



**HAL**  
open science

# Low-rank and sparse approximations for contact mechanics

Kiran Sagar Kollepara

► **To cite this version:**

Kiran Sagar Kollepara. Low-rank and sparse approximations for contact mechanics. Mechanical engineering [physics.class-ph]. École centrale de Nantes, 2022. English. ⟨NNT : 2022ECDN0030⟩. ⟨tel-03901111⟩

**HAL Id: tel-03901111**

**<https://theses.hal.science/tel-03901111v1>**

Submitted on 15 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# THÈSE DE DOCTORAT DE

L'ÉCOLE CENTRALE DE NANTES

ÉCOLE DOCTORALE N° 602

*Sciences pour l'Ingénieur*

Spécialité : *Mécanique des Solides, des Matériaux, des Structures et des Surfaces*

Par

**Kiran Sagar KOLLEPARA**

**Low-rank and sparse approximations for contact mechanics**

Thèse présentée et soutenue à l'Ecole Centrale de Nantes le 5 juillet 2022

Unité de recherche : UMR 6183, Institut de recherche en Génie Civil et Mécanique (GeM)

## Rapporteurs avant soutenance :

Virginie EHRLACHER  
David RYCKELYNCK

Professeure, Ecole des Ponts ParisTech, Marne la Vallée  
Professeur, Mines Paris PSL University

## Composition du Jury :

Président : Cyrille ALLERY  
Examineur : Yves LE GUENNEC  
Dir. de thèse : Luisa ROCHA DA SILVA  
Co-dir. de thèse : José Vicente AGUADO

Professeur des universités, Université de La Rochelle  
Docteur, IRT Jules Verne, Bouguenais  
Professeure, École Centrale de Nantes  
Maître de conférences, École Centrale de Nantes



ये घना निबिडाः अवयवसन्निवेशाः तैः विशिष्टेषु  
स्पर्शवत्सु द्रव्येषु वर्तमानः स्थितिस्थापकः स्वाश्रयमन्यथा  
कथमवनामितं यथावत् स्थापयति पुर्ववदृजुः करोति ।

Elasticity is a property in the bodies having densely populated elemental parts; when sheared by deforming forces, it helps the body recover its original state.

---

Sridhara Acharya (870 CE - 930 CE)

...dass zwei elastische isotrope Körper sich in einem sehr kleinen Theil ihrer Oberfläche benühren, und durch disen Theil einen endlichen Druck der eine auf den andern ausüben ...Die Fragen, deren Beantwortung uns naturgemäss zunächst obliegt, sind ...die Frage nach der Form und absoluten Grösse der Druckfigur.

... the case of two elastic isotropic bodies which touch each other over a very small part of their surface and exert upon each other a finite pressure, distributed over the common area of contact. ... The questions that we first seek answers to are ... What is the form and the absolute magnitude of the curve of pressure ?

---

Heinrich Hertz (1857-1894)



# Acknowledgement

The journey of this thesis began in about three and a half years ago. During this journey, many people who have supported me and my work, directly or indirectly.

Let me start by thanking all my advisors, José Aguado and Luisa Silva at ECN and Yves Le Guennec and Stephanie Colliou at IRT Jules Verne. I am fortunate to have José as an advisor during my early stages of my research career, as his mentorship has given me the confidence necessary to be a researcher. Technical discussions with him were certainly an enjoyable and enriching experience. Complementing the mentorship of Jose, was Luisa and Yves who shared invaluable knowledge and ideas during the course of the thesis. Last but not the least, I am thankful to each one of them for their support during the peak of the pandemic.

Certainly, I would not have been in a position to begin my research career without the support of my family. Since my high school days, my parents have always been supportive of my career choices at every stage, while also bearing the pain of me being far away from home. I would like to thank my family for their all love and support.

Life in a far away country would not have been easy without good company. For that, I would like to thank my colleagues and friends in the famous T building of ECN, who were a great support group during the last couple of years. In particular, I will never forget the support that received from a special few, namely, Alexia, Manisha, Paris and Ramiro (in no particular order). Moreover, my thanks to a lovely couple, Cécile and Dominique, who made me feel city of Nantes as a second home while also helping me with my transition to a semi-francophone.



# Contents

---

List of Figures	v
List of Algorithms	vii
List of Tables	ix
Nomenclature	xi
Abstract	1
Introduction	3
<b>1 State of the Art</b>	<b>5</b>
1.1 Overview of Contact Mechanics	5
1.1.1 Parametrized contact mechanics problems	6
1.1.2 Contact Formulations	7
1.1.2.1 Lagrange Multiplier Method	8
1.1.2.2 Penalty Method	8
1.1.2.3 Augmented Lagrange Multiplier Method	9
1.1.3 Discrete Formulations	10
1.1.3.1 Contact Detection and Pairing	10
1.1.3.2 Discretization using Finite Elements	11
1.1.4 Resolution schemes	12
1.1.4.1 Fixed point method	13
1.1.4.2 Newton-Raphson	14
1.2 Reduced Order Modelling for parametric problems	15
1.2.1 Low-rank methods	17
1.2.1.1 Proper Orthogonal Decomposition	18
1.2.1.2 Reduced Basis Methods	19
1.2.1.3 Computation of non-linear terms	19
1.3 ROMs for parametrized contact mechanics problems	20

1.3.1	Low-rank approach for contact problems . . . . .	22
1.4	Scope of the thesis . . . . .	25
<b>2</b>	<b>Limitations of low-rank approach to contact problems</b>	<b>27</b>
2.1	Low-rank contact problem using active set method . . . . .	28
2.2	Illustrative case . . . . .	29
2.3	Metrics for low-rank methods . . . . .	32
2.4	Numerical Examples . . . . .	34
2.4.1	Hertz problem . . . . .	34
2.4.2	Ironing problem . . . . .	36
2.5	Conclusions and Perspectives . . . . .	40
<b>3</b>	<b>Dictionary approach in contact mechanics problems</b>	<b>43</b>
3.1	Motivation . . . . .	43
3.2	Sparse methods with over-complete dictionaries . . . . .	44
3.2.1	Sparse regression in data approximation problems . . . . .	45
3.2.2	Application of sparse methods to unconstrained resolution problems	48
3.2.2.1	Diffusion-reaction problem . . . . .	49
3.3	Dictionary methods for contact problems . . . . .	50
3.3.1	A greedy active-set method for dual dictionary element selection .	52
3.3.1.1	Application to the Hertz problem . . . . .	54
3.3.1.2	Application to the Ironing problem . . . . .	57
3.3.2	A non-penetrating convex hull approach for monolithic dictionaries	63
3.3.2.1	Illustrative example: elastic rope-obstacle problem . . . . .	66
3.3.2.2	Demonstration on over-complete dictionaries of the Hertz	
problem . . . . .		69
3.3.2.3	The non-convex case of the Ironing problem . . . . .	69
3.4	Discussions . . . . .	71
3.5	Perspectives . . . . .	73
3.5.1	Efficient construction of constraints . . . . .	73
3.5.2	Application of sketching methods to dictionary-based approximation	73
3.5.3	Smart snapshot selection for Convex hull explorations . . . . .	74
<b>4</b>	<b>Exploring the contact pressure manifold</b>	<b>75</b>
4.1	Linear Subspaces vs. Low-dimensional Manifolds . . . . .	75
4.1.1	Contact pressure as a non-linear manifold . . . . .	77
4.2	Dimension warping . . . . .	78
4.2.1	Dynamic Time Warping . . . . .	78
4.2.2	Interpolating in warped space . . . . .	81
4.3	DTW-based enrichments for contact mechanics . . . . .	85
4.3.1	Numerical results . . . . .	89
4.4	Conclusions . . . . .	90
4.5	Persepctives . . . . .	90
4.5.1	The 3D case . . . . .	90

---

4.5.2	Efficient evaluation of non-linear terms . . . . .	91
4.5.3	Non-linear interpolations of primal dictionary . . . . .	92
<b>Conclusions and Perspectives</b>		<b>93</b>
<b>A Spurious penetrations</b>		<b>97</b>
<b>Bibliography</b>		<b>99</b>



# List of Figures

---

I	Decision-making process using numerical simulations. . . . .	4
1.1	Kinematic description of mechanical problems with possibility of contact . . . . .	6
1.2	Schematic of node-to-node contact pairing for conforming surface meshes . . . . .	10
1.3	Schematics of node-to-segment methods for finding contact pairs . . . . .	11
2.1	Sample snapshots for elastic rope-obstacle problem . . . . .	30
2.2	Primal reconstruction ( $\mathcal{H}^1$ ) errors for points in and outside the training set for the rope-obstacle problem. . . . .	31
2.3	Decay of singular values for the training set of the elastic rope-obstacle problem. . . . .	31
2.4	Sparsity pattern of dual snapshots for the rope-obstacle problem. . . . .	32
2.5	Hertz problem: two half cylinders loaded against each other and the resulting contact pressure snapshots. . . . .	34
2.6	Primal reconstruction ( $\mathcal{H}^1$ ) errors for the Hertz problem. . . . .	36
2.7	Sparsity pattern of dual snapshots for Hertz problem . . . . .	36
2.8	Compactness of reduced bases for Hertz problem . . . . .	36
2.9	Validation metrics for Hertz problem. . . . .	37
2.10	Ironing problem . . . . .	38
2.11	Sample snapshots of contact pressure for ironing problem . . . . .	38
2.12	A typical sparsity pattern of dual snapshots for ironing problem . . . . .	39
2.13	Compactness for reduced bases of ironing problem. . . . .	39
2.14	Projection error of nested level $n + 1$ snapshots on dual RB of level $n$ . . . . .	40
2.15	Validation metrics for the ironing problem. . . . .	40
3.1	Illustration of low-rank vs. dictionary-based approximation . . . . .	45
3.2	Two-dimensional domain for the transport problem (3.4) . . . . .	50
3.3	Snapshots of the transport problem (3.4) for randomly sampled parametric values . . . . .	51
3.4	Reconstruction errors of the transport problem using dictionary-based approximations. . . . .	52
3.5	A spy pattern of dual dictionary with 30 elements for the Hertz problem. . . . .	56

3.6	Greedy active-set reconstruction of contact pressure for certain parametric instances of the Hertz problem . . . . .	56
3.7	Reconstruction errors, primal rank, number of iterations and computation time vs. training set size for the Hertz problem . . . . .	58
3.8	Detailed primal and dual reconstruction errors for the Hertz problem . . . . .	59
3.9	Sparsity of dual reduced dofs for the validation set of Hertz problem . . . . .	60
3.10	Greedy active-set reconstruction of contact pressure for certain parametric instances of the ironing problem. . . . .	61
3.11	Sparsity of dual reduced dofs selected by greedy active-set method for the ironing problem . . . . .	61
3.12	Reconstruction errors, primal rank, number of iterations, and computation time vs. training set size for the Ironing problem . . . . .	62
3.13	Illustration of non-penetrating property of convex hull . . . . .	64
3.14	Illustration of the convex subset hypothesis test with a point mass moving in a convex feasible region. . . . .	67
3.15	Testing the convex subset hypothesis for the elastic rope-obstacle problem . . . . .	68
3.16	Convex hull reconstruction errors for elastic rope-obstacle problem. . . . .	68
3.17	Testing the convex subset hypothesis for the Hertz problem . . . . .	69
3.18	Convex hull reconstruction errors for the Hertz problem. . . . .	70
3.19	Mean number of iterations and computation times for Hertz problem using convex hull exploration. . . . .	71
3.20	Demonstration of the non-convex nature of the ironing problem. . . . .	72
4.1	Visualization of contact pressure non-linearity in ironing problem . . . . .	77
4.2	DTW illustration: Vectors $\lambda_1$ and $\lambda_2$ and their DTW alignments . . . . .	80
4.3	DTW illustration: $\tilde{\lambda}$ and $\tilde{x}$ vs. warped coordinate $\zeta$ . . . . .	80
4.4	Alignments and warped snapshots computed by DTW between two contact pressure snapshots of ironing problem . . . . .	81
4.5	Interpolation along DTW alignments and in warped space . . . . .	82
4.6	Interpolations of the illustrative example for $\alpha = 0.5$ in the spatial indices (corresponding to $x$ ) and the warped space $\zeta$ . . . . .	84
4.7	Linear and DTW-based interpolations of contact pressure for the elastic rope-obstacle problem . . . . .	84
4.8	Linear and DTW-based interpolations of contact pressure for the Hertz problem . . . . .	85
4.9	Interpolation of contact pressure snapshots using DTW for the ironing problem . . . . .	86
4.10	Comparison of the compactness for original and warped snapshots . . . . .	87
4.11	Projection error of nested level $n + 1$ snapshots on dictionary of level $n$ . . . . .	88
4.12	Comparison of reconstruction error for the Hertz and ironing problem using the original dual dictionary and DTW-adapted dual dictionary . . . . .	91
A.1	Reconstruction instance with spurious snapshot selection by the greedy-active set algorithm if non-penetration condition is applied “hardly” . . . . .	97

# List of Algorithms

---

2.1	Online phase	29
3.1	FOCUSS	48
3.2	Greedy active-set algorithm	55
4.1	dtw_interp	83
4.2	adapt_dictionary	87
4.3	DTW powered greedy active-set method	89



# List of Tables

---

2.1	Truncation tolerances and ranks for both bases for the rope-obstacle problem. . . . .	30
2.2	Truncation tolerances and ranks for both bases for Hertz problem. . . . .	35
3.1	Characteristics of the two dictionary-based approximation methods for contact mechanics problems . . . . .	53
3.2	Observed values of various compliances for the elastic rope-obstacle problem with parameter $\gamma_1$ . . . . .	68
4.1	Input variables values and the observed DTW alignments for the illustrative example . . . . .	79



# Nomenclature

---

$\mathbf{u}, \mathbf{v}$	Continuous/Discretized Displacement
$\lambda, \eta, \boldsymbol{\lambda}, \boldsymbol{\eta}$	Continuous/Discretized Lagrange Multiplier (Contact pressure)
$\mathcal{K}$	Feasible region of an inequality constrained problem
$\mathcal{L}$	Lagrangian functional
$\mathcal{V}, \mathcal{W}_+$	Displacement solution space and Contact pressure solution cone
$a, f$	Bilinear and Linear forms of energy functional
$b, d$	Bilinear and Linear forms of distance function in weak form
$k, g$	Distance functions
$\mathcal{P}$	Parametric space
$\mu$	Parameter
$\mathbf{N}, \mathbf{M}$	Displacement and contact pressure shape functions
$\mathbf{r}(\cdot)$	Residual of a discretized problem
$\langle \cdot \rangle^h$	Discrete version of the entity $\langle \cdot \rangle$
$\mathbf{C}, \mathbf{g}$	Discretized operators of bilinear and linear terms $k$ and $g$
$\mathbf{K}, \mathbf{f}$	Discretized operators of bilinear and linear terms $a$ and $f$
$\mathbf{K}_{\text{mono}}$	Operator to generate the residual with monolithic dictionary
$N_u^h, N_\lambda^h$	Number of FE dofs for $\mathbf{u}$ and $\lambda$
$(\mathbf{x}, \bar{\mathbf{x}})$	Contact pairs
$\mathcal{A}, \mathcal{I}$	Active sets

---

$\mathcal{A}^c, \mathcal{I}^c$	Complements of active sets
$\Gamma$	Potential contact surface
$\Gamma_c, \Gamma_{/c}$	Surface in contact, and out of contact (gap)
$\mathbf{d}, \tilde{\mathbf{d}}$	Element in a dictionary, Leave-one-out candidate in a dictionary
$\langle \cdot \rangle^r$	Reduced space of $\langle \cdot \rangle$
$\mathbf{B}$	An orthogonal basis of a given snapshot matrix or its randomized combination
$\mathbf{D}$	Dictionary, Monolithic Dictionary
$\mathbf{R}$	Matrix containing random entries between $[0, 1]$
$\mathbf{S}_{\text{tr}}$	Training set snapshots
$N_u^r, N_\lambda^r$	Number of reduced dofs for $\mathbf{u}$ and $\lambda$
$\Pi_\circ$	Projection operator on the subspace/subcone of basis $\circ$
$\mathbf{D}_p, \mathbf{D}_d$	Primal and Dual Dictionaries
$\Phi, \Theta$	Primal and Dual reduced bases
$\phi, \theta$	Primal and Dual reduced bases vectors/functions
$\widehat{\langle \cdot \rangle}$	Reduced operators and dofs
$\varepsilon_{\text{CHLS}}$	Convex hull least square error for the test of convex subset hypothesis
$\varepsilon_{\text{CrPen}}$	Cross penetration error in convexity test of the feasible region
$C(m)$	Compactness of rank- $m$
$G(m)$	Generalization ability of rank- $m$ basis
$H(n)$	Nested error between $n$ -th and $n + 1$ -th level nested training set
$S(m)$	Specificity of rank- $m$ basis
$\alpha, \gamma$	Coefficient vector multiplying with columns of a matrix/basis/dictionary
$\mathbf{x}$	Vector/Signal to be approximated
$\alpha$	Parameter for interpolating in high-dimensional space
$\circ_{\text{int}}$	Quantity $\circ$ interpolated in warped space
$\mathbf{D}_p^{\text{adapt}}$	Dictionary adapted/enriched using non-linear interpolations

---

$\varphi$	Non-linear transformation to a higher-dimensional warped space/domain
$\zeta$	Co-ordinate in the warped space
$\widetilde{\langle \cdot \rangle}$	Quantity $\langle \cdot \rangle$ in warped space
$I_k$	DTW alignments computed for a vector $\lambda_k$
$\mathbf{0}, \mathbf{1}$	Matrix or Vector consisting of zeros or ones of appropriate size
$\circ[:, S], \circ[S]$	Slicing of a vector/matrix $\circ$ using indices in set $S$ (Python style)
$\delta$	Truncation tolerance for an orthogonal basis
$\varepsilon$	Machine precision, Tolerance on residual
$(z)^-$	$z$ if $z < 0$ , 0 otherwise
$(z)^{+, \tau}$	$z$ if $z > \tau$ , 0 otherwise
$\langle \cdot \rangle^\dagger$	Moore-Penrose pseudo-inverse



# Abstract

---

Non-conformance decision-making processes in high-precision manufacturing of engineering structures are often delayed due to numerical simulations that are needed for analyzing the defective parts and assemblies. Various engineering assemblies often involve interfaces between parts that can only be simulated using the modeling of contact. Therefore, efficient parametric reduced order models (ROMs) are necessary for performing contact mechanics simulations in near real-time scenarios.

Typical strategies for reducing the computational cost of contact mechanics models use low-rank approximations. The underlying hypothesis is the existence of a low-dimensional subspace for the displacement field and a low-dimensional subcone for the contact pressure, as a result of non-negativity constraints. However, the contact pressure exhibits a local nature, as the position of contact can vary with parameters like loading or geometry. In this thesis, the adequacy of low-rank approximations for contact mechanics is investigated and alternative routes based on sparse regression techniques are explored.

It is shown that the local nature leads to loss of linear separability of contact pressure, thereby limiting the reconstruction accuracy of low-rank methods. The applicability of the low-rank assumption to contact pressure is analyzed using three different criteria, namely compactness, generalization and specificity.

Subsequently, the use of over-complete dictionaries containing a large number of snapshots to mitigate the inseparability issues is investigated. Two strategies to solve the dictionary-based approximation problem are devised: one based on a greedy active-set method where the elements from the contact pressure dictionary are selected greedily and another approach based on convex hull approximations eliminating the need to explicitly enforce non-penetration constraints in convex problems.

Lastly, Dynamic Time Warping (DTW) is studied as a possible non-linear interpolation method that permits the exploration of the non-linear manifold. This allows the synthesis of snapshots not computed in the training set with low complexity, thereby reducing the burden of creating over-complete dictionaries in the offline phase.



# Introduction

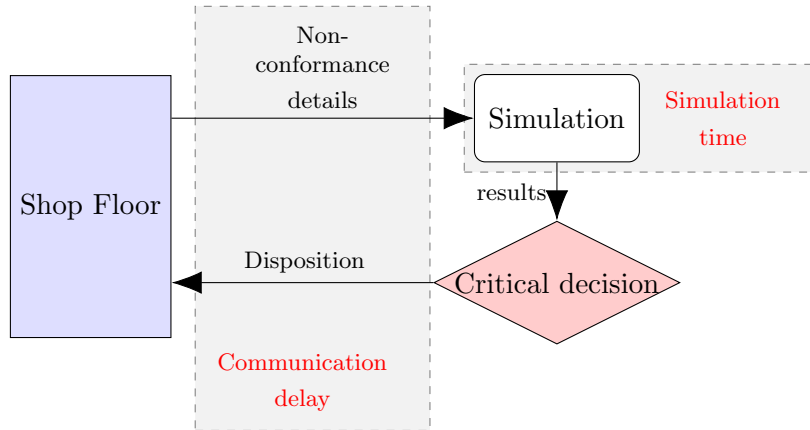
---

In the manufacturing of engineering structures, process variabilities often lead to deviations from the tolerance specifications on the features of a part or an assembly. These deviations can influence the stresses at the contact interfaces in the assembly, potentially leading to reduced performance or even failure of the assembled structure. Economic constraints often provide the motivation to salvage the “non-conforming” hardware. Depending on the nature of the deviations, a non-conforming part or assembly can be either: *accepted*, *reworked* to improve compliance or *rejected*. The decision is usually based on further analyses or tests on the non-conforming part. Naturally, our understanding of the mechanics of the contact interfaces in assemblies is crucial in this decision-making process.

Analytical and empirical solutions are usually not available for complex mechanical/structural configurations, while experimental testing is cost-prohibitive. Therefore, numerical modelling is often the most practical route to aid the decision-making process. However, the numerical simulations consume a significant amount of time and computational resources in building the model and subsequent resolution, especially for assemblies that include dynamic contacts generating additional complexities. Also, the communication delay between the shop floor and the simulation team adds to the delays. These bottlenecks have been schematically described in Fig. I. Therefore, it is necessary to accelerate assembly simulations in order to comply with desired decision times in the manufacturing cycle.

Real-time simulation tools have been of interest in various decision-making processes across fields like design and health-monitoring, among others. In such applications, fast simulations need to be performed for parametric values that are fed to the simulator in real-time. However, the complexity of a model usually increases exponentially with the number of parameters in the model. Therefore, high-fidelity models are often replaced with surrogate models that offer fast simulations, sometimes referred to as *digital twins*. Reduced Order Modelling (ROM) a.k.a. Model Order Reduction, is a well-known framework to create such models. ROMs are generally built to efficiently solve a parametric problem in a *multi-query* context. Therefore, ROM is a promising framework for moving towards efficient simulations of non-conforming engineering hardware.

For a real-time simulation of non-conforming assemblies, the task is two-fold. First is



**Figure I:** Decision-making process using numerical simulations. Additional time delays can be caused by communication and simulation. In an ideal situation: Communication delay + Simulation time  $\ll$  Manufacturing time

the modelling of geometric variabilities in individual parts. Shape parametrized ROMs have been explored [1–3] which can be used for parametrizing geometric variabilities. The second, and more challenging, is the application of ROM approach to contact mechanics. As contact mechanics deals with multi-body problems with interaction between the boundaries of each body, it introduces geometric non-linearities. These non-linearities are different in nature compared to, say, non-linear material behaviour or loading conditions which have been widely studied by the ROM community. Therefore, contact ROMs still need extensive research before they can be used in industrial real-time applications, such as the acceleration of the decision-making process discussed previously.

Therefore, in this thesis, the current limitations of contact ROMs have been investigated and corresponding mitigation strategies have been proposed. The thesis begins with an introduction to contact mechanics and the state of the art in ROMs applied to contact mechanics (Chapter 1). Then, the limitations of the so-called *low-rank* approach are investigated (Chapter 2). Strategies to counter these limitations are developed using over-complete dictionaries (Chapter 3) and non-linear interpolations (Chapter 4).

# 1 | State of the Art

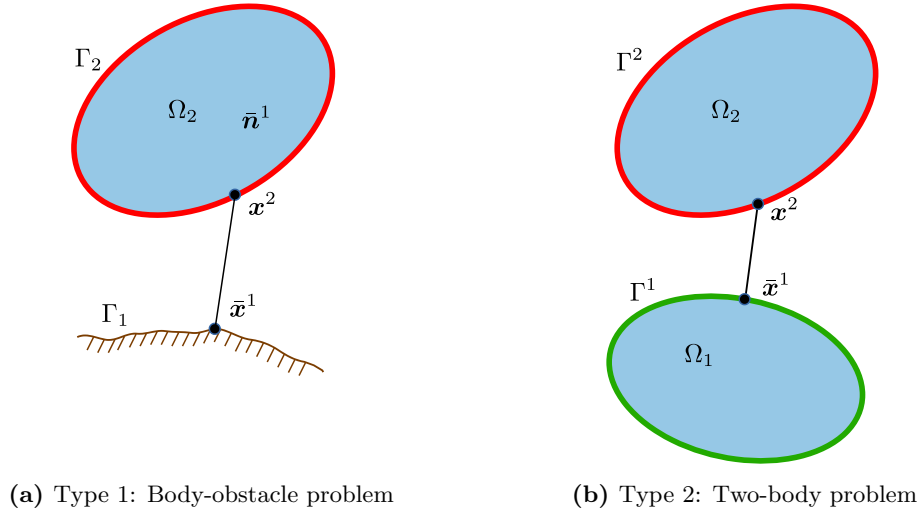
---

In this chapter, an overview of concepts and state of the art related to contact mechanics and reduced order modelling are presented. Ideas in contact mechanics are first introduced using a two-body problem where inequality constraints apply to the energy minimization statement of the standard mechanics problem. Some of the classical methods to solve inequality constrained problems are discussed. As one of the principal sources of complexity in such problems is contact detection, some methods applicable to contacts between one-dimensional boundaries are introduced. In the second part of the chapter, the concept of Reduced Order Modelling and its application on unconstrained problems is discussed. In the final part of the chapter, the two elements are combined, where the developments in reduced models of contact mechanics are presented.

## 1.1 Overview of Contact Mechanics

Many computational mechanics problems involve interaction between multiple bodies with various types of physics, say thermal or mechanical, where each kind of physical interaction adds a challenge to the modelling process. Contact mechanics, the field that considers mechanical interaction in multi-body problems, finds applications in a wide variety of fields as almost all mechanical system, natural or man-made, displays some sort of contact phenomena. A few examples that fall in this category include: foundation studies in civil engineering, human joints in biomechanics, roller joints in mechanisms, adhesive and frictional contacts in tribology and crash simulations in automotive industry.

Historical works in contact mechanics include studies by that of Hertz [4] on analytical solutions of contact between spherical frictionless surfaces of elastic bodies, that of Coulomb [5] on frictional contact, and many more. However, the work by Signorini [6] on the contact problems with elastic bodies, which consisted of the Boundary Value problem with non-penetration constraints, is central to approach of contact mechanics in the computational field. The problem was titled “problems with ambiguous boundary conditions.” by Signorini, aptly describing the nature of the problem. Analytical solutions to certain contact problems have been computed in [7–9].



**Figure 1.1:** Kinematic description of mechanical problems with possibility of contact. Points  $\mathbf{x}^2$  will possibly come into contact with point  $\bar{\mathbf{x}}^1$ , and hence considered as a contact pair. The distance between the two points in the initial and deformed configuration is given by  $g(\mu, \mathbf{v})(\mathbf{x}^2)$  and  $(g(\mu, \mathbf{v})(\mathbf{x}^2) - k(\mu, \mathbf{v}; \mathbf{v})(\mathbf{x}^2))$  respectively

### 1.1.1 Parametrized contact mechanics problems

Usually, an engineering problem is described as *parametrized* if a parametric solution is sought i.e. the solution of the problem is sought for a range of parameter values. For a mechanical problem, the parameters can be any attribute of the mechanical system such as material properties, geometry, loading or boundary conditions, etc. Parametric solutions usually find applications in design optimization, real-time monitoring and inverse problems (see Section 1.2). In this section, parametric mechanical problems with non-penetration constraints on displacement field are discussed. Linear elastic and small deformation problems are considered, keeping the evaluation of internal energy straightforward. The inequality constraints may be a result of the presence of either an obstacle or a second body in the domain, with possibility of contact between different surfaces in the deformed configuration.

For convenience, the contact problems will be referred to as Type-1, referring to a problem involving a deformable body and an immovable obstacle, and Type-2 referring to problems involving two deformable bodies. A schematic of the two types of contact problems is shown in Fig. 1.1. Contact phenomena like friction and adhesiveness are neglected. The resolution of contact mechanics problems depends on computing the distance between surface of bodies involved. As evaluation of distance is not a trivial task because of dynamic contact pairing, different strategies exist based on underlying simplifications. The details to these approaches can be found in [10–12].

The generic weak form of a mechanical problem with a parameter  $\mu$  involving contacts (for both Type-1 and Type-2) is described by the following inequality constrained

minimization problem:

$$\begin{aligned} \mathbf{u} &= \arg \min_{\mathbf{v} \in \mathcal{V}} \frac{1}{2} a(\mu; \mathbf{v}, \mathbf{v}) - f(\mu; \mathbf{v}) \\ \text{s.t. } \quad &k(\mu, \mathbf{v}; \mathbf{v})(\mathbf{x}) \leq g(\mu, \mathbf{v})(\mathbf{x}) \quad \text{on } \Gamma^2 \end{aligned} \quad (1.1)$$

where

- $a(\mu; \mathbf{v}, \mathbf{v})$  is strain energy function and  $f(\mu; \mathbf{v})$  is the work contribution by the external forces. The operators are defined  $\forall \mathbf{v} \in \mathcal{V}$  and  $\forall \mu \in \mathcal{P}$ , where  $\mathcal{V}$  be an appropriate function space for displacement  $\mathbf{v}$  and  $\mathcal{P}$  is the parametric space for the parameter  $\mu$ . Such problems with parametric dependence are usually referred to as parametrized problems. Note that  $\mu$  may not be necessarily a scalar and the parametric space  $\mathcal{P}$  can be possibly multidimensional.
- $k(\mu, \mathbf{v}; \mathbf{v})(\mathbf{x})$  and  $g(\mu, \mathbf{v})(\mathbf{x})$  are the distance functions that indicate separation between bodies, defined conveniently on surface  $\Gamma^2$ . The first term indicates the contribution of displacement field  $\mathbf{v}$  to the distance and the second term indicates the distance in the undeformed configuration. For a Type-2 problem, the distance functions can be expressed as:

$$\begin{aligned} k(\mu, \mathbf{v}; \mathbf{v})(\mathbf{x}) &= [\mathbf{v}^2(\mathbf{x}^2) - \mathbf{v}^1(\bar{\mathbf{x}}^1)] \cdot \mathbf{n} \\ g(\mu, \mathbf{v})(\mathbf{x}) &= [\mathbf{x}^2 - \bar{\mathbf{x}}^1] \cdot \mathbf{n} \end{aligned} \quad (1.2)$$

The pair  $(\mathbf{x}^2, \bar{\mathbf{x}}^1)$  and the normal  $\mathbf{n}$  are determined on the discrete domain using various methods, some of which are given in upcoming Section 1.1.3.1. For a Type-1 problem, the displacement field of the obstacle  $\mathbf{v}^1$  vanishes.

### Notation 1.1 :

Notations to distinguish strong and weak forms are adopted from [13]. Linear and bilinear operators that appear in weak formulations are expressed in the format  $o(\cdot; \cdot, \cdot)$ , where the non-linear dependencies and linear dependencies are separated by the semi-colon ';'. On the other hand, strong form functions are expressed in the format  $o(\cdot; \cdot, \cdot)(\cdot)$ , distinguished from weak forms using an additional argument which is the geometric position. The same applies to linear forms expressed as  $o(\cdot; \cdot)$  and  $o(\cdot; \cdot)(\cdot)$ , where the linear dependency is on the second argument.

## 1.1.2 Contact Formulations

To treat the contact problem (1.1) numerically, the contribution of contact to the total energy of the system is computed. This additional energy term is defined differently in each contact formulation, however, the basic idea is to introduce contact forces, directly or indirectly, that will prevent penetrations of contact surfaces. Among numerous formulations available, three of them are introduced in this section. Two of them are based on introducing the Lagrange Multiplier, which computes the contact forces directly, while the other method penalizes the penetrations.

### 1.1.2.1 Lagrange Multiplier Method

One of classical methods is an attempt to accurately satisfy the inequality constraints of the contact problem (or any minimization problem) is the Lagrange Multiplier Method (LMM). The resulting optimization problem can be expressed as a saddle point problem of the Lagrangian functional:

$$(\mathbf{u}, \lambda) = \arg \min_{\mathbf{v} \in \mathcal{V}, \eta \in \mathcal{W}_+} \max \mathcal{L}(\mathbf{v}, \eta) \quad (1.3)$$

where

$$\mathcal{L}(\mathbf{v}, \eta) = \frac{1}{2} a(\mu; \mathbf{v}, \mathbf{v}) - f(\mu; \mathbf{v}) + b(\mu, \mathbf{u}; \eta, \mathbf{u}) - d(\mu, \mathbf{u}; \eta) \quad (1.4)$$

where the distance function is expressed by following weak forms:

$$\begin{aligned} b(\mu, \mathbf{u}; \eta, \mathbf{u}) &= \int_{\Gamma^2} \eta k(\mu, \mathbf{u}; \mathbf{u})(\mathbf{x}) \partial\Gamma \\ d(\mu, \mathbf{u}; \eta) &= \int_{\Gamma^2} \eta g(\mu, \mathbf{u})(\mathbf{x}) \partial\Gamma \end{aligned} \quad (1.5)$$

where  $\mathcal{W}_+$  is an appropriate non-negative function cone for contact pressure defined on the surface  $\Gamma^2$ . The inequality  $\lambda \geq 0$  admits a physical meaning because  $\lambda$  is equivalent to contact pressure. A negative contact pressure implies traction, in other words, adhesion between contact surfaces which contradicts the simplifying assumptions.

The KKT conditions associated to the optimizer  $(\mathbf{u}, \lambda) \in \mathcal{V} \times \mathcal{W}_+$  of (1.3) can be expressed as follows:

$$a(\mu; \mathbf{v}, \mathbf{u}) - f(\mu; \mathbf{v}) + b(\mu, \mathbf{u}; \lambda, \mathbf{u}) = 0, \quad \mathbf{v} \in \mathcal{V} \quad (1.6a)$$

$$b(\mu, \mathbf{u}; \eta, \mathbf{u}) - d(\mu, \mathbf{u}; \eta) \leq 0, \quad \eta \in \mathcal{W}_+ \quad (1.6b)$$

$$b(\mu, \mathbf{u}; \lambda, \mathbf{u}) - d(\mu, \mathbf{u}; \lambda) = 0 \quad (1.6c)$$

Notice that the last KKT conditions (1.6c) implies that contribution of contact energy term will be zero at the solution. This makes sense because the non-penetration constraint can prevent the system from reaching the true minima of the energy functional, but it cannot add energy to the system.

### 1.1.2.2 Penalty Method

To constrain the solution in the feasible region defined by the inequality constraints, violations of inequality constraints can be penalized by adding a penalty term to the energy of the system. The resulting system is an unconstrained problem, unlike the LMM problem.

$$\mathbf{u} = \arg \min_{\mathbf{v} \in \mathcal{V}} \frac{1}{2} a(\mu; \mathbf{v}, \mathbf{v}) - f(\mu; \mathbf{v}) + \int_{\Gamma^2} \tau [(k(\mu, \mathbf{v}; \mathbf{v})(\mathbf{x}) - g(\mu, \mathbf{v})(\mathbf{x}))^+]^2 \partial\Gamma \quad (1.7)$$

where  $(z)^+ := \begin{cases} z & \text{if } z > 0 \\ 0 & \text{otherwise} \end{cases}$

The solution of (1.7) may not necessarily satisfy the inequality constraints exactly, but may be nearly compliant for sufficiently large values of the penalty parameter  $\tau$ . For very large values of  $\tau$ , the penalty solution should approach the LMM solution, however, the numerical system are ill-conditioned in this regime of  $\tau$ .

The penalty parameter  $\tau$  can be interpreted as a linear spring that resists penetration. Therefore, a natural variation of the penalty method can be built by using a non-linear penalty function where the penetration and the resultant contact pressure have a non-linear relation. This approach is useful to reduce the penetration effects, but also in cases where a constitutive law relating contact pressure and penetration is known. However, the problem of ill-conditioning still persists when the constitutive law enters a highly stiff regime [10, Chapter 5, 6].

### 1.1.2.3 Augmented Lagrange Multiplier Method

To circumvent the ill-conditioning problems of penalty method, Augmented Lagrange Multiplier Method [14] problems have been developed. This class of method use a combination of penalty and LMM. The basic idea is to establish contact forces using the penalty method and then transfer these forces, in multiple steps, into the Lagrange Multiplier, thereby reducing illegal penetrations. The advantage of using this method is that the solution will have high compliance to the constraints irrespective of the magnitude of the penalty parameter. At a given step  $k$ , the following problem is minimized:

$$\begin{aligned} \mathbf{u}^k = \arg \min_{\mathbf{v} \in \mathcal{V}} & \frac{1}{2} a(\mu; \mathbf{v}, \mathbf{v}) - f(\mu; \mathbf{v}) + \int_{\Gamma^2} \tau [(k(\mu, \mathbf{v}; \mathbf{v})(\mathbf{x}) - g(\mu, \mathbf{v})(\mathbf{x}))^+]^2 \partial\Gamma \\ & + b(\mu, \mathbf{v}; \lambda^k, \mathbf{v}) - d(\mu, \mathbf{v}; \lambda^k) \end{aligned}$$

with the update step:

$$\lambda^{k+1}(\mathbf{x}) = \lambda^k(\mathbf{x}) + \tau [(k(\mu, \mathbf{v}; \mathbf{v})(\mathbf{x}) - g(\mu, \mathbf{v})(\mathbf{x}))^+]$$

and initial conditions:

$$\lambda^0(\mathbf{x}) = 0$$

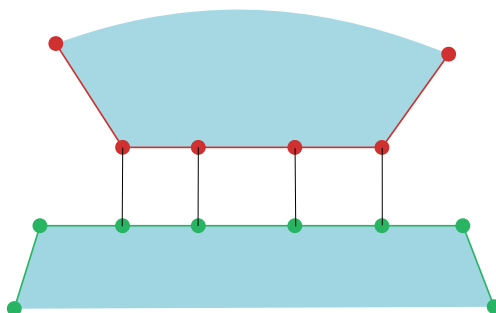
As  $\lambda^k$  does not contain the true Lagrange Multiplier during the intermediate steps, this method still allows some penetration. However, as the Lagrange Multipliers are updated with new contact pressure forces, the penetration reduces in the subsequent steps. With sufficient number of steps, penetration diminishes to very low-levels and the solution approaches that of the LMM solution.

### 1.1.3 Discrete Formulations

In this section, the usual discretization strategies based on Finite Element (FE) are briefly introduced. One of the foremost problems in contact mechanics is the detection of contact pairs. For large-scale problems, this step can prove to be expensive. Then comes the part of building the finite element (FE) operators that can be used to enforce or verify the constraints. As the contact problems in upcoming chapters will be solved by using the LMM approach, the discussion on discretization schemes will be limited to this method. Discretization of the penalty and Augmented LMM can be carried out in a similar manner. Details of all approaches can be found in [10, 12].

#### 1.1.3.1 Contact Detection and Pairing

Establishing the region of contact between two bodies is one of the primary challenges in contact problem. In simplified cases of small displacement, conforming meshes can be used, thereby allowing the use of *node-to-node* contacts. In this case, both points in the contact pair  $\boldsymbol{x}^2$  and  $\bar{\boldsymbol{x}}^1$  are nodes over respective surfaces, as shown in Fig. 1.2. Naturally, the contact pair can be predetermined based on the discretization. Obviously, the use of conforming mesh limits the applicability of this method. For example, it cannot be applied to sliding surfaces.



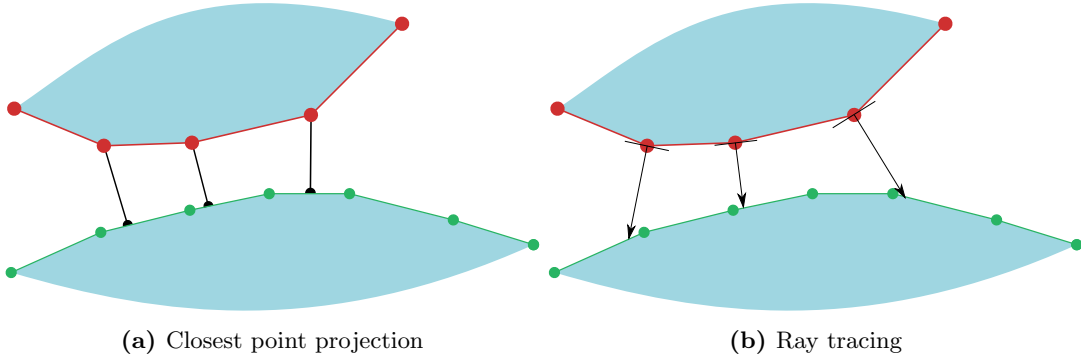
**Figure 1.2:** Schematic of node-to-node contact pairing for conforming surface meshes

*Node-to-segment* method is used in cases of non-conforming meshes with 1D boundaries, for e.g., 2D problems with possibility of contact on the edges of bodies. The contact pairings  $(\boldsymbol{x}^2, \bar{\boldsymbol{x}}^1)$  are not straightforward to compute as it was in node-to-node method. The contact pair evaluation is one of the sources of complexities in contact problems, especially for large displacement problems as contact pairs can be a strong function of displacement field. Established methods of computing contact pairs include:

- **Closest Point Projection:** For each node  $\boldsymbol{x}^2$  on  $\Gamma^2$ , the closest point  $\bar{\boldsymbol{x}}^1$  on  $\Gamma^1$  is chosen as the pair [15]. Naturally, the distance functions are evaluated using (1.2) with  $\boldsymbol{n} = \bar{\boldsymbol{n}}^1$ , the normal at the point  $\bar{\boldsymbol{x}}^1$ . A limitation of this method is the possibility of non-unique closest point in certain geometries [12, Chapter 2].

