

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Model Order Reduction and Data-Driven Computational Modeling for Linear and Nonlinear Solids

Permalink

<https://escholarship.org/uc/item/5pr4s2zb>

Author

He, Qizhi

Publication Date

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

**Model Order Reduction and Data-Driven Computational Modeling for Linear and
Nonlinear Solids**

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Structural Engineering with a Specialization in Computational Science

by

Qizhi He

Committee in charge:

Professor Jiun-Shyan Chen, Chair
Professor Yuri Bazilevs
Professor Michael Holst
Professor Petr Krysl
Professor Kenneth J. Loh
Professor Shantanu Sinha

2018

Copyright

Qizhi He, 2018

All rights reserved.

The Dissertation of Qizhi He is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2018

DEDICATION

This dissertation is dedicated to my mother

Shu-zhi Chen

for her love, endless support

and encouragement

EPIGRAPH

Never forget why you started, and your mission can be accomplished.

--- Shu-king

Above all else, guard your heart, for everything you do flows from it.

--- Proverbs 4:23

TABLE OF CONTENTS

SIGNATURE PAGE	iii
DEDICATION	iv
EPIGRAPH	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	x
LIST OF TABLES	xvii
ACKNOWLEDGEMENTS	xviii
VITA	xxi
ABSTRACT OF THE DISSERTATION	xxiii
Chapter 1 Introduction	1
1.1 Motivation	1
<i>1.1.1 Skeletal muscle modeling</i>	1
<i>1.1.2 Reduced-order modeling</i>	5
<i>1.1.3 Data-driven computational modeling</i>	7
1.2 Strategies and Objectives	8
1.3 Outline	11
Chapter 2 Literature Review	12
2.1 Machine Learning: Data Mining and Dimensionality Reduction	12
<i>2.1.1 Supervised learning</i>	14
<i>2.1.2 Unsupervised learning</i>	17
2.2 Model Reduction and Surrogate Modeling for Time-Critical Prediction	19
<i>2.2.1 Hierarchical model</i>	20
<i>2.2.2 Data-fit model</i>	21
<i>2.2.3 Reduced-order model (ROM)</i>	23
2.3 Data-Driven Engineering Science	27
<i>2.3.1 Hybrid approach: data-enhanced physics-based modeling paradigm</i>	28
2.4 Summary of Reduced-Order and Data-Driven Modeling	31
Chapter 3 Reduced-Order Modeling of Parameterized Systems	33
3.1 Introduction	33
<i>3.1.1 Problem formulation</i>	34

3.2 Model Order Reduction Methods for Linear Dynamical System	35
3.3 Selective Projection-Based Model Order Reduction Methods	36
3.3.1 Proper orthogonal decomposition.....	36
3.3.2 Reduced-basis method	42
3.3.3 Component mode synthesis.....	45
3.3.4 Enhanced POD.....	46
3.4 Hyper-Reduction Techniques for Nonlinear Systems.....	47
3.4.1 Lifting bottleneck.....	47
3.4.2 Interpolation-type strategy	49
3.4.3 Cubature-type strategy	58
3.4.4 Linearization-type strategy.....	62
3.5 Summary	63
Chapter 4 Model Order Reduction for Fracture Mechanics via Decomposed Projection	64
4.1 Introduction.....	64
4.2 Fine-Scale Modeling of Fracture Mechanics	68
4.2.1 Problem formulation based on decomposed solution scheme	68
4.2.2 ISBFM based Galerkin formulation	71
4.2.3 Comparison of standard Galerkin and ISBFM Galerkin methods for LEFM	74
4.3 Reduced-Order Modeling of Fracture Mechanics.....	76
4.3.1 Generic projection based MOR.....	77
4.3.2 MOR based on standard modal analysis	78
4.3.3 MOR based on decomposed projection	85
4.4 Numerical Examples: Reduced-Order Modeling of Fracture Mechanics	90
4.4.1 Comparison of MOR methods for a cracked beam Poisson's problem.....	91
4.4.2 Reduced-order modeling of a mode I line crack model.....	96
4.4.3 Reduced-order modeling of the mixed mode problem.....	102
4.5 Summary	105
4.6 Acknowledgement	107
4.7 Appendix.....	107
Chapter 5 Development of Manifold Learning Based Model Order Reduction for Nonlinear Mechanics	109
5.1 Introduction.....	109
5.2 Problem Formulation: Nonlinear Model Order Reduction	112
5.2.1 High-dimensional parameterized nonlinear discrete model	112

5.2.2 Proper orthogonal decomposition.....	114
5.2.3 Discrete empirical interpolation method (DEIM).....	115
5.2.4 Potential limitations of POD based model reduction methods.....	116
5.3 Manifold Learning via Graph Embedding Framework.....	118
5.3.1 Manifold learning.....	118
5.3.2 Laplacian eigenmaps (LE).....	120
5.3.3 Graph embedding and linearization approach.....	122
5.3.4 Linear graph embedding projection.....	125
5.4 Robust Nonlinear Model Reduction Based on LGE.....	130
5.4.1 Benchmark problem: 2D Parametrized function.....	133
5.4.2 Parameterized elliptic PDE.....	139
5.4.3 One-dimensional Burgers' problem.....	142
5.5 Summary.....	150
5.6 Acknowledgement.....	151
Chapter 6 Data-Driven Computational Simulation Based on Locally Convex Reconstruction for Noisy Database.....	152
6.1 Introduction.....	153
6.2 Data-Driven Computational Framework.....	156
6.2.1 Governing equations.....	156
6.2.2 Data-driven problem formulation.....	158
6.2.3 Data-driven solver.....	160
6.3 Locally Convex Data-Driven Scheme.....	166
6.3.1 Local convex hull construction based on k -nearest neighbors.....	166
6.3.2 Solving non-negative least squares.....	169
6.3.3 Locally convex data-driven algorithm.....	172
6.4 Application to Truss Structures.....	173
6.4.1 Example I: One-dimensional truss.....	174
6.4.2 Example II: Truss system.....	181
6.5 Application to Elasticity Problems.....	186
6.5.1 Discretization for elasticity problem.....	187
6.5.2 Example III: Elastic beam subjected to a shear load.....	188
6.6 Summary.....	195
6.7 Acknowledgement.....	198

Chapter 7 Nonlinear Model Order Reduction Based on Meshfree Nodal Integration Framework	199
7.1 Introduction.....	199
7.2 Meshfree Galerkin Formulation for Hyperelasticity.....	200
7.2.1 Stabilized conforming nodal integration (SCNI).....	203
7.3 Nonlinear MOR Based on Meshfree Galerkin Formulation.....	204
7.4 Numerical Example	207
7.5 Summary.....	211
7.6 Acknowledgement	212
Chapter 8 Application of Data-Driven Modeling to Biological Tissues	213
8.1 Data-Driven Computational Framework for Nonlinear Solids.....	213
8.2 Numerical Examples: Data-Driven Modeling for Nonlinear Solids.....	216
8.2.1 Large deflection of a cantilever beam subjected to a tip shear load.....	216
8.2.2 Uniaxial tension of hyper-elasticity material	218
8.3 Data-Driven Modeling: Porcine Atrioventricular Heart Valve Tissue	223
8.3.1 Biaxial mechanical experiments.....	223
8.3.2 Data-driven modeling.....	227
8.4 Summary.....	232
8.5 Acknowledgement	233
Chapter 9 Conclusions and Future Work.....	234
9.1 Conclusions.....	234
9.2 Recommendations for Future Research	236
Appendix A Reproducing Kernel Particle Method (RKPM).....	238
Appendix B Nonnegative Least Squares (NNLS) Solver.....	241
Bibliography	243

LIST OF FIGURES

Figure 1.1: Skeletal muscle structure: (a) The entire muscle is surrounded by the epimysium, which develops as perimysium covering the fascicles, and finally as endomysium which surrounds the individual muscle fibers (Gray & Carter 1995); (b) the honeycomb-like structure of endomysium in fiber level (Gillies & Lieber 2011).	2
Figure 1.2: Description of multiscale and multi-physics simulation of skeletal muscle.	3
Figure 2.1: Types of machine learning. (http://www.cognum.com/index.php/cognitive-platform/)	13
Figure 2.2: Schematic of general projection-based model order reduction for parameterized physical systems.....	25
Figure 2.3: Schematics of data-driven computational mechanics, where data-driven models instead of empirical models are used together with physics-based models to perform simulation.	31
Figure 3.1 Singular values between raw snapshots (left) and mean-centered snapshots (right). .	55
Figure 3.2 Comparison of DEIM between raw snapshots (left) and mean-centered snapshots (right).	56
Figure 3.3 The normalized singular value $\sigma_i^2 / (\sum_i \sigma_i^2)$ (left) and the first six POD basis vectors (right) of the of the snapshot matrix X_s corresponding to the nonlinear function g^1	57
Figure 3.4 First 12 DEIM points (left) selected by the greedy algorithm for the nonlinear function g^1 , and the comparison of average L^2 errors of POD and DEIM approximations (right) for the training data.	57
Figure 4.1: Geometry of a cracked domain	69
Figure 4.2: Normalized error in L_2 norm (left) and semi- H^1 norm (right) of the loaded line crack solution under different orders of Gaussian quadrature rule. The standard Galerkin and the proposed ISBFM Galerkin methods are compared under different orders of enrichment functions, i.e. $\hat{N} = 1, 2, 3$	75
Figure 4.3: Relative error of the Mode I stress intensity factor of the loaded line crack solution under different orders of Gaussian quadrature rule. The standard Galerkin and the proposed ISBFM Galerkin methods are compared under different orders of enrichment functions, i.e. $\hat{N} = 1, 2, 3$	76
Figure 4.4: A finite dimensional model of a line crack subjected to a far-field hydrostatic tension, where traction is prescribed on $\Gamma_1 \cup \Gamma_3 \cup \Gamma_4$ and displacement is imposed on Γ_2 according to the analytical solution of (England 2003). The adopted material properties of borosilicate glass are $E = 64,000 \text{ N/mm}^2$ and $\nu = 0.2$	81

Figure 4.5: Shapes of ROB functions of ISBFM-MA for the line crack model using $\hat{N} = 2$ non-scaled near-tip basis functions: (a) $w_x^r(\phi_j)$ and (b) $w_y^r(\phi_j)$. (The subtitle EV n of each plot denotes the n-th eigenfunction of the ISBFM discrete system \mathbf{K}) 83

Figure 4.6: Shapes of ROB functions of ISBFM-MA for the line crack model using $\hat{N} = 2$ scaled non-smooth basis functions with $\gamma = 50$: (a) $w_x^r(\phi_j)$ and (b) $w_y^r(\phi_j)$. (The subtitle EV n of each plot denotes the n-th eigenfunction of the ISBFM discrete system \mathbf{K}) 84

Figure 4.7: Deformation of the line crack model approximated by (a) fine-scale approximation, (b) ISBFM-MA ($k = 40$) using non-scaled near-tip basis functions, and (c) ISBFM-MA ($k = 40$) using scaled non-smooth basis functions with $\gamma = 50$. The blue circles denote the undeformed configuration and the red dots denote the deformed configuration. 85

Figure 4.8: Deformation for the line crack model obtained by (a) fine-scale approximation and (b) ISBFM-DSR ($k = 40$). The blue circles denote the undeformed configuration and the red dots denote the deformed configuration (The displacement solution is magnified by 5×10^3 times for visualization)..... 89

Figure 4.9: (a) Model of the cracked beam Poisson's problem; (b) fine-scale discretization. 91

Figure 4.10: Fine-scale solutions u^h , du^h/dx , and du^h/dy along $y = 0.001$ for the cracked beam problem solved with $\hat{N} = 2$ 94

Figure 4.11: Absolute error distribution along $y = 0.001$ of the reduced-order solutions u^r , du^r/dx , and du^r/dy by scaled ISBFM-MA ($\gamma = 50$) for the cracked beam problem solved with $\hat{N} = 2$ under different ratios of reduced dimension k/N 94

Figure 4.12: Absolute error distribution along $y = 0.001$ of the reduced-order solutions u^r , du^r/dx , and du^r/dy by ISBFM-DR for the cracked beam problem solved with $\hat{N} = 2$ under different ratios of reduced dimension k/N 95

Figure 4.13: Absolute error distribution along $y = 0.001$ of the reduced-order solutions u^r , du^r/dx , and du^r/dy by ISBFM-DSR for the cracked beam problem solved with $\hat{N} = 2$ under different ratios of reduced dimension k/N 95

Figure 4.14: Normalized error in L_2 norm (e_0) of the reduced-order solutions of ISBFM-MA under different scaling for the loading crack model with $\hat{N} = 2$. ISBFM-DR and ISBFM-DSR are also given for comparison. 96

Figure 4.15: Error distribution in Euclidian norm of the reduced-order solutions for the loading line crack model with the reduced dimension $k = 80$ ($k/N \approx 3.1\%$). ISBFM-MA ($\gamma = 50$) and ISBFM-DSR are compared under different orders of enrichment functions $\hat{N} = 1, 2, 5$ 99

Figure 4.16: Comparison of the fine-scale solution and the reduced-order solutions of u_2 along $x = 49.65$ mm for the loading line crack model are shown in the left column. The corresponding error distributions of the reduced-order solutions against the fine-scale solution near the crack tip region (denoted by the red box) are shown in the right column. 101

Figure 4.17: Comparison of the fine-scale solution and the reduced-order solutions of du_2/dy along $x = 50.35$ mm for the loading line crack model are shown in the left column. The corresponding error distributions of the reduced-order solutions against the fine-scale solution near the crack tip region (denoted by the red box) are shown in the right column. 102

Figure 4.18: Schematic model of the mixed-mode crack problem. The adopted material properties are $E = 10,000$ psi and $\nu = 0.3$ 103

Figure 4.19: Normalized error in L_2 norm (e_0) of the reduced-order solution for the mixed-mode crack model solved with $\hat{N} = 1$. ISBFM-MA ($\gamma = 50$) and ISBFM-DSR are compared under different ratios of reduced dimension k/N 105

Figure 5.1: (a) Two-dimensional manifold structure represented by the three-dimensional S-curve data set; (b) Two-dimensional embedding obtained by PCA; (c) Two-dimensional embedding obtained by a nonlinear dimensionality reduction technique, e.g. LSTA (Zhang & Zha 2004). 118

Figure 5.2: (a) Original dataset (contains two outliers) and the first principal component directions of POD (blue arrow) and LGE (red arrow). (b) The reconstruction solution by POD using one basis (the reconstruction error $\varepsilon_r = 3.57 \times 10^{-2}$). (c) The reconstruction solution by LGE using one basis, the inverse distance weighting with $p = 2$ and $k_n = 20$ (the reconstruction error $\varepsilon_r = 3.09 \times 10^{-2}$). 130

Figure 5.3: Function g^2 has a sharp peak in one of the two corners of the spatial domain. 133

Figure 5.4: The normalized singular value $\sigma_i^2 / (\sum_i \sigma_i^2)$ (left) and the first six corresponding POD basis vectors (right) of the of the snapshot matrix \mathbf{X}_s corresponding to the nonlinear function g^2 134

Figure 5.5: The distribution of first 12 DEIM points (left) selected by the greedy algorithm for the nonlinear function g^2 and the comparison of average relative errors of POD and DEIM approximations (right) for the training data. 134

Figure 5.6: The normalized singular value $\sigma_i^2 / (\sum_i \sigma_i^2)$ (left) and the first six corresponding LGE basis vectors (right) of the of the matrix $\mathbf{X}_s \mathbf{L} \mathbf{X}_s^T$ corresponding to the nonlinear function g^2 , where the quadratic inverse distance weighting and $k_n = 40$ are adopted. 135

Figure 5.7: The distribution of first 12 DEIM points selected by the greedy algorithm for the nonlinear function g^2 based on the LGE projection using inverse distances (a) $p = 1$ and (b) $p = 2$ as the weight function. 138

Figure 5.8: The new test data obtained by evaluating the nonlinear function g^2 at the (128 points) parameter distribution (left) is used to compare the performance (right) of POD and LGE projections, as well as the corresponding DEIM approximation with different dimensions of basis.....	139
Figure 5.9: The distribution of first 20 DEIM points selected by the greedy algorithm based on (a) POD basis and (b) LGE basis using linearly inverse distances and $k_n = 20$ as the weight function.	140
Figure 5.10: (a) The average relative error of the reduced-order approximation solved by different reduced systems for the elliptic PDE. (b) The average CPU time of each reduced Newton iteration for different reduced systems, which are normalized by the average CPU time for each Newton iteration in the full-scale model.....	141
Figure 5.11: Snapshot of (a) the state solution $u(x,t)$ and (b) the nonlinear function $f(x,t)$ of the Burgers' problem obtained by a high-dimensional finite volume model of size $\mathcal{N} = 201$..	143
Figure 5.12: First six POD basis modes of (a) the state snapshot matrix \mathbf{X}_s and (b) the nonlinear snapshot matrix \mathbf{X}_f	144
Figure 5.13: First six LGE basis modes of the nonlinear snapshot matrix \mathbf{X}_f under different weight functions and numbers of neighbor points.....	145
Figure 5.14: Snapshot of the reduced-order approximation of the state solution $u(x,t)$ and the nonlinear function $f(x,t)$ of the Burgers' problem obtained by (a) POD-Galerkin, (b) DEIM, and (c) G-DEIM reduced-order systems.....	149
Figure 6.1: Sketch of the projection \hat{s}_α^* (the blue square) on a convex hull \mathcal{E}_α (the region is depicted by red dashed lines) of k -NN points (the solid circles in black) when a local state s_α (the red star) locates (a) inside and (b) outside \mathcal{E}_α . Neighbor points of $k = 6$ are used for demonstration.....	169
Figure 6.2: One-bar truss structure under uniaxial load.	174
Figure 6.3: Comparison of the DMDD and LCDD solvers for database with mild noise ($P = 100$, $\chi = 0.05$).	176
Figure 6.4: Comparison of the DMDD and LCDD solvers for database with strong noise ($P = 100$, $\chi = 0.15$).	177
Figure 6.5: Comparison of the DMDD and LCDD solvers for database with an outlier ($P = 100$).	178
Figure 6.6: Comparison of the DMDD and LCDD solvers for noiseless sigmoid database ($P = 100$).	181
Figure 6.7: Schematic of a 15-bar truss structure with essential boundary conditions and externally applied loads, where $a = 4$ m, $h = 2$ m, $\bar{u}_x = 0.01$ m, and $F = 100$ kN	181

Figure 6.8: Three noisy material databases with different data sizes $P=10^2, 10^3, 10^4$ for truss members.....	182
Figure 6.9: Truss structure case. (a) Convergence plot of the normalized RMS strain error $\varepsilon_{(\%RMS)}$ against increasing number of sampling points. (b) The number of convergence steps against increasing number of sampling points.....	183
Figure 6.10: Truss structure case. Convergence plot of the normalized RMS errors, $\varepsilon_{(\%RMS)}$ and $\sigma_{(\%RMS)}$, against increasing number of sampling points for data-driven solvers using (a) the data sets with Gaussian random noise and (b) the corresponding noiseless data sets.....	184
Figure 6.11: Truss structure case. Plots of the data-driven solutions of (a) DMDD and (b) LCDD when using the noisy material data set ($P=100$).....	186
Figure 6.12: Schematic of (a) Voronoi diagram and (b) discretization of a beam model.	187
Figure 6.13: Schematic of a beam model subjected to a shear load, where $L=48$ m, $H=12$ m and $F=1000$ N.	188
Figure 6.14: Exemplary material data sets ($P=10^3$) used for data-driven solvers. (a) Noiseless data set; (b) Data set with random noise.	190
Figure 6.15: Shear beam model with noiseless data sets. (a) Convergence plot of the normalized RMS state error $\omega_{(\%RMS)}$ against increasing number of sampling points. (b) The number of convergence steps against increasing number of sampling points.....	192
Figure 6.16: Shear beam model with noisy data sets. (a) Convergence plot of the normalized RMS state error $\omega_{(\%RMS)}$ against increasing number of sampling points. (b) The number of convergence steps against increasing number of sampling points.....	193
Figure 6.17: Shear beam model. (a) DMDD displacement solution using the noiseless data set $P=80^3$. (b) LCDD displacement solution ($k=6$) using the noiseless data set $P=10^3$	194
Figure 6.18: Shear beam model. (a) DMDD displacement solution using the noisy data set $P=80^3$. (b) LCDD displacement solution ($k=6$) using the noisy data set $P=10^3$	194
Figure 6.19: Shear beam model. The comparison of σ_{xx} and σ_{xy} stress components between the reference solution and the LCDD solutions with $k=6$ using the noisy data set $P=10^3$	195
Figure 7.1 Illustration of Voronoi diagram for SCNI.	204
Figure 7.2 Meshfree RVE model of two-phase hyperelastic material. The integration cells of matrix material and inclusion material are depicted by white and blue, respectively.	208
Figure 7.3 Interpolation points given by DEIM based on Meshfree model. The red star denotes the selected degree of freedoms and the black circle denotes the integration cell that contributes to the selected degree of freedoms.....	210
Figure 8.1: Schematic of a beam model subjected to a shear load, where $L=20$ mm, $H=1$ mm.	216

Figure 8.2: An exemplary noiseless material data set ($P = 20^3$) used for data-driven solvers: strain components. (Note: The axis labels denote the Green strain components)	217
Figure 8.3: Comparison of model-based (black), DDCM (blue) and LCDD (red) solutions, where w is the tip-deflection.	218
Figure 8.4: Schematic of a specimen of hyper-elasticity material subjected to uniaxial tension.	218
Figure 8.5: Comparison of analytical solution and the model-based simulation solution (RKPM).	220
Figure 8.6: Exemplary noiseless material data sets ($P = 20^3$) used for data-driven solvers: strain components (left); stress components (right). (Note: The axis labels denote the Green strain components (left) and the 2 nd PK stress components (right)).....	220
Figure 8.7: Comparison of the proposed data-driven ($k = 6$) displacement solution against the model-based simulation solution by using the noiseless dataset.	221
Figure 8.8: The absolute errors of stress components S_{xx} and S_{yy} solved by DMDD and LCDD against the model-based simulation solution.	222
Figure 8.9: Schematic of Mitral Valve (MV) and Tricuspid Valve (TV) leaflets for biaxial mechanical testing (Jett et al. 2018). Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.	223
Figure 8.10: Experimental Images of the Experimental Setups for Biaxial Mechanical Testing, where C denotes circumferential direction and R denotes radial direction (Jett et al. 2018). Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.	224
Figure 8.11: Experimental Photos of Digital Image Correlation Based Tracking – Strain Calculation (Jett et al. 2018). Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.	226
Figure 8.12: Schematic of the force-controlled biaxial mechanical testing protocol (Jett et al. 2018). Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.	226
Figure 8.13: Schematic of the displacement-controlled biaxial mechanical testing protocol. Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.	227
Figure 8.14: Experiment results for mitral valve anterior leaflet (MVAL) under 5 different force-controlled biaxial mechanical testing protocols. Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.	228
Figure 8.15: Stress components of the material dataset collected by biaxial testing experiments under the 5 force-controlled loading and 4 displacement-controlled loading. (Note: The axis labels denote the 2 nd PK stress components)	228
Figure 8.16: (a) Schematic of a specimen of heart valve tissue subjected to the 1 st displacement-controlled testing protocol and (b) the associated mechanics model for data-driven modeling.	229
Figure 8.17: Predictions of 2 nd PK Stresses by the proposed data-driven (LCDD) solver.	230

Figure 8.18: Comparison of the experiment data and data-driven prediction for mitral valve anterior leaflet (MVAL) under 1st displacement-controlled biaxial mechanical testing protocol. 230

Figure 8.19: Comparison of data-driven solutions using different number of force-controlled datasets for the numerical modeling of the mitral valve anterior leaflet (MVAL) under 1st displacement-controlled biaxial mechanical testing protocol. 232

LIST OF TABLES

Table 2.1: Comparison of data-driven and model-driven approaches (+: good performance; -: bad performance).....	32
Table 3.1: Greedy Algorithm (Chaturantabut & Sorensen 2010).....	52
Table 4.1: Normalized error in L_2 norm (e_0) of the reduced-order solutions for the cracked beam problem. ISBFM-MA, ISBFM-DR and ISBFM-DSR are compared under different ratios of reduced dimension k/N and different orders of enrichment functions $\hat{N} = 1, 2, 5$	93
Table 4.2: Normalized error in L_2 norm (e_0) of the reduced-order solutions for the loading crack model. ISBFM-MA, ISBFM-DR and ISBFM-DSR are compared under different ratios of reduced dimension k/N and different orders of enrichment functions $\hat{N} = 1, 2, 5$	98
Table 4.3: Relative errors of SIFs of the reduced-order solutions for the mixed-mode crack model solved with $\hat{N} = 1$. ISBFM-MA and ISBFM-DSR are compared under different ratios of reduced dimension k/N	104
Table 5.1: Laplacian Eigenmaps Algorithm (Belkin & Niyogi 2001).....	122
Table 5.2: Reduced order modeling of elliptic partial differential equations using linear graph embedding projection for DEIM.....	132
Table 5.3: The average relative error of the reconstruction of training data by the LGE approximation and the corresponding DEIM (G-DEIM) approximation using 12 projection basis under different weight functions and neighbor numbers k_n . (The reconstruction errors of the standard POD and DEIM approximation are also given as reference).....	136
Table 5.4: The average relative error of the reconstruction of training data by different LGE approximations and the corresponding DEIM (G-DEIM) approximations using $\hat{k} = 20$ projection basis and $\hat{n} = 30$. (The reconstruction errors of the standard POD and DEIM approximation are also given as reference)	147
Table 5.5: The average relative error of the reconstruction of training data by different LGE approximations and the corresponding DEIM (G-DEIM) approximations using $\hat{k} = 30$ projection basis and $\hat{n} = 40$. (The reconstruction errors of the standard POD and DEIM approximation are also given as reference)	147
Table 6.1: Flow of the proposed LCDD solver.....	173
Table 7.1: Material parameters for matrix and inclusion.....	208
Table 7.2: The accuracy and efficiency performance of POD and POD-DEIM reduction methods.	211

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my supervisor, Prof. Jiun-Shyan (J.S.) Chen, for his constant support, guidance, and enormous patience in the past five years. It has been my great honor to work with J.S. not just because of his profound knowledge in mechanics, mathematics, and numerical methods that benefits my understanding of research work, but also his insight and vision that always inspire me to come up with creative ideas. More importantly, his devotion and enthusiasm to research as well as ceaselessly pursuit of excellence have greatly impressed me and inspired me to become a better person.

Second, I acknowledge the members of my Ph.D. dissertation defense and candidacy exam committees, Prof. Yuri Bazilevs, Prof. David Benson, Prof. Michael Holst, Prof. Petr Krysl, Prof. Kenneth J. Loh, and Prof. Shantanu Sinha, for their suggestions and helps for my study and research. Thanks are extended to Prof. Pui-Shum Shing for serving on my qualification exam committee. I also want to take this chance to thank Prof. Randolph Bank for arranging a mathematics comprehensive exam for me during summer. Prof. Xanthippi Markenscoff is also acknowledged for inviting me to attend her course on shock wave and to a nice Chinese restaurant.

Although the life as a PhD student is never easy, I am fortunate to have wonderful group members around. I am deeply grateful to them: Jonghyuk Baek, Dr. Ramya Rao Basava, Frank Beckwith, Xiaolong He, Dr. Michael Hillman, Tsung-Hui Huang, Jacob Koester, Dr. Shih-Po Lin, Dr. Camille Marodon, Julio Martinez, Marco Pasetto, Dr. Marcus Rüter, Karan Taneja, Dr. Haoyan Wei, Dr. Edouard Yreux, Dr. Yantao Zhang, Dr. Guohua Zhou, and many others. I thank all these members for their help and friendship. Particularly, I would like to thank Dr. Haoyan

Wei and Marco Pasetto who have been generously supporting me during the past two months of thesis writing. I also want to express my sincere appreciation to Prof. Xiaowei Deng, Prof. Yong Li, Prof. Jinhui Yan, Prof. Yewei Zheng, Dr. Meng Wang, Dr. Zongliang Du, Prof. Jeff Wang, Dr. Jimmy Ni, and Ruize Ma for their friendship and sharing wonderful moments.

Special thanks go to Chang Liu and Sherlock.

Eight years has passed since I first read the book *A first course in continuum mechanics* written by Prof. Y. C. Fung in my last month of college, which really motivated me to pursue my research career. Life is so magical that I unexpectedly came to UC San Diego for my Ph.D. study where Prof. Fung wrote his book. I thank everyone that's been part of my life's journey.

At last, I will be forever indebted to my parents and sister for their endless love. Without their support, I could never have walked this far.

Chapter 4, in part, is a reprint of the materials as they appear in: “He, Qizhi; Chen, Jiun-Shyan; Marodon, Camille. A decomposed subspace reduction for fracture mechanics based on the meshfree integrated singular basis function method. *Computational Mechanics* (2018): 1-22”. The dissertation author is the primary investigator and author of this paper.

Chapter 5, in part, is currently being prepared for submission for publication of the material as it may appear in “He, Qizhi; Chen, Jiun-Shyan. Nonlinear model order reduction based on linearized manifold learning for meshfree analysis”. The dissertation author was the primary investigator and author of this material.

Chapter 6, in part, is currently being prepared for submission for publication of the material as it may appear in “He, Qizhi; Chen, Jiun-Shyan. A data-driven computational

simulation based on locally convex reconstruction for noisy database”. The dissertation author was the primary investigator and author of this material.

Chapter 7, in part, is currently being prepared for submission for publication of the material as it may appear in “He, Qizhi; Chen, Jiun-Shyan. Nonlinear model order reduction based on linearized manifold learning for meshfree analysis”. The dissertation author was the primary investigator and author of this material.

Chapter 8, in part, is currently being prepared for submission for publication of the material as it may appear in “He, Qizhi; Chen, Jiun-Shyan. Robust data-driven computational modeling of biological material”. The dissertation author was the primary investigator and author of this material. The Biomechanics and Biomaterials Design Laboratory and Dr. Chung-Hao Lee at the University of Oklahoma are greatly acknowledged for providing experiment images and measurement data.

VITA

2010	B.S. in Engineering Mechanics, Wuhan University
2010-2013	Research Assistant, Dalian University of Technology
2013	M.S. in Computational Mechanics, Dalian University of Technology
2013	Research Assistant, University of California, Los Angeles
2016	M.A. in Mathematics (Applied), University of California San Diego
2014-2018	Research Assistant, University of California San Diego
2018	Ph.D. in Structural Engineering, University of California San Diego with a Specialization in Computational Science

PUBLICATIONS

- [1] He, Q., Chen, J. S. (2018). Robust data-driven computational modeling of biological material. (in preparation)
- [2] He, Q., Chen, J. S. (2018). Nonlinear model order reduction based on linearized manifold learning for meshfree analysis. (in preparation)
- [3] He, Q., Chen, J. S. (2018). A data-driven computational simulation based on locally convex reconstruction for noisy database. (in preparation)
- [4] He, Q., Chen, J. S., & Marodon, C. (2018). A decomposed subspace reduction for fracture mechanics based on the meshfree integrated singular basis function method. *Computational Mechanics*, 1-22.
- [5] He, Q., Wei, H., Chen, J. S., Wang, H. P., & Carlson, B. E. (2018). Analysis of hot cracking during lap joint laser welding processes using the melting state-based thermomechanical modeling approach. *The International Journal of Advanced Manufacturing Technology*, 94(9-12), 4373-4386.
- [6] Wei, H., He, Q., Chen, J. S., Wang, H. P., & Carlson, B. E. (2017). Coupled thermal-mechanical-contact analysis of hot cracking in laser welded lap joints. *Journal of Laser Applications*, 29(2), 022412.
- [7] He, Q., Kang, Z., & Wang, Y. (2014). A topology optimization method for geometrically nonlinear structures with meshless analysis and independent density field interpolation. *Computational Mechanics*, 54(3), 629-644.
- [8] Wang, Y., Kang, Z., & He, Q. (2014). Adaptive topology optimization with independent error control for separated displacement and density fields. *Computers & Structures*, 135, 50-61.

- [9] Wang, Y., Kang, Z., & He, Q. (2013). An adaptive refinement approach for topology optimization based on separated density field description. *Computers & Structures*, 117, 10-22.
- [10] He, Q. Z., Hu, H., Belouettar, S., Guinta, G., Yu, K., Liu, Y., Biscani, Y., Carrera, E. & Potier-Ferry, M. (2011). Multi-scale modelling of sandwich structures using hierarchical kinematics. *Composite Structures*, 93(9), 2375-2383.

PRESENTATIONS

- [1] He, Q., Chen, J. S., A machine learning enhanced data-driven simulation of solids and structures with noisy database, Abstract, USACM Thematic Conference on Meshfree and Particle Methods: Applications and Theory, Santa Fe, New Mexico, September 10-12, 2018.
- [2] He, Q., Marodon, C., Chen, J. S., A Decomposed Subspace Reduction for Fracture Mechanics based on ISBFM Galerkin Meshfree Method, Abstract, The 13th World Congress in Computational Mechanics (WCCMXIII), New York City, NY, Jul. 22-Jul. 27, 2018.
- [3] He, Q., Chen, J. S., A machine learning enhanced data-driven approach for computational mechanics, Abstract, 2018 EMI Conference, M.I.T., Cambridge, MA, May 29 –June 1, 2018.
- [4] He, Q., Marodon, C., Chen, J. S., Meshfree Reduced Order Modeling of Fracture, Abstract, 2017 EMI Conference, San Diego, California, USA, 06/2017.
- [5] Wei, H., He, Q., Chen, J. S., Coupled Thermal-Mechanical-Contact Analysis of Hot Cracking in Laser Welded Lap Joints, 35th International Congress on Applications of Lasers & Electro-Optics (ICALEO), San Diego, California, 10/2016.
- [6] He, Q., Marodon, C., Chen, J. S., Meshfree Reduced Order Modeling of Fracture, USACM Conference on Isogeometric Analysis and Meshfree Methods, La Jolla, California, 10/2016.
- [7] He, Q., Marodon, C., Chen, J. S., Model Order Reduction for Meshfree Analysis of Fracture Problems, 13th US National Congress on Computational Mechanics (USNCCM), San Diego, California, 07/2015.

ABSTRACT OF THE DISSERTATION

Model Order Reduction and Data-Driven Computational Modeling for Linear and Nonlinear Solids

by

Qizhi He

Doctor of Philosophy in Structural Engineering with a Specialization in Computational Science

University of California San Diego, 2018

Professor Jiun-Shyan Chen, Chair

Physics-based numerical simulation remains challenging as the complexity of today's high-fidelity models has dramatically increased. Model order reduction (MOR) and data-driven modeling, based on the emerging techniques of data learning and physical modeling, present a promising way to tackle the computational bottleneck related to the computational intensity and model complexity.

Nevertheless, MOR has proven to be significantly more difficult for parameterized mechanics systems that exhibit a wide variety of parameter-dependent nonlinear behaviors or that involve localized essential features. The first objective of this work is to develop robust, physics-preserving MOR methods. As constructing a low-dimensional MOR model can be considered as the hybrid data-physics approach, one can optimize it through a learning process

using both data and physical models. As such, we first propose a MOR method based on decomposed reduced-order projections that well preserve the essential near-tip characteristic for fracture mechanics. Moreover, we develop an enhanced reduced-order basis to construct a low-dimensional subspace, deriving from a generalized manifold learning framework that allows the employment of local information in the data structure during the learning phase. This approach can yield a robust reduced-order model against noise and outliers and is well suited for parameterized nonlinear physical systems. Finally, a nonlinear MOR for a meshfree Galerkin formulation with the stabilized conforming nodal integration (SCNI) scheme is developed to yield a pure node based MOR that is particularly effective for hyper-reduction techniques. A numerical example of two-phase hyperelastic solid with perturbed loading conditions is used to validate the effectiveness of the proposed reduction method.

The second goal of the dissertation is to develop a robust data-driven computational framework, which provides an alternative to conventional scientific computing for complex materials. This framework aims at performing physical simulation by directly interacting with material data via machine learning procedures instead of employing phenomenological constitutive models, and especially addressing the robustness issue associated with noisy and scarce data. To this end, we propose to search data solutions from a locally reconstructed convex hull associated with the k-nearest neighbor points, which leads to robustness to noisy data and ensures convergence stability. The accuracy and robustness of the proposed data-driven approach are demonstrated in the modeling of linear and nonlinear elasticity problems. In addition, we present a preliminary result of data-driven modeling of biological tissue using material data collected from laboratory testing on heart valve tissue, showing the potential of data-driven simulation by integrating physical modeling and machine learning techniques.

Chapter 1

Introduction

In this chapter, the motivation for the present research work is introduced in Section 1.1, followed by the strategies and objectives in Section 1.2. An outline of this dissertation is given in Section 1.3.

1.1 Motivation

1.1.1 Skeletal muscle modeling

The human musculoskeletal system composed of about 700 muscles provides support, stability, and movement to the body. It plays a very essential role in people's daily life. However, the performance or functioning of skeletal muscles suffers from ageing and various diseases due to the change of properties, microstructures, and morphologies of muscle tissues. For example, as we age, we progressively lose muscle mass in a process termed sarcopenia (Matthews et al. 2011), and we may suffer from dynopenia (Goodpaster et al. 2006), that is, the muscle strength is disproportionately greater reduced during this process. It has reported that 50 million people are currently affected by sarcopenia and this will rise to 200 million in the near future (Wang & Bai 2012). As a commonly disease on skeletal muscles, deep tissue injury (also known as pressure ulcer (PU)) that is a localized degeneration of skin and underlying muscle tissue, arises in the muscle layers adjacent to bony prominences as a result of sustained loading (NPUAP/EPUAP 2009). PUs are usually found in situations when a prolonged mechanical load is applied to soft

biological tissues, such as subjects who are bedridden and wheelchair bound, or wearing a prosthesis or orthosis for prolonged periods (Stekelenburg et al. 2007). PUs not only severely affect the patient's quality of life since they are painful but also become a financial burden for the nation. There are 2.5 million pressure ulcers are treated in US and cost 11 billion for treatment per year (Sen et al. 2009). Thus, it is desirable to understand the roles of different musculoskeletal factors (such as properties, microstructures, and morphologies of muscle tissues) in the functioning of skeletal muscles as well as the physiological causes of the change in those factors. Computational simulation and modeling of skeletal muscle has been considered as a powerful tool to enhance our understanding of the mechanism underlying the etiology and to achieve real-time detection and prediction for guiding better treatments.

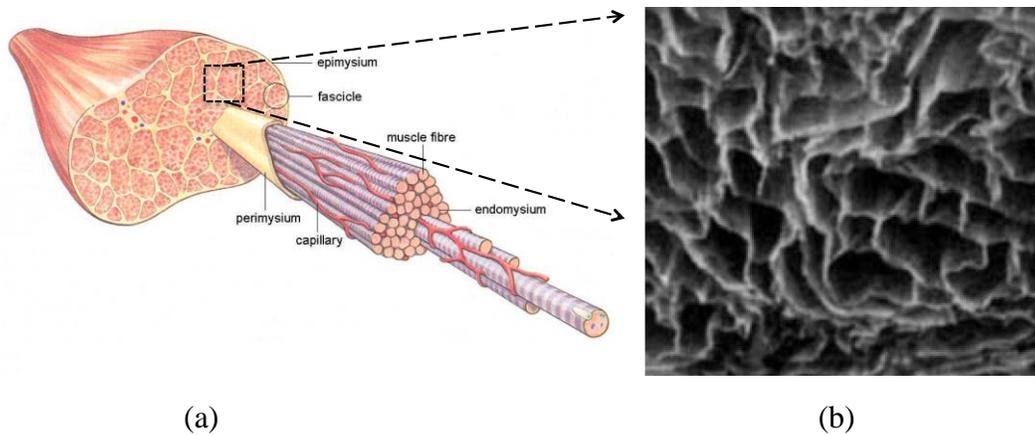


Figure 1.1: Skeletal muscle structure: (a) The entire muscle is surrounded by the epimysium, which develops as perimysium covering the fascicles, and finally as endomysium which surrounds the individual muscle fibers (Gray & Carter 1995); (b) the honeycomb-like structure of endomysium in fiber level (Gillies & Lieber 2011).

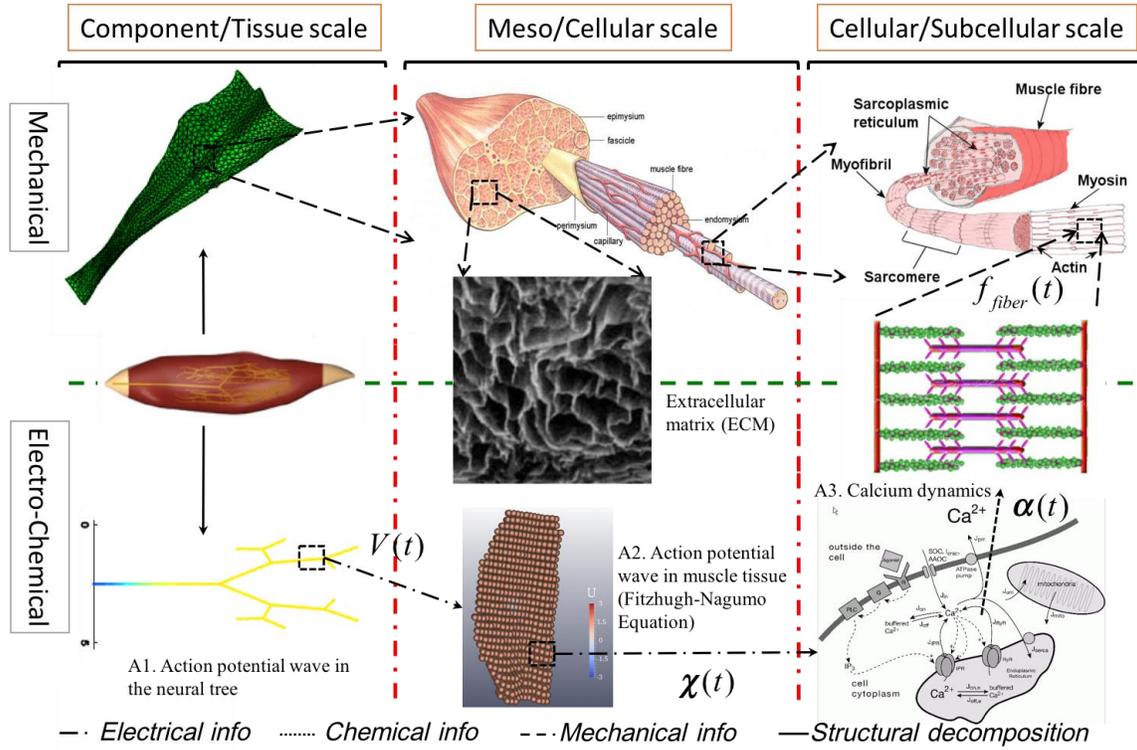


Figure 1.2: Description of multiscale and multi-physics simulation of skeletal muscle (Zhang 2015) ($V(t)$: action potential in neural system to yield an impulse; $\chi(t)$: pulse solved by the Fitzhugh-Nagumo equation to control the calcium concentration; $\alpha(t)$: activation factor solved by the calcium dynamics to control dynamic fiber force; $f_{fiber}(t)$: fiber force obtained from Huxley's cross-bridging model in sub-cellular scale).

A whole piece of muscle is usually composed of skeletal muscle tissue, connective tissues, nerve tissue and blood tissue or vascular tissue, as shown in Figure 1.1. To properly represent the across-scale process that electro-chemical excitation in neural system leading to mechanical contraction in muscle fibers, a coupled electro-chemical-mechanical multiscale computational framework based on image-based meshfree framework has been developed for modeling skeletal muscles (Basava 2015; Chen et al. 2016; Zhang 2015), and the proposed framework is summarized in Figure 1.2. Obviously, simulating skeletal muscles is extremely

complicated, computationally expensive, and sometimes even impractical due to the following Facts (i), (ii), and (iii):

- (i) The physics models governing electrical, chemical, and mechanical state variables are parameterized nonlinear partial differential equations (PDEs) that could exhibit a wide variety of parameter-dependent nonlinear behaviors. For example, the FitzHugh-Nagumo model (FHN) used to describe the propagation of potential actions in neural system is a nonlinear reaction-diffusion equation depending on an exterior electrical input parameter.
- (ii) The interact of different physical models arises at different length scales (subcellular, cellular, and component/tissue scales), and the parameter-dependent information need to transmit across scales. Thus, tedious upscale/downscale modeling is required. For example, computational homogenization methods (Feyel 2003; Kouznetsova, Brekelmans & Baaijens 2001) are used for multiscale mechanical analysis of muscle tissue, that can be considered as a simplified honey honeycomb-like structure (see Figure 1.1(b)) consisting of densely-distributed thin hyperelastic extensible fibers and a homogeneous tissue matrix (Humphrey & Yin 1987).
- (iii) Although numerous phenomenological and structural constitutive models (Chagnon, Rebouah & Favier 2015; Sacks & Sun 2003) have been developed, it remains challenging to effectively represent heterogeneous and anisotropic behavior of skeletal muscles with physical intuition and empiricism based models. Moreover, using complex material models brings about numerical difficulties to parameter identification and simulation in conventional numerical methods.

As a result, the potential of physics-based numerical simulation for providing a deeper understanding of complex skeletal systems and achieving time-critical prediction is restricted by the computational complexity and intensity.

1.1.2 Reduced-order modeling

To bridge the gap between the using high-fidelity models for accurate physics representation and the rapid time-to-solution demands of time-critical applications, model order reduction (MOR) has been considered as one promising strategy to tackle the computational burden problems related to the Facts (i) and (ii), given that most physical systems representing skeletal muscle behaviors are parameterized PDEs suited for reduced-order modeling (Benner, Gugercin & Willcox 2015). MOR is one approximation technique aiming at generating a low-dimensional, low-complexity model while accurately captures the map between system inputs (e.g., electro-chemical excitation, material parameters) and outputs represented by the original high-fidelity model (e.g., calcium concentration, force generation) in the regions of interest.

MOR has gained significant progress in a broad range of challenging engineering applications over the past two decades, including structures dynamics (Amsallem et al. 2009; Carlberg & Farhat 2011; Krysl, Lall & Marsden 2001; Lall, Krysl & Marsden 2003), system engineering (Antoulas & Sorensen 2001), fluid mechanics (Amsallem & Farhat 2008; Lieu, Farhat & Lesoinne 2006; Sirovich 1987), solid mechanics (Kerfriden et al. 2011; Millán & Arroyo 2013; Niroomandi et al. 2010; Radermacher & Reese 2016; Ryckelynck, Benziane & Paristech 2010), fracture mechanics (Akbari Rahimabadi, Kerfriden & Bordas 2015; He, Chen & Marodon 2018; Kerfriden et al. 2013), molecular dynamics (Lee & Chen 2013), and biological

modeling (Cueto & Chinesta 2014; Niroomandi et al. 2008; Rama & Skatulla 2018). Nonetheless, many outstanding challenges face the community, especially with respect to applying model reduction to parameterized nonlinear mechanics systems. The major challenges are listed as follows:

- **Applicability to nonlinear systems:** Since MOR methods usually realize computational savings by performing computations on a small subset of the underlying high-dimensional computational model when nonlinearities appear in the physical systems, one need to ensure the reduction procedures yielding sufficiently small error and applicable to original simulation code.
- **Physics preservation:** As most MOR methods are constructed based on spectral analysis, it is difficulty to preserve the essential characteristics of the physical system of interest in the reduced-order model, especially for the localized characteristics, e.g., singularities and discontinuities in fracture mechanics.
- **Robustness:** It is difficulty to ensure accurate and robust reduced-order solutions when a parameterized physical system exhibits a wide variety of nonlinear behaviors in terms of parameter changes.

Therefore, the development of a robust, physics-preserving MOR method for reduced-order modeling of parameterized PDEs is essential.

1.1.3 Data-driven computational modeling

Traditionally, the geometry model, physical models, and constitutive models that relate the kinematic and kinetic variables of the system are used to perform numerical simulation and predict a system response. However, one long-lasting debate is whether the constitutive models based on physical intuition and empiricism is still a good representation, especially for material that exhibits across-scale complex behaviors and sample-wise randomness.

On the other hand, with the significant advances in digital technologies and data-driven algorithms, and the proliferation of high-resolution datasets over the past decades, an emerging idea to enhance traditional scientific computing and engineering design procedures is becoming possible. This idea relies on utilizing both data and physical models simultaneously to provide more accurate representations of engineering or physical systems. Thus, it fosters an interdisciplinary research that fuses data processing, machine learning, and scientific computing that are apparently disparate research fields.

As machine learning models (Jordan & Mitchell 2015) have attained spectacular success in fields such as pattern recognition, language translation, and bioinformatics, it is natural to conjecture that we could abandon the traditional ways of physical-intuition driven model creation and instead use data directly for scientific simulation. Although the question is under debate, preliminary attempts to use inference and machine learning have been reported in many engineering and physical applications (Brunton, Proctor & Kutz 2016; Peherstorfer & Willcox 2016; Raissi & Karniadakis 2018; Raissi, Perdikaris & Karniadakis 2017; Schmidt & Lipson 2009; Williams, Kevrekidis & Rowley 2015).

We are particularly interested in addressing the difficulties caused by the employment of conventional constitutive models related to Fact (iii) in Section 1.1.1, by adopting the idea of blending data-driven models with the existing well-accepted mechanics models (Ibañez et al. 2016; Kirchdoerfer & Ortiz 2016; Lefik & Schrefler 2003; Matouš et al. 2017). It allows to perform simulation directly based on material data, bypassing phenomenological material models, and avoiding the associated issues of model construction, identification, and implementation. However, this novel but undeveloped computational paradigm of model-data fusion, called data-driven computational modeling, may suffer from lack of robustness and the “curse of dimensionality” as well as other data challenges due to the primitive data-driven techniques and the random nature of experimental data. Nevertheless, current research on the development of robust data-driven computational paradigm is still rare and deserves investigation.

1.2 Strategies and Objectives

The objective of this work is to further advance the current state of hybrid data-physics approaches in computational mechanics to address the computational limitations when dealing with nonlinear mechanics system and complex constitutive models, as presented in Section 1.1.2 and 1.1.3. This dissertation, thus, is divided into two parts, developing physics-preserving, robust MOR methods for parameterized PDEs, and data-driven computational framework for complex material modeling.

To accomplish this objective and avoid the drawbacks of currently used methods enumerated above, dimensionality reduction as well as other machine learning techniques are

introduced. Hence, these two topics are organically related since both are fundamentally based on the data learning and data-driven techniques. The specific developments are summarized as follows:

- 1) Development of an SVD-based MOR approach for reduced-order modeling of fracture mechanics that allows the preservation of essential characteristic, singularities and discontinuities of the original full-order model. In this approach, the reduced-order model is constructed based on the integrated singular basis function method (ISBFM) with meshfree approximation enriched by crack-tip basis functions. This framework allows a lower order integration of the Galerkin equation, but also yields a discrete system containing sparse sub-matrices for effective MOR procedures. As such, a decomposed reduced-order basis is developed to preserve the discontinuity and singularity characteristics of fracture. Several numerical examples are presented to examine the effectiveness of the proposed method for fracture modeling.
- 2) Development of a robust reduced-order model for parameterized nonlinear systems characterized by a wide variety of nonlinear behaviors in terms of parameter changes. The reduced-order basis used to construct the low-dimensional subspace is derived from a generalized manifold learning framework, called graph embedding, in conjunction with linearization techniques. This general framework allows to take a priori statistical knowledge of given data into account during the data learning phase. Thus, it yields a robust reduced-order model less sensitive to noise and outliers and well suited for nonlinear physical systems.
- 3) Development of nonlinear MOR for a meshfree Galerkin formulation based on the stabilized conforming nodal integration (SCNI) scheme, which has been demonstrated

effective for image-based modeling skeletal muscles. To extend the standard proper orthogonal decomposition (POD) method to nonlinear solid mechanics, an additional system approximation level achieved by interpolation-type methods based on the SCNI discretization is proposed. The procedures of applying the nonlinear MOR method for quasi-static analysis based on meshfree framework are presented and its effectiveness is demonstrated by solving a two-phase hyperelastic solid with perturbed loading conditions.

- 4) Development of a robust data-driven computational framework for small-deformation elastostatics problems, termed locally convex data-driven (LCDD) computing, that integrates machine learning techniques. This framework aims at performing physical simulation directly interacting with material data instead of employing constitutive models, especially tackling the robustness issue associated with such data sets of noise, outliers, and high-dimension. Specifically, LCDD searches for optimum data solutions from a locally reconstructed convex hull associated to the k -nearest neighbor (k -NN) points, which leads to less sensitivity to noisy data and ensures convergence stability. The numerical examples of structural and elasticity problems demonstrate the performance of the proposed LCDD method.
- 5) Development of the proposed robust data-driven computational framework for finite kinematics such that it allows to simulate biological materials under the same meshfree framework used for reduced-order modeling. The effectiveness of the proposed data-driven solver for nonlinear solids is verified by using two academic tests with synthetic material data. A preliminary study of applying data-driven simulation for a realistic

biological tissue with realistic material data collected by biaxial experiment testing is also given.

1.3 Outline

The remainder of this dissertation is organized as follows. In the next chapter, an overview on machine learning, model order reduction (MOR), and data-driven engineering science is given, along with the comparative discussion to show their intrinsic relation under the concept of model-driven and data-driven approaches. In Chapter 3, a detailed review of projection-based model order reduction techniques and their extensions to nonlinear mechanics systems is presented. In Chapter 4, a physics-preserving MOR method is proposed for fracture mechanics where solutions exhibit singularities and discontinuities. Moreover, a novel approach based on manifold learning to construct robust reduced order models for nonlinear partial differential equations is presented in Chapter 5. In Chapter 6, we develop a robust data-driven computational framework, which directly linking experimental data to physics laws for numerical simulation. In Chapter 7, a nonlinear MOR method for nonlinear solids is formulated under the SCNI (Stabilized Conforming Nodal Integration) framework. Furthermore, the proposed manifold learning-based data-driven solver is extended to nonlinear kinematics and applied to biological tissues modeling via directly utilizing material data measured by biaxial mechanical experiments, which are elucidated in Chapter 8. Finally, conclusions and discussions on future research directions are given in Chapter 9.

Chapter 2

Literature Review

We review several important concepts and methods related to machine learning, model reduction or surrogate modeling, and data-driven computational modeling. This chapter is, by no means, an exhaustive exposition of these topics. Its purpose is rather to provide enough background and the foundations for the developments of reduced-order modeling and data-driven modeling in the following chapters. Moreover, we provide a comparative discussion that lend insights to the intrinsic relation among them and potential advantages and disadvantages in applying each of the methods.

2.1 Machine Learning: Data Mining and Dimensionality Reduction

Machine learning has progressed dramatically over the past two decades, from laboratory curiosity to a practical technology in many real-world applications (Jordan & Mitchell 2015). It is motivated by the mission of extracting and processing information from massive data which challenges scientists and engineers in various fields such as data mining, image processing, speech recognition, computer vision, biological informatics, and medical diagnostics. Tasks of machine learning include regression, classification, clustering, dimension reduction, and feature selection. Machine learning algorithms aim at learning the underlying function relations (input and output) from collected data and making data-driven predictions or decisions through building a learning model from sample inputs. Due to the huge sample size and high dimensionality of

variables, achieving reliable machine learning models for real-life problems heavily relies on suitable statistical modeling and optimization methods to tackle the large-size problems.

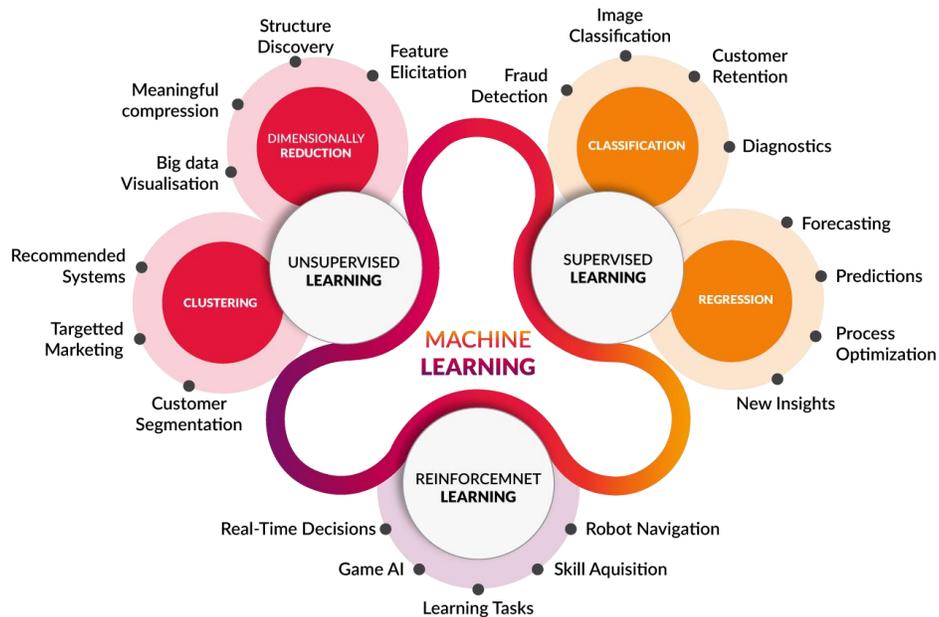


Figure 2.1: Types of machine learning. (<http://www.cognub.com/index.php/cognitive-platform/>)

In terms of tasks, there are essentially two main types of machine learning problems (see Figure 2.1): supervised learning, aiming to learn a mapping from inputs to target values, and unsupervised learning, whose goal is to find “interesting patterns” in the data. They are sometimes called data mining and knowledge discovery, respectively. Obviously, unsupervised learning is a much less well-defined problem since we have no idea of what kinds of patterns to look for and which error metric and criterion should be used. Machine learning is a huge research area and involves a very broad tasks ranging from data processing, learning, to interpretation of the results. We are only interested in the learning stage where systems are trained on the given data samples to generalize to new, unseen samples. In this section, we briefly review the basic concept of supervised and unsupervised learning and provide a few

learning methods that are used in the following sections. For a comprehensive understanding of this field, we refer (Bishop 2006; Hastie, Tibshirani & Friedman 2009; Murphy 2012).

2.1.1 Supervised learning

Given an input space $\mathcal{X} \subset \mathbb{R}^d$ for a learning problem, there is an underlying function $g: \mathcal{X} \rightarrow \mathcal{Y}$ mapping an element $\mathbf{x} \in \mathcal{X}$ to a point in the output space \mathcal{Y} . Assume we have sampled the function at the data points $\{\mathbf{x}_1, \dots, \mathbf{x}_M\} \subset \mathcal{X}$ and measured the corresponding outputs $\{y_1, \dots, y_M\} \subset \mathcal{Y}$, where the sampling process could be disturbed by noise. A supervised learning algorithm is to find a function $\hat{g}: \mathcal{X} \rightarrow \mathcal{Y}$ approximating the unknown function g to make prediction on unseen samples. The function \hat{g} is obtained by considering only the training data set $\mathcal{S} = \{(\mathbf{x}_i, y_i), i = 1, \dots, M\} \subset \mathcal{X} \times \mathcal{Y}$, but it should generalize to new data. The set of outputs y_i might be a set of quantitative measurements, i.e. $\mathcal{Y} \subset \mathbb{R}$, for regression problems (we consider only univariate regression for demonstration, but it is easy to extend to multivariate regression), or a finite set of qualitative labels or classes, i.e. $y_i \in \{1, \dots, C\}$ (e.g., species of fish), for classification problems, with C being the number of classes. Compared to quantitative outputs used in regression, there is no explicit ordering in the qualitative outputs, and they are discrete variables.

The prediction accuracy of a function \hat{g} can be measured quantitatively by means of a loss function: $L: \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}_+$ as $L(\hat{f}(\mathbf{x}), y)$ for an input-output pair $\{(\mathbf{x}, y) \in \mathcal{X} \times \mathcal{Y}\}$, where \mathbb{R}_+ denotes a set of non-negative real numbers. For regression, the *squared error loss*

$L(\hat{f}(\mathbf{x}), y) = (\hat{f}(\mathbf{x}) - y)^2$ is most often used and it gives the least squares error when the predicted value $\hat{f}(\mathbf{x})$ approximates the true output value y . Thus, one very fundamental type of supervised learning algorithm can be stated as empirical risk minimization (ERM):

$$\hat{f} = \arg \min \frac{1}{M} \sum_{i=1}^M L(\hat{f}(\mathbf{x}_i), y_i). \quad (2.1)$$

The extension of the standard ERM in (2.1) can be achieved by using regularization methods and reproducing kernel Hilbert spaces (RKHS) (Hastie, Tibshirani & Friedman 2009). For example, a general class of regularization problems has the form

$$\hat{f} = \arg \min_{\hat{f} \in \mathcal{H}} \frac{1}{M} \sum_{i=1}^M L(\hat{f}(\mathbf{x}_i), y_i) + \lambda J(\hat{f}), \quad (2.2)$$

where $\lambda > 0$ is a regularization parameter, $J(\hat{f})$ is a penalty function, and \mathcal{H} is a space of functions on which $J(\hat{f})$ is defined. It is noted that the regularized problem in (2.2) is a powerful algorithm for learning, and constitutes the building block for many advanced learning methods, such as multilayer neural network (Haykin 2009) and support vector machines (SVM) (Hastie, Tibshirani & Friedman 2009). Since it is very technical and beyond the scope of this these, we refer the interested readers to (Bishop 2006; Hastie, Tibshirani & Friedman 2009) and the references therein.

A simple but powerful prediction method is the linear model fit by the least squares. That is, the function \hat{f} is defined by a linear model

$$\hat{f}(\mathbf{x}) = \mathbf{x}^T \hat{\boldsymbol{\beta}}, \quad (2.3)$$

where $\hat{\boldsymbol{\beta}} \in \mathbb{R}^d$ are the coefficient vector to be solved. Then, inserting (2.3) into the ERM system (2.1), the coefficient $\hat{\boldsymbol{\beta}}$ can be obtained by solving a standard linear least square problem.

***k*-nearest neighbor (*k*-NN)**

Another powerful classification and regression technique but based on totally different philosophy from the linear model is the *k*-nearest-neighbor (*k*-NN) method (Cover & Hart 1967). It is one of the most fundamental non-parametric methods in pattern recognition and machine learning. Owing to its simplicity and flexibility, *k*-NN procedures had become the methods of choice in many scenarios (Anava & Levy 2016; Wu et al. 2008), especially in the case where the underlying model is complex. Moreover, it is fairly robust to errors given good cross-validation procedures (Ni & Nguyen 2009).

To approximate $\hat{g}(\mathbf{x})$, the *k*-NN prediction method simply uses those observations in the training set \mathcal{S} closest to \mathbf{x} in the input space \mathcal{X} to form $\hat{f}(\mathbf{x})$, that is

$$\hat{f}(\mathbf{x}) = \frac{1}{k} \sum_{\mathbf{x}_i \in \mathcal{N}_k(\mathbf{x})} y_i, \quad (2.4)$$

where $\mathcal{N}_k(\mathbf{x})$ is the neighborhood of \mathbf{x} defined by *k* nearest points \mathbf{x}_i in the training set. Compared to the linear model, the *k*-NN procedure does not rely on any stringent assumptions about the underlying data. Usually, a successful application of the *k*-NN algorithms requires a careful choice of the number of nearest neighbors *k* and the distance metric.

Remark 2.1: Following the discussion under the statistical perspective (Hastie, Tibshirani & Friedman 2009), both the linear model with least squares and the *k*-NN method

result in approximating conditional expectations by averages while their model assumptions are dramatically different. The former one assumes a global linear function to approximate $\hat{g}(\mathbf{x})$ whereas the latter one conjectures it can be well approximated by a locally constant function. Moreover, under mild regularity conditions on the joint probability distribution $P_r(\mathcal{X}, \mathcal{Y})$, one can show k -NN is a universal approximator for large training sample size.

However, k -NN suffers from the “curse of dimensionality (Bellman 1961)”. As the dimension d gets large, the metric used to measure neighborhood becomes problematic and the rate of convergence decreases significantly.

2.1.2 Unsupervised learning

We now consider unsupervised learning, where the training data contains only data points and no target values. The goal is to discover the hidden structures in the data. There are three major applications of unsupervised learning, including density estimation, clustering, and dimensionality reduction. We briefly review dimensionality reduction as it is closely related to the work of the thesis, and refer to (Murphy 2012) for the other two topics.

Linear and nonlinear dimensionality reduction

When dealing with high-dimensional data, it is desired to reduce the dimensionality by projecting the data to a lower dimensional subspace and find the latent variables that characterize the intrinsic structure of data. Although dimensionality reduction itself is an unsupervised problem, it plays an important role in data-mining, computer vision, and machine learning to

circumvent the “curse of dimensionality” and reduce computational burdens. Dimensionality reduction algorithms are usually based on spectral analysis (Zhang et al. 2009) to generate the mappings between the high- and low-dimensional spaces. Depending on whether the resulting mapping is linear or not, the algorithms can be classified into two groups: conventional linear dimensionality reduction algorithms and manifold learning-based algorithms.

Among the linear algorithms, principal component analysis (PCA) (Jolliffe 2002; Martinez & Kak 2001) and linear discriminant analysis (LDA) (Martinez & Kak 2001) are the two most popular ones and both are proposed based on Gaussian assumptions on the data distributions. PCA finds the low-dimensional embedding by constructing a linear representation of largest possible variance. This is a linear approach because it is a linear transformation from high-dimensional data to the embedding.

Assuming that data points are distributed in a nonlinear pattern, manifold learning techniques (also called nonlinear dimensionality reduction), such as kernel-PCA (Schölkopf, Smola & Müller 1998), locally linear embedding (LLE) (Roweis & Saul 2000), Laplacian eigenmap (LE) (Belkin & Niyogi 2001), and local tangent space alignment (LSTA) (Zhang & Zha 2004), among many other approaches, have been proposed to discover the low-dimensional manifold and remove correlations in data (Lee & Verleysen 2007; Van Der Maaten, Postma & Van Den Herik 2009). To construct the global low-dimensional representation without explicitly formulating the (nonlinear) mapping, most of these techniques are based on clustering neighbor data points and approximating the geometry (pairwise distances, local linear relationships, etc.) of the associated local manifold. Compared with linear approaches that usually apply global spectral analysis, local spectral analysis is used for manifold learning.

The techniques for dimensionality reduction have also been widely applied to many research areas in engineering science, such as model order reduction (Amsallem & Farhat 2011; Bhattacharjee & Matouš 2016; González et al. 2018; Millán & Arroyo 2013), data-driven computational mechanics (Ganapathysubramanian & Zabarar 2008; Ibañez et al. 2016) and computational materials engineering (Lopez et al. 2018).

2.2 Model Reduction and Surrogate Modeling for Time-Critical Prediction

Many modern mathematical models of real-world applications pose challenges when used in numerical simulations, due to the inherent complexity and large-scale (dimension) nature. Model reduction is a general strategy to construct a simplification of the original large-scale model, aiming to reducing solution time and reduce the computational burden, yet accurately preserve the essential behavior and dominant effects of the high-fidelity physics.

In fact, model reduction falls into a more general research area, called surrogate modeling, where the system complexity is alleviated by constructing approximation modes. Surrogate modeling is widely used to achieve *time-critical prediction* in applications such as design, control, optimization, and uncertainty quantification (see Figure 2.2). Specifically, the time-critical prediction is usually required in the following two relevant scenarios:

1. *Real-time or near-real-time simulation*: the objective is to compute outputs in a time less than certain critical value. Examples include embedded control and patient-specific simulator (Cueto & Chinesta 2014; González, Cueto & Chinesta 2016).
2. *Many-query applications*: the objective is to minimize the computational resources for each evaluation such that we can evaluate the model at as many points in the input

parameter space as possible given a fixed amount computing resource. Examples include uncertainty quantification, optimization problems, and design exploration (Amsallem et al. 2015).

As summarized in (Benner, Gugercin & Willcox 2015; Eldred & Dunlavy 2006), surrogate modeling is usually categorized into three primary strategies: hierarchical (physics simplification) type, data-fit type, reduced-order modeling. Based on the extent of exploitation of domain knowledge and data, the first one is considered as a physics-based approach, and the second one is a purely data-driven approach that directly constructs a model for the input–output map based on existing data using learning algorithms such as supervised learning. The last one, reduced-order modeling, can be viewed as a mixture of data-driven and physics-based approaches.

In this section, we first review the three general strategies of surrogate modeling. Then, we provide more discussion of reduced-order modeling in Chapter 3 as it is the main topic of the thesis.

2.2.1 Hierarchical model

Hierarchical surrogate models essentially are the corresponding low-fidelity models derived from high-fidelity models using approaches such as simplifying physics assumptions or neglecting physics, lower order or alternative basis approximation, coarser discretization or larger time step, and looser residual tolerances. It is also referred as to simplified physics approach, analogous to the traditional mathematical modeling approach. Thus, this approach is

not often the topic of discussion in the context of model reduction as it is a long-standing method in science, engineering and mathematics.

Hierarchical models have been applied to design optimization systems (Booker et al. 1999; Lewis & Nash 2000, 2005; March & Willcox 2012), solving linear inverse problem (Arridge et al. 2006), and so on. Because lower-fidelity models remain physics-based, this type can be expected to generate more accurate predictions than data-fitting techniques and not suffer from the curse of dimensionality. Moreover, this type does not require an offline training stage (refer to Section 2.2.3) to generate solution data for data learning since this approach is independent of system inputs. Thus, the simplified physics approach can also be categorized into the *a priori* approach that utilizes only existing domain knowledge.

However, this hierarchical type suffers from a major drawback that the resulting surrogate model is very difficult to achieve a reasonably accuracy with orders of magnitude speedup by simply neglecting some physics or coarsening the discretization (Carlberg 2011). Another challenge is how to maintain the physical consistency between low- and high-fidelity modes is not always clear, especially considering the variety of input. It requires a lot of domain knowledge that may not be available in many real-world problems.

2.2.2 Data-fit model

On the other hand, the data-fit type is a pure data-driven approach that applies supervised learning methods (interpolation, regression, or classification) to simulation data to fit a model for system output as a function of the input parameters and bypass the computation of the internal state entirely (Forrester 2008; Gorissen et al. 2010; Xiao et al. 2016), such as

response surface methods, black-box system identification techniques (Leontaritis & Billings 1985). In the statistical literature, Gaussian processes (Kennedy & O'Hagan 2001), polynomial response surfaces (Venter, Haftka & Starnes 1998), radial basis functions (Wild, Regis & Shoemaker 2008), Kriging models (Simpson et al. 2001), stochastic spectral approximations (Ghanem & Spanos 1991; Xiu & Karniadakis 2002), and other statistical and Bayesian inference techniques, have been used extensively as data-fit surrogates for complex computational models (Benner, Gugercin & Willcox 2015). Recently, the neural network-based techniques (Hambli, Katerchi & Benhamou 2011; Lefik & Schrefler 2003; Milano & Koumoutsakos 2002; Oishi & Yagawa 2017; Tracey, Duraisamy & Alonso 2015) are also receiving increasing attention due to its universal approximation and generalization properties.

The key feature of the data-fit type is that the online evaluations run in “black-box” mode, regardless of the underlying problem itself. Thus, this approach is code non-intrusive and able to achieve very low-complexity data-fitting models for problems where the input-output relation is relatively linear and continuous. However, the data-fit approach loses the flexibility to extract more known attributes of the system since it is ‘blind’ to the physics of the problem after the surrogate is built. Due to the ignorance of physics it can also lead to unacceptable prediction errors when the physics is sensitive to the input parameters, for example, handling dynamical evolution with a data-fit model is more challenging. Additionally, this approach may suffer from the curse of dimensionality (Queipo et al. 2005).

2.2.3 Reduced-order model (ROM)

The third type of surrogate modeling is reduced-order modeling (see Figure 2.2), which is based on a mathematical tool often called model order reduction (MOR). MOR seeks to achieve near real-time prediction by projecting the high-dimensional model (HDM) onto an associated low-dimensional subspace that spanned by properly constructed bases and solving the resulting low-dimensional system on that subspace. As the low-dimensional system is constructed by reducing the physical model's associated state dimension or degrees of freedom, it is commonly referred to as a reduced-order model (ROM). The terms reduced-order modeling and model order reduction are used interchangeably in literature.

On one hand, reduced-order modeling is a data-driven approach because the dimension reduction is achieved by learning the data usually collected from the precomputation of the HDM (see online/offline splitting). But it is also physics-based since ROM is built upon the governing equations that contain information about the underlying physics in terms of the input parameters. Thus, this hybrid approach is expected to generate online predictions of higher accuracy and robustness than the data-fit models that is a pure data-driven approach, especially for systems with a large variety of input parameters and rapidly changing physics within their range of applicability. In addition, it enables physical insight into the surrogate system and makes it more understandable to users compared to the data-fit system. Moreover, reduced-order modeling is potentially able to achieve much greater speedups than using hierarchical models (Carlberg 2011). If projection-based approaches are adopted for constructing ROM, i.e. projection-based MOR methods in Section 3.3, another advantage is that they permit a rigorous framework for deriving system-theoretic error bounds and error estimates (Benner, Gugercin & Willcox 2015; Hesthaven, Rozza & Stamm 2015; Quarteroni, Manzoni & Negri 2015).

However, one major disadvantage in the reduced-order models is that they tend to be code-intrusive since they typically require access to the system operators (e.g. residual, Jacobian) for the problem (refer to Chapter 3). Besides, reliable reduced-order models are often non-trivial to generate, especially in the case of highly nonlinear systems. Moreover, special techniques such as system approximation are needed for the extension to nonlinear systems, as shown in the next chapter.

Nevertheless, reduced-order modeling offers a promising mean to handle time-critical problems while preserving essential physics (Eldred & Dunlavy 2006). This thesis focuses on this approach and gives further review in Chapter 3. Before going to the next Chapter, some basic concepts for reduced-order modeling are given as follows.

Online/offline splitting

Generally, most MOR techniques admit an *offline-online* splitting, as shown in Figure 2.2. In the offline phase it might require solving the HDM for some chosen parameters to produce sufficient solution data, and the desired reduced-order basis and other reduced quantities used for constructing the associated ROM are computed based on the collected data. Apart from solving the HDM, in the offline stage the computation of determining the “suitable” parameters to be tested (known as sampling) is also an important issue in reduced-order modeling. Although the offline phase is computationally expensive, it is typically acceptable to consume time and effort constructing a model before deploying it in the online simulation. In the online phase, a rapid evaluation with new parameters can be obtained by the ROM. Under the context of machine learning, the offline phase is to generate training data and to construct the statistical

model learned from data, whereas the online phase is to perform the prediction based on test data.

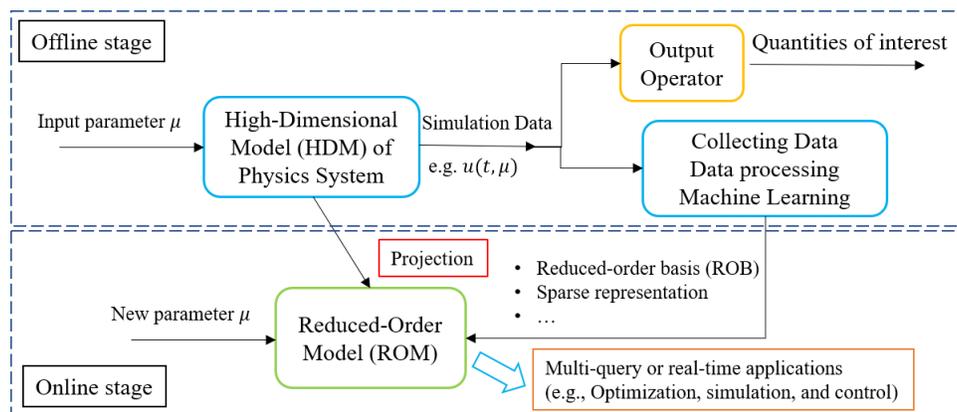


Figure 2.2: Schematic of general projection-based model order reduction for parameterized physical systems.

The *a priori* and *a posteriori* approaches

Depending on whether the offline stage is used to construct reduced basis and reduced model, MOR methods are categorized into two types (Galland et al. 2011): the *a posteriori* and *a priori* approaches. In the *a posteriori* approach, preliminary computations of the fine-scale model are performed in the *offline* phase to build the basis of a reduced subspace based on the precomputed solution data, and the reduced model constructed on the subspace is then used repeatedly in the *online* phase to obtain accelerated solution. The snapshot proper orthogonal decomposition (POD) method (Sirovich 1987), also called Karhunen-Loève expansion (Karhunen 1946; Loève 1955) or principal components analysis in other fields, and the reduced-basis method (RBM) (Prud'homme et al. 2002; Quarteroni, Rozza & Manzoni 2011; Rozza, Huynh & Patera 2008) are typical *a posteriori* approaches. On the contrary, the *a priori* approach

doesn't need the preliminary computations of solutions, and therefore no offline/online computational decomposition. The reduced basis is constructed by using some *a priori* knowledge about the particular problem. Examples include the *a priori* hyperreduction method (Ryckelynck 2005; Ryckelynck et al. 2006), the spectral decomposition (or known as model analysis) (Craig 1981; Dickens, Nakagawa & Wittbrodt 1997), component mode synthesis (CMS) methods (Hurty 1960, 1965), and the proper generalized decomposition (PGD) method (Chinesta, Ammar & Cueto 2010; Ladevèze, Passieux & Néron 2010; Nouy 2010) based on the reduced solution space by separation of variables.

Machine-learning enhanced reduced-order modeling

As pointed out before, reduced-order modeling intrinsically integrates the data-driven methodology as it applies statistical learning techniques to extract important physical features out of the precomputed data. Take the snapshot-based proper orthogonal decomposition (Sirovich 1987) as an example, a standard dimension reduction method equivalent to principal component analysis is used to construct low-dimensional solution subspace. While significant advances have been made in reduced-order modeling over the past fifteen years, there are still many outstanding challenges, especially with respect to applying MOR to parameterized nonlinear dynamical systems. To address these issues, some machine learning techniques like greedy algorithm (Barrault et al. 2004; Chaturantabut & Sorensen 2010), clustering (Peherstorfer et al. 2014), manifold learning (Millán & Arroyo 2013), and so on, were introduced to the MOR community.

Recently, the integration of data-driven or machine learning techniques with MOR has received increasing attention, yielding the potential to derive generalizable reduction models to

achieve time-critical prediction. Some exemplary studies are presented in (Bhattacharjee & Matouš 2016; Carlberg, Ray & van Bloemen Waanders 2015; Peherstorfer & Willcox 2015; Swischuk et al. 2018), to name a few. In (Swischuk et al. 2018), four different machine learning techniques (neural networks, multivariate polynomial regression, k -NN, and decision trees) were introduced into the proper orthogonal decomposition (POD), and the performance of prediction and preservation of physical constraints in their ROMs are compared.

More importantly, it is noted that the combination of data-driven techniques and physical models (data-model fusion) fosters a new appealing discipline in modern computational engineering science (Hambli, Katerchi & Benhamou 2011; Lefik & Schrefler 2003; Liu, Bessa & Liu 2016; Milano & Koumoutsakos 2002; Oishi & Yagawa 2017; Tracey, Duraisamy & Alonso 2015). In next section, we will briefly review the recent advances in data-driven techniques for physical modeling, and we can see a lot of similarities shared by both reduced-order modeling and data-driven modeling.

