REDUCED ORDER MULTISCALE MODELING OF NONLINEAR PROCESSES
IN HETEROGENEOUS MATERIALS

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by
Satyaki Bhattacharjee

_________________________________________
Karel Matouš, Director

Graduate Program in Aerospace and Mechanical Engineering
Notre Dame, Indiana
July 2017
Abstract

by

Satyaki Bhattacharjee

Predicting effective material properties of nonlinear heterogeneous materials from the knowledge of its micro-structure through numerical modeling and computational homogenization (CH) has many applications in engineering design. Direct numerical modeling (DNM) using finite element method (FEM) is capable of predicting material behavior accurately. Unfortunately, DNM and/or CH are computationally expensive methods. To address the computational complexity issue, many researchers have focused on reduced order modeling. However, most of the available schemes are only suited for linear and moderately nonlinear behavior in 2D setting. Moreover, these techniques do not preserve local micro-scale fields in localization processes and cannot accommodate generic loading conditions.

Addressing these shortcomings, this doctoral work presents a robust reduced order modeling technique for heterogeneous materials with nonlinear hyper-elastic constitutive behavior in a finite strain setting. The model is not only capable of predicting homogenized (overall) material properties, but also recovers the micro-fields (i.e., the local deformation gradients) with acceptable accuracy. In the present work, a novel manifold based reduced order model has been developed for nonlinear hyperelastic materials utilizing advanced machine learning techniques. This data-driven model can perform well in comparison with traditional CH.
Proposed technique extracts the pattern in the solution manifold, which consists of an ordered set of micro-scale deformation fields. Each point on the manifold represents data obtained from a detailed parallel finite element simulation of a representative micro-structure. Since essential micro-fields are invariant of macro-rotations, the parameter space is created based on the macro-scale stretch tensor, which is parameterized by three principal stretches and three rotation parameters which represent the corresponding principal directions. This parameterization leads to a 6-dimensional loading space. Also, a novel pattern/physics based sampling strategy has been introduced to construct a representative solution manifold with a few number of simulations. This graph-based technique essentially explores the rotational (principal direction) sensitivity (in terms of an approximate diameter of the submanifold) of the principal stretch vector with a few number of FE solutions and guides to achieve a representative HD manifold by eliminating the redundant expensive simulations. In this work, a global dimension reduction technique, Isomap, is used to understand the underlying pattern of the submanifolds. Isomap returns a reduced low-dimensional Euclidean space which approximately unfolds the HD manifold by preserving the geodesic distances. Next, a map between the reduced space and the macroscopic loading conditions has been established using a Neural Network. Finally, the micro-scale deformation field is obtained through a nonparametric regression model by exploiting the concept of reproducing kernel Hilbert space (RKHS).

This novel reduced order model is able to predict the macro-scale as well as micro-scale deformation field for any unknown loading condition without any expensive simulation. Furthermore, this model potentially can accelerate the traditional CH by providing an initial solution vector.
FIGURES

2.1 Schematic of macroscale and microscale domains in theory of computational homogenization. ........................................... 13

3.1 Schematic of nonlinear processes and corresponding manifold representation. (a) Plane pendulum. (b) One-dimensional bar subjected to axial load. ................................................................. 19

3.2 Schematic of CH and MNROM: Dashed lines represent solution paths for snapshots, and solid lines show solution paths for reduced analysis of query points. .................................................. 21

3.3 Intrinsic curvature: the hemisphere (left) is a manifold with a constant curvature and half-cylinder (right) is a flat manifold. ................................................................. 23

3.4 Reconstruction and convergence of RKKS based regression model. (a) Reconstructed Swiss-roll. (b) Convergence study of RKKS based regression. ................................................................. 30

3.5 Reconstruction of coordinates for outside convex-hull points with RKKS based regression. (a) Reconstructed coordinates for 10 unit radius disk. (b) Divergence of RKKS model as it moves away from the boundary of the convex-hull. .................................................. 31

3.6 Schematic of the different components of neural network and the concept of $L_2$ regularization. ................................................................. 33

4.1 Concept of the physics based construction of the manifold. ............ 36

4.2 Concept of the rotational sensitivity analysis. ............................................ 38

5.1 Statistical analysis of the RUC. (a) Isotropic second-order particle-to-particle probability function. (b) Directional covariance of the second-order particle-to-particle probability. ................................................................. 43

5.2 RUC with side length of $l_{RUC} = 10R = 5 \mu m$. The cell contains 95 spherical particles with 1 $\mu m$ diameter. ................................................................. 43

5.3 Schematic of the rotation parameters. .......................................................... 45

5.4 Isomap dimensionality ................................................................. 46

5.5 Distribution of the volume averaged error (homogenization) given by Equation (5.3) for randomly selected 10% of points from database (one-out cases). (a) Particle. (b) Matrix. ................................................................. 49
5.6 Distribution of the local error (localization) given by Equation (5.2) for randomly selected 10% of points from mode 3 (one-out cases). Mean error of the deformation gradient field. Standard deviation of the deformation gradient field. 

5.7 Distribution of $\|F\|_F$ from FEM and MNROM. Over Matrix. Over Particle. 

5.8 Distribution of Neural Network output error compared to reduced space over entire dataset. 

5.9 Distribution of the volume averaged error (homogenization) given by Equation (5.3) of the 100 randomly selected query points in each mode. Particle. Matrix. 

5.10 Distribution of the local error (localization) given by Equation (5.2) for the 100 randomly selected query points from mode 3. Mean error and standard deviation of the finite element field for 100 query points. 

5.11 Macroscale strain energy density function, MNROM versus FEM. The deviatoric potential, $\omega_W(C)$. The volumetric potential, $\omega_W(J)$. 

5.12 Least-square fit of the macroscale strain energy density function, MNROM versus FEM. The deviatoric potential, $\omega_W(C)$. The volumetric potential, $\omega_W(J)$. 

5.13 Macroscale first Piola stress tensor, $\sigma_P$, MNROM versus FEM. Frobenius norm of $\sigma_P$ against Frobenius norm of $\sigma_F$. Distribution of the homogenization error given by Eq. (5.5). 

5.14 Joint probability distribution (computed in discrete element-wise sense) of the local fields (in matrix phase). $E_m$ and $\|F\|_F$. $E_e$ and $\|e\|_F$. 

5.15 Visualization of the effective Almansi strain ($\|\varepsilon\|_F$) (Note that particles are rigid-like and are therefore removed for visualization purpose.). FEM simulation. MNROM analysis. 

5.16 Isomap dimensionality with step 1 data. 


5.18 Distribution of $n_{yi}$ for all steps with all modes. The size of the markers indicates the enrichment step. Mode 1. Mode 2. Mode 3. Mode 4. 

5.20 Convergence of adaptive data enrichment scheme in terms of the microscopic deformation gradient field. (a) matrix-mean for modes 1, 2 and 4. (b) matrix-max for modes 1, 2 and 4. (c) matrix-mean for mode 3. (d) matrix-max for mode 3. ................................. 76

5.21 Distribution of the volume averaged error (homogenization) given by Eq. (5.3) for randomly selected 10% of points from database (one-out cases). (a) Particle. (b) Matrix. .................. 78

5.22 Distribution of $\|F\|_\infty$ from FEM and MNROM. (a) Over Matrix. (b) Over Particle. ................................. 78

5.23 Distribution of the neural network output error compared to reduced space over entire dataset for different enrichment steps. The size of the markers indicates the enrichment step. (a) Mode 1. (b) Mode 2. (c) Mode 3. (d) Mode 4. ........................................ 80

5.24 Distribution of the volume averaged error (homogenization) given by Eq. (5.2) of the 200 randomly selected query points in each mode. (a) Particle. (b) Matrix. .................. 82

5.25 Macroscale strain energy density function MNROM versus FEM. (a) The deviatoric potential, $^0W_C(\hat{C})$. (b) The volumetric potential, $^0W_J(J)$. 83

5.26 Least-square fit of the macroscale strain energy density function MNROM versus FEM (a) The deviatoric potential, $^0W_C(\hat{C})$. (b) The volumetric potential, $^0W_J(J)$. ................................. 84

5.27 Macroscale first Piola stress tensor, $^0P$, MNROM versus FEM. (a) Frobenius norm of $^0P$ against Frobenius norm of $^0F$. (b) Distribution of the homogenization error given by Eq. (5.5) ................. 84

5.28 Joint probability distribution (computed in discrete element-wise sense) of the local fields (in matrix phase). (a) $E_m$ and $\|F_{FEM}\|_\infty$. (b) $E_e$ and $\|e_{FEM}\|_\infty$. ................................. 86

5.29 Visualization of the microscopic effective Almansi strain ($\|e\|_\infty$) for $\eta^0 = {^0\lambda_1 = 0.9284, ^0\lambda_2 = 0.9058, ^0\lambda_3 = 1.0838, \varphi^1 = 2.4217, \varphi^2 = 1.2802, \varphi^3 = 0.3165}$ (Note, particles are rigid-like and hence are removed for the visualization purpose). (a) FEM simulation. (b) MNROM analysis. ................................. 87
TABLES

5.1 MOONEY-RIVLIN MATERIAL PROPERTIES .......................... 44
5.2 DESCRIPTION OF THE MACROSCALE LOADING MODES ... 47
5.3 QUALITY OF THE REGRESSION MODEL FOR CONSTRUCTION
   OF THE INVERSE MAP, $f^{-1}$, FOR ALL MODES COMPUTED
   FROM EQUATION (5.3) .................................................. 52
5.4 NEURAL NETWORK REGRESSION PERFORMANCE ............. 53
5.5 QUALITY OF THE MNROM FOR ALL MODES COMPUTED FROM
   EQUATION (5.3) FOR 100 QUERY POINTS IN EACH MODE ... 57
5.6 LOCALIZATION ERROR CHARACTERISTICS OF MNROM FOR
   $\text{MEAN}(\varepsilon_m)$ ERROR FROM EQUATION (5.2) (SEE ALSO FIGURE 5.10)
   OF THE DEFORMATION GRADIENT FIELD FOR POPULATION
   OF 100 RANDOM QUERY POINTS FROM MODE 3 ............... 59
5.7 CALIBRATED HOMOGENIZED MATERIAL PROPERTIES ....... 62
5.8 DESCRIPTION OF THE MACROSCALE LOADING MODES WITH
   DATA ENRICHMENT STEPS ........................................... 68
5.9 PERCENTAGE OF THE OVERLAP FOR EACH MODE AND IN
   EACH STEP ............................................................. 73
5.10 QUALITY OF THE MNROM FOR ALL MODES COMPUTED FROM
   Eq. (5.3) FOR 200 QUERY POINTS IN EACH MODE ............ 81
5.11 CALIBRATED HOMOGENIZED MATERIAL PROPERTIES ....... 85
ACKNOWLEDGMENTS

The entire Ph.D. endeavor was an eventful journey for me. This journey not only enriched me by academic means, also has made my life multifaceted. During this period, first time I could experience the challenges of staying in altogether a different environment. I am grateful to the entire Notre Dame community for being cooperative throughout, which made me feel at home. I would like to express my sincere gratitude to my advisor, Dr. Karel Matouš for his constant suggestion and support during the development of my Ph.D. work. I would also like to thank him for his guidance and motivation through all the years, which inspired me to put my sincere efforts. I would like to be appreciative to all my committee members for bringing different perspectives of my research work. Of course, I cannot forget the effort of other professors who prepared my background with their interactive classroom teaching. I would like to thank my parents and sister for their love and moral support, without them, none of these would have been worth. Finally, I would like to express my special thanks to my sponsors for their generous financial support.
1.1 Motivation

Over time, a large number of scientific and engineering fields rely on computer simulations to handle complex real-world phenomena. Such complex phenomena can appear in any engineering field like very-large scale integration (VLSI) design \[14\], in chemical engineering \[56\], modeling of turbulence in aerospace or mechanical engineering design \[108\], just to name a few. Among many of these applications, modeling of complex materials also falls into the similar category. Scientific community has already come up with many delicate solution strategies to take a grip on such materials. Especially, computational homogenization (CH) is a very popular and effective method when material complexity involves multiple length scales \[48, 37, 69, 72, 31, 65, 37\]. Moreover, complexity enhances when deformation reaches high strain regime. Although problem definitions differ in different fields, broadly, these fall into the category of nonlinear optimization problem. CH poses extra complexity as it involves multiple length scales along with the costly functional evaluation to solve the inherent optimization/variational problem \[37, 69, 72, 31, 65\]. Unfortunately, computational costs associated with solving these complex problems limit the use of these high-fidelity simulations. The expenses are many folds. The first problem is it takes long time to produce results. For this reason, the use of high-performance computing (HPC) is unavoidable. Even after taking help of HPC, these simulations take significantly long time to run. This work deals with the concerns related to the
computational cost reduction in the context of CH for the heterogeneous particulate complex materials.

The scope of complex heterogeneous materials in engineering application is ubiquitous. Some practical examples in this category are reinforced rubbers [75, 85], soft biological materials [70, 49], solid rocket propellants [83, 82], just to name a few. Predicting the mechanical response of such materials from the knowledge of their microstructure and behavior of individual constituents has been a long standing endeavor. To this regard, multiscale methods play an essential role [104, 82]. In particular, CH is a popular method for bridging scales [48, 37, 69, 112, 66, 29, 33, 72]. Reviews providing more details on CH methods have been published by Geers et al. [37], Nguyen et al. [74] and very recently Matouš et al. [65].

For nonlinear heterogeneous materials, CH requires solution of two nested nonlinear boundary value problems (BVP). In short, the macroscopic BVP provides boundary conditions for the microscale, and the microscopic BVP yields the average flux (i.e. the macroscopic stress tensor at a microscopic material point) and the macroscopic (homogenized) consistent tangent. This nested solution structure can also be characterized in terms of information exchange. In particular, CH is composed of two basic multiscale processes: i) Homogenization and ii) Localization of the material response. The homogenization process returns effective material behavior through volume averaging of the local microscale fluctuations over a representative unit cell (RUC). Therefore, it reduces the information from the microscale (RUC) to a point at the macroscale (many to one). On the other hand, the localization process yields local (i.e. microscale) fields over the RUC based on the macro-stimuli of a macroscopic point (i.e. macro-deformations). Therefore, it expands the information from the macroscale (a point) to the microscale RUC (one to many). For real engineering applications, CH handles a large amount of information/data (in localization). To make the computation feasible, as mentioned earlier, the high performance comput-
ing is often used in CH implementation [72]. Unfortunately, the CH still demands a substantial amount of simulation time depending on the complexity of the problem. Therefore, the scientific community has explored the viable options of reduced order modeling.

As alluded in the previous discussion, CH involves large amounts of scientific data. In this doctoral work, patterns are sought within the CH solutions resulting from a few HPC simulations and this leads to development of a data-driven reduced order model for the multiscale modeling of heterogeneous materials. Here, a manifold based learning technique is adopted to find the low-dimensional representation of the high-dimensional data. Instead of performing expensive CH, localization and homogenization can be performed using precomputed solutions in much shorter time for the entire spectrum of loading cases.

1.2 Literature Review

CH is a very efficient technique and provides detailed solutions of both macro- and micro-domains (i.e. both homogenization and localization). However, it is also computationally intensive. Typically, solved numerically by employing the finite element method at both scales (i.e. FE$^2$ implementation) [31, 72]. To lower this computational expense, the research community has focused on model reduction as it can significantly reduce time and data storage requirements [109, 47, 27, 77, 36]. One particularly popular reduction method is based on proper orthogonal decomposition (POD) [55, 109, 47]. For example, both homogenization and localization of hyper-elastic composites using POD have been performed in 2D by Yvonnet and He [109]. Although very popular, POD is better suited for linear or weakly nonlinear processes (ordinary or partial differential equations) as it relies on a linear combination of few basis vectors [13, 8]. Moreover, POD models drift from the FE$^2$ solution, since the method is directly coupled with the macroscale in terms of the reduced basis. Other
popular methods are based on digital database and/or discrete material maps to construct the overall homogenized macro-potentials or to directly homogenize the stress tensor \([110, 98, 97]\). In this category, Temizer and Zhodi \([97]\) and Temizer and Wriggers \([98]\) have introduced a concept of discrete material maps based on virtual material experiments. Here, a range of tri-axial stretches and rotations is applied to create a material homogenization map of the second Piola stress tensor. However, this method can only recover the homogenized response for any new material point that is not stored in the digital database. Also, Neural Network (NN) techniques have been employed to obtain approximate constitutive behavior, see Ghaboussi et al. \([38]\), Unger and Konke \([30]\) and Le et al. \([1]\). Unfortunately, using the NN in high dimensions is impractical \([91]\). Model calibration is another popular method. In work of Andrade and Tu \([25]\), some key material parameters of a phenomenological plasticity model at the macroscale are calibrated on the fly from microscale computations. Limited mostly to small strain analysis, the Transformation Field Analysis (TFA) and its nonuniform extension (NTFA) have been popular as well \([24, 58, 35]\). Other attempts include proper generalized decomposition \([27]\) and techniques based on the phase-space simplicial interpolation \([57]\). Recently, novel Gaussian process emulator of parameterized partial differential equations for single scale analysis has been developed by Xing et al. \([107]\). In this work, the Isomap and kernel Isomap are used to embed the vectorized spatial and temporal outputs in a low dimensional (feature) space. Unfortunately, most of the works described above are in 2D or limited 3D setting (i.e. small RUC), cannot perform localization of the multiscale response, and/or are limited to small strain analysis. Furthermore, ROM demands extra care since the complexity increases with the number of simulations. Therefore, the data sampling aspect is worth emphasizing to eliminate the redundant simulations.