- **Ray Tracing:** At each node  $\mathbf{x}^2$  of  $\Gamma^2$ , the local normal  $\mathbf{n}^2$  is extended to intersect  $\Gamma^1$  at  $\bar{\mathbf{x}}^1$  [16]. Hence, distance functions are evaluated using  $\mathbf{n} = -\mathbf{n}^2$  for this method. This method may not find the nearest point, but avoids the problems of non-uniqueness and projecting on irregular points.

The two methods are shown in Fig. 1.3. The normals  $\mathbf{n}^1$  and  $\mathbf{n}^2$  are evaluated in deformed configuration. The pairing  $(\bar{\mathbf{x}}^1, \mathbf{x}^2)$  is hence dependent on the displacement field in both methods. This forces the distance function  $k$  to have an implicit non-linear dependence on displacement, apart from the explicit linear dependence that is evident in the expression.



**Figure 1.3:** Schematics of node-to-segment methods for finding contact pairs

### 1.1.3.2 Discretization using Finite Elements

The discrete equivalent of the Lagrangian functional in (1.4) can be derived by introducing finite element spaces  $\mathbf{V}^h \subset \mathbf{V}$  and  $\mathcal{W}_+^h \subset \mathcal{W}_+$ :

$$\mathcal{L}^h(\mathbf{v}, \boldsymbol{\eta}) = \frac{1}{2} \mathbf{v}^T \mathbf{K}(\boldsymbol{\mu}) \mathbf{v} - \mathbf{v}^T \mathbf{f}(\boldsymbol{\mu}) + \boldsymbol{\eta}^T [\mathbf{C}(\boldsymbol{\mu}, \mathbf{v}) \mathbf{v} - \mathbf{g}(\boldsymbol{\mu}, \mathbf{v})] \quad (1.8)$$

where  $\mathbf{v} \in \mathbb{R}^{N_u^h}$  and  $\boldsymbol{\eta} \in \mathbb{R}^{N_\lambda^h}$ . Here,  $N_u^h$  and  $N_\lambda^h$  indicate the number of respective finite element degrees of freedoms (dofs). Also, note that the notation  $\mathbf{u}$  and  $\mathbf{v}$  are used for both continuous and discrete versions of displacement fields.

Let  $\{\mathbf{N}_1, \mathbf{N}_2 \dots \mathbf{N}_{N_u^h}\}$  and  $\{M_1, M_2 \dots M_{N_\lambda^h}\}$  denote the basis for the FE space  $\mathbf{V}^h$  and cone  $\mathcal{W}_+^h$  respectively. Here, for a system with  $n$  displacement dofs per node,  $\mathbf{N}_i : \Omega \rightarrow \mathbb{R}^n$  and  $M_j : \Gamma \rightarrow \mathbb{R}$ . The discrete operators  $\mathbf{K}(\boldsymbol{\mu})$  and  $\mathbf{C}(\boldsymbol{\mu}, \mathbf{v})$  can be defined as:

$$\begin{aligned} \mathbf{K}(\boldsymbol{\mu})_{ij} &= a(\boldsymbol{\mu}; \mathbf{N}_i, \mathbf{N}_j) \\ \mathbf{C}(\boldsymbol{\mu}, \mathbf{v})_{ij} &= b(\boldsymbol{\mu}, \mathbf{v}; M_i, \mathbf{N}_j) \end{aligned}$$

and the vectors  $\mathbf{f}$  and  $\mathbf{g}$  are defined as:

$$\begin{aligned}\mathbf{f}(\mu)_i &= f(\mu; \mathbf{N}_i) \\ \mathbf{g}(\mu, \mathbf{v})_i &= d(\mu, \mathbf{v}; M_i)\end{aligned}$$

The bilinear and linear terms  $b$  and  $d$  are non-linear w.r.t. the displacement field  $\mathbf{v}$  and the parameter  $\mu$ , which is reflected in discrete form as  $\mathbf{C}(\mu, \mathbf{v})$  and  $\mathbf{g}(\mu, \mathbf{v})$  respectively. Using the above operators, the discrete form of the KKT conditions in (1.6) for the optimal solution  $(\mathbf{u}, \boldsymbol{\lambda}) \in \mathbb{R}^{N_u^h} \times \mathbb{R}^{N_\lambda^h}$  can be written as:

$$\mathbf{K}(\mu)\mathbf{u} - \mathbf{f}(\mu) + \mathbf{C}^T(\mu, \mathbf{u})\boldsymbol{\lambda} = \mathbf{0} \quad (1.9a)$$

$$\boldsymbol{\lambda} \geq \mathbf{0} \quad (1.9b)$$

$$\mathbf{C}(\mu, \mathbf{u})\mathbf{u} - \mathbf{g}(\mu, \mathbf{u}) \leq \mathbf{0} \quad (1.9c)$$

$$\boldsymbol{\lambda}^T[\mathbf{C}(\mu, \mathbf{u})\mathbf{u} - \mathbf{g}(\mu, \mathbf{u})] = 0 \quad (1.9d)$$

Apart from the non-linearity of the constraint operators  $\mathbf{C}$  and  $\mathbf{g}$ , additional non-linearities appear because the regions where the inequality constraints are active and inactive are unknown. Active regions correspond to the regions where contact is established, or in other words, constraints satisfy the equality. These non-linearities are potential contributors to the overall computational complexity.

**Stability of the saddle point problem:** The stability conditions of mixed formulations usually appear in the form of an inf-sup condition. The inf-sup constant of the bilinear operator  $b$  in (1.4) must be positive for every pair of test functions in the two solution spaces. Not every combination of finite-element spaces  $\mathbf{V}^h$  and  $\mathcal{W}_+^h$  satisfies the stability condition. In this thesis, linear elements for displacement and node-centered zero order (piecewise-constant) shape functions for contact pressure are used (based on Reference [13]), which generates smooth solutions for two-dimensional problems. More details on inf-sup stability of contact problems can be found in [10, 17].

#### 1.1.4 Resolution schemes

Even for contact problems involving small deformations with linear elastic constitutive law, geometric non-linearities can still occur if large displacements are involved. This is because the contact pairs change with the displacement field. Even for small displacement problems where contact pairs can be approximately precomputed in the reference configuration, the region of contact cannot be predetermined. Linearization schemes are needed to solve the discrete problem, that resolve both the contact region and the contact pairs in an iterative manner. Two such schemes: Fixed-point and Newton-Raphson iterations are described here. Python notation is used to indicate slicing operations.

### 1.1.4.1 Fixed point method

Fixed point schemes are used to find solutions to non-linear equations by posing them in the fixed-point form  $x = f(x)$ , leading to a fairly simple iterative form  $x^{p+1} = f(x^p)$ . For a discrete system with multiple dofs, the linearized form usually looks like  $\mathbf{A}(\mathbf{x}^p)\mathbf{x}^{p+1} = \mathbf{b}(\mathbf{x}^p)$ . These linearized forms can be iterated until convergence, without computing derivatives or gradients of the non-linear equation.

Fixed point methods for an inequality constrained problem can be implemented using active set approach [18, 19], in which the constraints are divided into active and inactive sets. For the discretized contact problem, the balance form (1.9a) can be used in the fixed-point method:

$$\mathbf{K}(\mu)\mathbf{u} + \mathbf{C}^T(\mu, \mathbf{u})\boldsymbol{\lambda} = \mathbf{f}(\mu) \quad (1.10)$$

Given the current estimate of active set  $\mathcal{A}$ , the activation state of the inequality constraints can be written as:

$$\boldsymbol{\lambda}[\mathcal{A}^c] \leftarrow \mathbf{0} \quad (1.11a)$$

$$\mathbf{C}_{\mathcal{A}}(\mu, \mathbf{u})\mathbf{u} - \mathbf{g}_{\mathcal{A}}(\mu, \mathbf{u}) = \mathbf{0} \quad (1.11b)$$

where

$$\mathbf{C}_{\mathcal{A}}(\mu, \mathbf{u}^p) = \mathbf{C}(\mu, \mathbf{u}^p)[\mathcal{A}, :]$$

$$\mathbf{g}_{\mathcal{A}}(\mu, \mathbf{u}^p) = \mathbf{g}(\mu, \mathbf{u}^p)[\mathcal{A}]$$

and  $\mathcal{A}^c$  is the complement of set  $\mathcal{A}$ .

These conditions apply because each entry of the vectors  $\mathbf{C}(\mu, \mathbf{u})\mathbf{u} - \mathbf{g}(\mu, \mathbf{v})$  and  $\boldsymbol{\lambda}$  correspond to a nodal constraint and a nodal Lagrange Multiplier respectively. As the active set contains the list of active nodes, it can be used for filtering the entries of constraint operators and enforcing the equality conditions i.e., where contact between bodies is established. The displacement field and the active Lagrange multipliers at iterative level  $p+1$  can be obtained using the linear system obtained from the equilibrium condition (1.10) and the active constraint (1.11b)

$$\begin{bmatrix} \mathbf{K}(\mu) & \mathbf{C}_{\mathcal{A}}^T(\mu, \mathbf{u}^p) \\ \mathbf{C}_{\mathcal{A}}(\mu, \mathbf{u}^p) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u}^{p+1} \\ \boldsymbol{\lambda}_{\mathcal{A}}^{p+1} \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\mu) \\ \mathbf{g}_{\mathcal{A}}(\mu, \mathbf{u}^p) \end{bmatrix} \quad (1.12)$$

and the discrete contact pressure can be updated as:

$$\begin{aligned} \boldsymbol{\lambda}^{p+1}[\mathcal{A}] &\leftarrow \boldsymbol{\lambda}_{\mathcal{A}}^{p+1} \\ \boldsymbol{\lambda}^{p+1}[\mathcal{A}^c] &\leftarrow \mathbf{0} \end{aligned}$$

The sign of entries in the expression  $\mathbf{C}(\mu, \mathbf{u})\mathbf{u} - \mathbf{g}(\mu, \mathbf{v})$  and in the vector  $\boldsymbol{\lambda}$  correspond to satisfaction or violation of inequality constraints at the respective node's support. The active set for the next iteration is obtained by verifying the constraints:

$$\mathcal{A}_i \leftarrow \{i \mid (\mathbf{C}(\mu, \mathbf{u}^p)\mathbf{u}^{p+1} - \mathbf{g}(\mu, \mathbf{u}^p))_i \geq 0 \text{ and } \lambda_i > 0\} \quad (1.13)$$

Moreover, the contact pairs are also be updated using the latest displacement field and the new operators  $\mathbf{C}(\mu, \mathbf{u}^{p+1})$  and  $\mathbf{g}(\mu, \mathbf{u}^{p+1})$  are computed. This process is continued until convergence.

#### 1.1.4.2 Newton-Raphson

The Newton-Raphson iterative method (a.k.a. Newton's method) is a derivative-based iterative method for root finding and optimization problems. For finding the roots of an equation  $f(x) = 0$ , the Newton-Raphson update at an iterative level  $p$  can be written as  $\Delta x^{p+1} = -\frac{f(x^p)}{f'(x^p)}$ , where  $\Delta x^{p+1} = x^{p+1} - x^p$ . For a discrete system with residual  $\mathbf{b}(\mathbf{x}) - \mathbf{A}(\mathbf{x})\mathbf{x}$ , the Newton-Raphson update step can be written as:

$$\mathbf{A}_T(\mathbf{x})\Delta \mathbf{x}^{p+1} = \mathbf{b}(\mathbf{x}^p) - \mathbf{A}(\mathbf{x}^p)\mathbf{x}^p$$

where  $\mathbf{A}_T(\mathbf{x})$  is the tangent matrix of the residual.

The Newton-Raphson method for contact problems is based on computing the tangent matrix  $\mathbf{K}_T^c$  of the residual of equilibrium equation (1.10) and then computing incremental updates. The contribution to the tangent matrix  $\mathbf{K}_T$  comes from of the contact part of the energy functional or the equilibrium equation. It consists of the second order terms that appear due to variations in contact pairing. The computation of tangent matrix for the non-linear system is quite detailed and can be found in Reference [10, Chapter 9, 10]. The iterative update at a level  $p + 1$  can be written as:

$$\begin{aligned} \begin{bmatrix} \mathbf{K}(\mu) + \mathbf{K}_T^c(\mu, \mathbf{u}^p) & \mathbf{C}_A^T(\mu, \mathbf{u}^p) \\ \mathbf{C}_A(\mu, \mathbf{u}^p) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{u}^{p+1} \\ \Delta \boldsymbol{\lambda}_A^{p+1} \end{bmatrix} \\ = \begin{bmatrix} \mathbf{f}(\mu) \\ \mathbf{g}_A(\mu, \mathbf{u}^p) \end{bmatrix} - \begin{bmatrix} \mathbf{K} & \mathbf{C}_A^T(\mu, \mathbf{u}^p) \\ \mathbf{C}_A(\mu, \mathbf{u}^p) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u}^p \\ \boldsymbol{\lambda}_A^p \end{bmatrix} \end{aligned}$$

The stiffness matrix  $\mathbf{K}(\mu)$  appears in addition to the contact tangent matrix because the stiffness matrix is the tangent of the elastic energy functional. After computing the increment, the state vector can be updated as:

$$\begin{bmatrix} \mathbf{u}^{p+1} \\ \boldsymbol{\lambda}_A^{p+1} \end{bmatrix} = \begin{bmatrix} \mathbf{u}^p \\ \boldsymbol{\lambda}_A^p \end{bmatrix} + \begin{bmatrix} \Delta \mathbf{u}^{p+1} \\ \Delta \boldsymbol{\lambda}_A^{p+1} \end{bmatrix}$$

This is followed by the active set update in (1.13).

## 1.2 Reduced Order Modelling for parametric problems

In recent applications of numerical methods to engineering design, decision-making processes are increasingly dependent on simulation results [20–23]. Various configurations of the product can be explored and compared by optimization algorithms to develop an optimal configuration using modern simulation tools [24–26]. For the integration of simulation tools into the decision-making process, simulation time for each configuration becomes a primary factor in determining the swiftness of the process. In this context of *multi-query* simulations, a high computational complexity of traditional simulation methods might discourage comprehensive exploration of the design space for economic reasons.

Traditional methods of simulations were restricted to the design phase of engineering developments due to the heavy computational costs. On the other hand, modern simulation tools capable of real-time simulations have promising applications not only in the design phase but also during the life-cycle of engineering products, for example, for operations and maintenance [27]. The idea behind such applications is to create real-time simulations whose predictions can be used to initiate corrective or preventive actions. At the heart of the development of real-time simulations lies the approach of *Reduced Order Modelling* (ROM).

Real-time simulators need to mimic the relevant physics of the system being studied taking into account the parameters that are fed to the simulator in real-time. High-fidelity numerical methods usually have very high computational complexity for such applications. The real-time query can be addressed by solving a high-dimensional problem, in which the real-time parameters are treated as coordinates of the problem (like space and time). Such a numerical model is usually referred to as a parametrized model and is usually very expensive to resolve using standard numerical methods. ROM is a framework for reducing the complexity of high-dimensional problems. ROMs have found successful applications in a wide variety of problems such as design optimization [28], uncertainty quantification [29], inverse problems [30], optimal control [31], real-time monitoring [21], and others.

Various methodologies have been adopted to create ROMs to address different bottlenecks. Some are based on reducing the number of dofs of the system, whereas others focus on the efficient evaluation of integrals over a small part of the domain. However, a common feature among most ROM methods has been the computation of a training set of *snapshots* i.e. a set of high-fidelity solutions based on which a reduced model is built. A brief description of a few classes of reduction methods is given here:

- *Low-rank methods*: These methods primarily focus on finding a low-rank solution subspace. Usually, this is done by computing the basis of the subspace using Singular Value Decomposition (SVD) or Proper Orthogonal Decomposition (POD). Low-rank methods are further detailed in Section 1.2.1.
- *Hyper-reduction methods*: Primary aim of hyper-reduction methods is to reduce the number of unknowns as well as quadrature points to compute weak form terms

by generating a reduced integration domain (RID), which is generated by selecting appropriate dofs [32]. This feature of hyper-reduction is used in [33] to reduce the effort in the computation of internal variables of non-linear constitutive models, where the RID is made up of regions containing maximum energy of POD basis vectors. On the other hand, the quadrature rules for RID in Reference [34], are computed using the least-square procedure so that all training samples are exactly integrated. Cheaper evaluation of weak forms over the RID leads to an efficient model for non-linear problems.

- *Non-intrusive methods*: Some methods have been developed that do not need access and modification of the high-fidelity source codes, and are termed as *non-intrusive* methods. These are highly useful if proprietary source codes or complex constitutive models are used. Also, it can be applied to different kind of physics without changing the non-intrusive framework. As the numerical model and its operators are not accessible, many non-intrusive methods rely on fitting an interpolation in the parametric domain. Radial basis functions (RBF) have been popularly used in this context. In these works, a POD-based solution subspace is computed and an RBF-based interpolation for the reduced coefficients is developed for the POD subspace [35–37]. Another approach is that of sparse subspace learning where a predefined set of hierarchical parametric basis functions with special quasi-interpolative properties, allowing enrichment of the reduced model at minimal cost [38]. The CUR decomposition method takes a different approach for generating the parametric interpolation. A set of selected spatial and temporal snapshots are used to build a surrogate model with coefficients that fit the snapshots in the training set [39].
- *Projection-based a priori method*: The Proper Generalized Decomposition (PGD) treats the parameters of the problem as coordinates and yet circumvents the *curse of dimensionality* by using separated representations for solving the PDE [40]. This is accomplished by means of a greedy approach, where the current solution is updated with a rank-1 tensor product of separated functions in each step. The rank-1 updates are computed using the separated form of PDE, one direction at a time, where each direction corresponds to one or more coordinates of the problem, for instance, position, time, or a parameter, or a combination of these. This is described as an *a priori* method because a parametric solution is built and stored, resulting in quick accessibility of the solution. Efficient computation of separated form of non-linear terms has been implemented in the PGD framework using cross-approximations [41].

Even though these approaches have different names and methodologies, all of them, directly or indirectly, assume the low-rank hypothesis. In the next section, low-rank methods are described for generic unconstrained problems.

### 1.2.1 Low-rank methods

In many numerical problems, the underlying structure lies in a subspace whose dimensionality is quite small compared to the full dimensionality of the solution space used in the numerical model. Examples include thermal problems, transport problems, elasticity and fluid flows among many others [42, 43]. Low-rank methods can be used to reduce the complexity of such problems. The idea behind low-rank ROMs is to split the cost of computation into two stages.

- The *offline* stage, where most of the computational complexity is resolved, consists of computing a set of solutions in the parametric space i.e. the training set, and then computing the reduced subspace that defines the span of training set. The idea is that space spanned by the training set snapshots is a good approximation to the true solution subspace. This is usually done by computing a reduced basis (RB) corresponding to the span of the training set. This step is typically performed only once to create the reduced model.
- The *online* step is performed every time a new query is submitted, in which the on-demand solution is computed using the reduced model which is cheaper to evaluate than the high-fidelity model.

Using notation from the previous section, let the high-fidelity (FE) solution space be denoted by  $\mathcal{V}^h$  and the reduced subspace be denoted by  $\mathcal{V}^r = \text{span}(\Phi)$ , where  $\Phi$  denotes the corresponding RB. In online stage, solutions in the subspace  $\mathcal{V}^r$  are sought to approximate the high-fidelity solution. In discrete sense,  $\Phi$  is a matrix with each column corresponding to a basis vector. Then, the approximation in the subspace can be expressed as:

$$\mathbf{u} \approx \Phi \hat{\mathbf{u}} \quad \hat{\mathbf{u}} \in \mathbb{R}^{N_u^r}$$

All reduced degrees of freedom and reduced operators will be indicated by a hat  $\hat{\cdot}$

For constructing the RB  $\Phi$  in offline stage, a finite subset of the parametric space  $\mathcal{P}_{\text{tr}} \in \mathcal{P}$  is explored and an approximate subspace  $\mathcal{V}^r$  is extracted from the set of high-fidelity snapshots from the training set. The number of degrees of freedom associated to the reconstruction problem in online stage is same as the cardinality of basis  $\Phi$ , and it influences the efficiency of online phase. This is the key to low-rank methods, as it is based on the assumption that the intrinsic structure of the system behaviour lies in a low-rank subspace. If this assumption does not hold, it may be difficult to generate a RB that approximates well the system behaviour in the entire parametric space.

Next, two well known methods of finding the low-rank subspace: the Proper Orthogonal Decomposition and the Reduced Basis Methods are described briefly, as these methods form the backbone of many reduced models [42, 43]. This is followed by the necessity and tools for efficient computation of terms that are non-linear w.r.t. the parameter and/or the solution.

### 1.2.1.1 Proper Orthogonal Decomposition

In Proper Orthogonal Decomposition (POD) framework, high-fidelity snapshots are generated over the training set  $\mathcal{P}_{\text{tr}}$  are collected and arranged as columns of a matrix. Let  $\mathbf{S}_{\text{tr}} = \{\mathbf{u}_s\}_{s=1}^{N_S} \in \mathbb{R}^{N_u^h \times N_S}$  be a matrix containing the training set snapshots. Here,  $N_S$  is the number of snapshots and  $N_u^h$  is the number of finite element dofs. Typically,  $N_S \ll N_u^h$  for two reasons: first is that the snapshot computation can be expensive and second is the fact that discovery of a low rank subspace only needs a relatively small number of snapshots well distributed in the parametric domain. The RB can generated using the Singular Value Decomposition (SVD) or the Principal Component Analysis (PCA) methods. The SVD can be expressed as:

$$\mathbf{S}_{\text{tr}} = \mathbf{\Phi} \mathbf{\Sigma} \mathbf{\Psi}^T$$

where  $\mathbf{\Phi} \in \mathbb{R}^{N_u^h \times N_S}$  and  $\mathbf{\Psi} \in \mathbb{R}^{N_S \times N_S}$  are unitary matrices containing orthonormal vectors.  $\mathbf{\Sigma} \in \mathbb{R}^{N_S \times N_S}$  is a diagonal matrix containing singular values.

The decomposition given by the SVD has many interesting properties. The columns of  $\mathbf{\Phi}$  form a orthonormal basis for column-space of  $\mathbf{S}_{\text{tr}}$  and therefore can be used to approximate the solution subspace. Moreover, the columns of  $\mathbf{\Phi}$ , called as singular vectors, are arranged in the decreasing order of contribution to the total energy in  $\mathbf{S}_{\text{tr}}$ . Their contribution is indicated by the singular values in  $\mathbf{\Sigma}$ , which can be used to truncate the basis  $\mathbf{\Phi}$ . In fact, the singular values  $\{\sigma_i\}_{i=1}^{N_S}$  are arranged in decreasing order, where  $\sigma_i$  is the  $i$ -th diagonal value of  $\mathbf{\Sigma}$ , exhibit many interesting properties in relation to the matrix  $\mathbf{S}_{\text{tr}}$ , some of which are listed below:

- The square of singular values sum up to the total energy in the snapshot matrix, i.e.  $\|\mathbf{S}_{\text{tr}}\|_F = \sum_{i=1}^{N_S} \sigma_i^2$
- The energy contribution by first  $n$  singular vectors to the snapshot matrix is given by  $\sum_{i=1}^n \sigma_i^2$ .
- If  $\mathbf{\Phi}'$  and  $\mathbf{\Psi}'$  contain the first  $n$  columns of  $\mathbf{\Phi}$  and  $\mathbf{\Psi}$ , and also  $\mathbf{\Sigma}' \in \mathbb{R}^{n \times n}$  contains the first  $n$  singular values  $\{\sigma_i\}_{i=1}^n$  in its diagonal; then the approximation

$$\mathbf{S}_{\text{tr}} \approx \mathbf{\Phi}' \mathbf{\Sigma}' \mathbf{\Psi}'^T$$

is the best  $n$ -rank approximation in the sense of Frobenius-norm, with an approximation error  $\sum_{i=n+1}^{N_S} \sigma_i^2$ . A related property is that  $\mathbf{\Phi}'$  and  $\mathbf{\Psi}'$  form the best  $n$ -rank subspaces of the column and row space, respectively, of the matrix  $\mathbf{S}_{\text{tr}}$  in sense of Frobenius-norm.

The above properties can be used for truncation of the basis  $\mathbf{\Phi}$  by applying a tolerance  $\delta$  to the fraction of energy contribution of first  $n$  singular vectors. The first  $n$  singular vectors of  $\mathbf{\Phi}$  that contribute more  $1 - \delta$  fraction of energy are kept. This number  $n$  can be calculated using:

$$\min_n 1 - \frac{\sum_{i=1}^n \sigma_i^2}{\sum_{i=1}^{N_S} \sigma_i^2} < \delta$$

The truncation leads to a subspace of the full solution space, whose rank is a function of tolerance  $\delta$ . This allows us to regulate the dimensionality but at the cost of losing diminished accuracy. Generation of the reduced model using the reduced basis  $\Phi$  will be shown in Section 1.3.1.

### 1.2.1.2 Reduced Basis Methods

Even though POD methods offer a robust technique to compute the RB, they do not offer a strategy to sample the parametric space. On the other hand, the Reduced Basis Methods (RBM) are usually equipped with error estimators that allow greedy sampling of the parametric space [44, 45]. The idea of greedy methods is to arrive at the objective by using the locally best way forward at each step, which is typically done by solving a minimization problem. An error estimator  $\Delta(\mu)$  is an analytical error bound for the true reconstruction error as follows:

$$\|\mathbf{u}(\mu) - \Phi \hat{\mathbf{u}}(\mu)\|_2 \leq \Delta(\mu) \quad \forall \mu \in \mathcal{P}$$

The error estimator is computed using the discrete residual  $\mathbf{r}_h$  of the high-fidelity system. For an unconstrained discrete problem with the residual:

$$\mathbf{r}_h(\mathbf{u}, \mu) = \mathbf{f}(\mu) - \mathbf{K}(\mu)\mathbf{u}$$

the error bound  $\Delta(\mu)$  can be expressed as:

$$\begin{aligned} \Delta(\mu) &= \|\mathbf{K}^{-1}(\mu)\mathbf{r}_h(\mathbf{u}, \Phi \hat{\mathbf{u}}(\mu))\|_2 \\ &\leq \|\mathbf{K}^{-1}(\mu)\|_2 \|\mathbf{r}_h(\mathbf{u}, \Phi \hat{\mathbf{u}}(\mu))\|_2 \\ &= \frac{1}{\sigma_{\min}(\mathbf{K}(\mu))} \|\mathbf{r}_h(\mathbf{u}, \Phi \hat{\mathbf{u}}(\mu))\|_2 \end{aligned}$$

Further details pertaining to efficient computation of the residual norm and the least singular value can be found in [42, Chapter 3]. The error estimator has negligible cost of evaluation compared to the true error, as the true numerical solution  $\mathbf{u}$  is unknown. Therefore, at given greedy level  $k$ , the estimated best value of the parameter to be added is given by  $\mu^{k+1} = \arg \max \Delta(\mu)$  and is cheaper to compute. Thus, the training step is enriched with a new snapshot  $\mathbf{u}(\mu^{k+1})$  and the RB  $\Phi$  is updated using methods like Gram-Schmidt process. This avoids the computation of the basis from scratch at each step  $k$ . The process can be stopped when  $\max \Delta(\mu)$  is under a certain tolerance. The RBMs are advantageous in terms of reliability and efficiency of the reduction process, though the performance of the process, of course, depends closeness of the error estimator with the true error (see [42, Chapter 7]).

### 1.2.1.3 Computation of non-linear terms

Operators such as  $\mathbf{K}(\mu)$  and  $\mathbf{f}(\mu)$  can be a non-linear function of  $\mu$  and therefore, their computation in each iterative step of a linearized reduced problem can prove to be expensive. Despite the introduction of low dimensional spaces  $\mathcal{V}^r$ , the complexity of computing

these operators depends on the number of dofs in high-fidelity model. Commonly used methods to efficiently compute non-linear terms include the computation of affine decomposition using the Empirical Interpolation Method (EIM) and the computation of an reduced integration domain (RID) using hyper reduction methods.

The affine decompositions of the non-linear terms can be written as:

$$\mathbf{K}(\mu) \approx \sum_{i=1}^{N_K} \alpha_i^K(\mu) \mathbf{K}_i$$

$$\mathbf{f}(\mu) \approx \sum_{i=1}^{N_f} \alpha_i^f(\mu) \mathbf{f}_i$$

where the scalar values  $\{\alpha_i^K\}$  and  $\{\alpha_i^f\}$  are the only quantities dependent on  $\mu$ . Therefore, the cost of online computation for each query of  $\mu$  is reduced to the cost of computing these scalar values.

It is possible that simple dependencies on  $\mu$  admit affine decompositions that are easier to compute (see elastic energy computation in [13]). However, this is not the case in general and also there are many cases where operators show dependence on the solution  $\mathbf{u}$  itself, for e.g. large deformation problems. In such cases, an affine decomposition is even more valuable as the resolution of the reduced problem will involve an iterative procedure demanding evaluation of non-linear terms in each iterative step. Usually, the EIM framework is used for computing the affine decomposition in such cases [46]. EIM based affine decomposition is also equipped with so-called magic points in the non-parametric domain. Evaluation of the non-linear term only at the magic points is sufficient for computing the coefficients  $\{\alpha_i^K\}$  and  $\{\alpha_i^f\}$ . The cost of computation is reduced drastically as the parameter independent modes  $\mathbf{K}_i$  and  $\mathbf{f}_i$  can be computed offline. For further details on computation of EIM based affine forms, see [42, Chapter 10].

Similar reduction in computation cost can also be achieved by hyper-reduction methods, where the non-linear terms are evaluated only a selected subset of the quadrature points in the entire integration domain. Recalling the property of standard Galerkin approach in finite elements, each row in matrix  $\mathbf{K}(\mu)$  (and each entry in  $\mathbf{f}(\mu)$ ) corresponds to integration over the local support of a particular node. The hyper-reduced evaluation of non-linear operators can be written as  $\mathbf{PK}(\mu)$  (and  $\mathbf{Pf}(\mu)$ ). The rectangular matrix  $\mathbf{P}$  each row containing only one non-zero entry equal to 1, at a position corresponding to a node in the RID [33]. In other words, weak forms are computed only on supports surrounding the nodes in the RID. The RID has be constructed by selecting specific dofs or quadrature points in the entire domain. Various methods have been applied to this end, including free energy indicators [33], optimized cubature methods [34, 47], and Discrete Empirical Interpolation Method [48].

### 1.3 ROMs for parametrized contact mechanics problems

ROMs have been applied extensively to unconstrained problems such as thermal and mechanics problems [42, 43] and also to mixed problems with equality constraints such

as the incompressibility condition in Stokes and Navier-Stokes flow problems [49]. Application of ROMs to variational problems with inequality constraints (also referred to as variational inequalities) has been more recent [13, 48, 50–53]. Inequality constraints appear in mechanical problems with obstacles or multi-body mechanical problems where there is a possibility of contact between bodies and obstacles, or with other bodies. Moreover, the region of contact is unknown a priori. Although the scope of this thesis is limited to contact mechanics, several other applications of variational inequalities are found in porous media flow problems [54], cavitation problems in lubrication systems [55], anti-plane frictional problems [56] and even in financial trading problems [53].

As seen in the Section 1.1.2, inequality constraint problems are often posed in mixed form with the KKT conditions that force the Lagrange Multipliers to be non-negative. Enforcing this condition in ROMs is non-trivial as traditional methods of computing RBs do not preserve the non-negative nature of the input information.

One of the first works on reducing the contact mechanics problem [50], proposed the idea of using the contact pressure snapshots directly to define a non-negative subcone. Both displacement and contact pressure snapshots were generated in a greedy fashion using error estimators. Fundamental aspects of the reduced problem, such as the existence and uniqueness of the solution and inf-sup stability were also explored.

Compression of snapshots to create a reduced basis was studied in [51], where Non-Negative Matrix Factorization (NNMF) was used to compute a basis with user-specified cardinality but does not provide any means to specify truncation tolerance for the reduced basis. Also, error estimators for greedy sampling of parametric space are developed.

Algorithms to sort a precomputed set of snapshots in order of importance to create a compact basis are explored using projection methods in [53] and [13], using an Angle-Greedy and Cone-Projected Greedy (CPG) procedures respectively. The former does not take into account non-negative restrictions, whereas the latter uses a *cone* projection involving non-negative coefficients. This makes CPG more efficient in capturing the contact pressure subcone.

The hyper-reduction approach in [48] defines a subdomain of the contact problem containing the most important points. The reduction is achieved with the usage of a POD basis for displacement and resolution of the weak form on the reduced integration domain. Reconstruction of displacement solution on the full domain is relatively straightforward using the POD basis. On the other hand, the reconstruction of contact pressure involves solving a non-negative least square problem using the snapshots.

A Craig-Bampton based resolution of the contact problem was discussed in [57] where the reduction of the displacement field was achieved using the Krylov subspace method. The Lagrange Multiplier method was not reduced under the assumption that the number of contact dofs remains small.

All of these works are based on Lagrange Multiplier approach to solve the inequality constrained reduced problem. An exception to this trend is [52] where penalty approach is used. Instead, EIM is used for efficient computation of barrier functions that are used to restrict the solution in the feasible domain.

Most contributions build reduced models on problems with static contact pairs, meaning contact pairs do not change with the state of the system. However, [13] considers dynamic contact pairs, with node-to-segment formulation. To efficiently evaluate non-linearities due to dynamic pairing, they use Empirical Interpolation Method to define an affine decomposition of distance functions.

### 1.3.1 Low-rank approach for contact problems

As seen above, many of the ROM approach to contact mechanics are based on the low-rank approach. As the validity of low-rank hypothesis will be discussed in detail in Chapter 2, a generic low-rank approach is detailed here, based on ideas of [13, 51]. The reduction of the displacement and contact pressure fields is discussed, followed by defining the reduced problem using the reduced spaces for the two fields. As the contact problem is a mixed problem, the displacement and contact pressure field will also be referred to as *primal* and *dual* fields

**Reduction of displacement:** The displacement field  $\mathbf{u}$  for a contact mechanics does not need any special treatment. Any methods that can be applied for reducing a mechanics problem without contacts can also be used for the displacement. Left singular vectors of displacement snapshots were used in many references in ROMs of contact mechanics. The same will be the case in this thesis unless specified, and the primal RB will be derived using POD.

**Reduction of contact pressure:** Computing the dual basis is more complicated than the primal basis. The dual field i.e. the Lagrange multiplier must satisfy the non-negativity constraints. In case of the contact problems without adhesive and cohesive surfaces, this constraint admits a physical meaning as negative contact pressure cannot be admitted, as discussed in Section 1.1.3.2. An orthogonal basis generated using projection based methods cannot satisfy such constraints, as it is surely bound to contain negative entries. To ensure such constraints, one way is to define a subcone for the dual field instead of a subspace. A function cone, unlike a function space, is spanned by a set of non-negative basis functions and non-negative coefficients. In fact,  $\mathcal{W}_+$  and  $\mathcal{W}_+^h$  in Section 1.1.3 are also cones. By extension, a subcone, unlike a subspace, must be equipped by a non-negative RB and must be spanned by non-negative coefficients.

$$\mathcal{W}_+^r = \text{span}^+(\Theta) = \sum_{i=1}^N \widehat{\lambda}_i \theta_i \quad , \quad \widehat{\lambda}_i \geq 0$$

where  $\Theta$  and  $\{\theta_i\}_{i=1}^N$  are the dual RB and dual RB functions.  $\widehat{\lambda}_i$ s are the dual reduced dofs.

Once the reduced subspace  $\mathcal{V}^r$  and subcone  $\mathcal{W}_+^r$  are available, the reduced problem can be generated from (1.6), replacing the continuous spaces  $\mathcal{V}$  and  $\mathcal{W}_+$  with the reduced spaces  $\mathcal{V}^r$  and  $\mathcal{W}_+^r$ . The reduced KKT conditions on the solution  $(\widehat{\mathbf{u}}, \widehat{\boldsymbol{\lambda}}) \in \mathbb{R}^{N_u^r} \times \mathbb{R}^{N_u^r}$  can be expressed as:

$$\widehat{\mathbf{K}}(\mu)\widehat{\mathbf{v}} - \widehat{\mathbf{v}}^T \widehat{\mathbf{f}}(\mu) + \widehat{\mathbf{C}}^T(\mu, \widehat{\mathbf{v}})\widehat{\boldsymbol{\lambda}} = \mathbf{0} \quad (1.14a)$$

$$\widehat{\boldsymbol{\lambda}} \geq \mathbf{0} \quad (1.14b)$$

$$\widehat{\mathbf{C}}(\mu, \widehat{\mathbf{u}})\widehat{\mathbf{u}} - \widehat{\mathbf{g}}(\mu, \widehat{\mathbf{u}}) \leq \mathbf{0} \quad (1.14c)$$

$$\widehat{\boldsymbol{\lambda}}^T (\widehat{\mathbf{C}}(\mu, \widehat{\mathbf{u}})\widehat{\mathbf{u}} - \widehat{\mathbf{g}}(\mu, \widehat{\mathbf{u}})) = 0 \quad (1.14d)$$

where the discrete reduced operators are built by introducing the reduced basis functions  $\phi$  and  $\theta$

$$\begin{aligned} \widehat{\mathbf{K}}(\mu)_{ij} &= a(\mu; \phi_i, \phi_j) \\ \widehat{\mathbf{C}}(\mu, \widehat{\mathbf{v}})_{ij} &= b(\mu, \mathbf{v}^r; \theta_i, \phi_j) \\ \widehat{\mathbf{f}}(\mu)_i &= f(\mu; \phi_i) \\ \widehat{\mathbf{g}}(\mu, \widehat{\mathbf{v}})_i &= d(\mu, \mathbf{v}^r; \theta_i) \end{aligned}$$

with

$$\mathbf{v}^r = \Phi \widehat{\mathbf{v}}$$

It is evident that the reduced form (1.14) has the same structure as the discrete form generated using finite elements in (1.9), as both equations are built using a Galerkin formulation and by introducing discrete solution spaces, albeit of different dimensionality.

Non-negative nature of the dual RB  $\Theta$  forbids the use of orthogonal decompositions of Lagrange Multiplier snapshot matrix  $\Lambda$ . To this end, non-negativity preserving decompositions have been explored by various authors. Non-negative matrix factorization (NNMF) method [58] can be used to decompose a non-negative matrix, such as  $\Lambda$  into two low-rank non-negative matrices  $\mathbf{W}$ ,  $\mathbf{H}$ , such that  $\Lambda \approx \mathbf{WH}$ . The left-hand matrix  $\mathbf{W}$  is used as the dual RB in [51].

Another possibility is to directly use the snapshot vectors as the basis vectors, instead of computing the NNMF. This was the approach of [50], where the possibility of non-unique dual solution is also discussed, as snapshot vectors are not guaranteed to be linearly independent. Also, the number of dual dofs increases with the number of snapshots, preventing optimal reduction of the system. A greedy snapshot selection method based on the criteria of maximizing the volume of the reduced cone, namely the Cone-Projected Greedy (CPG) algorithm, was proposed in [13]. This algorithm creates a dual RB by greedily selecting snapshots from the snapshot matrix, attempting to create a more compact basis in comparison to using the full snapshot matrix. The algorithm is based on projection  $\Pi_{\Theta}$  of a vector  $\boldsymbol{\lambda}$  on a vector cone  $\Theta$ ,

$$\Pi_{\Theta}(\boldsymbol{\lambda}) := \Theta \boldsymbol{\alpha}, \quad \text{where } \boldsymbol{\alpha} \text{ is } \arg \min_{\boldsymbol{\gamma}} \|\boldsymbol{\lambda} - \Theta \boldsymbol{\gamma}\| \quad \forall \boldsymbol{\gamma} \geq \mathbf{0} \quad (1.15)$$

The greedy algorithm evaluates the error between each snapshot and its projection on the cone spanned by previously selected snapshots, and then adds that snapshot with

maximum cone projection error. The process continues until the cone projection error is within a set tolerance. [53] had also proposed a similar algorithm called Angle-Greedy algorithm, but the projection error is calculated based on computation of the angle between the candidate snapshot vector and the space, and not the cone, spanned by the previously selected snapshots i.e. it does not place non-negativity constrain on the coefficients  $\gamma$  in (1.15). This is not the best way of selecting snapshots for variational inequality problems since solutions must be sought in the reduced cone, and not the entire reduced space.