2.3 Data-Driven Engineering Science

The explosion of data has radically reformed the nature of science and engineering into data-rich environment in recent years. Data science together with methodologies such as machine learning (Hastie, Tibshirani & Friedman 2009; Murphy 2012) permit the extraction of “knowledge” and relevant insightful information from large volumes of unstructured data (Larose 2014). The machine learning techniques have been widely applied to computational biological and medical image analysis (Angermueller et al. 2016; Litjens et al. 2016, 2017). One of the most appealing applications in engineering is material informatics, which employs the

principles of informatics and machine learning regression techniques to obtain predictive quantitative models (Olson 2000; Rajan 2005). These approaches have been extended to the extraction of constitutive correlations between microscopic and macroscopic material properties in multi-scale modelling (Gupta et al. 2015; Kalidindi, Niezgodá & Salem 2011). Another emerging application of machine learning is the extraction of physical models or governing equations purely from the collected data of system (Brunton, Proctor & Kutz 2016; Schmidt & Lipson 2009). Machine learning methods have also been adopted to develop data-driven surrogate models in fluid mechanics (Milano & Koumoutsakos 2002; Tracey, Duraisamy & Alonso 2015; Zhang & Duraisamy 2015) and advanced constitutive models in solid mechanics (Ghaboussi, Garrett & Wu 1991; Ghaboussi & Sidarta 1998; Lefik & Schrefler 2003). In conjunction with machine learning techniques such as manifold learning (Lee & Verleysen 2007) or neural networks (Haykin 2009), the recent studies (Bhattacharjee & Matouš 2016; Le, Yvonnet & He 2015) offer a new paradigm for the design of materials at larger scales under the data-driven computational homogenization framework. There is a vast body of literature devoted to these subjects, including the recent developments based on nonlinear dimensionality reduction (Bhattacharjee & Matouš 2016), nonlinear regression, deep learning (Oishi & Yagawa 2017; Stoecklein et al. 2017; Wang & Sun 2018), among others.

2.3.1 Hybrid approach: data-enhanced physics-based modeling paradigm

However, in the area of simulation-based engineering sciences (SBES) (Oden et al. 2006) where the mathematical models and the corresponding numerical tools are well established, the potential of data science has yet to be fully explored while it is commonly considered to

constitute a potential change of paradigm in SBES. One of the key debates when developing data-driven techniques for SBES is to what extent domain knowledge can and should be introduced into the simulation algorithms. One school of thought in machine learning community is that even with very little domain knowledge the algorithm should be capable of extracting patterns or governing laws on its own given sufficient data (Brunton, Proctor & Kutz 2016; Ling, Jones & Templeton 2016; Michalski 1986; Raissi & Karniadakis 2018; Schmidt & Lipson 2009). This is because if the domain knowledge contains imperfect knowledge and model arbitrariness, it can lead to diminished performance in the data-driven model

Nevertheless, in many physical systems, well-accepted and definitive domain knowledge deriving from physical laws does exist in the form of rigorous mathematical theory. On the other hand, useful data in SBES is actually not as rich as in the fields such as computer vision, image process and speech recognition, and they are very expensive to acquire (Ibañez et al. 2017; Raissi & Karniadakis 2018). As such, purely data-driven techniques present difficulties when the data is scarce relative to the complexity of the system. Since data cannot be an alternative for physical modeling, it is necessary to develop data-driven simulation approaches that can leverage the physical principles with limited data for highly complex systems.

Overall, a pragmatic solution to develop effective predictive models for complex real-world problems is to combine physics-based models with data-driven techniques under a hybrid computational framework. There are three hybrid approaches proposed as the data-enhanced physics-based modeling paradigm. The first approach enforces known physical constraints into data-driven models (Ling, Jones & Templeton 2016). This belongs to the data-fit type surrogate models in Section 2.2.2. In the second approach, the existing physical models are enriched by the information learned from data. This general framework can be used for obtaining data enhanced

physical models (Koscianski & De Cursi 2005), online updating dynamical system in a manner similar to data assimilation (Evensen 2010), or performing model reduction (Bessa et al. 2017; Liu, Bessa & Liu 2016; Matouš et al. 2017; Peherstorfer & Willcox 2015). The third approach is to apply data-driven models and physical models separately to approximate different aspects of the physical system and be connected to perform numerical simulation, as illustrated in Figure 2.3. More details of this approach for computational mechanics are given as follows.

Data-driven computational modeling

In the context of computational mechanics, it is crucial to distinguish two very different types of knowledge (Ibañez et al. 2017; Kirchdoerfer & Ortiz 2016). The first one is related to conservation laws of physics (momentum, mass, energy, etc.) that are axiomatic or epistemic. The second one consists of empirical models (e.g., material constitutive laws) based on experimental observation. Traditional methods for developing and tuning empirical models usually combine physical intuition with simple regression techniques on limited data sets. But the empirical models inevitably involve incomplete experimental information (Kirchdoerfer & Ortiz 2016, 2017), and the process of material parameter identification (Avril et al. 2008; Ben Azzouna, Feissel & Villon 2015; Bonne & Constantinescu 2005) remains numerically intractable.

However, under the hybrid framework the material relation can be described by data-driven models while the well-established partial differential equations (PDEs) are still used to represent conservation laws (see Figure 2.3). As the second main topic of this thesis, we focus on this third hybrid approach in order to overcome the issues of employing phenomenological

models in computational mechanics, with an particular emphasis on improving model robustness and consistency when dealing with noisy and high-dimensional material data, which will be further elaborated in Chapter 6.

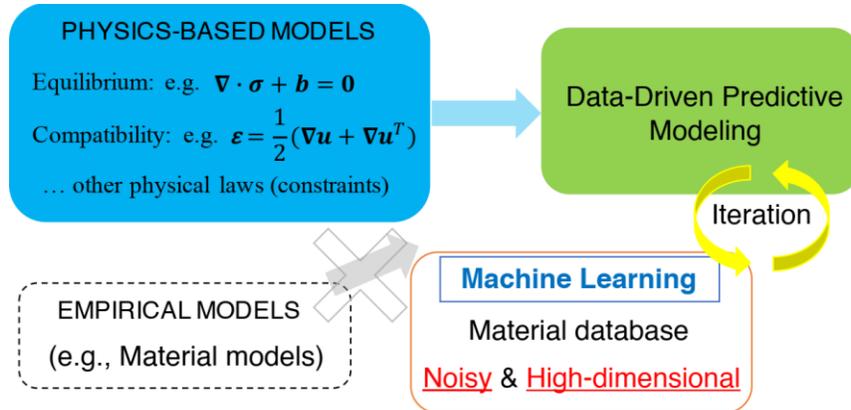


Figure 2.3: Schematics of data-driven computational mechanics, where data-driven models instead of empirical models are used together with physics-based models to perform simulation.

2.4 Summary of Reduced-Order and Data-Driven Modeling

This chapter provides a review of three general approaches for approximating a complicated simulation model: model-driven, data-driven, and the hybrid approaches. The model-driven approach uses the knowledge and empiricism about the physics of the phenomena involved, whereas the data-driven approach obtains a model representation by using data from observations. Apart from having different degrees of physics details and data volume, the advantages and shortcomings of these two approaches are summarized in Table 2.1. The hybrid approach is a mixture of the physical model with experimental data, aiming at optimizing the combination of the two approaches.

Table 2.1: Comparison of data-driven and model-driven approaches (+: good performance; -: bad performance)

	Physics agreement	Experimental data agreement	Extrapolation	Online prediction
Model-driven (MD)	+	-	+	-
Data-driven (DD)	-	+	-	+

MOR and data-driven modeling to be investigated in this thesis are closely related to the general machine learning and data-driven techniques. The dimensionality reduction technique serves as the basis in both applications due to its capability of reducing dimensions of data. To preserve the essential physics while make the best of use of data, our strategy to develop the enhanced MOR and data-driven modeling methods is based on the hybrid philosophy. Particularly, their robustness and accuracy are improved significantly by integrating proper machine learning techniques, as will be demonstrated in the following chapters.

Chapter 3

Reduced-Order Modeling of Parameterized Systems

This chapter presents a detailed review of model reduction techniques and their extensions to nonlinear systems. First, in next section an introduction is given, followed by a model problem for reduced-order modeling. In Section 3.2, we briefly review the classification of the model reduction techniques traditionally used in dynamical systems to motivate the SVD-based approaches, including proper orthogonal decomposition (POD). In Section 3.3, we introduce several widely-used projection-based model reduction methods for parametrized systems, such as POD and the reduced basis method (RBM) and discuss their basic properties. System approximation methods for addressing the bottleneck of standard projection-based methods when dealing with nonlinear systems are reviewed in Section 3.4.

3.1 Introduction

Reduced-order modeling techniques generate surrogate models from a high-dimensional model (HDM) via the use of a reduced basis (e.g., eigenmodes for modal analysis (Craig 1981; Dickens, Nakagawa & Wittbrodt 1997) or left singular vectors for proper orthogonal decomposition (POD) (Berkooz, Holmes & Lumley 1993; Rathinam & Petzold 2003; Sirovich 1987)). The construction of reduced-order model (ROM) relies on a *key assumption*: while the solution associated with a parametrized system, in theory, belongs to an infinite-dimensional space, it resides on a much lower-dimensional and typically very smooth manifold induced by

time evolution and parametric dependence. Since projection-based techniques are commonly used for constructing the reduced-order system, this type of reduction techniques is called *projection-based model reduction*.

While the application of model reduction to linear, time-dependent systems has reached a considerable level of maturity, effective methods for parameterized boundary value problems are limited (Antoulas & Sorensen 2001; Carlberg & Farhat 2011). This is because data in boundary value problems are computationally expensive to acquire, and only a small amount of information is available from each sample configuration. The utility of model reduction methods is also limited for time-critical analysis (real-time simulation or repeated analyses). However, numerous engineering problems demanding time-critical static analysis have received increasing attention and become an important and vibrant research area, such as fracture modeling, multiscale constitutive modeling under various loading conditions, and real-time surgical simulation. This chapter provides a review of some state-of-art model order reduction methodologies for static systems.

3.1.1 Problem formulation

In this chapter, we consider the discrete equilibrium equation of a general parameterized mechanical problem for demonstration:

$$\mathbf{f}_{\text{int}}(\mathbf{u}(\boldsymbol{\mu})) + \mathbf{f}_{\text{ext}}(\boldsymbol{\mu}) = \mathbf{0}, \quad (3.1)$$

where the vector \mathbf{u} represents the state solution of the mechanical system, typically representing displacement in solid mechanics, and $\boldsymbol{\mu}$ are generic parameters the physics system depends on,

such as time, material coefficients, boundary conditions, etc. The vector field $\mathbf{f}_{\text{int}} \in \mathbb{R}^{\mathcal{N}}$ represents the internal forces driven by the state \mathbf{u} and $\mathbf{f}_{\text{ext}} \in \mathbb{R}^{\mathcal{N}}$ the external forces applied on the system. Since the dimension of the system can be large, the concept of reduced-order modeling is to construct a surrogate system of smaller dimension, i.e.

$$\mathbf{f}_{\text{int}}^{\text{sur}}(\mathbf{u}^r(\boldsymbol{\mu})) + \mathbf{f}_{\text{ext}}^{\text{sur}}(\boldsymbol{\mu}) = \mathbf{0}, \quad (3.2)$$

that preserves the important features of (3.1) as well, where superscript “ r ” is used to denote the reduced quantities in this chapter.

3.2 Model Order Reduction Methods for Linear Dynamical System

Traditionally, for linear dynamical systems, projection-based model reduction is broadly classified into three different methods with regard to the way in obtaining the reduced basis vectors (Benner, Gugercin & Willcox 2015): rational interpolation methods, balanced truncation, and proper orthogonal decomposition (POD). Similarly, (Antoulas 2005; Antoulas & Sorensen 2001) suggest that model reduction techniques can be distinguished as SVD-based, Krylov-based, and the SVD-Krylov-based approximation methods. Generally, the SVD-based methods employ the singular value decomposition (SVD) based on balanced approximation and optimal Hankel-norm approximation techniques, such as balanced truncation (Gugercin & Antoulas 2004), leading to the stability preservation of the full-order system and providing global error bounds. The second type of methods are based on either moment matching techniques (Astolfi 2010) via Krylov iteration or Padé approximation that are relatively inexpensive computationally. They can be applied to very large dynamical system but usually lack global

error bounds and stability preservation as the SVD-based methods. The third approach, SVD-Krylov-based approximation, seeks to combine the best attributes of the first two. One should note that the POD method based on SVD could be added to the first category. We refer (Antoulas 2005; Antoulas & Sorensen 2001; Baur, Benner & Feng 2014; Benner, Gugercin & Willcox 2015) for more details of these model reduction techniques applied to linear dynamical equations.

3.3 Selective Projection-Based Model Order Reduction Methods

In this section, we review various projection-based methods commonly used in the reduced order modelling community, including (snapshot-based) POD, RBM, component mode synthesis, and some variants of the POD method. These methods are used to construct a set of basis vectors to approximate the state variables of interest on a low-dimensional subspace. This is not intended to be an exhaustive review and some other methods, such as the a priori proper generalized decomposition (PGD) method (Chinesta, Ammar & Cueto 2010; Ladevèze, Passieux & Néron 2010; Nouy 2010) based on the reduced solution space by separation of variables are not discussed in this work.

3.3.1 Proper orthogonal decomposition

Proper orthogonal decomposition (POD), also called Karhunen-Loève expansion (Karhunen 1946; Loève 1955) or empirical orthogonal functions, is a method to construct the reduced basis vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subset \mathbb{R}^N$ that span the reduced space (or reduced manifold) via

SVD of a set of snapshots taken from training tests computed during the offline stage, see (Berkooz, Holmes & Lumley 1993; Rathinam & Petzold 2003; Sirovich 1987). It is also referred to as principal components analysis (PCA) (Jolliffe 2002).

Note that POD is a posteriori approach, and the quality of the reduced space depends on the set of snapshots. There are many ways to construct snapshots, yielding different types of basis, such as Lagrange basis, Hermite basis, and Taylor basis. But the choice of snapshots is a separate issue and will not be discussed here. We refer the interested readers to (Benner, Gugercin & Willcox 2015; Carlberg 2011) and references therein for more details.

For general parameterized mechanical problems, the snapshots are usually sampled from the displacement solutions (the state solutions) at a series of parameter values, i.e. $\{\mathbf{u}(\boldsymbol{\mu}_1), \dots, \mathbf{u}(\boldsymbol{\mu}_{N_s})\}$, where $\mathbf{u}(\boldsymbol{\mu}_i) \in \mathbb{R}^{\mathcal{N}}$ denotes the discrete displacement solution solved by any full-order numerical methods, such as finite element method (FEM), at the input parameter $\boldsymbol{\mu}_i$, \mathcal{N} is the number of degrees of freedom, and N_s is the number of solutions (parameters) collected.

Let $\mathbf{X}_s = [\mathbf{u}(\boldsymbol{\mu}_1), \dots, \mathbf{u}(\boldsymbol{\mu}_{N_s})] \in \mathbb{R}^{\mathcal{N} \times N_s}$ be the snapshot matrix and $r = \text{rank}(\mathbf{X}_s)$ (r is also the dimension of the space \mathcal{X} spanned by the snapshots, i.e. $\mathcal{X}_s = \text{span}\{\mathbf{u}(\boldsymbol{\mu}_1), \dots, \mathbf{u}(\boldsymbol{\mu}_{N_s})\}$). Note that usually the snapshot matrix has been centered by subtracting the mean of the columns (Chatterjee 2000), called mean-centered snapshots. It removes one trivial mode due to the mean value, and in practice it helps to deal with the essential boundary condition and give a compact form for error estimate (Liang et al. 2002).

The POD method constructs a set of orthonormal basis vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subset \mathbb{R}^{\mathcal{N}}$ whose linear span, $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ ($k < r < \mathcal{N}$), best approximates the snapshot space \mathcal{X}_s in an 2-norm sense by solving the following minimization problem

$$\begin{aligned} \min_{\{\mathbf{v}_i\}_{i=1}^k} \sum_{j=1}^{N_s} \left\| \mathbf{u}(\boldsymbol{\mu}_j) - \sum_{i=1}^k (\mathbf{u}(\boldsymbol{\mu}_j)^T \mathbf{v}_i) \mathbf{v}_i \right\|_2^2, \\ \text{subject to } \mathbf{v}_i^T \mathbf{v}_j = \delta_{ij}, \quad i, j = 1, \dots, k. \end{aligned} \quad (3.3)$$

POD can also be interpreted as to find the optimal basis matrix $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_k] \in \mathbb{R}^{\mathcal{N} \times k}$ that allows the reconstruction $\tilde{\mathbf{X}}_s := \mathbf{V} \mathbf{V}^T \mathbf{X}_s \in \mathbb{R}^{\mathcal{N} \times N_s}$ on the reduced space $\tilde{\mathcal{X}}_s = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ best approximating the snapshot matrix \mathbf{X}_s in the Frobenius norm, i.e.

$$\begin{aligned} \arg \min_{\{\mathbf{v}_i\}_{i=1}^k, \mathbf{V}^T \mathbf{V} = \mathbf{I}} \sum_{j=1}^{N_s} \left\| \mathbf{u}(\boldsymbol{\mu}_j) - \sum_{i=1}^k (\mathbf{u}(\boldsymbol{\mu}_j)^T \mathbf{v}_i) \mathbf{v}_i \right\|_2^2 &= \arg \min_{\mathbf{V}^T \mathbf{V} = \mathbf{I}} \left\| \mathbf{X}_s - \mathbf{V} \mathbf{V}^T \mathbf{X}_s \right\|_F^2 \\ &= \arg \max_{\mathbf{V}^T \mathbf{V} = \mathbf{I}} \text{trace}(\mathbf{V}^T \mathbf{X}_s \mathbf{X}_s^T \mathbf{V}), \end{aligned} \quad (3.4)$$

where the last equality uses the orthonormal property of \mathbf{V} . Thus, it is easy to show by the Eckert-Young-Minsky theorem (low-rank approximation) that the basis vectors solution to (3.3) and (3.4) is provided by the left singular vectors of the snapshot matrix \mathbf{X}_s corresponding to the k largest singular values (Liang et al. 2002). Suppose the SVD of \mathbf{X}_s is

$$\begin{aligned} \mathbf{X}_s &= \mathbf{V}_r \boldsymbol{\Sigma}_r \mathbf{W}_r^T, \\ \boldsymbol{\Sigma}_r &= \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{r \times r}, \quad \text{with } \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0, \end{aligned} \quad (3.5)$$

where $\sigma_1, \dots, \sigma_r$ are the singular value of \mathbf{X}_s , and $\mathbf{V}_r = [\mathbf{v}_1, \dots, \mathbf{v}_r] \in \mathbb{R}^{\mathcal{N} \times r}$ is the matrix composed of left singular vectors. Then, the POD basis matrix \mathbf{V} is obtain by simply truncating the left

singular vectors V_r . Moreover, the approximation (reconstruction) error of the POD representation of the snapshots is given by an *a priori* error estimate

$$\sum_{j=1}^{N_s} \left\| \mathbf{u}(\boldsymbol{\mu}_j) - \sum_{i=1}^k (\mathbf{u}(\boldsymbol{\mu}_j)^T \mathbf{v}_i) \mathbf{v}_i \right\|_2^2 = \left\| \mathbf{X}_s - \mathbf{V} \mathbf{V}^T \mathbf{X}_s \right\|_F^2 = \sum_{i=k+1}^r \sigma_i^2. \quad (3.6)$$

We have shown that POD is essentially the same as singular value decomposition (SVD) in a finite dimensional space or in Euclidean space. In addition, POD is equivalent to PCA (Jolliffe 2002; Kerschen et al. 2005; Liang et al. 2002), as revealed in (3.4), but the name POD is usually used in the context of model reduction whereas PCA is used in the field of machine learning and statistics. Thus, the subspace $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ constructed by POD is not only the subspace with minimum reconstruction error but also the subspace with maximum variance.

As the POD method provides us with the reduced basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, the high-dimensional solution $\mathbf{u}(\boldsymbol{\mu}) \in \mathbb{R}^N$ for a parameter $\boldsymbol{\mu}$ is approximated by the *reduced approximation* $\tilde{\mathbf{u}}(\boldsymbol{\mu}) \in \mathbb{R}^N$ as follows

$$\mathbf{u}(\boldsymbol{\mu}) \approx \tilde{\mathbf{u}}(\boldsymbol{\mu}) = \mathbf{V} \mathbf{u}^r(\boldsymbol{\mu}) = \sum_{i=1}^k \mathbf{v}_i u_i^r(\boldsymbol{\mu}), \quad (3.7)$$

where $\mathbf{u}^r \in \mathbb{R}^k$ is the *reduced solution* (or *reduced coefficients*) of k dimension solved from the surrogate system (3.2) where Galerkin projection is usually applied.

Galerkin projection using POD

To solve the reduced solution $\mathbf{u}^r \in \mathbb{R}^k$ by the low-dimensional system, a dual reduced basis \mathbf{W} is also constructed such that Galerkin (or Petrov-Galerkin) projection can be used to

construct the parametric ROM by projecting the given parametric HDM (3.1) onto the global reduced basis (and its dual counterpart) as follows

$$\mathbf{W}^T \mathbf{f}_{\text{int}}(\mathbf{V}\mathbf{u}^r(\boldsymbol{\mu})) + \mathbf{W}^T \mathbf{f}_{\text{ext}}(\boldsymbol{\mu}) = \mathbf{0}. \quad (3.8)$$

Usually $\mathbf{W} = \mathbf{V}$ is selected as the standard Galerkin projection POD method (Berkooz, Holmes & Lumley 1993; Krysl, Lall & Marsden 2001; Sirovich 1987). But \mathbf{W} different from \mathbf{V} based on the Petrov-Galerkin concept has also been used to achieve real speedup or structure-preserving properties of ROM when dealing with nonlinear systems (Bui-Thanh, Willcox & Ghattas 2008; Carlberg, Barone & Antil 2017; Carlberg 2011). The Petrov-Galerkin projection is often applied together with greedy algorithms for reducing nonlinear systems, which is to be discussed in Section 3.4.

To achieve minimal marginal computational cost per input-output evaluation, the POD-Galerkin method is usually split into a computationally expensive *offline* and an efficient *online* stage. It works well for the linear operator or the operators that admits an affine decomposition (Grepl et al. 2007). For example, if \mathbf{f}_{int} is a linear operator, it means

$$\mathbf{f}_{\text{int}}(\mathbf{u}(\boldsymbol{\mu})) = \mathbf{A}\mathbf{u}(\boldsymbol{\mu}), \quad (3.9)$$

where \mathbf{A} is the Jacobian of \mathbf{f}_{int} independent of the input parameters $\boldsymbol{\mu}$. On the other hand, an affine decomposition for parameter-dependent operator should permit,

$$\mathbf{f}_{\text{int}}(\mathbf{u}(\boldsymbol{\mu})) = \mathbf{A}(\boldsymbol{\mu})\mathbf{u}(\boldsymbol{\mu}) = \left(\sum_{q=1}^Q \theta^q(\boldsymbol{\mu}) \mathbf{A}^q \right) \mathbf{u}(\boldsymbol{\mu}), \quad (3.10)$$

where $\theta^q(\boldsymbol{\mu})$ are the given scalar functions of the parameters $\boldsymbol{\mu}$. The key benefit is that, under this assumption, \mathbf{A}^q does not depend on $\boldsymbol{\mu}$, which means the reduced operators can be computed during the offline stage once and for all. That is, the snapshot matrix \mathbf{X}_s and the POD basis \mathbf{V} , as well as the reduced operator $\mathbf{V}^T \mathbf{A} \mathbf{V} \in \mathbb{R}^{k \times k}$ (or $\mathbf{V}^T \mathbf{A}^q \mathbf{V} \in \mathbb{R}^{k \times k}$, $q = 1, \dots, Q$, under the affine decomposition condition) are computed and stored in the offline phase. Then the reduced operator is directly used for the online evaluations of the associated ROM

$$(\mathbf{V}^T \mathbf{A} \mathbf{V}) \mathbf{u}^r(\boldsymbol{\mu}) + \mathbf{V}^T \mathbf{f}_{\text{ext}}(\boldsymbol{\mu}) = \mathbf{0}. \quad (3.11)$$

Note that the above reduced system assumes a parameterized linear system. Once \mathbf{u}^r is solved from (3.11), the reduced approximation for $\mathbf{u}(\boldsymbol{\mu}) \in \mathbb{R}^{\mathcal{N}}$ is retrieved by $\tilde{\mathbf{u}}(\boldsymbol{\mu}) = \mathbf{V} \mathbf{u}^r(\boldsymbol{\mu})$ in (3.7).

Error analysis for projection-based reduced order model (ROM)

Let $\mathbf{u}(\boldsymbol{\mu})$ be the high-fidelity solution solved by (3.1), $\mathbf{u}^r(\boldsymbol{\mu})$ the reduced solution of the projection-based ROM (3.8), and $\mathbf{V} \in \mathbb{R}^{\mathcal{N} \times k}$ a reduced projection matrix. Following (Rathinam & Petzold 2003), the ROM error of $\tilde{\mathbf{u}}(\boldsymbol{\mu}) = \mathbf{V} \mathbf{u}^r(\boldsymbol{\mu})$ can be decomposed into two contributions

$$\begin{aligned} \xi_{\text{ROM}}(\boldsymbol{\mu}) &= \mathbf{u}(\boldsymbol{\mu}) - \tilde{\mathbf{u}}(\boldsymbol{\mu}) = (\mathbf{u}(\boldsymbol{\mu}) - \Pi_{\mathbf{V}, \mathbf{V}} \mathbf{u}(\boldsymbol{\mu})) + (\Pi_{\mathbf{V}, \mathbf{V}} \mathbf{u}(\boldsymbol{\mu}) - \tilde{\mathbf{u}}(\boldsymbol{\mu})) \\ &= (\mathbf{I} - \Pi_{\mathbf{V}, \mathbf{V}}) \mathbf{u}(\boldsymbol{\mu}) + \mathbf{V} (\mathbf{V}^T \mathbf{u}(\boldsymbol{\mu}) - \mathbf{u}^r(\boldsymbol{\mu})) \\ &= \xi_{\mathbf{V}^\perp}(\boldsymbol{\mu}) + \xi_{\mathbf{V}}(\boldsymbol{\mu}) \end{aligned} \quad (3.12)$$

where $\Pi_{\mathbf{V}, \mathbf{V}} = \mathbf{V} \mathbf{V}^T \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$. Note that $\Pi_{\mathbf{W}, \mathbf{V}} = \mathbf{W} (\mathbf{V}^T \mathbf{W})^{-1} \mathbf{V}^T$ denotes a projector that projects a vector in $\mathbb{R}^{\mathcal{N}}$ onto the column space of \mathbf{W} ($\text{range}(\mathbf{W}) \subset \mathbb{R}^{\mathcal{N}}$) along the direction normal to \mathbf{V} .

As we can see, the first error $\xi_{\mathbf{V}^\perp}(\boldsymbol{\mu})$ is caused by the choosing of reduced projection basis \mathbf{V} , which means the exact solution trajectory with respect to $\boldsymbol{\mu}$ does not strictly resides on the subspace $\text{range}(\mathbf{V}) \subset \mathbb{R}^{\mathcal{N}}$. This error consists of two components: the first component is related to the quality of the snapshots (the sampling parameters) during the offline stage; the second one is due to the truncation error of the basis used in reduced approximation (see (3.6)). Thus, if the snapshot space is rich enough to cover the HDM solution space, and the POD basis vectors of full-rank order are used, the error $\xi_{\mathbf{V}^\perp}(\boldsymbol{\mu})$ is expected to vanish. Usually reducing the projection error $\xi_{\mathbf{V}^\perp}(\boldsymbol{\mu})$ is the key to decrease the total error of reduced-order modeling. One can either use better sampling strategies or more accurate projection basis. Moreover, $\xi_{\mathbf{V}^\perp}(\boldsymbol{\mu})$ is used to provide a lower bound for the reduced system, and it can be related to the upper bound if the Galerkin method is applied (Barrault et al. 2004; Porsching 1985).

The second error $\xi_{\mathbf{V}}(\boldsymbol{\mu})$ is caused by the fact that one is solving a different dynamical system (3.8) from (3.1) after the projection. It depends on how the reduced order model is constructed. For some dynamical system applications, Tröltzsch & Volkwein (Tröltzsch & Volkwein 2009) developed *a posteriori* error estimates for POD method.

3.3.2 Reduced-basis method

The reduced-basis methods (RBM) are widely used in the context of parameterized partial differential equations. They have been studied for both steady (Prud'homme et al. 2002; Rozza, Huynh & Patera 2008) and time dependent problems (Grepl & Patera 2005; Haasdonk & Ohlberger 2008). Compared to POD that requires to compute N_s solutions of the high-

dimensional system and conduct SVD analysis of the large snapshot matrix \mathbf{X}_s , RBM follows a different approach and constructs the reduced basis vectors directly from the high-fidelity solutions. This means, the reduced approximation of the high-fidelity solution $\mathbf{u}(\boldsymbol{\mu}) \in \mathbb{R}^N$ is defined as the linear combination of the orthogonal basis $\{\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_k\}$ which is obtained by orthogonalizing the precomputed solutions $\{\mathbf{u}(\boldsymbol{\mu}_1), \dots, \mathbf{u}(\boldsymbol{\mu}_{N_s})\}$, i.e.

$$\tilde{\mathbf{u}}(\boldsymbol{\mu}) = \sum_{i=1}^k \boldsymbol{\zeta}_i \beta_i \quad (3.13)$$

The key idea of RBM is to find the optimal set of parameters $\{\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_{N_s}\}$ that leads to the most accurate approximation of the original high-fidelity solution. This is achieved through a greedy algorithm using an *a posteriori* error estimator $\xi_k(\boldsymbol{\mu})$, where the subscript k shows the dependency on the order of the reduced basis approximation. The estimator $\xi_k(\boldsymbol{\mu})$ gives an upper bound for the error $\|\mathbf{u}(\boldsymbol{\mu}) - \tilde{\mathbf{u}}(\boldsymbol{\mu})\|$ of the reduced approximation solution $\tilde{\mathbf{u}}(\boldsymbol{\mu}) \in \text{span}\{\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_k\}$ with respect to the high-fidelity solution $\mathbf{u}(\boldsymbol{\mu}) \in \mathbb{R}^N$ in a certain norm. The error estimator is supposed to be valid for all chosen parameters $\{\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_{N_s}\}$ and all dimensions $i = 1, \dots, k$.

A simple procedure to construct RBM basis vectors is provided as follows. Let a set of parameters be $\mathcal{P} = \{\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_{N_s}\} \subset \mathcal{D}$ with the parameter domain $\mathcal{D} \subset \mathbb{R}^l$. Given an arbitrary starting parameter value $\boldsymbol{\mu}_1$, the first vector of the reduced basis $\mathbf{u}(\boldsymbol{\mu}_1)$ is computed. In the greedy algorithm, the basis is iteratively enriched by finding the parameter value $\boldsymbol{\mu}_j \in \mathcal{P}$ in \mathcal{P} that maximizes the error estimator ξ_k , assuming it corresponds to the most informative solution

$\mathbf{u}(\boldsymbol{\mu}_j)$. The chosen parameter is stored in a subset Ξ of the parameter space \mathcal{P} as a discrete surrogate of the continuous space \mathcal{D} . For example, at step j of the greedy algorithm, a new high-fidelity solution $\mathbf{u}(\boldsymbol{\mu}_j)$ is evaluated at the parameter given by

$$\boldsymbol{\mu}_j = \arg \max_{\boldsymbol{\mu} \in \Xi} \xi_j(\boldsymbol{\mu}), \quad (3.14)$$

and the reduced basis is enriched with $\mathbf{u}(\boldsymbol{\mu}_j)$. For conditioning purposes, the new introduced solution basis $\mathbf{u}(\boldsymbol{\mu}_j)$ is often orthogonalized to the existed basis vectors by performing the Gram-Schmidt orthogonalization.

One goal of the reduced basis community is to derive such error estimators for all kind of parameterized systems (Peherstorfer 2013) but it is still an open issue. For example, a posteriori error estimator (Grepl & Patera 2005) for parabolic problems with a affine parameter dependence has been proposed. The key idea of this work is to treat time as an additional parameter. Since the derivation of the error estimate $\xi_k(\boldsymbol{\mu})$ relies on the coercivity constant of the PDE systems. It is not straightforward to compute the coercivity constant if the associated weak form does not admit an affine expansion. We refer (Grepl et al. 2007; Nguyen 2007; Prud'homme et al. 2002; Quarteroni, Manzoni & Negri 2015; Quarteroni, Rozza & Manzoni 2011) and the references therein.

Although the reduced basis of RBM can also be constructed without this greedy approach and without using error estimators, a poor selection of the parameters might result in a set of reduced basis vectors with poor-representation, yielding disastrous approximation results. Compared to POD embedded with SVD procedure as an *a priori* error estimator, the success of

RBM relies on the rigorous, sharp, and efficient *a posteriori* error estimators as well as on the greedy strategy to select the reduced basis vectors.

3.3.3 Component mode synthesis

Component mode synthesis (CMS) methods (Hurty 1960, 1965) are widely used in finite element analysis of structural dynamics (Craig 1981; Klerk, Rixen & Voormeeren 2008). A given structure is subdivided into components or substructures, for each of which the associated natural frequencies and modes are independently analyzed and an approximation is made by projecting the state variables onto the pre-computed modes. The substructure mode shapes are then assembled to approximate the frequencies and mode shapes of the original structure.

The CMS methods can be seen as domain decomposition techniques or sub-structuring techniques which in fact date back to much earlier work (Klerk, Rixen & Voormeeren 2008). These methods are effective for the predictions of frequencies and mode shapes of complex structures without constructing the HDM. Additionally, parallelization is easily implemented due to the independence of each subdomain. However, it is noted that interface boundary handling between substructures has been an important issue in the CMS methods (Kim et al. 2017). Various other improved CMS methods have been proposed according to the compatibility of the interface DOF. Furthermore, CMS methods can be employed with other model reduction methods, such as dynamic condensation (Kim et al. 2017) and interpolation methods (Lee & Cho 2017), and the basis model can be enriched by static modes coming from constraints or attachments (Bathe & Dong 2014).

3.3.4 Enhanced POD

The standard POD introduced in Section 3.3.1 computes state vectors from snapshots and constructs the truncated bases that capture the dominant modes. This approach is referred to as *Lagrange* POD to distinguish from other model reduction methods that use the Taylor (Peterson 1989; Porsching 1985) and Hermite subspaces (Ito & Ravindran 1998), where the sensitivity derivatives of the state are included in the reduced bases. As pointed out in (Carlberg & Farhat 2008, 2011), Lagrange POD model reduction for static systems has several disadvantages. First, the POD bases are simply computed to minimize the average L^2 reconstruction error (see (3.3)) of the snapshots onto the subspace spanned by the bases, without using any information related to the input-output behavior of the system. Thus, the resulting POD bases might not be optimal in the sense of preserving the input-output map and representing the specific goal of model reduction. Additional knowledge about the system or data can be incorporated into the basis construction. Another shortcoming of POD is that it utilizes snapshot data that are computationally expensive to collect. Thus, extracting more data information from its derivatives in addition to the state variable is preferable in learning process. However, POD cannot be applied directly to the Taylor or Hermite bases due the different scaling of quantities.

A so-called compact POD method has been proposed in (Carlberg & Farhat 2008, 2011) to address the above problems. It essentially relies on a weighted version of POD method, where the snapshots, including the state vectors (0-order derivative) and the associated derivatives, are weighted appropriately to optimally represent the outputs of interest in a certain sense. Thus, the resulting reduced-order bases and ROM are goal-oriented. The weighted scheme of POD (Bistrrian & Susan-Resiga 2016; Carlberg & Farhat 2011; Peng & Mohseni 2016) shares a similar mathematical framework as the generalized (or weighted) PCA (Jolliffe 2002; Vidal, Ma &

Sastry 2016) that has been widely explored in the area of data mining and computer vision. More discussion about enhancing POD will be presented in Chapter 5.

3.4 Hyper-Reduction Techniques for Nonlinear Systems

In the previous sections, various projection-based reduction methods have proven effective in reducing the computational cost of a linear model by reducing the dependence of the state approximation on the large mesh discretization of the associated HDM. However, they all fail to achieve computational cost reduction (or speedup) when applying to nonlinear models (Chaturantabut & Sorensen 2010), as the cost of evaluation of the reduced nonlinear terms at each solver iteration still scales with the size of the underlying HDM rather than with that of the reduced subspace.

3.4.1 Lifting bottleneck

For demonstrating the inefficiency of projection-based methods, we assume the POD-Galerkin method has been applied for a general parameterized mechanical system in (3.1) and it results in the following reduced system

$$\mathbf{V}^T \mathbf{f}_{\text{int}}(\mathbf{V}\mathbf{u}^r(\boldsymbol{\mu})) + \mathbf{V}^T \mathbf{f}_{\text{ext}}(\boldsymbol{\mu}) = \mathbf{0}, \quad (3.15)$$

where $\mathbf{V} \in \mathbb{R}^{\mathcal{N} \times k}$ is the reduced basis matrix and $\mathbf{u}^r \in \mathbb{R}^k$ is the reduced variables, and let \mathbf{f}_{int} be a combination of a linear term and a nonlinear term, i.e.,

$$\mathbf{f}_{\text{int}}(\mathbf{u}(\boldsymbol{\mu})) = \mathbf{A}\mathbf{u}(\boldsymbol{\mu}) + \mathbf{f}(\mathbf{u}(\boldsymbol{\mu})), \quad (3.16)$$

where $\mathbf{A} \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ is a linear operator, and $\mathbf{f}(\mathbf{u}(\boldsymbol{\mu})) \in \mathbb{R}^{\mathcal{N}}$ denotes the nonlinear term. During the online reduced-order modeling (see (3.15)) the computational cost associated to the evaluation of the linear term scales only with the order of reduced basis k , since the corresponding reduced operators can be pre-computed offline, as discussed in Section 3.3.1. However, this is not the case for the evaluation of the reduced nonlinear term $\hat{\mathbf{f}}(\mathbf{u}^r) = \mathbf{V}^T \mathbf{f}(\mathbf{V}\mathbf{u}^r(\boldsymbol{\mu}))$. Take FE based solver as an example, the online computation of the reduced nonlinear term is carried out in the following manner

$$\hat{\mathbf{f}}(\mathbf{u}^r) := \mathbf{V}^T \mathbf{f}(\mathbf{V}\mathbf{u}^r) = \sum_{g=1}^{N_g} \mathbf{V}_{\mathcal{I}_g}^T \mathbf{f}_g(\mathbf{V}\mathbf{u}^r) w_g, \quad (3.17)$$

where N_g is the total number of Gauss integration points in the structure, w_g is the associated weights, \mathcal{I}_g is the set of the indices of the degree of freedoms (DOFs) corresponding to the Gauss point g , $|\mathcal{I}_g|$ is the cardinality of the set \mathcal{I}_g , $\mathbf{f}_g \in \mathbb{R}^{|\mathcal{I}_g|}$ is the contribution of the point g towards the vector \mathbf{f} , and the subscript \mathcal{I}_g of $\mathbf{V}_{\mathcal{I}_g} \in \mathbb{R}^{|\mathcal{I}_g| \times k}$ means the restriction of $\mathbf{V} \in \mathbb{R}^{\mathcal{N} \times k}$ to the rows indexed by \mathcal{I}_g . It is then easy to see from (3.17) that the computation complexity of the reduced nonlinear term $\hat{\mathbf{f}}(\mathbf{u}^r) \in \mathbb{R}^k$ is proportional to the total number of evaluation points N_g of the HDM discretization, regardless of the reduction of the number of state variables. This is known as the *lifting bottleneck*, and it degrades the performance of reduced models for nonlinear systems.

The above-mentioned scalability issue can be alleviated by introducing *hyper-reduction* techniques (i.e., another layer of reduced approximation on the underlying system in addition to the reduction of state solution), which aims to significantly reduce the cost in evaluating the

nonlinear terms by computing them at few, selected element or nodes of the discrete system, and effectively approximate the missing information. Depending on the selection strategy used to extract information from the nonlinear snapshots and the approximation strategy for reconstructing the nonlinear function, we classify the main methods found in the literature into two types: the *interpolation-type* and the *cubature-type* strategy.

3.4.2 Interpolation-type strategy

As previously highlighted, the direct projection-based MOR approach still depends on the dimension of full-scale spatial discretization. The essential idea of the interpolation-type hyper-reduction strategy is to optimally select some “important” grid points or DOFs of the original discrete system by using greedy algorithm, and to reconstruct the high-dimensional nonlinear function by only using those few selected points and an oblique projection.

The “gappy” POD method (Everson & Sirovich 1995), first introduced for image reconstruction and later applied in dynamical systems (Carlberg et al. 2013; Carlberg, Bou-Mosleh & Farhat 2011; Willcox 2006), was one of the first schemes of the interpolation-type hyper-reduction. The “empirical interpolation” method (EIM) (Barrault et al. 2004; Grepl et al. 2007), whose discrete variants are known as DEIM (Chaturantabut & Sorensen 2010), reduces the computational cost associated with the nonlinear terms by interpolating the nonlinear function at a few entries using an empirically derived basis. The variants includes the missing point estimation (MPE) (Astrid et al. 2008) and the “best” points interpolation method (Nguyen, Patera & Peraire 2008), to cite a few, which relies on the employment of POD basis. The term “hyper-reduction”, however, was coined in (Ryckelynck 2005), where the optimal reduced basis

is adaptively enriched by using the Krylov's subspaces generated by the governing equation residual, and few spatial integration points of the finite element model are selected to forecast the evolution of the reduced state variables. These different approaches are distinguished by using different measurements (norms) or regression form to select the small selected set of spatial grids or entries. Note that this technique is heavily related to the so-called sparse sampling (Sargsyan, Brunton & Kutz 2015).

Given the POD reduced nonlinear function in (3.17), the general form of the interpolation-type hyper-reduction is

$$\hat{f}(\mathbf{u}^r) = \mathbf{V}^T \mathbf{f}(\mathbf{V}\mathbf{u}^r) \approx \mathbf{V}^T \mathbf{M} \mathbf{f}_{\mathcal{P}}(\mathbf{V}\mathbf{u}^r) = \mathbf{V}^T \mathbf{M} \mathbf{P}^T \mathbf{f}(\mathbf{V}\mathbf{u}^r) = \sum_{g \in \text{supp}(\mathcal{P})} \mathbf{V}_{\mathcal{I}_g}^T \mathbf{M}_{\mathcal{I}_g} \mathbf{P}_{\mathcal{I}_g}^T \mathbf{f}_g(\mathbf{V}\mathbf{u}^r) w_g, \quad (3.18)$$

where $\mathbf{M} \in \mathbb{R}^{\mathcal{N} \times \hat{n}}$ is a matrix that maps a vector of dimension \hat{n} to the space $\mathbb{R}^{\mathcal{N}}$, \mathcal{P} is the set of the indices of the DOFs to be selected, $\mathbf{P} \in \mathbb{R}^{\mathcal{N} \times |\mathcal{P}|}$ is the associated selection matrix with one non-zero entry per column to select the set of chosen interpolation DOFs, and $g \in \text{supp}(\mathcal{P})$ means only the Gauss points within the support of the selected DOFs need to be evaluated. The construction of the projection \mathbf{M} and the interpolation indices \mathcal{P} is the key task in the interpolation-type hyper-reduction methods.

Discrete empirical interpolation method (DEIM)

DEIM (Chaturantabut & Sorensen 2010) is one of the most popular ‘‘Gappy’’-type empirical interpolation methods (EIM) (Barrault et al. 2004; Everson & Sirovich 1995) developed to overcome the weakness of the POD-Galerkin when the nonlinear operators exist. DEIM combines POD projection with interpolation. Specifically, the DEIM uses selected

interpolation indices to specify an interpolation-based projection for a nearly optimal l^2 subspace approximating the nonlinear vector.

In DEIM the collateral POD basis $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_{\hat{k}}] \in \mathbb{R}^{\mathcal{N} \times \hat{k}}$ ($\hat{k} \ll \mathcal{N}$) is constructed based on the snapshot matrix of the nonlinear function $\mathbf{X}_f = [\mathbf{f}(\mathbf{u}(\boldsymbol{\mu}_1)), \dots, \mathbf{f}(\mathbf{u}(\boldsymbol{\mu}_{N_s}))] \in \mathbb{R}^{\mathcal{N} \times N_s}$, where $\mathbf{f}(\mathbf{u}(\boldsymbol{\mu}_j))$ denotes the nonlinear vectors solved in HDM for the parameter $\boldsymbol{\mu}_j$ “offline”, see (3.1) and (3.16). As such, the nonlinear function at $\boldsymbol{\mu}$ is approximated by the linear combination of POD basis vectors

$$\mathbf{f}(\mathbf{u}(\boldsymbol{\mu})) \approx \mathbf{Z}\mathbf{c}(\boldsymbol{\mu}) = \sum_{i=1}^{\hat{k}} \mathbf{z}_i c_i(\boldsymbol{\mu}). \quad (3.19)$$

Since the system $\mathbf{f}(\mathbf{u}(\boldsymbol{\mu})) \approx \mathbf{Z}\mathbf{c}(\boldsymbol{\mu})$ to determine the coefficients $\mathbf{c}(\boldsymbol{\mu}) \in \mathbb{R}^{\hat{k}}$ is highly over-determined, an additional selection matrix $\mathbf{P} = [\mathbf{e}_{\varphi_1}, \dots, \mathbf{e}_{\varphi_{\hat{n}}}] \in \mathbb{R}^{\mathcal{N} \times \hat{n}}$ is used to restrict a vector of dimension \mathcal{N} onto its \hat{n} entries. In the offline stage, the DEIM interpolation points $\varphi_1, \dots, \varphi_{\hat{n}}$ are selected with a greedy algorithm, see Table 3.1. Suppose $\mathbf{P}^T \mathbf{Z} \in \mathbb{R}^{\hat{n} \times \hat{k}}$ is nonsingular (set $\hat{n} = \hat{k}$), $\mathbf{c}(\boldsymbol{\mu})$ can be uniquely defined from

$$\mathbf{P}^T \mathbf{f}(\mathbf{u}(\boldsymbol{\mu})) = \mathbf{P}^T \mathbf{Z}\mathbf{c}(\boldsymbol{\mu}), \quad (3.20)$$

that is, $\mathbf{c}(\boldsymbol{\mu}) = (\mathbf{P}^T \mathbf{Z})^{-1} \mathbf{P}^T \mathbf{f}(\mathbf{u}(\boldsymbol{\mu}))$, and then the DEIM approximation $\tilde{\mathbf{f}}$ of the original high-dimensional nonlinear vector \mathbf{f} is given by

$$\mathbf{f}(\mathbf{u}(\boldsymbol{\mu})) \approx \tilde{\mathbf{f}}(\mathbf{u}(\boldsymbol{\mu})) = \mathbf{Z}\mathbf{c}(\boldsymbol{\mu}) = \mathbf{Z}(\mathbf{P}^T \mathbf{Z})^{-1} \mathbf{P}^T \mathbf{f}(\mathbf{u}(\boldsymbol{\mu})) = \mathbf{Z}\hat{\mathbf{Z}}^{-1} \mathbf{P}^T \mathbf{f}(\mathbf{u}(\boldsymbol{\mu})), \quad (3.21)$$

where $\hat{\mathbf{Z}} = (\mathbf{P}^T \mathbf{Z}) \in \mathbb{R}^{\hat{k} \times \hat{k}}$. Let $\mathcal{Z} = \text{span}\{\mathbf{z}_1, \dots, \mathbf{z}_k\} \subset \mathbb{R}^{\mathcal{N}}$ and $\mathcal{P}^\perp = \text{span}\{\mathbf{e}_{\varphi_1}, \dots, \mathbf{e}_{\varphi_k}\} \subset \mathbb{R}^{\mathcal{N}}$. It is clear that $\Pi_{\mathcal{Z}, \mathcal{P}} = \mathbf{Z}(\mathbf{P}^T \mathbf{Z})^{-1} \mathbf{P}^T$ is an oblique projector that projects an arbitrary vector in $\mathbb{R}^{\mathcal{N}}$ onto \mathcal{Z} along \mathcal{P} . Compared with (3.18), we can see in DEIM $\mathbf{M} = \mathbf{Z}(\mathbf{P}^T \mathbf{Z})^{-1} \in \mathbb{R}^{\mathcal{N} \times \hat{k}}$.

Table 3.1: Greedy Algorithm (Chaturantabut & Sorensen 2010)

INPUT: $\{\mathbf{z}_l\}_{l=1}^m$ linearly independent POD basis

OUTPUT: $\vec{\varphi} = [\varphi_1, \dots, \varphi_m]^T \in \mathbb{R}^m$

1. $[\rho, \varphi_1] = \max\{|\mathbf{z}_1|\}$

2. $\mathbf{Z} = [\mathbf{z}_1]$, $\mathbf{P} = [\mathbf{e}_{\varphi_1}]$, $\vec{\varphi} = [\varphi_1]$

3. **for** $l = 2$ to m **do**

solve $(\mathbf{P}^T \mathbf{Z})\mathbf{c} = \mathbf{P}^T \mathbf{z}_l$ for \mathbf{c}

$\mathbf{r} = \mathbf{z}_l - \mathbf{Z}\mathbf{c}$

$[\rho, \varphi_l] = \max\{|\mathbf{r}|\}$

$\mathbf{Z} \leftarrow [\mathbf{Z} \ \mathbf{z}_l]$, $\mathbf{P} \leftarrow [\mathbf{P} \ \mathbf{e}_{\varphi_l}]$, $\vec{\varphi} \leftarrow \begin{bmatrix} \vec{\varphi}^T & \varphi_l \end{bmatrix}^T$

4. **end for**

Following the Lemma 3.2 in (Chaturantabut & Sorensen 2010), the error bound for the DEIM approximation is given as follows:

Lemma: Let $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_k] \in \mathbb{R}^{n \times m}$ be a given orthonormal matrix ($\mathbf{Z}^T \mathbf{Z} = \mathbf{I}_m$, $m < n$) and \mathbf{P} the index matrix computed by the greedy algorithm in Table 3.1. The DEIM approximation of an arbitrary vector $\mathbf{f} \in \mathbb{R}^{\mathcal{N}}$ is $\tilde{\mathbf{f}} = \mathbf{Z}(\mathbf{P}^T \mathbf{Z})^{-1} \mathbf{P}^T \mathbf{f} \in \mathbb{R}^{\mathcal{N}}$. An error bound for $\tilde{\mathbf{f}}$ is then given by

$$\|\mathbf{f} - \tilde{\mathbf{f}}\|_2 \leq c \xi_{\mathbf{Z}^\perp}(\mathbf{f}), \quad (3.22)$$

where $\xi_{\mathcal{Z}^\perp}(\mathbf{f}) = \|(\mathbf{I} - \mathbf{Z}\mathbf{Z}^T)\mathbf{f}\|_2$ and $c = \|(\mathbf{P}^T\mathbf{Z})^{-1}\|_2$. An a priori estimate for the magnification factor c is

$$c \leq (1 + \sqrt{2n})^{\hat{k}-1} \|\mathbf{z}_1\|_\infty^{-1} \leq \sqrt{n}(1 + \sqrt{2n})^{\hat{k}-1}. \quad (3.23)$$

As shown in the error bound, the lower bound of the DEIM error is the reconstruction error $\xi_{\mathcal{Z}^\perp}$ that measures the error of the optimal 2-norm approximation for \mathbf{f} on the range $\mathcal{Z} = \mathcal{R}(\mathbf{Z})$. Therefore, DEIM inevitably introduces extra projection error than the orthogonal projector $\Pi_{\mathcal{Z},\mathcal{Z}} = \mathbf{Z}\mathbf{Z}^T$ for approximating the snapshot data of nonlinear terms. The matrix norm $\|(\mathbf{P}^T\mathbf{Z})^{-1}\|_2$ depends on the selection matrix \mathbf{P} computed by the greedy algorithm, where each iteration aims to select an index to minimize the stepwise growth of the error bound. A hypothesis is that the range of given \mathbf{Z} is able to capture the system nonlinearity $\mathbf{f}(\boldsymbol{\mu})$ over the entire parameter space \mathcal{D} . As reported in (Drmac & Gugercin 2016) that the error bound in (3.22) is rather pessimistic, a new, robust selection procedure for DEIM based on the pivoted QR has been recently introduced therein.

Gappy-POD

Contrary to the interpolation in (3.21) the Gappy-POD (Carlberg et al. 2013; Carlberg, Bou-Mosleh & Farhat 2011; Everson & Sirovich 1995; Willcox 2006) uses regression to approximate the nonlinear function. That is, given the collateral POD basis $\mathbf{Z} \in \mathbb{R}^{\mathcal{N} \times \hat{k}}$, determine an index matrix $\mathbf{P} \in \mathbb{R}^{\mathcal{N} \times \hat{n}}$ such that

$$\mathbf{c} = \arg \min_{\mathbf{c} \in \mathbb{R}^{\hat{k}}} \|\mathbf{P}^T \mathbf{f} - \mathbf{P}^T \mathbf{Z} \mathbf{c}\|_2. \quad (3.24)$$

With $\hat{n} \geq \hat{k}$ as a necessary condition to guarantee a unique solution, the solution is given by $\mathbf{c} = (\mathbf{P}^T \mathbf{Z})^\dagger \mathbf{P}^T \mathbf{f}$, and the Gappy-POD approximation results in

$$\mathbf{f} \approx \mathbf{Z}\mathbf{c} = \mathbf{Z}(\mathbf{P}^T \mathbf{Z})^\dagger \mathbf{P}^T \mathbf{f}. \quad (3.25)$$

Here, there are more sampling points than POD modes, i.e. $\hat{n} \geq \hat{k}$, ‘ \dagger ’ denotes the pseudo-inverse, and the indices of \mathbf{P} is computed by the same greedy algorithm in Table 3.1.

Other interpolation-type methods are based on a very similar idea that combining the POD with greedy algorithm to select some “important” grid points or DOFs in the discrete system, such as the EIM, MPE, and the Gauss-Newton with approximated tensors (GNAT) method (Carlberg, Bou-Mosleh & Farhat 2011). For FE applications, Tiso and Rixen (Tiso & Rixen 2013) reported that the unassembled DEIM (UDEIM), by treating each displacement DOF separately without considering the element connectivity, is more efficient to approximate nonlinear force vector. However, it increases the storage burden in the offline phase.

1D nonlinear parametrized function

Consider a 1D nonlinear parametrized function (Chaturantabut & Sorensen 2010; Nguyen, Patera & Peraire 2008): $f : \Omega \times \mathcal{D} \rightarrow \mathbb{R}$ defined by

$$f(x; \mu) = (1-x) \cos(3\pi\mu(x+1)) e^{-(1+x)\mu} \quad (3.26)$$

where $x \in \Omega = [-1, 1]$ and $\mu \in \mathcal{D} = [1, \pi]$. The grid points of $n = 100$ are equidistantly positioned in the spatial domain Ω , denoted as $\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{R}^n$. Then, the nonlinear vector is defined as $\mathbf{f}(\mu) = [f(x_1; \mu), f(x_2; \mu), \dots, f(x_n; \mu)] \in \mathbb{R}^n$. Given $N_s = 51$ parameters uniformly taken from

the parameter domain \mathcal{D} , i.e. $\{\mu_i\}_{i=1}^{N_s}$, we collected 51 snapshots $f(\mu_i)$ to construct POD basis $\{\mathbf{z}_i\}_{i=1}^k \subset \mathbb{R}^n$. Figure 3.1 and Figure 3.2 show the differences of POD modes between the raw snapshots and the mean-centered snapshots. The plot of singular values in Figure 3.1 shows that the singular values of the mean-centered snapshots vanish as the number of POD basis increases. Figure 3.2 shows that by subtracting the mean value, the resulting POD modes represent the homogeneous boundary conditions. The mean vectors can be used to capture correct boundary conditions (Gunzburger, Peterson & Shadid 2007).

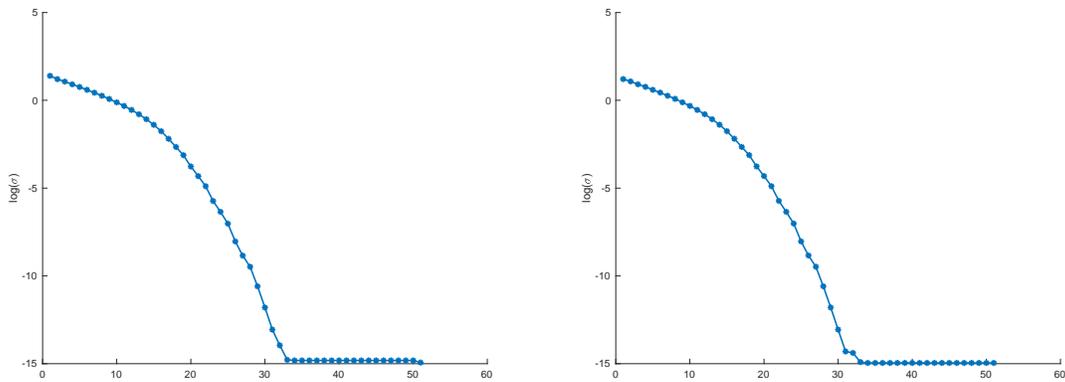


Figure 3.1 Singular values between raw snapshots (left) and mean-centered snapshots (right).

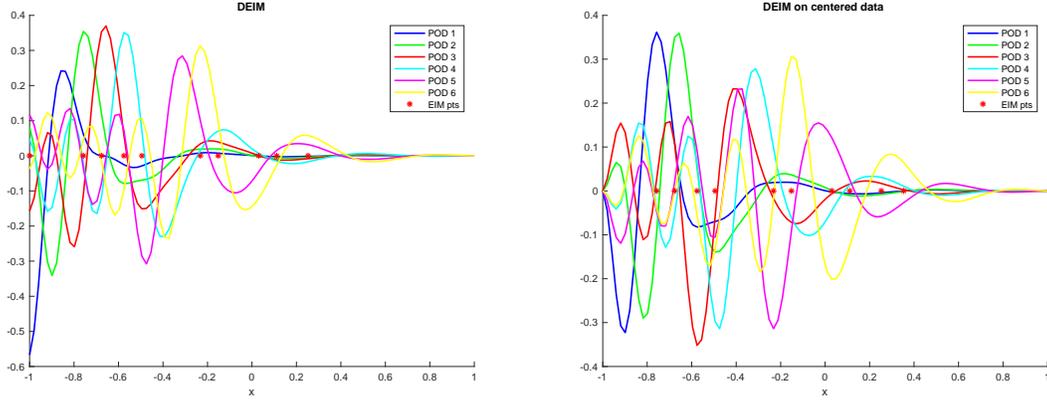


Figure 3.2 Comparison of DEIM between raw snapshots (left) and mean-centered snapshots (right).

2D nonlinear parametrized function

Following (Peherstorfer et al. 2014), a nonlinear parametrized function $g^1 : \Omega \times \mathcal{D} \mapsto \mathbb{R}$ defined in $\Omega = [0,1]^2 \subset \mathbb{R}^2$, where $\mathcal{D} = [0,1]^2 \subset \mathbb{R}^2$ is a parameter domain, is given by

$$g^1(\mathbf{x}; \boldsymbol{\mu}) = \frac{1}{\sqrt{((1-x_1) - (0.99\mu_1 - 1))^2 + ((1-x_2) - (0.99\mu_2 - 1))^2 + 0.1^2}}. \quad (3.27)$$

The parameter $\boldsymbol{\mu} = (\mu_1, \mu_2) \in \mathcal{D}$ controls the gradient of the peak near the corner (1,1) of the spatial domain Ω . The nonlinear functions g^1 are discretized by a 20×20 equidistant grid ($\mathcal{N} = 400$) in Ω and sampled on a 25×25 ($N_s = 625$) equidistant grid in \mathcal{D} . From the 625 snapshots of nonlinear function, stored in a snapshot matrix \mathbf{X}_s , we build the associated POD basis and DEIM approximations.

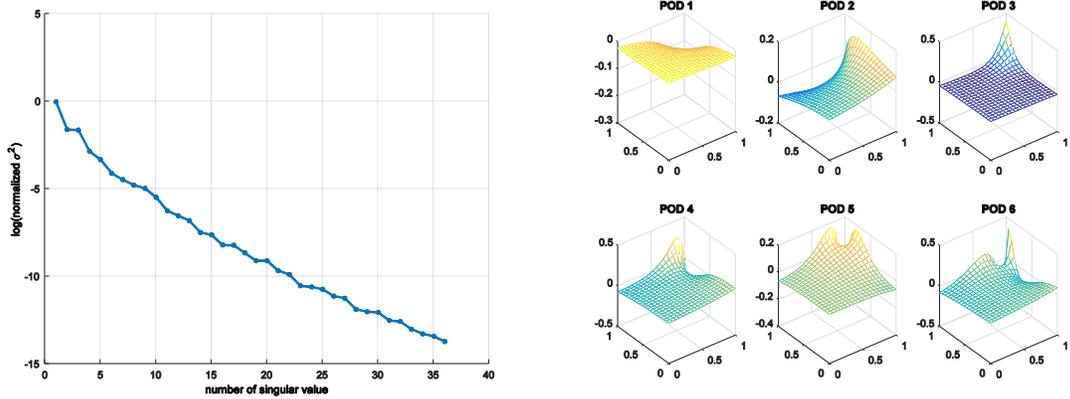


Figure 3.3 The normalized singular value $\sigma_i^2 / (\sum_i \sigma_i^2)$ (left) and the first six POD basis vectors (right) of the of the snapshot matrix \mathbf{X}_s corresponding to the nonlinear function g^1 .

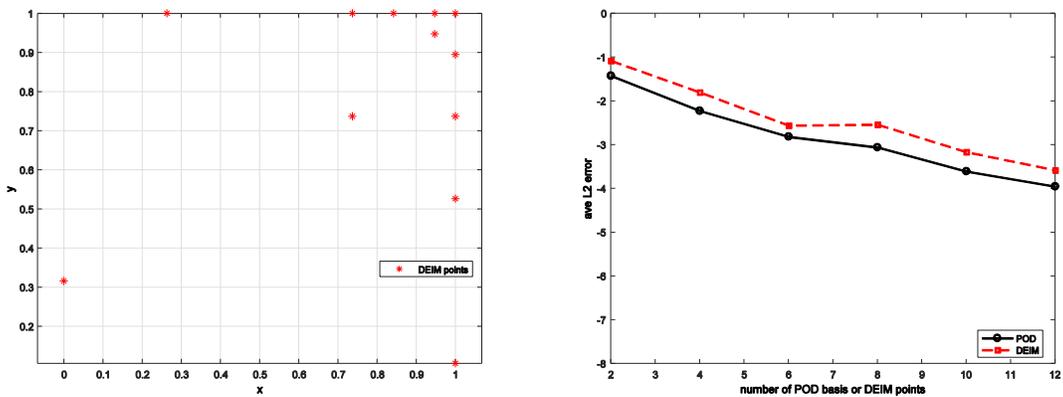


Figure 3.4 First 12 DEIM points (left) selected by the greedy algorithm for the nonlinear function g^1 , and the comparison of average L^2 errors of POD and DEIM approximations (right) for the training data.

Figure 3.3 shows the singular value and the first 6 POD basis vectors obtained by using SVD on the snapshot matrix of the nonlinear function g^1 . The singular value decays rapidly because the given function g^1 only exhibits a localized peak near the corner $\mathbf{x} = [1,1]$ and the first several POD basis vectors are able to well capture the localized features, as shown in Figure 3.3. Accordingly, the “sensor” points chosen by DEIM (see Figure 3.4 (left)) are primarily

distributed near the region with localized features as if the greedy algorithm is “aware” of the most crucial region and assign more sensors on that region. In consequence, the DEIM approximation shows a good reconstruction results as shown in Figure 3.4 (right) while only need $\hat{k} / \mathcal{N} \approx 3\%$ ($\hat{k} = 12$) evaluations compared to POD method.

3.4.3 Cubature-type strategy

Instead of directly approximating the assembled nonlinear vector $\mathbf{f} \in \mathbb{R}^{\mathcal{N}}$ with the interpolation indices and the empirical basis, an alternative hyper-reduction strategy to reduce the associated evaluation cost is to approximate the integral forming the nonlinear vector as a weighted sum of the integrand evaluated at optimal sampling points. Since it follows the classical recipe of Gaussian quadrature of polynomial functions, this type of hyper-reduction referred to as the *cubature* strategy. Recall the POD reduced term in (3.17), the idea here can be expressed as:

$$\hat{\mathbf{f}}(\mathbf{u}^r) = \mathbf{V}^T \mathbf{f}(\mathbf{V}\mathbf{u}^r) = \sum_{g=1}^{N_g} \mathbf{V}_{\mathcal{I}_g}^T \mathbf{f}_g(\mathbf{V}\mathbf{u}^r) w_g \approx \sum_{g \in \hat{\mathcal{G}}} \mathbf{V}_{\mathcal{I}_g}^T \mathbf{f}_g(\mathbf{V}\mathbf{u}^r) \hat{\zeta}_g, \quad (3.28)$$

where $\hat{\mathcal{G}}$ ($|\hat{\mathcal{G}}| = \hat{m}$) is the subset of the whole set of Gauss integration points \mathcal{G} to be selected, and $\hat{\zeta}_g$ are the associated positive weights. The other notations have been presented before.

The main idea behind the cubature hyper-reduction method is due to an important observation that each row of the reduced vector $\hat{\mathbf{f}} \in \mathbb{R}^k$ can be interpreted as the virtual work done by the force $\mathbf{f} \in \mathbb{R}^{\mathcal{N}}$ acting on each column of $\mathbf{V} \in \mathbb{R}^{\mathcal{N} \times k}$, seen as a virtual displacement. In the light of considering $\hat{\mathbf{f}}$ as global energy of only k dimension ($k \ll N_g$), one can

approximate it by the sum of the energy contributions of only a small subset of Gauss points, i.e. $\hat{\mathcal{G}}$, weighted by positive scalar accordingly.

To the best of the author' knowledge, the first instance of this type of model reduction was proposed in (An, Kim & James 2008) for the computer graphics applications, and Farhat and co-workers (Farhat et al. 2014; Farhat, Chapman & Avery 2015) recently introduced the similar idea in dynamic systems called energy-conserving mesh sampling and weighting (ECSW) hyper reduction. The cubature-type methods have also been applied to computational multiscale problems (Hernández, Caicedo & Ferrer 2017; Oliver et al. 2017; van Tuijl, Remmers & Geers 2018), where they are called empirical cubature or reduced optimal quadrature (ROQ) methods. The main idea of these methods is given as follows.

Optimal quadrature: a special least square problem

The indices $\hat{\mathcal{G}}$ and the associated weights $\{\hat{\zeta}_g\}_{g \in \hat{\mathcal{G}}}$ are the unknowns need to be determined. Let the integrand of the reduced nonlinear function $\hat{\mathbf{f}}(\boldsymbol{\mu}) = \mathbf{V}^T \mathbf{f}(\mathbf{V}\mathbf{u}^r(\boldsymbol{\mu}))$ be a k -dimensional parameterized function $\mathbf{g}(\mathbf{x}; \boldsymbol{\mu}) \in [L^2(\Omega)]^k \subset \mathbb{R}^k$, such that by using the Gaussian quadrature scheme the domain integration of $\mathbf{g}(\mathbf{x}; \boldsymbol{\mu}) = [g_1(\mathbf{x}; \boldsymbol{\mu}) \dots g_k(\mathbf{x}; \boldsymbol{\mu})]^T$ yields

$$\hat{f}_l(\boldsymbol{\mu}) = \int_{\Omega} g_l(\mathbf{x}; \boldsymbol{\mu}) d\Omega := \sum_{g=1}^{N_g} g_l(\mathbf{x}_g; \boldsymbol{\mu}) w_g, \quad (3.29)$$

where $f_l(\boldsymbol{\mu})$ denotes the l -th component of the integral output $\hat{\mathbf{f}}(\boldsymbol{\mu}) \in \mathbb{R}^k$, $\{w_g\}_{g=1}^{N_g}$ are the weights of Gaussian quadrature, and $\boldsymbol{\mu} \in \mathcal{D}$ denotes generic parameters that influence the

function. If we consider a set of training parameters $\{\boldsymbol{\mu}_i\}_{i=1}^P$, and denote $g_I^i(\mathbf{x}) = g_I(\mathbf{x}; \boldsymbol{\mu}_i)$ and $f_I^i = f_I(\boldsymbol{\mu}_i)$ for brevity, we have

$$f_I^i = \int_{\Omega} g_I^i(\mathbf{x}) d\Omega = \sum_{g=1}^{N_g} g_I^i(\mathbf{x}_g) w_g, \quad (3.30)$$

$i=1, \dots, P$ and $I=1, \dots, k$. These kP equalities can be obtained during the offline phase. As pointed out before, the idea of *Cubature integration* proposed by (An, Kim & James 2008) consists in approximating the integral of any $g_I^i: \Omega \rightarrow \mathbb{R}$ as the sum of positive, scalar weights multiplied by the function evaluated at appropriately chosen points, i.e.,

$$f_I^i = \int_{\Omega} g_I^i(\mathbf{x}) d\Omega \approx \sum_{g \in \hat{\mathcal{G}}} g_I^i(\mathbf{x}_g) \hat{\zeta}_g, \quad I=1, \dots, k \text{ and } i=1, \dots, P. \quad (3.31)$$

The locations of the integration points $\{\mathbf{x}_g\}_{g \in \hat{\mathcal{G}}}$ and their associated positive weights $\{\hat{\zeta}_g\}_{g \in \hat{\mathcal{G}}} \subset \mathbb{R}_{++}^{\hat{m}}$, where $|\hat{\mathcal{G}}| = \hat{m}$, are determined such that the integration error for all components is minimized. If we augment the cubature weights $\{\hat{\zeta}_g\}_{g \in \hat{\mathcal{G}}} \subset \mathbb{R}_{++}^{\hat{m}}$ with $(N_g - \hat{m})$ zero components and define those weights by a vector $\boldsymbol{\zeta} \in \mathbb{R}_+^{N_g}$, the regression problem in (3.31) to determine $\{\hat{\zeta}_g\}_{g \in \hat{\mathcal{G}}}$ with only \hat{m} non-zero coefficients can be cast in a *non-negative least squares* (NNLS) problem as follow:

Given a matrix $\mathbf{G} \in \mathbb{R}^{kP \times N_g}$ (usually $N_g \gg kP$) and a vector $\mathbf{b} \in \mathbb{R}^{kP}$, find a nonnegative vector $\boldsymbol{\zeta} \in \mathbb{R}_+^{N_g}$ to minimize the following functional, i.e.

$$\begin{aligned} & \arg \min_{\boldsymbol{\zeta} \in \mathbb{R}_+^{N_g}} \|\mathbf{G}\boldsymbol{\zeta} - \mathbf{b}\|, \\ & \text{subject to } \zeta_i \geq 0, \quad i=1, \dots, N_g, \end{aligned} \quad (3.32)$$

where $\|\cdot\|$ stands for the standard Euclidean norm, and

$$\mathbf{G} := \begin{bmatrix} \mathbf{g}^1(\mathbf{x}_1) & \mathbf{g}^1(\mathbf{x}_2) & \cdots & \mathbf{g}^1(\mathbf{x}_{N_g}) \\ \mathbf{g}^2(\mathbf{x}_1) & \mathbf{g}^2(\mathbf{x}_2) & \cdots & \mathbf{g}^2(\mathbf{x}_{N_g}) \\ \cdots & \cdots & \ddots & \cdots \\ \mathbf{g}^P(\mathbf{x}_1) & \mathbf{g}^P(\mathbf{x}_2) & \cdots & \mathbf{g}^P(\mathbf{x}_{N_g}) \end{bmatrix}, \mathbf{b} := \begin{bmatrix} \mathbf{f}^1 \\ \mathbf{f}^2 \\ \vdots \\ \mathbf{f}^P \end{bmatrix}, \quad (3.33)$$

in which $\mathbf{g}^i = [g_1^i, g_2^i, \dots, g_k^i]^T$ and $\mathbf{f}^i = [f_1^i, f_2^i, \dots, f_k^i]^T$, $i = 1, \dots, P$. The indices of the non-zero entries of $\boldsymbol{\zeta}$ is stored in a set $\hat{\mathcal{G}}$.

It should be noted that since the NNLS problem in (3.32) is underdetermined system due to $N_g \gg kP$, non-unique solutions are expected. An exact solution to (3.32) is given by the Gaussian weights $\{w_g\}_{g=1}^{N_g}$, i.e. $\zeta_g = w_g$, $g = 1, \dots, N_g$, but it is of no interest as it implies that the reduced integration set $\hat{\mathcal{G}}$ just match the full integration set \mathcal{G} . However, the NNLS problem with non-negative constraints yields a sparse solution with only a few non-zero entries.

In practice, we also want to impose an additional constrain $\sum_{g=1}^{N_g} \zeta_g = \sum_{g=1}^{N_g} w_g = \Omega$ for satisfying the volume conservation, which is important to maintain structure-preserving properties, such as the conservation of energy, momentum, and mass. To this end, the minimization problem in (3.32) is modified as

$$\begin{aligned} & \arg \min_{\boldsymbol{\zeta} \in \mathbb{R}^{N_g}} \|\mathbf{G}\boldsymbol{\zeta} - \mathbf{b}\|^2 + \lambda \|\mathbf{I}^T \boldsymbol{\zeta} - \Omega\|^2, \\ & \text{subject to } \zeta_i \geq 0, i = 1, \dots, N_g, \end{aligned} \quad (3.34)$$

where $\mathbf{I} = [1, 1, \dots, 1]^T \in \mathbb{R}^{N_g}$, Ω is a constant denoted as the volume of the physical domain, and λ is a penalty coefficient. The penalized minimization (3.35), in fact, can be recast as a standard NNLS by augmenting \mathbf{G} and \mathbf{b} with an additional row, i.e.

$$\begin{aligned} & \arg \min_{\zeta \in \mathbb{R}^{N_g}} \| \mathbf{G}^{aug} \zeta - \mathbf{b}^{aug} \|^2, \\ & \text{subject to } \zeta_i \geq 0, i = 1, \dots, N_g, \end{aligned} \quad (3.36)$$

where

$$\mathbf{G}^{aug} := \begin{bmatrix} \mathbf{G} \\ \sqrt{\lambda} \mathbf{I}^T \end{bmatrix} \in \mathbb{R}^{(kP+1) \times N_g}, \mathbf{b}^{aug} := \begin{bmatrix} \mathbf{b} \\ \sqrt{\lambda} \Omega \end{bmatrix} \in \mathbb{R}^{kP+1}. \quad (3.37)$$

The algorithm for solving the NNLS problem (3.37) and determining the sparse solution $\zeta \in \mathbb{R}_+^{N_g}$ with \hat{m} ($\hat{m} \ll N_g$) non-zero entries is presented in Appendix B.

A detailed comparison about accuracy and efficiency between interpolation-type and cubature-type hyper-reduction techniques in the context of micro-structural analyses has been made recently (van Tuijl, Remmers & Geers 2018). It shown that the interpolation-type approach usually yields less error and higher online efficiency due to the extra POD information from the nonlinear contribution, but the cubature-type approach naturally preserves the stability of the HDM without any additional stabilization needed for the interpolation-type approach. Moreover, the latter one is less sensitive to the extrapolation of the snapshot space.

3.4.4 Linearization-type strategy

There is, in fact, another strategy to overcome the computational inefficiency in reduced-order modeling of nonlinear system by approximating nonlinear function in conjunction with the POD-Galerkin projection method. This strategy can be viewed as approximating the parametric nonlinear function by using Taylor expansion, or assuming multilinear forms of polynomial nonlinearities (Bai 2002; Cardoso & Durlofsky 2010; Chen & White 2000; Phillips 2000). The

typical methods include the trajectory piecewise-linear approach (TPWL) (Rewiński & White 2003; Rewieński & White 2006), which is based on approximating a nonlinear function by a weighted combination of linearized models at selected parameter points along a state trajectory. These points are selected using prior knowledge from the training data of the high-dimensional nonlinear system. However, because the resulting ROM simply uses low degree piecewise polynomials (Chaturantabut & Sorensen 2010) and ignores the HDM information away from the linearization points, this method in principle lacks robustness for highly nonlinear problems such as the parametric systems arising from CFD applications and solid mechanics.

3.5 Summary

This chapter has reviewed the two key concepts of model order reduction, the projection-based reduction methods and the hyper-reduction techniques. The projection-based methods are mainly used for reduced approximation of state variables, and the reduced system associated to the original PDEs can be obtained by using Galerkin or Petrov-Galerkin method with the dual projection. The latter ones are used for conducting additional level of approximation, i.e. system approximation, when dealing nonlinear system. By combining these two-level approximations as well as taking the underlying discretization into account, many nonlinear model reduction methods have been developed for a variety of applications.

Chapter 4

Model Order Reduction for Fracture Mechanics via Decomposed Projection

The preservation of physical characteristic in reduced-order models remains challenging for most MOR methods, especially dealing with engineering problems where solutions exhibit singularity and discontinuities. In this chapter, we propose a SVD-based model order reduction (MOR) approach (He, Chen & Marodon 2018) for reduced-order modeling of fracture mechanics based on the integrated singular basis function method (ISBFM) with reproducing kernel approximation enriched by crack-tip basis functions. The reduced basis vectors used in the proposed MOR method is properly constructed based on the decomposition of the subspaces that characterize smooth and non-smooth solutions, respectively, and thus, it is called a *decomposition subspace reduction* (DSR) method.

4.1 Introduction

Repeated simulations of the problem of interest under various conditions are required in many engineering applications but are computationally unaffordable. Model order reduction (MOR) methods (Antoulas & Sorensen 2001; Schilders, van der Vorst & Rommes 2008) offer effective means in resolving the complexity issue while minimizing the loss of accuracy. However, for fracture problems with a large ratio between the scale of the structure and the scale

of cracks, the application of MOR methods remain difficult in preserving the local features in the reduced-order space.

The MOR strategies based on the projection of the fine-scale finite dimensional space onto a low-dimensional subspace have undergone significant development in recent years. They have been extensively studied in, for example, system engineering (Antoulas & Sorensen 2001), fluid mechanics (Amsallem & Farhat 2008; Lieu, Farhat & Lesoinne 2006; Sirovich 1987), and structures dynamics (Carlberg & Farhat 2011; Krysl, Lall & Marsden 2001; Lall, Krysl & Marsden 2003). Attempts were also made for highly nonlinear structural problems (Kerfriden et al. 2011; Millán & Arroyo 2013; Niroomandi et al. 2010; Ryckelynck, Benziane & Paristech 2010). These strategies rely on an assumption that the solution of the fine-scale discretization can be properly approximated in the low-dimensional subspace spanned by well-chosen projection bases (called Ritz vectors). Within this framework, the projection-based MOR methods have been proposed under two categories (Galland et al. 2011): the *a posteriori* and *a priori* approaches. In the *a posteriori* approach, preliminary computations of the fine-scale model are performed in the *offline* phase to build the basis of a reduced subspace based on the precomputed solution data, and the reduced model constructed on the subspace is then used repeatedly in the *online* phase to obtain accelerated solution. The snapshot proper orthogonal decomposition (POD) method, also called Karhunen-Loeve expansion, principal component analysis (PCA), and empirical orthogonal eigenfunctions, is a typical *a posteriori* approach (Everson & Sirovich 1995; Krysl, Lall & Marsden 2001; Sirovich 1987). On the contrary, the *a priori* approach doesn't need the preliminary computations of solutions, and therefore no offline/online computational decomposition. The reduced basis is constructed by using some *a priori* knowledge about the particular problem. Examples include the *a priori* hyperreduction method

(Ryckelynck 2005; Ryckelynck et al. 2006), the modal truncation method (Craig 1981; Dickens, Nakagawa & Wittbrodt 1997), and the proper generalized decomposition method (Chinesta, Ammar & Cueto 2010) based on the reduced solution space by separation of variables.

For fracture problems, the level of reducibility depends strongly on the size of the region of interest (Kerfriden et al. 2013), but this contradicts the fundamental approach of the reduced-order modeling. To circumvent the difficulties, the sub-structuring methods (Barbič & Zhao 2011; Kim & James 2012; Rixen 2004) based on the idea of domain partitioning, together with the projection based MOR methods, have been developed for fracture problems. The hybrid local/global MOR methods (Galland et al. 2011; Kerfriden et al. 2011; Kerfriden, Passieux & Bordas 2012; Niroomandi et al. 2012) were proposed for the simulation of localized failure, where the global (slave) problem is solved in a reduced subspace constructed by the classical POD method while the local (master) damage behavior are resolved in a fine-scale level. Kerfriden et al. (Kerfriden et al. 2013) developed a partitioned MOR approach using multiple reduced subspaces based on the Schur-based domain decomposition and adopted “gappy” approximations (Chaturantabut & Sorensen 2010; Everson & Sirovich 1995) to reduce the computational burden. Other methods based on the computational multiscale framework (Akbari Rahimabadi, Kerfriden & Bordas 2015; Oliver et al. 2017) to simulate fracture have also been proposed.