Since the data collection is computationally expensive and the optimal data distribution is not known in advance, data sampling is a grand challenge to maximize
the information gain [15]. Although there are several statistical techniques available to address this issue [10 15 67 53], none of these is universal. All these methods are categorized as classical response surface model (RSM) and space-filling method [15]. Generally, minimization of bias is more important for ROM for better predictability [80 63]. So uniformity of the points in the design space is the main objective. Uniformity is achieved either by maximizing the minimum distances among input points or by minimizing correlation measures among the sample data [51]. Although both the criteria are important and intuitively linked, achieving together is impossible in reality [51]. Practical implementation of space filling techniques include Orthogonal array [79], Latin Hypercube design [67], Minmax/Maxmin design [53], entropy based design [20] among many others. These techniques also have considerably intuitive appeal. The orthogonal arrays produce uniform samples but can generate particular forms of point replications. On the other hand, Latin Hypercube does not produce replicate but can lack uniformity. However, above mentioned methodologies have no control over the output space. It is true that uniformity in input space induces uniformity in the output space as well, but that requires a large number of sample points. Requirement of this number increases exponentially with the dimension of the input space [63]. Unfortunately, generation of such rich dataset is practically impossible. For practical implementations, one has to rely on a very few simulation points than requirement, which eventually maximize the information and that leads to an enriched ROM. However, in this particular scenario when extremely high-dimension is involved, uniform input space does not guarantee a uniform output space because of curse of dimensionality. This problem becomes more severe as the process of generating data is highly nonlinear like CH. When the process of generating data is expensive as in case of CH, sequential sampling technique is more efficient design [15 62]. The adaptive sequential sampling is superior than the technique without adaptation [62].
1.3 Novelty of this Doctoral Thesis

This work overcomes the lacunae of the previous developments towards the reduced order models dealing with complex heterogeneous materials by means of a data-driven modeling approach. The present model has been developed exploiting the principles of machine-learning [99, 60, 6, 96] and utilizing digital databases, generated from parallel simulations using the PGFem3D solver [66, 72] for heterogeneous hyper-elastic materials. As the data comes from a highly nonlinear process, here data idealization is a high-dimensional Riemannian manifold and the model is referred as manifold-based nonlinear reduced order model (MNROM). MNROM can handle any continuous field related to the problem in hand. However, a wise choice of field is important so that the other fields can easily be derived from the chosen one. In the context of CH, deformation and deformation gradient, both can be a prudent choice of field. However, there is a trade-off between these two fields. This work is based on nonlinear manifold learning and computational geometry [99, 60, 6, 96], which can perform both homogenization and localization of the material response in 3D finite strains setting with a realistic RUC by seeking underlying pattern in data from CH analysis. Moreover, this model incorporates a complete loading envelope in terms of the principal stretches and the orthogonal principal directions.

The size of the RUC is established based on statistical analysis of the morphology [43, 100]. Since the hyper-elastic materials modeling is highly nonlinear, a nonlinear manifold learning technique should be used instead of traditional Principal component analysis (PCA) or Multidimensional Scaling (MDS) for dimension reduction. Recently, several nonlinear dimension reduction techniques have been proposed. Among those, local techniques (i.e. Locally linear embedding (LLE) [96], Laplacian Eigenmaps [6]) preserve only local geometry and can fail to follow the global pattern, especially when the manifold contains noisy data [22, 102]. Moreover, some of the global techniques (i.e. Kernel PCA [88], diffusion maps [17]) are sensitive to the selection
of kernels. On the other hand, global technique like Isomap approximately preserves relevant geometrical scales [22]. Thus to construct the reduced space, Isomap is used. Isomap is a global technique for nonlinear manifold reduction that unfolds the solution manifold to a reduced space. Therefore, neighboring data points on the high-dimensional manifold $\mathcal{M}$ (where proximity is described by a geodesic distance) are also neighboring data points in the reduced space $\mathcal{A}$ (where proximity is described by a Euclidean distance). Due to the highly nonlinear nature of the manifold and curse of dimensionality, the map between the space of macroscopic loading parameters and the reduced space is established with NN [5, 101, 50, 95, 11, 18]. This map essentially links the macroscopic loading condition with the isometric reduced space. For the reconstruction purpose, an appropriate inverse map has been constructed. Note, the output space of the inverse map is extremely high dimensional. Fortunately, because of Isomap, the reduced space and the HD manifold are isometric. In this work, this isometric property has been utilized to construct the inverse map. The entire framework of this reconstruction map exploits the concept of the reproducing kernel Hilbert space (RKHS). The advantage of this reduced order model is in providing acceptably accurate microscale solutions without HPC simulation for any new loading case absent in the original dataset. Since MNROM framework is data-driven, an efficient sampling strategy is a key component towards its performance. In summary, the main components of MNROM are nonlinear manifold learning, especially, Isomap, which unfolds (i.e., projects to an Euclidean space) the solution manifold, a kernel based inverse map and neural network. The efficacy of the Isomap depends on how the limited simulation points can represent the manifold [7]. Since the CH is highly nonlinear process, a uniform distribution of points in the input space may lead to many pockets of extremely sparse regions in the manifold. On the other hand, some regions become unnecessarily dense. This fails to capture the global information of the manifold. Consequently, Isomap fails to perform as its a global learning technique.
and produces enormous noise \cite{99}. All other maps associated with MNROM are noise sensitive as well and that leads to a poorly performing ROM. Here, a deterministic strategy to efficiently sample in the input space has been proposed such that the sparse regions of high dimensional output manifold are minimized with the specific number of simulations. This novel sampling method is completely pattern/physics driven. In this method, the input and output spaces are coupled. This is based on the exploration of the data distribution pattern on the manifold and finding the physical parameters which guides the pattern. The enrichment of the complete manifold is accomplished by enriching the submanifolds. Some submanifolds require small data points. On the other hand, some submanifolds require more points depending on the volume of the submanifold. From the learned pattern, requirement of the number of points can easily be determined. This method distributes the snapshots more evenly on the output manifold. The present deterministic sampling technique is sequential and adaptive in nature. In every step, pattern is learned to enrich the manifold. However, this sampling strategy does not suggest to create a uniform input space. Although this technique has been applied in the context of CH, the application domain can be extended to many other mechanics related problems, especially, problem pertaining to solid mechanics.

In this work, I also address the issue of reducing the expense of costly Isomap as the enrichment technique proposed here is sequential in nature. For every enrichment step a highly expensive algorithm (i.e., Isomap) requires to reduce dimension. To mitigate this particular problem, a greedy algorithm can be used instead of full-scale Isomap to project any new point into the reduced space. This greedy algorithm is extremely effective if the current dataset is enough dense. The present indigenous physics guided sampling strategy is supportive to add this aspect of model reduction as well.
1.4 Dissertation Overview

Including this Introduction (Chapter 1), this dissertation contains total 6 chapters. All the relevant theoretical aspects are discussed in details in Chapters 2-4. Chapter 2 starts with an overview of the CH and then provides the mathematical background for the multi-scale analysis of randomly configured particulate materials. After development of the theoretical parts, chapter contains a succinct discussion mentioning some issues related to the CH. Chapter 3 begins with a brief motivation for ROM. The concept of MNROM has been elucidated with two very simple and common mechanical systems. Then the idea has been extended for multi-scale material modeling (CH). Next, all the maps (i.e., Isomap, Neural Network, Kernel inverse map) pertaining to MNROM have been presented one by one. In Section 3.1 theoretical part for Isomap has been provided with issues. Unfortunately, Isomap is an expensive algorithm. To circumvent this challenge, a greedy algorithm related to Isomap has been discussed in Section 3.1.1. In Section 3.2 different models of inverse map are presented by linking their equivalence. However, the RKHS based inverse map has been explained with more emphasis. Finally, different aspects of neural network has been catered in Section 3.3. As mentioned, the data-driven model needs an efficient sampling technique, in Chapter 4 various available sampling techniques are discussed with their limitations. To overcome those shortcomings, a physics guided sampling technique has been developed in Section 4.1 for the performance enhancement of the present reduced order model, MNROM. To substantiate the different aspects of newly developed reduced order model, two numerical experiments have been carried out for the multiscale response of a RUC with 95 stiff inclusions in Chapter 5. In Section 5.1 the main framework of MNROM has been verified without implementing any sampling technique. Then in Section 5.2 MNROM has been implemented in conjunction with the novel physics guided sequentially adaptive sampling technique. At the end, in Chapter 6 the entire work has been summarized with possible issues and scopes.
for future direction.
CHAPTER 2

COMPUTATIONAL HOMOGENIZATION

There are abundant applications in engineering where heterogeneous materials composed of different hyper-elastic phases experience nonlinear deformations that lead to very large strains. Some practical examples in this category are reinforced rubbers \[75, 85\], soft biological materials \[70, 49\], and solid rocket propellants \[83, 82\], just to name a few. The constituents (phases) of these heterogeneous engineering materials are distinguishable at much smaller scales than the scale of interest. Moreover, this can happen in multiple length-scales. Predicting material properties at the scale of interest from the known material properties of the identifiable phases is a standing endeavor for the scientific community. Although this smaller scale can be even less than the atomistic length scale \(10^{-9}\) m, in continuum setting, this scale is typically microscale \(10^{-6}\) m or larger. When dealing with the nonlinear heterogeneous materials under continuum framework, the main objective is to bridge the field of interest across the different length-scales to predict the effective material properties at the scale of interest, which is the macroscale \[48, 37, 69, 72, 31, 112, 66, 33, 104, 32, 74\]. In this regard, the homogenization technique is a well known to serve the purpose \[3\]. In homogenization technique, the effective field at the macroscale is the field average over the microscale \[48, 37, 69, 72, 66, 33, 65\]. Although the homogenization technique can handle multiple length scales, in this work, only three length-scales are considered: microscale, mesoscale and macroscale. The homogenization hinges on the concept of the Representative Volume Element (RVE) \[37, 65, 45, 78, 29\]. For the particulate composite, the same is termed as Representative Unit Cell (RUC)
Only two types of RVE/RUC are possible: a) periodic microstructure, which is extremely patterned and easy to handle, b) very large number of microscale phases are randomly scattered within the representative volume, retaining the statistical homogeneity and ergodicity \([65, 45, 78]\). Here, statistical homogeneity refers to the fact that the correlation functions remain same for any RVE/RUC at any point in the macroscale \([65, 45, 78, 43, 100, 29]\). The concept of the separation of length-scale plays a pivotal role in homogenization technique. It says, the characteristic length of the microscale \(l_{\text{micro}}\) is much smaller than the mesoscale \(l_{\text{meso}}\), which is again much much smaller than the macroscale \(l_{\text{macro}}\)(i.e., \(l_{\text{micro}} << l_{\text{meso}} << l_{\text{macro}}\)) \([65, 45, 78]\) (see Figure 2.1). This is known as the MMM principle in the literature \([45]\). Unless periodic microstructure, selection of RVE/RUC is a challenging part in homogenization. When the morphology of the microstructure is nearly periodic, the weak sense of MMM principle still holds. Otherwise, when microstructure is extremely random, stronger inequality of MMM principle needs to be applied. Homogenization methods seek the global governing behavior from the local boundary value problem (BVP) which is set up over the RVE/RUC. Other than a periodic unit cell and small strain problem, analytical solution is impossible \([32]\). Commonly, the BVPs are solved numerically, and often, finite element method is used. When nested two scale problems are numerically solved, the method is commonly termed as the computational homogenization (CH) \([65, 37]\). In this chapter, the theory of first-order computational homogenization with two-scales for quasi-static loading condition is briefly summarized. More details on this theory can be found in \([48, 37, 69, 72, 31]\).

2.1 Macroscale Problem

Consider a body \(\Omega_0 \subset \mathbb{R}^3\) consisting of material points \(X \in \Omega_0\) as shown in Figure 2.1. Let the boundaries \(\partial \Omega_0^u\) and \(\partial \Omega_0^t\) represent the boundary of applied displacement \(u\) and tractions \(t\), respectively. The boundary \(\partial \Omega_0\) is decomposed such
that \( \partial \Omega_0 = \partial \Omega^u_0 \cup \partial \Omega^t_0 \) with \( \partial \Omega^u_0 \cap \partial \Omega^t_0 = \emptyset \). Next, the deformation \( \chi^0 \) and the deformation gradient \( F^0 \) in the macroscale are defined as

\[
\begin{align*}
\chi^0(X) &= X + u^0(X) \quad \forall X \in \Omega_0, \\
F^0(X) &= \nabla X \chi^0 = 1 + \nabla X u^0 \quad \forall X \in \Omega_0,
\end{align*}
\]

where \( I \) is the second-order identity tensor, and \( u^0(X) \) is the macroscale displacement. Neglecting inertial forces, the macroscale boundary value problem is given by

\[
\begin{align*}
\nabla X \cdot P^0 + b^0 &= 0 \quad \text{in } \Omega_0, \\
P^0 \cdot N^0 &= t \quad \text{on } \partial \Omega^t_0, \\
u^0 &= u \quad \text{on } \partial \Omega^u_0,
\end{align*}
\]

Figure 2.1. Schematic of macroscale and microscale domains in theory of computational homogenization.

13
where $^0P = \partial^0W/\partial^0F$ is the macroscopic first Piola-Kirchhoff stress tensor, $^0W$ is the hyper-elastic macroscopic strain energy density function, $^0b$ is the macroscale body force, and $^0N$ is the unit normal to the surface $\partial\Omega^t_0$. Application of the standard variational procedures leads to the weak form: Find $^0u \in \mathcal{V}_{\Omega_0} \{^0u : \Omega_0 \to \mathbb{R}^3 | \det(^0F) > 0$ \}$ such that

$$
\mathcal{R}_u := \int_{\Omega_0} ^0P : (\nabla_X \delta^0u) \, dV - \int_{\partial\Omega_0} ^t \mathbf{t} \cdot \delta^0u \, dA - \int_{\Omega_0} ^0b \cdot \delta^0u \, dV = 0,
$$

(2.3)

holds for variations $\delta^0u(X) \in \overline{\mathcal{V}}_{\Omega_0} \{\delta^0u : \Omega_0 \to \mathbb{R}^3, \delta^0u|_{\partial\Omega_0^t} = 0\}$. As evident from Equations (2.2) and (2.3), the macroscale strain energy potential $^0W$ is required, which is determined by solving the microscale problem, in order to solve the macroscale problem.

2.2 Microscale Problem

A microstructure $\Theta_0 \subset \mathbb{R}^3$ consisting of microscale points $Y \in \Theta_0$ as shown in Figure [2.1] is locally attached to each macroscale point $X \in \Omega_0$. The microscale deformation $\chi$ and the deformation gradient $F$ are assumed to be functions of both macro and micro variables as

$$
\chi(X, Y) = ^0F(X)Y + ^1u(Y) \quad \forall Y \in \Theta_0,
$$

(2.4)

and

$$
F(X, Y) = ^0F(X) + \nabla_Y ^1u(Y) \quad \forall Y \in \Theta_0,
$$

where $^1u(Y)$ is the microscale displacement fluctuation. The microscale equilibrium boundary value problem, neglecting body forces and without prescribed tractions, is given by
\[
\n\nabla_Y \cdot (^{1}P) = 0 \quad \text{in } \Theta_0, \\
\chi = 0 \chi \quad \text{on } \partial \Theta_0.
\]

where \(^1P = \partial^1W/\partial F\) is the microscopic first Piola-Kirchhoff stress tensor. The microscale free energy density function, \(^1W\), describes material behavior of interest and is given. The weak form of Equation (2.5) is derived from the micro-to-macro transition with a particular set of physically applicable and mathematically admissible boundary conditions.

2.3 Micro-to-Macro Transition

In computational homogenization, the behavior of a material point at the macroscale is linked to the microscale through the Hill-Mandel stationarity condition \([48]\):

\[
\inf_{\delta u} \int_{\Theta_0} \left[ ^1W(0F + \nabla_Y \delta u) \right] d\Theta = 0.
\]

Expanding the variation of Equation (2.6) with respect to \(^1u\), provides the weak form of Equation (2.5). Find \(^1u \in V_{\Theta_0} \{ ^1u : \Theta_0 \to \mathbb{R}^3 \mid \det (F) > 0 \ \forall Y \text{ in } \Theta_0 \text{ and } ^1u|_{\partial \Theta_0} \text{ is periodic} \} \) such that

\[
\mathcal{R}_u := \int_{\Theta_0} ^1P : (\nabla_Y \delta \delta ^1u) d\Theta = 0,
\]

is stationary with variation \(\delta ^1u(Y) \in V_{\Theta_0} \{ \delta ^1u : \Theta_0 \to \mathbb{R}^3, \ \delta ^1u|_{\partial \Theta_0} \text{ is periodic} \} \).

