The Pennsylvania State University
The Graduate School

CONVEX APPROXIMATION OF CHANCE CONSTRAINED PROBLEMS: APPLICATION IS SYSTEMS AND CONTROL

A Dissertation in
Electrical Engineering
by
Ashkan M. Jasour

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2017

The dissertation of Ashkan M. Jasour was reviewed and approved* by the following:

Constantino Lagoa
Professor of Electrical Engineering
Dissertation Advisor, Chair of Committee

Necdet Serhat Aybat
Assistant Professor of Industrial Engineering
Dissertation Co-Advisor, Co-Chair of Committee

Vishal Monga
Associate Professor of Electrical Engineering

Minghui Zhu
Assistant Professor of Electrical Engineering

Antonios Armaou
Associate Professor of Chemical Engineering

Alexei Novikov
Professor of Mathematics

Kultegin Aydin
Professor and Department Head of Electrical Engineering

*Signatures are on file in the Graduate School.

Abstract

This dissertation concentrates on chance constrained optimization problems and their application in systems and control area. In chance optimization problems, we aim at maximizing the probability of a set defined by polynomial inequalities involving decision and uncertain parameters. These problems are, in general, nonconvex and computationally hard. With the objective of developing systematic numerical procedures to solve such problems, a sequence of convex relaxations based on the theory of measures and moments is provided, whose sequence of optimal values is shown to converge to the optimal value of the original problem. Indeed, we provide a sequence of semidefinite programs of increasing dimension which can arbitrarily approximate the solution of the original problem. In addition, we apply obtained results on chance optimization problems to challenging problems in the area of systems, control and data science. We consider the problem of probabilistic control of uncertain systems to ensure that the probability of defined failure/success is minimized/maximized. In particular, we consider the probabilistic robust control and chance constrained model predictive control problems. We also use the obtained results to analysis of stochastic and deterministic systems. More precisely, we address the problem of uncertainty set propagation and computing invariant robust set for uncertain systems and problem of computing region of attraction set for deterministic systems. In the problem of uncertainty propagation, we propagate the set of initial sets through uncertain dynamical systems and find the uncertainty set of states of the system for given time step. In the problem of region of attraction and invariant robust set, we aim at finding the largest set of all initial states whose trajectories converge to the origin. Moreover, we present the problem of corrupted and sparse data reconstruction where we want to complete the data with least possible complexity. In this thesis, to be able to efficiently solve the resulting large-scale problems, a first-order augmented Lagrangian algorithm is also implemented. Numerical examples are presented to illustrate the computational performance of the proposed approach.

Table of Contents

List of Figures viii

List of Tables x

List of Symbols xi

List of Abbreviation xiii

Acknowledgments xiv

Chapter 1
Introduction **1**
1.1 Classical Methods of Chance Optimization and Control 3
1.2 Contributions 8
1.3 The Sequel 10

Chapter 2
Preliminary Results on Measures, Polynomials, and Semidefinite Programs **11**
2.0.1 Polynomial Functions 11
2.0.2 Measures and Moments 12
2.0.3 Preliminary Results on Measures and Polynomials 16
2.0.4 Linear and Semidefinite Programming 18

Chapter 3
Chance Constrained Optimization **21**
3.1 Introduction 21
3.2 Chance Optimization over a Semialgebraic Set 22

3.2.1 An Equivalent Problem 23
3.2.2 Semidefinite Relaxations 24
3.2.3 Discussion on Improving Estimates of Probability 26
3.2.4 Simple Examples 28
3.2.5 Orthogonal Basis 31
3.2.6 Dual Convex Problem on Function Space 33
3.3 Chance Optimization over a Union of Sets 35
3.3.1 An Equivalent Problem 36
3.3.2 Semidefinite Relaxations 37
3.4 Implementation and Numerical Results 38
3.4.1 Regularized Chance Optimization Using Trace Norm 38
3.4.2 First-Order Augmented Lagrangian Algorithm 39
3.4.3 Numerical Examples 42
3.4.3.1 Monte Carlo Simulation 42
3.4.3.2 Example 1: A Simple Semialgebraic Set 43
3.4.3.3 Example 2: Union of Simple Sets 44
3.4.3.4 Example 3: Portfolio Selection Problem 45
3.4.3.5 Example 4: Run time 46
3.5 Conclusion 47
3.6 Appendix A: Proof of Theorem 10 48
3.7 Appendix B: Proof of Lemma 11 49
3.8 Appendix C: Proof of Theorem 12 49
3.9 Appendix D: Proof of Theorem 13 51
3.10 Appendix E: Proof of Theorem 14 52
3.11 Appendix F: Proof Of Theorem 15 54
3.12 Appendix G: Proof Of Theorem 16 54

Chapter 4
Convex Relaxation of Probabilistic Controller Design Problems 56
4.1 Introduction 56
4.2 Probabilistic Robust Control 57
4.2.1 Problem Statement 57
4.2.2 An Equivalent Problem 59
4.2.2.1 Set Invariant Control Laws 59
4.2.2.2 Maximizing Probability of Reaching χ_N 60
4.2.3 Semidefinite Relaxations 62
4.2.4 Numerical Results 63
4.2.4.1 Example 1: Nonlinear Control Problem 63
4.3 Chance Model Predictive Control 68
4.3.1 Problem Formulation 69

4.3.2 Equivalent Convex Problem on Measures 71
4.3.3 Semidefinite Programming Relaxations 73
4.4 Numerical results 74
4.5 Conclusion 76
4.5.1 Appendix A: Proof of Theorem 17 77
4.5.2 Appendix B: Proof of Theorem 18 78
4.5.3 Appendix C: Proof of Theorem 19 78
4.5.4 Appendix D: Proof of Theorem 20 78
4.5.5 Appendix E: Proof of Theorem 21 79
4.5.6 Appendix F: Proof of Theorem 22 80

Chapter 5
Uncertainty Propagation Through Uncertain Dynamical Systems 81
5.1 Problem Statement 81
5.2 Moment Information 82
5.3 Support Reconstruction 83
5.3.1 Convex Formulation 83
5.3.2 An Heuristic for Improved Performance 87
5.3.3 Support Reconstruction for Uniform Measures 89
5.4 Conclusion 92

Chapter 6
Constrained Volume Optimization Problem 93
6.1 Introduction 93
6.2 Problem Statement 95
6.3 Applications in Systems and Control 96
6.3.1 Region of Attraction 96
6.3.2 Maximal Invariant Set 97
6.3.3 Generalized Sum of Squares Problem 98
6.4 Equivalent Convex Problem on Measures and Moments 99
6.4.1 Linear Program on Measures 100
6.4.2 Finite Semidefinite Programming on Moments 102
6.4.3 Illustrative Example 105
6.5 Dual Convex Problem on Function Space 107
6.5.1 Illustrative Example 109
6.6 Implementation and Numerical Results 110
6.6.1 Example 1: ROA set of system 110
6.7 Conclusion 113
6.8 Appendix A: Proof of Theorem 25 113
6.9 Appendix B: Proof of Theorem 26 114

6.10 Appendix C: Proof of Theorem 27 115
6.11 Appendix D: Proof of Theorem 28 116
6.12 Appendix E: Proof of Theorem 29 116
6.13 Appendix F: Proof of Theorem 30 117

Chapter 7
Sparse Data Reconstruction in Sensory Networks **120**
7.1 Introduction 120
7.2 Problem Formulation 121
7.3 Hankel Matrix 122
7.3.1 n-Dimensional Block Hankel Matrix Decomposition 126
7.3.2 Row Permutation of Block Hankel Matrix 128
7.3.3 Rank of Block Hankel Matrix 134
7.4 Equivalent Problem and Convex Relaxation 135
7.5 Implementation and Numerical Results 136
7.5.1 First-Order Augmented Lagrangian Optimization Algorithm 137
7.5.2 Numerical Examples 138
7.6 Conclusion 140
7.7 Appendix A: Proof of Theorem 31 141
7.8 Appendix B: Proof of Theorem 32 145

Chapter 8
Conclusion and Discussion **153**

Bibliography **155**

List of Figures

3.1 **a)** Simple chance optimization problem over semialgebraic set $\mathcal{K}$ with random parameter q, and decision variable x, **b)** Equivalent problem in the measure space over probability measure μ_x as variable for given probability measure μ_q, **c)** Probability of given semi algebraic set $\mathcal{K}$ for a fixed μ_x is equal to the integral of $\mathcal{K}$ with respect to the measure $\mu_x \times \mu_q$, **d)** The probability is equal to the volume of the measure μ which is supported on the set $\mathcal{K}$ and has the same distribution as the measure $\mu_x \times \mu_q$ over its support 24

3.2 $\mathbf{P}_d$, $\mathbf{P'}_d$, and $\tilde{\mathbf{P}}_d$ for increasing relaxation order d 29

3.3 **a)** f^*: the degree-100 polynomial approximation to indicator function of $\mathcal{F}(\boldsymbol{x}^*)$, **b)** h^*: the degree-100 polynomial approximation of the piecewise-polynomial function $\max(0, \mathcal{P}(\boldsymbol{x}^*, q))$ 31

3.4 $\mathbf{P}_d$, $\mathbf{P'}_d$, $\tilde{\mathbf{P}}_d^{(\mathbf{1})}$, $\tilde{\mathbf{P}}_d^{(\mathbf{2})}$, and $\tilde{\mathbf{P}}_d$ for increasing relaxation order d 31

3.5 $\mathbf{P}_d$ for monomial and Chebyshev polynomial bases 32

3.6 Polynomial $\mathcal{P}_{\mathcal{W}}^d(x, q)$ obtained by SDP (3.15) for $d = 12$ 34

3.7 β and $\int_{\mathcal{Q}} \mathcal{P}_{\mathcal{W}}^d(x, q) d\mu_q$ obtained by SDP (3.15) for $d = 10$ 35

3.8 first-order Augmented Lagrangian algorithm for Conic Convex (ALCC) problems 41

4.1 The polynomial $P_d\,(b_1, b_2)$ of example 2 63

4.2 Example 2: trajectories of the uncertain system under obtained control input 66

4.3 Example 3: trajectories of the uncertain system under obtained control input 66

4.4 The trajectories of the uncertain system of Example 4 controlled by obtained state feedback 68

5.1 Result of SDP in (5.6) For Example 1 87

5.2 Result of SDP in (5.7) For Example 1 88

5.3 Result of SDP in (5.7) For Example 2 89

5.4 Result of SDP in (5.7) For Example 3 90

5.5 Result of SDP in (5.10) For Example 1 91
5.6 Result of SDP in (5.10) For Example 4 92

6.1 Sets $\mathcal{K}_1$ and $\mathcal{K}_2$ 105
6.2 Polynomial $\mathcal{P}^d_{\mathcal{A}}(a)$ obtained by SDP (6.19) for $d = 7$ 106
6.3 Polynomial $\mathcal{P}^d_{\mathcal{A}}(a)$ obtained by SDP (6.19) for $d = 2, 4, 6, 7$ 107
6.4 Polynomial $\mathcal{P}^d_{\mathcal{W}}(x, a)$ obtained by SDP (6.25) for $d = 12$ 109
6.5 β and $\int_\chi \mathcal{P}^d_{\mathcal{W}}(x, a)d\mu_x$ obtained by SDP (6.25) for $d = 12$ 110
6.6 The set $\left\{(a_1, a_2, a_3) : \mathcal{P}^d_{\mathcal{A}}(a) \leq 1 - \epsilon_{\mathcal{A}}\right\}$ obtained by SDP (6.19) for $d = 10$ and $\epsilon_{\mathcal{A}} = 0.02$ 111
6.7 The sets $\mathcal{S}_1(a) = \{x \in \chi : 0 \leq V(x, a) \leq 1\}$ and $\mathcal{S}_2(a) = \{x \in \chi : \dot{V}(x, a) \leq -\epsilon_r \|x\|_2^2\}$ for obtained a 112
6.8 The true and estimated ROA sets 112

7.1 Original and sparse signal of Example 1 139
7.2 Results of Example 1: a) Reconstructed signal, b) Singular values of Hankel matrix of reconstructed sparse signal 140
7.3 Example 2: a) Original and sparse signal, b) Singular values of Hankel matrix constructed by sparse signal 141
7.4 Example 2: a) Reconstructed signal, b) Singular values of Hankel matrix of reconstructed sparse signal 142
7.5 a) corrupted image, b) reconstructed image 143
7.6 Singular values of Hankel matrix of a) corrupted and b)reconstructed image 143
7.7 Indoor light and temperature sensory data 144
7.8 Ocean light and temperature sensory data 144
7.9 50% sparse Ocean temperature data reconstruction 149
7.10 50% sparse Ocean light data reconstruction 150
7.11 50% sparse indoor light data reconstruction 151
7.12 50% sparse indoor Temperature data reconstruction 152

List of Tables

3.1 ALCC and GloptiPoly results for Example 1 44
3.2 ALCC results for Example 2 45
3.3 ALCC and GloptiPoly results for Example 3 47
3.4 ALCC for increasing problem in Example 4 47

4.1 ALCC and GloptiPoly results for Example 1 65

7.1 First-Order Augmented Lagrangian Optimization Algorithm 137

List of Symbols

Sets

$\mathbb{R}$	Set of real numbers
$\mathbb{Z}$	Set of integers
$\mathbb{N}$	Set of natural numbers
$\mathbb{C}$	Set of complex numbers
$\mathbb{R}[x]$	Ring of real polynomials
$\mathbb{R}_d[x]$	Set of polynomials of degree at most d
$\mathbb{S}^2[x]$	Set of sum of squares polynomials
$\mathcal{M}$	Space of finite Borel measures
$\mathcal{M}_+$	Cone of finite nonnegative Borel measures

Vector Operations

$\mathbf{p}^T$	Transpose of a vector $\mathbf{p}$
$\mathbf{p}_i$	i^{th} element of vector $\mathbf{p}$
$\mathcal{QM}$	Quadratic module

Matrix Operations

$\mathbf{A}_{:,j}$	j^{th} column of matrix $\mathbf{A}$
$\mathbf{A}_{i,:}$	i^{th} row of matrix $\mathbf{A}$
$\sigma(A)$	is the singular value vector of the matrix A
$trace(A)$	trace of the matrix A, that is $trace(A) \doteq \sum_i A_{i,i}$
$\mathcal{N}(A)$	is null space of matrix A
$rank(A)$	is rank of matrix A
A^T	is transpose of the matrix A
A^*	complex conjugate transpose of matrix A
$A \geq 0$	denotes a positive semi definite matrix
$diag(\mathbf{g})$	returns a square diagonal matrix with the elements of vector $\mathbf{g}$
$a_{i,j}$	denotes the ith row and the jth column of matrix A

Norms

$\|x\|_2$	is l_2 norm of the vector x, that is: $\|x\|_2 \doteq \sqrt{x^T x}$
$\|A\|_*$	is nuclear norm of the matrix A, that is: $\|A\|_* \doteq \sum_i \sigma_i(A)$

Optimization

$\min_x$	Function minimization over x, optimal function value is returned
argmin_x	Function minimization over x, optimal value of x is returned
$s.t.$	Subject to the constraints

List of Abbreviation

SDP	Semidefinite Programming
LP	Linear Program
MPC	Model Predictive Control
ROA	Region Of Attraction

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my advisor Professor Constantino Lagoa for the continuous support of my Ph.D study and research and for his inspiration and motivation. He has been a distinguished advisor and an true friend. I never would have made my Ph.D at Penn State without his support.

My sincere thanks also goes to my co-advisor Prof. Necdet Serhat Aybat who have provided me with valuable help. Besides my advisors, I would like to thank the members of my dissertation committee, Prof. Vishal Monga, Prof. Minghui Zhu, Prof. Antonios Armaou and Prof. Alexei Novikov, for their time, reviewing my dissertation and comments.

I am thankful to my lab-mates Korkut Bekiroglu and Emil Laftchiev for their help and friendship. I also would like to thank the wonderful friends who became my second family in State College. My special thanks goes to Sam and Gelareh, Bijan and Sharareh, Hesam and Azadeh, Roozbeh and Farima, Saba and Navid, Shahrzad Fadaei, Shideh Shams, Mahsa Masoudi, Farid Tayari, and Hojjat Mousavi.

Last but not the least, I would like to thank my family: my mother Fahimeh Kami who I owe my success to her prayers, my father Professor Davoud Jasour who is my role model in my life and my lovely sister, Sheida. I owe any success I achieve in my life to their support. Thank you for your unconditional love and support throughout my life.

Dedication

To my mom...

Chapter 1

Introduction

The work presented in this thesis addresses the risk bounded control and planning problems. In uncertain environments and systems, not only we need to specify the desired behavior and control objectives, but also the level of risk that we are willing to accept. In such environments, achieving predefined control objectives for all possible uncertainties results in conservative approaches. Instead of achieving control objectives for all possible scenarios, we aim at increasing the probability of achieving these predefined goals. For this purpose, we use a probabilistic approach to represent uncertainty and risk and introduce chance constrained optimization problems. In chance constrained optimization problem, we aim at minimizing the probability of failure or, on the other hand, maximizing the probability of success. Probability of success and failure are defined in terms of constraints involving decision and uncertain parameters with known probability distributions.

The potential application area of this problem class is quite large and encompasses many well-known problems in different areas as special cases. For example, designing probabilistic robust controllers, model predictive controllers in the presence of random disturbances, and optimal path planning and obstacle avoidance problems in robotics can be cast as special cases of this framework. Moreover, problems in the areas of economics, finance, and trust design can also be formulated as chance optimization problems. These problems are, in general, nonconvex and computationally hard. In this thesis, with the objective of developing systematic numerical procedures to solve such problems, a sequence of convex relaxations based on the theory of measures and moments is provided, whose sequence of optimal values is shown to converge to the optimal value of the original problem. Indeed, we provide a sequence of semidefinite programs of increasing dimension which can arbitrarily approximate the solution of the original problem.

We apply obtained results on chance optimization problem to design probabilistic robust controllers for uncertain dynamical systems. Probabilistic control formulations are used in different areas to deal with uncertain systems in order to ensure that the probability of failure/success is minimized/maximized. For example, minimizing probability of obstacle collision in motion planning of robotic systems under environment uncertainty can be formulated as instances of probabilistic control problems. In this thesis, we provide results aimed at designing robust controllers that maximize the probability of reaching a given target set. More precisely, we start with an uncertain polynomial system subjected to external perturbations for which we know the probability distribution of the initial state, the uncertainty and the disturbances. Then, given a target set defined by polynomial inequalities and number of steps N, we provide algorithms for designing a nonlinear state feedback control law that i) makes the target set a robustly invariant set and ii) maximizes the probability of reaching the target set in N steps.

Also, we provide chance constrained model predictive control problems whose objective is to obtain finite-horizon optimal control of dynamical systems subject to probabilistic constraints. The control laws provided are designed to have precise bounds on the probability of achieving the desired objectives. More precisely, consider a polynomial dynamical system subject to external perturbation and assume that the probability distribution of the disturbances at each time is known. Then, given a desired set defined by polynomial inequalities and a polynomial cost function defined in terms of states and control input of the system, we aim at designing a controller to i) minimize the expected value of given cost function over the finite horizon and ii) reach the given desired set with high probability. For this purpose, at each sampling time we solve a convex optimization problem that minimizes the expected value of cost function subject to probabilistic constraints over the finite horizon.

We also, address the problem of uncertainty propagation through dynamical systems where the aim is to find uncertainty set of states of the system for given time. This set can be employed for analysis of uncertain systems and robust control purposes.

Moreover, we generalize the obtained results on chance constrained problem and introduce constrained semialgebraic volume optimization problem. This framework enables us to find a convex equivalent problem for different problems with deterministic or probabilistic nature. In this problem, we aim at maximizing the volume of a semialgebraic set under some semialgebraic constraints. We show that many well-known problems can be formulated as a constrained volume optimization problem. In particular, we address the problem of computing region of attraction set and maximal invariant set for uncertain systems.

At the end, we show the application of convex methods in data science and address the problem of reconstructing noisy sparse data. In this problem, we aim at finding the missing part of given data with least possible complexity. This problem arises in different areas like sensory networks, where the sensors do not completely cover the area of interest; Hence, the sampled data are usually inadequate. Moreover, reconstruction of corrupted image or videos can can be cast in this framework. In this thesis, we propose a Hankel based approach to address this problem, where the data is completed such that the number of exponential signals that can describe the data, becomes minimum.

1.1 Classical Methods of Chance Optimization and Control

Several approaches have been proposed to solve chance constrained problems. The main idea behind most of the proposed methods is to find a tractable approximation for chance constraints. One particular method is the so-called *scenario approach;* see [1, 2, 3, 4, 5] and the references therein. In this approach, the probabilistic constraint is replaced by a (large) number of deterministic constraints obtained by drawing independent identically distributed (iid) samples of random parameters. Being a randomized approach, there is always a positive probability of failure (perhaps small). In [6, 7, 8, 9, 10], robust optimization is used to deal with uncertain linear programs (LP). In this method, the uncertain LP is replaced by its robust counterpart, where the worst case realization of uncertain data is considered. The proposed method is not computationally tractable for every type of uncertainty set. A specific case that is tractable is LP with ellipsoidal uncertainty set [10]. In [11, 12, 13], an alternative approach is proposed where one analytically determines an upper bound on the probability of constraint violation. Although this method does provide a convex approximation, it can only be applied to specific uncertainty structures. In [14, 15] the authors propose the so-called Bernstein approximation where a convex conservative approximation of chance constraints is constructed using generating functions. Although approximation is efficiently computable, it is only applicable to problems with convex constraints that are affine in random vector $q \in \mathbb{R}^m$. Moreover, components of q need to be independent and have computable finite generating functions. In [16, 17, 18] convex relaxations of chance constrained problems are presented. The concept of polynomial kinship function is used to estimate an upper bound on the probability of constraint violation. Solutions to a sequence of relaxed problems are shown to converge to a solution of the original problem as the degree of the polynomial kinship function increases along the sequence. In [18, 19], an equivalent convex formulation is provided based on the theory of moments. In this method the probability of a polynomial being negative is approximated by computing polynomial approximations for univariate indicator functions [19].

Distributionally robust chance constrained programming – (see [20, 21, 22, 23, 24]), is another popular tool for dealing with uncertainty in the problem, where only a finite number of moments m_α of the underlying measure $\bar{\mu}_q$ are assumed to be known, i.e., $\{m_\alpha\}_{\alpha\in A}$ is known for $A \subset \mathbb{N}^m$ such that $|A| < \infty$. In this approach robust chance constraints are formulated by considering the worst case measure within a family of measures with moments equal to $\{m_\alpha\}_{\alpha\in A}$. However, proposed methods in this literature are mainly limited to linear chance constraints and/or to specific types of uncertainty distributions. For instance, in [20], under the assumption $\bar{m} = E_{\bar{\mu}_q}[q]$ and $\bar{S} = E_{\bar{\mu}_q}[(q-\bar{m})(q-\bar{m})^T]$ are known, the linear chance constraint of the form $\bar{\mu}_q\left(\{q : \ q^T x \geq 0\}\right) \geq 1-\epsilon$ is replaced by its robust counterpart: $\inf_{\mu_q \in \mathcal{M}} \mu_q\left(\{q : \ q^T x \geq 0\}\right) \geq 1-\epsilon$, where $\mathcal{M}$ is the set of finite (positive) Borel measures on $\bar{\Sigma}_q$ with their means and covariances equal to $\bar{m}$ and $\bar{S}$, respectively; and it is shown that these robust constraints can be represented as second-order cone constraints for a wide class of probability distributions. In [21], the authors has reviewed and developed different approximation methods for problems with joint chance constraints. In the proposed method, joint chance constraints are decomposed into individual chance constraints, and classical robust optimization approximation is used to deal with the new constraints. In [22] a tractable approximation method for probabilistically dependent linear chance constraints is presented. In [23] linear chance constraints with Gaussian and log-concave uncertainties are addressed, and it is shown that they can be reformulated as semi-infinite optimization problems; moreover, tight probabilistic bounds are provided for the resulting *comprehensive robust optimization* problems [25, 26]. In [24] an SDP formulation is provided to approximate distributionally robust chance constraints where only the support of $\bar{\mu}_q$, and its first and second order moments are known.

Probabilistic control methods are used in different areas to deal with uncertain systems in order to ensure that the probability of defined failure/success is minimized/maximized. Several approaches have been proposed to involve statistics of uncertainty in control procedure of uncertain systems. The main approaches are (i) adding probabilistic constraints on states and inputs of system (e.g., see [27, 28]), (ii) minimizing the expected value of the objective function (e.g., see [27, 28]). The main problem in the formulation of chance constrained control is the efficient evaluation of the probabilistic constraints. Hence, several approaches have been proposed to provide tractable approximations of the chance constraints involved in a probabilistic control design problem. One of such methods is the so-called *randomized approach;* see [28, 28, 29, 30, 31] and references therein. In this case, the probabilistic constraint is replaced by a (large) number of deterministic constraints obtained by drawing iid samples of the random parameters. Being a randomized approach, there is always a (perhaps small) probability of failure of the algorithm. In [32, 33, 34, 35, 36, 37], an alternative approach is proposed where one analytically determines an upper bound on the probability of constraint violation. In [37] expected value of uncertain objective function is proposed using the notion of particles. It tries to approximate the distribution of the system state using a finite number of particles.

Probabilistic formulations of model predictive control can be used in different areas to deal with systems subject to disturbances. The MPC method is an optimal control based method, which a finite cost function is optimized at every sampling time under imposed constraints. At each sampling time, MPC needs to predict the future states of the system over the finite horizon using the dynamic of the system. To deal with uncertain parameters of the system and disturbance, several approaches have been proposed. In ([38, 39]), robust MPC for linear and polynomial systems are proposed where robust constraints are employed. In this method, MPC is formulated considering the a bundle of trajectories for all possible realizations of the uncertainty. The robust MPC methods are conservative, due to the requirement of robust feasibility for all disturbance realizations. In ([40, 41, 42, 43, 44]), adaptive MPC are provided where neural networks are used to predict the future behavior of the system. Using the online training algorithm, robustness against changes in the robot parameters is obtained. In ([36, 45, 14]), to deal with model uncertainty the probabilistic constraints are used. In ([36, 45]) probabilistic constraints for linear systems are replaced with hard constrained assuming the Gaussian distribution for uncertainty. In [14], a semialgabriac approximation of the probabilistic constraints are obtained.

In dynamical systems with uncertain parameters and initial states, the states of the system at each time are uncertain. This uncertainty set is a result of propagation of initial states set through uncertain dynamical system. Computing this set of uncertainty for states of system could be used for analysis and robust control of uncertain systems. This set can be obtained using the information of the probability distribution of system states when probabilistic representation of uncertainty is considered. More precisely, to compute such set of uncertainty one needs to reconstruct the support of probability distribution of system states. Several approaches have been proposed to construct the support from the moments information. In [46] an approach to exact reconstruction of convex polytope supports is proposed, which is based on the collection of moment formulas combined with Vandermonde factorization of finite rank Hankel matrices. In [47], a method to reconstruct planar semi-analytic domains from their moments is proposed based on the diagonal Pade approximation where it can approximate arbitrarily closely any bounded domain. The approach in [48] provides a method to obtain a polynomial that vanishes on the boundary of support. However, obtained polynomial may vanish inside the support or on any other points as well. Hence, one can not use the polynomial to reconstruct the support.

In this thesis, we extend the results obtained for chance optimization and develop a framework that enables us to find convex formulation for well known problems in system and control. Building on theory of measures and moments as well as theory of sum of squares polynomials, many approaches have been proposed to reformulate different problems in the area of control and system as a convex optimization problems with Linear Matrix Inequalities (LMI). The methods in ([49, 50, 51]), provide hierarchy of finite dimensional LMI relaxations to compute polynomial outer approximation of the region of attraction set and maximum controlled invariant set for polynomial systems. The concept of occupation measure is used to reformulate the original problem to truncated moment problem. It is shown that the optimal value of the provided LMI converges to the volume of desired sets and on the other hand the optimal value of the dual problem in continuous function space converges to the polynomial outer approximation of the set. This proposed method is modified in [50] to obtain a inner approximation for region of attraction set for finite time-horizon polynomial systems. In this case, outer approximations of complement of region of attraction set is computed. In ([52, 53]) sum of square formulation to find a suitable Lyapunov function for dynamical system and approximation of region of attraction set is provided. In provided approach, one needs to look for SOS Lyapunov function whose negative derivative is also SOS. Also, a repetitive control design approach is provided where enables to maximize the ROA set of the system.

We also show the application of convex methods in data science and address the problem of reconstructing noisy sparse data. This problem could arise in data networks [54, 55] and sensory networks. For example, incomplete sensory networks, where the sensors do not completely cover the area of interest; Hence, the sampled data are usually inadequate. Moreover, reconstruction of corrupted image or videos can can be cast in this framework. Several approaches to construct the data from spars measurement have been proposed. The main idea in the most of proposed methods is to use interpolation methods to setup an approximating statistical model for measured data. For example, the methods K-Nearest Neighbors [56] is a local interpolation method that uses the nearest neighbors to estimated the missing data. It uses Bayes decision procedure to estimated the missing data and bound the probability of error. This approach does not depend on any assumptions about the underlying statistics for its application. On the other hand, the Delaunay Triangulation [57] method is a global interpolation methods that use vertices of data and consider the global error in interpolation. In [58] a data adaptive method called

Multi-channel Singular Spectrum Analysis (MSSA) is proposed which works based on the embedded lag-covariance matrix of measured data.

These proposed interpolation methods needs large number of measurements and fail when the number of missing data grow. To solve this problem, a method called compressed sensing [59] is proposed. In this method the number of measurements can be dramatically smaller than size of missing data. This method uses the weighted linear combination of samples to reconstruct the data. Using the randomly selected samples, a underdetermined linear system is built whose solution is used to construct the data. In solving linear system, the assumption of sparsity for the initial signal is made. However, for accurate environment reconstruction, this method requires some inherent structure [60, 61]. The sampling matrices constructed by measurement samples are required to satisfy certain conditions, for example the mutual coherence of matrices should have a small mutual coherence, the largest absolute and normalized inner product between different columns in matrix. Actually, the optimal matrices where the CS algorithm performs well are Gaussian iid matrices, uniform random ortho-projectors, or Bernoulli matrices.

In [62, 63, 64] the proposed methods seeks low-rank structure for estimating the missing data. [62] proposes a approach based on compressed sensing to reconstruct the massive missing data. It develops an environmental space time improved compressed sensing (ESTICS) algorithm for estimating the missing data, where it computes low-rank approximations of the incomplete matrix called environment matrix. In [63, 64] a Hankel operator based approach to the problems of texture modeling and in-painting is proposed. It models textured images as the output of an unknown, periodic, linear shift invariant operator in response to a suitable input and solves sequence of rank minimization problems.

1.2 Contributions

In this thesis, we take a different approach to deal with chance constrained problems [65, 66]. The proposed method is based on volume approximation results in [67] and the theory of measures and moments [68, 69]. In [67], a hierarchy of SDP problems is proposed to compute the volume of a given compact semialgebraic set. It is shown that the volume of a semialgebraic set can be computed by solving a maximization problem over finite (positive) Borel measures supported on the given set, and restricted by the Lebesgue measure on a simple set containing the semialgebraic set of interest. Building on volume approximation results, we propose a convex approximation method to address chance optimization problems over semialgebraic sets. In particular, we address the problem of probability maximization over the union of semialgebraic sets defined by intersections of a finite number of polynomial inequalities. Here, one needs to search for the (positive) Borel measure with maximum possible mass on the given semialgebraic set while simultaneously searching for an upper bound probability measure over a simple set containing the semialgebraic set and restricting the Borel measure. Using the theory of moments, we provide sequence of semidefinite programs (SDP's) where one need to look for moment sequences of measures of chance optimization problem. To solve resulted SDP's, a first-order augmented Lagrangian algorithm is implemented that enables us to solve large scale chance optimization problems.