**Remark 1.1 : Constraint satisfying primal subspace in flow problems**

If we shift our gaze to other types of constrained ROMs, an interesting treatment of the mixed problems appears in incompressible flow problems. In such problems, the velocity snapshots are divergence-free and the divergence operator in the mass conservation equation is linear; offering the possibility to compute a divergence-free subspace for the velocity field. Consequently, the reduced incompressible flow problem is as an unconstrained problem, as all candidates in velocity subspace naturally satisfy the constraints [59, 60]. Apart from naturally satisfying the constraints, some elegant properties of vector algebra<sup>a</sup> result in disappearance of bilinear term that includes pressure.

Unfortunately, it is not straightforward to extend the same idea to contact mechanics problems (or variational inequalities), i.e. to compute a primal subspace that satisfies the non-penetration (or the inequality) constraints. This is because linear combinations of non-penetrating snapshots do not satisfy the non-penetration condition in general. As a consequence, the inequality constraints need be enforced explicitly while solving the reduced problem, in general.

<sup>a</sup>The product of a divergence-free velocity field and the gradient of pressure vanishes due to the orthogonality of the solenoidal and irrotational vector fields

**Construction of non-linear operators:** For an efficient ROM, the construction of nonlinear operators must also be cheap. The construction of operators  $\mathbf{C}(\mu, \hat{\mathbf{u}})$  and  $\mathbf{g}(\mu, \hat{\mathbf{u}})$  in each iterative step can prove expensive. Affine decompositions of the nonlinear constraint operators using the EIM are discussed in [13]. The affine form was achieved by computing the EIM based interpolation of the distance functions  $k(\mu, \mathbf{v}, \cdot)(\mathbf{x})$  and  $g(\mu, \mathbf{v})(\mathbf{x})$  (recall Notation 1.1):

$$k(\mu, \mathbf{v}; \cdot)(\mathbf{x}) \approx \sum_{p=1}^{N_k} \alpha_p^k(\mu, \mathbf{v}) \kappa_p(\cdot)(\mathbf{x})$$

$$g(\mu, \mathbf{v})(\mathbf{x}) \approx \sum_{q=1}^{N_g} \alpha_q^g(\mu, \mathbf{v}) \gamma_q(\mathbf{x})$$

where the interpolation is exact at the set of magic points  $\{\mathbf{x}_i^k\}$  and  $\{\mathbf{x}_i^g\}$ . Note that  $\kappa_p$  is linear w.r.t. its argument which can be any member in  $\mathcal{V}$ . Therefore, the coefficients

can be computed so that they satisfy the following conditions:

$$\begin{aligned} k(\mu, \mathbf{v}; \cdot)(\mathbf{x}_m^k) &= \sum_{i=1}^{N_k} \alpha_p^k(\mu, \mathbf{v}) \kappa_p(\cdot)(\mathbf{x}_m^k) \quad \forall m = \{1, 2, \dots, N_k\} \\ g(\mu, \mathbf{v})(\mathbf{x}_n^g) &= \sum_{q=1}^{N_g} \alpha_q^g(\mu, \mathbf{v}) \gamma_q(\mathbf{x}_n^g) \quad \forall n = \{1, 2, \dots, N_g\} \end{aligned} \quad (1.16)$$

Therefore, the distance function needs to be evaluated at most  $N_k + N_g$  points, instead of the whole contact surface. Also,  $\kappa_i(\cdot)$  and  $\gamma_i$  are computed just once using the training set and the corresponding discrete operators can also be built during the offline stage using the definition of weak forms in (1.5).

$$\begin{aligned} \widehat{\mathbf{C}}_{ij}^p &= \int_{\Gamma^2} \theta_i \kappa_p(\phi_j)(\mathbf{x}) \partial\Gamma \\ \widehat{\mathbf{g}}_i^q &= \int_{\Gamma^2} \theta_i \gamma_q(\mathbf{x}) \partial\Gamma \end{aligned} \quad (1.17)$$

where  $b_p$  and  $d_q$  correspond to the weak forms of individual EIM modes  $\kappa_p$  and  $\gamma_q$  that are independent of the state of the system. Thus, during the online stage, the non-linear operators can be approximated as:

$$\begin{aligned} \widehat{\mathbf{C}}(\mu, \mathbf{v}) &= \sum_{p=1}^{N_k} \alpha_p^k(\mu, \mathbf{v}) \widehat{\mathbf{C}}^p \\ \widehat{\mathbf{g}}(\mu, \mathbf{v}) &= \sum_{q=1}^{N_g} \alpha_q^g(\mu, \mathbf{v}) \widehat{\mathbf{g}}^q \end{aligned}$$

where only  $\alpha_p^k$  and  $\alpha_q^g$  are evaluated efficiently in online stage using (1.16).

Another approach for efficient computation of non-linear operators is the hyper-reduction approach [48]. In this approach, a small subset of the full domain  $\Omega_A \subset \Omega_1 \cup \Omega_2$  is computed that serves as the reduced integration domain (RID). The RID  $\Omega_A$  consists of the most ‘‘important’’ integration points computed by performing EIM on the displacement RB. Then, integration of the operators  $\widehat{\mathbf{C}}(\mu, \mathbf{v})$  and  $\widehat{\mathbf{g}}(\mu, \mathbf{v})$  is carried out on a subset of the potential contact surface, computed using the intersection of the RID and the contact surface  $\Omega_A \cap \Gamma_2$ . Thus, the distance functions are computed only on a small subset of the contact surface.

## 1.4 Scope of the thesis

The current literature consists of various methodologies to reduce the contact mechanics problem, especially to the Lagrange Multiplier approach. In this thesis, the following contributions are made:

- As seen in this chapter, the reducibility of contact problems has not been sufficiently investigated in current literature. This is particularly important as contact mechanics problems exhibit a local nature in terms of the contact zone. As the region of contact is dynamic for many contact problems, significant changes in the contact zone can prove consequential to reducibility. Therefore, in Chapter 2, we explore the validity of so-called *low-rank hypothesis* and robustness of the reduced models using various metrics. The chapter concludes that the contact pressure field lacks linear separability, an essential feature for reducibility.
- The lack of low-rank structure implies that the offline phase might be expensive as small training sets will be unable to explore the underlying structure sufficiently. Even if the user is willing to compute large training sets, it is still not possible to compute a robust low-rank subspace for the contact pressure. In Chapter 3, we focus on algorithms to handle dictionaries containing a large number of snapshots using ideas from sparse regression methods.
- As the linear inseparability of contact pressure field is established, we explore the idea of non-linear interpolations between contact pressure snapshots in Chapter 4. Interpolation is done in a warped space generated using Dynamic Time Warping algorithm. We show that reconstructions of contact pressure can be more reliably performed with smaller training sets, demonstrating the potential of non-linear dimensionality reduction methods for contact mechanics problems.

## 2 | Limitations of low-rank approach to contact problems

---

Strategies for reducing complex numerical models to facilitate building parametric models requires the acquisition of underlying physics using precomputed data. A brief overview of these methods was given in Section 1.2. Commonly used ROM methods are based on the paradigm of the so-called *linear dimensionality reduction*. The foundation of this framework rests on the low-rank hypothesis, which states that for a high-dimensional quantity of interest, often a representative low-dimensional space exists and can be extracted using linear transformations. The successful applications of ROM methods like POD, based on linear dimensionality reduction methods like PCA, to many parametric models can be attributed to the underlying low-dimensional behaviour of these systems.

In Sections 1.2 and 1.3, the wide variety of low-rank applications were mentioned, including constrained problems such as incompressible flows and contact mechanics problems. Reduced models for contact mechanics are important because the inequality constraints are geometrically non-linear, thereby introducing additional computational challenges (see Sections 1.1.2 to 1.1.3). However, the assumption of low-rank behaviour has implications on reduced models of contact mechanics problems, which are a class of variational inequality problems.

This chapter focuses on applicability of the low-rank hypothesis to the variational inequality problems, specifically contact mechanics problems. The effect of this hypothesis on various contact mechanics problems are demonstrated. A generic methodology of low-rank approach in contact mechanics problems was discussed in Section 1.3. In this chapter, the limitation of low-rank approach will be demonstrated on reduced models of various contact problems. It will be shown that the contact pressure field is linearly inseparable and therefore does not satisfy low-rank hypothesis. To endorse this argument quantitatively, the so-called *validation metrics* will be introduced to used to endorse the arguments about linear inseparability using specific numerical examples. Finally, con-

cluding arguments about linear inseparability and a few perspectives on circumventing the limitations due to lack of low-rank structure are given.

## 2.1 Low-rank contact problem using active set method

The non-linear reduced problem in (1.14) is based on the low-rank hypothesis, as it introduces the low-rank primal and dual bases. This approach is similar to [13, 50, 51]. Two sources of non-linearities exist in this problem: the resolution of active/inactive constraints and the dynamic contact pairs. Irrespective of the linearization algorithm utilized, both non-linearities depend on the current estimate of the displacement field at a given linearization step.

Here, (1.14) is solved using the fixed-point iterations and the active-set method applied to the reduced constraints. The linearized problem at a fixed-point iteration level  $p$  can be expressed as follows:

$$\begin{bmatrix} \widehat{\mathbf{K}} & \widehat{\mathbf{C}}_{\mathcal{A}}^T(\mu, \widehat{\mathbf{u}}^p) \\ \widehat{\mathbf{C}}_{\mathcal{A}}(\mu, \widehat{\mathbf{u}}^p) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{u}}^{p+1} \\ \widehat{\boldsymbol{\lambda}}_{\mathcal{A}}^{p+1} \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{f}} \\ \widehat{\mathbf{g}}_{\mathcal{A}}(\mu, \mathbf{u}^p) \end{bmatrix} \quad (2.1)$$

The subscript  $A$  indicates the active set of reduced contact constraints, i.e. the constraints that satisfy the equality and force the solution to lie on the boundary of the feasible region. Note that (2.1) is similar to the fixed-point system in (1.12). The fixed point algorithm stated in Alg. 2.1, was implemented in Python using NumPy package [61] for vectorial operations.

**Construction of operators:** The implication of dynamic contact pairing between two bodies of a contact problem is that the non-linear operators that enforce the non-penetration constraint must be updated for each new estimate of displacement field. On the other hand, for generating an efficient ROM, the construction of nonlinear operators must be inexpensive. This is not the case in Alg. 2.1, where the full order operators  $\mathbf{C}(\mu, \widehat{\mathbf{u}})$  and  $\mathbf{g}(\mu, \widehat{\mathbf{u}})$  are constructed in each iteration (Step 5). Affine decompositions of the nonlinear constraint operators using the EIM are discussed in [13]. EIM decomposition of the constraint operators further splits computational complexity into offline and online stages, by selecting the so-called “magic points” that are a small subset of all potential contact nodes where the nonlinear terms are evaluated in the online phase. The hyper-reduction method also permits computation of operators using fewer points [48]. This is crucial to solve problems where the contact pairs are strongly dependent on the displacement field. However, the efficient construction of constraint operators is not covered in this work, and the focus is maintained on the non-negative dual space and its low-rank approximation.

**Algorithm 2.1** Online phase

- 
- 1: Input: Queried value of parameter  $\mu$
  - 2: Given: Primal basis  $\Phi$  and dual basis  $\Theta$   
           Reduced operators  $\widehat{\mathbf{K}}, \widehat{\mathbf{f}}$  (possible to build offline)
  - 3: Initiate boolean array  $\mathcal{A}$  with one random element set to True.
  - 4: **while**  $\widehat{\mathbf{u}}$  not converge **do**
  - 5:     Build constraint operators  $\mathbf{C}(\mu, \widehat{\mathbf{u}}^p), \mathbf{g}(\mu, \widehat{\mathbf{u}}^p)$  using FEM
  - 6:     Project constraint operators on RBs  
           
$$\widehat{\mathbf{C}}(\mu, \widehat{\mathbf{u}}^p) = \Theta^T \mathbf{C}(\mu, \widehat{\mathbf{u}}^p) \Phi$$

$$\widehat{\mathbf{g}}(\mu, \widehat{\mathbf{u}}^p) = \Theta^T \mathbf{g}(\mu, \widehat{\mathbf{u}}^p)$$
  - 7:     Filter-out rows that are not in active set:  
           
$$\widehat{\mathbf{C}}_{\mathcal{A}}(\mu, \widehat{\mathbf{u}}^p) = \widehat{\mathbf{C}}(\mu, \widehat{\mathbf{u}}^p)[\mathcal{A}, :]$$

$$\widehat{\mathbf{g}}_{\mathcal{A}}(\mu, \widehat{\mathbf{u}}^p) = \widehat{\mathbf{g}}(\mu, \widehat{\mathbf{u}}^p)[\mathcal{A}]$$
  - 8:     Solve system (2.1)
  - 9:     Set  $\widehat{\boldsymbol{\lambda}}^{p+1}[\mathcal{A}] \leftarrow \widehat{\boldsymbol{\lambda}}_{\mathcal{A}}^{p+1}$  and  $\widehat{\boldsymbol{\lambda}}^{p+1}[\mathcal{A}^c] \leftarrow \mathbf{0}$
  - 10:    Update active constraints set  
           
$$\mathcal{A}[i] = \begin{cases} \text{False} & \text{if } \widehat{\lambda}_i^{p+1} < 0 \\ \text{True} & \text{if } (\widehat{\mathbf{C}}\widehat{\mathbf{u}}^{p+1} - \widehat{\mathbf{g}} \geq 0)_i \end{cases} \quad \triangleright \text{violations of (1.14b)}$$
           and (1.14c)
  - 11: **end while**
  - 12: Reconstruct  $\mathbf{u} = \Phi \widehat{\mathbf{u}}$  and  $\boldsymbol{\lambda} = \Theta \widehat{\boldsymbol{\lambda}}$
  - 13: Output:  $\mathbf{u}, \boldsymbol{\lambda}$
- 

## 2.2 Illustrative case

In this section, the Type-1 problem of a 1D elastic rope-obstacle problem from [50–52] is considered. The geometry of the problem enjoys further simplification of static contact pairs. In other words, each point on the elastic rope can come in contact with a unique point on the obstacle, thereby stripping the distance functions  $k$  and  $g$  of their non-linear dependence on the displacement field  $u$ .

The parametrized model of the elastic rope with obstacle function can be expressed as:

$$\begin{aligned} \nu \nabla^2 u(x) &= f & \text{on } x \in [0, 1] \\ u(0) &= u(1) = 0 \\ u(x) &\geq g(x, \gamma) & \text{on } \gamma \in [-0.5, 0.5] \end{aligned} \quad (2.2)$$

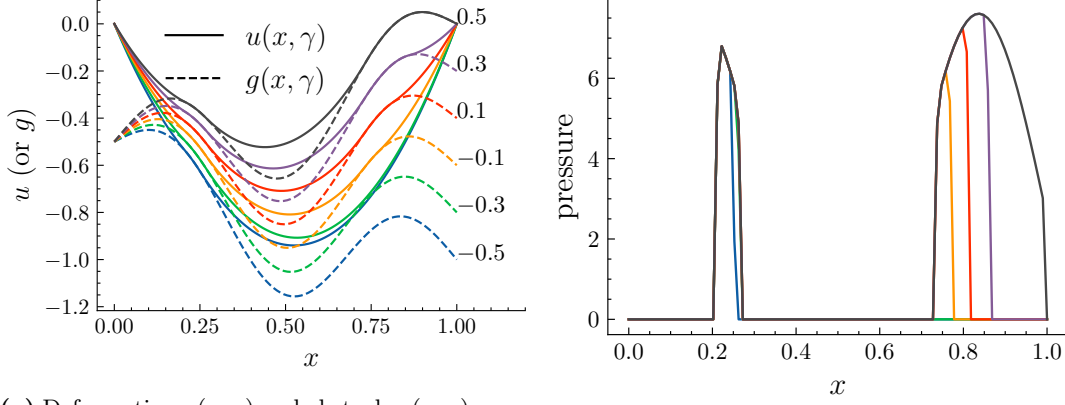
where the parameterized obstacle function is defined as:

$$g(x, \gamma) = -0.2(\sin(\pi x) - \sin(3\pi x)) - 0.5 + \gamma x \quad (2.3)$$

and the quantities  $\nu = 30$  and  $f = 250$  are independent of the parameter  $\gamma$ .

Snapshots are generated in the training set,  $\mathcal{P}_{\text{tr}} \in \mathcal{P}$  consisting of 10 equidistant points in the parametric space  $\mathcal{P}$  defined in (2.3). The deformation and contact pressure

snapshots are shown in Fig. 2.1. The primal RB is built using the POD approach and dual RB is created using the cone-projected greedy algorithm of [13].



(a) Deformation  $u(x, \gamma)$  and obstacle  $g(x, \gamma)$  snapshots

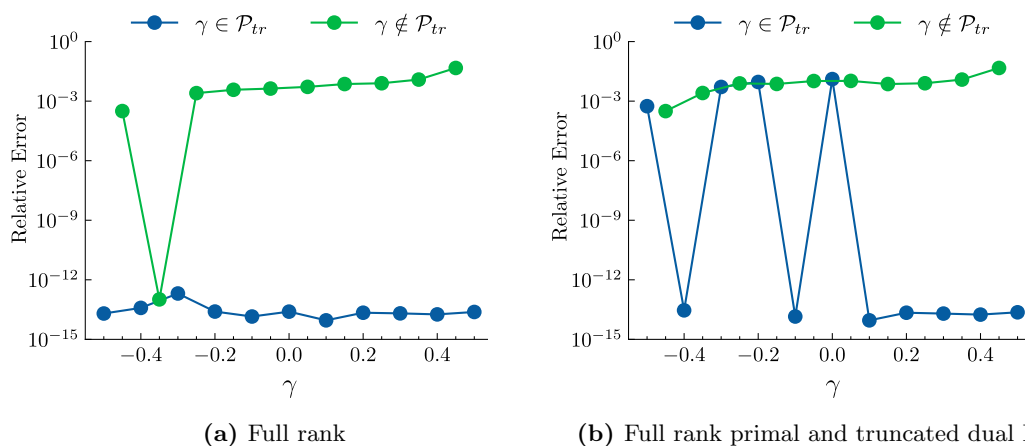
(b) Contact pressure snapshots

**Figure 2.1:** Sample snapshots for elastic rope-obstacle problem with  $\gamma \in \mathcal{P}_{\text{tr}} \subset [-0.5, 0.5]$ .

The reconstruction errors are shown in Fig. 2.2 for the two cases of full and truncated dual RB. As expected, the points in the training set show very low reconstruction error, and the points outside the training set show a moderate error. The only exception at  $\gamma = -0.35$  where reconstruction with full RB is highly accurate because the contact area does not change in the regime  $\gamma \in [-0.4, -0.3]$  and the training set has two snapshots in this region. When dual RB is truncated, some points in the training set achieve the same error level as that of points outside the training set. These points correspond to the same dual snapshots that were eliminated by the truncation procedure of the CPG algorithm. Another relevant observation is that the number of active dofs of the reduced dual solution  $\hat{\lambda}$  is small for all reconstruction cases. For points in the training set,  $\hat{\lambda}$  has exactly 1 active dof, whereas for the points outside the training set, it has a maximum of 2 active dofs. This is due to the fact that the dual basis is composed primarily of snapshots and does not undergo any compression like in the case of the primal basis generated by POD.

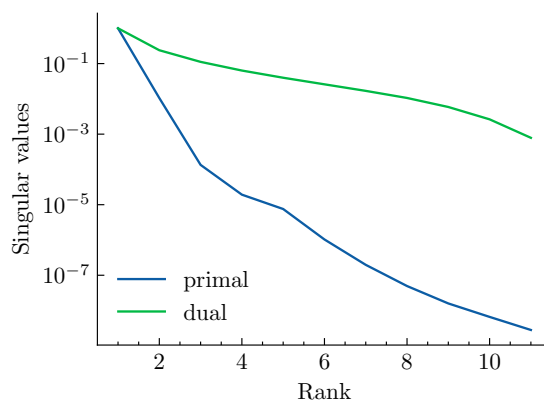
		Primal basis (POD)	Dual basis (CP-greedy)
Full	Tolerance	$\varepsilon$	$\varepsilon$
	Rank	11	11
Truncated	Tolerance	$\varepsilon$	$2 \times 10^{-1}$
	Rank	11	7

**Table 2.1:** Truncation tolerances and ranks for both bases for the rope-obstacle problem.  $\varepsilon$  indicates numerical precision



**Figure 2.2:** Primal reconstruction ( $\mathcal{H}^1$ ) errors for points in and outside the training set for the rope-obstacle problem.

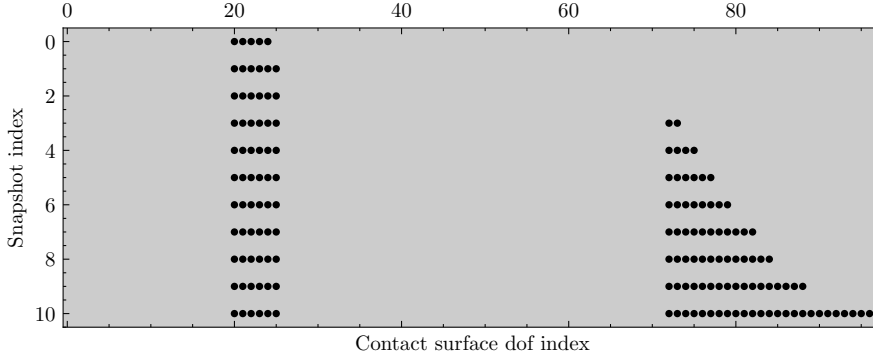
In general, any efficiently generated reduced basis has its vectors arranged in order of importance. In case the basis consists of left singular vectors, generated by SVD, the decreasing importance of basis vectors is indicated by the singular values. This is a very useful indication of the rank of the subspace. Even though the dual basis is not computed using SVD, it can still be useful to compute singular values and assess their decay. This evolution of singular values, seen in Fig. 2.3, shows that the decay is slow compared to the primal variable.



**Figure 2.3:** Decay of singular values (cumulative and normalized) for the training set of the elastic rope-obstacle problem. A similar figure is also given in Reference [48, Figure 4]

Also, the discussions in previous sections highlight the fact that dual basis  $\Theta$  cannot be truncated without a significant loss of accuracy. This is also evident from Table 2.1, even a truncation tolerance as high as 0.2 leads to truncation of just four vectors in dual RB.

The above observations are explainable by the hypothesis that the dual snapshots are highly “inseparable” and do not lie in a low-dimensional subcone. Each contact pressure snapshot in Fig. 2.1 is linearly independent of the other due to difference in the contact area. Hence, if any snapshot is picked randomly from this figure, it cannot be reasonably approximated by any linear combination of the rest of the snapshots. In other words, each snapshot vanishes outside the contact zone, and this zone is almost “unique” for each snapshot. For a particular snapshot to be reasonably approximated, other snapshots with similar contact zones are needed. Therefore, the elimination of snapshots from the dual basis leads to a drastic increase in the reconstruction error for the corresponding parametric values in the training set. The dual variable, therefore, not only needs special treatment due to its positivity constraints but is also highly inseparable because of sensitivity to contact position and area. The inseparability is also evident in the non-zero pattern (“spy” plot) of dual snapshots visualized in Fig. 2.4, where most rows have an almost unique sparsity pattern. It is unlikely that a subspace whose members show varying sparsity patterns will admit a low-dimensional behaviour.



**Figure 2.4:** Sparsity pattern of dual snapshots for the rope-obstacle problem. Each row corresponds to a snapshot

More numerical examples are explored to study the lack of low-rank behaviour in the Section 2.4.

### 2.3 Metrics for low-rank methods

In this section, we introduce the validation metrics [62], that provide a quantitative evaluation of the robustness of low-rank models. Validation metrics defined in [62] identify three different measures for this purpose:

- *Compactness*: As the name suggests, this metric is a measure of dimensionality of the reduced space. Reference [62] defines compactness as the squared sum of the first  $m$  singular values of the snapshot matrix. This metric is similar to the decay of singular values studied in the previous section in Fig. 2.3. However, an alternative definition given by (2.4) is used here. The new definition allows

extending the concept of compactness to CPG reduced bases, where equivalents of singular values are undefined. Also, in case of an orthogonal basis, this definition is equivalent to the normalized singular values used in Fig. 2.3.

$$C(m) := \frac{\|\mathbf{S}_{\text{tr}} - \Pi_{\Psi[:m]}(\mathbf{S}_{\text{tr}})\|_F}{\|\mathbf{S}_{\text{tr}}\|_F} \quad (2.4)$$

where  $\mathbf{S}_{\text{tr}}$  is a matrix containing (either primal or dual) snapshots in  $\mathcal{P}_{\text{tr}}$  arranged column-wise.  $\Pi_{\Psi[:m]}(\mathbf{S}_{\text{tr}})$  is a projection operators that projects each column in  $\mathbf{S}_{\text{tr}}$  on the first  $m$  vectors of the basis  $\Psi$  (indicated in Python notation by  $\Psi[:m]$ ). The snapshot matrix  $\mathbf{S}_{\text{tr}}$ , projection operator  $\Pi$  and the basis  $\Psi$  are generic symbols and are used in the following combinations, indicated by the following labels:

Label in figures	Snapshots in $\mathbf{S}_{\text{tr}}$	Projection operator $\Pi$	Basis $\Psi$
primal orth	Primal	Orthogonal	POD primal basis $\Phi$
dual orth	Dual	Orthogonal	POD dual basis $\Upsilon$
dual cone	Dual	Cone	CPG dual basis $\Theta$

The Frobenius norm is chosen in (2.4), because it is a natural choice for a POD/SVD basis that contains the best rank-wise approximations in the sense of Frobenius norm, but other norms can be also be used. This is actually equivalent to the original definition of compactness in [62], because in the case  $\Psi$  consists of the left singular vectors, compactness corresponds to cumulative energy of singular values.

For the dual field, the metric is given using two kinds of projections: orthogonal and CPG. Though the dual CPG metric is more relevant to the inequality constrained problems, the dual orthogonal metric is also included as it is more intuitive due to correspondence with the singular values.

- *Generalization Ability:* This metric measures the ability of the reduced basis to approximate parametric instances that are not in the training set. It is computed by using a leave-one-out approach on the training set to generate a  $m$ -rank basis. Then, the reconstruction error of the eliminated snapshot is used to compute generalization ability as given by (2.5),

$$G(m) = \frac{1}{N} \sum_{i=1}^N \frac{\|\mathbf{s}_i - \Pi_{\Psi(\mathbf{S}_{\text{tr}} \setminus \mathbf{s}_i)[:m]}(\mathbf{s}_i)\|}{\|\mathbf{s}_i\|} \quad (2.5)$$

where  $\Psi(\mathbf{S}_{\text{tr}} \setminus \mathbf{s}_i)$  is the basis created using the snapshot subset  $\mathbf{S}_{\text{tr}}$  with  $\mathbf{s}_i$  removed (i.e.  $\mathbf{S}_{\text{tr}} \setminus \mathbf{s}_i$ ). The symbols  $\mathbf{S}_{\text{tr}}$ ,  $\Pi$  and  $\Psi$  follow the definition in (2.4).

- *Specificity:* The subspace spanned by the reduced basis is expected to contain elements similar to that of the training set. This metric measures the extent of dissimilarity between randomly picked elements from the subspace and the snapshots in the training set. To compute specificity, a set of random shapes in the associated subspace are generated using random coefficients  $R = \{\Psi \alpha_i\}_{i=1}^N$ , where

each vector  $\alpha_i$  is drawn randomly within a predefined range, and compared to the closest snapshot in the training set. The mean value of this error is defined as Specificity.

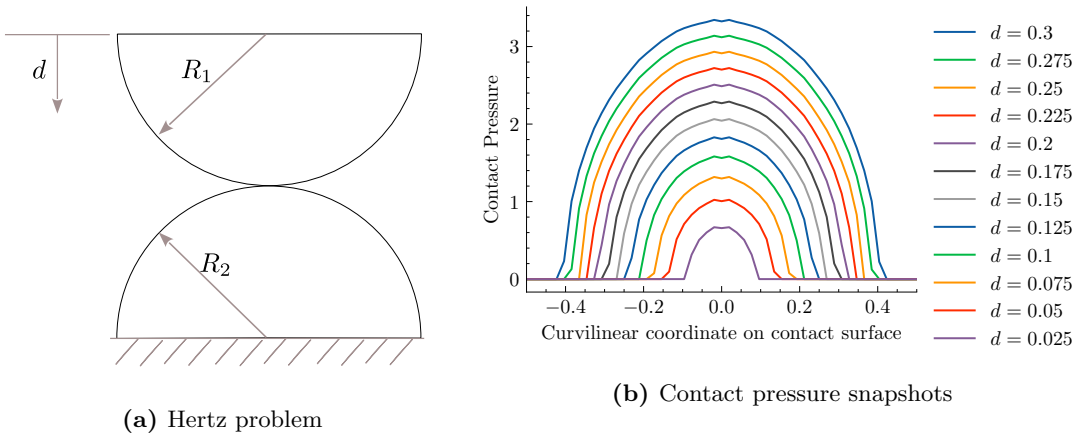
$$S(m) = \frac{1}{N} \sum_{i=1}^N \min_{s_j \in \mathbf{S}_{\text{tr}}} \frac{\|s_j - \Psi[:, m] \alpha_i\|}{\|s_j\|} \quad (2.6)$$

The vectors  $\{\alpha_i \mid \alpha_i \in \mathbb{R}^m\}$  are drawn randomly in the range defined by  $\mu_\alpha \pm \sigma_\alpha \subset \mathbb{R}^m$ . The vectors  $\mu_\alpha$  and  $\sigma_\alpha$  contain the element-wise mean and standard deviation of vectors  $\{\gamma_k\}_{k=1}^m$ , where  $\gamma_k$  is the reduced coordinate of the training set snapshot  $k$ . (e.g. if  $\Psi$  is an orthogonal basis, the vectors  $\gamma_k$  are the columns of matrix  $\Gamma := \Psi^T \mathbf{S}_{\text{tr}}$ ). In the specific case where the basis  $\Psi$  is the  $\Theta$ , the CPG basis, only non-negative entries are allowed in  $\{\alpha_i\}$ . The symbols  $\mathbf{S}_{\text{tr}}$ ,  $\Psi$  follows the definition in (2.4).

## 2.4 Numerical Examples

### 2.4.1 Hertz problem

This section deals with the reduced model of a Type-2 contact problem, namely the Hertz contact problem, of two half-cylinders loaded against each other, as shown in Fig. 2.5a. A *loading parametrization* problem is solved, where the imposed displacement  $d \in (0, 0.3)$  is the parametric space, with  $R_1 = R_2 = 1.0$ . The contact pressure snapshots for various imposed displacement values are shown in Fig. 2.5b.



**Figure 2.5:** Hertz problem: two half cylinders loaded against each other and the resulting contact pressure snapshots. Displacement  $d$ , imposed on the top cylinder, is treated as parameter in the reduced model.

**High-fidelity model:** A finite element model of the geometry in Fig. 2.5a is created using a quad mesh. Each half-cylinder has 513 nodes and 466 elements of which 78

elements lie of the potential contact surface i.e. the semi-circular edges. Linear shape functions are used to discretize the displacement field, while contact pressure is discretized using piece-wise constant shape functions centred at the surface nodes of the quad mesh (collocation method, as described in [13]). For surface integrals, a single-point gauss quadrature centred on the node is used. This formulation gives mostly smooth contact pressure profiles, except near the peak pressure.

**Reduced Model:** Snapshots are generated in the training set,  $\mathcal{P}_{\text{tr}} \in \mathcal{P}$  consisting of 12 equidistant points in the parametric space. The full rank and truncated dual RB (using CPG) is considered, as given in Table 2.2, whereas the primal RB is not truncated. The full RB has 12 basis vectors for both primal and dual fields, whereas the truncated dual RB has 8 vectors. Note that dual RB truncates only 4 out of 12 basis vectors for a high truncation tolerance of 0.05.

The reconstruction errors are shown in Fig. 2.6 for full and truncated dual RBs. As expected, the points in the training set show very low reconstruction error, and the points outside the training set show a relatively high error. When the dual RB is truncated, points corresponding to truncated dual snapshots have a high error level, in the same order of error for points outside the training set. Like the illustrative case in Section 2.2, the non-zero pattern in snapshots in Fig. 2.7 also show a unique sparsity pattern for each snapshot. Also, the compactness shown in Fig. 2.8 displays a slow decay of the dual orth and dual cone, indicating a high-rank behaviour of the dual variable. The slightly slower dual cone curve than the dual orth was expected because of the non-negativity constraint that appears in approximation using the cone of the dual basis. These observations again reinforce the proposition that the dual basis does not admit a low-rank behaviour.

		Primal basis (POD)	Dual basis (CP-greedy)
Full	Tolerance	$\varepsilon$	$\varepsilon$
	Rank	12	12
Truncated	Tolerance	$\varepsilon$	$5 \times 10^{-2}$
	Rank	12	8

**Table 2.2:** Truncation tolerances and ranks for both bases for Hertz problem.  $\varepsilon$  indicates numerical precision

Generalization ability and Specificity metrics introduced in Section 2.3 are plotted in Fig. 2.9. The generalization ability for primal field is much better than the dual field, another reflection of separability issues with the dual field. On the other hand, primal and dual specificity are about the same order, though the primal specificity is consistently lower.

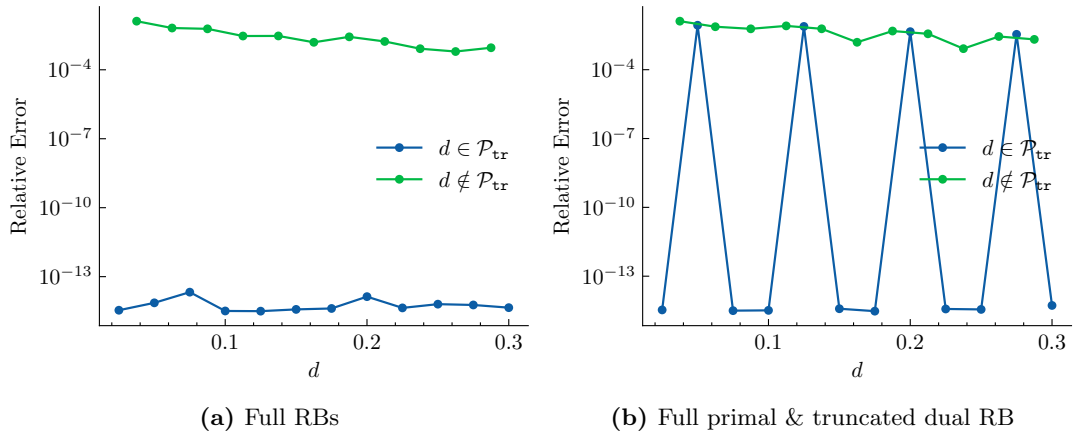


Figure 2.6: Primal reconstruction ( $\mathcal{H}^1$ ) errors for the Hertz problem.

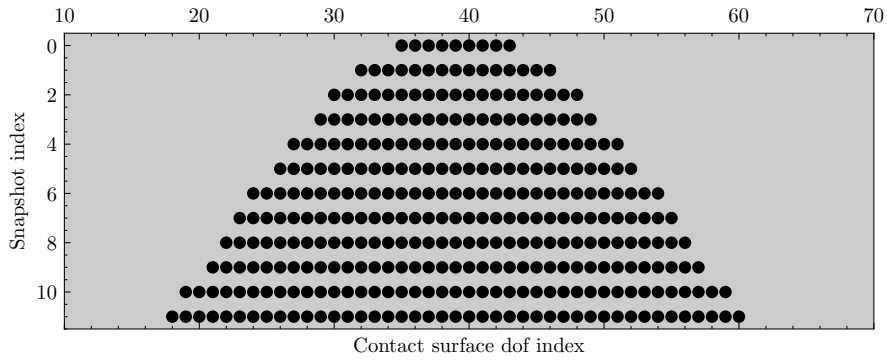


Figure 2.7: Sparsity pattern of dual snapshots for Hertz problem

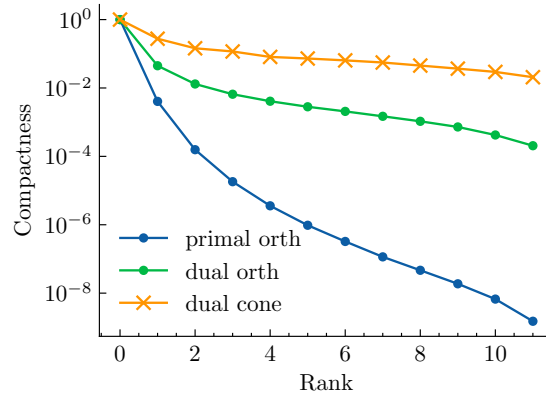
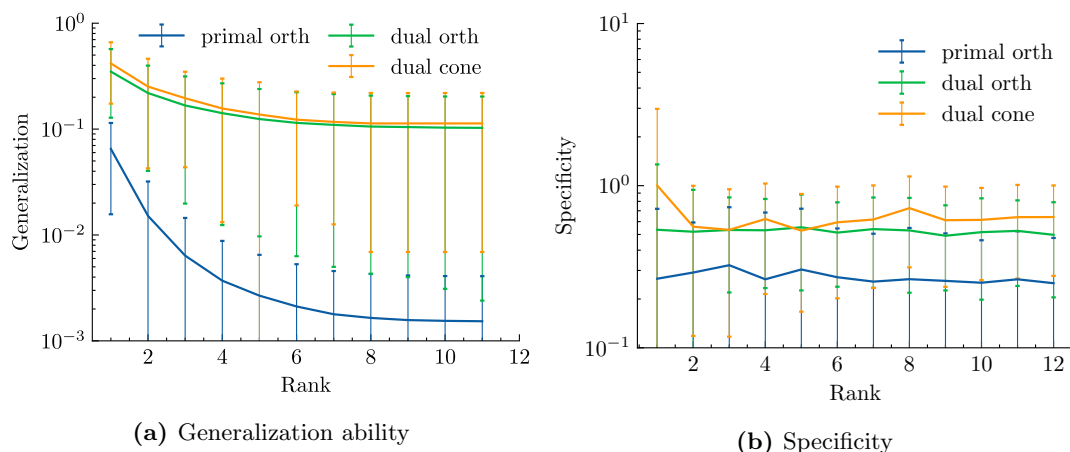


Figure 2.8: Compactness of reduced bases for Hertz problem

### 2.4.2 Ironing problem

A Type-2 problem where the contact area changes quite significantly, namely the ironing problem [11], is considered. The ironing problem consists of an iron block pressed against



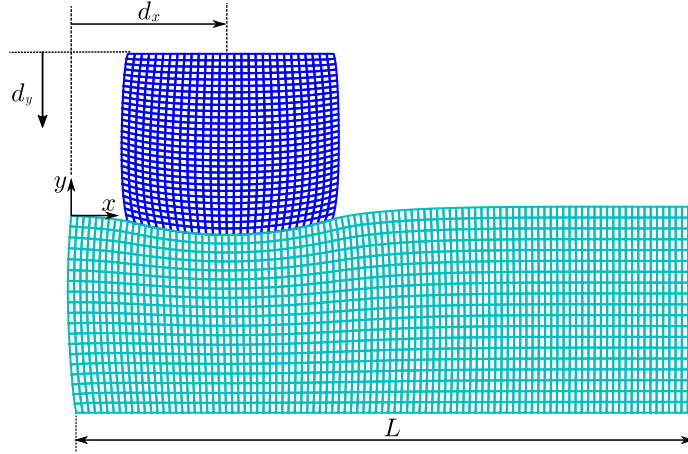
**Figure 2.9:** Validation metrics for Hertz problem. The bars indicate  $\pm 1\sigma$  interval of corresponding metrics

a flat slab and moved along the length of the slab (Fig. 2.10). The problem is simplified with two more assumptions: the first is that iron moves slowly enough that the problem can be considered to be *quasi-static* and the second is that surfaces are frictionless. The horizontal position  $d_x$  of the iron is taken as the parameter for the reduced model. In this problem, as the potential contact surface on the slab is quite larger than the actual position of contact, the contact pressure snapshots display large changes in contact position, shown in Fig. 2.11. The same is reflected in **spy** pattern of the contact pressure snapshot matrix, shown in Fig. 2.12

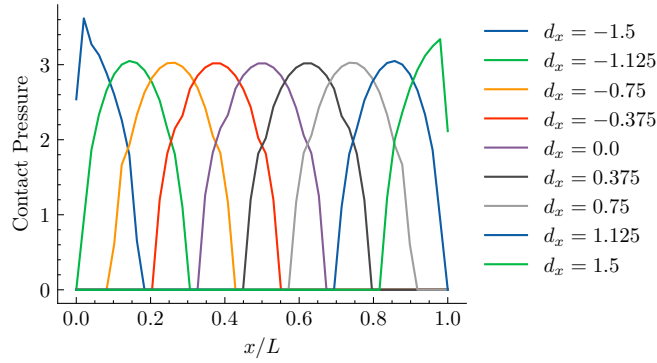
**Details of FE model:** A finite element model of the ironing problem is created using a structured quad mesh for both the iron and the slab. Like the Hertz problem, the displacement field is approximated using linear shape functions and contact pressure using piecewise-constant shape functions centred at the surface nodes. To better demonstrate the inseparability issues, two meshes: coarse and fine, are considered for iron and slab.