In the computational homogenization, the periodic boundary conditions for the cell fluctuations, \(^1u\) is typical (see Geers et al., \([37]\)).
2.4 Discussion

The CH solves two weak forms in a nested manner (Typically, by FEM): the macroscale problem given in Equation (2.3) determines the deformation gradient, which is used as an input for microscale problem given in Equation (2.7), which in return gives back the constitutive relation for the macroscale problem. Issues related to practical implementation of CH can be found in [65, 37, 72]. The results produced by CH very closely depend on the selection of RVE/RUC [37, 65, 45, 78, 29]. Although the periodic boundary condition is a natural choice, CH results vary with the different homogeneous boundary conditions [73, 78]. The incompressible materials are common in engineering applications. However, the theory provided here does not encompass the incompressibility criteria. In that particular situation, the main framework provided here remains the same. Furthermore, bifurcation is a common phenomenon in non-linear continuum mechanics and appears even in seemingly simple problems involving fairly standard constitutive models [68, 64]. Capturing limit and bifurcation points requires a special numerical treatment [106]. When working with random materials with many inclusions, the microstructure introduces geometrical (morphological) imperfections that guide the solution path. Moreover, in displacement driven CH, the macro-deformation is uniquely prescribed. Generally for CH, high performance computing is necessary, which is expensive [72, 65]. Moreover, Computational cost increases with the complexity of the problem.
CHAPTER 3

MANIFOLD BASED REDUCED ORDER MODEL

Solutions of many parameter driven nonlinear continuum processes reside on an extremely high-dimensional (HD) manifold. However, in several cases, parameter space is a low dimensional space. Although the ambient space of those manifolds is very high, these manifolds inherit the topological dimension of the parameter space. This is justifiable by Whitney embedding theorem [41]. Moreover, Sard’s theorem suggests that the possibility of intersection reduces when a low dimensional manifold is embedded in a very high-dimensional Euclidean space [41]. To exemplify this notion, a plane pendulum, see Figure 3.1a, although moves in a 2D plane, motion can easily be described by a single parameter (angle/geodesic distance). Similarly, when an one dimensional bar is solved by finite element method (FEM) under an axial load, the dimension of the solution manifold is equal to its degrees of freedom (see Figure 3.1b). However, the solution manifold evolves only by a single parameter, that is the axial load. This entails the intrinsic dimension of the solution manifold is only one. When solving the model (typically a differential equation) to capture certain physics, although the solution manifold is intrinsically low dimensional, need to solve all the degrees of freedom. This leads to an extremely expensive process. Furthermore, it aggravates with the complexity of the problem. To ease this computational burden, the concept of reduced order model (ROM) has gained attention [55, 109, 47]. There are several techniques available in the literature to address this problem. Here, a novel data driven manifold based reduced order model has been introduced, where a low-dimensional structure is extracted from the precomputed high-dimensional so-
olutions and numerically constructing the maps to predict any other solution which is not in the dataset. To elucidate this idea, the bar problem is solved by FEM for finite sequence \((N)\) of axial force \((F)\). This \(N\) number of displacement solutions can be represented as a HD manifold (dimension depends on the finite element discretization). Next, this HD manifold is approximately unfolded to its intrinsic dimensional Euclidean space by preserving some invariant properties (i.e., geodesic). For the bar problem, low-dimensional space is a line (intrinsic dimension is one). Now, constructing a map between the axial load and the reduced space is much easier. Moreover, the reduced Euclidean space of the HD manifold preserves certain geometric characteristics of the nonlinear manifold. If this map is a bijection, an inverse map can be numerically established. For any other loading condition, the solution can be predicted without solving for expensive full-scale models. This concept can be extended for the highly complex problems as well by addressing the complexities involved in the process. In this work, the similar idea has been implemented for the complex materials in the context of CH.

In Chapter 2, the relevant theory for two-scale CH process for nonlinear heterogeneous hyper-elastic materials has been established by providing the governing equations. In a nutshell, for nonlinear heterogeneous materials, CH requires solution of two nested nonlinear boundary value problems (BVP). The macroscopic BVP provides boundary conditions for the microscale (localization), and the microscopic BVP yields the average flux (homogenization, i.e. the macroscopic stress tensor at a macroscale material point) and the macroscopic (homogenized) consistent tangent. In this chapter, the macro and microscale coupling is presented in a novel way from nonlinear manifold point of view. Here, a nonlinear manifold is constructed out of the microscale simulations which are also referred to as snapshots. This novel modeling concept described as a manifold-based nonlinear reduced order model (MNROM), illustrated in Figure 3.2. The MNROM is constructed from a digital database of \(N\) de-
Figure 3.1. Schematic of nonlinear processes and corresponding manifold representation. (a) Plane pendulum. (b) One-dimensional bar subjected to axial load.
tailed microscale simulations (denoted as $\mathcal{Z}$) and the set of their associated macroscale loading parameters (denoted as $\Xi$). An entry in the database is identified by the pair $(\eta^i, \xi^i)$ for $i = 1, \cdots, N$, where $\eta^i$ denotes the $i^{th}$ macroscale loading case and $\xi^i$ denotes data from the $i^{th}$ microscale simulation. When creating the digital database, any relevant continuum field can be stored, but deformation or deformation gradient is more fundamental in a sense that any other fields can directly be determined from these two fields. This suggests that the choice of the field for the manifold construction is also a key issue. In this work, model has been implemented and verified for both the deformation and deformation gradient. There is a natural trade-off between these two fields. If deformation is used, ROM complexity is less as the dimension of the HD manifold is lower than by constructing the manifold by deformation gradient. On the other hand, to compute any other field of interest (i.e., stress, energy), first, deformation gradient needs to be computed.

When considering hyper-elastic response, and if the choice of field is the deformation gradient, the database stores the nine components of $\mathbf{F}$ (see Equation (2.4)) for $N_e$ finite elements of the microscale discretization (discretized by the FEM, for example). Thus, each entry $(\eta^i, \xi^i)$ in the database contains $D = 9 \times N_e$ numbers, and the total size of the digital database is $N \times D$. Using the digital database, a high-dimensional manifold $\mathcal{M} \subset \mathbb{R}^D$ (gray region in Figure 3.2) is built containing the set $\mathcal{Z} = \{\xi^1, \cdots, \xi^N\}$ associated with its corresponding set of macroscale loading conditions $\Xi = \{\eta^1, \cdots, \eta^N\} \subset \mathbb{R}^{\tilde{d}}$, where $\tilde{d}$ is the number of the physical macroscale loading/controlling parameters. The microscale simulation points, $\xi^i$, are given by the ordered set of the deformation gradients in each element, i.e. $\xi^i = \{\mathbf{F}^1, \cdots, \mathbf{F}^{N_e}\} \in \mathcal{M}$. On the other hand, if the choice of field is the deformation, a nonlinear manifold is constructed using the microscale displacement field, as mentioned before, this substantially reduces the dimensionality of the problem (i.e., $\mathbf{F}$ is a second order tensor $(3 \times 3)$ and $\mathbf{u}$ is a vector $(3 \times 1)$). By vector-
izing the snapshots, again, a high-dimensional manifold \( \mathcal{M} \subset \mathbb{R}^D \) (gray region in Fig. 3.2) can be constructed, containing the set \( \mathfrak{z} = \{ \xi_1, \cdots, \xi_N \} \). The point on the manifold is given by the ordered set of the microscale displacements at each node, \( \xi = \{ \chi_1, \cdots, \chi_{N_n} \} \in \mathcal{M} \). This set governs the dimension of the embedding/ambient space of the manifold \( D = 3 \times N_n \). Here, \( N_n \) is the number of nodes in FE discretization. Again, the manifold points are associated with its corresponding macroscale loading conditions, \( \Xi = \{ \eta_1, \cdots, \eta_N \} \subset \mathbb{R}^\tilde{d} \), where \( \tilde{d} \) is the number of the physical macroscale loading/controlling parameters. For nearly incompressible materials the representation of the manifold will remain similar. Generally, a mixed finite element method is used to handle the nearly incompressible material response \[93, 64\]. In that situation, pressure/volume variables have to be included in \( \xi \).

Figure 3.2. Schematic of CH and MNROM: Dashed lines represent solution paths for snapshots, and solid lines show solution paths for reduced analysis of query points.
Although mathematically there is a map $F : \mathcal{T} \mapsto \mathcal{M}$ (the elliptic Eq. (2.5)) that defines the manifold $\mathcal{M}$ parametrically in terms of the input space $\mathcal{T}$, it is not a trivial task to construct this map analytically or even by deep learning [90], especially when the map is extremely nonlinear and involving very high dimension. For the well-posed problems [84], if the parameter space $\mathcal{T}$ is compact, the manifold $\mathcal{M}$ is also compact and inherits the topological dimension of the parameter space. As discussed before, after constructing the high-dimensional manifold, a low-dimensional Euclidean space $\mathcal{A} \subset \mathbb{R}^d$ (orange region in Figure [3.2]) is constructed. This space of lower dimension contains the set of points, $\mathcal{A} = \{\zeta_1, \cdots, \zeta_N\}$, generated from the limited set of simulations $\mathcal{Z}$, which are associated with the essential macroscale influence parameters ($\eta^j \in \mathcal{T}$). The map $F$ is established as a composition of two maps $f^{-1}$ and $\mathcal{P}$ ($F = f^{-1} \circ \mathcal{P}$). Thus, the MNROM consists of three individual components: i) dimension reduction (map $f$), ii) linkage of parameter space to the reduced space (map $\mathcal{P}$), and iii) reconstruction of the microscale solution (map $f^{-1}$) [9]. For dimension reduction, Isomap [99] is used as the data idealization is a nonlinear manifold. This idealization is consistent since the displacement/ deformation gradient field is highly nonlinear over the macroscopic loading space. Isomap globally unfolds the manifold and returns an Euclidean space, $\mathcal{A} = \{\zeta_1, \cdots, \zeta_N\} \subset \mathbb{R}^d$, with dimensionality $d$. This is the intrinsic dimension of the manifold ($d$). Isomap uses graph techniques and preserves the geodesics, thereby preserving the global structure of $\mathcal{M}$ in $\mathcal{A}$. Thus, near and far points on $\mathcal{M}$ (by geodesic distance) are also near and far on $\mathcal{A}$ (by Euclidean distance). The relationship between the loading space $\mathcal{T}$ and the reduced space, $\mathcal{A}$ is constructed by the neural network (NN), which is referred here as $\mathcal{P}$. To complete the correspondence between the manifold and the reduced space an appropriate inverse map, $f^{-1}$ is required, which is established by nonparametric regression. After establishment of all the spaces and maps, both homogenization and localization of the microscale response for new macroscale loading conditions,
denoted as a macroscale query point $\eta^q \not\in \Sigma$, can be performed. Moreover, this is accomplished without any additional large-scale HPC simulations. In what follows, a brief mathematical theories and principles are provided here to construct a reliable MNROM.

### 3.1 Isomap and Kernel Isomap

As mentioned, the map $f$ has been established using Isomap, which assumes that $M = f^{-1}(A)$ is globally isometric to $A$ \[99\]. Mathematically, this is only possible for flat manifolds. Which means, the intrinsic curvature of the manifold is zero. In other words, the Jacobian at every point is up to the rotation. Example of a flat manifold is a half-cylinder (see Fig. 3.3) and can easily be isometrically unfolded to a 2D Euclidean space. However the hemisphere in Fig. 3.3 is not a flat manifold. Consequently, cannot be isometrically projected to its intrinsic dimensional Euclidean space. Most of the manifolds do not satisfy this criteria. However, flat approximation of the manifold is possible \[23, 86\]. Especially, when the manifold is represented by few points, Riemannian structure is already lost. For the manifolds with intrinsic curvature, Isomap can produce globally optimal low-dimensional Euclidean representation \[99\].

![Image of Isomap and Kernel Isomap](image)

**Figure 3.3.** Intrinsic curvature: the hemisphere (left) is a manifold with a constant curvature and half-cylinder (right) is a flat manifold.
In the context of CH, in the limit, $N \to \infty$, the data-representation is a smooth manifold $\mathcal{M}$, without holes, as the parametric space $\mathcal{X}$ is without singularity. Since the map $f$ is globally isometric and $\mathcal{M}$ is compact, and without holes, the Euclidean space $\mathcal{A}$ is convex. For the smooth manifold, $\mathcal{M}$, the geodesic distance, $D_M(\xi^i, \xi^j)$, between all pair of points $\xi^i$ and $\xi^j$ is defined as follows:\[59\]:

$$D_M(\xi^i, \xi^j) = \inf_{\gamma} \{ \text{length}(\gamma) : \gamma : [0, 1] \mapsto \mathcal{M} \text{ and } \gamma(0) = \xi^i, \gamma(1) = \xi^j \}. \quad (3.1)$$

The Isomap exploits the concept of the distance kernel trick with the geodesic kernel $\mathcal{K}(\xi^i, \xi^j) = -D_M^2(\xi^i, \xi^j)$ that is conditionally positive definite (CPD) and returns the parameter space $\mathcal{A}$ [87]. The manifold is represented by the graph, $G$, constructed from the points, $\xi^i$, on the manifold $\mathcal{M}$ using the $k$-rule, where $k$ nearest neighbors are connected through the graph edges. Isomap involves the computation of an affinity matrix, $D_{ij} = D_G^2(\xi^i, \xi^j)$, from the data set $\mathcal{Z}$. Then, the matrix $\mathcal{D}$ is transformed into a positive definite (PD) double-centered matrix, $\tilde{\mathcal{D}}$, as follows:

$$\tilde{\mathcal{D}} = -\frac{1}{2} H \mathcal{D} H, \quad (3.2)$$

where $H = I - \frac{1}{N} \mathbf{1} \mathbf{1}^T$ is a centering matrix, $I$ is the $N \times N$ identity matrix, and $\mathbf{1}$ is a $N$-vector with all elements equal 1. Note, $\tilde{\mathcal{D}}$ is a PD Gram matrix [87].

To construct the parameter space, $\mathcal{A}$, Isomap utilizes the Mercer theorem [71]: Let $\mathcal{M} \subset \mathbb{R}^D$ be closed with a strictly positive Borel measure on $\mathcal{M}$, and $\mathcal{K}$ be a continuous function on $\mathcal{M} \times \mathcal{M}$ which is PD and square integrable. Then,

$$\mathcal{K}(\xi^i, \xi^j) = \sum_{k=1}^{\infty} \sigma_k \mathbf{v}(\xi^i) \mathbf{v}(\xi^j). \quad (3.3)$$
Here, the series converges absolutely for each pair \((\xi^i, \xi^j) \in \mathcal{M} \times \mathcal{M}\) and uniformly on each compact subset of \(\mathcal{M}\). Mercer’s theorem still holds pointwise if \(\mathcal{M}\) is a finite set. Exploiting the Eq. (3.3), the embedding is constructed by

\[ \zeta^i_k = \sqrt{\sigma_k} v^k_i, \]  

(3.4)

where \(\sigma_k\) is the \(k^{th}\) eigenvalue and \(v^k\) is the corresponding eigenvector of \(\tilde{D}\). Here, \(\zeta^i_k\) is the \(k^{th}\) component of \(\zeta^i\) out of \(d\) components, and \(v^k_i\) is the \(i^{th}\) component of \(v^k\) out of \(N\) components. Truncation in the decaying spectrum is done by applying the Beardwood-Halton-Hammersley (BHH) theorem [19]. Once \(d\) is computed, embedding can be constructed by taking first \(d\) eigenvalues and eigenvectors. It can be shown that convergence takes place asymptotically for an isometric map.

One of the major components of Isomap is computation of approximate geodesic distances in terms of all-pair shortest paths. Here, Dijkstra algorithm [92] is used to compute the shortest path. The approximate geodesic distances computed over the graph \(G\) obey the following theorem: Let \(\mathcal{M}\) be a compact manifold, embedded in \(\mathbb{R}^D\), isometrically equivalent to a convex domain \(\mathcal{A} \subset \mathbb{R}^d\) and let some parameters \(\omega_1, \omega_2\) and \(\mu\) be given in \((0, 1)\). Neglecting the boundary effects, it can be asserted that with the probability at least \(1 - \mu\) the following inequalities

\[ (1 - \omega_1)D_M(\xi^i, \xi^j) \leq D_G(\xi^i, \xi^j) \leq (1 + \omega_2)D_M(\xi^i, \xi^j), \]  

(3.5)

hold on \(\mathcal{M}\) for all \(\xi^i, \xi^j\). Here, \(D_G(\xi^i, \xi^j)\) is the approximate geodesic computed from the graph, \(G\). Equation (3.5) guarantees the asymptotic convergence of the computed approximate geodesics in limiting conditions \((N \rightarrow \infty)\). Finally, the steps associated with Isomap is provided below.

The inequality (3.5) indicates the possibility of noise associated with the com-
Algorithm 1 Isomap

1: Construct neighborhood graph, $G$: Compute $L_2$ distance of all points $\xi_i$. Apply $k$-rule to select the $k$-neighbors of $\xi_i$. Every $\xi_i$ will have $k$-edges with Euclidean distance as the edge weight.

2: Compute all-pair shortest paths in $G$: Use Floyd/Dijkstra algorithm \cite{92, 34} to compute all-pair shortest paths and store the squared value of each shortest path in $D$ matrix. Compute double centered matrix $\tilde{D}$ from $D$ given by Equation (3.2).

3: Determine the intrinsic dimension: Use BHH theorem and follow Equations (3.3)-(3.4) to determine the intrinsic dimension of the manifold, $d$.

4: Construct the reduced-dimensional space: Apply classical MDS technique on $\tilde{D}$. Compute $d$ largest eigenvalues, $\sigma_k$, and corresponding eigenvectors, $v^k$ of $\tilde{D}$. The reduced dimensional space is obtained as $\zeta^k_i = \sqrt{\sigma_k} v^k_i$.

Computed geodesic distances. This implies, the Gram matrix, $\tilde{D}$, may not be PD and violates the fundamentals of the kernel operations \cite{16}. Although Isomap replaces the negative eigenvalue with zero, which is mathematically sound, this can underestimate the dimension. To circumvent this situation kernel version has been proposed. Kernel version employs the analytical framework of the additive constant problem and transforms the kernel into a PD one \cite{12}. Although topological instability can be an issue for Isomap \cite{4}, it can be circumvented by carefully constructing a finite number of disjoint submanifolds and analyzing those separately \cite{9}.

3.1.1 Greedy Algorithm

The Isomap is an expensive algorithm, since it involves computation of all shortest paths (approximate geodesics) in conjunction with the eigenvalue analysis. However, once the manifold reaches certain density, further enrichment of the manifold needs to be very quick. This enrichment should allow adding new snapshots to the manifold without recomputing all existing geodesics. Therefore, to avoid the computational cost, a greedy algorithm can be used for any new snapshot, $\xi^{\text{new}}$, added to the manifold by projecting it to the current reduced space (in terms of the existing eigenvalues and eigenvectors) as follows \cite{7}:
\[
\zeta_k(\xi^{new}) = \frac{1}{2\sqrt{\sigma_k}} \sum v_k^i (\mathbb{E}[\tilde{D}^2(\xi^{new}, \xi^i)] - \tilde{D}^2(\xi^{new}, \xi^i)).
\] (3.6)

Here the geodesic distances for newly added points, \(\xi^{new}\) need to be computed and \(\mathbb{E}\) is the averaging operator. Note, the value of the geodesic distances can change drastically for manifold that is not well populated. This implies that the accuracy of this greedy technique depends on the current density of the manifold data.