To show the application of chance optimization problems in control and systems, we consider probabilistic control of uncertain systems [70, 71]. In the problem of probabilistic control, we incorporate the probability directly in the objective function and aim at maximizing the probability of desired defined control objectives. We, also, consider the problem of uncertainty set propagation through stochastic dynamical systems where enables us to compute uncertainty at each time step. Building on theory of measures and moment, we provide semidefinite program to approximate the set of uncertainty [72].

Moreover, we generalize the results obtained for chance optimization problems and define constrained volume optimization problems where enables us to obtain convex formulation of different challenging problems in system and control [73]. We reformulate problems of computing region of attraction and invariant set of dynamical systems as a particular case of constrained volume optimization problems.

Finally, we address the problem of data reconstruction and motivated by low-rank structure methods, we propose a novel approach to reconstruct a noisy sparse data with least possible complexity. To obtain the complete data, we look for minimum rank block Hankel matrix associated with given sparse and noisy data. We show that minimizing the rank of constructed block Hankle matrix is equivalent to minimizing the number of exponential signals that describes the data. The proposed method, with out making any assumption on the structure of data and data loss pattern, could reconstruct the complete data.

1.3 The Sequel

The outline of this thesis is as follows: in Chapter 2, the notation adopted in the paper and preliminary results on measure and polynomial theory and also linear and semidefinite programs are presented; in Chapter 3 we address the chance optimization problem. We first start with semialgebraic set involving intersection of polynomials and provide equivalent convex problem in measure and moment spaces. We then extend the result for more general case with semialgebraic set involving union and intersection of polynomials. To be able to solve large problems we present First order Lagrangian algorithm in this chapter. In Chapter 4, the problem of probabilistic robust controller design and also chance constrained model predictive controllers are addressed. In this chapter, semidefinite programs are provided to solve the original problems. In Chapter 5, the problem of uncertainty propagation through dynamical systems are presented. In Chapter 6, we introduce the problem of constrained volume optimization problem and address the problem of computing the region of attraction and also robust invariant set of uncertain systems as a particular case of this optimization problem. The problem of noisy and sparse data reconstruction is presented in Chapter 7 and finally in Chapter 8, we give concluding remarks and some future direction of this research.

Chapter 2

Preliminary Results on Measures, Polynomials, and Semidefinite Programs

In this thesis, building on the theory of measure and moments as well as theory of polynomials, we develop our semidefinite programs to approximate the optimal solution of the original problems. Hence, in this chapter the mathematical background and some basic definitions on polynomial and measure theory as well as linear and semidefinite programming are presented.

2.0.1 Polynomial Functions

Let $\mathbb{R}[x]$ be the ring of real polynomials in the variables $x \in \mathbb{R}^n$. Given $\mathcal{P} \in \mathbb{R}[x]$, we represent $\mathcal{P}$ as $\sum_{\alpha\in\mathbb{N}^n} p_\alpha x^\alpha$ using the standard basis $\{x^\alpha\}_{\alpha\in\mathbb{N}^n}$ of $\mathbb{R}[x]$, and $\mathbf{p} = \{p_\alpha\}_{\alpha\in\mathbb{N}^n}$ denotes the polynomial coefficients. We assume that the elements of the coefficient vector $\mathbf{p} = \{p_\alpha\}_{\alpha\in\mathbb{N}^n}$ are sorted according to grevlex order on the corresponding monomial exponent α. Given n and d in $\mathbb{N}$, we define $S_{n,d} := \binom{d+n}{n}$ and $\mathbb{N}^{\mathrm{n}}_d := \{\alpha \in \mathbb{N}^n : \|\alpha\|_1 \leq d\}$. Let $\mathbb{R}_{\mathrm{d}}[x] \subset \mathbb{R}[x]$ denote the set of polynomials of degree at most $d \in \mathbb{N}$, which is indeed a vector space of dimension $S_{n,d}$. Similarly to $\mathcal{P} \in \mathbb{R}[x]$, given $\mathcal{P} \in \mathbb{R}_{\mathrm{d}}[x]$, $\mathbf{p} = \{p_\alpha\}_{\alpha\in\mathbb{N}^{\mathrm{n}}_d}$ is sorted such that $\mathbb{N}^{\mathrm{n}}_d \ni \mathbf{0} = \alpha^{(1)} <_g \ldots <_g \alpha^{(S_{n,d})}$, where $S_{n,d}$ is the number of components in $\mathbf{p}$.

Now, consider the following definitions on polynomials.

Sum of Squares Polynomials: Let $\mathbb{S}^2[x] \subset \mathbb{R}[x]$ be the set of sum of squares (SOS) polynomials. Polynomial $s : \mathbb{R}^n \to \mathbb{R}$ is an SOS polynomial if it can be written as a sum of *finitely* many squared polynomials, i.e., $s(x) = \sum_{j=1}^{\ell} h_j(x)^2$ for some $\ell < \infty$ and $h_j \in \mathbb{R}[x]$ for $1 \le j \le \ell$, ([74, 75]).

Quadratic Module: For a given set of polynomials $\mathcal{P}_j(x) \in \mathbb{R}[x], j = 1, \ldots, \ell$, the quadratic module generated by these polynomials is denoted by $\mathcal{QM}(\mathcal{P}_1, \ldots, \mathcal{P}_\ell) \subset \mathbb{R}[x]$ and defined as ([69, 74])

$$\mathcal{QM}(\mathcal{P}_1, \ldots, \mathcal{P}_\ell) := s_0(x) + \sum_{j=1}^{\ell} s_j(x)\mathcal{P}_j, \ \{s_j\}_{j=0}^{\ell} \subset \mathbb{S}^2[x] \tag{2.1}$$

Putinar's property: A closed semialgebraic set $\mathcal{K} = \{x \in \mathbb{R}^n : \mathcal{P}_j(x) \ge 0, \ j = 1, 2, \ldots, \ell \ \}$ defined by polynomials $\mathcal{P}_j \in \mathbb{R}[x]$ satisfies *Putinar's property* [76] if there exists $\mathcal{U} \in \mathbb{R}[x]$ such that $\{x : \mathcal{U}(x) \ge 0\}$ is compact and $\mathcal{U} = s_0 + \sum_{j=1}^{\ell} s_j \mathcal{P}_j$ for some SOS polynomials $\{s_j\}_{j=0}^{\ell} \subset \mathbb{S}^2[x]$ – see [76, 69, 77]. Putinar's property holds if the level set $\{x : \mathcal{P}_j(x) \ge 0\}$ is compact for some j, or if all $\mathcal{P}_j$ are affine and $\mathcal{K}$ is compact - see [77]. Putinar's property is not a geometric property of the semi-algebraic set $\mathcal{K}$, but rather an algebraic property related to the representation of the set by its defining polynomials. Hence, if there exits $M > 0$ such that the polynomial $\mathcal{P}_{\ell+1}(x) := M - \|x\|^2 \ge 0$ for all $x \in \mathcal{K}$, then the *new representation* of the set $\mathcal{K} = \{x \in \mathbb{R}^n : \mathcal{P}_j(x) \ge 0, \ j = 1, 2, \ldots, \ell + 1 \ \}$ satisfies Putinar's property, [66].

Orthogonal Polynomials: A set of polynomials are orthogonal if inner product of any two different polynomials is zero, i.e., $\int \mathcal{P}_n(x)\mathcal{P}_m(x)dx = 0, \quad m \ne n$. For example, univariate Chebyshev polynomials, defined as $\mathcal{P}_0(x) = 1, \mathcal{P}_1(x) = x, \mathcal{P}_{n+1}(x) = 2x\mathcal{P}_n(x) - \mathcal{P}_{n-1}(x)$ are orthogonal.

2.0.2 Measures and Moments

Let $\mathcal{M}(\chi)$ be the space of finite Borel measures and $\mathcal{M}_+(\chi)$ be the cone of finite nonnegative Borel measures μ such that $supp(\mu) \subset \chi$, where $supp(\mu)$ denotes the support of the measure μ; i.e., the smallest closed set that contains all measurable sets with strictly positive μ measure. Also, let $C \subset \mathbb{R}^n$, $\Sigma(C)$ denotes the Borel σ-algebra over C. Given two measures μ_1 and μ_2 on a Borel σ-algebra Σ, the notation

$\mu_1 \preccurlyeq \mu_2$ means $\mu_1(S) \leq \mu_2(S)$ for any set $S \in \Sigma$. Moreover, if μ_1 and μ_2 are both measures on Borel σ-algebras Σ_1 and Σ_2, respectively, then $\mu = \mu_1 \times \mu_2$ denotes the product measure satisfying $\mu(S_1 \times S_2) = \mu_1(S_1)\mu_2(S_2)$ for any measurable sets $S_1 \in \Sigma_1$, $S_2 \in \Sigma_2$ [67].

Let $\mathbb{R}^{\mathbb{N}}$ denote the vector space of real sequences. Given $\mathbf{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n} \subset \mathbb{R}^{\mathbb{N}}$, let $L_{\mathbf{y}} : \mathbb{R}[x] \to \mathbb{R}$ be a linear map defined as ([68, 69])

$$\mathcal{P} \quad \mapsto \quad L_{\mathbf{y}}(\mathcal{P}) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha y_\alpha, \quad \text{where} \quad \mathcal{P}(x) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha x^\alpha \tag{2.2}$$

A sequence $\mathbf{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}}$ is said to have a *representing measure*, if there exists a finite Borel measure μ on $\mathbb{R}^n$ such that $y_\alpha = \int x^\alpha d\mu$ for every $\alpha \in \mathbb{N}^n$ – see ([68, 69]). In this case, $\mathbf{y}$ is called the moment sequence of the measure μ.

Given two square symmetric matrices A and B, the notation $A \succcurlyeq 0$ denotes that A is positive semidefinite, and $A \succcurlyeq B$ stands for $A - B$ being positive semidefinite.

Moment Matrix: Given $r \geq 1$ and the sequence $\{y_\alpha\}_{\alpha \in \mathbb{N}^n}$, the moment matrix $M_r(\mathbf{y}) \in \mathbb{R}^{S_{n,r} \times S_{n,r}}$, containing all the moments up to order $2r$, is a symmetric matrix and its (i,j)-th entry is defined as follows ([68, 69]):

$$M_r(\mathbf{y})(i,j) := L_{\mathbf{y}}\left(x^{\alpha^{(i)}+\alpha^{(j)}}\right) = y_{\alpha^{(i)}+\alpha^{(j)}} \tag{2.3}$$

where $1 \leq i, j \leq S_{n,r}$, $\mathbb{N}_r^{\mathrm{n}} \ni \mathbf{0} = \alpha^{(1)} <_g \ldots <_g \alpha^{(S_{n,2r})}$ and $S_{n,2r}$ is the number of moments in $\mathbb{R}^n$ up to order $2r$. Let $\mathcal{B}_r^T = \left[x^{\alpha^{(1)}}, \ldots, x^{\alpha^{(S_{n,r})}}\right]^T$ denote the vector comprised of the monomial basis of $\mathbb{R}_{\mathrm{r}}[x]$. Note that the moment matrix can be written as $M_r(\mathbf{y}) = L_{\mathbf{y}}\left(\mathcal{B}_r \mathcal{B}_r^T\right)$; here, the linear map $L_{\mathbf{y}}$ operates componentwise on the matrix of polynomials, $\mathcal{B}_r \mathcal{B}_r^T$. For instance, let $r = 2$ and $n = 2$; the moment matrix containing moments up to order $2r$ is given as

$$M_2(\mathbf{y}) = \left[\begin{array}{c|cc|ccc} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ \hline y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ \hline y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{array}\right] \tag{2.4}$$

Localizing Matrix: Given a polynomial $\mathcal{P} \in \mathbb{R}[x]$, let $\mathbf{p} = \{p_\gamma\}_{\gamma\in\mathbb{N}^n}$ be its coefficient sequence in standard monomial basis, i.e., $\mathcal{P}(x) = \sum_{\alpha\in\mathbb{N}^n} p_\alpha x^\alpha$, the (i,j)-th entry of the *localizing matrix* $M_r(\mathbf{y};\mathcal{P}) \in \mathbb{R}^{S_{n,r}\times S_{n,r}}$ with respect to $\mathbf{y}$ and $\mathbf{p}$ is defined as follows ([68, 69]):

$$M_r(\mathbf{y};\mathcal{P})(i,j) := L_{\mathbf{y}}\left(\mathcal{P}x^{\alpha^{(i)}+\alpha^{(j)}}\right) = \sum_{\gamma\in\mathbb{N}^n} p_\gamma y_{\gamma+\alpha^{(i)}+\alpha^{(j)}} \tag{2.5}$$

where, $1 \le i,j \le S_{n,d}$. Equivalently, $M_r(\mathbf{y},\mathcal{P}) = L_{\mathbf{y}}\left(\mathcal{P}\mathcal{B}_r\mathcal{B}_r^T\right)$, where $L_{\mathbf{y}}$ operates componentwise on $\mathcal{P}\mathcal{B}_r\mathcal{B}_r^T$. For example, given $\mathbf{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^2}$ and the coefficient sequence $\mathbf{p} = \{p_\alpha\}_{\alpha\in\mathbb{N}^2}$ corresponding to polynomial $\mathcal{P}$,

$$\mathcal{P}(x_1,x_2) = bx_1 - cx_2^2, \tag{2.6}$$

the localizing matrix for $r = 1$ is formed as follows

$$M_1(\mathbf{y};\mathcal{P}) = \begin{bmatrix} by_{10} - cy_{02} & by_{20} - cy_{12} & by_{11} - cy_{03} \\ by_{20} - cy_{12} & by_{30} - cy_{22} & by_{21} - cy_{13} \\ by_{11} - cy_{03} & by_{21} - cy_{13} & by_{12} - cy_{04} \end{bmatrix} \tag{2.7}$$

Orthogonal Basis: One can represent the moment and localization matrices in terms of given orthogonal basis. Let $\{b_i\}_{i\in\mathbb{N}}$ be an orthogonal basis of *univariate* polynomials on $[-1,1]$, i.e., $\int_{[-1,1]} b_i(t)b_j(t)\ dt = 0$ for all $i \neq j$. Without loss of generality, suppose that the degree of b_i is equal to i for all $i \in \mathbb{N}$. Given $n \ge 1$, for all $\alpha \in \mathbb{N}^n$, define $b_\alpha : \mathbb{R}^n \to \mathbb{R}$ such that $b_\alpha(x) := \prod_{i=1}^n b_{\alpha_i}(x_i)$, where α_i and x_i are the i-th components of $\alpha \in \mathbb{N}^n$ and $x \in \mathbb{R}^n$, respectively. Clearly $\{b_\alpha : \alpha \in \mathbb{N}^{\mathrm{n}}_d\}$ is an orthogonal basis of multivariate polynomials on $[-1,1]^n$ with degree at most d, i.e., $\int_{[-1,1]^n} b_{\alpha^{(i)}}(x)\ b_{\alpha^{(j)}}(x)\ dx = 0$ for all $1 \leq i \neq j \leq S_{n,d}$. Let $\mathcal{B}^o_d$ denote the vector of polynomials in $\mathbb{R}_{\mathrm{d}}[x]$ defined as $\mathcal{B}^{0^T}_d = \left[b_{\alpha^{(1)}}(x),\ b_{\alpha^{(2)}}(x),\ \ldots, b_{\alpha^{(S_{n,d})}}(x)\right]$; and $T_d \in \mathbb{R}^{S_{n,d}\times S_{n,d}}$ denote the one-to-one correspondence such that $\mathcal{B}^o_d = T_d\mathcal{B}_d$. Moreover, for a given sequence $\mathbf{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$, let $L^o_{\mathbf{y}} : \mathbb{R}[x] \to \mathbb{R}$ be a linear map defined as

$$\mathcal{P} \quad \mapsto \quad L^o_{\mathbf{y}}(\mathcal{P}) = \sum_{\alpha\in\mathbb{N}^n} p^o_\alpha y_\alpha, \quad \text{where} \quad \mathcal{P}(x) = \sum_{\alpha\in\mathbb{N}^n} p^o_\alpha b_\alpha(x). \tag{2.8}$$

Given $y \in \mathbb{R}^{S_{n,2d}}$ such that $y^T = \left[y_{\alpha^{(1)}}, \ldots, y_{\alpha^{(S_{n,2d})}}\right]^T$, define its extension $\mathbf{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$ such that $y_\alpha = 0$ for all $\alpha \in \mathbb{N}^n$ with $\|\alpha\|_1 > 2d$. For $\bar{y} := T^{-1}_{2d}y$, define its extension $\bar{\mathbf{y}}$ similarly. Then for all $\mathcal{P} \in \mathbb{R}_d[x]$, we have $L^o_{\mathbf{y}}(\mathcal{P}) = L_{\bar{\mathbf{y}}}(\mathcal{P})$. In the rest of the chapter, we abuse the notation and write $\bar{\mathbf{y}} = T^{-1}_{2d}\mathbf{y}$. Then the moment matrix operator, $M^o_d(\mathbf{y})$, for the given orthogonal basis is defined as

$$M^o_d(\mathbf{y}) := L^o_{\mathbf{y}}\left(\mathcal{B}^o_d\ \mathcal{B}^{oT}_d\right) = L_{T^{-1}_{2d}\mathbf{y}}\left(T_d\mathcal{B}_d\ {\mathcal{B}_d}^T T^T_d\right) = T_d M_d\left(T^{-1}_{2d}\mathbf{y}\right) T^T_d. \tag{2.9}$$

For example for $d = 2$ and $n = 2$, the moment matrix under the orthogonal basis formed by Chebyshev polynomials of the first kind can be written as follows

$$M^o_2(\mathbf{y}) = \begin{bmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{10} & \frac{y_{00}+y_{20}}{2} & y_{11} & \frac{y_{10}+y_{30}}{2} & \frac{y_{01}+y_{21}}{2} & y_{12} \\ y_{01} & y_{11} & \frac{y_{00}+y_{02}}{2} & y_{21} & \frac{y_{10}+y_{12}}{2} & \frac{y_{01}+y_{03}}{2} \\ y_{20} & \frac{y_{10}+y_{30}}{2} & y_{21} & \frac{y_{00}+y_{40}}{2} & \frac{y_{11}+y_{31}}{2} & y_{22} \\ y_{11} & \frac{y_{01}+y_{21}}{2} & \frac{y_{10}+y_{12}}{2} & \frac{y_{11}+y_{31}}{2} & \frac{y_{00}+y_{20}+y_{02}+y_{22}}{4} & \frac{y_{11}+y_{13}}{2} \\ y_{02} & y_{12} & \frac{y_{01}+y_{03}}{2} & y_{22} & \frac{y_{11}+y_{13}}{2} & \frac{y_{00}+y_{04}}{2} \end{bmatrix}. \tag{2.10}$$

Let $\mathcal{P} \in \mathbb{R}[x]$ be a given polynomial with degree δ, and $\mathbf{p} = \{p_\alpha\}_{\alpha\in\mathbb{N}^n}$ denote its coefficient sequence with respect to the standard monomial basis, i.e., $\mathcal{P}(x) = \sum_{\alpha\in\mathbb{N}^n} p_\alpha x^\alpha$. For a given orthogonal basis, the localization matrix operator is defined as

$$M_d^o(\mathbf{y};\mathbf{p}) := L_{\mathbf{y}}^o\left(\mathcal{P}\mathcal{B}_d^o\ \mathcal{B}_d^{oT}\right) = L_{T_{2d+\delta}^{-1}\mathbf{y}}\left(T_d\mathcal{P}\mathcal{B}_d\ \mathcal{B}_d^T T_d^T\right) = T_d M_d\left(T_{2d+\delta}^{-1}\mathbf{y};\mathbf{p}\right) T_d^T. \tag{2.11}$$

Let $r := \lceil\frac{\delta}{2}\rceil$. It is important to note that since T_{2d} is invertible, $\{\mathbf{y}:\ M_d^o(\mathbf{y}) \succeq 0,\ M_{d-r}^o(\mathbf{y};\mathbf{p}) \succeq 0\}$ and $\{\mathbf{y}:\ M_d(\mathbf{y}) \succeq 0,\ M_{d-r}(\mathbf{y};\mathbf{p}) \succeq 0\}$ are *isomorphic.*

2.0.3 Preliminary Results on Measures and Polynomials

In this section, we state some standard results found in the literature that will be referred to later in this thesis.

Moment Condition: The following lemmas give necessary, and sufficient conditions for sequence of moments $\mathbf{y}$ to have a representing measure μ – for details see [67, 78, 69].

Lemma 1. *Let μ be a finite Borel measure on $\mathbb{R}^n$, and $\mathbf{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$ such that $y_\alpha = \int x^\alpha d\mu$ for all $\alpha \in \mathbb{N}^n$. Then $M_d(\mathbf{y}) \succcurlyeq 0$ for all $d \in \mathbb{N}$.*

Lemma 2. *Let $\mathbf{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$ be a real sequence. If $M_d(\mathbf{y}) \succcurlyeq 0$ for some $d \geq 1$, then*

$$|y_\alpha| \leq \max\left\{y_0,\ \max_{i=1,\ldots,n} L_{\mathbf{y}}\left(x_i^{2d}\right)\right\} \quad \forall\alpha \in \mathbb{N}_{2d}^n.$$

Lemma 3. *If there exist a constant $c > 0$ such that $M_d(\mathbf{y}) \succcurlyeq 0$ and $|y_\alpha| \leq c$ for all $d \in \mathbb{N}$ and $\alpha \in \mathbb{N}^n$, then there exists a representing measure μ with support on $[-1, 1]^n$.*

Lemma 4. *Let μ be a Borel probability measure supported on the hyper-cube $[-1, 1]^n$. Its moment sequence $\mathbf{y} \in \mathbb{R}^{\mathbb{N}}$ satisfies $\|\mathbf{y}\|_\infty \leq 1$.*

Proof. Since *supp*(μ) $\subset [-1, 1]^n$ and μ is a probability measure, we have $|y_\alpha| \leq \int |x^\alpha| d\mu \leq \int |x| d\mu \leq 1$ for each $\alpha \in \mathbb{N}^n$. Hence, $\|\mathbf{y}\|_\infty \leq 1$. □

Given polynomials $\mathcal{P}_j \in \mathbb{R}[x]$, let $\mathbf{p}_j$ be its coefficient sequence in standard monomial basis for $j = 1, 2, \ldots, \ell$; consider the semialgebraic set $\mathcal{K}$ defined as

$$\mathcal{K} = \{x \in \mathbb{R}^n : \mathcal{P}_j(x) \geq 0,\ j = 1, 2, \ldots, \ell\ \}. \tag{2.12}$$

The following lemma gives a necessary and sufficient condition for $\mathbf{y}$ to have a representing measure μ supported on $\mathcal{K}$ – see [67, 78, 69, 68].

Lemma 5. *If $\mathcal{K}$ defined in* (2.12) *satisfies Putinar's property, then the sequence* $\mathbf{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$ *has a* representing *finite Borel measure μ on the set $\mathcal{K}$, if and only if*

$$M_d(\mathbf{y}) \succcurlyeq 0, \quad M_d(\mathbf{y}; \boldsymbol{p}_j) \succcurlyeq 0, \quad j = 1, \ldots, \ell, \text{ for all } d \in \mathbb{N}.$$

Measure of Compact Set: The following lemma, proven in [67], shows that the Borel measure of a compact set is equal to the optimal value of an infinite dimensional LP problem.

Lemma 6. *Let Σ be the Borel σ-algebra on $\mathbb{R}^n$, and μ_1 be a measure on a compact set $\mathcal{B} \subset \Sigma$. Then for any given $\mathcal{K} \in \Sigma$ such that $\mathcal{K} \subseteq \mathcal{B}$, one has*

$$\mu_1(\mathcal{K}) = \int_{\mathcal{K}} d\mu_1 = \sup_{\mu_2 \in \mathcal{M}(\mathcal{K})} \left\{ \int d\mu_2 : \mu_2 \preccurlyeq \mu_1 \right\},$$

where $\mathcal{M}(\mathcal{K})$ is the set of finite Borel measures on $\mathcal{K}$.

SOS Representation: The following lemma gives a sufficient condition for $f \in \mathbb{R}[x]$ to be nonnegative on the set $\mathcal{K}$– see [79, 78, 68, 69].

Lemma 7. *Assume $\mathcal{K}$ defined in* (2.12) *satisfies Putinar's property. If $\mathcal{P} \in \mathbb{R}[x]$ is strictly positive on $\mathcal{K}$, then $\mathcal{P} \in \mathcal{QM}(\{\mathcal{P}_j\}_{j=1}^{\ell})$. Hence,*

$$\mathcal{P} = s_0 + \sum_{j=1}^{\ell} s_j \mathcal{P}_j, \; s_j \in \mathbb{S}^2[x], \; j = 0, ..., \ell$$

Duality: The following theorems show the relationship between measures, continuous functions and polynomials:

i) Stone-Weierstrass Theorem: Every continuous function defined on a closed set can be uniformly approximated as closely as desired by a polynomial function, [80].

ii) Riesz Representation Theorem: Let $\mathcal{C}(\chi)$ be the Banach space of continuous functions on χ with associated norm $\|f\| := \sup_{x\in\chi} |f(x)|$ for $f \in \mathcal{C}$ and $\mathcal{C}_+(\chi) := \{f \in \mathcal{C} : f \geq 0 \text{ on } \chi\}$ be the cone of nonnegative continuous functions.

The cone of nonnegative measures is dual to the cone of nonnegative continuous functions with inner product $\langle \mu, f \rangle := \int_\chi f d\mu$, $\mu \in \mathcal{M}_+(\chi)$, $f \in \mathcal{C}_+(\chi)$; i.e., any $\mu \in \mathcal{M}_+(\chi)$ belongs to the space of all linear functional on $\mathcal{C}_+(\chi)$ - see ([81], Section 21.5, [79, 82]).

2.0.4 Linear and Semidefinite Programming

In this section preliminary results on linear program and semidefinite programs are presented.

Consider the linear programming (LP) problem in standard form

$$\mathbf{P}^* := \max\langle x, c \rangle \tag{2.13}$$

$$\text{s.t.} \quad Ax \leq b \tag{2.13a}$$

$$x \geq 0. \tag{2.13b}$$

where, $x \in \mathbb{R}^n$ is variable vector, $A : \mathbb{R}^n \to \mathbb{R}^m$ is the linear operator, $b \in \mathbb{R}^m$ are real matrices and vector.Also, $\langle x, c \rangle = c^T x$. Based on standard results on LP [79, 82], the dual problem of (2.13) reads as

$$\mathbf{P}^*_{\mathbf{Dual}} := \min\langle b, y \rangle \tag{2.14}$$

$$\text{s.t.} \quad A^* y \geq c \tag{2.14a}$$

$$y \geq 0. \tag{2.14b}$$

where, $A^* : \mathbb{R}^m \to \mathbb{R}^n$ denotes the adjoint operator of A, i.e., $\langle A^* y, x \rangle = \langle y, Ax \rangle$. The following theorem shows the relationship of primal and dual problems.

Theorem 8. ***Strong Duality:*** *If in problem* (2.13), $\langle x, c \rangle$ *is finite value and the set* $\{(Ax, \langle x, c \rangle) : x \geq 0)\}$ *is closed, then there is no duality gap between* (2.13) *and* (2.14), *i.e.,* $\mathbf{P}^* = \mathbf{P}^*_{\mathbf{Dual}}$, *([79], Theorem 3.10, [82], Theorem 7.2)*

Consider the semidefinite programming (SDP) problem in standard form

$$\mathbf{P}^* := \min\langle C, X \rangle \tag{2.15}$$

$$\text{s.t.} \quad \langle A_i, X \rangle = b_i, \; i = 1, ..., m \tag{2.15a}$$

$$X \succcurlyeq 0. \tag{2.15b}$$

where, $A_i, C \in \mathbb{R}^n \times \mathbb{R}^n$, vector $b \in \mathbb{R}^m$, and $X \in \mathbb{R}^n \times \mathbb{R}^n$, $\langle C, X \rangle = trace(CX)$. Based on standard results on SDP [83, 84], the dual problem of (2.15) reads as

$$\mathbf{P}^*_{\mathbf{Dual}} := \max b^T y \tag{2.16}$$

$$\text{s.t.} \quad C - \sum_{i=1}^{m} A_i y_i \succcurlyeq 0 \tag{2.16a}$$

Example: Consider following SDP for $n = 2$,$m = 2$ and given matrices:

$$A_1 = \begin{bmatrix} 1 & 3 \\ 0 & 5 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 2 \\ 7 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 4 \\ 1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \tag{2.17}$$

The variable matrix is symmetric matrix as

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \tag{2.18}$$

Hence, the standard SDP reads as

$$\min 3x_{11} + 5x_{12} + x_{22} \tag{2.19}$$

$$\text{s.t.} \quad x_{11} + 3x_{12} + 5x_{22} = 2 \tag{2.19a}$$

$$x_{11} + 9x_{12} + 4x_{22} = 1 \tag{2.19b}$$

$$X \succcurlyeq 0. \tag{2.19c}$$

The dual problem is as

$$\max 2y_1 + 1y_2 \tag{2.20}$$

$$\text{s.t.} \begin{bmatrix} 3 & 4 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 3 \\ 0 & 5 \end{bmatrix} y_1 - \begin{bmatrix} 1 & 2 \\ 7 & 4 \end{bmatrix} y_2 \succcurlyeq 0 \tag{2.20a}$$

The following theorem shows the relationship of primal and dual problems.

Theorem 9. ***Slater's sufficient condition:*** *if the feasible set of strictly positive matrices in constraint of primal SDP is nonempty, then there is no duality gap* *between* (2.15) *and* (2.16), *i.e.,* $\mathbf{P}^* = \mathbf{P}^*_{\mathbf{Dual}}$, *([83, 84])* .

Chapter 3

Chance Constrained Optimization

3.1 Introduction

In this chapter, we aim at solving *chance optimization problems;* i.e., problems which involve maximization of the probability of a semialgebraic set defined by polynomial inequalities [65, 66]. More precisely, given a probability space $\left(\mathbb{R}^m, \bar{\Sigma}_q, \bar{\mu}_q\right)$ with $\bar{\Sigma}_q$ denoting the Borel σ-algebra of $\mathbb{R}^m$ and $\bar{\mu}_q$ denoting a finite (nonnegative) Borel measure on $\bar{\Sigma}_q$, we focus on the problem given in (3.1) over decision variable $x \in \mathbb{R}^n$.

$$\mathbf{P}^* := \sup_{x\in\mathbb{R}^n} \; \bar{\mu}_q \left(\bigcup_{k=1,\ldots,N} \bigcap_{j=1,\ldots,\ell_k} \left\{ q \in \mathbb{R}^m : \; \mathcal{P}_j^{(k)}(x,q) \geq 0 \right\} \right), \tag{3.1}$$

where $\mathcal{P}_j^{(k)} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $j = 1, 2, \ldots, \ell_k$ and $k = 1, \ldots, N$ are given polynomials. Let $\mathcal{K}_k := \left\{ (x,q) : \mathcal{P}_j^{(k)}(x,q) \geq 0, \; j = 1, \ldots, \ell_k \right\}$ and $\mathcal{K} := \bigcup_{k=1}^{N} \mathcal{K}_k$. Under the assumption that $\mathcal{K}$ is bounded, we show that by solving a sequence of semidefine programming (SDP) problems of growing dimension, we can construct a sequence $\{\mathbf{y}_{\mathbf{x}}^d\}_{d\in\mathbb{Z}_+} \subset \mathbb{R}^{\mathbb{N}}$ that has an accumulation point in the weak-$\star$ topology of ℓ_∞, and for every accumulation point $\mathbf{y}_{\mathbf{x}}^* \in \mathbb{R}^{\mathbb{N}}$, there is a representing finite (positive) Borel measure μ_x^* such that any $x^* \in supp(\mu_{\mathbf{x}}^*)$ is an optimal solution to (3.1), i.e., the supremum $\mathbf{P}^*$ is attained at x^*, where $\mathbb{R}^{\mathbb{N}}$ denotes the vector space of real sequences. Note that the problem of interest in (3.1), when reformulated in *hypograph* form, can be equivalently written as a chance constrained optimization problem: $\sup_{x\in R^n, \gamma\in\mathbb{R}} \left\{ \gamma : \; \bar{\mu}_q \left(\bigcup_{k=1,\ldots,N} \bigcap_{j=1,\ldots,\ell_k} \left\{ q \in \mathbb{R}^m : \; \mathcal{P}_j^{(k)}(x,q) \geq 0 \right\} \right) \geq \gamma \right\}$.

