Just as the FEM, the RB model is evaluated at kinematic boundary conditions corresponding to Hencky strains that result from CS. But in contrast to the FE model, the numbers $N_{\text{dir}}^{\text{eval}}$ and $N_{\text{mag}}^{\text{eval}}$ are chosen significantly larger for the call of the **RB_{-N}** Algorithm 8.4.

As a rule of thumb, the number of boundary conditions, $N_{\text{dir}}^{\text{eval}} N_{\text{mag}}^{\text{eval}}$, applied to **RB_{-N}** can be greater than that of **FE_{-N}** by the same factor as the speed-up of the RB with respect to the FEM. This way, the computational work load associated with these two stages of the overall homogenization process is comparable. Positive experiences were obtained by increasing $N_{\text{dir}}^{\text{eval}}$ by a factor of 5 and $N_{\text{mag}}^{\text{eval}}$ by a factor of 2, corresponding to the speed-up factor of approximately 10.

The choice of these parameters strongly depends on the size of the RB, N_{RB} . On the one hand, a smaller basis is advantageous in that a greater number of boundary conditions can be applied, i.e. a greater sampling resolution can be achieved. On the other hand, the

Algorithm 8.4: Reduced Basis sub-routine **RB_{-N}**

Input : Reduced Basis \mathcal{B} with $|N_{\text{RB}}|$;
 numbers of load directions $N_{\text{dir}}^{\text{eval}}$ and of load magnitudes $N_{\text{mag}}^{\text{eval}}$;
 scaling factor J^* of the Hencky strain basis;
 maximum load magnitude T_{max} ;
 load directions $\{\underline{N}^{(n)}\}_{n=1, \dots, N_{\text{dir}}^{\text{eval}}} \subset \mathbb{S}^6$;
 load magnitudes $\{t_m\}_{m=1, \dots, N_{\text{mag}}^{\text{eval}}} \subset [0, T_{\text{max}}]$;
 RVE $\Omega_0^{(N)}$

Output: Concentric Interpolation support data \mathcal{D}

- 1 **initialize** CI support data $\mathcal{D} = \{\}$; extrapolation counter $N_{\text{ext}}^{\text{eval}} = 0$
- // loop directions
- 2 **for** $n = 1, \dots, N_{\text{dir}}^{\text{eval}}$ **do**
- // loop magnitudes
- 3 **for** $m = 1, \dots, N_{\text{mag}}^{\text{eval}}$ **do**
- // set boundary condition
- 4 $\underline{U}^{(N+1)} = \exp\left(t_m \sum_{i=1}^6 N_i^{(n)} \underline{Y}^{(i)}\right)$ ($\underline{Y}^{(i)} \in \mathcal{Y}_{J^*}$)
- 5 try to solve minimization task (8.25) with boundary condition $\underline{U}^{(N+1)}$
 using \mathcal{B}
- 6 **if** convergence **then**
- 7 compute effective stress $\underline{S}_{\text{RB}}^{(N+1)}$ according to (8.27) and effective
 stiffness $\underline{C}_{\text{RB}}^{(N+1)}$ according to (8.28)
- 8 **else**
- 9 linearly extrapolate $\underline{S}_{\text{RB}}^{(N+1)}$ and $\underline{C}_{\text{RB}}^{(N+1)}$ in t from magnitudes t_{m-2} ,
 t_{m-1} (assume convergence at least for $m = 2$)
- 10 subtract initial stiffness (i.e. at $t = 0$) from current stiffness:
 $\underline{C}_{\text{RB}}^{(N+1)} = \underline{C}_{\text{RB}}^{(N+1)} - \underline{C}_{\text{RB}}^{\text{ini}}$
- 11 include vectorized support data in list,
 $\mathcal{D} \leftarrow \mathcal{D} \cup \left\{ \left(\underline{S}_{\text{RB}}^{(N+1)}; \underline{C}_{\text{RB}}^{(N+1)} \right) \in \mathbb{R}^{6+21} \right\}$

quality of this sampling data is increased by a larger basis. It is a complex task to choose all parameters (of the RB but also of the overall method) in an optimal way.

Importantly, the robustness of the RB method is greatly increased in contrast to the FEM. Hence, the maximum magnitude T_{max} may be chosen much larger for the RB than for the FE model. However, one should keep in mind that the reliability of the RB results is

then decreased as the RB is evaluated far outside the domain of the snapshot data. Also, the enhanced robustness is still limited and an occasional lack of convergence for (large) magnitudes of certain kinds of deformations is to be expected.

Convergence in this context is defined as the RB residual, [Kunc & Fritzen, 2019b, (28)], reaching a norm of less than $5 \cdot 10^{-7} \text{MPa} \cdot L^2$ (where L is the side length of the cubic RVE). Alternatively, convergence is assumed when the ratio of the residual to the current increment's initial residual is below 10^{-7} .

In contrast to the **FE**_{*N*}, where partly unsuccessfully applied boundary conditions simply reduced the number of snapshots, this is a more severe issue for the **RB**_{*N*}: the outputs of this stage serve as inputs to the subsequent CI scheme. The latter necessarily requires data at all support points (with the current assumption of piecewise linear interpolants with identical support positions along all directions). We choose to linearly extrapolate incomplete RB data along the respective direction. The number of such interpolated RB results is counted in $N_{\text{ext}}^{\text{eval}}$. Alternatively, any other means of data completion could be employed, e.g. my means of specific choices of the radial data functions $p_i(t)$ from (8.32).

Note that just like the FE method, the RB method is suitable for trivial parallelization by simply running multiple instances of the program at the same time, each with different sets of boundary conditions.

User-defined material function based on Concentric Interpolation

As the last step before the completion of one homogenization loop in Figure 8.1, the **CI**_{*N*} scheme is set up according to (8.35). The interpolant \tilde{f} is incorporated in the user-defined material Algorithm 8.5. Lastly, the scale counter \tilde{N} is incremented by one and Algorithm 8.1 proceeds correspondingly.

8.4 Numerical examples

8.4.1 Three-scale homogenization of fractal structures on standard workstations

Goals

The main goal of the following considerations is to provide an example of a three-scale simulation based on Algorithm 8.1, cf. Figure 8.1. This is meant as a proof of concept only and thorough investigations of important details, such as the influences of different kinds of parameters, are not within the current scope. The geometries of the RVE's on Scales 0 and 1 are identical. On Scale 2, the same geometry is again utilized. In this

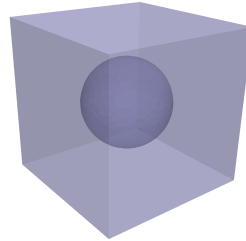
Algorithm 8.5: User-defined material function of $\mathbf{FE}_{(N+1)}$ and of $\mathbf{RB}_{(N+1)}$, wrapping \mathbf{CI}_N

Input : right Cauchy-Green tensor $\mathbf{C}^{(N+1)}$;
initial stiffness $\mathbf{C}_{\text{RB}}^{\text{ini}}{}^{(N+1)}$ (constant);
scale factor J^* of the basis $\mathcal{Y}(J^*)$ (constant)

Output: stress $\underline{\mathbf{S}}_{\text{CI}}^{(N+1)}$ and stiffness $\underline{\mathbf{C}}_{\text{CI}}^{(N+1)} = \underline{\mathbf{C}}_{\text{CI}}^{(N+1)} + \underline{\mathbf{C}}_{\text{RB}}^{\text{ini}}{}^{(N+1)}$
// Hencky strain

- 1 compute $\mathbf{E}^{(N+1)} = \frac{1}{2} \log(\mathbf{C}^{(N+1)})$
// direction and magnitude w.r.t. basis $\mathcal{Y}(J^*)$
- 2 transform $\mathbf{E}^{(N+1)}$ by (8.4) and (8.5) $\rightarrow \underline{\mathbf{N}}, t$
// stress and shifted stiffness (cf. Algorithm 8.4)
- 3 evaluate the \mathbf{CI}_N interpolant $\tilde{\mathbf{f}}$ at $(\underline{\mathbf{N}}, t) \rightarrow \underline{\mathbf{S}}_{\text{CI}}^{(N+1)}, \underline{\mathbf{C}}_{\text{CI}}^{(N+1)}$

Figure 8.3: Geometry of the cubic structure with a centered spherical void. The volume fraction of the void is 10%.



sense, the geometric setup (cf. Table 8.1, *top*) is of self-similar fractal kind. Roughly comparable scenarios were recently considered in other works, [Yang et al., 2020].

If the geometry on Scale 2 is again considered an RVE and no body forces are applied, the resulting local stress field may be volume averaged and then interpreted as a fourth-scale homogenized quantity. This subsection ends with selected evaluations of such stresses on a hypothetical Scale 3.

Geometry and Scale 0 model

The geometry considered on all three scales is a cube with a centered spherical void, cf. Figure 8.3. The mesh contains 4399 second-order TET10 elements with a total of 7531 nodes resulting in almost 22600 DOF. The mesh, i.e. the solid phase, occupies a volume of 0.90 within the unit cube.

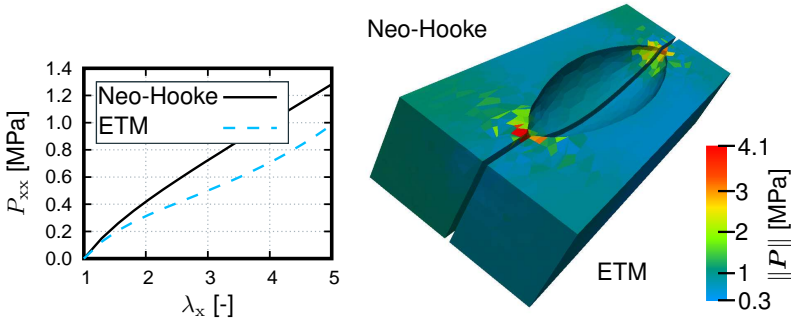


Figure 8.4: Comparison of the Neo-Hookean and the Extended Tube Model under isochoric uniaxial extension. As for the ETM, the Neo-Hookean material parameter is chosen from a comparative study, [Marckmann & Verron, 2006, Table III]. *Left:* evolution of the principal stress over the principal stretch. *Right:* composite image consisting of two cut-out quarters of the deformed voided microstructure for comparison of the influences of the different material laws. Coloring indicates the magnitude of the first Piola-Kirchhoff stress per integration point.

For the material law on Scale 0, the Extended Tube Model (ETM), [Kaliske & Heinrich, 1999, Eq. (22)], together with a mixed quadratic-logarithmic volumetric model, [Doll & Schweizerhof, 1999], is utilized:

$$W_C(\mathbf{C}) = W_{\text{ETM}}(J^{-2/3}\mathbf{C}) + \frac{K}{4}((J-1)^2 + \log(J)^2) \quad (8.37)$$

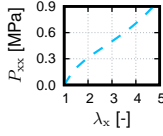
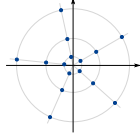
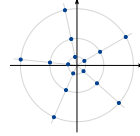
As found by a comparative study, [Marckmann & Verron, 2006, Table III], the ETM is a realistic model of rubber-like materials for the parameters $G_c = 0.202$ MPa, $G_e = 0.153$ MPa, $\beta = 0.178$ MPa, $\delta = 0.0856$ MPa, which are therefore employed in the present work. The bulk modulus-like parameter K is empirically chosen as $K = 10$ MPa in order to set a high ratio to the other parameters and, thereby, to approach quasi-incompressible behavior. A nonlinearity that is more pronounced with the ETM than with the widely used Neo-Hookean model is visible in Figure 8.4.

Three-scale setup

According to Algorithm 8.1 and Figure 8.1, the homogenization of Scale 0 and of Scale 1 is performed in a staged manner. All parameters employed throughout this homogenization process are listed in Table 8.1.

The kernel width parameter γ is determined via optimization of the stresses $\bar{\mathbf{S}}_{\text{CI}}$ with respect to validation stresses $\bar{\mathbf{S}}_{\text{RB}}$ on a large set of validation Hencky strains: $N_{\text{dir}}^{\text{eval}} = 1000$, $N_{\text{mag}}^{\text{eval}} = 100$, $J^* = 1.2$, $T_{\text{max}} = 1.0$. For comparison, the optimization of

Table 8.1: Three-scale setup: geometry and parameters

	Scale 0	Scale 1	Scale 2
field variables	E, S, C, \dots	$\bar{E}, \bar{S}, \bar{C}, \dots$	$\bar{\bar{E}}, \bar{\bar{S}}, \bar{\bar{C}}, \dots$
boundary condition	$\bar{U} = \exp(\bar{E})$	$\bar{U} = \exp(\bar{E})$	$\bar{\bar{U}} = \exp(\bar{\bar{E}})$ or arbitrary
material law	ETM 	CI_0 	CI_1 
FE parameters	FE_0: $N_{dir}^{eval} = 100$ $N_{mag}^{eval} = 10$ $J^* = 1.2$ $T_{max} = 1.0$	FE_1: $N_{dir}^{eval} = 100$ $N_{mag}^{eval} = 10$ $J^* = 1.2$ $T_{max} = 0.3$	
POD parameters	POD_0: $N_{snap} = 956$ $N_{RB} = 30$	POD_1: $N_{snap} = 957$ $N_{RB} = 30$	
RB parameters	RB_0: $N_{dir}^{eval} = 500$ $N_{mag}^{eval} = 20$ $J^* = 1.2$ $T_{max} = 1.0$	RB_1: $N_{dir}^{eval} = 500$ $N_{mag}^{eval} = 20$ $J^* = 1.2$ $T_{max} = 1.0$	
RB results extrapolation	$N_{ext}^{eval} = 94$	$N_{ext}^{eval} = 1438$	
CI parameters	CI_0: $N_{dir}^{supp} = 500$ $N_{mag}^{supp} = 20$ $\gamma = 3.12$	CI_1: $N_{dir}^{supp} = 500$ $N_{mag}^{supp} = 20$ $\gamma = 3.12$	
CI action	$\bar{E} \mapsto \bar{S}_{CI}, \bar{C}_{CI}$	$\bar{E} \mapsto \bar{S}_{CI}, \bar{C}_{CI}$	

the kernel parameter γ yields an optimal value of $\gamma = 3.56$ when the discrepancy between the stiffnesses \bar{C}_{CI} and \bar{C}_{RB} is minimized. As the stress approximation directly influences the balance equation, the result obtained from the stress optimization is employed. The same value of γ is used for the interpolation scheme **CI_1** without further optimization.

Another noteworthy parameter is T_{max} , which is 1.0 throughout Scale 0. However, it is reduced for **FE_1** to $T_{max} = 0.3$ due to severe convergence issues, cf. Section 9. The enhanced robustness of **RB_1**, cf. 5, allows for the choice of $T_{max} = 1.0$ again. At

this stage, $14\% \approx N_{\text{ext}}^{\text{eval}} / (N_{\text{dir}}^{\text{eval}} N_{\text{mag}}^{\text{eval}})$ of the applied boundary condition lacked RB convergence.