### 3.2 Reconstruction of Microscale Solution

The map \(f\) is globally isometric [99], which leads to a bijection since the geodesic distance on \(\mathcal{M}\) is equal to the Euclidean distance on \(\mathcal{A}\) and distance between two distinct points cannot be zero. Also entire framework assumes that the embedding, \(f^{-1}\), is a diffeomorphism. Although the isometric map can be constructed explicitly [60], in this work, the map \(f\) is without an explicit form and does not have an explicit inverse. Moreover, the computed Euclidean distance does not exactly correspond to the geodesic distance over manifold \(\mathcal{M}\) [9]. In order to predict the micro-structural response from the reduced dimensional parameters, the inverse map has been constructed numerically through regression [9, 36, 105, 42]. For the completeness of the presentation, here different inverse map models are provided with relevant issues.

Consider a simulation point \(\xi^i \in \mathcal{M}\) and its corresponding MNROM point \(\zeta^i \in \mathcal{A}\), where \(\zeta^i = f(\xi^i)\). In regression setting, it is assumed that the discrete map \(f_N\) converges to the limiting map, i.e. \(f_N(\xi^i) = f(\xi^i)\) for \(i = 1, \ldots, N\). Thus, the approximate inverse, \(f_N^\dagger\), is constructed from a finite set of points such that \(\forall \zeta \in f(\mathcal{M})\) the inverse map \(f_N^\dagger\) converges to the exact inverse \(f^{-1}\) at the limit \((\lim_{N \to \infty} f_N^\dagger(\zeta) = f^{-1}(\zeta))\). The approximate inverse map for a point of interest \(\zeta^q\) (query point) is denoted as \(\xi^q = f_N^\dagger(\zeta^q)\) and is constructed in the sense of non-
parametric regression \[42\].

One of the basic approaches for creating this reconstruction map over the manifold \(\mathcal{M}\) is to use the conditional probability density \(P^i(\xi^i|\zeta^q)\):

\[
\xi^q = \sum_{i=1}^{N} \xi^i P^i(\xi^i|\zeta^q).
\]  

(3.7)

Many models emanate from Equation (3.7) since it provides generic description \[42\].

One simple approach of finding inverse is by finding the set of \(k\)-nearest neighbors (points) \(\mathcal{S}^q = \{\zeta^q,1,\ldots,\zeta^q,k\}\) of \(\zeta^q\) and the corresponding set of simulation points \(\mathcal{S}^i = f^{-1}_N(\mathcal{S}^q) = \{\xi^q,1,\ldots,\xi^q,k\}\) on the manifold \(\mathcal{M}\). Using the fact that \(\mathcal{A}\) is isometrically embedded in \(\mathcal{M}\), the point \(\xi^q \in \mathcal{M}\) is obtained by (see reference \[36\])

\[
\xi^q = \frac{\sum_{i=1}^{k} \xi^{q,i}(1/l^{q,i})}{\sum_{i=1}^{k}(1/l^{q,i})},
\]  

(3.8)

where \(l^{q,i} = \|\zeta^q - \zeta^{q,i}\|_{L^2}\). Note that Equation (3.8) is equivalent to Equation (3.7) with \(P^i(\xi^i|\zeta^q) = 1/k^{q,i}/\sum_{i=1}^{k}(1/l^{q,i})\) and sum is truncated within \(k\)-nearest points. This method is commonly known as local linear interpolation. Another simple, but often used interpolation method, is nearest neighbor averaging. For this method, \(P^i(\xi^i|\zeta^q) = 1/k\) with summation truncation up to \(k\)-nearest points. In the Nadaraya Watson approach \[2\], regression is the probability estimation of the data as follows

\[
\xi^q = \frac{\sum_{i=1}^{N} \xi^i R \left( \frac{\zeta^i - \zeta^q}{h(N)} \right)}{\sum_{i=1}^{N} R \left( \frac{\zeta^i - \zeta^q}{h(N)} \right)},
\]  

(3.9)

where \(R(.,.)\) is the kernel density and \(h(N)\) is the scale associated with the distribution of data. The factor \(R \left( \frac{\zeta^i - \zeta^q}{h(N)} \right) / \sum_{i=1}^{N} R \left( \frac{\zeta^i - \zeta^q}{h(N)} \right)\) can be interpreted as the conditional
probability density \( P^i(\xi^i|\zeta^q) \). Unfortunately, the Nadaraya-Watson type regression is highly sensitive to kernel selection. Also most of the generic kernels are highly scale \((h_N)\) dependent. Moreover, this technique suffers when neighborhood of the query point is very sparse.

Also one can exploit the concept of Reproducing Kernel Hilbert Space (RKHS) to construct the inverse map, \( f^{-1} \). For every RKHS, there is a corresponding unique PD kernel. However, most of the radial functions do not correspond to the PD kernel \[103\]. The generalized version of RKHS is the Reproducing Kernel Krein Space (RKKS), where the associated kernel is CPD. Note the geodesic kernel, \( K(\xi^i,\xi^j) = -D_M(\xi^i,\xi^j) \), is a CPD one \[87\]. This allows to represent each component as \[89, 76\]

\[
ξ^q_j = \sum_{i=1}^{N} θ^i_j K(\xi^i, ξ^q).
\]

(3.10)

Here, \( ξ^q_j \) is the \( j^{th} \) component of the solution vector \( (ξ^q) \) corresponding to its reduced space representation \( ζ^q \). The reproducing kernel, \( K(\xi^i, ξ^q) = -D_M(\xi^i, ξ^j) = -\|ξ^i - ξ^q\|_{L_2} \), is a scale free radial basis function (RBF) (note, \( \|ξ^i - ξ^q\|_M = \|ξ^i - ξ^q\|_{L_2} \), due to isometry.). The coefficients \( θ^i_j \) are determined using the training data set \((ξ^i, ξ^i)\) as follows

\[
\begin{bmatrix}
θ^1_1 & \cdots & θ^1_D \\
\vdots & \ddots & \vdots \\
θ^N_1 & \cdots & θ^N_D
\end{bmatrix}
= \begin{bmatrix}
R(ξ^1, ξ^1) & \cdots & R(ξ^1, ξ^N) \\
\vdots & \ddots & \vdots \\
R(ξ^N, ξ^1) & \cdots & R(ξ^N, ξ^N)
\end{bmatrix}^{-1}
\begin{bmatrix}
ξ^1_1 & \cdots & ξ^1_D \\
\vdots & \ddots & \vdots \\
ξ^N_1 & \cdots & ξ^N_D
\end{bmatrix}.
\]

(3.11)

However, it is worth mentioning that the solution in RKKS may not represent the global minima of the cost function \[76\]. Moreover, Bayesian framework \[61\] is very expensive in case of very high-dimensional regression and hence avoided in this work.
The successful implementation of this technique for the reconstruction of an extremely high dimensional manifold can be found in [9].

![Reconstructed Swiss-roll](image)

Figure 3.4. Reconstruction and convergence of RKKS based regression model. (a) Reconstructed Swiss-roll. (b) Convergence study of RKKS based regression.

To verify the ability of RKKS based regression towards reconstruction of the coordinates, three cases are tested: a cube \((20 \times 20 \times 20)\), a disk (radius 10) and the Swiss Roll (see Fig. 3.5a). In Figure 3.5b the convergence with the number of data points, \(N\) is provided. For the convergence study for all three cases, the common error measure, mean squared error (MSE) has been used. It is noticeable that for \(N = 100\), the error already has been reduced significantly except Swiss roll. Note, the scale of Swiss roll is much higher than the other two examples and also the geodesics computed here in an approximate sense (shortest path). Furthermore, it is worth noting that this is a pure interpolation technique. If the points are outside the convex-hull, the prediction error is high and it diverges as the point moves away from the boundary.
Figure 3.5. Reconstruction of coordinates for outside convex-hull points with RKKS based regression. (a) Reconstructed coordinates for 10 unit radius disk. (b) Divergence of RKKS model as it moves away from the boundary of the convex-hull.

of the convex-hull (see Fig. 3.5b, the error computed here is pointwise $L_2$ error.). Moreover, all these out of domain points gets projected into the boundary of the convex-hull (see Fig. 3.5a, for the disk). This is not surprising since the coefficients $\theta_i^j$ in Eq. (3.10) retains the convexity criteria. More details related to this topic can be found in [21].

3.3 Neural Network

The map $\Psi$ is essentially the relationship between the macroscale loading parameter space $\mathcal{X}$ and the reduced space $\mathcal{A}$ (see Fig. 3.2). The connection between the reduced space, $\mathcal{A}$, and the space of physical macroscale parameters, $\mathcal{X}$, is not straightforward for nonlinear problems. Since feed-forward neural network is a universal approximator [50, 95], in this work, the map between the reduced space, $\mathcal{A}$, and the macroscale loading parameter space, $\mathcal{X}$ is established through use of NN by minimizing the mean-square error [5, 50]. The neural network consists of an input
layer, hidden layers and the output layer (see Fig. 3.6).

One of the key issues in NN is to select the number of hidden layers and also the number of nodes in each layer. Each node in the hidden layer represents a feature. If NN architecture contains many redundant nodes, that automatically leads to overfitting, which is a very common problem with NN [11, 5]. To overcome this issue, in this work, Bayesian regularization [39, 46, 5] is used (the concept of $L_2$ regularization is given in Fig. 3.6). This concept is also known as weight decay. Another important issue with NN is the possibility of getting stuck to the local minima. Although the objective function (i.e., mean-square error) is convex, the convexity of the feature space is not guaranteed. Note, the global optima is guaranteed for convex function only when the feasible solution space is also convex. There are several techniques available to circumvent this issue. In this work, an ensemble NN is used to limit this problem [44]. There are two ways to implement ensemble NN. Very common one is to take the direct average over the output of the candidate NNs. The second one is more robust in a sense that the average is done in a weighted manner. The weights are determined by the singular value decomposition of the error matrix formed by the error associated with the candidate NNs. Also weights can be determined using genetic algorithm [111].

3.4 Discussion

The MNROM is an interpolation based reduced order model. It is evident from Figure 3.5b that in case of extrapolation the MNROM solution drifts away from the FEM solution. Also all the maps are constructed in an approximate sense. So, there are errors associated with each map. All the errors get accumulated when predicting for a query point. However, all the maps used in MNROM has convergence property. Unfortunately, it is impossible to construct MNROM by large number of data and of course that contradicts the primary objective of the reduced order model. Since
the entire quality of any data driven modeling rests on the ensemble nature of the
data, the grand challenge lies in how to construct a representative manifold, $\mathcal{M}$, by
limited number ($N$) of data points. This is an emphasis of this work. In Chapter 4, a
novel physics-driven sequentially adaptive technique has been discussed to construct
the solution manifold by sampling only non-redundant simulations. As indicated in
Chapter 2, CH solution can bifurcate under compressive loading condition. In such
situation, multiple MNROMs would have to be constructed and analyzed to study
the solutions corresponding to multiple paths.
CHAPTER 4

SEQUENTIALLY ADAPTIVE SAMPLING

The data driven models, as indicated in the previous chapter, rely on the ensemble nature of the data. Typically regression models improve with the size of the data. Unfortunately, all data driven models need to work with much less number of data than the requirement. So the main objective is to create an optimal input space such that the model shows the best predictability with the limited data set. Usually, for reduced order models, data comes from the computer simulation with low noise. In regression, the quadratic error function can be decomposed into two parts, bias and variance. Although there is a natural trade-off between bias and variance, lower bias is more expected for ROM. However, this can produce large variance. Bias term reduces with the complexity of the model used. On the other hand, variance decreases with the size of the data. With the limited number of data points in hand, the balance between bias and variance is maintained through regularization. Also bias error can be reduced when the sample points are uniformly distributed. Practical implementations in this regard involves Orthogonal array [79], Latin Hypercube design [67], Minmax/Maxmin design [53], entropy based design [20]. However, this is not correct when some parts of the input space show more sensitivity than the others. Moreover, when the output space is extremely high-dimensional and highly nonlinear, uniformity in the input space does not lead to uniformity in the output space because of the curse of dimensionality. All the sampling technique can be either an one-stage or sequential. The added advantage in sequential sampling technique is one can monitor the error stepwise. As the sample size reaches to produce desired error level, the data
enrichment process can be stopped. This also provides scope to select optimal data set. The main objective of the sequential sampling is to use the current information to select new data points. This adds adaptive nature in the sampling technique. Of course the sequential sampling without adaptation is no good than a one-stage sampling. In this work, a novel sampling technique has been proposed addressing curse of dimensionality, sensitivity of the input space for MNROM. This sampling technique is sequential in nature. Moreover, the adaptation technique is physics guided.

4.1 Physics Guided Sampling

As discussed before, the importance of the representativeness of the manifolds $\mathcal{M}$ is crucial. Therefore, a sequentially adaptive technique has been proposed to construct the solution manifold. In doing so, the available snapshot patterns are explored. This novel sampling strategy substantially limits the number of snapshots needed for accurate MNROM representation. Moreover, it can be used in several similar model reduction schemes. Any sampling technique relies on the definition of the design space. In this work, the parameter space, $\mathcal{X}$ stores the macroscopic deformation gradients $\mathbf{F}$. Using the polar decomposition, $\mathbf{F} = \mathbf{R}^{\mathbf{U}}$, and removing the pure rotation part, the stretch tensor, $\mathbf{U}$ can be represented as

$$\mathbf{U} = \lambda_1 (e^1 \otimes e^1) + \lambda_2 (e^2 \otimes e^2) + \lambda_3 (e^3 \otimes e^3),$$

(4.1)

where $\lambda_i \in [0, \infty)$ are principal stretches and $e^i$ are the orthogonal principal directions generated by three angles $\varphi^i$ (the axis-angle representation). Axis is characterized by the angles $\varphi^1 \in [0, 2\pi]$, $\varphi^2 \in [0, \frac{\pi}{2}]$, and the rotation about the axis $\varphi^3 \in [0, \frac{\pi}{2}]$. The parameter space $\mathcal{X}$ can be expressed as a cartesian product of two spaces: i) the stretch space $\mathcal{X}_\lambda$, and ii) the rotational space $\mathcal{X}_\varphi$ as depicted in Figure 4.1. In the
numerical implementation, stretches $\lambda_i$ are from closed interval, and thus, the parameter space, $\mathfrak{T}$, is a compact subset of Euclidean space. The objective of this analysis is to better understand the distribution pattern for the rotation parameters.

![Diagram](image)

Figure 4.1. Concept of the physics based construction of the manifold.

### 4.1.1 Rotational Sensitivity

A concept of the rotational sensitivity analysis is illustrated in Figure 4.2. In this analysis, the stretch space is a cube with the line of volumetric deformations (the dotted line in Figure 4.2) as the main diagonal of the cube. All the points on this line manifest pure volumetric deformation ($\lambda_1 = \lambda_2 = \lambda_3$). From the physical perspective, the pure volumetric deformations are invariant of rotation. Therefore, if the stretch vector is near the volumetric deformation line, its rotational dependency will be less. Moreover, volumetric part of the energy density in Eq. (2.6) is more
dominant in the high strain regime. In rotational sensitivity analysis, the space $\Sigma_{\lambda}$ is covered by the finite number of open sets $U_{\Sigma_{\lambda}}^i$ (see Figure 4.1). Thus, the parameter space $\Sigma$ can be represented as

$$\Sigma = (\cup U_{\Sigma_{\lambda}}^i) \times \Sigma_{\varphi} = \cup(U_{\Sigma_{\lambda}}^i \times \Sigma_{\varphi}).$$

(4.2)

Since the manifold $\mathcal{M}$ is compact, the submanifolds, $\mathcal{M}_{\lambda}^i$, can be defined as

$$\mathcal{M}_{\lambda}^i = \mathcal{F}(U_{\Sigma_{\lambda}}^i \times \Sigma_{\varphi}),$$

(4.3)

where $\mathcal{M} = \bigcup \mathcal{M}_{\lambda}^i$. Note, these submanifolds also have the same dimension, $d$, since the $\mathcal{M}$ is smooth. In the discrete setting, submanifolds are constructed as a subset of points with all rotations for a specific stretch vector $\vec{\lambda}^i$ (see Fig. 4.1). Therefore, the discrete space $\mathcal{M}_{\lambda}^i$ reads

$$\mathcal{M}_{\lambda}^i = \mathcal{F}(\vec{\lambda}^i \times \Sigma_{\varphi}),$$

(4.4)

for the finite number of stretch points. Note, the number of generating parameters for $\mathcal{M}_{\lambda}^i$ is 3, which implies $\mathcal{M}_{\lambda}^i$ is a 3-manifold. Here, the subgraph, $G_r^i$, which graphically represents the submanifold, $\mathcal{M}_{\lambda}^i$ (see Fig. 4.2). The graph diameter, $d_r^i$, represents the diameter of the minimum covering ball, $B_{\lambda}^i$, for the corresponding submanifold $\mathcal{M}_{\lambda}^i$ (see Fig. 4.2).

Physically, it is expected that the rotational dependency will increase with the perpendicular distance, $r_{\lambda}^i$, from any point on the line of the volumetric deformation (see Figure 4.2). It is also important to note that the effect is not the same for
Figure 4.2. Concept of the rotational sensitivity analysis.

all volumetric deformations. As the volumetric strain energy increases, the graph diameter should be a function of the volumetric deformation (third invariant of the strain tensor, $I_3$) and the distance, $r^i_\lambda$, of the stretch vector from the corresponding pure volumetric point. In the present implementation, this coupling is neglected.