First, the emphasis will be placed on the following special case of (3.1), where $N = 1$,

$$\mathbf{P}^* := \sup_{x \in \mathbb{R}^n} \bar{\mu}_q\Big(\{q \in \mathbb{R}^m : \ \mathcal{P}_j(x,q) \geq 0, \quad j = 1,\ldots,\ell\}\Big), \tag{3.2}$$

and then all the results derived for the special case (3.2) will be extended to the case where $N > 1$.

Although, in some particular cases, the problem in (3.1) is convex (e.g., see [27, 28]), in general, chance constrained problems are not convex; e.g., see [27] for non-convex chance constrained linear programs. In this chapter, we use results on moments of measures (e.g., see [68, 69, 74]) to develop a sequence of SDP problems, known as Lasserre's hierarchy [69], whose solutions converge to the solution of (3.1).

3.2 Chance Optimization over a Semialgebraic Set

In this section we focus on the *chance optimization* problem stated in (3.2). We first provide an equivalent problem over finite (positive) Borel measures, and then we consider its relaxations in the moment space. Given polynomials $\mathcal{P}_j : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ with degree δ_j for $j = 1,\ldots,\ell$, we define

$$\mathcal{K} = \{(x,q) \in \mathbb{R}^n \times \mathbb{R}^m : \ \mathcal{P}_j(x,q) \geq 0, \ j = 1,2,\ldots,\ell\}. \tag{3.3}$$

Assumption 1. *$\mathcal{K}$ satisfies Putinar's property.*

Remark 3.2.1. *Assumption 1 implies that $\mathcal{K}$ is a compact set; hence the projections of $\mathcal{K}$ onto x-coordinates and onto q-coordinates, i.e., $\Pi_1 =: \{x \in \mathbb{R}^n : \exists q \in \mathbb{R}^m \text{ s.t. } (x,q) \in \mathcal{K}\}$ and $\Pi_2 =: \{q \in \mathbb{R}^m : \ \exists x \in \mathbb{R}^n \text{ s.t. } (x,q) \in \mathcal{K}\}$, are also compact. Therefore, after rescaling of polynomials, we assume without loss of generality that $\Pi_1 \subset \chi := [-1,1]^n$ and $\Pi_2 \subset \mathcal{Q} := [-1,1]^m$ and also the set $(\chi \times \mathcal{Q}) \backslash \mathcal{K} = \{(x,q) \in \chi \times \mathcal{Q} : (x,q) \notin \mathcal{K}\}$ has a nonempty interior. Furthermore, instead of working on the original probability space $(\mathbb{R}^m, \bar{\Sigma}_q, \bar{\mu}_q)$, we can adopt a smaller probability space $(\mathcal{Q}, \Sigma_q, \mu_q)$, where $\Sigma_q := \{S \cap \mathcal{Q} : \ S \in \bar{\Sigma}_q\}$ and $\mu_q(S) := \frac{\bar{\mu}_q(S)}{\bar{\mu}_q(\mathcal{Q})}$ for all $S \in \Sigma_q$. Therefore, we can take for granted that $\mu_q \in \mathcal{M}(\mathcal{Q})$, where $\mathcal{M}(\mathcal{Q})$ is the set of finite Borel measures μ_q such that $supp(\mu_q) \subset \mathcal{Q}$. We also assume that moments of any order of μ_q can be computed.*

3.2.1 An Equivalent Problem

As an intermediate step in the development of convex relaxations of the original problem, a related infinite dimensional problem in the measure space is provided below:

$$\mathbf{P}^*_{\mu_\mathbf{q}} := \sup_{\mu,\mu_x} \int d\mu, \tag{3.4}$$

$$\text{s.t.} \quad \mu \preccurlyeq \mu_x \times \mu_q, \tag{3.4a}$$

$$\mu_x \text{ is a probability measure}, \tag{3.4b}$$

$$\mu_x \in \mathcal{M}(\chi), \quad \mu \in \mathcal{M}(\mathcal{K}). \tag{3.4c}$$

Theorem 10. *The optimization problems in* (3.2) *and* (3.4) *are equivalent in the following sense:*

i) The optimal values are the same, i.e., $\mathbf{P}^* = \mathbf{P}^*_{\mu_\mathbf{q}}$.

ii) If an optimal solution to (3.4) *exists, call it* μ_x^**, then any* $x^* \in supp(\mu_x^*)$ *is an optimal solution to* (3.2).

iii) If an optimal solution to (3.2) *exists, call it* x^**, then* $\mu_x = \delta_{x^*}$*, Dirac measure at* x^**, and* $\mu = \delta_{x^*} \times \mu_q$ *is an optimal solution to* (3.4).

Proof. See Appendix A. □

As an example, consider the following chance constrained problem corresponding to the semialgebraic set $\mathcal{K}$, displayed in Fig 3.1.a, in the space of $(x, q) \in \mathbb{R} \times \mathbb{R}$. Our objective is to compute an optimal decision x^* that attains $\mathbf{P}^* = \sup_{x \in [-1,1]} \mu_q(\mathcal{F}(x))$, in presence of random variable q with known probability measure μ_q supported on $[-1, 1]$. In other words, x^* should be chosen such that the probability of the random point (x^*, q) belonging to $\mathcal{K}$ becomes maximum. Fig 3.1.b shows the problem in the measure space, where a probability measure μ_x is assigned to decision variable x. If $x \in [-1, 1]$ is chosen randomly according to fixed μ_x, then to calculate the probability of the random event $(x, q) \in \mathcal{K}$, one should compute an integral with respect to measure $\mu_x \times \mu_q$ over the set $\mathcal{K}$ as in (3.31) – see (Fig 3.1). This integral is equal to the volume of a measure which is supported on $\mathcal{K}$ and has the same distribution as $\mu_x \times \mu_q$ on $\mathcal{K}$ – see (Fig 3.1.d). Hence, for fixed μ_x, one needs to

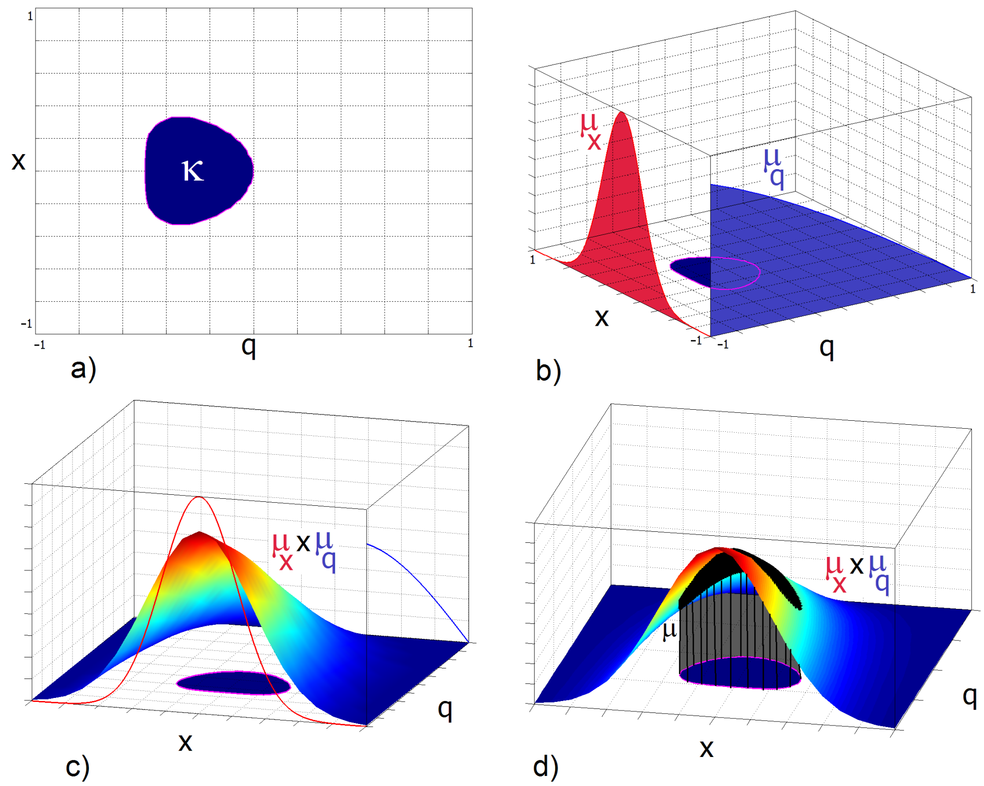

Figure 3.1: **a)** Simple chance optimization problem over semialgebraic set $\mathcal{K}$ with random parameter q, and decision variable x, **b)** Equivalent problem in the measure space over probability measure μ_x as variable for given probability measure μ_q, **c)** Probability of given semi algebraic set $\mathcal{K}$ for a fixed μ_x is equal to the integral of $\mathcal{K}$ with respect to the measure $\mu_x \times \mu_q$, **d)** The probability is equal to the volume of the measure μ which is supported on the set $\mathcal{K}$ and has the same distribution as the measure $\mu_x \times \mu_q$ over its support

look for the measure μ supported on $\mathcal{K}$ with maximum volume, and bounded above with measure $\mu_x \times \mu_q$. Therefore, searching for μ_x and μ simultaneously leads to the optimization problem (3.4) in the measure space.

3.2.2 Semidefinite Relaxations

In this section, we provide an infinite dimensional SDP of which feasible region is defined over real sequences in $\mathbb{R}^{\mathbb{N}}$. Unlike the problem (3.4) in which we are looking for a measure, in the SDP formulation given in (3.5), we aim at finding a sequence of moments corresponding to a measure that is optimal to (3.4). After proving the equivalence of (3.4) and (3.5), we next provide a sequence of finite dimensional SDPs and show that the corresponding sequence of optimal solutions can arbitrarily approximate the optimal solution of (3.5), which characterizes the optimal solution of (3.4).

Consider the following infinite dimensional SDP:

$$\mathbf{P}^*_{\mathbf{y_q}} := \sup_{\mathbf{y}, \mathbf{y}_x \in \mathbb{R}^{\mathbb{N}}} (\mathbf{y})_0, \tag{3.5}$$

$$\text{s.t.} \quad M_\infty(\mathbf{y}) \succcurlyeq 0,\ M_\infty(\mathbf{y}; \mathbf{p}_j) \succcurlyeq 0, \quad j = 1, \ldots, \ell, \tag{3.5a}$$

$$M_\infty(\mathbf{y_x}) \succcurlyeq 0,\ \|\mathbf{y_x}\|_\infty \leq 1,\ (\mathbf{y_x})_0 = 1, \tag{3.5b}$$

$$M_\infty(\mathbf{A}\mathbf{y_x} - \mathbf{y}) \succcurlyeq 0, \tag{3.5c}$$

where $\mathbf{A} : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ is a linear map depending only on μ_q. Indeed, let $\mathbf{y_q} := \{y_{q_\beta}\}_{\beta \in \mathbb{N}^m}$ be the moment sequence of μ_q. Then for any given $\mathbf{y_x} = \{y_{x_\alpha}\}_{\alpha \in \mathbb{N}^n}$, $\mathbf{A}\mathbf{y_x}$ is equal to $\mathbf{y}$ such that $y_\theta = y_{q_\beta} y_{x_\alpha}$ for all $\theta = (\beta, \alpha) \in \mathbb{N}^m \times \mathbb{N}^n$. Given $\mathbf{y} \in \mathbb{R}^{\mathbb{N}}$, $M_\infty(\mathbf{y}) \succcurlyeq 0$ means that $M_d(\mathbf{y}) \succcurlyeq 0$ for all $d \in \mathbb{Z}_+$.

The following lemma establishes the equivalence of (3.4) and (3.5).

Lemma 11. *Suppose that $\mathcal{K}$ satisfies Assumption 1. If an optimal solution to* (3.4) *exists, call it* $(\mu^*,\ \mu^*_x)$, *then their moment sequences* $(\mathbf{y}^*, \mathbf{y}^*_x)$ *is an optimal solution to* (3.5). *Conversely, if an optimal solution to* (3.5) *exists, call it* $(\mathbf{y}^*, \mathbf{y}^*_x)$, *then there exists* representing *measures* μ^* *and* μ^*_x *such that* $(\mu^*,\ \mu^*_x)$ *is optimal to* (3.4). *Moreover, the optimal values of* (3.4) *and* (3.5) *are the same, i.e.,* $\mathbf{P}^*_{\mu_\mathbf{q}} = \mathbf{P}^*_{\mathbf{y_q}}$.

Proof. See Appendix B. □

In order to have tractable approximations to the infinite dimensional SDP in (3.5), we consider the following sequence of SDPs, known as Lasserre's hierarchy [69], defined below:

$$\mathbf{P}_d := \sup_{\mathbf{y} \in \mathbb{R}^{S_{n+m,2d}},\ \mathbf{y_x} \in \mathbb{R}^{S_{n,2d}}} (\mathbf{y})_0, \tag{3.6}$$

$$\text{s.t.} \quad M_d(\mathbf{y}) \succcurlyeq 0,\ M_{d-r_j}(\mathbf{y}; \mathbf{p}_j) \succcurlyeq 0, \quad j = 1, \ldots, \ell, \tag{3.6a}$$

$$M_d(\mathbf{y_x}) \succcurlyeq 0,\ \|\mathbf{y_x}\|_\infty \leq 1,\ (\mathbf{y_x})_0 = 1, \tag{3.6b}$$

$$M_d(A_d\mathbf{y_x} - \mathbf{y}) \succcurlyeq 0, \tag{3.6c}$$

where δ_j is the degree of $\mathcal{P}_j$, $r_j := \left\lceil \frac{\delta_j}{2} \right\rceil$ for all $1 \leq j \leq \ell$, and $A_d : \mathbb{R}^{S_{n,2d}} \to \mathbb{R}^{S_{n+m,2d}}$ is defined similarly to $\mathbf{A}$ in (3.5). Indeed, let $\mathbf{y_q} := \{y_{q_\beta}\}_{\beta \in \mathbb{N}^m_{2d}}$ be the truncated moment sequence of μ_q. Then for any given $\mathbf{y_x} = \{y_{x_\alpha}\}_{\alpha \in \mathbb{N}^n_{2d}}$, $\mathbf{y} = A_d \mathbf{y_x}$ such that $y_\theta = y_{q_\beta} y_{x_\alpha}$ for all $\theta = (\beta, \alpha) \in \mathbb{N}^{n+m}_{2d}$.

In the following theorem, it is shown that the sequence of optimal solutions to the SDPs in (3.6) converges to the solution of the infinite dimensional SDP in (3.5). In essence, the following theorem is similar to Theorem 3.2 in [67]; however, for the sake of completeness we give its proof below.

Theorem 12. *For all $d \geq 1$, there exists an optimal solution $(\mathbf{y}^d, \mathbf{y}^d_\mathbf{x}) \in \mathbb{R}^{S_{n+m,2d}} \times \mathbb{R}^{S_{n,2d}}$ to* (3.6) *with the optimal value $\mathbf{P}_d$. Let $\mathcal{S} := \{(\mathbf{y}^d, \mathbf{y}^d_\mathbf{x})\}_{d \in \mathbb{Z}_+} \subset \mathbb{R}^\mathbb{N} \times \mathbb{R}^\mathbb{N}$ be such that each element of $\mathcal{S}$ is obtained by zero-padding, i.e., $\left(\mathbf{y}^d\right)_\alpha = 0$ for all $\alpha \in \mathbb{N}^{n+m}$ such that $\|\alpha\|_1 > 2d$, and $\left(\mathbf{y}^d_\mathbf{x}\right)_\alpha = 0$ for all $\alpha \in \mathbb{N}^n$ such that $\|\alpha\|_1 > 2d$. Then $\{\mathbf{P}_d\}_{d \in \mathbb{Z}_+}$ and $\mathcal{S}$ have the following properties:*

i) $\lim_{d \in \mathbb{Z}_+} \mathbf{P}_d = \mathbf{P}^*$*, the optimal value of* (3.2)*,*

ii) There exists an accumulation point of $\mathcal{S}$ in the weak-$\star$ topology of ℓ_∞ and every accumulation point of $\mathcal{S}$ is an optimal solution to (3.5)*. Hence, there exists corresponding representing measures (μ^*, μ^*_x) that is optimal to* (3.4) *and any $x^* \in supp(\mu^*_x)$ is optimal to* (3.2)*.*

Proof. See Appendix C. □

3.2.3 Discussion on Improving Estimates of Probability

In our numerical experiments, we have observed that the convergence of the upper bound $\mathbf{P}_d$ to the optimum probability $\mathbf{P}^*$ was slow in d when we solved the sequence of SDP relaxations in (3.6). Suppose that the semi-algebraic set $\mathcal{K} := \{(x,q) : \mathcal{P}_j(x,q) \geq 0,\ j = 1, \ldots, \ell\}$ satisfies Putinar's property. The procedure detailed below helped us to get better estimates on the optimum probability $\mathbf{P}^*$. To make the upcoming discussion easier we make the following assumptions: i) there is a *unique* $x^* \in \mathbf{relint}\, \Pi_1$ such that $\mu_q(\mathcal{F}(x^*)) = \mathbf{P}^*$, where $\mathcal{F}$ is defined in (3.29), and $\Pi_1 := \{x \in \mathbb{R}^n :\ \exists q \in \mathbb{R}^m \text{ s.t. } (x,q) \in \mathcal{K}\} \subset \chi := [-1,1]^n$; and ii) $\mu_q \in \mathcal{M}(\mathcal{Q})$ has the following "*continuity*" property: if $\{S_k\} \subset \Sigma_q$ such that $\lim_{k\to\infty} S_k = S^*$ in the Hausdorff-metric, then $\lim_{k\to\infty} \mu_q(S_k) = \mu_q(S^*)$. Let $(\mathbf{y}^d, \mathbf{y}^d_\mathbf{x})$ denote an optimal solution to the SDP relaxation in (3.6), and form $x^d \in \mathbb{R}^n$ using the components of $(\mathbf{y}_{\mathbf{x}}^d)_\alpha$ such that $\|\alpha\|_1 = 1$. Clearly, $x^d \in \chi$. Since $\mu_q \in \mathcal{M}(\mathcal{Q})$ is given, we approximate the volume $\int_{\mathcal{F}(x^d)} d\mu_q$ as described in [67] by solving an SDP relaxation for

$$\bar{\mathbf{P}}_d := \sup_{\mu' \in \mathcal{M}(\mathcal{F}(x^d))} \int d\mu' \text{ s.t. } \mu' \preceq \mu_q. \tag{3.7}$$

Let $\mathbf{P}'_d$ denote the optimal value of the volume approximation SDP corresponding to (3.7) with relaxation order d. Clearly, $\bar{\mathbf{P}}_d = \mu_q\left(\mathcal{F}(x^d)\right) \geq 0$, and for all d we have $\mathbf{P}_d \geq \mathbf{P}^* \geq \bar{\mathbf{P}}_d$, and $\mathbf{P}_d \geq \mathbf{P}'_d \geq \bar{\mathbf{P}}_d$. Note that since x^* is the unique optimal solution (*assumption i*), Theorem 12 implies that $\lim_{d\to\infty}(\mathbf{y}_{\mathbf{x}}^d)_\alpha = (\mathbf{y}_{\mathbf{x}}^*)_\alpha$ for all $\alpha \in \mathbb{N}^n$ such that $\mathbf{y}_{\mathbf{x}}^*$ is the moment sequence corresponding to Dirac measure at x^*. Therefore, from the definition of x^d, it follows that $\lim_{d\to\infty} x^d = x^*$. Also note that since $\mathcal{K}$ is compact (from Putinar's property) and $\mathcal{P}_j$ is a polynomial in (x, q) for all $j = 1, \ldots, \ell$, it follows that the multifunction $\mathcal{F} : \chi \to \Sigma_q$ such that $\mathcal{F}(x) = \{q \in \mathcal{Q} : \ (x, q) \in \mathcal{K}\}$ with $\mathbf{dom}\,\mathcal{F} = \Pi_1$ is locally bounded, closed-valued, and $\lim_{d\to\infty} \mathcal{F}(x^d) = \mathcal{F}(x^*)$ in Hausdorff metric. Hence, *assumption ii* implies that $\lim_{d\to\infty} \bar{\mathbf{P}}_d = \lim_{d\to\infty} \mu_q\left(\mathcal{F}(x^d)\right) = \mathbf{P}^*$. Moreover, since $\lim_{d\to\infty} \mathbf{P}_d = \mathbf{P}^*$ (from Theorem 12), and $\mathbf{P}_d \geq \mathbf{P}'_d \geq \bar{\mathbf{P}}_d$ for all d, we can conclude that $\lim_{d\to\infty} \mathbf{P}'_d = \mathbf{P}^*$ as well.

We noticed in our numerical experiments that although $\{\mathbf{P}'_d\}_{d\in\mathbb{Z}_+}$ is closer to $\mathbf{P}^*$ when compared to $\{\mathbf{P}_d\}_{d\in\mathbb{Z}_+}$, the convergence of $\mathbf{P}'_d$ to $\mathbf{P}^*$ was still slow in practice as d increases. This phenomena may partly be explained as in [67] by considering the dual problem. Let $\mathcal{C}$ be the Banach space of continuous functions on $\mathcal{Q}$ such that $\|f\| := \sup_{q\in\mathcal{Q}} f(q)$ for $f \in \mathcal{C}$, and $\mathcal{C}_+ := \{f \in \mathcal{C} : \ f \geq 0 \text{ on } \mathcal{Q}\}$. The Lagrangian dual of (3.7) is given below:

$$\begin{aligned} \bar{\mathbf{P}}_d^{\mathbf{Dual}} := \inf_{f\in\mathcal{C}_+} \int f \ d\mu_q&, \\ \text{s.t.} \quad f \geq 1 \text{ on } \mathcal{F}(x^d)&. \end{aligned} \tag{3.8}$$

Moreover, *assumption ii* ("continuity" of μ_q) and Urysohn's Lemma together imply that $\bar{\mathbf{P}}_d^{\mathbf{Dual}} = \bar{\mathbf{P}}_d$ for all d. Indeed, solving the SDP relaxation of (3.7) corresponds to approximating the indicator function of the semi-algebraic set $\mathcal{F}(x^d)$ in dual space, which is *discontinuous* on the boundary of the set. Therefore, although there exists a minimizing sequence of functions belonging to $\mathcal{C}_+$ that approximates the indicator function of $\mathcal{F}(x^d)$ from above, the discontinuity on the boundary $\mathcal{F}(x^d)$, causing numerical problems in $\mathbf{P}'_d$ computation, might be an important factor leading to slow convergence of $\{\mathbf{P}'_d\}_{d\in\mathbb{Z}_+}$ to $\mathbf{P}^*$.

Let $\mathcal{G}^d : \mathcal{Q} \to \mathbb{R}$ such that $\mathcal{G}^d(q) := \prod_{j=1}^{\ell} \mathcal{P}_j(x^d, q)$. To deal with the numerical problems caused by approximating the discontinuous indicator function, we propose to solve

$$\sup_{\tilde{\mu}\in\mathcal{M}(\mathcal{F}(x^d))} \int \mathcal{G}^d \, d\tilde{\mu} \text{ s.t. } \tilde{\mu} \preceq \mu_q. \tag{3.9}$$

Let μ_d^* denote the optimal solution to (3.9). Note that $\mathcal{G}^d$ is continuous on the boundary of $\mathcal{F}(x^d)$. Moreover, "continuity" of μ_q in *assumption ii* implies that $\mathcal{G}^d$ is strictly positive almost everywhere on $\mathcal{F}(x^d)$. Hence, μ_d^* is clearly also optimal to (3.7). Therefore, $\mu_d^*\left(\mathcal{F}(x^d)\right) = \mu_q(\mathcal{F}(x^d)) = \bar{\mathbf{P}}_d \to \mathbf{P}^*$ as $d \to \infty$. But most importantly, solving (3.9) corresponds to approximating the continuous function $\max(\mathcal{G}^d(q), 0)$ from above on $\mathcal{F}(x^d)$. These properties of (3.9) motivated us to numerically investigate the behaviour of $\{\tilde{\mathbf{P}}_d\}_{d\in\mathbb{Z}_+}$ sequence, where $\tilde{\mathbf{P}}_d := (\tilde{\mathbf{y}}^d)_0$ and $\tilde{\mathbf{y}}^d$ denotes an optimal solution to the SDP relaxation for (3.9) with order d. In our numerical experiments we observed that $\tilde{\mathbf{P}}_d \to \mathbf{P}^*$; however, this time with a faster convergence rate. To illustrate this behavior numerically, we considered two simple example problems in Section 3.2.4.

3.2.4 Simple Examples

In this section, we present two simple example problems that illustrate the effectiveness of the proposed methodology to solve the chance optimization problem in (3.2). The decision variables and the uncertain problem parameters in these examples are low dimensional for illustrative purposes. In the first example, we considered a problem over a semialgebraic set defined by a single polynomial:

$$\sup_{x\in\mathbb{R}} \mu_q\left(\{q \in \mathbb{R} : \ \mathcal{P}(x,q) \geq 0 \ \}\right), \tag{3.10}$$

where

$$\mathcal{P}\left(x,q\right) = 0.5 \ q\left(q^2 + (x-0.5)^2\right) - \left(q^4 + q^2(x-0.5)^2 + (x-0.5)^4\right). \tag{3.11}$$

The uncertain parameter $q \in \mathbb{R}$ has a uniform distribution on [-1,1]. To obtain an approximate solution, we solve the SDP in (3.6) with the minimum relaxation order $d = 2$ since the degree of the polynomial in (3.11) is 4. The moment vectors $\mathbf{y_q}$, $\mathbf{y_x}$, and $\mathbf{y}$ for the measures μ_q and μ_x, and μ up to order four are

$$\mathbf{y_q}^T = \left[1,\ 0,\ \tfrac{1}{3},\ 0,\ \tfrac{1}{5}\right], \quad \mathbf{y_x}^T = [1,\ y_{x_1},\ y_{x_2},\ y_{x_3},\ y_{x_4}],$$

$$\mathbf{y}^T = [y_{00} \mid y_{10},\ y_{01} \mid y_{20},\ y_{11},\ y_{02} \mid y_{30},\ y_{21},\ y_{12},\ y_{03} \mid y_{40},\ y_{31},\ y_{22},\ y_{13},\ y_{04}].$$

Given moment vectors $\mathbf{y_q}$, the moment vector $\bar{\mathbf{y}}$ for the measure $\overline{\mu} = \mu_x \times \mu_q$ has the form

$$\begin{aligned}\bar{\mathbf{y}}^T &= [1|y_{x_1}, y_{q_1}|y_{x_2}, y_{x_1}y_{q_1}, y_{q_2}|y_{x_3}, y_{x_2}y_{q_1}, y_{x_1}y_{q_2}, y_{q_3}|y_{x_4}, y_{x_3}y_{q_1}, y_{x_2}y_{q_2}, y_{x_1}y_{q_3}, y_{q_4}],\\ &= \left[1 \mid y_{x_1},\ 0 \mid y_{x_2},\ 0,\ \tfrac{1}{3} \mid y_{x_3},\ 0,\ \tfrac{1}{3}y_{x_1},\ 0 \mid y_{x_4},\ 0,\ \tfrac{1}{3}y_{x_2},\ 0,\ \tfrac{1}{5}\right].\end{aligned}$$

SDP in (3.6) with $d = 2$ is solved using SeDuMi [85], which is an interior-point solver add-on for Matlab, and the following solution was obtained:

$$\mathbf{y}^{*T} = [0.66, 0.3, 0.14, 0.16, 0.07, 0.1, 0.08, 0.03, 0.05, 0.04, 0.04, 0.02, 0.02, 0.02, 0.02],$$

$$\mathbf{y_x^*}^T = [1, 0.50, 0.25, 0.13, 0.85].$$

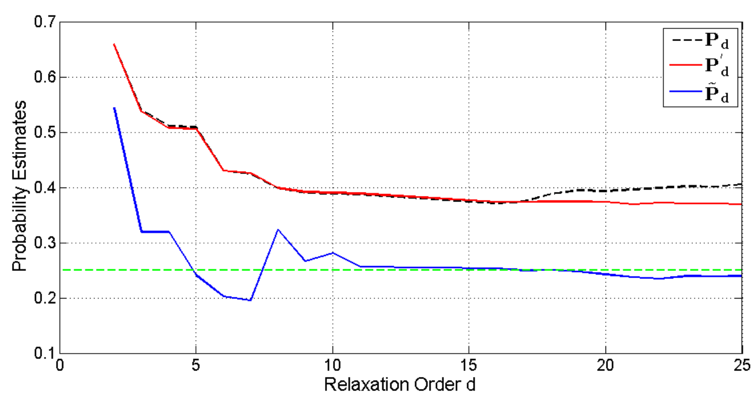

Figure 3.2: $\mathbf{P}_d$, $\mathbf{P'}_d$, and $\tilde{\mathbf{P}}_d$ for increasing relaxation order d

We approximate the solution to (3.2) with $y^*_{x_1} = 0.5$ (in Section 3.2.3 we make a case for this approximation under some simplifying assumptions), and estimate the optimal probability $\mathbf{P}^*$ with $\mathbf{P}_2 = y^*_{00} = 0.66$. To test the accuracy of the results obtained, we used Monte Carlo simulation to estimate an optimal solution to (3.10). The details of the Monte Carlo simulation are discussed in Section 3.4.3.1. This computationally intensive method estimated that $x^* = 0.5$ with optimal probability of 0.25. To obtain better estimates of the optimum probability, one needs to increase the relaxation order d. Figure 3.2 displays the three sequences defined in Section 3.2.3: $\{\mathbf{P}_d\}_{d\in\mathbb{Z}_+}$, $\{\mathbf{P}'_d\}_{d\in\mathbb{Z}_+}$, and $\{\tilde{\mathbf{P}}_d\}_{d\in\mathbb{Z}_+}$, against the optimal probability $\mathbf{P}^*$ denoted by the green dashed line.