Hardware setup and runtimes

In order to accelerate the setup phase (or offline phase), the FE, POD and RB computations were performed on a standard workstation with an Intel(R) Xeon(R) E5-2643 v3 CPU (12 threads) and 256 Gb of RAM. The memory intensive POD never came close to requiring all memory.

An exact quantitative comparison of the various runtimes is not within the scope of the current work. Depending on the method and on the stage, different kinds of overhead data were created and stored during the multi-scale homogenization process, e.g. visualization data, statistics, and debugging information. Also, important meta-parameters were chosen as sophisticated guesses, e.g. stopping criteria, maximum number of Newton iterations, behavior in case of convergence failure. Furthermore, the FE and the RB simulations were conducted with multiple instances of the same program running in parallel, each with a different set of boundary conditions. Most importantly, the computations for the current case study were not always conducted on a dedicated workstation. For these reasons, quantitatively exactly reproducible runtime values are set aside for the moment.

The overall wall time to conduct the whole three-scale homogenization procedure on the workstation, i.e. for **FE_0**, **POD_0**, **RB_0**, **FE_1**, **POD_1** and **RB_1** together, was approximately 7 hours. The simulations on the third scale by means of **FE_2** were conducted on a laptop computer and took approximately 2-3 minutes per load path.

Results

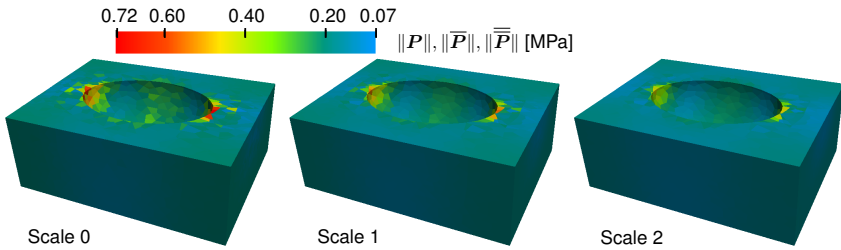
As the overall porosity of the structure increases with each homogenized scale, a **scale softening effect** is observed in the initial stiffnesses, i.e. in the stiffness tensors at zero magnitude, $t = 0$, cf. Table 8.2.

Next, the geometrically identical structures on Scales 0, 1 and 2 are subjected to the same boundary condition, $\bar{U} = \bar{U} = \bar{U}$, with the principal stretches $\lambda_x = \lambda_z = 0.9035$, $\lambda_y = 1.225$. The resulting deformations and stress distributions are depicted in Figure 8.5. The effect of the scale softening is clearly visible which is in accordance with Table 8.2.

We now address the question of **how Reduced Bases of two scales compare to each other**. Although they are returned from identical routines **POD_0** and **POD_1**, their

Table 8.2: Initial stiffness matrices ($t = 0$) of the ETM and as output of **RB_0** and **RB_1**.

$\underline{\underline{C}}_{\text{ETM}}^{\text{ini}} [\text{MPa}] =$	10.5	9.77	9.77	0	0	0
		10.5	9.77	0	0	0
			10.5	0	0	0
				0.704	0	0
		sym			0.704	0
					0.704	
$\overline{\underline{\underline{C}}}_{\text{RB}}^{\text{ini}} [\text{MPa}] =$	3.77	3.09	3.09	$1.22 \cdot 10^{-3}$	$7.74 \cdot 10^{-4}$	$3.74 \cdot 10^{-4}$
		3.77	3.09	$8.58 \cdot 10^{-4}$	$1.10 \cdot 10^{-3}$	$1.18 \cdot 10^{-3}$
			3.77	$6.38 \cdot 10^{-4}$	$9.37 \cdot 10^{-4}$	$8.44 \cdot 10^{-4}$
				0.654	$-4.02 \cdot 10^{-5}$	$2.15 \cdot 10^{-4}$
		sym			0.653	$3.21 \cdot 10^{-4}$
					0.653	
$\overline{\underline{\underline{C}}}_{\text{RB}}^{\text{ini}} [\text{MPa}] =$	2.34	1.69	1.69	$8.89 \cdot 10^{-4}$	$5.94 \cdot 10^{-4}$	$5.48 \cdot 10^{-5}$
		2.34	1.69	$5.85 \cdot 10^{-4}$	$8.68 \cdot 10^{-4}$	$7.74 \cdot 10^{-4}$
			2.34	$3.97 \cdot 10^{-4}$	$7.43 \cdot 10^{-4}$	$4.68 \cdot 10^{-4}$
				0.602	$-5.57 \cdot 10^{-5}$	$9.85 \cdot 10^{-5}$
		sym			0.600	$2.62 \cdot 10^{-4}$
					0.601	

**Figure 8.5:** Half sections of the porous structures on three scales, subject to isochoric uniaxial extension with the principal stretches $\lambda_x = \lambda_z = 0.9035$, $\lambda_y = 1.225$. Coloring indicates the stress magnitude per integration point.

underlying snapshot data differs in terms of the material law and, most significantly, the parameter T_{max} . Since the eigenvectors of the snapshot correlation matrix are ordered descendingly with respect to the corresponding eigenvalues, the returned $N_{\text{RB}} = 30$ RB

elements represent the most significant information contained in the snapshots. Out of this set, the first nine functions are depicted in Figure 8.6 for both scales.

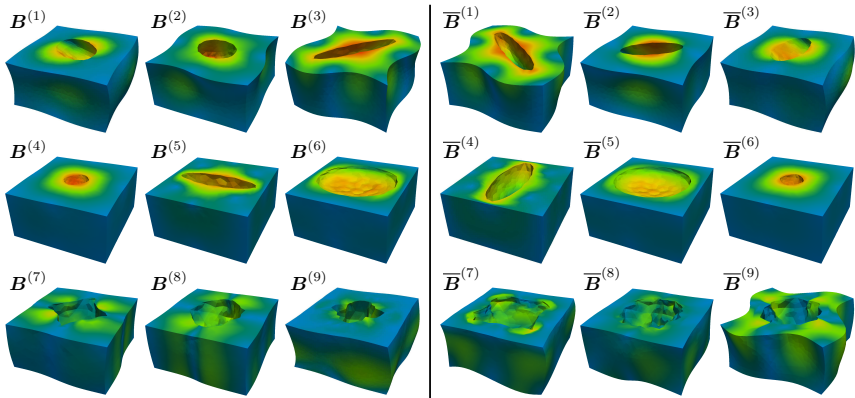


Figure 8.6: Lower halves of the deformation patterns corresponding to the first nine of the 30 Reduced Basis functions on Scale 0 (left) and on Scale 1 (right). Coloring indicates displacement magnitude. See also Figure 8.7.

A systematic comparison of the two bases is possible when considering the **absolute correlation of the Reduced Bases**, i.e. the mutual projections of the RB elements onto each other, $\left| \langle B^{(i)} \cdot \bar{B}^{(j)} \rangle \right|$. This is well-defined, cf. (8.10), when considering that the geometries of the RVE are identical, $\bar{\Omega}_0 = \bar{\bar{\Omega}}_0$. The RB correlations are visualized in Figure 8.7

These correlation values reveal some similarities that one might expect from visual inspection of Figure 8.6, e.g. $B^{(4)} \leftrightarrow \bar{B}^{(6)}$ and $B^{(6)} \leftrightarrow \bar{B}^{(5)}$. Moreover, some less

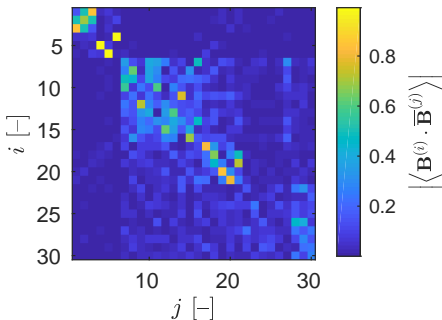


Figure 8.7: Correlations of the Reduced Bases $B^{(1)}, \dots, B^{(30)}$ and $\bar{B}^{(1)}, \dots, \bar{B}^{(30)}$ of Scales 0 and 1, respectively. See also Figure 8.6.

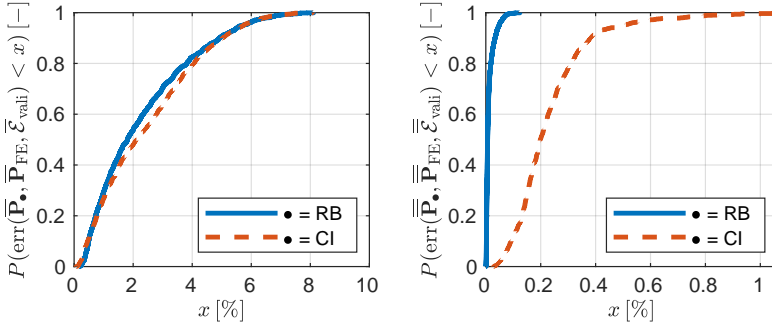


Figure 8.8: Empirical distribution functions of the relative error of the first Piola-Kirchhoff stress on Scale 1 (*left*) and on Scale 2 (*right*) as a homogenized results of the \mathbf{FE}_N , \mathbf{RB}_N , and \mathbf{CI}_N methods with $N = 0, 1$, respectively.

obvious relationship are found, e.g. $\mathbf{B}^{(5)} \leftrightarrow \overline{\mathbf{B}}^{(4)}$. One should bear in mind that the visualization in Figure 8.6 might appear differently if another clipping plane was chosen.

Next, the **accuracy** of \mathbf{RB}_N and \mathbf{CI}_N is compared against \mathbf{FE}_N for $N = 0$ and $N = 1$. Recall the definition of the homogenized stress from (8.20). For $N = 0$, a set of CS boundary conditions $\mathcal{E}_{\text{vali}}$ are chosen with the parameters $T_{\text{max}} = 0.3$, $J^* = 1.2$, $N_{\text{dir}}^{\text{eval}} = 150$ and $N_{\text{mag}}^{\text{eval}} = 10$, i.e. $|\overline{\mathcal{E}}| = 1500$. The same CS parameters are chosen for $\overline{\mathcal{E}}_{\text{vali}}$ and $N = 1$. However, this set is reduced to contain only those boundary conditions for which \mathbf{FE}_1 successfully converges, wherefore effectively $|\overline{\mathcal{E}}_{\text{vali}}| = 1437$. The set of relative errors

$$\text{err}(\overline{\mathbf{P}}_{\bullet}, \overline{\mathbf{P}}_{\text{FE}}, \overline{\mathcal{E}}_{\text{vali}}) = \left\{ \left\| \frac{\overline{\mathbf{P}}_{\bullet}(\overline{\mathbf{E}}) - \overline{\mathbf{P}}_{\text{FE}}(\overline{\mathbf{E}})}{\overline{\mathbf{P}}_{\text{FE}}(\overline{\mathbf{E}})} \right\| \mid \overline{\mathbf{E}} \in \overline{\mathcal{E}}_{\text{vali}} \right\} \quad (8.38)$$

($N = 1, 2$; $\bullet = \text{RB}, \text{CI}$) is considered. The corresponding empirical distribution functions are shown in Figure 8.8. One should keep in mind that \mathbf{FE}_1 (and \mathbf{RB}_1) utilizes \mathbf{CI}_0 as a material law. Hence, the comparison in Figure 8.8 *right* is with respect to an approximate reference.

In any case, it is observed that the error in the stress approximation is within acceptable bounds ($\leq 8\%$ on Scale 1, $\leq 1\%$ on Scale 2). This has to be seen in the context of actual engineering applications where many other additional uncertainties come into play, such as the material models on Scale 0 and the geometry of the considered RVE's.

Eventually, a couple of **four-scale simulations** are conducted. Isochoric uniaxial extension boundary conditions of the kind

$$\overline{\mathbf{U}} = \overline{\overline{\mathbf{U}}} = \overline{\overline{\overline{\mathbf{U}}}} = \exp(t \mathbf{Y}^{(1)}) \quad (\mathbf{Y}^{(1)} \leftrightarrow \underline{\underline{\mathbf{Y}}}^{(1)}, \text{ cf. (8.2)}), \quad (8.39)$$

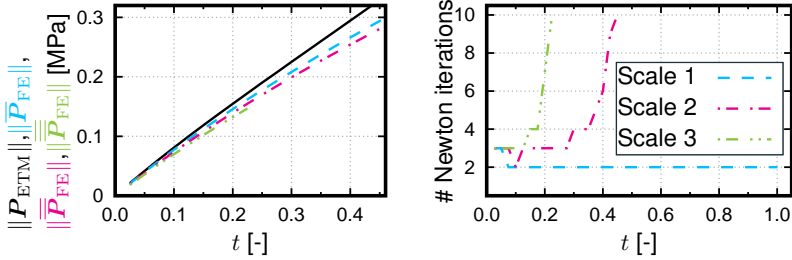


Figure 8.9: Evolution of stress magnitudes on Scales 0–3 (*left*) and of the number of Newton iterations of the corresponding FEM (*right*) over the magnitude t of the uniaxial isochoric extension boundary condition.

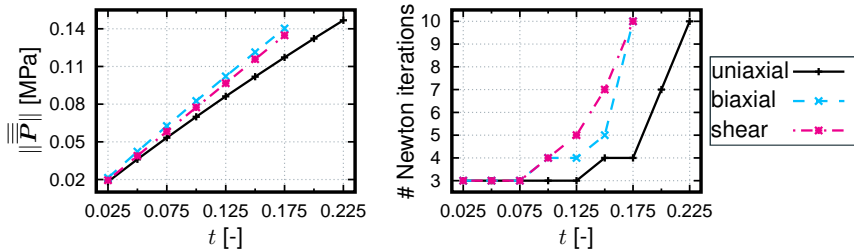


Figure 8.10: Evolution of the effective stress (*left*) and of the number of Newton iterations of the FEM (*right*) on Scale 3 over the boundary condition magnitude.

with $t \in [0, 1]$, are applied to the three Scales 0, 1 and 2. The respective FE methods are invoked and the resulting stress fields $\mathbf{P}(\mathbf{X})$, $\overline{\mathbf{P}}(\overline{\mathbf{X}})$ and $\overline{\overline{\mathbf{P}}}(\overline{\overline{\mathbf{X}}})$, respectively, are computed. Subsequently, these are volume averaged, leading to $\overline{\mathbf{P}}$, $\overline{\overline{\mathbf{P}}}$ and $\overline{\overline{\overline{\mathbf{P}}}}$ which are quantities on Scales 1, 2 and 3, respectively. The magnitudes of these effective stresses are depicted in Figure 8.9 *left*. The amount of corresponding Newton iterations are graphed in Figure 8.9 *right*.