This rotational sensitivity analysis allows optimal distribution of rotational points for the corresponding number of stretches. Lower diameter will suggest less rotational dependency and the higher diameter will represent the higher rotational dependency of that particular point in the stretch space. From this relationship, one can decide the number of rotational points require for a particular point in the stretch space or equivalently construct the representative submanifold $\mathcal{M}^i_\lambda$. It is worth mentioning that the approximate diameter $d^i_r$ computed over the graph $G^i_r$ obeys the inequality given in Eq. (3.5). The proposed adaptive data enrichment technique utilizes the concept of rotational sensitivity analysis to represent the manifold by enriching the submanifolds and eliminating possible redundant points.
4.1.2 Adaptive Snapshot Redistribution

In the first step, the process starts with a few simulations, \( N_0 \). After determining all diameters, \( d_i^\lambda \), a relationship can be established with the distance, \( r_i^\lambda \). This function allows to adaptively add extra points in the stretch space, \( \Xi_\lambda \). To do so, the average number of points, \( \bar{n}_g \), in the \( \epsilon \) neighborhood over all points in the entire dataset is defined as follows

\[
\bar{n}_g = \frac{1}{N} \sum_{i=1}^{N} n_{gi}, \quad (4.5)
\]

where \( n_{gi} \) is the number of points within \( \epsilon \) neighborhood centering any point \( \xi_i \) on the manifold, \( \mathcal{M} \). Next, the \( \bar{n}_g \) is computed for all data points associated with each closed domain, \( U_{\Xi_\lambda}^i \), for the same \( \epsilon \) neighborhood. If \( \bar{n}_g \) is less than \( \text{max}(\bar{n}_g, k) \), an extra stretch point can be added in the center of the \( U_{\Xi_\lambda}^i \) (see a filled triangle in the 2D representation in Fig. 4.1). Here, \( k \), is the \( k \)-nearest neighbor of the graph representation of the manifold, \( \mathcal{M} \), and \( \bar{n}_g \) measures the average data density. This strategy helps to add data points in the sparse regions of the manifold by increasing the average local data density, \( \bar{n}_g \) to more than \( k \) and \( \bar{n}_g \). In this work, the space \( U_{\Xi_\lambda}^i \) is constructed using the tetrahedral refinement. Note, any extra points \( (N_e) \) will be added to the manifold, \( \mathcal{M} \), by enriching the submanifolds, \( \mathcal{M}_i^\lambda \). Requirement of the number of enrichment points are proportional to the volume of the submanifolds, \( V_{\mathcal{M}_i^\lambda} \). To accomplish this purpose, the volume of these submanifolds need to be computed. However, it is difficult to compute \( V_{\mathcal{M}_i^\lambda} \) exactly from a few points. Therefore, the tacit assumption is the volume \( V_{\mathcal{M}_i^\lambda} \) is proportional to the volume of its minimum covering ball, \( B_i^\lambda \), which yields
Here, $\Gamma$ is the gamma function and $\pi = 3.14159\ldots$ is a constant. The number of submanifolds $N_\lambda$ is the same as the number of the points in the stretch space. The number of enrichment points $N^i_\varphi$ in the rotational space, $\Sigma_\varphi$, for the corresponding stretch point is calculated as follows

$$N^i_\varphi = \frac{V^i_i M_\lambda}{\sum_{i=1}^{N_\lambda} V^i_i M_\lambda} N_e = \frac{V^i_i B_\lambda}{\sum_{i=1}^{N_\lambda} V^i_i B_\lambda} N_e.$$  (4.7)

$N^i_\varphi$ takes the next integer value as computed from Eq. (4.7). The tacit assumption behind the above algorithm is that the manifold $\mathcal{M}$ is the disjoint union of the $N_\lambda$ number of submanifolds ($\mathcal{M} = \bigsqcup \mathcal{M}_\lambda$). To verify this assumption, it is better to compute the degree of overlap $\mathcal{P}_O$ as follows

$$\mathcal{P}_O = \left(1 - \frac{V_I}{V_M}\right) \times 100\%,$$  (4.8)

where $V_I$ is the volume of the data associated with the interface of the closed domains, $U^i_{\Sigma_\lambda}$, with all attached rotations. Note, that the interface regions are counted twice during volume calculations.

4.2 Discussion

To the best of my knowledge, the proposed technique is not implemented in any other work. The novelty of this technique is many folds. First of all, it can handle extremely high-dimensional space. Most importantly, it can identify the sensitive
parts of the design space from the available information and distribute the new points accordingly. However, determination of the optimal number of simulations for the reduced order model can only be done partially. Finding that number still remains an open challenge.
CHAPTER 5

NUMERICAL EXAMPLE

Here, the MNROM framework is tested for predicting the behavior of randomly configured (statistically isotropic) particulate materials. As indicated in Chapter 2 for CH, the definition of RUC is very important. In order to select the size of the representative microscopic cell, a statistical approach \[43, 100\] is adopted. Three different cell sizes, 8R, 10R and 16R, generated using the packing algorithm Rocpack \[94\], where R is the radius of the inclusion. Here, the inclusions are monodisperse spheres with radius R=0.5 \(\mu\)m.

Figure 5.1 shows the isotropic second order probability function, \(S_{pp}\) (i.e. particle-particle probability function), and its directional covariance. As typical, larger cells preserve more structure and are more representative. From this statistical analysis, a RUC with a side length of \(l_{RUC} = 10R = 5\mu\)m containing 95 particles is selected with the covariance error below 5%. More details on construction of a RUC can be found in \[43, 100\]. Figure 5.2 shows the geometry of the cell used in this work. After a convergence study, the RUC is discretized with 96,252 nodes and 486,051 elements, which results in 250,035 nonlinear degrees of freedom.

With the discretized RUC, appropriate constitutive models at the microscale are selected. Here, an isotropic hyper-elastic strain-energy density function similar to the Mooney-Rivlin hyper-elastic model is employed with the volumetric and deviatoric
Figure 5.1. Statistical analysis of the RUC. (a) Isotropic second-order particle-to-particle probability function. (b) Directional covariance of the second-order particle-to-particle probability.

Figure 5.2. RUC with side length of $l_{RUC} = 10R = 5 \mu m$. The cell contains 95 spherical particles with 1 \( \mu m \) diameter.
split, $^1W(C) = ^1W_C(\hat{C}) + ^1W_J(J)$, with

$$^1W_C(\hat{C}) = \mu_{10} \left[ \text{tr}(\hat{C}) - 3 \right] + \frac{\mu_{01}}{2} \left[ (\text{tr}\hat{C})^2 - \hat{C} : \hat{C} - 6 \right],$$

$$^1W_J(J) = \frac{\kappa}{2} \left[ \exp(J - 1) - \ln(J) - 1 \right].$$

In Equation (5.1), $J = \det(F)$ is the Jacobian of the deformation, $\hat{C} = J^{-2/3}C$ is the deviatoric right Cauchy-Green deformation tensor, $\mu_{10}$ and $\mu_{01}$ are the shear moduli, and $\kappa$ is the bulk modulus. The associated material properties of the microscale constituents are provided in Table 5.1. With this microstructure, next the MNROM framework has been verified for both with and without sampling technique.

**TABLE 5.1**

<table>
<thead>
<tr>
<th>Material</th>
<th>E [MPa]</th>
<th>$\nu$</th>
<th>$\kappa$ [MPa]</th>
<th>G [MPa]</th>
<th>$\mu_{01}$ [MPa]</th>
<th>$\mu_{10}$ [MPa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Particle</td>
<td>5e3</td>
<td>0.25</td>
<td>3.33e3</td>
<td>2e3</td>
<td>5e2</td>
<td>5e2</td>
</tr>
<tr>
<td>Matrix</td>
<td>50</td>
<td>0.35</td>
<td>55.56</td>
<td>18.52</td>
<td>4.63</td>
<td>4.63</td>
</tr>
</tbody>
</table>

5.1 One-stage Design of MNROM

The parameter space for MNROM is provided in detail in Section 4.1 of Chapter 4. In order to parameterize the principal directions associated with the principal stretches, the axis $^0e^1$ is constructed from the position vectors of points on the unit hemisphere ($X^3 \geq 0$ in Figure 5.3), which are characterized by the angles $\varphi^1 \in [0, 2\pi]$,
\( \varphi^2 \in [0, \frac{\pi}{2}] \). HEALPix grid \([40]\) is used to obtain a uniform discretization of points on the hemisphere. For the rotation about the axis \( \mathbf{e}_1 \) (see Figure 5.3), the angle \( \varphi^3 \) is selected in \([0, \frac{\pi}{2}]\). Thus, the set \( \mathcal{X} = \{0_{\lambda_1}, 0_{\lambda_2}, 0_{\lambda_3}, \varphi^1, \varphi^2, \varphi^3\} \) represents the space of \( \mathbf{U} \) and each loading condition \( \eta^i \in \mathcal{X} \) is described by six generators.

Since CH is a highly nonlinear process, two different loading conditions can produce the similar microscale solutions or snapshots. Similarity of the microscale data can lead to misrepresentation of the manifold, which is referred to as the topological instability \([4]\). To avoid any potential topological instability, in this work, the digital database is divided into four separate loading segments (referred to as modes 1-4) and is constructed as described in Table 5.2. Here, mode 1 is purely tensile and mode 4 is purely compressive. On the other hand, mode 2 and 3 are mixed ones. Next, each of these manifolds are constructed separately for the reduced order model.

![Figure 5.3. Schematic of the rotation parameters.](image)

After a discretization study, the principal stretch space can be approximated with a \( 4 \times 3 \times 3 \) uniform grid over 10\% stretch interval. This leads to large strain conditions.
Moreover, 112 distinct rotations are used to generate the bases $\begin{bmatrix} 0 & e_1 & e^2 & 0 & e^3 \end{bmatrix}$. Note that this discretization leads to 4032 solutions of the microscale boundary value problem for each mode ($N = 4 \times 4032$). After creating the digital database, the next step is to construct the HD manifold $\mathcal{M}$. Although any field could have been used, for this one-stage construction of MNROM, the manifold has been represented based on the deformation gradient $F$. More detail on manifold representation can be found in Chapter 3.

Although this is a moderately large number of simulations, all simulations are executed in parallel using the parallel $PGFem3D$ solver [66, 72]. Thus, the whole database can be generated quickly given current HPC capabilities. Although no numerical difficulties have been encountered while solving the nonlinear BVP involved
TABLE 5.2

DESCRIPTION OF THE MACROSCALE LOADING MODES

<table>
<thead>
<tr>
<th>Loading Case</th>
<th>Description</th>
<th>No. Simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode 1</td>
<td>$\lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0$</td>
<td>4032</td>
</tr>
<tr>
<td>Mode 2</td>
<td>$\lambda_1 \leq 0, \lambda_2 \geq 0, \lambda_3 \geq 0$</td>
<td>4032</td>
</tr>
<tr>
<td>Mode 3</td>
<td>$\lambda_1 \leq 0, \lambda_2 \leq 0, \lambda_3 \geq 0$</td>
<td>4032</td>
</tr>
<tr>
<td>Mode 4</td>
<td>$\lambda_1 \leq 0, \lambda_2 \leq 0, \lambda_3 \leq 0$</td>
<td>4032</td>
</tr>
</tbody>
</table>

in CH for this example. However, when such numerical difficulties occur, the manifold would experience coverage deficiency that would have to be mitigated by conducting extra numerical simulations in regions of convergence impediment. Note that even this seemingly large number of simulations still leads to a very sparse manifold $M$ considering the high-dimensionality ($D = 9 \times N$), which is commonly referred to as the curse of dimensionality. As discussed in Chapter 3, to establish the map, $f$, Isomap is used for MNROM. Isomap gives a reduced dimensionality of $d = 6$ for each loading mode down from the complete solution manifold with the dimension of $D = 4,374,459$. The residual correlation coefficient between the Euclidean and the approximate geodesic distances of the complete solution manifold are below 0.1 for all modes (see Figure 5.4). Moreover, near perfect exponential convergence is obtained. Note that modes 2 and 3 are more nonlinear and have initially higher residual variance. As anticipated, multiple loading conditions can have the similar solution which leads to non-uniqueness of the deformation gradient. For example, pure volumetric deformations are rotation invariant. Although, from Isomap point of view redundant data points are ineffectual, all non-unique solutions are retained for the enrichment of the model.
Since all the maps associated with the MNROM are constructed in an approximate sense, verification of each map is extremely important from data-driven modeling perspective. As a next step, all the maps (i.e., inverse map, NN map) are verified individually.

5.1.1 Verification of Inverse Map, $f^{-1}$

In order to understand the errors associated with $f^{-1}$, the common leave-one-out procedure is employed from the machine learning community [26]. This procedure consists of removing a single point from the manifold construction, interpolating this removed point from the reduced space, and comparing the interpolated point to the known one. 10% of the total 4,032 points (selected randomly) for each loading mode are validated using this method.

The error associated with removing one of these random points from the manifold is quantified by considering the local (localization) error ($\mathcal{E}_r$) computed at micro-points $Y$ (i.e. each finite element in the RUC discretization). This error measure is defined as

$$\mathcal{E}_r = \frac{\|F^{\text{ROM}}_r - F^{\text{FEM}}_r\|_F}{\|F^{\text{FEM}}_r\|_F} \times 100 \,[\%], \quad (5.2)$$

where $F^{\text{ROM}}_r$ and $F^{\text{FEM}}_r$ are the deformation gradients computed from the regression model and the FEM simulation, respectively. Here, $(\| \cdot \|_F)$ is the Frobenius norm. The subscript $r$ represents a material phase (particle or matrix). In order to quantify the cumulative error for a single simulation point removed from the manifold, the following volume averaged (homogenization) error, $^0\mathcal{E}_r$, reads

48
\[ 0\epsilon_r = \frac{1}{\Theta_0} \int_{\Theta_0} \epsilon_r d\Theta \parallel 0\mathbf{F}_r \parallel_F, \]  

(5.3)

where \( 0\mathbf{F}_r = \frac{1}{\Theta_0} \int_{\Theta_0} \mathbf{F}_r d\Theta \).

Figures 5.5a and 5.5b show the cumulative error (homogenization) of the leave-one-out cases for the particles and matrix, respectively. Error characteristics of the inverse map are shown in Table 5.3. Note that the largest error over all one-out points occurs for mode 3 (circular markers in Figure 5.5a and 5.5b). Therefore, further localization error analysis is conducted only for mode 3.

![Figure 5.5a](image1.png)

![Figure 5.5b](image2.png)

Figure 5.5. Distribution of the volume averaged error (homogenization) given by Equation (5.3) for randomly selected 10% of points from database (one-out cases). (a) Particle. (b) Matrix.

The distributions of the element-wise (localization) error, are shown in Figures 5.6a
and 5.6b for the matrix and particles, respectively. To simplify the interpretation of the results, Figure 5.6 shows the mean and standard deviation of the localization error, $\varepsilon_r$, computed over the finite element field (i.e. deformation gradient) given in Equation (5.2) from leave-one-out points. As can be seen, the maximum localization error occurs in the compliant matrix and is more uniformly distributed. Moreover, all errors (both localization and homogenization) are below 1.5%, which demonstrates good construction of this inverse map.

<table>
<thead>
<tr>
<th></th>
<th>Matrix</th>
<th>Particles</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean($\varepsilon_r$) [%]</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>Frequency [%]</td>
<td>0.5</td>
<td>0.75</td>
</tr>
<tr>
<td>std($\varepsilon_r$) [%]</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Frequency [%]</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Figure 5.6. Distribution of the local error (localization) given by Equation (5.2) for randomly selected 10% of points from mode 3 (one-out cases). (a) Mean error of the deformation gradient field. (b) Standard deviation of the deformation gradient field.

Finally, the local distribution of the deformation gradient magnitude $\|F\|_F$ is shown in Figures 5.7a and 5.7b. This distribution is generated for data point, $\xi^i$, with mean($\varepsilon_r$) = 1.2367% error from mode 3. Note that these distributions obtained
through the regression model match the simulation data extremely well. This indicates that an accurate microscale fields are obtained from the reduced space using this approximate inverse map, $f^{-1}$.

5.1.2 Verification of Map, $\mathfrak{F}$

To complete the MNROM framework, a map is established between the reduced space $\mathcal{A}$ and the space of macroscopic loading parameters $\mathcal{X}$. This map is denoted as $\mathfrak{F}$ (see Figure 3.2). Since the process of generating data is highly nonlinear and the neighborhood structure is destroyed due to the curse of dimensionality (recall that $D = 9 \times Ne = O(10^6)$), a multilayer feedforward network is employed which is very efficient at establishing a large class of measurable maps [50].

Since NN is not an explicit map, the performance of the model depends on many parameters. The number of hidden layers and the number of nodes are important issues for better generalization of the model. The architecture of NN used in this

\begin{figure}[h]
\centering
\begin{subfigure}{0.45\textwidth}
\includegraphics[width=\textwidth]{figure_a}
\caption{(a) Over Matrix.}
\end{subfigure}
\begin{subfigure}{0.45\textwidth}
\includegraphics[width=\textwidth]{figure_b}
\caption{(b) Over Particle.}
\end{subfigure}
\caption{Distribution of $\|F\|_F$ from FEM and MNROM. (a) Over Matrix. (b) Over Particle.}
\end{figure}
TABLE 5.3

QUALITY OF THE REGRESSION MODEL FOR CONSTRUCTION OF THE INVERSE MAP, $f^{-1}$, FOR ALL MODES COMPUTED FROM EQUATION (5.3)

<table>
<thead>
<tr>
<th>Loading Cases</th>
<th>Min. [%]</th>
<th>Max. [%]</th>
<th>Mean [%]</th>
<th>Median [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>1.8410e-07</td>
<td>0.3161</td>
<td>0.0314</td>
<td>0.0028</td>
</tr>
<tr>
<td>Matrix</td>
<td>4.6437e-07</td>
<td>0.9589</td>
<td>0.1586</td>
<td>0.0314</td>
</tr>
<tr>
<td>Mode 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>2.6032e-06</td>
<td>0.5846</td>
<td>0.0424</td>
<td>0.0226</td>
</tr>
<tr>
<td>Matrix</td>
<td>4.8607e-06</td>
<td>1.1823</td>
<td>0.2560</td>
<td>0.2395</td>
</tr>
<tr>
<td>Mode 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>1.6000e-03</td>
<td>0.6870</td>
<td>0.0530</td>
<td>0.0381</td>
</tr>
<tr>
<td>Matrix</td>
<td>2.7100e-02</td>
<td>1.2367</td>
<td>0.4002</td>
<td>0.3827</td>
</tr>
<tr>
<td>Mode 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>3.0786e-04</td>
<td>0.3472</td>
<td>0.0386</td>
<td>0.0181</td>
</tr>
<tr>
<td>Matrix</td>
<td>5.7000e-03</td>
<td>1.2755</td>
<td>0.2309</td>
<td>0.1553</td>
</tr>
</tbody>
</table>
work is two hidden layers with 70 nodes per layer for modes 1, 2 and 4. Mode 3 contains 75 nodes per layer. This selection is guided by the fact that in general a two-layer NN is sufficient to capture nonlinear features \[95\].