	Iron	Slab
Coarse Mesh	$30 \times 30$	$20 \times 100$
Fine Mesh	$60 \times 60$	$40 \times 200$

The compactness metric, which is equivalent to decay of truncation error as discussed before, is shown in Fig. 2.13 for a coarse and fine mesh, created using 128 snapshots. The contrasting part of the two graphs is the green curve representing the dual orth compactness, as this curve vanishes at rank 100 for the coarse mesh, but not for the fine mesh. This is natural, since coarse mesh has only 100 nodes on the surface of the slab and thus allowing a maximum rank of 100, whereas the fine mesh has around 200 nodes. Since the orthogonally computed dual (dual orth) compactness is also a measure for determining the true rank of the snapshot matrix, the green curve naturally vanishes at a rank of 100. However, the most interesting part of this figure is the blue curve



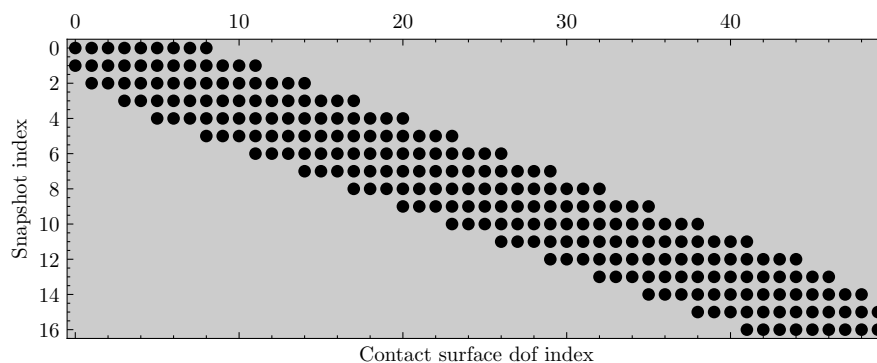
**Figure 2.10:** Ironing problem: Iron block is pressed against the flat slab by a displacement  $d_y$  and moved horizontally. The horizontal displacement of the iron block  $d_x \in [0, L]$  is treated as the parameter in the reduced model



**Figure 2.11:** Sample snapshots of contact pressure for ironing problem

representing the truncation computed using cone-projection. Even for the coarse mesh, whose true rank is bounded by the number of surface nodes, i.e. 100, the dual cone compactness doesn't decay well beyond this rank. This is another interesting behaviour of the low-rank approach since the dual solution is sought in a subcone that involves positivity constraints, further degrading compactness.

Apart from compactness, another way to demonstrate the dimensionality of the dual subcone is by using the projection error of nested training sets, which will be referred to as nested error subsequently. To compute the nested error, snapshots are computed in a nested set of points in the parametric space, so that  $n$ -th level training set is a subset of the  $(n + 1)$ -th level training set. In each nested level,  $2^n + 1$  points are uniformly distributed in the parametric space  $[0, L]$ . This means, every alternate point in  $(n + 1)$ -th level is the mid-point of two consecutive points in  $n$ -th level. The nested error  $H(n)$  can

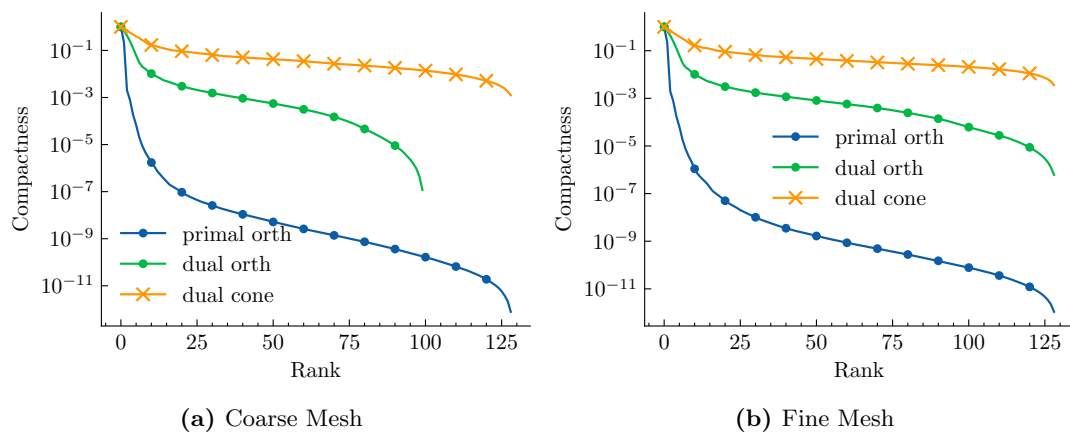


**Figure 2.12:** A typical sparsity pattern of dual snapshots for ironing problem

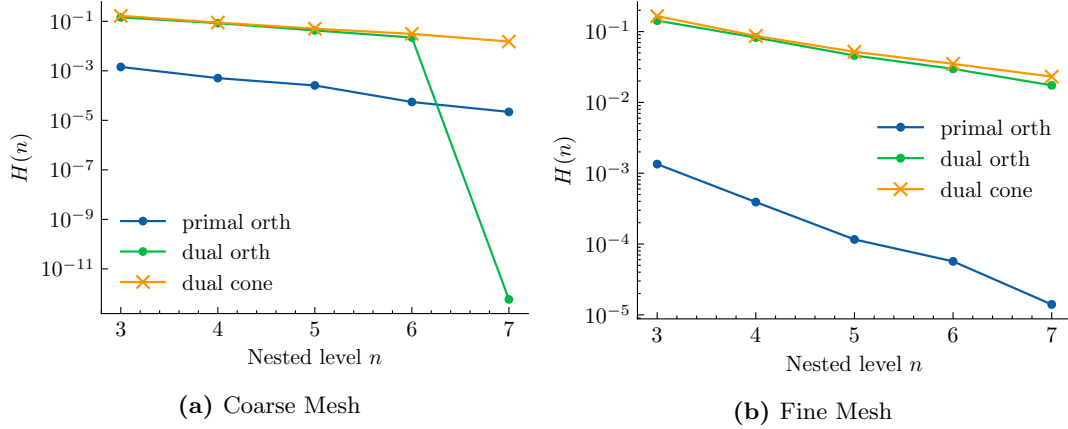
then be defined as:

$$H(n) = \frac{\|\mathbf{S}_{\text{tr}}^{n+1} - \Pi_{\Psi^n}(\mathbf{S}_{\text{tr}}^{n+1})\|_F}{\|\mathbf{S}_{\text{tr}}^{n+1}\|_F} \quad (2.7)$$

where the symbols  $\Psi$  and  $\mathbf{S}_{\text{tr}}$  carry same meaning defined in Section 2.3.  $\|\cdot\|_F$  indicates the Frobenius norm. The RB computed using the  $n$ -th level training set is indicated by  $\Psi^n$ , and snapshot matrix for  $(n+1)$ -th level is indicated by  $\mathbf{S}_{\text{tr}}^{n+1}$ . Nested training sets where each level contains nearly twice the number of points compared to previous level are created and the dual RB for each level is computed. The slope of nested errors computed using orthogonal and cone projections indicate that the offline stage will require a very large number of snapshots to explore the dual subcone, which is another indication of its high-dimensionality, as shown in Fig. 2.14. The effects of mesh size are similar to those seen in compactness, as nested error using dual cone projection does not decay even after training set is full rank in case of coarse mesh.

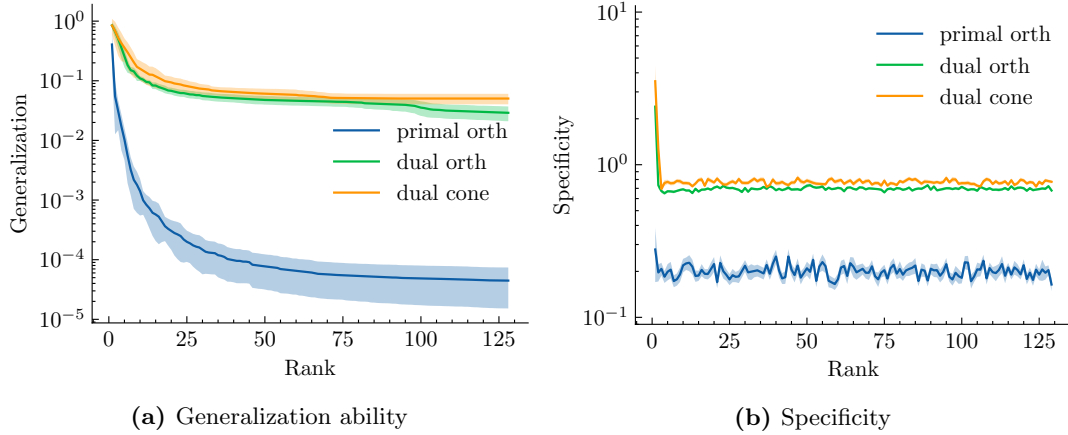


**Figure 2.13:** Compactness for reduced bases of ironing problem.



**Figure 2.14:** Projection error of nested level  $n + 1$  snapshots on dual RB of level  $n$ .

Generalization ability and Specificity metrics are plotted in Fig. 2.15. The generalization ability shows a similar trend as that of nested error, as they are similar quantities measuring reconstruction errors outside the training set. The primal basis outperforms its dual counterparts in this metric, as expected. The specificity metric is quite higher for dual quantities than the primal ones, which is because random shapes generated using the RB are quite different.



**Figure 2.15:** Validation metrics for the ironing problem (for fine mesh). The fill regions indicate  $\pm 1\sigma$  interval of corresponding metrics.

## 2.5 Conclusions and Perspectives

In this chapter, the limitations of low-rank ROMs to contact mechanics problems were discussed. Demonstrations of this approach show moderate reconstruction errors for parametric points outside the training set. To explore the possible sources of this error,

---

we focus on the key assumption of the low-rank hypothesis i.e. the solution subspace is low-dimensional. It turns out that the Lagrange Multiplier associated with the inequality constraints, i.e. the contact pressure, does not admit a low-rank subspace. Qualitative and quantitative assessments are provided to support this argument. The linear inseparability of dual subspace/subcone is demonstrated using various metrics that show high-dimensionality of contact pressure compared to the relatively low-dimensionality of the displacement field.

To circumvent the limitations due to the lack of low-rank structure, several potential approaches need to be explored. One possibility is the usage of dictionary-based approximation with a relatively large number of snapshots spanning the dual subspace efficiently. Another possibility is resorting to either a non-linear transformation; potentially transforming the dual snapshots in a way that they lie in a low-dimensional subspace.



# 3 | Dictionary approach in contact mechanics problems

---

## 3.1 Motivation

The contact mechanics problem has a feature, the contact zone, that exhibits a local nature. The contact zone is a function of the state of the system, and hence, may strongly depend on the loading, geometry, and other physical parameters. In general, for variational inequalities, only a subset of the constraints are usually active. An inequality constraint is referred to as active only if the current state of the problem satisfies the equality. Due to the last KKT condition in (1.14d), the Lagrange multipliers  $\lambda$  assume a non-zero value only for active constraints i.e. the part of surface where contact is established. The strong dependence of the active zone on the state of the system leads to the linear inseparability of the Lagrange multiplier field. It has been shown in Chapter 2, that the local characteristics of contact pressure can render subspace learning ineffective and cause difficulties in creating efficient reduced models of contact problems. The separability issues may cause the following problems:

1. **Complexity of subspace learning:** It is desirable to keep the complexity of the offline stage at affordable levels. Usually, this is achievable because a relatively low number of snapshots compared to the dofs of the high-fidelity problem are sufficient to create a reliable reduced basis, thanks to the low-rank nature of many resolution problems. However, given the inseparability of the contact pressure field, a large training set may be necessary to discover the entire solution subspace to a reasonable accuracy.
2. **Lack of low-rank subspace:** Discovery of a low-rank subspace is key to creating a low-cost reduced model, as the number of reduced dofs scales with the rank of the discovered solution subspace. While low-rank methods for contact problems [13, 51] define a contact pressure subcone, it is shown in Chapter 2 using various metrics that compression of contact pressure data into a low-rank subcone or subspace leads to loss of useful information. Therefore, despite the creation of a large training

set, as mentioned in the previous point, extracting a low-rank subspace might not be possible.

The degree of these effects depends on the nature of the contact problem. In this chapter, the approximation using a dictionary instead of a low-rank subcone is explored, addressing the second point as the dictionary is made of snapshots without applying any compression, like NNMF or CPG. In effect, the high complexity of the offline stage is accepted as a compromise for creating a reliable reduced model. The ideas behind over-complete dictionaries are expanded in the coming sections.

## 3.2 Sparse methods with over-complete dictionaries

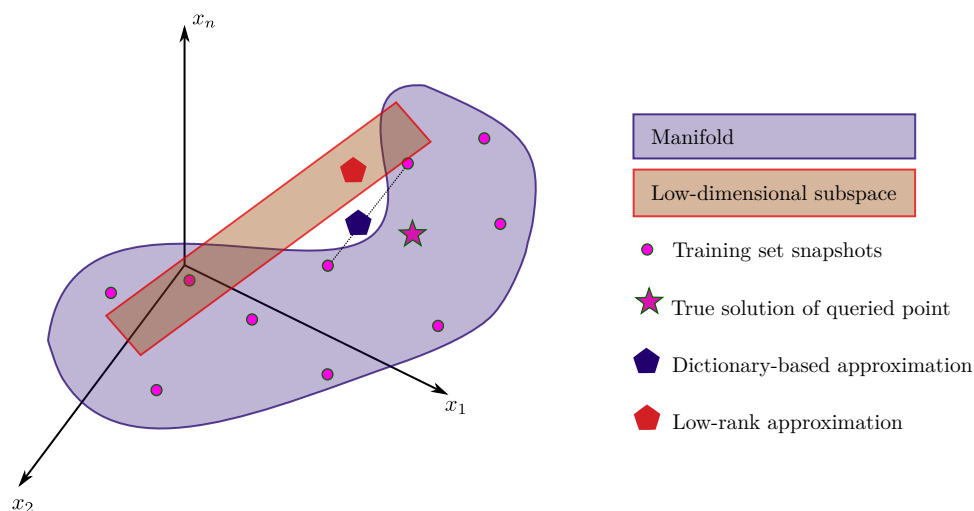
The primary motivation for using a dictionary-based approximation is to mitigate the limitations due to linear inseparability, discussed in Section 3.1. The word “mitigate” is explicitly used here to indicate that the strategy developed in this chapter does not resolve the inseparability issue at a fundamental level, rather it attempts to limit the effects of inseparability. The intention is to accommodate the cost of computing snapshots over a large training set and focus mainly on addressing the point 2 of Section 3.1 using an algorithm that can handle large dictionaries efficiently. A matrix containing the large set of snapshots is referred to as an over-complete dictionary  $\mathbf{D}$ . The term “over-complete” is used loosely in this context, as it is impossible to sufficiently sample a truly inseparable subspace for a given approximation error target and is simply used to imply that the dictionary has numerous snapshots, possibly larger than the number of contact dofs in the high-fidelity problem.

The dictionary-based approximation can be visualized in an abstract  $n$ -dimensional space in Fig. 3.1. Typically, in cases where low-rank hypothesis is not satisfied, the underlying physics of the problem assume a manifold structure. A manifold can be thought of as a low-dimensional entity that does not satisfy the properties of a linear space. The idea expressed in this figure is that a manifold, like a hypersurface with non-zero curvature, can be approximated better using a large cloud of points lying on it. If the point cloud is rich enough, it is more likely that this approximation will be better than a low-dimensional subspace which is a hyperplane<sup>1</sup>. It is this large cloud of points that is referred to as over-complete dictionary in this chapter.

In dictionary-based approximation, unlike the low-rank approach [13, 51], no compression of the contact pressure snapshots is applied. The dual solution can simply be restricted to the column-cone<sup>2</sup> of the dictionary  $\mathbf{D}$ , which might be reasonably rich to approximate the true contact pressure subcone. In the online phase, only a few snapshots will be selected to estimate the contact pressure field. This is justified because not all information in the dictionary is necessary for the reconstruction of a particular instance and hence, the dual reduced dofs will admit a sparse solution. This feature is also seen

<sup>1</sup>The terms hyperplane and hypersurface refer to a plane and surface in dimensions higher than 3 respectively.

<sup>2</sup>column-cone is the subset of the column-space restricted to non-negative coefficients.



**Figure 3.1:** Illustration of low-rank vs. dictionary-based approximation. Snapshots (circles) in an  $n$ -dimensional system lie on a manifold of dimension lower than  $n$ . The low-dimensional subspace computed using low-rank methods has a dimension  $\ll n$ . A dictionary-based approximation of the queried point (star-shaped) can be computed as linear combination of the nearest snapshots (shown by a line connecting the nearest candidates). The low-rank approximation is simply the projection of the queried point on the low-dimensional subspace.

in case of the low-rank model generated using snapshots selected by CPG. Therefore, enough motivation exists to use sparsity enforcing methods to choose a few columns of a dictionary  $\mathbf{D}$  to approximate the dual solution. Keeping these ideas in mind, sparse regression techniques are briefly introduced at this stage and a few of these methods from current literature are discussed. Following this, their application to resolution problems will be briefly discussed.

### 3.2.1 Sparse regression in data approximation problems

Sparse regression techniques have been under development in recent decades across various fields, especially in signal processing and statistics. The main idea behind these techniques is to approximate a signal with the few most suitable elements of a precomputed dictionary of signals [63], or to develop the simplest predictive model [64]. A typical sparse regression can be written as the following optimization statement:

$$\begin{aligned} \min & \|\boldsymbol{\alpha}\|_p \\ \text{s.t.} & \|\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}\| < \varepsilon \end{aligned} \quad (3.1)$$

where  $x$  is the signal being approximated,  $\mathbf{D}$  could either be a dictionary of signals or a basis that possibly admits sparse representation,  $\|\cdot\|_p$  is a sparsity inducing norm, and  $\varepsilon$  is the tolerance on the approximation error. It is also possible to write other forms of sparse regression problems where the approximation error is minimized and  $p$ -norm

term is constrained to stay under a specific tolerance of sparsity, or the minimization of a penalty form with a weighted sum of the two terms.

It is common knowledge in the communities using sparse regression methods that the  $\ell_0$ -norm is a measure of cardinality, hence (3.1) with  $p = 0$  is the natural choice in a theoretical sense. However, this problem is NP-hard in general, and therefore, it is more common to use  $\ell_p$ -norm with  $p \in ]0, 1]$ . A detailed explanation of sparsity inducing properties of  $\ell_p$ -norms with  $p \in ]0, 1]$  is given in [65]. A few sparse regression methods are briefly detailed here:

1. **LASSO** [66]: The specific case of (3.1) with  $p = 1$  is known as the LASSO is one of the most widely known sparse regression methods. The  $\ell_1$ -norm, even though not a true measure of cardinality, has a constant “driving force” or more technically speaking, a gradient towards the origin that is constant in magnitude everywhere. This drives many entries of  $\boldsymbol{\alpha}$  to zero, unlike the  $\ell_2$  norm that can reduce the magnitude of  $\boldsymbol{\alpha}$  entries, but not set them to zero. Also, the  $\ell_1$ -norm allows solving (3.1) in polynomial time.
2. **Dantzig Selector** [67]: This method allows selection of a sparse  $\boldsymbol{\alpha}$  using a modified version of (3.1), that can be converted into a linear programming problem after with algebraic manipulations. This is achieved with  $p = 1$  and the  $\ell_\infty$  norm for the residual. (3.1) is recast as:

$$\begin{aligned} \min & \|\boldsymbol{\alpha}\|_1 \\ \text{s.t.} & \|\mathbf{D}^T(\mathbf{x} - \mathbf{D}\boldsymbol{\alpha})\|_\infty < \varepsilon \end{aligned}$$

which can be expressed in a linear programming form as:

$$\begin{aligned} \min & \sum_i \gamma_i \\ \text{s.t.} & -\boldsymbol{\gamma} < \boldsymbol{\alpha} < \boldsymbol{\gamma} \\ & -\varepsilon \mathbf{1} < \mathbf{D}^T(\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}) < \varepsilon \mathbf{1} \end{aligned}$$

where  $\boldsymbol{\gamma}$  is a dummy variable obtained by opening the absolute operator  $\|\boldsymbol{\alpha}\|$ .

3. **Orthogonal Matching Pursuit** [68]: The OMP method aims to approximate the solution of (3.1) with  $p = 0$  using a greedy technique. At each greedy step  $k$ , the OMP algorithm searches for the column of dictionary  $\mathbf{D}$  that is the least orthogonal to the current residual  $(\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}^{k-1})$  and adds to the set of indices previously selected. This process is carried out until the residual is below the tolerance  $\varepsilon$ . This is done by projecting the residual vector on each column and selecting the largest projection.

$$\mathcal{I} \leftarrow \mathcal{I} \cup \arg \max_j \left| \mathbf{D}[:, j]^T (\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}^{k-1}) \right|$$

where  $\mathcal{I}$  is the current set of selected indices and columns of dictionary  $\mathbf{D}$  are normalized. Then the new set of coefficients  $\boldsymbol{\alpha}^k$  are computed using the least square solution:

$$\boldsymbol{\alpha}_{\mathcal{I}}^k = (\mathbf{D}_{\mathcal{I}}^T \mathbf{D}_{\mathcal{I}})^{-1} \mathbf{D}_{\mathcal{I}}^T \mathbf{x}$$

with  $\boldsymbol{\alpha}_{\mathcal{I}} = \boldsymbol{\alpha}[\mathcal{I}]$  and  $\mathbf{D}_{\mathcal{I}} = \mathbf{D}[:, \mathcal{I}]$ .

Update of  $\mathcal{I}$  is locally optimal, but the above least-square update of  $\boldsymbol{\alpha}^k$ , the OMP solution is optimal w.r.t to the currently selected subset  $\mathbf{D}_{\mathcal{I}}$  of the dictionary. This is in contrast to the predecessor of OMP, the Matching Pursuit [69], which used locally optimal updates for  $\boldsymbol{\alpha}^k$  at the current iteration.

4. **FOCUSS** [70]: The FOCUSS method finds a sparse approximation using an iterative process that starts from a fully dense solution and progresses towards “localized energy solutions”. The FOCUSS method is broadly based on the following ideas:

- For a full rank dictionary of size, say  $m \times n$  ( $m < n$ ), the system  $\mathbf{D}\boldsymbol{\alpha} = \mathbf{x}$  is under-determined. For such a system, the closed form solution for (3.1) with  $p = 2$  is given by  $\boldsymbol{\alpha} = \mathbf{D}^\dagger \mathbf{x}$ , where  $\mathbf{D}^\dagger$  is the Moore-Penrose pseudo-inverse. Note that this solution is not sparse in general, as  $\ell_2$ -norm does not have sparsity inducing properties.
- A modified minimization problem can be defined by replacing the  $\|\boldsymbol{\alpha}\|_2$  with a weighted norm  $\|\mathbf{W}^{-1}\boldsymbol{\alpha}\|_2$ , or more generically  $\|\mathbf{W}^\dagger\boldsymbol{\alpha}\|_2$  if  $\mathbf{W}$  is singular. The authors of [70] argue that by changing  $\mathbf{W}$ , every possible solution of the under-determined system  $\mathbf{D}\boldsymbol{\alpha} = \mathbf{x}$  can be obtained.
- The basis of the algorithm lies in iterating towards a sparse solution using a weight  $\mathbf{W}^k$  which induces sparsity. For the iteration  $k$ , the weight is chosen as  $\mathbf{W}_k = \text{diag}(\boldsymbol{\alpha}^{k-1})$ . The trick lies in the fact that algorithm ends up minimizing the following weighted norm:

$$\|\mathbf{W}_k^\dagger \boldsymbol{\alpha}\|_2^2 = \sum_i \left( \frac{\alpha_i^k}{\alpha_i^{k-1}} \right)^2$$

due to which smaller entries of  $\alpha^{k-1}$  tend to diminish further.

On combining these ideas, the FOCUSS algorithm ends up with quite a simple implementation, given in Alg. 3.1. Usually, the algorithm is initialized with the  $\boldsymbol{\alpha}^0$  containing all non-zeros. In [70], the  $\|\boldsymbol{\alpha}\|_2$  minimizing solution  $\boldsymbol{\alpha}^0 = \mathbf{D}^\dagger \mathbf{x}$  was used for initialization.

**Algorithm 3.1** FOCUS

---

```

1: Inputs:  $\mathbf{D}$ ,  $\mathbf{x}$ ,  $\boldsymbol{\alpha}^0$ 
2: Initialize  $k = 1$ 
3: while  $\boldsymbol{\alpha}$  not converge do
4:    $\mathbf{W}^k = \text{diag}(\boldsymbol{\alpha}^{k-1})$ 
5:    $\boldsymbol{\alpha}^k = \mathbf{W}^k (\mathbf{D}\mathbf{W}^k)^\dagger \mathbf{x}$ 
6:    $k \leftarrow k + 1$ 
7: end while

```

---

**3.2.2 Application of sparse methods to unconstrained resolution problems**

Sparse regression methods, though originally developed for application on pure data approximation or data compression problems, have seen recent interest in applications towards physics problems [64, 71, 72]. The methods that have been discussed previously are also meant for data approximation. At this point, application of such methods on unconstrained resolution problems is discussed. A resolution problem can be framed as residual minimization problem in order to use dictionary-based approximation and sparse regression. Consider a discretized resolution problem with the state vector  $\mathbf{v}$  and residual  $\mathbf{r}(\mathbf{v})$ . The basic idea is to approximate the solution  $\mathbf{v} \approx \mathbf{D}\hat{\mathbf{v}}$  using the elements of dictionary such that the residual of the numerical model  $\mathbf{r}(\mathbf{D}\hat{\mathbf{v}})$  is below a certain tolerance, while minimizing the  $\ell_p$ -norm of coefficient vector, as shown in (3.2). Sparsity is imposed on the coefficients of the dictionary, meaning that only a few dictionary columns are selected.

$$\begin{aligned} \min & \|\hat{\mathbf{v}}\|_p \\ \text{s.t.} & \|\mathbf{r}(\mathbf{D}\hat{\mathbf{v}})\| < \varepsilon \end{aligned} \quad (3.2)$$

Dictionary-based sparse approximation for resolution problems has been discussed in [73], where the solution space is set to the span of dictionary columns. They also discuss the application of random sketching to residual evaluation, which provides an efficient way of solving the minimization problem. This is similar to the approach of randomized SVD proposed in [74]. The idea is to project the residual on a low-dimensional subspace in  $\text{colsp}(\mathbf{D})$ .<sup>3</sup> The matrix  $\mathbf{B}$ , containing a basis corresponding to a random combination of dictionary columns, is used for this projection in (3.3). This essentially reduces the complexity of enforcing the residual constraint. Given a dictionary  $\mathbf{D}$  of size  $m \times n$ , a smaller matrix that has information from all columns of  $\mathbf{D}$  can be generated by multiplying a matrix  $\mathbf{R}$  of size  $n \times L$  with  $L \ll \min(m, n)$ . After orthogonalization,

---

<sup>3</sup>colsp indicates column-space

this matrix  $\mathbf{B}$  can be used for projecting the residual  $\mathbf{r}$ .

$$\min \|\widehat{\mathbf{v}}\|_p \quad (3.3a)$$

$$\text{s.t. } \|\mathbf{B}^T \mathbf{r}(\mathbf{D}\widehat{\mathbf{v}})\| < \varepsilon \quad (3.3b)$$

$$\text{where } \mathbf{B} = \text{orth}(\mathbf{DR}), \quad \mathbf{R} \in \mathcal{U}_{[0,1]}^{n \times L} \quad (3.3c)$$

The number  $L$  can be tuned to not only control the complexity of the minimization problem, but also to crudely constrain sparsity of the solution. For  $L \ll n$ ,  $\text{colsp}(\mathbf{B}) \subset \text{colsp}(\mathbf{D})$  is low-dimensional and hence, the projection  $\mathbf{B}^T \mathbf{r}$  is a “weaker” evaluation of  $\mathbf{r}$ , which allows for a sparser  $\widehat{\mathbf{v}}$  at the price of permitting residuals orthogonal to  $\mathbf{B}$ .

Another related approach in which snapshots are directly used for reducing the model without any explicit compression is the CUR decomposition [39]. In this method, not only columns, but also rows from the snapshot matrix are used for developing a parametric regression. It works by algorithmically selecting a column and row subset from the snapshot matrix, and using them for developing a regression model. A contrasting feature of the CUR approach is the use of regression without invoking the partial differential equation or its weak form, resulting in a non-intrusive method.

### 3.2.2.1 Diffusion-reaction problem

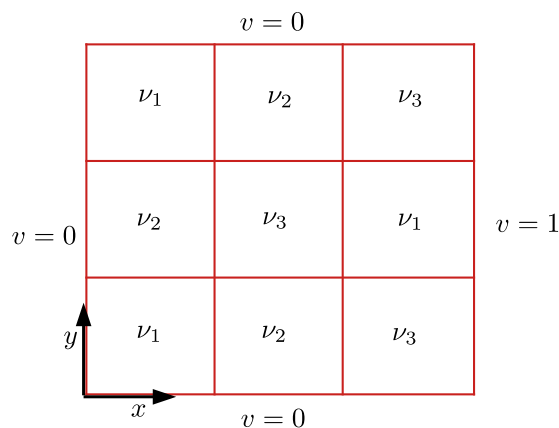
A transport problem of a quantity  $v$  with two-dimensional domain of size  $3 \times 3$  units, divided into a grid with cells of size  $1 \times 1$  is considered. The cells have a piecewise uniform diffusivity value. A three-dimensional parametrization is considered with diffusivity values  $\nu_1, \nu_2$  and  $\nu_3$ , as shown in Fig. 3.2. The reaction coefficient  $\sigma$  is considered to be uniform across all sectors.

$$\begin{aligned} -\nu(x, y) \nabla^2 v(x, y) + \sigma v(x, y) &= f \\ v(x, 0) &= 0 \\ v(x, 3) &= 0 \\ v(0, y) &= 0 \\ v(3, y) &= 1 \end{aligned} \quad (3.4)$$

where  $\nu(x, y)$  describes the piecewise uniform diffusivity values.

The 2D domain is discretized into a structured  $15 \times 15$  mesh. The diffusivity of each block  $\nu_i \in (1, 150)$  is considered for the parametric model. The reaction coefficient is chosen as  $\sigma = 10$ . The reaction is included mainly to add some local effects to an otherwise globally diffusive problem. A dictionary of 320 snapshots is created using randomly sampled diffusivity values. Fig. 3.3 shows some sample snapshots from the training set.

Fig. 3.4 shows reconstruction errors using the dictionary-based approximation methods using (3.3). FOCUSS and OMP are used for sparse regression. As expected, the training set snapshots are almost accurately reconstructed with a single non-zero coefficient. In the validation set, reconstruction errors are observed in the order of  $10^{-2}$  to



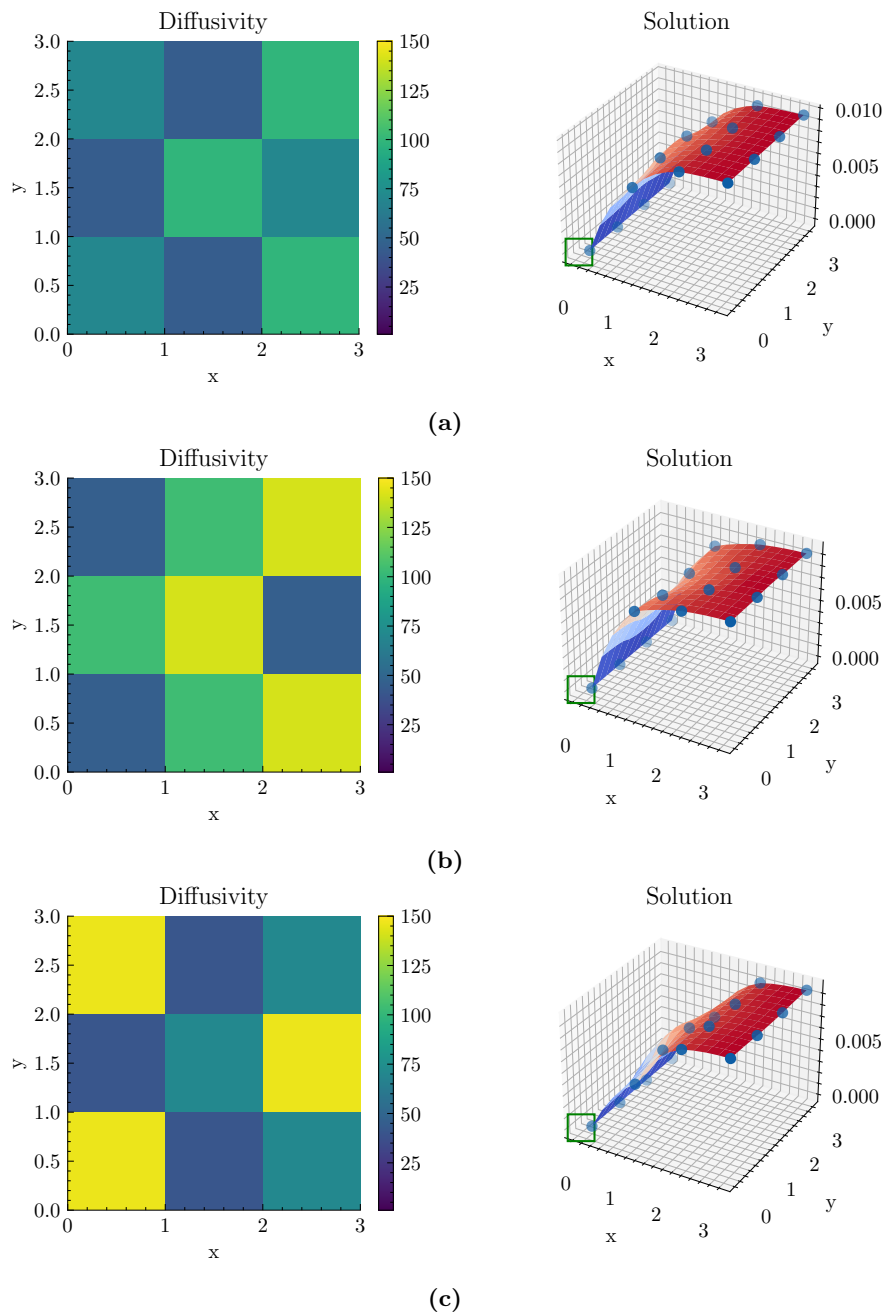
**Figure 3.2:** Two-dimensional domain for the transport problem (3.4)

$10^{-1}$  for sparsity target  $L = 10$  and errors improve about an two orders for  $L = 30$ . The outcome of sparsity for each  $L$ , shown using a cross ( $\times$ ), is almost the same as the parameter  $L$  for the validation set. This is due to the algorithm trying to select as many vectors permitted by the user.

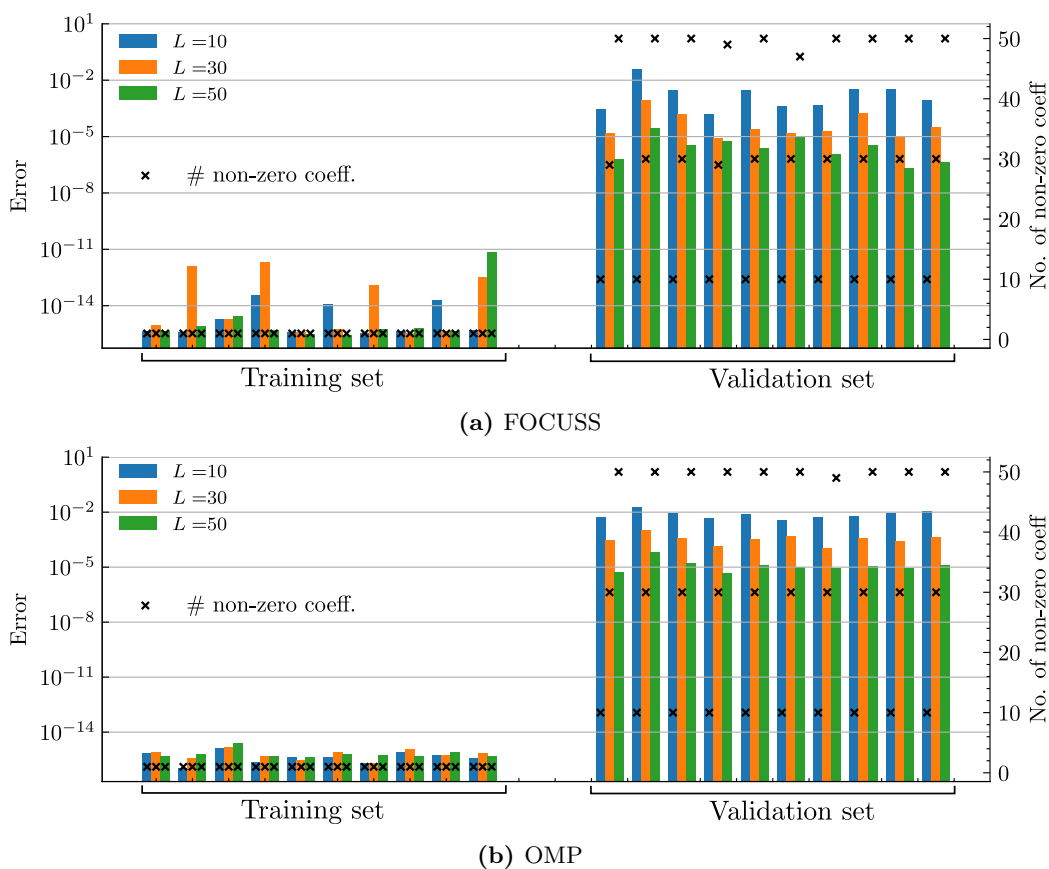
Now that the dictionary methods have been demonstrated on unconstrained methods, we would like to extend these methods to inequality-constrained problems. However, it is not straightforward to apply random sketching to problems involving inequality constraints. The projection of inequality constraints on a randomized matrix like  $\mathbf{B}$  creates linear combinations of constraints, which makes it harder to identify which constraints are being violated and which are not. In upcoming sections, two dictionary-based approximation schemes will be studied, but the projection of residual on  $\mathbf{B}$  is applied in only one of them.

### 3.3 Dictionary methods for contact problems

Now that the feasibility of dictionary-based approximations have been established in an unconstrained problem, we move towards inequality constrained problems. To this end, two dictionary-based approximation methods for contact mechanics problems have been formulated, namely the *greedy active-set* approach and *convex-hull approximation* approach. To provide a broad overview of the features of these methods, Table 3.1 gives a list of features of these methods. The important differences between the two methods are bases types and the manner of enforcing constraint. These details will be discussed along with their formulations and application to contact problems are discussed in this section.



**Figure 3.3:** Snapshots of the transport problem (3.4) for randomly sampled parametric values



**Figure 3.4:** Reconstruction errors (bars) of the transport problem using dictionary-based approximations and number of non-zero entries ( $\times$ ) on training and validation sets.  $L$  indicates the target sparsity

### 3.3.1 A greedy active-set method for dual dictionary element selection

To solve the contact problem using a dictionary of contact pressure snapshots, a greedy method inspired from the OMP and active set method is devised. From this point onwards,  $\mathbf{D}_d$  indicates a dictionary of dual snapshots of the contact problem. Consider the following sparse problem:

$$\min \|\widehat{\boldsymbol{\lambda}}\|_0 \quad (3.5a)$$

$$\text{s.t.} \quad \begin{bmatrix} \boldsymbol{\Phi}^T \mathbf{K} \boldsymbol{\Phi} & \boldsymbol{\Phi}^T \mathbf{C}^T \mathbf{D}_{d,\mathcal{I}} \\ \mathbf{D}_{d,\mathcal{I}}^T \mathbf{C} \boldsymbol{\Phi} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{u}} \\ \widehat{\boldsymbol{\lambda}}_{\mathcal{I}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}^T \mathbf{f} \\ \mathbf{D}_{d,\mathcal{I}}^T \mathbf{g} \end{bmatrix} \quad (3.5b)$$

$$\text{and } \mathbf{D}_d^T (\mathbf{C} \boldsymbol{\Phi} \widehat{\mathbf{u}} - \mathbf{g}) \leq \tau \quad (3.5c)$$

$$\widehat{\boldsymbol{\lambda}} \geq 0 \quad (3.5d)$$

where  $\mathcal{I}$  indicates the active set and  $\mathbf{D}_{d,\mathcal{I}} = \mathbf{D}_d[:, \mathcal{I}]$

	Greedy active-set	Convex hull approximation
Primal Basis type	Truncated low-rank	Dictionary
Dual Basis type	Dictionary	Dictionary
Monolithic dictionary	No	Yes
Iteration method	Fixed point	Fixed point
Constraint enforcement	Active-set	Convex combinations of non-penetrating snapshots
Sparsity induction	Greedy activation and deactivation of dual dofs	Non-negative FOCUSS
Additional assumptions	None	Convexity of the feasible region

**Table 3.1:** Characteristics of the two dictionary-based approximation methods for contact mechanics problems

(3.5b), (3.5c) and (3.5d) are based on the same principle as the reduced KKT conditions (1.14), but applied over a dictionary  $\mathbf{D}_d$  rather than a low-rank dual basis. Comparing (3.5) with the generic form of sparse regression statement (3.2), the balance equation is imposed as an equality constraint, instead of an inequality constraint where the residual norm is limited to certain tolerance. Imposing limits on the residual norm is usually more relevant when using random sketching, like in (3.3) when the system is under-determined. Also, it is not straightforward to use standard sparse regression algorithms in presence of the KKT inequality constraints (3.5c) and (3.5d). Equality constraint can be used as long as the system is determined, which is currently the case in (3.5b). In (3.5c),  $\tau$  indicates a small penetration that has been permitted in the algorithm, which is necessary to ignore spurious penetrations created by the low-rank displacement basis  $\Phi$ . Unless specified, the  $\tau$  is set to a value in the same order of truncation tolerance  $\delta$  (see Appendix A).