By the number of iterations and by the maximum value of the magnitude t for which FE convergence is obtained, one can clearly see a trend of deteriorating FE convergence as more scales are homogenized, cf. Section 9. Also, the scale softening effect is again noticeable.

More kinds of fourth scale boundary conditions $\overline{\overline{\overline{\mathbf{U}}}} = \exp(t\mathbf{Y}^{(i)})$ are applied to Scale 2. The evolution of the volume-averaged stresses for the choices $i = 1, 2, 3$ is shown in Figure 8.10 *left*, with the corresponding Newton iteration count on the *right*.

As the maximum magnitude is comparatively low, $t = 0.175$ respectively $t = 0.225$, the resulting stress curves appear almost linear. Nonetheless, the capability

of the proposed scheme to perform **four-scale simulations on laptop computers** is successfully demonstrated.

8.4.2 Homogenized stress response of stiffening microstructure via CI

Goals

The aims of this numerical example are to

- prove the suitability of the CI method for the homogenization of a highly anisotropic RVE, i.e. homogenize one scale only,
- modify the CI interpolant $\tilde{f} : \mathbb{R}^6 \rightarrow \mathbb{R}^9$ such that it interpolates the stress $\overline{\mathbf{P}}$ as a function of the Hencky strain $\overline{\mathbf{E}}$
- setup the CI interpolant directly on FE data, avoiding the offline stages POD and RB as far as possible,
- gain insights into the influence of the parameters $N_{\text{dir}}^{\text{supp}}$ and γ on the accuracy of the stress interpolation,
- study the influence of the parameter $N_{\text{dir}}^{\text{supp}}$ on the runtime of the CI scheme.

This subsection does *not* strictly employ Algorithm 8.1 but instead investigates the possible shortcut FE \rightarrow CI. As pointed out in Section 8.2.2, the FE method is not nearly as well suited as the RB for providing the tangent modulus, hence we now focus on the stress only.

Model

The RVE under consideration is depicted in Figure 8.11. The FE mesh of the non-cubic RVE consists of quadratic TET10 elements with a total of 33,923 nodes, resulting in over 10^5 total DOF. The hash-shaped inclusion phase has a volume fraction of 13.3% and is the source of significant geometric stiffening effects under certain boundary conditions, as was investigated previously, [Kunc & Fritzen, 2019b].

An isochoric Neo-Hookean material law together with a mixed logarithmic-quadratic volumetric part, [Doll & Schweizerhof, 1999], of the form

$$W_C(\mathbf{C}) = \frac{G}{2} \left(\text{trace}(J^{-2/3} \mathbf{C}) - 3 \right) + \frac{K}{4} \left((J - 1)^2 + \log(J)^2 \right) \quad (8.40)$$

is employed for both the inclusion and the surrounding matrix phase. While $G_m = 10$ MPa and $K_m = 3000$ MPa are chosen for the latter, the former is described by the constants $G_i = 100$ MPa and $K_m = 3000$ MPa, i.e. the phase contrast is 10 for

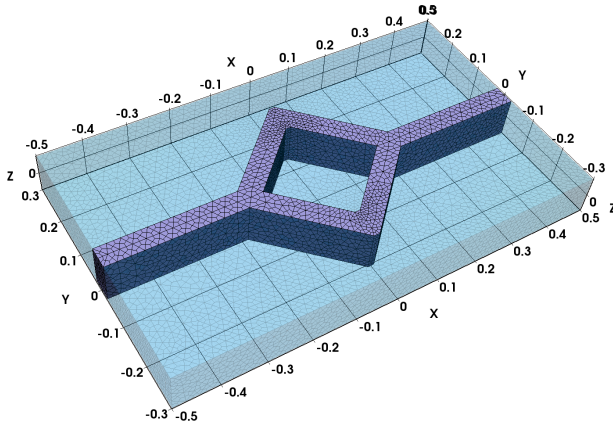


Figure 8.11: Geometry and graded FE mesh of a periodic microstructure with hash-shaped, stiff inclusion.

the shear modulus. These material parameters roughly correspond a soft polymer matrix and a stiffer polymer inclusion.

As before, it is important to emphasize that the choice of the hyperelastic material model is completely arbitrary and serves as an example only. The method is general with respect to this choice.

Design of the interpolation scheme

For this example, the domain of interest within the Hencky space is defined via $T_{\max} = 1$ and $J^* = 1.005$. Initially, it is assumed that a sufficiently accurate interpolant $\bar{\mathbf{P}}_{\text{CI}}$ of the effective stress $\bar{\mathbf{P}}$ should be attainable with $N_{\text{dir}}^{\text{supp}}$ between 50 and 1000. The supporting magnitudes are chosen at $N_{\text{mag}}^{\text{supp}} = 10$ values, wherefore the total numbers of supporting points lie between 500 and 10000. If instead, for example, regular Cartesian grids were employed, the same total numbers of points p^6 would imply as few as $p \approx 2.8$ to $p \approx 4.7$ points per axis.

The $N_{\text{mag}}^{\text{supp}} = 10$ supporting magnitudes are chosen as 0.1, ..., 0.9, 1.0. The number of supporting directions $N_{\text{dir}}^{\text{supp}}$ takes the values 50, 100, 200, 250, 500, 1000. These sets of directions are mutually independent, i.e. no intersections and especially no hierarchies are enforced.

In this example, a direct setup of the CI scheme is pursued, i.e. based on FE data in order to circumvent the offline stages POD and RB. However, due to the complexity of the microstructure, only the supporting points of the CI schemes with $N_{\text{dir}}^{\text{supp}} = 50, 100, 200$ are computed by means of the FEM.

For the larger values of $N_{\text{dir}}^{\text{supp}}$, the original Algorithm 8.1 is made use of. For the corresponding POD, the support data of the CI scheme with $N_{\text{dir}}^{\text{supp}} = 200$ is taken as snapshot data \mathcal{S} . The results are again truncated at $N_{\text{RB}} = 30$, which is based on experience with an RVE of the same geometry, [Kunc & Fritzen, 2019b]. Some of the RB elements are visualized in Figure 8.12. This RB model is employed for the creation of the supporting data of the CI schemes with $N_{\text{dir}}^{\text{supp}} = 250, 500, 1000$.

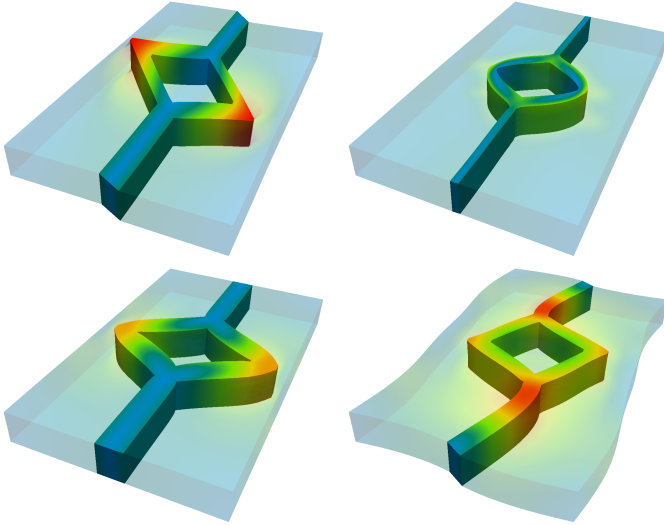


Figure 8.12: Displacement fields corresponding to some Reduced Basis elements. Coloring indicates displacement magnitude.

Evaluation and results

A validation set of response stresses $\overline{\mathbf{P}}_{\text{FE}}$ is created at additional Hencky strain sampling sites (i.e. disjoint from the training sites) parametrized by $N_{\text{dir}}^{\text{eval}} = 150$ and $N_{\text{mag}}^{\text{eval}} = 16$, where the magnitudes are of the form $t^{(i)} = i/16$, $i = 1, \dots, N_{\text{mag}}^{\text{eval}}$. Of these $150 \cdot 16 = 2400$ evaluation points, 17 are excluded due to a lack of convergence of the FEM, cf. Section 9. Therefore, here, the validation set $\overline{\mathcal{E}}_{\text{vali}}$ consists of 2383 FE-homogenized stresses corresponding to an incomplete set of concentric samples of Hencky strains $\overline{\mathbf{E}}$.

In order to assess the accuracy of the interpolant $\overline{\mathbf{P}}_{\text{CI}}$ compared to the result $\overline{\mathbf{P}}$ of the FEM, the error measure

$$\text{err}_{\text{P}} = \frac{\|\overline{\mathbf{P}}_{\text{CI}} - \overline{\mathbf{P}}_{\text{FE}}\|}{\|\overline{\mathbf{P}}_{\text{FE}}\|} \quad (8.41)$$

is evaluated on $\bar{\mathcal{E}}_{\text{vali}}$. With this at hand, the quality of a given kernel parameter γ can be judged. Figure 8.13 visualizes the influence of this parameter on the overall error for each of the chosen number of supporting directions.

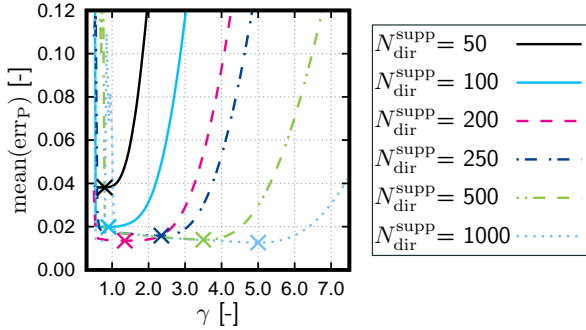


Figure 8.13: Mean value of the stress error function of Equation (8.41) on the set of 2383 validation points.

One can observe an increase of the minimum mean error from $N_{\text{dir}}^{\text{supp}} = 200$ to $N_{\text{dir}}^{\text{supp}} = 250$. This observation is interpreted as the effect of the additional error introduced by the RB approximation in the latter case.

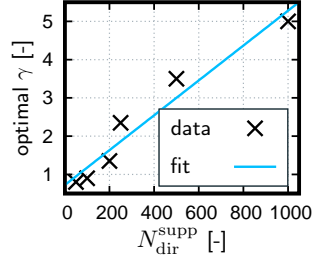
As is well-known, [Kunc & Fritzen, 2019c, Sec. 7], very small values of γ render the method numerically unstable and lead to volatile results. For small numbers of supporting directions, the optimum value marks the beginning of the numerically stable domain, e.g. as is the case for $N_{\text{dir}}^{\text{supp}} = 50$. As the number of supporting directions increases, the best value of γ increases, shifting into the interior of the numerically stable domain, e.g. as is the case for $N_{\text{dir}}^{\text{supp}} = 1000$.

An important property of the CI method is that changes to the minimum mean error value with respect to the kernel parameter are marginal within a notably large neighborhood of the optimum. This confirms an empirical finding of the original study of the small strain equivalent method, [Fritzen & Kunc, 2018c, p. 212]: the kernel parameter may be fixed at $\gamma = 2.0$ for a wide range of $N_{\text{dir}}^{\text{supp}}$ without pronounced loss of accuracy. The found optimum values of the present examples are visualized in Figure 8.14. A linear least squares regression of these data points yields the approximate relationship

$$\gamma(N_{\text{dir}}^{\text{supp}}) \approx 0.00456 N_{\text{dir}}^{\text{supp}} + 0.7208. \quad (8.42)$$

Next, the distribution of the function err_P is studied for these optimum γ values. To this end, Figure 8.15 shows the empirical distribution functions (EDF's) of this error quantity for each value of $N_{\text{dir}}^{\text{supp}}$ and for each evaluation magnitude. In other words, the EDF's measure the probability, for each evaluation magnitude separately, of the error lying within certain bounds when evaluated at any of the $N_{\text{dir}}^{\text{eval}} = 150$ directions.

Figure 8.14: Values of the kernel parameter γ as results from the conducted optimizations (*points*, cf. Figure 8.13) and the result of a linear least squares regression (*line*, cf. Equation (8.42))



One can observe that the mean of the error err_P is below 2% for as few as $N_{\text{dir}}^{\text{supp}} = 100$ supporting directions. Also, the individual EDF's are comparatively “smooth”, i.e. they do not exhibit significant jumps or plateaus. This confirms a lack of severe outliers of the error function (8.41).

It is worth noticing that, just as with the results depicted in Figure 8.13, the mean of the error seems to be decreasing monotonically with an increasing number of supporting directions, except for the transition from $N_{\text{dir}}^{\text{supp}} = 200$ to $N_{\text{dir}}^{\text{supp}} = 250$. This is very likely due to the additional error introduced by the RB, slightly deteriorating the quality of the data at the supporting points. Still, a further increasing number of supporting directions leads, again, to a monotonic decrease of the mean error. In terms of the mean error, parity of the RB-based CI scheme with the finest FE-based CI scheme is only reached with $N_{\text{dir}} > 500$. However, even with $N_{\text{dir}}^{\text{supp}} = 1000$ the ROM-based CI scheme is notably more prone to the maximum error than the FEM-based one with just one fifth as many supporting points.

Finally, the average CPU time of all CI evaluations on the validation set $\bar{\mathcal{E}}_{\text{vali}}$ are depicted in Figure 8.16 as a function of $N_{\text{dir}}^{\text{supp}}$. The average is taken with respect to 452770 evaluations of the stress interpolant $\bar{\mathbf{P}}_{\text{CI}}$ (2383 evaluation Hencky strains and 190 different values of γ , cf. Figure 8.13). For these computations, a standard laptop computer was used. No parallelization was employed.

An almost linear relationship between the average CPU time T and $N_{\text{dir}}^{\text{supp}}$ can be observed, $T \approx 1.91 \cdot 10^{-7} \text{s} \cdot N_{\text{dir}}^{\text{supp}} + 2.91 \cdot 10^{-6} \text{s}$. The actual relation barely notably tends to superlinearity.

These CPU times correspond to more than 5100 evaluations per second for $N_{\text{dir}}^{\text{supp}} = 1000$ and 75000 evaluations per second for $N_{\text{dir}}^{\text{supp}} = 50$. In particular, CI enables 1 million evaluations of the homogenized finite strain constitutive model in only ~ 200 seconds ($N_{\text{dir}}^{\text{supp}} = 10000$) and in 14 seconds ($N_{\text{dir}}^{\text{supp}} = 50$), respectively.