<table>
<thead>
<tr>
<th>Loading Case</th>
<th>Regression Coefficient</th>
<th>Performance</th>
<th>Target Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode 1</td>
<td>0.99879</td>
<td>5.3934</td>
<td>([-150, 150])</td>
</tr>
<tr>
<td>Mode 2</td>
<td>0.99894</td>
<td>9.3900</td>
<td>([-220, 220])</td>
</tr>
<tr>
<td>Mode 3</td>
<td>0.99940</td>
<td>5.3042</td>
<td>([-220, 220])</td>
</tr>
<tr>
<td>Mode 4</td>
<td>0.99914</td>
<td>3.8544</td>
<td>([-150, 150])</td>
</tr>
</tbody>
</table>

In this work, the NN Toolbox in MATLAB is utilized [5]. The target-function/error is computed by a mean-square error (MSE). Common regularization techniques are adopted to optimize the regression model [5]. Furthermore, the Bayesian model for multilayer feedforward networks [11, 5] is explored to train the data. The quality of the regression is summarized in Table 5.4 for all loading modes. Note that the linear regression coefficient is close to one for all cases indicating a good fit. The performance of a NN should be judged relative to the range of the target space, and the MSE performance is on the order \(O(10^{-4})\) in the range \([-1, 1]\).

To investigate further the quality of this map, the error between \(\mathcal{A}\) and \(\overline{\mathcal{A}}\), which is the output of NN, is illustrated for each of the loading modes in Figure 5.8. The geometric interpretation of this error is simply the Euclidean distance between points.
Figure 5.8. Distribution of Neural Network output error compared to reduced space over entire dataset.

in $\mathcal{A}$ and the corresponding points in $\overline{\mathcal{A}}$. Figure 5.8 also shows that the average shift of the MNROM solution from the corresponding FEM along the manifold $\mathcal{M}$ is around 5 units (around 7 units in mode 2) and 4.57% of points are more than 10 units in terms of geodesic distance. Moreover, the spread of the data is large as indicated in Table 5.4 (Note the range of the six-dimensional space $\mathcal{A}$ (Target Range)) and this difference is accumulated over $D = 4, 374, 459$ components. Therefore, the map established by means of NN provides a useful transformation of the macroscale physical parameters to the reduced space and it will not deteriorate the physics unless the process is highly sensitive to the macroscale loading parameters.

5.1.3 Verification of Complete MNROM Framework

After creating the digital database and establishing maps between the high dimensional space of microscale data, the reduced space, and the macroscale loading parameters, the MNROM framework (see Figure 3.2) can be used to compute macroscale
(homogenization) as well as microscale (localization) fields for a given macroscale input that is not present in the digital database. The input query is denoted as $\eta^q$, which corresponds to a point in $\mathcal{T}$ (the macroscale loading parameters). This query point is mapped to the reduced space $\mathcal{A}$ using the map $\mathcal{P}$ derived from the NN. Finally, the microscale data is obtained using the approximate inverse, $f^{-1}$.

Now the entire MNROM framework is checked by computing the microscale response, $\xi^q$, of 100 random query points, $\eta^q$, for each mode taken within the range of the $\mathcal{T}$ space. A total of 400 random query points are generated (i.e. 100 in each mode) from a uniform distribution over the macroscopic loading parameter space $\mathcal{T}$. To ensure that the query points are well spaced and uncorrelated with the existing dataset, I first generate 10,000 samples from a uniform distribution in the macroscopic loading space and refer this set as $\mathcal{T}^*$. Then I pick the query point $\eta^q$ from $\mathcal{T}^*$ which produce the maximum Euclidean distance from the existing loading set, $\mathcal{T}$. After picking $\eta^q$, I remove this point from $\mathcal{T}^*$ and add it to $\mathcal{T}$. This process is iterated for 100 times. Once this process ends, all these query points are taken out from $\mathcal{T}$ and I construct the set of query points, $\mathcal{T}^q$. The idea is similar to maxmin design.

Recall that $\xi^q$ is an ordered list of the microscale deformation gradient, $F$, for each finite element in the discretized microstructure (RUC). The error of the interpolated microscale field is evaluated by comparison with simulation data using Equation (5.3) and (5.2), respectively.

The homogenization error is shown in Figures 5.9a and 5.9b for particle and matrix respectively. Figure 5.9 shows that the homogenization error in $F_m$ is higher (Figure 5.9b), but only less than 13% of query points have a volume averaged error above 5% considering all modes. This is strong performance considering that no FEM analysis was performed, and the deformation gradient field (represented by $Ne = 486,051$ finite elements) is approximated using MNROM. On the other hand, some of the query points (less then 1.25%) have a homogenization error above 10%.
Figure 5.9. Distribution of the volume averaged error (homogenization) given by Equation (5.3) of the 100 randomly selected query points in each mode. (a) Particle. (b) Matrix.

These are usually query points close to the loading envelope or in a region where the manifold data density is substantially low.

To understand these errors, note that Isomap is not based on exact geodesics. Therefore, the reduced space $\mathcal{A}$ does not unfold the manifold $\mathcal{M}$ perfectly. This indicates that some degree of noise is involved in the data associated with $\mathcal{A}$. Since the accuracy of the computed geodesics directly depends on the density of data, the degree of noise indicates the amount of data sparsity. Insufficient data density also leads to inadequate NN learning \cite{101,18} and deterioration of $\mathcal{M}$ map. An efficient sampling technique is needed to reduce this error. Nevertheless, these first results indicate a strong potential of this novel manifold-based reduced order method.

Minimum, maximum, mean and median of the volume averaged error over all query points for all modes are listed in Table 5.5. Note that the data reconstruction process is preserved using MNROM with the mean error less than 3.5% (median error...
<table>
<thead>
<tr>
<th>Loading Cases</th>
<th>Min. [%]</th>
<th>Max. [%]</th>
<th>Mean [%]</th>
<th>Median [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mode 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>0.0464</td>
<td>4.1540</td>
<td>0.2860</td>
<td>0.2469</td>
</tr>
<tr>
<td>Matrix</td>
<td>0.2413</td>
<td>11.5384</td>
<td>2.2622</td>
<td>1.8902</td>
</tr>
<tr>
<td><strong>Mode 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>0.0275</td>
<td>0.7621</td>
<td>0.1737</td>
<td>0.1235</td>
</tr>
<tr>
<td>Matrix</td>
<td>0.3989</td>
<td>9.6881</td>
<td>2.5506</td>
<td>1.8707</td>
</tr>
<tr>
<td><strong>Mode 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>0.0310</td>
<td>0.9997</td>
<td>0.2308</td>
<td>0.1274</td>
</tr>
<tr>
<td>Matrix</td>
<td>0.4140</td>
<td>11.8873</td>
<td>3.3761</td>
<td>1.9683</td>
</tr>
<tr>
<td><strong>Mode 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>0.0905</td>
<td>2.4268</td>
<td>0.3194</td>
<td>0.2869</td>
</tr>
<tr>
<td>Matrix</td>
<td>0.6626</td>
<td>10.2583</td>
<td>2.2163</td>
<td>1.7815</td>
</tr>
</tbody>
</table>
less than 2\%) for all cases. It is notable that among all modes, the homogenization error from 100 query points is higher for mode 3. Therefore, I analyze the mode 3 error in more depth.

The mean and standard deviation of the localization error defined in Equation \((5.2)\) computed over the finite element field (i.e. deformation gradient) for 100 random query points from mode 3 are shown in Figures 5.10\(\text{a}\) and 5.10\(\text{b}\). Minimum, maximum, mean and median of the mean \((\mathcal{E}_m)\) error evaluated over the deformation gradient field for the population of 100 random query points from mode 3 are listed in Table 5.6. Similar to the homogenization process (See Figure 5.9), the localization error is of the same order. Moreover, since \(\mathbf{F}\) is equal to \(\mathbf{U}\) in this analysis, the quality of the microscopic strain field can be judged from Figure 5.10.
TABLE 5.6

LOCALIZATION ERROR CHARACTERISTICS OF MNROM FOR MEAN($\varepsilon_m$) ERROR FROM EQUATION (5.2) (SEE ALSO FIGURE 5.10) OF THE DEFORMATION GRADIENT FIELD FOR POPULATION OF 100 RANDOM QUERY POINTS FROM MODE 3

<table>
<thead>
<tr>
<th></th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.4519</td>
<td>12.5747</td>
<td>3.6266</td>
<td>2.1200</td>
</tr>
</tbody>
</table>

5.1.3.1 Homogenization of Material Response

To understand further the quality of MNROM solution, the homogenized strain energy density function is computed. Figure 5.11 shows the deviatoric (Figure 5.11a) and volumetric (Figure 5.11b) parts of the strain energy density function. The volumetric potential, $^0W_J(J)$, shows remarkable correlation between detailed parallel simulations using PGFem3D and MNROM (see Figure 5.11b). A moderate error from MNROM is noticeable, especially for large volume changes ($J < 0.8$ and $J > 1.2$). The deviatoric potential, $^0W_C(\hat{C})$, comparison is more difficult to interpret since it depends on a second-order tensor. Therefore, I plot $^0W_C(\hat{C})$ as a function of $\|\hat{C}\|$ in Figure 5.11a. Furthermore, to carry out an error estimate for the homogenized potential energy, an error measure can be defined as

$$\mathcal{E}_{0W} = \left| \frac{^0W_{ROM} - ^0W_{FEM}}{^0W_{FEM}} \right| \times 100\%.$$  \hspace{1cm} (5.4)

As observed in Figure 5.11, the total potential energy, $^0W$, ranges from 0.05 MPa to 3 MPa. The median error ($\mathcal{E}_{0W}$) is 10%, and 75% of the cases are below 20% error.
Figure 5.11. Macroscale strain energy density function, MNROM versus FEM. (a) The deviatoric potential, $^0W_C(\hat{C})$. (b) The volumetric potential, $^0W_J(J)$.

Figure 5.12. Least-square fit of the macroscale strain energy density function, MNROM versus FEM. (a) The deviatoric potential, $^0W_C(\hat{C})$. (b) The volumetric potential, $^0W_J(J)$. 

60
as computed from Equation (5.4). Moreover, points with high error (> 20%) are concentrated in low to moderate potential energy regions. From Figure 5.11 it can be seen that the volumetric energy is higher compared to the deviatoric one for large strains. Considering the mode construction (recall Table 5.2), the principal stretches are either all positive or all negative in modes 1 and 4, respectively. Therefore, the portions of high strain and potential energy are associated with modes 1 and 4 (see Figure 5.11b). Moreover, the points with increased errors are close to the loading envelopes of modes 1 and 4. Points with low to moderate potential energy and low volumetric deformation are mostly in modes 2 and 3 (see Table 5.2), where the signs of the principal stretches are mixed. The data pattern for modes 2 and 3 is quite complex, and data is sparse compared to modes 1 and 4 (see Target Range in Table 5.4). Moreover, the degree of nonlinearity of modes 2 and 3 is higher (see decay of residual variance in Figure 5.4). Overall, the results suggest an acceptable agreement between detailed FEM simulations and MNROM. Once more, this is an one-stage construction and data has not been efficiently sampled. Nonetheless, Figure 5.11 shows that the MNROM discovers with reasonable accuracy the homogenized constitutive laws for a large number of query points. Finally, the functional form of these macroscale potentials is not assumed *a priori* and is discovered from our data without any additional large parallel simulations.

Since the spread in the data complicates the analysis (especially for the deviatoric portion of the potential function), a model calibration is conducted to construct the homogenized potentials with volumetric and deviatoric split as $^0W(C) = ^0W_C(\hat{C}) + ^0W_J(J)$. Although the functional form of the macro-potential can be obtained from results in Figure 5.11 the expressions of the micro-potential functions in Equation (5.1) are considered for simplicity. Figure 5.12 depicts model calibration using the volumetric (Figure 5.12b) and deviatoric (Figure 5.12a) functions (see Equation (5.1)) for both the FEM and MNROM data from Figure 5.11. Both the
FEM and MNROM parameterizations show excellent agreement for both the volumetric and the deviatoric potentials. Finally, the homogenized material constants for both FEM and MNROM models are listed in Table 5.7.

### TABLE 5.7

**CALIBRATED HOMOGENIZED MATERIAL PROPERTIES**

<table>
<thead>
<tr>
<th>Source</th>
<th>$\kappa$ [MPa]</th>
<th>$\mu_{10}$ [MPa]</th>
<th>$\mu_{01}$ [MPa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNROM</td>
<td>105</td>
<td>13.45</td>
<td>13.45</td>
</tr>
<tr>
<td>FEM</td>
<td>94</td>
<td>12.40</td>
<td>12.40</td>
</tr>
</tbody>
</table>

Although one can derive the homogenized stress after learning the functional form of the homogenized potential, here I also like to check the quality of the homogenized stress computed from the microscale field using MNROM. I compare this stress with the FEM. In Figure 5.13(a) I plot the Frobenius norm of the macroscopic first Piola stress tensor, $^0P = \frac{1}{|\Theta_0|} \int_{\Theta_0} P \, d\Theta$, with the corresponding Frobenius norm of macroscopic deformation gradient, $^0F$, for all four modes. Note, $^0F$ is the energy conjugate to $^0P$. It can be seen that the points corresponding to $\| ^0P \|$ computed by MNROM are very close to the points which are calculated using the FEM. To close the comparison, I compute the error as follows

$$
\mathcal{E}_{^0P} = \frac{\| ^0P^{ROM} - ^0P^{FEM} \|}{\| ^0P^{FEM} \|} \times 100\%.
$$

(5.5)
However, it is not easy to interpret a second-order tensor from its norm alone. Therefore, the error plot is shown in Figure 5.13b where 99.5% of the query points have less than $E_{0P} = 5\%$ error, and 67.25% of all query points have less than 1% error for all modes.

5.1.3.2 Localization of Material Response

In the previous section, homogenization properties of MNROM are investigated. Here, the result related to localization accuracy of MNROM and its ability to predict the local (element-wise) engineering fields derived from $F$ are presented. The Almansi strain, $e$, is defined as

$$e = \frac{1}{2}(1 - F^{-1}F^{-T}).$$  \hspace{1cm} (5.6)
To quantify the strain error between the FEM and MNROM solutions at each local point \( Y \) in the microstructure (i.e. each finite element in RUC), the effective strain error can be defined as

\[
\mathcal{E}_e = \frac{\| e_{\text{ROM}} \|_F - \| e_{\text{FEM}} \|_F}{\| e_{\text{FEM}} \|_F} \times 100\%.
\] (5.7)

In order to understand representative loading conditions, a query point from mode 3 with the homogenization error of \( \mathcal{E}_p = 0.3316\% \), and \( \mathcal{E}_m = 4.9335\% \) is investigated. Looking at Figure 5.9 (also see Table 5.5) and quantifying the percentage, the solution quality of 87.75\% of the query points is better than this selected point while considering all modes. Therefore, this selection is representative of a large range of different loading conditions (i.e. not present in the original dataset).

To get a clear understanding about both the magnitude of field and distributions of error with respect to the local deformation extent within each finite element, two joint probability distributions are good representations. Figure 5.14a shows the joint probability distribution of \( \mathcal{E}_m \) in Equation (5.3) and \( \| F_{\text{FEM}} \|_F \), and Figure 5.14b displays the joint probability distribution of \( \mathcal{E}_e \) in Equation (5.7) and \( \| e_{\text{FEM}} \|_F \). Here I view \( \| e \|_F \) like the effective Almansi strain. Since particles behave as rigid-like, deformation in particles is negligible, and I focus on the matrix. The joint probability distribution is constructed from a cloud of points representing individual finite elements within the matrix phase (285,380 elements in the matrix).

Figure 5.14 shows that the most finite elements in the matrix carry a mean localization error of 5.3786\% for this loading case and that the deformation is concentrated between \( \| F_{\text{FEM}} \|_F = 1.58 \) and \( \| F_{\text{FEM}} \|_F = 1.75 \) (Note that \( \| 1 \|_F = \sqrt{3} = 1.732 \)). Moreover, Figure 5.14a shows that the error \( \mathcal{E}_m \) is concentrated between \( \sim 4 - 7\% \) and decays quickly. Figure 5.14b reveals that the mean effective Almansi strain error calculated at every finite element in the matrix is 13.6321\%. Note that the regions of
Figure 5.14. Joint probability distribution (computed in discrete element-wise sense) of the local fields (in matrix phase). (a) $E_m$ and $\|F^{FEM}\|_F$. (b) $E_e$ and $\|e^{FEM}\|_F$.

Figure 5.15. Visualization of the effective Almansi strain ($\|e\|_F$) (Note that particles are rigid-like and are therefore removed for visualization purpose.). (a) FEM simulation. (b) MNROM analysis.
high strains are concentrated between errors of 5 – 15%. Moreover, 50% of the finite elements have errors below 10% (median error is 9.9678%) and 78% of finite elements in the matrix have errors below 20%. Finally, it can be seen that the largest errors are at small to moderate strains. This is not surprising since the Almansi strain is derived from \( \mathbf{F} \) (see Equation (5.6)) and any small noise near the unloaded state (\( \mathbf{F} = 1 \)) is contributing to this relative error measure. These results show good localization characteristics of the MNROM for both \( \mathbf{F} \) as well as engineering measures, such as the effective Almansi strain (defined as \( \| \mathbf{e} \|_F \) in this work).