The primary notion behind greedy algorithms is to enrich the approximation by a rank-1 term in each step. For standard sparse regression problems like (3.2), the OMP algorithm starts with an empty coefficient vector and adds a single non-zero entry in one step. The position of the new non-zero entry is chosen by searching the vector in the dictionary with the highest correlation to the residual. The greedy active-set algorithm devised here performs the greedy enrichment in an almost similar way but with a modification, i.e. the orthogonality with the violation of inequality constraint (3.5c) is used instead of the residual of system (3.5b). More precisely, the algorithm starts with a zero  $\hat{\lambda}$  and in each enrichment step, the dictionary column with highest correlation to the violations of the non-penetration condition is added in the greedy enrichment. Thus, enrichment index of the dictionary  $p_{\text{add}}$  can be computed as:

$$p_{\text{add}} := \arg \max_p \mathbf{D}_d[\cdot, p]^T \left( \mathbf{C}\Phi\hat{\mathbf{u}}^k - \mathbf{g} \right)^{+, \tau} \quad (3.6)$$

where  $(\mathbf{z})_i^{+, \tau} := \begin{cases} z_i & \text{if } z_i > \tau \\ 0 & \text{otherwise} \end{cases}$ , which is a hard-thresholding operator.

(3.6) can be computed more efficiently using the reduced operators  $\widehat{\mathbf{C}}(\mu, \hat{\mathbf{u}}^{k-1})$  and  $\widehat{\mathbf{g}}(\mu, \hat{\mathbf{u}}^{k-1})$ :

$$p_{\text{add}} := \arg \max_p \left( \widehat{\mathbf{C}}\hat{\mathbf{u}}^k - \widehat{\mathbf{g}} \right)^{+, \tau} [p] \quad (3.7)$$

where

$$\begin{aligned} \widehat{\mathbf{C}}(\mu, \hat{\mathbf{u}}^{k-1}) &= \mathbf{D}_d^T \mathbf{C}(\mu, \hat{\mathbf{u}}^{k-1}) \Phi \\ \widehat{\mathbf{g}}(\mu, \hat{\mathbf{u}}^{k-1}) &= \mathbf{D}_d^T \mathbf{g}(\mu, \hat{\mathbf{u}}^{k-1}) \end{aligned}$$

However, to ensure that inequality constraints (3.5d) that prohibits non-negative pressures, it is necessary to not only enrich but also to eliminate terms that violate the constraints in each greedy step, which is simply done by eliminating the largest negative  $\widehat{\boldsymbol{\lambda}}$  i.e. setting the largest negative entry to zero. The elimination index  $p_{\text{rem}}$  is computed as:

$$p_{\text{rem}} = \arg \min_p \left( \widehat{\boldsymbol{\lambda}}^k \right)^- [p] \quad (3.8)$$

where  $(\mathbf{z})_i^- := \begin{cases} z_i & \text{if } z_i < 0 \\ 0 & \text{otherwise} \end{cases}$

Though the motivation for this algorithm was based on the OMP to perform the dual dictionary element selection, the greedy active-set algorithm ends up being quite similar to the fixed-point active-set method. In this approach, only a maximum of one vector is activated in a given iteration thereby maintaining a small active set in intermediate iterations. The algorithm is given in Alg. 3.2, where Steps 8 to 10 contain the enrichment and the elimination process. In each iterative step, only one of the two operations is performed. Due to the architecture of the algorithm, the only steps that scale with the size of the dictionary are enrichment/elimination Steps 8 to 10. The size of mixed system in Step 6 is not significantly influenced by the size of active set during the intermediate iterations, thanks to the small sizes of active set  $\mathcal{I}$  which is closely related to the sparsity of  $\widehat{\boldsymbol{\lambda}}$ .

### 3.3.1.1 Application to the Hertz problem

This section deals with the implementation of the greedy active-set algorithm for the Hertz contact problem discussed in Section 2.4.1 of previous chapter (Fig. 2.5a). The parametrization also remains same, i.e. the imposed displacement on the top cylinder

**Algorithm 3.2** Greedy active-set algorithm

- 
- 1: Input: Queried value of parameter  $\mu$
  - 2: Given: Primal basis  $\Phi$  and dual Dictionary  $\mathbf{D}_d$   
           Reduced operators  $\Phi^T \mathbf{K} \Phi$  and  $\Phi^T \mathbf{f}$  ▷ can be built offline
  - 3: Initialize:  $k = 0$ ,  $\mathcal{I} = \emptyset$
  - 4: **while**  $\hat{\mathbf{u}}$  and  $\mathcal{I}$  not converge **do**
  - 5:     Build reduced constraint operators  $\hat{\mathbf{C}}(\mu, \hat{\mathbf{u}}^{k-1})$  and  $\hat{\mathbf{g}}(\mu, \hat{\mathbf{u}}^{k-1})$
  - 6:     Solve the linear system (3.5b)
  - 7:     Set  $\hat{\lambda}[\mathcal{I}] \leftarrow \hat{\lambda}_{\mathcal{I}}$  and  $\hat{\lambda}[\mathcal{I}^c] \leftarrow \mathbf{0}$  ▷  $\mathcal{I}^c$  is the complementary set
  
  - Compute the largest constraint violations and enrich/eliminate accordingly:
  - 8:     **if**  $\hat{\lambda}^k \geq \mathbf{0}$  : **then**

$$p_{\text{add}} = \arg \max_p \left( \hat{\mathbf{C}} \Phi \hat{\mathbf{u}}^k - \hat{\mathbf{g}} \right)^{+, \tau}$$

$$\mathcal{I} \leftarrow \mathcal{I} \cup \{p_{\text{add}}\}$$
  - 9:     **else:**

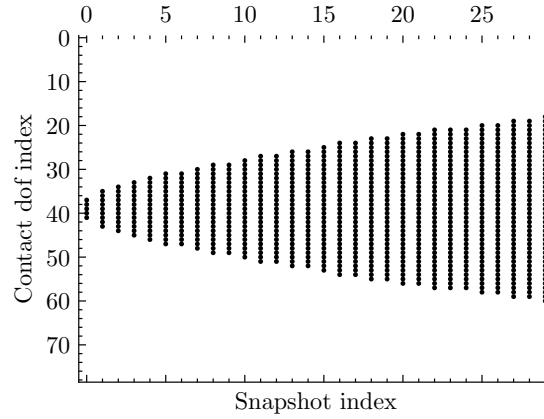
$$p_{\text{rem}} = \arg \min_p \left( \hat{\lambda}^k \right)^- [p]$$

$$\mathcal{I} \leftarrow \mathcal{I} \setminus \{p_{\text{rem}}\}$$
  - 10:    **end if**
  - 11:     $k \leftarrow k + 1$
  - 12: **end while**
  - 13: Reconstruct  $\mathbf{u} = \Phi \hat{\mathbf{u}}$  and  $\lambda = \mathbf{D}_d \hat{\lambda}$
  - 14: Output:  $\mathbf{u}, \lambda$
- 

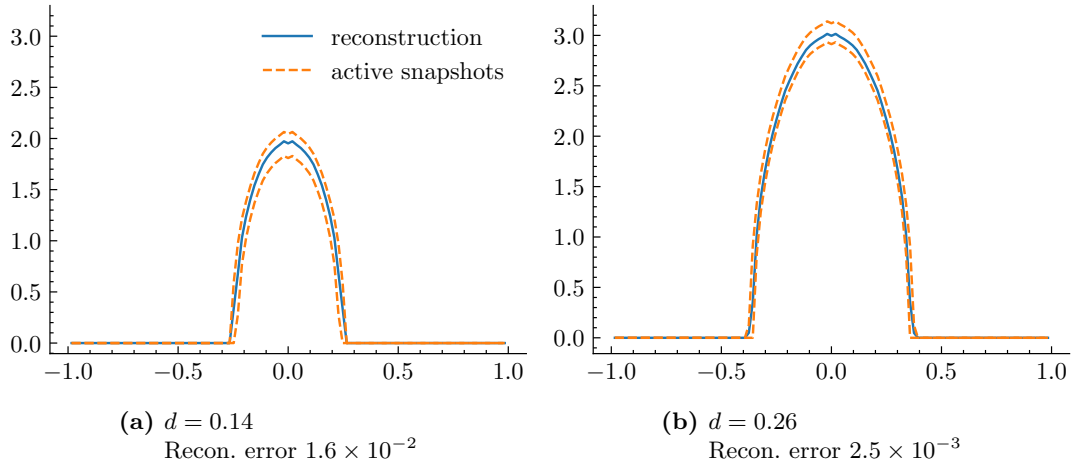
$d \in (0, 0.3)$ . The finite element model also remains the same.

**Reduced model:** Primal and dual snapshots on three training sets of sizes 12, 30, 60 and 120 are computed in the offline stage. The parametric points in each training set  $\mathcal{P}_{\text{tr}}$  are distributed uniformly in parametric space  $\mathcal{P} = ]0, 0.3]$ . The primal basis is computed using SVD and truncated using cumulatively normalized singular values using a tolerance  $\delta$ . For the dual variable, a dictionary consisting contact pressure snapshots is used instead of a reduced basis. A spy plot of the dual dictionary for the training set of size 30 is shown in Fig. 3.5 for visualization purposes. A validation set  $\mathcal{P}_{\text{val}}$  with 119 points, which are the mid-points of the 4th training set is used to compute reconstruction errors.

Two instances of contact pressure reconstructions by the greedy active-set algorithm for parametric points outside the training set are shown in Fig. 3.6. The dictionary snapshots selected by the algorithm for the reconstruction are also shown. As expected, only a few dictionary elements, two in this case, are selected. The two selected snapshots are the ones whose contact position is similar to the required contact position for the specific instance. It indicates the greedy enrichment and elimination algorithm is selecting the correct snapshots for reconstruction.



**Figure 3.5:** A spy pattern of dual dictionary with 30 elements for the Hertz problem. The number of rows, 79, is same as the number of dofs on the potential contact surface. The dictionary columns are arranged in the increasing order of loading parameter  $d$ .



**Figure 3.6:** Greedy active-set reconstruction of contact pressure for certain parametric instances outside the training set of the Hertz problem. The elements of dictionary chosen by the algorithm are shown and the reconstruction errors are given. Dictionary of size 12 is used in these examples.

Mean reconstruction errors over the validation set vs. size of the training set are shown in Fig. 3.7a, along with the rank of primal basis. The reconstruction errors improve by more than an order of magnitude as dictionary size is increased from 12 to 120 for the case  $\delta = 10^{-10}$ . However, such gains with over-complete dictionaries diminish for the case  $\delta = 10^{-6}$ , especially the primal error improves only slightly. This is probably because at  $\delta = 10^{-6}$ , the primal truncation starts to become dominant source of error, rather than just the linear inseparability of contact pressure which seems to be the case for  $\delta = 10^{-10}$ .

The computational time is computed as the total time taken for computation of

Alg. 3.2, excluding the time taken in Step 5. The operator construction time is excluded because their efficient construction is not a part of this work. Mean computational time on the validation set (Fig. 3.7c) shows an initial increase with the dictionary size, but later settles down, but it is evident that the computational time is mostly influenced primarily by the number of iterations (Fig. 3.7b). This can also be seen in Fig. 3.7d, where the computational time per iteration is more or less flat w.r.t. the size of dictionary. The small increases in computational time per iteration with dictionary size can be explained due to the increase in the dimension of primal basis  $\Phi$ , especially the case of  $\delta = 10^{-10}$  where the primal rank is strongly dependent on the training set size.

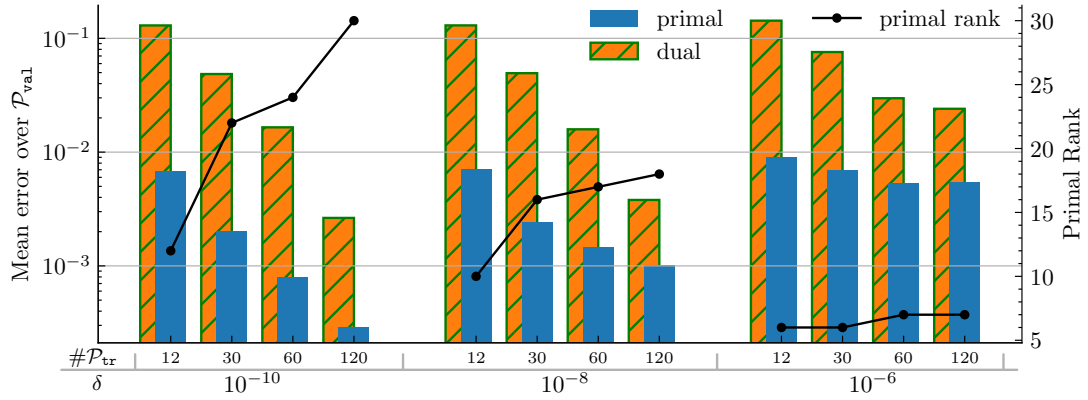
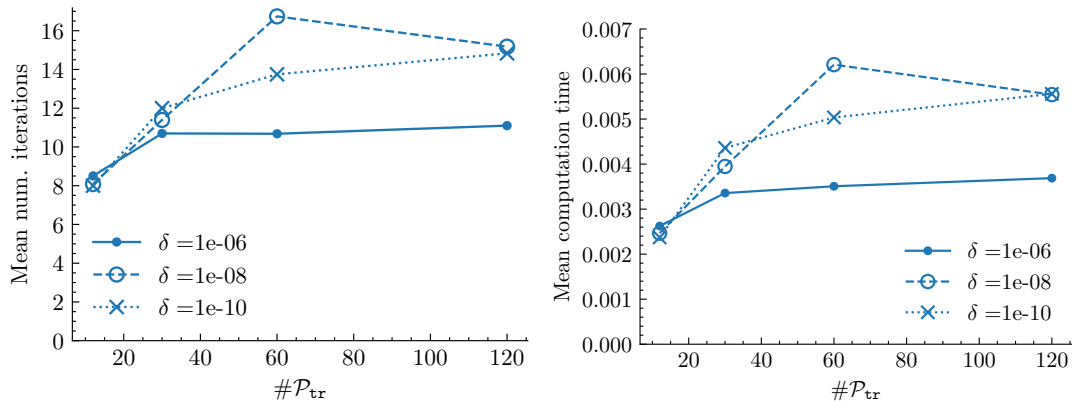
Detailed plots of reconstruction errors over the validation set  $\mathcal{P}_{\text{val}}$  are shown in Fig. 3.8 for primal truncation tolerance  $\delta \in \{10^{-10}, 10^{-8}, 10^{-6}\}$ . Each curve corresponds to one of the training sets. Naturally, a larger training set is expected to have a lower (or at least the same) reconstruction error level. This anticipation is marginally satisfied in the case of  $\delta = 10^{-10}$ , but not in the case of  $\delta = 10^{-6}$ , where there are certain regions where the dual error for larger training sets is higher than a smaller training set. This happens due to the spurious penetrations discussed previously (and in Appendix A) and implies that even after relaxing the constraint by value of  $\tau$  (see (3.5c)), some spurious selections are still possible for higher values of  $\delta$ . Also, for all cases, the errors tend to be larger in general near  $d = 0.0$ , as the contact is sensitive to the parameter  $d$  around this region.

Sparsity pattern of the dual reduced dofs  $\hat{\lambda}$  (i.e. the coefficients of the dual dictionary  $\mathbf{D}_d$ ) resulting from the greedy active-set approach for the training set of size 30 is shown in Fig. 3.9. As the snapshots in  $\mathbf{D}_d$  are arranged in increasing order of  $d$ , most of the columns of the sparsity pattern have two close non-zeros, indicating that nearest two snapshots to the targeted reconstruction were chosen by the algorithm. Spurious selections are seen in some rows, but the values of corresponding coefficients are quite small and do not have significant impact on solution.

### 3.3.1.2 Application to the Ironing problem

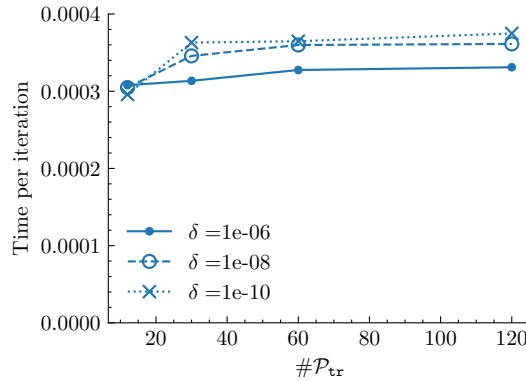
Unlike the Hertz problem, other contact problems might exhibit larger changes in the contact zone, like the ironing problem from Section 2.4.2 (Fig. 2.10). The effects of the strong variations of contact zone on the inseparability of the contact pressure were also discussed. Therefore, over-complete dictionaries might be more significant in generating an effective reduced model.

The parametrization of the horizontal displacement  $d_x$  of the iron over the slab is considered for the reduced model. The vertical displacement  $d_y = 0.3$  is a fixed value. Given length of the slab  $L$ , the parametric space is defined by  $\mathcal{P} := \{d_x \mid 0 \leq d_x \leq L\}$ . To study the influence of size of training sets on the reconstruction quality, training sets  $\mathcal{P}_{\text{tr}}$  are built in a nested manner which allows studying the evolution of reconstruction errors as a function of training set size i.e. size of the over-complete dictionary. Each nested level has uniformly distributed points over the parametric space  $\mathcal{P}$ . The number of points is nearly doubled in the next nested level which also contain the points from the current level. Therefore, at the  $n$ -th nested level,  $2^n + 1$  snapshots are computed,

(a) Mean errors (bar plot, left  $y$ -axis) and primal rank (line plot, right  $y$ -axis)

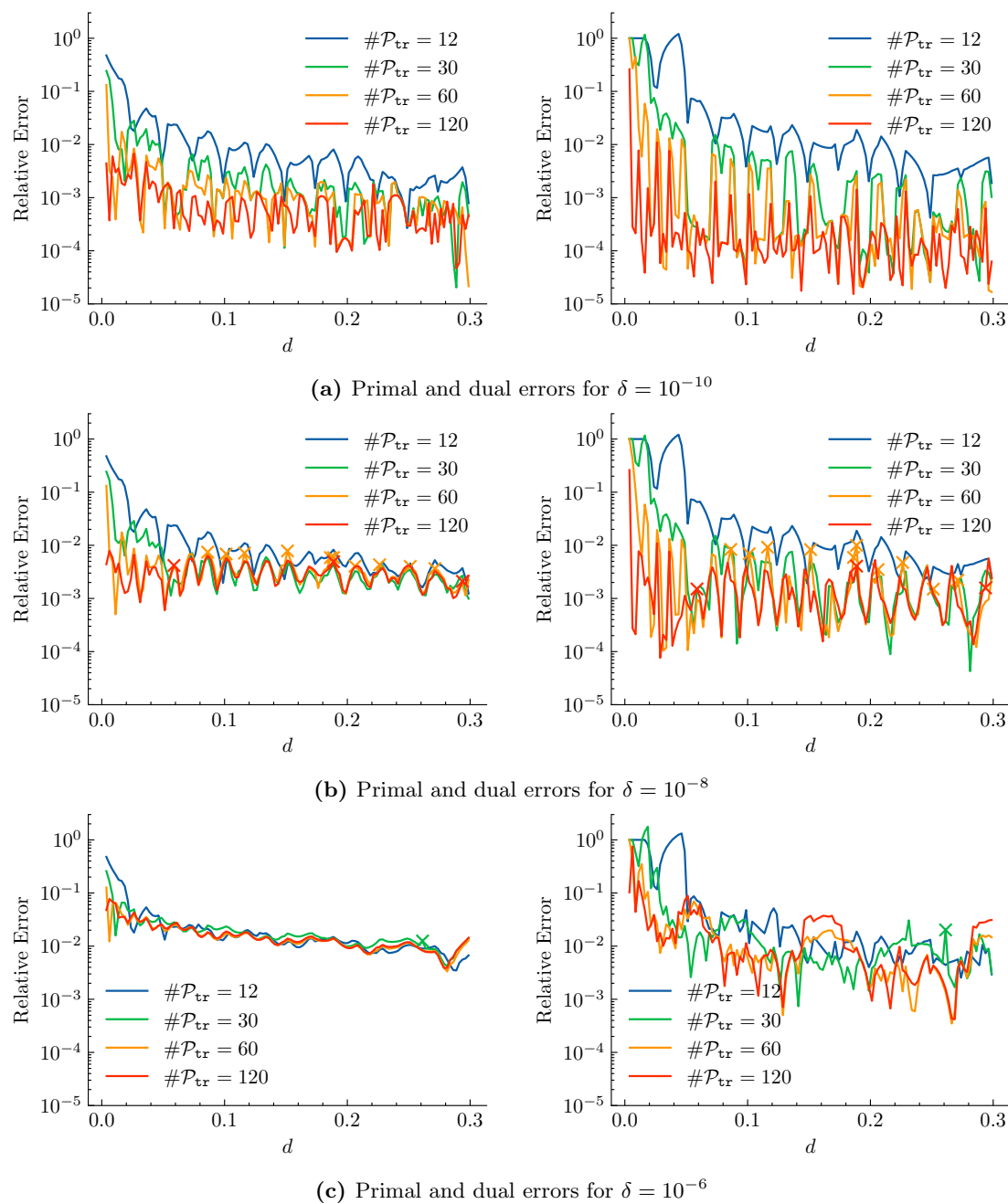
(b) Mean number of iterations

(c) Mean computation time

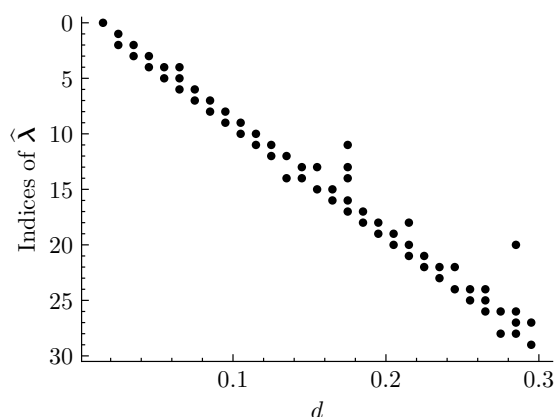


(d) Mean computation time per iteration

**Figure 3.7:** Evolution of (a) Mean reconstruction (relative) errors and primal rank, (b) Mean number of iterations, (c) Mean computation time and (d) Mean computation time per iteration with training set size for the Hertz problem. The mean reconstruction error is computed by taking mean of reconstruction errors over the validation set  $\mathcal{P}_{val}$ . Primal and dual errors are computed using  $\mathcal{H}^1$  and  $\mathcal{L}^2$  norm, respectively. Curves correspond to different truncation tolerance  $\delta$  for the primal basis, for values  $10^{-6}$ ,  $10^{-8}$  and  $10^{-10}$ . Note that computation time excludes time for construction of non-linear operators.



**Figure 3.8:** Detailed primal and Dual reconstruction errors for the Hertz problem using the training sets 1,2 and 3. Validation set is common to all curves, consisting of 60 equispaced points in  $\mathcal{P}$ . Each subfigure corresponds to a different truncation tolerance  $\delta$  for the primal basis. Primal and dual errors are computed using  $\mathcal{H}^1$  and  $\mathcal{L}^2$  norm, respectively. Cross-marks  $\times$  indicate points where the greedy active-set algorithm did not converge to the defined tolerance  $10^{-5}$



**Figure 3.9:** Sparsity of dual reduced dofs  $\hat{\lambda}$  selected by greedy active-set method for reconstructions in validation set of the Hertz problem using dual dictionary of size 30. Dots indicate the non-zero positions.

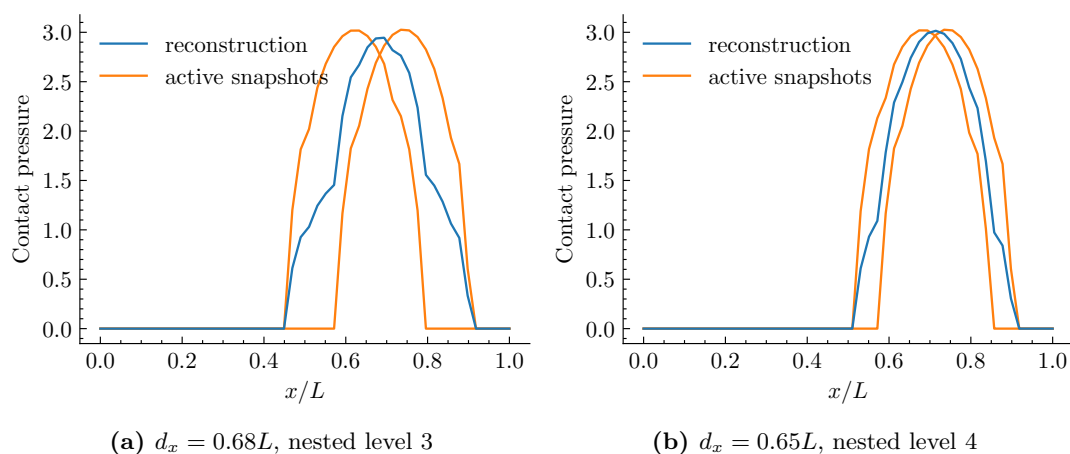
and the contact pressure snapshots constitute the dual dictionary  $\mathbf{D}_d$ . The primal basis, however, is truncated by filtering the singular values by a fixed tolerance  $\delta$ , and therefore its rank must be less than  $2^n$ . Training sets of nested levels  $n = 3$  to  $n = 7$  are used to build the primal RB and the dual dictionary. The validation set  $\mathcal{P}_{\text{val}}$  contains all points from nested level  $n = 8$  that are not present in level  $n = 7$ .

Fig. 3.10 shows two instances of reconstruction of contact pressure along with the dictionary elements that were chosen by the algorithm. The algorithm selects snapshots that are closest to the required contact position. This can also be seen in Fig. 3.11, where the sparsity of dual dofs shows a clear selection pattern. Unlike the sparsity pattern of Hertz problem (Fig. 3.9), the ironing problem does not display spurious snapshot selections; possibly due to the stronger dependence of contact position of the ironing problem; unlike Hertz problem where contact pressure curves were centered around the same position. Therefore, the activation process is less ambiguous and less sensitive to primal truncations for the case of ironing problem.

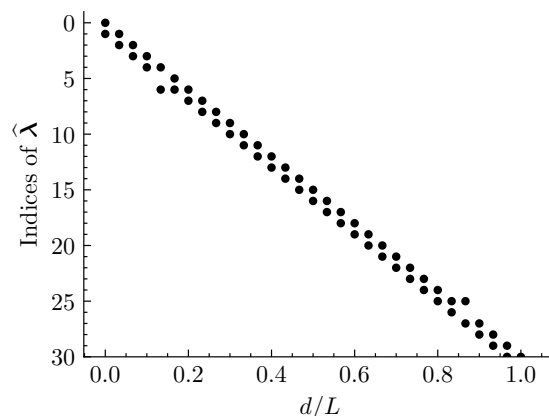
Fig. 3.12 shows the evolution of reconstruction (relative) errors, primal basis rank, number of iterations and computation time over the validation set  $\mathcal{P}_{\text{val}}$  vs. training set nested levels ( $\log_2(\#\mathcal{P}_{\text{tr}})$ ). Means of the reconstruction related quantities over the validation set is shown for primal truncation tolerances of  $10^{-6}$ ,  $10^{-8}$  and  $10^{-10}$ . The dual error improves, although slowly, with increasing size of dictionary, but the primal error shows signs of stagnation near level 6 and 7 and even earlier for the case of  $\delta = 10^{-6}$ . Errors for  $\delta = 10^{-8}$  and  $10^{-10}$  are almost the same, possibly because the inseparability of contact pressure and the finite size of dual dictionary are important sources of error. Conversely, in case  $\delta = 10^{-6}$ , the primal truncation seems to have more influence as primal error stagnates early, similar to the observations in case of the Hertz problem.

Moreover, the number of iterations and consequently the computation time seems to increase significantly with the dictionary size, in contrast to the Hertz problem where

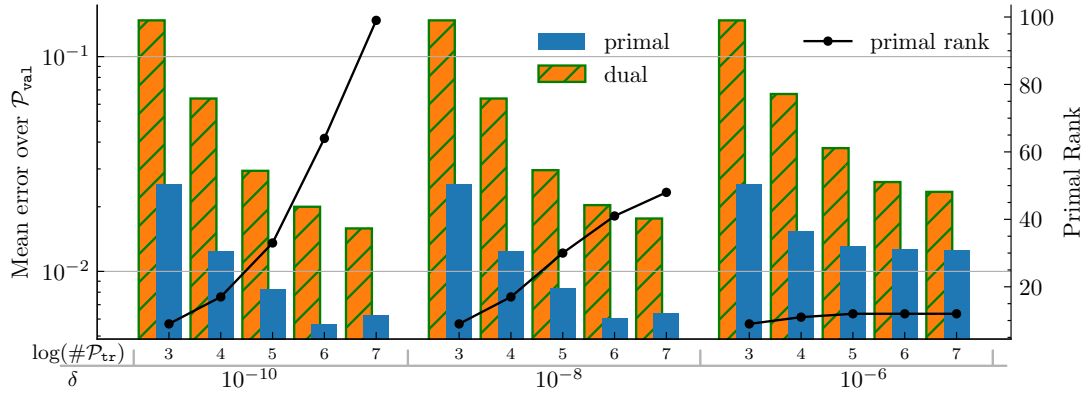
this relation was comparatively flat. A possible explanation for this behaviour could be that the larger dictionaries of the ironing problem, coupled with stronger parametric dependence of contact pairs, forces the greedy active-set to perform more iterations to find the appropriate dual dictionary candidates and contact pairs. Mean computation time per iteration shows trends that are a bit similar to the primal rank i.e. both primal rank and computational time per iteration are flat for  $\delta = 10^{-6}$ , whereas both are monotonically increasing for  $\delta = 10^{-10}$ . This was an expected observation since a large part of the system (3.5b) is related to the dimensionality of the primal basis  $\Phi$ . The size of the dual dictionary, even though not truncated, does not seem to have any significant influence on the computation time per iteration.



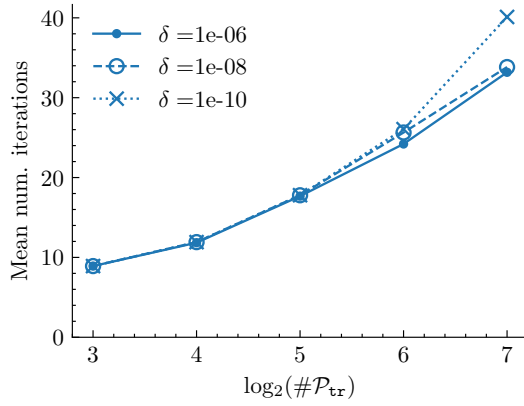
**Figure 3.10:** Greedy active-set reconstruction of contact pressure for certain parametric instances of the ironing problem. The dictionary nested level, dictionary elements chosen by the algorithm and the reconstruction errors are also shown.



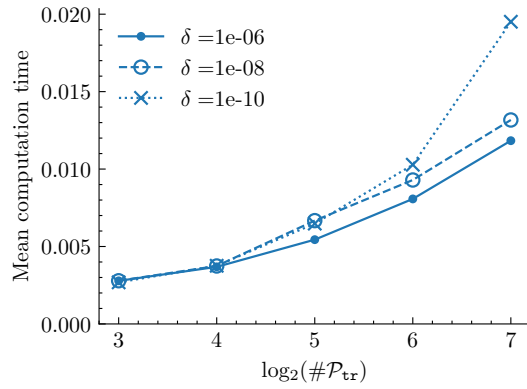
**Figure 3.11:** Sparsity of dual reduced dofs  $\hat{\lambda}$  selected by greedy active-set method for various reconstructions of ironing problem using dual dictionary of nested level 5



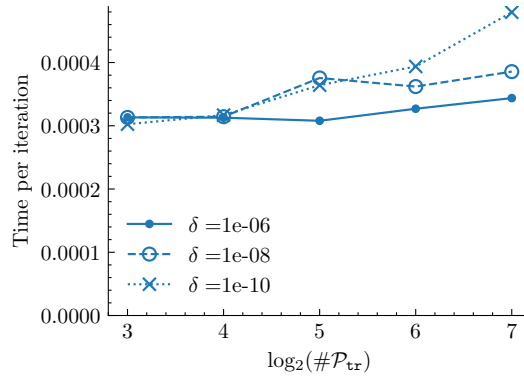
(a) Mean reconstruction errors and primal rank



(b) Mean number of iterations



(c) Mean computation time



(d) Mean computation time per iteration

**Figure 3.12:** Evolution of (a) mean reconstruction (relative) errors and rank of primal basis (b) mean number of iterations, (c) mean computation time and (d) mean computation time per iteration with training set size for the Ironing problem. The mean reconstruction error is computed by taking mean of reconstruction errors over the validation set  $\mathcal{P}_{\text{val}}$ . Primal and dual errors are computed using  $\mathcal{H}^1$  and  $\mathcal{L}^2$  norm, respectively. Curves correspond to different truncation tolerance  $\delta$  for primal basis, for values  $10^6$ ,  $10^{-8}$  and  $10^{-10}$ . Note that computation time excludes time for construction of non-linear operators.

### 3.3.2 A non-penetrating convex hull approach for monolithic dictionaries

As discussed in Remark 1.1, an interesting feature of mixed ROMs for incompressible flow problems is the computation of divergence-free RB. The advantage of this feature is that incompressibility constraint does not need to be explicitly enforced in the reduced problem. Moreover, orthogonality of a divergence-free velocity subspace and the irrotational pressure gradient term causes the pressure terms in the reduced problem to completely vanish (see [59, 60] for details). Computation of a divergence-free RB does not require any special treatment, as the divergence-free property of the training set snapshots is easily preserved in the resulting reduced subspace.

In this section, we attempt to extend a similar notion of constraint satisfying solution space to contact mechanics. But the approach is quite different compared to the incompressible flow problem because the nature of constraint is different. The incompressibility constraint, an equality constraint, applies globally over the domain; whereas the non-penetration constraint, an inequality constraint, applies locally on the contact surfaces and can be either active or inactive locally. Also, the displacement and contact pressure do not possess any special mathematical properties that cause the pressure term to vanish, as in the case of incompressible flow. In fact, the approach presented in this section, a reduced solution “space” is defined that naturally satisfies the non-penetration condition, but does not result in elimination of the contact pressure.

Unlike the greedy active-set algorithm previously explored which used POD basis for displacement and a dictionary for contact pressure, the following approach utilises a monolithic dictionary whose columns are snapshots consisting of both displacement and contact pressure. The reason for using a monolithic dictionary will become evident in the formulation of this approach. Moreover, this method allows the residual to be projected on a low-dimensional space like it was done in the unconstrained problem in Section 3.2.2. It also permits the utilisation of sparse regression methods discussed in Section 3.2.1, provided the non-negativity constraint can be enforced. In this work, the non-negative version of the FOCUSS algorithm is used.

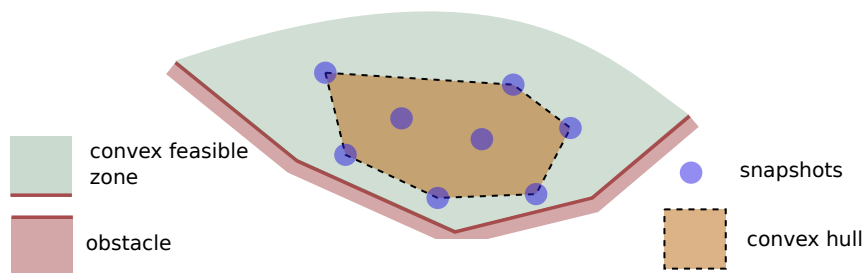
Consider a contact problem where the contact pairs are constant and independent of the parameter of the problem. The discrete inequality constraint can be expressed using matrices that do not change with the displacement field. The feasible region  $\mathcal{K}$  in the displacement solution space can be defined as:

$$\mathcal{K} := \{\mathbf{u} \in \mathcal{V} \mid \mathbf{C}\mathbf{u} - \mathbf{g} \leq 0\} \quad (3.9)$$

As the operators  $\mathbf{C}$  and  $\mathbf{g}$  are linear, it is easy to demonstrate the convexity of  $\mathcal{K}$ . This would mean that a convex combination<sup>4</sup> of non-penetrating snapshots would naturally satisfy the inequality constraint. The implication here is that if the reduced model explores the convex hull<sup>5</sup> of the training set snapshots, the inequality constraint need

<sup>4</sup>A convex combination is a linear combination with non-negative coefficients that sum up to unity

<sup>5</sup>Convex hull of a set of points (or snapshots) is the set containing all convex combinations of the points. In other words, it is also the smallest convex set that envelopes the set of points.



**Figure 3.13:** Illustration of non-penetrating property of convex hull. The non-penetration constraint in (3.9) is linear, convex and is made of segments corresponding to each element of the constraints  $\mathbf{C}\mathbf{u} = \mathbf{g}$ . In a high-dimensional problem, each segment of the constraint can be thought of as a hyperplane. The convex hull, which is simply a convex enclosure of snapshots, can also be thought of as a high-dimensional set.

not be explicitly enforced. This notion is visually shown in Fig. 3.13, where the convex hull of training set snapshots does not violate the convex non-penetration constraint. This property of convexity can be exploited to build efficient reduced models, if the following hypothesis holds:

### Hypothesis 3.1 : Convex Subset hypothesis

Given a parametrized high-dimensional inequality-constrained problem with a convex feasible set  $\mathcal{K}$ , its solutions lie in a low-dimensional convex subset inside  $\mathcal{K}$ .

The Convex Subset hypothesis is an extension of the idea behind the low-rank hypothesis to inequality-constrained problems with convex feasible regions. This hypothesis allows us to explore a solution set limited to the convex hull defined by the snapshots without worrying about the non-penetration conditions. Note that the convexity of feasible region assures that any solution in the convex hull of displacement snapshots still lie inside the feasible region, independent of the Hypothesis 3.1. Therefore, in this approach based on the convex hull of the dictionary, the displacement field candidates built using convex combinations of the primal dictionary  $\mathbf{D}_p$  will be explored:

$$\begin{aligned} \mathbf{u} &\approx \mathbf{D}_p \hat{\mathbf{u}} \\ \text{s.t. } \mathbf{1}^T \hat{\mathbf{u}} &= 1 \\ \hat{\mathbf{u}} &\geq \mathbf{0} \end{aligned}$$

As the non-penetration constraint will be satisfied naturally by all candidates in the convex hull of dictionary snapshots, a dictionary-based approximation can be computed using coefficients  $\hat{\mathbf{u}}$  that satisfy equilibrium equations. Therefore, the objective is to minimize the following residual:

$$\mathbf{r}(\hat{\mathbf{u}}, \hat{\boldsymbol{\lambda}}) = \mathbf{K}(\mu) \mathbf{D}_p \hat{\mathbf{u}} + \mathbf{C}^T(\mu, \mathbf{u}) \mathbf{D}_d \hat{\boldsymbol{\lambda}} - \mathbf{f}(\mu)$$

As mentioned before, a monolithic dictionary will be used in this problem, in which each column consists of contact pressure snapshot stacked under the displacement snapshot. This means the coefficients  $\widehat{\mathbf{u}}$  and  $\widehat{\boldsymbol{\lambda}}$  will be same and will be denoted by  $\boldsymbol{\alpha}$  here onwards. Use of monolithic dictionary not only reduces the number of unknowns, but also makes the problem of minimizing the residual  $\|\mathbf{r}(\boldsymbol{\alpha}, \boldsymbol{\alpha})\|$  well-defined. Moreover, the residual can be projected on a low-dimensional subspace of the  $\text{colsp}(\mathbf{D}_p)$ , described by the matrix  $\mathbf{B}$  containing its basis, while also adding sparsity constraints to the unknown  $\boldsymbol{\alpha}$ , like in (3.3). Thus, the problem can be stated as:

$$\min \|\boldsymbol{\alpha}\|_p \quad (3.10a)$$

$$\text{s.t. } \|\mathbf{B}^T \mathbf{r}(\boldsymbol{\alpha}, \boldsymbol{\alpha})\| < \varepsilon \quad (3.10b)$$

$$\mathbf{1}^T \boldsymbol{\alpha} = 1 \quad (3.10c)$$

$$\boldsymbol{\alpha} \geq \mathbf{0} \quad (3.10d)$$

As the constraint (1.9c) is not imposed directly in this approach, the non-penetration condition is satisfied only because of the convexity of the feasible region. Thus, the equality condition  $\mathbf{C}_A(\mu, \mathbf{u})\mathbf{u} = \mathbf{g}_A(\mu, \mathbf{u})$  for active constraints is also not imposed explicitly. Consequently, the complementary slackness KKT condition (1.9d) may not be satisfied exactly. This will be seen in the upcoming numerical examples.