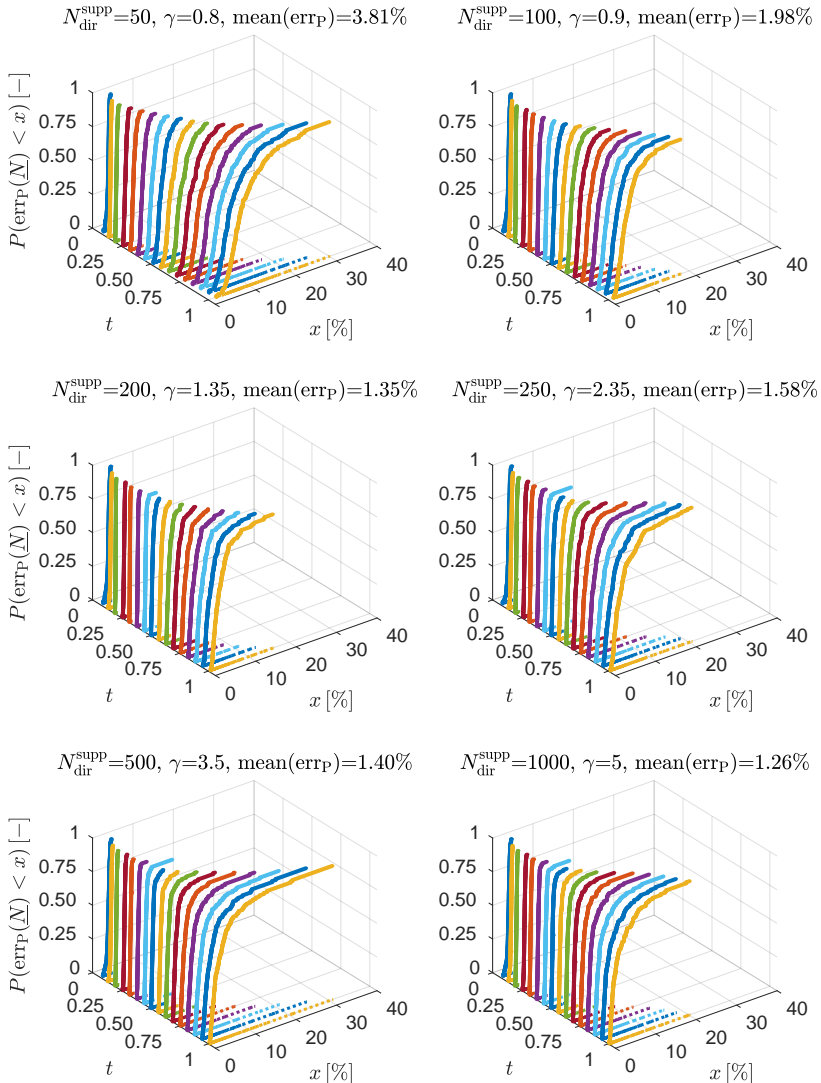


Figure 8.15: EDF's of stress error err_P of hash microstructure.

8.5 Summary and discussion

8.5.1 Summary

A novel method for the homogenization of many spatial scales in finite strain hyperelasticity was successfully demonstrated. The combination of moderate, simple paral-

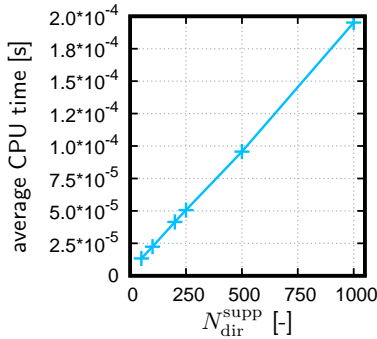


Figure 8.16: Average CPU time of a single evaluation of the stress interpolant $\bar{\mathbf{P}}_{\text{CI}}$, as a function of the number of supporting directions $N_{\text{dir}}^{\text{supp}}$. The average is taken with respect to 452770 evaluations.

lization on standard workstations and efficient numerical methods rendered the offline homogenization across three scales possible within a couple of hours. The online phase was conducted on a laptop computer within a few minutes of compute time. Many thousands of evaluations of the numerical surrogate for the upscaled material response are evaluated per second.

Qualitatively, the observed results matched the expectations, i.e. nested scales with porous RVE's resulted in a significant scale-softening effect. Also, a highly anisotropic RVE was homogenized with good accuracy.

The main challenge was a deteriorating convergence behavior for the larger scales. This is a widely-known issue that, to the best knowledge of the authors, is present in most if not all computational multi-scale methods. Additionally, the interplay of the many parameters for the main algorithm does not contribute to user-friendliness.

The workflow can possibly be significantly simplified and accelerated if one settles with interpolation of data generated directly by the FEM. This is especially viable when effective stiffnesses are not sought-after. As far as usability is concerned, it must be emphasized that the scheme is non-intrusive in the sense that it can be utilized with any simulation software that provides an interface to user-defined material routines.

8.5.2 Interpolation given different strain measures

Sticking to the Hencky strain space for the sampling comes with many advantages. Most importantly, the Hencky strain space is isomorphic to the full space \mathbb{R}^6 while the spaces of other deformation or strain measures are isomorphic to nonlinear manifolds embedded into \mathbb{R}^6 (e.g. for \mathbf{C} , \mathbf{U}) and \mathbb{R}^9 (e.g. for \mathbf{F}) due to constraints on the tensors such as positive definiteness or preserved orientations. More precisely, any point in \mathbb{R}^6 will yield an admissible Hencky strain while for other deformation measures (independent of the chosen basis) additional checks on the admissibility are required. Further, the Hencky strain allows to explicitly control the sampling of the determinant of \mathbf{F} in a straightforward manner, which is a tricky procedure for other strain measures. It should be

noted that regardless of what input strain is delivered, as long as the sampling is effected with respect to strain states corresponding to the CS points in the Hencky strain space, concentric interpolation can be performed after transforming the input to the Hencky domain. However, this does not imply that the program calling the CI-based scheme must work with the Hencky strain. Note also that the validity of the samples in more obvious deformation measures (such as the right Cauchy Green strain) yields points for which the requested samples cannot be obtained, as discussed in other works, [Yvonnet et al., 2013]. Therefore, the authors decided to use and also to promote the use of CS in the Hencky strain space and the related CI although this choice is not intrinsically unique and explicitly not a rigorous requirement of CI.

8.5.3 Number of sampling directions

From the results of Section 8.4.2, it is concluded that the gains in accuracy by a larger number of supporting directions is partly offset by the additional modeling error of the RB scheme. Moreover, this error comes along with significant additional offline costs for the identification and evaluation of the RB. Future investigations of the shortcut of directly interpolating high-fidelity data – if available – might be fruitful. The trade-off with respect to the resolution of anisotropic effects may be a delicate choice. Note that by anisotropy, not just explicitly given material laws are meant but also geometry-induced effects on higher scales.

8.5.4 Number and position of sampling magnitudes

Choosing the distribution of the $N_{\text{mag}}^{\text{supp}}$ radial supporting points relies on empiricism at the current state of the art. Variations of the qualitative distribution as well as variations of the number $N_{\text{mag}}^{\text{supp}}$ were exemplary treated previously, [Fritzen & Kunc, 2018c, Kunc & Fritzen, 2019c]. It was found that CI is rather insensitive to the distribution of the supporting points and that an increase of their amount correlates with the quality of the interpolation. However, neither the positioning of the points nor the determination of their quantity can be rigorously guided at the moment, e.g. by means of an error estimation.

8.5.5 Kernel parameter

Also, the findings of Section 8.4.2 suggest that the determination of the optimum value of the kernel parameter γ may be omitted if the value is chosen according to Equation (8.42). The error induced by this possibly sub-optimal γ value may be justifiable depending on the context. An optimization of the kernel parameter should generally be conducted on data that is primary to the balance equations, i.e. on stresses and not on tangent moduli or on energy densities. However, fitting to modulus data yielded a comparable result in the present study.

8.5.6 Accuracy

Errors propagate exponentially through scales. As far as errors stemming from the data-driven approach are concerned, these might be assessed by performing unreduced FE² simulations – for a single scale transition. Concerning three or more scales, such unmodified schemes are practically impossible with meaningful spatial resolution at each scale, as the computational effort also grows exponentially with the number of scales. Further, convergence issues with classical methods remain a serious hurdle, as was also experienced in the present work. Thus, the application of the RB method reduces the computational effort to a manageable level not just because of the numerical economy but also due to the increased numerical robustness. Because of the latter, larger load increments are possible and, as we experienced, certain load levels become computationally feasible. For growing load magnitudes, the accuracy of the CI-model tends to decrease. This effect was investigated in previous works, [Kunc & Fritzen, 2019c, Fig. 12]. There, it was also shown that an increase of $N_{\text{mag}}^{\text{supp}}$ generally improves the accuracy – with only very minor effects on the computational effort during the online evaluation of the surrogate model.

8.5.7 Generalization to other material models

In principle, a CI surrogate can be set up for any material model with exclusive state-dependency. Gradient elasticity, where the energy density function also depends on the second order deformation gradient, $W = W(\mathbf{F}, \nabla_{\mathbf{X}}\mathbf{F})$, is a candidate that belongs to this category, [Forest, 2009]. In such a case, the interpolation domain should be re-considered since the choice of the Hencky strain space was not made with higher gradients in mind.

Dissipative material laws might be covered or at least assisted by CI. For instance, standard viscoelastic material models, [Simo, 1987], contain a hyperelastic term that can be directly treated by the CI method as in the present paper. History dependence might be treated by means of a decisive super-model switching between multiple CI schemes, each covering a certain region of the historic state space. However, the high dimensionality of this space would drastically increase the computational effort. To begin with, the numerical surrogate would be more complex. But even more severely, the state history would need to be tracked at each integration point on the current Scale N , as is the case in many classical dissipative material models involving internal variables. It could be an option for the super-model to decide to retreat to unreduced FE simulations on Scale $N-1$ in case the state evolves along a path leaving the CI-covered domain, [Fritzen et al., 2019]. Future works might reveal alternatives pathways for the consideration of dissipative effects with notable contributions of CI for attaining computational efficiency.

Author contributions

Conceptualization, O.K. and F.F.; Methodology, O.K. and F.F.; Software, O.K.; Graphics, O.K.; Writing, O.K.; Supervision, F.F.; Project administration, F.F.; Funding acquisition, F.F.

Funding information

This work is funded by the German Research Foundation (DFG) – project number 257987586, grant DFG FR2702/6 within the scope of the Emmy-Noether program. This work is funded by the German Research Foundation (DFG) – project number 406068690, grant DFG FR2702/8 within the scope of the Heisenberg program. This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2075 – project number 390740016.

Conflict of interest

The authors declare no potential conflict of interests.

Chapter 9:

Outlook

9.1 Possible advancements of the current method

The demonstrated computability of fourth-scale results on laptop computers suggests that the *application phase* of the proposed numerical scheme is sufficiently efficient. Still, further improvements could possibly be achieved on the implementation side. For instance, the efficient usage of platform-specific SIMD instructions (Single Instruction, Multiple Data) is a popular method to comparatively easily boost algebraic operations. In the published code, such an advanced set of instructions is not yet employed.

In contrast, it is desirable that the *setup phase* could be conducted in a more efficient manner. A major bottleneck is the POD during which the pre-computed FE results are searched for correlations. This memory-intensive operation could be alternatively conducted by the established Hierarchical Approximate POD (HAPOD) method, [Himpe et al., 2018]. There, the monolithical POD approach is split into POD's of smaller subsets. This allows for parallelization and/or a decrease of the peak memory requirement.

Another significant bottleneck of the *setup phase* is the RB method. The major runtime is spent on field operations, such as the evaluation of the material laws and the integration over the RVE domain. This is somewhat disappointing when thinking of the tremendous reduction of degrees of freedom from the FE model to the RB model. It is a well-known drawback of the classical RB method that the complexity of the original domain prevails. This problem was specifically targeted, e.g., by the Empirical Cubature Methode (ECM), [Hernández et al., 2017]. It is to be expected that a reduced integration procedure similar to this would lead to significant speed-ups. Comparisons of various reduction techniques were conducted, e.g., in Brands et al. [2019]. It might also be worth to explore the possibility to interpolate the RB coefficients via CI, based upon previous RB evaluations. This could lead to quicker convergence due to better initial guesses of the coefficients. Also, the option to migrate the implementation of the RB method onto a massively parallel GPU framework, e.g. similar to Fritzen & Hodapp [2016], shall be mentioned.

In terms of *usability* and *robustness*, the current method may possibly be improved if the amount of parameters could be reduced. For example, when thinking of Table 8.1, it is suggestive that some relationships among the many parameters might be existent. It could therefore be rewarding to conduct parameter studies in order reveal dependencies or to at least be able to provide guidance as for how to choose the parameters.

9.2 Extension to other non-dissipative material classes

In principle, the CI-method is suitable for any non-dissipative material class. Since such models have a unique relationship between the kinematic state variables and the material response, they fall within the regime of moderate-dimensional functions to which the CI-method was proven applicable.

One example of a more general, non-dissipative class of materials is electro-elasticity, cf. [Keip et al., 2014]. There, the material model consists of a potential that is dependent on both the right Cauchy-Green tensor and the vectorial electric field, [Keip et al., 2014, Equation (30)].

Another example is the class of magneto-active materials, cf. [Javili et al., 2013, Keip & Rambausek, 2016]. There, the material model ([Javili et al., 2013, Equation (27)], [Keip & Rambausek, 2016, Equation (56)]) is of extended hyperelasticity-kind, too. Besides the right Cauchy-Green tensor only the vectorial magnetic field introduces an additional dependency of the potential.

In both of these examples, the dimension of the input space of the model is nine. Since the CI-method has demonstrated black-box capabilities for an arbitrarily chosen, eight-dimensional scalar function, it is suggestive that it is well applicable to these electro- or magneto-active cases, too. More specifically, one could apply Concentric Sampling and subsequent Concentric Interpolation of the effective response of an RVE with respect to the respective nine-dimensional input space.

9.3 Extension to dissipative material classes

It is fundamental to the present two-staged scheme that the relationship between strain and the hyperelastic response is a mathematical function, cf. (4.26):

$$\tilde{f} : \mathbf{E}_H \mapsto \mathbf{S}, \mathbf{C} \quad (9.1)$$

General material laws do not exhibit such a functional relationship. Path-dependency, i.e. the significance of the load path through the space of state variables, requires knowledge of the load history. It is thus necessary to increase the dimension of the domain of the function \tilde{f} . The number of required additional dimensions may be significant, as multiple historic values might be required for each component of the strain.

A common and general modeling approach to path-dependent material behavior is the concept of *internal variables*, cf., e.g., [Maugin & Muschik, 1994]. An internal variable is a quantity that is not directly measurable but nonetheless affects the state of the material. Past values of kinematic states, hardening variables, and viscous strains are just few examples of this concept.