Figure 5.15 compares the effective Almansi strain within the matrix for both the FEM (left) and MNROM (right) analysis. As can be seen, the MNROM captures well highly localized strain regions (see top cell surface and cut vertical surface) as well as the overall strain distribution (see the rest of the cell).

Note, the results provided here is based on one-stage 0 of MNROM. No sampling technique is used. However, no additional HPC simulations were performed for the case computed from the MNROM. In the next section, I would like to show how the results improve with less number of snapshots when an effective sampling technique is employed. In contrast to this section, the sequentially adaptive sampling technique is used in conjunction with the 0 based manifold construction. Moreover, as mentioned before, the deformation based manifold construction has added advantage over the deformation gradient based manifold construction as the dimensionality to handle is much lower.

5.2 Sequentially Adaptive Design of MNROM

For this study, I use the same RUC as given in Figure 5.2. For this analysis, again 10% maximum principal stretch in each direction is used. In doing so, the novel physics-guided sampling technique is implemented for the performance enhancement of the ROM. This technique is deterministic and sequential in nature.
To begin with this sequentially adaptive sampling technique, initially (step 0), I take a uniform grid \((3 \times 3 \times 3)\) to create the stretch space, \(\mathcal{T}_\lambda\). This leads to 27 points. Next, to construct the rotational space, \(\mathcal{T}_\varphi\), associated with each stretch point I take 15 randomly generated (uniform distribution) well spaced points. To create these well spaced points, I first generate 10,000 random points inside the rotational space and call this set as \(\mathcal{T}^*_\varphi\). Next, I pick any point from \(\mathcal{T}^*_\varphi\) to get the first point of \(\mathcal{T}_\varphi\). Then, the concept of Hausdorff distance is used and pick the next point from \(\mathcal{T}^*_\varphi\), which produces farthest distance from \(\mathcal{T}_\varphi\). This process is continued for specified times (i.e., 14 here) to generate 15 points to construct the rotational space. Since the parameter space is represented as a product of the stretch and rotational spaces \((\mathcal{X} = \mathcal{T}_\lambda \times \mathcal{T}_\varphi)\), the total number of initial snapshots for step 0 is 405 (see Table 5.8).

Figure 5.16. Isomap dimensionality with step 1 data.

After the initial snapshot set is defined, I employ both regular and kernel Isomap to estimate the dimension of the manifold. Both Isomap representations unfold the
manifold into the 6-dimensional Euclidean space, which is the dimension of the parameter space, \( \mathcal{X} \). Here, the reduced space is constructed by the regular Isomap.

**TABLE 5.8**

DESCRIPTION OF THE MACROSCALE LOADING MODES WITH DATA ENRICHMENT STEPS

<table>
<thead>
<tr>
<th>Loading Case</th>
<th>Description</th>
<th>Step 0</th>
<th>Step 1</th>
<th>Step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode 1</td>
<td>( \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0 )</td>
<td>405</td>
<td>530</td>
<td>730</td>
</tr>
<tr>
<td>Mode 2</td>
<td>( \lambda_1 \leq 0, \lambda_2 \geq 0, \lambda_3 \geq 0 )</td>
<td>405</td>
<td>1133</td>
<td>2770</td>
</tr>
<tr>
<td>Mode 3</td>
<td>( \lambda_1 \leq 0, \lambda_2 \leq 0, \lambda_3 \geq 0 )</td>
<td>405</td>
<td>1060</td>
<td>2471</td>
</tr>
<tr>
<td>Mode 4</td>
<td>( \lambda_1 \leq 0, \lambda_2 \leq 0, \lambda_3 \leq 0 )</td>
<td>405</td>
<td>524</td>
<td>790</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1620</td>
<td>3247</td>
<td>6761</td>
</tr>
</tbody>
</table>

Although number of enrichment steps is not limited, in this study, I only show the results for two steps (step 1 and step 2). Also, one can enrich manifold \( \mathcal{M} \) with any number of points in each step. Here, I double the total number of snapshots in every step to understand the convergence properties of the enrichment (see Table 5.8). In each step, I compute the volume of each manifold corresponding to each mode (See Eq. (4.6)) and proportionally distribute the added points among these four manifolds. After determining the number of extra points in each mode, I first enrich the stretch space \( \mathcal{X}_\lambda \) by exploiting the density analysis as described in Figure 4.1 (red marker). Then, I employ the rotational sensitivity analysis and decide the number of rotational
points for each stretch vector, which is proportional to the volume of the submanifolds, 
\( M^i_\lambda \) (see Eq. (4.7)).

Figure 5.17. Rotational sensitivity analysis with step 0 and step 1 data. (a) Mode 1. (b) Mode 2. (c) Mode 3. (d) Mode 4.

For the first enrichment step (step 1), the density analysis is conducted on the step 0 dataset as described in section 4.1.2 The average density, \( \bar{\eta}_g \) is calculated by
Eq. (4.5), on the entire dataset for a particular $\epsilon$-0. In this work, $\epsilon = (1/4) d_G$ is used for all modes. Here, $d_G$ is the diameter of the complete graph corresponding to the entire manifold, $\mathcal{M}$. Next, I consider each closed domain, $U_{\lambda}^i$, in stretch space (see Fig. 4.1). I construct $i^{th}$ stretch space grid by tetrahedral elements covering $U_{\lambda}^i$. Then $\pi^i_g$ is computed by taking the data associated to the $i^{th}$ element (see the gray triangle in the stretch space in Fig. 4.1) with all associated rotations. If $\pi^i_g$ is less than $\max(\pi_g, k)$, extra stretch point is added in the center of that element (see added red data point in the stretch space in Fig. 4.1). This ensures that after the enrichment step, the sparse regions in the manifold have at least $k$ neighboring points (in average sense, in $\epsilon$-0). In this work, I have used $k = 21$. This completes the stretch space enrichment for step 1.

Next, a pattern is explored using rotational sensitivity analysis to distribute the rotational points in the newly constructed stretch space. To distribute the rotational points, the submanifolds, $\mathcal{M}_{\lambda}^i$ are separated out. This submanifolds are approximated in terms of the subgraphs, $G_{\lambda}^i$. All subgraphs initially contain 15 rotational points (step 0). Next, I compute the diameter, $d_{\lambda}^i$, for each subgraph which is the longest shortest path (i.e., the longest graph geodesic) in the graph [28]. Although diameters are function of both $r_{\lambda}^i$ and $I_{3}^i$, it is noticeable that the effect of $I_{3}^i$ is not very significant. Therefore, I plot the graph diameters $d_{\lambda}^i$ as a function of $r_{\lambda}^i$ in Figure 5.17 for all four modes. Since the graph diameters are clustered around a line, the diameter is learned as the linear least-square fit. The small scatter of the diameters of the submanifolds is due to the effect of the third invariant, $I_3$, and noise associated with the geodesic computations. Using this relationship (see Fig. 5.17), the diameter can be estimated for any new stretch point, which is used to compute the volume by Eq. (4.6) with $d = 3$. Next I distribute new points among the submanifolds $\mathcal{M}_{\lambda}^i$ by Eq. (4.7). This completes the first enrichment process (step 1).

For the second enrichment step (step 2), I use the current dataset (step 1) with
total 3,247 snapshots (distributed among the modes, see Table 5.8) and repeat the process for further enrichment. As listed in Table 5.8, mode 2 and mode 3 take most of the extra points. On the other hand, mode 1 and mode 4 enrichment is not that significant. Note, mode 2 and mode 3 are mixed modes and more deviatoric in nature, which is consistent with the rotational sensitivity analysis. Also it is indicative that the manifolds corresponding to mode 2 and mode 3 are more nonlinear than mode 1 and mode 4, respectively. To summarize this novel physics guided enrichment strategy, a complete algorithm is provided, which contains the steps discussed above.

Algorithm 1 Adaptive Physics-guided MNROM

1: Start with \( N_0 \) number of uncorrelated snapshots.
2: Compute all-pair geodesic/shortest paths (see Eq. (3.1)) and carry out density distribution analysis (see Eq. (4.5)) to predict the number of stretch points required and their locations. The current number of stretch points is \( N_\lambda \).
3: Carry out rotational sensitivity analysis and learn the diameter of the submanifolds, \( d_i^\lambda \), as a function of the parameter, \( r_i^\lambda \).
4: From rotational sensitivity analysis, determine the diameter of the submanifold corresponding to any stretch point. Next, compute the volume of those submanifolds by Eq. (4.6). Then, determine the number of the rotational points corresponding to each stretch point, given the total number of extra points by following Eq. (4.7).
5: Any further extra snapshot will be added to the database based on the pattern of the current enrichment state.

To understand the quality of the enrichment scheme for two steps, in Fig. 5.18, I present the distribution of the number of points, \( n_{gi} \), within \( \epsilon = 0 \) (density) with all steps for all four modes. In Fig. 5.18, one see that the distributions become smoother with this enrichment scheme. Moreover, the distributions have distinct tail, which indicates smaller data redundancy. Since the sparsity of each neighborhood decreases in the process, in effect, it causes higher possibility to capture the local features/scales...
Figure 5.18. Distribution of $n_{gi}$ for all steps with all modes. The size of the markers indicates the enrichment step. (a) Mode 1. (b) Mode 2. (c) Mode 3. (d) Mode 4.
of the manifold. This is especially important for mixed modes 2 and 3, which manifest more nonlinearity.

In each step, also the percentage of the overlap, $P_O$ is computed for each mode (see Table 5.9). The percentage of overlap estimates the degree of redundancy of the data. Note, the percentage of the overlap remains almost unchanged with the manifold enrichment. This suggests that the adaptive enrichment technique helps to distribute the data on the manifold evenly.

<table>
<thead>
<tr>
<th>TABLE 5.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERCENTAGE OF THE OVERLAP FOR EACH MODE AND IN EACH STEP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$P_O$</th>
<th>Step 0</th>
<th>Step 1</th>
<th>Step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode 1</td>
<td>3.65</td>
<td>3.99</td>
<td>4.49</td>
</tr>
<tr>
<td>Mode 2</td>
<td>2.49</td>
<td>2.66</td>
<td>2.66</td>
</tr>
<tr>
<td>Mode 3</td>
<td>2.52</td>
<td>2.62</td>
<td>2.72</td>
</tr>
<tr>
<td>Mode 4</td>
<td>3.67</td>
<td>3.88</td>
<td>4.64</td>
</tr>
</tbody>
</table>

For this refinement strategy, Isomap is used at every enrichment step. However, as mentioned previously, Isomap algorithm is expensive. To avoid full-Isomap, a greedy algorithm is proposed in Section 3.1.1 where newly added points are projected into the existing reduced space $\mathcal{A}$. To verify the greedy algorithm, I first take the step 2 dataset and construct the reduced space, $\mathcal{A}_2$, by using the regular Isomap. The subscript 2 indicates that the reduced space involves all data in step 2. Similarly,
Figure 5.19. Efficacy of the greedy algorithm: enrichment points in step 2 dataset are projected to the step 0 and step 1 reduced spaces by Eq. (3.6). 
(a) Mode 1. (b) Mode 2. (c) Mode 3. (d) Mode 4.
I have constructed $\mathbf{A}_0$ and $\mathbf{A}_1$ with step 0 and step 1 dataset, respectively. Next, the extra snapshots are projected in step 2 dataset, which are not in step 0 dataset by using the Eq. (3.6) and denote this reduced space as $\mathbf{A}_2(0)$. Same way, $\mathbf{A}_2(1)$ is constructed, where all extra snapshots are being projected using the eigenvalues and eigenvectors computed from the step 1 data. To compare $\mathbf{A}_2$ with $\mathbf{A}_2(0)$ and $\mathbf{A}_2(1)$, I first vectorize these matrices into $\hat{\mathbf{\zeta}}_2$, $\hat{\mathbf{\zeta}}_2(0)$, and $\hat{\mathbf{\zeta}}_2(1)$, respectively. Note, the reduced space, $\mathbf{A}$ is a $6 \times N$ matrix, $d = 6$. However, in continuum sense ($N \to \infty$) these vectors can be viewed as the finite set of points constructed from the scalar field, $\hat{\mathbf{\zeta}}$. To verify the greedy algorithm, I plot the distribution of the scalar field $\hat{\mathbf{\zeta}}_2$, $\hat{\mathbf{\zeta}}_2(0)$, and $\hat{\mathbf{\zeta}}_2(1)$ in Figure 5.19. Figure 5.19 shows that the enrichment points on the manifold have been projected to the reduced space satisfactorily as the $\hat{\mathbf{\zeta}}_2(1)$ (projected) converges to $\hat{\mathbf{\zeta}}_2$ (directly computed). This is also indicative that the step 1 dataset is enriched enough and any future snapshot can be projected to the reduced space, $\mathbf{A}_1$, by avoiding the expensive Isomap computation of all-pair shortest paths and eigenvalue analysis.

For error analysis, I first establish all the maps associated with the MNROM framework for step 0, step 1 and step 2 dataset. Next, I compute the microscopic deformations for 200 query points in each mode, which are well spaced and uncorrelated with the data set. After computing the displacement field for all the query points by MNROM framework, I can determine the deformation gradient locally and compute the volume averaged error by Eq. (5.3) in comparison with the deformation gradient computed from the FEM framework. Next, I plot the mean and the maximum error in Figure 5.20. Note that in step 1, the maximum error has decreased to below 2%. However, from the step 1 to step 2 error decay has saturated. Considering the enrichment result presented in this section, I proceed in further studies with the dataset from step 1 since this manifold is well populated. Next, similar to one-stage design, the verification of each map is provided which associated with MNROM individually.
Figure 5.20. Convergence of adaptive data enrichment scheme in terms of the microscopic deformation gradient field. (a) matrix-mean for modes 1, 2 and 4. (b) matrix-max for modes 1, 2 and 4. (c) matrix-mean for mode 3. (d) matrix-max for mode 3.
for step 1 data set.

5.2.1 Verification of Inverse Map, $f^{-1}$

To check the accuracy of the inverse map, $f^{-1}: \mathcal{A} \mapsto \mathcal{M}$, a standard leave-one-out experiment has been carried out on 10% randomly selected (uniform distribution) points from each loading mode like the one-stage example. For each of these points, I compare inverse map solution with the FEM solution in terms of the volume averaged error for the matrix and particles (see Eq. (5.3)) of the deformation gradient field, $\mathbf{F}$. The distribution of the volume averaged error is presented in Figure 5.21. The particle error is less than 0.11% (see Fig. 5.21(a)) and the matrix error is bounded by 1.8% (see Fig. 5.21(b)). To check the quality of the localized MNROM solution, I show the distribution of $\|\mathbf{F}\|_F$ for the particle in Figure 5.22(b) and for the matrix in Figure 5.22(a). Comparison with the FEM solution yields the highest volume averaged error in the matrix (i.e., $0\mathcal{E}_m = 1.789\%$). It is worth commenting that the RKKS based inverse map can predict the localized field remarkably well.

5.2.2 Verification of Map, $\mathcal{P}$

To establish a link between the macroscopic loading space and the reduced space, the map $\mathcal{P}$ is built using neural network. However, NN is very prone to get trapped to a local minima. To circumvent this issue, here, ensemble NN is used. The NNs are trained 10 times for each mode and the direct average is taken as the output. Also to overcome overfitting, Bayesian regularization technique [39, 46, 5] has been employed. Then, I reconstruct the reduced space as the NN output for the entire dataset. This reconstructed reduced space, $\mathcal{A}$, deviates from the actual reduced space, $\mathcal{A}$. In Figure 5.23 I plot the distribution of the error, which is calculated as the Euclidean distance between the same index vectors $\zeta^i$ and $\bar{\zeta}^i$ from the two sets $\mathcal{A}$ and $\overline{\mathcal{A}}$, respectively. One can observe that the NN error has been reduced drastically
Figure 5.21. Distribution of the volume averaged error (homogenization) given by Eq. (5.3) for randomly selected 10% of points from database (one-out cases). (a) Particle. (b) Matrix.

Figure 5.22. Distribution of $\|F\|_F$ from FEM and MNROM. (a) Over Matrix. (b) Over Particle.
from step 0 to step 1 (see Fig. 5.23). However, error reduction is not noticeable from
step 1 to step 2. This suggests that the step 1 data is enriched enough. The NN
map has been improved significantly over one in [9] with this novel physics-guided
sampling technique. The average error for all modes is around 3 unit with step 1
data, which is $O(10^{-3})$ in $[0, 1]$ range. The physical interpretation of this error is the
drift of the predicted MNROM solution from the FE solution along the manifold. It
is worth mentioning that the effect of this error on the local micro-field is extremely
small. Also I like to comment that NN map has been improved significantly than the
one-stage case. Moreover, it happened with much smaller number of simulations.

5.2.3 Verification of Complete MNROM Framework

After verification of all maps, the MNROM framework is now used to compute
the macroscale (homogenization) as well as the microscale (localization) fields for any
given loading conditions of the unit cell that are bounded by 10% principal stretch
in each direction. Next, the microscale deformation fields for 200 query points are
predicted for each four modes (i.e., 800 total query points).

I compute the microscale deformation gradient and compare with the computed
deformation gradient using FEM. First, I compute the volume averaged error for 200
query points considering all four modes by Eq. (5.3). The volume averaged error
is shown in Figure 5.24(a) and 5.24(b) for the particles and matrix, respectively.
Minimum, maximum, mean and median of the volume averaged error over all query
points for all modes are listed in Table 5.10. Figure 5.24 shows that the error in
the matrix is higher than the particles. In Figure 5.24(b), it can be observed that
the maximum error goes to 1.67%, but only less than 10% of query points have a
volume averaged error above 1% considering all modes. The query points associated
with the higher error are either close to the loading envelope or in the regions where
the manifold data density is still low. Next, the quality of MNROM is analyzed for

79
Figure 5.23. Distribution of the neural network output error compared to reduced space over entire dataset for different enrichment steps. The size of the markers indicates the enrichment step. (a) Mode 1. (b) Mode 2. (c) Mode 3. (d) Mode 4.
the homogenization and the localization for different fields of interest.