The **nnFOCUSS** Algorithm: To solve the convex hull approximation problem (3.10), the non-negative version of the FOCUSS, the nnFOCUSS algorithm [47, 75] is used. The nnFOCUSS algorithm works by computing an appropriate relaxation parameter for each iterative update that maintains non-negativity. After computation of new coefficients at Step 5 of the Alg. 3.1, the relaxation step of nnFOCUSS algorithm is performed as follows:

```

if  $\min(\boldsymbol{\alpha}^k) < 0$  then
     $\Delta\boldsymbol{\alpha} = -(\boldsymbol{\alpha}^k - \boldsymbol{\alpha}^{k-1})^-$ 
     $\boldsymbol{\alpha}^k \leftarrow \boldsymbol{\alpha}^{k-1} + \min\left(\frac{\boldsymbol{\alpha}^{k-1}}{\Delta\boldsymbol{\alpha}}\right) \Delta\boldsymbol{\alpha}$  ▷ Element wise division
end if

```

The algorithm must be initialized using a non-negative coefficients  $\boldsymbol{\alpha}^0$ , which is computed using the non-negative least squares [76] solution (which is not sparse, in general). The rest of the algorithm is same as standard FOCUSS. The problem (3.10) can be solved by plugging the following inputs in the nnFOCUSS algorithm:

$$\mathbf{D} = \begin{bmatrix} \mathbf{B}^T \mathbf{K}_{\text{mono}} \\ \mathbf{1}^T \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \mathbf{B}^T \mathbf{f}(\mu) \\ 1 \end{bmatrix}, \quad \boldsymbol{\alpha}^0 = \text{nnls}(\mathbf{D}, \mathbf{x})$$

where  $\mathbf{K}_{\text{mono}}$  is the operator for the monolithic residual  $\mathbf{r}(\boldsymbol{\alpha}, \boldsymbol{\alpha})$ :

$$\mathbf{K}_{\text{mono}} = \mathbf{K}(\mu)\mathbf{D}_p + \mathbf{C}^T(\mu, \mathbf{u})\mathbf{D}_d$$

`nls` indicates the non-negative least squares. An open source implementation of `nls` provided in SciPy package [77] is used.

In this thesis, analytical validation for the convexity of feasible set and Hypothesis 3.1 will not be given. Instead, the following tests, that apply the leave-one-out approach on the training set snapshots, are proposed to analyze if these conditions hold within a reasonable error:

Test A. Even if the operators  $\mathbf{C}(\mu, \mathbf{u})$  and  $\mathbf{g}(\mu, \mathbf{u})$  are non-linear and the feasible region  $\mathcal{K}(\mu)$  shows a dependence on the parameter  $\mu$ , it may be reasonable to assume a nearly convex feasible set if the contact pairs do not change drastically. For such cases, the convexity of feasible set can be checked numerically using the training set snapshots and the corresponding non-linear operators. Given the constraint,  $\mathbf{C}(\mu, \mathbf{u})\mathbf{u} - \mathbf{g}(\mu, \mathbf{u}) \leq 0$ , the cross penetration (CrPen) error of the  $k$ -th dictionary element  $\mathbf{d}$  across all constraints encountered in training set is computed as follows:

$$\epsilon_{\text{CrPen}}[k] = \max [\mathbf{C}(\mu, \tilde{\mathbf{d}})\tilde{\mathbf{d}} - \mathbf{g}(\mu, \tilde{\mathbf{d}})]^{+,0} \quad \forall \tilde{\mathbf{d}} \in \tilde{\mathbf{D}}_p \quad (3.11)$$

with  $\mathbf{d} = \mathbf{D}_p[:, k]$  and  $\tilde{\mathbf{D}}_p = \mathbf{D}_p \setminus \mathbf{d}$ . If all values of  $\epsilon_{\text{CrPen}}$  are sufficiently small, it would be reasonable to assume the feasible set  $\mathcal{K}(\mu)$  to be convex. The notation  $[\cdot]^{+,0}$  has the same meaning as in (3.6).

Test B. The Hypothesis 3.1 can be verified by approximating the left out snapshot with the convex hull of the rest of the snapshots using least-square criteria. The approximation error, referred to here as the convex hull least square (CHLS) error, of the leave-one-out approach can be used to assess the existence of hypothetical low-dimensional convex set. Each dictionary element  $\mathbf{d}$  is approximated using the convex hull of  $\tilde{\mathbf{D}}_p$  and the error is recorded:

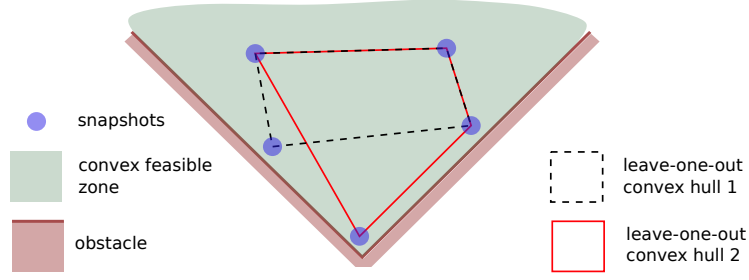
$$\epsilon_{\text{CHLS}}[k] = \min_{\boldsymbol{\alpha}} \frac{\left\| \tilde{\mathbf{D}}_p \boldsymbol{\alpha} - \mathbf{d} \right\|_2}{\|\mathbf{d}\|_2} \quad (3.12)$$

s.t.  $\mathbf{1}^T \boldsymbol{\alpha} = 1$

A high value of  $\epsilon_{\text{CHLS}}$  might occur due to lack of low-dimensional convexity or due to scattered training set data. The  $\epsilon_{\text{CHLS}}$  can be visualized using a low-dimensional problem shown in Fig. 3.14, where a point mass moves is permitted to move between a V-shaped obstacle. When leave-one-out approach is applied to a set of five snapshots,  $\epsilon_{\text{CHLS}}$  is high when the snapshot in the corner is left out. On the other hand, a low value of  $\epsilon_{\text{CHLS}}$  assures that it is possible to reconstruct reasonably using convex hull of the dictionary.

### 3.3.2.1 Illustrative example: elastic rope-obstacle problem

The convex hull approximation is first demonstrated in a problem where constraint operators are constant, so that convexity of feasible space is assured. The elastic rope-obstacle problem from Section 2.2 is considered, but with a different parametrization.



**Figure 3.14:** Illustration of the Test B with a point mass moving in a V-shaped obstacle forming a convex feasible region. Two leave-one-out instances and the convex hull of the corresponding  $\widetilde{\mathbf{D}}_p$  are shown.  $\epsilon_{\text{CHLS}}$  is a measure of distance between the left out snapshot and the convex hull of  $\widetilde{\mathbf{D}}_p$ . The leave-one-out instance 1 has a higher  $\epsilon_{\text{CHLS}}$  than the instance 2.

In this case, the elasticity  $\nu(x)$  of the rope is a function of parameter  $\gamma_1$ .

$$\begin{aligned} \nu(x)\nabla^2 u(x) &= f \quad \text{on } x \in [0, 1] \\ u(0) &= u(1) = 0 \\ u(x) &\geq -0.2(\sin(\pi x) - \sin(3\pi x)) - 0.5 \end{aligned} \quad (3.13)$$

where

$$\nu(x) = \begin{cases} \gamma_1 & \text{if } x < 0.5 \\ 30 & \text{otherwise} \end{cases} \quad \text{on } \gamma_1 \in [10, 50]$$

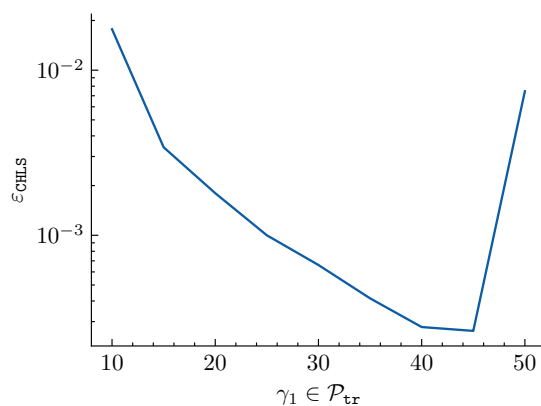
Training set:  $\gamma_1 \in \mathcal{P}_{\text{tr}} = \{10, 15, 20 \dots 45, 50\}$

Validation set:  $\gamma_1 \in \mathcal{P}_{\text{val}} = \{12.5, 17.5, 22.5 \dots 42.5, 47.5\}$

The monolithic dictionary with 9 snapshots corresponding to training set defined above is computed in offline stage. The constraint is linear and independent of the parameter, and therefore, the feasible region is convex, which eliminates the necessity of Test A. Hence, only the Hypothesis 3.1 is tested using leave-one-out approach (Test B). The test results are shown in Fig. 3.15, indicating that hypothesis holds to a reasonable level.

The low-rank matrix  $\mathbf{B}$  on which the residual will be projected is computed using left singular vectors of the truncated SVD  $\mathbf{B} \leftarrow \text{svd}(\mathbf{D}_p, \delta)$ , with tolerance  $\delta = 10^{-7}$ . The reconstruction errors and resultant sparsity of the solution in the training set and a validation set is shown in Fig. 3.16. The training set is reconstructed within near numerical precision. In the validation set, the algorithm chooses the training set points that are nearest to  $\gamma_1$  to create the best possible reconstruction with the given dictionary.

As discussed earlier, the KKT conditions, except the non-negativity of contact pressure, is not explicitly imposed. However, in this particular case with perfectly convex feasible region, one would expect that at least the non-penetration would be satisfied accurately. However, the linear combinations are not perfectly convex as the condition (3.10c) is satisfied only to a certain precision. The observed orders of these quanti-

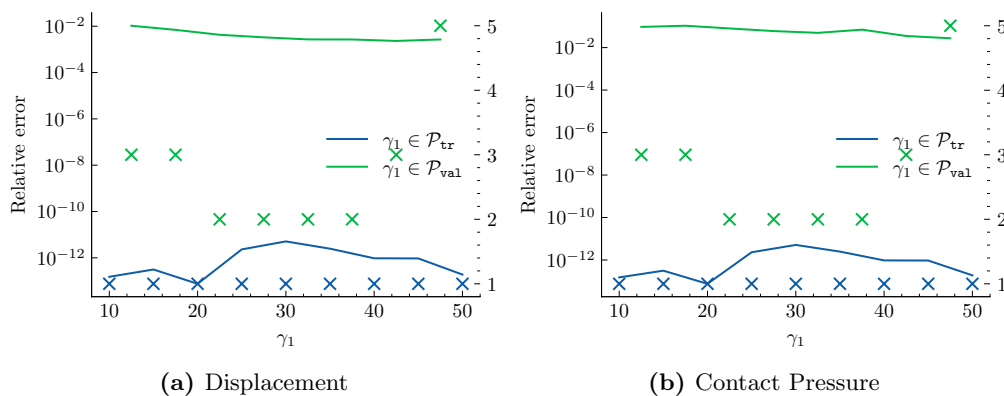


**Figure 3.15:** Test **B**: testing the Hypothesis 3.1 for the given training set of elastic rope-obstacle problem (3.13). Test **A** for convexity of feasible set is not necessary as the constraints are linear.

ties during reconstruction are given in Table 3.2. The validation set reconstructions in fact show penetration and non-zero complementary slack in the order similar as that of the deviation from convex combinations i.e.  $|\mathbf{1}^T \boldsymbol{\alpha} - 1|$ .

	$\gamma_1 \in \mathcal{P}_{\text{tr}}$	$\gamma_1 \in \mathcal{P}_{\text{val}}$
$ \mathbf{1}^T \boldsymbol{\alpha} - 1 $	$\mathcal{O}(10^{-13})$	$\mathcal{O}(10^{-3})$
Penetration	$\mathcal{O}(10^{-16})$	$\mathcal{O}(10^{-4})$
Complementary Slack	$\mathcal{O}(10^{-15})$	$\mathcal{O}(10^{-3})$

**Table 3.2:** Observed values of various compliances for the elastic rope-obstacle problem with parameter  $\gamma_1$ .



**Figure 3.16:** Convex hull reconstruction errors for elastic rope-obstacle problem with parameter  $\gamma_1$  using  $\mathbf{B} \leftarrow \text{svd}(\mathbf{D}_p, \delta = 10^{-7})$ . Crosses  $\times$  indicate achieved sparsity (w.r.t. right y-axis). Primal and dual sparsity are naturally equal as the dictionary is monolithic.

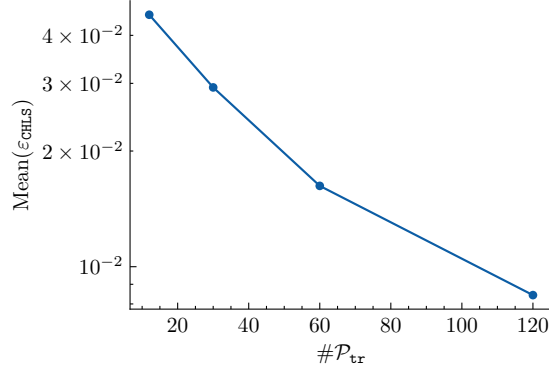
### 3.3.2.2 Demonstration on over-complete dictionaries of the Hertz problem

Hertz problem with same parametrization and training sets from Section 3.3.1.1 is solved using the convex hull approximation. For projecting the residual on a low-rank subspace, the matrix  $\mathbf{B}$  is computed in two different ways:

$$\mathbf{B} \leftarrow \text{svd}(\mathbf{D}_p, \delta) \quad \text{or} \quad \mathbf{B} \leftarrow \text{orth}(\mathbf{D}_p \mathbf{R})$$

where  $\mathbf{R}$  is a matrix with random entries from a uniform distribution (3.3c). As discussed in Section 3.2.2, this matrix is used to create a smaller matrix by randomly combining dictionary columns.

As the constraint operators  $\mathbf{C}$  and  $\mathbf{g}$  show non-linear dependence w.r.t.  $\mathbf{u}$ , it is necessary to test the convexity of the feasible set  $\mathcal{K}(\mu)$  (Test A). All CrPen errors ( $\epsilon_{\text{CrPen}}$ ) were found to be quite small, in the order of  $\mathcal{O}(10^{-7})$  for all combinations of  $\mathbf{d}$  and  $\tilde{\mathbf{d}}$  in (3.11). Test B is performed to check the Hypothesis 3.1, with results shown in Fig. 3.17. The reconstruction errors for the reduced model can be expected to be more than the  $\epsilon_{\text{CHLS}}$  shown in the figure.

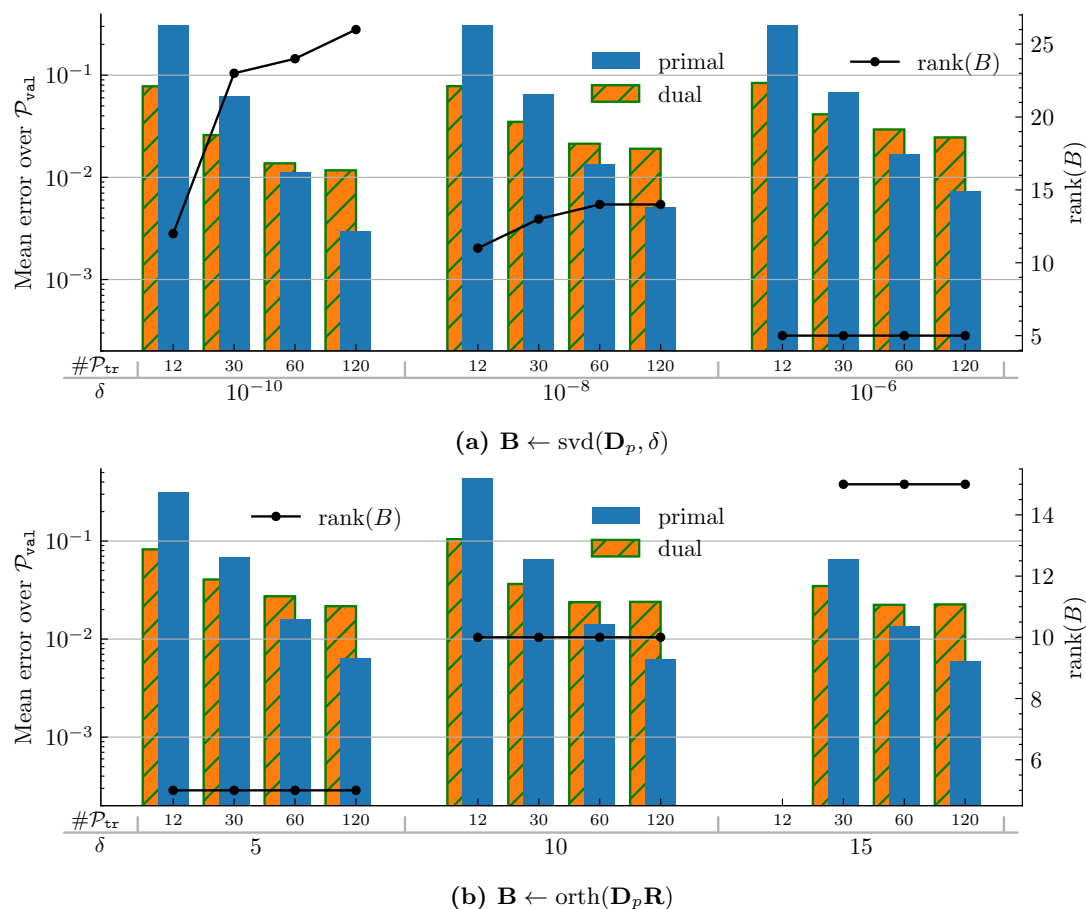


**Figure 3.17:** Test B: testing the Hypothesis 3.1 for the various training sets of the Hertz problem

For large dictionaries, it makes sense to compare the performance of the method using various  $\delta$  and sizes of  $\mathbf{R}$ . The mean reconstruction errors over the validation set vs size of dictionaries can be seen in Fig. 3.18. The use of randomized matrices  $\mathbf{R}$  naturally generates more error than truncated SVD, as the truncated SVD selects the best representative subspace for the given training set. Similar to the greedy active-set method of Alg. 3.2, the improvement in error flattens out for large dictionaries. On the other hand, the number of iterations and computation time seems to be nearly independent of the dictionary size.

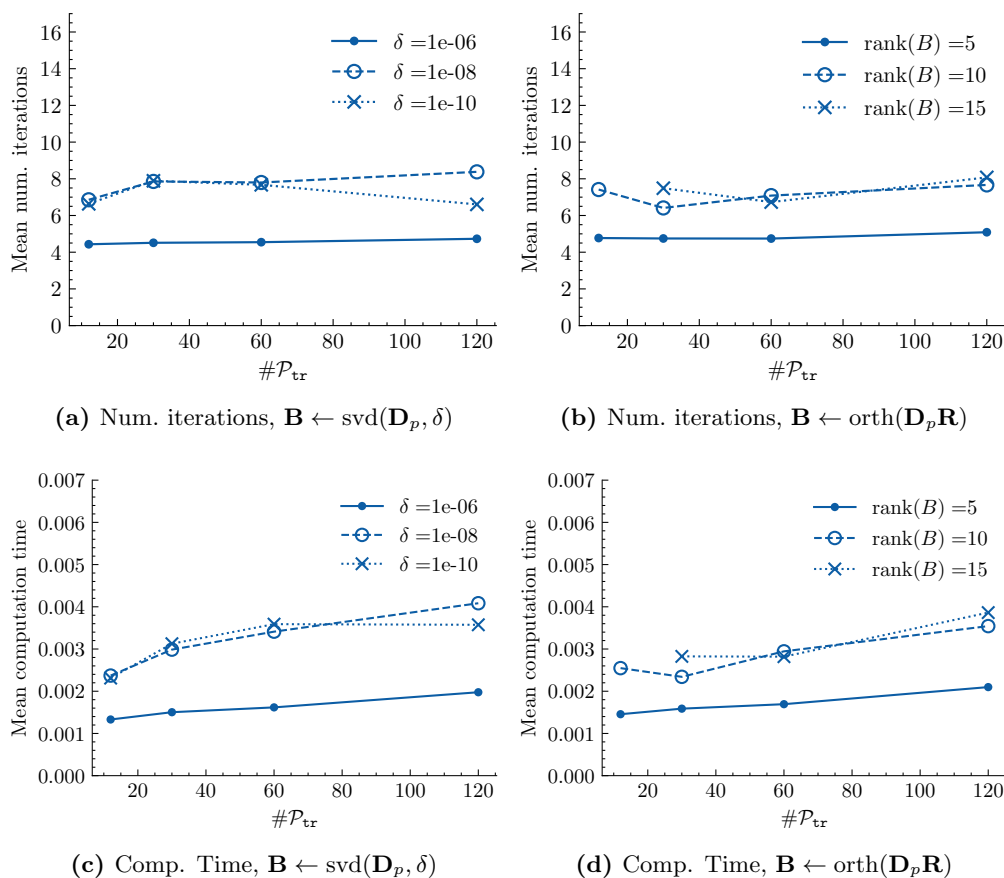
### 3.3.2.3 The non-convex case of the Ironing problem

The convexity test for the feasible set (Test A) is performed on the Ironing problem. Since, the contact zone changes significantly, the test performed using the training set



**Figure 3.18:** Convex hull reconstruction errors for the Hertz problem using different dimensions of  $\mathbf{B}$ . The penetration in all reconstruction cases is observed to be in order of  $10^{-8}$ . The case of  $\text{rank}(B) = 15$  and  $\#\mathcal{P}_{\text{tr}}$  is not possible, hence not shown in (b). Primal and dual errors are computed using  $\mathcal{H}^1$  and  $\mathcal{L}^2$  norm, respectively.

snapshots fails as expected. The CrPen error was found in  $\mathcal{O}(10^{-1})$ . Therefore, the ironing problem cannot be solved using the current formulation of convex hull approach. The non-convexity can also be demonstrated by a convex combination that violates the non-penetration condition significantly, as shown in Fig. 3.20. For the two snapshots shown in this figure, the slab surface is deformed at different positions. The convex combination where the two displacement field snapshots were equally weighted, generated a deformed configuration with the slab being deformed at the two locations corresponding to each of the snapshots. But no deformation was generated at the position of the iron, causing penetration at that position. In this case, local behaviour of the displacement field leads to the non-convexity of the ironing problem.



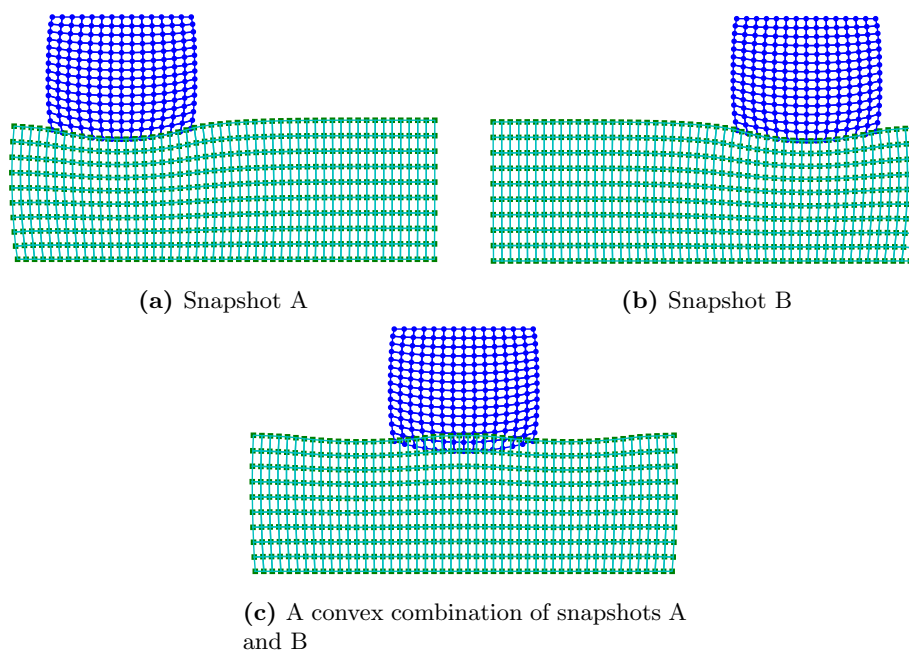
**Figure 3.19:** Mean number of iterations and computation times for Hertz problem using convex hull exploration with different dimensions of  $\mathbf{B}$ . Note that computation time excludes time for construction of non-linear operators.

### 3.4 Discussions

The application of dictionary methods was proposed to mitigate the challenges posed by linear inseparability of the contact pressure. As the lack of local information prevents the creation of a robust RB, over-complete dictionaries have the potential to preserving local information computed in the offline stage, thereby providing a better chance at resolving the position of contact more effectively. The same is demonstrated by building reasonably accurate reconstructions using dictionary methods.

For the reconstructions of Hertz problem using the greedy active-set algorithm, the reconstruction errors decay consistently when richer training sets are used, and the computational time per iteration seems to be fairly constant, but change slightly with the primal rank.

The case of ironing problem where the contact position changes significantly is also considered. However, the decay of reconstruction errors vs. size of dictionary curves were



**Figure 3.20:** Demonstration of the non-convex nature of the ironing problem. Two snapshots from training set of the ironing problem and a convex combination of the snapshots that violates the non-penetration condition

found to be flatter compared to the Hertz problem. This is probably due to stronger dependence of contact position on the parameter of the problem. Though the dictionary size didn't have any significant influence on the computational time per iteration, the number of iterations increased significantly for larger dictionaries in case of ironing problem.

Comparing the two dictionary methods: the greedy active-set algorithm and the convex hull approximations, the reconstruction errors of the Hertz problem were better using the greedy active-set algorithm (see Figs. 3.7a and 3.18). This observation can be possibly attributed to the use of the monolithic dictionary in case of the convex hull approximations; which forced common interpolation coefficients for the displacement as well as contact pressure. In particular, the convex hull approach shows large errors for small training sets, as the convex hull of a small set of snapshots might be a very small subregion in the entire solution space.

Another possible source of error in convex hull approximations is that the snapshots in the dictionary may not be the best ones to explore the low-dimensional convex subset. On the other hand, the computational time (and number of iterations) for the convex hull approach seems to be more independent of dictionary size, compared to the greedy active-set algorithm (see Figs. 3.7b, 3.7c and 3.19), thanks to the elimination of active set selection process.

In case of the convex hull approach, the reconstruction error seems to be more sen-

sitive to the size of dictionary than the truncation tolerance/rank of the orthogonal basis  $\mathbf{B}$ . Also, for the greedy active-set method, the sensitivity to truncation seems to be low between  $10^{-10}$  and  $10^{-8}$ . Moreover, in both approaches, the computational time per iteration is nearly flat when the rank is low. Thus, one can infer that that large dictionaries and moderate truncation tolerances / rank of  $\mathbf{B}$  are an optimal choice.

One could argue that the number of iterations increases significantly with size of dictionary in case of the ironing problem solved using the greedy active-set method (Figs. 3.12b and 3.12c), however, this is related to the inherent non-linearity of the contact problem. The number of iterations could possibly be reduced by using smarter initialization of the contact pairs according to the queried parametric value rather than the initializing with reference state. For the convex hull approach, the primary disadvantage is the assumption of convexity of the feasible region, limiting the applicability to convex problems. The non-convex nature of ironing problem meant that convex hull approximation method could not be used in this case.

## 3.5 Perspectives

### 3.5.1 Efficient construction of constraints

In the Alg. 3.2, Step 5 can be the most expensive step if full operators are built from scratch and then projected on the reduced bases. The constraint operators  $\hat{\mathbf{C}}$  and  $\hat{\mathbf{g}}$  show non-linear dependence on the displacement field  $\hat{\mathbf{u}}$  and therefore must be reconstructed in every iteration. An efficient construction of the constraint operators for low-rank methods is discussed in [13] using the Empirical Interpolation Method (EIM), where affine decompositions of the distance functions are computed offline. This enables efficient construction of constraint operators in online phase by sampling distance functions at a few points on the contact surface, followed by computing the coefficients of the affine form.

The distance function  $k(\mu, \mathbf{v}, \mathbf{v})$  of (1.2) when expressed in discrete form (as discussed in [13]) turns out to be a sparse vector. The `spy` pattern of this vector can be a strong function of the parameter when contact area changes significantly. This is caused by the dependence of the distance function on the contact pairs established using the current configuration. Hence, it is bound to exhibit some form of linear inseparability just like contact pressure. Therefore, it is possible that linear inseparability issues might also appear in generating an effective affine decomposition.

### 3.5.2 Application of sketching methods to dictionary-based approximation

As discussed briefly in Section 3.2.2 and in detail in [73], random sketching can be used to project the residual of a dictionary-based approximation problem on a randomly sampled subspace of the dictionary. By controlling the size of the subspace, one can control the cost of solving the dictionary-based approximation problem, relatively independent

of the dictionary size. The difficulties in applying these techniques to contact pressure dictionary appear due to the inequality constraints. If random sketching methods are applied, the inequality constraints are linearly combined in random fashion leading to loss of information on violation (penetration) and satisfaction (non-penetration) of constraints.

### **3.5.3 Smart snapshot selection for Convex hull explorations**

RB methods equipped with error estimators sample the parametric space efficiently. Similarly, it might be possible to extend the same notion to efficiently sampling the convex feasible set. The convex hull approximations might be more robust if the snapshots lie on the edges/vertices of the low-dimensional convex subset presumed in Hypothesis 3.1. For e.g. convex hull approximation for the low-dimensional problem in Fig. 3.14, it is important to include the snapshot near the corner of the feasible region in the training set for a robust reconstruction. However, deducing the corners in a high-dimensional problem may not be geometrically intuitive as the low-dimensional problem in the figure.

# 4 | Exploring the contact pressure manifold

---

In Section 3.1, the philosophy of dictionary-based approximation methods for contact problems was discussed. The principal argument in favour of over-complete dictionaries was the lack of low-rank structure due to linear inseparability of the contact pressure field, which has two consequences: (a) a large number of snapshots might be necessary to truly discover the underlying structure of the contact pressure field, and (b) linear model order reduction methods cannot capture local effects of contact pressure in a low-rank subspace. The focus of Chapter 3 was on efficient handling large dictionaries of contact pressure snapshots in the online phase, but does not address the high offline cost of creating such dictionaries. This chapter focuses on the use of non-linear dimensionality reduction methods as means to improve reconstruction errors without paying the large computational cost of creating over-complete dictionaries.

## 4.1 Linear Subspaces vs. Low-dimensional Manifolds

Many of the dimensionality reduction methods work on the principal assumption that the underlying physics occupies a subset of the solution space defined in the full model. At this point, the question arises about the nature of the aforementioned subset of solution space. Low-rank methods assume that this subset is a low-dimensional linear subspace<sup>1</sup> of the full solution space. The assumption that such a low-dimensional subspace exists implies that the solution of any point in the particular parametric space may be approximated using a linear combination of a limited number of precomputed basis vectors that naturally lie in this subspace. As seen in Chapter 1, methods like Proper Orthogonal Decomposition (POD) are capable of extracting a basis for the linear subspace that describes the underlying physics.

However, in some problems the parametric solutions occupy a low-dimensional manifold, but not necessarily linear subspace. Consequently, it is not possible to compute a robust low-rank basis, as shown for the contact pressure field in Chapter 2. Nonetheless,

---

<sup>1</sup>usually the term “linear” is dropped in unambiguous cases

there exists a paradigm called non-linear dimensionality reduction that aims at computing the underlying structure by finding the associated non-linear behaviour. A subclass of non-linear dimensionality reduction known as manifold learning is quite extensively used in the analysis and visualisation of high dimensional datasets [78, 79]. The fundamental idea behind manifold learning is similar to low-rank methods and is known as the manifold hypothesis [80]. Roughly speaking, a manifold can be described as an  $n$ -dimensional space that is globally non-linear in general but locally resembles a linear vector space.

The manifold hypothesis assumes that high-dimensional data sourced from practical applications often lie in low-dimensional manifolds. The low-rank hypothesis is a special case of the manifold hypothesis involving the assumption of the low-dimensional manifold being a linear subspace. Manifold learning methods rely on finding the non-linear low-dimensional manifold within the high-dimensional data. A well-known application of manifold learning is the case of a data set containing gray-scale images of an object rotated by various angles. Linear interpolation methods like PCA fail to reduce this type of data, but manifold learning methods can extract the inherent one-dimensional space [78, 79].

To better approximate the manifold structure of a parametrized problem, partitioning of the parametric domain [81] and a local reduced basis methods [82] have been explored. Interpolation in the tangent space to the Grassmann manifold structure have been used to construct reduced models with a wide parametric range [83].

One of the most commonly applied non-linear dimensionality reduction techniques in the parametrized ROMs is the kernel methods [84–86]. Kernel interpolation methods along with greedy algorithms to create nested training sets to create efficient surrogate models are discussed in [84]. The kernel-PCA (kPCA) approach can also be used for the computation of intrinsic parameters that are not explicitly included in the mathematical model [85]. An application of kPOD to an advection-diffusion problem with inseparable nature can be found in [86], where the effectiveness of kernel-based method is shown using the higher concentration of variance in the first singular values of kPOD compared to the case of linear POD.

A known drawback of kernel-based methods and some other manifold learning techniques is that backward mapping, i.e. reconstruction of a high-dimensional entity from low-dimensional data is not as trivial as it is done in linear dimensionality reduction methods such as PCA. Usual the reconstruction is an approximation to the local tangent space of the manifold, computed using a linear combination of snapshots with weights computed by employing a minimization of a discrepancy functional criteria [87]. An improved approximation based on a quadratic tangent space is discussed in [86].

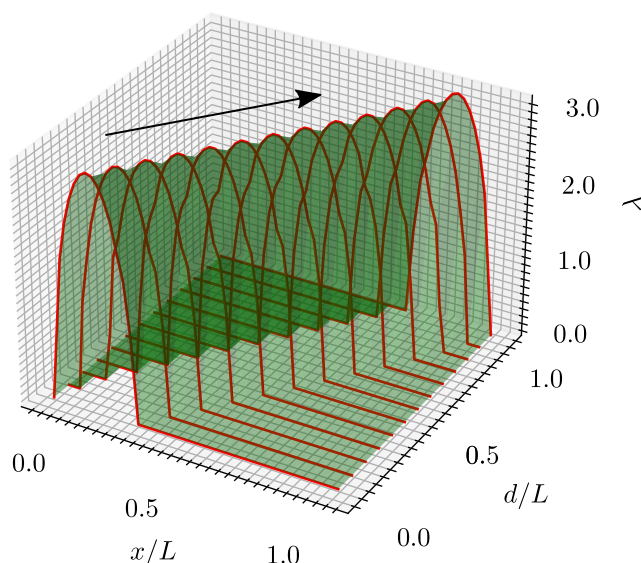
Hyperbolic systems are often associated with moving discontinuities, which also leads to lack of a linear subspace. An approach involving the transformation of the physical variable has been used in such problems to separate the discontinuous and continuous behaviour [88]. A related approach aimed at generating transformed snapshots thereby aligning the discontinuities has been demonstrated on Burger’s equation [89]. As it will be seen, the ideas given in this chapter have similar motivations of aligning the contact

pressure snapshots, but without transforming the physical variable.

### 4.1.1 Contact pressure as a non-linear manifold

In Chapter 2, qualitative and quantitative arguments about the linear inseparability of contact pressure have been made. The lack of low-dimensional linear subspace, i.e. a low-rank subspace for contact pressure was demonstrated. In this section, we discuss the idea of contact pressure lying in a low-dimensional non-linear manifold. Looking back at Fig. 2.5b and Fig. 2.11, the snapshots in respective figures have similarities that are evident intuitively. Given these similarities, there is a strong possibility that these snapshots exist in a low-dimensional non-linear manifold.

This intuitive hypothesis can be supported by Fig. 4.1. Contact pressure snapshots of the ironing problem are shown by the lines as a function of the parameter  $d$  (the horizontal displacement of the iron), and the green translucent surface is generated using the surface plot of the same data. One can intuitively see that the contact pressure is a non-linear function of just one parameter, which happens to be also the physical parameter  $d$  in this case. The fallacy of low-rank methods is that they attempt to perform a linear interpolation in direction  $d$ , when in fact the relationship is non-linear. The idea that will be exploited in this chapter is that an appropriate non-linear transformation applied on this surface in a way that allows interpolation along the “aligned” direction shown in the figure. This will possibly improve the linear separability and reducibility.



**Figure 4.1:** Visualization of contact pressure non-linearity in ironing problem. The dimension  $x \in \Gamma$  is the coordinate on contact surface, and  $d \in \mathcal{P}$  is the coordinate on the parametric domain. The arrow shown in the figure points to a direction in which snapshots are “aligned” along which the snapshots are, possibly, linearly separable.

## 4.2 Dimension warping

The core idea in manifold learning methods (such as kPCA) is that linearly inseparable data can be appropriately mapped into a higher dimension where it is relatively more separable. The mapping to high-dimensional space has to be essentially non-linear, since a linear mapping will not resolve the inseparability issues.

Consider a non-linear mapping from the spatial domain of contact surface  $\Gamma$  to a warped domain  $\tilde{\Gamma}$ . The warped domain  $\tilde{\Gamma}$  need not admit a physical interpretation.

$$\begin{aligned} \varphi(\mu, \cdot) : \Gamma &\mapsto \tilde{\Gamma} \\ \zeta &\leftarrow \varphi(\mu, x) \end{aligned} \quad (4.1)$$

where  $\zeta$  is a coordinate in  $\tilde{\Gamma}$ . The notation  $\tilde{x}(\mu, \zeta)$  is used to indicate  $x$  that maps to  $\zeta$ .

A qualitative argument can be made regarding an effective linear behaviour in  $\tilde{\Gamma}$  space. The mapping  $\varphi$  must warp the spatial dimension so that the regions with non-zero contact pressure must be mapped to a common warped region, and the gap (non-contact) regions must be mapped to other regions in warped space. In other words, let the warped space be a union of disjoint sets,  $\tilde{\Gamma} = \tilde{\Gamma}_c \cup \tilde{\Gamma}_{/c}$ , then for every pair  $(x, \zeta)$  the mapping  $\varphi$  must satisfy the following:

$$\varphi(\mu, \cdot) : \Gamma \mapsto \tilde{\Gamma} \mid \begin{cases} \zeta \in \tilde{\Gamma}_c & \text{if } \lambda(\mu, x) > 0 \\ \zeta \in \tilde{\Gamma}_{/c} & \text{if } \lambda(\mu, x) = 0 \end{cases} \quad (4.2)$$

where  $\tilde{\Gamma}_c$  and  $\tilde{\Gamma}_{/c}$  are the regions where contact and no-contact zones are mapped respectively. Any function  $f$  defined on  $\Gamma$  has an equivalent  $\tilde{f}$  defined on  $\tilde{\Gamma}$  domain and can be related to each other with the following relation:

$$\tilde{f}(\zeta) = f(x)$$

### Remark 4.1 : Preserving the split of contact and gap zones

Extending the above qualitative argument to the case where the contact pressure snapshots have multiple disjoint regions of contact and gap zones, each subregion should be mapped to corresponding subregions in warped space. For e.g., in the ironing problem, each snapshot can be split into three zones: one contact zone and two gap zones. Thus, there should be three disjoint zones in warped space corresponding to each zone of the original space.

### 4.2.1 Dynamic Time Warping

It is not straightforward to compute the mapping  $\varphi(\mu, \cdot)$  and the warped space  $\tilde{\Gamma}$  with the desired properties discussed in the previous section. Algorithms that identify regions of similarities between two signals can prove to be useful to generate mappings that

improve separability. One such method, known as Dynamic Time Warping (DTW), is an algorithm built to measure similarity between two discrete signals containing similar but displaced features. The algorithm is named so because it was built for measuring similarities in temporal sequences involving speech or motion [90, 91]. Other methods exist that deals with sequence similarities, such as elastic matching [92]; but they are not studied in this thesis. In this section, we illustrate the application of DTW to generate mappings that satisfy the conditions specified in (4.2).<sup>2</sup> Though only unidimensional DTW has been explored here, it is extendable to higher dimensions [93].