Internal variables increase the dimension of the domain of the function \tilde{f} and may even require dedicated modeling assumptions. In [Fritzen & Leuschner, 2013], for instance, the chosen material models allow for an efficient treatment of the internal variables by means of non-kinematic RB's. It remains an open task to generalize this method to the case of geometric nonlinearity.

An application of the CI method in the context of path-dependent three-dimensional problems would face many additional issues, including but not limited to:

- A careful choice of the state variables (as far as possible) in order to mitigate the effects of the curse of dimensionality.
- Especially the setup phase with Concentric Sampling would be sensitive to an excessive growth of the dimensionality, assuming a classical high-fidelity method is utilized to compute the solutions at the sampling sites.
- It may be necessary to fix the temporal discretization, especially for rate-dependent problems. An individual instance of CI may be necessary for each discretization of time.
- The physical interpretation of the angular distance between sampling directions in the space of state variables is lost. While in the present scheme this distance measures how much loads differ in kind, such an interpretation is not readily available if non-strain variables are included in the sampling space. Therefore, it may possibly be advantageous to choose distributions of sampling directions that are *not* homogeneous.
- Consequently, the spherical kernel interpolation would have to be adapted by means of a suitable metric. A prototype of such an interpolation scheme is existent but has not been published yet.

At this point, a sophisticated estimation of the suitability of the CI method for such generalized applications is not made. It is recommended to apply CI to one-dimensional path-dependent constitutive laws first in order to conduct initial proof-of-concept studies.

However, it is straightforward to assist homogenization by means of the present two-staged method in other cases involving path-dependency: certain finite strain visco-elastic material laws involve a hyperelastic part, cf., e.g., [Simo, 1987, Section 1.2], [Hartmann, 2002, Section 2.2]. If the behavior on a scale $N > 0$ was sought to be described by a model of such kind, the current method could be directly utilized to identify the hyperelastic part.

Appendix A:

Appendix

The compatibility conditions (2.8) and (2.9) require the following additional standard notation. Firstly, the Levi-Civita symbol is defined as

$$e_{ijk} = \begin{cases} 1 & \text{if } (i, j, k) \in \{(1, 2, 3), (2, 3, 1), (3, 1, 2)\} \\ -1 & \text{if } (i, j, k) \in \{(3, 2, 1), (2, 1, 3), (1, 3, 2)\} \\ 0 & \text{else} \end{cases}, \quad (\text{A.1})$$

cf. [Lublinter, 2008, p. 2]. Secondly, let $\mathbf{a} = \mathbf{a}(\mathbf{X})$ be a twice differentiable second order tensor field on Ω_0 . Then the first derivative of a component a_{ij} of \mathbf{a} with respect to the k -th spatial coordinate is denoted as $a_{ij,k}$ ($i, j, k \in \{1, 2, 3\}$). Likewise, $a_{ij,kl}$ denotes the differentiation of $a_{ij,k}$ with respect to the l -th spatial coordinate ($l \in \{1, 2, 3\}$).

The ij -component of the second order tensor on the left-hand side of the compatibility conditions (2.8) is defined via

$$(\nabla_{\mathbf{X}} \times \mathbf{F})_{ij} = \sum_{k,l=1}^3 e_{klj} F_{il,k}, \quad (\text{A.2})$$

where the terms F_{il} coin the components of the deformation gradient and $i, j \in \{1, 2, 3\}$.

The ij -component of the second order tensor on the left-hand side of the compatibility conditions (2.9) is defined via

$$\nabla_{\mathbf{X}} \times (\nabla_{\mathbf{X}} \times \boldsymbol{\varepsilon})_{ij} = \sum_{k,l,m,n=1}^3 e_{mki} e_{nlj} \varepsilon_{mn,kl}, \quad (\text{A.3})$$

where the terms ε_{il} denote the components of the small strain tensor for $i, j \in \{1, 2, 3\}$.

Bibliography

- Abiodun, O. I.; Jantan, A.; Omolara, A. E.; Dada, K. V.; Mohamed, N. A. & Arshad, H.: State-of-the-art in artificial neural network applications: A survey. *Heliyon* **4** (2018), e00938, Elsevier BV, doi:10.1016/j.heliyon.2018.e00938.
- Acharya, A.: On compatibility conditions for the left cauchy-green deformation field in three dimensions. *Journal of Elasticity* **56** (1999), 95–105, Springer Science and Business Media LLC, doi:10.1023/a:1007653400249.
- Agoras, M.; Avazmohammadi, R. & Ponte-Castañeda, P.: Incremental variational procedure for elasto-viscoplastic composites and application to polymer and metal-matrix composites reinforced by spheroidal elastic particles. *International Journal of Solids and Structures* **97-98** (2016), 668–686, Elsevier BV, doi:10.1016/j.ijsolstr.2016.04.008.
- Agoras, M. & Castañeda, P. P.: Homogenization estimates for multi-scale nonlinear composites. *European Journal of Mechanics - A/Solids* **30** (2011), 828–843, Elsevier BV, doi:10.1016/j.euromechsol.2011.05.007.
- Akkari, N.; Casenave, F. & Moureau, V.: Time stable reduced order modeling by an enhanced reduced order basis of the turbulent and incompressible 3d navier–stokes equations. *Mathematical and Computational Applications* **24** (2019), 45, MDPI AG, doi:10.3390/mca24020045.
- Altenbach, H.: *Kontinuumsmechanik*. Springer Berlin Heidelberg 2018, doi:10.1007/978-3-662-57504-8.
- An, S.; Kim, T. & James, D. L.: Optimizing cubature for efficient integration of subspace deformations. *ACM Transactions on Graphics* **27** (2008), 1–10, Association for Computing Machinery (ACM), doi:10.1145/1409060.1409118.
- Andelfinger, P.; Jünemann, K. & Hartenstein, H.: Parallelism Potentials in Distributed Simulations of Kademia-based Peer-to-peer Networks. In *Proceedings of the Seventh International Conference on Simulation Tools and Techniques*, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) 2014, SIMUTools '14, pp. 41–50, doi:10.4108/icst.simutools.2014.254609.
- Ball, J. M.: Convexity Conditions and Existence Theorems in Nonlinear Elasticity. *Archive for Rational Mechanics and Analysis* **63** (1976), 337–403, Springer Science and Business Media LLC, doi:10.1007/bf00279992.
- Ballinger, B.; Blekherman, G.; Cohn, H.; Giansiracusa, N.; Kelly, E. & Schürmann, A.: Experimental study of energy-minimizing point configurations on

- spheres. *Experimental Mathematics* **18** (2009), 257–283, Taylor & Francis, doi:10.1080/10586458.2009.10129052.
- Barkaoui, A.; Chamekh, A.; Merzouki, T.; Hambli, R. & Mkaddem, A.: Multiscale approach including microfibril scale to assess elastic constants of cortical bone based on neural network computation and homogenization method. *International Journal for Numerical Methods in Biomedical Engineering* **30** (2013), 318–338, Wiley, doi:10.1002/cnm.2604.
- Barrault, M.; Maday, Y.; Nguyen,Ngoc Cuong & Patera, A. T.: An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique* **339** (2004), 667–672, Elsevier BV, doi:10.1016/j.crma.2004.08.006.
- Bensoussan, A.; Lions, J.-L. & Papanicolaou, G.: *Asymptotic Analysis for Periodic Structures*. North-Holland 1978, doi:10.1016/s0168-2024(08)x7015-8.
- Bertram, A.: *Elasticity and Plasticity of Large Deformations*. Springer Berlin Heidelberg 2008, doi:10.1007/978-3-540-69400-7.
- Bhattacharjee, S. & Matouš, K.: A nonlinear manifold-based reduced order model for multiscale analysis of heterogeneous hyperelastic materials. *Journal of Computational Physics* **313** (2016), 635–653, Elsevier BV, doi:10.1016/j.jcp.2016.01.040.
- Bilger, N.; Auslender, F.; Bornert, M.; Michel, J.-C.; Moulinec, H.; Suquet, P. & Zaoui, A.: Effect of a nonuniform distribution of voids on the plastic response of voided materials: a computational and statistical analysis. *International Journal of Solids and Structures* **42** (2005), 517–538, Elsevier BV, doi:10.1016/j.ijsolstr.2004.06.048.
- Bilger, N.; Auslender, F.; Bornert, M.; Moulinec, H. & Zaoui, A.: Bounds and estimates for the effective yield surface of porous media with a uniform or a nonuniform distribution of voids. *European Journal of Mechanics - A/Solids* **26** (2007), 810–836, Elsevier BV, doi:10.1016/j.euromechsol.2007.01.004.
- Bishop, J. F. W. & Hill, R.: XLVI. A theory of the plastic distortion of a polycrystalline aggregate under combined stresses. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **42** (1951), 414–427, Informa UK Limited, doi:10.1080/14786445108561065.
- Boas, W. & Schmid, E.: Zur berechnung physikalischer konstanten quasiisotroper vielkristalle. *Helvetica Physica Acta* **7** (1934), 628–632, doi:10.5169/seals-110392.
- Braides, A.: Loss of polyconvexity by homogenization. *Archive for Rational Mechanics and Analysis* **127** (1994), 183–190, Springer Science and Business Media LLC, doi:10.1007/bf00377660.

- Brands, B.; Davydov, D.; Mergheim, J. & Steinmann, P.: Reduced-order modelling and homogenisation in magneto-mechanics: A numerical comparison of established hyper-reduction methods. *Mathematical and Computational Applications* **24** (2019), 20, MDPI AG, doi:10.3390/mca24010020.
- Brauchart, J.; Hardin, D. & Saff, E.: The next-order term for optimal Riesz and logarithmic energy asymptotics on the sphere. *Contemporary Mathematics* **578** (2012), 31–61, American Mathematical Society, doi:10.1090/conm/578/11483.
- Brauchart, J. S. & Grabner, P. J.: Distributing many points on spheres: Minimal energy and designs. *Journal of Complexity* **31** (2015), 293–326, Elsevier BV, doi:10.1016/j.jco.2015.02.003.
- Brough, D. B.; Wheeler, D. & Kalidindi, S. R.: Materials knowledge systems in python—a data science framework for accelerated development of hierarchical materials. *Integrating Materials and Manufacturing Innovation* **6** (2017), 36–53, ISSN 2193-9772, Springer Science and Business Media LLC, doi:10.1007/s40192-017-0089-0.
- Castañeda, P. P. & Suquet, P.: Nonlinear composites. Elsevier 1997, vol. 34 of *Advances in Applied Mechanics*, pp. 171–302, doi:https://doi.org/10.1016/S0065-2156(08)70321-1.
- Chen, X.; Womersley, R. S. & Ye, J. J.: Minimizing the condition number of a Gram matrix. *SIAM Journal on optimization* **21** (2011), 127–148, SIAM.
- Christensen, J. P. R.: On Some Measures Analogous to Haar Measure. *Mathematica Scandinavica* **26** (1970), 103–106, Aarhus University Library, doi:10.7146/math.scand.a-10969.
- Clément, A.; Soize, C. & Yvonnet, J.: Computational nonlinear stochastic homogenization using a nonconcurrent multiscale approach for hyperelastic heterogeneous microstructures analysis. *International Journal for Numerical Methods in Engineering* **91** (2012), 799–824, Wiley, doi:10.1002/nme.4293.
- Cohn, H. & Kumar, A.: Universally optimal distribution of points on spheres. *Journal of the American Mathematical Society* **20** (2007), 99–148, American Mathematical Society (AMS), doi:10.1090/s0894-0347-06-00546-7.
- Cubitt, T. S.; Pérez-García, D. & Wolf, M.: The un(solv)able problem. *Scientific American* **319** (2018), 28–37, Springer Science and Business Media LLC, doi:10.1038/scientificamerican1018-28.
- Damelin, S. B.; Hickernell, F. J.; Ragozin, D. L. & Zeng, X.: On energy, discrepancy and group invariant measures on measurable subsets of euclidean space. *Journal of Fourier Analysis and Applications* **16** (2010), 813–839, Springer Science and Business Media LLC, doi:10.1007/s00041-010-9153-2.

- Dietrich, S.; Gebert, J.-M.; Stasiuk, G.; Wanner, A.; Weidenmann, K.; Deutschmann, O.; Tsukrov, I. & Piat, R.: Microstructure characterization of CVI-densified carbon/carbon composites with various fiber distributions. *Composites Science and Technology* **72** (2012), 1892–1900, Elsevier BV, doi:10.1016/j.compscitech.2012.08.009.
- Doll, S. & Schweizerhof, K.: On the Development of Volumetric Strain Energy Functions. *Journal of Applied Mechanics* **67** (1999), 17–21, ASME International, doi:10.1115/1.321146.
- Drozдов, G.; Ostanin, I. & Oseledets, I.: Time- and memory-efficient representation of complex mesoscale potentials. *Journal of Computational Physics* **343** (2017), 110–114, Elsevier BV, doi:10.1016/j.jcp.2017.04.056.
- Dvorak, G.; Bahei-El-Din, Y. & Wafa, A.: The modeling of inelastic composite materials with the transformation field analysis. *Modelling and Simulation in Materials Science and Engineering* **2** (1994), 571–586, IOP Publishing, doi:10.1088/0965-0393/2/3a/011.
- Dvorak, G. & Benveniste, Y.: On transformation strains and uniform fields in multiphase elastic media. *Proceedings of the Royal Society of London A* **437** (1992), 291–310, The Royal Society, doi:10.1098/rspa.1992.0062.
- Esmailbeigi, M.; Chatrabgoun, O. & Shafa, M.: Scattered data fitting of hermite type by a weighted meshless method. *Advances in Computational Mathematics* **44** (2017), 673–691, Springer Science and Business Media LLC, doi:10.1007/s10444-017-9555-7.
- Faraut, J.: *Analysis on Lie Groups*. Cambridge University Press 2008, doi:10.1017/cbo9780511755170.
- Fasshauer, G. & McCourt, M.: *Kernel-based Approximation Methods using MATLAB*, vol. 19. WORLD SCIENTIFIC 2015, doi:10.1142/9335.
- Fasshauer, G. E. & Schumaker, L. L.: Scattered data fitting on the sphere. *Mathematical Methods for Curves and Surfaces II* (1998), 117–166, Vanderbilt University Press.
- Feyel, F.: Multiscale FE^2 elastoviscoplastic analysis of composite structures. *Computational Materials Science* **16** (1999), 344–354, Elsevier BV, doi:10.1016/S0927-0256(99)00077-4.
- Flory, P.: Thermodynamic relations for high elastic materials. *Transactions of the Faraday Society* **57** (1961), 829–838, Royal Society of Chemistry, doi:10.1039/tf9615700829.
- Forest, S.: Micromorphic approach for gradient elasticity, viscoplasticity, and damage. *Journal of Engineering Mechanics* **135** (2009), 117–131, American Society of Civil Engineers (ASCE), doi:10.1061/(asce)0733-9399(2009)135:3(117).