We employed Monte Carlo simulation to compute $\mathbf{P}^*$ – see Section 3.4.3.1 for details of the simulation; and for increasing relaxation orders $d = 2, ..., 25$, we adopted SeDuMi [85] to compute $\mathbf{P}_d$ and $\mathbf{P}'_d$, the optimal values of the SDP in (3.6), and of the SDP relaxation for the volume problem in (3.7) with relaxation order d, respectively; and also to compute $\tilde{\mathbf{P}}_d = (\tilde{\mathbf{y}}^d)_0$. Similar to the results in [67], Figure 3.2 shows a faster convergence to $\mathbf{P}^*$ for the case when $\int \mathcal{G}^d\ d\tilde{\mu}$ is maximized as in (3.9).As discussed in Section 3.2.3, $\max\{\mathcal{G}^d, 0\}$ is continuous while the indicator function of $\mathcal{F}(x^d)$ is discontinuous on the boundary; and this might be a factor affecting the convergence speed. Indeed, Figure 3.3.a displays the degree-100 polynomial approximation f^* to the indicator function of the set $\mathcal{F}(x^*)$, i.e., f^* is a minimizer to $\inf_{f\in\mathbb{R}_\mathrm{d}[x]}\{\int f\ d\mu_q :\ f \geq 0 \text{ on } \mathcal{Q},\ f \geq 1 \text{ on } \mathcal{F}(x^*)\}$ for $d = 100$. Note that this problem is a restriction of the Lagrangian dual problem for $\sup\{\int\ d\mu' :\ \mu \preceq \mu_q,\ \mu' \in \mathcal{M}(\mathcal{F}(x^*))\}$ –indeed, dual variable $f \in \mathcal{C}$ is restricted to be in $\mathbb{R}_\mathrm{d}[x]$. On the other hand, Figure 3.3.b displays the degree-100 polynomial approximation h^* to the piecewise-polynomial function $\mathcal{G}(q) = \mathcal{P}(x^*, q)$, where h^* is a minimizer to $\inf_{h\in\mathbb{R}_\mathrm{d}[x]}\{\int h\ d\mu_q :\ h \geq 0 \text{ on } \mathcal{Q},\ h \geq \mathcal{G} \text{ on } \mathcal{F}(x^*)\}$ for $d = 100$. Similarly, this problem is a restriction of the Lagrangian dual problem for $\sup\{\int \mathcal{G}\ d\tilde{\mu} :\ \tilde{\mu} \preceq \mu_q,\ \tilde{\mu} \in \mathcal{M}(\mathcal{F}(x^*))\}$. Note that Figure 3.3 shows that it is easier to approximate the *continuous* function $\max\{\mathcal{G}, 0\}$ than the *discontinuous* indicator function of $\mathcal{F}(x^*)$.

Next, we considered a problem over a semialgebraic set defined by an intersection of two polynomials:

$$\sup_{x\in\mathbb{R}} \mu_q\left(\{q \in \mathbb{R} :\ \mathcal{P}_1(x, q) \geq 0,\ \mathcal{P}_2(x, q) \geq 0\ \}\right), \tag{3.12}$$

where

$$\mathcal{P}_1(x, q) = 0.1275 + 0.7x - x^2 - q^2, \quad \mathcal{P}_2(x, q) = -0.1225 + 0.7x + q - x^2 - q^2. \tag{3.13}$$

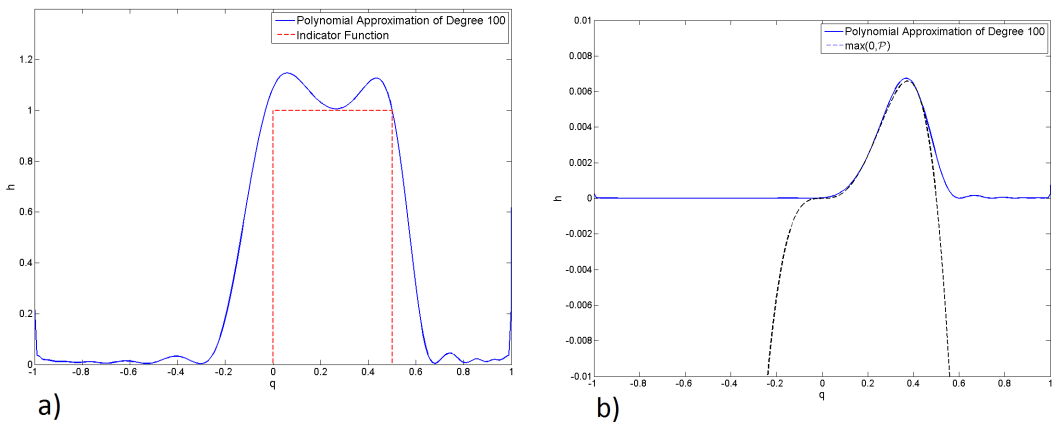

Figure 3.3: **a)** f^*: the degree-100 polynomial approximation to indicator function of $\mathcal{F}(\boldsymbol{x}^*)$, **b)** h^*: the degree-100 polynomial approximation of the piecewise-polynomial function $\max(0, \mathcal{P}(\boldsymbol{x}^*, q))$

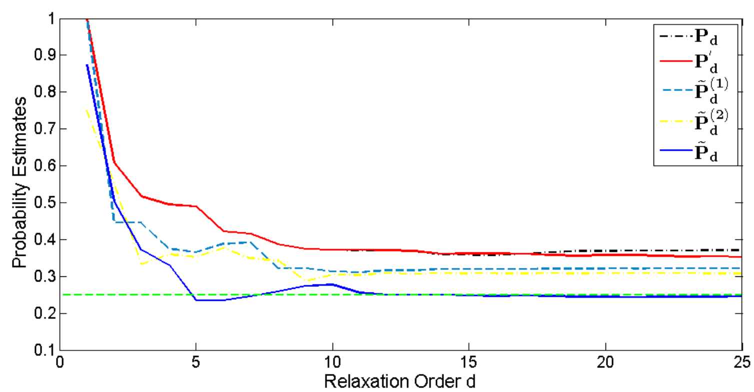

Figure 3.4: $\mathbf{P}_d$, $\mathbf{P'}_d$, $\tilde{\mathbf{P}}^{(\mathbf{1})}_d$, $\tilde{\mathbf{P}}^{(\mathbf{2})}_d$, and $\tilde{\mathbf{P}}_d$ for increasing relaxation order d

The uncertain parameter $q \in \mathbb{R}$ has a uniform distribution on [-1,1]. Figure 3.4 displays two other sequences, $\{\tilde{\mathbf{P}}^{(\mathbf{1})}_d\}_{d\in\mathbb{Z}_+}$ and $\{\tilde{\mathbf{P}}^{(\mathbf{2})}_d\}_{d\in\mathbb{Z}_+}$, in addition to the three sequences defined in Section 3.2.3: $\{\mathbf{P}_d\}_{d\in\mathbb{Z}_+}$, $\{\mathbf{P}'_d\}_{d\in\mathbb{Z}_+}$, and $\{\tilde{\mathbf{P}}_d\}_{d\in\mathbb{Z}_+}$. Here, $\tilde{\mathbf{P}}^{(\mathbf{1})}_d$ and $\tilde{\mathbf{P}}^{(\mathbf{2})}_d$ are defined similarly to $\tilde{\mathbf{P}}_d = (\tilde{\mathbf{y}}^{\mathbf{d}})_0$ by replacing $\mathcal{G}^d(q) = \mathcal{P}_1(x^d, q)\mathcal{P}_2(x^d, q)$ in (3.9) with $\mathcal{P}_1(x^d, q)$, and $\mathcal{P}_2(x^d, q)$, respectively.

3.2.5 Orthogonal Basis

In this work, all polynomials are expanded in the usual monomial basis, and the SDPs are therefore formulated as optimization problems over ordinary monomial moments.

However, one can improve the numerical performance as in [67] by employing an orthogonal basis of polynomials. Hence, one can reformulate the SDP relaxation in (3.6) using the new moment and localization matrix operators defined in (2.9) and (2.11), respectively; and the resulting problem stated in the given orthogonal basis is equivalent to (3.6). In order to illustrate the effect of orthogonal polynomial basis on the numerical behavior of the proposed method, we compared the two formulations of the simple example in (3.10): the first formulation is given in (3.6) using monomial basis, and the second formulation is obtained by replacing $M_d(.)$ and $M_{d-r_j}(.;\mathbf{p}_j)$ in (3.6) with $M_d^o(.)$ and $M_{d-r_j}^o(.;\mathbf{p}_j)$, i.e., moment and localizing matrices in Chebyshev polynomial basis representations. In order to avoid matrix inversions as in (2.9) and in (2.11), we used Chebfun package [86], which can efficiently manipulate *univariate* Chebyshev polynomials, to form $M_d^o(.)$ and $M_{d-r_j}^o(.;\mathbf{p}_j)$ that use *multivariate* Chebyshev polynomials in a *numerically stable* way; and solved the resulting SDP problems represented in the Chebyshev polynomial basis using SeDuMi. Figure 3.5 shows that the approximations to the optimal probability $\mathbf{P}^*$ converge faster when Chebyshev polynomial basis is used as opposed to the standard monomial basis as relaxation order d increases. For the problems in Chebyshev basis, the approximation $(x^o)^d$ to the optimal decision x^* is formed similarly as x^d – see Section 3.2.3. For this example x^d and $(x^o)^d$ sequences were close.

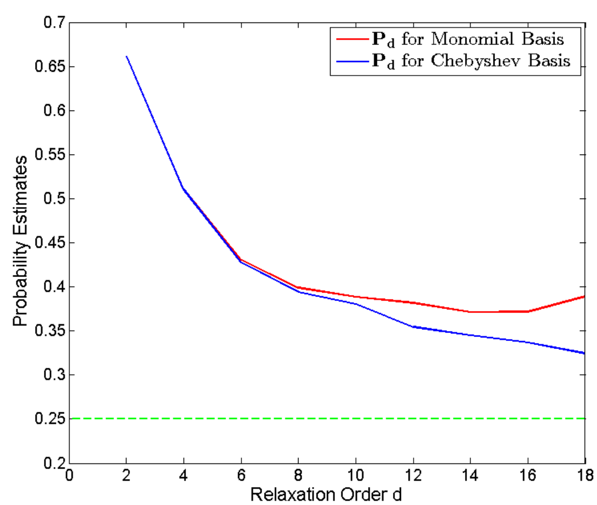

Figure 3.5: $\mathbf{P}_d$ for monomial and Chebyshev polynomial bases

3.2.6 Dual Convex Problem on Function Space

In this section, we provide an infinite LP on continuous functions which is dual to the infinite LP on measure in (3.4). To obtain a dual problem to the infinite LP in (3.4), let $\mathcal{C}(\chi \times \mathcal{Q})$ be the Banach space of continuous functions on $\chi \times \mathcal{Q}$. Then, Lagrangian dual of (3.4) is:

$$\mathbf{P}^*_{\mathbf{Dual}} := \inf_{\beta \in \mathbb{R}, \mathcal{W} \in \mathcal{C}(\chi \times \mathcal{Q})} \beta, \tag{3.14}$$

$$\text{s.t.} \quad \mathcal{W}(x,q) \geq 1 \quad \text{on} \quad \mathcal{K}, \tag{3.14a}$$

$$\beta - \int_{\mathcal{Q}} \mathcal{W}(x,q) d\mu_q \geq 0 \quad \text{on} \quad \chi, \tag{3.14b}$$

$$\mathcal{W}(x,q) \geq 0, \; \beta \geq 0. \tag{3.14c}$$

where, $\mathcal{K}$ is defined as (3.3), μ_q is a given Borel measure. We can interpret the obtained dual problem as follow. If we assume that x is given, then the optimal solution for $\mathcal{W}(x,q)$ is the indicator function of the set $\mathcal{K}$ and the optimal value $\mathbf{P}^*_{\mathbf{Dual}}$ is the volume of the set $\mathcal{K}$, i.e., $\mathbf{P}^*_{\mathbf{Dual}} = \beta = \int_{\mathcal{Q}} \mathcal{W}(x,q) d\mu_q$. Otherwise, $\int_{\mathcal{Q}} \mathcal{W}(x,q) d\mu_q$ is an upper bound for the volume of the set $\mathcal{K}$.

The following theorem establish the equivalence of problems in (3.4) and (3.14).

Theorem 13. *There is no duality gap between the infinite LP on measure in* (3.4) *and infinite LP on continuous function in* (3.14) *in the sense that the optimal values are the same, i.e.,* $\mathbf{P}^*_{\mu_\mathbf{q}} = \mathbf{P}^*_{\mathbf{Dual}}$

Proof. See Appendix D. □

To be able to obatin a tractable relaxation of infinite LP in (3.14), we use polynomial approximation of continuous function $\mathcal{W}$ and use SOS relaxation to satisfy the nonnegativity constraints, where results in following finite SDP on polynomials:

$$\mathbf{P}^*_{\mathbf{d}} := \min_{\beta \in \mathbb{R}, \mathcal{P}^d_{\mathcal{W}} \in \mathbb{R}_d[x,q]} \beta, \tag{3.15}$$

$$\text{s.t.} \quad \mathcal{P}^d_{\mathcal{W}}(x,q) - 1 \in \mathcal{QM}\left(\{\mathcal{P}_j\}_{j=1}^l\right), \tag{3.15a}$$

$$\beta - \int_{\mathcal{Q}} \mathcal{P}^d_{\mathcal{W}}(x,q) d\mu_q \in \mathcal{QM}\left(\{1 - x_i^2\}_{i=1}^n\right), \tag{3.15b}$$

$$\mathcal{P}^d_{\mathcal{W}}(x,q) \geq 0, \; \beta \geq 0. \tag{3.15c}$$

where, $\mathcal{P}_{\mathcal{W}}^{d}(x,q) \in \mathbb{R}_d[x,q]$, μ_q is a given finite Borel measure and $\mathcal{QM}$ defined in (2.1) is quadratic module generated by polynomials. According to the Lemma 7, constraints (3.15a) and (3.15b) imply that polynomials $\mathcal{P}_{\mathcal{W}}^{d}(x,q) - 1$ and $\beta - \int_{\mathcal{Q}} \mathcal{P}_{\mathcal{W}}^{d}(x,q) d\mu_q$ are positive on the sets $\mathcal{K}$ in (3.3) and χ, respectively. Problem in (3.15) is a SDP, where objective function is a linear and constraints are convex linear matrix inequalities in terms of coefficients of polynomial $\mathcal{P}_{\mathcal{W}}^{d}$.

The following theorem establish the equivalence of problems in (3.6) and (3.15).

Theorem 14. *There is no duality gap between the finite SDP on moments in* (3.6) *and finite SDP on polynomials in* (3.15) *in the sense that the optimal values are the same.*

Proof. See Appendix E. □

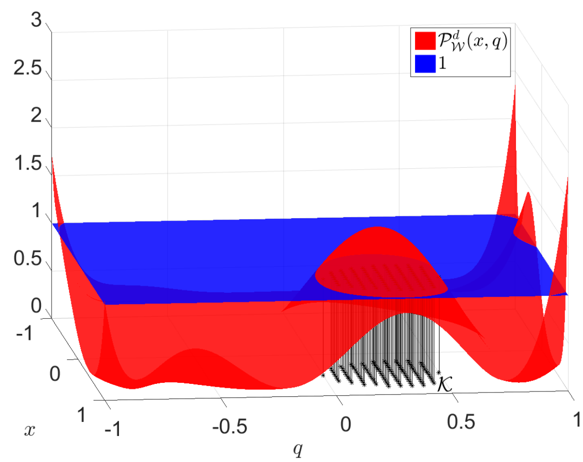

Figure 3.6: Polynomial $\mathcal{P}_{\mathcal{W}}^{d}(x,q)$ obtained by SDP (3.15) for $d = 12$

Remark 3.2.2. *In low dimensional problems, we can replace the global positivity condition in* (3.15c) *with local constraint as* $\{\mathcal{P}_{\mathcal{W}}^{d}(q,x) \geq 0 \quad on \quad \chi \times \mathcal{Q}\}$ *to improve the obtained results.*

Illustrative Example Consider the simple example (3.10) provided in section 3.2.4. Here, to obtain an approximate solution, we solve the dual problem provided in finite SDP (3.15). We solve SDP in (3.15) for polynomial order $d = 12$ by Yalmip. Figure 3.6 displays obtained $\mathcal{P}_{\mathcal{W}}^{d}(x,q)$ which is greater than 1 on the set $\mathcal{K}$ and is positive on $\chi \times \mathcal{Q} = [-1,1]^2$ as in constraint (3.15a). Figure 3.7 displays obtained β and also $\int_{\mathcal{Q}} \mathcal{P}_{\mathcal{W}}^d(x,q)d\mu_q$. As in constraint (3.15b) β is greater than $\int_{\mathcal{Q}} \mathcal{P}_{\mathcal{W}}^d(x,q)d\mu_q$ on the set χ. Based on obtained β and $\mathcal{P}_{\mathcal{W}}^{10}(x,q)$, we approximate the solution to the volume optimization problem with $x = 0.5$ that maximizes polynomial $\int_{\mathcal{Q}} \mathcal{P}_{\mathcal{W}}^d(x,q)d\mu_q$ on the χ and estimate the optimal volume $\mathbf{P_d^*}$ with $\mathbf{P_d} = \beta = 0.51$. Based on the Theorem 14, the obtained solution by solving dual SDP in (3.15) matches the solution obtained by SDP in (3.6).

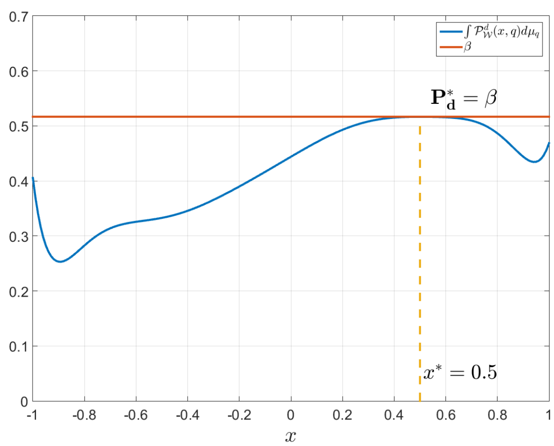

Figure 3.7: β and $\int_{\mathcal{Q}} \mathcal{P}_{\mathcal{W}}^d(x,q)d\mu_q$ obtained by SDP (3.15) for $d = 10$

3.3 Chance Optimization over a Union of Sets

We now focus on the more general setting of the chance optimization problem in (3.1). Given polynomials $\mathcal{P}_j^k : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ with degree $\delta_j^{(k)}$ for $j = 1, \dots, \ell_k$ and $k = 1, \dots, N$, the semi-algebraic set of interest is $\mathcal{K} = \cup_{k=1}^N \mathcal{K}_k$, where

$$\mathcal{K}_k = \left\{ (x,q) \in \mathbb{R}^n \times \mathbb{R}^m : \mathcal{P}_j^{(k)}(x,q) \geq 0,\ j = 1, \dots, \ell_k \right\}, \quad k = 1, \dots, N. \quad (3.16)$$

Similar to the previous section, we need *Putinar's property* to hold for $\mathcal{K}_k$ for all $k = 1, \dots, N$. With the following assumption, we can ensure this.

Assumption 2. $\mathcal{K} = \cup_{k=1}^N \mathcal{K}_k$ *is bounded, where* $\mathcal{K}_k$ *is defined in* (3.16).

Hence, as discussed in Remark 3.2.1, we can assume without loss of generality that $\mathcal{K} \subseteq \chi \times \mathcal{Q}$ and the probability measure $\mu_q \in \mathcal{M}(\mathcal{Q})$, where $\chi = [-1,1]^n$ and

$\mathcal{Q} = [-1, 1]^m$. Therefore, for all $(x, q) \in \mathcal{K}$, we have $\|x\|_2^2 + \|q\|_2^2 \leq m + n$. Define $\mathcal{P}_0^{(k)}(x, q) := m + n - \sum_{i=1}^n x_i^2 - \sum_{i=1}^m q_i^2$ for all $k = 1, \ldots, N$. $\mathcal{K}_k$ can be represented as $\mathcal{K}_k = \left\{(x, q) : \mathcal{P}_j^{(k)}(x, q) \geq 0,\ j = 0, \ldots, \ell_k\right\}$ –note that index j starts from 0. Since polynomials are continuous in (x, q), the new representation of $\mathcal{K}_k$ satisfies *Putinar's property* for each k and we still have $\mathcal{K} = \cup_{k=1}^N \mathcal{K}_k$.

The objective of this section is to provide a sequence of SDP relaxations to the chance optimization problem in (3.1) with $N > 1$, and show that the results presented in the previous sections can be easily extended for this case. More precisely, we start by providing an equivalent problem in the measure space and then develop relaxations based on moments of measures.

3.3.1 An Equivalent Problem

As an intermediate step in the development of convex relaxations of (3.1), an equivalent problem in the measure space is provided below.

$$\mathbf{P}^*_{\mu_\mathbf{q}} := \sup_{\mu_k,\ \mu_x} \sum_{k=1}^N \int d\mu_k, \tag{3.17}$$

$$\text{s.t.} \quad \sum_{k=1}^N \mu_k \preccurlyeq \mu_x \times \mu_q, \tag{3.17a}$$

$$\mu_x \text{ is a probability measure,} \tag{3.17b}$$

$$\mu_x \in \mathcal{M}(\chi), \quad \mu_k \in \mathcal{M}(\mathcal{K}_k) \quad k = 1, \ldots, N. \tag{3.17c}$$

This problem is equivalent to the problem addressed in this work in the following sense.

Theorem 15. *The optimization problems in* (3.1) *and* (3.17) *are equivalent in the following sense:*

i) The optimal values are the same, i.e. $\mathbf{P}^* = \mathbf{P}^*_{\mu_\mathbf{q}}$.

ii) If an optimal solution to (3.17) *exists, call it* μ_x^**, then any* $x^* \in supp(\mu_x^*)$ *is an optimal solution to* (3.1).

iii) If an optimal solution to (3.1) *exists, call it* x^**, then Dirac measure at* x^**,* $\mu_x = \delta_{x^*}$ *and* $\mu = \delta_{x^*} \times \mu_q$ *is an optimal solution to* (3.17).

Proof. See Appendix F. □

3.3.2 Semidefinite Relaxations

In this section, a sequence of semidefinite programs is provided which can arbitrarily approximate the optimal solution of (3.17). As before, this is done by considering moments of measures instead of the measures themselves. Define the following optimization problem indexed by the relaxation order d.

$$\mathbf{P_d} := \sup_{\mathbf{y}_k \in \mathbb{R}^{S_{n+m,2d}},\ \mathbf{y_x} \in \mathbb{R}^{S_{n,2d}}} \sum_{k=1}^{N} (\mathbf{y}_k)_0 \,, \tag{3.18}$$

$$\text{s.t.} \quad M_d(\mathbf{y}_k) \succcurlyeq 0,\ M_{d-r_j^{(k)}}\left(\mathbf{y}_k; \mathbf{p}_j^{(k)}\right) \succcurlyeq 0, \quad j = 1, \ldots, l_k, \quad k = 1, \ldots, N \tag{3.18a}$$

$$M_d(\mathbf{y_x}) \succcurlyeq 0,\ \|\mathbf{y_x}\|_\infty \leq 1,\ (\mathbf{y_x})_0 = 1, \tag{3.18b}$$

$$M_d\left(A_d \mathbf{y_x} - \sum_{k=1}^{N} \mathbf{y}_k\right) \succcurlyeq 0, \tag{3.18c}$$

where $\delta_j^{(k)}$ is the degree of $\mathcal{P}_j^{(k)}$, $r_j^{(k)} := \left\lceil \frac{\delta_j^{(k)}}{2} \right\rceil$ for all $1 \leq j \leq \ell_k$ and $1 \leq k \leq N$; and $A_d : \mathbb{R}^{S_{n,2d}} \to \mathbb{R}^{S_{n+m,2d}}$ is defined similarly to $\mathbf{A}$ in (3.5). Indeed, let $\mathbf{y_q} := \{y_{q_\beta}\}_{\beta \in \mathbb{N}_{2d}^m}$ be the truncated moment sequence of μ_q. Then for any given $\mathbf{y_x} = \{y_{x_\alpha}\}_{\alpha \in \mathbb{N}_{2d}^n}$, $\mathbf{y} = A_d \mathbf{y_x}$ such that $y_\theta = y_{q_\beta} y_{x_\alpha}$ for all $\theta = (\beta, \alpha) \in \mathbb{N}_{2d}^{n+m}$.

Next, we show that the sequence of optimal solutions to the SDPs in (3.18) converges to the solution of the infinite dimensional SDP in (3.17). More precisely, we have the following result.

Theorem 16. *For all $d \geq 1$, there exists an optimal solution $\left(\{\mathbf{y}_k^d\}_{k=1}^N, \mathbf{y_x}^d\right)$ to* (3.18) *with the optimal value $\mathbf{P}_d$. Moreover,*

i) $\lim_{d \in \mathbb{Z}_+} \mathbf{P}_d = \mathbf{P}^*$*, the optimal value of* (3.1).

ii) Let $\mathcal{S} := \left\{\left(\{\mathbf{y}_k^d\}_{k=1}^N, \mathbf{y_x}^d\right)\right\}_{d \in \mathbb{Z}_+}$ such that each element is obtained by zero-padding $\mathbf{y}^d$ and $\mathbf{y}_k^d$ for $1 \leq k \leq N$. There exists an accumulation point of $\mathcal{S}$ in the weak-$\star$ topology of ℓ_∞, and for every accumulation point of $\mathcal{S}$, there exists corresponding representing measures $\left(\{\mu_k^*\}_{k=1}^N, \mu_x^*\right)$ that is optimal to (3.17) *and any $x^* \in supp(\mu_x^*)$ is optimal to* (3.1).

Proof. See Appendix G. □

3.4 Implementation and Numerical Results

In previous sections, we showed that chance optimization problem in (3.1) can be relaxed to a sequence of SDPs. In this section, we go one step further to improve approximation quality of the relaxed problems in practice and implement an efficient first-order algorithm to solve the resulting SDP relaxations.

3.4.1 Regularized Chance Optimization Using Trace Norm

As shown in Theorem 10 and Theorem 15, if the chance optimization problems in (3.2) and (3.1) have unique optimal solution x^*, then the optimal distribution μ_x^* is a Dirac measure whose mass is concentrated on the single point x^*, i.e., its support is the singleton $\{x^*\}$. Such distributions, have moment matrices with rank one. To improve the solution quality of the algorithm, one can incorporate this observation in the formulation of the relaxed problem. For the sake of notational simplicity, in this section we will consider the regularized version of chance optimization problem (3.6) for presenting the algorithm:

$$\min_{\mathbf{y}\in\mathbb{R}^{S_{n+m,2d}},\ \mathbf{y_x}\in\mathbb{R}^{S_{n,2d}}} \omega_r\,\mathbf{Tr}(M_d(\mathbf{y_x}))-(\mathbf{y})_0 \quad \text{subject to} \quad (3.6a),\ (3.6b),\ (3.6c) \quad (3.19)$$

for some $\omega_r > 0$, where $\mathbf{Tr}(.)$ denotes the trace function. Our objective is to achieve the maximum probability with a low-rank moment matrix $M_d(\mathbf{y}_\mathbf{x}^*)$, hopefully with rank 1. To this end, we regularize the objective with trace norm. Since $M_d(\mathbf{y}_\mathbf{x}^*) \succcurlyeq 0$, $\mathbf{Tr}(M_d(\mathbf{y}_\mathbf{x}^*))$ is equal to sum of singular values of $M_d(\mathbf{y}_\mathbf{x}^*)$, which is called the nuclear norm of $M_d(\mathbf{y}_\mathbf{x}^*)$. This is a well known approach for obtaining low-rank solutions. Indeed, the nuclear norm is the convex envelope of the rank function and, in practice, produces good results; see [87] and [88] for details.

To be able to solve the SDP in (3.19) involving large scale matrices in practice, one need to implement an efficient convex optimization algorithm. Recently, a first-order augmented Lagrangian algorithm ALCC has been proposed in [89] to deal with regularized conic convex problems. We will adapt this algorithm to solve SDPs of the form in (3.19). In the following section, we briefly discuss the algorithm ALCC.

3.4.2 First-Order Augmented Lagrangian Algorithm

Consider the optimization problem:

$$(P):\ p^* = \min\{\rho(x) + \gamma(x) : \quad A(x) - b \in \mathcal{C}\}, \tag{3.20}$$

where $\gamma : \mathbb{R}^n \to \mathbb{R}$ is a convex function such that $\nabla\gamma$ is Lipschitz continuous with constant L_γ, $\rho : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a closed convex function such that $\Delta := \mathbf{dom}(\rho)$ is convex compact set, $A : \mathbb{R}^n \to \mathbb{R}^m$ is a *linear map*, and $\mathcal{C} \subset \mathbb{R}^m$ is a closed convex cone. Let $\mathcal{C}^* := \{\theta \in \mathbb{R}^n : \langle z, \theta\rangle \geq 0,\ \forall z \in C\}$ denote the dual cone of $\mathcal{C}$, and $B > 0$ denote the diameter of Δ, i.e., $B = \max\{\|x - y\|_2 :\ x, y \in \Delta\}$; and we assume that B is given. Given a penalty parameter $\nu > 0$ and Lagrangian dual multiplier $\theta \in \mathcal{C}^*$, the augmented Lagrangian for (P) in (3.20) is given by

$$\mathcal{L}(x; \nu, \theta) := \tfrac{1}{\nu}\left(\rho(x) + \gamma(x)\right) + \tfrac{1}{2} d_{\mathcal{C}}(A(x) - b - \theta)^2, \tag{3.21}$$

where $d_{\mathcal{C}} : \mathbb{R}^m \to \mathbb{R}$ denotes the distance function to cone $\mathcal{C}$, i.e., $d_{\mathcal{C}}(\bar{z}) := \|\bar{z} - \Pi_{\mathcal{C}}(\bar{z})\|_2$, and $\Pi_{\mathcal{C}}(\bar{z}) := \text{argmin}\{\|z - \bar{z}\|_2 :\ z \in \mathcal{C}\}$ denotes the Euclidean projection of $\bar{z}$ onto $\mathcal{C}$. Given $\nu_k > 0$ and $\theta_k \in \mathcal{C}^*$, we define $\mathcal{L}_k(x) := \mathcal{L}(x; \nu_k, \theta_k)$ and $\mathcal{L}_k^* := \min_x \mathcal{L}_k(x)$. Let $f_k : \mathbb{R}^n \to \mathbb{R}$ such that $f_k(x) := \frac{1}{\nu_k}\gamma(x) + \frac{1}{2} d_{\mathcal{C}}(A(x) - b - \theta)^2$; hence, $\mathcal{L}_k^* = \min_x \frac{1}{\nu_k}\rho(x) + f_k(x)$. It is important to note that f_k is a convex function with Lipschitz continuous gradient $\nabla f_k(x) = \frac{1}{\nu_k}\gamma(x) - A^*\left(\Pi_{\mathcal{C}^*}(\theta_k + b - A(x))\right)$; and the Lipschitz constant of ∇f_k is equal to $L_k := \frac{1}{\nu_k} L_\gamma + \sigma_{\max}^2(A)$, where $A^* : \mathbb{R}^m \to \mathbb{R}^n$ denotes the adjoint operator of $A : \mathbb{R}^n \to \mathbb{R}^m$, and $\sigma_{\max}(A)$ denotes the maximum singular value of the linear map A. Therefore, given $\epsilon_k > 0$, an ϵ_k-optimal solution, $\tilde{x}_k$, to $\mathcal{L}_k^* := \min_x \mathcal{L}_k(x)$ can be efficiently computed such that $\mathcal{L}_k(\tilde{x}_k) - \mathcal{L}_k^* \leq \epsilon_k$ using an Accelerated Proximal Gradient (APG) algorithm [90, 91, 92, 93] within $\ell_k^{\max}(\epsilon_k) := B\sqrt{\frac{2L_k}{\epsilon_k}}$ APG iterations. In each APG iteration, ∇f_k, $\Pi_{\mathcal{C}^*}$ and proximal map of ρ are all evaluated *once*.