The DTW algorithm works on two given signals and computes similarities based on distances between pairs of points. It matches each point on the first signal to at least one point on the second, and vice-versa. A detailed explanation of DTW can be found in [90] and Python implementations can be found in the `dtw` module [94]. The mapping  $\varphi(\mu_1, x)$  and  $\varphi(\mu_2, x)$  generated by DTW is described by indices of the discretization of the coordinate  $x$ . This is illustrated using a simplified example.

**An illustration of DTW:** Consider a set of  $n$  discrete points  $\mathbf{x}$  in  $[0, 1]$  and two discrete vectors  $\boldsymbol{\lambda}_k \in \mathbb{R}^n, k \in \{1, 2\}$ . For a simple illustration, the vectors are chosen such that each has a peak (with value 1.0) at different positions surrounded by vanishing areas (values are given in Table 4.1a). The DTW alignments (computed using [94]) are given by indices  $I_k \in \mathbb{R}^{\tilde{n}}$ . The index at  $I_1[m]$  is matched to the index at  $I_2[m]$  of respective  $\boldsymbol{\lambda}$  vectors (in Python notation). Usually  $\tilde{n} \geq n$ , where the equality holds if alignment happens to be a one-to-one mapping.

The alignments are given in Table 4.1b and are shown graphically in Fig. 4.2. One can see that the alignments connect not only the peaks of the two vectors, but also the vanishing areas on the either sides as well. The alignments  $I_k$  are equivalent to  $\varphi(\mu_k, \cdot)$  of (4.3), the mapping to warped coordinate  $\zeta$ . The warped snapshots  $\tilde{\boldsymbol{\lambda}}_k = \boldsymbol{\lambda}_k[I_k]$  and the warped spatial coordinate  $\tilde{\mathbf{x}}_k = \mathbf{x}[I_k]$  of dimension  $\tilde{n}$  are shown in Fig. 4.3. The DTW warped coordinate  $\zeta$  can be simply expressed as set of indices  $\{1, 2 \dots \tilde{n}\}$ .

Variables	Input Values	Alignments	Observed values
$\mathbf{x}$	[0, 0.25, 0.5, 0.75, 1]	$I_1$	[1, 1, 1, 2, 3, 4, 5]
$\boldsymbol{\lambda}_1$	[0, 1.0, 0, 0, 0]	$I_2$	[1, 2, 3, 4, 5, 5, 5]
$\boldsymbol{\lambda}_2$	[0, 0, 0, 1.0, 0]		

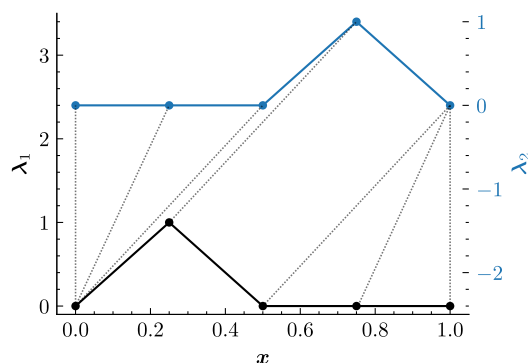
(a) Data
(b) DTW alignments

**Table 4.1:** Input variables values and the observed DTW alignments for the illustrative example

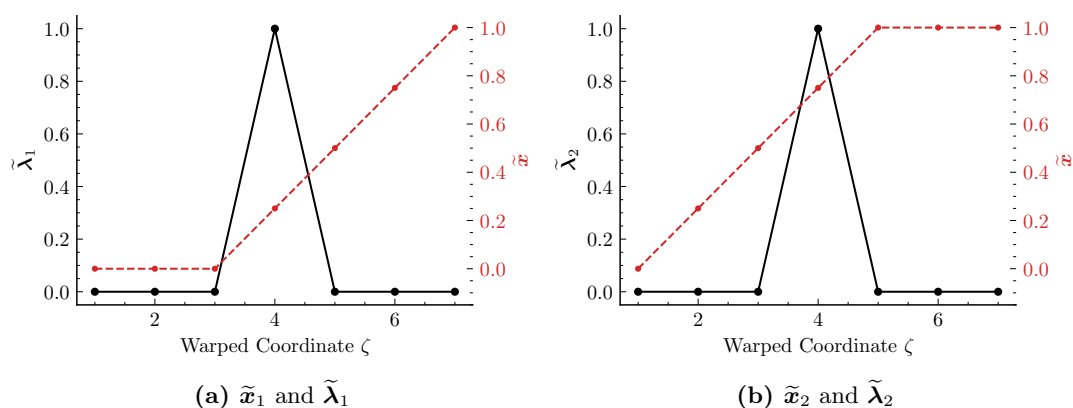
Though a detailed explanation of the DTW algorithm is beyond the scope of this thesis, the major features of uni-dimensional DTW are listed here:

1. The first (and the last) points of the two input vectors are always connected, even

<sup>2</sup>Though the DTW algorithm was originally intended for temporal data, here it will be used to warp spatial dimension.



**Figure 4.2:** DTW illustration: Vectors  $\lambda_1$  and  $\lambda_2$  and their DTW alignments (dotted). Note that the vectors are offset in  $y$ -direction for illustration purposes.



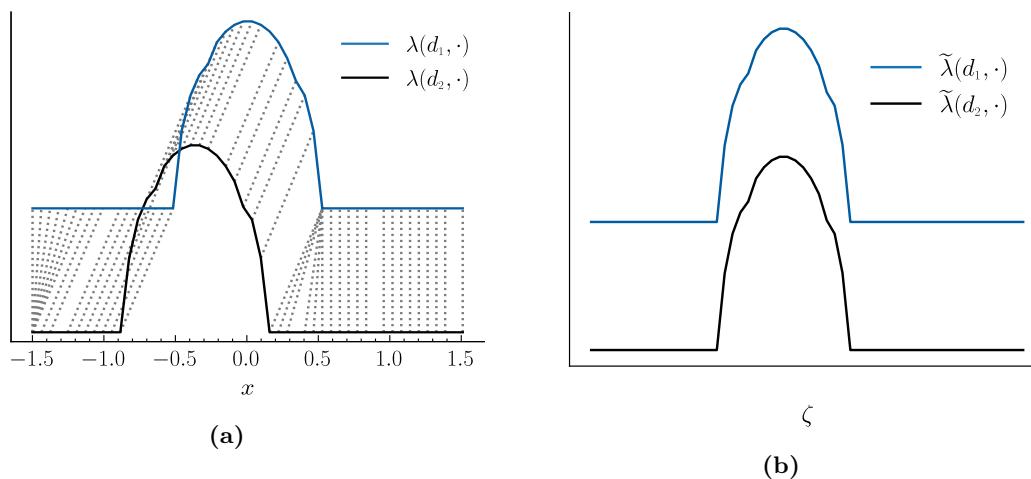
**Figure 4.3:** DTW illustration:  $\tilde{\lambda}$  and  $\tilde{x}$  vs. warped coordinate  $\zeta$

if they are connected to other points. In other words,  $I_k[1] = 0$  and  $I_k[\tilde{n}] = n$  for any inputs.

2. The alignments give indices in a monotonically non-decreasing manner i.e.  $I_k[j+1] \geq I_k[j]$ , meaning that the alignments shown in Fig. 4.2 never cross each other.
3. Roughly speaking, the alignments are generated such that the distance between warped vectors  $\tilde{\lambda}_1$  and  $\tilde{\lambda}_2$  is minimized. This is evident in the alignments and the warped vectors shown in Fig. 4.2 and Fig. 4.3.

Now, the application of DTW algorithm to the contact pressure curves of the ironing problem is considered. The alignments for two snapshots corresponding to parametric values  $d_1$  and  $d_2$  are shown in Fig. 4.4a and the warped contact pressure curves are shown in Fig. 4.4b. Notice that the alignments were able to connect respective contact and no-contact zones of both curves. Also, in the warped spatial dimension, respective contact and no-contact zones of both the warped snapshots are the same. In other words, condition (4.2) is satisfied. Also, looking back at the “aligned” direction in Fig. 4.1, we

can say that the warping has rendered this direction orthogonal to the warped coordinate  $\zeta$ , thereby making linear combinations useful.



**Figure 4.4:** (a) Alignments and (b) warped snapshots computed by DTW between two contact pressure snapshots of ironing problem for parametric values  $\mu_1 = d_1$  and  $\mu_2 = d_2$ . Notice that alignments in original space are not bijective.  $y$ -offset between the two curves is given only for illustrative purpose. Values on  $\zeta$  axis are not given as they do not carry a physical significance, and also because DTW computes warped space on indices of the discretization rather than physical dimension.

#### Remark 4.2 : Non-linear nature of DTW mapping

The mapping generated by DTW is linear w.r.t. individual snapshots. However, the mapping is not only different for each snapshot, it depends on the pair of snapshots taken into account. These properties mean that the mapping is essentially non-linear.

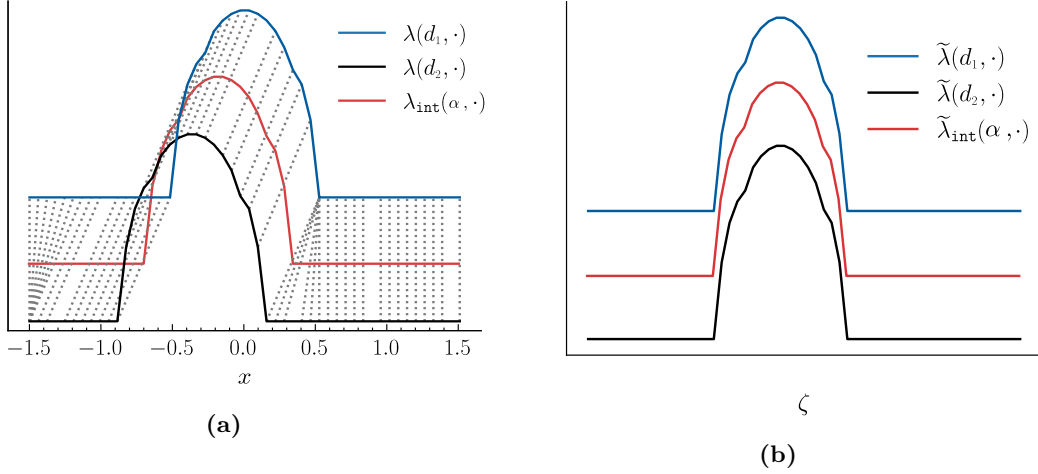
### 4.2.2 Interpolating in warped space

Although the DTW algorithm is built to compute similarities between signals, we intend to use it as an interpolation tool. Interpolations using DTW have been previously performed to reconstruct missing frames in fast motion detection [95]. The objective of performing DTW on pairs of contact pressure snapshots is to use the DTW alignments and warped space for exploring the contact pressure manifold without generating intermediate snapshots directly with FE solver.

Fig. 4.5 shows a demonstration of an interpolation in the original and the warped spatial dimension. Though interpolating along the alignments in the original dimension looks intuitive (Fig. 4.5a), only the interpolation in warped space is a straightforward linear operation (Fig. 4.5b). Coefficients  $\alpha$  and  $1 - \alpha$  are used as weights for each

snapshot in this operation, and the case for  $\alpha = 0.5$  is shown in the figure. In principle, the value of the parameter  $\mu$  that corresponds to the interpolated snapshot is not known.

As interpolation is performed in the warped dimension, it means that a backward mapping is also needed to transform the interpolated curve back into the original spatial dimension. However, the mapping  $\varphi(\mu, \cdot)$  is a function of the parameter and the mapping for the interpolated curve needs to be computed.



**Figure 4.5:** Interpolation along (a) DTW alignments and (b) interpolation in warped space, for parametric values  $\mu_1 = d_1$  and  $\mu_2 = d_2$ . Notice that unlike in (a), the interpolation in (b) is a linear operation.

Formally, the interpolations on the two snapshots, and the mappings for parametric values  $\mu_1$  and  $\mu_2$  can be written as:

$$\tilde{\lambda}_{\text{int}}(\alpha, \zeta) := \alpha \tilde{\lambda}(\mu_1, \zeta) + (1 - \alpha) \tilde{\lambda}(\mu_2, \zeta) \quad (4.3a)$$

$$\tilde{x}_{\text{int}}(\alpha, \zeta) := \alpha \tilde{x}(\mu_1, \zeta) + (1 - \alpha) \tilde{x}(\mu_2, \zeta) \quad (4.3b)$$

$$\varphi_{\text{int}}(\alpha, x) := \alpha \varphi(\mu_1, x) + (1 - \alpha) \varphi(\mu_2, x) \quad (4.3c)$$

where  $\lambda_{\text{int}}$  and  $\varphi_{\text{int}}$  are interpolated contact pressure and mapping respectively and  $\alpha$  is the linear interpolation parameter between the two snapshots. The interpolated contact pressure in original and warped space can be related as:

$$\lambda_{\text{int}}(\alpha, x) = \lambda_{\text{int}}(\alpha, \varphi_{\text{int}}^{-1}(\alpha, \zeta)) = \tilde{\lambda}_{\text{int}}(\alpha, \zeta) \quad (4.4)$$

Once DTW alignments are available, (4.3) and (4.4) can be easily implemented in a short routine `dtw_interp` to perform the interpolation.

#### Notation 4.1 :

There is a slight abuse of notation in (4.3), (4.4) and in interpolated quantities here onwards. The first argument of quantities  $\varphi_{\text{int}}(\cdot, \cdot)$  is the interpolation parameter  $\alpha$  and not the physical parameter  $\mu$ .

**Notation 4.2 :**

Interpolated quantities  $\circ_{\text{int}}(\cdot, \cdot)$  also depend implicitly on  $\mu_1$  and  $\mu_2$ , which are not shown in the arguments to simplify the notation.

`dtw_interp` is given in Alg. 4.1. It takes two discrete contact pressure snapshots  $\lambda_k \in \mathbb{R}^n$ ,  $k \in \{1, 2\}$  and the corresponding spatial coordinate  $\mathbf{x} \in \mathbb{R}^n$  i.e. there are  $n$  discrete dofs on the contact surface. The DTW package computes alignments and a warped space of dimension  $\tilde{n}$ . Indices  $I_k$  provide the mapping to the warped space for respective  $\lambda_k$ . The warped vectors  $\tilde{\lambda}_k \in \mathbb{R}^{\tilde{n}}$  can be linearly combined to create the interpolated vector  $\tilde{\lambda}_{\text{int}} \in \mathbb{R}^{\tilde{n}}$  and the corresponding spatial coordinate  $\tilde{\mathbf{x}}_{\text{int}} \in \mathbb{R}^{\tilde{n}}$ . The `interp1d` function in Step 5 is borrowed from the Python module `scipy.interpolate` [77]. This function is used for changing the discretization from  $\tilde{\mathbf{x}}_{\text{int}}$  to  $\mathbf{x}$  using linear piecewise interpolations. Steps 3 and 4 corresponds to (4.3a) and (4.3b) and Step 5 corresponds to (4.4).

**Algorithm 4.1 dtw\_interp**

- 
- |   |  |
|---|--|
| 1: Input: $\mathbf{x}, \lambda_1, \lambda_2, \alpha$  |  |
| 2: $I_1, I_2 = \text{dtw}(\lambda_1, \lambda_2)$  | ▷ $I_1$ and $I_2$ map indices to warped dimension                |
| 3: $\lambda_{\text{int}} = \alpha \lambda_1[I_1] + (1 - \alpha) \lambda_2[I_2]$   | ▷ $\lambda_j[I_j]$ is the warped vector $\tilde{\lambda}_j$      |
| 4: $\tilde{\mathbf{x}}_{\text{int}} = \alpha \mathbf{x}[I_1] + (1 - \alpha) \mathbf{x}[I_2]$                                    | ▷ interpolating in warped spatial coordinate                     |
| 5: $\lambda_{\text{int}} \leftarrow \text{interp1d}(\tilde{\lambda}_{\text{int}}, \tilde{\mathbf{x}}_{\text{int}})(\mathbf{x})$ | ▷ linear interpolation from $\tilde{\mathbf{x}}$ to $\mathbf{x}$ |
| 6: Output: $\lambda_{\text{int}}$   |  |
- 

The application of `dtw_interp` is demonstrated on the illustrative example of Table 4.1 in Section 4.2.1. The interpolated quantities for this example are shown in Fig. 4.6.

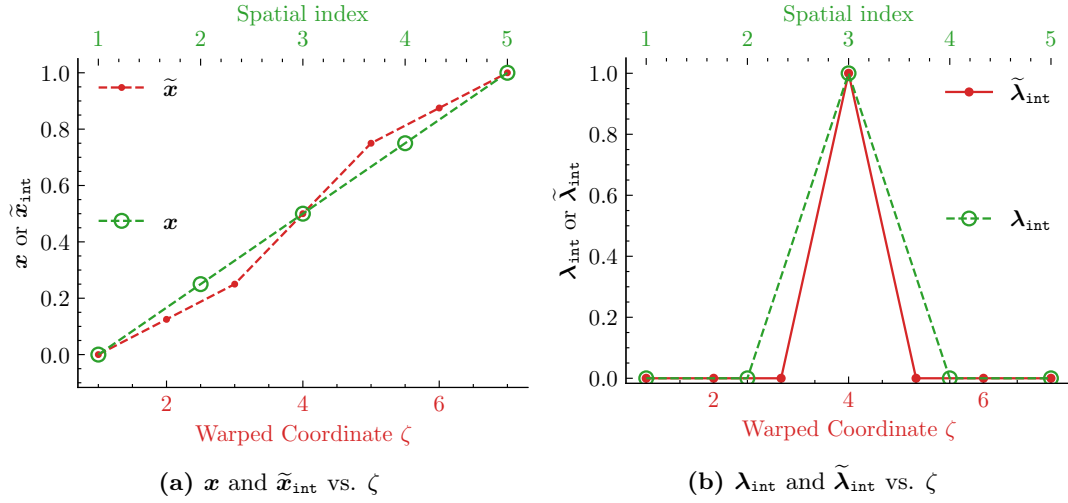
**Remark 4.3 : Interpolation using the alignments  $I_k$** 

In Alg. 4.1, the computation of  $\lambda_{\text{int}}$  from  $\tilde{\lambda}_{\text{int}}$  can be implemented in an alternate way. The interpolation of the mapping (4.3c) can be computed by interpolating between  $I_1$  and  $I_2$  to generate a set of interpolated indices. The steps would read as:

$$I_{\text{int}} = \lfloor (\alpha I_1 + (1 - \alpha) I_2) + \frac{1}{2} \rfloor$$

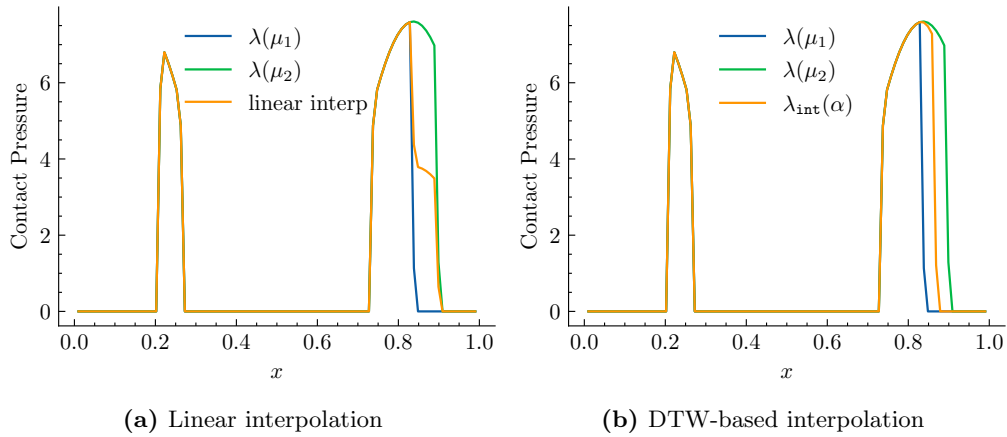
$$\lambda_{\text{int}}[I_{\text{int}}] \leftarrow \tilde{\lambda}_{\text{int}} \quad \triangleright \text{same as applying inverse mapping } \varphi_{\text{int}}^{-1}$$

involving a round off operation to the nearest integer, expressed here using the greatest integer function  $\lfloor \cdot \rfloor$  and the offset of  $\frac{1}{2}$ . This is equivalent to using the inverse mapping  $\varphi^{-1}$  in (4.4). However, this approach is not followed in Alg. 4.1, as the round off operation is sensitive to truncation at machine precision level, especially for  $\alpha = 0.5$ , leading to improper interpolation. Instead,  $\tilde{\mathbf{x}}_{\text{int}}$  that corresponds to the warped coordinate  $\zeta$  is used for interpolating back to  $\mathbf{x}$ . Also, the resulting  $I_{\text{int}}$  contains non-unique entries, leading to non-injective inverse map  $\varphi_{\text{int}}^{-1}$ .

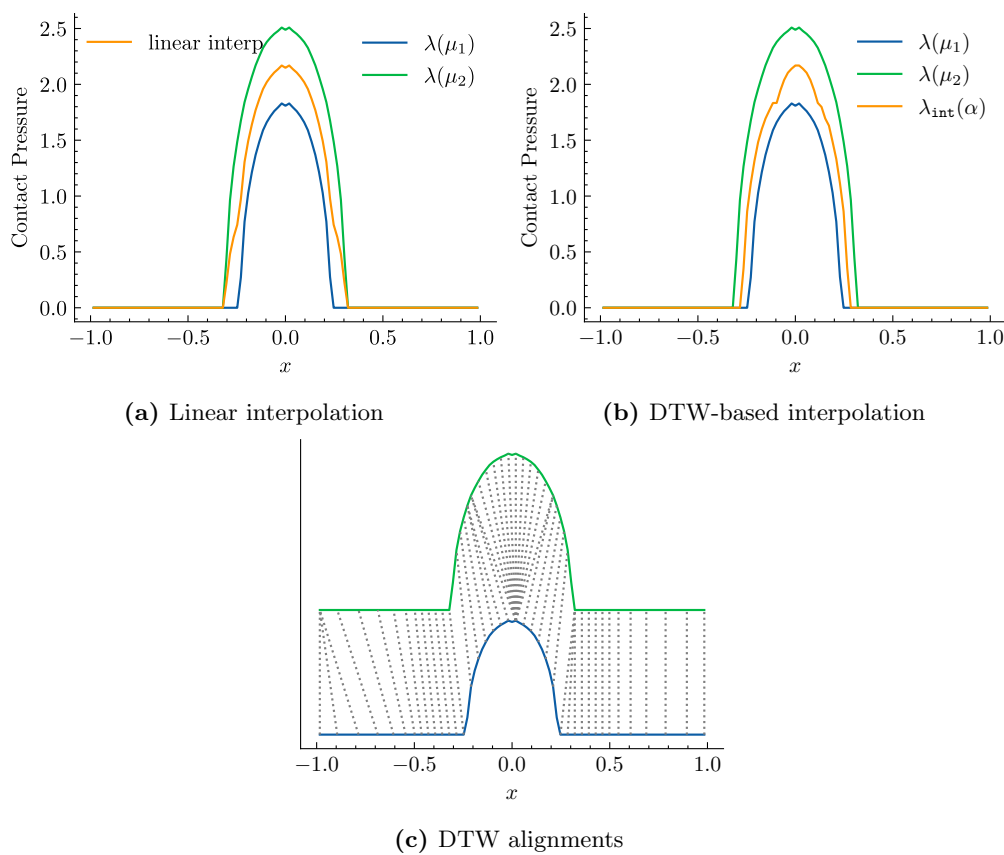


**Figure 4.6:** Interpolations of the illustrative example for  $\alpha = 0.5$  in the spatial indices (corresponding to  $x$ ) and the warped space  $\zeta$

The quality of interpolation using DTW on the ironing problem is quite good as the snapshots have a similar contact pressure profile, with only the contact zone being translated in the  $x$ -direction. DTW-based interpolations of the elastic rope-obstacle problem and the Hertz problem are also shown in Figs. 4.7 and 4.8. In these cases contact zone does not translate but magnifies. DTW is able to interpolate the contact zone correctly, but the interpolated pressure profile has some aberrations near the peak pressure of the Hertz case. This is explained by the concentration of the DTW alignments near the peak i.e., many points from  $\lambda(\mu_2)$  are mapped to the peak of  $\lambda(\mu_1)$ .



**Figure 4.7:** Linear and DTW-based interpolations of contact pressure for the elastic rope-obstacle problem.



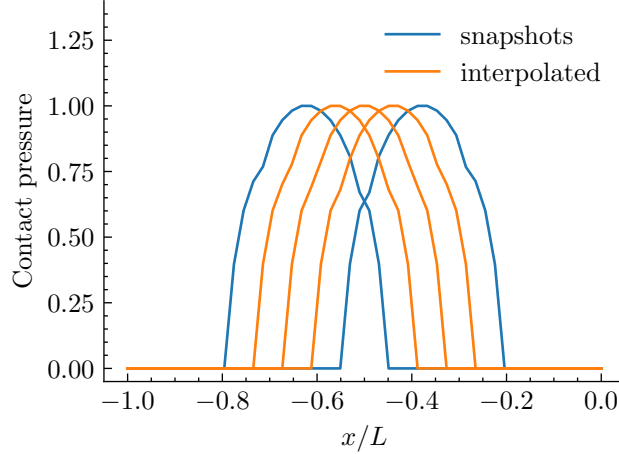
**Figure 4.8:** (a) Linear and (b) DTW-based interpolations of contact pressure for the Hertz problem. To explain distorted interpolation, DTW alignments is also shown in (c).

### 4.3 DTW-based enrichments for contact mechanics

As mentioned at the beginning of the chapter, the idea is to use manifold exploration to reduce the offline cost. Since the main cause of linear inseparability is the geometrically local nature of contact pressure field, the ability to scale and translate local effects across the domain of interest seems promising. Also, the displacement field does not have linear separability issues and a robust linear subspace seems to be readily available (Fig. 2.14) even if parametric space is relatively unexplored in offline stage. In an optimistic scenario, if the warped interpolations can reduce the size of training set to the same level necessary for capturing the displacement field, the bottleneck posed by linear inseparability will be effectively removed.

We propose the use of DTW as a warped interpolation tool that can generate intermediate contact pressures that haven't been explored in the offline stage. This is an inexpensive way of enriching the contact pressure dictionary; since the computation cost of warped interpolations is negligible compared to generating proper snapshots. Fig. 4.9 shows a demonstration of such an interpolation for three different values of  $\alpha$ . It is no-

ticeable that the peak position of contact pressure of the interpolated curves relative to the peaks of the original snapshots seem to have a nearly linear relation with  $\alpha$ , though this is a particular feature of the ironing problem and may not apply in general.



**Figure 4.9:** Interpolation of contact pressure snapshots using DTW for the ironing problem. Interpolated curves for  $\alpha \in \{0.25, 0.5, 0.75\}$  are shown

For an effective enrichment, the interpolation should not be done on every pair of contact pressure curves. Rather, each pair of curves should be reasonably similar in some sense. Similarity can be computed using a measure of distance which can be either the  $\mathcal{L}^2$  norm of the difference or the distance measure defined by the DTW algorithm. Snapshots that can be considered similar according to the distance measure will be referred to as *neighbours* here onwards. For each pair of neighbours, the DTW alignments are computed and interpolations are performed with a set of  $M$  parameter values  $\{\alpha_m\}_{m=1}^M$ . This process is shown in Alg. 4.2, where neighbours of currently active vectors are computed using the distances between vectors of the dual dictionary in a matrix  $\text{Dist}$ . The entry  $\text{Dist}[i, j]$  has the distance between  $\mathbf{D}_d[:, i]$  and  $\mathbf{D}_d[:, j]$ , and hence,  $\text{Dist}$  is a symmetric matrix with main diagonal containing only zeros.

The effectiveness of DTW-based enrichments can be shown using the compactness metric. Recall that the “dual cone” compactness metric defined in Section 2.3 can be used for as a measure of linear separability. The comparison of this metric for the original and enriched CPG dual basis is shown in Fig. 4.10. The compactness of the enriched basis was computed by performing 19 equispaced DTW interpolations (i.e. with  $\alpha_m = 0.05m$  and  $M = 19$  as input for Alg. 4.2) between each pair of neighbours in the original dual basis. The compactness of the enriched basis is considerably lower than the original, indicating that DTW enrichment has introduced the information that was missing in the cone of the original basis.

**Algorithm 4.2** `adapt_dictionary`

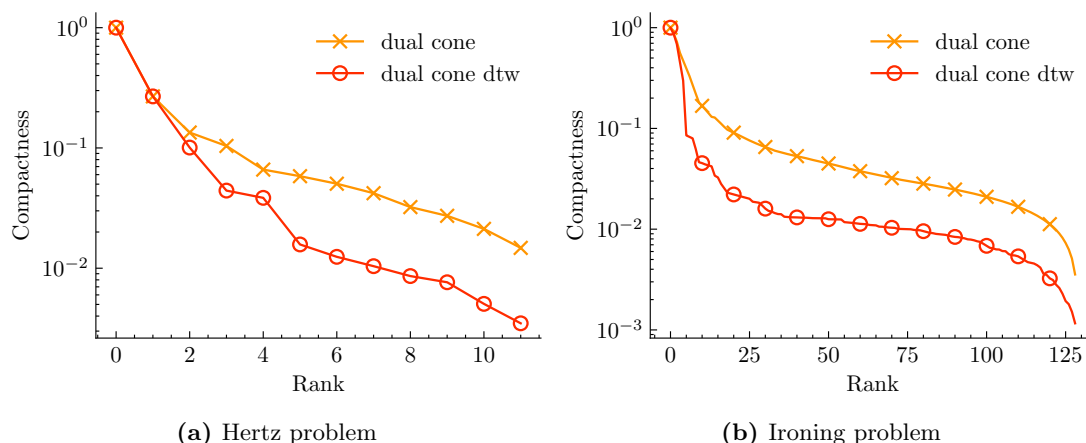

---

```

1: Input:  $\mathbf{D}_d, \mathcal{I}, \{\alpha_m\}_{m=1}^M$   $\triangleright \alpha_m \in ]0, 1[$ 
2: Parameters:  $\text{Dist}, N$ 
3: Initialize:  $\mathbf{D}_d^{\text{adapt}} = \mathbf{D}_d[:, \mathcal{I}], i = 0, j = 0$   $\triangleright \mathcal{I}$  is the current active set
4: for  $i \in \mathcal{I}$  do
5:    $\text{nbrs} \leftarrow$  indices of the least  $N$  positive entries of  $\text{Dist}[i, :]$ 
6:   Initialize  $\mathbf{D}_d^{\text{int}} = \{.\}$ 
7:   for  $j \in \text{nbrs}$  do  $\triangleright \text{nbrs}$  contains neighbouring indices
8:     if pair  $(i, j)$  is not repeated then  $\triangleright (i, j)$  is an unordered pair
9:        $\mathbf{D}_d^{\text{int}} \leftarrow \mathbf{D}_d^{\text{int}} \cup \left[ \bigcup_{m=1}^M \text{dtw\_interp}(\mathbf{D}_d[:, j], \mathbf{D}_d[:, i], \alpha_m) \right]$   $\triangleright$  Alg. 4.1
10:    end if
11:  end for
12:   $\mathbf{D}_d^{\text{adapt}} \leftarrow \mathbf{D}_d^{\text{adapt}} \cup \mathbf{D}_d[:, \text{nbrs}] \cup \mathbf{D}_d^{\text{int}}$ 
13: end for
14: Output:  $\mathbf{D}_d^{\text{adapt}}$ 

```

---

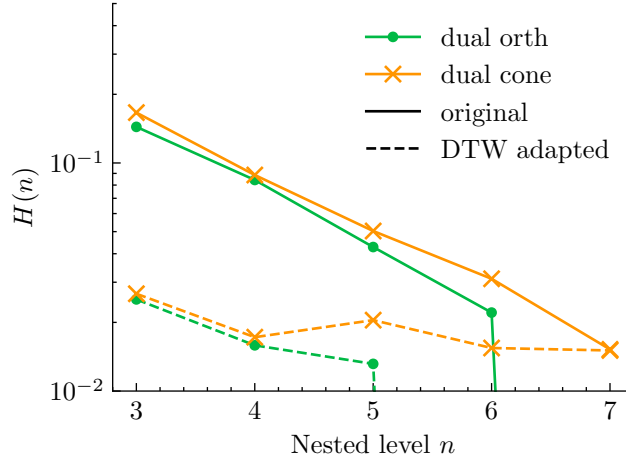


**Figure 4.10:** Comparison of the compactness for original and warped snapshots. Note that the curve labelled “dual cone” are the same curves from Figs. 2.8 and 2.13b, respectively. Curves labelled “dual cone dtw” show the compactness of the enriched basis.

#### Remark 4.4 : Indefiniteness of the DTW kernel

A more appropriate way of testing the separability of given data in warped space is by observing the decay of singular values (equivalent to the dual orth compactness in Chapter 2). However, computation of singular values requires that the DTW mapping generates a symmetric positive definite (SPD) Gramian matrix  $G_{ij} = \langle \tilde{\lambda}_i, \tilde{\lambda}_j \rangle$ , where each entry of  $G$  contains an inner product of a pair of warped snapshots. Such a comparison would have been possible if SPD kernels are used for computation of the Gramian matrix, as done in [86] for an advection-diffusion problem. However, instances of Gramian matrix computed using DTW kernel being indefinite are known [96]. The Gramian matrix computed for this case also resulted in negative eigenvalues, hence the dual orth compactness cannot be compared.

The richness of  $\mathbf{D}_d^{\text{adapt}}$  w.r.t  $\mathbf{D}_d$  can also be compared using the nested errors defined for the ironing problem (see Section 2.4.2, Fig. 2.14). Let us recall that the nested error (2.7) was defined as the projection error of  $(n + 1)$ -th level basis on  $n$ -th level basis. The comparison of nested errors with and without DTW interpolation is shown in Fig. 4.11. For each level  $n$ , the dictionary neighbours are interpolated using Alg. 4.2 with  $\mathcal{I} = \{1, 2, 3, \dots, \#\mathcal{P}_{\text{tr}}\}$ ,  $N = 2$ ,  $M = 1$  and  $\alpha_1 = 0.5$ . The ironing problem has a peculiar property where the position of contact pressure peak is more or less same as the parameter  $d$ . Therefore, the chosen  $N, M$  and  $\alpha_1$  form a near best case scenario of the approximation because it suits the distribution of nested training set points i.e.  $(n + 1)$ -th level contain mid-points of the points in  $n$ -th level. Thus, a considerable decrease in the nested errors is seen for smaller training sets. Naturally, the gains in dual orth error due to DTW decreases as the training set becomes larger, as the snapshots in large training sets are closer to the target solutions.



**Figure 4.11:** Projection error of nested level  $n + 1$  snapshots on dictionary of level  $n$ . Note that solid dual orth and dual cone lines are similar to the ones in Fig. 2.14. The dual orth points not seen in the figure drop to the order of  $10^{-14}$  (nearly machine precision)

However, in a general scenario, the value of an optimum  $\alpha$  may not be trivial. Therefore, to arrive at an appropriate value of  $\alpha$ , it is necessary to explore multiple DTW-based interpolations corresponding to different  $\alpha$ 's. This is the reason a set of  $M$  values  $\{\alpha_m\}$  is supplied to Alg. 4.2. The enrichment process can be strategised in two different ways:

- **Number of interpolation points ( $M$ ):** For a given set of snapshots generated in offline phase, warped interpolation is performed between every pair of neighbouring snapshots with a certain resolution of in  $]0, 1[$ . For a uniformly distributed points,  $\alpha_m = m/(M + 1)$ . The interpolations in Fig. 4.9 is such a case with  $M = 3$ .
- **Number of enrichment levels ( $L$ ):** It is possible to continue the enrichment process further by creating multiple levels of interpolated vectors. Suppose, the original dictionary contains two vectors  $\lambda_1$  and  $\lambda_2$  and interpolation is performed

with  $M = 1$  and  $\alpha_1 = 0.5$ . Thus, an interpolated vector, let us call it  $\boldsymbol{\lambda}_{1.5}$ , can be computed using DTW interpolation in the first level. In a second level, further interpolation can be performed on the pairs  $\{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_{1.5}\}$  and  $\{\boldsymbol{\lambda}_{1.5}, \boldsymbol{\lambda}_2\}$  to generate  $\boldsymbol{\lambda}_{1.25}$  and  $\boldsymbol{\lambda}_{1.75}$ .

Further, these enrichments can be done during either the online or offline stage, depending on the desired distribution of the computational cost. If done in the offline stage, warped interpolations between every possible pair of neighbours need to be computed. On the other hand, in the online stage, interpolations can be computed on a smaller set of vectors, i.e. the active dictionary vectors and their neighbours. In this thesis, the enrichment in the online stage is studied, as the focus is on the enhancement of the reconstruction accuracy using DTW and the study of computational cost is reserved for future study. However, a preliminary comment can be made in this aspect, that the DTW-based enrichment is not a computationally intensive task. As the Alg. 3.2 generates a small active set, the number of interpolations performed is not large. The resulting algorithm is given in Alg. 4.3, with a parameter  $L$  denoting the number of enrichment levels. In the numerical results that follow, the parameters used are  $L = 5$  and  $M = 4$  with the interpolation points  $\{0.2, 0.4, 0.6, 0.8\}_{m=1}^M$ .

---

**Algorithm 4.3** DTW powered greedy active-set method

---

```

1: Input : Queried value of parameter  $\mu$ 
2: Parameters :  $\{\alpha_m\}_{m=1}^M, L$ 
3: Given: Primal basis  $\Phi$  and dual Dictionary  $\mathbf{D}_d$ 
           Reduced operators  $\Phi^T \mathbf{K} \Phi$  and  $\Phi^T \mathbf{f}$ 
4: Initialize:  $l = 0, \mathcal{I} = \{\cdot\}, \boldsymbol{\lambda} = \mathbf{0}$ 
5: while  $\boldsymbol{\lambda}$  not converge and  $l \leq L$  do
6:   Compute  $\hat{\mathbf{u}}, \hat{\boldsymbol{\lambda}}, \mathcal{I}$  using current dictionary  $\mathbf{D}_d$            ▷ Steps 4 to 12 of Alg. 3.2
7:   Reconstruct  $\boldsymbol{\lambda} = \mathbf{D}_d \hat{\boldsymbol{\lambda}}$ 
8:    $\mathbf{D}_d \leftarrow \text{adapt\_dictionary}(\mathbf{D}_d, \mathcal{I}, \{\alpha_m\})$            ▷ Alg. 4.2
9:    $\mathcal{I} = \{\cdot\}$            ▷ Reset active set after adapting  $\mathbf{D}_d$ 
10:   $l \leftarrow l + 1$ 
11: end while
12: Reconstruct  $\mathbf{u} = \Phi \hat{\mathbf{u}}$ 
13: Output:  $\mathbf{u}, \boldsymbol{\lambda}$ 

```

---

### 4.3.1 Numerical results

For the Hertz problem and the ironing problem training sets defined in Chapter 3, reconstruction errors with enriched dual dictionaries using Alg. 4.3 are shown in Fig. 4.12. For the ironing problem (Fig. 4.12b), the DTW-based interpolation improves the quality of reconstruction for smaller training sets. The improvement asymptotically dampens out as the training set size increases. This is natural as snapshots come quite close in large training sets and DTW-based interpolations do not generate any useful information.

Moreover, the improvement in reconstruction is seen more in the dual field rather than the primal. This is an expected observation as any improvement in primal reconstruction due to dual dictionary enrichments is caused indirectly through the mixed system (3.5b).

On the other hand, there is no improvement in the reconstruction of Hertz problem (Fig. 4.12a). One likely reason could have been that the contact pressure snapshots are relatively close, unlike the ironing problem, and therefore the DTW-based interpolations are unable to offer any significant new information. However, this contradicts the observations in Fig. 4.10a where improvements in dual cone compactness is observed. A more likely reason could be, that to improve in reconstruction of Hertz problem, the primal dictionary also needs enrichment rather than just the dual dictionary. However, artificial enrichment could not be applied to the primal dictionary (as it needs a multidimensional DTW implementation), and the dual dictionary enrichments alone are ineffective at improving the errors. This reasoning is more consistent with observation that reconstruction error improves only with increase in training set size and not with DTW-based enrichments. Another possible source of error is the distorted DTW-based interpolation of the contact pressure snapshots, as shown in Fig. 4.8. Though, in reality, the distortion is not as pronounced as shown in the figure, since the interpolation done in Alg. 4.3 is carried out between snapshots that are closer.