- Fritzen, F. & Böhlke, T.: Influence of the type of boundary conditions on the numerical properties of unit cell problems. *Technische Mechanik* **30** (2010), 354–363.
- Fritzen, F.; Fernández, M. & Larsson, F.: On-the-fly adaptivity for nonlinear twoscale simulations using artificial neural networks and reduced order modeling. *Frontiers in Materials* **6** (2019), Frontiers Media SA, doi:10.3389/fmats.2019.00075.
- Fritzen, F.; Forest, S.; Böhlke, T.; Kondo, D. & Kanit, T.: Computational homogenization of elasto-plastic porous metals. *International Journal of Plasticity* **29** (2012), 102–119, Elsevier BV, doi:10.1016/j.ijplas.2011.08.005.
- Fritzen, F. & Hodapp, M.: The Finite Element Square Reduced (FE^{2R}) method with GPU acceleration: towards three-dimensional two-scale simulations. *International Journal for Numerical Methods in Engineering* **107** (2016), 853–881, Wiley, doi:10.1002/nme.5188.
- Fritzen, F.; Hodapp, M. & Leuschner, M.: GPU accelerated computational homogenization based on a variational approach in a reduced basis framework. *Computer Methods in Applied Mechanics and Engineering* **278** (2014), 186–217, Elsevier BV, doi:10.1016/j.cma.2014.05.006.
- Fritzen, F. & Kunc, O.: Github repository ConcentricInterpolation (2018a).
- Fritzen, F. & Kunc, O.: Github repository MinimumEnergyPoints (2018b).
- Fritzen, F. & Kunc, O.: Two-stage data-driven homogenization for nonlinear solids using a reduced order model. *European Journal of Mechanics - A/Solids* **69** (2018c), 201–220, Elsevier BV, doi:10.1016/j.euomechsol.2017.11.007.
- Fritzen, F. & Leuschner, M.: Reduced basis hybrid computational homogenization based on a mixed incremental formulation. *Computer Methods in Applied Mechanics and Engineering* **260** (2013), 143–154, Elsevier BV, doi:10.1016/j.cma.2013.03.007.
- Fritzen, F. & Leuschner, M.: Nonlinear reduced order homogenization of materials including cohesive interfaces. *Computational Mechanics* **56** (2015), 131–151, Springer Berlin Heidelberg, doi:10.1007/s00466-015-1163-0.
- Fuselier, E.; Hangelbroek, T.; Narcowich, F. J.; Ward, J. & Wright, G.: Localized bases for kernel spaces on the unit sphere. *SIAM Journal on Numerical Analysis* **51** (2013), 2538–2562, Society for Industrial & Applied Mathematics (SIAM), doi:10.1137/120876940.
- Galli, M.; Cugnoni, J. & Botsis, J.: Numerical and statistical estimates of the representative volume element of elastoplastic random composites. *European Journal of Mechanics - A/Solids* **33** (2012), 31–38, Elsevier BV, doi:10.1016/j.euomechsol.2011.07.010.
- Geers, M. & Yvonnet, J.: Multiscale modeling of microstructure-property relations. *MRS Bulletin* **41** (2016), 610–616, Cambridge University Press, doi:10.1557/mrs.2016.165.

- Ghaboussi, J.; Garrett, J. H. & Wu, X.: Knowledge-based modeling of material behavior with neural networks. *Journal of Engineering Mechanics* **117** (1991), 132–153, American Society of Civil Engineers (ASCE), doi:10.1061/(asce)0733-9399(1991)117:1(132).
- Gitman, I.; Askes, H. & Sluys, L.: Representative volume: Existence and size determination. *Engineering Fracture Mechanics* **74** (2007), 2518–2534, Elsevier BV, doi:10.1016/j.engfracmech.2006.12.021.
- Grasedyck, L.: Hierarchical Singular Value Decomposition of Tensors. *SIAM Journal on Matrix Analysis and Applications* **31** (2010), 2029–2054, Society for Industrial & Applied Mathematics (SIAM), doi:10.1137/090764189.
- Gurson, A.: Continuum theory of ductile rupture by void nucleation and growth: Part I—yield criteria and flow rules for porous ductile media. *Journal for Engineering Materials and Technology* **99** (1977), 2–15, ASME International, doi:10.1115/1.3443401.
- Göküzüm, F. S.; Nguyen, L. T. K. & Keip, M.-A.: An artificial neural network based solution scheme for periodic computational homogenization of electrostatic problems. *Mathematical and Computational Applications* **24** (2019), 40, MDPI AG, doi:10.3390/mca24020040.
- Hackett, R. M.: *Hyperelasticity primer*. Springer International Publishing 2016, doi:10.1007/978-3-319-23273-7.
- Hardin, D. & Saff, E.: Minimal Riesz energy point configurations for rectifiable d -dimensional manifolds. *Advances in Mathematics* **193** (2005), 174–204, Elsevier BV, doi:10.1016/j.aim.2004.05.006.
- Hardin, R. H.; Sloane, N. J. A. & Smith, W. D.: Minimal energy arrangements of points on a sphere, last modified june 1 1997 (1997).
- Hartmann, S.: Computation in finite-strain viscoelasticity: Finite elements based on the interpretation as differential-algebraic equations. *Computer Methods in Applied Mechanics and Engineering* **191** (2002), 1439–1470, Elsevier BV, doi:10.1016/s0045-7825(01)00332-2.
- Hashin, Z. & Shtrikman, S.: A variational approach to the theory of the elastic behaviour of multiphase materials. *Journal of the Mechanics and Physics of Solids* **11** (1963), 127–140, Elsevier BV, doi:10.1016/0022-5096(63)90060-7.
- Heida, M.; Kornhuber, R. & Podlesny, J.: Fractal homogenization of multiscale interface problems. *Multiscale Modeling & Simulation* **18** (2020), 294–314, Society for Industrial & Applied Mathematics (SIAM), doi:10.1137/18m1204759.
- Hernández, J.; Caicedo, M. & Ferrer, A.: Dimensional hyper-reduction of nonlinear finite element models via empirical cubature. *Computer Methods in Applied Mechanics and Engineering* **313** (2017), 687–722, Elsevier BV, doi:10.1016/j.cma.2016.10.022.

- Hesse, K.; Sloan, I. H. & Womersley, R. S.: Numerical integration on the sphere. In *Handbook of Geomathematics*. Springer Berlin Heidelberg 2010, pp. 1185–1219, doi:10.1007/978-3-642-01546-5_40.
- Hickernell, F. J. & Hon, Y.: Radial basis function approximations as smoothing splines. *Applied Mathematics and Computation* **102** (1999), 1–24, ISSN 0096-3003, Elsevier BV, doi:10.1016/S0096-3003(98)10012-7.
- Hill, R.: The elastic behaviour of a crystalline aggregate. *Proceedings of the Physical Society. Section A* **65** (1952), 349–354, IOP Publishing, doi:10.1088/0370-1298/65/5/307.
- Hill, R.: New horizons in the mechanics of solids†. *Journal of the Mechanics and Physics of Solids* **5** (1956), 66–74, Elsevier BV, doi:10.1016/0022-5096(56)90009-6.
- Hill, R.: Elastic properties of reinforced solids: Some theoretical principles. *Journal of the Mechanics and Physics of Solids* **11** (1963), 357–372, Elsevier BV, doi:10.1016/0022-5096(63)90036-X.
- Hill, R.: A self-consistent mechanics of composite materials. *Journal of the Mechanics and Physics of Solids* **13** (1965), 213–222, Elsevier BV, doi:10.1016/0022-5096(65)90010-4.
- Hill, R.: On constitutive inequalities for simple materials-i. *Journal of the Mechanics and Physics of Solids* **16** (1968), 229–242, Elsevier BV, doi:10.1016/0022-5096(68)90031-8.
- Himpe, C.; Leibner, T. & Rave, S.: Hierarchical approximate proper orthogonal decomposition. *SIAM Journal on Scientific Computing* **40** (2018), A3267–A3292, Society for Industrial & Applied Mathematics (SIAM), doi:10.1137/16m1085413.
- Hoger, A.: The stress conjugate to logarithmic strain. *International Journal of Solids and Structures* **23** (1987), 1645–1656, Elsevier BV, doi:10.1016/0020-7683(87)90115-6.
- Jänicke, R.; Quintal, B. & Steeb, H.: Numerical homogenization of mesoscopic loss in poroelastic media. *European Journal of Mechanics - A/Solids* **49** (2015), 382–395, Elsevier BV, doi:10.1016/j.euromechsol.2014.08.011.
- Jarroudi, M. E. & Er-Riani, M.: Homogenization of elastic materials containing self-similar rigid micro-inclusions. *Continuum Mechanics and Thermodynamics* **31** (2018), 457–474, Springer Science and Business Media LLC, doi:10.1007/s00161-018-0700-4.
- Javili, A.; Chatzigeorgiou, G. & Steinmann, P.: Computational homogenization in magneto-mechanics. *International Journal of Solids and Structures* **50** (2013), 4197–4216, Elsevier BV, doi:10.1016/j.ijsolstr.2013.08.024.

- Kabel, M.; Böhlke, T. & Schneider, M.: Efficient fixed point and Newton–Krylov solvers for FFT-based homogenization of elasticity at large deformations. *Computational Mechanics* **54** (2014), 1497–1514, Springer Science and Business Media LLC, doi:10.1007/s00466-014-1071-8.
- Kalidindi, S. R.; Niezgodá, S. R.; Landi, G.; Vachhani, S. & Fast, T.: A novel framework for building materials knowledge systems. *Computers, Materials, & Continua* **17** (2010), 103–125, Tech Science Press, doi:10.3970/cmc.2010.017.103.
- Kaliske, M. & Heinrich, G.: An extended tube-model for rubber elasticity: Statistical-mechanical theory and finite element implementation. *Rubber Chemistry and Technology* **72** (1999), 602–632, Rubber Division, ACS, doi:10.5254/1.3538822.
- Kanit, T.; Forest, S.; Galliet, I.; Mounoury, V. & Jeulin, D.: Determination of the size of the representative volume element for random composites: statistical and numerical approach. *International Journal of Solids and Structures* **40** (2003), 3647–3679, Elsevier BV, doi:10.1016/S0020-7683(03)00143-4.
- Kanouté, P.; Boso, D.; Chaboche, J. & Schrefler, B.: Multiscale methods for composites: A review. *Archives of Computational Methods in Engineering* **16** (2009), 31–75, Springer Science and Business Media LLC, doi:10.1007/s11831-008-9028-8.
- Karhunen, K.: *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, vol. 37 of *Al. Math. Phys. Ann. Acad. Sci. Fennicae* 1946.
- Keip, M.-A. & Rambausek, M.: A multiscale approach to the computational characterization of magnetorheological elastomers. *International Journal for Numerical Methods in Engineering* **107** (2016), 338–360, Wiley, doi:10.1002/nme.5178.
- Keip, M.-A.; Steinmann, P. & Schröder, J.: Two-scale computational homogenization of electro-elasticity at finite strains. *Computer Methods in Applied Mechanics and Engineering* **278** (2014), 62–79, Elsevier BV, doi:10.1016/j.cma.2014.04.020.
- Kim, H. J. & Swan, C. C.: Algorithms for automated meshing and unit cell analysis of periodic composites with hierarchical tri-quadratic tetrahedral elements. *International Journal for Numerical Methods in Engineering* **58** (2003), 1683–1711, Wiley, doi:10.1002/nme.828.
- Klusemann, B.; Böhm, H. & Svendsen, B.: Homogenization methods for multi-phase elastic composites with non-elliptical reinforcements: Comparisons and benchmarks. *European Journal of Mechanics - A/Solids* **34** (2012), 21–37, Elsevier BV, doi:10.1016/j.euromechsol.2011.12.002.
- Kochmann, J.; Wulfinghoff, S.; Ehle, L.; Mayer, J.; Svendsen, B. & Reese, S.: Efficient and accurate two-scale FE-FFT-based prediction of the effective material behavior of elasto-viscoplastic polycrystals. *Computational Mechanics* **61** (2017), 751–764, Springer Science and Business Media LLC, doi:10.1007/s00466-017-1476-2.

- Kouznetsova, V.; Brekelmans, W. A. M. & Baaijens, F. P. T.: An approach to micro-macro modeling of heterogeneous materials. *Computational Mechanics* **27** (2001), 37–48, Springer Science and Business Media LLC, doi:10.1007/s004660000212.
- Kruch, S. & Chaboche, J.-L.: Multi-scale analysis in elasto-viscoplasticity coupled with damage. *International Journal of Plasticity* **27** (2011), 2026–2039, Elsevier BV, doi:10.1016/j.ijplas.2011.03.007.
- Kühnel, W.: *Matrizen und Lie-Gruppen*. Vieweg+Teubner 2011, doi:10.1007/978-3-8348-9905-7.
- Kuijlaars, A. B. J. & Saff, E. B.: Asymptotics for minimal discrete energy on the sphere. *Transactions of the American Mathematical Society* **350** (1998), 523–538, American Mathematical Society (AMS), doi:10.1090/s0002-9947-98-02119-9.
- Kunc, O.: GitHub repository ReducedBasisDemonstrator (2019).
- Kunc, O. & Fritzen, F.: Efficient assembly of linearized equations in nonlinear homogenization. *PAMM* **19** (2019a), Wiley, doi:10.1002/pamm.201900322.
- Kunc, O. & Fritzen, F.: Finite strain homogenization using a reduced basis and efficient sampling. *Mathematical and Computational Applications* **24** (2019b), MDPI, doi:10.3390/mca24020056.
- Kunc, O. & Fritzen, F.: Generation of energy-minimizing point sets on spheres and their application in mesh-free interpolation and differentiation. *Advances in Computational Mathematics* **45** (2019c), 3021–3056, Springer Science and Business Media LLC, doi:10.1007/s10444-019-09726-5.
- Kunc, O. & Fritzen, F.: Many-scale finite strain computational homogenization via Concentric Interpolation. *International Journal for Numerical Methods in Engineering* (2020), Wiley, doi:10.1002/nme.6454.
- Lahellec, N. & Suquet, P.: On the effective behavior of nonlinear inelastic composites: I. Incremental variational principles. *Journal of the Mechanics and Physics of Solids* **55** (2007a), 1932–1963, Elsevier BV, doi:10.1016/j.jmps.2007.02.003.
- Lahellec, N. & Suquet, P.: On the effective behavior of nonlinear inelastic composites: II: A second-order procedure. *Journal of the Mechanics and Physics of Solids* **55** (2007b), 1964–1992, Elsevier BV, doi:10.1016/j.jmps.2007.02.004.
- Landkof, N. S.: *Foundations of Modern Potential Theory*, vol. 180 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag Berlin Heidelberg 1972, doi:10.1007/978-3-642-65183-0.
- Le, B. A.; Yvonnet, J. & He, Q.-C.: Computational homogenization of nonlinear elastic materials using neural networks. *International Journal for Numerical Methods in Engineering* **104** (2015), 1061–1084, Wiley, doi:10.1002/nme.4953.