ALCC algorithm proposed in [89] can generate a minimizing sequence $\{x_k\}$ to (P) in (3.20) by *inexactly* solving a sequence of subproblems $\min_x \mathcal{L}_k(x)$. In particular, given inexact computation parameters $\alpha_k > 0$ and $\eta_k > 0$, x_k is computed such that either one of the following conditions holds:

$$\mathcal{L}_k(x_k) - \mathcal{L}_k^* \leq \tfrac{\alpha_k}{\nu_k}, \tag{3.22}$$
$$\exists s_k \in \partial\mathcal{L}_k(x_k) \quad \text{such that} \quad \|s_k\|_2 \leq \tfrac{\eta_k}{\nu_k}, \tag{3.23}$$

where $\partial\mathcal{L}_k(x_k)$ denotes the subdifferential of $\mathcal{L}_k$ at x_k. Then dual Lagrangian multiplier is updated: $\theta_{k+1} = \frac{\nu_k}{\nu_{k+1}}\Pi_{C^*}(\theta_k + b - A(x_k))$. For given $c, \beta > 1$, fix the parameter sequence as follows: $\nu_k = \beta^k\nu_0$, $\alpha_k = \frac{1}{k^{2(1+c)}\beta^k}\alpha_0$, and $\eta_k = \frac{1}{k^{2(1+c)}\beta^k}\eta_0$ for all $k \geq 1$; and let $\{x_k, \theta_k\} \subset \Delta \times C^*$ be the primal-dual ALCC iterate sequence. **Theorem 3.10** in [89] shows that $\lim_k \theta_k\nu_k$ exists and it is an optimal solution to the dual problem. Moreover, **Theorem 3.8** shows that for all $\epsilon > 0$, x_k is ϵ-feasible, i.e., $d_{\mathcal{C}}(Ax_k - b) \leq \epsilon$, and ϵ-optimal, i.e., $|\rho(x_k) + \gamma(x_k) - p^*| \leq \epsilon$ within $\log(1/\epsilon)$ ALCC iterations, i.e., $k = \mathcal{O}(\log(1/\epsilon))$, which requires $\mathcal{O}(\epsilon^{-1}\log(\epsilon^{-1}))$ APG iterations in total. Moreover, every limit point of $\{x_k\}$ is optimal (when $A \in \mathbb{R}^{m\times n}$ is surjective, the techniques used for proving **Theorem 4** in [94] can be used to improve the rate result to $\mathcal{O}(1/\epsilon)$).

Now consider the following problem $p^* = \min_{x\in\Delta}\{\gamma(x) : \ A(x) - b \in \mathcal{C}\}$, where $\Delta \subset \mathbb{R}^n$ is a compact convex set. Note that this problem can be written as a special case of (3.20) by setting $\rho(x) = \mathbf{1}_\Delta(x)$, the indicator function of the set Δ. In Figure 3.8, we present the ALCC algorithm customized to solve $p^* = \min_{x\in\Delta}\{\gamma(x) : A(x) - b \in \mathcal{C}\}$. Note that Step 11 and 12 in Figure 3.8 are the bottleneck steps (one $\nabla\gamma$ evaluation and two projections: one onto $\mathcal{C}^*$, and one onto Δ) – in Step 11 ∇f_k is evaluated at $x_\ell^{(2)}$, and then in Step 12 $x_\ell^{(1)}$ is computed via a projected gradient step of length $1/L_k$.

In this customized version, ALCC iterate x_k is set to $x_\ell^{(1)}$ whenever either $\ell > \ell_k^{\max}$ or $\|x_\ell^{(1)} - x_\ell^{(2)}\|_2 \leq \frac{\eta_k}{\nu_k}$. Note that $\ell_k^{\max} := k^{1+c}\beta^k B\sqrt{\frac{2\nu_0 L_k}{\alpha_0}}$, which is equal to $\ell_k^{\max}(\epsilon_k)$ when $\epsilon_k = \frac{\alpha_k}{\nu_k}$. Therefore, if $\ell > \ell_k^{\max}$, then $\mathcal{L}_k(x_k) - \mathcal{L}_k^* \leq \frac{\alpha_k}{\nu_k}$ – this follows from the complexity of Accelerated Proximal Gradient algorithm (lines 9-19 in Figure 5.1) running on $\min\mathcal{L}_k(x)$; next we'll show that if $\|x_\ell^{(1)} - x_\ell^{(2)}\|_2 \leq \frac{1}{2L_k}\frac{\eta_k}{\nu_k}$, then (3.23) holds. For $\rho(x) = \mathbf{1}_\Delta(x)$, we have $\mathcal{L}_k(x) = \rho(x) + f_k(x)$. Suppose that for some ℓ, $\|x_\ell^{(1)} - x_\ell^{(2)}\|_2 \leq \frac{1}{2L_k}\frac{\eta_k}{\nu_k}$ holds. Note that g_ℓ computed in Line 11 is equal to $\nabla f_k(x_\ell^{(2)})$; thus $x_\ell^{(1)}$ computed in Line 12 is equal to $\Pi_\Delta(x_\ell^{(2)} - \nabla f_k(x_\ell^{(2)})/L_k)$, where $L_k := \frac{1}{\nu_k}L_\gamma + \sigma_{\max}^2(A)$ is the Lipschitz constant of ∇f_k. One can easily show that

$x_\ell^{(2)} - \nabla f_k(x_\ell^{(2)})/L_k - x_\ell^{(1)} \in \partial\rho(x_\ell^{(1)})$; and since ρ is the indicator function, we also have $L_k\left(x_\ell^{(2)} - x_\ell^{(1)}\right) - \nabla f_k(x_\ell^{(2)}) \in \partial\rho(x_\ell^{(1)})$. Hence, $s_k := L_k\left(x_\ell^{(2)} - x_\ell^{(1)}\right) + \nabla f_k(x_\ell^{(1)}) - \nabla f_k(x_\ell^{(2)}) \in \partial P_k(x_\ell^{(1)})$. Since ∇f_k is Lipschitz continuous, we have $\|\nabla f_k(x_\ell^{(1)}) - \nabla f_k(x_\ell^{(2)})\|_2 \leq L_k\|x_\ell^{(2)} - x_\ell^{(1)}\|_2$. Therefore, we have $\|s_k\|_2 \leq 2L_k\|x_\ell^{(2)} - x_\ell^{(1)}\|_2 \leq \frac{\eta_k}{\nu_k}$.

Algorithm ALCC $(x_0, \nu_0, \alpha_0, L_\gamma, B)$

$k \leftarrow 1$, $\theta_1 \leftarrow \mathbf{0}$
$\eta_0 \leftarrow 0.5\ \|\nabla\gamma(x_0) - \nu_0 A^*\left(\Pi_{C^*}\left(b - A(x_0)\right)\right)\|_2$
while $k \geq 1$ **do**
  $\ell \leftarrow 0$, $t_1 \leftarrow 1$,
  $x_0^{(1)} \leftarrow x_{k-1}$, $x_1^{(2)} \leftarrow x_{k-1}$
  $L_k \leftarrow \frac{1}{\nu_k}L_\gamma + \sigma_{\max}^2(A)$, $\ell_k^{\max} \leftarrow k^{1+c}\beta^k B\sqrt{\frac{2\nu_0 L_k}{\alpha_0}}$
  $\nu_k \leftarrow \beta^k\nu_0$, $\alpha_k \leftarrow \dfrac{1}{k^{2(1+c)}\beta^k}\alpha_0$, $\eta_k \leftarrow \dfrac{1}{k^{2(1+c)}\beta^k}\eta_0$
  STOP $\leftarrow$ **false**
  while STOP $=$ **false do**
    $\ell \leftarrow \ell + 1$
    $g_\ell \leftarrow \frac{1}{\nu_k}\nabla\gamma\left(x_\ell^{(2)}\right) - A^*\left(\Pi_{C^*}\left(\theta_k + b - A\left(x_\ell^{(2)}\right)\right)\right)$
    $x_\ell^{(1)} \leftarrow \Pi_\Delta\left(x_\ell^{(2)} - g_\ell/L_k\right)$
    if $\|x_\ell^{(1)} - x_\ell^{(2)}\|_2 \leq \dfrac{1}{2L_k}\dfrac{\eta_k}{\nu_k}$ **or** $\ell > \ell_k^{\max}$ **then**
      STOP $\leftarrow$ **true**
      $x_k \leftarrow x_\ell^{(1)}$
    end if
    $t_{\ell+1} \leftarrow \left(1 + \sqrt{1 + 4\ t_\ell^2}\right)/2$
    $x_{\ell+1}^{(2)} \leftarrow x_\ell^{(1)} + \left(\frac{t_\ell - 1}{t_{\ell+1}}\right)\left(x_\ell^{(1)} - x_{\ell-1}^{(1)}\right)$
  end while
  $\theta_{k+1} \leftarrow \frac{\nu_k}{\nu_{k+1}}\Pi_{C^*}\left(\theta_k + b - A(x_k)\right)$
end while

Figure 3.8: first-order Augmented Lagrangian algorithm for Conic Convex (ALCC) problems

Semidefinite program of (3.19) is a special case of the conic convex problem in (3.20), where $\gamma(\mathbf{y_x}, \mathbf{y}) = c_r^T\mathbf{y_x} + c_p^T\mathbf{y}$ for some $c_r \in \mathbb{R}^{S_{n,2d}}$ and $c_p \in \mathbb{R}^{S_{n+m,2d}}$ since the objective of (3.19) is linear in $(\mathbf{y}, \mathbf{y_x})$; hence, $L_\gamma = 0$, the conic constraint $A(.) - b \in C$ in (3.20) is a linear matrix inequality (LMI), with $\mathcal{C} = \mathcal{C}^*$ being the cone of positive semidefinite matrices $\mathbb{S}_+$, and the compact set $\Delta = \{(\mathbf{y}, \mathbf{y_x}) :\ \|\mathbf{y}\|_\infty \leq 1,\ \|\mathbf{y_x}\|_\infty \leq 1,\ (\mathbf{y_x})_0 = 1\}$. Hence, $\Pi_{\mathcal{C}}(.) = \Pi_{\mathcal{C}^*}(.)$ can be computed using one eigenvalue decomposition, and $\Pi_\Delta(.)$ is very efficient and can be computed in linear time. In our numerical experiments in Section 3.4.3, we used $\|x_k - x_{k-1}\|_2/(1 + \|x_{k-1}\|_2) \leq \text{tol}$ as the stopping condition for ALCC.

3.4.3 Numerical Examples

In this section, four numerical examples are presented that illustrate the performance of the proposed methodology, discussed in Sections 3.2 and 3.3, coupled with the augmented Lagrangian algorithm presented in Section 3.4.2 in finding approximate solutions to the chance constrained problems in (3.1) and (3.2) by solving their regularized semidefinite relaxations in (3.19). In all the tables, for problems of the form (3.2), i.e., $N = 1$, $\mathbf{P}_d$, $\mathbf{P'}_d$, $\bar{\mathbf{P}}_d$, and $\tilde{\mathbf{P}}_d$ denote the optimal probability estimates defined similarly as in Section 3.2.3 for x^d obtained by solving the regularized problem in (3.19); for problems of the form (3.1), i.e., $N > 1$, these estimates can be defined naturally using $(\mathbf{y}^d, \mathbf{y}^d_{\mathbf{x}})$ with $\mathbf{y}^d := \sum_{k=1}^N \mathbf{y}^d_k$; and $d \in \mathbb{Z}_+$ denotes the relaxation order. In order to compute $\mathbf{P}^*$ and $\bar{\mathbf{P}}_d$, we used Monte Carlo simulation discussed in Section 3.4.3.1. In all the tables, **iter** denotes the total number of algorithm iterations, and **cpu** denotes the computing time in *seconds* required for computing $\mathbf{P}_d$; $\mathbf{n}_{\text{var}}$ denotes the number of variables, i.e., total number of moments used. For ALCC **iter** is the total number of APG iterations, and for GloptiPoly it denotes the total number of SeDuMi [85] iterations.

3.4.3.1 Monte Carlo Simulation

To test the accuracy of the results obtained using ALCC and GloptiPoly, we used Monte Carlo integration to estimate an optimal solution and the corresponding optimal probability. Let $\mathcal{K} \subset \mathbb{R}^n \times \mathbb{R}^m$ be the given semialgebraic set such that $\Pi_1 := \{x \in \mathbb{R}^n : \exists q \in \mathbb{R}^m \text{ s.t. } (x, q) \in \mathcal{K}\} \subset \chi := [-1, 1]^n$, and $\Pi_2 := \{q \in \mathbb{R}^m : \exists q \in \mathbb{R}^m \text{ s.t. } (x, q) \in \mathcal{K}\} \subset \mathcal{Q} := [-1, 1]^m$. Define $\mathcal{F} : \chi \to \Sigma_q$,

$$\mathcal{F}(x) := \{q \in \mathcal{Q} : (x, q) \in \mathcal{K}\}. \tag{3.24}$$

First, we uniformly grid χ into $\bar{N}$ grid-points ($\bar{N}$ depending on the desired precision). Let $\{x^{(i)}\}_{i=1}^{\bar{N}} \subset \chi$ denote the points in the uniform grid. Next, for each grid point $x^{(i)}$, we sample from the distribution induced by the given finite Borel measure μ_q supported on $\mathcal{Q}$. Let $\{q^{(i,k)}\}_{k=1}^{N_i}$ be N_i i.i.d. sample of random parameter q. Then we approximate $\mu_q(\mathcal{F}(x^{(i)}))$ by

$$P_{N_i}^{(i)} := \frac{1}{N_i}\sum_{k=1}^{N_i} \mathbf{1}_{\mathcal{K}}\left(x^{(i)}, q^{(i,k)}\right), \quad \text{where} \quad \mathbf{1}_{\mathcal{K}}(x,q) = \begin{cases} 1, & \text{if } (x,q) \in \mathcal{K}; \\ 0, & \text{otherwise.} \end{cases}$$

Because of law of large numbers, $\lim_{N_i \nearrow \infty} P_{N_i}^{(i)} = \mu_q(\mathcal{F}(x^{(i)}))$. For each $x^{(i)}$, we chose sample size N_i such that $P_{N_i}^{(i)}$ becomes stagnant to further increase in N_i. Finally, we approximate x^* by $x^{(i^*)}$, where $i^* \in \operatorname{argmax}\{P_{N_i}^{(i)} : 1 \leq i \leq \bar{N}\}$. It is clear that what we used is a *naive* method, and it can be made much more efficient by using an adaptive gridding scheme on χ. On the other hand, as the dimensions n and m are very small for the problems discussed in the numerical section, this naive method served its purpose.

3.4.3.2 Example 1: A Simple Semialgebraic Set

Consider the chance optimization problem

$$\sup_{x \in \mathbb{R}^5} \mu_q\left(\{q \in \mathbb{R}^5 : \ \mathcal{P}(x,q) \geq 0 \ \}\right), \tag{3.25}$$

where

$$\begin{aligned} \mathcal{P}(x,q) = \ & 0.185 + 0.5x_1 - 0.5x_2 + x_3 - x_4 + 0.5q_1 - 0.5q_2 + q_3 - q_4 - x_1^2 - 2x_1q_1 - x_2^2 \\ & -2x_2q_2 - x_3^2 - 2x_3q_3 - x_4^2 - 2x_4q_4 - x_5^2 + 2x_5q_5 - q_1^2 - q_2^2 - q_3^2 - q_4^2 - q_5^2, \end{aligned}$$

and the uncertain parameters q_1, q_2, q_3, q_4, q_5 have a uniform distribution: $q_1 \sim U[-1,0]$, $q_2 \sim U[0,1]$, $q_3 \sim U[-0.5,1]$, $q_4 \sim U[-1,0.5]$, $q_5 \sim U[0,1]$ – $U[a,b]$ denotes the uniform distribution between a and b. The k-th moment of uniform distribution U[a,b] is $(\mathbf{y_q})_k = \frac{b^{k+1}-a^{k+1}}{(b-a)(k+1)}$. The optimum solution and corresponding optimal probability are obtained by Monte Carlo method: $x_1^* = 0.75$, $x_2^* = -0.75$, $x_3^* = 0.25$, $x_4^* = -0.25$, $x_5^* = 0.5$, and $P^* = 0.75$. To obtain an approximate solution, we solve the SDP in (3.6) using GloptiPoly and ALCC. For ALCC, we set ν_0 to 1, 5×10^{-2} and 5×10^{-3} when d is equal to 1, 2, and 3, respectively, and tol $= 1 \times 10^{-2}$. The results for relaxation order $d = 1,2,3$ are shown in Table 3.1. As in Figure 3.2, when compared to $\mathbf{P}_d$, $\tilde{\mathbf{P}}_d$ approximates $\mathbf{P}^*$ better, i.e., when $\max\{\int \mathcal{P}(x^d,q)\ d\tilde{\mu} : \ \tilde{\mu} \preceq \mu_q,\ \tilde{\mu} \in \mathcal{M}(\mathcal{F}(x^d))\}$ is solved instead of $\max\{\int d\mu' : \ \mu' \preceq \mu_q,\ \mu' \in \mathcal{M}(\mathcal{F}(x^d))\}$. We reported results up to order $d = 3$,

ALCC				GloptiPoly			
d	1	2	3	**d**	1	2	3
$\mathbf{n}_{\text{var}}$	87	1127	8463	$\mathbf{n}_{\text{var}}$	87	1127	8463
iter	169	624	1207	**iter**	18	25	41
cpu	0.9	28.1	785.9	**cpu**	0.5	12.3	15324.3
$\mathbf{x_1}$	0.742	0.745	0.757	$\mathbf{x_1}$	0.467	0.710	0.742
$\mathbf{x_2}$	-0.777	-0.701	-0.721	$\mathbf{x_2}$	-0.467	-0.710	-0.742
$\mathbf{x_3}$	0.213	0.226	0.216	$\mathbf{x_3}$	0.163	0.245	0.249
$\mathbf{x_4}$	-0.239	-0.250	0.236	$\mathbf{x_4}$	-0.163	-0.245	-0.249
$\mathbf{x_5}$	0.500	0.551	0.557	$\mathbf{x_5}$	0.319	0.475	0.495
$\mathbf{P}_d$	0.991	0.971	0.961	$\mathbf{P}_d$	1	1	1
$\mathbf{P}'_d$	1	1	1	$\mathbf{P}'_d$	1	1	1
$\tilde{\mathbf{P}}_d$	0.996	0.7739	0.6919	$\tilde{\mathbf{P}}_d$	0.9652	0.7768	0.7031
$\bar{\mathbf{P}}_d$	0.7504	0.7459	0.7459	$\bar{\mathbf{P}}_d$	0.5067	0.7484	0.7535

Table 3.1: ALCC and GloptiPoly results for Example 1

because for larger d, GloptiPoly did not terminate in 24 hours.

3.4.3.3 Example 2: Union of Simple Sets

Given the following polynomials

$$\begin{aligned}\mathcal{P}^{(1)}(x,q) =& -0.263 + 0.4x_1 - 0.4x_2 + 0.8x_3 - 0.8x_4 + 1.2x_5 + 0.1q_1 + 0.08q_2 + 0.04q_3 \\ & + 0.4q_4 + 0.6q_5 - x_1^2 - x_2^2 - x_3^2 - x_4^2 - x_5^2 - 0.5q_1^2 - 0.4q_2^2 - 0.1q_3^2 - q_4^2 - q_5^2, \\ \mathcal{P}^{(2)}(x,q) =& -2.06 + 0.4x_1 - 0.8x_2 + 3.2x_3 - 1.6x_4 + 3.6x_5 - 0.4q_1 - 0.4q_2 - 0.2q_3 \\ & - 0.2q_4 - 0.8q_5 - x_1^2 - 2x_2^2 - 4x_3^2 - 2x_4^2 - 3x_5^2 - q_1^2 - q_2^2 - q_3^2 - q_4^2 - q_5^2,\end{aligned}$$

consider the chance optimization problem

$$\sup_{\mathrm{x}\in\mathbb{R}^5} \mu_q \left(\bigcup_{j=1,2} \left\{ q \in \mathbb{R}^5 : \ \mathcal{P}^{(j)}(x,q) \geq 0 \right\} \right), \tag{3.26}$$

where $q_i \sim U[-0.5, 0.5]$ for all $i = 1, \ldots, 5$, i.e., the uncertain parameters q_i are uniformly distributed on $[-0.5, 0.5]$. The optimum solution and corresponding optimal probability are obtained by Monte Carlo method: $x_1^* = 0.2$, $x_2^* = -0.2$, $x_3^* = 0.4$, $x_4^* = -0.4$, $x_5^* = 0.6$, and $\mathbf{P}^* = 0.80$. To obtain an approximate solution, we

ALCC			
d	1	2	3
$\mathbf{n}_{\text{var}}$	153	2128	16478
iter	979	1467	1875
cpu	6.5	102.2	434.7
$\mathbf{x_1}$	0.209	0.328	0.201
$\mathbf{x_2}$	-0.202	-0.174	-0.201
$\mathbf{x_3}$	0.397	0.466	0.430
$\mathbf{x_4}$	-0.400	-0.405	-0.401
$\mathbf{x_5}$	0.667	0.638	0.591
$\mathbf{P}_d$	1	0.997	0.981
$\mathbf{P}'_d$	1	1	1
$\tilde{\mathbf{P}}_d$	0.9973	0.8610	0.8926
$\bar{\mathbf{P}}_d$	0.8937	0.8745	0.8984

Table 3.2: ALCC results for Example 2

solve the SDP in (3.18) using ALCC, where we set ν_0 to 1, 1×10^{-1} and 1×10^{-3} when d is equal to 1, 2, and 3, respectively, and tol $= 1 \times 10^{-2}$. The results for relaxation order $d = 1, 2, 3$ are shown in Table 3.2. Let $\mathcal{F}^{(k)}(x) =: \{q \in \mathcal{Q} : \ \mathcal{P}^{(k)}(x, q) \geq 0\}$ for $k = 1, 2$. The probability estimates $\tilde{\mathbf{P}}_d$ reported in Table 3.2 are computed by solving the SDP relaxation for

$$\max \int \mathcal{P}^{(1)}(x^d, q)\ d\tilde{\mu}_1 + \int \mathcal{P}^{(2)}(x^d, q)\ d\tilde{\mu}_2 :$$

$$\tilde{\mu}_1 + \tilde{\mu}_2 \preceq \mu_q,\ \ \tilde{\mu}_1 \in \mathcal{M}(\mathcal{F}^{(1)}(x^d)),\ \ \tilde{\mu}_2 \in \mathcal{M}(\mathcal{F}^{(2)}(x^d)).$$

For this example, GloptiPoly fails to extract the optimum solution.

3.4.3.4 Example 3: Portfolio Selection Problem

We aim at selecting a portfolio of financial assets to maximize the probability of achieving a return higher than a specified amount r^*. Suppose that for each asset $i = 1, ..., N$, its uncertain rate of return is a random variable $\xi_i(q)$; and let $(\mathcal{Q}, \Sigma_q, \mu_q)$ denote the underlying probability space. In this context x_i denotes the percentage of money invested in asset i. More precisely, we solve the following problem:

$$\sup_{x\in\mathbb{R}^N} \mu_q\left(\left\{q \in \mathbb{R}^N :\ \sum_{i=1}^{N} \xi_i(q)x_i \geq r^*\right\}\right) \text{ s.t. } \sum_{i=1}^{N} x_i \leq 1, \quad x_i \geq 0 \quad \forall\, i \in \{1,\ldots,N\}. \tag{3.27}$$

In our example problem, $r^* = 1.5$, $N = 4$, $\xi_1(q) = 1 + q_1$, $\xi_2(q) = 1 + q_2$, $\xi_3(q) = 0.9+q_3$, $\xi_4(q) = 0.9+q_4$, where $\{q_i\}_{i=1}^4$ are independent, and $q_1 \sim \text{Beta}(3-\sqrt{2}, 3+\sqrt{2})$, $q_2 \sim \text{Beta}(4,4)$, $q_3 \sim \text{Beta}(3+\sqrt{2}, 3-\sqrt{2})$, $q_4 \sim U[0.5,1]$. The k-th moment of Beta distribution $\text{Beta}(\alpha,\beta)$ over [0,1] is $y_k = \frac{\alpha+k-1}{(\alpha+\beta+k-1)}y_{k-1}$ and $y_0 = 1$. We will solve an equivalent problem in the form of (3.2) with $\ell = 7$, where $\mathcal{P}_j(x,q) = x_j$ for $j = 1,\ldots,4$, $\mathcal{P}_5(x,q) = 1 - \sum_{i=1}^4 x_i$, $\mathcal{P}_6(x,q) = 8 - \sum_{i=1}^4 x_i^2 - \sum_{i=1}^4 q_i^2$, and $\mathcal{P}_7(x,q) = \sum_{i=1}^4 \xi_i(q)x_i - r^*$. Since any $(x,q) \in \mathcal{K}$ satisfies $x \in \chi$ and $q \in \mathcal{Q}$, we added polynomial $\mathcal{P}_6(x,q)$ to assure that the resulting representation of the semialgebraic set $\mathcal{K}$ satisfies Putinar's property. The optimum solution and the corresponding optimal probability are computed approximately by Monte Carlo method: $x_1^* = 0$, $x_2^* = 0$, $x_3^* = 0.3$, $x_4^* = 0.7$, and $P^* = 0.89$. To obtain an approximate solution, we solve the SDP relaxation in (3.6) using GloptiPoly and ALCC. For ALCC, we set ν_0 to 1×10^{-2}, 1×10^{-2} and 1×10^{-3} when d is equal to 1, 2, and 3, respectively, and $\text{tol} = 1 \times 10^{-3}$. The results for relaxation order $d = 1,2,3$ are shown in Table 3.3. We reported results up to order $d = 3$, because for larger d, GloptiPoly did not terminate in 24 hours.

3.4.3.5 Example 4: Run time

In this example, for fixed degree of the relaxation order d, we examined how the run times of ALCC algorithm scale as the problem size increases. For this purpose, we consider the following problem: Given $n \geq 1$, we set $\mathcal{P} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\mathcal{P}(x,q) = 0.81 - \sum_{i=1}^n (x_i - q_i)^2$; and solve

$$\sup_{x\in R^n} \mu_q\left(\{q \in R^n :\ \mathcal{P}(x,q) \geq 0\ \}\right). \tag{3.28}$$

The numerical results for increasing n and fixed relaxation order $d = 1$ are displayed in Table 3.4. For each n, ALCC recovered the optimal decision value: $x^* = 0$.

ALCC			
d	1	2	3
$\mathbf{n}_{\text{var}}$	60	565	3213
iter	573	388	2227
cpu	3.625	16.426	756.798
$\mathbf{x_1}$	0.004	0.009	0.002
$\mathbf{x_2}$	0.012	0.009	0.006
$\mathbf{x_3}$	0.438	0.449	0.299
$\mathbf{x_4}$	0.5007	0.522	0.677
$\mathbf{P}_d$	0.996	0.994	0.980
$\mathbf{P}'_d$	1	1	0.9716
$\tilde{\mathbf{P}}_d$	0.7928	0.8177	0.8220
$\bar{\mathbf{P}}_d$	0.7405	0.8655	0.8422

GloptiPoly			
d	1	2	3
$\mathbf{n}_{\text{var}}$	60	565	3213
iter	15	20	48
cpu	0.509	2.617	1025.045
$\mathbf{x_1}$	0.133	0.0462	0.003
$\mathbf{x_2}$	0.192	0.154	0.075
$\mathbf{x_3}$	0.295	0.297	0.210
$\mathbf{x_4}$	0.325	0.493	0.710
$\mathbf{P}_d$	1	1	0.999
$\mathbf{P}'_d$	0.9071	0.9997	0.9896
$\tilde{\mathbf{P}}_d$	0.3808	0.7753	0.8395
$\bar{\mathbf{P}}_d$	0.3865	0.8267	0.8675

Table 3.3: ALCC and GloptiPoly results for Example 3

ALCC									
n	5	10	20	30	40	50	60	70	80
d	1	1	1	1	1	1	1	1	1
$\mathbf{n}_{\text{var}}$	10	20	40	60	80	100	120	140	160
iter	82	140	97	182	201	175	191	186	208
cpu	0.3969	1.5349	3.5542	14.2899	27.7978	37.2624	60.4454	83.3669	122.7844

Table 3.4: ALCC for increasing problem in Example 4

3.5 Conclusion

In this chapter, “chance optimization” problems are introduced, where one aims at maximizing the probability of a set defined by polynomial inequalities. These problems are, in general, nonconvex and computationally hard. A sequence of semidefinite relaxations is provided whose sequence of optimal values is shown to converge to the optimal value of the original problem. To solve the semidefinite programs of increasing size obtained by relaxing the original chance optimization problem, a first-order augmented Lagrangian algorithm is implemented which enables us to solve much larger size semidefinite programs that interior point methods can deal with. Numerical examples are provided that show that one can obtains reasonable approximations to the optimal solution and the corresponding optimal probability even for lower order relaxations. In the next chapter, we show the application of chance constrained problems in system and control. We formulate the problem of controller design for uncertain systems as a chance optimization problem and building on the result obtained in this chapter, sequence of SDP's are developed.

3.6 Appendix A: Proof of Theorem 10

Let $(\mathcal{Q}, \Sigma, \mu_q)$ be the probability space defined in Remark 3.2.1. Note that since $\mathcal{P}_j(x,q)$ is a polynomial in random vector $q \in \mathbb{R}^m$ for all $x \in \mathbb{R}^n$, it is continuous in q; hence $\mathcal{P}_j(x,.)$ is Borel measurable for all $x \in \mathbb{R}^n$ and $j = 1, \ldots, \ell$. As discussed in Remark 3.2.1, it can be assumed that $\mathcal{K} \subset \chi \times \mathcal{Q} = [-1,1]^n \times [-1,1]^m$. Define $\mathcal{F} : \mathbb{R}^n \to \Sigma$ as follows

$$\mathcal{F}(x) := \{q \in \mathbb{R}^m : \mathcal{P}_j(x,q) \geq 0,\ j = 1, 2, \ldots, \ell\}, \tag{3.29}$$

and consider the following problem over the probability measures in $\mathcal{M}(\chi)$:

$$\mathbf{P} := \sup_{\mu_x \in \mathcal{M}(\chi)} \left\{ \int_\chi \mu_q(\mathcal{F}(x))\ d\mu_x :\ \mu_x(\chi) = 1 \right\}. \tag{3.30}$$

Note that the optimal value of (3.2) can be written as $\mathbf{P}^* = \sup_{x\in\chi} \mu_q(\mathcal{F}(x))$.

Let μ_x be a feasible solution to (3.30). Since $\mu_q(\mathcal{F}(x)) \leq \mathbf{P}^*$ for all $x \in \chi$, we have $\int \mu_q(\mathcal{F}(x))\ d\mu_x \leq \mathbf{P}^*$. Thus, $\mathbf{P} \leq \mathbf{P}^*$. Conversely, let $x \in \mathbb{R}^n$ be a feasible solution to the problem in (3.2) and δ_x denote the Dirac measure at x. The objective value of x in (3.2) is equal to $\mu_q(\mathcal{F}(x))$. Moreover, $\mu_x = \delta_x$ is a feasible solution to the problem in (3.4) with objective value equal to $\mu_q(\mathcal{F}(x))$. This implies that $\mathbf{P}^* \leq \mathbf{P}$. Hence, $\mathbf{P}^* = \mathbf{P}$, and (3.30) can be rewritten as

$$\mathbf{P}^* = \sup_{\mu_x \in \mathcal{M}(\chi)} \left\{ \int_\chi \int_{\mathcal{F}(x)} d\mu_q d\mu_x :\ \mu_x(\chi) = 1 \right\} = \sup_{\mu_x \in \mathcal{M}(\chi)} \left\{ \int_{\mathcal{K}} d\mu_x \mu_q :\ \mu_x(\chi) = 1 \right\}, \tag{3.31}$$

and using the epigraph formulation shown in Lemma 6, we finally obtain

$$\mathbf{P}^* = \sup_{\mu_x \in \mathcal{M}(\chi)} \sup_{\mu \in \mathcal{M}(\mathcal{K})} \int d\mu \quad \text{s.t.} \quad \mu \preccurlyeq \mu_x \times \mu_q,\ \mu_x(\chi) = 1.$$

Therefore, $\mathbf{P}^* = \mathbf{P}^*_{\boldsymbol{\mu_q}}$.