Different from the idea of domain decomposition in the above strategies, the present approach explicitly represents the crack characteristics by a decomposition of solution into smooth and non-smooth parts, and constructs a projection-based MOR framework to preserve the discontinuity and singularity characteristics of fracture in the reduced-order model (ROM). In this study, the integrated singular basis function method (ISBFM) (Chen, Marodon & Hu 2015;

Georgiou, Olson & Smyrlis 1996; Olson, Georgiou & Schultz 1991) is introduced to obtain the fine-scale solution of the fracture problem. Under the ISBFM framework, a set of enrichment functions are derived from a local asymptotic solution near the crack tip to meet the equilibrium and traction-free conditions on the crack surfaces, leading to a weak formulation where the domain integral involving the near-tip enrichment functions is transformed to the boundary integrals without containing the crack surfaces. As such, the need of high-order quadrature scheme for domain integration and boundary integration on crack surfaces is avoided, and it results in a sparser discrete matrix system to allow effective MOR procedures. More specifically, the proposed reduction method considers two decomposed projections constructed based on the Schur decomposition of the fine-scale system, allowing the low-rank representation of the smooth subspace while optimally preserving the non-smooth subspace, termed the decomposed subspace reduction (DSR) method. It is shown that the standard model reduction technique, whose reduced-order projection is obtained directly from the modal analysis of the fine-scale discrete system (also called the mode displacement method in dynamics (Craig 1981; Dickens, Nakagawa & Wittbrodt 1997)), performs poorly due to the loss of singularity and discontinuity characteristics. The error analysis also shows that the modal basis of the standard reduction method is sensitive to the scaling of the enrichment functions, whereas the DSR reduced approximation is invariant to the scaling effect and the non-smooth behavior can be robustly captured.

4.2 Fine-Scale Modeling of Fracture Mechanics

In this section, the solution decomposition scheme together with the integrated singular basis function method (ISBFM) (Chen, Marodon & Hu 2015; Georgiou, Olson & Smyrlis 1996; Olson, Georgiou & Schultz 1991) are introduced to construct the full-order model and obtain the fine-scale solution of linear elastic fracture mechanics (LEFM).

4.2.1 Problem formulation based on decomposed solution scheme

The LEFM problem as illustrated in Figure 4.1 is described by a boundary value problem as follows

$$\begin{aligned}
 \sigma_{ij,j} &= 0 & \text{in } \Omega \\
 u_i &= g_i & \text{on } \Gamma_{g_i} \\
 \sigma_{ij}n_j &= h_i & \text{on } \Gamma_{h_i}
 \end{aligned} \tag{4.1}$$

In above equations, u_i is the displacement vector component, σ_{ij} is the stress tensor component expressed as $\sigma_{ij} = C_{ijkl}u_{(k,l)}$, C_{ijkl} is the elasticity tensor component, $u_{(i,j)} = (u_{i,j} + u_{j,i})/2$ is the strain tensor component, $u_{i,j} = \partial u_i / \partial x_j$, Ω is the problem domain, Γ_{g_i} and Γ_{h_i} are the Dirichlet and Neumann boundaries, respectively, and g_i and h_i are the associated prescribed displacement and traction components, respectively. Noted that in Figure 4.1, the crack surfaces Γ_{C+} and Γ_{C-} are traction free and $\Gamma_{C+} \cup \Gamma_{C-} = \Gamma_C \subset \Gamma_{h_i}$.

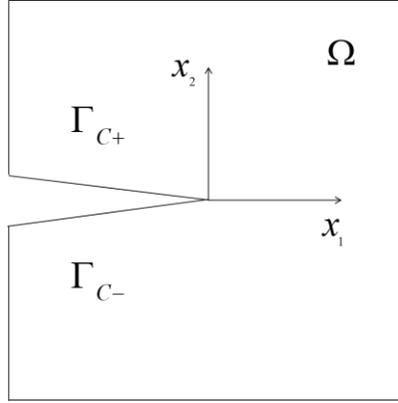


Figure 4.1: Geometry of a cracked domain

The potential energy functional corresponds to (4.1) with Nitsche's method (Fernández-Méndez & Huerta 2004; Nitsche 1971) employed for imposing Dirichlet boundary conditions on Γ_{g_i} is expressed as

$$\begin{aligned} \Pi = & \frac{1}{2} \int_{\Omega} u_{(i,j)} C_{ijkl} u_{(k,l)} d\Omega - \int_{\Gamma_{h_i}} u_i h_i d\Gamma s \\ & - \int_{\Gamma_{g_i}} \sigma_{ij} n_j (u_i - g_i) d\Gamma + \frac{\beta}{2} \int_{\Gamma_{g_i}} (u_i - g_i)(u_i - g_i) d\Gamma \end{aligned} \quad (4.2)$$

where β is a positive parameter to ensure the coercivity. The variational form of (4.2) is

$$\begin{aligned} & \int_{\Omega} \delta u_{(i,j)} C_{ijkl} u_{(k,l)} d\Omega - \int_{\Gamma_{g_i}} \delta \sigma_{ij} n_j u_i d\Gamma - \int_{\Gamma_{g_i}} \delta u_i \sigma_{ij} n_j d\Gamma + \beta \int_{\Gamma_{g_i}} \delta u_i u_i d\Gamma \\ & = \int_{\Gamma_{h_i}} \delta u_i h_i d\Gamma - \int_{\Gamma_{g_i}} \delta \sigma_{ij} n_j g_i d\Gamma + \beta \int_{\Gamma_{g_i}} \delta u_i g_i d\Gamma \end{aligned} \quad (4.3)$$

Let u_i^h be the approximation of u_i , which is decomposed into a smooth part \bar{u}_i^h and a non-smooth part \hat{u}_i^h as

$$u_i^h = \bar{u}_i^h + \hat{u}_i^h \quad (4.4)$$

In this work, the reproducing kernel (RK) approximation is employed to discretize the

smooth part of the solution \bar{u}_i^h as follows

$$\bar{u}_i^h(\mathbf{x}) = \sum_{I=1}^{\bar{N}} \Psi_I(\mathbf{x}) \bar{d}_{Ii} \text{ or } \bar{\mathbf{u}}^h(\mathbf{x}) = \sum_{I=1}^{\bar{N}} \Psi_I \bar{\mathbf{d}}_I \quad (4.5)$$

where \bar{N} is the number of discrete points, \bar{d}_{Ii} is the nodal coefficient for the smooth approximation associated with node I , and Ψ_I is the RK shape function. The cubic B-spline with C^2 continuity is used as the kernel function in this work. To properly account for the non-convex geometry and introduce the discontinuities across crack surfaces, visibility criterion (Belytschko et al. 1996) is used to truncate the kernel function supports.

In general, the non-smooth solution \hat{u}_i^h can be approximated by a set of near-tip basis functions $\{F_{ji}^s, F_{ji}^{as}\}_{j=1}^{\hat{N}}$,

$$\hat{u}_i^h = \sum_{J=1}^{\hat{N}} F_{ji}^s \hat{d}_j^s + \sum_{J=1}^{\hat{N}} F_{ji}^{as} \hat{d}_j^{as} \text{ or } \hat{\mathbf{u}}^h = \sum_{J=1}^{\hat{N}} \mathbf{F}_J \hat{\mathbf{d}}_J \quad (4.6)$$

where $\hat{\mathbf{d}}_J = [\hat{d}_j^s \quad \hat{d}_j^{as}]^T$ and $\mathbf{F}_J = [\mathbf{F}_J^s \quad \mathbf{F}_J^{as}]$ with superscripts ‘‘s’’ and ‘‘as’’ denote the symmetric and anti-symmetric properties, respectively. Here, the near-tip basis functions are required to satisfy the homogeneous equilibrium and the traction free conditions on crack surfaces, i.e.,

$$\begin{aligned} C_{ijkl} F_{J(k,l),j}^s &= 0 \quad \text{and} \quad C_{ijkl} F_{J(k,l),j}^{as} = 0 \quad \text{in } \Omega \\ C_{ijkl} F_{J(k,l)}^s n_j &= 0 \quad \text{and} \quad C_{ijkl} F_{J(k,l)}^{as} n_j = 0 \quad \text{on } \Gamma_C \end{aligned} \quad (4.7)$$

To fulfill these conditions, the local asymptotic basis functions derived from William’s solution (Williams 1952, 1957) are introduced here as follows

$$\begin{aligned}
\mathbf{F}_J^s &= \begin{bmatrix} F_{J1}^s(r, \theta) \\ F_{J2}^s(r, \theta) \end{bmatrix} = \begin{bmatrix} r^{J/2} \left[\left(\kappa + \frac{J}{2} + (-1)^J \right) \cos\left(\frac{J}{2}\right) - \frac{J}{2} \cos\left(\left(\frac{J}{2}-2\right)\theta\right) \right] \\ r^{J/2} \left[\left(\kappa - \frac{J}{2} - (-1)^J \right) \sin\left(\frac{J}{2}\right) + \frac{J}{2} \sin\left(\left(\frac{J}{2}-2\right)\theta\right) \right] \end{bmatrix} \\
\mathbf{F}_J^{as} &= \begin{bmatrix} F_{J1}^{as}(r, \theta) \\ F_{J2}^{as}(r, \theta) \end{bmatrix} = \begin{bmatrix} -r^{J/2} \left[\left(\kappa + \frac{J}{2} - (-1)^J \right) \sin\left(\frac{J\theta}{2}\right) - \frac{J}{2} \sin\left(\left(\frac{J}{2}-2\right)\theta\right) \right] \\ r^{J/2} \left[\left(\kappa - \frac{J}{2} + (-1)^J \right) \cos\left(\frac{J\theta}{2}\right) + \frac{J}{2} \cos\left(\left(\frac{J}{2}-2\right)\theta\right) \right] \end{bmatrix}
\end{aligned} \tag{4.8}$$

where $J = 1, 2, \dots, \hat{N}$, and κ is the Kolosov's constant for plane problems. These local asymptotic basis functions are employed for the non-smooth approximation (4.6) under the ISBFM framework to be introduced in Section 4.2.2.

Note that the near-tip basis functions corresponding to even integers of J in (4.8) can be reproduced by the monomial bases and thus are not considered in approximation since they are linearly dependent with the smooth part basis functions in (4.5). Moreover, the first symmetric and anti-symmetric near-tip basis functions, i.e. \mathbf{F}_1^s and \mathbf{F}_1^{as} , represent exactly the crack opening displacements of Mode I and the sliding Mode II. In general, both symmetric and anti-symmetric basis functions are considered in the non-smooth solution \hat{u}_i^h . However, if the symmetry characteristic of the solution is known *a priori*, only the basis functions with proper symmetry are employed in the non-smooth approximation.

4.2.2 ISBFM based Galerkin formulation

With the solution decomposition in (4.4) and considering arbitrary variation of $\delta \bar{u}_i^h$ and $\delta \hat{u}_i^h$, the *standard Galerkin formulation* of (4.3) can be obtain as follows

$$\begin{aligned}
& \int_{\Omega} \delta \bar{u}_{(i,j)}^h C_{ijkl} (\bar{u}_{(k,l)}^h + \hat{u}_{(k,l)}^h) d\Omega \\
& - \int_{\Gamma_{g_i}} \delta \bar{\sigma}_{ij}^h n_j (\bar{u}_i^h + \hat{u}_i^h) d\Gamma - \int_{\Gamma_{g_i}} \delta \bar{u}_i^h C_{ijkl} (\bar{u}_{(k,l)}^h + \hat{u}_{(k,l)}^h) n_j d\Gamma + \beta \int_{\Gamma_{g_i}} \delta \bar{u}_i^h (\bar{u}_i^h + \hat{u}_i^h) d\Gamma \quad (4.9) \\
& = \int_{\Gamma_{h_i}} \delta \bar{u}_i^h h_i d\Gamma - \int_{\Gamma_{g_i}} \delta \bar{\sigma}_{ij}^h n_j g_i d\Gamma + \beta \int_{\Gamma_{g_i}} \delta \bar{u}_i^h g_i d\Gamma
\end{aligned}$$

$$\begin{aligned}
& \int_{\Omega} \delta \hat{u}_{(i,j)}^h C_{ijkl} (\bar{u}_{(k,l)}^h + \hat{u}_{(k,l)}^h) d\Omega \\
& - \int_{\Gamma_{g_i}} \delta \hat{\sigma}_{ij}^h n_j (\bar{u}_i^h + \hat{u}_i^h) d\Gamma - \int_{\Gamma_{g_i}} \delta \hat{u}_i^h C_{ijkl} (\bar{u}_{(k,l)}^h + \hat{u}_{(k,l)}^h) n_j d\Gamma + \beta \int_{\Gamma_{g_i}} \delta \hat{u}_i^h (\bar{u}_i^h + \hat{u}_i^h) d\Gamma \quad (4.10) \\
& = \int_{\Gamma_{h_i}} \delta \hat{u}_i^h h_i d\Gamma - \int_{\Gamma_{g_i}} \delta \hat{\sigma}_{ij}^h n_j g_i d\Gamma + \beta \int_{\Gamma_{g_i}} \delta \hat{u}_i^h g_i d\Gamma
\end{aligned}$$

where $\bar{\sigma}_{ij}^h = C_{ijkl} \bar{u}_{(k,l)}^h$ and $\hat{\sigma}_{ij}^h = C_{ijkl} \hat{u}_{(k,l)}^h$. Following (Chen, Marodon & Hu 2015; Georgiou, Olson & Smyrlis 1996; Olson, Georgiou & Schultz 1991), if the near-tip basis functions, such as those given in (10), for the non-smooth solution \hat{u}_i^h are so selected to satisfy the homogeneous equilibrium equation, i.e. $\hat{\sigma}_{ij,j}^h = C_{ijkl} \hat{u}_{(k,l),j}^h = 0$ in Ω , the domain integral involving the non-smooth approximation functions \hat{u}_i^h can be transformed to a boundary integral by means of integration by parts and divergence theorem to yield

$$\int_{\Omega} \delta u_{(i,j)} C_{ijkl} \hat{u}_{(k,l)}^h d\Omega = \int_{\Gamma_{g_i} + \Gamma_{h_i}} \delta u_i C_{ijkl} \hat{u}_{(k,l)}^h n_j d\Gamma \quad (4.11)$$

If the non-smooth solution \hat{u}_i^h further satisfies the homogeneous Neumann boundary conditions on crack boundaries accordingly, i.e. $C_{ijkl} \hat{u}_{(k,l)}^h n_j = 0$ on Γ_C , such as those in (4.8), equations (4.9) and (4.10) can be rearranged as

$$\begin{aligned}
& \int_{\Omega} \delta \bar{u}_{(i,j)}^h C_{ijkl} \bar{u}_{(k,l)}^h d\Omega - \int_{\Gamma_{g_i}} \left(\delta \bar{u}_i^h (C_{ijkl} \bar{u}_{(k,l)}^h n_j - \beta \bar{u}_i^h) + \delta \bar{\sigma}_{ij}^h n_j \bar{u}_i^h \right) d\Gamma \\
& + \int_{\Gamma_{h_i}} \delta \bar{u}_i^h C_{ijkl} \hat{u}_{(k,l)}^h n_j d\Gamma - \int_{\Gamma_{g_i}} \left(\delta \bar{\sigma}_{ij}^h n_j \hat{u}_i^h - \delta \bar{u}_i^h \beta \hat{u}_i^h \right) d\Gamma \quad (4.12) \\
& = \int_{\Gamma_{h_i}} \delta \bar{u}_i^h h_i d\Gamma - \int_{\Gamma_{g_i}} \left(\delta \bar{\sigma}_{ij}^h n_j g_i - \delta \bar{u}_i^h \beta g_i \right) d\Gamma
\end{aligned}$$

$$\begin{aligned}
& \int_{\bar{\Gamma}_{h_i}} \delta \hat{u}_{(i,j)}^h C_{ijkl} \bar{u}_k^h n_l d\Gamma - \int_{\Gamma_{s_i}} \left(\delta \hat{\sigma}_{ij}^h n_j \bar{u}_i^h - \delta \hat{u}_i^h \beta \bar{u}_i^h \right) d\Gamma \\
& + \int_{\bar{\Gamma}_{h_i}} \delta \hat{u}_i^h C_{ijkl} \hat{u}_{(k,l)}^h n_j d\Gamma - \int_{\Gamma_{s_i}} \left(\delta \hat{\sigma}_{ij}^h n_j \hat{u}_i^h - \delta \hat{u}_i^h \beta \hat{u}_i^h \right) d\Gamma \\
& = \int_{\bar{\Gamma}_{h_i}} \delta \hat{u}_i^h h_i d\Gamma - \int_{\Gamma_{s_i}} \left(\delta \hat{\sigma}_{ij}^h n_j g_i - \delta \hat{u}_i^h \beta g_i \right) d\Gamma
\end{aligned} \tag{4.13}$$

where $\bar{\Gamma}_{h_i} = \Gamma_{h_i} \setminus \Gamma_C$. In the above two equations, the domain integral simply involves the derivatives of the smooth solution \bar{u}_i^h , whereas the non-smooth solution \hat{u}_i^h and its derivatives only need to be integrated on the boundaries away from the crack tip that exhibits singularities. As a result, this ISBFM Galerkin formulation can be numerically integrated using a much lower-order quadrature scheme compared with that needed for the standard Galerkin formulation in (4.9) and (4.10). The comparison between the standard Galerkin and ISBFM Galerkin formulations with different orders of Gauss quadrature scheme in fracture modeling is made in Section 4.2.3.

By introducing the RK approximation in (4.5) enriched with the near-tip basis functions in (4.6) into the ISBFM Galerkin equations in (4.12) and (4.13), the ISBFM discrete system is obtained as

$$\mathbf{Kd} = \begin{bmatrix} \bar{\mathbf{K}} & \hat{\mathbf{K}} \\ \hat{\mathbf{K}}^T & \hat{\mathbf{K}} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{d}} \\ \hat{\mathbf{d}} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{f}} \\ \hat{\mathbf{f}} \end{bmatrix} = \mathbf{f} \tag{4.14}$$

where $\bar{\mathbf{d}} \in \mathbb{R}^{2\bar{N}}$ and $\hat{\mathbf{d}} \in \mathbb{R}^{2\hat{N}}$ are the coefficient vectors to be determined. The components of the stiffness matrix and the force vector in (4.14) are shown in the Appendix 4.7. The stiffness matrix $\mathbf{K} \in \mathbb{R}^{N \times N}$ is a symmetric positive definite (SPD) matrix, where $N = 2(\bar{N} + \hat{N})$. The sub-matrix $\bar{\mathbf{K}} \in \mathbb{R}^{2\bar{N} \times 2\bar{N}}$ is a sparse SPD matrix due to the compactness of RK shape functions, while the sub-matrix $\hat{\mathbf{K}} \in \mathbb{R}^{2\hat{N} \times 2\hat{N}}$ is a dense SPD matrix but low-dimensional, i.e. $\hat{N} \ll \bar{N}$. Note that

due to the treatment of the domain integral with the non-smooth near-tip bases, the coupling sub-matrix $\widehat{\mathbf{K}} \in \mathbb{R}^{2\bar{N} \times 2\hat{N}}$ becomes much sparser under the ISBFM Galerkin framework compared with the counterpart obtained from the standard Galerkin formulation. This sparsity of ISBFM discrete system allows effective MOR methods as introduced in the following sections.

4.2.3 Comparison of standard Galerkin and ISBFM Galerkin methods for LEFM

We first verify the performance of ISBFM Galerkin formulation for the line crack fracture problem in Figure 4.4. The fine-scale solution \mathbf{u}^h of the standard Galerkin and ISBFM Galerkin methods are obtained using a uniform discretization of 48×48 ($\bar{N} = 2304$) nodes. Gaussian quadrature with uniform 48×48 domain integration cells is employed for the integration of the Galerkin methods. To evaluate the performance of the ISBFM fine-scale solutions for fracture problems, the normalized errors in L_2 norm, semi- H^1 norm of the solution \mathbf{u}^h , and the relative absolute error of the Mode I stress intensity factor (SIF) K_I^h , are defined as

$$e_0 = \frac{\|\mathbf{u} - \mathbf{u}^h\|_{L_2}}{\|\mathbf{u}\|_{L_2}}, \quad e_1 = \frac{\|\nabla^s \mathbf{u} - \nabla^s \mathbf{u}^h\|_{L_2}}{\|\nabla^s \mathbf{u}\|_{L_2}}, \quad e_{KI} = \frac{|K_I - K_I^h|}{|K_I|} \quad (4.15)$$

where $\|\mathbf{u}\|_{L_2} = \left(\int_{\Omega} u_i u_i d\Omega \right)^{1/2}$ and $\|\nabla^s \mathbf{u}\|_{L_2} = \left(\int_{\Omega} u_{i,j} u_{i,j} d\Omega \right)^{1/2}$.

Figure 4.2 and Figure 4.3 present the errors e_0 , e_1 and e_{KI} of the approximations obtained from the standard Galerkin and ISBFM Galerkin formulations under different orders of Gaussian quadrature rule and near-tip basis functions. Figure 4.2 shows that when the first order

near-tip basis function ($\hat{N} = 1$) is used, the standard Galerkin method requires up to 12th and 8th order quadrature rules for displacement and strain solutions, respectively, to attain a similar accuracy as that solved by the ISBFM Galerkin method with just 2nd order quadrature rule. With the employment of higher order near-tip basis functions, e.g. $\hat{N} = 2$ and $\hat{N} = 3$, the increased superiority of the ISBFM Galerkin solution over the standard Galerkin solution indicates the need of using quadrature rule. On the other hand, by using the ISBFM Galerkin method the non-smooth solution is integrated only at the boundaries away from the crack tip, allowing a sufficiently accurate result to be achieved by using 4th order quadrature rule. Due to the enhanced capability of capturing the singularity and discontinuity of fracture, the ISBFM Galerkin formulation also gives better SIF solution as shown in Figure 4.3.

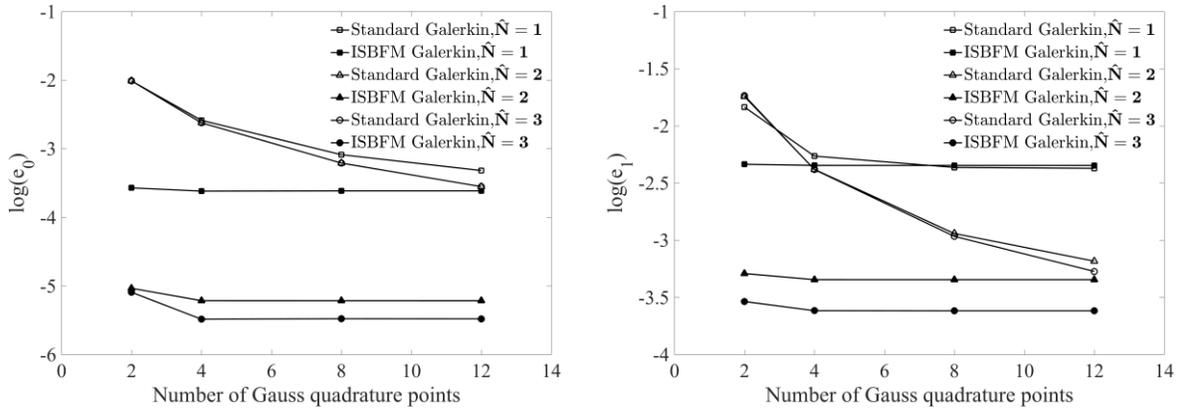


Figure 4.2: Normalized error in L_2 norm (left) and semi- H^1 norm (right) of the loaded line crack solution under different orders of Gaussian quadrature rule. The standard Galerkin and the proposed ISBFM Galerkin methods are compared under different orders of enrichment functions, i.e. $\hat{N} = 1, 2, 3$.

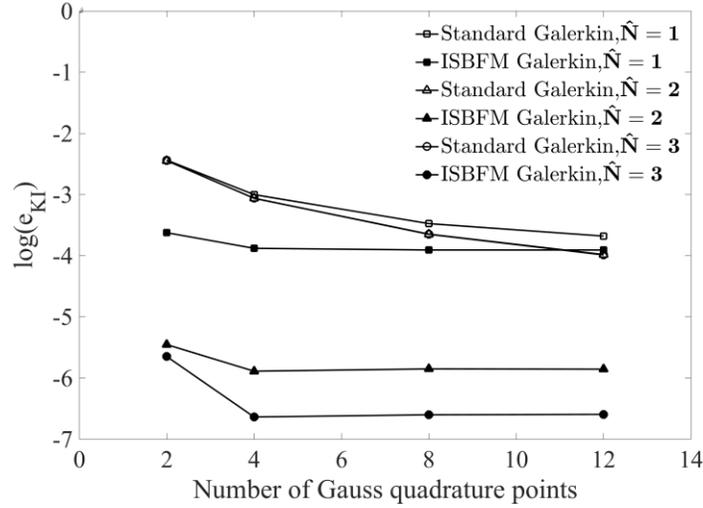


Figure 4.3: Relative error of the Mode I stress intensity factor of the loaded line crack solution under different orders of Gaussian quadrature rule. The standard Galerkin and the proposed ISBFM Galerkin methods are compared under different orders of enrichment functions, i.e. $\hat{N} = 1, 2, 3$.

4.3 Reduced-Order Modeling of Fracture Mechanics

In this section, a projection-based MOR approach using the basis vectors obtained by *standard modal analysis* is first reviewed. We then introduce the proposed decomposed projection bases such that the low-rank representation of the smooth subspace is properly constructed while preserving the non-smooth subspace. The following notations will be used throughout this section. The variable k denotes the subspace rank, the subspace reduced quantities are denoted with a superposed tilde “ \sim ”, and the reduced-order approximation is denoted by a superscript “ r ”.

4.3.1 Generic projection based MOR

For a given fine-scale solution $\mathbf{d} \in \mathbb{R}^N$ to the ISBFM discrete system (4.14), an optimal reduced-order solution $\mathbf{d}^r = \mathbf{P}\boldsymbol{\alpha} \in \mathbb{R}^N$ can be obtained by minimizing the defined energy error norm

$$\mathcal{E}(\boldsymbol{\alpha}) = (\mathbf{d} - \mathbf{d}^r)^T \mathbf{K}(\mathbf{d} - \mathbf{d}^r) \quad (4.16)$$

where $\mathbf{P} \in \mathbb{R}^{N \times k}$ ($k \ll N$) spans the k -rank subspace to approximate the fine-scale solution. The stationary condition to (4.16) leads to the projection based reduced-order model (ROM) as follows

$$\mathbf{P}^T \mathbf{K} \mathbf{P} \boldsymbol{\alpha} = \mathbf{P}^T \mathbf{f} \quad (4.17)$$

where the equality $\mathbf{P}^T \mathbf{K} \mathbf{d} = \mathbf{P}^T \mathbf{f}$ is used. The above equation (4.17) can be written in a compact form of ROM

$$\tilde{\mathbf{K}} \boldsymbol{\alpha} = \tilde{\mathbf{f}} \quad (4.18)$$

where $\boldsymbol{\alpha} \in \mathbb{R}^k$ is the reduced coefficient vector that minimizes \mathcal{E} in (4.16), $\tilde{\mathbf{K}} := \mathbf{P}^T \mathbf{K} \mathbf{P} \in \mathbb{R}^{k \times k}$ is the reduced stiffness matrix, and $\tilde{\mathbf{f}} := \mathbf{P}^T \mathbf{f} \in \mathbb{R}^k$ is the reduced force vector. After solving the ROM problem (4.18), the reduced-order approximation can be obtained by relating the reduced coefficient vector to the high-dimensional space

$$\mathbf{d}^r = \mathbf{P} \boldsymbol{\alpha} \quad (4.19)$$

where $\mathbf{d}^{rT} = [\bar{\mathbf{d}}^{rT} \hat{\mathbf{d}}^{rT}] \in \mathbb{R}^{2\bar{N}+2\hat{N}}$, $\bar{\mathbf{d}}^{rT} = [\bar{\mathbf{d}}_1^{rT} \bar{\mathbf{d}}_2^{rT} \dots \bar{\mathbf{d}}_{\bar{N}}^{rT}] \in \mathbb{R}^{2\bar{N}}$, $\hat{\mathbf{d}}^{rT} = [\hat{\mathbf{d}}_1^{rT} \hat{\mathbf{d}}_2^{rT} \dots \hat{\mathbf{d}}_{\hat{N}}^{rT}] \in \mathbb{R}^{2\hat{N}}$, and $\boldsymbol{\alpha} = [\alpha_1 \alpha_2 \dots \alpha_k]^T$ is to be solved from the reduced-order discrete equation (4.18). By means of the decomposition into the smooth and non-smooth approximations in (4.5) and (4.6),

respectively, the reduced-order displacement \mathbf{u}^r associated with the reduced coefficient \mathbf{d}^r in (4.19) can be obtained as follows

$$\mathbf{u}^r = \sum_{I=1}^{\bar{N}} \Psi_I(\mathbf{x}) \bar{\mathbf{d}}_I^r + \sum_{J=1}^{\hat{N}} \mathbf{F}_J \hat{\mathbf{d}}_J^r \quad (4.20)$$

The remaining issue is how to define the reduced-order projection \mathbf{P} for the ISBFM discrete system of interest.

4.3.2 MOR based on standard modal analysis

The selection of the reduced-order projection \mathbf{P} can be based on modal analysis or singular value decomposition (SVD) of the fine-scale system, which is also referred to as the mode displacement method in dynamics (Craig 1981; Dickens, Nakagawa & Wittbrodt 1997). We denote the employment of the reduction method based on the standard modal analysis together with the ISBFM discrete system (4.14) as ISBFM-MA. In this manner, $\mathbf{P} = [\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_k] \in \mathbb{R}^{N \times k}$ is composed of the orthonormal eigenvectors of the stiffness matrix \mathbf{K} corresponding to the first k smallest eigenvalues that are in ascending order, i.e. $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_k$. Then, the reduced stiffness matrix in the corresponding ROM in (4.18) becomes

$$\tilde{\mathbf{K}} = \text{diag}([\lambda_1 \ \lambda_2 \ \dots \ \lambda_k]) \quad (4.21)$$

And the reduced-order solution \mathbf{d}^r is given with respect to the selected eigenpairs $\{\boldsymbol{\phi}_j, \lambda_j\}_{j=1}^k$ and the input data force vector \mathbf{f} ,

$$\mathbf{d}^r = \mathbf{P}\boldsymbol{\alpha} = \sum_{j=1}^k \boldsymbol{\phi}_j \alpha_j = \sum_{j=1}^k \boldsymbol{\phi}_j \boldsymbol{\phi}_j^T \mathbf{f} / \lambda_j \quad (4.22)$$

Note that the following relationship has been used,

$$\alpha_j = \boldsymbol{\phi}_j^T \mathbf{f} / \lambda_j, \quad j = 1, \dots, k \quad (4.23)$$

Consequently, the energy error defined in (4.16) for ISBFM-MA becomes

$$\begin{aligned} \mathcal{E} &= \left(\sum_{j=k+1}^N \boldsymbol{\phi}_j \alpha_j \right)^T \mathbf{K} \left(\sum_{l=k+1}^N \boldsymbol{\phi}_l \alpha_l \right) = \sum_{j=k+1}^N \alpha_j^2 \lambda_j = \sum_{j=k+1}^N (\boldsymbol{\phi}_j^T \mathbf{f})^2 / \lambda_j \\ &\leq \sum_{j=k+1}^N \lambda_j^{-1} \|\boldsymbol{\phi}_j\|^2 \|\mathbf{f}\|^2 \leq C_1 \|\mathbf{f}\|_{\infty}^2 \lambda_{k+1}^{-1} \end{aligned} \quad (4.24)$$

where $C_1 = N(N-k)$. It can be observed that the error is related to the smallest eigenvalues of the discarded eigenmodes not considered in the projection \mathbf{P} . This is consistent to the fact that the energy of the system is primarily carried by the low modes (Shabana 1991). In addition, as the norm $\|\mathbf{f}\|_{\infty}$ represents the largest component of the force vector $\mathbf{f} = [\bar{\mathbf{f}}^T \hat{\mathbf{f}}^T]^T$, it indicates that the performance of ISBFM-MA could be influenced radically by the scale of the near-tip basis functions (4.8) used for constructing the non-smooth part of the force vector $\hat{\mathbf{f}}$ in (4.46). In next section, we will discuss the susceptibility of modal analysis and show how it fails to preserve the near-tip characteristics in reduced-order approximation.

Difficulties in the standard modal analysis for fracture mechanics

We examine the reduced-order displacements \mathbf{u}^r in (4.20) associated with each projection basis vector $\boldsymbol{\phi}_j = [\bar{\boldsymbol{\phi}}_j^T \hat{\boldsymbol{\phi}}_j^T]^T \in \mathbb{R}^{2\bar{N}+2\hat{N}}$ (the j -th eigenmode of \mathbf{K}) as given by

$$\mathbf{u}^r = \sum_{j=1}^k \mathbf{w}^r(\phi_j) \alpha_j \quad (4.25)$$

where $\mathbf{w}^r(\phi_j)$ is the generalized shape function associated with each reduced coefficient α_j , which is also called the reduced-order basis (ROB) functions (Barrault et al. 2004) and expressed as

$$\mathbf{w}^r(\phi_j) = \sum_{I=1}^{\bar{N}} \Psi_I(\mathbf{x}) \bar{\phi}_{Ij} + \sum_{J=1}^{\hat{N}} \mathbf{F}_J \hat{\phi}_{Jj} \quad (4.26)$$

where $\bar{\phi}_{Ij} \in \mathbb{R}^2$ and $\hat{\phi}_{Jj} = [\hat{\phi}_{Jj}^s \quad \hat{\phi}_{Jj}^{as}]^T \in \mathbb{R}^2$ are the components of $\bar{\phi}_j \in \mathbb{R}^{2\bar{N}}$ and $\hat{\phi}_j \in \mathbb{R}^{2\hat{N}}$, respectively.

For demonstration purpose we consider a symmetric line crack model subjected to a far-field hydrostatic tension load $\sigma_\infty = 1$. The corresponding half model used for numerical study is shown in Figure 4.4, where traction is prescribed on $\Gamma_1 \cup \Gamma_3 \cup \Gamma_4$ and displacement is imposed on Γ_2 according to the analytical solution in (England 2003). A uniform discretization of 24×24 nodes ($2\bar{N} = 1152$) with $\hat{N} = 2$ near-tip basis functions are used to construct the ISBFM discrete system (4.14) based on 5th order and 8th order quadrature rules for domain integrals and boundary integrals, respectively. Note that only the symmetric part of the near-tip basis functions \mathbf{F}_J^s in (4.8) is employed for the non-smooth approximation due to the Mode I symmetry.

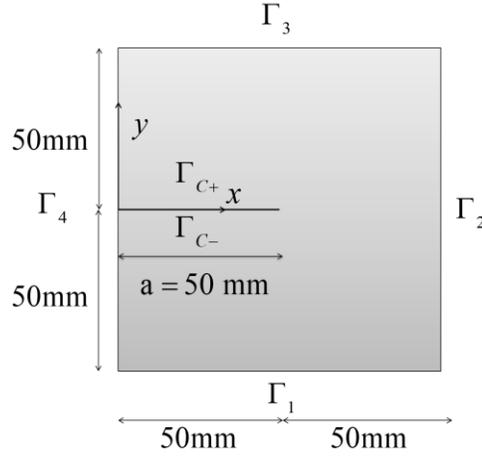


Figure 4.4: A finite dimensional model of a line crack subjected to a far-field hydrostatic tension, where traction is prescribed on $\Gamma_1 \cup \Gamma_3 \cup \Gamma_4$ and displacement is imposed on Γ_2 according to the analytical solution of (England 2003). The adopted material properties of borosilicate glass are $E = 64,000 \text{ N/mm}^2$ and $\nu = 0.2$.

Some representative ROB functions $\mathbf{w}^r(\phi_j)$ associated with low-frequency and high-frequency modes are plotted in Figure 4.5. As can be seen, for this line crack model, the ROB functions representing part of the near-tip behaviors are associated with two highest modes (i.e. EV 1153-1154) resulting from the two local asymptotic basis functions \mathbf{F}_j^s in (4.8). As such, the reduced-order approximation by ISBFM-MA misses the near-tip behavior as shown in Figure 4.7(b).

A simple scaling approach can be applied to the near-tip basis functions to mitigate this issue, where the basis functions \mathbf{F}_J in (4.8) are scaled by $\chi = (\gamma l_c^{J/2})^{-1}$ as follows

$$\mathbf{F}_J^* = \chi \mathbf{F}_J(r^{J/2}) = \frac{1}{\gamma l_c^{J/2}} \mathbf{F}_J(r^{J/2}), \quad J = 1, 2, \dots, \hat{N} \quad (4.27)$$

Here the characteristic length l_c is chosen as the domain dimension and the parameter γ is a scaling constant. Using the scaled basis functions in (4.27) with $\gamma = 50$ for the non-smooth approximation, the primary near-tip behavior is captured by the low modes, as shown by EV 4 and EV 6 in Figure 4.6. As a result, a more reasonable reduced-order deformation is obtained in Figure 4.7(c). Clearly, the ISBFM-MA reduced-order solution corrected by the scaling of near-tip enrichment functions is still not very satisfactory in representing the crack-tip behavior and the inhomogeneous Dirichlet boundary conditions on Γ_2 . These issues can be improved by the decomposed subspace reduction method to be discussed in the next section.

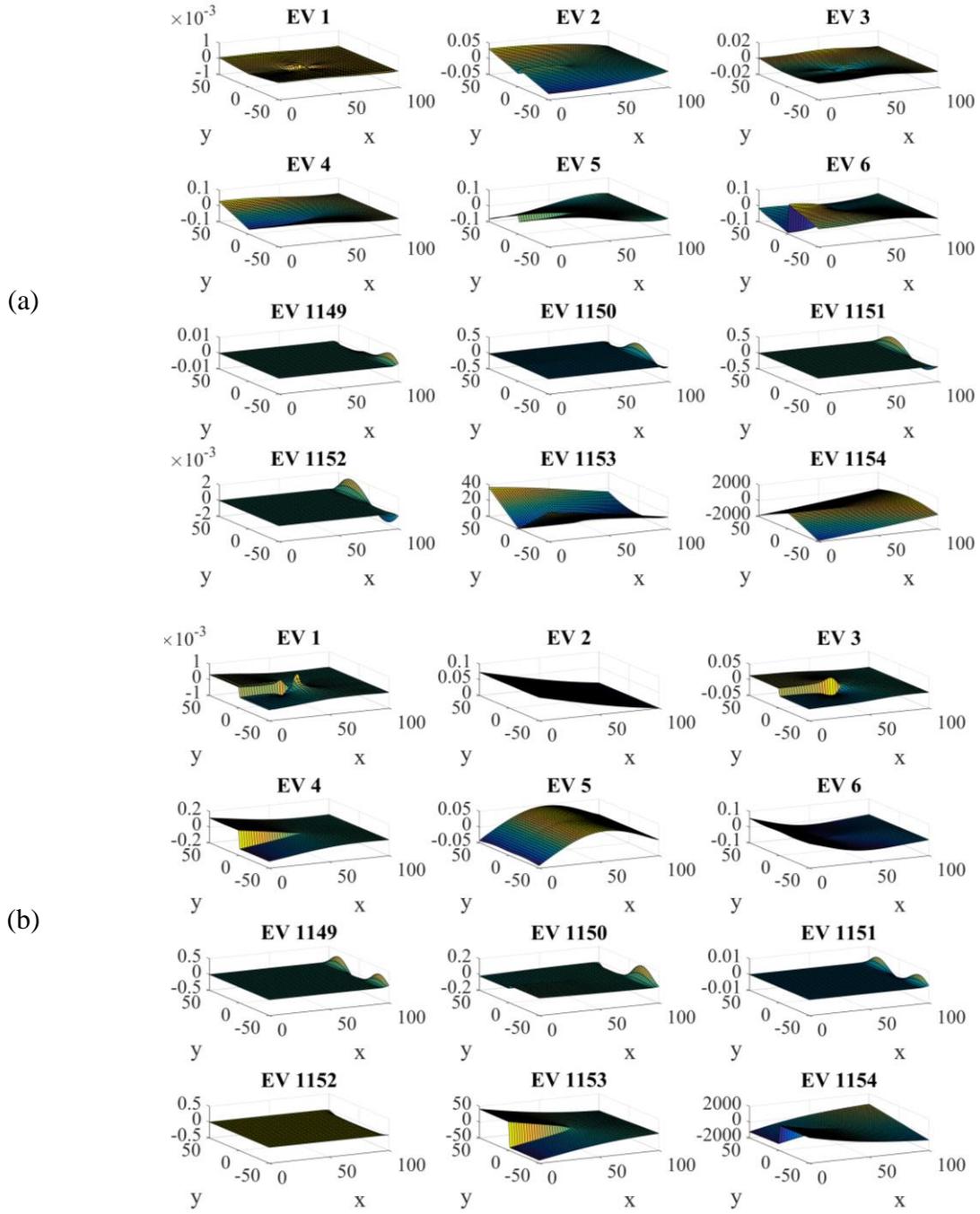


Figure 4.5: Shapes of ROB functions of ISBFM-MA for the line crack model using $\hat{N} = 2$ non-scaled near-tip basis functions: (a) $w_x^r(\phi_j)$ and (b) $w_y^r(\phi_j)$. (The subtitle EV n of each plot denotes the n-th eigenfunction of the ISBFM discrete system \mathbf{K})

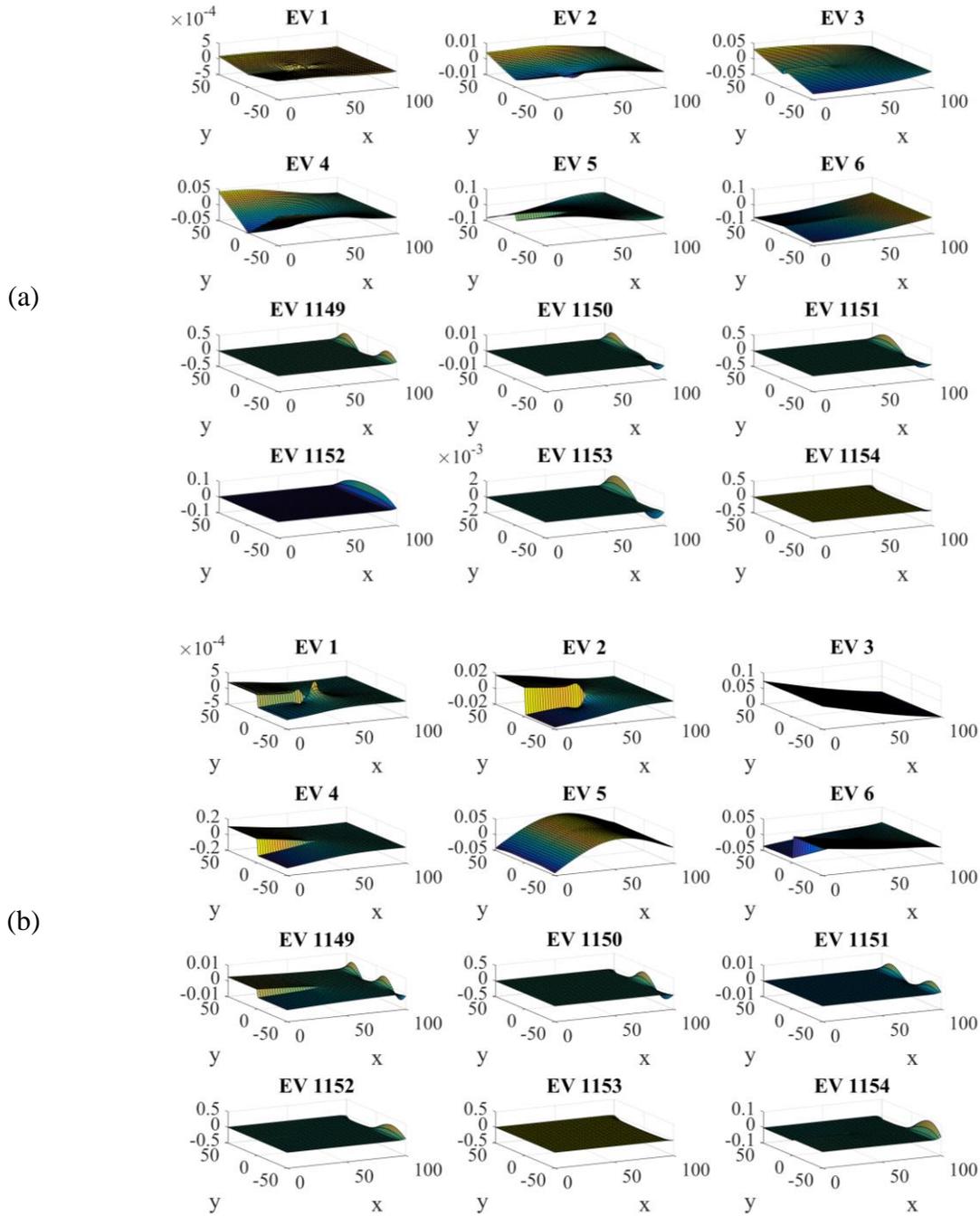


Figure 4.6: Shapes of ROB functions of ISBFM-MA for the line crack model using $\hat{N} = 2$ scaled non-smooth basis functions with $\gamma = 50$: (a) $w'_x(\phi_j)$ and (b) $w'_y(\phi_j)$. (The subtitle EV n of each plot denotes the n-th eigenfunction of the ISBFM discrete system \mathbf{K})

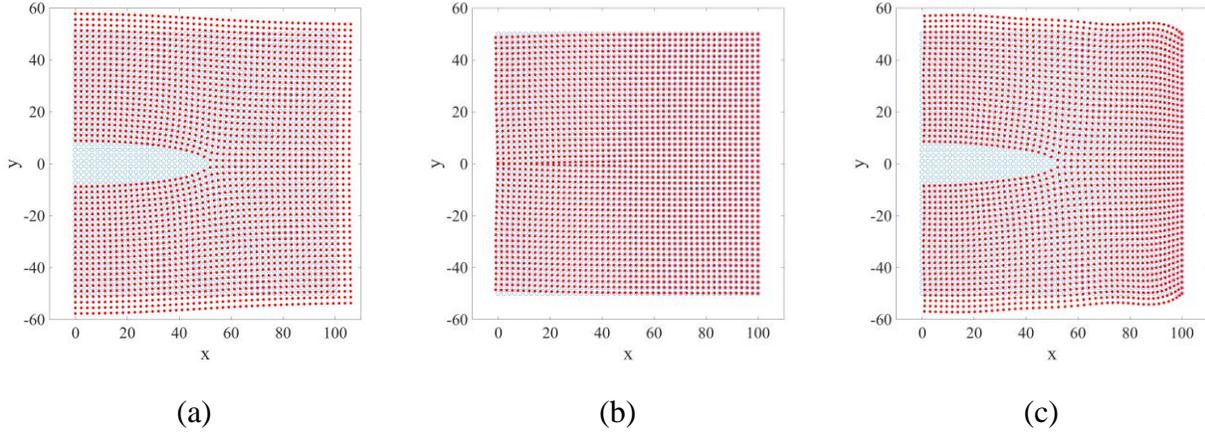


Figure 4.7: Deformation of the line crack model approximated by (a) fine-scale approximation, (b) ISBFM-MA ($k = 40$) using non-scaled near-tip basis functions, and (c) ISBFM-MA ($k = 40$) using scaled non-smooth basis functions with $\gamma = 50$. The blue circles denote the undeformed configuration and the red dots denote the deformed configuration (The displacement solution is magnified by 5×10^3 times for visualization).

4.3.3 MOR based on decomposed projection

To robustly preserve the near-tip non-smooth characteristics, we adopt a decomposed projection approach proposed in (Chen, Marodon & Hu 2015) where the reduced-order approximation is expressed as

$$\mathbf{d}^r = \begin{bmatrix} \bar{\mathbf{d}}^r \\ \hat{\mathbf{d}}^r \end{bmatrix} = \mathbf{P}\boldsymbol{\alpha}, \quad \mathbf{P} = \begin{bmatrix} \bar{\mathbf{P}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{P}} \end{bmatrix} \in \mathbb{R}^{2(\bar{N}+\hat{N}) \times (\bar{k}+\hat{k})} \quad (4.28)$$

where $\boldsymbol{\alpha} = [\bar{\boldsymbol{\alpha}}^T \ \hat{\boldsymbol{\alpha}}^T]^T$, $\bar{\mathbf{P}} \in \mathbb{R}^{2\bar{N} \times \bar{k}}$ and $\hat{\mathbf{P}} \in \mathbb{R}^{2\hat{N} \times \hat{k}}$ are the smooth part and the non-smooth part of the projection, respectively, and $\bar{\boldsymbol{\alpha}} \in \mathbb{R}^{\bar{k}}$ and $\hat{\boldsymbol{\alpha}} \in \mathbb{R}^{\hat{k}}$ are the associated reduced coefficient vectors. The total reduced dimension in the decomposed reduction is $k = \bar{k} + \hat{k}$. In this manner, the ROM of the fine-scale system (4.14) is given as

$$\begin{bmatrix} \bar{\mathbf{P}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{P}} \end{bmatrix}^T \begin{bmatrix} \bar{\mathbf{K}} & \hat{\mathbf{K}} \\ \hat{\mathbf{K}}^T & \hat{\mathbf{K}} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{P}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{P}} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{a}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{P}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{P}} \end{bmatrix}^T \begin{bmatrix} \bar{\mathbf{f}} \\ \hat{\mathbf{f}} \end{bmatrix} \quad (4.29)$$

which is rewritten as

$$\tilde{\mathbf{K}}\tilde{\mathbf{a}} := \begin{bmatrix} \tilde{\mathbf{K}} & \tilde{\mathbf{K}} \\ \tilde{\mathbf{K}}^T & \tilde{\mathbf{K}} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{a}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{f}} \\ \tilde{\mathbf{f}} \end{bmatrix} := \tilde{\mathbf{f}} \quad (4.30)$$

where the sub-matrices $\tilde{\mathbf{K}}^{\bar{\bar{k}}}$, $\tilde{\mathbf{K}}^{\hat{k}}$, and $\tilde{\mathbf{K}}^{\hat{k}\hat{k}}$ of the reduced stiffness matrix $\tilde{\mathbf{K}}$ are

$$\tilde{\mathbf{K}}^{\bar{\bar{k}}} = \bar{\mathbf{P}}^T \bar{\mathbf{K}} \bar{\mathbf{P}} \in \mathbb{R}^{\bar{k} \times \bar{k}}, \quad \tilde{\mathbf{K}}^{\hat{k}} = \bar{\mathbf{P}}^T \hat{\mathbf{K}} \hat{\mathbf{P}} \in \mathbb{R}^{\bar{k} \times \hat{k}}, \quad \tilde{\mathbf{K}}^{\hat{k}\hat{k}} = \hat{\mathbf{P}}^T \hat{\mathbf{K}} \hat{\mathbf{P}} \in \mathbb{R}^{\hat{k} \times \hat{k}} \quad (4.31)$$

Similarly, the sub-vectors $\tilde{\mathbf{f}}^{\bar{\bar{k}}}$ and $\tilde{\mathbf{f}}^{\hat{k}}$ of the reduced force vector $\tilde{\mathbf{f}}$ are

$$\tilde{\mathbf{f}}^{\bar{\bar{k}}} = \bar{\mathbf{P}}^T \bar{\mathbf{f}} \in \mathbb{R}^{\bar{k}}, \quad \tilde{\mathbf{f}}^{\hat{k}} = \hat{\mathbf{P}}^T \hat{\mathbf{f}} \in \mathbb{R}^{\hat{k}} \quad (4.32)$$

Since the motivation of the decomposed reduction is to optimally preserve near-tip characteristics of fracture in the solution subspace, the dimension reduction for non-smooth quantities can be ignored given the employment of very small number of enrichment functions, i.e. $\hat{N} \ll \bar{N}$. Thus, we let $\hat{k} = 2\hat{N}$, $\hat{\mathbf{P}} = \mathbf{I}_{2\hat{N}}$, and $\hat{\mathbf{a}} = \hat{\mathbf{d}}^r$. The reduced system (4.29) becomes

$$\begin{bmatrix} \bar{\mathbf{P}}^T \bar{\mathbf{K}} \bar{\mathbf{P}} & \bar{\mathbf{P}}^T \hat{\mathbf{K}} \\ \hat{\mathbf{K}}^T \bar{\mathbf{P}} & \hat{\mathbf{K}} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{a}} \\ \hat{\mathbf{d}}^r \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{P}}^T \bar{\mathbf{f}} \\ \hat{\mathbf{f}} \end{bmatrix} \quad (4.33)$$

Remark 4.1: One feature inherited from ISBFM Galerkin formulation in equations (4.12) and (4.13) is that it yields a sparse sub-matrix $\hat{\mathbf{K}}$ in (4.14) due to minimal interaction between the near-tip basis functions and the smooth RK functions on the boundaries that are not connected to the crack tip, while the standard Galerkin formulation in (4.9) and (4.10) yields a

fully dense $\widehat{\mathbf{K}}$. As such, the complexity of operations involved with $\overline{\mathbf{P}}^T \widehat{\mathbf{K}}$ is significantly reduced using the ISBFM formulation.

Decomposed subspace reduction (DSR)

The reduced system (4.33) can be expressed as

$$\begin{aligned}\widetilde{\mathbf{K}}\bar{\mathbf{a}} + \overline{\mathbf{P}}^T \widehat{\mathbf{K}}\hat{\mathbf{d}}^r &= \widetilde{\mathbf{f}} \\ \widehat{\mathbf{K}}^T \overline{\mathbf{P}}\bar{\mathbf{a}} + \widehat{\mathbf{K}}\hat{\mathbf{d}}^r &= \hat{\mathbf{f}}\end{aligned}\quad (4.34)$$

where $\widetilde{\mathbf{K}}$ and $\widetilde{\mathbf{f}}$ are defined in (4.31) and (4.32), respectively. By static condensation we have the following condensed system for the non-smooth degrees of freedom (DOF),

$$\widehat{\mathbf{K}}_c^* \hat{\mathbf{d}}^r = \hat{\mathbf{f}}_c^* \quad (4.35)$$

where

$$\widehat{\mathbf{K}}_c^* = \widehat{\mathbf{K}} - \widehat{\mathbf{K}}^T \overline{\mathbf{P}} \widetilde{\mathbf{K}}^{-1} \overline{\mathbf{P}}^T \widehat{\mathbf{K}}, \quad \hat{\mathbf{f}}_c^* = \hat{\mathbf{f}} - \widehat{\mathbf{K}}^T \overline{\mathbf{P}} \widetilde{\mathbf{K}}^{-1} \overline{\mathbf{P}}^T \widetilde{\mathbf{f}} \quad (4.36)$$

The matrix $\widehat{\mathbf{K}}_c^*$ is the Schur complement which is SPD by the *Schur complement lemma*.

After solving for $\hat{\mathbf{d}}^r$, the smooth DOF can be retrieved by the supplement equation

$$\bar{\mathbf{a}} = \widetilde{\mathbf{K}}^{-1} (\widetilde{\mathbf{f}} - \overline{\mathbf{P}}^T \widehat{\mathbf{K}}\hat{\mathbf{d}}^r) \quad (4.37)$$

In this condensation procedure, only the reduced matrix $\widetilde{\mathbf{K}} \in \mathbb{R}^{\bar{k} \times \bar{k}}$ needs to be inverted, which is operated in the low-dimensional subspace of $\overline{\mathbf{P}}$. Considering that $\widehat{\mathbf{K}}$ is sparse and the non-smooth dimension is low, the construction effort of the condensed system (4.35) is low.

It is observed that the expression $\bar{\mathbf{P}}\tilde{\mathbf{K}}^{-1}\bar{\mathbf{P}}^T \in \mathbb{R}^{2\bar{N} \times 2\bar{N}}$ behaves like the low-rank approximation of $\bar{\mathbf{K}}^{-1}$. Thus, we construct the projection $\bar{\mathbf{P}}$ by using the smallest \bar{k} eigenvectors of the smooth sub-matrix $\bar{\mathbf{K}}$, whose eigenpairs are denoted as $\{\boldsymbol{\gamma}_j, \nu_j\}_{j=1}^{2\bar{N}}$. Consequently, the reduced smooth sub-matrix becomes $\tilde{\mathbf{K}} = \text{diag}([\nu_1 \ \nu_2 \ \dots \ \nu_{\bar{k}}])$. Since the decomposed reduced system (4.34) has been projected onto the subspace of $\bar{\mathbf{P}}$ and remains in low dimension, we can directly solve this reduced system as the alternative of solving the decoupled condensed system given in (4.35) and (4.37). We term this reduction method the *decomposed subspace reduction* (DSR) to distinguish the *decomposed reduction* (DR) used in (Chen, Marodon & Hu 2015) based on a different projection.

Remark 4.2: As presented in (4.34), ISBFM-DSR only reduces the smooth subspace while preserves the non-smooth features. Since the associated eigenmodes of $\bar{\mathbf{P}}$ are obtained from the smooth sub-matrix $\bar{\mathbf{K}}$ instead of the Schur complement $\bar{\mathbf{K}}_c = \bar{\mathbf{K}} - \hat{\mathbf{K}}\hat{\mathbf{K}}^{-1}\hat{\mathbf{K}}^T$ used in (Chen, Marodon & Hu 2015), it yields a better low-rank approximation for the smooth subspace. It can be also easily shown that the reduced-order solution \mathbf{u}^r from ISBFM-DSR is invariant to arbitrary scaling of enrichment functions since the construction of $\bar{\mathbf{P}}$ is independent of the non-smooth quantities. Moreover, the computational complexity of ISBFM-DSR is in the same order as that of the ISBFM-DR proposed earlier (Chen, Marodon & Hu 2015).

In last section we showed that a proper scaling is needed in ISBFM-MA to preserve the near-tip characteristics in the reduced order solution for fracture mechanics. Re-examining the line crack problem in Figure 4.4, the ISBFM-DSR method, on the other hand, correctly preserves

the near-tip modes in the reduced-order subspace as shown in Figure 4.8(b) due to the employment of decomposed projection.

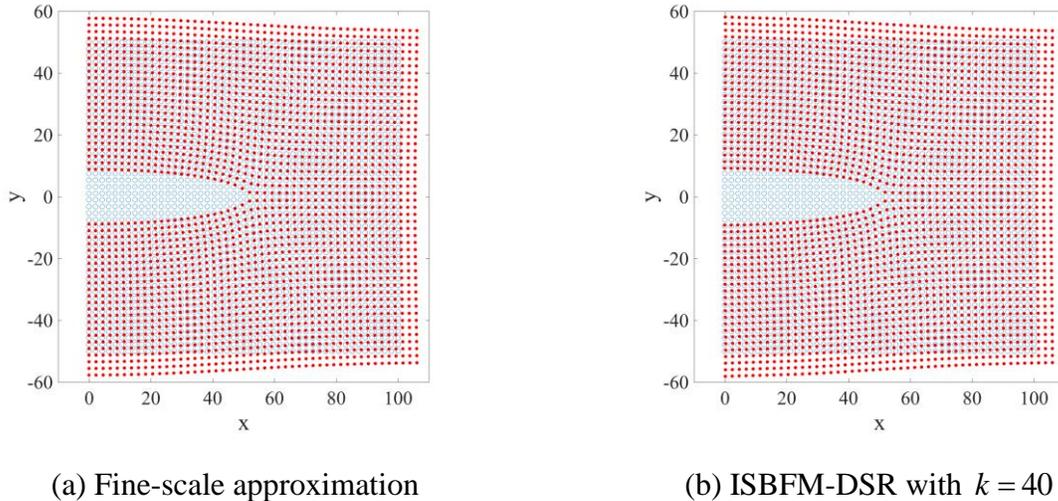


Figure 4.8: Deformation for the line crack model obtained by (a) fine-scale approximation and (b) ISBFM-DSR ($k = 40$). The blue circles denote the undeformed configuration and the red dots denote the deformed configuration (The displacement solution is magnified by 5×10^3 times for visualization).

Issues resulting from the inhomogeneous Dirichlet boundary conditions

In addition to the failure in preserving the near-tip behavior, the standard modal analysis may also suffer from errors in representing the *inhomogeneous Dirichlet boundary conditions (IDBC)*, i.e. $g_i \neq 0$ in (4.1) as shown in Figure 4.7(c) where a non-zero Dirichlet boundary condition on Γ_2 in Figure 4.4. are not properly represented in the ISBFM-MA reduced-order solution. This is due to the penalty coefficient β used in Nitsche's method, which yields high ROB modes associated with IDBC while the low modes correspond to homogeneous Dirichlet boundary conditions. This issue in ISBFM-MA cannot be resolved by the scaling technique as

used for the near-tip modes. ISBFM-DSR, on the other hand, correctly represents IDBC in the reduced-order solution as shown in Figure 4.8. This is because the non-smooth basis functions are inhomogeneous on the problem boundaries and they are fully preserved in the decomposed subspaces reduction.

Remark 4.3: Loss of mechanical constrains in projection based reduced-order models (ROMs) is a common issue and needs to be addressed carefully (Lall, Krysl & Marsden 2003). The imposition of IDBC in ROM has been investigated (Cosimo, Cardona & Idelsohn 2016; Gunzburger, Peterson & Shadid 2007). However, these techniques were developed under the POD-based MOR and were applied to the problems without singularities and discontinuities.

4.4 Numerical Examples: Reduced-Order Modeling of Fracture Mechanics

Several numerical examples are analyzed to illustrate the effectiveness of the proposed ISBFM based MOR method for fracture problems. In the following study, RK functions using the linear basis and cubic B-spline kernel function with a normalized support size of $a/h = 1.71$ are selected for the smooth part of the solution. For Nitsche's treatment of Dirichlet boundary conditions (Fernández-Méndez & Huerta 2004), the penalty parameter $\beta = 100E/h$ is employed, where E is the Young's modulus and h is the minimal spacing of the discretization nodes.

4.4.1 Comparison of MOR methods for a cracked beam Poisson's problem

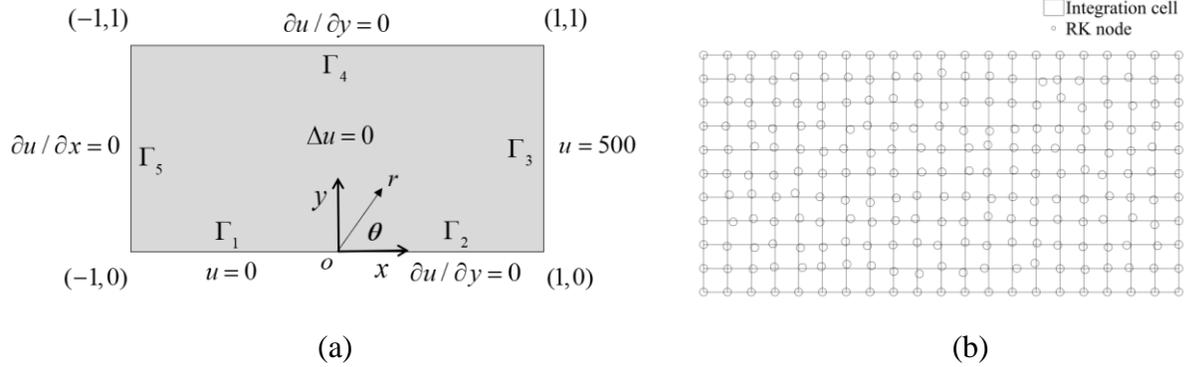


Figure 4.9: (a) Model of the cracked beam Poisson's problem; (b) fine-scale discretization.

To compare the proposed DSR method with modal analysis (MA) and decomposed reduction (DR) introduced in (Chen, Marodon & Hu 2015), we consider a cracked beam Poisson's problem in Figure 4.9(a), where the discontinuity of boundary condition at $(0,0)$ yields a singularity of order $O(r^{1/2})$. The non-smooth basis functions of this problem are given as

$$F_J(r, \theta) = r^{\frac{2J-1}{2}} \cos\left(\frac{2J-1}{2}\theta\right), \quad J = 1, 2, \dots, \hat{N} \quad (4.38)$$

where $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$ and $\theta = \tan^{-1}((y-y_0)/(x-x_0))$. These basis functions satisfy the conditions required for ISBFM method, i.e. the homogeneous equilibrium equation $\Delta F_J = 0$, the conditions $F_J = 0$ on Γ_1 , and $\partial F_J / \partial y = 0$ on Γ_2 . Since this is a scalar value problem, we only have one set of \hat{N} non-smooth basis functions and $N = \bar{N} + \hat{N}$. The interested readers can refer to (Chen, Marodon & Hu 2015; Li, Mathon & Sermer 1987) for more details. The fine-scale solution is obtained using a non-uniform discretization 21×11 ($\bar{N} = 231$) with random coefficient 0.1 for the interior points as shown in Figure 4.9(b). The same normalized norm e_0

defined in (4.15) is used to measure the error between the fine-scale displacement solution u^h and the reduced-order displacement solution u^r .

The error e_0 of different MOR methods for the cracked beam problem with various orders of enrichment functions, $\hat{N} = 1, 2, 5$, are compared in Table 3.1. As discussed in Section 0, the standard modal analysis (ISBFM-MA) with direct employment of the enrichment functions yields worst results. When using the scaled enrichment functions in (4.27) with $\gamma = 50$, the error e_0 is reduced by one order. The remaining error is primarily due to the inaccurate representation of IDBC at the boundary Γ_3 in the reduced-order solution (see u^r in Figure 4.11).

The performances of the proposed ISBFM-DSR method (Section 0) and the ISBFM-DR method introduced in the earlier study (Chen, Marodon & Hu 2015) are compared in Table 3.1. Overall, both ISBFM-DSR and ISBFM-DR yield satisfactory reduced-order results compared to the ISBFM-MA approach. It is noticed that ISBFM-DR with two enrichment functions ($\hat{N} = 2$) does not necessary yield a better accuracy than the case with one enrichment function ($\hat{N} = 1$), and the errors decrease slowly as the reduced dimension increases. On the hand, the errors in ISBFM-DSR decrease quickly as the order of near-tip basis functions (\hat{N}) or the ratio of reduced dimension (k/N) increases.

Figure 4.11 to Figure 4.13 show the absolute error distributions of different MOR methods against the fine-scale solution (Figure 4.10) under different ratios of reduced dimension k/N along $y=0.001$ near the singularity point. ISBFM-MA yields largest errors among all methods under consideration especially on the Dirichlet boundary Γ_3 (Figure 4.11). On the other

hand, ISBFM-DR (Figure 4.12) and ISBFM-DSR (Figure 4.13) both work well in this problem in capturing the singularity in derivative fields (du^r/dx and du^r/dy) with less than 5% error. Nevertheless, ISBFM-DSR yields the best reduced-order approximation overall, and the error is reduced uniformly with increasing the dimension reduction ratio.

Table 4.1: Normalized error in L_2 norm (e_0) of the reduced-order solutions for the cracked beam problem. ISBFM-MA, ISBFM-DR and ISBFM-DSR are compared under different ratios of reduced dimension k/N and different orders of enrichment functions $\hat{N} = 1, 2, 5$.

	k (k/N)	10 (4%)	20 (8%)	40 (16%)	80 (32%)
ISBFM-MA (non-scaled)	$\hat{N} = 1$	1.12×10^0	1.11×10^0	1.12×10^0	1.12×10^0
	$\hat{N} = 2$	1.02×10^0	1.00×10^0	1.00×10^0	0.99×10^0
	$\hat{N} = 5$	1.00×10^0	9.93×10^{-1}	9.91×10^{-1}	9.85×10^{-1}
ISBFM-MA (scaled $\gamma = 50$)	$\hat{N} = 1$	6.06×10^{-1}	5.17×10^{-1}	4.50×10^{-1}	3.76×10^{-1}
	$\hat{N} = 2$	6.06×10^{-1}	5.17×10^{-1}	4.51×10^{-1}	3.77×10^{-1}
	$\hat{N} = 5$	6.06×10^{-1}	5.51×10^{-1}	4.51×10^{-1}	3.78×10^{-1}
ISBFM-DR	$\hat{N} = 1$	6.66×10^{-2}	7.03×10^{-2}	3.61×10^{-2}	2.60×10^{-2}
	$\hat{N} = 2$	6.42×10^{-2}	5.89×10^{-2}	5.44×10^{-2}	5.23×10^{-2}
	$\hat{N} = 5$	5.06×10^{-3}	5.06×10^{-3}	4.92×10^{-3}	4.91×10^{-3}
ISBFM-DSR	$\hat{N} = 1$	5.60×10^{-2}	3.43×10^{-2}	2.30×10^{-2}	1.40×10^{-2}
	$\hat{N} = 2$	2.02×10^{-2}	8.94×10^{-3}	5.79×10^{-3}	3.15×10^{-3}
	$\hat{N} = 5$	3.75×10^{-3}	2.71×10^{-3}	1.95×10^{-3}	1.17×10^{-3}

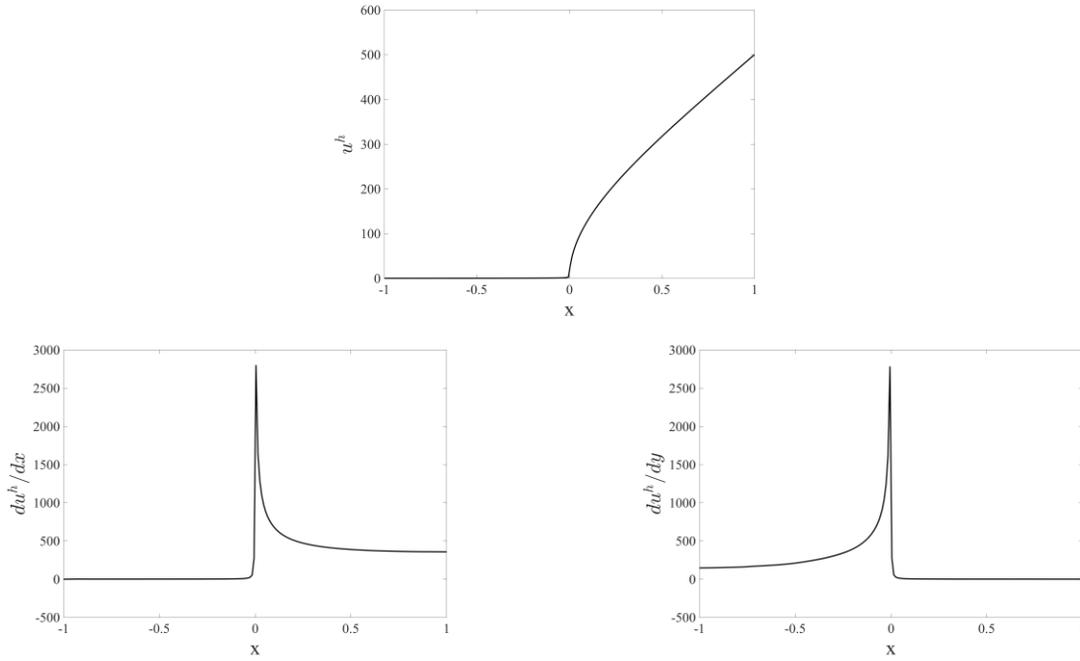


Figure 4.10: Fine-scale solutions u^h , du^h/dx , and du^h/dy along $y=0.001$ for the cracked beam problem solved with $\hat{N} = 2$.

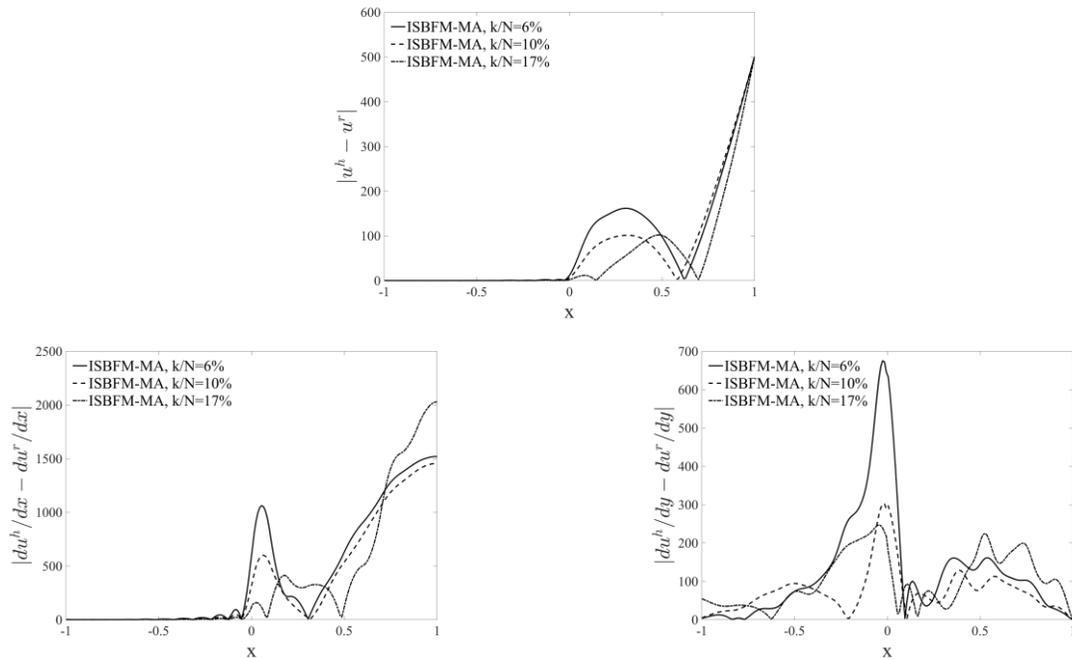


Figure 4.11: Absolute error distribution along $y=0.001$ of the reduced-order solutions u^r , du^r/dx , and du^r/dy by scaled ISBFM-MA ($\gamma=50$) for the cracked beam problem solved with $\hat{N} = 2$ under different ratios of reduced dimension k/N .

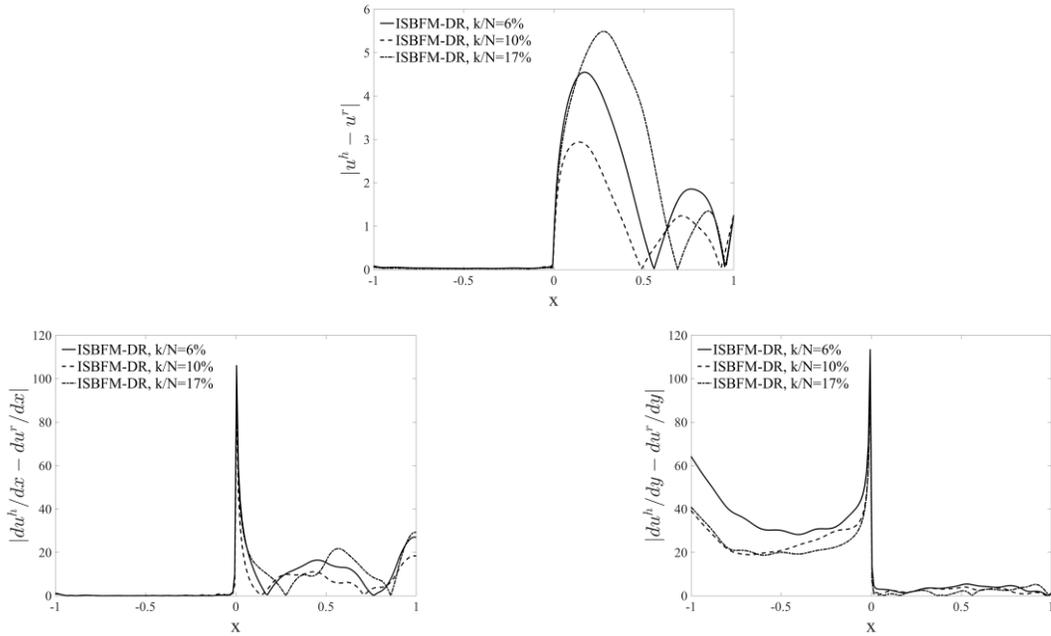


Figure 4.12: Absolute error distribution along $y=0.001$ of the reduced-order solutions u^r , du^r/dx , and du^r/dy by ISBFM-DR for the cracked beam problem solved with $\hat{N}=2$ under different ratios of reduced dimension k/N .

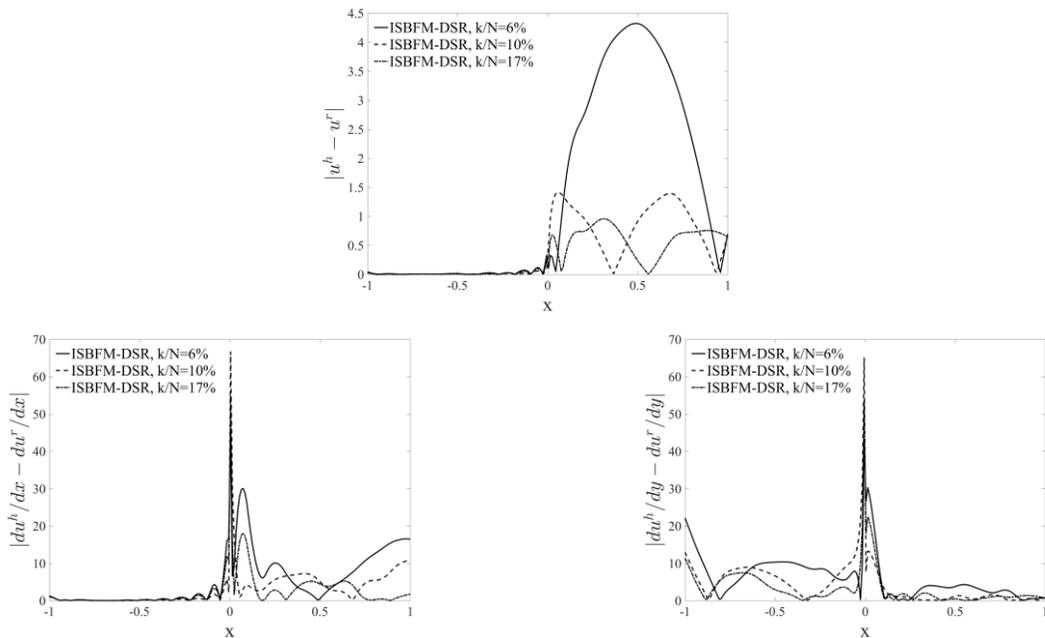


Figure 4.13: Absolute error distribution along $y=0.001$ of the reduced-order solutions u^r , du^r/dx , and du^r/dy by ISBFM-DSR for the cracked beam problem solved with $\hat{N}=2$ under different ratios of reduced dimension k/N .

4.4.2 Reduced-order modeling of a mode I line crack model

Here we perform a more detailed analysis of the mode I line crack problem given in Figure 4.4, with a uniform discretization 36×36 nodes ($\bar{N} = 1296$) employed to obtain the fine-scale solutions based on the ISBFM Galerkin formulation with 5th order and 8th order quadrature rules for domain integrals and boundary integrals, respectively.

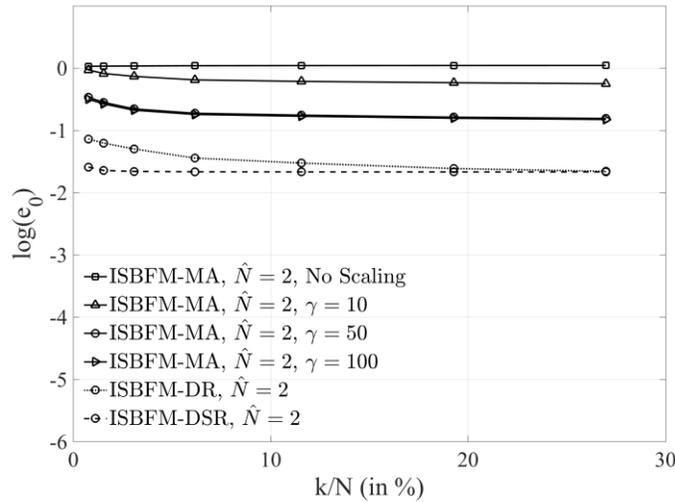


Figure 4.14: Normalized error in L_2 norm (e_0) of the reduced-order solutions of ISBFM-MA under different scaling for the loading crack model with $\hat{N} = 2$. ISBFM-DR and ISBFM-DSR are also given for comparison.