- Leopardi, P.: A partition of the unit sphere into regions of equal area and small diameter. *Electronic Transactions on Numerical Analysis* **25** (2006), 309–327.
- Leopardi, P.: Discrepancy, separation and riesz energy of finite point sets on compact connected riemannian manifolds. *Dolomites Research Notes on Approximation* **6** (2013), 120–129, Padova University Press, doi:10.14658/pupj-drna-2013-Special_Issue-12.
- Leuschner, M. & Fritzen, F.: Fourier-Accelerated Nodal Solvers (FANS) for homogenization problems. *Computational Mechanics* **62** (2018), 359–392, Springer Science and Business Media LLC, doi:10.1007/s00466-017-1501-5.
- Lišner, J. & Fritzen, F.: Data-driven microstructure property relations. *Mathematical and Computational Applications* **24** (2019), 57, MDPI AG, doi:10.3390/mca24020057.
- Liu, Z.; Bessa, M. & Liu, W. K.: Self-consistent clustering analysis: An efficient multi-scale scheme for inelastic heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering* **306** (2016), 319–341, Elsevier BV, doi:10.1016/j.cma.2016.04.004.
- Loève, M.: *Probability Theory*. The University Series in Higher Mathematics, Van Nostrand, Princeton, NJ 1963.
- Lovisolò, L. & Da Silva, E.: Uniform distribution of points on a hyper-sphere with applications to vector bit-plane encoding. *IEE Proceedings - Vision, Image, and Signal Processing* **148** (2001), 187–193, Institution of Engineering and Technology (IET), doi:10.1049/ip-vis:20010361.
- Lubliner, J.: *Plasticity Theory*. Dover Publications 2008.
- Ma, J.; Wang, C. & Shene, C.-K.: Coherent view-dependent streamline selection for importance-driven flow visualization. In *Visualization and Data Analysis*, SPIE 2013, doi:10.1117/12.2001887.
- Mandel, J.: *Plasticité classique et viscoplasticité*. Tech. rep., International Center for Mechanical Sciences, Udine (Italy) 1972.
- Manzhos, S. & Carrington Jr, T.: A random-sampling high dimensional model representation neural network for building potential energy surfaces. *The Journal of Chemical Physics* **125** (2006), 084109–1 – 084109–14, AIP Publishing, doi:10.1063/1.2336223.
- Marckmann, G. & Verron, E.: Comparison of hyperelastic models for rubber-like materials. *Rubber Chemistry and Technology* **79** (2006), 835–858, Rubber Division, ACS, doi:10.5254/1.3547969.
- Mardia, K. V. & Jupp, P. E. (eds.): *Directional Statistics*. John Wiley & Sons, Inc. 1999, doi:10.1002/9780470316979.

- Marsaglia, G.: Choosing a point from the surface of a sphere. *The Annals of Mathematical Statistics* **43** (1972), 645–646, Institute of Mathematical Statistics, doi:10.1214/aoms/1177692644.
- Martini, E.; Carli, G. & Maci, S.: A domain decomposition method based on a generalized scattering matrix formalism and a complex source expansion. *Progress In Electromagnetics Research B* **19** (2010), 445–473, EMW Publishing, doi:10.2528/pierb10012110.
- Marzo, J. & Ortega-Cerdà, J.: Equidistribution of feketete points on the sphere. *Constructive Approximation* **32** (2009), 513–521, Springer Science and Business Media LLC, doi:10.1007/s00365-009-9051-5.
- Matouš, K.; Geers, M. G.; Kouznetsova, V. G. & Gillman, A.: A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials. *Journal of Computational Physics* **330** (2017), 192–220, Elsevier BV, doi:10.1016/j.jcp.2016.10.070.
- Maugin, G. A. & Muschik, W.: Thermodynamics with internal variables. part i. general concepts. *Journal of Non-Equilibrium Thermodynamics* **19** (1994), Walter de Gruyter GmbH, doi:10.1515/jnet.1994.19.3.217.
- Mckay, M. D.; Beckman, R. J. & Conover, W. J.: A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. *Technometrics* **42** (2000), 55–61, Informa UK Limited, doi:10.1080/00401706.2000.10485979.
- Michel, J.; Galvanetto, U. & Suquet, P.: Constitutive relations involving internal variables based on a micromechanical analysis. In Drouot, R.; Maugin, G. & Sidoroff, F. (eds.): *Continuum Thermomechanics*. Kluwer Academic Publishers, the Netherlands 2002, pp. 301–312.
- Michel, J. & Suquet, P.: Nonuniform transformation field analysis. *International Journal of Solids and Structures* **40** (2003), 6937–6955, Elsevier BV, doi:10.1016/s0020-7683(03)00346-9.
- Michel, J. & Suquet, P.: Computational analysis of nonlinear composite structures using the nonuniform transformation field analysis. *Computer Methods in Applied Mechanics and Engineering* **193** (2004), 5477–5502, Elsevier BV, doi:10.1016/j.cma.2003.12.071.
- Miehe, C.: Numerical computation of algorithmic (consistent) tangent moduli in large-strain computational inelasticity. *Computer Methods in Applied Mechanics and Engineering* **134** (1996), 223–240, Elsevier BV, doi:10.1016/0045-7825(96)01019-5.
- Miehe, C.: Strain-driven homogenization of inelastic microstructures and composites based on an incremental variational formulation. *International Journal for Numerical Methods in Engineering* **55** (2002), 1285–1322, Wiley, doi:10.1002/nme.515.

- Miehe, C.: Computational micro-to-macro transitions for discretized micro-structures of heterogeneous materials at finite strains based on the minimization of averaged incremental energy. *Computer Methods in Applied Mechanics and Engineering* **192** (2003), 559–591, Elsevier BV, doi:10.1016/S0045-7825(02)00564-9.
- Miehe, C.; Schotte, J. & Schröder, J.: Computational micro-macro transitions and overall moduli in the analysis of polycrystals at large strains. *Computational Materials Science* **16** (1999a), 372–382, Elsevier BV, doi:10.1016/S0927-0256(99)00080-4.
- Miehe, C.; Schröder, J. & Schotte, J.: Computational homogenization analysis in finite plasticity Simulation of texture development in polycrystalline materials. *Computer Methods in Applied Mechanics and Engineering* **171** (1999b), 387–418, Elsevier BV, doi:10.1016/S0045-7825(98)00218-7.
- Mindlin, R. & Eshel, N.: On first strain-gradient theories in linear elasticity. *International Journal of Solids and Structures* **4** (1968), 109–124, Elsevier BV, doi:10.1016/0020-7683(68)90036-x.
- Mishra, N.; Vondřejc, J. & Zeman, J.: A comparative study on low-memory iterative solvers for FFT-based homogenization of periodic media. *Journal of Computational Physics* **321** (2016), 151–168, Elsevier BV, doi:10.1016/j.jcp.2016.05.041.
- Morawiec, A.: *Orientations and Rotations*. Springer Berlin Heidelberg 2004, doi:10.1007/978-3-662-09156-2.
- Morris, M. D.; Mitchell, T. J. & Ylvisaker, D.: Bayesian design and analysis of computer experiments: Use of derivatives in surface prediction. *Technometrics* **35** (1993), 243–255, [Taylor & Francis, doi:10.2307/1269517.
- Moulinec, H. & Suquet, P.: A numerical method for computing the overall response of nonlinear composites with complex microstructure. *Computer Methods in Applied Mechanics and Engineering* **157** (1998), 69–94, Elsevier BV, doi:10.1016/s0045-7825(97)00218-1.
- Moulinec, H. & Suquet, P.: Intraplase strain heterogeneity in nonlinear composites: a computational approach. *European Journal of Mechanics - A/Solids* **22** (2003), 751–770, Elsevier BV, doi:10.1016/S0997-7538(03)00079-2.
- Narcowich, F. J.; Petrushev, P. & Ward, J. D.: Localized Tight Frames on Spheres. *SIAM Journal on Mathematical Analysis* **38** (2006a), 574–594, Society for Industrial & Applied Mathematics (SIAM), doi:10.1137/040614359.
- Narcowich, F. J.; Sun, X.; Ward, J. D. & Wendland, H.: Direct and inverse Sobolev error estimates for scattered data interpolation via spherical basis functions. *Foundations of Computational Mathematics* **7** (2006b), 369–390, Springer Science and Business Media LLC, doi:10.1007/s10208-005-0197-7.

- Neff, P.; Eidel, B. & Martin, R. J.: Geometry of Logarithmic Strain Measures in Solid Mechanics. *Archive for Rational Mechanics and Analysis* **222** (2016), 507–572, Springer Science and Business Media LLC, doi:10.1007/s00205-016-1007-x.
- Nemat-Nasser, S. & Hori, M.: *Micromechanics: overall properties of heterogeneous materials*. Elsevier, Amsterdam 1999, 2., rev. ed. edn., doi:10.1016/c2009-0-09128-4.
- Nerattini, R.; Brauchart, J. S. & Kiessling, M. K.-H.: Optimal N -Point Configurations on the Sphere: “Magic” Numbers and Smale’s 7th Problem. *Journal of Statistical Physics* **157** (2014), 1138–1206, Springer Science and Business Media LLC, doi:10.1007/s10955-014-1107-7.
- Néron, D. & Ladevèze, P.: Proper Generalized Decomposition for Multiscale and Multiphysics Problems. *Archives of Computational Methods in Engineering* **17** (2010), 351–372, Springer Science and Business Media LLC, doi:10.1007/s11831-010-9053-2.
- Nguyen, V.-D.; Lani, F.; Pardoën, T.; Morelle, X. & Noels, L.: A large strain hyperelastic viscoelastic-viscoplastic-damage constitutive model based on a multi-mechanism non-local damage continuum for amorphous glassy polymers. *International Journal of Solids and Structures* **96** (2016), 192–216, Elsevier BV, doi:10.1016/j.ijsolstr.2016.06.008.
- Nguyen-Thanh, V. M.; Nguyen, L. T. K.; Rabczuk, T. & Zhuang, X.: A surrogate model for computational homogenization of elastostatics at finite strain using the hdmr-based neural network approximator. *International Journal for Numerical Methods in Engineering* **121** (2020), 4811–4842, Wiley, doi:10.1002/nme.6493.
- Nikolov, S.; Petrov, M.; Lymperakis, L.; Friák, M.; Sachs, C.; Fabritius, H.-O.; Raabe, D. & Neugebauer, J.: Revealing the Design Principles of High-Performance Biological Composites Using Ab initio and Multiscale Simulations: The Example of Lobster Cuticle. *Advanced Materials* **22** (2010), 519–526, Wiley, doi:10.1002/adma.200902019.
- Ogden, R. W.: *Non-linear elastic deformations*. Ellis Horwood, Chichester, U.K. 1984.
- Oliver, X. & Agelet de Saracibar, C.: *Continuum Mechanics for Engineers. Theory and Problems*. 2017, 2nd. edn., doi:10.13140/RG.2.2.25821.20961.
- Ostoja-Starzewski, M.: Material spatial randomness: From statistical to representative volume element. *Probabilistic Engineering Mechanics* **21** (2006), 112–132, Elsevier BV, doi:10.1016/j.probingmech.2005.07.007.
- Özdemir, I.; Brekelmans, W. & Geers, M.: Fe^2 computational homogenization for the thermo-mechanical analysis of heterogeneous solids. *Computer Methods in Applied Mechanics and Engineering* **198** (2008), 602–613, Elsevier BV, doi:10.1016/j.cma.2008.09.008.

- Pérez-Ramírez, Ú.; López-Orive, J. J.; Arana, E.; Salmerón-Sánchez, M. & Moratal, D.: Micro-computed tomography image-based evaluation of 3D anisotropy degree of polymer scaffolds. *Computer Methods in Biomechanics and Biomedical Engineering* **18** (2013), 446–455, Taylor & Francis, doi:10.1080/10255842.2013.818663.
- Ponte-Castañeda, P.: The effective mechanical properties of nonlinear isotropic composites. *Journal of the Mechanics and Physics of Solids* **39** (1991), 45–71, Elsevier BV, doi:10.1016/0022-5096(91)90030-R.
- Ponte-Castañeda, P.: New variational principles in plasticity and their application to composite materials. *Journal of the Mechanics and Physics of Solids* **40** (1992), 1757–1788, Elsevier BV, doi:10.1016/0022-5096(92)90050-C.
- Ponte-Castañeda, P.: Second-order homogenization estimates for nonlinear composites incorporating field fluctuations: I – theory. *Journal of the Mechanics and Physics of Solids* **50** (2002), 737–757, Elsevier BV, doi:10.1016/S0022-5096(01)00099-0.
- Pruchnicki, E.: Hyperelastic homogenized law for reinforced elastomer at finite strain with edge effects. *Acta Mechanica* **129** (1998), 139–162, Elsevier BV, doi:10.1007/BF01176742.
- Quarteroni, A.; Manzoni, A. & Negri, F.: *Reduced Basis Methods for Partial Differential Equations: An Introduction*. La matematica per il 3 + 2, vol. 92, Springer 2016, doi:10.1007/978-3-319-15431-2.
- Radermacher, A. & Reese, S.: Proper orthogonal decomposition-based model reduction for non-linear biomechanical analysis. *International Journal of Materials Engineering Innovation* **4** (2013), 149–165, Inderscience Publishers, doi:10.1504/ijmatei.2013.054393.
- Radermacher, A. & Reese, S.: POD-based model reduction with empirical interpolation applied to nonlinear elasticity. *International Journal for Numerical Methods in Engineering* **107** (2015), 477–495, Wiley, doi:10.1002/nme.5177.
- Rall, L. B. (ed.): *Automatic Differentiation: Techniques and Applications*. Springer Berlin Heidelberg 1981, doi:10.1007/3-540-10861-0.
- Rambausek, M.; Göküzüm, F. S.; Nguyen, L. T. K. & Keip, M.-A.: A two-scale FE-FFT approach to nonlinear magneto-elasticity. *International Journal for Numerical Methods in Engineering* **117** (2019), 1117–1142, Wiley, doi:10.1002/nme.5993.
- Ramírez-Torres, A.; Penta, R.; Rodríguez-Ramos, R.; Merodio, J.; Sabina, F. J.; Bravo-Castillero, J.; Guinovart-Díaz, R.; Preziosi, L. & Grillo, A.: Three scales asymptotic homogenization and its application to layered hierarchical hard tissues. *International Journal of Solids and Structures* **130-131** (2018), 190–198, Elsevier BV, doi:10.1016/j.ijsolstr.2017.09.035.