**TABLE 5.10**

QUALITY OF THE MNROM FOR ALL MODES COMPUTED FROM Eq. (5.3) FOR 200 QUERY POINTS IN EACH MODE

<table>
<thead>
<tr>
<th>Loading Cases</th>
<th>Min. [%]</th>
<th>Max. [%]</th>
<th>Mean [%]</th>
<th>Median [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>0.0038</td>
<td>0.0284</td>
<td>0.0127</td>
<td>0.0124</td>
</tr>
<tr>
<td>Matrix</td>
<td>0.1080</td>
<td>0.8224</td>
<td>0.4093</td>
<td>0.4013</td>
</tr>
<tr>
<td>Mode 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>0.0069</td>
<td>0.0541</td>
<td>0.0239</td>
<td>0.0214</td>
</tr>
<tr>
<td>Matrix</td>
<td>0.1009</td>
<td>0.8058</td>
<td>0.3866</td>
<td>0.3810</td>
</tr>
<tr>
<td>Mode 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>0.0101</td>
<td>0.1083</td>
<td>0.0330</td>
<td>0.0295</td>
</tr>
<tr>
<td>Matrix</td>
<td>0.1168</td>
<td>1.6605</td>
<td>0.4612</td>
<td>0.3796</td>
</tr>
<tr>
<td>Mode 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle</td>
<td>0.0100</td>
<td>0.0722</td>
<td>0.0344</td>
<td>0.0334</td>
</tr>
<tr>
<td>Matrix</td>
<td>0.1557</td>
<td>1.2821</td>
<td>0.6322</td>
<td>0.6335</td>
</tr>
</tbody>
</table>
5.2.3.1 Homogenization of Material Response

In this section, I investigate the capability of MNROM to predict the homogenized response for 800 query points (200 in each mode) with step 1 dataset. Similar to the one-stage design, first, the homogenized strain energy density is computed, which is the volume average of the local strain energy density. In Figure 5.25, I plot the deviatoric (Fig. 5.25(a)) and the volumetric (Fig. 5.25(b)) part of the strain energy density. Looking at both the deviatoric and volumetric strain energy, one can observe a very high correlation between the FEM and the MNROM predictions. Moreover, I compute the error associated with the homogenized potential energy. Since mode 1 and mode 4 are purely tensile and compressive, and also more volumetric in nature, these are the modes contributing to the moderate to high volumetric strain energy in Figure 5.25(b). On the other hand, mode 2 and mode 3 are more deviatoric and contribute more to the deviatoric energy, which is reflected in Figure 5.25(a). In the entire analysis with 800 query points, 94.12% of the points have less than 10% error which is computed from Eq. (5.4). Only a few points (around 5%) show more than
10% error. This is a large improvement over one-stage design. This indicates the strength of the physics guided manifold construction.

![Graphs showing strain energy density function MNROM versus FEM.](image)

**Figure 5.25.** Macroscale strain energy density function MNROM versus FEM. (a) The deviatoric potential, $^{0}W_{C}(\hat{C})$. (b) The volumetric potential, $^{0}W_{J}(J)$.

Furthermore, I like to learn the functional form of the homogenized potential energy, $^{0}W(C)$ as before. It is done separately for the volumetric part, $^{0}W_{J}(J)$, and the deviatoric part, $^{0}W_{C}(\hat{C})$. One could use machine learning technique for this calibration, however, here I simply take the functional form given in Eq. (5.1) and learn it as a least square fit like one-stage design. Figure 5.26 depicts model calibration using the volumetric (Fig. 5.26(b)) and deviatoric (Fig. 5.26(a)) functions for both the FEM and MNROM data. The FEM and MNROM parameterizations show an excellent agreement for both the volumetric and the deviatoric potentials. Finally, the homogenized material constants (i.e., the coefficients of least-square fit) for both FEM and MNROM models are presented in Table 5.11.
Figure 5.26. Least-square fit of the macroscale strain energy density function MNROM versus FEM. (a) The deviatoric potential, $W_C(\hat{C})$. (b) The volumetric potential, $W_J(J)$.

Figure 5.27. Macroscale first Piola stress tensor, $P$, MNROM versus FEM. (a) Frobenius norm of $P$ against Frobenius norm of $F$. (b) Distribution of the homogenization error given by Eq. (5.5).
TABLE 5.11

CALIBRATED HOMOGENIZED MATERIAL PROPERTIES

<table>
<thead>
<tr>
<th>Source</th>
<th>$\kappa$ [MPa]</th>
<th>$\mu_{10}$ [MPa]</th>
<th>$\mu_{01}$ [MPa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNROM</td>
<td>96.37</td>
<td>11.01</td>
<td>11.01</td>
</tr>
<tr>
<td>FEM</td>
<td>96.36</td>
<td>11.74</td>
<td>11.74</td>
</tr>
</tbody>
</table>

In Figure 5.27(a), I plot the Frobenius norm of the macroscopic first Piola stress tensor, $^0\mathbf{P} = \frac{1}{|\Theta_0|} \int_{\Theta_0}^1 \mathbf{P} \, d\Theta$, with the corresponding Frobenius norm of macroscopic deformation gradient, $^0\mathbf{F}$, for all four modes. It can be seen that the points corresponding to $\|^0\mathbf{P}\|$ computed by MNROM are very close to the points which are calculated using the FEM. Finally, the error plot is shown in Figure 5.27(b) where the maximum error is $E_{^0\mathbf{P}} = 0.38\%$. This is again a large improvement over one-stage design.

5.2.3.2 Localization of Material Response

In this section, the ability of the adaptive data enrichment scheme towards the improvement of the MNROM in terms of the local (element-wise) engineering fields derived from the microscopic deformation is explored. In this study, I consider the Almansi strain, $e$ (see Eq. (5.6)) to compare FEM and MNROM solutions by quantifying the strain error (Equation (5.7)) at each local (microscopic) point $\mathbf{Y}$ in the microstructure (i.e. each finite element in RUC).

Since mode 3 is relatively more complex in nature, I chose a query point from mode 3 to judge the quality of the localization process, which corresponds to a very high strain. Then, I plot two joint probability distributions: Figure 5.28(a) shows the joint
probability distribution of $\mathcal{E}_m$ and $\|\mathbf{F}^{FEM}\|_\mathcal{F}$, and Figure 5.28(b) displays the joint probability distribution of $\mathcal{E}_e$ and $\|\mathbf{e}^{FEM}\|_\mathcal{F}$. Here I consider $\|\mathbf{e}\|_\mathcal{F}$ as the measure of the microscopic effective Almansi strain. Since the deformation in particles is negligible, I focus on the matrix. The joint probability distribution is constructed from a cloud of points representing individual finite elements within the matrix phase (285,380 elements in the matrix).

Figure 5.28. Joint probability distribution (computed in discrete element-wise sense) of the local fields (in matrix phase). (a) $\mathcal{E}_m$ and $\|\mathbf{F}^{FEM}\|_\mathcal{F}$. (b) $\mathcal{E}_e$ and $\|\mathbf{e}^{FEM}\|_\mathcal{F}$.

Figure 5.28(a) shows that most finite elements in the matrix carry a mean localization error of 0.99% for this loading case. Although the range of deformations is $[1.55, 2.15]$, the deformation is concentrated between $\|\mathbf{F}^{FEM}\|_\mathcal{F} = 1.5$ and $\|\mathbf{F}^{FEM}\|_\mathcal{F} = 1.65$, which characterizes very high compressive strain (Note that $\|\mathbf{1}\|_\mathcal{F} = \sqrt{3} = 1.732$). However, this particular loading case also induces very high tensile strains locally. Figure 5.28(a) shows that the error $\mathcal{E}_m$ is concentrated between
Figure 5.29. Visualization of the microscopic effective Almansi strain (\(\|e\|_F\)) for \(\eta = \{0, \lambda_1 = 0.9284, \lambda_2 = 0.9058, \lambda_3 = 1.0839, \varphi^1 = 2.4217, \varphi^2 = 1.2802, \varphi^3 = 0.3165\}\) (Note, particles are rigid-like and hence are removed for the visualization purpose). (a) FEM simulation. (b) MNROM analysis.

\(~ 0.5 – 1.5\%\). Figure 5.28(b) reveals that the mean microscopic effective Almansi strain error calculated at every finite element in the matrix is 4.53\%. Note that the regions of high strains are concentrated between errors of 3 – 5\%. Moreover, 70.46\% of the finite elements have errors below 5\% (median error is 2.84\%), and all of the finite elements in the matrix have errors below 15\%. These results show an excellent localization prediction of the MNROM. Moreover, all higher error points are associated with the negligible effective strain (see strain region \(\|e^{FEM}\|_F = 0.03\) in Fig. 5.28(b)). This is a substantial improvement over the one-stage design of MNROM.

Finally, Figure 5.29 provides a visual comparison of the microscopic effective Almansi strain in the matrix between the FEM (left: Fig. 5.29(b)) and MNROM (right: Fig. 5.29(a)) analysis. One can observe that the MNROM is able to capture the overall strain distribution extremely well including the localized features.
CHAPTER 6

CONCLUSION

In this work, a novel manifold-based reduced order model (MNROM) for multiscale modeling of nonlinear hyper-elastic materials in finite strain setting has been proposed and verified. This reduced order model is built based on the regression framework. However, interpolation in high-dimensional space is not straightforward. The main challenge regarding the interpolation is the curse of dimensionality. Due to the curse of dimensionality, the neighborhood structure gets distorted. Furthermore, any sophisticated technique, like deep-learning poses extra challenges when learning involves very high dimension. To circumvent these issues, this novel technique takes the help of dimension reduction, which compress the data retaining the useful informations. This provides the scope to apply NN like learning technique. In this work, the Isomap algorithm is used, which provides dimensionality reduction from the high-dimensional data stemming from large parallel computational homogenization simulations. Isomap, is a manifold learning method and idealizes the HD data as a Riemannian manifold. This method approximately unfolds the HD manifold to its intrinsic-dimensional Euclidean reduced space. Since Isomap is isometric, it is bijective and consequently invertible. The inverse map has been constructed by exploiting the concept of Reproducing Kernel Hilbert space (RKHS). RKHS demands a positive definite (PD) kernel. However, in distance based learning setting, generating a PD kernel is difficult. In this work, instead of RKHS, more generalized concept, that is the concept of RKKS is used. Conditionally positive definite kernel supports the concept of RKKS. The map between the set of macroscopic loading parameters
and the reduced space is accomplished by the Neural Network. Once constructed, the manifold-based reduced order model provides both homogenization and localization of the multiscale solution for complex three-dimensional material domains. Thus, solution of the large microscale boundary value problem to arbitrary multiscale loading conditions is approximated quickly without the need for additional large parallel finite element simulations.

Since MNROM is a data-driven model, all the maps associated with MNROM are noise sensitive. Fortunately, in this work, data comes from the computer simulation, which possesses low noise. However, the noise generates during the application of different maps. When Isomap is used, which computes the all-pair geodesics in an approximate sense. If the data density is low, the approximate geodesics are erroneous and provides noise for NN and inverse map. An obvious way to reduce this noise is to increase the number of simulations/snapshots. Unfortunately, this is not a feasible solution. First of all, simulations are expensive and most importantly, complexity of ROM increases with the number of simulations. Thus, ROM needs to rely on a very few number of simulations or snapshots. To improve the efficacy of the model, a good sampling technique is required. Although there are several sampling methods available, most of these techniques do not address the inherent challenges. To keep low bias, these methods seek a uniform distribution in the input space. Unfortunately, when dealing with HD space, uniformity in the input space does not guaranty uniformity in the HD space. Even in low or moderate dimensional output space, if there are pockets of sensitive regions, the uniformity is not a good choice. Addressing all these issues, in this work, a novel physics guided sampling technique has been proposed. This sequentially adaptive sampling technique improves the MNROM many folds.

The proposed model reduction technique has been implemented and verified for a particulate RUC. In the first example, deformation gradient field is used to construct the manifold. Although no sampling technique is used, the result produced is promis-
ing. In the second numerical example, the novel physics guided sampling technique is used and that improved the proposed model multiple times. Moreover, the proposed manifold-based reduced order model can be used for material classification and lends itself to predictive scientific studies as well as Virtual Materials Testing.

6.1 One-stage MNROM

In order to build the manifold, I use the large parallel three-dimensional finite strains solver, \textit{PGFem3D}, which performs several parallel simulations using a statistically representative unit cell. This statistically representative unit cell is generated by a packing algorithm. The multiscale loading conditions are simulated in terms of the macroscopic principal stretches as well as the orthogonal principal directions. Therefore, arbitrary multiscale loading conditions can be applied. In order to obtain a uniform discretization for rotational parameters, the HEALPix grid has been employed. The manifold-based reduced order model is verified step by step using traditional machine-learning procedures. Next, I have performed homogenization and localization of the multiscale solution for data points not in the digital database. Rigorous assessment of both homogenization and localization errors is performed. The manifold-based reduced order model discovers the homogenized material response as well as provides microscopic fields of interest. In this work, the microscopic deformation gradient is used as the building block of the manifold, since the other physical fields (i.e. Almansi strain, Piola-Kirchhoff stress) can be easily derived.

6.2 MNROM with Sequentially Adaptive Sampling

In this part, I have designed and implemented a novel physics-guided adaptive sequential sampling technique in the context of manifold-based reduced order model, MNROM, for multiscale modeling of nonlinear hyper-elastic materials. Although MNROM is a promising data-driven approach, it requires an efficient data sampling
technique as the method involves an extremely high-dimensional output space. As mentioned earlier, there are several statistical techniques available, which are intended to distribute data points effectively, but fails to perform as dimension of the problem increases as in case of MNROM. This novel physics-driven sampling strategy mitigates the issue related to curse of dimensionality. This is a stepwise data enrichment method, intended to minimize the sparse regions of the HD manifolds and in effect maximizes information in each step. The method couples input and output space and explores the pattern in the output space, which is directly guided by the inherent physics in the input space, the macroscale loading condition. The pattern is learned in each step based on the current data set and extra simulations are decided and added to the data base by the help of this pattern learned. In this work, the multiscale loading conditions are simulated in terms of the macroscopic principal stretches as well as the orthogonal principal directions and the HD manifold has been constructed from microscopic displacement fields. Here, the extent of the effectiveness of the proposed sampling technique by meticulous numerical experimentations has been demonstrated. In the numerical example, it is observed that the data distribution becomes smooth with stepwise data enrichment and accordingly, all the maps improved significantly. This physics-driven deterministic sampling strategy also appreciably reduced the number of required simulations by eliminating redundant data points. Hence, the computational complexity of the ROM is reduced enormously. Also with the proper enrichment, the additional simulation can be added to the dataset seamlessly. Moreover, it can be observed that even localization process produced small error with a few number of simulations. This sampling strategy is not only helpful for CH, it can also be applied to other physical problems by identifying the influential parameters.
6.3 Future Directions

*CH Acceleration:* Although MNROM itself provides extremely good solutions both for homogenization and localization, still one can rely on complete CH process. As described, CH solves two nested BVPs in macroscale and microscale. The deformation gradient at the macroscale material point is used as an input for the microscale RUC, where material properties of constituents are known. The microscale BVP performs the localization process and returns the volume averaged macroscale energy density, stress and tangent stiffness matrix for that material point at the macroscale. Since both macroscale and microscale BVPs are highly coupled, a nonlinear nested solver is preferred as a solution method for CH. Typically, a nonlinear solver requires an initial guess to start the solution process. Predicting an initial guess is often difficult especially when the number of degrees of freedom is large. Therefore, it is important to have a good initial guess. An arbitrary initial guess can take many iterations to converge to the microscale solution. On the other hand, MNROM can predict an approximate solution with high fidelity which is close to the microscale localization solution. It is reasonable to consider that the MNROM solution belongs to the tangent space of the exact microscale solution. Moreover, the MNROM framework avoids expensive numerical iteration. Therefore, it is expected that the nonlinear solver with MNROM as an initial guess will converge fast or with large speedup to the traditional CH implementation.

*Dissipative Process:* For inelastic deformations, additional evolution equations that describe the state variables of plastic flow and/or damage, for example, are solved simultaneously. In the context of the MNROM framework, the solution manifold evolves with time, and the evolution of the solution on the manifold is highly path-dependent. Furthermore, an inelastic process dissipates the energy differently for unique loading paths. Due to this inherent path dependent challenge, the deformation cannot be captured by a single manifold as in hyper-elastic case. Thus, a direct
approach tends to be highly expensive. Therefore, an innovative way to circumvent this challenge will be incrementally projecting the hyper-elastic behavior to the actual solution in the predictor-corrector fashion, where the MNROM will be built for hyper-elastic material as a surrogate model.

Uncertainty Quantification: So far, MNROM has been developed and verified considering deterministic processes. In any physical scenario, it is beyond expectation that material properties of the constituents will remain uniform in RUC. The source of uncertainty can be the morphology of the microstructure or the material properties of the constituents. For example, particle stiffness can be random. Since the material properties of the constituents are spatially distributed within the microstructure, the input uncertainty lies in high-dimensional manifold. Handling high-dimensional input uncertainty is really cumbersome in terms of the computation. To avoid this issue, the high-dimensional input space can be reduced to its low-dimensional representation using dimension-reduction technique. Since all other mechanical response variables (i.e. stress, strain) can be derived from the displacement field, the microscale displacement field will be the primary response field. Therefore, MNROM framework can be utilized as a surrogate for the uncertainty quantification.