## 4.4 Conclusions

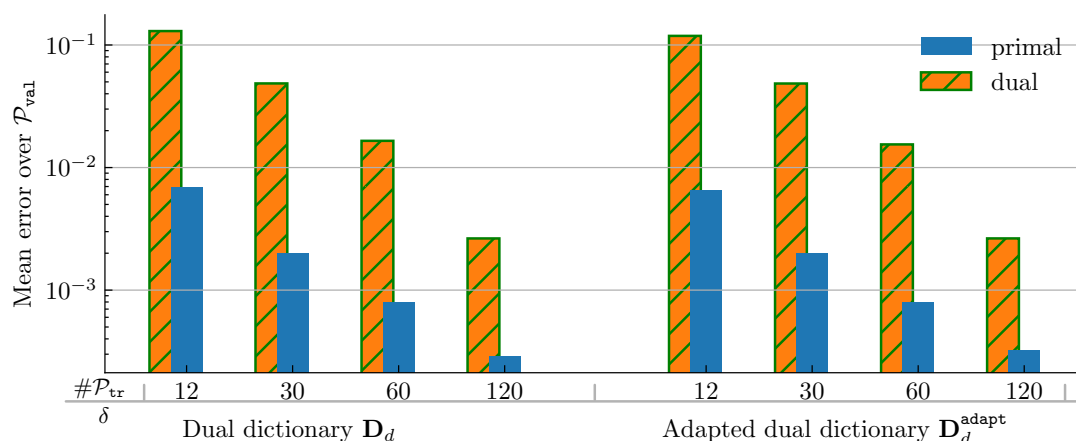
The DTW enrichments proved to be useful in improving contact pressure reconstructions for a problem with higher variations in contact area (ironing problem), especially when the dictionary is not rich enough to resolve the local effects. The improvements obtained using DTW diminish with increasing training set size, as the local information related to contact pressure start appearing in the dictionary directly. For problems with less variations in contact pressure (Hertz problem), DTW does not provide any improvements in reconstruction errors.

Overall, it is fair to say that non-linear interpolations can prove useful to train models for problems with large variability in contact positions, especially if the complexity of high-fidelity problem makes it difficult to compute a large training set. This is a strong motivator to test other non-linear reduction techniques such as kPOD on such problems. However, whether it is trivial to enforce inequality constraints in non-linear frameworks or not is yet to be explored.

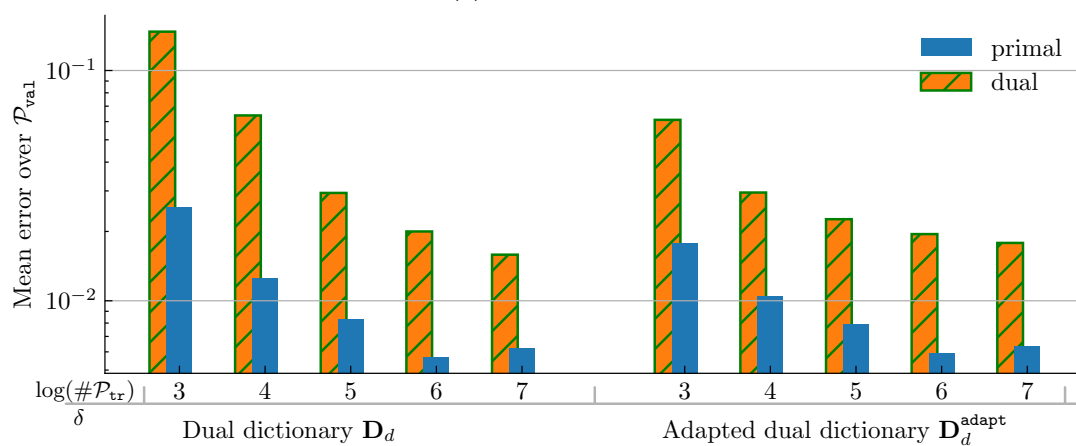
## 4.5 Perspectives

### 4.5.1 The 3D case

Extension of DTW based interpolation to 3D problems (with 2D contact surfaces) may not be straightforward. Though the DTW algorithm was originally built for a 1D signals, but a 2D version has been developed [93]. However, the 2D DTW is limited to structured meshes, thereby limiting its application. Even if the geometry of the contact



(a) Hertz problem



(b) Ironing problem

**Figure 4.12:** Comparison of reconstruction error for the Hertz and ironing problem using the original dual dictionary  $\mathbf{D}_d$  and DTW-adapted dual dictionary  $\mathbf{D}_d^{\text{adapt}}$ . Reconstructions performed using Alg. 3.2 and Alg. 4.3 for respective dictionaries. DTW enrichments seem to be more impactful in the ironing problem than Hertz problem. Primal and dual errors are computed using  $\mathcal{H}^1$  and  $\mathcal{L}^2$  norm, respectively.

surfaces allow discretization using a structured mesh, it may not be the optimal mesh for generating the snapshots using finite elements. It may however, be possible to compute the FE snapshots using a generic mesh and then interpolating the solution on a structured mesh during a post-processing, thereby making it suitable for DTW-based interpolations.

#### 4.5.2 Efficient evaluation of non-linear terms

To solve the contact problem using the dictionary method, it is necessary to compute the operators  $\widehat{\mathbf{C}}$  and  $\widehat{\mathbf{g}}$ , which are dependent on the dual dictionary  $\mathbf{D}_d$ . However, if

DTW based interpolation is used, the dictionary  $\mathbf{D}_d$  is dynamic. This can create additional difficulties in devising affine decompositions of these operators. If EIM approach from [13] is used, computation of the independent terms  $\widehat{\mathbf{C}}^p$  and  $\widehat{\mathbf{g}}^q$  in (1.17) becomes dependent on the dual basis/dictionary. An efficient way of handling such situations is necessary. A straightforward but less efficient way might be to build an affine decomposition of high-fidelity operators  $\mathbf{C}$  and  $\mathbf{g}$ , and update the reduced operators whenever the operators or the dictionary is updated.

### 4.5.3 Non-linear interpolations of primal dictionary

As seen in the case of Hertz problem, the dual dictionary enrichment is not effective at improving reconstruction errors. Also, in case of the ironing problem, mainly the dual reconstruction errors show improvement than the primal one. This observation opens up the possibility of enriching the primal dictionary using non-linear methods. This will also be useful to generate enriched monolithic dictionary for convex hull based approximation.

# Conclusions and Perspectives

---

## Conclusions

In Chapter 2, the robustness of low-rank hypothesis in the context of contact mechanics was analyzed. The premise of this analysis was based on the local nature of the contact pressure field which leads to an inherent lack of linear separability. As the underlying physics exhibited a high intrinsic dimensionality, reconstructions performed using ROMs had a moderate error, as expected. The robustness of contact pressure reduced bases computed using SVD and CPG methods were shown to have a slower decay of error (i.e. higher values of compactness) than the displacement field. The lack of low-rank structure had also negatively influenced the generalization ability of the reduced basis.

The linear inseparability of the contact pressure makes it harder to find a robust low-dimensional structure in the framework of linear dimensionality reduction. In Chapter 3, it is shown that a large number of snapshots can provide some improvements to the approximate solution subspace that is exploited by reduced models. This is achieved by using large dictionaries of snapshots, and the reconstructed solution is restricted to a sparse linear combination of these snapshots. It is seen that dictionary-based approximations are quite useful for contact problems with large changes in contact position like the ironing problem, as local effects are quite significant in these problems.

Two dictionary-based approximation strategies were devised. The first is based on the active-set approach implemented greedily, and the second is based on exploring the convex-hull of the snapshots to reconstruct the solution under the condition that the feasible (non-penetrative) region of the displacement space is convex. The computational time for these methods seems to be primarily dependent on the number of iterations for convergence, which may or may not change significantly with the size of the dictionary depending on the contact problem. In terms of reconstruction accuracy, the greedy active-set approach with separate dictionaries for displacement and contact pressure outperforms the convex-hull approximations with monolithic dictionaries. Moreover, the convex-hull approximations could not be applied to non-convex problems such as the ironing problem.

Non-linear dimensionality reduction aspects are particularly interesting for a problem lacking low-rank structure and are explored in Chapter 4. As the computational costs of

generating a large dictionary of snapshots can be very expensive, non-linear methods are used to reduce the need for a large number of contact pressure snapshots. Dynamic Time Warping (DTW) based non-linear interpolations were used to synthesize snapshots that were not present in the training set, thereby enriching the dictionary without paying the large cost of full simulations. This is found to be beneficial in cases of large variations in contact position.

## Perspectives

1. The problems arising from linear inseparability of contact pressure have been addressed using over-complete dictionaries and DTW-based enrichments in this thesis. However, distance functions that are used to compute the non-penetration conditions (i.e. the inequality constraints) are likely to have similar local effects and consequent linear inseparability issues. Also, the current methods of computing affine decompositions of distance functions (like EIM) are based on low-rank approximations requiring linear separability. Therefore, similar strategies might be required for the efficient construction of the reduced inequality constraint operators.
2. Random sketching methods were applied only to the convex hull approach as the method was free of inequality constraints. It would be interesting if the random sketching could be applied to other dictionary-based strategies as well, including the greedy active-set method.
3. The dictionary method based on the convex hull approach is currently built for monolithic dictionaries, which forces the reduced coefficients to be the same for both primal and dual variables. An approach with two uncoupled dictionaries must be explored for potential improvements in reconstruction quality. Also, an efficient sampling strategy for the offline phase is necessary, where snapshots placed at corners and edges of the presumed low-dimensional convex subset are sampled with priority.
4. As DTW-based non-linear interpolation was found useful in particular cases with large variations of contact positions, application of other non-linear methods like kPCA/kPOD should be explored. These methods might be more relevant for three-dimensional contact problems with two-dimensional contact surfaces, as multidimensional DTW is limited to structured meshes.
5. Currently, non-linear interpolation is only applied to contact pressure. It would be interesting to see if non-linear interpolation of displacement snapshots provides any advantages. DTW-based enrichment of displacement snapshots, along with the contact pressure, might be beneficial for creating monolithic dictionaries which need the same number of primal and dual snapshots.

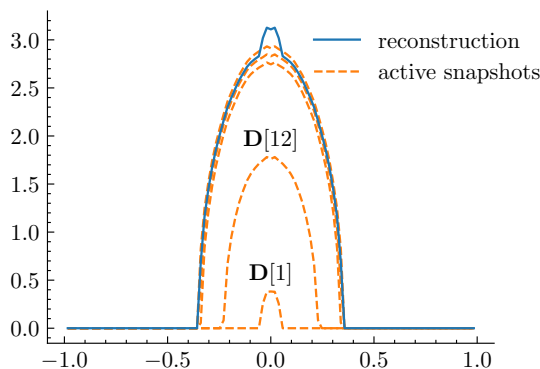
6. When the dual dictionary is updated using DTW-based enrichments, a fresh computation of the reduced inequality constraint operators is required. This will not be tenable if the DTW-based enrichments are to be deployed in a real-life application of the contact ROM. A strategy for efficient computation of these operators for an updated dictionary is necessary.



# A | Spurious penetrations

---

In the greedy active-set method, the snapshots of the dual dictionary are selected based on the violations of the non-penetration constraints. The dictionary vector with the highest correlation with the current state of penetration is chosen. However, instead of zero penetration, a small value of penetration  $\tau$  is allowed (see (3.5c) and (3.6)). This is done to avoid spurious selection of dictionary elements, which happens due to a truncated primal basis. Such a reconstruction example of Hertz problem for parameter value  $d = 0.25$  is shown in Fig. A.1, where a spurious peak in the contact pressure is evident. In this case, a dual dictionary of size 30 and primal basis truncated at  $10^{-6}$  are used. The algorithm selected the appropriate snapshots from the dual dictionary, but also selected some unnecessary snapshots: the first ( $\mathbf{D}_d[1]$ ) and twelfth ( $\mathbf{D}_d[12]$ ) dictionary vectors. Importantly, the  $\mathbf{D}_d[1]$  vector contributes to large error in reconstruction as evident in the figure. The false selection of this snapshot is linked to the truncation of primal basis, and therefore, can be avoided by setting the value of  $\tau$  same as  $\delta$ .



**Figure A.1:** Reconstruction instance of Hertz problem with  $d = 0.25$  using  $\#\mathcal{P}_{\text{tr}} = 30$  and primal basis truncated at  $\delta = 10^{-6}$ , where few dictionary snapshots were spuriously selected by the greedy active-set algorithm if non-penetration condition is applied “hardly”. The falsely selected snapshots are indicated by their index in the dual dictionary. The snapshot which contributes to most of the error is  $\mathbf{D}_d[1]$ , corresponds to loading parameter  $d = 0.01$



# Bibliography

---

- [1] N. Lauzeral, D. Borzacchiello, M. Kugler, D. George, Y. Rémond, A. Hostettler, and F. Chinesta. A model order reduction approach to create patient-specific mechanical models of human liver in computational medicine applications. *Computer Methods and Programs in Biomedicine*, 170:95–106, 2019. ISSN 0169-2607. doi:[10.1016/j.cmpb.2019.01.003](https://doi.org/10.1016/j.cmpb.2019.01.003).
- [2] J. Marconi, P. Tiso, and F. Braghin. A nonlinear reduced order model with parametrized shape defects. *Computer Methods in Applied Mechanics and Engineering*, 360:112785, 2020. ISSN 0045-7825. doi:[10.1016/j.cma.2019.112785](https://doi.org/10.1016/j.cma.2019.112785).
- [3] J. M. Navarro-Jiménez, J. V. Aguado, G. Bazin, V. Albero, and D. Borzacchiello. Reconstruction of 3D surfaces from incomplete digitisations using statistical shape models for manufacturing processes. *Journal of Intelligent Manufacturing*, mar 2022. ISSN 0956-5515. doi:[10.1007/s10845-022-01918-z](https://doi.org/10.1007/s10845-022-01918-z).
- [4] H. Hertz. Ueber die Berührung fester elastischer Körper. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1882(92):156–171, jan 1882. ISSN 0075-4102. doi:[10.1515/crll.1882.92.156](https://doi.org/10.1515/crll.1882.92.156).
- [5] C. Coulomb. Essai sur une application des regles de maximis et minimis a quelques problemes de statique relatifs a l’architecture. pages 343–382, 1773.
- [6] A. Signorini. Questioni di elasticità non linearizzata e semilinearizzata. *Rend. Mat. Appl., V. Ser.*, 18:95–139, 1959. ISSN 0034-4427.
- [7] K. L. Johnson. *Contact Mechanics*. Cambridge University Press, may 1985. ISBN 9780521255769. doi:[10.1017/CBO9781139171731](https://doi.org/10.1017/CBO9781139171731).
- [8] K. L. Johnson. *Introduction to Contact Mechanics: a summary of the principal formulae*, pages 589–603. Springer Netherlands, Dordrecht, 1992. ISBN 978-94-011-2811-7. doi:[10.1007/978-94-011-2811-7.30](https://doi.org/10.1007/978-94-011-2811-7.30).
- [9] I. G. Goryacheva. *Contact Mechanics in Tribology*. Springer, 1988. ISBN 978-90-481-5102-8. doi:[10.1007/978-94-015-9048-8](https://doi.org/10.1007/978-94-015-9048-8).

- [10] P. Wriggers. *Computational Contact Mechanics*. Springer Berlin Heidelberg, 2006. ISBN 978-3-540-32608-3. doi:[10.1007/978-3-540-32609-0](https://doi.org/10.1007/978-3-540-32609-0).
- [11] K. A. Fischer and P. Wriggers. Mortar based frictional contact formulation for higher order interpolations using the moving friction cone. *Computer Methods in Applied Mechanics and Engineering*, 195(37-40):5020–5036, 2006. doi:[10.1016/j.cma.2005.09.025](https://doi.org/10.1016/j.cma.2005.09.025).
- [12] V. A. Yastrebov. *Numerical Methods in Contact Mechanics*. John Wiley & Sons, Inc., Hoboken, NJ USA, feb 2013. ISBN 9781118647974. doi:[10.1002/9781118647974](https://doi.org/10.1002/9781118647974).
- [13] A. Benaceur, A. Ern, and V. Ehrlacher. A reduced basis method for parametrized variational inequalities applied to contact mechanics. *Int. J. Numer. Methods Eng.*, 121(6):1170–1197, mar 2020. doi:[10.1002/nme.6261](https://doi.org/10.1002/nme.6261).
- [14] M. R. Hestenes. Multiplier and gradient methods. *Journal of Optimization Theory and Applications*, 4(5):303–320, nov 1969. doi:[10.1007/BF00927673](https://doi.org/10.1007/BF00927673).
- [15] J. Hallquist, G. Goudreau, and D. Benson. Sliding interfaces with contact-impact in large-scale lagrangian computations. *Computer Methods in Applied Mechanics and Engineering*, 51(1):107–137, 1985. doi:[10.1016/0045-7825\(85\)90030-1](https://doi.org/10.1016/0045-7825(85)90030-1).
- [16] K. Poullos and Y. Renard. An unconstrained integral approximation of large sliding frictional contact between deformable solids. *Computers & Structures*, 153:75–90, 2015. doi:[10.1016/j.compstruc.2015.02.027](https://doi.org/10.1016/j.compstruc.2015.02.027).
- [17] N. El-Abbasi and K.-J. Bathe. Stability and patch test performance of contact discretizations and a new solution algorithm. *Computers & Structures*, 79(16):1473–1486, 2001. doi:[10.1016/S0045-7949\(01\)00048-7](https://doi.org/10.1016/S0045-7949(01)00048-7).
- [18] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer New York, 2006. ISBN 978-0-387-30303-1. doi:[10.1007/978-0-387-40065-5](https://doi.org/10.1007/978-0-387-40065-5).
- [19] D. G. Luenberger and Y. Ye. *Linear and Nonlinear Programming*, volume 228 of *International Series in Operations Research & Management Science*. Springer International Publishing, Cham, 2016. ISBN 978-3-319-18841-6. doi:[10.1007/978-3-319-18842-3](https://doi.org/10.1007/978-3-319-18842-3).
- [20] R. F. Guratzsch and S. Mahadevan. Structural health monitoring sensor placement optimization under uncertainty. *AIAA Journal*, 48(7):1281–1289, 2010. doi:[10.2514/1.28435](https://doi.org/10.2514/1.28435).
- [21] D. Modesto, S. Zlotnik, and A. Huerta. Proper generalized decomposition for parameterized helmholtz problems in heterogeneous and unbounded domains: Application to harbor agitation. *Computer Methods in Applied Mechanics and Engineering*, 295:127 – 149, 2015. doi:[10.1016/j.cma.2015.03.026](https://doi.org/10.1016/j.cma.2015.03.026).

- [22] A. Mosavi. *Engineering Design and Decision Making Models*. PhD thesis, 07 2015.
- [23] C. Bigoni and J. S. Hesthaven. Simulation-based anomaly detection and damage localization: An application to structural health monitoring. *Computer Methods in Applied Mechanics and Engineering*, 363:112896, 2020. ISSN 0045-7825. doi:[10.1016/j.cma.2020.112896](https://doi.org/10.1016/j.cma.2020.112896).
- [24] F. Chinesta, A. Leygue, F. Bordeu, J. V. Aguado, E. Cueto, D. Gonzalez, I. Alfaro, A. Ammar, and A. Huerta. PGD-Based Computational Vademecum for Efficient Design, Optimization and Control. *Archives of Computational Methods in Engineering*, 20(1):31–59, mar 2013. ISSN 1134-3060. doi:[10.1007/s11831-013-9080-x](https://doi.org/10.1007/s11831-013-9080-x).
- [25] A.-T. Nguyen, S. Reiter, and P. Rigo. A review on simulation-based optimization methods applied to building performance analysis. *Applied Energy*, 113:1043–1058, 2014. ISSN 0306-2619. doi:[10.1016/j.apenergy.2013.08.061](https://doi.org/10.1016/j.apenergy.2013.08.061).
- [26] E. T. Woldemariam and H. G. Lemu. Numerical Simulation-Based Effect Characterization and Design Optimization of a Micro Cross-Flow Turbine. *Strojniški vestnik - Journal of Mechanical Engineering*, 65(6):386–397, jun 2019. ISSN 00392480. doi:[10.5545/sv-jme.2018.5901](https://doi.org/10.5545/sv-jme.2018.5901).
- [27] D. Hartmann, M. Herz, and U. Wever. *Model Order Reduction a Key Technology for Digital Twins*. Springer International Publishing, Cham, 2018. doi:[10.1007/978-3-319-75319-5\\_8](https://doi.org/10.1007/978-3-319-75319-5_8).
- [28] S. McBane and Y. Choi. Component-wise reduced order model lattice-type structure design. *Computer Methods in Applied Mechanics and Engineering*, 381:113813, 2021. doi:[10.1016/j.cma.2021.113813](https://doi.org/10.1016/j.cma.2021.113813).
- [29] F. D. Freitas, R. Pulch, and J. Rommes. Fast and accurate model reduction for spectral methods in uncertainty quantification. *International Journal for Uncertainty Quantification*, 6(3):271–286, 2016.
- [30] D. González, F. Masson, F. Poulhaon, A. Leygue, E. Cueto, and F. Chinesta. Proper generalized decomposition based dynamic data driven inverse identification. *Mathematics and Computers in Simulation*, 82(9):1677–1695, 2012. doi:[10.1016/j.matcom.2012.04.001](https://doi.org/10.1016/j.matcom.2012.04.001).
- [31] M. Kärcher, Z. Tokoutsis, M. A. Grepl, and K. Veroy. Certified reduced basis methods for parametrized elliptic optimal control problems with distributed controls. *Journal of Scientific Computing*, 75(1):276–307, Aug 2017. doi:[10.1007/s10915-017-0539-z](https://doi.org/10.1007/s10915-017-0539-z).
- [32] D. Ryckelynck. A priori hyperreduction method: an adaptive approach. *Journal of Computational Physics*, 202(1):346–366, 2005. doi:[10.1016/j.jcp.2004.07.015](https://doi.org/10.1016/j.jcp.2004.07.015).

- [33] D. Ryckelynck. Hyper-reduction of mechanical models involving internal variables. *International Journal for Numerical Methods in Engineering*, 77(1):75–89, 2009. doi:[10.1002/nme.2406](https://doi.org/10.1002/nme.2406).
- [34] J. Hernández, M. Caicedo, and A. Ferrer. Dimensional hyper-reduction of nonlinear finite element models via empirical cubature. *Computer Methods in Applied Mechanics and Engineering*, 313:687–722, 2017. ISSN 0045-7825. doi:[10.1016/j.cma.2016.10.022](https://doi.org/10.1016/j.cma.2016.10.022).
- [35] C. Audouze, F. De Vuyst, and P. B. Nair. Reduced-order modeling of parameterized pdes using time–space-parameter principal component analysis. *International Journal for Numerical Methods in Engineering*, 80(8):1025–1057, 2009. doi:[10.1002/nme.2540](https://doi.org/10.1002/nme.2540).
- [36] C. Audouze, F. De Vuyst, and P. B. Nair. Nonintrusive reduced-order modeling of parametrized time-dependent partial differential equations. *Numerical Methods for Partial Differential Equations*, 29(5):1587–1628, 2013. doi:[10.1002/num.21768](https://doi.org/10.1002/num.21768).
- [37] D. Xiao, F. Fang, C. Pain, and G. Hu. Non-intrusive reduced-order modelling of the navier–stokes equations based on rbf interpolation. *International Journal for Numerical Methods in Fluids*, 79(11):580–595, 2015. doi:[10.1002/flid.4066](https://doi.org/10.1002/flid.4066).
- [38] D. Borzacchiello, J. V. Aguado, and F. Chinesta. Non-intrusive sparse subspace learning for parametrized problems. *Archives of Computational Methods in Engineering*, Jul 2017. doi:[10.1007/s11831-017-9241-4](https://doi.org/10.1007/s11831-017-9241-4).
- [39] Y. Le Guennec, J. Brunet, F. Daim, M. Chau, and Y. Tourbier. A parametric and non-intrusive reduced order model of car crash simulation. *Computer Methods in Applied Mechanics and Engineering*, 338:186 – 207, 2018. doi:[10.1016/j.cma.2018.03.005](https://doi.org/10.1016/j.cma.2018.03.005).
- [40] A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modelling of complex fluids. *Journal of Non-Newtonian Fluid Mechanics*, 139(3):153–176, 2006. doi:[10.1016/j.jnnfm.2006.07.007](https://doi.org/10.1016/j.jnnfm.2006.07.007).
- [41] J. V. Aguado, D. Borzacchiello, K. S. Kollépara, F. Chinesta, and A. Huerta. Tensor Representation of Non-linear Models Using Cross Approximations. *J. Sci. Comput.*, Feb 2019. doi:[10.1007/s10915-019-00917-2](https://doi.org/10.1007/s10915-019-00917-2).
- [42] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced Basis Methods for Partial Differential Equations*, volume 92 of *UNITEXT*. Springer International Publishing, Cham, 2016. ISBN 978-3-319-15430-5. doi:[10.1007/978-3-319-15431-2](https://doi.org/10.1007/978-3-319-15431-2).
- [43] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*. SpringerBriefs in Mathematics. Springer International Publishing, Cham, 2016. ISBN 978-3-319-22469-5. doi:[10.1007/978-3-319-22470-1](https://doi.org/10.1007/978-3-319-22470-1).

- [44] C. Prud'homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, and G. Turinici. Reliable Real-Time Solution of Parametrized Partial Differential Equations: Reduced-Basis Output Bound Methods. *Journal of Fluids Engineering*, 124(1):70–80, 11 2001. doi:[10.1115/1.1448332](https://doi.org/10.1115/1.1448332).
- [45] C. Prud'homme, D. V. Rovas, K. Veroy, and A. T. Patera. A Mathematical and Computational Framework for Reliable Real-Time Solution of Parametrized Partial Differential Equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 36(5):747–771, sep 2002. doi:[10.1051/m2an:2002035](https://doi.org/10.1051/m2an:2002035).
- [46] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. An ‘empirical interpolation’ method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathématique*, 339(9):667–672, nov 2004. ISSN 1631073X. doi:[10.1016/j.crma.2004.08.006](https://doi.org/10.1016/j.crma.2004.08.006).
- [47] N. Lauzeral. *Reduced order and sparse representations for patient-specific modeling in computational surgery*. PhD thesis, École centrale de Nantes, dec 2019. URL <http://www.theses.fr/2019ECDN0062>.
- [48] J. Fauque, I. Ramière, and D. Ryckelynck. Hybrid hyper-reduced modeling for contact mechanics problems. *Int. J. Numer. Methods Eng.*, 115(1):117–139, Jul 2018. doi:[10.1002/nme.5798](https://doi.org/10.1002/nme.5798).
- [49] A. Quarteroni and G. Rozza. Numerical solution of parametrized navier–stokes equations by reduced basis methods. *Numerical Methods for Partial Differential Equations*, 23(4):923–948, 2007. doi:[10.1002/num.20249](https://doi.org/10.1002/num.20249).
- [50] B. Haasdonk, J. Salomon, and B. Wohlmuth. A Reduced Basis Method for Parametrized Variational Inequalities. *SIAM J. Numer. Anal.*, 50(5):2656–2676, jan 2012. doi:[10.1137/110835372](https://doi.org/10.1137/110835372).
- [51] M. Balajewicz, D. Amsallem, and C. Farhat. Projection-based model reduction for contact problems. *Int. J. Numer. Methods Eng.*, 106(8):644–663, may 2016. doi:[10.1002/nme.5135](https://doi.org/10.1002/nme.5135).
- [52] E. Bader, Z. Zhang, and K. Veroy. An empirical interpolation approach to reduced basis approximations for variational inequalities. *Math. Comput. Model. Dyn. Syst.*, 22(4):345–361, 2016. doi:[10.1080/13873954.2016.1198388](https://doi.org/10.1080/13873954.2016.1198388).
- [53] O. Burkovska, B. Haasdonk, J. Salomon, and B. Wohlmuth. Reduced basis methods for pricing options with the black–scholes and heston models. *SIAM Journal on Financial Mathematics*, 6(1):685–712, 2015. doi:[10.1137/140981216](https://doi.org/10.1137/140981216).
- [54] J. Oden and N. Kikuchi. Theory of variational inequalities with applications to problems of flow through porous media. *International Journal of Engineering Science*, 18(10):1173–1284, 1980. doi:[10.1016/0020-7225\(80\)90111-1](https://doi.org/10.1016/0020-7225(80)90111-1).

- [55] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and Their Applications*. Society for Industrial and Applied Mathematics, 2000. doi:[10.1137/1.9780898719451](https://doi.org/10.1137/1.9780898719451).
- [56] M. Sofonea and A. Matei. *Variational Inequalities with Applications*, volume 18 of *Advances in Mechanics and Mathematics*. Springer New York, New York, NY, 2009. doi:[10.1007/978-0-387-87460-9](https://doi.org/10.1007/978-0-387-87460-9).
- [57] D. Manvelyan, B. Simeon, and U. Wever. An efficient model order reduction scheme for dynamic contact in linear elasticity. *Computational Mechanics*, 68(6):1283–1295, dec 2021. doi:[10.1007/s00466-021-02068-4](https://doi.org/10.1007/s00466-021-02068-4).
- [58] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791, oct 1999. doi:[10.1038/44565](https://doi.org/10.1038/44565).
- [59] K. Veroy and A. T. Patera. Certified real-time solution of the parametrized steady incompressible navier–stokes equations: rigorous reduced-basis a posteriori error bounds. *International Journal for Numerical Methods in Fluids*, 47(8-9):773–788, 2005. doi:[10.1002/flid.867](https://doi.org/10.1002/flid.867).
- [60] E. Liberge and A. Hamdouni. Reduced order modelling method via proper orthogonal decomposition (pod) for flow around an oscillating cylinder. *Journal of Fluids and Structures*, 26(2):292–311, 2010. ISSN 0889-9746. doi:[10.1016/j.jfluidstructs.2009.10.006](https://doi.org/10.1016/j.jfluidstructs.2009.10.006).
- [61] C. R. Harris, K. J. Millman, S. J. van der Walt, and et. al. Array programming with NumPy. *Nature*, 585(7825):357–362, Sept. 2020. doi:[10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2).
- [62] R. H. Davies. *Learning shape: optimal models for analysing natural variability*. PhD thesis, University of Manchester, 2002. URL <http://www.bmva.org/theses/2002/2002-davies.pdf>.
- [63] M. Elad. *Sparse and Redundant Representations*. Springer New York, New York, NY, 2010. ISBN 978-1-4419-7010-7. doi:[10.1007/978-1-4419-7011-4](https://doi.org/10.1007/978-1-4419-7011-4).
- [64] S. L. Brunton, J. L. Proctor, and J. N. Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, apr 2016. doi:[10.1073/pnas.1517384113](https://doi.org/10.1073/pnas.1517384113).
- [65] J. Mairal, F. Bach, and J. Ponce. Sparse modeling for image and vision processing. *Foundations and Trends in Computer Graphics and Vision*, 8(2-3):85–283, 2014. doi:[10.1561/06000000058](https://doi.org/10.1561/06000000058).
- [66] R. Tibshirani. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, jan 1996. doi:[10.1111/j.2517-6161.1996.tb02080.x](https://doi.org/10.1111/j.2517-6161.1996.tb02080.x).

- [67] E. Candes and T. Tao. The Dantzig selector: Statistical estimation when  $p$  is much larger than  $n$ . *The Annals of Statistics*, 35(6):2313–2351, dec 2007. doi:[10.1214/009053606000001523](https://doi.org/10.1214/009053606000001523).
- [68] Y. Pati, R. Rezaifar, and P. Krishnaprasad. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, pages 40–44 vol.1, 1993. doi:[10.1109/ACSSC.1993.342465](https://doi.org/10.1109/ACSSC.1993.342465).
- [69] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, 1993. doi:[10.1109/78.258082](https://doi.org/10.1109/78.258082).
- [70] I. Gorodnitsky and B. Rao. Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm. *IEEE Transactions on Signal Processing*, 45(3):600–616, mar 1997. doi:[10.1109/78.558475](https://doi.org/10.1109/78.558475).
- [71] W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi. Predicting catastrophes in nonlinear dynamical systems by compressive sensing. *Phys. Rev. Lett.*, 106:154101, Apr 2011. doi:[10.1103/PhysRevLett.106.154101](https://doi.org/10.1103/PhysRevLett.106.154101).
- [72] H. Schaeffer, R. Caffisch, C. D. Hauck, and S. Osher. Sparse dynamics for partial differential equations. *Proceedings of the National Academy of Sciences*, 110(17):6634–6639, 2013. doi:[10.1073/pnas.1302752110](https://doi.org/10.1073/pnas.1302752110).
- [73] A. Balabanov, O. and Nouy. Randomized linear algebra for model reduction—part II: minimal residual methods and dictionary-based approximation. *Advances in Computational Mathematics*, 47(2):26, apr 2021. doi:[10.1007/s10444-020-09836-5](https://doi.org/10.1007/s10444-020-09836-5).
- [74] N. Halko, P. G. Martinsson, and J. A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53(2):217–288, 2011. doi:[10.1137/090771806](https://doi.org/10.1137/090771806).
- [75] M. Manucci, J. V. Aguado, and D. Borzacchiello. Sparse data-driven quadrature rules via  $\ell^p$ -quasi-norm minimization. *Computational Methods in Applied Mathematics*, 2022. doi:[10.1515/cmam-2021-0131](https://doi.org/10.1515/cmam-2021-0131).
- [76] C. L. Lawson and R. J. Hanson. *Solving Least Squares Problems*. Society for Industrial and Applied Mathematics, jan 1995. ISBN 978-0-89871-356-5. doi:[10.1137/1.9781611971217](https://doi.org/10.1137/1.9781611971217).
- [77] P. Virtanen, R. Gommers, T. Oliphant, and et. al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi:[10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2).
- [78] C. Bregler and S. Omohundro. Nonlinear image interpolation using manifold learning. In G. Tesauro, D. Touretzky, and T. Leen, editors, *Advances in Neural Information Processing Systems*, volume 7. MIT

- Press, 1995. URL <https://proceedings.neurips.cc/paper/1994/file/d516b13671a4179d9b7b458a6ebdeb92-Paper.pdf>.
- [79] A. Ghodsi. Dimensionality reduction a short tutorial. 2006.
- [80] C. Fefferman, S. Mitter, and H. Narayanan. Testing the manifold hypothesis. *Journal of the American Mathematical Society*, 29(4):983–1049, 2016. doi:[10.1090/jams/852](https://doi.org/10.1090/jams/852).
- [81] M. Dihlmann, M. Drohmann, and B. Haasdonk. Model reduction of parametrized evolution problems using the reduced basis method with adaptive time-partitioning. *Proc. of ADMOS*, 2011:64, 2011.
- [82] D. Amsallem, M. J. Zahr, and C. Farhat. Nonlinear model order reduction based on local reduced-order bases. *International Journal for Numerical Methods in Engineering*, 92(10):891–916, 2012. doi:[10.1002/nme.4371](https://doi.org/10.1002/nme.4371).
- [83] D. Amsallem and C. Farhat. Interpolation method for adapting reduced-order models and application to aeroelasticity. *AIAA Journal*, 46(7):1803–1813, 2008. doi:[10.2514/1.35374](https://doi.org/10.2514/1.35374).
- [84] B. Haasdonk and G. Santin. Greedy Kernel Approximation for Sparse Surrogate Modeling. In *Reduced-Order Modeling (ROM) for Simulation and Optimization*, pages 21–45. Springer International Publishing, Cham, 2018. doi:[10.1007/978-3-319-75319-5\\_2](https://doi.org/10.1007/978-3-319-75319-5_2).
- [85] D. González, J. V. Aguado, E. Cueto, E. Abisset-Chavanne, and F. Chinesta. kPCA-Based Parametric Solutions Within the PGD Framework. *Archives of Computational Methods in Engineering*, 25(1):69–86, jan 2018. ISSN 1134-3060. doi:[10.1007/s11831-016-9173-4](https://doi.org/10.1007/s11831-016-9173-4).
- [86] P. Díez, A. Muixí, S. Zlotnik, and A. García-González. Nonlinear dimensionality reduction for parametric problems: A kernel proper orthogonal decomposition. *International Journal for Numerical Methods in Engineering*, 122(24):7306–7327, 2021. doi:[10.1002/nme.6831](https://doi.org/10.1002/nme.6831).
- [87] A. García-González, A. Huerta, S. Zlotnik, and P. Díez. A kernel Principal Component Analysis (kPCA) digest with a new backward mapping (pre-image reconstruction) strategy. 2020. URL <http://arxiv.org/abs/2001.01958>.
- [88] H. Bansal, S. Rave, L. Iapichino, W. Schilders, and N. van de Wouw. Model Order Reduction Framework for Problems with Moving Discontinuities. In F. J. Vermolen and C. Vuik, editors, *Numerical Mathematics and Advanced Applications ENUMATH 2019*, pages 83–91, Cham, 2021. Springer International Publishing. ISBN 978-3-030-55874-1.
- [89] G. Welper. Transformed snapshot interpolation. may 2015. URL <http://arxiv.org/abs/1505.01227>.

- 
- [90] L. Rabiner, L. Rabiner, and B. Juang. *Pattern-Comparison Techniques*. Prentice-Hall Signal Processing Series: Advanced monographs. PTR Prentice Hall, 1993. ISBN 9780130151575.
- [91] M. Müller. *Dynamic Time Warping*, pages 69–84. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. ISBN 978-3-540-74048-3. doi:[10.1007/978-3-540-74048-3.4](https://doi.org/10.1007/978-3-540-74048-3.4).
- [92] I. Talmi, R. Mechrez, and L. Zelnik-Manor. Template matching with deformable diversity similarity. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1311–1319, 2017. doi:[10.1109/CVPR.2017.144](https://doi.org/10.1109/CVPR.2017.144).
- [93] M. Shokoohi-Yekta, B. Hu, H. Jin, J. Wang, and E. Keogh. Generalizing DTW to the multi-dimensional case requires an adaptive approach. *Data Mining and Knowledge Discovery*, 31(1):1–31, Jan. 2017. ISSN 1384-5810, 1573-756X. doi:[10.1007/s10618-016-0455-0](https://doi.org/10.1007/s10618-016-0455-0).
- [94] T. Giorgino. Computing and visualizing dynamic time warping alignments in r: The dtw package. *Journal of Statistical Software*, 31(7):1–24, 2009. doi:[10.18637/jss.v031.i07](https://doi.org/10.18637/jss.v031.i07).
- [95] A. Almog, A. Levi, and A. Bruckstein. Spatial de-interlacing using dynamic time warping. In *IEEE International Conference on Image Processing 2005*, volume 2, pages II–1010, 2005. doi:[10.1109/ICIP.2005.1530229](https://doi.org/10.1109/ICIP.2005.1530229).
- [96] H. Lei and B. Sun. A study on the dynamic time warping in kernel machines. In *2007 International Conference on Signal-Image Technology & Internet-Based Systems*, pages 839–845, Los Alamitos, CA, USA, dec 2007. IEEE Computer Society. doi:[10.1109/SITIS.2007.112](https://doi.org/10.1109/SITIS.2007.112).

---

**Titre :** Approximations parcimonieuses et de rang faible pour la mécanique des contacts

**Mot clés :** modèles d'ordre réduit, mécanique des contacts, approximations dictionnaires, réduction non-linéaire de dimensions

**Résumé :** Les stratégies typiques de modélisation réduite de la mécanique des contacts utilisent des approximations de rang inférieur. L'hypothèse sous-jacente est l'existence d'un sous-espace et d'un sous-cône de faible dimension pour les champs de déplacement et de pression de contact, respectivement. Cependant, la pression de contact présente un caractère local, car la zone de contact peut varier en fonction de paramètres tels que le chargement ou la géométrie.

Dans cette thèse, la pertinence des approximations de rang faible pour la mécanique des contacts est étudiée et des voies alternatives fondées sur des techniques de régression parcimonieuses sont explorées. Il est montré que la nature locale de la pression de contact conduit à une perte de la séparabilité, limitant ainsi la pré-

cision des méthodes de rang faible.

Par la suite et pour atténuer les problèmes d'in-séparabilité, les approximations utilisant des dictionnaires contenant un grand nombre d'atomes sont étudiées. Deux stratégies sont établies à l'aide de méthodes d'approximation parcimonieuses. L'une est fondée sur une approche d'ensemble actif où les éléments du dictionnaire sont sélectionnés de manière gloutonne, et l'autre sur des approximations d'enveloppes convexes qui éliminent le besoin d'appliquer explicitement des contraintes de non-pénétration.

Enfin, les méthodes de réduction non linéaires sont explorées. Elles permettent d'enrichir le dictionnaire en utilisant des interpolations non-linéaires, réduisant ainsi le coût de calcul associé à la construction du dictionnaire dans la phase hors-ligne.

---

**Title:** Low-rank and sparse approximations for contact mechanics

**Keywords:** model order reduction, contact mechanics, dictionary-based approximation, non-linear dimensionality reduction

**Abstract:** Typical strategies for reduced modelling of contact mechanics use low-rank approximations. The underlying hypothesis is the existence of low-dimensional subspace and subcone for the displacement and contact pressure fields, respectively. However, the contact pressure exhibits a local nature, as the contact position can vary with parameters like loading or geometry.

The adequacy of low-rank approximations for contact mechanics is studied and alternative routes based on sparse regression techniques are explored. It is shown that the local nature of contact pressure leads to a loss in linear separability, thereby limiting the accuracy of low-rank methods.

Subsequently, approximations using overcomplete dictionaries to mitigate the inseparability issues is investigated. Two strategies are devised using sparse approximation methods. One is based on an active-set approach where the dictionary elements are selected greedily, and another is based on convex hull approximations where the non-penetration constraints need not be explicitly enforced.

Lastly, the non-linear dimensionality reduction framework is explored. The snapshot set computed in the training phase is enriched at a low complexity using non-linear interpolations, thereby reducing the burden of creating overcomplete dictionaries in the offline phase.