In this problem, ISBFM-MA employs the scaled basis functions defined in (4.27) with the characteristic length $l_c = 50\sqrt{2}$ mm (the maximum distance of material points to the crack tip). For different scaling cases, such as non-scaled, $\gamma = 10$, $\gamma = 50$, and $\gamma = 100$, the associated condition numbers of the stiffness matrix are 9.27×10^{13} , 2.53×10^7 , 5.63×10^7 , and 1.76×10^8 ,

respectively. In this problem, using $\gamma = 50$ is most effective to obtain improved reduced-order solutions of ISBFM-MA while keeping a well-conditioned stiffness matrix. As shown in Figure 4.14, while ISBFM-MA requires proper scaling of the enrichment functions for an enhanced accuracy, the ISBFM-DR and ISBFM-DSR results appear to be fairly independent to the scaling of near-tip enrichment functions.

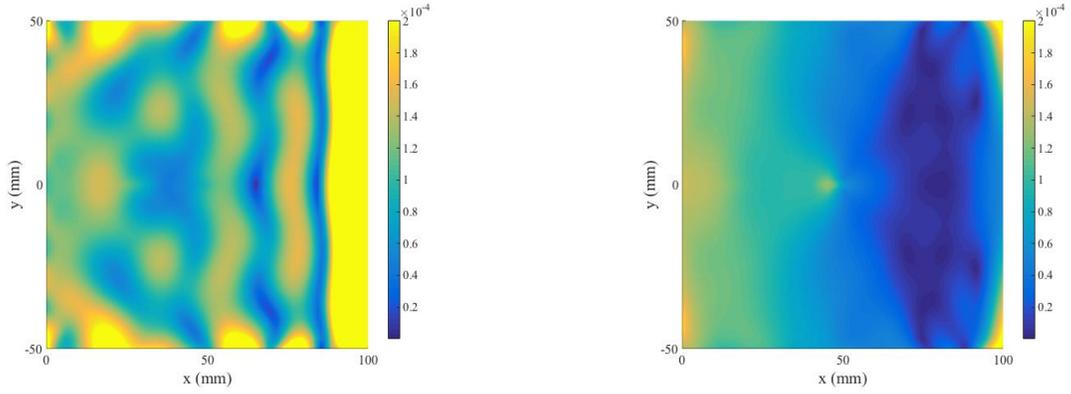
Table 4.2 summarizes the errors e_0 of the ISBFM-MA, ISBFM-DR and ISBFM-DSR solutions under different orders of enrichment functions. Taking $\hat{N} = 5$ as an example, the reduced-order solution from ISBFM-DSR yields 2 orders of magnitude reduction in errors compared to that of the ISBFM-MA approach and with less than 50% error compared to the ISBFM-DR approach. Moreover, ISBFM-DSR consistently gives improved results with increasing as the ratio of reduced dimension or the order of enrichment basis functions, whereas ISBFM-DR yields worse reduced-order approximation using $\hat{N} = 2$ than using $\hat{N} = 1$. This inconsistency in ISBFM-DR may be due to the suboptimal selection of the decomposed projection. Thus, hereafter, we only consider the comparison of ISBFM-DSR method and the ISBFM-MA method with the scaled near-tip basis functions with $\gamma = 50$ for the remaining numerical studies.

Figure 4.15 shows the error distribution in Euclidian error norm, $e_E(\mathbf{x}) = |\mathbf{u}^h(\mathbf{x}) - \mathbf{u}^r(\mathbf{x})|$, of the reduced-order solutions obtained by the ISBFM-MA and the ISBFM-DSR methods under the same reduced dimension $k = 80$ ($k/N \approx 3.1\%$) and with different orders of symmetric near-tip basis functions ($\hat{N} = 1, 2, 5$). Similar to the cracked beam problem in Figure 4.9, the ISBFM-MA method poorly approximates the behaviors near the crack tip at $(x, y) = (50, 0)$, whereas ISBFM-DSR yields a much better accuracy due to the preservation of the near-tip

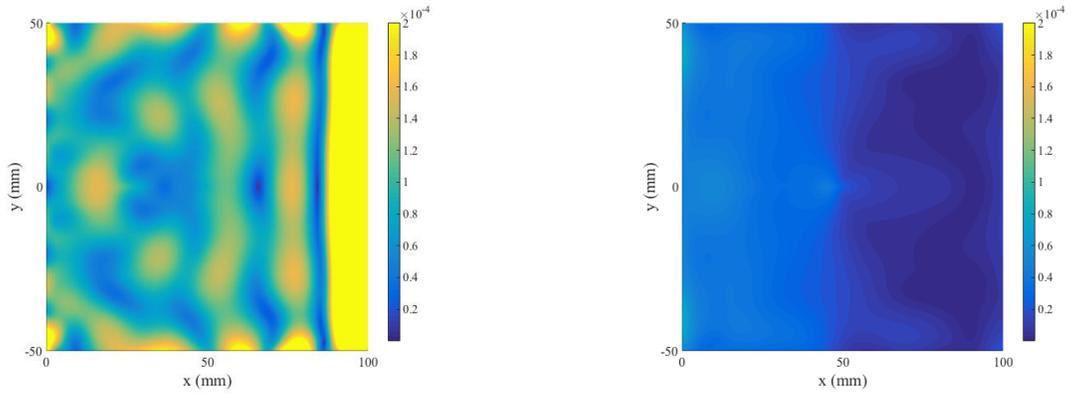
characteristics. As the order of enrichment functions increases from $\hat{N} = 1$ to $\hat{N} = 5$, the error of the ISBFM-DSR solution reduces significantly but the ISBFM-MA solution gives nearly the same level of error without improvement (See Table 4.2). Moreover, we observe that ISBFM-MA exhibits large errors near the right edge where the inhomogeneous Dirichlet boundary conditions are imposed as have been discussed in Section 4.3. On the other hand, ISBFM-DSR gives a better accuracy on that boundary, especially when using higher order enrichment functions (Figure 4.15(c)).

Table 4.2: Normalized error in L_2 norm (e_0) of the reduced-order solutions for the loading crack model. ISBFM-MA, ISBFM-DR and ISBFM-DSR are compared under different ratios of reduced dimension k / N and different orders of enrichment functions $\hat{N} = 1, 2, 5$.

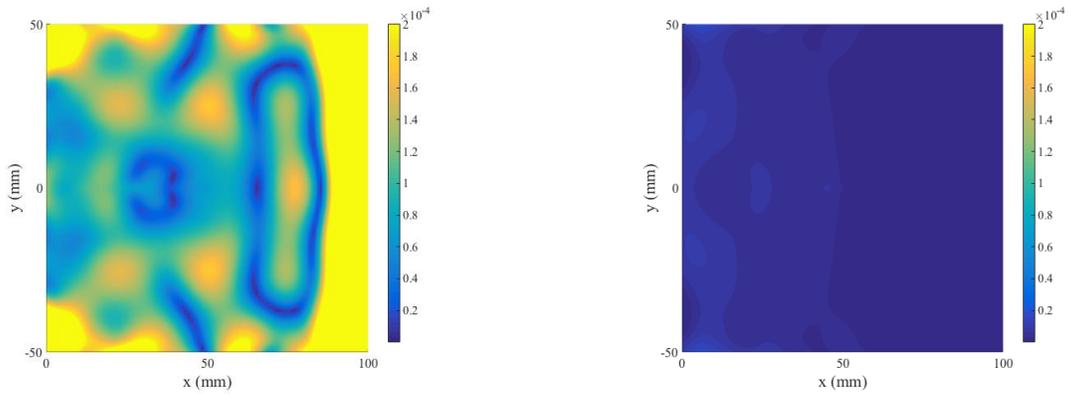
	k / N	0.77%	1.54%	3.09%	6.17%
ISBFM-MA (scaled $\gamma = 50$)	$\hat{N} = 1$	3.5×10^{-1}	2.9×10^{-1}	2.2×10^{-1}	1.9×10^{-1}
	$\hat{N} = 2$	3.4×10^{-1}	2.8×10^{-1}	2.2×10^{-1}	1.9×10^{-1}
	$\hat{N} = 5$	3.8×10^{-1}	2.9×10^{-1}	2.4×10^{-1}	2.0×10^{-1}
ISBFM-DR	$\hat{N} = 1$	5.4×10^{-2}	3.6×10^{-2}	3.4×10^{-2}	2.6×10^{-2}
	$\hat{N} = 2$	7.3×10^{-2}	6.2×10^{-2}	5.0×10^{-2}	3.6×10^{-2}
	$\hat{N} = 5$	8.0×10^{-3}	9.2×10^{-3}	9.5×10^{-3}	9.5×10^{-3}
ISBFM-DSR	$\hat{N} = 1$	7.2×10^{-2}	6.6×10^{-2}	6.4×10^{-2}	6.3×10^{-2}
	$\hat{N} = 2$	2.6×10^{-2}	2.3×10^{-2}	2.2×10^{-2}	2.1×10^{-2}
	$\hat{N} = 5$	4.4×10^{-3}	4.3×10^{-3}	3.5×10^{-3}	2.7×10^{-3}



(a) ISBFM-MA (left) and ISBFM-DSR (right) with $k = 80$, $\hat{N} = 1$



(b) ISBFM-MA (left) and ISBFM-DSR (right) with $k = 80$, $\hat{N} = 2$



(c) ISBFM-MA (left) and ISBFM-DSR (right) with $k = 80$, $\hat{N} = 5$

Figure 4.15: Error distribution in Euclidian norm of the reduced-order solutions for the loading line crack model with the reduced dimension $k = 80$ ($k / N \approx 3.1\%$). ISBFM-MA ($\gamma = 50$) and ISBFM-DSR are compared under different orders of enrichment functions $\hat{N} = 1, 2, 5$.

The closer observations on the reduced-order solutions u_2' along $x = 49.65$ mm and the derivative du_2'/dy along $x = 50.35$ mm under different ratios of reduced dimension ($k/N = 0.8\%$, 3.1% , and 6.2%) are presented in Figure 4.16 and Figure 4.17, respectively. Compared to the ISBFM-MA approach, ISBFM-DSR produces better results in capturing the displacement discontinuity across the crack surfaces (Figure 4.16) and the strain singularity near the crack tip (Figure 4.17), as well as approximating the solution near boundaries. The error distributions of the reduced-order solutions near the crack tip region (indicated by the red box, i.e. $y \in [-15, 15]$) are also provided in Figure 4.17. It shows that the reduced-order solutions of ISBFM-DSR converge quickly while the ISBFM-MA reduction does not guarantee improved results as the reduced dimension increases. Moreover, it is observed that by using ISBFM-DSR the near-tip error is well-controlled by the non-smooth bases and less sensitive to the smooth subspace approximation. Thus, using the higher order enrichment functions (e.g., $\hat{N} = 5$) in ISBFM-DSR significantly improves the near-tip accuracy in the reduced-order approximation even with a very low ratio of reduced dimension (e.g., $k/N = 0.8\%$). The ISBFM-MA reduction, however, does not show increased near-tip solution improvement when using higher order enrichment functions.

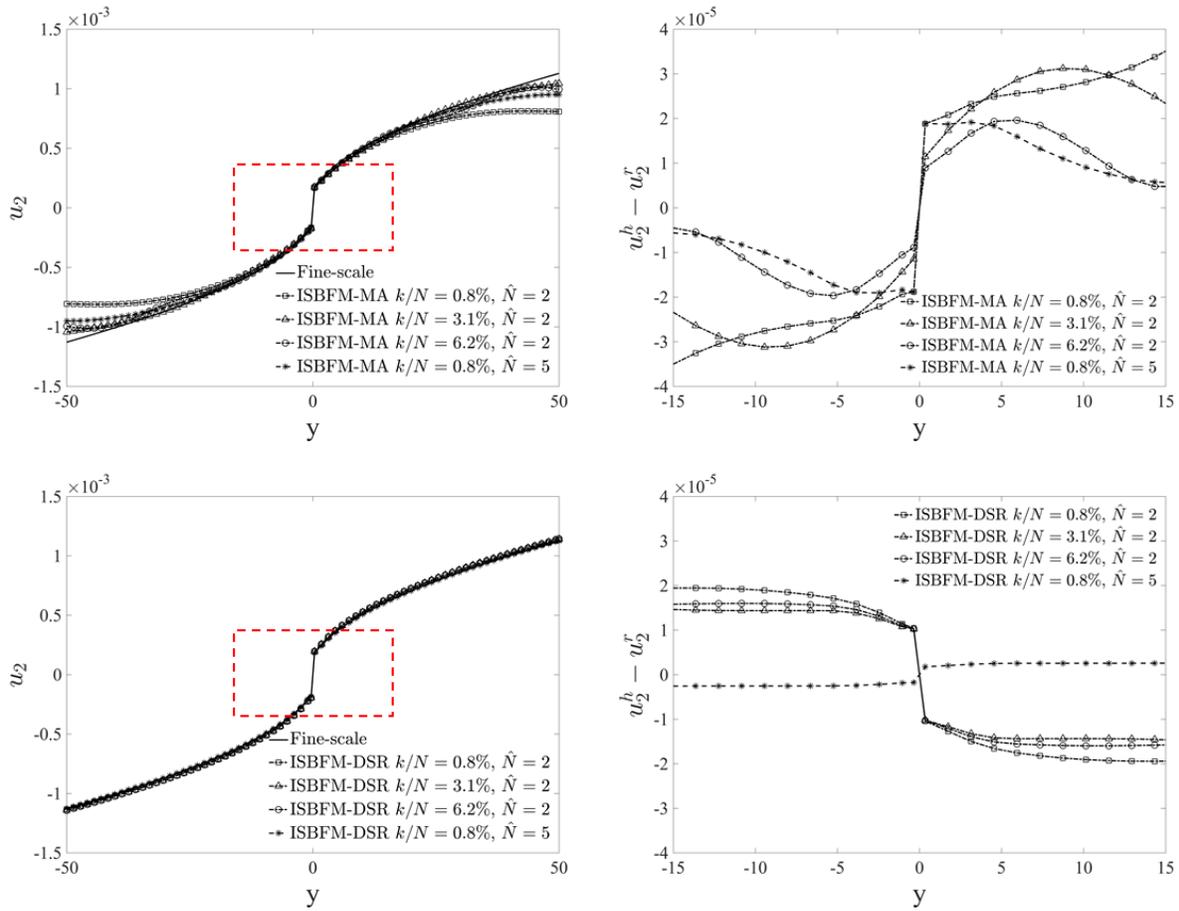


Figure 4.16: Comparison of the fine-scale solution and the reduced-order solutions of u_2 along $x = 49.65$ mm for the loading line crack model are shown in the left column. The corresponding error distributions of the reduced-order solutions against the fine-scale solution near the crack tip region (denoted by the red box) are shown in the right column.

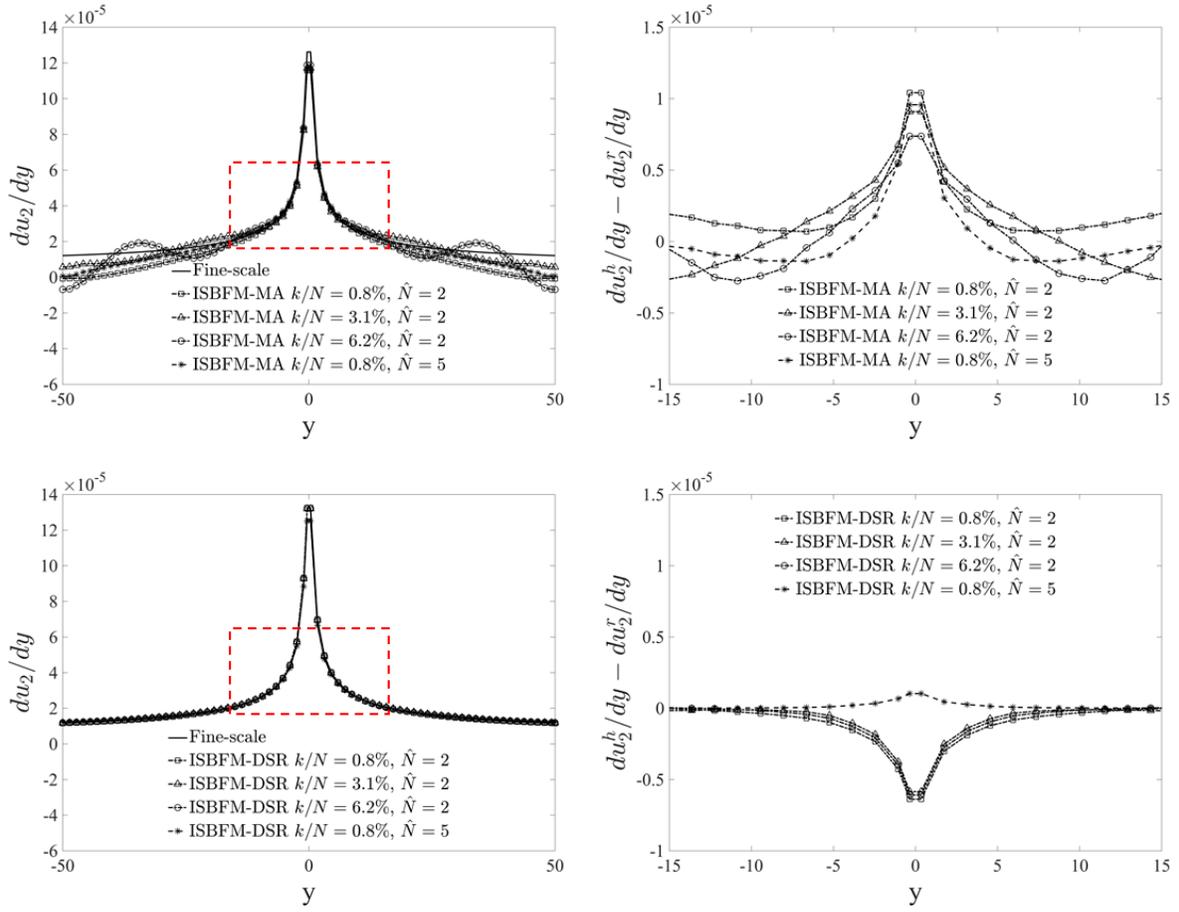


Figure 4.17: Comparison of the fine-scale solution and the reduced-order solutions of du_2/dy along $x=50.35$ mm for the loading line crack model are shown in the left column. The corresponding error distributions of the reduced-order solutions against the fine-scale solution near the crack tip region (denoted by the red box) are shown in the right column.

4.4.3 Reduced-order modeling of the mixed mode problem

Here, we consider a plane-strain mixed-mode problem with an edge crack, and the top surface is subjected to a shear load as shown in Figure 4.18. The reference stress intensity factor (SIF) solutions (Yau, Wang & Corten 1980) are $K_I = 34.0 \text{ psi}\sqrt{\text{in}}$ and $K_{II} = 4.55 \text{ psi}\sqrt{\text{in}}$. The fine-scale solution is obtained from the ISBFM Galerkin formulation using a uniform discretization with 30×48 ($\bar{N} = 1440$) nodes. Considering the mixed-mode situation, the first

order symmetric and anti-symmetric near-tip basis functions in (4.8) are used. The SIFs are computed by a mixed mode J -integral formulation (Shih & Asaro 1988; Yau, Wang & Corten 1980) that uses the near crack tip fields as auxiliary fields. The following relative errors of SIFs are used to measure the performance of reduced-order solutions,

$$e_{KI} = \frac{|K_I^h - K_I^r|}{|K_I^h|}, \quad e_{KII} = \frac{|K_{II}^h - K_{II}^r|}{|K_{II}^h|} \quad (4.39)$$

where K_I^h and K_{II}^h are the Mode I and Mode II SIFs solved by fine-scale approximation, respectively, and K_I^r and K_{II}^r are the Mode I and Mode II SIFs solved by reduced-order approximation, respectively.

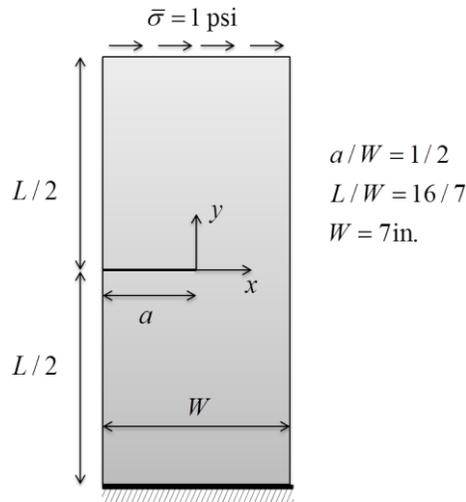


Figure 4.18: Schematic model of the mixed-mode crack problem. The adopted material properties are $E = 10,000$ psi and $\nu = 0.3$.

The errors e_{KI} and e_{KII} of the reduced-order solutions under different ratios of reduced dimension are presented in Table 4.3. The ISBFM-MA approach yields large errors of $O(10^{-2})$ for K_I^r and of $O(10^{-1})$ for K_{II}^r even when the ratio of reduced dimension k/N reaches 90%.

Compared to ISBFM-MA, the SIF results computed by ISBFM-DSR show better agreement with the fine-scale approximation. Moreover, as shown in Figure 4.19, while ISBFM-MA using the properly scaled near-tip basis functions ($\gamma = 50$) gives an acceptable results ($O(10^{-1})$), the reduced-order solutions from ISBFM-DSR not only have better accuracy but also converge faster to the fine-scale solution as the reduced dimension increases.

Table 4.3: Relative errors of SIFs of the reduced-order solutions for the mixed-mode crack model solved with $\hat{N} = 1$. ISBFM-MA and ISBFM-DSR are compared under different ratios of reduced dimension k / N .

Reference stress intensity factors: $K_I = 34.0 \text{ psi}\sqrt{\text{in}}$, $K_{II} = 4.55 \text{ psi}\sqrt{\text{in}}$

k / N	e_{KI}		e_{KII}	
	ISBFM-MA	ISBFM-DSR	ISBFM-MA	ISBFM-DSR
2.8%	6.0×10^{-2}	1.7×10^{-2}	4.1×10^{-1}	1.2×10^{-1}
38.2%	5.3×10^{-2}	4.3×10^{-3}	9.8×10^{-2}	4.0×10^{-2}
90.2%	1.5×10^{-2}	8.9×10^{-6}	9.5×10^{-2}	4.1×10^{-4}

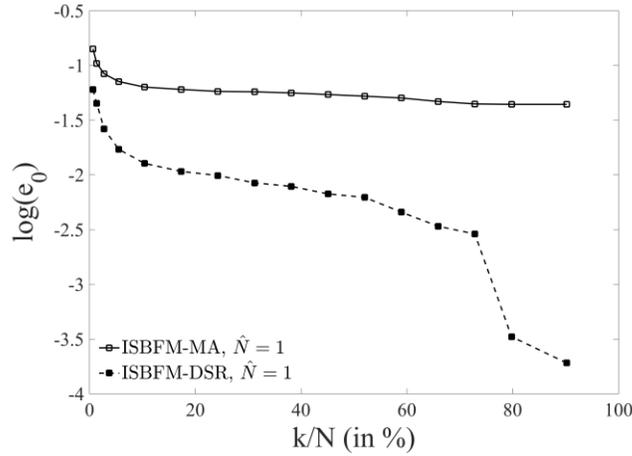


Figure 4.19: Normalized error in L_2 norm (e_0) of the reduced-order solution for the mixed-mode crack model solved with $\hat{N}=1$. ISBFM-MA ($\gamma=50$) and ISBFM-DSR are compared under different ratios of reduced dimension k/N .

4.5 Summary

The development of the decomposed subspace reduction (DSR) method for ISBFM based Galerkin meshfree solution of linear elastic fracture mechanics have been presented. The employment of ISBFM based Galerkin Meshfree method for the fine-scale fracture model not only allows a lower order integration of the Galerkin equation, it also yields a discrete system containing sparse sub-matrices for effective MOR procedures. With this computational framework for solving the fine-scale solution of fracture problems, the proposed ISBFM-DSR low-rank representation optimally preserves the discontinuity and singularity characteristics of fracture.

In this work, the properties of various MOR methods have been investigated and compared. It has been shown that the standard modal analysis approach for constructing reduced-order approximation of ISBFM fine scale model is ineffective with a direct employment of the

near-tip basis functions due to its inappropriate scaling effect to the stiffness matrix. As such, the low energy eigenmodes do not properly capture the near-tip responses. Further, it is also shown that the inhomogeneous Dirichlet boundary conditions (IDBC) are not correctly represented by the modal analysis approach. On the other hand, by means of the decomposed projection, the proposed DSR method preserves the essential crack-tip behaviors as well as IDBC in the low-rank approximation.

This study also shows that ISBFM-DSR performs better than the decomposed ISBFM-DR method previously proposed for scalar cracked beam problem (Chen, Marodon & Hu 2015). The difference in performance in these two methods is due to the employment of projection matrix for the smooth solution. In ISBFM-DSR, the projection is based on the sub-matrix of stiffness associated with the smooth degrees of freedom instead of the Schur complement obtained from the condensed sub-matrix of stiffness associated with both smoothed and non-smooth degrees of freedom in ISBFM-DR. It also shows that the reduced-order solution from ISBFM-DSR is not affected by the arbitrary scaling of enrichment functions since the project matrix $\bar{\mathbf{P}}$ is independent of the non-smooth bases.

Although this study focuses on the static linear elastic fracture mechanics problems, the proposed ISBFM-DSR framework can be extended to crack propagation problems by further introducing techniques such as the local/global strategy (Kerfriden, Passieux & Bordas 2012; Passieux et al. 2013) or the subdomain method (Kerfriden et al. 2013; Radermacher & Reese 2014). For fracture problems with material and geometric nonlinearities, while the same near-tip basis functions can be employed, the reduced-order basis will need to be updated at certain load steps based on some nonlinearity measures, and suitable hyperreduction schemes (Chaturantabut

& Sorensen 2010; Kerfriden et al. 2013; Ryckelynck, Benziane & Paristech 2010) could be applied to further reduce computational cost.

4.6 Acknowledgement

This chapter, in part, is a reprint of the materials as they appear in: “He, Qizhi; Chen, Jiun-Shyan; Marodon, Camille. A decomposed subspace reduction for fracture mechanics based on the meshfree integrated singular basis function method. *Computational Mechanics* (2018): 1-22”. The dissertation author is the primary investigator and author of this paper.

4.7 Appendix

The discrete ISBFM matrix equation is obtained by introducing RK approximation together with enrichment functions into the ISBFM Galerkin weak formulations (4.12) and (4.13):

$$\begin{bmatrix} \bar{\mathbf{K}} & \hat{\mathbf{K}} \\ \hat{\mathbf{K}}^T & \hat{\mathbf{K}} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{d}} \\ \hat{\mathbf{d}} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{f}} \\ \hat{\mathbf{f}} \end{bmatrix} \quad (4.40)$$

The components of $\bar{\mathbf{K}} \in \mathbb{R}^{2\bar{N} \times 2\bar{N}}$, $\hat{\mathbf{K}} \in \mathbb{R}^{2\hat{N} \times 2\hat{N}}$ and $\hat{\mathbf{K}} \in \mathbb{R}^{2\bar{N} \times 2\hat{N}}$ in 2-dimension are defined as follows

$$\begin{bmatrix} \bar{K}_{(2I-1)(2J-1)} & \bar{K}_{(2I-1)(2J)} \\ \bar{K}_{(2I)(2J-1)} & \bar{K}_{(2I)(2J)} \end{bmatrix} = \int_{\Omega} \bar{\mathbf{B}}_I^T \mathbf{C} \bar{\mathbf{B}}_J d\Omega - \int_{\Gamma_{sk}} \left(\Psi_I \mathbf{e}_k \mathbf{e}_k^T \mathbf{N} \mathbf{C} \bar{\mathbf{B}}_J + \bar{\mathbf{B}}_I^T \mathbf{C}^T \mathbf{N}^T \mathbf{e}_k \mathbf{e}_k^T \Psi_J \right) d\Gamma + \beta \int_{\Gamma_{sk}} \Psi_I \Psi_J \mathbf{e}_k \mathbf{e}_k^T d\Gamma \quad (4.41)$$

$$\begin{bmatrix} \hat{K}_{(2I-1)(2J-1)} & \hat{K}_{(2I-1)(2J)} \\ \hat{K}_{(2I)(2J-1)} & \hat{K}_{(2I)(2J)} \end{bmatrix} = \int_{\bar{\Gamma}_{h_k}} \mathbf{F}_I^T \mathbf{e}_k \mathbf{e}_k^T \mathbf{N} \mathbf{C} \hat{\mathbf{B}}_J d\Gamma - \int_{\Gamma_{gk}} \hat{\mathbf{B}}_I^T \mathbf{C}^T \mathbf{N}^T \mathbf{e}_k \mathbf{e}_k^T \mathbf{F}_J d\Gamma + \beta \int_{\Gamma_{gk}} \mathbf{F}_I^T \mathbf{e}_k \mathbf{e}_k^T \mathbf{F}_J d\Gamma \quad (4.42)$$

$$\begin{bmatrix} \hat{K}_{(2I-1)(2J-1)} & \hat{K}_{(2I-1)(2J)} \\ \hat{K}_{(2I)(2J-1)} & \hat{K}_{(2I)(2J)} \end{bmatrix} = \int_{\bar{\Gamma}_{h_k}} \Psi_I \mathbf{e}_k \mathbf{e}_k^T \mathbf{N} \mathbf{C} \hat{\mathbf{B}}_J d\Gamma - \int_{\Gamma_{gk}} \bar{\mathbf{B}}_I^T \mathbf{C}^T \mathbf{N}^T \mathbf{e}_k \mathbf{e}_k^T \mathbf{F}_J d\Gamma + \beta \int_{\Gamma_{gk}} \Psi_I \mathbf{e}_k \mathbf{e}_k^T \mathbf{F}_J d\Gamma \quad (4.43)$$

where $\bar{\mathbf{B}}_I$, $\hat{\mathbf{B}}_J$, and \mathbf{N} are

$$\bar{\mathbf{B}}_I = \begin{bmatrix} \Psi_{I,1} & 0 \\ 0 & \Psi_{I,2} \\ \Psi_{I,2} & \Psi_{I,1} \end{bmatrix}, \hat{\mathbf{B}}_J = \begin{bmatrix} F_{J1,1}^s & F_{J1,1}^{as} \\ F_{J2,2}^s & F_{J2,2}^{as} \\ F_{J1,2}^s + F_{J2,1}^s & F_{J1,2}^{as} + F_{J2,1}^{as} \end{bmatrix}, \mathbf{N} = \begin{bmatrix} n_1 & 0 & n_2 \\ 0 & n_2 & n_1 \end{bmatrix} \quad (4.44)$$

and \mathbf{e}_k represents the unit normal outward vector on the boundary. Accordingly, the components

of the force vectors $\bar{\mathbf{f}} \in \mathbb{R}^{2\bar{N}}$ and $\hat{\mathbf{f}} \in \mathbb{R}^{2\hat{N}}$ are

$$\begin{bmatrix} \bar{f}_{2I-1} \\ \bar{f}_{2I} \end{bmatrix} = \int_{\bar{\Gamma}_{h_k}} \Psi_I \mathbf{e}_k h_k d\Gamma - \int_{\Gamma_{gk}} \bar{\mathbf{B}}_I^T \mathbf{C}^T \mathbf{N}^T \mathbf{e}_k g_k d\Gamma + \beta \int_{\Gamma_{gk}} \Psi_I \mathbf{e}_k g_k d\Gamma \quad (4.45)$$

$$\begin{bmatrix} \hat{f}_{2I-1} \\ \hat{f}_{2I} \end{bmatrix} = \int_{\bar{\Gamma}_{h_k}} \mathbf{F}_I^T \mathbf{e}_k h_k d\Gamma - \int_{\Gamma_{gk}} \hat{\mathbf{B}}_I^T \mathbf{C}^T \mathbf{N}^T \mathbf{e}_k g_k d\Gamma + \beta \int_{\Gamma_{gk}} \mathbf{F}_I^T \mathbf{e}_k g_k d\Gamma \quad (4.46)$$

Chapter 5

Development of Manifold Learning Based Model Order Reduction for Nonlinear Mechanics

This chapter presents a new approach to construct a robust reduced order model for nonlinear partial differential equations (PDEs) using a manifold based projection, called linear graph embedding (LGE) projection. The LGE projection is derived from a generalized manifold learning framework, called graph embedding, in conjunction with linearization techniques. The employment of LGE allows to introduce *a priori* statistical knowledge as well as local information of data in the projection construction. This study shows that the proposed projection better preserves the underlying data structure than the canonical POD projection for systems characterized by a wide variety of nonlinear behaviors, and it yields a robust reduced-order model less sensitive to noise and outliers. It also shows that the manifold learning technique used to determine robust projections can be viewed as a generalization of the weighted POD method.

5.1 Introduction

In Chapter 3, we introduce a typical nonlinear model order reduction method based on POD and DEIM for nonlinear parameterized systems. The POD-DEIM scheme or the similar variant has been widely applied to reduced-order modeling in fluid mechanics (Carlberg et al. 2013; Carlberg, Bou-Mosleh & Farhat 2011) and solid mechanics, such as fracture problems

(Goury et al. 2016; Kerfriden et al. 2013), hyperelastic structure (Radermacher & Reese 2016), elasto-plastic structure (Ghavamian, Tiso & Simone 2017; Ryckelynck, Lampoh & Quilicy 2016), and computational multiscale modeling (Goury et al. 2016; Hernandez et al. 2014; Soldner et al. 2017). However, since the truncated POD basis is very sensitive to the magnitude of the snapshots, the POD-DEIM scheme lacks robustness for many other applications that typically involve a wide range of parameter changes (Amsallem & Farhat 2008; Carlberg & Farhat 2008), such as design optimization, control and data-driven systems. Intuitively, the online computations of such systems call for adaptively using new reduced-order modes (ROMs) each time a sensitive physical or modeling parameter is varied. However, reconstructing a ROM online is usually computationally expensive, and might compromise the gain from using MOR methods.

There are at least three approaches that have been proposed to address the aforementioned adaptation issue in the context of the POD method (Peng & Mohseni 2016): global POD (GPOD), local POD, and adaptive POD. Global POD approximates the solution of interest using a global basis obtained based on the idea of data weighting to enhance robustness (Carlberg & Farhat 2011; Christensen, Brøns & Sørensen 1999; Schmit & Glauser 2004; Taylor & Glauser 2004). In the local POD approach, the precomputed snapshots usually can be clustered into multiple subdomains, either through parameter (or time) domain partitions (Eftang & Stamm 2012; Haasdonk, Dihlmann & Ohlberger 2011; Peng & Mohseni 2014), or space domain partitions (Amsallem, Zahr & Farhat 2012; Burkardt, Gunzburger & Lee 2006; Kerfriden et al. 2013; Peherstorfer et al. 2014). Then POD is performed at each subdomain of the snapshots to construct the local reduced-order basis to project the original system onto the associated subspace. During the online computation, an appropriate subdomain along with its local

subspace is selected for reduced modeling, depending on the current state (parameters) of the system. In contrary to local POD, adaptive POD utilizes all the precomputed snapshot data and constructs adaptive reduced basis through subspace interpolation methods, such as angle interpolation (Lieu, Farhat & Lesoinne 2006; Lieu & Lesoinne 2004) and differential geometry-based interpolation (Amsallem & Farhat 2008, 2011).

The local approach based on partitions of snapshots (Amsallem, Zahr & Farhat 2012) or parameter space (Peherstorfer et al. 2014), however, could encounter difficulties such as unstable clustering behavior for large numbers of clusters if the clustering method and its classification parameters are not carefully chosen (Von Luxburg 2007). Thus, it may still need a large number of subdomains due to the poor division or clustering. Furthermore, numerical instabilities during the online reduced order computing could be introduced when the selected reduced model is switched from one to another (Amsallem, Zahr & Farhat 2012; Idelsohn & Cardona 1985) if the state solution lies just between two partitions. If there are strong bifurcation within the solution or parameter space (Brunton et al. 2014; Sargsyan, Brunton & Kutz 2015), information across several neighbor partitions may be needed. On the other hand, the subspace interpolation approach is usually a low-order interpolation method and might be restricted to the interpolation between two precomputed reduced-order bases corresponding to two parameters that are sufficiently close. Moreover, adaptive POD usually requires constructing a reduced system during the online stage rather than the offline stage such that it is less efficient compared to the other two strategies.

Given only a few sets of snapshots are provided in most solid mechanics problems of interest compared to time-dependent systems, we prefer to take advantage of all data and thus, the clustering is not an appropriate solution. To handle the issue of irregular distribution of

snapshots, we present a new model reduction method inspired from the Graph embedding framework (Belkin & Niyogi 2003; Yan et al. 2005, 2007) to constructed robust reduced-order projections. The proposed method can be related to the so-called weighted POD technique, which has been demonstrated effective in the tests of some parameter-dependent systems (Bistrián & Susan-Resiga 2016; Carlberg & Farhat 2011; Christensen, Brøns & Sørensen 1999).

This chapter is organized as follows. The formulations of standard nonlinear model order reduction based on POD and DEIM are briefly summarized in Section 5.2. To over the limitation of POD (or known as principal component analysis (PCA) in machine learning area) for linear manifold, manifold learning techniques used to discover nonlinear data structure are reviewed in Section 5.3. Then, a general framework based on graph embedding is introduced to develop the proposed LGE projection by using local information of data, which is shown less sensitive to outliers or other irregular data. Section 5.4 presents examples to demonstrate the effectiveness of LGE. Section 5.5 gives the concluding remarks of this Chapter.

5.2 Problem Formulation: Nonlinear Model Order Reduction

5.2.1 High-dimensional parameterized nonlinear discrete model

We consider a generic parameterized partial differential equations (PDEs), where the parameters may represent different physical quantities the system depends on, for example, material properties, system configuration, initial conditions, and boundary conditions. Let Ω denote the problem spatial domain, $\mathcal{D} \subset \mathbb{R}^l$ denote a predefined parameter domain, and $\mathcal{R} \subset \mathbb{R}^{\mathcal{N}}$ denote a solution manifold embedded in \mathcal{N} -dimensional Euclidean space, where \mathcal{N} is the number of degree of freedoms (DOFs) of the associated discretized system. By spatial

discretization (e.g., finite difference, finite element, or meshfree methods), a parameterized steady PDE in discrete manner can be expressed as

$$\mathbf{A}\mathbf{u} + \mathbf{f}(\boldsymbol{\mu}, \mathbf{u}) = \mathbf{0}, \quad (5.1)$$

where $\mathbf{A} \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ and $\mathbf{f} : \mathcal{D} \times \mathcal{R} \rightarrow \mathbb{R}^{\mathcal{N}}$ correspond to the linear and the nonlinear term of the PDE, respectively. \mathbf{A} is a constant matrix, but the function \mathbf{f} evaluated at all spatial grid points nonlinearly depends on the parameters $\boldsymbol{\mu}$ and the current state \mathbf{u} . The solution (or state vector) $\mathbf{u} : \Omega \times \mathcal{D} \rightarrow \mathbb{R}^{\mathcal{N}}$ is implicitly defined by Equation Online/offline splitting given any fixed input parameters $\boldsymbol{\mu} \in \mathcal{D}$. Thus, the state is dependent on the input parameters, and we denote $\mathbf{u} = \mathbf{u}(\boldsymbol{\mu})$ with some abuse of notation.

Similarly, a spatial discretization of a time dependent PDE results in a system of ordinary differential equation (ODE) as follows

$$\dot{\mathbf{u}} = \mathbf{A}\mathbf{u} + \mathbf{f}(t, \boldsymbol{\mu}, \mathbf{u}), \quad (5.2)$$

with an initial condition $\mathbf{u}(0, \boldsymbol{\mu}) = \mathbf{u}_0$, where $\mathbf{f} : \mathcal{T} \times \mathcal{D} \times \mathcal{R} \rightarrow \mathbb{R}^{\mathcal{N}}$ denotes the discretized nonlinear term, where $\mathcal{T} = [0, T] \subset \mathbb{R}$ denotes a time domain. For formulating both time dependent and steady parameterized PDEs under the same computational framework, $\boldsymbol{\mu} \in \mathcal{D}$ is also used to represent the Cartesian product $(t, \boldsymbol{\mu}) \in \mathcal{T} \times \mathcal{D}$ in (5.2) with some abuse of notation. As such, $\boldsymbol{\mu}$ denotes generic parameters that could be time, material coefficients, boundary conditions, etc.

The Jacobian of the above parameterized systems in Online/offline splitting and (5.2) is given by

$$\mathbf{J}(\boldsymbol{\mu}, \mathbf{u}) = \mathbf{A} + \mathbf{J}_f(\boldsymbol{\mu}, \mathbf{u}), \quad (5.3)$$

where $\mathbf{J}(\boldsymbol{\mu}, \mathbf{u}) \in \mathbb{R}^{N \times N}$ is a function of $\boldsymbol{\mu}$ and \mathbf{u} .

5.2.2 Proper orthogonal decomposition

The POD method introduced in Chapter 3 is used to construct the reduced-order basis matrix \mathbf{V} for the dimensionality reduction of the state unknown. The collection of discrete snapshots (Sirovich 1987) are precomputed by the original high-dimensional system (either Online/offline splitting or (5.2)). Let $\mathbf{X}_s = [\mathbf{u}_1, \dots, \mathbf{u}_{N_s}] \in \mathbb{R}^{N \times N_s}$ be a snapshot matrix, where each column $\mathbf{u}_i = \mathbf{u}(\mu_i)$ is the solution corresponding to input parameters μ_i . The key idea of POD is to find a k -dimensional subspace $\tilde{\mathcal{X}}_s = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ that best fits the snapshot space $\mathcal{X}_s = \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_{N_s}\}$. As such, the solution $\mathbf{u}(\boldsymbol{\mu})$ is approximated on the k -dimensional subspace given by the ROB basis

$$\tilde{\mathbf{u}}(\boldsymbol{\mu}) = \bar{\mathbf{u}} + \mathbf{V}\mathbf{u}^r(\boldsymbol{\mu}), \quad (5.4)$$

where $\mathbf{u}^r \in \mathbb{R}^k$ denote the reduced coefficients (the state variable in the subspace coordinate system), $\bar{\mathbf{u}}$ is usually the mean value of snapshots in boundary value problems (Chatterjee 2000) or the initial condition \mathbf{u}_0 in initial-value problems (Amsallem, Zahr & Farhat 2012). A debating point in the community is whether the “ensemble” mean should be subtracted from the snapshot matrix for the application of POD. Most previous studies in literature adopted the mean subtraction. A discussion of this point is given in (Chatterjee 2000). In this study, POD is applied to the centered state snapshot matrix \mathbf{X}_s by subtracting $\bar{\mathbf{u}}$.

Following Galerkin projection and projecting the steady system Online/offline splitting onto $\tilde{\mathcal{X}}_s$, it yields the corresponding reduced-order model

$$\mathbf{V}^T \mathbf{A} \mathbf{V} \mathbf{u}^r + \mathbf{V}^T \mathbf{f}(\mu, \mathbf{V} \mathbf{u}^r) = \mathbf{0}. \quad (5.5)$$

Analogously, a reduced system of a time dependent PDE (5.2) is

$$\dot{\mathbf{u}}^r = \mathbf{V}^T \mathbf{A} \mathbf{V} \mathbf{u}^r + \mathbf{V}^T \mathbf{f}(\mu, \mathbf{V} \mathbf{u}^r). \quad (5.6)$$

The associated reduced Jacobian is

$$\tilde{\mathbf{J}}(\mu, \mathbf{u}^r) = \mathbf{V}^T \mathbf{A} \mathbf{V} + \mathbf{V}^T \mathbf{J}_f(\mu, \mathbf{V} \mathbf{u}^r) \mathbf{V}. \quad (5.7)$$

As pointed out in Chapter 3, despite its low dimension, the resulting nonlinear ROM does not guarantee computational efficiency. This is because the online evaluation of the nonlinear function not only scales with the dimension of the reduced state variable, but also the size of the underlying high-dimensional model (HDM). This is called lifting bottleneck.

5.2.3 Discrete empirical interpolation method (DEIM)

Many nonlinear MOR methods equipped with a *hyper reduction* technique have been proposed to tackle this lifting bottleneck at the cost of introducing an additional layer of approximation. Because the proposed manifold-based projection can be combined with DEIM to solve nonlinear systems, we briefly review DEIM in this section. More details about DEIM can be found in Section 3.4.2. The key idea behind DEIM is to compute the nonlinear term at a small selected set of spatial grid points and extrapolate its value at all the other points. Introducing DEIM, the nonlinear vector can be approximated by

$$\mathbf{f}(\mathbf{u}(\mu)) \approx \tilde{\mathbf{f}}(\mathbf{u}(\mu)) = \mathbf{Z}(\mathbf{P}^T \mathbf{Z})^\dagger \mathbf{P}^T \mathbf{f}(\mathbf{u}(\mu)) = \mathbf{Z} \hat{\mathbf{Z}}^\dagger \mathbf{P}^T \mathbf{f}(\mathbf{u}(\mu)), \quad (5.8)$$

where $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_{\hat{k}}] \in \mathbb{R}^{\mathcal{N} \times \hat{k}}$ denotes the collateral POD basis of the nonlinear term snapshot $\mathbf{f}(\mu, \mathbf{u}(\mu))$ and $\mathbf{P} \in \mathbb{R}^{\mathcal{N} \times \hat{n}}$ is the selection matrix obtained by a greedy algorithm (Chaturantabut & Sorensen 2010). It is aware that when $\hat{n} > \hat{k}$ it is also referred to as Gappy-POD (see Section 3.4.2).

As a result, we can obtain the POD-DEIM reduced systems of the two PDEs in Online/offline splitting and (5.2) as follows

$$\mathbf{V}^T \mathbf{A} \mathbf{V} \mathbf{u}^r + \mathbf{V}^T \mathbf{Z} \hat{\mathbf{Z}}^\dagger \mathbf{P}^T \mathbf{f}(\mu, \mathbf{V} \mathbf{u}^r) = \mathbf{0}, \quad (5.9)$$

$$\dot{\mathbf{u}}^r = \mathbf{V}^T \mathbf{A} \mathbf{V} \mathbf{u}^r + \mathbf{V}^T \mathbf{Z} \hat{\mathbf{Z}}^\dagger \mathbf{P}^T \mathbf{f}(\mu, \mathbf{V} \mathbf{u}^r), \quad (5.10)$$

and the associated reduced Jacobian is

$$\hat{\mathbf{J}}(\mu, \mathbf{u}^r) = \mathbf{V}^T \mathbf{A} \mathbf{V} + \mathbf{V}^T \mathbf{Z} \hat{\mathbf{Z}}^\dagger \mathbf{P}^T \mathbf{J}_f(\mu, \mathbf{V} \mathbf{u}^r) \mathbf{V}. \quad (5.11)$$

Notice that $\mathbf{V}^T \mathbf{A} \mathbf{V} \in \mathbb{R}^{k \times k}$ and $\mathbf{V}^T \mathbf{Z} \hat{\mathbf{Z}}^\dagger \in \mathbb{R}^{k \times \hat{n}}$ are calculated only once at the offline stage, and $\mathbf{P}^T \mathbf{f}(\mu, \mathbf{V} \mathbf{u}^r) \in \mathbb{R}^{\hat{n}}$ requires evaluation only at \hat{n} entries online of $\mathbf{f}(\mu, \mathbf{V} \mathbf{u}^r)$.

5.2.4 Potential limitations of POD based model reduction methods

It has been reported that if the nonlinear system exhibits a wide variety of behaviors over the range of parameter domain, the local features are hard to be represented in a single global basis constructed by POD learning approach (Amsallem, Zahr & Farhat 2012; Peherstorfer et al. 2014; Peng & Mohseni 2016; Sargsyan, Brunton & Kutz 2015). This is because POD based on least squares minimization tends to preserve the faraway data information and is easily misled

by the “outliers”. From statistical perspective, POD method only works well for data that is Gaussian or lying on a “flat” manifold. The drawbacks of POD will be further explained in Section 5.3.

Compared to the POD approximation of state solutions, it is even worse for the DEIM approximation of nonlinear functions because those nonlinear data in PDEs usually are less smooth and nonlinear over the parameter domain than the state solution. Take the one-dimensional Burgers’ problem in Section 5.4.3 as example, the state solution over the parameter of time \mathcal{T} exhibits as a moving “step” function along the spatial domain, whereas the nonlinear solution behaves like a traveling Dirac Delta function. Therefore, for DEIM approximation a relatively large number of POD basis vectors and associated interpolation points are needed in order to accurately represent $f(\mu)$ over all possible situations of interest. It may lead to a more time-consuming evaluation in the reduced system than the full system.

This difficulty motivates the present study to develop an alternative subspace learning method to preserve better nonlinear structure of data than the standard POD method and provide an outlier-insensitive projection for nonlinear model reduction via DEIM.

5.3 Manifold Learning via Graph Embedding Framework

5.3.1 Manifold learning

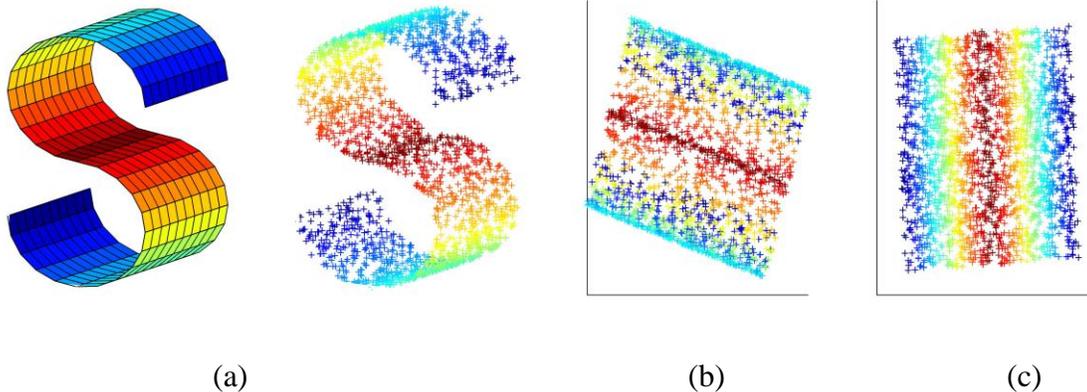


Figure 5.1: (a) Two-dimensional manifold structure represented by the three-dimensional S-curve data set; (b) Two-dimensional embedding obtained by PCA; (c) Two-dimensional embedding obtained by a nonlinear dimensionality reduction technique, e.g. LSTA (Zhang & Zha 2004).

POD is essentially identical to Principal Component Analysis (PCA) (Jolliffe 2002), which is a classical dimensionality reduction technique for manifold learning. PCA is designed to find a set of mutually orthogonal basis functions representing the directions of maximum variance of the original n -dimensional data so that the pairwise Euclidean distances can be best preserved in the d ($\ll n$)-dimensional linear subspace. If the data is embedded linearly or almost linearly in the ambient space, PCA is guaranteed to discover the dimensionality of the manifold and provide a “best” low-dimensional representation. In many real-world problems, however, data may be not sampled from a linear subspace. For example, it is always believed that the face images are sampled from a nonlinear low-dimensional manifold (He et al. 2005), or the snapshots gathered from nonlinear PDEs in this study. PCA has difficulty in discovering the

underlying nonlinear structure in these cases. For example, the S-curve (Figure 5.1(a)) represented by data is a two-dimensional nonlinear manifold embedded in 3D Euclidean space and can be characterized by two global internal coordinates, as shown in Figure 5.1(c). But the covariance matrix of data has full-rank, and thus, PCA is unable to find the exact two-dimensional embedding (Figure 5.1(b)).

Many nonlinear dimensionality reduction (NLDR), or called manifold learning techniques have been proposed to discover the nonlinear structure of the manifold (e.g., Figure 5.1(c)), including locally linear embedding (LLE) (Roweis & Saul 2000), ISOMAP (Tenenbaum, de Silva & Langford 2000), Laplacian eigenmaps (LE) (Belkin & Niyogi 2001), and local tangent space alignment (LTSA) (Zhang & Zha 2004), to name a few. Although the manifold learning methods were motivated and formulated differently, they can be unified in a general framework: graph embedding (Yan et al. 2005, 2007). These manifold learning methods are distinct in using different local information of data, such as pairwise distances, local neighborhoods, local linear relationships, but they all form a final optimization to search directly for a global low-dimensional coordinates through the integration of all local information based on the alignment trick (Zhang et al. 2009; Zhang & Zha 2004). For a more comprehensive review of other manifold learning methods, we refer the reader to (Lee & Verleysen 2007; Van Der Maaten, Postma & Van Den Herik 2009).

In the following sections, we first briefly review the graph embedding framework, which is developed based on Laplacian Eigenmap. Then, we introduce the linearization techniques (He & Niyogi 2004; Min, Lu & He 2004; Park et al. 2004; Xiaofei He et al. 2005; Yan et al. 2005) that lead to the graph-based projections that is used for robust reduced order modeling.

5.3.2 Laplacian eigenmaps (LE)

Laplacian eigenmaps (LE) (Belkin & Niyogi 2001) aims to capture the local information of the manifold by finding a low-dimensional representation in a way that nearby points in the manifold are mapped to nearby points in the low-dimensional embedding. Note that in this section, we follow the conventional notations in manifold learning community.

More specifically, if given a data matrix $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbb{R}^{n \times N}$ with column centering, where $\mathbf{x}_i \in \mathbb{R}^n$, n is the feature dimension, and N is the sample number, LE finds a low-dimensional ($d \ll n$) embedding $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_N] \in \mathbb{R}^{d \times N}$ (i.e., the reduced coefficients in the field of model order reduction) such that if \mathbf{x}_i and \mathbf{x}_j are close to each other, then so are \mathbf{y}_i and \mathbf{y}_j . This can be achieved by minimizing the Hall's energy (Koren, Carmel & Harel 2002):

$$\mathcal{E}(\mathbf{Y}) = \sum_{i,j}^N w_{ij} \|\mathbf{y}_i - \mathbf{y}_j\|^2, \quad (5.12)$$

subject to appropriate constraints on \mathbf{Y} that ensure the nontrivial solution. Letting $\mathbf{W} \in \mathbb{R}^{N \times N}$ be the matrix of weights whose entries w_{ij} measure the affinity between two points \mathbf{x}_i and \mathbf{x}_j , and \mathbf{D} be a diagonal matrix with diagonal entries $d_{jj} = \sum_i w_{ij}$, the objective function can be rewritten as

$$\begin{aligned} \mathcal{E}(\mathbf{Y}) &= \sum_{i,j}^N w_{ij} (\|\mathbf{y}_i\|^2 + \|\mathbf{y}_j\|^2 - 2\mathbf{y}_i^T \mathbf{y}_j) = 2 \sum_j^N d_{jj} \mathbf{y}_j^T \mathbf{y}_j - 2 \sum_{i,j}^N w_{ij} \mathbf{y}_i^T \mathbf{y}_j \\ &= 2\text{trace}(\mathbf{Y}\mathbf{D}\mathbf{Y}^T) - 2\text{trace}(\mathbf{Y}\mathbf{W}\mathbf{Y}^T) = 2\text{trace}(\mathbf{Y}\mathbf{L}\mathbf{Y}^T), \end{aligned} \quad (5.13)$$

where $L = D - W$ is called the Laplacian matrix that is a symmetric, positive semidefinite matrix. The optimal Y that minimizes the objective function subject to the constraints $YDY^T = I$ and $YDI = \mathbf{0}$ is given by the 2nd to $(d+1)$ th minimum eigenvalue solution to the generalized eigenvalue problem (5.13), where $I = [1, 1, \dots, 1]^T$. The constraint $YDY^T = I$ removes an arbitrary scaling factor in the embedding, and $YDI = \mathbf{0}$ avoids the constant embedding, which can be interpreted as removing a translation invariance in Y .

The standard LE algorithm is summarized in Table 3.1. An alternative approach to construct the adjacency graph is by ϵ -neighborhoods, where nodes i and j are defined as connected if $\|\mathbf{x}_i - \mathbf{x}_j\|^2 < \epsilon$ via the usual Euclidean norm. Since it is difficult to choose a proper parameter ϵ and easily leads to disconnected graph (Belkin & Niyogi 2001), we adopt the k -NN in the algorithm. It is known that based on spectral graph theory (Chung & Graham 1997), the embedding obtained by the LE algorithm optimally preserves local information. Moreover, Belkin and Niyogi (2003) pointed out that the Laplacian matrix L is analogous to the Laplace Beltrami operator on compact Riemannian manifolds, and the choice of the Gaussian weighted matrix (5.14) is related to the Green's function for the standard heat equation. Thus, the LE algorithm capable of discovering the optimal low-dimensional representation of nonlinear manifold structure has been justified.

Table 5.1: Laplacian Eigenmaps Algorithm (Belkin & Niyogi 2001)

Input: A set of points $\{\mathbf{x}_j\}_{j=1}^N$ in a manifold \mathcal{M} , integers k and d .

Output: A set of points (embedding) $\{\mathbf{y}_j\}_{j=1}^N$ lying in \mathbb{R}^d .

Constructing the adjacency graph by k nearest neighbors (k -NN): For each data point \mathbf{x}_j , search the k nearest neighbors, which are clustered into $\{\mathbf{x}_i\}$, $i \in \mathcal{N}_k(\mathbf{x}_j)$, according to some distance measurement $dist: \mathcal{M} \rightarrow \mathbb{R}$.

Choosing the weights: Define a matrix of weight $\mathbf{W} \in \mathbb{R}^{N \times N}$ whose entries w_{ij} measure the affinity between two connected points \mathbf{x}_i and \mathbf{x}_j . For example, w_{ij} can be defined as a heat kernel (with a given parameter $\sigma > 0$):

$$w_{ij} = \begin{cases} \exp\left(-\frac{dist(\mathbf{x}_i, \mathbf{x}_j)^2}{2\sigma^2}\right), & \text{if } j \in \mathcal{N}_k(\mathbf{x}_i) \text{ or } i \in \mathcal{N}_k(\mathbf{x}_j), \\ 0 & , \text{ else.} \end{cases} \quad (5.14)$$

Eigenmaps: Find a matrix $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_N] \in \mathbb{R}^{d \times N}$ whose rows are the d generalized eigenvectors of the pair (\mathbf{L}, \mathbf{D}) associated with its 2nd to $(d+1)$ th smallest generalized eigenvalues. That is, solve for \mathbf{Y} from:

$$\mathbf{Y}\mathbf{L} = \mathbf{A}\mathbf{Y}\mathbf{D}, \quad (5.15)$$

where \mathbf{A} is a diagonal matrix with the generalized eigenvalues.

5.3.3 Graph embedding and linearization approach

One of the major limitations of most NLDR methods is that they do not generally provide a functional mapping between the high and low dimensional spaces, which are applicable both on the training data and the novel test data. There are some approaches that attempt to address this issue by either the linearization techniques (Cai et al. 2007; He & Niyogi 2004; Min, Lu & He 2004; Park et al. 2004; Xiaofei He et al. 2005) or providing an out-of-sample extension (Bengio et al. 2003; Strange & Zwigelaar 2010) of the NLDR methods.

Graph embedding (Yan et al. 2005, 2007) is a general framework along with its linearization and kernelization, which theoretically unifies most previous algorithms. The essence of the graph embedding is to represent each vertex (each data points) of a graph by preserving the similarities of vertex pairs measured by the graph weight matrix, which characterizes desired statistic or geometry property of the data set. Graph embedding is developed based on LE starting with an undirected weighted graph $G = \{\mathbf{X}, \mathbf{W}\}$, where $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$ is a set of N vertices and \mathbf{W} is a similarity matrix. Recalling LE in (5.13), the optimal graph embedding that preserves the similarity of a graph is obtain by minimizing the *graph preserving criterion* as follows

$$\mathbf{Y}^* = \arg \min_{\mathbf{Y}\mathbf{B}\mathbf{Y}^T = \mathbf{C}} \sum_{i,j}^N w_{ij} \|\mathbf{y}_i - \mathbf{y}_j\|^2 = \arg \min_{\mathbf{Y}\mathbf{B}\mathbf{Y}^T = \mathbf{C}} \text{trace}(\mathbf{Y}\mathbf{L}\mathbf{Y}^T), \quad (5.16)$$

where \mathbf{B} and \mathbf{C} are the constraint matrices that impose certain conditions on the embedding \mathbf{Y}^* . Although the minimization problem provides the graph embedding for the data set, it cannot directly present the low dimensional representation for new testing data due to the lack of explicit mapping.

The linearization technique is introduced by assuming the low dimensional representation can be obtained from linear projections as $\mathbf{Y} = \mathbf{Z}^T \mathbf{X}$. Thus, the minimization problem is changed to

$$\mathbf{Z}^* = \arg \min_{\mathbf{Z}^T \mathbf{X}\mathbf{B}\mathbf{X}^T \mathbf{Z} = \mathbf{C}} \sum_{i,j}^N w_{ij} \|\mathbf{Z}^T \mathbf{x}_i - \mathbf{Z}^T \mathbf{x}_j\|^2 = \arg \min_{\mathbf{Z}^T \mathbf{X}\mathbf{B}\mathbf{X}^T \mathbf{Z} = \mathbf{C}} \text{trace}(\mathbf{Z}^T \mathbf{X}\mathbf{L}\mathbf{X}^T \mathbf{Z}). \quad (5.17)$$

If $\mathbf{X}\mathbf{B}\mathbf{X}^T = \mathbf{C} = \mathbf{I}$ is used, the constraint ensures \mathbf{Z} an orthonormal projection. To extend the linear projections to nonlinear case, one can utilize the kernel trick (Schölkopf, Smola & Müller

1998; Yan et al. 2005), which maps the data from the original input space to a higher dimensional Hilbert space and then the algorithm is performed in this new feature space. More details about kernelization can be found in (Ham et al. 2004; Yan et al. 2005). But the current study is limited to the standard linearization technique for simplicity.

It has been shown in (Yan et al. 2005) that many NLRD methods, such as ISOMAP, LLE, and LE, and the associated linear versions are just special cases of the graph-based formulation (5.16) and (5.17) with specific similarity and constraint matrices. Many other frameworks that unified different dimensionality reduction approaches have also been developed in statistics (machine) learning community, such as weighted or generalized PCA (Jolliffe 2002; Koren & Carmel 2004; Vidal, Ma & Sastry 2016), kernel methods (Ham et al. 2004), maximum entropy unfolding (Lawrence 2012), and patch alignment (Zhang et al. 2009).

Link to standard PCA/POD

PCA (refer to Section 3.3.1) can be written within the graph embedding framework as follows

$$\mathbf{Z}_{\text{PCA}}^* = \arg \max_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}} \text{trace}(\mathbf{Z}^T \mathbf{X} \mathbf{H} \mathbf{X}^T \mathbf{Z}), \quad (5.18)$$

where $\mathbf{H} = \mathbf{I} - \frac{1}{N} \mathbf{I} \mathbf{I}^T$ is a mean operator to ensure the sum of columns of $\bar{\mathbf{X}} = \mathbf{X} \mathbf{H}$ is a zero vector (note that $\mathbf{H} \mathbf{H} = \mathbf{H}$ and $\bar{\mathbf{X}} \mathbf{H} = \bar{\mathbf{X}}$). Thus, we can see that when the weight of all pair of points are chosen as a constant, i.e. $w_{ij} = 1/N$ for all $i \neq j$ (thus, $\mathbf{L} = \mathbf{H}$), and the constraint is defined as $\mathbf{X} \mathbf{B} \mathbf{X}^T = \mathbf{C} = \mathbf{I}$, the linearization of graph embedding in (5.17) can be used to find the standard PCA (or POD) projection in Section 3.3.1. It should be also noted that maximizing of

the objective is used in (5.18) because PCA/POD seeks for projection directions with maximal variance. Moreover, according to the formulation in (5.17) PCA can be expressed as

$$\mathbf{Z}_{\text{PCA}}^* = \arg \max_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}} \sum_{i \neq j} \frac{1}{N} \|\mathbf{Z}^T \mathbf{x}_i - \mathbf{Z}^T \mathbf{x}_j\|^2. \quad (5.19)$$

It indicates that the PCA projection is the one that maximally preserves the pairwise distances of data in the low-dimensional representation.

5.3.4 Linear graph embedding projection

The graph embedding framework along with its linearization allows to recover most popular NLRD and linear dimensionality reduction methods (Belhumeur, Hespanha & Kriegman 1997; Wright et al. 2010; Zhang et al. 2009) by properly adjusting the weight matrix and the graph locality. It has also been shown to achieve impressive performance superior to the standard PCA (or POD) in computer vision and pattern recognition applications (He et al. 2005; Yan et al. 2007) for non-Gaussian distributed data. Inspired by the graph embedding framework, we develop a linear graph embedding (LGE) projection as a more optimal candidate for nonlinear MOR as the priori knowledge of the data system can be utilized during the projection construction.

Let \mathbf{X} be a centered snapshot matrix of the quantities of interest (the column centering is not necessary to conduct beforehand because it could be achieved by using uniform weighting in LGE). Following the graph embedding framework in Section 5.3.3, the LGE projection $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_d] \in \mathbb{R}^{n \times d}$ that maps the high n -dimensional features to the low d -dimensional embedding is formally obtained by the following graph embedding framework

$$\mathbf{Z}_{\text{LGE}}^* = \arg \max_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}} \sum_{i,j}^N w_{ij} \left\| \mathbf{Z}^T \mathbf{x}_i - \mathbf{Z}^T \mathbf{x}_j \right\|^2 = \arg \max_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}} \text{trace}(\mathbf{Z}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{Z}), \quad (5.20)$$

where $\mathbf{L} = \mathbf{D} - \mathbf{W}$, and \mathbf{W} is a user-defined similarity matrix that better represents the nonlinear structure of input data with components defined as

$$w_{ij} = \begin{cases} w(\text{dist}(\mathbf{x}_i, \mathbf{x}_j)), & \text{if } j \in \mathcal{N}_{k_n}(\mathbf{x}_i) \text{ or } i \in \mathcal{N}_{k_n}(\mathbf{x}_j), \\ 0 & , \text{ else.} \end{cases} \quad (5.21)$$

As we can see, LGE is to find a projection to construct an affine subspace where the weighted pairwise distances of the low-dimensional data representation are maximally preserved. In LGE, there are three adjustable functions and parameters, which allow one to design a better projection for reduced order modeling. First, k_n denotes the number of neighbors to construct the adjacent graph, which controls the locality of the projection. Second, the distance function $\text{dist}(\mathbf{x}_i, \mathbf{x}_j)$ is usually defined as l^2 norm in Euclidean space, i.e. $\text{dist}(\mathbf{x}_i, \mathbf{x}_j) = \|\mathbf{x}_i - \mathbf{x}_j\|_2$. In many engineering applications, the data \mathbf{x}_i is extracted from a parametrized system with respect to certain parameters μ , i.e. $\mathbf{x}_i = \mathbf{x}(\mu_i)$. In this manner, $\text{dist}(\mu_i, \mu_j)$ can be used as the distance function to replace $\text{dist}(\mathbf{x}_i, \mathbf{x}_j)$. Lastly, $w: \mathbb{R} \rightarrow \mathbb{R}$ is a function of distance to assign proper affinity between two neighbor data points, which has significant impact on constructing a projection that preserves the localities and nonlinearities of the data structure. Some optional choices of weight function are

1. Inverse distance weighting: $w(\text{dist}(\mathbf{x}_i, \mathbf{x}_j)) = \frac{1}{\text{dist}(\mathbf{x}_i, \mathbf{x}_j)^p}, \quad (5.22)$

2. Gaussian weighting: $w(\text{dist}(\mathbf{x}_i, \mathbf{x}_j)) = \exp\left(-\frac{\text{dist}(\mathbf{x}_i, \mathbf{x}_j)^2}{2\sigma^2}\right), \quad (5.23)$

where σ can be defined as the maximum distance to normalize the pairwise distance.

Link to weighted POD method

Here, we show that the proposed LGE approach can be related to the weighted POD approaches recently used in reduced-order modeling (Bistrrian & Susan-Resiga 2016; Carlberg & Farhat 2011; Peng & Mohseni 2016) for better capturing local information with respect to parameter variations in the snapshot data. Essentially, the weighted POD approaches search for the best projection based on either goal-oriented least squares (Carlberg & Farhat 2011) or weighted snapshot matrix (Peng & Mohseni 2016). They all lead to the same weighted version of PCA (Zhang & Zha 2003): Given a data set $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbb{R}^{n \times N}$ and weights $\{\gamma_j\}_{j=1}^N$, we want to best fit the data set in a weighted sense using an affine mapping

$$\mathbf{x}_i \approx \bar{\mathbf{x}}_\gamma + \mathbf{Z}\mathbf{y}_i, \quad i = 1, \dots, N, \quad (5.24)$$

where $\mathbf{Z} \in \mathbb{R}^{n \times d}$ denotes the orthonormal basis of the affine subspace, $\bar{\mathbf{x}}_\gamma \in \mathbb{R}^n$ is the constant shift vector of the affine space, and \mathbf{y}_i represent the low-dimensional embedding coordinates of the points \mathbf{x}_i . To this end, the following weighted least squares problem is given

$$\min_{\bar{\mathbf{x}}_\gamma, \mathbf{Y}, \mathbf{Z}^T \mathbf{Z} = \mathbf{I}} \sum_{j=1}^N \gamma_j^2 \left\| \mathbf{x}_j - (\bar{\mathbf{x}}_\gamma + \mathbf{Z}\mathbf{y}_j) \right\|_2^2, \quad (5.25)$$

where $\{\gamma_j\}_{j=1}^N$ weight on the snapshot reconstruction. The optimal solution can be easily shown as

$$\bar{\mathbf{x}}_\gamma = \sum_i \gamma_i^2 \mathbf{x}_i / \sum_i \gamma_i^2, \quad \mathbf{Z} = [\mathbf{v}_1, \dots, \mathbf{v}_d], \quad \mathbf{y}_i = \mathbf{Z}^T (\mathbf{x}_i - \bar{\mathbf{x}}_\gamma), \quad (5.26)$$

where $\bar{\mathbf{x}}_\gamma$ is the weighted mean of snapshot data, and $\mathbf{v}_1, \dots, \mathbf{v}_d$ are the largest singular vectors of $(\mathbf{X} - \bar{\mathbf{x}}_\gamma \mathbf{1}^T) \mathbf{\Gamma}$ with $\mathbf{\Gamma} = \text{diag}([\gamma_1, \dots, \gamma_N])$. Let $\bar{\mathbf{X}}_\gamma = (\mathbf{X} - \bar{\mathbf{x}}_\gamma \mathbf{1}^T)$, the weighted PCA in (5.25) can be also formulated in the graph embedding framework by derivations as follows

$$\begin{aligned}
& \arg \min_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}} \sum_{j=1}^N \gamma_j^2 \left\| \mathbf{x}_j - (\bar{\mathbf{x}}_\gamma + \mathbf{Z} \mathbf{y}_j) \right\|_2^2 \\
&= \arg \min_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}} \left\| \bar{\mathbf{X}}_\gamma \mathbf{\Gamma} - \mathbf{Z} \mathbf{Z}^T \bar{\mathbf{X}}_\gamma \mathbf{\Gamma} \right\|_F^2 = \arg \min_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}} \left\| (\mathbf{I} - \mathbf{Z} \mathbf{Z}^T) \bar{\mathbf{X}}_\gamma \mathbf{\Gamma} \right\|_F^2 \\
&= \arg \max_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}} \text{trace}(\mathbf{Z}^T \bar{\mathbf{X}}_\gamma \mathbf{\Gamma}^2 \bar{\mathbf{X}}_\gamma^T \mathbf{Z}) = \arg \max_{\mathbf{Z}^T \mathbf{Z} = \mathbf{I}} \text{trace}(\mathbf{Z}^T \mathbf{X} \mathbf{H}_\gamma \mathbf{\Gamma}^2 \mathbf{H}_\gamma^T \mathbf{X}^T \mathbf{Z}),
\end{aligned} \tag{5.27}$$

where $\mathbf{H}_\gamma = \mathbf{I} - \gamma^2 \mathbf{I}^T / \mathbf{I}^T \gamma^2$ and $\gamma^2 = [\gamma_1^2, \dots, \gamma_N^2]^T$. Thus, the Laplacian matrix \mathbf{L} in (5.20) corresponding to the weighted PCA is defined as

$$\mathbf{L} = \mathbf{H}_\gamma \mathbf{\Gamma}^2 \mathbf{H}_\gamma^T = \mathbf{\Gamma}^2 - \gamma^2 (\gamma^2)^T / \mathbf{I}^T \gamma^2, \tag{5.28}$$

and the associated similarity matrix \mathbf{W} becomes

$$w_{ij} = \begin{cases} \gamma_i^2 \gamma_j^2 / \sum_{i=1}^N \gamma_i^2, & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases} \tag{5.29}$$

Conversely, if a symmetric matrix \mathbf{W} with a full graph (i.e., all components are non-zero except for the diagonal ones) are defined in (5.21), $\gamma_j = (\sum_i w_{ij})^{1/2}$ is defined to ensure an equivalent weighted PCA. Therefore, compared to weighted POD (Bistrián & Susan-Resiga 2016; Carlberg & Farhat 2011; Peng & Mohseni 2016), the proposed LGE seems to be a more generalization approach since it allows to control the local support of neighbor graphs \mathcal{N}_{k_n} .

LGE: a robust projection

As pointed out before, the POD learning methods based on least squares estimation are known for its sensitivity to “outliers”, yet outliers are common in realistic training sets. The term outlier refers to data that does not conform to the assumed statistical model. In some realistic applications, however, outliers are not necessary “noise” in a traditional sense but rather are violations of highly simplified mathematical models of the world, such as in computer vision (Torre & Black 2003). It motivates the development of LGE projection to improve the outlier robustness. It is also shown that outlier removal and noise reduction help to better preserve the topological structures of the nonlinear manifold in a neighborhood graph constructed from a finite set of sample points (Zhang & Zha 2003). Note that different from using weighted local learning techniques to remove outlier (Zhang & Zha 2003), this chapter focuses on developing a “robust” projection for reduced-order modeling that can tolerate some percentage of outlying data without having the solution severely skewed from the desired solution.

To illustrate the improvement of LGE over POD, a simple synthetic example is given in Figure 5.2, where a two-dimensional data set with $N = 50$ points, comprising 48 normally-distributed points and two outlying points. As can be seen in the figure, the first POD basis projects the data in a direction that emphasizes the outliers while ignoring almost the whole structure. This is because POD strives to maximize the sum of all squared distances and the pairwise distances involving outliers are significantly larger than the other pairwise distances, as suggested by (5.19). In contrast to the POD projection deceived by the outliers, the LGE projection is less sensitive to outliers while able to capture the global structure of the data, by underweighting distant data elements and using adjacent graphs with local support. Moreover, LGE yields lower reconstruction error measured in the *average relative error* defined as

$$\varepsilon_r = \frac{1}{N} \sum_{i=1}^N \frac{\|\mathbf{x}(\mu_i) - \tilde{\mathbf{x}}(\mu_i)\|}{\|\mathbf{x}(\mu_i)\|}, \quad (5.30)$$

where $\tilde{\mathbf{x}}$ denote the low-rank reconstruction of the original data \mathbf{x} using reduced-order basis.

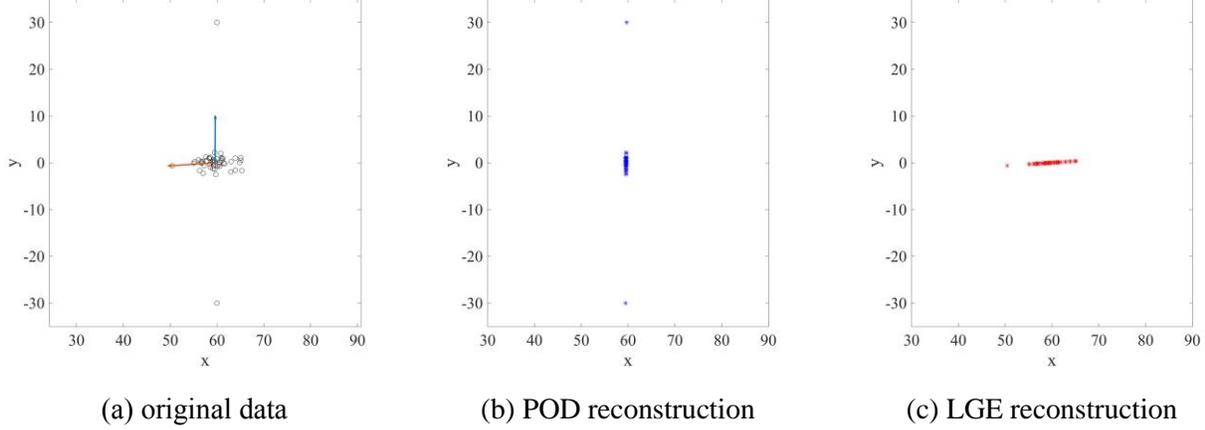


Figure 5.2: (a) Original dataset (contains two outliers) and the first principal component directions of POD (blue arrow) and LGE (red arrow). (b) The reconstruction solution by POD using one basis (the reconstruction error $\varepsilon_r = 3.57 \times 10^{-2}$). (c) The reconstruction solution by LGE using one basis, the inverse distance weighting with $p = 2$ and $k_n = 20$ (the reconstruction error $\varepsilon_r = 3.09 \times 10^{-2}$).

5.4 Robust Nonlinear Model Reduction Based on LGE

The application of the LGE projection for dimensionality reduction is the same as the standard POD projection. For example, the projection basis \mathbf{V} for subspace approximation of the state variables (see (5.4)) and \mathbf{Z} for the nonlinear vectors (see (5.8)) can be constructed based on LGE in (5.20) instead of POD in (5.19). It should be borne in mind that if the solution manifold exhibits nearly Gaussian distribution, POD yields an optimal low-rank approximation. Therefore, in the following numerical examples, the conventional POD basis is still chosen for the reduced-

order approximation of state solution because its manifold is not as severely nonlinear or localized as the nonlinear counterpart. But if the state data can be told by domain expertise to be non-Gaussian or with some outlying information, LGE basis is also suggested for state solution reduction.

We take the reduced-order modeling of elliptic PDEs in (5.9) to illustrate the procedures of using LGE for DEIM for nonlinear MOR, as shown in Table 5.2. The implementation for time dependent dynamic problem can be easily achieved by modifying the online stage to a time integration scheme. In the offline stage, the POD basis and the reduced linear operator are computed in step 1 and 2, respectively. The LGE basis of nonlinear snapshots used for the DEIM approximation is obtained in step 3. In the online stage, steps 5.1-5.3 form the main loop of the online computation using the reduced coordinates, and their computational complexity is independent of the underlying discretization. Here, we apply the LGE basis to the analysis of parameterized systems to compare its quality with that of the standard POD basis.

Table 5.2: Reduced order modeling of elliptic partial differential equations using linear graph embedding projection for DEIM

Input: Precomputed solution snapshots $\mathbf{X}_s \in \mathbb{R}^{N \times N_s}$, nonlinear snapshots $\mathbf{X}_f \in \mathbb{R}^{N \times N_f}$

Output: The approximation solution in the original coordinate system $\tilde{\mathbf{u}}(\boldsymbol{\mu}) = \bar{\mathbf{u}} + \mathbf{V}\mathbf{u}^r(\boldsymbol{\mu})$

Offline:

1. Compute the POD basis matrix $\mathbf{V} \in \mathbb{R}^{N \times k}$ for the solution snapshots matrix \mathbf{X}_s .
2. Use the Galerkin projection to compute $\mathbf{V}^T \mathbf{A} \mathbf{V}$ for the linear operator.
3. Create DEIM approximation for nonlinear operator
 - 3.1 Construct the similarity matrix $\mathbf{W} \in \mathbb{R}^{N_f \times N_f}$ by choosing the distance function $dist(\cdot, \cdot)$, weighting function w , and the neighbor number k_n , as (5.21).
 - 3.2 Compute the LGE basis matrix $\mathbf{Z} \in \mathbb{R}^{N \times \hat{k}}$ by solving the generalized eigenvalue decomposition problem $(\mathbf{X}_f \mathbf{L} \mathbf{X}_f^T) \mathbf{Z} = \mathbf{Z} \mathbf{A}$ induced from (5.20).
 - 3.3 Use the Greedy algorithm based on \mathbf{Z} to obtain the DEIM indices \mathcal{P} .
 - 3.4 Compute $\mathbf{M} = \mathbf{V}^T \mathbf{Z} (\mathbf{P}^T \mathbf{Z})^\dagger \in \mathbb{R}^{k \times \hat{k}}$ for the nonlinear term and Jacobian.