3.7 Appendix B: Proof of Lemma 11

Suppose that (μ, μ_x) is feasible to (3.4). Let $\mathbf{y}$ and $\mathbf{y_x}$ be the moment sequences corresponding to μ and μ_x, respectively. Lemma 5 implies (3.5a); Lemma 1 and Lemma 4 imply (3.5b). Moreover, let $\bar{\mathbf{y}} = \{\bar{y}_\alpha\}_{\alpha \in \mathbb{N}^{n+m}}$ be the moment sequence corresponding to the product measure $\bar{\mu} := \mu_x \times \mu_q$. (3.4a) implies that $\bar{\mu} - \mu$ is a measure; hence, Lemma 1 implies $M_\infty(\bar{\mathbf{y}} - \mathbf{y}) \succcurlyeq 0$. Moreover, the definition of $\mathbf{A}$ implies that $\bar{\mathbf{y}} = \mathbf{A}\mathbf{y_x}$, which gives (3.5c). Since $\mathbf{y}$ is chosen to be the moment sequence of μ, we have $\int d\mu = y_{\mathbf{0}}$. This shows that for each (μ, μ_x) feasible to (3.4), one can construct a feasible solution to (3.5) with the same objective value. Therefore, $\mathbf{P}^*_{\mathbf{y_q}} \geq \mathbf{P}^*_{\mu_\mathbf{q}}$. Note that Assumption 1 is not used for this argument.

Next, suppose that $(\mathbf{y}, \mathbf{y_x})$ is a feasible solution to (3.5). Since $\mathcal{K}$ satisfies Assumption 1, (3.5a) and Lemma 5 together imply that $\mathbf{y}$ has a representing finite Borel measure μ supported on $\mathcal{K}$, i.e., $\mu \in \mathcal{M}(\mathcal{K})$. Moreover, (3.5b) and Lemma 3 together imply that $\mathbf{y_x}$ has a representing probability measure μ_x supported on hyper-cube χ, i.e., $\mu_x \in \mathcal{M}(\chi)$ such that $\mu_x(\chi) = 1$. Hence, the sequence $A\mathbf{y_x}$ has a representing measure $\bar{\mu}$ which is the product measure of μ_x and μ_q, i.e., $\bar{\mu} = \mu_x \times \mu_q$. Furthermore, since $\mathcal{K} \subset \chi \times \mathcal{Q} = [-1, 1]^{n+m}$, (3.5c) implies that $\mu \preceq \bar{\mu}$, which is (3.4a). Finally, the fact that μ is a representing measure of $\mathbf{y}$ implies that $\int d\mu = y_{\mathbf{0}}$. Therefore, $\mathbf{P}^*_{\mathbf{y_q}} \leq \mathbf{P}^*_{\mu_\mathbf{q}}$. Combining this with the above result gives us $\mathbf{P}^*_{\mathbf{y_q}} = \mathbf{P}^*_{\mu_\mathbf{q}}$.

3.8 Appendix C: Proof of Theorem 12

First, we will show that for all $d \geq 1$, the corresponding feasible region of (3.6) is bounded . Fix $d \geq 1$. Let $(\mathbf{y}, \mathbf{y_x})$ be a feasible solution to (3.6). Then from (3.6b), we have $\|\mathbf{y_x}\|_\infty \leq 1$. Since μ_q is a probability measure supported on $\mathcal{Q} = [-1, 1]^m$, Lemma 4 implies that $\|\mathbf{y_q}\|_\infty \leq 1$ as well. Moreover, the definition of A_d further implies that $\|A_d\mathbf{y_x}\|_\infty \leq 1$. Let $\bar{\mathbf{y}} := A_d\mathbf{y_x}$. It follows from (3.6c) that the diagonal elements of $M_d(\bar{\mathbf{y}} - \mathbf{y})$ are nonnegative, i.e., $(\bar{\mathbf{y}})_{2\alpha} - (\mathbf{y})_{2\alpha} \geq 0$ for all $\alpha \in \mathbb{N}^{n+m}_d$. This implies that

$$\max\left\{y_0, \max_{i=1,\ldots,n+m} L_\mathbf{y}\left(x_i^{2d}\right)\right\} \leq \max_{\alpha \in \mathbb{N}^{n+m}_d} y_{2\alpha} \leq \max_{\alpha \in \mathbb{N}^{n+m}_d} \bar{y}_{2\alpha} \leq \|\bar{\mathbf{y}}\|_\infty \leq 1, \tag{3.32}$$

where the first inequality follows from the fact that

$$\{y_0\} \cup \left\{L_{\mathbf{y}}\left(x_i^{2d}\right):\ i = 1, \ldots, n+m\right\} \subset \{y_{2\alpha}:\ \alpha \in \mathbb{N}_d^{n+m}\}.$$

From (3.6a), we have $M_d(\mathbf{y}) \succcurlyeq 0$. Hence, using Lemma 2, (3.32) implies that $|y_\alpha| \leq \|\bar{\mathbf{y}}\|_\infty \leq 1$ for all $\alpha \in \mathbb{N}_{2d}^{n+m}$. Therefore, the feasible region is bounded. Since the cone of positive semidefinite matrices is a closed set and all the mappings in (3.6) is linear, we also conclude that the feasible region is compact. Hence, there exists an optimal solution $(\mathbf{y}^d, \mathbf{y}_{\mathbf{x}}^d)$ to the problem (3.6) for all $d \geq 1$.

Fix $d \geq 1$. Clearly, for any given feasible solution $(\mathbf{y}, \mathbf{y}_{\mathbf{x}})$ to (3.5), by truncating the both sequences to vectors $\mathbf{y} \in \mathbb{R}^{S_{n+m,2d}}$ and $\mathbf{y}_{\mathbf{x}} \in \mathbb{R}^{S_{n,2d}}$, we can construct a feasible solution to (3.6) with the same objective value. Hence, it can be concluded that $\mathbf{P}_d \geq \mathbf{P}^*_{\mathbf{y_q}}$ for all $d \geq 1$. Moreover, the same argument also shows that $\mathbf{P}_d \geq \mathbf{P}_{d'}$ for all $d' \geq d$. Hence, $\{\mathbf{P}_d\}_{d \in \mathbb{Z}_+}$ is a decreasing sequence bounded below by $\mathbf{P}^*_{\mathbf{y_q}}$. Therefore, it is convergent and has a limit such that $\lim_{k \in \mathbb{Z}_+} \mathbf{P}_k \geq \mathbf{P}^*_{\mathbf{y_q}}$.

In order to collect all the optimal solutions corresponding to different d in one space, we extend $(\mathbf{y}^d, \mathbf{y}_{\mathbf{x}}^d) \in \mathbb{R}^{S_{n+m,2d}} \times \mathbb{R}^{S_{n,2d}}$ to vectors in ℓ_∞ (the Banach space of bounded sequences equipped with the sup-norm) by zero-padding, i.e., we set $(\mathbf{y}^d)_\alpha = 0$ for all $\alpha \in \mathbb{N}^{n+m}$ such that $\|\alpha\|_1 > 2d$, and $\left(\mathbf{y}_{\mathbf{x}}^d\right)_\alpha = 0$ for all $\alpha \in \mathbb{N}^n$ such that $\|\alpha\|_1 > 2d$. Note that ℓ_∞ is the dual space of ℓ_1, which is separable; hence, sequential Banach-Alaoglu theorem states that the closed unit ball of ℓ_∞, denoted by $\mathcal{B}_\infty$, is weak-$\star$ sequentially compact. Since $\{\mathbf{y}^d\}_{d \in \mathbb{Z}_+} \subset \mathcal{B}_\infty$ and $\{\mathbf{y}_{\mathbf{x}}^d\}_{d \in \mathbb{Z}_+} \subset \mathcal{B}_\infty$, there exists a subsequence $\{d_k\} \subset \mathbb{Z}_+$ such that $\{\mathbf{y}^{d_k}\}_{k \in \mathbb{Z}_+}$ and $\{\mathbf{y}_{\mathbf{x}}^{d_k}\}_{k \in \mathbb{Z}_+}$ converge weak-$\star$ to $\mathbf{y}^* \in \mathcal{B}_\infty$ and $\mathbf{y}_{\mathbf{x}}^* \in \mathcal{B}_\infty$ in the weak-$\star$ topology, respectively. Hence,

$$\lim_{k \in \mathbb{Z}_+} \left(\mathbf{y}^{d_k}\right)_\alpha = \left(\mathbf{y}^*\right)_\alpha, \quad \forall\, \alpha \in \mathbb{N}^{n+m}, \qquad \lim_{k \in \mathbb{Z}_+} \left(\mathbf{y}_{\mathbf{x}}^{d_k}\right)_\alpha = \left(\mathbf{y}_{\mathbf{x}}^*\right)_\alpha, \quad \forall\, \alpha \in \mathbb{N}^n. \tag{3.33}$$

Fix $d \geq 1$, then for all $k \in \mathbb{Z}_+$ such that $d_k \geq d$, we have

$$\begin{aligned}
&M_d(\mathbf{y}^{d_k}) \succcurlyeq 0,\ M_{d-r_j}(\mathbf{y}^{d_k}; \mathbf{p}_j) \succcurlyeq 0, \quad j = 1, \ldots, \ell,\\
&M_d(\mathbf{y}_{\mathbf{x}}^{d_k}) \succcurlyeq 0,\ \|\mathbf{y}_{\mathbf{x}}^{d_k}\|_\infty \leq 1,\ \left(\mathbf{y}_{\mathbf{x}}^{d_k}\right)_0 = 1,\\
&M_d(\mathbf{A}\mathbf{y}_{\mathbf{x}}^{d_k} - \mathbf{y}^{d_k}) \succcurlyeq 0.
\end{aligned}$$

Since $d \in \mathbb{Z}_+$ is arbitrary, by taking the limit as $k \to \infty$, we see that $(\mathbf{y}^*, \mathbf{y}_{\mathbf{x}}^*)$ satisfies all the constraints in (3.5). Therefore, $(\mathbf{y}^*)_0 \leq \mathbf{P}^*_{\mathbf{y_q}}$. On the other hand, $(\mathbf{y}^*)_0 = \lim_{k\in\mathbb{Z}_+}\left(\mathbf{y}^{d_k}\right)_0 = \lim_{k\in\mathbb{Z}_+}\mathbf{P}_{d_k}$. Moreover, since every subsequence of a convergent sequence converges to the same point, we have $\lim_{k\in\mathbb{Z}_+}\mathbf{P}_k = \lim_{k\in\mathbb{Z}_+}\mathbf{P}_{d_k} = \mathbf{P}^*_{\mathbf{y_q}}$. This shows that the subsequential limit $(\mathbf{y}^*, \mathbf{y}^*_{\mathbf{x}})$ is an optimal solution to (3.5). The rest of the claims follow from our previous results: Theorem 10 and Lemma 11.

3.9 Appendix D: Proof of Theorem 13

The LP in (3.4) can be rewritten as

$$\mathbf{P}^*_{\mathbf{1}} := \sup\langle\gamma, c\rangle \tag{3.34}$$

$$\text{s.t.}\quad A^*\gamma = b \tag{3.34a}$$

$$\gamma \in \mathcal{M}_+(\mathcal{K}) \times \mathcal{M}_+(\chi). \tag{3.34b}$$

where, $\gamma := (\mu, \mu_x) \in \mathcal{M}_+(\mathcal{K}) \times \mathcal{M}_+(\chi)$ is the variable vector, and $c := (1, 0) \in \mathcal{C}_+(\mathcal{K}) \times \mathcal{C}_+(\chi)$, so objective function is $\langle\gamma, c\rangle = \int d\mu$. Also, $A^* : \mathcal{M}_+(\mathcal{K}) \times \mathcal{M}_+(\chi) \to \mathcal{M}_+(\mathcal{Q} \times \chi) \times \mathbb{R}_+$ is the linear operator that is defined by $A^*\gamma := (\mu - \mu_x \times \mu_q, \int_\chi d\mu_x)$ and $b := (0, 1) \in \mathcal{M}_+(\mathcal{Q} \times \chi) \times \mathbb{R}_+$,([49], Theorem 2, [79, 82]). The problem in (3.34) is infinite LP defined in cone of nonnegative measures. The cone of nonnegative continuous functions are dual to cone of nonnegative measures. Based on standard results on LP ([49], Theorem 2, [79, 82]), the dual problem of (3.34) reads as

$$\mathbf{P}^*_{\mathbf{2}} := \inf\langle b, z\rangle \tag{3.35}$$

$$\text{s.t.}\quad Az - c \in \mathcal{C}_+(\mathcal{K}) \times \mathcal{C}_+(\chi) \tag{3.35a}$$

where, $z := (\mathcal{W}(x, q), \beta) \in \mathcal{C}_+(\mathcal{Q} \times \chi) \times \mathbb{R}_+$ is the variable vector, so the objective function is $\langle b, z\rangle = \beta$. The linear operator $A : \mathcal{C}_+(\mathcal{Q} \times \chi) \times \mathbb{R}_+ \to \mathcal{C}_+(\mathcal{K}) \times \mathcal{C}_+(\chi)$ satisfies adjoint relation$\langle A^*\gamma, z\rangle = \langle\gamma, Az\rangle$; hence, is defined by $Az := (\mathcal{W}(x, q), \beta - \int_\mathcal{Q} \mathcal{W}(x, q)d\mu_q)$. As a result, the dual problem (3.35) is equal to the problem (3.34).

If problem in (3.34) is consistent with finite value and the set

$$D := \{(A^*\gamma, \langle\gamma, c\rangle) : \gamma \in \mathcal{M}_+(\mathcal{K}) \times \mathcal{M}_+(\chi)\}$$

is closed, then there is no duality gap between (3.34) and (3.35), ([79], Theorem 3.10, [82], Theorem 7.2). The support of measures in (3.34) are compact. Also, the measure μ is constrained by the measure $\mu_x \times \mu_q$ in which, measure μ_x is probability measure; i.e., $\mu_x(\chi) = 1$, and μ_q is finite Borel measure defined on compact set $\mathcal{Q}$. Hence, $\mathbf{P_1^*} = \sup \int d\mu < \infty$. Also, the feasible set of (3.34) is nonempty for instance $(\delta_x \times \mu_q, \delta_x)$ for $x \in \chi$ is a feasible solution; therefor $0 \leq \mathbf{P_1^*} = \sup \int d\mu < \infty$. Using sequential Banach−Alaoglu theorem [80] and weak-$\star$ continuity of the A^*, there exist an accumulation point of $\gamma_k = (\mu_k, \mu_{xk})$ in the weak-$\star$ topology of nonnegative measures such that $\lim_{k\to\infty} ((A^*\gamma_k, \langle\gamma_k, c\rangle)) \in D$; hence, D is closed, ([49], Theorem 2).

3.10 Appendix E: Proof of Theorem 14

Matrices of the problem (3.6) can be rewritten as follow, ([67, 68]). $M_r(\mathbf{y}) = \sum_\alpha A_\alpha y_\alpha$ and $M_{d-r_j}(\mathbf{y}; \mathcal{P}_j) = \sum_\alpha B_\alpha^j y_\alpha$. Also, $M_d(\mathbf{y_x}) = \sum_\alpha D_\alpha y_{x\alpha}$, $M_{d-r_i}(\mathbf{y_x}; \{1 - x_i^2\}_{i=1}^n)) = \sum_\alpha E_\alpha^i y_{a\alpha}$, and $M_d(\mathbf{y_x} \times \mathbf{y_q} - \mathbf{y}) = \sum_\alpha F_\alpha y_{x\alpha} - \sum_\alpha A_\alpha y_\alpha$ for appropriate real symmetric matrices $(A_\alpha, \{B_\alpha^j\}_{j=1}^l,\ D_\alpha, \{E_\alpha^i\}_{i=1}^n, F_\alpha)$ and $0 \leq |\alpha| \leq 2d$. Let, $\gamma = (\mathbf{y} \in \mathbb{R}^{S_{n+m,2d}},\ \mathbf{y_x} \in \mathbb{R}^{S_{n,2d}})$. Then problem in (3.6) can be rewritten as a standard form as follow:

$$\mathbf{P_r^*} := \sup_\gamma b^T \gamma, \tag{3.36}$$

$$\text{s.t.}\quad C_1 + \sum_\alpha \hat{A}_\alpha \gamma_\alpha \succcurlyeq 0, \tag{3.36a}$$

$$C_2^j + \sum_\alpha \hat{B}_\alpha^j \gamma_\alpha \succcurlyeq 0, \quad j = 1, \ldots, l \tag{3.36b}$$

$$C_3 - \sum_\alpha \hat{C}_\alpha \gamma_\alpha \succcurlyeq 0, \tag{3.36c}$$

$$C_4 + \sum_\alpha \hat{D}_\alpha \gamma_\alpha \succcurlyeq 0, \tag{3.36d}$$

$$C_5^j + \sum_\alpha \hat{E}_\alpha^j \gamma_\alpha \succcurlyeq 0, \quad j = 1, \ldots, n \tag{3.36e}$$

$$C_6 + \sum_\alpha \hat{F}_\alpha \gamma_\alpha \succcurlyeq 0, \tag{3.36f}$$

where, $b = (1, \mathbf{0}) \in \mathbb{R}^{S_{n+m,2d}+S_{m,2d}}$, $(C_1, C_2, C_4, C_5, C_6)$ are zero matrices, $(\hat{A}_\alpha, \{\hat{B}_\alpha^j\}_{j=1}^l$ $, \hat{D}_\alpha, \{\hat{E}_\alpha^j\}_{j=1}^n, \hat{F}_\alpha)$ are real symmetric matrices, $C_3 = 1$, and $\hat{C}^T = (\mathbf{0} \in \mathbb{R}^{S_{n+m,2d}}, 1, \mathbf{0} \in$

$\mathbb{R}^{S_{m,2d-1}}) \in \mathbb{R}^{S_{n+m,2d}+S_{m,2d}}$. Based on standard results on duality of SDP, the dual problem to (3.36) reads as ([83, 84])

$$\mathbf{P_d^*} = \inf_{\{X^j\}_{j=0}^l, \{Y^j\}_{j=0}^n, Z, \beta} \left\langle C_1, X^0 \right\rangle + \sum_{j=1}^{l} \left\langle C_2^j, X^j \right\rangle + \langle C_3, \beta\rangle + \left\langle C_4, Y^0 \right\rangle + \sum_{j=1}^{n} \left\langle C_5^j, Y^j \right\rangle + \langle C_6, Z\rangle \tag{3.37}$$

$$\text{s.t.} \quad \beta - \left\langle A_\alpha, X^0 \right\rangle - \sum_{j}^{l} \left\langle B_\alpha^j, X^j \right\rangle - \left\langle D_\alpha, Y^0 \right\rangle - \sum_{j}^{n} \left\langle E_\alpha^j, Y^j \right\rangle - \langle F_\alpha, Z\rangle = b_\alpha, \quad \alpha = 0, \tag{3.37a}$$

$$- \left\langle A_\alpha, X^0 \right\rangle - \sum_{j}^{l} \left\langle B_\alpha^j, X^j \right\rangle - \left\langle D_\alpha, Y^0 \right\rangle - \sum_{j}^{n} \left\langle E_\alpha^j, Y^j \right\rangle - \langle F_\alpha, Z\rangle = b_\alpha, \quad 0 < |\alpha| \leq 2d, \tag{3.37b}$$

$$X^0, \{X^j\}_{j=1}^l, Y^0, \{Y^j\}_{j=1}^n, Z, \beta \succcurlyeq 0 \tag{3.37c}$$

where, $\langle X, Y\rangle = \mathrm{trace}(XY)$. This problem is equal to the problem in (3.15). Based on the defined matrices and vectors, the cost function of (3.37) is equal to β. Also, let $\mathcal{B}_d$ denote the vector comprised of the monomial basis of $\mathbb{R}_{\mathrm{d}}[x, q]$. We can represent the polynomials of (3.15) as $\mathcal{P}_{\mathcal{W}}^d(q, x) = \mathcal{B}_d^T X^0 \mathcal{B}_d$, $\mathcal{QM}\left(\{\mathcal{P}_j\}_{j=1}^{l_1}\right) = \sum_j^{l_1} \mathcal{B}_d^T X^j \mathcal{B}_d$, $\int \mathcal{P}_{\mathcal{W}}^d(q, x) d\mu_q = \mathcal{B}_d^T Y^0 \mathcal{B}_d$, $\mathcal{QM}\left(\{1 - x_j^2\}_{j=1}^n\right) = \sum_j^n \mathcal{B}_d^T Y^j \mathcal{B}_d$, and $\hat{\mathcal{P}}_{\mathcal{W}}^d(q, x) = \mathcal{B}_d^T Z \mathcal{B}_d$. Then constraints (3.37a) and (3.37b) are conditions for α-th coefficient of polynomial $\mathcal{P}_{\mathcal{W}}^d(q, x)$ so that as constraints (3.15a) and (3.15b), $\mathcal{P}_{\mathcal{W}}^d(q, x) - 1 \in \mathcal{QM}\left(\{\mathcal{P}_j\}_{j=1}^{l_1}\right)$, $\beta - \hat{\mathcal{P}}_{\mathcal{W}}^d(q, x) \in \mathcal{QM}\left(\{1 - x_j^2\}_{j=1}^n\right)$, and $\hat{\mathcal{P}}_{\mathcal{W}}^d(q, x) = \int_{\mathcal{Q}} \mathcal{P}_{\mathcal{W}}^d(x, q) d\mu_q$ are satisfied.

Based om Slater's sufficient condition ([83, 84]) if the feasible set of strictly positive matrices in constraint of primal SDP is nonempty, then there is no duality gap. Consider SDP in (3.6). Let μ_x be uniform measure on χ and $\mu = \mu_x \times \mu_q$. Since set $\mathcal{K}$ and χ have a nonempty interior, then $M_d(\mathbf{y}) \succ 0$, $M_{d-r_j}(\mathbf{y}; \mathcal{P}_j) \succ 0, j = 1, \ldots, l$, $M_d(\mathbf{y_x}) \succ 0$, and $M_{d-r_j}(\mathbf{y_x}; \{1 - x_j^2\}) \succ 0, j = 1, \ldots, n$. Based on Remark 3.2.1, $\chi \times \mathcal{Q} \setminus \mathcal{K}$ has nonempty interior; hence $M_r(\mathbf{y_x} \times \mathbf{y_q} - \mathbf{y}) \succ 0$. Therefore, Slater's condition holds, (see [67, 68] for similar setup). Also, Theorem 13 can be proved based strong duality condition provided in [83], (e.g., see [49, 95]).

3.11 Appendix F: Proof Of Theorem 15

Let $\mathbf{P}^*$ denote the optimal value of (3.1), and $\mathcal{K} = \cup_{k=1}^{N} \mathcal{K}_k$, where $\mathcal{K}_k$ is defined in (3.16). It can be proven as in Theorem 10 that

$$\mathbf{P}^* = \sup_{\mu_x \in \mathcal{M}(\chi)} \sup_{\mu \in \mathcal{M}(\mathcal{K})} \int d\mu \quad \text{s.t.} \quad \mu \preccurlyeq \mu_x \times \mu_q,\ \mu_x(\chi) = 1. \tag{3.38}$$

Let $\{\mu_k\}_{k=1}^N$ and μ_x be a feasible solution to (3.17) with objective value P. Since $\mu_k \in \mathcal{M}(\mathcal{K}_k) \subset \mathcal{M}(\mathcal{K})$ for all $k = 1, \ldots, N$, we have $\sum_{k=1}^N \mu_k \in \mathcal{M}(\mathcal{K})$. Hence, $\left(\sum_{k=1}^N \mu_k,\ \mu_x\right)$ is a feasible solution to (3.38) with objective value P, as well. Clearly, this shows that $\mathbf{P}^*_{\boldsymbol{\mu}_\mathbf{q}} \leq \mathbf{P}^*$, where $\mathbf{P}^*_{\boldsymbol{\mu}_\mathbf{q}}$ denotes the optimal value of (3.17).

Suppose that $(\mu,\ \mu_x)$ is a feasible solution to (3.38) with objective value P. Define $\{\mu_k\}_{k=1}^N$ as follows

$$\mu_k(S) := \mu\left(S \cap \left(\mathcal{K}_k \setminus \bigcup_{j=0}^{k-1} \mathcal{K}_j\right)\right), \quad \forall S \in \Sigma(\mathcal{K}), \tag{3.39}$$

for all $k = 1, \ldots, N$, where $\mathcal{K}_0 := \emptyset$ and $\Sigma(\mathcal{K})$ denotes the Borel σ-algebra over $\mathcal{K}$. Definition in (3.39) implies that $\mu_k \in \mathcal{M}(\mathcal{K}_k)$ for all $k = 1, \ldots, N$, and $\sum_{k=1}^N \mu_k(S) = \mu(S)$ for all $S \in \Sigma(\mathcal{K})$. Hence, $\{\mu_k\}_{k=1}^N$ and μ_x form a feasible solution to (3.4) with objective value equal to P. Therefore, $\mathbf{P}^*_{\boldsymbol{\mu}_\mathbf{q}} = \mathbf{P}^*$.

3.12 Appendix G: Proof Of Theorem 16

Let $\{\mathbf{y}_k\}_{k=1}^N \subset \mathbb{R}^{S_{n+m,2d}}$ and $\mathbf{y_x} \in \mathbb{R}^{S_{n,2d}}$ be a feasible solution to (3.18). As in Theorem 12, it can be shown that

$$\max\left\{(\mathbf{y})_0\,,\ \max_{i=1,\ldots,n+m} L_\mathbf{y}\left(x_i^{2d}\right)\right\} \leq 1, \tag{3.40}$$

where $\mathbf{y} := \sum_{k=1}^N \mathbf{y}_k$. Note that $L_\mathbf{y}\left(x_i^{2d}\right) = \sum_{k=1}^N L_{\mathbf{y}_k}\left(x_i^{2d}\right)$, and $\{L_{\mathbf{y}_k}\left(x_i^{2d}\right)\}_{i=1}^{n+m}$ is a subset of diagonal elements of $M_d(\mathbf{y}_k) \succeq \mathbf{0}$ for each $k \in \{1, \ldots, N\}$. Hence, $L_{\mathbf{y}_k}\left(x_i^{2d}\right) \geq 0$ for all $i \in \{1, \ldots, n+m\}$ and $k \in \{1, \ldots, N\}$. Therefore, (3.40) implies that

$$\max\left\{(\mathbf{y}_k)_0\,,\max_{i=1,\ldots,n+m} L_{\mathbf{y}_k}\left(x_i^{2d}\right)\right\}\leq 1 \tag{3.41}$$

for all $k \in \{1,\ldots,N\}$. Lemma 2 implies that $|(y_k)_\alpha| \leq 1$ for all $\alpha \in \mathbb{N}_{2d}^{n+m}$. Therefore, the feasible region is bounded. The rest of the proof is exactly the same as in Theorem 12.

Chapter 4

Convex Relaxation of Probabilistic Controller Design Problems

4.1 Introduction

In this chapter, we address the application of chance optimization algorithms in the control of stochastic systems. For this purpose, we consider the problems of probabilistic robust controller design and chance constrained model predictive control [70, 71]. In the problem of designing probabilistic robust controllers, we aim at designing robust controllers that maximize the probability of reaching a given target set. More precisely, given probability distributions for the initial state, uncertain parameters and disturbances, we develop algorithms for designing a control law that i) maximizes the probability of reaching the target set in N steps and ii) makes the target set robustly positively invariant.

In the chance constrained model predictive control problem, we aim at finding optimal control input for given disturbed dynamical system to minimize expected value of a given cost function subject to probabilistic constraints, over a finite horizon. The control laws provided have a predefined (low) risk of not reaching the desired target set.

Building on the theory of measures and moments, a sequence of finite semidefinite programmings are provided, whose solution is shown to converge to the optimal solution of the original problems. Numerical examples are presented to illustrate the computational performance of the proposed approach.

4.2 Probabilistic Robust Control

In this section, we provide results aimed at designing robust controllers that maximize the probability of reaching a given target set. More precisely, we start with an uncertain polynomial system subjected to external perturbations for which we know the probability distribution of the initial state, the uncertainty and the disturbances. Then, given a target set defined by polynomial inequalities and number of steps N, we provide algorithms for designing a nonlinear state feedback control law that i) makes the target set a robustly invariant set and ii) maximizes the probability of reaching the target set in N steps. It is assumed that a static polynomial state feedback control law exists that makes the target set robustly invariant. In the provided method, we incorporate the probability directly in the objective function and aim at maximizing the probability of desired defined control objectives. The proposed method is based on results on semialgebraic chance optimization provided in Chapter 3. Being, in general, a non-convex problem, a hierarchy of semidefinite relaxations for the approximation of the solution was proposed. These results provide the main motivation for the approach taken in this section.

In the next section, an explicit definition of chance robust control problem is given. Then, a sequence of convergent convex relaxations is provided.

4.2.1 Problem Statement

Consider the following discrete-time stochastic dynamic system

$$x(k+1) = f(x(k), u(k), \delta, \omega(k)) \tag{4.1}$$

where $f : R^{n+2m+p} \to R^n$ is a polynomial function, $x(k) \in \chi \subseteq R^n$ is the system state, $u(k) \in \psi \subseteq R^m$ is the control input, $\delta \in \Delta \subseteq R^p$ is the uncertain model parameter and $\omega(k) \in \Omega \subseteq R^m$ is the disturbance, at time step k.

The initial state $x(0) \in \chi_0 \subseteq \chi$, model parameter δ, and disturbance $\omega(k)$ at time k are independent random variables having probability measure μ_{x_0}, μ_δ, and μ_{ω_k}, with compact supports $supp(\mu_{x_0}) \subseteq \chi_0$, $supp(\mu_\delta) \subseteq \Delta$ and $supp(\mu_{\omega_k}) \subseteq \Omega$, respectively. We assume that χ_0, Δ, Ω are compact semialgebraic sets of the form $\chi_0 = \{x : g_0(x) \geq 0\}$, $\Delta = \{\delta : g_\delta(\delta) \geq 0\}$, $\Omega = \{\omega : g_\omega(\omega) \geq 0\}$ for given polynomials g_0, g_δ, g_ω. Although each of these sets is defined by just one polynomial, the approach proposed in this work can be extended to more complex semialgebraic sets. This assumption is only done to simplify the exposition.

Let N be a given integer. The desired terminal set at time step N is defined as the compact semialgebraic set

$$\chi_N = \{x : g_N(x) \geq 0\}.$$

We aim at finding a polynomial state feedback control input

$$u(x) = \sum_{\|i\|_1 \leq n_u} b_i x^i$$

where $u : R^n \to R^m$ is polynomial of order no more that n_u and $\underline{b} \in B$ is a vector of coefficients b_i, such that χ_N is an invariant set and maximizes the probability of reaching χ_N in N steps. Terminal set χ_N is invariant under control law if

$$f(x, u(x), \delta, \omega) \in \chi_N$$

for all

$$x \in \chi_N, \quad \delta \in \Delta, \quad \omega \in \Omega.$$

Under the definitions provided above, the stochastic control problem can be stated as follows

Problem 1: Solve,

$$\mathbf{P^1}^* = \max_{\underline{b}} Prob_{\mu_{x_0}, \mu_\delta, \underline{\mu_\omega}} \{g_N(x(N)) \geq 0\} \tag{4.2}$$

subject to,

$$x(k+1) = f(x(k), u(k), \delta, \omega(k))$$

$$u(k) = \sum_{\|i\|_1 \leq n_u} b_i x^i(k)$$

$$x_0 \sim \mu_{x_0}, \delta \sim \mu_\delta, \omega(k) \sim \mu_{\omega_k}, \underline{\mu_\omega} = [\mu_{\omega_0}, ..., \mu_{\omega_N}]$$

$$f(x, u(x), \delta, \omega) \in \chi_N \text{ for all } x \in \chi_N, \delta \in \Delta, \omega \in \Omega$$

4.2.2 An Equivalent Problem

As mentioned before we address this problem in two steps. First we determine a set of control laws that renders the set χ_N robustly positively invariant. Then, we search for a control law in this set that maximizes the probability of reaching χ_N in N steps.