- Rao, C. & Liu, Y.: Three-dimensional convolutional neural network (3d-CNN) for heterogeneous material homogenization. *Computational Materials Science* **184** (2020), 109850, Elsevier BV, doi:10.1016/j.commatsci.2020.109850.
- Rashidinia, J.; Fasshauer, G. & Khasi, M.: A stable method for the evaluation of Gaussian radial basis function solutions of interpolation and collocation problems. *Computers & Mathematics with Applications* **72** (2016), 178–193, Elsevier BV, doi:10.1016/j.camwa.2016.04.048.
- Relun, N.; Néron, D. & Boucard, P. A.: A model reduction technique based on the PGD for elastic-viscoplastic computational analysis. *Computational Mechanics* **51** (2013), 83–92, Springer Science and Business Media LLC, doi:10.1007/s00466-012-0706-x.
- Rendek, M. & Lion, A.: Amplitude dependence of filler-reinforced rubber: Experiments, constitutive modelling and fem – implementation. *International Journal of Solids and Structures* **47** (2010), 2918–2936, Elsevier BV, doi:10.1016/j.ijstr.2010.06.021.
- Reuß, A.: Berechnung der Fließgrenze von Mischkristallen auf Grund der Plastizitätsbedingung für Einkristalle. *Zeitschrift für Angewandte Mathematik und Mechanik* **9** (1929), 49–58, Wiley, doi:10.1002/zamm.19290090104.
- Rippa, S.: An algorithm for selecting a good value for the parameter c in radial basis function interpolation. *Advances in Computational Mathematics* **11** (1999), 193–210, Springer Science and Business Media LLC, doi:10.1023/A:1018975909870.
- Roşca, D.: New uniform grids on the sphere. *Astronomy and Astrophysics* **520** (2010), A63, EDP Sciences, doi:10.1051/0004-6361/201015278.
- Ryckelynck, D.: A priori hyperreduction method: an adaptive approach. *Journal of Computational Physics* **202** (2005), 346–366, Elsevier BV, doi:10.1016/j.jcp.2004.07.015.
- Sachs, G.: Zur ableitung einer fließbedingung. In *Mitteilungen der deutschen Materialprüfungsanstalten*. Springer Berlin Heidelberg 1929, pp. 94–97, doi:10.1007/978-3-642-92045-5_12.
- Saeb, S.; Steinmann, P. & Javili, A.: Aspects of Computational Homogenization at Finite Deformations: A Unifying Review From Reuss’ to Voigt’s Bound. *Applied Mechanics Reviews* **68** (2016), 050801–050801–33, American Society of Mechanical Engineers, doi:10.1115/1.4034024.
- Saff, E. B. & Kuijlaars, A. B. J.: Distributing many points on a sphere. *The Mathematical Intelligencer* **19** (1997), 5–11, Springer Science and Business Media LLC, doi:10.1007/bf03024331.
- Schneider, M.: Beyond polyconvexity: an existence result for a class of quasiconvex hyperelastic materials. *Mathematical Methods in the Applied Sciences* **40** (2016), 2084–2089, Wiley, doi:10.1002/mma.4123.

- Schneider, M.; Merkert, D. & Kabel, M.: FFT-based homogenization for microstructures discretized by linear hexahedral elements. *International Journal for Numerical Methods in Engineering* **109** (2017), 1461–1489, Wiley, doi:10.1002/nme.5336.
- Schröder, J.: A numerical two-scale homogenization scheme: the FE²-method. In *Plasticity and Beyond*. Springer Vienna 2014, pp. 1–64, doi:10.1007/978-3-7091-1625-8_1.
- Seth, B. R.: *Generalized strain measure with applications to physical problems*. techreport 248, Mathematics Research Center, United States Army, University of Wisconsin 1961.
- Shen, W.; Shao, J.-F.; Dormieux, L. & Kondo, D.: Approximate criteria for ductile porous materials with compressible plastic matrix: application to double porous media. *Computational Materials Science* **62** (2012), 189–194, Elsevier BV, doi:10.1016/j.commatsci.2012.05.021.
- Shewchuk, J. R.: Triangle: Engineering a 2d quality mesh generator and delaunay triangulator. In *Applied Computational Geometry Towards Geometric Engineering*. Springer Berlin Heidelberg 1996, pp. 203–222, doi:10.1007/bfb0014497.
- Si, H.: TetGen, a delaunay-based quality tetrahedral mesh generator. *ACM Transactions on Mathematical Software* **41** (2015), 1–36, Association for Computing Machinery (ACM), doi:10.1145/2629697.
- Simo, J.: On a fully three-dimensional finite-strain viscoelastic damage model: Formulation and computational aspects. *Computer Methods in Applied Mechanics and Engineering* **60** (1987), 153–173, Elsevier BV, doi:10.1016/0045-7825(87)90107-1.
- Simo, J.: A framework for finite strain elastoplasticity based on maximum plastic dissipation and the multiplicative decomposition: Part i. continuum formulation. *Computer Methods in Applied Mechanics and Engineering* **66** (1988), 199–219, Elsevier BV, doi:10.1016/0045-7825(88)90076-X.
- Sirovich, L.: Turbulence and the Dynamics of Coherent Structures. Part 1: Coherent structures. *Quarterly of Applied Mathematics* **45** (1987), 561–571, Brown University, Division of Applied Mathematics, doi:10.1090/qam/910462.
- Sloan, I. H. & Womersley, R. S.: Constructive Polynomial Approximation on the Sphere. *Journal of Approximation Theory* **103** (2000), 91–118, Elsevier BV, doi:10.1006/jath.1999.3426.
- Smale, S.: Mathematical problems for the next century. *The Mathematical Intelligencer* **20** (1998), 7–15, Springer Science and Business Media LLC, doi:10.1007/bf03025291.

- Smit, R. J. M.; Brekelmans, W. A. M. & Meijer, H. E. H.: Prediction of the mechanical behavior of nonlinear heterogeneous systems by multi-level finite element modeling. *Computer Methods in Applied Mechanics and Engineering* **155** (1998), 181–192, Elsevier BV, doi:10.1016/S0045-7825(97)00139-4.
- Soldner, D.; Brands, B.; Zabihyan, R.; Steinmann, P. & Mergheim, J.: A numerical study of different projection-based model reduction techniques applied to computational homogenisation. *Computational Mechanics* **60** (2017), 613–625, Springer Science and Business Media LLC, doi:10.1007/s00466-017-1428-x.
- Sommariva, A. & Womersley, R.: Integration by RBF over the sphere. *Applied Mathematics Report AMR05/17* (2005), University of New South Wales.
- Spahn, J.; Andrä, H.; Kabel, M. & Müller, R.: A multiscale approach for modeling progressive damage of composite materials using fast fourier transforms. *Computer Methods in Applied Mechanics and Engineering* **268** (2014), 871–883, Elsevier BV, doi:10.1016/j.cma.2013.10.017.
- Staniforth, A. & Thuburn, J.: Horizontal grids for global weather and climate prediction models: a review. *Quarterly Journal of the Royal Meteorological Society* **138** (2011), 1–26, Wiley, doi:10.1002/qj.958.
- Suquet, P.: Elements of homogenization for inelastic solid mechanics. In *Homogenization Techniques for Composite Media*. Springer Berlin Heidelberg 1987, vol. 272 of *Lecture Notes in Physics*, pp. 193–198, doi:10.1007/3-540-17616-0_15.
- Takano, N. & Okuno, Y.: Three-scale finite element analysis of heterogeneous media by asymptotic homogenization and mesh superposition methods. *International Journal of Solids and Structures* **41** (2004), 4121–4135, Elsevier BV, doi:10.1016/j.ijsolstr.2004.02.049.
- Taylor, G. I.: Plastic strain in metals. *J. Inst. Metals* **62** (1938), 307–324.
- Temizer, İ. & Wriggers, P.: An adaptive method for homogenization in orthotropic nonlinear elasticity. *Computer Methods in Applied Mechanics and Engineering* **196** (2007), 3409–3423, Elsevier BV, doi:10.1016/j.cma.2007.03.017.
- Temizer, İ. & Wriggers, P.: On the computation of the macroscopic tangent for multiscale volumetric homogenization problems. *Computer Methods in Applied Mechanics and Engineering* **198** (2008), 495–510, Elsevier BV, doi:10.1016/j.cma.2008.08.018.
- Temizer, İ. & Zohdi, T.: A numerical method for homogenization in non-linear elasticity. *Computational Mechanics* **40** (2007), 281–298, Springer-Verlag, doi:10.1007/s00466-006-0097-y.
- Torquato, S.: *Random Heterogeneous Materials*. Springer 2006, doi:10.1007/978-1-4757-6355-3.

- Tumanov, A.: Minimal biquadratic energy of five particles on a 2-sphere. *Indiana University Mathematics Journal* **62** (2013), 1717–1731, Indiana University Mathematics Department, doi:10.1512/iumj.2013.62.5148.
- Voigt, W.: Ueber die Beziehung zwischen den beiden Elasticitätsconstanten isotroper Körper. *Annalen der Physik* **274** (1889), 573–587, Wiley, doi:10.1002/andp.18892741206.
- Voigt, W.: *Lehrbuch der Kristallphysik*, Teubner, Berlin 1910.
- Wang, Y. G.; Gia, Q. T. L.; Sloan, I. H. & Womersley, R. S.: Fully discrete needlet approximation on the sphere. *Applied and Computational Harmonic Analysis* **43** (2017), 292–316, Elsevier BV, doi:10.1016/j.acha.2016.01.003.
- Weyl, H.: Über die Gleichverteilung von Zahlen mod. Eins. *Mathematische Annalen* **77** (1916), 313–352, Springer Science and Business Media LLC, doi:10.1007/bf01475864.
- Womersley, R.: Minimum energy points on the sphere s^2 , last updated 24-jan-2003 (2003).
- Womersley, R. S. & Sloan, I. H.: How good can polynomial interpolation on the sphere be? *Advances in Computational Mathematics* **14** (2001), 195–226, Springer Science and Business Media LLC, doi:10.1023/a:1016630227163.
- Wriggers, P.: *Nonlinear finite element methods*. Springer, Berlin 2008, doi:10.1007/978-3-540-71001-1.
- Wright, G. B. & Fornberg, B.: Stable computations with flat radial basis functions using vector-valued rational approximations. *Journal of Computational Physics* **331** (2017), 137–156, Elsevier BV, doi:10.1016/j.jcp.2016.11.030.
- Wu, M.; Liu, J.; Lv, X.; Shi, D. & Zhu, Z.: A study on homogenization equations of fractal porous media. *Journal of Geophysics and Engineering* **15** (2018), 2388–2398, Oxford University Press (OUP), doi:10.1088/1742-2140/aac4c1.
- Wu, X.; J.H., J. G. & Ghaboussi, J.: Representation of material behavior: neural network-based models. In *1990 IJCNN International Joint Conference on Neural Networks*, IEEE 1990, doi:10.1109/ijcnn.1990.137574.
- Xia, L. & Breitkopf, P.: Multiscale structural topology optimization with an approximate constitutive model for local material microstructure. *Computer Methods in Applied Mechanics and Engineering* **286** (2015), 147–167, Elsevier BV, doi:10.1016/j.cma.2014.12.018.
- Yang, Z.; Long, C. & Sun, Y.: A high-order three-scale reduced asymptotic approach for thermo-mechanical problems of nonlinear heterogeneous materials with multiple spatial scales. *European Journal of Mechanics - A/Solids* **80** (2020), 103905, Elsevier BV, doi:10.1016/j.euromechsol.2019.103905.

- Yavari, A.: Compatibility equations of nonlinear elasticity for non-simply-connected bodies. *Archive for Rational Mechanics and Analysis* **209** (2013), 237–253, Springer Science and Business Media LLC, doi:10.1007/s00205-013-0621-0.
- Yu, C.; Kafka, O. L. & Liu, W. K.: Self-consistent clustering analysis for multiscale modeling at finite strains. *Computer Methods in Applied Mechanics and Engineering* (2019), Elsevier BV, doi:10.1016/j.cma.2019.02.027.
- Yuan, Z. & Fish, J.: Multiple scale eigendeformation-based reduced order homogenization. *Computer Methods in Applied Mechanics and Engineering* **198** (2009), 2016–2038, Elsevier BV, doi:10.1016/j.cma.2008.12.038.
- Yvonnet, J.: *Computational Homogenization of Heterogeneous Materials with Finite Elements*. Solid Mechanics and Its Applications, Springer Nature 2019, doi:10.1007/978-3-030-18383-7.
- Yvonnet, J.; Gonzalez, D. & He, Q.-C.: Numerically explicit potentials for the homogenization of nonlinear elastic heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering* **198** (2009), 2723–2737, Elsevier BV, doi:10.1016/j.cma.2009.03.017.
- Yvonnet, J. & He, Q.-C.: The reduced model multiscale method (R3M) for the nonlinear homogenization of hyperelastic media at finite strains. *Journal of Computational Physics* **223** (2007), 341–368, Elsevier BV, doi:10.1016/j.jcp.2006.09.019.
- Yvonnet, J.; Monteiro, E. & He, Q.-C.: Computational homogenization method and reduced database model for hyperelastic heterogeneous structures. *Journal for Multiscale Computational Engineering* **11** (2013), 201–225, Begell House, doi:10.1615/IntJMultCompEng.2013005374.
- Zeman, J.; Vondřejc, J.; Novák, J. & Marek, I.: Accelerating a FFT-based solver for numerical homogenization of periodic media by conjugate gradients. *Journal of Computational Physics* **229** (2010), 8065–8071, Elsevier BV, doi:10.1016/j.jcp.2010.07.010.

A novel computational homogenization method is proposed for hyperelastic materials on an arbitrary number of spatial scales. The scheme consists of two stages on each scale. First, a projection based reduced order model is constructed via a Reduced Basis (RB) approximation of kinematic non-displacement quantities. The computation of the effective material behavior is conducted significantly more efficiently than with the unreduced, original model. Second, samples of the homogenized material response are interpolated by means of Concentric Interpolation (CI). The support data for the CI is created at the first stage whereas the RB setup is based on solution data of any high-fidelity method, such as Finite Elements. An original Concentric Sampling (CS) strategy alleviates the curse of dimensionality and is applied to the kinematic state space at each stage and on every scale. The application of the overall method to fibrous and porous micro-structures is demonstrated successfully on up to four scales. Only laptop computers and standard workstations are utilized. It is exemplified in eight dimensions that CI is applicable in a black-box manner to more general mathematical functions. Source code is provided under a permissive license.