Online:

4. Given \mathbf{u}_0^r as the starting point in the reduced coordinate system for a new input $\boldsymbol{\mu}$.
5. **for** $j = 1, \dots$, (until convergence in certain criterion) **do**
 - 5.1 By only evaluating the selected component related to \mathcal{P} , compute the DEIM approximation of the reduced residual vector and the reduced Jacobian,

$$\begin{aligned} \hat{\mathbf{R}}(\boldsymbol{\mu}, \mathbf{u}_{(j)}^r) &= \mathbf{V}^T \mathbf{A} \mathbf{V} \mathbf{u}_{(j)}^r + \mathbf{M} \mathbf{f}_{\mathcal{P}}(\boldsymbol{\mu}, \mathbf{V} \mathbf{u}_{(j)}^r), \\ \hat{\mathbf{J}}(\boldsymbol{\mu}, \mathbf{u}_{(j)}^r) &= \mathbf{V}^T \mathbf{A} \mathbf{V} + \mathbf{M} \mathbf{P}^T \mathbf{J}_f(\boldsymbol{\mu}, \mathbf{V} \mathbf{u}_{(j)}^r) \mathbf{V}. \end{aligned} \quad (5.31)$$

5.2 Solve $\hat{\mathbf{J}}(\boldsymbol{\mu}, \mathbf{u}_{(j)}^r) \delta \mathbf{u}^r = -\hat{\mathbf{R}}(\boldsymbol{\mu}, \mathbf{u}_{(j)}^r)$.

5.3 Update $\mathbf{u}_{(j+1)}^r = \mathbf{u}_{(j)}^r + \delta \mathbf{u}^r$.

end

5.4.1 Benchmark problem: 2D Parametrized function

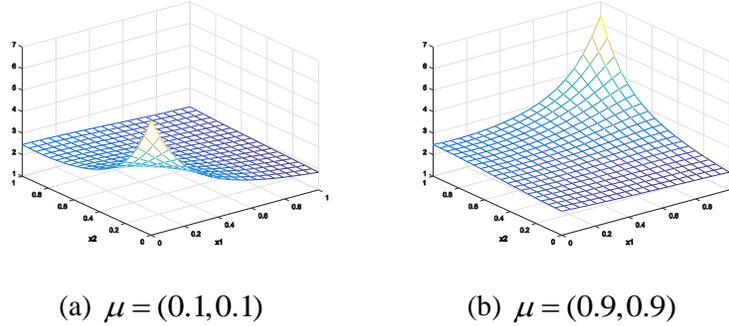


Figure 5.3: Function g^2 has a sharp peak in one of the two corners of the spatial domain.

If the system exhibits a wide range variation of behaviors, many DEIM basis vectors and interpolation points are required to accurately approximate the nonlinear term. This is because the global subspace spanned by POD basis tries to preserve the global structure of the snapshot data at the expense of losing the local varieties over a range of bifurcation parameters. On the other hand, POD is not robust in the sense that outlying measurement can significantly skew the solution from the desired solution (Ruppert 1987; Torre & Black 2003), as illustrated in Figure 5.2. The lack of robustness may lead to a POD projection basis that over-emphasizes certain “abnormal” data and loses the representation of more general solution behavior.

We demonstrate these issues on an example modified from the 2D nonlinear parametrized function g^1 in Section 3.4.2 as follows

$$g^2(\mathbf{x}; \mu) = g^1(x_1, x_2; \mu_1, \mu_2) + g^1(1 - x_1, 1 - x_2; 1 - \mu_1, 1 - \mu_2). \quad (5.32)$$

The same parameter setting in Section 3.4.2 is used here, i.e. g^2 are discretized by a 20×20 ($\mathcal{N} = 400$) equidistant grid in Ω and sampled on a 25×25 ($N_s = 625$) equidistant grid in \mathcal{D} .

Depending on the parameter μ , the function g^2 may trigger a high-gradient localization in one of the two corners of Ω near to $\mathbf{x} = [0,0]$ or $\mathbf{x} = [1,1]$, as shown by the examples in Figure 5.4.

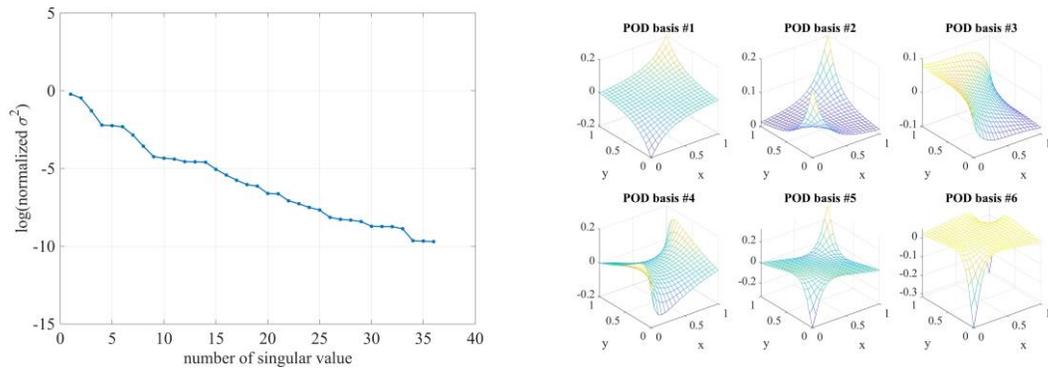


Figure 5.4: The normalized singular value $\sigma_i^2 / (\sum_i \sigma_i^2)$ (left) and the first six corresponding POD basis vectors (right) of the of the snapshot matrix \mathbf{X}_s corresponding to the nonlinear function g^2 .

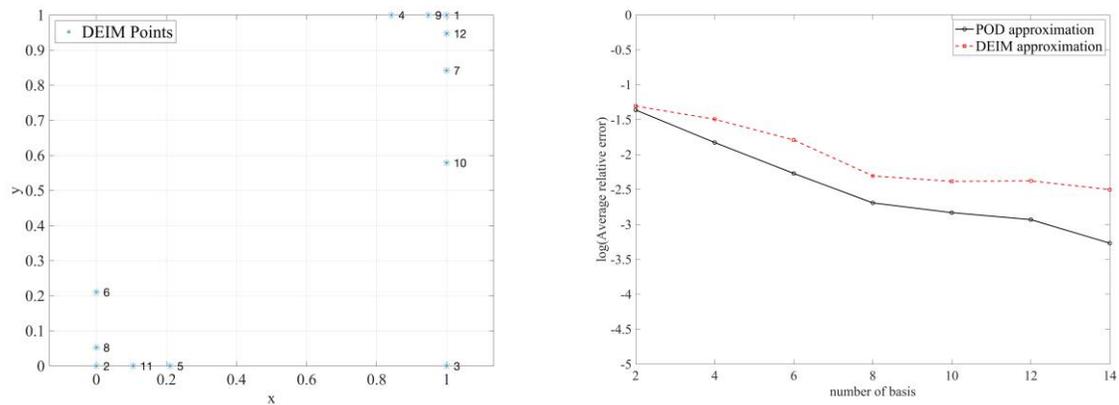


Figure 5.5: The distribution of first 12 DEIM points (left) selected by the greedy algorithm for the nonlinear function g^2 and the comparison of average relative errors of POD and DEIM approximations (right) for the training data.

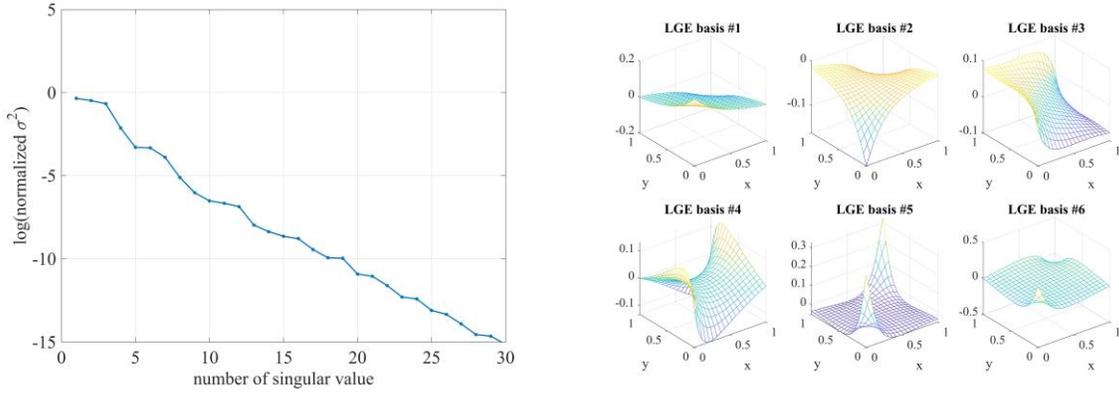


Figure 5.6: The normalized singular value $\sigma_i^2 / (\sum_i \sigma_i^2)$ (left) and the first six corresponding LGE basis vectors (right) of the of the matrix $\mathbf{X}_s \mathbf{L} \mathbf{X}_s^T$ corresponding to the nonlinear function g^2 , where the quadratic inverse distance weighting and $k_n = 40$ are adopted.

Although g^2 is simply combination of g^1 functions with variable transformations, we observe a much slower decay of the singular values of the snapshot matrix corresponding to g^2 , as shown in Figure 5.4, implying the overall snapshot behavior across the sampled parameters is difficult to capture by the POD basis. Similar to the previous example in Section 3.4.2, the DEIM points learned from training data of g^2 function are majorly positioned near the two corners where a sharp peak may occur (see Figure 5.5). However, this nature may cause the lack of robustness as the DEIM algorithm overly favor the data exhibits extreme localizations such that it simply places the sensing points locally at the sensitive region and leave the remaining large are undetected and ignored. Unless sufficiently large size of basis and interpolation points are used, this is dangerous for reduced-order modeling of real engineering applications since new test parameters are usually not a priori known and they may trigger critical mechanical response in the area unseen by the DEIM-points.

Now we examine the quality of the proposed LGE projection. The corresponding normalized singular values and LGE basis vectors are plot in Figure 5.6, where we can observe the singular values of LGE decay much faster than POD in Figure 5.4 due to the weighting operator in $\mathbf{X}_s \mathbf{L} \mathbf{X}_s^T$ (refer to (5.20)). To measure the approximation performance by DEIM based on different projection basis (POD or LGE), the average relative error defined in (7.19) is also used. To distinguish from the standard DEIM based on POD basis (see Section 5.2.3), the one based on the LGE projection to identify interpolation points and perform DEIM approximation for nonlinear functions is named as Graph-DEIM or G-DEIM.

Table 5.3: The average relative error of the reconstruction of training data by the LGE approximation and the corresponding DEIM (G-DEIM) approximation using 12 projection basis under different weight functions and neighbor numbers k_n . (The reconstruction errors of the standard POD and DEIM approximation are also given as reference)

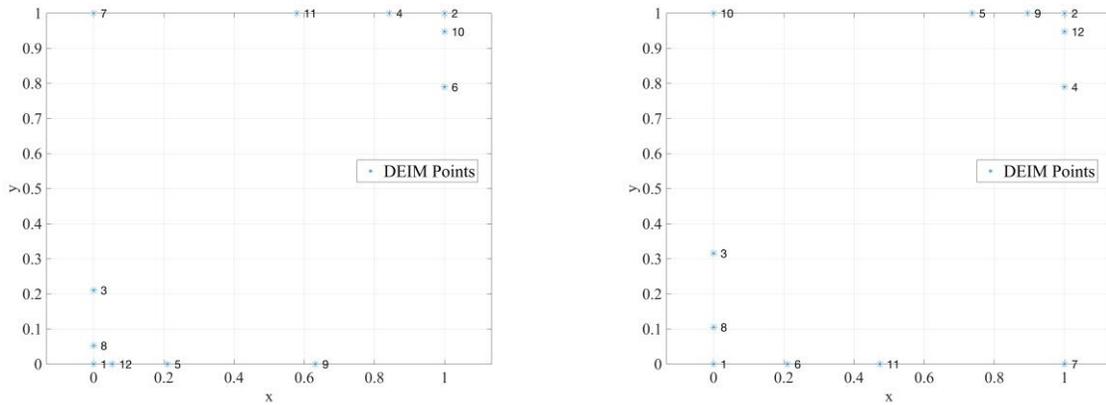
k_n	LGE approximation			G-DEIM approximation		
	20	40	80	20	40	80
Constant, $w_{ij} = 1$	2.31×10^{-3}	1.92×10^{-3}	1.43×10^{-3}	6.83×10^{-3}	4.71×10^{-3}	3.68×10^{-3}
Inverse distance, $p = 1$	1.48×10^{-3}	1.09×10^{-3}	0.81×10^{-3}	5.03×10^{-3}	4.22×10^{-3}	2.45×10^{-3}
Inverse distance, $p = 2$	0.83×10^{-3}	0.71×10^{-3}	0.63×10^{-3}	2.63×10^{-3}	2.45×10^{-3}	1.79×10^{-3}
Reference	POD: 1.17×10^{-3}			DEIM: 4.20×10^{-3}		

A parameter study of LGE approach is present in Table 5.3, where LGE and the LGE-based DEIM (G-DEIM) are used to approximate the training data (or called reconstruction) using $\hat{k} = 12$ projection basis (also 12 DEIM points). It can be seen from Table 5.3 that as more neighbor points, e.g., from $k_n = 20$ to $k_n = 80$, are used to construct the LGE projection, higher reconstruction accuracy can be obtained by both directly projection approximation, i.e.

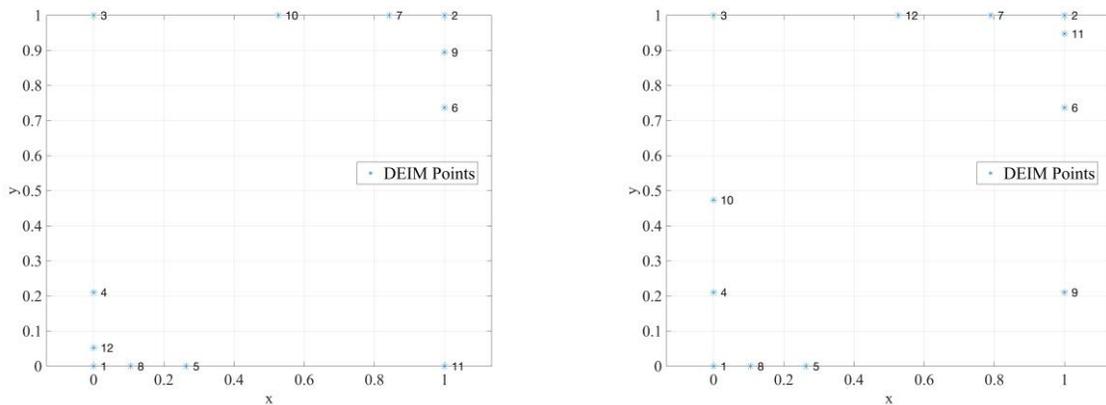
$\tilde{\mathbf{X}}_s = \mathbf{Z}\mathbf{Z}^T \mathbf{X}_s$ and G-DEIM approximation, $\tilde{\mathbf{X}}_s = \mathbf{Z}(\mathbf{P}^T \mathbf{Z})^{-1} \hat{\mathbf{Z}}^{-1} \mathbf{P}^T \mathbf{X}_s$. This is can be explained by the fact that more information of data used for learning helps to better capture the global structure. Moreover, we observe that the approximation by LGE with quadratically inverse weight function (i.e., $w_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|^{-2}$) yields better results than the cases using constant weight function and linearly inverse weight function (i.e., $w_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|^{-1}$), which indicates stronger weighting is effective to subside the outlying effect and better preserve the global structure. The distribution of DEIM points selected by the greedy algorithm corresponding to some LGE projections used in Table 5.3 are exemplified in Figure 5.7. It shows that with larger k_n and more heavily penalized weight function, the employment of LGE projection allows DEIM to place interpolation points more evenly over the whole domain compared to the Figure 5.5, and thus, it enhances the robustness. Essentially, the robustness is restored in LGE by underweighting the data points that are significantly different from others (outliers), i.e. faraway from others in a geometric sense.

Lastly, we compare the average relative errors of the approximation obtained by POD and LGE, as well as their associated DEIM approximation, for new test dataset shown in Figure 5.8 (left). This data points are chosen particularly outside the critical parameter region that leads to sharp peaks. The comparison in Figure 5.8 shows that the proposed LGE projection with quadratically inverse weight function and its G-DEIM yield better results than the standard POD projection and DEIM, respectively. It should be noted that the performance of DEIM approximation is highly correlated with the quality of projection basis. That is, when a better projection basis is designed by LGE, the associated DEIM approximation also achieves better representation of the data structure. Thus, through introducing local data information by weight

function, the LGE projection is expected to be more robust and able to prevent DEIM points from overly localizing due to sensitivity to “abnormal” data, leading to better representation of overall solution behavior.



(a) G-DEIM points with inverse distance $p = 1$ and $k_n = 20$ (left) and $k_n = 80$ (right)



(b) G-DEIM points with inverse distance $p = 2$ and $k_n = 20$ (left) and $k_n = 80$ (right)

Figure 5.7: The distribution of first 12 DEIM points selected by the greedy algorithm for the nonlinear function g^2 based on the LGE projection using inverse distances (a) $p = 1$ and (b) $p = 2$ as the weight function.

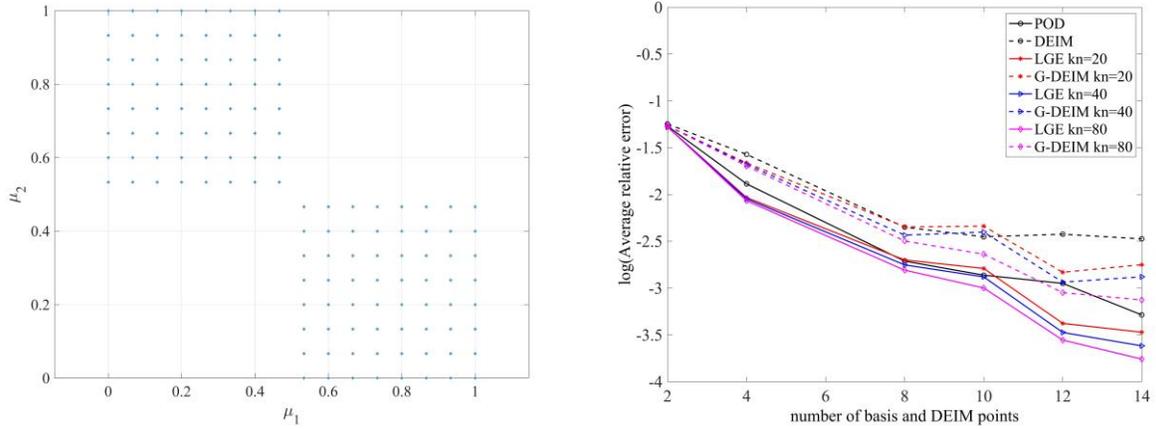


Figure 5.8: The new test data obtained by evaluating the nonlinear function g^2 at the (128 points) parameter distribution (left) is used to compare the performance (right) of POD and LGE projections, as well as the corresponding DEIM approximation with different dimensions of basis.

5.4.2 Parameterized elliptic PDE

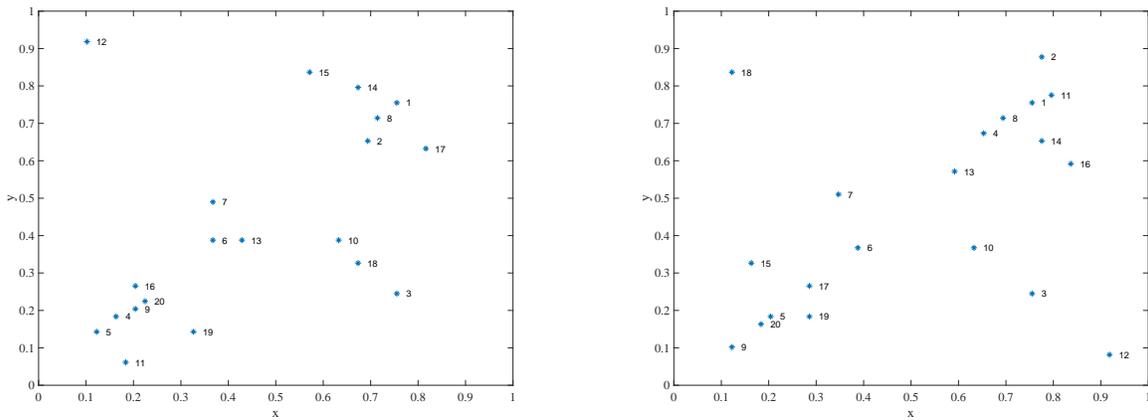
The DEIM based on LGE (G-DEIM) is used for reduced-order modeling of an elliptic PDE problem (Chaturantabut & Sorensen 2010; Grepl et al. 2007) given by

$$-\nabla^2 u(x, y) + \frac{\mu_1}{\mu_2} (e^{\mu_2 u} - 1) = 100 \cos(2\pi x) \cos(2\pi y), \quad \forall (x, y) \in \Omega, \quad (5.33)$$

where the spatial domain $\Omega = (0,1)^2$, the parameters satisfy $\mu = (\mu_1, \mu_2) \in \mathcal{D} = [0.01, 10]^2$, with homogeneous Dirichlet boundary conditions, i.e., $u(0, y) = u(1, y) = u(x, 0) = u(x, 1) = 0$. The reference full-scale solution is solved by applying Newton's method with a finite difference discretization. The spatial grid points (x_i, y_j) are equally spaced in Ω for $i, j = 1, \dots, 51$, and the full dimension of state variable is then $\mathcal{N} = 2601$. Let $f = \mu_1 / \mu_2 (e^{\mu_2 u} - 1)$ be the nonlinear function to be collected. With using 12×12 uniformly selected parameters from \mathcal{D} , the snapshot

matrices for state solution $u(\mu)$ and $f(\mu, u(\mu))$ can be obtained, and denoted as $\mathbf{X}_s \in \mathbb{R}^{N \times N_s}$ and $\mathbf{X}_f \in \mathbb{R}^{N \times N_f}$ ($N_f = N_s$), respectively, where $N_s = 144$.

As shown in Table 5.2, the POD-DEIM approach is used for reduced-order modeling of the system in (5.33), where the state variable is approximated by POD while the nonlinear function is approximated by either POD or LGE for comparison. That is, in the online stage, the DEIM approximation based on the associated POR or LGE basis for nonlinear snapshots is used to construct reduced systems. The selection of first $\hat{k} = 20$ DEIM points based on POD and LGE basis vectors are shown in Figure 5.9. The average relative errors of the associated DEIM reconstruction for the snapshot matrix \mathbf{X}_f are 5.45×10^{-4} and 1.98×10^{-4} , respectively, indicating the better performance of LGE in approximating the training data.



(a) DEIM (based on POD basis)

(b) G-DEIM (based on LGE basis)

Figure 5.9: The distribution of first 20 DEIM points selected by the greedy algorithm based on (a) POD basis and (b) LGE basis using linearly inverse distances and $k_n = 20$ as the weight function.

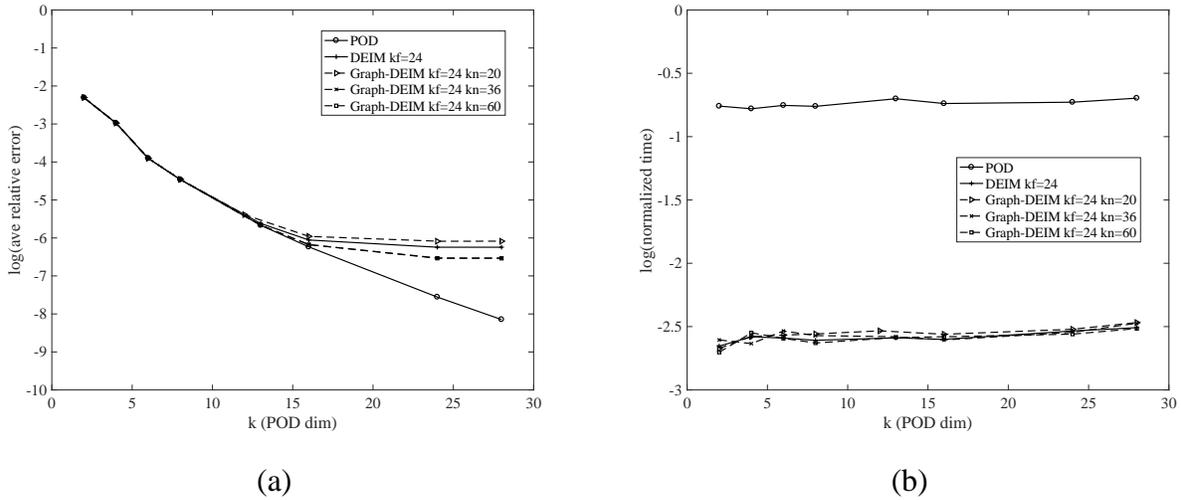


Figure 5.10: (a) The average relative error of the reduced-order approximation solved by different reduced systems for the elliptic PDE. (b) The average CPU time of each reduced Newton iteration for different reduced systems, which are normalized by the average CPU time for each Newton iteration in the full-scale model.

Based on 225 uniformly selected parameters that were not used to obtain the sample snapshots, Figure 5.10(a) plots the relative errors of the reduced-order approximation solved by different reduced systems for the elliptic PDE in (5.33), where the subspace dimension of nonlinear term is fixed as $\hat{k} = 24$ to provide a good approximation. Several observations can be gleaned from this figure. First, when the reduced dimension of nonlinear term is sufficiently large compared to that of state variable (e.g. $k \leq 10$), all DEIM approximations achieve nearly same accuracy as the POD-Galerkin system. Second, although all DEIM approximations obtain the same accuracy, it seems enlarging neighbor graph of LGE improves the performance of the associated DEIM reduced-order modeling. But it should be noted that the goal of the development of LGE is to enhance robustness against outliers or other unexplored test data rather than simply improving accuracy. Moreover, it provides a more general framework where one can introduce a priori knowledge of data for projection construction.

Figure 5.10(b) provides the normalized running times of each reduced Newton iteration for different reduced systems. By using the hyper-reduction by DEIM, the reduced-order systems achieve speedups of more than 100 times. Since reduced-order methods may affect convergence properties, more iterations are usually required. But overall the computational effort is greatly reduced. Note that the proposed LGE projection has the same computational complexity as POD in the online stage.

5.4.3 One-dimensional Burgers' problem

Consider the following one-dimensional Burgers' problem that describes the movement of a shock

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial F(u)}{\partial x} &= s(x), \quad x \in [0,100], \quad t \in [0,50], \\ u(x,0) &= 1, \\ u(0,t) &= \sqrt{5}, \end{aligned} \tag{5.34}$$

where $u(t,x)$ is the unknown representing a conserved quantity, the flux function is defined as $F(u) = 0.5u^2$, and the source term is $s(x) = 0.02e^{0.02x}$. The HDM reference solution to (5.34) is solved by the Godunov-type finite volume method (LeVeque 2002) with a spatial discretization of size $\mathcal{N} = 201$ and a time step $\Delta t = 0.1s$ for time integration. For constructing DEIM approximation, the gradient flux, i.e. $f(x,t) = \partial F(u)/\partial x$, is considered as the nonlinear function to be sampled for nonlinear snapshot matrix. Figure 5.11 plots six snapshots of the state solution and the nonlinear function, respective, solved by the HFM model at six different time instances.

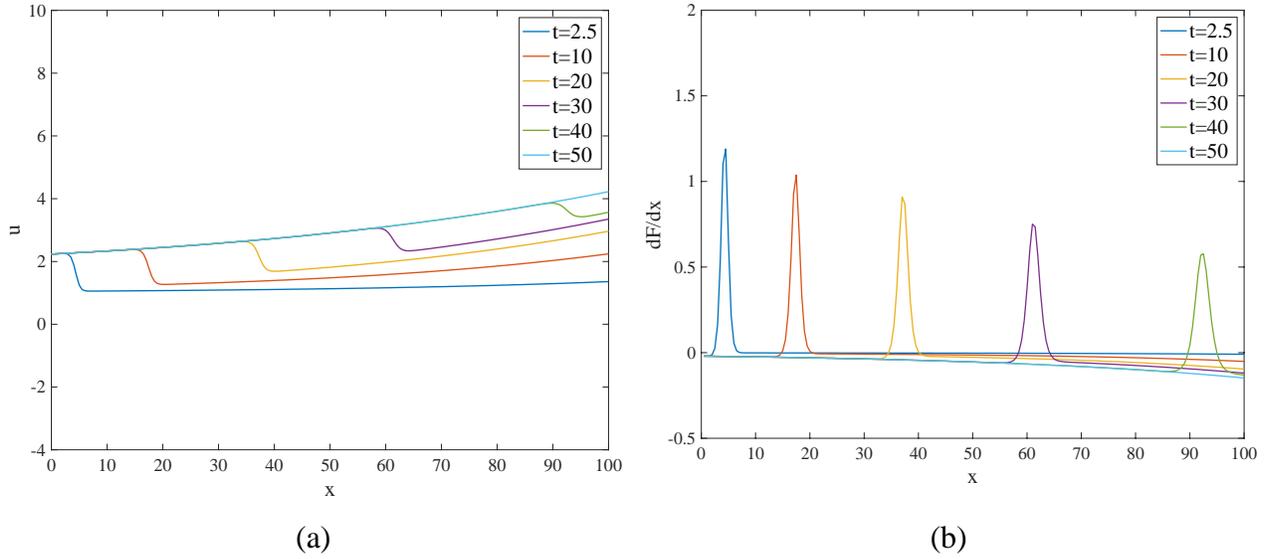
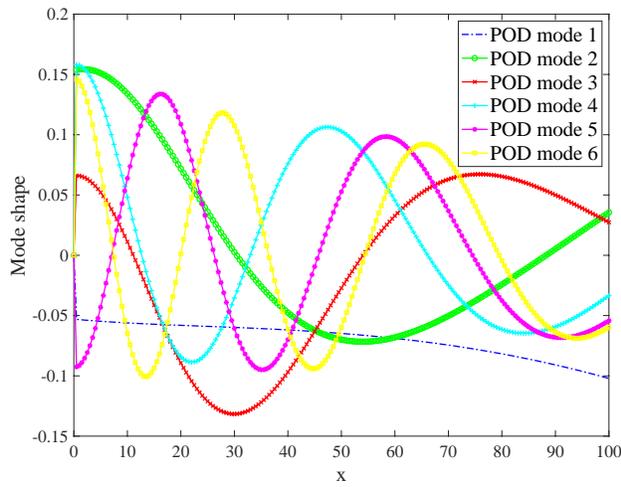
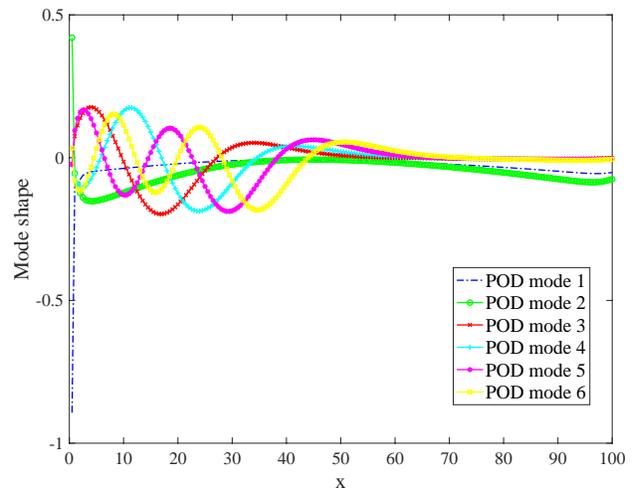


Figure 5.11: Snapshot of (a) the state solution $u(x,t)$ and (b) the nonlinear function $f(x,t)$ of the Burgers' problem obtained by a high-dimensional finite volume model of size $\mathcal{N} = 201$.

In offline stage, the HFM solution at every 5 time-steps is collected, and thus, it results in a total number of $N_s = 100$ snapshots. The first six POD modes for the state snapshot matrix \mathbf{X}_s and the nonlinear snapshot matrix \mathbf{X}_f are given in Figure 5.12. Note that the 1st POD model is representing a sharp peak at the boundary as it is an outlying behavior. Compared to the POD modes for state snapshots (Figure 5.12(a)), the POD modes for nonlinear snapshots (Figure 5.12(b)) are more localized at the side where the shock starts to propagate, which represents the bias of the POD approach in capturing the overall behavior of the nonlinear function in a large time range as it travels as a singular function, as shown in Figure 5.11(b). This could be an issue for the associated reduced-order modeling when the shock wave marches into the region that the POD modes are unable to represent, leading to erroneous results and even unstable responses.



(a) POD of state snapshot matrix



(b) POD of nonlinear snapshot matrix

Figure 5.12: First six POD basis modes of (a) the state snapshot matrix \mathbf{X}_s and (b) the nonlinear snapshot matrix \mathbf{X}_f .

Instead of POD, the LGE method in (5.20) is adopted to construct the projection basis of the nonlinear snapshot matrix \mathbf{X}_f . Different weight functions and numbers of neighbor for LGE are tested, and the corresponding first six basis modes are presented in Figure 5.13. By weighting the pair-wise relation of snapshot data, the LGE construction generates the projection basis modes that spread widely across the spatial domain and prevent from over, e.g., the modes shown in Figure 5.13(b), compared to that of POD in Figure 5.12. We observe that the outlying effect at the boundary side is greatly suppressed and the resulting LGE modes are shifted to the distant end as the power of inverse weight function (p) or the graph neighbors (k_n) is increased to exert stronger weighting, e.g., the LGE modes in Figure 5.13(a) and in Figure 5.13(f) cover distinct spatial regions.

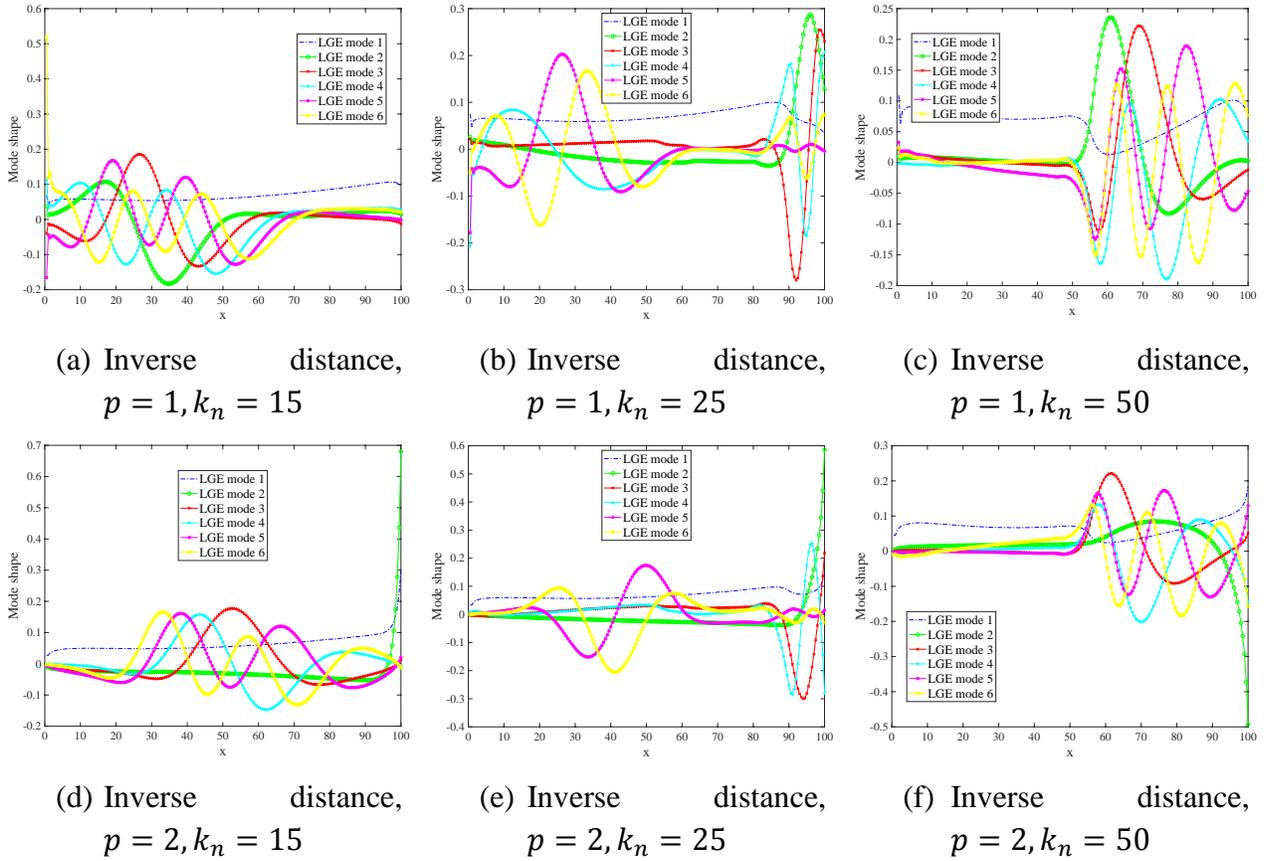


Figure 5.13: First six LGE basis modes of the nonlinear snapshot matrix \mathbf{X}_f under different weight functions and numbers of neighbor points.

To demonstrate the approximation performance of LGE for the nonlinear function, Table 5.4 and Table 5.5 provide the average relative error (see (7.19)) between the snapshot data \mathbf{X}_f and the reconstruction data $\tilde{\mathbf{X}}_f$ approximated by the direct LGE projection and the corresponding DEIM approximation when using the subspace dimension of $\hat{k} = 20$ and $\hat{k} = 30$, respectively. The reconstruction errors by the standard POD and DEIM approximations are also given as reference. In these cases, the number of interpolation points is selected as $\hat{n} = \hat{k} + 10$ to ensure a good DEIM approximation. It shows that higher reconstruction accuracy is obtained by using LGE over POD for both direct projection and DEIM approximations. In contrast to the

results of 2D benchmark problem in Table 5.3, the LGE projection with more neighbor points leads to worse reconstruction results for Burgers' problem. This could be due to the local characteristic of the nonlinear snapshots, as shown in Figure 5.11(b), and thus, a more local graph helps to preserve the global data structure. This example shows the importance of introducing the a priori knowledge of data to construct a better projection basis, which is the flexibility offered by the LGE approach.

For reduced-order modeling of the initial-boundary-value problem (5.34) in the online stage, the POD-DEIM reduced-order model in (5.10) is employed. We emphasize again that since the state solution can be well represented by POD basis (see Figure 5.12(a)), the state variables follows the POD subspace approximation in (5.4) and the POD-Galerkin ROM is constructed based on (5.6). On the other hand, for reduced approximation of the nonlinear function, both POD basis and LGE basis are used for DEIM for comparison. The numerical tests show that the reconstruction error for nonlinear snapshots serves as an indicator of the performance of the constructed LGE projection basis for reduced-order modeling. The linearly inverse weight function (i.e., $p = 1$) and $k_n = 15$, therefore, are selected for the LGE projection basis. Using $k = 30$, $\hat{k} = 30$, and $\hat{n} = 40$, the snapshots of the reduced-order solution solved by DEIM (using POD basis) and G-DEIM (using LGE basis) are plot in Figure 5.14, where the approximation of POD-Galerkin is also given. Compared with the HFM solution in Figure 5.11, the POD-Galerkin certainly yields the best reduced-order solution because no hyper-reduction is performed, whereas it is computationally expensive. The nonlinear model reduction method based on LGE shows better results than the counterpart of POD. Especially, for the instance of $t = 50s$ (the green line), DEIM losses the accurate representation of the tail behavior, which is

also reflected by the approximation of the nonlinear function, while G-DEIM robustly capture this nonlinear behavior.

Table 5.4: The average relative error of the reconstruction of training data by different LGE approximations and the corresponding DEIM (G-DEIM) approximations using $\hat{k} = 20$ projection basis and $\hat{n} = 30$. (The reconstruction errors of the standard POD and DEIM approximation are also given as reference)

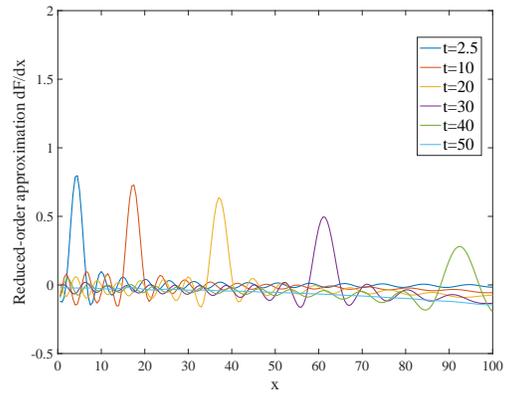
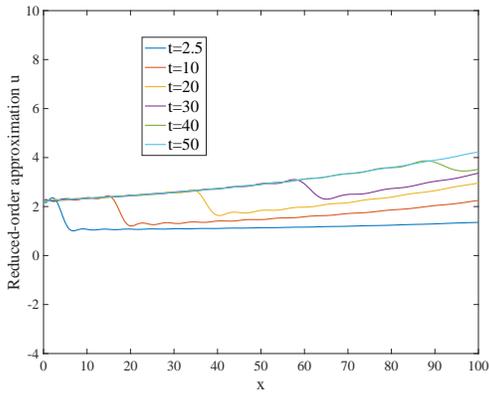
k_n	LGE approximation			G-DEIM approximation		
	15	25	50	15	25	50
Inverse distance, $p = 1$	5.20×10^{-1}	5.29×10^{-1}	5.41×10^{-1}	6.07×10^{-1}	6.13×10^{-1}	6.23×10^{-1}
Inverse distance, $p = 2$	5.18×10^{-1}	5.23×10^{-1}	5.33×10^{-1}	6.03×10^{-1}	6.02×10^{-1}	6.24×10^{-1}
Reference	POD: 5.48×10^{-1}			DEIM: 6.33×10^{-1}		

Table 5.5: The average relative error of the reconstruction of training data by different LGE approximations and the corresponding DEIM (G-DEIM) approximations using $\hat{k} = 30$ projection basis and $\hat{n} = 40$. (The reconstruction errors of the standard POD and DEIM approximation are also given as reference)

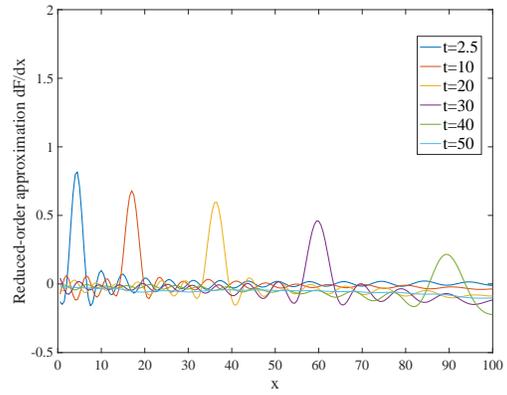
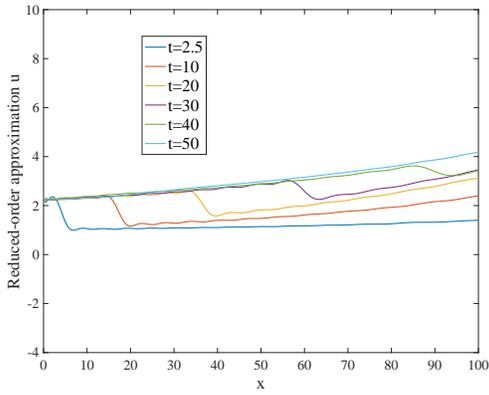
k_n	LGE approximation			G-DEIM approximation		
	15	25	50	15	25	50
Inverse distance, $p = 1$	3.73×10^{-1}	3.78×10^{-1}	3.87×10^{-1}	4.33×10^{-1}	4.51×10^{-1}	4.75×10^{-1}
Inverse distance, $p = 2$	3.74×10^{-1}	3.76×10^{-1}	3.77×10^{-1}	4.52×10^{-1}	4.57×10^{-1}	4.46×10^{-1}
Reference	POD: 4.01×10^{-1}			DEIM: 4.77×10^{-1}		

Note that the subspace approximation of the state solution is critical to reduced-order modeling since the reduced system is solved on the low-dimension subspace constructed based on its projection basis. It controls the quality of representation of the reduced solution. LGE can be used for approximating the state solution as well. However, numerical tests show that with less than $k = 20$ basis, poor overall results are given. This issue encourages to combine local

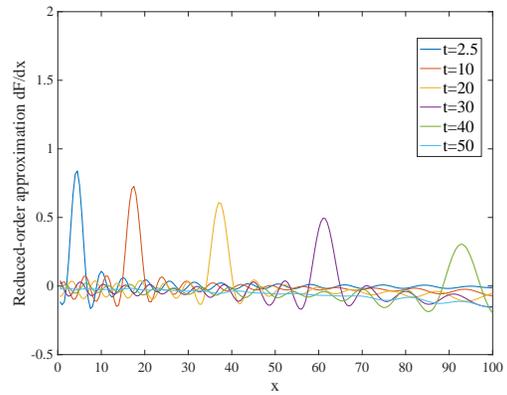
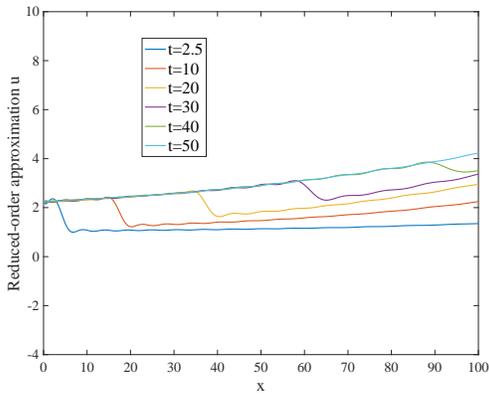
approaches (Amsallem, Zahr & Farhat 2012; Peherstorfer et al. 2014; Peng & Mohseni 2016) for further improvement.



(a) POD-Galerkin: Reduced approximation of u (left) and $f = \partial F(u)/\partial x$ (right)



(b) DEIM: Reduced approximation of u (left) and $f = \partial F(u)/\partial x$ (right)



(c) G-DEIM: Reduced approximation of u (left) and $f = \partial F(u)/\partial x$ (right)

Figure 5.14: Snapshot of the reduced-order approximation of the state solution $u(x,t)$ and the nonlinear function $f(x,t)$ of the Burgers' problem obtained by (a) POD-Galerkin, (b) DEIM, and (c) G-DEIM reduced-order systems.

5.5 Summary

In this chapter, we have developed a new technique, called linear graph embedding (LGE), to construct reduced projection for reduction of parameterized PDEs, based on the graph embedding that is a general framework to unify most manifold learning methods. The unique feature of this framework is that it allows to introduce *a priori* information of data for constructing a reliable projection through reweighting the relationship between pair-wise data points. It has also been shown that the standard POD and the weighted POD are special cases under this framework. In contrast to POD that is easily misled by outlying data and ignore the whole data structure, LGE can extract essential information from irregular data, and therefore, yield more robust reduced order model (insensitive to outliers) without sacrificing much accuracy.

Thus, LGE is very suited for model reduction of systems that exhibit a large range of variations in terms of parameters changes. In this chapter, the LGE approach has been tested for the elliptic PDE and hyperbolic PDE, and relatively better performance in capturing the global behavior of nonlinear solution has been demonstrated. However, thorough study on LGE has not been done yet and a real-world example is required to confirm the advantage of LGE. Moreover, it is interesting to study the automated strategy to select a proper weight function for the problem of interest. These are important topics to be further investigated in future study.

On the other hand, while LGE is expected to be less sensitive to outliers and better reveal the nonlinear data structure, it may still need a relatively high-dimensional subspace to represent all possible solution variations well if the original system shows a severe variability. This effect

is considerably increased when dealing with dynamical problems with significant solution variations in time (e.g., Burgers' problem). Local subspace methods (Amsallem, Zahr & Farhat 2012; Peherstorfer et al. 2014; Peng & Mohseni 2016; Sargsyan, Brunton & Kutz 2015) may be good candidate to address this issue. As the weighted POD is usually used together with local approach for reduced-order modeling of nonlinear system (Carlberg & Farhat 2011; Peng & Mohseni 2016), the proposed LGE projection can be readily implemented in the same computational structure with data clustering for local subspace construction as well as subspace selection (Amsallem, Zahr & Farhat 2012; Peherstorfer et al. 2014). The combination with local approaches will be presented in our future research.

5.6 Acknowledgement

This Chapter, in part, is currently being prepared for submission for publication of the material as it may appear in "He, Qizhi; Chen, Jiun-Shyan. Nonlinear model order reduction based on linearized manifold learning for meshfree analysis". The dissertation author was the primary investigator and author of this material.

Chapter 6

Data-Driven Computational Simulation Based on Locally Convex Reconstruction for Noisy Database

Although most simulation methods for materials under large deformation, such as bio-material, rely on representative constitutive models, the formulation of constitutive models as well as the calibration of material parameters for such complex biological material remain challenging. As the second objective of the dissertation, we develop a robust data-driven computational framework for scientific computing by directly linking experimental data to physics laws in order to perform numerical simulation.

In this Chapter, we present a new data-driven simulation approach embedded with manifold learning techniques, termed locally convex data-driven (LCDD) computing, for elastostatics problems, aiming to enhance robustness against noise and outliers in data sets that are especially limited and high-dimensional. In contrast to most existing data-driven methods based on either the direct distance measurement to a single data point or the calculation of a tangent matrix, LCDD searches for optimum data solutions from a locally reconstructed convex hull associated to the *k-nearest neighbor* (*k*-NN) points, which leads to less sensitivity to noisy data and ensures convergence stability. By using the penalty relaxation for convexity constraints, the acquisition of optimal data within the locally convex space is related to a classical *non-negative least squares* (NNLS) problem and thus, can be solved efficiently. In addition, LCDD performs well for high-dimensional data sets where data are scarce relative to the complexity of

the system, attributing to the inherent manifold learning of NNLS that admits a locally linear regression. Numerical tests in truss problems and 2D elasticity mechanics are given to validate the robustness, accuracy, and convergence properties of the proposed approach. The result shows that LCDD enhances nearly one order of accuracy compared to the standard distance-minimization data-driven scheme when dealing with noisy database, and it yields a linear exactness to the canonical numerical reference solution when the database is well sampled from the graph of constitutive relations.

6.1 Introduction

The present study focuses on the development of a new data science paradigm for computational mechanics, referred to as *data-driven computing* (Kirchdoerfer & Ortiz 2016), in which a given material database is used to replace the constitutive relations needed in the conventional computational mechanics whereas the conservation laws are still utilized to avoid the need of learning physical models from huge amount of data. In this manner, automated machine learning techniques on material data are carried out during the computation of the initial-boundary-value problem, and it bypasses the traditional construction of constitutive models. Thus, data-driven computing opens a door for new approaches in computational mechanics, especially for modelling those complex material behaviours that are extremely difficult to explicitly formalizing a constitutive law.

The pioneering work by Kirchdoerfer and Ortiz (Kirchdoerfer & Ortiz 2016) introduced the so called distance-minimizing data-driven computing (DMDD), where the constitutive law is relaxed by a minimization of distance between the computed material states (strain and stress

hereof) and the optimum data solution given by a material data set. A closely relevant study (Ibañez et al. 2016) proposed to applying manifold learning techniques (specifically, locally linear embedding (Roweis & Saul 2000)) to the database for constructing the local tangent approximation, with which the convergent solution could be attained by using directional search solvers such as Newton linearization technique. This idea has been recently extended to dynamics and nonlinear problems (González, Chinesta & Cueto 2018; Kirchdoerfer & Ortiz 2018; Nguyen & Keip 2018). Overall, the common idea of the abovementioned methods is to seek the intersection of the hidden constitutive (or material) manifold and the equilibrium (or physical) manifold by iterative processes considering appropriate search directions.

Despite the major advancement made in the field, one challenge is to be amenable to perturbations by the data inputs and enhanced the capability of handling uncertainties in input data (Darema & Rotea 2006). The standard DMDD paradigm (Kirchdoerfer & Ortiz 2016) is shown sensitive to noisy data and outliers (Ayensa-Jiménez et al. 2018; Kirchdoerfer & Ortiz 2017), while the approaches based on manifold learning (Ibañez et al. 2016) or local regression (Kanno 2017) may fail to converge due to the over-relaxed manifold construction. To enhance robustness, the DMDD approach was extended to the max-ent data driven computing (Kirchdoerfer & Ortiz 2017), which utilizes entropy estimation to consider the statistics information of data. But the resulting optimization problem becomes intractable and need to employ simulated annealing schedule. Alternatively, Ayensa-Jimenez et al. (Ayensa-Jiménez et al. 2018) modified the standard DMDD approach by explicitly incorporating statistical quantities into the cost function and defined a stochastic analogous problem. However, the issue related to noisy experimental data hasn't been fully addressed, especially when the available data is scarce relative to the complexity of the system.

In this work, we propose a new data-driven paradigm that integrates DMDD with a special local manifold learning technique, allowing to approximate the convex envelop of the cluster of data points to represent the underlying constitutive manifold. By means of locally convex reconstruction of the given material data set, the solution space for searching optimum data point is regularized into a bounded, continuous, and convex subset (polytope), which ensures robustness and convergence stability for data-driven computing. This proposed approach is referred to as locally convex data-driven computing (LCDD). In this approach, the data cluster for each local state (e.g., the pair of strain and stress) is first identified by the *k-nearest neighbor* (*k*-NN) algorithm. The optimum data solution to a given local state is searched within the associated locally convex hull instead of the discrete material set. To solve this local search problem efficiently, we recast it into a *non-negative least squares* (NNLS) problem by using a penalty relaxation method. Besides, because of the inherited manifold learning capacity in the NNLS solvers, the proposed LCDD permits the locally linear approximation for the underlying material manifold, which means that LCDD could reproduce the same state solutions as conventional computing if the material data represents a locally linear pattern. LCDD can also be viewed as a generalization of DMDD by equipping manifold learning techniques that naturally takes the probabilistic perspective into account and retains a simple computing framework.

The objective of the present work is to study the main issues of data-driven approaches when dealing with noisy data in high-dimensional space. The paper is organized as follows. In Section 6.2, a generalized data-driven computational formalism is laid out. In this section, a close relation to the material parameter identification problem is outlined to provide context for the data-driven methodology. In Section 6.3, the proposed LCDD approach by introducing *k*-NN and formulating the manifold learning under the NNLS framework is presented. Section 4 provides

numerical tests of truss structures to demonstrate the effectiveness of LCDD against noisy data. In Section 6.5, continuum mechanics with elastic solid is considered to assess the accuracy and convergence properties of LCDD when the noisy data is of high-dimensional phase space. Finally, concluding remarks and discussions are given in Section 6.6.

6.2 Data-Driven Computational Framework

6.2.1 Governing equations

The deformation of a solid occupying the domain Ω bounded by Neumann boundary Γ_t and Dirichlet boundary Γ_u can be described by two basic laws:

1. Kinetics (equilibrium)

$$\begin{aligned}\operatorname{div}\boldsymbol{\sigma}(\mathbf{x}) + \mathbf{b}(\mathbf{x}) &= \mathbf{0}, \quad \mathbf{x} \in \Omega, \\ \boldsymbol{\sigma}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) &= \bar{\mathbf{t}}(\mathbf{x}), \quad \mathbf{x} \in \Gamma_t,\end{aligned}\tag{5.35}$$

in which $\boldsymbol{\sigma}$ is the symmetric Cauchy stress tensor, \mathbf{b} is a distributed body force, and $\bar{\mathbf{t}}$ is the traction applied on Γ_t , with surface outward unit normal vector \mathbf{n} .

2. Kinematics (compatibility)

$$\begin{aligned}\boldsymbol{\varepsilon}(\mathbf{x}) &= \frac{1}{2}(\nabla\mathbf{u}(\mathbf{x}) + \nabla\mathbf{u}^T(\mathbf{x})), \quad \mathbf{x} \in \Omega, \\ \mathbf{u}(\mathbf{x}) &= \bar{\mathbf{u}}(\mathbf{x}), \quad \mathbf{x} \in \Gamma_u,\end{aligned}\tag{5.36}$$

where $\boldsymbol{\varepsilon}$ is the strain tensor related to the displacement vector \mathbf{u} over Ω , and displacement $\bar{\mathbf{u}}$ is prescribed on Γ_u . In this study, the state of the system is defined by the pairs of strain and stress, denoted as the *local state* $\mathbf{s}(\mathbf{x}) = (\boldsymbol{\varepsilon}(\mathbf{x}), \boldsymbol{\sigma}(\mathbf{x}))$, and the appropriate space characterizing the local state over Ω is referred to as the *phase space* \mathcal{Z} . For the following discussions, we define the

sets \mathcal{C}_σ and \mathcal{C}_u for *equilibrium admissible* stress fields and *kinematically admissible* displacement fields, respectively, i.e.:

$$\mathcal{C}_\sigma = \{ \boldsymbol{\tau} \in \mathcal{V}_\sigma \mid \operatorname{div} \boldsymbol{\tau} + \mathbf{b} = \mathbf{0} \text{ in } \Omega, \text{ and } \boldsymbol{\tau} \cdot \mathbf{n} = \bar{\mathbf{t}} \text{ on } \Gamma_t \}, \quad (5.37)$$

$$\mathcal{C}_u = \{ \mathbf{v} \in \mathcal{V}_u \mid \mathbf{v} = \bar{\mathbf{u}} \text{ on } \Gamma_u \}, \quad (5.38)$$

where $\mathcal{V}_\sigma = [L^2(\Omega)]^6$ is the space in which the symmetric stress field is being sought, and $\mathcal{V}_u = [H^1(\Omega)]^3$ is the space for the displacement field. In addition, we define an admissible set $\mathcal{C} \subset \mathcal{Z}$ (the *physical set*) consisting of all states $s = (\boldsymbol{\varepsilon}, \boldsymbol{\sigma})$ that meet the compatibility and equilibrium equations in (5.35) and (5.36), respectively, and the state $s \in \mathcal{C}$ is called the *mechanics state*.

Seeking the solution fields $(\boldsymbol{\varepsilon}, \boldsymbol{\sigma})$ over the physical set \mathcal{C} , conventional computation usually requires constitutive equations to relate the two fields. For instance, a generic constitutive equation for elastic material is:

$$\boldsymbol{\sigma} - \frac{\partial \mathcal{A}}{\partial \boldsymbol{\varepsilon}} = \mathbf{0}, \quad (5.39)$$

where \mathcal{A} is the strain energy density function, and (5.39) is usually called the constitutive law. When a relationship in (5.39) such as Hooke's law or the hyperelasticity hypothesis is assumed, the solution that satisfies (5.35), (5.36) and (5.39) can be uniquely obtained.

While (5.35) and (5.36) are well-established in mechanics, the parameters in the stress-strain constitutive relation in (5.39) are usually characterized by experimental discrete data with noise and uncertainty. Errors arise during the procedures of i) determining the material functional form and ii) calibrating parameters in the constitutive law in (5.39) by experimental data fitting

(Kirchdoerfer & Ortiz 2016). Moreover, constitutive parameter identification for complex material systems remains challenging due to the difficulties in solving inverse problems. To circumvent these issues, data-driven computing that directly utilizes a material database and avoids a pre-defined constitutive relationship was proposed to solve boundary-value problems (Ibañez et al. 2016; Kirchdoerfer & Ortiz 2016).

Different from the employment of constitutive models, in the data-driven scenario the material behavior is characterized by a set $\mathbb{E}_{loc} = \{(\hat{\boldsymbol{\varepsilon}}^i, \hat{\boldsymbol{\sigma}}^i), i = 1, \dots, P\}$ of data points, called the *local material data set*, resulting from experimental measurement. The corresponding *global material data set* $\mathbb{E} \subset \mathcal{Z}$ is identified as the ensemble of the local material sets \mathbb{E}_{loc} over Ω . In this manner, the solution of data-driven computing is essentially the intersection of the global material data set \mathbb{E} and the physical constraint set \mathcal{C} , i.e. $\mathbb{E} \cap \mathcal{C}$.

6.2.2 Data-driven problem formulation

Data-driven computing and many other data-enabled applications, such as DDDAS (Dynamics Data-Driven Application Systems) (Darema & Rotea 2006) and parameter identification for pre-defined material models (Avril et al. 2008; Ben Azzouna, Feissel & Villon 2015; Bonne & Constantinescu 2005), share a similar idea that incorporating additional data into the modelling process of a given system. The main difference is that data-driven computing is developed for predictive simulation (a forward problem) while parameter identification is for material calibration (an inverse problem). We refer interested readers to the papers (Avril et al. 2008; Ben Azzouna, Feissel & Villon 2015; Bonne & Constantinescu 2005) for more details of parameter identification.

Inspired by the problem setting in parameter identification, data-driven computing can be stated as the following generic minimizing problem:

$$\hat{\boldsymbol{s}}^* = \arg \min_{\hat{\boldsymbol{s}} \in \mathbb{E}} J(\hat{\boldsymbol{s}}) \text{ with} \quad (5.40a)$$

$$J(\hat{\boldsymbol{s}}) = \min_{(\boldsymbol{u}, \boldsymbol{\sigma}) \in \mathcal{C}_u \times \mathcal{C}_\sigma} \mathcal{H}(\boldsymbol{u}, \boldsymbol{\sigma}, \hat{\boldsymbol{s}}), \quad (6.6b)$$

where $\hat{\boldsymbol{s}} = (\hat{\boldsymbol{\varepsilon}}, \hat{\boldsymbol{\sigma}})$ denotes the material data associated to the global material database \mathbb{E} , and $\hat{\boldsymbol{s}}^* = (\hat{\boldsymbol{\varepsilon}}^*, \hat{\boldsymbol{\sigma}}^*)$ is the optimum material data closest to the computed state $\boldsymbol{s} = (\boldsymbol{\varepsilon}, \boldsymbol{\sigma})$ under a certain “distance” measurement defined by a given functional \mathcal{H} , which is to be elaborated in next section. Here, we use the “hat” symbol to distinguish the material data $\hat{\boldsymbol{s}}$ from the mechanics state \boldsymbol{s} that is sought to satisfy the physical laws in (5.35) and (5.36). In computation, strain $\boldsymbol{\varepsilon}$ is computed from the displacement $\boldsymbol{u} \in \mathcal{C}_u$ using the strain-displacement relationship in (5.36), and denoted by $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}[\boldsymbol{u}]$. With the material state $\hat{\boldsymbol{s}} = (\hat{\boldsymbol{\varepsilon}}, \hat{\boldsymbol{\sigma}})$ in \mathbb{E} and the mechanics state $\boldsymbol{s} = (\boldsymbol{\varepsilon}[\boldsymbol{u}], \boldsymbol{\sigma})$ in \mathcal{C} , the data-driven problem (5.40) is equivalent to two compact forms as follows:

$$\min_{\hat{\boldsymbol{s}} \in \mathbb{E}} \min_{(\boldsymbol{u}, \boldsymbol{\sigma}) \in \mathcal{C}_u \times \mathcal{C}_\sigma} \mathcal{H}(\boldsymbol{u}, \boldsymbol{\sigma}, \hat{\boldsymbol{s}}) = \min_{(\boldsymbol{u}, \boldsymbol{\sigma}) \in \mathcal{C}_u \times \mathcal{C}_\sigma} \min_{\hat{\boldsymbol{s}} \in \mathbb{E}} \mathcal{H}(\boldsymbol{u}, \boldsymbol{\sigma}, \hat{\boldsymbol{s}}). \quad (5.41)$$

It is worth to emphasize that the minimizer used in (5.40a) is to relax the overly-rigid condition of the intersection solution $\mathbb{E} \cap \mathcal{C}$ since it is likely empty due to the discrete nature of the material data set (Conti, Müller & Ortiz 2018; Kirchdoerfer & Ortiz 2016). As shown in (5.41), data-driven computing is to find the mechanics states from the physical set \mathcal{C} while minimizing the “distance” to the material data set \mathbb{E} such that the data-driven solution is determined directly from the material data set without specifying any constitutive models such as (5.39). From an optimization perspective, the solution procedures of this data-driven problem

involve an alternate-direction search where a minimization with respect to $(\mathbf{u}, \boldsymbol{\sigma})$ is followed by a minimization with respect to $\hat{\mathbf{s}} = (\hat{\boldsymbol{\varepsilon}}, \hat{\boldsymbol{\sigma}})$, denoted as a *global step* and a *local step*, respectively.

6.2.3 Data-driven solver

For numerical implementation, state variables are presented in a discrete manner such as $\mathbf{s}(\mathbf{x}_\alpha) \equiv \mathbf{s}_\alpha = (\boldsymbol{\varepsilon}_\alpha, \boldsymbol{\sigma}_\alpha) \in \mathbb{R}^s \times \mathbb{R}^s$, $\alpha = 1, \dots, m$, where $\mathbf{x}_\alpha \in \Omega$ are the coordinates of the m integration points, and $s=3$ and $s=6$ for symmetric 2D and 3D elasticity, respectively. Accordingly, the material data is assumed to be available at each integration point, $\hat{\mathbf{s}}(\mathbf{x}_\alpha) \equiv \hat{\mathbf{s}}_\alpha = (\hat{\boldsymbol{\varepsilon}}_\alpha, \hat{\boldsymbol{\sigma}}_\alpha) \in \mathbb{R}^s \times \mathbb{R}^s$, $\alpha = 1, \dots, m$. Given a homogeneous material, the material behavior at each \mathbf{x}_α is characterized by the same local material data set, i.e. $\mathbb{E}_\alpha = \mathbb{E}_{loc}$, $\alpha = 1, \dots, m$, and thus, the discrete counterpart of the global material data set is defined as $\mathbb{E}^h = \prod_{\alpha=1}^m \mathbb{E}_{loc}$.

Local step of data-driven solver

Based on the previous work (Kirchdoerfer & Ortiz 2016), a local distance function $d(\mathbf{s}_\alpha, \hat{\mathbf{s}}_\alpha)$ is used to measure the distance between the material data $\hat{\mathbf{s}}_\alpha$ and the mechanics state \mathbf{s}_α at each integration point \mathbf{x}_α , defined as

$$d^2(\mathbf{s}_\alpha, \hat{\mathbf{s}}_\alpha) = d^{\varepsilon^2}(\boldsymbol{\varepsilon}_\alpha, \hat{\boldsymbol{\varepsilon}}_\alpha) + d^{\sigma^2}(\boldsymbol{\sigma}_\alpha, \hat{\boldsymbol{\sigma}}_\alpha), \quad (5.42)$$

where

$$\begin{aligned}
d^{\varepsilon^2}(\boldsymbol{\varepsilon}_\alpha, \hat{\boldsymbol{\varepsilon}}_\alpha) &= \frac{1}{2}(\boldsymbol{\varepsilon}_\alpha - \hat{\boldsymbol{\varepsilon}}_\alpha)^T \mathbf{M}^\varepsilon (\boldsymbol{\varepsilon}_\alpha - \hat{\boldsymbol{\varepsilon}}_\alpha), \\
d^{\sigma^2}(\boldsymbol{\sigma}_\alpha, \hat{\boldsymbol{\sigma}}_\alpha) &= \frac{1}{2}(\boldsymbol{\sigma}_\alpha - \hat{\boldsymbol{\sigma}}_\alpha)^T \mathbf{M}^\sigma (\boldsymbol{\sigma}_\alpha - \hat{\boldsymbol{\sigma}}_\alpha).
\end{aligned} \tag{5.43}$$

Note that the coefficient matrices $\mathbf{M}^\varepsilon \in \mathbb{R}^{s \times s}$ and $\mathbf{M}^\sigma \in \mathbb{R}^{s \times s}$ (usually $\mathbf{M}^\sigma = \mathbf{M}^{\varepsilon^{-1}}$) are regarded as the multi-dimensional weights. They need not represent actual material properties but instead can be any positive-definite matrices (Ayensa-Jiménez et al. 2018; Kirchdoerfer & Ortiz 2016). The choice of the coefficient matrices affects the performance of data-driven computing due to the distance scaling effect. One approach for selecting the weighted coefficients is by computing the covariance of the material data set and using the so-called Mahalanobis distance for multivariate data, as proposed in (Ayensa-Jiménez et al. 2018).

With the distance function in (5.42), the functional \mathcal{H} used in the data-driven problem (Equation (5.40) or (5.41)) can be defined as the following integral form

$$\mathcal{H}(\mathbf{u}, \boldsymbol{\sigma}, \hat{\mathbf{s}}) = \int_{\Omega} d^2(s(\mathbf{x}), \hat{\mathbf{s}}(\mathbf{x})) d\Omega \approx \sum_{\alpha=1}^m d^2(s_\alpha, \hat{\mathbf{s}}_\alpha) V_\alpha, \tag{5.44}$$

where $\{V_\alpha\}_{\alpha=1}^m$ are the weights associated with the m integration points. Considering the functional \mathcal{H} is simply a summation of local operations at each integration point, the local step of (5.40a) can be decomposed into m independent local minimization problems: find the optimum local material data $\hat{\mathbf{s}}_\alpha^* \in \mathbb{E}_{loc}$ such that minimizes the distance to a given local mechanics state s_α , i.e.

$$\hat{\mathbf{s}}_\alpha^* = \arg \min_{\hat{\mathbf{s}}_\alpha \in \mathbb{E}_{loc}} d^2(s_\alpha, \hat{\mathbf{s}}_\alpha) = \arg \min_{\hat{\mathbf{s}}_\alpha \in \mathbb{E}_{loc}} d^{\varepsilon^2}(\boldsymbol{\varepsilon}_\alpha, \hat{\boldsymbol{\varepsilon}}_\alpha) + d^{\sigma^2}(\boldsymbol{\sigma}_\alpha, \hat{\boldsymbol{\sigma}}_\alpha), \tag{5.45}$$

for $\alpha = 1, \dots, m$. Thus, the collection of optimum local material data $\{\hat{\mathbf{s}}_\alpha^*\}_{\alpha=1}^m \in \mathbb{E}^h$ approximates the solution $\hat{\mathbf{s}}^* \in \mathbb{E}$ in (5.40a).

Global step of data-driven solver

Considering the functional \mathcal{H} defined in (5.42) and (5.44), the global step of the data-driven problem (5.40b) is reformulated as:

$$\begin{aligned} J(\hat{\mathbf{s}}) &= \min_{\mathbf{u} \in \mathcal{C}_u, \boldsymbol{\sigma} \in \mathcal{V}_\sigma} \int_{\Omega} \left(d^{\varepsilon^2}(\boldsymbol{\varepsilon}[\mathbf{u}], \hat{\boldsymbol{\varepsilon}}) + d^{\sigma^2}(\boldsymbol{\sigma}, \hat{\boldsymbol{\sigma}}) \right) d\Omega, \\ \text{subject to: } \quad & \operatorname{div} \boldsymbol{\sigma} + \mathbf{b} = \mathbf{0} \quad \text{in } \Omega, \\ & \boldsymbol{\sigma} \cdot \mathbf{n} = \bar{\mathbf{t}} \quad \text{on } \Gamma_t. \end{aligned} \quad (5.46)$$

Here, the equilibrium admissible set \mathcal{C}_σ has been represented by the constraints explicitly. This global step searches for the physically admissible mechanics state $\mathbf{s} = (\boldsymbol{\varepsilon}[\mathbf{u}], \boldsymbol{\sigma})$ with the minimal distance to a material data $\hat{\mathbf{s}} = (\hat{\boldsymbol{\varepsilon}}, \hat{\boldsymbol{\sigma}})$ given by the previous local step (5.45).

Enforcing the equilibrium constraints by means of Lagrange multipliers, the associated Lagrangian of (5.46) becomes

$$\begin{aligned} \mathcal{L}_{DD}(\mathbf{u}, \boldsymbol{\sigma}, \boldsymbol{\lambda}) &= \int_{\Omega} \left(d^{\varepsilon^2}(\boldsymbol{\varepsilon}[\mathbf{u}], \hat{\boldsymbol{\varepsilon}}) + d^{\sigma^2}(\boldsymbol{\sigma}, \hat{\boldsymbol{\sigma}}) \right) d\Omega \\ &+ \int_{\Omega} \boldsymbol{\lambda} \cdot (\operatorname{div} \boldsymbol{\sigma} + \mathbf{b}) d\Omega + \int_{\Gamma_t} \boldsymbol{\eta} \cdot (\boldsymbol{\sigma} \cdot \mathbf{n} - \bar{\mathbf{t}}) d\Gamma, \end{aligned} \quad (5.47)$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ are the Lagrange multipliers in proper function spaces. Considering $\boldsymbol{\eta} = -\boldsymbol{\lambda}$ on Γ_t , the variational form is

$$\begin{aligned} \delta \mathcal{L}_{DD} &= \int_{\Omega} \left(\delta \boldsymbol{\varepsilon}[\mathbf{u}] : \mathbf{M}^\varepsilon : (\boldsymbol{\varepsilon}[\mathbf{u}] - \hat{\boldsymbol{\varepsilon}}) + \delta \boldsymbol{\sigma} : (\mathbf{M}^\sigma : (\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}})) \right) d\Omega \\ &- \int_{\Omega} \delta \boldsymbol{\sigma} : \boldsymbol{\varepsilon}[\boldsymbol{\lambda}] d\Omega + \int_{\Gamma_u} (\delta \boldsymbol{\sigma} \cdot \mathbf{n}) \cdot \boldsymbol{\lambda} d\Gamma - \int_{\Omega} \delta \boldsymbol{\varepsilon}[\boldsymbol{\lambda}] : \boldsymbol{\sigma} d\Omega + \int_{\Omega} \delta \boldsymbol{\lambda} \cdot \mathbf{b} d\Omega + \int_{\Gamma_t} \delta \boldsymbol{\lambda} \cdot \bar{\mathbf{t}} d\Gamma, \end{aligned} \quad (5.48)$$

where $\boldsymbol{\varepsilon}[\boldsymbol{\lambda}] = 1/2(\nabla\boldsymbol{\lambda} + \nabla\boldsymbol{\lambda}^T)$. It is observed that $\boldsymbol{\lambda} = \mathbf{0}$ on Γ_u analogous to the homogeneous boundary condition for the test functions used in Galerkin approximation. Then, the following weak form variational equations associated to (5.47) are obtained as follows

$$\begin{aligned}\delta_u \mathcal{L}_{DD} = 0 &\Rightarrow \int_{\Omega} \delta \boldsymbol{\varepsilon}[\mathbf{u}] : \mathbf{M}^\varepsilon : \boldsymbol{\varepsilon}[\mathbf{u}] d\Omega = \int_{\Omega} \delta \boldsymbol{\varepsilon}[\mathbf{u}] : \mathbf{M}^\varepsilon : \hat{\boldsymbol{\varepsilon}} d\Omega, \\ \delta_\lambda \mathcal{L}_{DD} = 0 &\Rightarrow \int_{\Omega} \delta \boldsymbol{\varepsilon}[\boldsymbol{\lambda}] : \boldsymbol{\sigma} d\Omega = \int_{\Omega} \delta \boldsymbol{\lambda} \cdot \mathbf{b} d\Omega + \int_{\Gamma_f} \delta \boldsymbol{\lambda} \cdot \bar{\mathbf{t}} d\Gamma := P^{\text{ext}}(\delta \boldsymbol{\lambda}), \\ \delta_\sigma \mathcal{L}_{DD} = 0 &\Rightarrow \int_{\Omega} \delta \boldsymbol{\sigma} : (\mathbf{M}^\sigma : \boldsymbol{\sigma} - \boldsymbol{\varepsilon}[\boldsymbol{\lambda}]) d\Omega = \int_{\Omega} \delta \boldsymbol{\sigma} : \mathbf{M}^\sigma : \hat{\boldsymbol{\sigma}} d\Omega.\end{aligned}\quad (5.49)$$

The solutions to (5.49) can be approximated by using mixed-form numerical methods with respect to the field variables $(\mathbf{u}, \boldsymbol{\sigma}, \boldsymbol{\lambda})$.

By employing proper approximation functions based on Galerkin method in conjunction with an integration scheme for (5.49), we obtain the following discrete equations

$$\sum_{J=1}^N \left(\sum_{\alpha=1}^m V_\alpha \mathbf{B}_{\alpha J}^T \mathbf{M}^\varepsilon \mathbf{B}_{\alpha J} \right) \mathbf{d}_J = \sum_{\alpha=1}^m V_\alpha \mathbf{B}_{\alpha I}^T \mathbf{M}^\varepsilon \hat{\boldsymbol{\varepsilon}}_\alpha, \quad I = 1, \dots, N, \quad (5.50a)$$

$$\sum_{\alpha=1}^m V_\alpha \mathbf{B}_{\alpha I}^T \boldsymbol{\sigma}_\alpha = \mathbf{f}_I, \quad I = 1, \dots, N, \quad (6.16b)$$

$$\mathbf{M}_\alpha^\sigma \boldsymbol{\sigma}_\alpha - \sum_{I=1}^N \mathbf{B}_{\alpha I} \mathbf{A}_I = \mathbf{M}^\sigma \hat{\boldsymbol{\sigma}}_\alpha, \quad \alpha = 1, \dots, m, \quad (6.16c)$$

where N is the number of discretization nodes, $\{\mathbf{d}_I\}_{I=1}^N$ are the nodal displacement vectors, $\{\mathbf{A}_I\}_{I=1}^N$ are the nodal Lagrange multiplier vectors implying the displacement adjustment, $\{\mathbf{f}_I\}_{I=1}^N$ are the nodal force vectors, $\mathbf{B}_{\alpha I}$ is the strain-displacement matrix that encodes the connectivity and geometry corresponding to the adopted numerical discretization method. The last equation in (5.50) takes advantage of that the stress approximation is evaluated by a integration scheme such

that the third equation in (5.49) can be disassembled into a point-wise system at the integration points $\{\mathbf{x}_\alpha\}_{\alpha=1}^m$. Then, the coupled system (5.50b) and (5.50c) can be equivalently recast as

$$\sum_{J=1}^N \left(\sum_{\alpha=1}^m V_\alpha \mathbf{B}_{\alpha I}^T \mathbf{M}^{\sigma-1} \mathbf{B}_{\alpha J} \right) \mathbf{A}_J = \mathbf{f}_I - \sum_{\alpha=1}^m V_\alpha \mathbf{B}_{\alpha I}^T \hat{\boldsymbol{\sigma}}_\alpha, \quad I = 1, \dots, N, \quad (5.51a)$$

$$\boldsymbol{\sigma}_\alpha = \hat{\boldsymbol{\sigma}}_\alpha + \mathbf{M}^{\sigma-1} \sum_{I=1}^N \mathbf{B}_{\alpha I} \mathbf{A}_I, \quad \alpha = 1, \dots, m, \quad (6.17b)$$

in which the unknowns $\{\mathbf{A}_I\}_{I=1}^N$ and $\{\boldsymbol{\sigma}_\alpha\}_{\alpha=1}^m$ are solved in sequence.

It can be seen from (5.50a) and (5.51b) that the displacement solution $\{\mathbf{d}_I\}_{I=1}^N$ is solved to comply with compatibility provided by the strain data $\{\hat{\boldsymbol{\epsilon}}_\alpha\}_{\alpha=1}^m$, while the the displacement adjustment $\{\mathbf{A}_I\}_{I=1}^N$ is driven by the force residuals between the external force and the internal force given by the stress data $\{\hat{\boldsymbol{\sigma}}_\alpha\}_{\alpha=1}^m$. Thus, Equations (5.50a), (5.51a) and (5.51b) solve the global step of the data-driven problem in (5.40b) (or (5.46)) with the material data $\{\hat{\mathbf{s}}_\alpha\}_{\alpha=1}^m$ as input, yielding the discrete solution of the mechanics state, $\{\mathbf{s}_\alpha\}_{\alpha=1}^m$, that conforms to the physical constraints.