4.2.2.1 Set Invariant Control Laws

In first step, we are looking for a set of parameters of control laws that render desired terminal set χ_N invariant. In this work, we approximate this set by a semialgebraic set $P_d \subseteq B$ of the form

$$P_d = \{\underline{b} : p_d(\underline{b}) \geq 0\}$$

where the p_d is a polynomial of order d of the form

$$p_d(\underline{b}) = \sum_{j \in \mathbb{N}^n} \lambda_j \underline{b}^j \in \mathbb{R}[\underline{b}]_{\mathrm{d}}.$$

To determine λ_j the coefficients of this polynomial, one needs to solve the following optimization problem involving SOS polynomials, which can be easily done using semidefinite programming.

$$min_{\lambda_j, \sigma_0, \sigma_1, \sigma_2, \sigma_3} \sum_{j \in \mathbb{N}^n} \gamma_j \lambda_j \tag{4.3}$$

subject to

$$g_N(f(x, \sum_{i=0}^{m} b_i x^i, \delta, \omega)) - \sum_{j \in \mathbb{N}^n} \lambda_j b^j = \sigma_0(x, \underline{b}, \delta, \omega)$$

$$+\sigma_1(x, \underline{b}, \delta, \omega) g_N(x) + \sigma_2(x, \underline{b}, \delta, \omega) g_\omega(\omega) + \sigma_3(x, \underline{b}, \delta, \omega) g_\delta(\delta)$$

where,γ_j is j-th moment of uniform probability measure over the set B of parameters of the control law and $\sigma_0, \sigma_1, \sigma_2, \sigma_3 \subset \sum^2[x, \underline{b}, \delta, \omega]$, are finite degree SOS polynomials such that $deg(\sigma_0) \leqslant d$, $deg(\sigma_1 g_N) \leqslant d$, $deg(\sigma_2 g_\omega) \leqslant d$, $deg(\sigma_3 g_\delta) \leqslant d$.

We then have the following result:

Theorem 17. *Let $p_d(\underline{b})$ be a polynomial constructed by solution of the optimization problem (4.3). Then, for any*

$$\underline{b} \in P_d = \{\underline{b} : p_d(\underline{b}) \geq 0\}$$

the corresponding control law

$$u(x) = \sum_{i \in N^n_{n_u}} b_i x^i$$

renders the set χ_N positively invariant. Moreover, define the set

$$P_{total} = \{\underline{b} : u(x) \text{ renders the set } \chi_N \text{ positively invariant}\}.$$

Then

$$\lim_{d \to \infty} \mu_B(P_{total} - P_d) = 0$$

where $P_{total} - P_d$ denotes the elements of P_{total} not in P_d.

Proof. See Appendix A. □

4.2.2.2 Maximizing Probability of Reaching χ_N

Now that we have an estimate of the set of control laws that render the set χ_N positively invariant, we can now address the problem of maximizing the probability of reaching the target set in at most N steps. Note that this is equivalent to maximizing the probability of $x(N) \in \chi_N$ since we have restricted the control laws to those that make the set χ_N invariant.

Define the function h as

$$x(N) = h(x_0, \underline{b}, \delta, \underline{\omega})$$

as the value of the state at time N when the value of the uncertain parameters is δ, the disturbances are

$$\underline{\omega} = [\omega_0, ..., \omega_N],$$

the control has coefficients $\underline{b}$ and the initial condition is x_0. Note that since one has a polynomial system and a polynomial control law, h is a polynomial. Additionally define the semialgebraic set

$$\mathcal{K}_1 = \{(x_0, \underline{b}, \delta, \underline{\omega}) : g_N(h(x_0, \underline{b}, \delta, \underline{\omega}) \geq 0\}$$

which represents all the values of the variables that will result in $x(N) \in \chi_N$ and the semialgebraic set

$$\mathcal{K}_2 = P_d(\underline{b}) \cap B$$

of control laws that render the set χ_N invariant. Define the following problem

Problem 2: Solve

$$\mathbf{P^2}^* = \max_{\mu, \mu_b} \int d\mu \tag{4.4}$$

subject to

$$\mu \preccurlyeq \mu_b \times \mu_{x_0} \times \mu_\delta \times \prod_{k=0}^{N} \mu_{\omega_k}$$

$$\mu_b \text{ is a probability measure}$$

$$supp(\mu) \subseteq \mathcal{K}_1$$

$$supp(\mu_b) \subseteq \mathcal{K}_2$$

This problem is equivalent to the problem addressed in this work in the following sense.

Theorem 18. *Problem 2 is equivalent to Problem 1 in the following sense: Let's restrict our attention to control laws that have*

$$\underline{b} \in P_d$$

Under this addition restriction one has

1. *The optimal values are the same.*

2. *If μ_b^* be a solution of Problem 2, then, any $\underline{b}^* \in supp(\mu_b^*)$ is a solution of Problem 1.*

3. *If $\underline{b}^*$ be a solution of Problem 1, then $\mu_b^* = \delta_{\underline{b}^*}$ is a solution of Problem 2*

Proof. See Appendix B. □

4.2.3 Semidefinite Relaxations

In this section, a sequence of semidefinite programs is provided which can arbitrarily approximate the optimal solution of Problem 2. Unlike Problem 2 in which we are looking for a measure, in the provided semidefinite program, we aim at finding a sequence of moments of a measure that satisfies the criteria of Problem 2. One should note that looking for a sequence of moments associated with one measure, is equivalent to looking for the measure itself. Proceeding as in Chapter 3.3.2, this leads to the following finite dimensional approximation.

Problem 3: Let $\mathbf{y} = (y_\alpha)$, $\mathbf{y_b} = (y_{b_\alpha})$ be a sequence with appropriate dimension, and defined semialgebraic sets $\mathcal{K}_1, \mathcal{K}_2$. Consider the sequence of semidefinite programs as:

$$\mathbf{P^{3^{*i}}} = \sup_{\mathbf{y}, \mathbf{y}_b} y_0 \tag{4.5}$$

subject to

$$M_i(\mathbf{y}) \succcurlyeq 0, M_{i-r_j}(g(h(.))\mathbf{y}) \succcurlyeq 0$$

$$M_i(\mathbf{y_b}) \succcurlyeq 0, M_{i-r_j}(p(.)\mathbf{y}_b) \succcurlyeq 0$$

$$M_i(\widehat{\mathbf{y}} - \mathbf{y}) \succcurlyeq 0$$

where, $\widehat{\mathbf{y}} = (\hat{y}_\alpha)$ are the moments of measure $\widehat{\mu} = \mu_b \times \mu_{x_0} \times \mu_\delta \times \prod_{k=0}^{N} \mu_{\omega_k}$, if one assumes that $\mathbf{y_b}$ are the moments of a measure μ_b. Given the fact that, $\mathbf{y_{x_0}}, \mathbf{y}_\delta, \mathbf{y}_{\omega_\mathbf{k}}$, sequence of moments of measures $\mu_{x_0}, \mu_\delta, \mu_{\omega_k}$, are given $\widehat{\mathbf{y}}$ is a linear transformation of $\mathbf{y_b}$. The sequence of problems provided above converges to the solution of the original problem. More precisely, we have the following result.

Theorem 19. *Optimal value of problem* $\mathbf{P^{3^i}}$ *converges to optimal value of problem* $\mathbf{P^2}$ *as* $\mathrm{i} \to \infty$.

Proof. See Appendix C. □

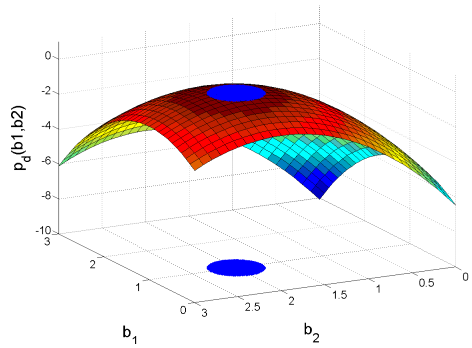

Figure 4.1: The polynomial $P_d\left(b_1, b_2\right)$ of example 2

4.2.4 Numerical Results

We now present two numerical examples that illustrate that the proposed semidefinite relaxations are effective in finding an appropriate control input even with lower order relaxations. Further research is needed to develop more efficient numerical implementations and study their behavior. Matlab toolboxes Yalmip [96] and Gloptipoly [97] are employed to solve the semidefinite programs 4.3 and 4.5, respectively.

4.2.4.1 Example 1: Nonlinear Control Problem

In this example, we consider the controller design problem for the following uncertain nonlinear dynamical system. For a given control parameter vector $K \in \mathbb{R}^3$, let the system $x(k)^T = [x_1(k),\ x_2(k),\ x_3(k)] \in \mathbb{R}^3$ satisfy

$$\begin{aligned} u(k) = &\quad K_1 x_1(k) + K_2 x_2(k) + K_3 x_3(k), \\ x_1(k+1) = &\quad \Delta\ x_2(k), \\ x_2(k+1) = &\quad x_1(k)\ x_3(k), \\ x_3(k+1) = &\quad 1.2\ x_1(k) - 0.5\ x_2(k) + x_3(k) + u(k), \end{aligned} \tag{4.6}$$

for $k = 0, 1$, where $x_1(0) \sim U[-1,1]$, $x_2(0) \sim U[-1,1]$, $x_3(0) \sim U[-1,1]$, $\Delta \sim U[-0.4, 0.4]$, i.e., initial state vector $x(0)$, and model parameter Δ are uncertain and uniformly distributed. The objective is to lead the system using state feedback control $u(k)$ to the cube centered at the origin with the edge length of 0.2 in at most 2 steps by properly choosing the control decision variables $\{K_i\}_{i=1}^3$ such that

$-1 \leq K_i \leq 1$. The equivalent chance problem is stated in (4.7), where $\mathbf{e}^T = [1, 1, 1]$.

$$\begin{aligned} \sup_{K \in \mathbb{R}^3} \; & \mu_q \left(\left\{ \left(x(0), \Delta \right) : \; -0.1\mathbf{e} \leq x(2) \leq 0.1\mathbf{e} \right\} \right), \\ \text{s.t. } & \{x(k), u(k)\}_{k=0}^{2} \text{ satisfy (4.26)}, \\ & -\mathbf{e} \leq K \leq \mathbf{e}. \end{aligned} \tag{4.7}$$

The following optimal solution and the corresponding optimal probability are computed by Monte Carlo method: $K_1^* = -1$, $K_2^* = 0.5$, $K_3^* = -0.9$, and $\mathbf{P}^* = 0.84$. To obtain an equivalent SDP formulation for the chance constrained problem in (4.7), $x(2)$ is explicitly written in terms of control vector $K \in \mathbb{R}^3$ and uncertain parameters, $x(0)$ and Δ, using the dynamic system given in (4.26):

$$\begin{aligned} x_1(2) = \; & \Delta \; x_1(0)x_3(0), \\ x_2(2) = \; & (1.2 + K_1)\Delta \; x_1(0)x_2(0) + (K_2 - 0.5)\Delta \; x_2(0)^2 + (1 + K_3)\Delta \; x_2(0)x_3(0), \\ x_3(2) = \; & (1 + 2K_3 + K_3^2) \; x_3(0) + (K_2 - 0.5K_3 - 0.5 + 1.2\Delta + K_1\Delta + K_2K_3) \; x_2(0) \\ & + (1.2 + K_1 + 1.2K_3 + K_1K_3) \; x_1(0) + (K_2 - 0.5) \; x_1(0)x_3(0). \end{aligned}$$

Based on the obtained polynomials, the minimum relaxation order for this problem is 2. To obtain an approximate solution, we solve the SDP in (3.6) using GloptiPoly and ALCC. For ALCC, we set ν_0 to 5×10^{-3}, 5×10^{-3} and 1×10^{-3} when d is equal to 2, 3 and 4, respectively, and tol $= 1 \times 10^{-3}$. The results for relaxation order $d = 2, 3, 4$ are shown in Table 4.1.

Example 2: Consider the uncertain systems as:

$$x_1(k+1) = \delta x_2(k) \tag{4.8}$$

$$x_2(k+1) = x_1(k) + 2x_2(k) + u(k) + \omega(k)$$

$$x_0 \sim U[-10, 10]^2, \delta \sim U[-0.5, 0.5], \omega(k) \sim U[-0.4, 0.4]$$

where, uncertain initial state x_0, model parameter δ, and disturbance ω has uniform probability distribution U. We aim at leading the system using state feedback control $u(k) = -b_1x_1(k) - b_2x_2(k)$ to the unit circle centered at the origin in at most 2 steps, in presence of uncertainties. Solving the semidefinite program 4.3, the semialgebraic

	ALCC				GloptiPoly		
d	2	3	4	**d**	2	3	4
$\mathbf{n}_{\mathrm{var}}$	365	1800	6600	$\mathbf{n}_{\mathrm{var}}$	365	1800	6600
iter	416	4300	5325	**iter**	19	26	36
cpu	14.934	897.708	5318.387	**cpu**	1.3	99.2	10389.8
$\mathbf{K_1}$	0	-0.244	-0.683	$\mathbf{K_1}$	0	-0.492	-0.796
$\mathbf{K_2}$	0	0.468	0.476	$\mathbf{K_2}$	0	0.439	0.487
$\mathbf{K_3}$	0	-0.868	-0.868	$\mathbf{K_3}$	0	-0.823	-0.891
$\mathbf{P}_d$	0.238	0.996	0.983	$\mathbf{P}_d$	1	1	1
$\mathbf{P}'_d$	0.65	0.9	0.982	$\mathbf{P}'_d$	0.65	0.959	0.999
$\bar{\mathbf{P}}_d$	0.061	0.445	0.685	$\bar{\mathbf{P}}_d$	0.061	0.508	0.766

Table 4.1: ALCC and GloptiPoly results for Example 1

set P_d for $d = 2$ is obtained as:

$$P_d = \{\underline{b} : -6.46 + 2.45b_1 + 5.35b_2 - 1.22b_1^2 - 0.009b_1b_2 - 1.33b_2^2 \geq 0\}$$

Solving semidefinite program 4.5 with relaxation order $i = 6$, the obtained optimal probability is 1. Using the obtained optimal $\mathbf{y}, \mathbf{y}_b$, the control is

$$u(k) = -0.99x_1(k) - 1.99x_2(k)$$

Where, $[0.99, 1.99]$ are the moments order one, from the moments sequence $\mathbf{y}_b$. Applying the obtained control input to the uncertain system, with probability one, the trajectories of the system for all initial states from the box $[-10, 10]^2$ will reach and remain in a unit ball, in presence of model uncertainty and disturbances; see Fig. 4.2.

Example 3: Consider the uncertain systems as:

$$x_1(k+1) = x_2(k) \tag{4.9}$$

$$x_2(k+1) = x_1(k)x_2(k) + u(k) + \omega(k)$$

$$x_0 \sim U[-5, 5]^2, \omega(k) \sim U[-0.5, 0.5]$$

where, uncertain initial state x_0, model parameter δ, and disturbance ω has uniform probability distribution U. We aim at leading the state of the system to unit box

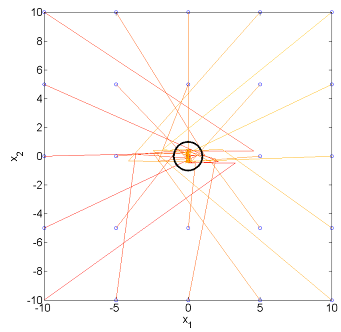

Figure 4.2: Example 2: trajectories of the uncertain system under obtained control input

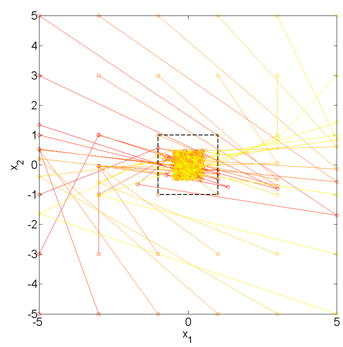

Figure 4.3: Example 3: trajectories of the uncertain system under obtained control input

$[-1, 1]^2$ at most in 2 steps, in presence of uncertainties. Since, the system consists of polynomial with order two, we use a state feedback control of the form $u(k) = b_1 x_1(k)^2 + b_2 x_1(k) x_2(k) + b_3 x_2(k)^2$ to control the system. Solving the semidefinite programs 4.5, with relaxation order $i = 6$, the obtained optimal probability is 1. Using the obtained optimal $\mathbf{y}, \mathbf{y}_b$, the control input is:

$$u(k) = 0.98x_1(k)^2 - 0.94x_1(k)x_2(k) - 0.98x_2(k)^2$$

State response of the system under obtained control input is provided in Fig 4.3.

Example 4: Consider the uncertain nonlinear system as

$$\begin{aligned} x_1(k+1) = &\ \ \delta x_2(k), \\ x_2(k+1) = &\ \ x_1(k)\ x_3(k), \\ x_3(k+1) = &\ \ x_1(k) - x_2(k) + x_3(k) + u(k) \end{aligned} \tag{4.10}$$

where, initial system states $x_1(0) \sim U[-1,1]$, $x_2(0) \sim U[-1,1]$, $x_3(0) \sim U[-1,1]$, and model parameter $\delta \sim U[-0.2, 0.2]$ are uncertain and uniformly distributed. Also, there is a sphere shaped obstacle centered at $(-0.5, -0.5, 0)$ with radius of 0.3 in the state space.

The objective is to find the state feedback control of the form $u(k) = a_1 x_1(k) + a_2 x_2(k) + a_3 x_3(k)$ to lead the states of the system to the cube centered at the origin with the edge length of 0.2 in at most 3 steps and at the same time to avoid the obstacle at each time k with high probability. In other word, we want to maximize the probability of semialgebraic sets $\chi_3 = \{-0.1 \leq x_1(3) \leq 0.1,\ -0.1 \leq x_2(3) \leq 0.1,\ -0.1 \leq x_3(3) \leq 0.1\}$ and $\chi_{x_k} = \{(x_1(k)+0.5)^2 + (x_2(k)+0.5)^2 + x_3(k)^2 - 0.3^2 \geq 0\}, k = 1, 2$. Hence, the semialgebraic of chance optimization problem reads as

$$\left\{(x_0, \delta) : \{-0.1 \leq \mathcal{P}_{x_i(3)} \leq 0.1\}_{i=1}^{3},\ \left\{(\mathcal{P}_{x_1(k)} + 0.5)^2 + (\mathcal{P}_{x_2(k)} + 0.5)^2 + \mathcal{P}^2_{x_3(k)} - 0.3^2 \geq 0\right\}_{k=1}^{2}\right\} \tag{4.11}$$

where, $\{x_i(k) = \mathcal{P}_{x_i(k)}(x_0, \delta, a)\}_{k=1}^{3},\ i = 1, 2, 3$, are states of the system in terms of control coefficients vector a, initial states x_0, uncertain parameters δ that is derived by dynamic of the system given in (4.10). The maximum degree of polynomials defining semialgebraic set is 8. Using Monte Carlo method, we obtain the optimal solution as $(a_1^*, a_2^*, a_3^*) = (-0.5, 1, -1)$ and the corresponding optimal probability as 1. To obtain an approximate solution, we solve SDP in (3.6).

Based on moments of uniform measures, we construct the matrices in constraints of SDP (3.6) in terms of unknown moment vectors $\mathbf{y} \in \mathbb{R}^{S_{7,2d}}$ and $\mathbf{y_a} \in \mathbb{R}^{S_{3,2d}}$. The SDP in (3.6) with $d = 7$ is solved using GloptiPoly. Based on obtained solution for moment vectors, we approximate the (a_1, a_2, a_3) with the first order moments of vector y_a as $(y_{a_{100}}, y_{a_{010}}, y_{a_{001}}) = (-0.2820, 0.4766, -0.8602)$ and also we approximate the probability with zero moment of vector y as $y_{0000000} = 1$. Using Monte Carlo method, the true probability for obtained solution $(a_1, a_2, a_3) = (-0.2820, 0.4766, -0.8602)$ is computed as 0.95. To improve the estimated probability and also to estimate the probability of reaching to target set, i.e.,

$$\text{Prob}\left\{(x_0, \delta) : \{-0.1 \leq \mathcal{P}_{x_i(3)} \leq 0.1\}_{i=1}^{3}\right\}$$

and probability of avoiding the obstacle, i.e.,

$$\text{Prob}\left\{(x_0, \delta) : \left\{(\mathcal{P}_{x_1(k)} + 0.5)^2 + (\mathcal{P}_{x_2(k)} + 0.5)^2 + \mathcal{P}^2_{x_3(k)} - 0.3^2 \geq 0\right\}_{k=1}^{2}\right\}$$

separately, we solve the SDP suggested in chapter 3 for obtained points $(a_1, a_2, a_3) = (-0.2820, 0.4766, -0.8602)$. By solving the SDP with relaxation order of 7 for the set $\left\{(x_0, \delta) : \{-0.1 \leq \mathcal{P}_{x_i(3)} \leq 0.1\}_{i=1}^{3}\right\}$, the estimated probability of 0.994 is obtained, while the true one computed by Monte Carlo is 0.95. Also, the estimated probabilty for the set $\left\{(x_0, \delta) : \left\{(\mathcal{P}_{x_1(k)} + 0.5)^2 + (\mathcal{P}_{x_2(k)} + 0.5)^2 + \mathcal{P}_{x_3(k)}^2 - 0.3^2 \geq 0\right\}_{k=1}^{2}\right\}$ is obtained as 1, while the true one computed by Monte Carlo is 1. Figure 4.4 shows the trajectories of the uncertain system (4.10) controlled by obtained state feedback $u(k) = -0.2820x_1(k) + 0.4766x_2(k) - 0.8602x_3(k)$ for different initial points. Note that, although states of the system $x(k)$ avoids the obstacle, the trajectories between the points $x(k), k = 0, ..., 3$ may collide the obstacle.

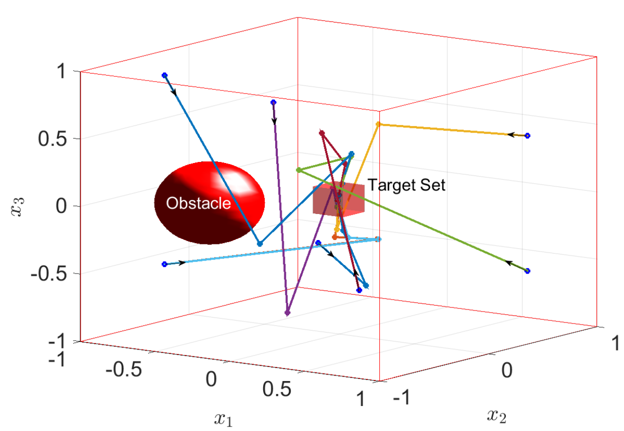

Figure 4.4: The trajectories of the uncertain system of Example 4 controlled by obtained state feedback

4.3 Chance Model Predictive Control

In this section, we aim at solving chance constrained model predictive control problems whose objective is to obtain finite-horizon optimal control of dynamical systems subject to probabilistic constraints. The control laws provided are designed to have precise bounds on the probability of achieving the desired objectives. More precisely, consider a polynomial dynamical system subject to external perturbation and assume that the probability distribution of the disturbances at each time is known. Then, given a desired set defined by polynomial inequalities and a polynomial cost function defined in terms of states and control input of the system, we aim at designing a controller to i) minimize the expected value of given cost function over the finite horizon and ii) reach the given desired set with high probability. For this purpose, at each sampling time we solve a convex optimization problem that minimizes the expected value of cost function subject to probabilistic constraints over the finite horizon.

In the next section, we precisely define the chance constrained MPC problem. Next, we provide equivalent infinite dimensional convex problem one measure and a semidefinite program on moments to solve obtained convex problem on measures.

4.3.1 Problem Formulation

We consider *chance constrained model predictive control* problem defined as follows. Consider the following discrete-time stochastic dynamical system

$$x_{k+1} = f(x_k, u_k, \omega_k) \tag{4.12}$$

where $f : \mathbb{R}^{n_x+n_u+n_\omega} \rightarrow \mathbb{R}^{n_x}$ is a polynomial function, $x_k \in \chi \subseteq \mathbb{R}^{n_x}$ is system state, $u_k \in \psi \subseteq \mathbb{R}^{n_u}$ is control input, and $\omega_k \in \Omega \subseteq R^{m_\omega}$ is disturbance, at time step k. The disturbances ω_k at time k are independent random variables with probability measure μ_{ω_k} supported on Ω, respectively. We assume that Ω is compact semialgebraic set of the form $\Omega = \{\omega \in \mathbb{R}^{n_\omega} : \mathcal{P}_\omega(\omega) \geq 0\}$ for given polynomial $\mathcal{P}_\omega$. Also, let χ_N be a given desired set defined by the compact semialgebraic sets as

$$\chi_D = \{x \in \chi : \mathcal{P}_{\chi_D}(x) \leq 0\} \tag{4.13}$$

In this work we aim at solving following problem.

Problem 1: For a given stochastic dynamical system in (4.12), find an optimal control u to:

i) Reach the desired set χ_D with high probability,

ii) Minimize the expected value of given cost function in terms of states and inputs of the system.

To obtain such control input, at each sampling time k, we solve the following optimization problem:

$$\mathbf{P}^*_{\mathbf{MPC}} := \min_{u \in \mathcal{U}} \mathrm{E}\left[\mathcal{P}_{cost}\left(\{x_i\}_{i=k+1}^{k+N_p}, \{u_i\}_{i=k}^{k+N_p}\right)\right] \tag{4.14}$$

s.t.

$$\mathrm{Prob}_{\mu_{\omega_k}}\{\mathcal{P}_{\chi_D}(x_{k+1}) \geq \alpha \mathcal{P}_{\chi_D}(x_k)\} \geq 1 - \beta \mathcal{P}_{\chi_D}(x_k) \tag{4.14a}$$

$$x_{k+1} = f(x_k, u_k, \omega_k), \quad \{\omega_i \sim \mu_{\omega_i}\}_{i=k}^{k+N_p-1} \tag{4.14b}$$

where, $u = \{u_i\}_{i=k}^{k+N_p} \in \mathcal{U} \subset \mathbb{R}^{N_p}$ is sequence of inputs, $E[.] = \int (.) d\mu_{\omega_k} ... d\mu_{\omega_{k+N_p-1}}$ is expected value operator, $N_p \geq 1 \in \mathbb{N}$ is prediction horizon. $0 < \alpha < 1$ and $0 < \beta < 1$ such that $0 \leq \beta \mathcal{P}_{\chi_D}(x) < 1$ for all $x \in \chi$. Polynomial $\mathcal{P}_{cost}\left(\{x_i\}_{i=k+1}^{k+N_p}, \{u_i\}_{i=k}^{k+N_p}\right)$ is defined cost function in terms of states and control input of the system over control and prediction horizon. We assume that the set of feasible control input $\mathcal{U}$ is a semialgebraic set defined as

$$\mathcal{U} := \left\{u = (u_k, ..., u_{k+N_p}) : \mathcal{P}_{\mathcal{U}}(u) \geq 0\right\} \tag{4.15}$$

Also, using the dynamic of the system in (4.12), $\{x_i\}_{i=k+1}^{k+N_p}$, sequences of system states over the prediction horizon, can be explicitly expressed in terms of disturbance and input of the system as

$$x_i = \mathcal{P}_{x_i}(\{u_j\}_{j=k}^{i-1}, \{\omega_j\}_{j=k}^{i-1}) \quad i = k+1, ..., N_p \tag{4.16}$$

Then, expected value in the cost function (4.14) can be rewritten in terms of inputs as

$$\mathrm{E}\left[\mathcal{P}_{cost}\left(\{x_i\}_{i=k+1}^{k+N_p}, \{u_i\}_{i=k}^{k+N_p}\right)\right] = \mathcal{P}_E(u) \tag{4.17}$$

where, $\mathcal{P}_E : \mathbb{R}^{N_p} \to \mathbb{R}$ is a polynomial function and $u = \{u_i\}_{i=k}^{k+N_p}$.

By solving problem in (4.14), we find sequence of control inputs $\{u_i\}_{i=k}^{k+N_p}$ that minimizes expected value of defined cost function over the finite horizon with respect to the chance constraint (4.14a). Chance constraint (4.14a) implies that the probability of getting closer to the desired set at next sampling time $k+1$ is bounded with respect to $\mathcal{P}_{\chi_D}(x_k)$, the distance of states of the system to the desired set at current time k. At each sampling time k, the first element of the obtained control input u is applied to the system. The implemented chance constraint (4.14a) depend only on u_k; hence, is recursively feasible.

Assumption: We assume that for every $x \in \chi$, there exist a u such that the probability constraint (4.14a) is satisfied. Hence, problem (4.14) is always feasible.

The following theorem holds true.

Theorem 20. *Given an initial state $x_0 \in \chi$ and $\epsilon > 0$ there exist a $\hat{k}(\epsilon, \alpha, \beta)$ and $\hat{P}(\epsilon, \alpha, \beta)$ such that*

$$Prob\left\{\mathcal{P}_{\chi_D}(x_k) \leq \epsilon,\ \forall k \geq \hat{k}(\epsilon, \alpha, \beta)\right\} \geq \hat{P}(\epsilon, \alpha, \beta) \tag{4.18}$$

where,

$$\hat{k}(\epsilon, \alpha, \beta) \geq \frac{ln(\epsilon) - ln(\mathcal{P}_{\chi_D}(x_0))}{ln(\alpha)} \tag{4.19}$$

$$\hat{P}(\epsilon, \alpha, \beta) = \prod_{i=0}^{\hat{k}-1}(1 - \beta\alpha^i) > 0 \tag{4.20}$$

Proof. See Appendix D. $\square$

The probability lower bound (4.20) is a convergent product and converges to a non-zoro constant. For example, consider the cases that $(\alpha, \beta) = (0.8, 0.05)$. For this case, $\hat{P}$ converges to 0.8169 for $\hat{k} \geq 36$. In the section 4.4, where numerical examples are presented, we consider this case for α and β.

Remark: The lower bound probability (4.20) is conservative bound and the actual probability of reaching the ϵ level set of $\mathcal{P}_{\chi_D}$ is greater than provided $\hat{P}(\epsilon, \alpha, \beta)$. However, lower bound (4.20) is useful for controller design purposes and shows that the probability of reaching the set is nonzero.

The provided problem in (4.14) is in general nonconvex and hard to solve. In the next section, we provide a convex equivalent problems to the problem (4.14).

4.3.2 Equivalent Convex Problem on Measures

As an intermediate step in the development of finite convex relaxations of the original problem in (4.14), a related infinite dimensional convex problem on measures is provided as follows. Let μ_u and μ be the finite nonnegative Borel measures and also the set $\mathcal{K}$ be defined as

$$\mathcal{K} := \{(u_k, \omega_k) : \mathcal{P}_{\chi_D}(x_{k+1}) - \alpha \mathcal{P}_{\chi_D}(x_k) \geq 0\} = \{(u_k, \omega_k) : \mathcal{P}_{\mathcal{K}}(u_k, \omega_k) \geq 0\} \quad (4.21)$$

where, polynomial $\mathcal{P}_{\mathcal{K}}$ can be obtained using system dynamics and and polynomial $\mathcal{P}_{\chi_D}$. Consider the following convex problem on measures:

$$\mathbf{P}^*_{\mathbf{measure}} := \sup_{\mu, \mu_u} \int \mathcal{P}_E(u) d\mu_u, \quad (4.22)$$

$$\text{s.t.} \int d\mu \geq 1 - \beta \mathcal{P}_{\chi_D}(x_k) \quad (4.22a)$$

$$\mu \preccurlyeq \mu_u \times \Pi_{i=k}^{k+N_p-1} \mu_{\omega_i}, \quad (4.22b)$$

$$\int \mu_u = 1, \quad (4.22c)$$

$$\mu \in \mathcal{M}_+(\mathcal{K}),\ \mu_u \in \mathcal{M}_+(\mathcal{U}). \quad (4.22d)$$

where, measures μ and μ_u are supported on the sets $\mathcal{U}$ and $\mathcal{K}$ defined as (4.15) and (4.21).