While the procedures for solving the data-driven problem in (5.40) have been presented in a similar way in literature (Ayensa-Jiménez et al. 2018; Kirchdoerfer & Ortiz 2016), we provide a simple sketch here for completeness of the study. Given a material data solution $\{\hat{\mathbf{s}}_\alpha^{(\nu)}\}_{\alpha=1}^m$ of the ν -th previous iteration, iterate the following global step and local step until convergence:

- 1) *Global Step*. Input: $\{\hat{\mathbf{s}}_\alpha^{(\nu)}\}_{\alpha=1}^m \rightarrow$ Output: $\{\mathbf{s}_\alpha^{(\nu)}\}_{\alpha=1}^m$

1.1 Solve Equations (5.50a) for $\{\mathbf{d}_I^{(\nu)}\}_{I=1}^N$ and (5.51a) for $\{\mathcal{A}_I^{(\nu)}\}_{I=1}^N$.

1.2 Update mechanics states for $\{\mathbf{s}_\alpha^{(\nu)}\}_{\alpha=1}^m = \{(\boldsymbol{\varepsilon}_\alpha^{(\nu)}, \boldsymbol{\sigma}_\alpha^{(\nu)})\}_{\alpha=1}^m$ via

$$\boldsymbol{\varepsilon}_\alpha^{(\nu)} = \sum_{I=1}^N \mathbf{B}_{\alpha I} \mathbf{d}_I^{(\nu)} \text{ and } \boldsymbol{\sigma}_\alpha^{(\nu)} = \hat{\boldsymbol{\sigma}}_\alpha^{(\nu)} + \mathbf{M}^{\sigma-1} \sum_{I=1}^N \mathbf{B}_{\alpha I} \mathcal{A}_I^{(\nu)} \text{ in (5.51b)}$$

2) *Local Step*. Input: $\{\mathbf{s}_\alpha^{(\nu)}\}_{\alpha=1}^m \rightarrow$ Output: $\{\hat{\mathbf{s}}_\alpha^{(\nu+1)}\}_{\alpha=1}^m$

for $\alpha = 1, \dots, m$, solve Equation (5.45) for $\hat{\mathbf{s}}_\alpha^{(\nu+1)}$

Remark 6.1: While the distance-minimizing relaxation in (5.45) ensures the existence of data-driven solution, it is observed that this standard distance-minimizing data-driven computing (DMDD) solver proposed in (Kirchdoerfer & Ortiz 2016) is sensitive to data noise and outliers (Ayensa-Jiménez et al. 2018; Kirchdoerfer & Ortiz 2017) because the local minimization stage in (5.45) simply finds the nearest data point from the given material data set regardless of any latent data structure. For instance, the data-driven solution could be overwhelmingly influenced by the outliers who locate near to the physical manifold \mathcal{C} but do not conform to the overall material data pattern (or the latent statistical model) of \mathbb{E} . Due to the lack of knowledge of the underlying data manifold, it requires a large amount of data to achieve sufficiently accurate predictions. The useful data, however, is expensive to collect in the area of SBES.

To address these issues, a new data-driven method is developed in next section, where the data-driven local solver is integrated with manifold learning techniques in order to remedy the non-convex and non-smooth properties in material data sets and reduce the susceptibility to noise and outliers in the material data set.

6.3 Locally Convex Data-Driven Scheme

In this section, we introduce a locally convex data-driven computing (LCDD) approach which utilizes the *clustering technique* and a local reconstruction of *convex envelop* based on manifold learning techniques. The proposed approach is based on the assumptions that there exists an underlying low-dimensional manifold, and that the relationship between the material data and the manifold coordinates is locally linear and smoothly varying.

6.3.1 Local convex hull construction based on k-nearest neighbors

In the subsequent discussion, the local state $\mathbf{s}_\alpha = (\boldsymbol{\varepsilon}_\alpha, \boldsymbol{\sigma}_\alpha) \in \mathbb{R}^s \times \mathbb{R}^s$ is recast into a vector format for ease of exposition, i.e. $\mathbf{s}_\alpha = [\boldsymbol{\varepsilon}_\alpha^T \boldsymbol{\sigma}_\alpha^T]^T \in \mathbb{R}^{2s}$. Accordingly, an induced norm “ $\|\cdot\|$ ” based on (5.42) is defined as

$$\|\mathbf{s}_\alpha\|^2 = \frac{1}{2} \mathbf{s}_\alpha^T \text{diag}([\mathbf{M}^\varepsilon, \mathbf{M}^\sigma]) \mathbf{s}_\alpha. \quad (5.52)$$

Given a local mechanics state \mathbf{s}_α , the k nearest neighbor (k -NN) material data points in \mathbb{E}_{loc} are first identified using the same metric induced by the given norm “ $\|\cdot\|$ ”, and collected as $\{\hat{\mathbf{s}}_\alpha^i\}_{i \in N_k(\mathbf{s}_\alpha)}$, in which the indices are stored in a set $N_k(\mathbf{s}_\alpha)$. Note that specifying the number of k -NN points, in general, provides an opportunity to incorporate a priori knowledge about the material database and therefore, enhance the robustness of data learning (Saul & Roweis 2003).

To recover the convexity, we extend the associated k -NN data points $\{\hat{\mathbf{s}}_\alpha^i\}_{i \in N_k(s_\alpha)}$ to a local convex set $\mathcal{E}(s_\alpha)$, i.e., $\text{Conv}:\{\hat{\mathbf{s}}_\alpha^i\}_{i \in N_k(s_\alpha)} \subset \mathbb{E}_{loc} \rightarrow \mathcal{E}(s_\alpha)$, where

$$\mathcal{E}(s_\alpha) = \text{Conv}(\{\hat{\mathbf{s}}_\alpha^i\}_{i \in N_k(s_\alpha)}) = \left\{ \sum_{i \in N_k(s_\alpha)} w_i \hat{\mathbf{s}}_\alpha^i \mid \sum_{i \in N_k(s_\alpha)} w_i = 1 \text{ and } w_i \geq 0 \right\}, \quad (5.53)$$

or concisely denoted as \mathcal{E}_α . Accordingly, the minimization problem for the local step in (5.45) is modified as: Given $\{\hat{\mathbf{s}}_\alpha^i\}_{i \in N_k(s_\alpha)} \subset \mathbb{E}_{loc}$ for s_α , solve $\hat{\mathbf{s}}_\alpha^*$ such that

$$\hat{\mathbf{s}}_\alpha^* = \arg \min_{\hat{\mathbf{s}}_\alpha \in \mathcal{E}_\alpha} \|\| s_\alpha - \hat{\mathbf{s}}_\alpha \|\|^2, \quad (5.54)$$

for $\alpha = 1, \dots, m$. By comparing (5.45) and (5.54), we can observe that the solution space \mathbb{E}_{loc} used in the standard data-driven schemes in (Ayensa-Jiménez et al. 2018; Kirchdoerfer & Ortiz 2016) is simply replaced by the associated convex hull \mathcal{E}_α that is locally reconstructed around the input s_α by learning techniques, allowing to capture the local material manifold. As a result, the optimum local material data $\hat{\mathbf{s}}_\alpha^*$ is sought from the reconstruction set \mathcal{E}_α that enjoys convexity and smoothness. With the definition in (5.53), the solution of the minimization problem (5.54) is given by solving an equivalent form as follows:

$$\begin{aligned} \mathbf{w}^* &= \arg \min_{\mathbf{w} \in \mathbb{R}^k} \|\| s_\alpha - \sum_{i \in N_k(s_\alpha)} w_i \hat{\mathbf{s}}_\alpha^i \|\|^2, \\ \text{subject to: } &\sum_{i \in N_k(s_\alpha)} w_i = 1, \\ &w_i \geq 0, \forall i \in N_k(s_\alpha), \end{aligned} \quad (5.55a)$$

and then the optimum data solution $\hat{\mathbf{s}}_\alpha^*$ can be retrieved by using the convex combination of

$\{\hat{\mathbf{s}}_\alpha^i\}_{i \in N_k(s_\alpha)}$ with the computed weight vector $\mathbf{w}^* \in \mathbb{R}^k$, i.e.

$$\hat{\mathbf{s}}_\alpha^* = \hat{\mathbf{S}}_\alpha \mathbf{w}^*, \quad (6.21b)$$

where $\hat{\mathbf{S}}_\alpha \in \mathbb{R}^{2s \times k}$ is the corresponding matrix composed of the k -NN data points $\{\hat{\mathbf{s}}_\alpha^i\}_{i \in N_k(s_\alpha)}$.

Remark 6.2: Equation (range(\mathbf{W}) $\subset \mathbb{R}^{\mathcal{N}}$ a) is a *constrained regression* or *constrained least-squares* problem under a *invariance* constraint and a *non-negative* constraint. The invariance constraint imposes the partition of unity on the weight array \mathbf{w} , i.e. $\sum_{i=1}^k w_i = \mathbf{I}^T \mathbf{w} = 1$, where $\mathbf{I} = [1, 1, \dots, 1]^T \in \mathbb{R}^k$. It ensures the invariance of the reconstruction weights \mathbf{w}^* to rotations, rescaling, and translations of the same k -NN data points, and thus, the weights characterize geometric properties independent of a particular frame of reference (Roweis & Saul 2000; Saul & Roweis 2003). Moreover, the constraint guarantees the linear approximation property such that $\hat{\mathbf{s}}_\alpha^*$ is in the affine space spanned by $\{\hat{\mathbf{s}}_\alpha^i\}_{i \in N_k(s_\alpha)}$. When we further consider the non-negative constraint, the approximation $\hat{\mathbf{s}}_\alpha^*$ is restricted to the convex hull of its neighbors $\{\hat{\mathbf{s}}_\alpha^i\}_{i \in N_k(s_\alpha)}$ (see Figure 6.1). The imposed convexity tends to enhance the robustness of linear regressions to outliers (Cevikalp & Triggs 2010; Saul & Roweis 2003), and overcome numerical instability across different clusters of neighbor points during data-driven iterations. Moreover, it is well known that the non-negative constraint naturally imposes sparseness on the coefficient solution \mathbf{w}^* .

Essentially, the local minimization problem in (5.54) or (5.55) is an application of manifold learning involving locally linear embedding (LLE) (Roweis & Saul 2000) in conjunction with convex reconstructions (Saul & Roweis 2003), which aims to identify the low-dimensional manifold of high-dimensional datasets. So the modified local step of data-driven

problem in (5.54) can be interpreted as a process of seeking the data approximation based on the low-dimensional manifold \mathcal{E}_α associated to the given input mechanics state \mathbf{s}_α . From a geometrical point of view, it searches the projection (i.e. the optimum local material data $\hat{\mathbf{s}}_\alpha^* \in \mathcal{E}_\alpha$) of the given point \mathbf{s}_α onto the associated convex set \mathcal{E}_α . If \mathbf{s}_α locates inside \mathcal{E}_α , the projection is represented by \mathbf{s}_α itself (Figure 6.1(a)). Otherwise, the projection is attained at a boundary point of \mathcal{E}_α closest to the local state \mathbf{s}_α (Figure 6.1(b)).

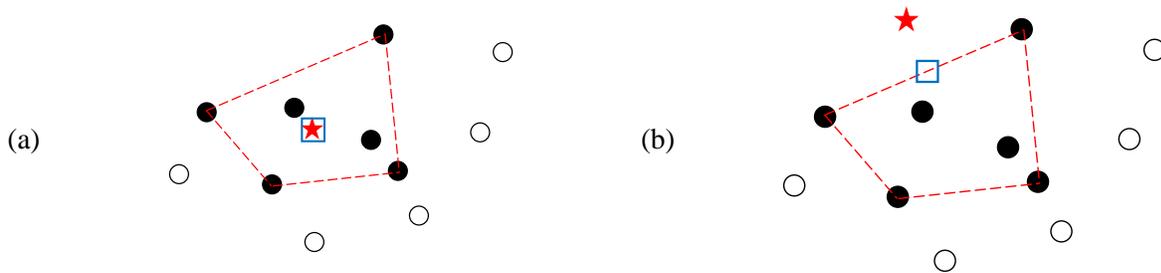


Figure 6.1: Sketch of the projection $\hat{\mathbf{s}}_\alpha^*$ (the blue square) on a convex hull \mathcal{E}_α (the region is depicted by red dashed lines) of k-NN points (the solid circles in black) when a local state \mathbf{s}_α (the red star) locates (a) inside and (b) outside \mathcal{E}_α . Neighbor points of $k = 6$ are used for demonstration.

6.3.2 Solving non-negative least squares

This section shows that the minimization problem in (5.55a) is related to a *non-negative least squares* (NNLS) problem, and thus can be solved by efficient NNLS solvers. Let us recall a standard NNLS problem (Appendix B): given a matrix $\mathbf{A} \in \mathbb{R}^{n \times p}$ (usually $p > n$) and a observed vector $\mathbf{z} \in \mathbb{R}^n$, find a nonnegative vector $\mathbf{y}^* \in \mathbb{R}_+^p$ to minimize the following function,

$$\begin{aligned} \mathbf{y}^* &= \arg \min_{\mathbf{x} \in \mathbb{R}^p} \|\mathbf{A}\mathbf{y} - \mathbf{z}\|, \\ \text{subject to: } & y_i \geq 0, \quad i = 1, \dots, p, \end{aligned} \quad (5.56)$$

where $\|\cdot\|$ stands for the standard Euclidean norm.

However, to solve the optimization problem (5.55) under the NNLS framework (5.56), the partition of unity constraint need to be taken into account. Thus, we propose to employ the quadratic penalty method (Boyd & Vandenberghe 2004) such that the minimization (5.55) is reformulated by penalizing the residuals of the partition of unity constraint in the auxiliary objective as follows

$$\begin{aligned} \mathbf{w}^* &= \arg \min_{\mathbf{w} \in \mathbb{R}^k} \|\|\mathbf{s}_\alpha - \hat{\mathbf{S}}_\alpha \mathbf{w}\|^2 + \kappa \|\mathbf{I}^T \mathbf{w} - \mathbf{1}\|^2, \\ \text{subject to: } & w_i \geq 0, \quad i = 1, \dots, k, \end{aligned} \quad (5.57)$$

where $\kappa > 0$ is a regularized coefficient to impose the associated constraint. Note that to conform with the Euclidean metric used in the standard NNLS solver in (5.56), the local state \mathbf{s}_α and the data matrix $\hat{\mathbf{S}}_\alpha$ are rescaled to $\tilde{\mathbf{s}}_\alpha$ and $\tilde{\mathbf{S}}_\alpha = [\tilde{\mathbf{s}}^1, \tilde{\mathbf{s}}^2, \dots, \tilde{\mathbf{s}}^k] \in \mathbb{R}^{2s \times k}$, respectively, under the transformation with a kernel matrix $\mathbf{W}_\alpha = \text{diag}([\mathbf{M}_\alpha^\varepsilon, \mathbf{M}_\alpha^\sigma])$ induced by the given norm “ $\|\|\cdot\|\|$ ”, e.g. $\|\|\mathbf{s}_\alpha\|\|^2 = \mathbf{s}_\alpha^T \mathbf{W}_\alpha \mathbf{s}_\alpha = \|\tilde{\mathbf{s}}_\alpha\|^2$. As a result, the optimization problem (5.57) can be recast into a NNLS form as shown in (5.56) by augmenting $\tilde{\mathbf{S}}_\alpha$ and $\tilde{\mathbf{s}}_\alpha$ with an additional row, i.e.

$$\begin{aligned} \mathbf{w}^* &= \arg \min_{\mathbf{w} \in \mathbb{R}^k} \|\tilde{\mathbf{S}}_\alpha^{\text{aug}} \boldsymbol{\alpha} - \tilde{\mathbf{s}}_\alpha^{\text{aug}}\|^2, \\ \text{subject to: } & w_i \geq 0, \quad i = 1, \dots, k, \end{aligned} \quad (5.58)$$

where

$$\tilde{\mathbf{S}}_\alpha^{\text{aug}} := \begin{bmatrix} \tilde{\mathbf{S}}_\alpha \\ \sqrt{\kappa} \mathbf{I}^T \end{bmatrix} \in \mathbb{R}^{(2s+1) \times k}, \quad \tilde{\mathbf{s}}_\alpha^{\text{aug}} := \begin{bmatrix} \tilde{\mathbf{s}}_\alpha \\ \sqrt{\kappa} \end{bmatrix} \in \mathbb{R}^{2s+1}. \quad (5.59)$$

To properly imposed the penalty term, we set $\kappa = \bar{\kappa} \text{tr}(\tilde{\mathbf{S}}_\alpha^T \tilde{\mathbf{S}}_\alpha)/k$, where $\bar{\kappa}$ is a large parameter.

The problem in $\mathcal{X} \subset \mathbb{R}^d$ can be used by the NNLS solver in Appendix B. After obtaining \mathbf{w}^* , the optimum local data solution is given by $\hat{\mathbf{s}}_\alpha^* = \hat{\mathbf{S}}_\alpha \mathbf{w}^*$, and the local cost function is obtained

$$\hat{F}_\alpha(\mathbf{s}_\alpha) = \min_{\hat{\mathbf{s}}_\alpha \in \mathcal{E}_\alpha} \|\mathbf{s}_\alpha - \hat{\mathbf{s}}_\alpha\|^2 = \|\mathbf{s}_\alpha - \hat{\mathbf{s}}_\alpha^*\|^2 = \|\mathbf{s}_\alpha - \hat{\mathbf{S}}_\alpha \mathbf{w}^*\|^2. \quad (5.60)$$

Nevertheless, the constrained least squares in (5.58) may still suffer numerical instability due to rank deficiency when the number of neighbors is larger than the rank of the neighborhood, i.e. $k > \text{rank}(\tilde{\mathbf{S}}_\alpha^{\text{aug}})$. Some further regularization should be introduced to the NNLS problem. In this study, a commonly used ridge regression (Hoerl & Kennard 1970), or called Tikhonov regularization, is applied to address the ill-posed issues, and the NNLS problem (5.58) is modified as

$$\begin{aligned} \mathbf{w}^* &= \arg \min_{\mathbf{w} \in \mathbb{R}^k} \|\tilde{\mathbf{S}}_\alpha^{\text{aug}} \mathbf{w} - \tilde{\mathbf{s}}_\alpha^{\text{aug}}\|^2 + \mu \|\mathbf{w}\|^2, \\ \text{subject to: } & w_i \geq 0, \quad i = 1, \dots, k, \end{aligned} \quad (5.61)$$

where the regularized coefficient is

$$\mu = \bar{\mu} \text{tr}(\tilde{\mathbf{S}}_\alpha^T \tilde{\mathbf{S}}_\alpha)/k, \quad (5.62)$$

where $\bar{\mu}$ is a small constant (set as 10^{-4}) such that the regularization affects negligibly the solution \mathbf{w}^* and the reconstruction $\hat{\mathbf{s}}_\alpha^*$. It is also shown that (Hastie, Tibshirani & Friedman 2009) this regularization imposes certain smoothness on the solution and guarantees a unique solution.

It has been shown (Tropp & Gilbert 2007; Yaghoobi & Davies 2015) that the Lawson-Hanson method used for the NNLS solver converges in a finite number of iterations less than the

size of the output coefficient vector, which is the size of k -NN in LCDD. Moreover, considering the small size of the local matrix $\hat{\mathbf{S}}_\alpha \in \mathbb{R}^{2s \times k}$, where $k, s \ll \min(N, m)$, the additional computational cost in solving the NNLS problem in (5.58) or (5.61) is negligible compared to solving the linear system (5.50).

Remark 6.3: As discussed in (Saul & Roweis 2003), the size range of k -NN points depends on various features of the data, such as the manifold geometry and the sampling density. In principle k should be greater than the underlying manifold dimensionality of the material data set \mathbb{E}_{loc} in order to explore the data structural and prevent overwhelming influence from outliers/noise. Moreover, the resultant neighborhoods should be localized enough to ensure the validity of locally linear approximation.

6.3.3 Locally convex data-driven algorithm

Given two high-dimensional sets, i.e. the physical set \mathcal{C}^h and the global material data set $\mathbb{E}^h = \prod_{\alpha=1}^m \mathbb{E}_{loc}$ are the discrete counterpart of \mathcal{C} and \mathbb{E} , the main idea of the proposed LCDD is to use the k -NN clustering and the NNLS for locally approximating the convex hull, allowing the data-driven iterative solver to search a reliable intersection of the physical set \mathcal{C}^h and the convex approximation of the global material data set \mathbb{E}^h . A simple algorithm for the proposed LCDD solver is listed in Table 3.1.

Table 6.1: Flow of the proposed LCDD solver

Given a convergence tolerance TOL and material databases \mathbb{E}_e , $e = 1, \dots, m$, then
<ol style="list-style-type: none"> 1. Initialize $\hat{\mathbf{s}}_e^{*(0)} = [\hat{\boldsymbol{\varepsilon}}_e^{*(0)T} \hat{\boldsymbol{\sigma}}_e^{*(0)T}]^T$, $e = 1, \dots, m$ randomly, and $\nu = 0$. 2. While $\max_{e=1, \dots, m} \ \hat{\mathbf{s}}_e^{*(\nu)} - \hat{\mathbf{s}}_e^{*(\nu-1)}\ < TOL$ <ol style="list-style-type: none"> a. Solve equations (5.50), and output $\{\mathbf{s}_e^{(\nu)}\}_{e=1}^m$ b. Construct k-NN neighborhood $N_k(\mathbf{s}_e^{(\nu)})$ for each local state $\mathbf{s}_e^{(\nu)}$, and obtain $\hat{\mathbf{S}}_e$. c. Solve NNLS (5.58) (or the stabilized NNLS (5.61)) by Algorithm 1 in Appendix B d. Output $\hat{\mathbf{s}}_e^{*(\nu+1)}$. e. Update: $\nu \leftarrow \nu + 1$ 3. Solution is $\mathbf{s}_e = [\boldsymbol{\varepsilon}_e^T \boldsymbol{\sigma}_e^T]^T \leftarrow \mathbf{s}_e^{(\nu)} = [\boldsymbol{\varepsilon}_e^{(\nu)T} \boldsymbol{\sigma}_e^{(\nu)T}]^T$, $e = 1, \dots, m$.

6.4 Application to Truss Structures

The proposed locally convex data-driven (LCDD) approach against the distance-minimizing data-driven (DMDD) approach (Kirchdoerfer & Ortiz 2016) are compared for truss structures in this section and for elastic solid in Section 6.5. To show the basis qualities, we use synthetic noisy data instead of real data in order to avoid potential issues linked to data generation from experimental tests. For the sake of simplicity, we assume the material databases $\{\mathbb{E}_\alpha\}_{\alpha=1}^m$ for all material/integration points are the same in the following numerical study in Section 6.4 and 6.5, i.e. $\mathbb{E}_\alpha = \{\hat{\mathbf{s}}^i = (\hat{\boldsymbol{\varepsilon}}^i, \hat{\boldsymbol{\sigma}}^i) \in \mathbb{R}^{2s}\}_{i=1}^P$.

For truss structures, the material behavior of the e -th bar member is characterized by a simple relation between the uniaxial strain $\boldsymbol{\varepsilon}_\alpha$ and uniaxial stress $\boldsymbol{\sigma}_\alpha$. As such, the local state

vector is defined as $\mathbf{s}_\alpha = [\varepsilon_\alpha \ \sigma_\alpha]^T \in \mathbb{R}^2$, i.e. $s=1$, and the associated norm to measure the distances of local states is defined as

$$\|\mathbf{s}_\alpha\| = \left(\frac{1}{2} M_\alpha \varepsilon_\alpha^2 + \frac{1}{2} M_\alpha^{-1} \sigma_\alpha^2 \right)^{1/2}, \quad (5.63)$$

where M_α is the a certain positive constant analogous to the Young's modulus of the linear reference material.

6.4.1 Example I: One-dimensional truss

As discussed in Remark 6.1, the main issues of data-driven computing appear when dealing with irregular material data that exhibits noise and outliers. To exemplify these pathologies, a basic one-bar truss problem is used for the first numerical study, as schematized in Figure 6.2, where a force of 10 kN is applied at the end and the cross-section area is $A = 200 \text{ cm}^2$.

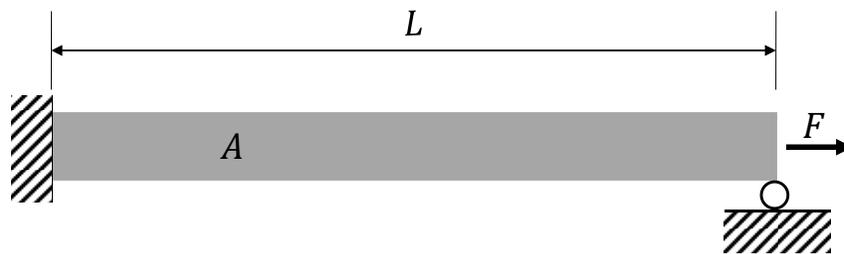


Figure 6.2: One-bar truss structure under uniaxial load.

Database with different level of random noise or outliers

We first consider two different databases $\mathbb{E}_\alpha = \{(\hat{\varepsilon}^i, \hat{\sigma}^i)\}_{i=1}^P$ with mild Gaussian random noise ($\chi = 0.05$) and strong Gaussian random noise ($\chi = 0.15$), as shown in Figure 6.3 and Figure 6.4, respectively. Data generation is as follows

$$\begin{aligned}\hat{\varepsilon}^i &= \bar{\varepsilon}^i + \mathcal{N}(0, \chi \bar{\varepsilon}_{\max}), \quad i = 1, \dots, P, \\ \hat{\sigma}^i &= \bar{\sigma}^i + \mathcal{N}(0, \chi \bar{\sigma}_{\max}), \quad i = 1, \dots, P,\end{aligned}\tag{5.64}$$

where $\bar{\varepsilon}_{\max} = 0.01$, $\bar{\varepsilon}^i \sim \mathcal{U}(-\bar{\varepsilon}_{\max}, \bar{\varepsilon}_{\max})$, $\bar{\sigma}^i = M_\alpha \bar{\varepsilon}^i$, $\bar{\sigma}_{\max} = M_\alpha \bar{\varepsilon}_{\max}$, and the coefficient constant $M_\alpha = 100$ MPa is used for data-driven solvers. The external force is incrementally loaded via five equal steps, such that there are five incremental data-driven solutions (i.e. 5 consecutive time steps for the local state $s_\alpha = [\varepsilon_\alpha \ \sigma_\alpha]^T$) as depicted by the star points in the result plots. And the associated 5 optimum data solutions (i.e. $\hat{s}_\alpha^* = [\hat{\varepsilon}_\alpha^* \ \hat{\sigma}_\alpha^*]^T$) are denoted by the square points.

As shown in Figure 6.3, both DMDD and LCDD can converge to the final equilibrium state point at the physical manifold of $\sigma = 0.5$ Mpa where it is close to the given data set (denoted by the circle points) in the (ε, σ) phase space. Although the randomness is mild, DMDD seems to converge to a local minimum. However, LCDD allows the data-driven solution to attain the intersection of the physical manifold and the conjectural material graph (see Figure 6.3(b)), which is a preferable solution. It suggests that LCDD is a more robust scheme to capture the local data structure and overcome local-minimum wells.

When the given material database presents much stronger randomness, as shown in Figure 6.4, DMDD is severely sensitive to the noisy data as the incremental solutions are susceptibly arrived at some local minimum points (Figure 6.4(a)). It is shown in Figure 6.4(a)

that DMDD converges to a suboptimal solution outside the data set, while we can observe from this 1D case that there exist some better data solutions locating at the physical manifold that lead to less distances or lower material costs. Again, LCDD yields more desirable incremental results (see Figure 6.4(b)), where the local state solutions move consistently to an intersection of the equilibrium manifold and the underlying material submanifold as the applied load linearly increases.

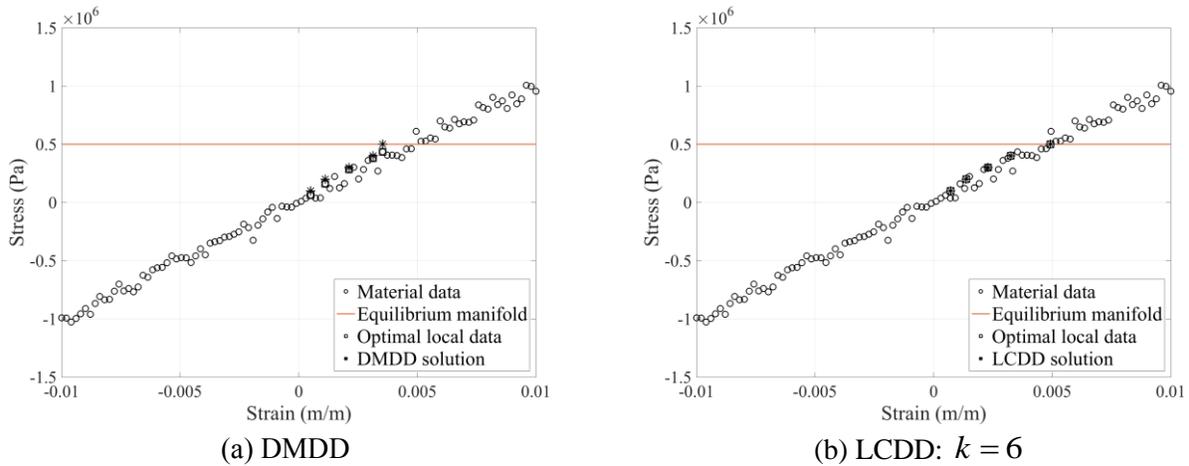


Figure 6.3: Comparison of the DMDD and LCDD solvers for database with mild noise ($P = 100$, $\chi = 0.05$).

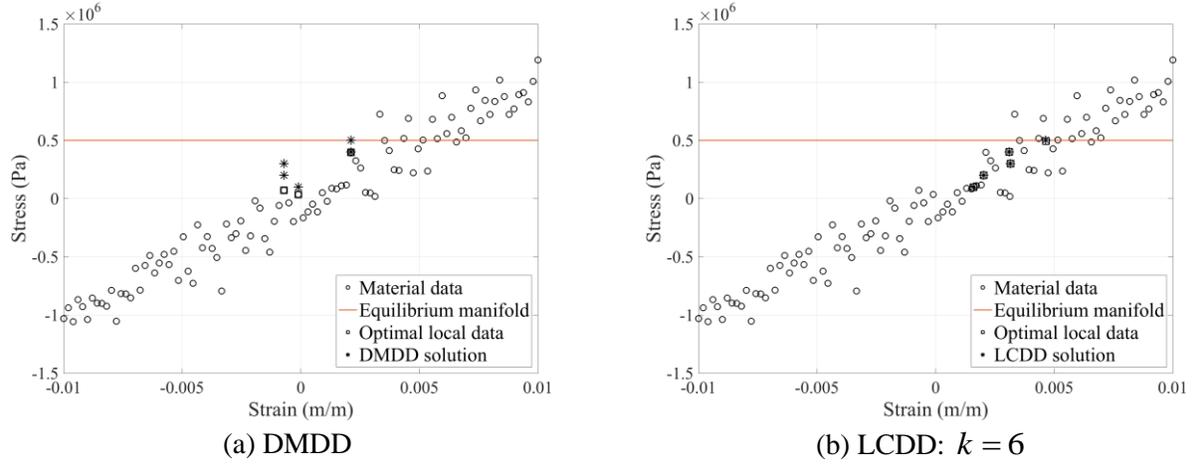
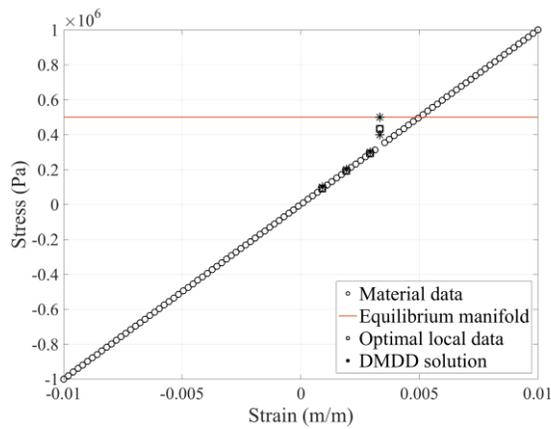
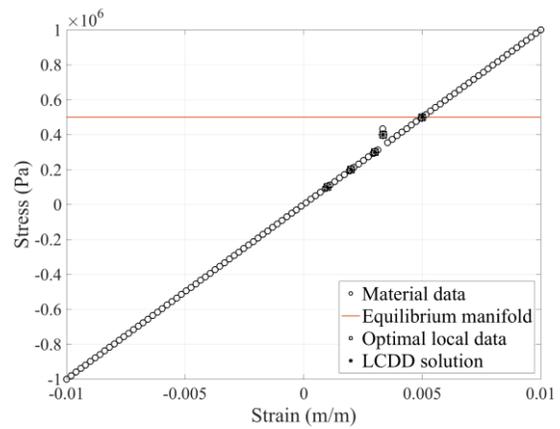


Figure 6.4: Comparison of the DMDD and LCDD solvers for database with strong noise ($P=100$, $\chi=0.15$).

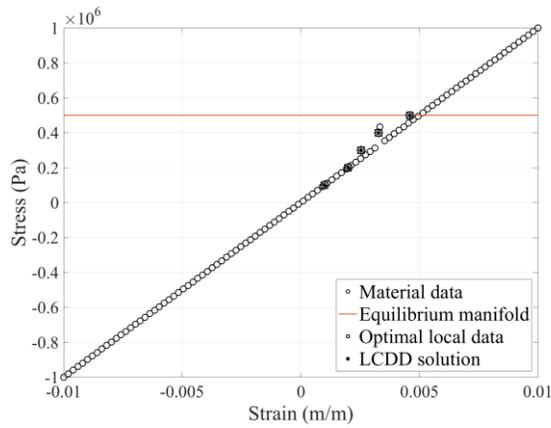
The comparison of results in Figure 6.3 and Figure 6.4 shows that LCDD is robust against large noise and can yield similar pattern of convergence for material databases under different random level, whereas the solution path of DMDD is strongly affected by the strength and distribution of noise. It is also observed that the optimum data solutions $(\hat{\varepsilon}_\alpha^*, \hat{\sigma}_\alpha^*)$ obtained from LCDD usually overlap with the data-driven solutions $(\varepsilon_\alpha, \sigma_\alpha)$ even though no experiment data in \mathbb{E}_α is exactly at those locations. It indicates that LCDD implicitly constructs a local material graph (the convex hull) associated to the k -NN points and allows the graph to be the solution space (interpolation) for searching the optimum local data points. This unique feature provides not only more robustness against noise due to clustering analysis, but also the reproducibility to a locally linear manifold if the data is well sampled, which will be further discussed in following examples.



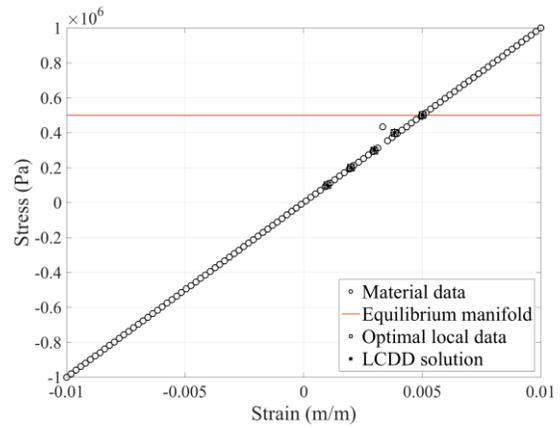
(a) DMDD



(b) LCDD: $k = 6$



(c) LCDD: $k = 12$



(d) LCDD: $k = 6, \bar{\mu} = 0.01$

Figure 6.5: Comparison of the DMDD and LCDD solvers for database with an outlier ($P = 100$).

Now consider a material database, as shown in Figure 6.5, approximating to an underlying linear graph, i.e. $F(\hat{\varepsilon}^i, \hat{\sigma}^i) = \hat{\sigma}^i - M_\alpha \hat{\varepsilon}^i = 0$ with a slope of $M_\alpha = 100$ Mpa, except for an outlier positioned near to the equilibrium manifold. The data-driven results show that the DMDD solutions are strongly misled by the outlier (Figure 6.5(a)), whereas the LCDD solutions converge to a reasonable location at the material manifold (Figure 6.5(b)). The LCDD solver is also tested under different procedure parameters and two typical results are presented in Figure 6.5(c) and (d). As shown in Figure 6.5(c), the data-driven solutions using more neighbor points

($k=12$) appears to be slightly influenced by the outlier during iterations, it remains robust against the outlier and yields consistent solutions as the one using $k=6$. This robustness is gained from the clustering analysis in LCDD based on multiple data points that prevents the dominance of outliers. The study of the effect of the regularization coefficient $\bar{\mu}$ on LCDD is also shown in Figure 6.5(d). It shows that with larger $\bar{\mu}$ the optimum data solution constructed within the associated convex hull favors the region with higher data density due to the stronger penalization of the large reconstruction weights (refer to (5.61)). Thus, increasing the regularization coefficient can enhance the robustness. There are many other regularization methods or robust penalty functions that allow further suppressing the influence of noise or outliers, e.g. the Huber penalty function. But the investigation of regularization is out of the scope of the current paper and interested readers are encouraged to consult the reference (Hastie, Tibshirani & Friedman 2009).

Database associated with a nonlinearly elastic material

To show another pathology caused by the discrete nature of data set for distance-minimizing approaches, consider a “nonlinear” database \mathbb{E}_α generated by a sigmoid function. As shown in Figure 6.6, the plot of data points in phase space transits from a nearly linear stage of slope $M_\alpha = 100$ Mpa to a plateau of $\sigma = 0.51$ Mpa, approximating an uniaxial perfect plasticity behavior.

Although no noise is presented, DMDD still converges to a suboptimal solution (see Figure 6.6(a)). It indicates a limitation of DMDD that the data-driven solver purely uses the discrete information from database. Since the solution is close to the flat plateau of the material

graph, the corresponding gradient descent (linearly scaled with the coefficient matrix M_α) associated with the data-driven iterations approaches to a direction normal to the plateau and becomes very stiff. Hence, the resulting displacement adjustment in (5.51) is not sufficient to drive the local solver to find the neighbor material data points, and the DMDD solver would just converge to this suboptimal location. This issue is attributed to the non-continuous nature of discrete data, resulting in a susceptibility to the selection of weighted coefficient M_α and the metric norm used to measure distance in phase space, and the density and the underlying structure of data.

Nevertheless, LCDD converges to a preferable solution (Figure 6.6(b)) where the physical and material manifolds intersect. This because the locally linear approximation provided by the LCDD solver can learn the data structure and implicitly construct a smooth constitutive submanifold (i.e. the convex envelop) based on the cluster of discrete material data, allowing the iterative scheme to find the data-driven solution in accordance with the classical solution using the corresponding material model. It is rooted in the manifold learning technique introduced in (Roweis & Saul 2000). Thus, LCDD is able to reproduce a locally linear material model based on the sampled data points. It should be emphasized that this reproducibility is very attractive in dealing with higher-dimensional phase space (such as elasticity problems in Section 6.5) when a limited amount of data is available.

Moreover, the proposed LCDD approach also overcomes the non-convergence issue that usually appears in regression based data-driven methods (Ibañez et al. 2016; Kanno 2017). This is because the introduction of local convexity prevents the solution space of searching optimum local data from being over-large.

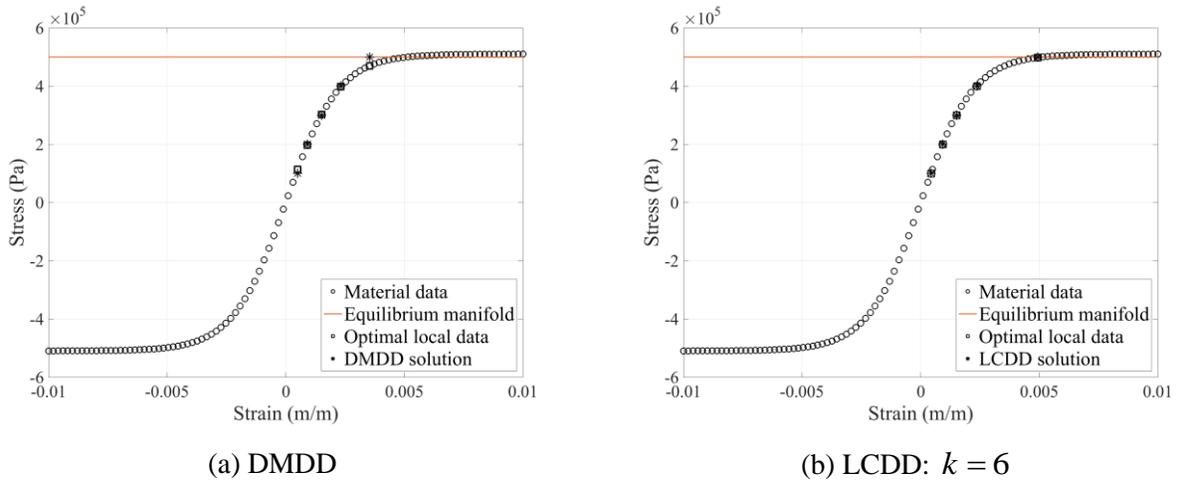


Figure 6.6: Comparison of the DMDD and LCDD solvers for noiseless sigmoid database ($P=100$).

6.4.2 Example II: Truss system

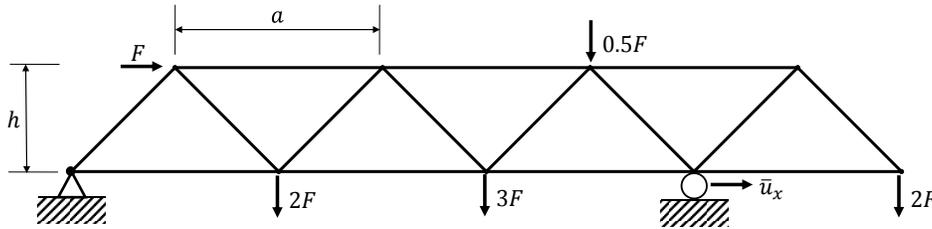


Figure 6.7: Schematic of a 15-bar truss structure with essential boundary conditions and externally applied loads, where $a=4$ m, $h=2$ m, $\bar{u}_x=0.01$ m, and $F=100$ kN.

To examine the convergence behavior with respect to the material data set size, we consider a 15-bar truss structure ($m=15$), illustrated in Figure 6.7, with unity cross-sectional area. A displacement $\bar{u}_x=0.01$ m is applied at one support such that the structure is indeterminate. The solution obtained from different data-driven solver (DMDD or LCDD) are compared against the referent solution under the following normalized root-mean-square (%RMS) state errors

$$\varepsilon_{(\% \text{RMS})} = \frac{1}{\varepsilon_{\max}^{\text{ref}}} \left(\frac{1}{m} \sum_{\alpha=1}^m l_{\alpha} (\varepsilon_{\alpha} - \varepsilon_{\alpha}^{\text{ref}})^2 \right)^{1/2}, \quad (65a)$$

$$\sigma_{(\% \text{RMS})} = \frac{1}{\sigma_{\max}^{\text{ref}}} \left(\frac{1}{m} \sum_{\alpha=1}^m l_{\alpha} (\sigma_{\alpha} - \sigma_{\alpha}^{\text{ref}})^2 \right)^{1/2}, \quad (32b)$$

where $\{l_{\alpha}\}_{\alpha=1}^m$ are the length of the bars, $\{(\varepsilon_{\alpha}, \sigma_{\alpha})\}_{\alpha=1}^m$ are the data-driven solutions for all bar members, $\{(\varepsilon_{\alpha}^{\text{ref}}, \sigma_{\alpha}^{\text{ref}})\}_{\alpha=1}^m$ are the strain and stress reference solutions solved by FEM with the corresponding synthetic material law, and $(\varepsilon_{\max}^{\text{ref}}, \sigma_{\max}^{\text{ref}})$ are the largest absolute values of strain and stress among all bar members.

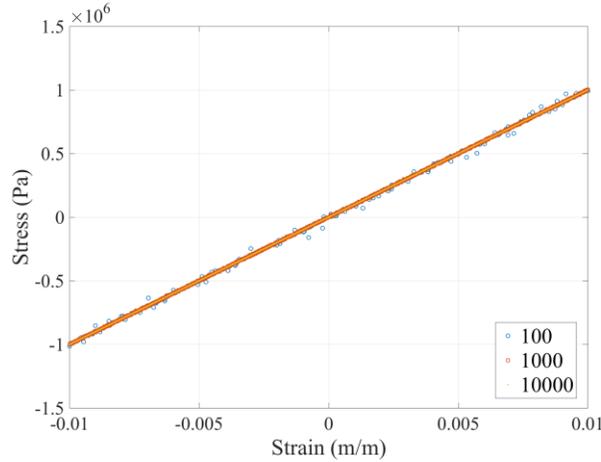


Figure 6.8: Three noisy material databases with different data sizes $P=10^2, 10^3, 10^4$ for truss members.

Three material data sets with different sizes (i.e. $P=10^2, 10^3, 10^4$) used for numerical tests are shown in Figure 6.8. The data sets uniformly converge to a linear curve with a slope of $M_{\alpha} = 100 \text{ Mpa}$, but include random noise in inverse proportion to the data set size P . The same Gaussian noise generation in (5.64) with $\chi = 2P^{-1}$ is used to perturb the data.

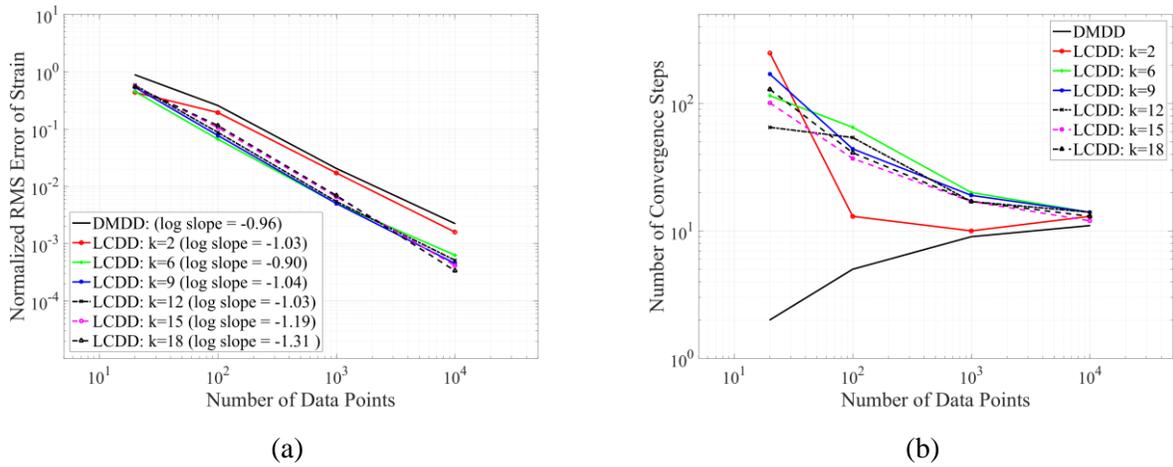


Figure 6.9: Truss structure case. (a) Convergence plot of the normalized RMS strain error $\varepsilon_{(\%RMS)}$ against increasing number of sampling points. (b) The number of convergence steps against increasing number of sampling points.

The convergence results of different data-driven solvers in terms of the normalized RMS strain error $\varepsilon_{(\%RMS)}$ are shown in Figure 6.9(a), where the data-driven solutions converge towards the classical solution with a rate close to 1 in accordance with the estimate in (Conti, Müller & Ortiz 2018; Kirchdoerfer & Ortiz 2016). Although it has been shown that DMDD can converge to the classical solution when the data set approximates a limiting constitutive law with increasing fidelity, the accuracy of DMDD is less than satisfaction. However, it shows that the solutions of LCDD are generally better with up to nearly 1 order accuracy. It is also observed that the LCDD solver is stable over a range of neighborhood sizes, i.e. from $k=6$ to $k=18$, obtaining the solutions under a similar level of accuracy (Figure 6.9(a)) and converged in a similar amount of steps (Figure 6.9(b)). Surprisingly, the analysis in Figure 6.9(b) suggests that the LCDD solutions converge in fewer steps when the data set size increases as the random noise

reduces. This is significantly distinct from the DMDD solver (Kirchdoerfer & Ortiz 2016) and the max-ent data-driven solver (Kirchdoerfer & Ortiz 2017) that require more iterations for larger data sets. It is due to the capability of local manifold learning of LCDD that allows a better representation of the underlying manifold if the data points are well sampled. Moreover, when using $k = 2$, the intrinsic dimension of the data sets, LCDD appears to lose high accuracy and yields solutions approaching to that from DMDD (shown in Figure 6.9(a)), implying that LCDD would recover DMDD in the limit of using few neighbors.

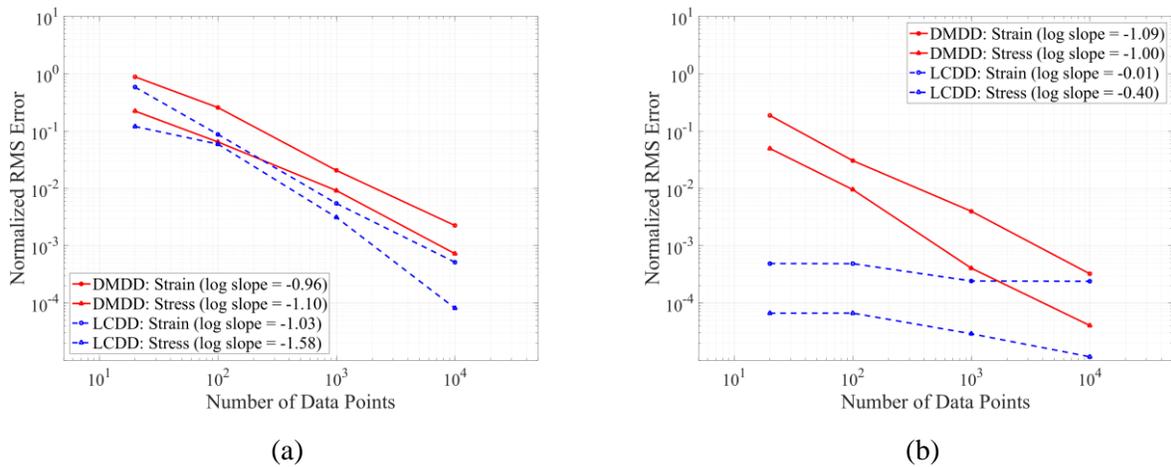


Figure 6.10: Truss structure case. Convergence plot of the normalized RMS errors, $\varepsilon_{(\%RMS)}$ and $\sigma_{(\%RMS)}$, against increasing number of sampling points for data-driven solvers using (a) the data sets with Gaussian random noise and (b) the corresponding noiseless data sets.

To show the effect of noise on data-driven solutions, the convergence results of data-driven solvers using the corresponding noiseless data sets (i.e. no random perturbation applied on the uniformly distributed data points that are given by $\hat{\sigma}^i = M_\alpha \hat{\varepsilon}^i$) are provided in Figure 6.10(b) for comparison. For noiseless databases, DMDD still yields an approximately linear convergence

under the same level of accuracy as for noisy databases in Figure 6.9(a). Differently, LCDD (with $k = 12$) exactly approximate, up to the level of the convergence tolerance, to the classical FEM solutions irrespective of the data set size, suggesting the reproducibility of locally linear model due to manifold learning. As shown in Figure 6.10(b), by only using 20 data points LCDD obtains a desirable solution as accurate as that of DMDD using 10^4 data points. We argue that this inherent manifold learning ability also contributes to the improved accuracy against noisy data (Figure 6.10(a)) when using LCDD.

A close comparison between the data-driven solutions of DMDD and LCDD using the data set with 100 material points (see Figure 6.8) are conducted in Figure 6.11, where the reference solution computed by FEM are also plotted in the (ϵ, σ) phase space. Since the FEM solver utilizes the synthetic linear model ($\sigma = M_\alpha \hat{\epsilon}$), the reference solutions (red triangles) of the 15 bar members locate exactly at the synthetic graph (dashed line) in the phase space. As can be seen from Figure 6.11(a), the variations of the noisy data set substantially influence the DMDD performance as several local state solutions (the star points in a dashed box) are badly sought such that deviate from the linear graph. They converge to some local minima caused by data noise, resulting in a poor overall performance of DMDD that is also suggested in the convergence study. By contrast, the LCDD solver overcomes such issues and the solutions agree well with the references as indicated in Figure 6.11(b). It demonstrates that the proposed data-driven solver can robustly capture the underlying data manifold against the superimposed noise and attain preferable solutions close to the material manifold.

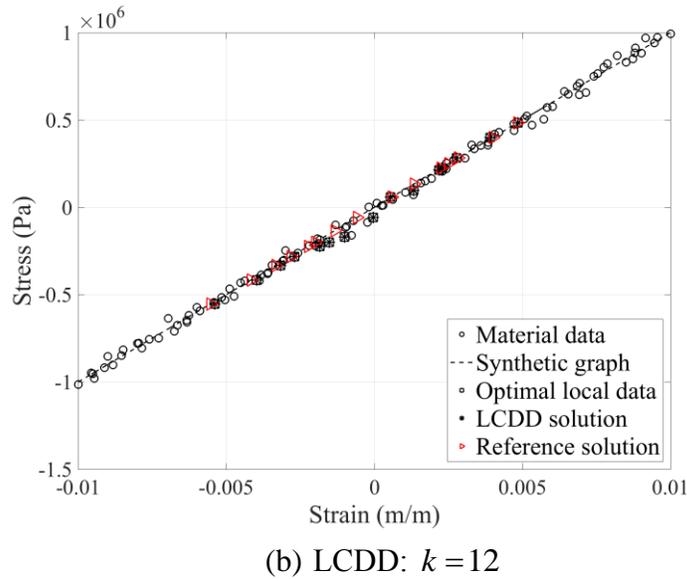
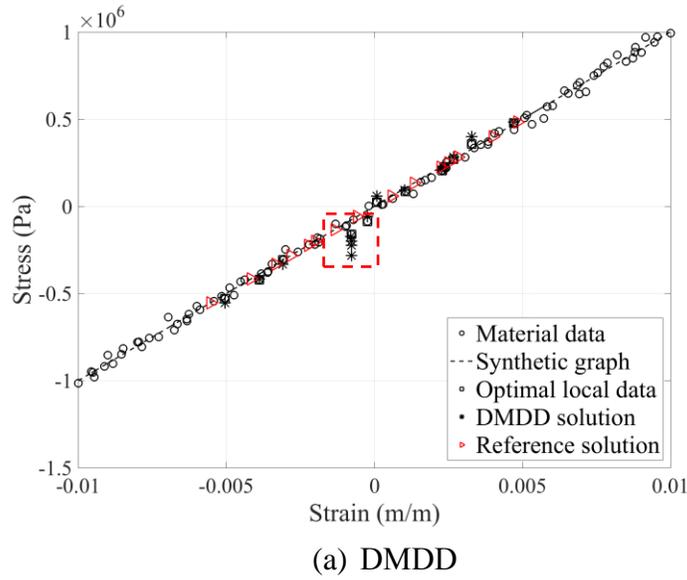


Figure 6.11: Truss structure case. Plots of the data-driven solutions of (a) DMDD and (b) LCDD when using the noisy material data set ($P = 100$).

6.5 Application to Elasticity Problems

In this section, we test a two-dimensional elasticity problem ($2s = 6$). The dimensionality of the associated phase space is high enough to start raising the “curse of dimensionality”, which

challenges the accuracy and robustness of data-driven solvers in measuring and approximating local states under such high-dimensional space while the material data is relatively sparse.

6.5.1 Discretization for elasticity problem

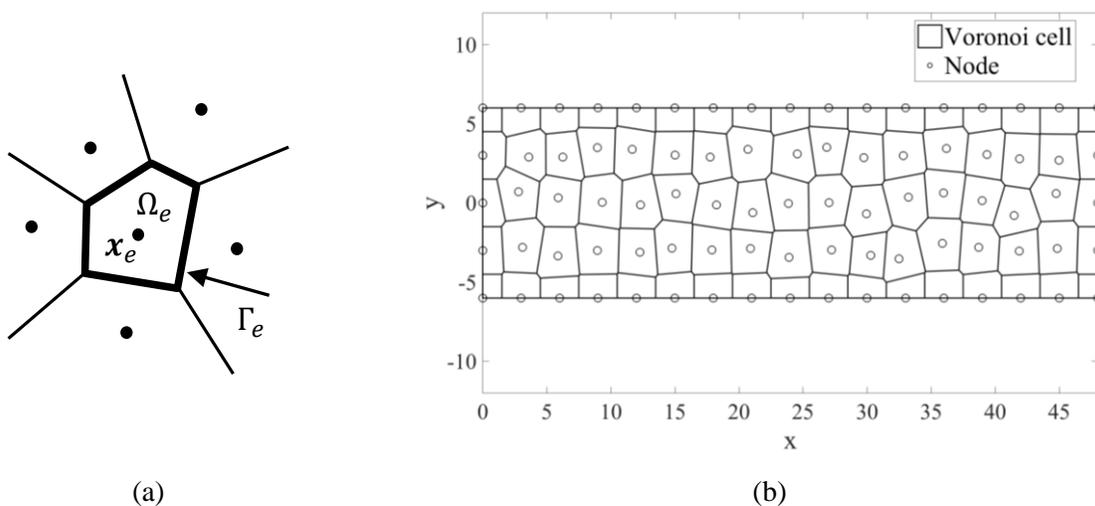


Figure 6.12: Schematic of (a) Voronoi diagram and (b) discretization of a beam model.

In this study, we employ a meshfree method (See Appendix A) for the field approximation, where the displacement is approximated by reproducing kernel (RK) shape functions (Liu, Jun & Zhang 1995), and a stabilized conforming nodal integration (SCNI) approach (Chen, Wu & Yoon 2001) is used for the integration of the weak form (5.49). Note that other numerical methods such as FEM for boundary value problems of continuum mechanics can also be applied. We apply the particle-based meshfree method is due to the simple nodal representation of both displacement and state field variables, i.e. $N = m$. As such, the continuum domain is partitioned by a Voronoi diagram (see Figure 6.12), and both the nodal displacement vectors $\{\mathbf{u}_I\}_{I=1}^N = \{\mathbf{u}(\mathbf{x}_I)\}_{I=1}^N$ and the state variables $\{(\boldsymbol{\varepsilon}_\alpha, \boldsymbol{\sigma}_\alpha)\}_{\alpha=1}^N = \{(\boldsymbol{\varepsilon}(\mathbf{x}_\alpha), \boldsymbol{\sigma}(\mathbf{x}_\alpha))\}_{\alpha=1}^N$ are carried

by the set of nodes at $\{\mathbf{x}_\alpha\}_{\alpha=1}^N$. Besides, the weight V_α in (5.50) is just the volume of each Voronoi cells Ω_α under this integration scheme.

It should be emphasized that the data-driven formulations (5.49) can be applied for nonlinear-elastic solid directly, whereas for finite strain problems the compatibility constraint and the strain-displacement matrix need modifications.

6.5.2 Example III: Elastic beam subjected to a shear load

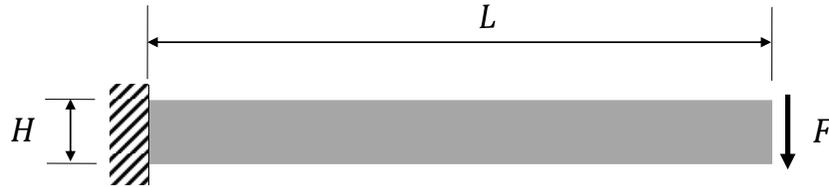


Figure 6.13: Schematic of a beam model subjected to a shear load, where $L=48$ m, $H=12$ m and $F=1000$ N.

We consider a plane stress beam model composed of homogeneous elastic solid material and subjected to a shear load, as given in Figure 6.13. The synthetic elastic material model is given by the classical Hook's law in two dimensional solid with Young's modulus $E=30$ Mpa and Poisson's ratio $\nu=0.3$. Therefore, the coefficient matrix used for norm (5.52) at each integration point is defined as

$$\mathbf{M}_\alpha = \mathbf{M}_\alpha^\varepsilon = \mathbf{M}_\alpha^{\sigma^{-1}} = \frac{E}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & (1-\nu)/2 \end{bmatrix}. \quad (5.66)$$

To evaluate the performance of data-driven solvers, the following normalized root-mean-square (%RMS) state error is re-defined for high-dimensional state (i.e. $s = 3$ in this problem),

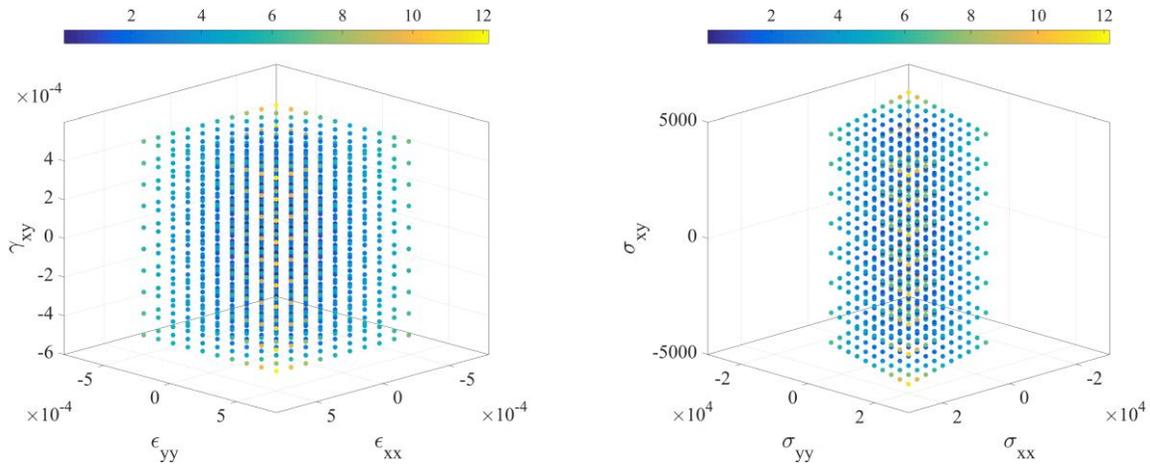
$$\omega_{(\% \text{RMS})} = \left(\frac{\sum_{\alpha=1}^N V_{\alpha} \|\mathbf{s}_{\alpha} - \mathbf{s}_{\alpha}^{\text{ref}}\|^2}{\sum_{\alpha=1}^N V_{\alpha} \|\mathbf{s}_{\alpha}^{\text{ref}}\|^2} \right)^{\frac{1}{2}}, \quad (5.67)$$

where $\mathbf{s}_{\alpha}^{\text{ref}} = [\boldsymbol{\varepsilon}_{\alpha}^{\text{ref}T} \boldsymbol{\sigma}_{\alpha}^{\text{ref}T}]^T$ denotes the nodal strain and stress reference solutions solved by the meshfree method (Chen, Wu & Yoon 2001) using the synthetic material model, $\mathbf{s}_{\alpha} = [\boldsymbol{\varepsilon}_{\alpha}^T \boldsymbol{\sigma}_{\alpha}^T]^T$ denotes the solutions solved by data-driven solvers using a given material data set. The discretization shown in Figure 6.12(b) is employed for the following numerical study.

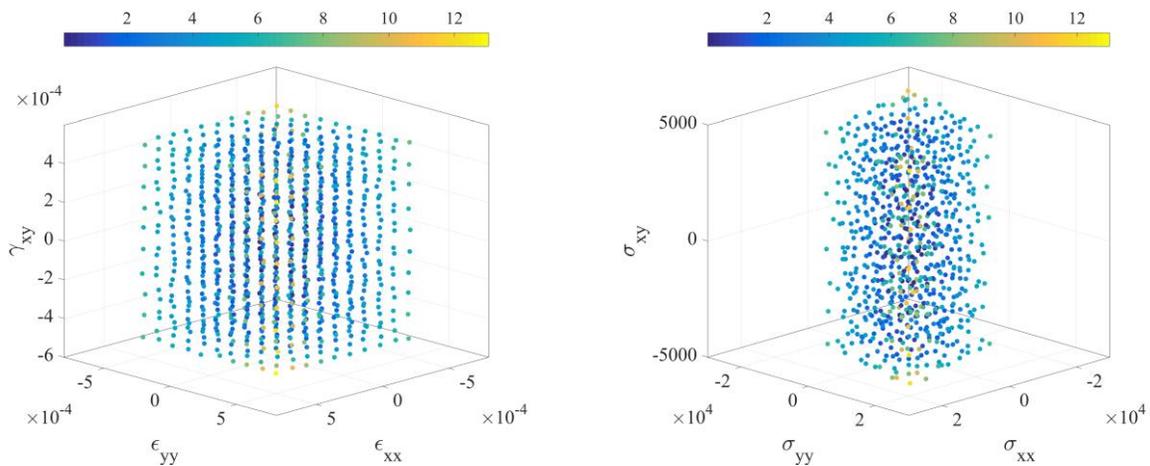
The procedures of the noiseless and noisy data generation are given as follows. First, the uniformly distributed strain vectors $\bar{\boldsymbol{\varepsilon}}^i = [\bar{\varepsilon}_{xx}^i \ \bar{\varepsilon}_{yy}^i \ \bar{\gamma}_{xy}^i]^T$, $i = 1, \dots, P$, are generated, e.g. the case in Figure 6.14(a) where 10 points are evenly distributed along each axis. The associate noiseless stress vector $\bar{\boldsymbol{\sigma}}^i = [\bar{\sigma}_{xx}^i \ \bar{\sigma}_{yy}^i \ \bar{\sigma}_{xy}^i]^T$ is obtained by left-multiplying the strain vector $\bar{\boldsymbol{\varepsilon}}^i$ by the coefficient matrix \mathbf{M}_{α} in (5.66). Then, the correlated strain and stress vectors are paired as $\bar{\mathbf{s}}^i = [\bar{\boldsymbol{\varepsilon}}^{iT} \ \bar{\boldsymbol{\sigma}}^{iT}]^T$ and stored into a noiseless data set $\mathbb{E}_{\alpha}^{\text{noiseless}}$. To generate the noisy data points $\{\hat{\mathbf{s}}^i\}_{i=1}^P$ in a data set $\mathbb{E}_{\alpha}^{\text{noisy}}$, the uniformly distributed random perturbations scaled by a factor $0.4/\sqrt[3]{P} \bar{s}_{\max}$ are independently superimposed on each component of the noiseless data points $\{\bar{\mathbf{s}}^i\}_{i=1}^P$, where \bar{s}_{\max} is the maximum value for each component among the data set $\{\bar{\mathbf{s}}^i\}_{i=1}^P$.

Four material data sets in various size (i.e. $P = 10^3, 20^3, 40^3, 80^3$) are considered for the beam model. For illustration of the randomness, one of the data sets with $P = 10^3$ points is

provided in Figure 6.14(b) as an example, where the strain components and the stress components of data points are plotted separately in two phase spaces for the sake of visualization.



(a) Noiseless data points $\mathbb{E}_\alpha^{noiseless}$: strain components (left); stress components (right)



(b) Noisy data points $\mathbb{E}_\alpha^{noisy}$: strain components (left); stress components (right)

Figure 6.14: Exemplary material data sets ($P = 10^3$) used for data-driven solvers. (a) Noiseless data set; (b) Data set with random noise.

The performance of the data-driven solvers using the noiseless data sets and the noisy data sets shown in Figure 6.15 and Figure 6.16, respectively, is demonstrated herein. Again, consistent to the convergence estimate in (Conti, Müller & Ortiz 2018; Kirchdoerfer & Ortiz 2016), the DMDD solutions converge linearly to the reference solution against the cubic root of the number of data points, regardless of using noiseless or noisy databases. On the other hand, LCDD (e.g. with $k = 6, 9$) using noiseless data sets yields almost accurate solutions (Figure 6.15(a)) up to the error tolerance defined in the iteration scheme. It implies that LCDD perfectly capture the underlying linear material graph even in such a high dimensional phase space. The convergence study with noisy data sets (Figure 6.16(a)) shows that the LCDD solution using the sparse data set ($P = 10^3$) is able to achieve higher accuracy than the DMDD solution using very dense data set ($P = 80^3$), suggesting the superiority of LCDD over DMDD. Considering the difficulty to obtain a database with sufficiently dense data for high-dimensional spaces, the proposed LCDD approach appears to be even more attractive when solving mechanics problems and overcoming the potential “curse of dimensionality”. As it is known that the intrinsic dimensionality of a linear elastic database is $d = 2$, it is interesting to observe from Figure 6.15(a) and Figure 6.16(a) that the LCDD solutions obtained by using $k = 3$ ($d < k < 2s$) present a blending phenomenon between the DMDD solution (i.e. $k = 1$) and the other LCDD solutions (when using $k \geq 2s = 6$). It indicates that one need to use large enough neighborhood to fully preserve the manifold learning capacity in LCDD.

The associated analyses of number of convergence steps for the data-driven solvers are also presented in Figure 6.15(b) and Figure 6.16(b), respectively. In contrast to data size that dramatically increases with a cubic power due to $s = 3$, the number of iterations for the

convergence of the LCDD solver does not evidently increase. Moreover, the comparison of the LCDD results in Figure 6.15(b) and Figure 6.16(b) shows that only few more convergence steps are needed for data-driven computing when applying the noisy data sets than the noiseless data sets. The above results suggest that the convergence property of LCDD is not sensitive to the database size as well as the data sampling quality, and thus, it has a good scalability in terms of increasing data set. Nevertheless, the iteration step to attain convergence by DMDD substantially changes in accordance to the size of data sets, and the solutions are much worse than the counterparts obtained from the LCDD solver.

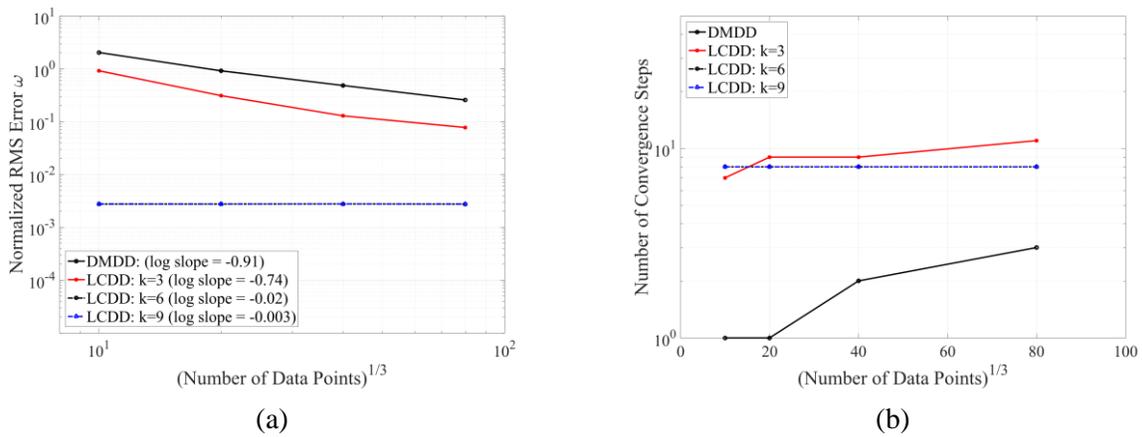


Figure 6.15: Shear beam model with noiseless data sets. (a) Convergence plot of the normalized RMS state error $\omega_{(\% \text{RMS})}$ against increasing number of sampling points. (b) The number of convergence steps against increasing number of sampling points.

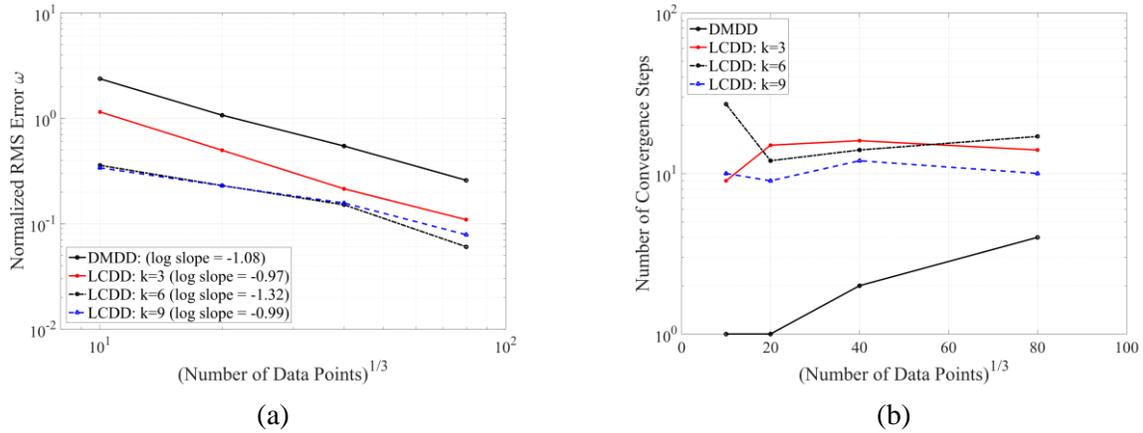
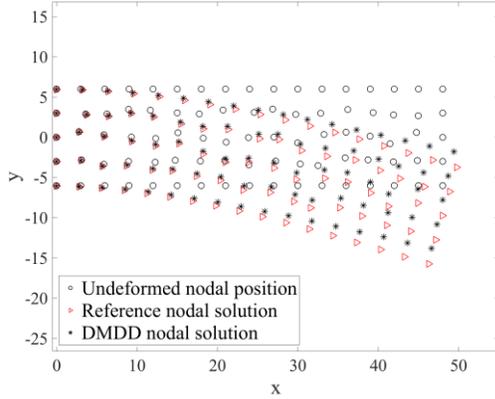
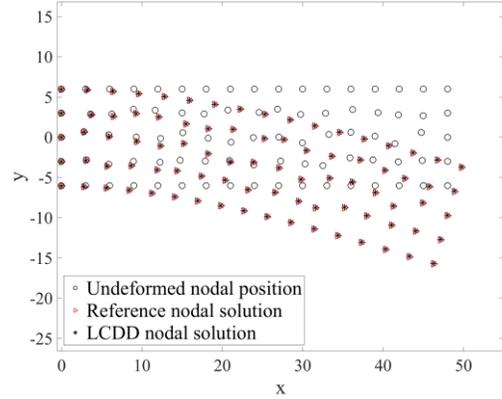


Figure 6.16: Shear beam model with noisy data sets. (a) Convergence plot of the normalized RMS state error $\omega_{(\%RMS)}$ against increasing number of sampling points. (b) The number of convergence steps against increasing number of sampling points.

The displacement solutions computed by the data-driven solvers are also compared in Figure 6.17 and Figure 6.18. In this case, DMDD performs poorly (Figure 6.17(a) and Figure 6.18(a)) due to the susceptibility to noise and local minimum issues as pointed out in Section 6.4. The nature of high-dimensional phase space in elasticity problems aggravates these issues. However, the LCDD nodal displacement solutions, with $k = 6$ for the k -NN as an example, seem to exactly agree with the reference solutions when using noiseless database (Figure 6.17(b)), while only show negligibly deviations from the reference when using the noisy database (Figure 6.18(b)). Figure 6.19 presents the stress component (σ_{xx} and σ_{xy}) results of the LCDD solution and the reference, showing LCDD can obtain desirable local state solutions across the problem domain for a noisy database. It demonstrates that the LCDD approach remains robust against noise for solving elasticity problems.

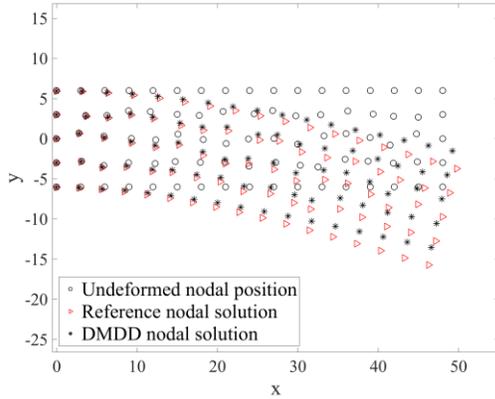


(a)

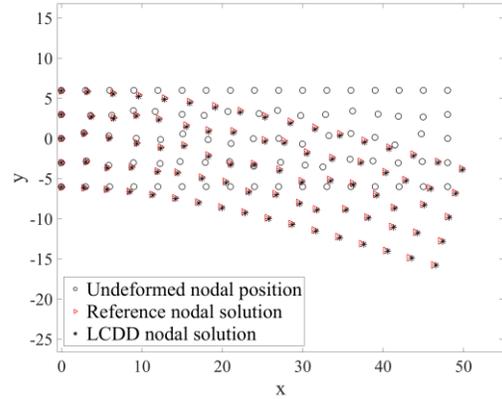


(b)

Figure 6.17: Shear beam model. (a) DMDD displacement solution using the noiseless data set $P = 80^3$. (b) LCDD displacement solution ($k = 6$) using the noiseless data set $P = 10^3$.



(a)



(b)

Figure 6.18: Shear beam model. (a) DMDD displacement solution using the noisy data set $P = 80^3$. (b) LCDD displacement solution ($k = 6$) using the noisy data set $P = 10^3$.

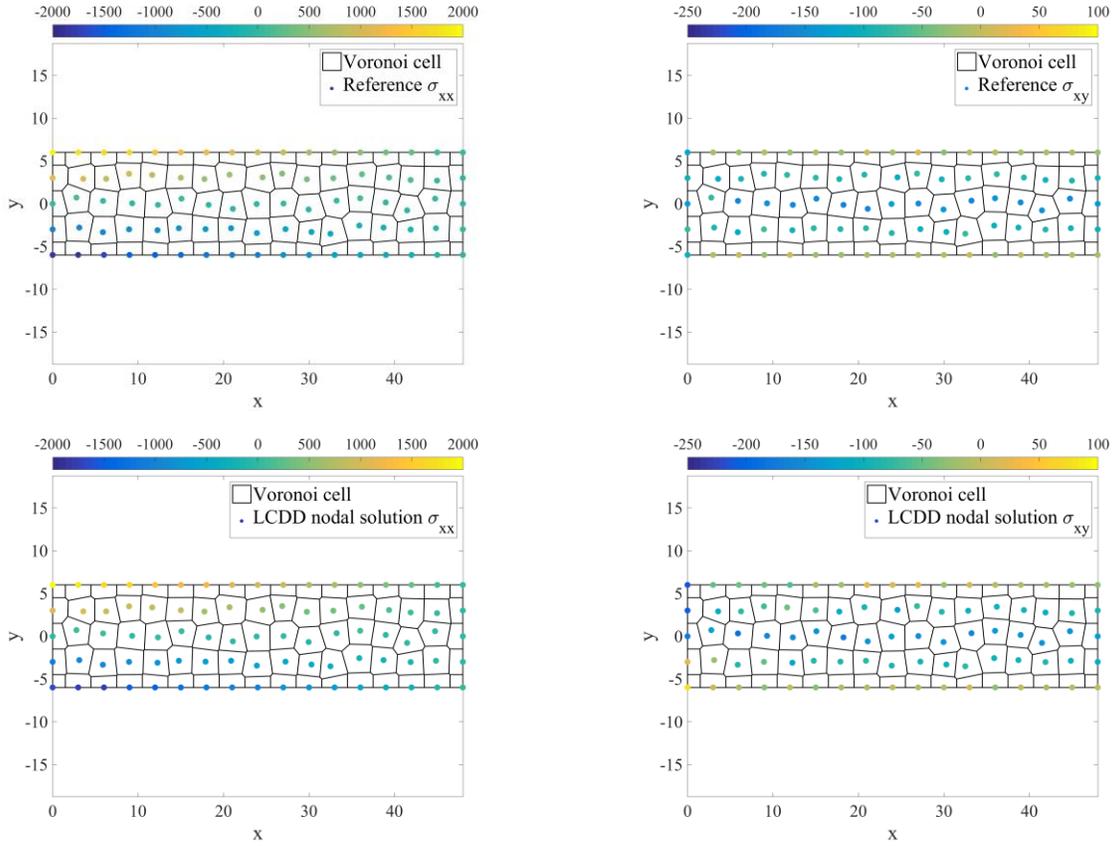


Figure 6.19: Shear beam model. The comparison of σ_{xx} and σ_{xy} stress components between the reference solution and the LCDD solutions with $k = 6$ using the noisy data set $P = 10^3$.

6.6 Summary

We have formulated a locally convex data-driven (LCDD) solver, a new data-driven computing paradigm integrated with manifold learning techniques that generalizes the distance-minimizing data-driven (DMDD) computing proposed in (Kirchdoerfer & Ortiz 2016), in order to enhance robustness and accuracy of data-driven solvers against noise/outliers. The proposed method selects some nearest data points adaptively and searches for optimum data solutions from a bounded solution space constructed by a convex combination of the selected k -NN data points. This local data searching procedure has been related to a standard *non-negative least squares*

(NNLS) solver and can be solved efficiently in few iterations. By means of the clustering analysis based on k -NN and the reconstruction of convex envelop for data points, LCDD guarantees enhanced robustness and convergence stability.

From the *pure data-driven approach* point of view (refer to (Ayensa-Jiménez et al. 2018)), LCDD is inspired by measuring the distance to a nearest neighbor convex set \mathcal{E}_e instead of a discrete data set \mathbb{E}_e , aim to enhance the robustness against noise and prevent undesirable local minima. From the *fitted data-driven (or linearization) approach* point of view (Ayensa-Jiménez et al. 2018), on the other hand, LCDD relies on the approximation of locally linear material graph by using manifold learning methodologies (Lee & Verleysen 2007; Saul & Roweis 2003) in order to capture the global structure via local data information, but it takes the additional convexity condition on the material graph into account. LCDD is distinguished from the other manifold learning based data-driven approach (Ibañez et al. 2016) in that the iteration process of LCDD neither explicitly uses the constitutive manifold nor suffers from any convergence issues due to non-convexity. Lastly, it is easy to show that LCDD can be reduced to the standard DMDD approach (Kirchdoerfer & Ortiz 2016) when using only one neighbor data point, i.e. $k = 1$.

It has also been shown that the embedded NNLS solver essentially seeks for the projection point of a given computed state onto a nearby material graph that is implicitly constructed based on the selected neighbor points, allowing achieving the reproducibility of locally linear approximation. Hence, LCDD not just achieve significantly improved robustness and accuracy when dealing with noisy database, but also yield nearly exact approximation to the numerical reference solution when the database is well sampled for a constitutive relation that shows locally linear behavior. The proposed method has been developed for truss problems and

continuum elasticity problems in order to demonstrate the effectiveness of LCDD in terms of robustness, convergence, accuracy, and the generalization to non-trivial computational mechanics cases with high-dimensional phase space.

This paper has been intended to introduce manifold learning techniques, or dimensionality reduction (Lee & Verleysen 2007), into the data-driven computing framework while remaining a simple computational structure as (Kirchdoerfer & Ortiz 2016). Our numerical studies show that it is very crucial to possess manifold learning capability when dealing with high-dimensional database, because the “curse of dimensionality” makes data points extremely sparse in high-dimensional spaces and the acquisition of sufficient data is not practical. Thus, it demands dimensionality reduction to identify and extract the essential information from database. The elasticity example in Section 6.5 demonstrates that LCDD really outperforms DMDD due to the inherent manifold learning.

In general, we envision that the proposed data-driven framework could be most useful in cases where one would like to learn from noisy experimental data and a governing equation is known. Since the present work is a preliminary study, there are a number of possible enhancements as well as worthwhile applications that ask for further research. For example, some other robust techniques or advanced clustering algorithms (Vidal, Ma & Sastry 2016) to suppress outliers that are well established in machine learning area can be applied under the proposed LCDD framework. The way to identify the optimal number of neighbors in the k -NN algorithm is still an open question (Anava & Levy 2016). Moreover, to consider uncertain or stochastic nature of experimental data, reliability based quantities (Ayensa-Jiménez et al. 2018) that reflect the statistics of data variations need to be taken into account. Due to the merit of manifold learning that implicitly constructs the local tangent space of the underlying material

manifold, it seems promising to extend the proposed method for nonlinear elasticity (Nguyen & Keip 2018) under finite deformation. And the effectiveness of LCDD for high-dimensional data may contribute to the inverse problems (Ibañez et al. 2017; Leygue et al. 2018) that aims to identify the strain-stress relation of nonlinear solid by data-driven computing without any underlying constitutive model.

6.7 Acknowledgement

This chapter, in part, is currently being prepared for submission for publication of the material as it may appear in “He, Qizhi; Chen, Jiun-Shyan. A data-driven computational simulation based on locally convex reconstruction for noisy database”. The dissertation author was the primary investigator and author of this material.

Chapter 7

Nonlinear Model Order Reduction Based on Meshfree Nodal Integration Framework

In this chapter nonlinear model order reduction, which combines POD and a system approximation approach, is developed under a meshfree nodal integration framework for hyperelasticity.

7.1 Introduction

Many simulations in computational science and engineering are parameterized input-output problems formulated under a state space. For example, consider a computational homogenization problem of a heterogeneous material, the inputs are the parameters, e.g. the boundary conditions and the elasticity coefficients of the material, the state is the displacement field, and the output of interest is the macroscopic material response. Although the nonlinear model order reduction based on the POD-Galerkin method together with the interpolation-type hyper-reduction (e.g., DEIM) has been widely applied to fluid systems (Chaturantabut & Sorensen 2011; Hochman, Bond & White 2011), its application in the field of nonlinear solid mechanics is quite rare (Radermacher & Reese 2016).

To the best of the author's knowledge, most nonlinear model order reduction of static solid systems are based on FEM (Ghavamian, Tiso & Simone 2017; Radermacher & Reese

2016; Ryckelynck & Missoum Benziane 2016) while the meshfree approximation-based applications remains unexplored. The aim of this section is to apply the POD-DEIM based nonlinear MOR method to solid mechanics with nonlinearities under the meshfree framework proposed for hyperelasticity (nonlinear elasticity) and to study its performance and efficiency of the proposed method. Meshfree approximation offers higher-order continuity, and generally for smooth systems it is effective to construct the reduced order models (Haller & Ponsioen 2017; Szalai 2015). This can also be conceptually explained by the implicit function theorem that a system of equations with a mild requirement on the partial derivatives locally defines a graph of a function.

In this section, we first introduce the meshfree Galerkin formulation based on the stabilized conforming nodal integration (SCNI) scheme, which has been demonstrated effective in many engineering applications (Chen et al. 1996, 1997; Chen, Wu & Yoon 2001; J S Chen et al. 2000; Jiun Shyan Chen et al. 2000). Next, we introduce the procedures of applying nonlinear MOR (POD-DEIM) to quasi-static large deformation analysis based on meshfree framework. Last, we test the performance of POD-DEIM by modeling a two-phase hyperelastic solid with perturbed loading conditions.

7.2 Meshfree Galerkin Formulation for Hyperelasticity

Total Lagrangian formulation is usually adopted for simulation of biological material characterized by hyperelastic models. Let the d -dimensional problem domain in the undeformed configuration be denoted by Ω_x , with the corresponding displacement boundary and traction boundary denoted by Γ_x^g and Γ_x^h , respectively. Let $\mathcal{V}_u = \{\mathbf{v} \in H^1(\Omega)^d \mid \mathbf{v} = \mathbf{g} \text{ on } \Gamma_x^g\}$ be the

kinematically admissible trial space, where \mathbf{g} is the displacement enforced on Γ_X^g , and \mathcal{V}_0 be the associated test space. The displacement-based Galerkin weak form for solving quasi-static problems is given by: find $\mathbf{u} \in \mathcal{V}_u$ such that

$$\int_{\Omega_x} \delta F_{ij} P_{ij} d\Omega = \int_{\Omega_x} \delta u_i b_i d\Omega + \int_{\Gamma_x^h} \delta u_i h_i d\Gamma, \text{ for } \delta \mathbf{u} \in \mathcal{V}_0, \quad (7.1)$$

where P_{ij} is the first Piola-Kirchhoff stress deriving from the strain energy density function $W(\mathbf{u})$, i.e. $P_{ij} = dW(\mathbf{u})/dF_{ij}$, $F_{ij} = \delta_{ij} + \partial u_i / \partial X_j$ is the deformation gradient, b_i is the body force per unit undeformed volume, and h_i is the prescribed traction on the natural boundary Γ_X^h . Due to the geometric or material nonlinearities, Newton's method is introduced and the solution to the nonlinear equation (7.1) is obtained by iteratively solving a sequence of linear models. Let n and v denote n -th load step and the v -th iteration step, respectively, the linearized equation of (7.1) is given by:

$$\int_{\Omega_x} \delta F_{ij} (C_{ijkl})_{n+1}^v \Delta F_{kl} d\Omega = \int_{\Omega_x} \delta u_i (b_i)_{n+1} d\Omega + \int_{\Gamma_x^h} \delta u_i (h_i)_{n+1} d\Gamma - \int_{\Omega_x} \delta F_{ij} (P_{ij})_{n+1}^v d\Omega, \quad (7.2)$$

where C_{ijkl} is the first elasticity tensor. The displacement is iteratively updated by $(\mathbf{u}_i)_{n+1}^{v+1} = (\mathbf{u}_i)_{n+1}^v + \Delta \mathbf{u}_i$ until convergence.

To avoid the complexity related to meshing in FE method, the meshfree modeling technique based on a reproducing kernel particle method (RKPM) (Chen et al. 1996; Liu, Jun & Zhang 1995) is used for biological material modeling, for example, using images directly from CT scans (Basava et al. 2014; Chen et al. 2016). Moreover, the application of RK approximation allows for a smooth transition of material properties without exhibiting sharp discontinuities at the interfaces between different material components in the muscle (Basava et al. 2014; Chen et

al. 2016), which is consistent to the observations in many studies (Tidball 1984) that the change of material properties from skeletal muscle to tendon is a smooth transition. To this end, the displacement, its variation and the incremental displacement are approximated by the RK shape functions as follows

$$u_i(\mathbf{x}) = \sum_{I=1}^N \Psi_I(\mathbf{x}) d_{ii}, \quad \delta u_i(\mathbf{x}) = \sum_{I=1}^N \Psi_I(\mathbf{x}) \delta d_{ii}, \quad \Delta u_i(\mathbf{x}) = \sum_{I=1}^N \Psi_I(\mathbf{x}) \delta d_{ii}, \quad (7.3)$$

where N denotes the number of RK nodes, and the construction of RK shape functions Ψ_I is presented in Appendix A. The number of total DOFs of the system is denoted as $\mathcal{N} = dN$.

With the RK approximation for displacement solutions given in (7.3) the Galerkin weak form in (7.1) leads to the following discrete vector equation

$$\mathbf{f}^{\text{int}} = \mathbf{f}^{\text{ext}}, \quad (7.4)$$

where $\mathbf{f}^{\text{int}}, \mathbf{f}^{\text{ext}} \in \mathbb{R}^{\mathcal{N}}$ are expressed as

$$\begin{aligned} \mathbf{f}_I^{\text{int}} &= \int_{\Omega_x} \mathbf{B}_I^T \mathbf{P} d\Omega, \\ \mathbf{f}_I^{\text{ext}} &= \int_{\Omega_x} N_I \mathbf{b} d\Omega + \int_{\Gamma_x^h} N_I \mathbf{h} d\Gamma, \end{aligned} \quad (7.5)$$

where \mathbf{B}_I^T is the gradient matrix associated to the incremental deformation gradient in vector form, i.e.

$$\Delta \mathbf{F} = \sum_{I=1}^N \mathbf{B}_I \Delta \mathbf{d}_I, \quad (7.6)$$

and $N_I = \Psi_I \mathbf{I}_{d \times d}$. Similarly, the linearized equation (7.2) yields

$$\mathbf{K}(\mathbf{d}_{n+1}^v) \Delta \mathbf{d} = \mathbf{f}_{n+1}^{\text{ext}} - \mathbf{f}^{\text{int}}(\mathbf{d}_{n+1}^v), \quad (7.7)$$

where

$$\begin{aligned}
\mathbf{K}_{IJ}(\mathbf{d}_{n+1}^v) &= \int_{\Omega_X} \mathbf{B}_I^T \mathbf{C}(\mathbf{d}_{n+1}^v) \mathbf{B}_J d\Omega, \\
\mathbf{f}_I^{\text{int}}(\mathbf{d}_{n+1}^v) &= \int_{\Omega_X} \mathbf{B}_I^T \mathbf{P}(\mathbf{d}_{n+1}^v) d\Omega, \\
(\mathbf{f}_I^{\text{ext}})_{n+1} &= \int_{\Omega_X} N_I \mathbf{b}_{n+1} d\Omega + \int_{\Gamma_X^b} N_I \mathbf{h}_{n+1} d\Gamma.
\end{aligned} \tag{7.8}$$

\mathbf{C} is a matrix form corresponding to the elasticity tensor C_{ijkl} , and $\mathbf{d}_{n+1}^v \in \mathbb{R}^{\mathcal{N}}$ is the vector of nodal displacements at $(n+1)$ -th time step and v -th Newton iteration step. We refer to (Chen et al. 1996, 1997, 2016; J S Chen et al. 2000; Jiun Shyan Chen et al. 2000) for more details of the nonlinear formulations.

7.2.1 Stabilized conforming nodal integration (SCNI)

The method of stabilized conforming nodal integration (SCNI) (Chen, Yoon & Wu 2002; Chen, Wu & Yoon 2001) is used for integrating the discrete equation in (7.7) to achieve computational efficiency and accuracy. Based on the smoothed derivative of the approximation function together with a nodal integration scheme, SCNI ensures the linear consistency of the meshfree Galerkin approximation and reduces greatly computational cost under its numerical integration scheme compared to the conventional Gauss quadrature rule. The key idea of SCNI is to approximate the deformation gradient in (7.6), or other derivatives of displacement, by using smoothed derivative instead of direct derivative, which is given by

$$\bar{\nabla} u^h(\mathbf{x}_L) = \frac{1}{A_L} \int_{\Omega_L} \nabla u^h d\Omega = \frac{1}{A_L} \int_{\partial\Omega_L} u^h \mathbf{n} d\Gamma, \tag{7.9}$$

where $A_L = \int_{\Omega_L} d\Omega$ is the volume of a conforming smoothing domain associated to the node \mathbf{x}_L .

As shown in Figure 7.1, the continuum domain Ω_X is partitioned into N conforming cells by

Voronoi diagram, and both the nodal displacement vectors and the state variables (e.g., stress, strain) are carried by the set of nodes at $\{\mathbf{x}_L\}_{L=1}^N$. Since the smoothed derivative computation in (7.9) in conjunction with the Voronoi discretization, as shown in Figure 7.1, naturally satisfies the so called integration constraint (Chen, Wu & Yoon 2001), the linear patch-test is passed and a quadratic rate of convergence is achieved for linear solid problems solved by meshfree Galerkin methods.

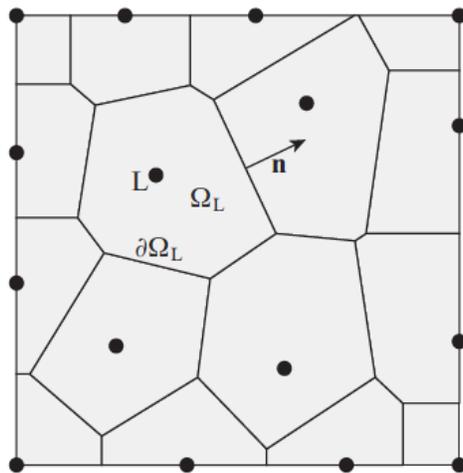


Figure 7.1 Illustration of Voronoi diagram for SCNI.

7.3 Nonlinear MOR Based on Meshfree Galerkin Formulation

This section will demonstrate how to apply nonlinear MOR for hyperelasticity problem. Based on the meshfree computational framework introduced in Section 7.2, the equilibrium of a generic parameterized quasi-static system of a solid is given by the following discrete nonlinear vector equation

$$\mathbf{R}(\mathbf{d}; \boldsymbol{\mu}) := \mathbf{f}^{\text{ext}}(\boldsymbol{\mu}) - \mathbf{f}^{\text{int}}(\mathbf{d}; \boldsymbol{\mu}) = \mathbf{0}, \quad (7.10)$$

where the parameters $\boldsymbol{\mu}$ can be any physical quantities the system depends on. Thus, the nodal displacement solution $\boldsymbol{d} \in \mathbb{R}^{\mathcal{N}}$ relies on the given input parameters $\boldsymbol{\mu}$, denoted as $\boldsymbol{d}(\boldsymbol{\mu})$.

To reduce the HDM system (7.10), the displacement \boldsymbol{u} as well as the nonlinear term $\boldsymbol{f}^{\text{int}}$ are approximated by a reduced order projection onto a low-dimensional subspaces

$$\boldsymbol{d}(\boldsymbol{\mu}) \approx \tilde{\boldsymbol{d}}(\boldsymbol{\mu}) = \boldsymbol{V}\boldsymbol{d}^r(\boldsymbol{\mu}), \quad (7.11)$$

and

$$\boldsymbol{f}^{\text{int}}(\boldsymbol{d}(\boldsymbol{\mu})) \approx \tilde{\boldsymbol{f}}^{\text{int}}(\boldsymbol{d}(\boldsymbol{\mu})) = \boldsymbol{Z}\boldsymbol{c}(\boldsymbol{\mu}), \quad (7.12)$$

where $\boldsymbol{V} \in \mathbb{R}^{\mathcal{N} \times k}$ consists of k basis vectors representing the displacement field, whereas $\boldsymbol{Z} \in \mathbb{R}^{\mathcal{N} \times \hat{k}}$ contains \hat{k} basis vectors characterizing the nonlinear term $\boldsymbol{f}^{\text{int}}$. The vectors $\boldsymbol{d}^r(\boldsymbol{\mu}) \in \mathbb{R}^k$ and $\boldsymbol{c}(\boldsymbol{\mu}) \in \mathbb{R}^{\hat{k}}$ are the low-dimensional unknowns to be determined in the reduced system. The POD approach (Section 3.3.1) is employed to construct \boldsymbol{V} and \boldsymbol{Z} , as shown in Section 3.4.2. To collect snapshot data, the problem (7.10) without any reduction is precomputed in terms of different parameters. The precomputations can be performed for different material parameters, geometry, or loading conditions. The snapshot matrices containing the precomputed displacement vectors and nonlinear vectors are denoted as \boldsymbol{X}_s and as \boldsymbol{X}_f , respectively.

Recall the classical *POD-Galerkin reduction model* in (3.8) using the reduced approximation of the displacements together with a Galerkin projection is given

$$\tilde{\boldsymbol{R}}(\boldsymbol{d}^r; \boldsymbol{\mu}) := \boldsymbol{V}^T \boldsymbol{f}^{\text{ext}}(\boldsymbol{\mu}) - \boldsymbol{V}^T \boldsymbol{f}^{\text{int}}(\boldsymbol{V}\boldsymbol{d}^r(\boldsymbol{\mu})) = \mathbf{0}. \quad (7.13)$$

Following the Newton-Raphson method in (7.7), the POD-Galerkin reduced system is solved by the iterative scheme as follows

$$\begin{aligned} \mathbf{V}^T \mathbf{K}(\tilde{\mathbf{d}}_{n+1}^v) \mathbf{V} \Delta \mathbf{d}^r &= \mathbf{V}^T \mathbf{f}_{n+1}^{\text{ext}}(\boldsymbol{\mu}) - \mathbf{V}^T \mathbf{f}^{\text{int}}(\tilde{\mathbf{d}}_{n+1}^v), \\ \tilde{\mathbf{d}}_{n+1}^{v+1} &= \tilde{\mathbf{d}}_{n+1}^v + \mathbf{V} \Delta \mathbf{d}^r. \end{aligned} \quad (7.14)$$

Although this POD-Galerkin reduced system only involves k unknowns, the computation of (7.14) depends still on the original dimension \mathcal{N} due to the lifting bottleneck discussed in Section 3.4. This limits the possible reduction in computational cost. To overcome this issue, the idea of DEIM or Gappy-POD (refer to Section 3.4.2) is introduced to approximate the nonlinear term (7.12) by means of an empirical interpolation.

To this end, an additional selection matrix $\mathbf{P} = [\mathbf{e}_{\varrho_1}, \dots, \mathbf{e}_{\varrho_n}] \in \mathbb{R}^{\mathcal{N} \times \hat{n}}$ is constructed offline by applying the Greedy algorithm to the POD matrix \mathbf{Z} . Recall Equation (3.21) or (3.25), the reduced approximation of nonlinear function in (7.12) can be further expressed as

$$\mathbf{f}^{\text{int}}(\mathbf{d}(\boldsymbol{\mu})) \approx \tilde{\mathbf{f}}^{\text{int}}(\mathbf{d}(\boldsymbol{\mu})) = \mathbf{Z}(\mathbf{P}^T \mathbf{Z})^{-1} \mathbf{P}^T \mathbf{f}^{\text{int}}(\mathbf{d}(\boldsymbol{\mu})). \quad (7.15)$$

This is called the DEIM approximation of the nonlinear term. For simplicity, we denote $\mathbf{M} = \mathbf{Z}(\mathbf{P}^T \mathbf{Z})^{-1} \in \mathbb{R}^{\mathcal{N} \times \hat{k}}$. Note that if the number of selected interpolation points is greater than the basis dimension in \mathbf{Z} , i.e. $\hat{n} > \hat{k}$, it is also referred to as ‘‘Gappy’’ POD, where $\mathbf{M} = \mathbf{Z}(\mathbf{P}^T \mathbf{Z})^\dagger \in \mathbb{R}^{\mathcal{N} \times \hat{n}}$. But we don’t distinguish these two ideas in this study.

By introducing the POD approximation of the displacement (7.11) and the DEIM approximation (7.15) into the nonlinear vector and using a Galerkin projection, we obtain a *POD-DEIM reduced system* of the HDM in (7.10):

$$\tilde{\mathbf{R}}(\mathbf{d}^r; \boldsymbol{\mu}) := \mathbf{V}^T \mathbf{f}^{\text{ext}}(\boldsymbol{\mu}) - \mathbf{V}^T \mathbf{M} \mathbf{P}^T \mathbf{f}^{\text{int}}(\mathbf{V} \mathbf{d}^r(\boldsymbol{\mu})) = \mathbf{0}, \quad (7.16)$$

where $\mathbf{V}^T \mathbf{M}$ is constant matrix and can be precomputed. The corresponding iterative scheme is given

$$\begin{aligned}
(\mathbf{V}^T \mathbf{M}) \mathbf{P}^T \mathbf{K}(\tilde{\mathbf{d}}_{n+1}^v) \mathbf{V} \Delta \mathbf{d}^r &= \mathbf{V}^T \mathbf{f}_{n+1}^{\text{ext}}(\boldsymbol{\mu}) - (\mathbf{V}^T \mathbf{M}) \mathbf{P}^T \mathbf{f}^{\text{int}}(\tilde{\mathbf{d}}_{n+1}^v), \\
\tilde{\mathbf{d}}_{n+1}^{v+1} &= \tilde{\mathbf{d}}_{n+1}^v + \mathbf{V} \Delta \mathbf{d}^r.
\end{aligned} \tag{7.17}$$

Due to the selection matrix \mathbf{P} , less evaluations are needed to assemble the reduced tangent stiffness matrix and reduced internal force vector compared with the POD approach.

It is noted that the DEIM approximation leads to a non-symmetric reduced stiffness matrix $(\mathbf{V}^T \mathbf{M}) \mathbf{P}^T \mathbf{K}(\tilde{\mathbf{d}}_{n+1}^v) \mathbf{V}$. Moreover, to avoid singular stiffness matrices, the unphysical restriction $\hat{k} > k$ would be needed. Besides, the interpolation point must not be less than the basis dimension, i.e. $\hat{n} \geq \hat{k}$, to prevent singular of \mathbf{M} . It has also been shown that using more sampling points than the basis order benefits the conditioning of reduced tangent, leading to a more stable reduced system (Cosimo, Cardona & Idelsohn 2014). A few studies (Hochman, Bond & White 2011; Radermacher & Reese 2016) have been conducted to enhance the stability of DEIM reduced system by adding a linearized stabilization term.

7.4 Numerical Example

Consider a representative volume element (RVE) of two-phase hyperelastic material described by St. Venant-Kirchhoff model, as depicted in Figure 7.2. The material parameters for the matrix and the inclusion are provided in Table 7.1. The matrix and inclusion are used to mimic connective tissues and muscle fibers that constitutes the cellular structure of skeletal muscle tissue (Zhang 2015). The RVE domain is discretized by Voronoi diagram and the computations are carried out by meshfree approximation with SCNI under a plane strain configuration. As shown in Figure 7.2, $N = 441$ nodes are used for domain discretization.

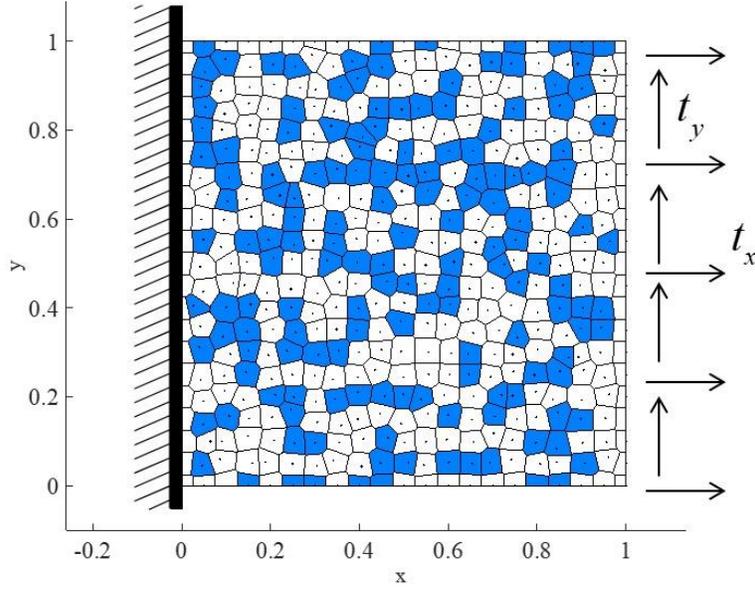


Figure 7.2 Meshfree RVE model of two-phase hyperelastic material. The integration cells of matrix material and inclusion material are depicted by white and blue, respectively.

Table 7.1: Material parameters for matrix and inclusion.

	Matrix	Inclusion
Young's Modulus (Pa)	$E_m = 3.0 \times 10^7$	$E_i = 1.5 \times 10^6$ Pa
Poisson's ratio	$\nu_m = 0.49$	$\nu_i = 0.4$

As mentioned in Section 7.3, a snapshot POD is used to construct the reduced-order basis for both the displacement field and the nonlinear term selected from the HDM system. Hence, a training set $\mathcal{D}_{\text{train}} \in \mathcal{D}$ is necessary for which the HDM simulation is performed to obtain snapshot data. In this case, we define

$$\begin{aligned} \boldsymbol{\mu} &= (t_x, t_y) \in \mathcal{D}, \\ \mathcal{D} &= [-1 \times 10^4, 1 \times 10^4] \times [-2 \times 10^5, 2 \times 10^5] \subset \mathbb{R}^2, \end{aligned} \tag{7.18}$$

where t_x and t_y denote the components of traction applied on the right edge of RVE. The training set $\mathcal{D}_{\text{train}}$ is constructed by randomly selecting 20 points from \mathcal{D} . The intermediate solutions which do not satisfy balanced state during the Newton-Raphson procedure can also be collected into the snapshot data to enrich the representation of the solution space.

Based on the training set the POD basis matrices, \mathbf{V} and \mathbf{Z} , used for the approximation of displacement (7.11) and nonlinear term (7.12) are obtained, and the distribution of the selected interpolation DOFs are depicted in Figure 7.3. As shown in Figure 7.3, the DEIM algorithm applied to SCNI based meshfree model selects the interpolation DOFs (denoted by red stars) and the neighboring integration cells (denoted by black circles) for assembling the reduced tangent stiffness matrix and the internal force vector. The ratio between the number of integration cells used for assembling the DEIM reduced system and the total number of integration cells is denoted as $r_{DEIM} = N_{DEIM}/N$. For this case, the $N_{DEIM}/N = 13\%$ as shown in Figure 7.3.

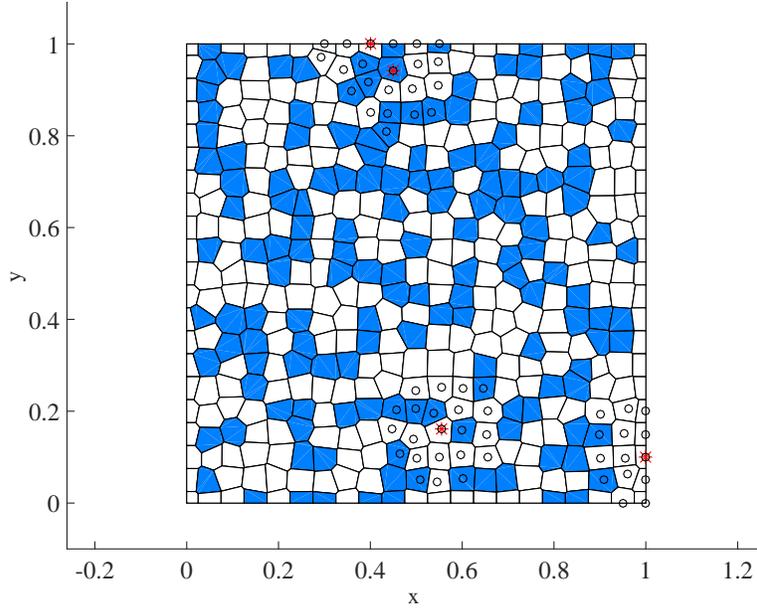


Figure 7.3 Interpolation points given by DEIM based on Meshfree model. The red star denotes the selected degree of freedoms and the black circle denotes the integration cell that contributes to the selected degree of freedoms.

The aforementioned POD-DEIM reduction in (7.17) and POD reduction in (7.14) are tested for various dimensionalities of \mathbf{V} and \mathbf{Z} using a test set $\mathcal{D}_{\text{test}}$ composed of 10 loading cases selected from \mathcal{D} , which is different to the set $\mathcal{D}_{\text{train}}$. Define an *average relative error* measure as

$$\varepsilon_r = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \frac{\|\mathbf{d}(\mu_i) - \tilde{\mathbf{d}}(\mu_i)\|}{\|\mathbf{d}(\mu_i)\|}, \quad (7.19)$$

where N_{test} denotes the total number of snapshots collected from the numerical test using $\mathcal{D}_{\text{test}}$ (for each loading case, multiple snapshots can be obtained if the load is applied incrementally), \mathbf{d} denotes the reference solution solved original high-dimensional system, and $\tilde{\mathbf{d}}$ denote the approximation obtained from the reduced order system by POD or POD-DEIM. Table 7.2 gives

the average relative errors of POD reduction and POD-DEIM reduction, and the associated ratio of CUP time normalized by the cost of HDM. It shows that increasing the dimensions of the POD basis, V and Z significantly improves the reduced-order approximation. Moreover, it is observed that the employment of DEIM further reduces the computational cost because less integration cells need to be evaluated. Since the cost reduction by DEIM is correlated to the ratio r_{DEIM} , the support size in meshfree approximation is also a factor affecting the performance.

Table 7.2: The accuracy and efficiency performance of POD and POD-DEIM reduction methods.

	$k = 2, \hat{k} = 6, N_{DEIM}/N = 13\%$		$k = 4, \hat{k} = 10, N_{DEIM}/N = 21\%$	
	Normalized CPU time	Average relative error	Normalized CPU time	Average relative error
POD	38.3%	1.52×10^{-1}	38.7%	3.62×10^{-2}
POD-DEIM	15.2%	1.97×10^{-1}	19.9%	8.93×10^{-2}

7.5 Summary

In this chapter, we focus on the development of nonlinear model reduction method based on a meshfree framework for hyperelasticity. In the numerical example, the meshfree based nonlinear model order reduction framework is used to simulation the two-phase hyperelastic material and satisfactory results have been obtained. We want to point out that this framework is not limited to DEIM. Since the SCNI based meshfree framework is carried out by nodal integration and both nodes and integration points are represented by a same set of points, the cubature-type hyper-reduction techniques are expected to suit well the meshfree framework as well. This is because these techniques rely on the selection of integration points instead of DOFs

of the state nodal vector. The unique feature of nodal integration offers convenience in constructing the hyper-reduction model. Thus, introducing meshfree approximation into cubature-type hyper-reduction techniques is a promising way for nonlinear model order reduction.

7.6 Acknowledgement

This chapter, in part, is currently being prepared for submission for publication of the material as it may appear in “He, Qizhi; Chen, Jiun-Shyan. Nonlinear model order reduction based on linearized manifold learning for meshfree analysis”. The dissertation author was the primary investigator and author of this material.

Chapter 8

Application of Data-Driven Modeling to Biological Tissues

In this chapter, the proposed robust, manifold learning-based data-driven solver is extended to finite deformation kinematics for biological tissues modeling via directly utilizing the material data. It is noted that the use of experimental data as a basis for developing constitutive models of biological materials is classical and remains important, and enormous progress has been made during recent years in the phenomenological modeling of biological tissues (Fung & Skalak 1981; Holzapfel, Gasser & Ogden 2000; Humphrey 2003; Zhou & Fung 1997). However, constitutive models are difficult to construct for complex biological material, and the employment of data-driven approach in minimizing material modeling empiricism and preventing the loss of experimental information offers an alternative for this class of problems.

8.1 Data-Driven Computational Framework for Nonlinear Solids

To consider large deformation response of biological material, we reformulate the data-driven computational framework based on finite strain theory. Let us consider the Green strain tensor \mathbf{E} and the 2nd-PK stress tensor \mathbf{S} as the state variables due to their symmetry. Given $(\hat{\mathbf{E}}, \hat{\mathbf{S}})$ as material data provide by the local step, the global step of data-driven solver in Section 6.2.3 is rewritten as

$$\begin{aligned}
& \min_{\mathbf{u}, \mathbf{S}} \int_{\Omega^x} \left(d^{\varepsilon^2}(\mathbf{E}[\mathbf{u}] - \hat{\mathbf{E}}) + d^{\sigma^2}(\mathbf{S} - \hat{\mathbf{S}}) \right) d\Omega, \\
& \text{subject to: } \quad \text{DIV}(\mathbf{F} \cdot \mathbf{S}) + \bar{\mathbf{b}} = \mathbf{0} \quad \text{in } \Omega^x, \\
& \quad \quad \quad \mathbf{F} \cdot \mathbf{S} \cdot \mathbf{N} = \bar{\mathbf{t}} \quad \text{on } \Gamma_t^x,
\end{aligned} \tag{8.1}$$

where $\bar{\mathbf{b}}$ and $\bar{\mathbf{t}}$, respectively, are the body force and the traction under initial configuration, and the superscript “X” denotes the initial configuration. Following the same procedures conducted in Chapter 6 for linear kinematics, the following Lagrangian is obtained by using Lagrange multipliers:

$$\begin{aligned}
\mathcal{L}_{DD}(\mathbf{u}, \mathbf{S}, \boldsymbol{\lambda}) &= \int_{\Omega^x} \left(d^{\varepsilon^2}(\mathbf{E}[\mathbf{u}] - \hat{\mathbf{E}}) + d^{\sigma^2}(\mathbf{S} - \hat{\mathbf{S}}) \right) d\Omega \\
&+ \int_{\Omega^x} \boldsymbol{\lambda} \cdot (\text{DIV}(\mathbf{F} \cdot \mathbf{S}) + \bar{\mathbf{b}}) d\Omega - \int_{\Gamma_t^x} \boldsymbol{\lambda} \cdot (\mathbf{F} \cdot \mathbf{S} \cdot \mathbf{N} - \bar{\mathbf{t}}) d\Gamma.
\end{aligned} \tag{8.2}$$

By means of the integration by parts, we obtain an equivalent form as follows

$$\begin{aligned}
\mathcal{L}_{DD}(\mathbf{u}, \mathbf{S}, \boldsymbol{\lambda}) &= \int_{\Omega^x} \left(d^{\varepsilon^2}(\mathbf{E}[\mathbf{u}] - \hat{\mathbf{E}}) + d^{\sigma^2}(\mathbf{S} - \hat{\mathbf{S}}) \right) d\Omega \\
&- \int_{\Omega^x} \nabla \boldsymbol{\lambda} : (\mathbf{F} \cdot \mathbf{S}) - \boldsymbol{\lambda} \cdot \bar{\mathbf{b}} d\Omega + \int_{\Gamma_t^x} \boldsymbol{\lambda} \cdot \bar{\mathbf{t}} d\Gamma + \int_{\Gamma_u^x} \boldsymbol{\lambda} \cdot (\mathbf{F} \cdot \mathbf{S} \cdot \mathbf{N}) d\Gamma.
\end{aligned} \tag{8.3}$$

For the sake of simplicity, we set $\boldsymbol{\lambda} = \mathbf{0}$ on Γ_u^x . The variational form of (8.3) becomes

$$\begin{aligned}
\delta \mathcal{L}_{DD}(\mathbf{u}, \mathbf{S}, \boldsymbol{\lambda}) &= \int_{\Omega} \left(\delta \mathbf{E}[\mathbf{u}] : \mathbf{M}^\varepsilon : (\mathbf{E}[\mathbf{u}] - \hat{\mathbf{E}}) + \delta \mathbf{S} : (\mathbf{M}^\sigma : (\mathbf{S} - \hat{\mathbf{S}})) \right) d\Omega \\
&- \int_{\Omega^x} \left(\nabla \boldsymbol{\lambda} : (\delta \mathbf{F}[\mathbf{u}] \cdot \mathbf{S}) + \delta \nabla \boldsymbol{\lambda} : (\mathbf{F}[\mathbf{u}] \cdot \mathbf{S}) + \nabla \boldsymbol{\lambda} : (\mathbf{F}[\mathbf{u}] \cdot \delta \mathbf{S}) \right) d\Omega \\
&+ \int_{\Omega^x} \delta \boldsymbol{\lambda} \cdot \bar{\mathbf{b}} d\Omega + \int_{\Gamma_t^x} \delta \boldsymbol{\lambda} \cdot \bar{\mathbf{t}} d\Gamma.
\end{aligned} \tag{8.4}$$

Then, the following weak form variational equations are obtained:

$$\delta_u \mathcal{L}_{DD} = 0 \Rightarrow \int_{\Omega^x} \delta \mathbf{E}[\mathbf{u}] : \mathbf{M}^\varepsilon : (\mathbf{E}[\mathbf{u}] - \hat{\mathbf{E}}) d\Omega = \int_{\Omega^x} (\delta \mathbf{F}^T[\mathbf{u}] \cdot \nabla \boldsymbol{\lambda}) : \mathbf{S} d\Omega, \tag{8.5a}$$

$$\delta_\lambda \mathcal{L}_{DD} = 0 \Rightarrow \int_{\Omega^x} \delta \nabla \boldsymbol{\lambda} : (\mathbf{F}[\mathbf{u}] \cdot \mathbf{S}) d\Omega = \int_{\Omega^x} \delta \boldsymbol{\lambda} \cdot \bar{\mathbf{b}} d\Omega + \int_{\Gamma_t^x} \delta \boldsymbol{\lambda} \cdot \bar{\mathbf{t}} d\Gamma := P^{ext}(\delta \boldsymbol{\lambda}), \tag{8.5b}$$

$$\delta_S \mathcal{L}_{DD} = 0 \Rightarrow \int_{\Omega^x} \delta \mathbf{S} : (\mathbf{M}^\sigma : \mathbf{S} - \mathbf{F}^T[\mathbf{u}] \cdot \nabla \lambda) d\Omega = \int_{\Omega^x} \delta \mathbf{S} : \mathbf{M}^\sigma : \hat{\mathbf{S}} d\Omega. \quad (8.5c)$$

Compared to (6.15) in Section 6.2.3, an additional term appears in the first equation of (8.5a) due to the nonlinear kinematics, while the Equations (8.5b) and (8.5c) characterizing equilibrium are similar except for using different strain and stress measures.

Again, if the stress approximation under numerical discretization is evaluated by either static condensation at the element level or nodal integration schemes, Equation (8.5c) can be disassembled into independent element-wise or point-wise systems given as

$$\mathbf{S} = \hat{\mathbf{S}} + \mathbf{M}^{\sigma^{-1}} : (\mathbf{F}^T[\mathbf{u}] \cdot \nabla \lambda), \quad (8.6)$$

By inserting the equation (8.6) into Equations ((5.49)a) and ((5.49)b), we obtain

$$\begin{aligned} J_u &:= \int_{\Omega^x} \delta \mathbf{E}[\mathbf{u}] : \mathbf{M}^\varepsilon : \mathbf{E}[\mathbf{u}] d\Omega - \int_{\Omega^x} (\delta \mathbf{F}^T[\mathbf{u}] \cdot \nabla \lambda) : \mathbf{M}^{\sigma^{-1}} : (\mathbf{F}^T[\mathbf{u}] \cdot \nabla \lambda) d\Omega \\ &\quad - \int_{\Omega^x} \delta \mathbf{E}[\mathbf{u}] : \mathbf{M}^\varepsilon : \hat{\mathbf{E}} d\Omega - \int_{\Omega^x} (\delta \mathbf{F}^T[\mathbf{u}] \cdot \nabla \lambda) : \hat{\mathbf{S}} d\Omega = 0, \\ J_\lambda &:= \int_{\Omega^x} (\mathbf{F}^T[\mathbf{u}] \cdot \delta \nabla \lambda) : \mathbf{M}^{\sigma^{-1}} : (\mathbf{F}^T[\mathbf{u}] \cdot \nabla \lambda) d\Omega + \int_{\Omega^x} (\mathbf{F}^T[\mathbf{u}] \cdot \delta \nabla \lambda) : \hat{\mathbf{S}} d\Omega = P^{ext}(\delta \lambda). \end{aligned} \quad (8.7)$$

In contrast to the data-driven system in (6.16), the weak form system in (8.7) need to be solved by Newton's method.

Once the convergent state solutions (\mathbf{E}, \mathbf{S}) are solved from (8.7) by a iterative scheme, the local step of the data-driven solver in (6.11) is readily used to update the material data.

$$(\hat{\mathbf{E}}_\alpha^*, \hat{\mathbf{S}}_\alpha^*) = \arg \min_{(\mathbf{E}_\alpha, \mathbf{S}_\alpha) \in \mathcal{E}_\alpha} d^{\varepsilon^2}(\mathbf{E}_\alpha, \hat{\mathbf{E}}_\alpha) + d^{\sigma^2}(\mathbf{S}_\alpha, \hat{\mathbf{S}}_\alpha), \quad \alpha = 1, \dots, m, \quad (8.8)$$

where the material solution space is denoted as \mathcal{E}_α , suggesting that the proposed Locally Convex Data-Driven (LCDD) solver can be applied. Similar to Chapter 6, the coefficient matrices

$\mathbf{M}^\varepsilon \in \mathbb{R}^{s \times s}$ and $\mathbf{M}^\sigma \in \mathbb{R}^{s \times s}$ (usually $\mathbf{M}^\sigma = \mathbf{M}^{\varepsilon^{-1}}$) are used in the metrics, d^ε and d^σ , respectively, to measure distance.

8.2 Numerical Examples: Data-Driven Modeling for Nonlinear Solids

In this section, two demonstration examples are used to examine the capability and reliability of the data-driven computational framework for modelling nonlinear solids. Since the focus of the present work is to verify the proposed mathematical structure, in this section we generate the material database for data-driven solvers through the phenomenological constitutive models that are used to compute the reference solutions. We will then employ the experimental data for the data-driven modeling of biological materials in the next section.

8.2.1 Large deflection of a cantilever beam subjected to a tip shear load

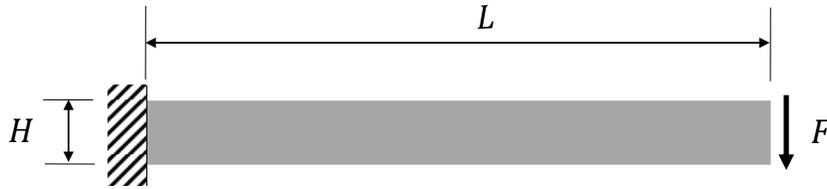


Figure 8.1: Schematic of a beam model subjected to a shear load, where $L = 20$ mm, $H = 1$ mm .

Consider a cantilever beam model subjected to a shear load (Chen, Yoon & Wu 2002), as shown in Figure 8.1. The beam is made of an elastic material (Saint Venant-Kirchhoff model), where Young's modulus $E = 4.8 \times 10^3$ N/mm² and Poisson's ratio $\nu = 0.0$. A 41×5 randomly spaced nodes are used to discretise the domain. The analysis is carried out using 10 equal loading

steps under plane strain condition, and the model-based reference solution to this problem can be computed by using the meshfree method introduced in Section 7.1. On the other hand, the coefficient matrix at each integration point is defined as

$$\mathbf{M}_\alpha = \mathbf{M}_\alpha^\varepsilon = \mathbf{M}_\alpha^{\sigma^{-1}} = E \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}. \quad (8.9)$$

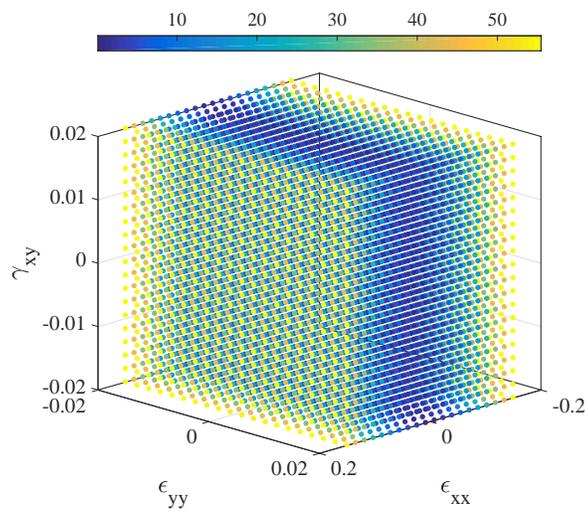


Figure 8.2: An exemplary noiseless material data set ($P = 20^3$) used for data-driven solvers: strain components. (Note: The axis labels denote the Green strain components)

Following the same procedure of data generation in Section 6.5.2, a noiseless material data set (e.g., Figure 8.2) is provided for data-driven computing. The solutions obtained from the distance-minimizing data-driven (DMDD) approach and the proposed locally convex data-driven (LCDD) approach are provided in Figure 8.3, where the reference solution by the canonical numerical solver based on constitutive models is also given. As shown in Figure 8.3, LCDD (with $k = 6$) obtains satisfactory results at every loading step, while DMDD diverges badly at

some steps. However, the results suggest that distinct from model-based numerical simulations, data-driven computing is relatively insensitive to solving process, and it may still arrive at a reasonable final solution regardless of misleading intermediate results. This may be because data-driven solution relies on the discrete material data instead of continuous function forms of material model.

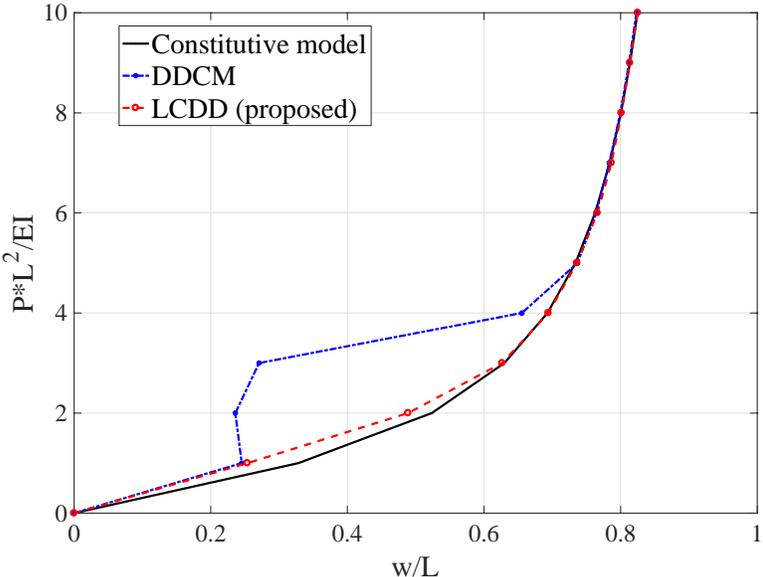


Figure 8.3: Comparison of model-based (black), DDCM (blue) and LCDD (red) solutions, where w is the tip-deflection.

8.2.2 Uniaxial tension of hyper-elasticity material

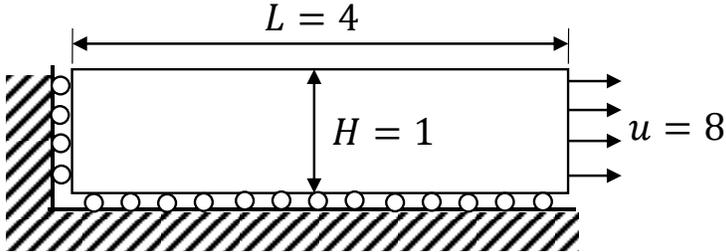


Figure 8.4: Schematic of a specimen of hyper-elasticity material subjected to uniaxial tension.

Figure 8.4 shows the initial configuration of a specimen to be stretched in x direction. In this case, we adopt an incompressible polynomial hyperelastic model (the generalized Mooney-Rivlin model) such that the strain energy density per unit volume is expressed as

$$W(I_1, I_2, I_3) = \sum_{i+j=1}^2 A_{ij} (I_1 - 3)^i (I_2 - 3)^j, \quad I_3 = 1, \quad (8.10)$$

where I_j ($j=1,2,3$) are the invariants of the right Cauchy-Green tensor, and the material parameters are chosen as $C_{10} = 0.373$, $C_{20} = -0.031$, and $C_{20} = 0.005$. For analysis, plane stress condition is assumed. The analytical solution for incompressible material was given in (Rivlin 1948; Rivlin & Saunders 1951),

$$\sigma_{xx} = 2(\lambda^2 - \lambda^{-1}) \left(\frac{\partial W}{\partial I_1} + \lambda^{-1} \frac{\partial W}{\partial I_2} \right), \quad (8.11)$$

where σ_{xx} is the normal Cauchy stress component along x direction and λ is the associated stretch ratio in x direction. With simple derivation, we have the 2nd Piola-Kirchhoff (2nd-PK) stress component $S_{xx} = \lambda^{-2} \sigma_{xx}$. The stress-strain plots given by the analytical solution and the RKPM simulation solution are given in Figure 8.5. The agreement demonstrates the accuracy of the model-based simulation solution. In the following study, the model-based simulation solution is used as the reference for the data-driven solutions.

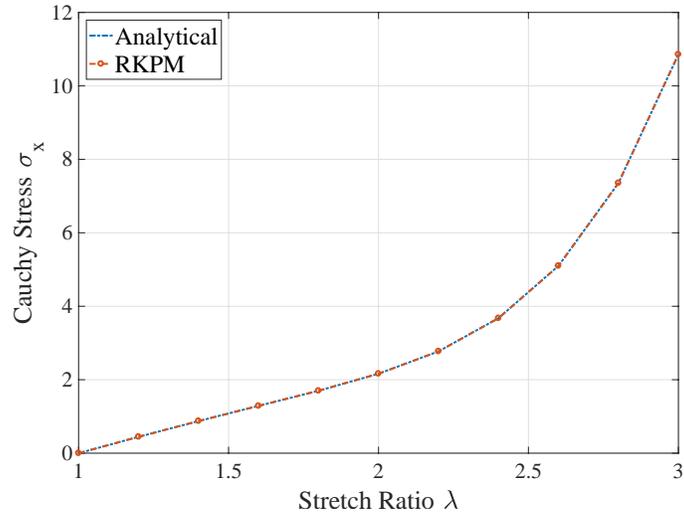


Figure 8.5: Comparison of analytical solution and the model-based simulation solution (RKPM).

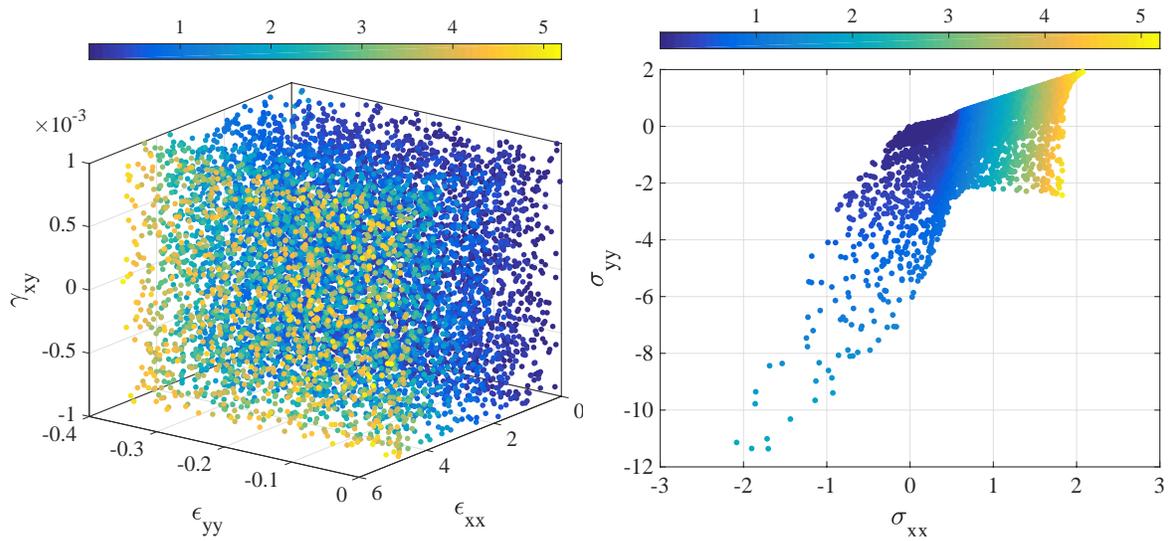


Figure 8.6: Exemplary noiseless material data sets ($P = 20^3$) used for data-driven solvers: strain components (left); stress components (right). (Note: The axis labels denote the Green strain components (left) and the 2nd PK stress components (right))

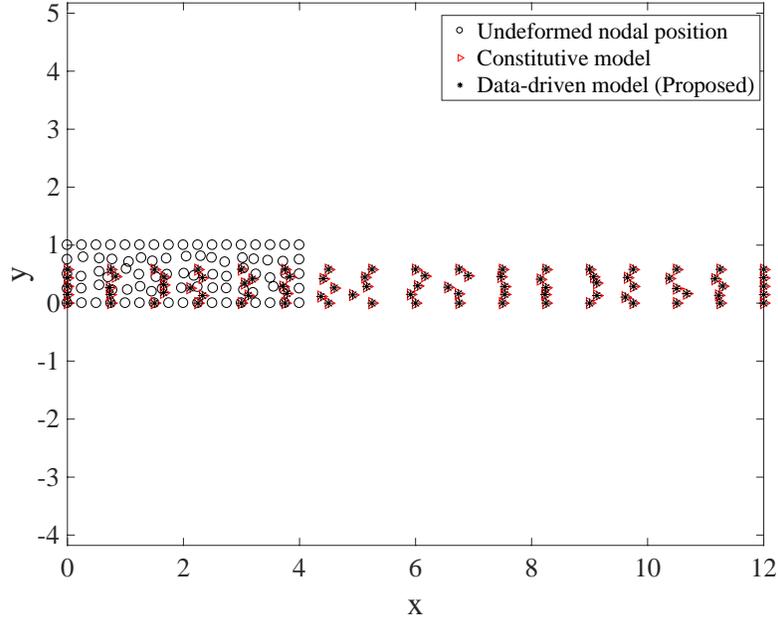
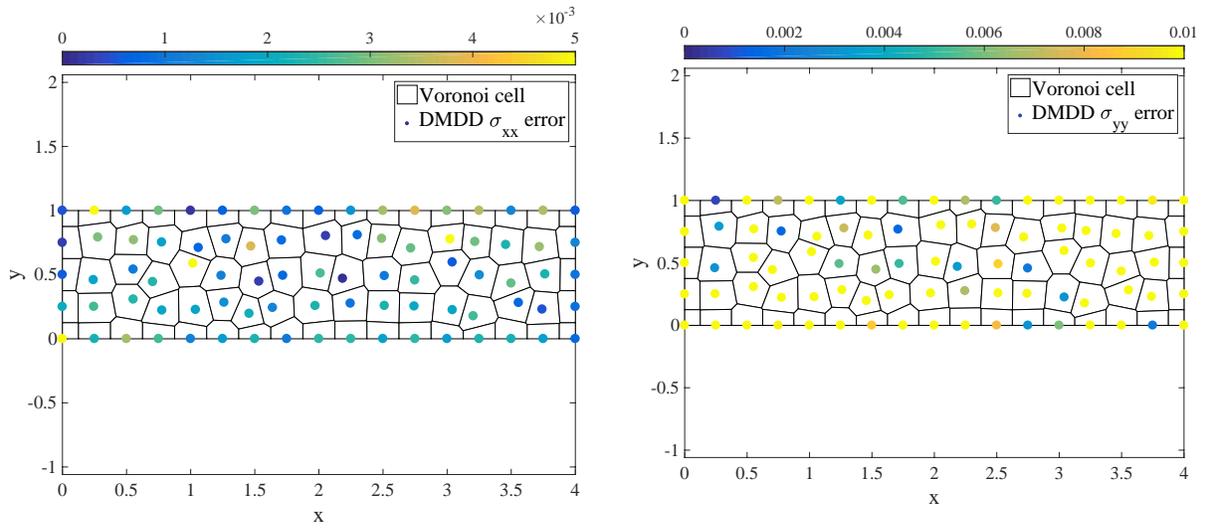
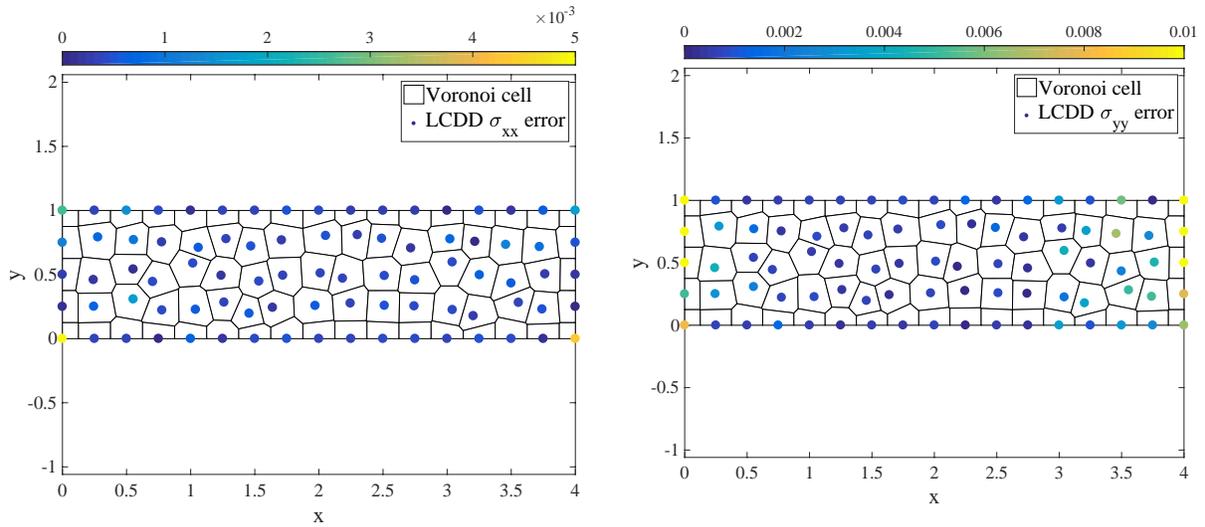


Figure 8.7: Comparison of the proposed data-driven ($k = 6$) displacement solution against the model-based simulation solution by using the noiseless dataset.

Based on the hyperelastic model in (8.10), we produce the material database (Figure 8.6) for data-driven simulation. As shown in Figure 8.7, the displacement result given by the proposed LCDD solver agree well with the model-based reference solution. The absolute errors of the 2nd PK stress components are also provided in Figure 8.8. Although the standard data-driven solver, DDCM, yields acceptable prediction of S_{xx} , it performs poorly in approximating S_{yy} that is supposed to be vanishing under this uniaxial tension condition. On the other hand, the proposed LCDD approximates both S_{xx} and S_{yy} accurately and yields negligible errors.



(a) DMDD: the error of 2nd PK stress S_{xx} and S_{yy}



(b) LCDD: the error of 2nd PK stress S_{xx} and S_{yy}

Figure 8.8: The absolute errors of stress components S_{xx} and S_{yy} solved by DMDD and LCDD against the model-based simulation solution.

8.3 Data-Driven Modeling: Porcine Atrioventricular Heart Valve Tissue

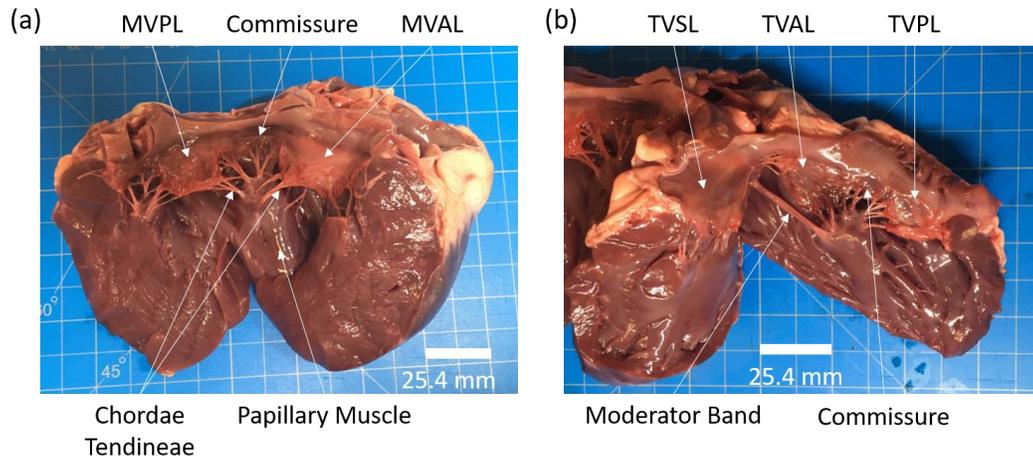


Figure 8.9: Schematic of Mitral Valve (MV) and Tricuspid Valve (TV) leaflets for biaxial mechanical testing (Jett et al. 2018). Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.

As the proposed data-driven solver has been verified in last section, here it is used to simulate a biological tissue with realistic experimental data, provided by our collaborators in Biomechanics and Biomaterials Design Laboratory at the University of Oklahoma (Jett et al. 2018). The material data set was collected by performing biaxial mechanical experiments with extensive testing protocols to examine the mechanical behaviors of the mitral valve and tricuspid valve leaflets (see Figure 8.9).

8.3.1 Biaxial mechanical experiments

As most of the constitutive models calibrated from uniaxial tension tests cannot accurately characterize the material responses under complex deformation states (Sacks 2000), it is a long-standing challenge to predict the general constitutive behaviors of biological tissues

under finite deformation. Hence, planar biaxial testing experiments with multiple loading ratios is utilized to investigate the anisotropic material response of each mitral and tricuspid valve leaflet. A simple description of the biaxial testing experiments is provided as follows. More information of the experiment setting is referred to (Jett et al. 2018).

For biaxial mechanical testing, a square specimen ($8 \text{ mm} \times 8 \text{ mm}$) was dissected from the central region of each valve leaflet, as shown in Figure 8.10(a) and Figure 8.10(b). The thickness of the specimen was measured at three different locations to obtain the average value. Then, the square tissue specimen (Figure 8.10(c)) was mounted on a commercial biaxial mechanical testing system, BioTester (CellScale, Waterloo, ON, Canada), for various loading tests.

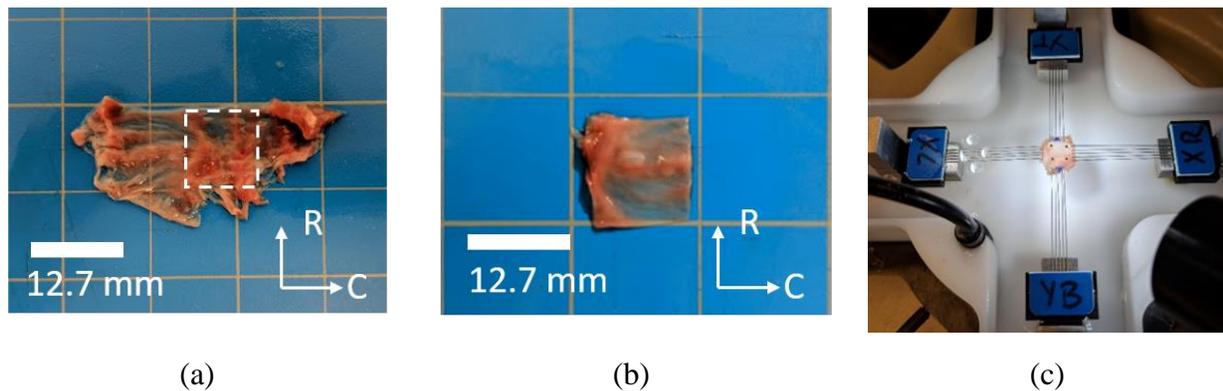


Figure 8.10: Experimental Images of the Experimental Setups for Biaxial Mechanical Testing, where C denotes circumferential direction and R denotes radial direction (Jett et al. 2018). Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.

To measure the in-plane deformation and strain of the specimen under loading, a non-contact digital image correlation (DIC) based techniques (Pan et al. 2009) is employed. As shown in Figure 8.11, four fiducial markers (with diameters of $300\text{-}500 \mu\text{m}$) were attached in the interior region of the specimen to prevent the Saint-Venant effects on tissue deformations during

biaxial mechanical testing (Humphrey, Vawter & Vito 1987). The schematics of the force-controlled and displacement-controlled biaxial mechanical testing protocols are illustrated in Figure 8.12 and Figure 8.13, respectively. Before performing loading tests, the tissues are loaded and unloaded 10~15 cycles as a “preconditioning” step to bring the tissue back to its functional state because the heart valve tissue just excised from the heart were typically not chosen as the reference configuration. Under the process of loading, a series of images of the deformed tissue specimen were collected by the high-resolution CCD camera (The Imaging Source LLC, Charlotte, NC), and the time-dependent positions of the four markers were tracked based on the acquired images. Applying finite element discretization, the in-plane deformation gradient \mathbf{F} associated to the tissue specimen can be computed by using the displacements of the four markers. As a result, the right Cauchy-Green deformation tensor $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ and the Green strain tensor $\mathbf{E} = 1/2(\mathbf{C} - \mathbf{I})$ can be obtained. On the other hand, the first Piola-Kirchhoff (1st-PK) stress tensor \mathbf{P} was computed from the applied membrane tensions, T_C and T_R in two orthogonal directions, as follows

$$\mathbf{P} = \frac{1}{t} \begin{bmatrix} T_C & 0 \\ 0 & T_R \end{bmatrix}, \quad (8.12)$$

where t is the tissue undeformed thickness.

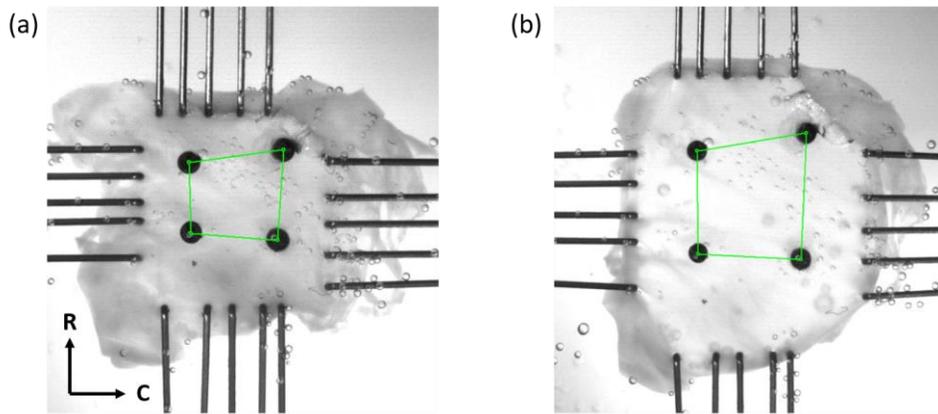


Figure 8.11: Experimental Photos of Digital Image Correlation Based Tracking – Strain Calculation (Jett et al. 2018). Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.

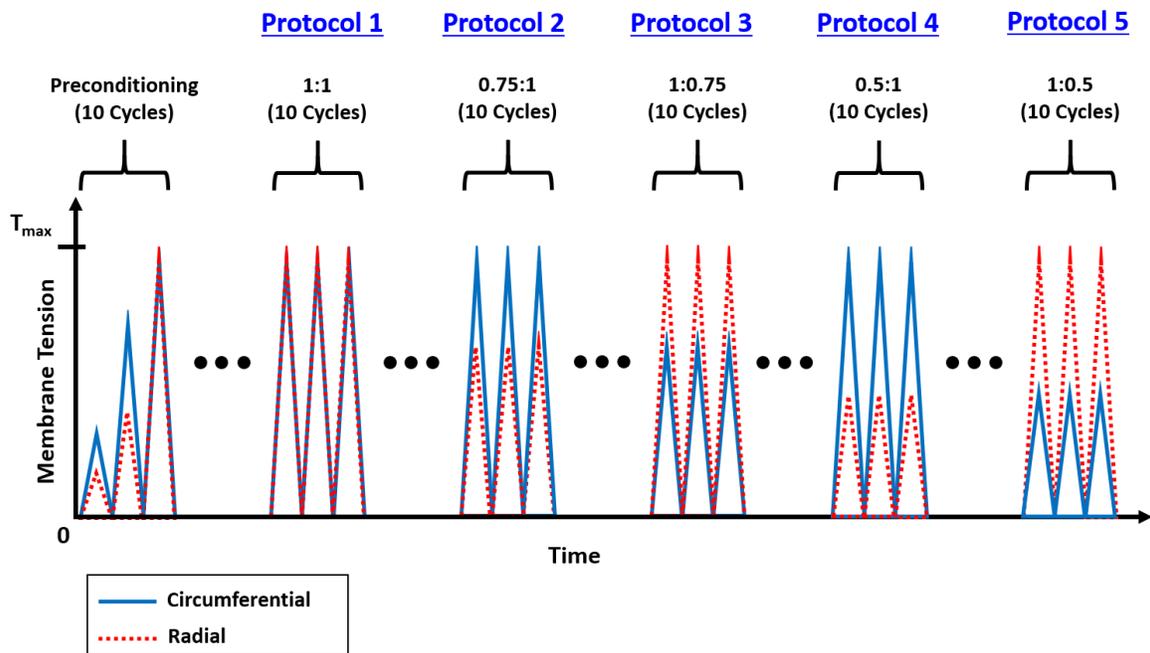


Figure 8.12: Schematic of the force-controlled biaxial mechanical testing protocol (Jett et al. 2018). Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.

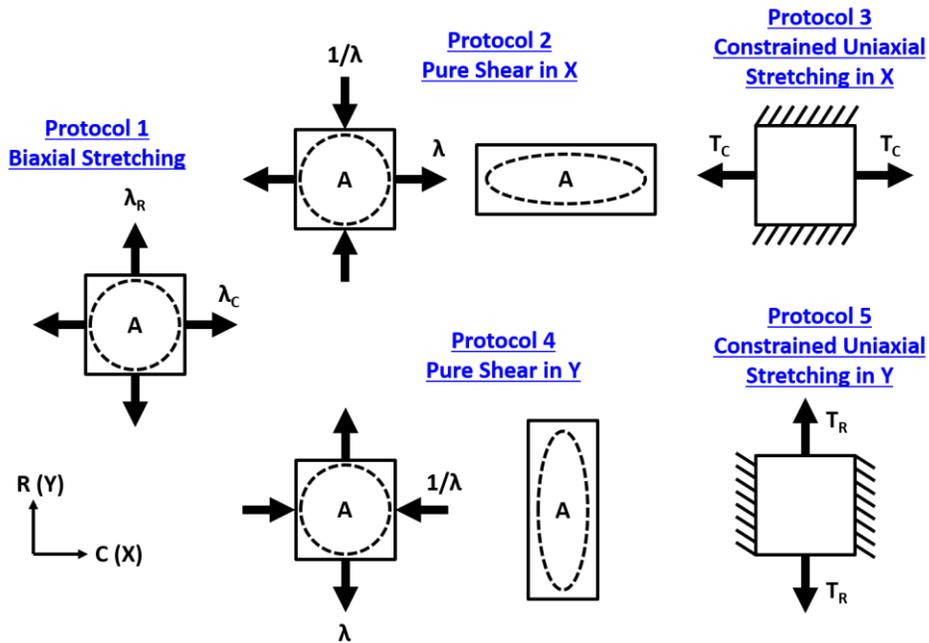


Figure 8.13: Schematic of the displacement-controlled biaxial mechanical testing protocol. Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.

8.3.2 Data-driven modeling

For data-driven modeling, we generate a material database acquired from experiment tests introduced in Section 8.3.1 for mitral valve anterior leaflet (MVAL) under all 5 loading protocols ($T_{C,\max} : T_{R,\max} = 1:1, 0.75:1, 1:0.75, 0.5:1$, and $1:0.5$ as shown in Figure 8.12) and 4 displacement protocols (2nd -5th protocols as shown in Figure 8.13). The biaxial mechanical testing results under the 5 loading protocols are given in Figure 8.14 as an example. The stress components of the combinatorial material dataset are plotted in Figure 8.15. Compared to the previous synthesized material data sets, the data collected from experimental measurement seems more irregular in the phase space. To validate data-driven method, a specimen of heart valve tissue (Figure 8.16(a)) under the 1st displacement-controlled testing protocol is not used in training data.

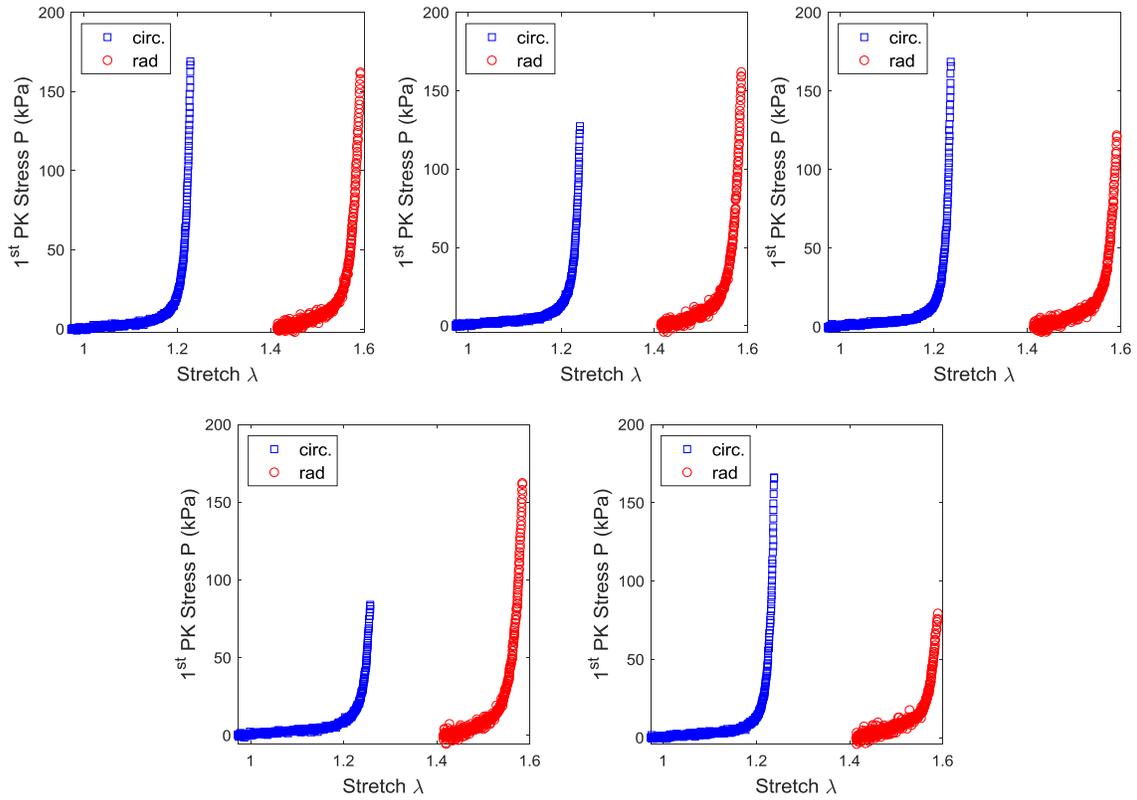


Figure 8.14: Experiment results for mitral valve anterior leaflet (MVAL) under 5 different force-controlled biaxial mechanical testing protocols. Courtesy of Professor Chung-Hao Lee at The University of Oklahoma.

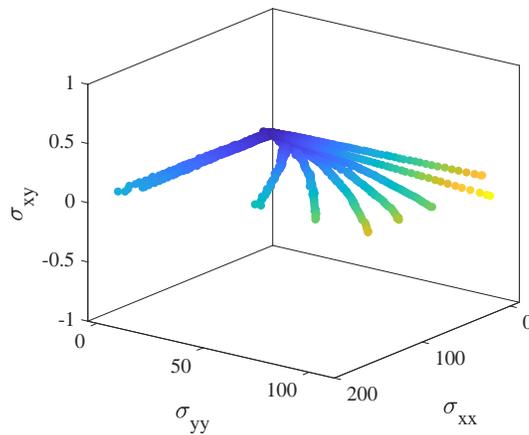


Figure 8.15: Stress components of the material dataset collected by biaxial testing experiments under the 5 force-controlled loading and 4 displacement-controlled loading. (Note: The axis labels denote the 2nd PK stress components)

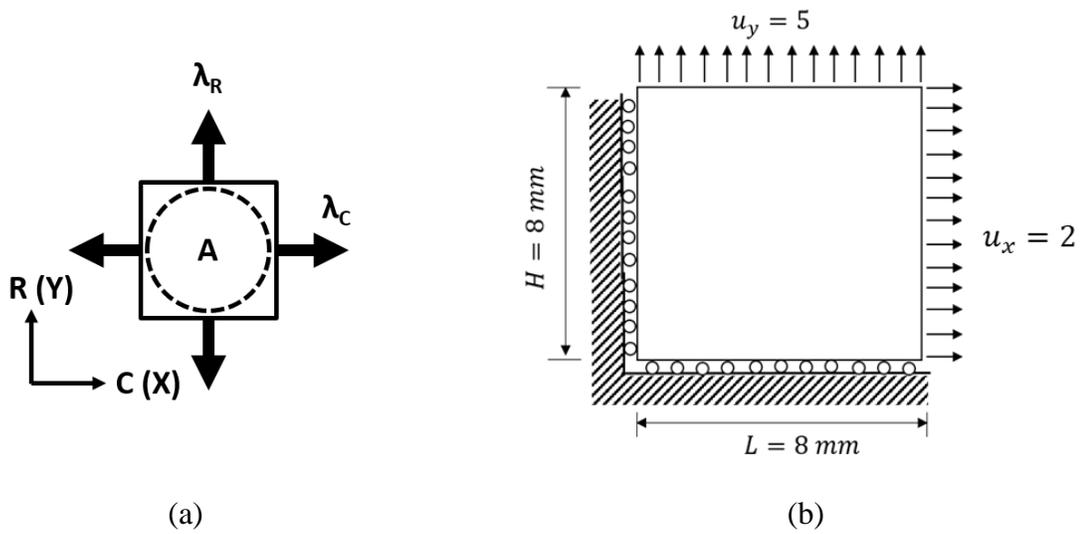


Figure 8.16: (a) Schematic of a specimen of heart valve tissue subjected to the 1st displacement-controlled testing protocol and (b) the associated mechanics model for data-driven modeling.

The proposed LCDD solver using the given material data is employed to simulate the response of the model subjected to the biaxial tension defined in Figure 8.16(b). The data-driven prediction of the 2nd-PK stress components S_{xx} and S_{yy} is given in Figure 8.17. Despite the irregularities in the material dataset, the proposed data-driven solver yields nearly homogeneous stress distributions over the specimen domain except for the boundary areas. By taking average of the data-driven strain and stress results over the domain, we compare the data-driven prediction to the experimental data for the mitral valve anterior leaflet (MVAL) under 1st displacement-controlled biaxial mechanical testing protocol, as shown in Figure 8.18.

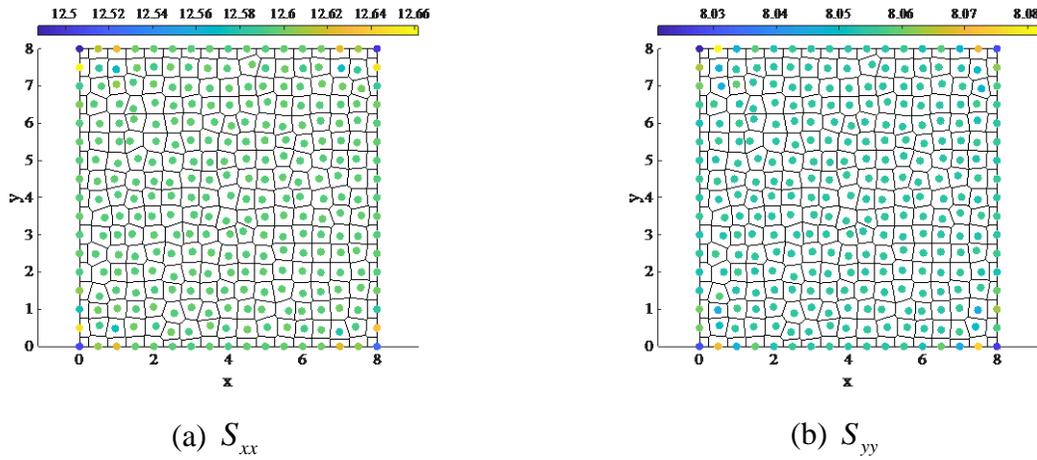


Figure 8.17: Predictions of 2nd PK Stresses by the proposed data-driven (LCDD) solver.

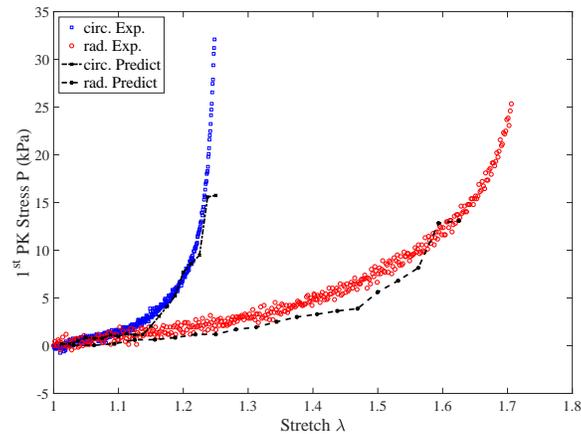
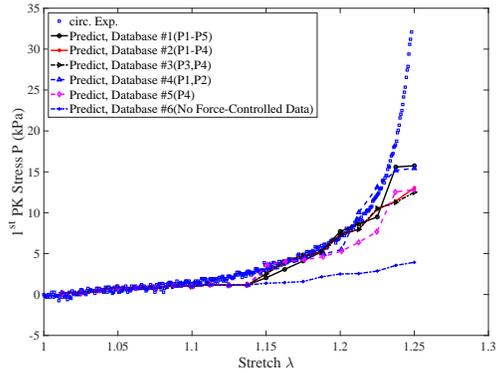


Figure 8.18: Comparison of the experiment data and data-driven prediction for mitral valve anterior leaflet (MVAL) under 1st displacement-controlled biaxial mechanical testing protocol.

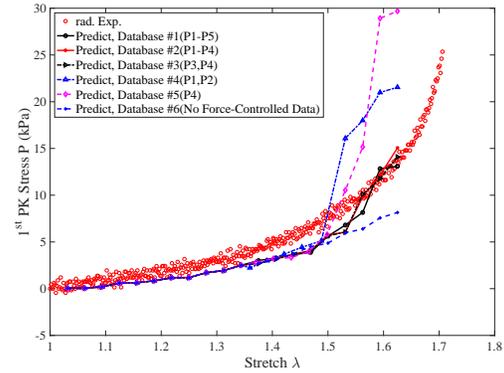
As shown in Figure 8.18, the data-driven prediction of stretch against 1st PK stress generally agrees well with experimental data in both the circumferential and radial direction. The mismatch observed in the radial direction could be due to the different boundary conditions applied to the simulation model (Figure 8.16(b)) and the experimental set-up, and the boundary could have strong influence on the behavior of biological tissues (Humphrey, Vawter & Vito 1987). Moreover, it should be noted that the mechanical behavior of biological tissues is

sensitive to loading ratios, fiber orientation, temperature, etc., and it has a great variety from specimen to specimen. Thus, it is expected that the material database collected from different specimens of MVAL for training purposes could be different from the materials selected for validation.

The sensitivity of data-driven solutions to the size of material database is analyzed by varying the number of force-controlled protocols used to constitute the database for data-driven modeling while remaining all the datasets obtained from the four displacement-controlled protocols in the database. As shown in Figure 8.19, the data-driven solver yields worse predictions as more force-controlled datasets are removed from the material database. The data-driven predictions with less force-controlled datasets cannot capture the turning point well, and result in large errors especially for the large strain region. For example, the data-driven prediction given by database #5 (P4), where only the dataset of the 4th force-controlled protocol (see Figure 8.14) is collected into the material database, shows large discrepancies at the large strain range even though it agrees with the experimental results at the early stage of loading. The results suggest that a rich data with respect to more loading scenarios is essential for reliable data-driven solutions, which is the fundamentals of any data-driven approaches. Lastly, due to the scattering distribution of data (see Figure 8.15), a learning technique with better conditioning on data (e.g., a better selection of coefficient matrix \mathbf{M}_α) may help to improve the results.



(a) Circumferential direction



(b) Radial direction

Figure 8.19: Comparison of data-driven solutions using different number of force-controlled datasets for the numerical modeling of the mitral valve anterior leaflet (MVAL) under 1st displacement-controlled biaxial mechanical testing protocol.

8.4 Summary

In this chapter, we showed the extension of the locally convex data-driven (LCDD) solver for nonlinear kinematics such that it allows to simulate biological materials under finite deformation. Several demonstration tests with synthetic material data have been conducted to verify the effectiveness of the proposed data-driven solver for nonlinear solids. Furthermore, we presented a preliminary study of applying data-driven simulation for a realistic biological tissue, where the material data is collected by biaxial experiment testing. It shows that data-driven computing is a promising predictive simulation tool for complex materials, although its performance is subjected to the quality of data. More investigations on data collection, data processing, and physics-informed machine learning are essential to advance this field.

8.5 Acknowledgement

This chapter, in part, is currently being prepared for submission for publication of the material as it may appear in “He, Qizhi, Chen; Jiun-Shyan. Robust data-driven computational modeling of biological material”. The dissertation author was the primary investigator and author of this material. The Biomechanics and Biomaterials Design Laboratory and Dr. Chung-Hao Lee at the University of Oklahoma are greatly acknowledged for providing experiment images and measurement data.

Chapter 9

Conclusions and Future Work

9.1 Conclusions

In this dissertation, we investigated the applications of model order reduction (MOR) and data-driven computational modeling to a variety of linear and nonlinear mechanics problems. These two methodologies are shown closely related since they are both based on a hybrid approach that integrates physical models with statistical learning techniques, e.g. dimensionality reduction. The objective of this work is to further advance the current state of hybrid data-model approaches in computational mechanics to address the computational limitations when dealing with nonlinear mechanics system and complex constitutive models. Thus, we proposed to enhance the performance of MOR and data-driven modeling by introducing suitable machine learning techniques as well as physics-preserving methods. This work includes two parts, the development of robust, physics-preserving MOR methods for parameterized PDEs, and the development of data-driven computational framework for complex material modeling.

In the context of MOR, we first proposed a decomposed subspace reduction (DSR) method to preserve the essential near-tip characteristics, singularities and discontinuities, of the original full-order model for fracture mechanics. In this approach, the reduced-order model is constructed based on the integrated singular basis function method (ISBFM) with meshfree approximation enriched by crack-tip basis functions, which allows a lower order integration of

the Galerkin equation, but also yields a discrete system containing sparse sub-matrices for physics-preserving MOR procedures.

Next, we developed a robust reduced-order model for parameterized nonlinear systems characterized by a wide variety of nonlinear behaviors in terms of parameter changes. The reduced-order basis used to construct the low-dimensional subspace is derived from a generalized manifold learning framework in conjunction with linearization techniques, coined as linear graph embedding (LGE). This general framework allows the utilization of a priori statistical knowledge, such as local geometry, of given data during the construction of reduced-order basis vectors. Thus, it yields a robust reduced-order model less sensitive to noise and outliers and is well suited for nonlinear physical systems. The advantages of the proposed LGE based MOR has been demonstrated by testing the parameterized elliptic PDE and Burgers' problem.

Furthermore, a nonlinear MOR for a meshfree Galerkin formulation based on the stabilized conforming nodal integration (SCNI) scheme is developed, which yields a pure node based MOR that is particularly effective for hyper-reduction techniques. A numerical example of two-phase hyperelastic solid with perturbed loading conditions is used to validate the effectiveness of the proposed reduction method.

In the context of data-driven modeling, we developed an accurate, robust data-driven computational framework to provides an alternative to conventional physical modeling of complex materials. This framework allows physical simulation to directly interact with material data through machine learning procedures instead of employing phenomenological constitutive models. To address the “curse of dimensionality” and the robustness issue associated with noisy

and limited data, we proposed a manifold learning enhanced data-driven solver where it searches data solutions from a locally reconstructed convex hull associated to the k -nearest neighbor (k -NN) points, resulting in robustness in dealing with noisy data and ensuring stability. It has shown that the proposed data-driven solver yields a linear exactness to the canonical numerical reference solution when the database is well sampled from the graph of constitutive relations.

The proposed data-driven computational framework has also been extended to finite deformation mechanics for simulating biological materials under the same meshfree framework used for reduced-order modeling. The accuracy and robustness of the proposed data-driven approach are demonstrated in the modeling of linear and nonlinear elasticity problems. In addition, a preliminary result of data-driven modeling of biological tissue has been presented utilizing material data collected from laboratory testing on heart valve tissue, showing the potential of data-driven simulation by integrating physical modeling and machine learning techniques.

9.2 Recommendations for Future Research

This work is a first step in the direction of using the hybrid data-physics approach to address the computational limitations in engineering applications when dealing with nonlinear mechanics system and complex constitutive models. There are a number of possible enhancements as well as challenging applications that ask for further study. Some potential topics for future research are summarized as follows:

1. The proposed LGE is robust and suitable for model reduction of systems that exhibit a wide range of variations in terms of parameters changes. However, the weight function is

selected empirically for the demonstration problems. Thus, it is of interest to study the automated strategy to select a proper weight function for the problem of interest based on iterative approaches, such as the expectation-maximization (EM) algorithm (Vidal, Ma & Sastry 2016). Moreover, for nonlinear systems with a severe variability, the introduction of local subspace methods (Amsallem, Zahr & Farhat 2012; Peherstorfer et al. 2014; Peng & Mohseni 2016; Sargsyan, Brunton & Kutz 2015) into the proposed LGE projection can be used to construct a more effective MOR model.

2. The data-enhanced physics-based computational paradigm can be further enhanced. For example, some other robust techniques or advanced clustering algorithms to suppress outliers that are well established in machine learning area can be applied under the proposed LCDD framework. The way to identify the optimal number of neighbors in the k-NN algorithm is still an open issue (Anava & Levy 2016).
3. The proposed data-driven framework is inspired by manifold learning techniques that utilize the local information of material data to form the local manifold. But this framework is not limited to the manifold learning structure. The neural network-based techniques (Ghaboussi, Garrett & Wu 1991; Goodfellow et al. 2016; Haykin 2009; MacKay & Mac Kay 2003) can also be applied to the material data stage and carry out the simulation together with physics-based models.

Appendix A

Reproducing Kernel Particle Method (RKPM)

Since the reproducing kernel particle method (Chen et al. 1996; Liu, Jun & Zhang 1995) is adopted for numerical discretization and solution approximation in most of the numerical studies of this thesis, we briefly review the construction of the reproducing kernel (RK) shape functions and its properties in this section.

The RK approximation of a function u , denoted as u^h , is expressed as a linear combination of N RK shape functions as follows

$$u^h(\mathbf{x}) = \sum_{I=1}^N \Psi_I(\mathbf{x}) d_I \quad (\text{A.1})$$

where \mathbf{x}_I are the nodal coordinates of the N RK particles, d_I are the nodal coefficients, and Ψ_I are the RK shape functions expressed as a corrected version of a compactly supported kernel

$$\Psi_I(\mathbf{x}) = C(\mathbf{x}; \mathbf{x} - \mathbf{x}_I) \Phi_a(\mathbf{x} - \mathbf{x}_I) \quad (\text{A.2})$$

The kernel function Φ_a with a compact support size “ a ” defines the locality of the shape functions Ψ_I . In this manner, the shape function associated with a node interacts only with a small group of neighboring nodes, yielding better sparseness as well as conditioning in the stiffness matrix for solving linear system. The kernel function Φ_a also determines the smoothness of the RK shape function. For example, the cubic-B splines commonly used as kernel functions provides C^2 continuity, defined as

$$\Phi_a(z) = \begin{cases} 2/3 - 4z^2 + 4z^3 & \text{for } 0 \leq z < 1/2 \\ 4/3 - 4z + 4z^2 - 4/3 z^3 & \text{for } 1/2 \leq z < 1 \\ 0 & \text{for } z \geq 1 \end{cases} \quad (\text{A.3})$$

where $z = \|\mathbf{x} - \mathbf{x}_I\|/a$.

To impose the n -th order of completeness in the approximation (A.2), the correction function $C(\mathbf{x}; \mathbf{x} - \mathbf{x}_I)$ is defined as

$$C(\mathbf{x}; \mathbf{x} - \mathbf{x}_I) = \sum_{|\alpha| \leq n} (\mathbf{x} - \mathbf{x}_I)^\alpha b_\alpha(\mathbf{x}) = \mathbf{H}^T(\mathbf{x} - \mathbf{x}_I) \mathbf{b}(\mathbf{x}) \quad (\text{A.4})$$

in which $\mathbf{H}(\mathbf{x} - \mathbf{x}_I)$ is a vector consisting of n th-order monomial basis functions. The coefficient vector \mathbf{b} is solved by enforcing the following reproducing conditions

$$\sum_{I=1}^N \Psi_I(\mathbf{x}) \mathbf{x}_I^\alpha = \sum_{I=1}^N \mathbf{H}^T(\mathbf{x} - \mathbf{x}_I) \mathbf{b}(\mathbf{x}) \Phi_a\left(\frac{\|\mathbf{x} - \mathbf{x}_I\|}{a}\right) \mathbf{x}_I^\alpha = \mathbf{x}^\alpha, \quad |\alpha| \leq n \quad (\text{A.5})$$

After some mathematical derivation, the coefficient vector is given by

$$\mathbf{b}(\mathbf{x}) = \mathbf{M}^{-1}(\mathbf{x}) \mathbf{H}(\mathbf{0}) \quad (\text{A.6})$$

where

$$\mathbf{M}(\mathbf{x}) = \sum_{I=1}^N \mathbf{H}(\mathbf{x} - \mathbf{x}_I) \mathbf{H}^T(\mathbf{x} - \mathbf{x}_I) \Phi_a\left(\frac{\|\mathbf{x} - \mathbf{x}_I\|}{a}\right) \quad (\text{A.7})$$

is the moment matrix. The RK shape functions are then obtained as

$$\Psi_I(\mathbf{x}) = \mathbf{H}^T(\mathbf{0}) \mathbf{M}^{-1}(\mathbf{x}) \mathbf{H}(\mathbf{x} - \mathbf{x}_I) \Phi_a\left(\frac{\|\mathbf{x} - \mathbf{x}_I\|}{a}\right) \quad (\text{A.8})$$

The construction based on reproducing conditions allows the RK approximation to exactly reproduce monomials up to degree n and therefore, the partition of unity property holds.

It should be noted that the RK shape functions do not possess the Kronecker delta property, i.e. $\Psi_I(\mathbf{x}_J) \neq \delta_{IJ}$, and thus, $u^h(\mathbf{x}_I) \neq d_I$. Special treatment is needed for the imposition of essential boundary conditions while solving boundary value problems based on Galerkin method. A number of methods have been proposed to address this issue, such as Lagrange multipliers method (Belytschko, Lu & Gu 1994), Nitsche's method (Fernández-Méndez & Huerta 2004; Nitsche 1971), boundary singular kernel method (Chen & Wang 2000), and transformation methods (Chen et al. 1996; Chen & Wang 2000). More discussions on the mathematical properties of the reproducing kernel approximation and the imposition of essential boundary conditions can be found in the recent review paper (Chen, Hillman & Chi 2017).

Appendix B

Nonnegative Least Squares (NNLS) Solver

Let us recall a standard NNLS problem: *given a matrix $\mathbf{A} \in \mathbb{R}^{n \times p}$ (usually $p > n$) and a observed vector $\mathbf{z} \in \mathbb{R}^n$, find a nonnegative vector $\mathbf{y}^* \in \mathbb{R}_+^p$ to minimize the following function,*

$$\begin{aligned} \mathbf{y}^* &= \arg \min_{\mathbf{x} \in \mathbb{R}^p} \|\mathbf{A}\mathbf{y} - \mathbf{z}\|, \\ \text{subject to: } &y_i \geq 0, \quad i = 1, \dots, p, \end{aligned} \tag{B.1}$$

where $\|\cdot\|$ stands for the standard Euclidean norm.

A variety of methods have been applied to tackle the NNLS problem since 1980s. Those algorithms in general can be roughly categorized into active-set methods and iterative approaches (Chen & Plemmons 2010). Lawson and Hanson (Lawson & Hanson 1987) seems to propose the first standard algorithm to solve NNLS problem (B.1). Their method is essentially an active set method (Gill, Murray & Wright 1981), which is based on the observation that only a small subset of the non-negative constraints are usually active at the solution. It shows in (Lawson & Hanson 1987) that the iteration in the active set method converges and terminates without any cutoff in iterations. Since most research studies in computational mechanics are based on the active-set method (Lawson & Hanson 1987) to solve the *Cubature problem*, we only review this method in Algorithm 1.

Algorithm 1 Non-negative least squares solver: $\mathbf{y}^* \leftarrow \text{NNLS}(\mathbf{A}, \mathbf{z}, \text{TOL})$

Input: $\mathbf{A} \in \mathbb{R}^{n \times p}$, $\mathbf{z} \in \mathbb{R}^n$, TOL

Output: $\mathbf{y}^* \succeq \mathbf{0}$ such that $\mathbf{y}^* = \arg \min \|\mathbf{A}\mathbf{y} - \mathbf{z}\|^2$

Initialization: $\mathcal{Z} \leftarrow \emptyset$, $\mathcal{Y} \leftarrow \{1, 2, \dots, p\}$, $\mathbf{y} = \mathbf{0}$, $\mathbf{r} \leftarrow \mathbf{z}$

WHILE $\|\mathbf{r}\|/\|\mathbf{z}\| > \text{TOL}$ and $\mathcal{Y} \neq \emptyset$, DO

$\mathbf{q} \leftarrow \mathbf{A}^T(\mathbf{z} - \mathbf{A}\mathbf{y})$, $j = \arg \max_{i=1, \dots, p} (q_i)$

Include the index j in \mathcal{Z} and remove it from \mathcal{Y}

$\mathbf{s}_{\mathcal{Z}} \leftarrow (\mathbf{A}_{\mathcal{Z}}^T \mathbf{A}_{\mathcal{Z}})^{-1} \mathbf{A}_{\mathcal{Z}}^T \mathbf{z}$, $\mathbf{s}_{\mathcal{Y}} \leftarrow \mathbf{0}$, where $\mathbf{A}_{\mathcal{Z}} \in \mathbb{R}^{n \times |\mathcal{Z}|}$

WHILE $\min_{i \in \mathcal{Z}} s_{\mathcal{Z}} \leq 0$, DO

$\alpha = -\min_{i \in \mathcal{Z}} y_i / (y_i - s_i)$

$\mathbf{y} \leftarrow \mathbf{y} + \alpha(\mathbf{s} - \mathbf{y})$

Update \mathcal{Y} with zero value indices of \mathbf{y} and \mathcal{Z} with the positive indices of \mathbf{y}

$\mathbf{s}_{\mathcal{Z}} \leftarrow (\mathbf{A}_{\mathcal{Z}}^T \mathbf{A}_{\mathcal{Z}})^{-1} \mathbf{A}_{\mathcal{Z}}^T \mathbf{z}$, $\mathbf{s}_{\mathcal{Y}} \leftarrow \mathbf{0}$

$\mathbf{y} \leftarrow \mathbf{s} = \mathbf{s}_{\mathcal{Z}} \cup \mathbf{s}_{\mathcal{Y}}$

$\mathbf{r} \leftarrow \mathbf{z} - \mathbf{A}\mathbf{y}$

Return $\mathbf{y}^* \leftarrow \mathbf{y}$

Bibliography

- Akbari Rahimabadi, A., Kerfriden, P. & Bordas, S. 2015, 'Scale selection in nonlinear fracture mechanics of heterogeneous materials', *Philosophical Magazine*, vol. 95, no. 28–30, pp. 3328–47.
- Amsallem, D., Cortial, J., Carlberg, K. & Farhat, C. 2009, 'A method for interpolating on manifolds structural dynamics reduced-order models', *International Journal for Numerical Methods in Engineering*.
- Amsallem, D. & Farhat, C. 2008, 'Interpolation Method for Adapting Reduced-Order Models and Application to Aeroelasticity', *AIAA Journal*, vol. 46, no. 7, pp. 1803–13.
- Amsallem, D. & Farhat, C. 2011, 'An Online Method for Interpolating Linear Parametric Reduced-Order Models', *SIAM Journal on Scientific Computing*, vol. 33, no. 5, pp. 2169–98.
- Amsallem, D., Zahr, M., Choi, Y. & Farhat, C. 2015, 'Design optimization using hyper-reduced-order models', *Structural and Multidisciplinary Optimization*.
- Amsallem, D., Zahr, M.J. & Farhat, C. 2012, 'Nonlinear model order reduction based on local reduced-order bases', *International Journal for Numerical Methods in Engineering*, vol. 92, no. 10, pp. 891–916.
- An, S.S., Kim, T. & James, D.L. 2008, 'Optimizing cubature for efficient integration of subspace deformations', *ACM SIGGRAPH Asia 2008 papers on - SIGGRAPH Asia '08*, p. 1.
- Anava, O. & Levy, K. 2016, 'k*-Nearest Neighbors: From Global to Local', *NIPS Advances in Neural Information Processing Systems*, no. Nips, pp. 4916–24.
- Angermueller, C., Pärnamaa, T., Parts, L. & Stegle, O. 2016, 'Deep learning for computational biology', *Molecular Systems Biology*.
- Antoulas, A.C. 2005, 'An overview of approximation methods for large-scale dynamical systems', *Annual Reviews in Control*, vol. 29, no. 2, pp. 181–90.
- Antoulas, A.C. & Sorensen, D.C. 2001, 'Approximation of large-scale dynamic systems: an overview', *International Journal of Applied Mathematics and Computer Science*, vol. 11, pp. 1093–121.
- Arridge, S.R., Kaipio, J.P., Kolehmainen, V., Schweiger, M., Somersalo, E., Tarvainen, T. & Vauhkonen, M. 2006, 'Approximation errors and model reduction with an application in optical diffusion tomography', *Inverse Problems*.
- Astolfi, A. 2010, 'Model reduction by moment matching for linear and nonlinear systems', *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2321–36.

- Astrid, P., Weiland, S., Willcox, K. & Backx, T. 2008, ‘Missing Point Estimation in Models Described by Proper Orthogonal Decomposition’, *IEEE Transactions on Automatic Control*, vol. 53, no. 10, pp. 2237–51.
- Avril, S., Bonnet, M., Bretelle, A.S., Grédiac, M., Hild, F., Jenny, P., Latourte, F., Lemosse, D., Pagano, S., Pagnacco, E. & Pierron, F. 2008, ‘Overview of identification methods of mechanical parameters based on full-field measurements’, *Experimental Mechanics*, vol. 48, no. 4, pp. 381–402.
- Ayensa-Jiménez, J., Doweidar, M.H., Sanz-Herrera, J.A. & Doblaré, M. 2018, ‘A new reliability-based data-driven approach for noisy experimental data with physical constraints’, *Computer Methods in Applied Mechanics and Engineering*, vol. 328, pp. 752–74.
- Ben Azzouna, M., Feissel, P. & Villon, P. 2015, ‘Robust identification of elastic properties using the Modified Constitutive Relation Error’, *Computer Methods in Applied Mechanics and Engineering*, vol. 295, pp. 196–218.
- Bai, Z. 2002, ‘Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems’, *Applied Numerical Mathematics*.
- Barbič, J. & Zhao, Y. 2011, ‘Real-time large-deformation substructuring’, *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4, p. 91.
- Barrault, M., Maday, Y., Nguyen, N.C. & Patera, A.T. 2004, ‘An “empirical interpolation” method: application to efficient reduced-basis discretization of partial differential equations’, *Comptes Rendus Mathématique*, vol. 339, no. 9, pp. 667–72.
- Basava, R.R. 2015, ‘Meshfree Image-Based Reduced Order Modeling of Multiple Muscle Components with Connective Tissue and Fat’, UC San Diego.
- Basava, R.R., Chen, J.S., Zhang, Y., Sinha, S., Sinha, U., Hodgson, J., Csapo, R. & Malis, V. 2014, ‘Pixel based meshfree modeling of skeletal muscles’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.
- Bathe, K.J. & Dong, J. 2014, ‘Component mode synthesis with subspace iterations for controlled accuracy of frequency and mode shape solutions’, *Computers and Structures*.
- Baur, U., Benner, P. & Feng, L. 2014, ‘Model Order Reduction for Linear and Nonlinear Systems: A System-Theoretic Perspective’, *Archives of Computational Methods in Engineering*.
- Belhumeur, P.N., Hespanha, J.P. & Kriegman, D.J. 1997, ‘Eigenfaces vs. fisherfaces: Recognition using class specific linear projection’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–20.

- Belkin, M. & Niyogi, P. 2001, ‘Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering’, *Nips*, vol. 14, pp. 585–91.
- Belkin, M. & Niyogi, P. 2003, ‘Laplacian Eigenmaps for Dimensionality Reduction and Data Representation’, *Neural Computation*, vol. 15, no. 6, pp. 1373–96.
- Bellman, R.E. 1961, ‘Adaptive control processes: A guided tour’, *Princeton University Press*.
- Belytschko, T., Lu, Y.Y. & Gu, L. 1994, ‘Element- free Galerkin methods’, *International Journal for Numerical Methods in Engineering*, vol. 37, no. 2, pp. 229–56.
- Belytschko, T.B., Krongauz, Y., Organ, D., Fleming, M. & Krysl, P. 1996, ‘Meshless Methods: An Overview and Recent Developments’, *Computer Method in Applied Mechanics and Engineering*, vol. 139, pp. 3–47.
- Bengio, Y., Paiement, J.-F., Vincent, P., Delalleau, O., Roux, N. Le & Ouimet, M. 2003, ‘Out-of-sample extensions for lle, isomap, mds, eigenmaps, and spectral clustering’, *Advances in Neural Information Processing Systems*, no. April 2016, pp. 177–84.
- Benner, P., Gugercin, S. & Willcox, K. 2015, ‘A survey of projection-based model reduction methods for parametric dynamical systems’, *SIAM Review*, vol. 57, no. 4, pp. 1–49.
- Berkooz, G., Holmes, P. & Lumley, J.L. 1993, ‘The proper orthogonal decomposition in the analysis of turbulent flows’, *Annual Review of Fluid Mechanics*.
- Bessa, M.A., Bostanabad, R., Liu, Z., Hu, A., Apley, D.W., Brinson, C., Chen, W. & Liu, W.K. 2017, ‘A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality’, *Computer Methods in Applied Mechanics and Engineering*.
- Bhattacharjee, S. & Matouš, K. 2016, ‘A nonlinear manifold-based reduced order model for multiscale analysis of heterogeneous hyperelastic materials’, *Journal of Computational Physics*, vol. 313, pp. 635–53.
- Bishop, C.M. 2006, *Pattern Recognition and Machine Learning*, *Pattern Recognition*, Springer.
- Bistrián, D.A. & Susan-Resiga, R.F. 2016, ‘Weighted proper orthogonal decomposition of the swirling flow exiting the hydraulic turbine runner’, *Applied Mathematical Modelling*.
- Bonne, M. & Constantinescu, A. 2005, ‘Inverse problems in elasticity’, *Inverse Problems*, vol. 21, no. 2, pp. R1–50.
- Booker, A.J., Dennis, J.E., Frank, P.D., Serafini, D.B., Torczon, V. & Trosset, M.W. 1999, ‘A rigorous framework for optimization of expensive functions by surrogates’, *Structural Optimization*.
- Boyd, S. & Vandenberghe, L. 2004, *Convex optimization*, Cambridge university press.

- Brunton, S.L., Proctor, J.L. & Kutz, J.N. 2016, ‘Discovering governing equations from data: Sparse identification of nonlinear dynamical systems’, *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–7.
- Brunton, S.L., Tu, J.H., Bright, I. & Kutz, J.N. 2014, ‘Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems’, *SIAM Journal on Applied Dynamical Systems*, vol. 13, no. 4, pp. 1716–32.
- Bui-Thanh, T., Willcox, K. & Ghattas, O. 2008, ‘Model reduction for large-scale systems with high-dimensional parametric input space’, *SIAM Journal on Scientific Computing*.
- Burkardt, J., Gunzburger, M. & Lee, H.C. 2006, ‘POD and CVT-based reduced-order modeling of Navier-Stokes flows’, *Computer Methods in Applied Mechanics and Engineering*.
- Cai, D., He, X., Han, J. & others 2007, ‘Isometric projection’, *AAAI*, pp. 528–33.
- Cardoso, M.A. & Durlofsky, L.J. 2010, ‘Linearized reduced-order models for subsurface flow simulation’, *Journal of Computational Physics*.
- Carlberg, K., Barone, M. & Antil, H. 2017, ‘Galerkin v. least-squares Petrov–Galerkin projection in nonlinear model reduction’, *Journal of Computational Physics*, vol. 330, pp. 693–734.
- Carlberg, K., Bou-Mosleh, C. & Farhat, C. 2011, ‘Efficient non-linear model reduction via a least-squares Petrov-Galerkin projection and compressive tensor approximations’, *International Journal for Numerical Methods in Engineering*, vol. 86, no. 2, pp. 155–81.
- Carlberg, K. & Farhat, C. 2008, ‘A compact proper orthogonal decomposition basis for optimization-oriented reduced-order models’, *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, p. 5964.
- Carlberg, K. & Farhat, C. 2011, ‘A low-cost, goal-oriented “compact proper orthogonal decomposition” basis for model reduction of static systems’, *International Journal for Numerical Methods in Engineering*, vol. 86, no. 3, pp. 381–402.
- Carlberg, K., Farhat, C., Cortial, J. & Amsallem, D. 2013, ‘The GNAT method for nonlinear model reduction: Effective implementation and application to computational fluid dynamics and turbulent flows’, *Journal of Computational Physics*, vol. 242, pp. 623–47.
- Carlberg, K., Ray, J. & van Bloemen Waanders, B. 2015, ‘Decreasing the temporal complexity for nonlinear, implicit reduced-order models by forecasting’, *Computer Methods in Applied Mechanics and Engineering*.
- Carlberg, K.T. 2011, *Model reduction of nonlinear mechanical systems via optimal projection and tensor approximation*, Stanford University.
- Cevikalp, H. & Triggs, W. 2010, ‘Face recognition based on image sets’, *Computer Vision and Pattern Recognition (CVPR)*, vol. 365, IEEE Conference, pp. 2567–73.

- Chagnon, G., Rebouah, M. & Favier, D. 2015, 'Hyperelastic Energy Densities for Soft Biological Tissues: A Review', *Journal of Elasticity*.
- Chatterjee, A. 2000, 'An introduction to the proper orthogonal decomposition', *Current science*.
- Chaturantabut, S. & Sorensen, D.C. 2010, 'Nonlinear Model Reduction via Discrete Empirical Interpolation', *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2737–64.
- Chaturantabut, S. & Sorensen, D.C. 2011, 'Application of POD and DEIM on dimension reduction of non-linear miscible viscous fingering in porous media', *Mathematical and Computer Modelling of Dynamical Systems*.
- Chen, D. & Plemmons, R.J. 2010, 'Nonnegativity constraints in numerical analysis', *The birth of numerical analysis*, World Scientific, pp. 109–39.
- Chen, J.-S., Hillman, M. & Chi, S.-W. 2017, 'Meshfree Methods: Progress Made after 20 Years', *Journal of Engineering Mechanics*, vol. 143, no. 4, p. 04017001.
- Chen, J.-S., Pan, C., Wu, C.-T. & Liu, W.K. 1996, 'Reproducing Kernel Particle Methods for large deformation analysis of non-linear structures', *Computer Methods in Applied Mechanics and Engineering*, vol. 139, no. 1–4, pp. 195–227.
- Chen, J.-S., Yoon, S. & Wu, C.-T. 2002, 'Non-linear version of stabilized conforming nodal integration for Galerkin mesh-free methods', *International Journal for Numerical Methods in Engineering*, vol. 53, no. 12, pp. 2587–615.
- Chen, J., Wu, C. & Yoon, S. 2001, 'A stabilized conforming nodal integration for Galerkin mesh-free methods', *Int. J. Numer. Meth. Eng.*, vol. 50, no. 2, pp. 435–66.
- Chen, J.S., Basava, R.R., Zhang, Y., Csapo, R., Malis, V., Sinha, U., Hodgson, J. & Sinha, S. 2016, 'Pixel-based meshfree modelling of skeletal muscles', *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*.
- Chen, J.S., Han, W., Wu, C.T. & Duan, W. 1997, 'On the perturbed Lagrangian formulation for nearly incompressible and incompressible hyperelasticity', *Computer Methods in Applied Mechanics and Engineering*.
- Chen, J.S., Marodon, C. & Hu, H.Y. 2015, 'Model order reduction for meshfree solution of Poisson singularity problems', *International Journal for Numerical Methods in Engineering*, vol. 102, no. 5, pp. 1211–37.
- Chen, J.S. & Wang, H.P. 2000, 'New boundary condition treatments in meshfree computation of contact problems', *Computer Methods in Applied Mechanics and Engineering*.
- Chen, J.S., Wang, H.P., Yoon, S. & You, Y. 2000, 'Some recent improvements in meshfree methods for incompressible finite elasticity boundary value problems with contact', *Computational Mechanics*.

- Chen, J.S., Yoon, S., Wang, H.P. & Liu, W.K. 2000, 'An improved reproducing kernel particle method for nearly incompressible finite elasticity', *Computer Methods in Applied Mechanics and Engineering*.
- Chen, Y. & White, J. 2000, 'A Quadratic Method for Nonlinear Model Order Reduction', *Technical Proceedings of the 2000 International Conference on Modeling and Simulation of Microsystems*, pp. 477–80.
- Chinesta, F., Ammar, A. & Cueto, E. 2010, 'Recent Advances and New Challenges in the Use of the Proper Generalized Decomposition for Solving Multidimensional Models', *Archives of Computational Methods in Engineering*, vol. 17, no. 4, pp. 327–50.
- Christensen, E., Brøns, M. & Sørensen, J. 1999, 'Evaluation of Proper Orthogonal Decomposition--Based Decomposition Techniques Applied to Parameter-Dependent Nonturbulent Flows', *SIAM Journal on Scientific Computing*, vol. 21, no. 4, pp. 1419--1434.
- Chung, F.R. & Graham, F.C. 1997, *Spectral Graph Theory*, American Mathematical Soc.
- Conti, S., Müller, S. & Ortiz, M. 2018, 'Data-Driven Problems in Elasticity', *Archive for Rational Mechanics and Analysis*, vol. 229, no. 1, pp. 79–123.
- Cosimo, A., Cardona, A. & Idelsohn, S. 2014, 'Improving the k-compressibility of Hyper Reduced Order Models with moving sources: Applications to welding and phase change problems', *Computer Methods in Applied Mechanics and Engineering*.
- Cosimo, A., Cardona, A. & Idelsohn, S. 2016, 'General treatment of essential boundary conditions in reduced order models for non-linear problems', *Advanced Modeling and Simulation in Engineering Sciences*.
- Cover, T. & Hart, P. 1967, 'Nearest neighbor pattern classification', *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–7.
- Craig, R.R. 1981, *Structural dynamics: an introduction to computer methods*, John Wiley & Sons Inc.
- Cueto, E. & Chinesta, F. 2014, 'Real time simulation for computational surgery: a review', *Advanced Modeling and Simulation in Engineering Sciences*, vol. 1, no. 1, pp. 1–10.
- Darema, F. & Rotea, M. 2006, 'Dynamic data-driven applications systems', *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, SC'06*.
- Dickens, J.M., Nakagawa, J.M. & Wittbrodt, M.J. 1997, 'A critique of mode acceleration and modal truncation augmentation methods for modal response analysis', *Computers & Structures*, vol. 62, no. 6, pp. 985–98.
- Drmac, Z. & Gugercin, S. 2016, 'A new selection operator for the discrete empirical

- interpolation method---Improved a priori error bound and extensions’, *SIAM Journal on Scientific Computing*, vol. 38, no. 2, pp. A631--A648.
- Eftang, J.L. & Stamm, B. 2012, ‘Parameter multi-domain “hp” empirical interpolation’, *International Journal for Numerical Methods in Engineering*, vol. 90, no. 4, pp. 412–28.
- Eldred, M.S. & Dunlavy, D.M. 2006, ‘Formulations for Surrogate-Based Optimization with Data Fit, Multifidelity, and Reduced-Order Models’, *11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Portsmouth, Virginia*, no. September, p. AIAA-2006-7117.
- England, A.H. 2003, *Complex Variable Methods in Elasticity*, Courier Corporation.
- Evensen, G. 2010, *Data assimilation: The ensemble kalman filter*, *Data Assimilation: The Ensemble Kalman Filter*, Springer Science & Business Media.
- Everson, R. & Sirovich, L. 1995, ‘Karhunen-Loeve Procedure for Gappy Data’, *Journal of the Optical Society of America a-Optics Image Science and Vision*, vol. 12, no. 8, pp. 1657–64.
- Farhat, C., Avery, P., Chapman, T. & Cortial, J. 2014, ‘Dimensional reduction of nonlinear finite element dynamic models with finite rotations and energy-based mesh sampling and weighting for computational efficiency’, *International Journal for Numerical Methods in Engineering*, vol. 98, no. 9, pp. 625–62.
- Farhat, C., Chapman, T. & Avery, P. 2015, ‘Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models’, *International Journal for Numerical Methods in Engineering*, vol. 102, no. 5, pp. 1077–110.
- Fernández-Méndez, S. & Huerta, A. 2004, ‘Imposing essential boundary conditions in mesh-free methods’, *Computer Methods in Applied Mechanics and Engineering*, vol. 193, no. 12–14, pp. 1257–75.
- Feyel, F. 2003, ‘A multilevel finite element method (FE2) to describe the response of highly non-linear structures using generalized continua’, *Computer Methods in Applied Mechanics and Engineering*, vol. 192, no. 28–30, pp. 3233–44.
- Forrester, A.I.J. 2008, *Engineering Design via Surrogate Modelling: A Practical Guide - Constructing a Surrogate*, *Engineering Design via Surrogate Modelling: A Practical Guide*.
- Fung, Y.C. & Skalak, R. 1981, *Biomechanics: Mechanical Properties of Living Tissues*, *Journal of Biomechanical Engineering*, vol. 103, Springer Science & Business Media.
- Galland, F., Gravouil, A., Malvesin, E. & Rochette, M. 2011, ‘A global model reduction approach for 3D fatigue crack growth with confined plasticity’, *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 5–8, pp. 699–716.

- Ganapathysubramanian, B. & Zabarar, N. 2008, 'A non-linear dimension reduction methodology for generating data-driven stochastic input models', *Journal of Computational Physics*, vol. 227, no. 13, pp. 6612–37.
- Georgiou, G.C., Olson, L. & Smyrlis, Y.S. 1996, 'A singular function boundary integral method for the Laplace equation', *Communications in Numerical Methods in Engineering*, vol. 12, no. August 1995, pp. 127–34.
- Ghaboussi, J., Garrett, J.H. & Wu, X. 1991, 'Knowledge- Based Modeling of Material Behavior with Neural Networks', *Journal of Engineering Mechanics*, vol. 117, no. 1, pp. 132–53.
- Ghaboussi, J. & Sidarta, D.E. 1998, 'New Nested Adaptive Neural Networks (NANN) for Constitutive Modeling', *Computers and Geotechnics*, vol. 22, no. 1, pp. 29–52.
- Ghanem, R.G. & Spanos, P.D. 1991, *Stochastic Finite Elements: A Spectral Approach*, Dover publications.
- Ghavamian, F., Tiso, P. & Simone, A. 2017, 'POD–DEIM model order reduction for strain softening viscoplasticity', *Computer Methods in Applied Mechanics and Engineering*, vol. 317, pp. 458–79.
- Gill, P.E., Murray, W. & Wright, M.H. 1981, *Practical optimization*, Emerald Group Publishing Limited.
- Gillies, A.R. & Lieber, R.L. 2011, 'Structure and function of the skeletal muscle extracellular matrix', *Muscle {&} nerve*.
- González, D., Aguado, J. V., Cueto, E., Abisset-Chavanne, E. & Chinesta, F. 2018, 'kPCA-Based Parametric Solutions Within the PGD Framework', *Archives of Computational Methods in Engineering*.
- González, D., Chinesta, F. & Cueto, E. 2018, 'Thermodynamically consistent data-driven computational mechanics', *Continuum Mechanics and Thermodynamics*, pp. 1–15.
- González, D., Cueto, E. & Chinesta, F. 2016, 'Computational Patient Avatars for Surgery Planning', *Annals of Biomedical Engineering*.
- Goodfellow, I., Bengio, Y., Courville, A. & Bengio, Y. 2016, *Deep learning*, vol. 1, MIT press Cambridge.
- Goodpaster, B.H., Park, S.W., Harris, T.B., Kritchevsky, S.B., Nevitt, M., Schwartz, A. V., Simonsick, E.M., Tylavsky, F.A., Visser, M. & Newman, A.B. 2006, 'The loss of skeletal muscle strength, mass, and quality in older adults: The Health, Aging and Body Composition Study', *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*.
- Gorissen, D., Couckuyt, I., Demeester, P., Dhaene, T. & Crombecq, K. 2010, 'A surrogate

- modeling and adaptive sampling toolbox for computer based design', *The Journal of Machine Learning Research*.
- Goury, O., Amsallem, D., Bordas, S.P.A., Liu, W.K. & Kerfriden, P. 2016, 'Automatised selection of load paths to construct reduced-order models in computational damage micromechanics: from dissipation-driven random selection to Bayesian optimization', *Computational Mechanics*.
- Gray, H. & Carter, H. V 1995, *Anatomy: Descriptive and Surgical*, Parker and Son.
- Grepl, M. a., Maday, Y., Nguyen, N.C. & Patera, A.T. 2007, 'Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations', *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 41, no. 3, pp. 575–605.
- Grepl, M.A. & Patera, A.T. 2005, 'A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations', *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 39, no. 1, pp. 157–81.
- Gugercin, S. & Antoulas, A.C. 2004, 'A survey of model reduction by balanced truncation and some new results', *International Journal of Control*.
- Gunzburger, M.D., Peterson, J.S. & Shadid, J.N. 2007, 'Reduced-order modeling of time-dependent PDEs with multiple parameters in the boundary data', *Computer Methods in Applied Mechanics and Engineering*, vol. 196, no. 4–6, pp. 1030–47.
- Gupta, A., Cecen, A., Goyal, S., Singh, A.K. & Kalidindi, S.R. 2015, 'Structure-property linkages using a data science approach: Application to a non-metallic inclusion/steel composite system', *Acta Materialia*, vol. 91, no. 0, pp. 239–54.
- Haasdonk, B., Dihlmann, M. & Ohlberger, M. 2011, 'A training set and multiple bases generation approach for parameterized model reduction based on adaptive grids in parameter space', *Mathematical and Computer Modelling of Dynamical Systems*, vol. 17, no. 4, pp. 423–42.
- Haasdonk, B. & Ohlberger, M. 2008, 'Reduced basis method for finite volume approximations of parametrized linear evolution equations', *ESAIM: Mathematical Modelling and Numerical Analysis*.
- Haller, G. & Ponsioen, S. 2017, 'Exact model reduction by a slow–fast decomposition of nonlinear mechanical systems', *Nonlinear Dynamics*, vol. 90, no. 1, pp. 617–47.
- Ham, J.H., Lee, D.D., Mika, S. & Schölkopf, B. 2004, 'A kernel view of the dimensionality reduction of manifolds', *Proceedings of the 21st International Conference on Machine Learning*, no. 47, p. 47.
- Hambli, R., Katerchi, H. & Benhamou, C.-L. 2011, 'Multiscale methodology for bone remodelling simulation using coupled finite element and neural network computation',

- Biomechanics and Modeling in Mechanobiology*, vol. 10, no. 1, pp. 133–45.
- Hastie, T., Tibshirani, R. & Friedman, J. 2009, *The Elements of Statistical Learning*, vol. 1, Springer series in statistics, New York.
- Haykin, S. 2009, *Neural Networks and Learning Machines*, Pearson Prentice Hall New Jersey USA 936 pLinks, vol. 3, Pearson, NJ, USA.
- He, Q., Chen, J.-S. & Marodon, C. 2018, ‘A decomposed subspace reduction for fracture mechanics based on the meshfree integrated singular basis function method’, *Computational Mechanics*.
- He, X. & Niyogi, P. 2004, ‘Locality preserving projections’, *Neural information processing systems*, vol. 16, p. 153.
- He, X., Yan, S., Hu, Y., Niyogi, P. & Zhang, H.-J.J. 2005, ‘Face recognition using laplacianfaces’, *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 3, pp. 328–40.
- Hernández, J.A., Caicedo, M.A. & Ferrer, A. 2017, ‘Dimensional hyper-reduction of nonlinear finite element models via empirical cubature’, *Computer Methods in Applied Mechanics and Engineering*, vol. 313, pp. 687–722.
- Hernandez, J.A., Oliver, J., Huespe, A.E., Caicedo, M.A. & Cante, J.C. 2014, ‘High-performance model reduction techniques in computational multiscale homogenization’, *Computer Methods in Applied Mechanics and Engineering*, vol. 276, pp. 149–89.
- Hesthaven, J.S., Rozza, G. & Stamm, B. 2015, *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, Springer.
- Hochman, A., Bond, B.N. & White, J.K. 2011, ‘A stabilized discrete empirical interpolation method for model reduction of electrical, thermal, and microelectromechanical systems’, *Proceedings of the 48th Design Automation Conference on - DAC '11*.
- Hoerl, A.E. & Kennard, R.W. 1970, ‘Ridge Regression: Biased Estimation for Nonorthogonal Problems’, *Technometrics*, vol. 12, no. 1, pp. 55–67.
- Holzapfel, G.A., Gasser, T.C. & Ogden, R.W. 2000, ‘A new constitutive framework for arterial wall mechanics and a comparative study of material models’, *Journal of Elasticity*.
- Humphrey, J.D. 2003, ‘Continuum biomechanics of soft biological tissues’, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*.
- Humphrey, J.D., Vawter, D.L. & Vito, R.P. 1987, ‘Quantification of strains in biaxially tested soft tissues’, *Journal of Biomechanics*.

- Humphrey, J.D. & Yin, F. 1987, 'On constitutive relations and finite deformations of passive cardiac tissue - I. A pseudostrain-energy function', *Journal of Biomechanical Engineering*.
- Hurty, W.C. 1960, 'Vibrations of Structural Systems by Component Mode Synthesis', *Journal of the Engineering Mechanics Division*, vol. 86, no. 4, pp. 51–70.
- Hurty, W.C. 1965, 'Dynamic analysis of structural systems using component modes', *AIAA Journal*, vol. 3, no. 4, pp. 678–85.
- Ibañez, R., Abisset-Chavanne, E., Aguado, J.V., Gonzalez, D., Cueto, E. & Chinesta, F. 2016, 'A Manifold Learning Approach to Data-Driven Computational Elasticity and Inelasticity', *Archives of Computational Methods in Engineering*, pp. 1–11.
- Ibañez, R., Borzacchiello, D., Aguado, J.V., Abisset-Chavanne, E., Cueto, E., Ladeveze, P. & Chinesta, F. 2017, 'Data-driven non-linear elasticity: constitutive manifold construction and problem discretization', *Computational Mechanics*, pp. 1–14.
- Idelsohn, S.R. & Cardona, A. 1985, 'A reduction method for nonlinear structural dynamic analysis', *Computer Methods in Applied Mechanics and Engineering*, vol. 49, no. 3, pp. 253–79.
- Ito, K. & Ravindran, S.S. 1998, 'A Reduced-Order Method for Simulation and Control of Fluid Flows', *Journal of Computational Physics*.
- Jett, S., Laurence, D., Kunkel, R., Babu, A.R., Kramer, K., Baumwart, R., Towner, R., Wu, Y. & Lee, C.-H. 2018, 'An investigation of the anisotropic mechanical properties and anatomical structure of porcine atrioventricular heart valves', *Journal of the mechanical behavior of biomedical materials*, vol. 87, pp. 155–71.
- Jolliffe, I. T. 2002, *Principal Component Analysis, Springer Series in Statistics*, Springer-Verlag, New York.
- Jordan, M.I. & Mitchell, T.M. 2015, 'Machine learning: Trends, perspectives, and prospects', *Science*.
- Kalidindi, S.R., Niezgodá, S.R. & Salem, A.A. 2011, 'Microstructure informatics using higher-order statistics and efficient data-mining protocols', *JOM*, vol. 63, no. 4, pp. 34–41.
- Kanno, Y. 2017, 'Data-Driven Approach to Computational Mechanics via Local Robust Regression', *arXiv preprint arXiv:1708.05794*, no. ii, pp. 1–15.
- Karhunen, K. 1946, 'Zur spektraltheorie stochastischer', *Annales Academiae Scientiarum Fennicae Series A*, vol. 1, p. 34.
- Kennedy, M.C. & O'Hagan, A. 2001, 'Bayesian calibration of computer models', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*.

- Kerfriden, P., Gosselet, P., Adhikari, S. & Bordas, S. 2011, 'Bridging Proper Orthogonal Decomposition methods and augmented Newton-Krylov algorithms: an adaptive model order reduction for highly nonlinear mechanical problems', *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 5, pp. 850–66.
- Kerfriden, P., Goury, O., Rabczuk, T. & Bordas, S.P.A. 2013, 'A partitioned model order reduction approach to rationalise computational expenses in nonlinear fracture mechanics', *Computer Methods in Applied Mechanics and Engineering*, vol. 256, pp. 169–88.
- Kerfriden, P., Passieux, J.C. & Bordas, S.P.A. 2012, 'Local/global model order reduction strategy for the simulation of quasi-brittle fracture', *International Journal for Numerical Methods in Engineering*, vol. 89, no. 2, pp. 154–79.
- Kerschen, G., Golinval, J.C., Vakakis, A.F. & Bergman, L.A. 2005, 'The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: An overview', *Nonlinear Dynamics*, vol. 41, no. 1–3, pp. 147–69.
- Kim, J.G., Park, Y.J., Lee, G.H. & Kim, D.N. 2017, 'A general model reduction with primal assembly in structural dynamics', *Computer Methods in Applied Mechanics and Engineering*.
- Kim, T. & James, D.L. 2012, 'Physics-based character skinning using multidomain subspace deformations', *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, pp. 1228–40.
- Kirchdoerfer, T. & Ortiz, M. 2016, 'Data-driven computational mechanics', *Computer Methods in Applied Mechanics and Engineering*, vol. 304, pp. 81–101.
- Kirchdoerfer, T. & Ortiz, M. 2017, 'Data Driven Computing with noisy material data sets', *Computer Methods in Applied Mechanics and Engineering*, vol. 326, pp. 622–41.
- Kirchdoerfer, T. & Ortiz, M. 2018, 'Data-driven computing in dynamics', *International Journal for Numerical Methods in Engineering*, vol. 113, no. 11, pp. 1697–710.
- Klerk, D. De, Rixen, D.J. & Voormeeren, S.N. 2008, 'General Framework for Dynamic Substructuring: History, Review and Classification of Techniques', *AIAA Journal*.
- Koren, Y. & Carmel, L. 2004, 'Robust linear dimensionality reduction', *IEEE Transactions on Visualization and Computer Graphics*, vol. 10, pp. 459–70.
- Koren, Y., Carmel, L. & Harel, D. 2002, 'ACE: A fast multiscale eigenvectors computation for drawing huge graphs', *Proceedings - IEEE Symposium on Information Visualization, INFO VIS*.
- Koscianski, A. & De Cursi, S. 2005, 'Physically Constrained Neural Networks and Regularization of Inverse Problems', *6 th World Congresses of Structural and Multidisciplinary Optimization*.

- Kouznetsova, V., Brekelmans, W.A.M. & Baaijens, F.P.T. 2001, ‘Approach to micro-macro modeling of heterogeneous materials’, *Computational Mechanics*.
- Krysl, P., Lall, S. & Marsden, J.E. 2001, ‘Dimensional model reduction in non-linear finite element dynamics of solids and structures’, *International Journal for Numerical Methods in Engineering*, vol. 51, no. 4, pp. 479–504.
- Ladevèze, P., Passieux, J.C. & Néron, D. 2010, ‘The LATIN multiscale computational method and the Proper Generalized Decomposition’, *Computer Methods in Applied Mechanics and Engineering*, vol. 199, no. 21–22, pp. 1287–96.
- Lall, S., Krysl, P. & Marsden, J.E. 2003, ‘Structure-preserving model reduction for mechanical systems’, *Physica D: Nonlinear Phenomena*, vol. 184, no. 1–4, pp. 304–18.
- Larose, D.T. 2014, *Discovering knowledge in data: an introduction to data mining*, Statistics, John Wiley & Sons.
- Lawrence, N.D. 2012, ‘A Unifying Probabilistic Perspective for Spectral Dimensionality Reduction: Insights and New Models’, *Journal of Machine Learning Research*, vol. 13, pp. 1609–38.
- Lawson, C.L. & Hanson, R.J. 1987, *Solving Least Squares Problems*, Society for Industrial and Applied Mathematics.
- Le, B.A., Yvonnet, J. & He, Q.-C. 2015, ‘Computational homogenization of nonlinear elastic materials using neural networks’, *International Journal for Numerical Methods in Engineering*, vol. 104, no. 12, pp. 1061–84.
- Lee, C.H. & Chen, J.S. 2013, ‘Proper orthogonal decomposition-based model order reduction via radial basis functions for molecular dynamics systems’, *International Journal for Numerical Methods in Engineering*.
- Lee, J. & Cho, M. 2017, ‘An interpolation-based parametric reduced order model combined with component mode synthesis’, *Computer Methods in Applied Mechanics and Engineering*.
- Lee, J.A. & Verleysen, M. 2007, *Nonlinear Dimensionality Reduction, Machine Learning*, Springer Science & Business Media.
- Lefik, M. & Schrefler, B.A. 2003, ‘Artificial neural network as an incremental non-linear constitutive model for a finite element code’, *Computer Methods in Applied Mechanics and Engineering*, vol. 192, no. 28–30, pp. 3265–83.
- Leontaritis, I.J. & Billings, S.A. 1985, ‘Input-output parametric models for non-linear systems Part I: Deterministic non-linear systems’, *International Journal of Control*.
- LeVeque, R.J. 2002, *Finite Volume Methods for Hyperbolic Problems*, vol. 31, Cambridge university press.

- Lewis, R. & Nash, S. 2000, ‘A multigrid approach to the optimization of systems governed by differential equations’, *8th Symposium on Multidisciplinary Analysis and Optimization*.
- Lewis, R.M. & Nash, S.G. 2005, ‘Model problems for the multigrid optimization of systems governed by differential equations’, *Siam Journal On Scientific Computing*.
- Leygue, A., Coret, M., Réthoré, J., Stainier, L. & Verron, E. 2018, ‘Data-based derivation of material response’, *Computer Methods in Applied Mechanics and Engineering*, vol. 331, pp. 184–96.
- Li, Z.-C., Mathon, R. & Sermer, P. 1987, ‘Boundary methods for solving elliptic problems with singularities and interfaces’, *SIAM Journal on Numerical Analysis*, vol. 24, no. 3, pp. 487–98.
- Liang, Y.C., Lee, H.P., Lim, S.P., Lin, W.Z., Lee, K.H. & Wu, C.G. 2002, ‘Proper orthogonal decomposition and its applications - Part I: Theory’, *Journal of Sound and Vibration*, vol. 252, no. 3, pp. 527–44.
- Lieu, T., Farhat, C. & Lesoinne, M. 2006, ‘Reduced-order fluid/structure modeling of a complete aircraft configuration’, *Computer Methods in Applied Mechanics and Engineering*, vol. 195, no. 41–43, pp. 5730–42.
- Lieu, T. & Lesoinne, M. 2004, ‘Parameter adaptation of reduced order models for three-dimensional flutter analysis’, *AIAA Paper*, no. January, pp. 1–9.
- Ling, J., Jones, R. & Templeton, J. 2016, ‘Machine learning strategies for systems with invariance properties’, *Journal of Computational Physics*, vol. 318, pp. 22–35.
- Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B. & Sánchez, C.I. 2017, ‘A survey on deep learning in medical image analysis’, *Medical Image Analysis*.
- Litjens, G., Sánchez, C.I., Timofeeva, N., Hermsen, M., Nagtegaal, I., Kovacs, I., Hulsbergen-Van De Kaa, C., Bult, P., Van Ginneken, B. & Van Der Laak, J. 2016, ‘Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis’, *Scientific Reports*.
- Liu, W.K., Jun, S. & Zhang, Y.F. 1995, ‘Reproducing kernel particle methods’, *International Journal for Numerical Methods in Fluids*, vol. 20, no. 8–9, pp. 1081–106.
- Liu, Z., Bessa, M.A. & Liu, W.K. 2016, ‘Self-consistent clustering analysis: An efficient multi-scale scheme for inelastic heterogeneous materials’, *Computer Methods in Applied Mechanics and Engineering*.
- Loève, M. 1955, *Probability theory: foundations, random sequences*, van Nostrand Princeton, NJ.

- Lopez, E., Gonzalez, D., Aguado, J. V., Abisset-Chavanne, E., Cueto, E., Binetruy, C. & Chinesta, F. 2018, 'A Manifold Learning Approach for Integrated Computational Materials Engineering', *Archives of Computational Methods in Engineering*.
- Von Luxburg, U. 2007, 'A tutorial on spectral clustering', *Statistics and Computing*, vol. 17, no. 4, pp. 395–416.
- Van Der Maaten, L.J.P., Postma, E.O. & Van Den Herik, H.J. 2009, 'Dimensionality Reduction: A Comparative Review', *Journal of Machine Learning Research*, vol. 10, pp. 1–41.
- MacKay, D.J.C. & Mac Kay, D.J.C. 2003, *Information theory, inference and learning algorithms*, Cambridge university press.
- March, A. & Willcox, K. 2012, 'Provably Convergent Multifidelity Optimization Algorithm Not Requiring High-Fidelity Derivatives', *AIAA Journal*.
- Martinez, A.M. & Kak, A.C. 2001, 'PCA versus LDA', *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Matouš, K., Geers, M.G.D., Kouznetsova, V.G. & Gillman, A. 2017, 'A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials', *Journal of Computational Physics*.
- Matthews, G.D.K., Huang, C.L.H., Sun, L. & Zaidi, M. 2011, 'Translational musculoskeletal science: Is sarcopenia the next clinical target after osteoporosis?', *Annals of the New York Academy of Sciences*.
- Michalski, R.S. 1986, 'Understanding the Nature of Learning: Issues and Research Directions', *Machine Learning: An Artificial Intelligence Approach (vol. 2)*.
- Milano, M. & Koumoutsakos, P. 2002, 'Neural network modeling for near wall turbulent flow', *Journal of Computational Physics*, vol. 182, no. 1, pp. 1–26.
- Millán, D. & Arroyo, M. 2013, 'Nonlinear manifold learning for model reduction in finite elastodynamics', *Computer Methods in Applied Mechanics and Engineering*, vol. 261, pp. 118–32.
- Min, W., Lu, K. & He, X. 2004, 'Locality pursuit embedding', *Pattern Recognition*, vol. 37, no. 4, pp. 781–8.
- Murphy, K.P. 2012, *Machine Learning: A Probabilistic Perspective*, MIT Press, Cambridge: MIT press, Cambridge.
- Nguyen, L.T.K. & Keip, M.-A. 2018, 'A data-driven approach to nonlinear elasticity', *Computers & Structures*, vol. 194, pp. 97–115.
- Nguyen, N.C. 2007, 'A posteriori error estimation and basis adaptivity for reduced-basis

- approximation of nonaffine-parametrized linear elliptic partial differential equations’, *Journal of Computational Physics*, vol. 227, no. 2, pp. 983–1006.
- Nguyen, N.C., Patera, A.T. & Peraire, J. 2008, ‘A “best points” interpolation method for efficient approximation of parametrized functions’, *International Journal for Numerical Methods in Engineering*.
- Ni, K.S. & Nguyen, T.Q. 2009, ‘An adaptable k-nearest neighbors algorithm for MMSE image interpolation’, *IEEE Transactions on Image Processing*, vol. 18, no. 9, pp. 1976–87.
- Niroomandi, S., Alfaro, I., Cueto, E. & Chinesta, F. 2008, ‘Real-time deformable models of non-linear tissues by model reduction techniques’, *Computer Methods and Programs in Biomedicine*, vol. 91, no. 3, pp. 223–31.
- Niroomandi, S., Alfaro, I., Cueto, E. & Chinesta, F. 2010, ‘Model order reduction for hyperelastic materials’, *International Journal for Numerical Methods in Engineering*, vol. 81, no. 9, pp. 1180–206.
- Niroomandi, S., Alfaro, I., González, D., Cueto, E. & Chinesta, F. 2012, ‘Real-time simulation of surgery by reduced-order modeling and X-FEM techniques’, *International Journal for Numerical Methods in Biomedical Engineering*, vol. 28, no. 5, pp. 574–88.
- Nitsche, J. 1971, ‘Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind’, *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 36, no. 1, pp. 9–15.
- Nouy, A. 2010, ‘A priori model reduction through Proper Generalized Decomposition for solving time-dependent partial differential equations’, *Computer Methods in Applied Mechanics and Engineering*, vol. 199, no. 23–24, pp. 1603–26.
- NPUAP/EPUAP 2009, ‘Prevention and Treatment of Pressure Ulcers: Clinical Practice Guideline’, *Washington, DC: NPUAP, National Pressure Ulcer Advisory Panel, Washington DC*.
- Oden, J.T., Belytschko, T., Fish, J., Hughes, T.J.R., Johnson, C., Keyes, D., Laub, A., Petzold, L., Srolovitz, D. & Yip, S. 2006, ‘Simulation-based engineering science: Revolutionizing engineering science through simulation’, *National Science Foundation*, pp. 1–88.
- Oishi, A. & Yagawa, G. 2017, ‘Computational mechanics enhanced by deep learning’, *Computer Methods in Applied Mechanics and Engineering*.
- Oliver, J., Caicedo, M., Huespe, A.E., Hernández, J.A. & Roubin, E. 2017, ‘Reduced order modeling strategies for computational multiscale fracture’, *Computer Methods in Applied Mechanics and Engineering*, vol. 313, pp. 560–95.
- Olson, G.B. 2000, ‘Designing a New Material World’, *Science*, vol. 288, no. 5468, pp. 993–8.

- Olson, L.G., Georgiou, G.C. & Schultz, W.W. 1991, 'An efficient finite element method for treating singularities in Laplace's equation', *J. Comput. Phys.*, vol. 96, no. 2, pp. 391–410.
- Pan, B., Qian, K., Xie, H. & Asundi, A. 2009, 'Two-dimensional digital image correlation for in-plane displacement and strain measurement: A review', *Measurement Science and Technology*.
- Park, J., Zhang, Z., Zha, H. & Kasturi, R. 2004, 'Local smoothing for manifold learning', *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 2, pp. 452–9.
- Passieux, J.C., Rethore, J., Gravouil, A. & Baietto, M.C. 2013, 'Local/global non-intrusive crack propagation simulation using a multigrid X-FEM solver', *Computational Mechanics*, vol. 52, no. 6, pp. 1381–93.
- Peherstorfer, B. 2013, 'Model order reduction of parametrized systems with sparse grid learning techniques', Technische Universität München.
- Peherstorfer, B., Butnaru, D., Willcox, K. & Bungartz, H. 2014, 'Localized discrete empirical interpolation method', *Journal of scientific computing*, vol. 36, no. 1, pp. 168–92.
- Peherstorfer, B. & Willcox, K. 2015, 'Dynamic data-driven reduced-order models', *Computer Methods in Applied Mechanics and Engineering*, vol. 291, pp. 21–41.
- Peherstorfer, B. & Willcox, K. 2016, 'Data-driven operator inference for nonintrusive projection-based model reduction', *Computer Methods in Applied Mechanics and Engineering*.
- Peng, L. & Mohseni, K. 2014, 'An Online Manifold Learning Approach for Model Reduction of Dynamical Systems', *SIAM Journal of Numerical Analysis*.
- Peng, L. & Mohseni, K. 2016, 'Nonlinear model reduction via a locally weighted POD method', *International Journal for Numerical Methods in Engineering*, vol. 106, no. 5, pp. 372–96.
- Peterson, J.S. 1989, 'The Reduced Basis Method for Incompressible Viscous Flow Calculations', *SIAM Journal on Scientific and Statistical Computing*.
- Phillips, J.R. 2000, 'Projection frameworks for model reduction of weakly nonlinear systems', *Proceedings of the 37th conference on Design automation - DAC '00*.
- Porsching, T.A. 1985, 'Estimation of the error in the reduced basis method solution of nonlinear equations', *Mathematics of Computation*, vol. 45, no. 172, pp. 487–96.
- Prud'homme, C., Rovas, D. V., Veroy, K., Machiels, L., Maday, Y., Patera, A.T. & Turinici, G. 2002, 'Reliable Real-Time Solution of Parametrized Partial Differential Equations: Reduced-Basis Output Bound Methods', *Journal of Fluids Engineering*.
- Quarteroni, A., Manzoni, A. & Negri, F. 2015, *Reduced basis methods for partial differential*

equations: an introduction, vol. 92, Springer.

- Quarteroni, A., Rozza, G. & Manzoni, A. 2011, ‘Certified reduced basis approximation for parametrized partial differential equations and applications’, *Journal of Mathematics in Industry*.
- Queipo, N. V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R. & Kevin Tucker, P. 2005, ‘Surrogate-based analysis and optimization’, *Progress in Aerospace Sciences*.
- Radermacher, A. & Reese, S. 2014, ‘Model reduction in elastoplasticity: proper orthogonal decomposition combined with adaptive sub-structuring’, *Computational Mechanics*, vol. 54, no. 3, pp. 677–87.
- Radermacher, A. & Reese, S. 2016, ‘POD-based model reduction with empirical interpolation applied to nonlinear elasticity’, *International Journal for Numerical Methods in Engineering*.
- Raissi, M. & Karniadakis, G.E. 2018, ‘Hidden physics models: Machine learning of nonlinear partial differential equations’, *Journal of Computational Physics*, vol. 357, pp. 125–41.
- Raissi, M., Perdikaris, P. & Karniadakis, G.E. 2017, ‘Inferring solutions of differential equations using noisy multi-fidelity data’, *Journal of Computational Physics*.
- Rajan, K. 2005, ‘Materials informatics’, *Materials Today*, vol. 8, no. 10, pp. 38–45.
- Rama, R.R. & Skatulla, S. 2018, ‘Towards real-time cardiac mechanics modelling with patient-specific heart anatomies’, *Computer Methods in Applied Mechanics and Engineering*, vol. 328, pp. 47–74.
- Rathinam, M. & Petzold, L.R. 2003, ‘A New Look at Proper Orthogonal Decomposition’, *SIAM Journal on Numerical Analysis*, vol. 41, no. 5, pp. 1893–925.
- Rewienski, M. & White, J. 2003, ‘A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices’, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, no. 2, pp. 155–70.
- Rewieński, M. & White, J. 2006, ‘Model order reduction for nonlinear dynamical systems based on trajectory piecewise-linear approximations’, *Linear Algebra and Its Applications*, vol. 415, no. 2–3, pp. 426–54.
- Rivlin, R.S. 1948, ‘Large Elastic Deformations of Isotropic Materials. I. Fundamental Concepts’, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*.
- Rivlin, R.S. & Saunders, D.W. 1951, ‘Large Elastic Deformations of Isotropic Materials. VII. Experiments on the Deformation Rubber’, *Philosophical Transactions of the Royal Society*

A: Mathematical, Physical and Engineering Sciences.

- Rixen, D.J. 2004, 'A dual Craig-Bampton method for dynamic substructuring', *Journal of Computational and Applied Mathematics*, vol. 168, no. 1–2, pp. 383–91.
- Roweis, S.T. & Saul, L.K. 2000, 'Nonlinear dimensionality reduction by locally linear embedding.', *Science*, vol. 290, no. 5500, pp. 2323–6.
- Rozza, G., Huynh, D.B.P. & Patera, A.T. 2008, 'Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics', *Archives of Computational Methods in Engineering*.
- Ruppert, D. 1987, 'Robust Statistics: The Approach Based on Influence Functions', *Technometrics*.
- Ryckelynck, D. 2005, 'A priori hyperreduction method: an adaptive approach', *Journal of Computational Physics*, vol. 202, no. 1, pp. 346–66.
- Ryckelynck, D., Benziane, D.M. & Paristech, M. 2010, 'Multi-level a priori hyper reduction of mechanical models involving internal variables', *Comput. Methods Appl. Mech. Eng.*, vol. 199, no. 17, pp. 1134–42.
- Ryckelynck, D., Chinesta, F., Cueto, E. & Ammar, A. 2006, 'On the a priori model reduction: Overview and recent developments', *Archives of Computational Methods in Engineering*, pp. 91–128.
- Ryckelynck, D., Lampoh, K. & Quilicy, S. 2016, 'Hyper-reduced predictions for lifetime assessment of elasto-plastic structures', *Meccanica*, vol. 51, no. 2, pp. 309–17.
- Ryckelynck, D. & Missoum Benziane, D. 2016, 'Hyper-reduction framework for model calibration in plasticity-induced fatigue', *Advanced Modeling and Simulation in Engineering Sciences*, vol. 3, no. 1.
- Sacks, M.S. 2000, 'Biaxial mechanical evaluation of planar biological materials', *Journal of Elasticity*.
- Sacks, M.S. & Sun, W. 2003, 'Multiaxial Mechanical Behavior of Biological Materials', *Annual Review of Biomedical Engineering*.
- Sargsyan, S., Brunton, S.L. & Kutz, J.N. 2015, 'Nonlinear model reduction for dynamical systems using sparse sensor locations from learned libraries', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 92, no. 3.
- Saul, L.K.L. & Roweis, S.S.T. 2003, 'Think globally, fit locally: unsupervised learning of low dimensional manifolds', *The Journal of Machine Learning Research*, vol. 4, no. 1999, pp. 119–55.

- Schilders, W.H., van der Vorst, H.A. & Rommes, J. 2008, *Model Order Reduction: Theory, Research Aspects and Applications*, Theory, Research Aspects and Applications, Springer, Berlin, Germany.
- Schmidt, M. & Lipson, H. 2009, 'Distilling free-form natural laws from experimental data', *Science*, vol. 324, no. 5923, pp. 81–5.
- Schmit, R.F. & Glauser, M.N. 2004, 'Improvements in Low Dimensional Tools for Flow-Structure Interaction Problems: Using Global POD', *42th Aerospace Science Meeting & Exhibit, January, 2004, Reno, Nevada, AIAA Paper 2004--889*.
- Schölkopf, B., Smola, A. & Müller, K.-R. 1998, 'Nonlinear Component Analysis as a Kernel Eigenvalue Problem', *Neural Computation*, vol. 10, no. 5, pp. 1299–319.
- Sen, C.K., Gordillo, G.M., Roy, S., Kirsner, R., Lambert, L., Hunt, T.K., Gottrup, F., Gurtner, G.C. & Longaker, M.T. 2009, 'Human skin wounds: A major and snowballing threat to public health and the economy', *Wound Repair and Regeneration*, vol. 17, no. 6, pp. 763–71.
- Shabana, A.A. 1991, 'Theory of Vibration, vol. 2', *I, Mechanical Engineering Series, Springer*.
- Shih, C.F. & Asaro, R.J. 1988, 'Elastic-Plastic Analysis of Cracks on Bimaterial Interfaces: Part I—Small Scale Yielding', *Journal of Applied Mechanics*, vol. 55, no. 2, p. 299.
- Simpson, T.W., Mauery, T.M., Korte, J. & Mistree, F. 2001, 'Kriging models for global approximation in simulation-based multidisciplinary design optimization', *AIAA Journal*.
- Sirovich, L. 1987, 'Turbulence and the dynamics of coherent structures I: coherent structures', *Quarterly of Applied Mathematics*, vol. 45, no. 3, pp. 561–71.
- Soldner, D., Brands, B., Zabihiyan, R., Steinmann, P. & Mergheim, J. 2017, 'A numerical study of different projection-based model reduction techniques applied to computational homogenisation', *Computational Mechanics*.
- Stekelenburg, A., Strijkers, G.J., Parusel, H., Bader, D.L., Nicolay, K. & Oomens, C.W. 2007, 'Role of ischemia and deformation in the onset of compression-induced deep tissue injury: MRI-based studies in a rat model', *Journal of Applied Physiology*, vol. 102, no. 5, pp. 2002–11.
- Stoecklein, D., Lore, K.G., Davies, M., Sarkar, S. & Ganapathysubramanian, B. 2017, 'Deep Learning for Flow Sculpting: Insights into Efficient Learning using Scientific Simulation Data', *Scientific Reports*.
- Strange, H. & Zwigelaar, R. 2010, 'A Generalised Solution to the Out-of-Sample Extension Problem in Manifold Learning', *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, pp. 471–6.

- Swischuk, R., Mainini, L., Peherstorfer, B. & Willcox, K. 2018, 'Projection-based model reduction: Formulations for physics-based machine learning', *Computers & Fluids*.
- Szalai, R. 2015, *Model reduction of non-densely defined piecewise-smooth systems in Banach spaces*.
- Taylor, J.A. & Glauser, M.N. 2004, 'Towards Practical Flow Sensing and Control via POD and LSE Based Low-Dimensional Tools', *Journal of Fluids Engineering*.
- Tenenbaum, J.B., de Silva, V. & Langford, J.C. 2000, 'A global geometric framework for nonlinear dimensionality reduction.', *Science (New York, N.Y.)*, vol. 290, no. 5500, pp. 2319–23.
- Tidball, J.G. 1984, 'Myotendinous junction: Morphological changes and mechanical failure associated with muscle cell atrophy', *Experimental and Molecular Pathology*.
- Tiso, P. & Rixen, D.J. 2013, 'Discrete empirical interpolation method for finite element structural dynamics', *Conference Proceedings of the Society for Experimental Mechanics Series*.
- Torre, F. Dela & Black, M.J. 2003, 'A Framework for Robust Subspace Learning', *International Journal of Computer Vision*, vol. 54, no. 1, pp. 117–42.
- Tracey, B.D., Duraisamy, K. & Alonso, J.J. 2015, 'A Machine Learning Strategy to Assist Turbulence Model Development', *53rd AIAA Aerospace Sciences Meeting*.
- Tröltzsch, F. & Volkwein, S. 2009, 'POD a-posteriori error estimates for linear-quadratic optimal control problems', *Computational Optimization and Applications*.
- Tropp, J.A. & Gilbert, A.C. 2007, 'Signal recovery from random measurements via orthogonal matching pursuit', *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–66.
- van Tuijl, R.A., Remmers, J.J.C. & Geers, M.G.D. 2018, 'Integration efficiency for model reduction in micro-mechanical analyses', *Computational Mechanics*, vol. 62, no. 2, pp. 151–69.
- Venter, G., Haftka, R.T. & Starnes, J.H.J. 1998, 'Construction of response surface approximations for design optimization', *AIAA Journal*.
- Vidal, R., Ma, Y. & Sastry, S.S. 2016, *Generalized Principal Component Analysis, Book*, Springer.
- Wang, C. & Bai, L. 2012, 'Sarcopenia in the elderly: Basic and clinical issues', *Geriatrics and Gerontology International*.
- Wang, K. & Sun, W.C. 2018, 'A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning', *Computer Methods in Applied Mechanics*

and Engineering.

- Wild, S.M., Regis, R.G. & Shoemaker, C.A. 2008, 'ORBIT: Optimization by Radial Basis Function Interpolation in Trust-Regions', *SIAM Journal on Scientific Computing*.
- Willcox, K. 2006, 'Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition', *Computers and Fluids*, vol. 35, no. 2, pp. 208–26.
- Williams, M.L. 1952, 'Stress singularities resulting from various boundary conditions in angular corners of plates in extension', *Journal of Applied Mechanics*, vol. 19, no. 4, pp. 526–8.
- Williams, M.L. 1957, 'On the Stress State at the Base of a Stationary Crack', *Journal of Applied Mechanics*, vol. 24, no. march, pp. 109–14.
- Williams, M.O., Kevrekidis, I.G. & Rowley, C.W. 2015, 'A Data-Driven Approximation of the Koopman Operator: Extending Dynamic Mode Decomposition', *Journal of Nonlinear Science*.
- Wright, J., Ma, Y., Mairal, J., Sapiro, G., Huang, T.S. & Yan, S. 2010, 'Sparse representation for computer vision and pattern recognition', *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1031–44.
- Wu, X., Kumar, V., Ross, Q.J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G.J., Ng, A., Liu, B., Yu, P.S., Zhou, Z.H., Steinbach, M., Hand, D.J. & Steinberg, D. 2008, 'Top 10 algorithms in data mining', *Knowledge and Information Systems*.
- Xiao, D., Yang, P., Fang, F., Xiang, J., Pain, C.C. & Navon, I.M. 2016, 'Non-intrusive reduced order modelling of fluid-structure interactions', *Computer Methods in Applied Mechanics and Engineering*.
- Xiaofei He, Deng Cai, Shuicheng Yan & Hong-Jiang Zhang 2005, 'Neighborhood preserving embedding', *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, p. 1208–1213 Vol. 2.
- Xiu, D. & Karniadakis, G.E. 2002, 'The Wiener--Askey Polynomial Chaos for Stochastic Differential Equations', *SIAM Journal on Scientific Computing*.
- Yaghoobi, M. & Davies, M.E. 2015, 'Fast non-negative orthogonal least squares', *2015 23rd European Signal Processing Conference, EUSIPCO 2015*, pp. 479–83.
- Yan, S., Xu, D., Zhang, B. & Zhang, H.-J. 2005, 'Graph embedding: a general framework for dimensionality reduction', *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, no. 1, pp. 830–7 vol. 2.
- Yan, S., Xu, D., Zhang, B., Zhang, H.J., Yang, Q. & Lin, S. 2007, 'Graph embedding and extensions: A general framework for dimensionality reduction', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 40–51.

- Yau, J.F., Wang, S.S. & Corten, H.T. 1980, 'A Mixed-Mode Crack Analysis of Isotropic Solids Using Conservation Laws of Elasticity', *Journal of Applied Mechanics*, vol. 47, no. 2, p. 335.
- Zhang, T., Tao, D., Li, X. & Yang, J. 2009, 'Patch alignment for dimensionality reduction', *IEEE Transactions on Knowledge and Data Engineering*.
- Zhang, Y. 2015, 'A Coupled Electro-Chemical-Mechanical Multi-Scale Computational Framework for Simulation of Skeletal Muscles', UC San Diego.
- Zhang, Z. & Zha, H. 2003, 'Local linear smoothing for nonlinear manifold learning', *CSE-03--003, Technical Report*.
- Zhang, Z. & Zha, H. 2004, 'Principal manifolds and nonlinear dimensionality reduction via tangent space alignment', *SIAM journal on scientific computing*, vol. 26, no. 1, pp. 313–38.
- Zhang, Z.J. & Duraisamy, K. 2015, 'Machine Learning Methods for Data-Driven Turbulence Modeling', *22nd AIAA Computational Fluid Dynamics Conference*.
- Zhou, J. & Fung, Y.C. 1997, 'The degree of nonlinearity and anisotropy of blood vessel elasticity', *Proceedings of the National Academy of Sciences*.