Assume that there exist a unique solution $u^* \in \mathcal{U}$ to the problem in (4.14). Then, following theorem shows the equivalency of the problem in (4.22) and the original volume problem in (4.14).

Theorem 21. *Assume that μ_u^*, the solution of the problem (4.22), is a delta distribution whose mass is concentrated on a single point u^*. Then, optimization problem in (4.14) is equivalent to the infinite LP in (4.22) in the following sense:*

i) The optimal values are the same, i.e., $\mathbf{P}^*_{\mathbf{MPC}} = \mathbf{P}^*_{\mathbf{measure}}$.

ii) $u^* \in supp(\mu_u^*)$ is an optimal solution to (4.14).

iii) If an optimal solution to (4.14) exists, call it u^*, then $\mu_u = \delta_{u^*}$, delta measure at u^*, and $\mu = \delta_{u^*} \times \Pi_{i=k}^{k+N_p-1} \mu_{\omega_i}$ is an optimal solution to (4.22).

Proof. See Appendix E. □

In the next section, we provide the tractable finite relaxations to the problem (4.22).

4.3.3 Semidefinite Programming Relaxations

In this section, we provide an finite dimensional semidefinite programming (SDP) of which feasible region is defined over real sequences. We show that the corresponding sequence of optimal solutions can arbitrarily approximate the optimal solution of (4.22), which characterizes the optimal solution of original problem in (4.14). Unlike the problem (4.22) in which we are looking for measures, in the SDP formulation given in (4.23), we aim at finding moment sequences corresponding to measures that are optimal to (4.22). Consider the following finite dimensional SDP:

$$\mathbf{P}_{\mathbf{r}}^{*} := \sup_{\mathbf{y}\in\mathbb{R}^{S_{(N_p-1)n_\omega+N_p,2r}},\ \mathbf{y_u}\in\mathbb{R}^{S_{N_p,2r}}} L_{\mathbf{y_u}}(\mathcal{P}_E\,(u))\,, \tag{4.23}$$

$$\text{s.t.}\quad M_r(\mathbf{y}) \succcurlyeq 0,\ M_{r-r_{\mathcal{K}}}(\mathbf{y};\mathcal{P}_{\mathcal{K}}) \succcurlyeq 0, \tag{4.23a}$$

$$(\mathbf{y})_{\mathbf{0}} \geq 1-\beta\mathcal{P}_{\chi_D}(x_k),\ (\mathbf{y_u})_{\mathbf{0}} = 1, \tag{4.23b}$$

$$M_r(\mathbf{y_u}) \succcurlyeq 0,\ M_{r-r_{\mathcal{U}}}(\mathbf{y_u};\mathcal{P}_{\mathcal{U}}) \succcurlyeq 0, \tag{4.23c}$$

$$M_r(\mathbf{y_u}\times\Pi_{i=k}^{k+N_p-1}\mathbf{y}_{\omega_\mathbf{i}} - \mathbf{y}) \succcurlyeq 0. \tag{4.23d}$$

where $L_{\mathbf{y_u}}$ is the linear map defined in (2.2). $(\mathbf{y})_{\mathbf{0}}$ and $(\mathbf{y_u})_{\mathbf{0}}$ are first element of the sequences $\mathbf{y}$ and $\mathbf{y_u}$, respectively. Polynomials $\mathcal{P}_{\mathcal{U}}$ and $\mathcal{P}_{\mathcal{K}}$ are defined in (4.15) and (4.21). $r \in \mathbb{Z}_+$ is relaxation order of matrices, $d_{\mathcal{K}}$ and $d_{\mathcal{U}}$ are the degree of polynomial $\mathcal{P}_{\mathcal{K}}$ and $\mathcal{P}_{\mathcal{U}}$, $r_{\mathcal{K}} := \lceil \frac{d_{\mathcal{K}}}{2} \rceil$ and $r_{\mathcal{U}} := \lceil \frac{d_{\mathcal{U}}}{2} \rceil$. Also, $\mathbf{y_u}\times\Pi_{i=k}^{k+N_p-1}\mathbf{y}_{\omega_\mathbf{i}}$ is truncated moment sequence of measure $\mu_u \times \Pi_{i=k}^{k+N_p-1}\mu_{\omega_i}$. $M_{r-r_{\mathcal{K}}}(\mathbf{y};\mathcal{P}_{\mathcal{K}})$ and $M_{r-r_{\mathcal{U}}}(\mathbf{y_u};\mathcal{P}_{\mathcal{U}})$ are localization matrices constructed by polynomials $\mathcal{P}_{\mathcal{K}}$ and $\mathcal{P}_{\mathcal{U}}$.

Now, consider the following theorem.

Theorem 22. *The sequence of optimal solutions to the finite SDP in* (4.23) *converges to the moment sequence of measures that are optimal to the infinite LP in* (4.22)*. Hence,* $lim_{r\to\infty}\mathbf{P}_{\mathbf{r}}^{*} = \mathbf{P}_{\mathbf{measures}}^{*}$.

Proof. See Appendix F. □

As in Theorem 21 and Theorem 22, if equivalent problem on measures has delta distribution solution μ_u^*, then problems on measures and moments in (4.22) and (4.23) are equivalent to the chance constraint problem (4.14) and the optimal distribution μ_u^* is a delta distribution whose mass is concentrated on the single point u^*, i.e., its support is the singleton $\{u^*\}$. Such distributions, have moment matrices with rank one. Hence, we incorporate this observation in the formulation of the relaxed problem (4.23) as follows:

$$\mathbf{P}^*_{\mathbf{trace}} := \min_{\mathbf{y},\ \mathbf{y_u}} L_{\mathbf{y_u}}(\mathcal{P}_E\,(u)) + \omega_r \mathbf{Tr}(M_r(\mathbf{y_u})), \tag{4.24}$$

$$\text{s.t.} \quad (4.23\text{a}), (4.23\text{b}), (4.23\text{c}), (4.23\text{d}) \tag{4.24a}$$

where, $\mathbf{Tr}(.)$ is the trace function and $\omega_r > 0$. We want to minimize the expected value with a low rank momnet matrix $M_r({\mathbf{y}^*}_{\mathbf{u}})$. For this, we use the trace norm (nuclear norm) which is the convex envelope of the rank function, ([88, 98]). Since, $M_r({\mathbf{y}^*}_{\mathbf{u}}) \succcurlyeq 0$, $\mathbf{Tr}(M_r({\mathbf{y}^*}_{\mathbf{u}}))$ is equal to sum of singular values of $M_r({\mathbf{y}^*}_{\mathbf{u}})$.

Remark To be able to apply the provided chance constrained model predictive control to large scale systems, we can implement Fast MPC approach [99] where, one needs to compute the control input u_k offline for all possible states x_k. Then, the online controller can be implemented as a lookup table, (see [99] for more details).

4.4 Numerical results

In this section, two numerical examples are presented that illustrate the performance of the proposed method. To solve proposed SDP in (4.23), GloptiPoly is employed which is a Matlab-based toolbox aimed at optimizing moments of measures [97]. Using GloptiPoly, we call Mosek, which is an interior-point solver add-on for Matlab.

Example 1: Consider the unstable nonlinear system as

$$\begin{aligned} x_1(k+1) = &\ \ x_2(k), \\ x_2(k+1) = &\ \ x_1(k)x_2(k) + \omega(k) + u(k) \end{aligned} \tag{4.25}$$

where, $\chi = [-1,1]^2$ and disturbance $\omega_k \sim U[-0.5, 0.5]$ are uniformly distributed. The desired set is a circle centered at the origin with radius 0.2; hence $\chi_D = \{x \in \chi : \mathcal{P}_{\chi_D}(x) = x_1^2 + x_2^2 - 0.2^2 \leq 0\}$. The finite cost function is defined as $\mathcal{P}_{cost} = \sum_{i=k}^{k+N_p} \|x(i)\|_2^2 + \sum_{i=k}^{k+N_p} \|u(i)\|_2^2$, where $\|.\|_2$ is L-2 norm and $N_p = 3$ To obtain control input, we solve the SDP in (4.24) for $\alpha = 0.8$, $\beta = 0.0510$, $\omega_r = 1$, and relaxation order $r = 5$. The obtained control input at each time k for the initial condition $x_0 = (1, 1)$ is

$$u_k = [-0.5634, -0.4647, 0.0007]$$

where results in the trajectory of

$$x_1(k) = [1, 1, 0.878, -0.0430]$$

$$x_2(k) = [1, 0.878, -0.0430, -0.168]$$

Hence in 3 steps the trajectory of the system under control reaches the desired set. The observed disturbance is $\omega_k = [0.4421, -0.4570, -0.1315]$. Also, by applying the obtained control input u_k, the cost function at time k, $\|x(k)\|_2^2 + \|u(k)\|_2^2$ is as $[3.11, 2.56, 0.408]$ and also the trace of the moment matrix is as $[1.58, 1.37, 1.00]$. Moreover, the lower bound probability $1 - \beta\mathcal{P}_D(x_k)$ and the obtained probability of the event $\{\mathcal{P}_{\chi_D}(x_{k+1}) \geq \alpha\mathcal{P}_{\chi_D}(x_k)\}$ is as $[0.5, 0.558, 0.812]$. Note that, we stop the optimization problem and input control by reaching the desired set. We can add extra constraint that makes the given desired set, an invariant set; hence the trajectories of the system remains in the set despite all disturbance and uncertainties, (See Section 4.2.2.1 for more details).

Example 2:

Consider the uncertain nonlinear system as

$$\begin{aligned} x_1(k+1) = &\;\; x_2(k), \\ x_2(k+1) = &\;\; x_1(k)\; x_3(k), \\ x_3(k+1) = &\;\; x_1(k) - x_2(k) + x_3(k) + \omega(k) + u(k) \end{aligned} \tag{4.26}$$

where, $\chi = [-1, 1]^3$ and disturbances $\omega(k) \sim U[-0.5, 0.5]$ are uniformly distributed. Also, The desired set is a circle centered at the origin with radius 0.2.The finite cost function is defined as $\mathcal{P}_{cost} = \sum_{i=k}^{k+N_p} \|x(i)\|_2^2 + \sum_{i=k}^{k+N_p} \|u(i)\|_2^2$, where $\|.\|_2$ is L-2 norm and $N_p = 3$. To obtain control input, we solve the SDP in (4.24) for $\alpha = 0.9$, $\beta = 0.2027$, $\omega_r = 1$, and relaxation order $r = 5$. The obtained control input at each time k for the initial condition $x_0 = (1, 1, 1)$ is

$$u_k = [-0.227, -0.219, -0.325, -0.196, -0.215, -0.605, 0.550]$$

where results in the trajectory of

$$x_1(k) = [1, 1, 1, 0.752, 0.892, 0.417, -0.101, 0.0487]$$

$$x_2(k) = [1, 1, 0.752, 0.892, 0.417, -0.101, 0.0487, 0.041]$$

$$x_3(k) = [1, 0.752, 0.892, 0.554, -0.113, 0.116, -0.410, 0.171]$$

Hence in 7 steps the trajectory of the system under control reaches the desired set. The observed disturbance is

$$\omega_k = [-0.020, 0.359, -0.260, -0.332, -0.028, -0.440, 0.182]$$

Also, by applying the obtained control input u_k, the cost function at time k, $\|x(k)\|_2^2 + \|u(k)\|_2^2$ is as $[7.61, 5.33, 5.86, 2.95, 1.6, 1.61, 1.45]$ and also the trace of the moment matrix is as $[1.26, 1.22, 1.34, 1.16, 1.12, 1.65, 1.68]$. Moreover, the lower bound probability $1 - \beta\mathcal{P}_D(x_k)$ and the obtained probability of the event $\{\mathcal{P}_{\chi_D}(x_{k+1}) \geq \alpha\mathcal{P}_{\chi_D}(x_k)\}$ is as $[0.5, 0.57, 0.6, 0.72, 0.84, 0.973, 0.976]$.

4.5 Conclusion

In this chapter, we presented a novel approach based on chance optimization results to the chance constrained controller design when the objective is to reach a given target set with high probability. More precisely, given a target set defined by polynomial inequalities and number of steps N, we provide algorithms for designing a nonlinear state feedback control law that i) makes the target set a robustly invariant set and ii) maximizes the probability of reaching the target set in N steps. Also, we provide chance constrained model predictive control problems whose objective is to obtain finite-horizon optimal control of dynamical systems subject to probabilistic constraints. The control laws provided are designed to have precise bounds on the probability of achieving the desired objectives.

These problems are, in general, nonconvex and computationally hard. Using theory of measures and moments, a sequence of semidefinite relaxations is provided whose sequence of optimal values is shown to converge to the optimal value of the original problem. Numerical examples are provided that show that one can obtains reasonable approximations to the optimal solution. In dynamical systems with uncertain parameters and initial states, states of the system at each time step are uncertain. In the next Chapter, we consider the problem computing uncertainty set for states of the system where we have probabilistic representation of uncertainty.

4.5.1 Appendix A: Proof of Theorem 17

Recall that

$$\chi_N = \{x : g_N(x) \geq 0\}.$$

Define function

$$p^*(\underline{b}) = \min_{x \in \chi_N, \delta \in \Delta, \omega \in \Omega} \left\{ g_N(f(x, \sum_i b_i x^i, \delta, \omega)) \right\} \tag{4.27}$$

Then,

$$u(x) = \sum_i b_i x^i$$

renders the set invariant if and only if

$$p^*(\underline{b}) \geq 0.$$

Now, the results on robust polynomial optimization in [100, 77, 101] show that

$$p_d(\underline{b}) \leq p^*(\underline{b}) \text{ for all } \underline{b} \in B.$$

Hence, for any $\underline{b} \in P_d$, we have

$$0 \leq p_d(\underline{b}) \leq p^*(\underline{b})$$

and the corresponding control law makes the set χ_N invariant. This proves the first part of the theorem.

The second part of the theorem is a consequence of the fact that

$$\int |p_d(\underline{b}) - p^*(\underline{b})| d\mu_B(\underline{b}) \to 0 \text{ as } d \to \infty$$

a result that has also been proven in [100, 77, 101].

Therefore, $\underline{b}$ parameters of control law should belong to the semialgebraic set P_d, otherwise the trajectories of system may not remain inside the desired terminal set.

4.5.2 Appendix B: Proof of Theorem 18

Problem 1 is a semialgebraic chance constrained optimization, where $[x_0, \delta, \omega_0, ..., \omega_N]$ is a random vector with probability measure $[\mu_{x_0}, \mu_\sigma, \mu_{\omega_0}, ..., \mu_{\omega_N}]$, and $[b_0, ..., b_m] \in \mathcal{K}_2$ is our decision variable vector, and $g_N(h(x_0, \underline{b}, \delta, \underline{\omega}))$ is a polynomial. Therefore, based on (4.2), the problem of maximizing probability of reaching the target set in N steps is equivalent to problem (4.4). This is a consequence of the results of Theorems 10 and 15.

4.5.3 Appendix C: Proof of Theorem 19

It can be shown that $\mathbf{y}, \mathbf{y}_b$ in Problem 3 are bounded and converge to the sequence of moments of measures μ and μ_b satisfying the optimal value of Problem 3 in the weak * topology $\sigma(l_\infty, l_1)$ sense; see Theorems 12 and 16.

4.5.4 Appendix D: Proof of Theorem 20

Given the system in (4.12), the desired set χ_D, and the initial state $x_0 \in \chi$, the condition $\text{Prob}_{\mu_{\omega_k}} \{\mathcal{P}_{\chi_D}(x_{k+1}) \leq \alpha \mathcal{P}_{\chi_D}(x_k)\} \geq 1 - \beta \mathcal{P}_{\chi_D}(x_k)$ is satisfied at each sampling time k, where $0 < \alpha, \beta < 1$. For a given $\hat{k}$, we define the events χ_1 and χ_2 as follow:

$$\chi_1 = \{(x_0, ..., x_{\hat{k}}) : \mathcal{P}_{\chi_D}(x_{\hat{k}}) \leq \epsilon\} \tag{4.28}$$

$$\chi_2 = \{(x_0, ..., x_{\hat{k}}) : \mathcal{P}_{\chi_D}(x_{i+1}) \leq \alpha \mathcal{P}_{\chi_D}(x_i),\ i = 0, ..., \hat{k} - 1\} \tag{4.29}$$

where, $\alpha \mathcal{P}_{\chi_D}(x_{\hat{k}-1}) \leq \epsilon$ and; hence, $\alpha^{\hat{k}} \mathcal{P}_{\chi_D}(x_0) \leq \epsilon$. This implies that given x_0, ϵ, and α, the time $\hat{k}$ for which $\mathcal{P}_{\chi_D}(x_{\hat{k}}) \leq \epsilon$ has lower bound of

$$\hat{k} \geq \frac{\ln(\epsilon) - \ln(\mathcal{P}_{\chi_D}(x_0))}{\ln(\alpha)} \tag{4.30}$$

Also, $\chi_2 \subset \chi_1$ and thus $\text{Prob}(\chi_2) \leq \text{Prob}(\chi_1)$. Since, the distribution of the uncertain parameters and disturbance at each time k are independent, the stochastic model (4.12) has Markov property; hence, the probability of the event χ_2 is

$$\text{Prob}\{\chi_2\} = \prod_{i=0}^{\hat{k}-1} \text{Prob}\{\mathcal{P}_{\chi_D}(x_{i+1}) \leq \alpha \mathcal{P}_{\chi_D}(x_i) | x_i\} \tag{4.31}$$

The probability in (4.31) has lower bound as

$$\text{Prob}\{\chi_2\} \geq \prod_{i=0}^{\hat{k}-1} (1 - \beta \mathcal{P}_{\chi_D}(x_i)) \geq \prod_{i=0}^{\hat{k}-1} (1 - \beta \alpha^i) \tag{4.32}$$

where, $\mathcal{P}_{\chi_D}(x_i) \leq \alpha^i \mathcal{P}_{\chi_D}(x_0)$. Hence, the lower bound of probability read as

$$\hat{P}(\epsilon, \alpha, \beta) = \prod_{i=0}^{\hat{k}-1} (1 - \beta \alpha^i) \tag{4.33}$$

This is a convergent product and converges to nonzero constant as $\hat{k} \to \infty$. As $\epsilon \to 0$, by (4.30) $\hat{k} \to \infty$; hence, $\hat{P}$ is non-zero and bounded.

4.5.5 Appendix E: Proof of Theorem 21

Consider the following problem over the measures μ_u

$$\mathbf{P}_{\mu_\mathbf{u}} := \min_{\mu_u \in \mathcal{M}_+(\mathcal{U})} \int_{\mathcal{U}} \mathcal{P}_E(u) d\mu_u \tag{4.34}$$

s.t.

$$\int_{\mathcal{U}} \text{Prob}\{\mathcal{P}_{\chi_D}(x_{k+1}) \geq \alpha \mathcal{P}_{\chi_D}(x_k)\} \, d\mu_u \geq \int_{\mathcal{U}} (1 - \beta \mathcal{P}_{\chi_D}(x_k)) d\mu_u \tag{4.34a}$$

$$\int d\mu_u = 1, \quad \{\omega_i \sim \mu_{\omega_i}\}_{i=k}^{k+N_p-1} \tag{4.34b}$$

We first want to show that $\mathbf{P}^*_{\mathbf{MPC}} = \mathbf{P}_{\mu_\mathbf{u}}$. Let μ_u be a feasible solution to (4.34), i.e., $\int_{\mathcal{U}} \text{Prob}\{\mathcal{P}_{\chi_D}(x_{k+1}) \geq \alpha \mathcal{P}_{\chi_D}(x_k)\} \, d\mu_u \geq 1 - \beta \mathcal{P}_{\chi_D}(x_k))$. Then for any u in support of measure μ_u $\text{Prob}\{\mathcal{P}_{\chi_D}(x_{k+1}) \geq \alpha \mathcal{P}_{\chi_D}(x_k)\} \geq 1 - \beta \mathcal{P}_{\chi_D}(x_k))$, i.e, the feasible set of problem (4.14). Also, Since, $\mathcal{P}_E(u) \leq \mathbf{P}^*_{\mathbf{MPC}}$ for all $u \in \mathcal{U}$, we have$\int_{\mathcal{U}} \mathcal{P}_E(u) d\mu_u \leq \mathbf{P}^*_{\mathbf{MPC}}$. Thus,

$\mathbf{P}_{\mu_\mathbf{u}} \leq \mathbf{P}^*_{\mathbf{MPC}}$. Conversely, let $u \in \mathcal{U}$ be a feasible solution to the problem in (4.14). Let δ_u denotes the Dirac measure at u. Then the δ_u belongs to the feasible set of problem (4.34). The objective value of u in (4.14) is equal to $\mathcal{P}_E(u)$. Moreover, $\mu_u = \delta_u$ is a feasible solution to the problem in (4.34) with objective value equal to $\mathcal{P}_E(u)$. This implies that $\mathbf{P}^*_{\mathbf{MPC}} \leq \mathbf{P}_{\mu_\mathbf{u}}$. Hence, $\mathbf{P}^*_{\mathbf{MPC}} = \mathbf{P}_{\mu_\mathbf{u}}$, and (4.34) can be rewritten as

$$\mathbf{P}_{\mu_\mathbf{u}} := \min_{\mu_u \in \mathcal{M}_+(\mathcal{U})} \int_{\mathcal{U}} \mathcal{P}_E(u) d\mu_u \tag{4.35}$$

s.t.

$$\int_{\mathcal{U}} \int_{\mathcal{K}} d\mu_u d\mu \geq 1 - \beta \mathcal{P}_{\chi_D}(x_k) \tag{4.35a}$$

$$\int d\mu_u = 1, \quad \{\omega_i \sim \mu_{\omega_i}\}_{i=k}^{k+N_p-1} \tag{4.35b}$$

where, set $\mathcal{K}$ is defined in (4.21). Using the Lemma 6, we obtain

$$\mathbf{P}^*_{\mathbf{measure}} := \min_{\mu, \mu_u} \int \mathcal{P}_E(u) d\mu_u, \tag{4.36}$$

$$\text{s.t.} \int d\mu \geq (1 - \beta \mathcal{P}_{\chi_D}(x_k)) \tag{4.36a}$$

$$\mu \preccurlyeq \mu_u \times \Pi_{i=k}^{k+N_p-1} \mu_{\omega_i}, \tag{4.36b}$$

$$\int d\mu_u = 1, \tag{4.36c}$$

$$\mu \in \mathcal{M}_+(\mathcal{K}), \; \mu_u \in \mathcal{M}_+(\mathcal{U}). \tag{4.36d}$$

Note that, if there exist delta solution μ_u^* for the problem (4.22) whose mass is concentrated on a single point u^*, the $\int d\mu$ in constraint (4.36a) implies the probability of event $\{\mathcal{P}_{\chi_D}(x_{k+1}) \geq \alpha \mathcal{P}_{\chi_D}(x_k)\}$ for a control input u^*.

4.5.6 Appendix F: Proof of Theorem 22

Using Lemma (4) and (5), the constraints of problem (4.23) implies that the sequence of $\mathbf{y}$ and $\mathbf{y_u}$ are the moment sequence of the measures of problem (4.22). For more details, see Theorems 12 and 16.

Chapter 5

Uncertainty Propagation Through Uncertain Dynamical Systems

Given an uncertain dynamical system and a set of initial conditions with known probability distribution, we want to propagate the set of initial condition through uncertain dynamical system and find the set of uncertainty for given time step. To obtain such set, we reconstruct the support of probability distribution of states of the system using the information of the moments of distribution [72]. In this chapter, we first obtain the moment information of probability distribution of system states at time k. Next, we address the problem of reconstruction of support of a positive finite Borel measure from its moments. More precisely, given a finite subset of the moments of a measure, we develop a semidefinite program for approximating the support of measure using level sets of polynomials. To solve this problem, a sequence of convex relaxations is provided, whose optimal solution is shown to converge to the support of measure of interest. Moreover, the provided approach is modified to improve the results for uniform measures. Numerical examples are presented to illustrate the performance of the proposed approach.

5.1 Problem Statement

Consider the following discrete-time stochastic dynamic system

$$x(k+1) = f(x(k), u(k), \delta, \omega(k)) \tag{5.1}$$

where $f: R^{n+2m+p} \rightarrow R^n$ is a polynomial function, $x(k) \in \chi \subseteq R^n$ is the system state, $u(k) \in \psi \subseteq R^m$ is the control input, $\delta \in \Delta \subseteq R^p$ is the uncertain model parameter and $\omega(k) \in \Omega \subseteq R^m$ is the disturbance, at time step k. The initial state $x(0) \in \chi_0 \subseteq \chi$, model parameter δ, and disturbance $\omega(k)$ at time k are independent random variables having probability measure μ_{x_0}, μ_δ, and μ_{ω_k}, with compact supports $supp(\mu_{x_0}) \subseteq \chi_0$, $supp(\mu_\delta) \subseteq \Delta$ and $supp(\mu_{\omega_k}) \subseteq \Omega$, respectively.

Due to the uncertainty in the system, state of the system at time k is uncertain with probability measure μ_{x_k} supported on $\chi_k \in \chi$, i.e., $x_k \in \chi_k$, $\forall x_0 \in \chi_0, \delta \in \Delta, \omega \in \Omega$. In fact, the support set χ_k is the set of uncertainty obtained by propagation of initial state set χ_0 through uncertain system 5.1.

In this chapter, we aim at finding the support set χ_k using the moment information of the probability distribution μ_{x_k} associated to states x_k. For given k, we first obtain the moment information of μ_{x_k} and then, we provide the semidefinite program to reconstruct the support set χ_k, using the moment information.

5.2 Moment Information

Consider the following discrete-time stochastic dynamic system 5.1. States of the system at time k, can be explicitly written in terms of uncertain parameters x_0, δ, and $\{\omega_i\}_{i=1}^{k-1}$ using the dynamic of the system, i.e., $x_k = \mathcal{P}_{x_k}(x_0, \delta, \{\omega_i\}_{i=1}^{k-1})$ where $\mathcal{P}_{x_k}$ is a polynomial function. Hence, α-th moment associated with probability distribution of states x_k can be written in terms of known moments of uncertain parameters as:

$$m_\alpha(k) = E[x_k^\alpha] = E[\mathcal{P}_{x_k}^\alpha(x_0, \delta, \{\omega_i\}_{i=1}^{k-1})] = \sum_{i,j,k,l,\dots,p} a_i m_{x_{0_j}} m_{\delta_k} m_{\omega_{0l}} \dots m_{\omega_{k-1_p}}. \quad (5.2)$$

Where, a_i are known coefficients, $m_{x_{0_j}}$, m_{δ_k},$m_{\omega_{0l}}$, and $m_{\omega_{k-1_p}}$ are the moments of uncertain parameters x_0, δ, and ω_0, and ω_{k-1}, respectively.

For example, consider the following uncertain system $x(k+1) = \delta x^2(k) + \omega(k)$. We aim at finding the moment of stets at time $k = 2$. States of the system at time $k = 2$ can be written in terms of uncertain parameters as $x_2 = \delta^3 x_0^4 + \delta\omega_0^2 + 2\delta^2 x_0^2 \omega_0 + \omega_1$. Hence, the α-th moment reads as $m_{x_{2\alpha}} = E[x_2^\alpha] = E[(\delta^3 x_0^4 + \delta\omega_0^2 + 2\delta^2 x_0^2\omega_0 + \omega_1)^\alpha]$.

For instance, the moment of order α =2 reads as:

$$m_{x2_1} = E[x_2] = E[\delta^3 x_0^4 + \delta\omega_0^2 + 2\delta^2 x_0^2 \omega_0 + \omega_1] = m_{\delta_3} m_{x0_4} + m_{\delta_1} m_{\omega0_2} + 2m_{\delta_2} m_{x0_2} m_{\omega0_1} + m_{\omega1_1}$$

5.3 Support Reconstruction

In this section, we aim at solving the problem of reconstructing of support of a measure using only its moments. More precisely, we consider the following problem.

Problem Given the moment sequence of a measure μ, find a polynomial $\mathcal{P}$: $\mathbb{R}^n \to \mathbb{R}$ such that the set

$$\mathcal{K} = \{x \in \mathbb{R}^n : \mathcal{P}(x) \geq 1\}$$

coincides with the support set of the measure μ.

To reconstruct the support of the measure of interest from its moments, we develop a sequence of semidefinite programming (SDP) problems whose solutions converge to the solution of Problem 5.3. The proposed method relies on results on Sum of Squares (SOS) polynomials and also, results on necessary and sufficient condition for moment sequence to have a representing measure. A hierarchy of semidefinite relaxations for approximation of the support set is proposed.

5.3.1 Convex Formulation

The approach presented in this work relies on finding polynomial approximations of the indicator function of the support set of the measure of interest. More precisely, let $\mathcal{K}$ represent the support set of a given measure μ. The results in this work aim at finding polynomial approximations of

$$\mathbb{I}_{\mathcal{K}}(x) \doteq \begin{cases} 1 & \text{if } x \in \mathcal{K} \\ 0 & \text{otherwise.} \end{cases}$$

and use the level sets of these polynomials to approximate $\mathcal{K}$. In order to approximate the indicator function above consider the following optimization problem.

Problem Let d be a given integer. Moreover, let $\mathcal{B}$ be a known (simple) set containing the support set $\mathcal{K}$ and $\mu_{\mathcal{B}}$ be the Lebesgue measure supported on the set $\mathcal{B}$. Solve

$$\mathbf{P_2^*} := \min_{\mathcal{P}_d(x)\in\mathbb{R}_\mathrm{d}[x]} \int \mathcal{P}_d(x) d\mu_{\mathcal{B}} \tag{5.3}$$

$$\text{s.t.} \quad \mathcal{P}_d(x) \geq 0, \text{for all } x \in \mathcal{B} \tag{5.3a}$$

$$\mathcal{P}_d(x) \geq 1, x \in \mathcal{K}. \tag{5.3b}$$

For every d, the problem above provides a polynomial $\mathcal{P}_d^*$ with the smallest ℓ_1-norm on $\mathcal{B}$ that is i) positive in the (simple) set $\mathcal{B}$ and ii) larger than one inside the support set $\mathcal{K}$. For this (infinite dimensional) optimization problem we have the following result.

Theorem 23. *For a given integer d, let*

$$\mathcal{K}_d \doteq \{x \in \mathbb{R}^n : \mathcal{P}_d^*(x) \geq 1\}$$

be the semialgebraic set constructed using the solution $\mathcal{P}_d^*$ of the problem (5.3). Then

$$\lim_{\mathrm{d}\to\infty} \mu_{\mathcal{B}}(\mathcal{K}_d - \mathcal{K}) = 0.$$

Sketch of proof: As in [102] one can show that $\mathcal{P}_d^*$ converges almost uniformly (with respect to measure $\mu_{\mathcal{B}}$) to the indicator function $\mathbb{I}_{\mathcal{K}}$. Moreover, one has $\mathcal{K} \subseteq \mathcal{K}_d$ for all d. These two facts imply that

$$\lim_{\mathrm{d}\to\infty} \mu_{\mathcal{B}}(\mathcal{K}_d - \mathcal{K}) = 0$$

which completes the proof.

In the optimization problem above, one approximates the indicator function of the set $\mathcal{K}$ by using the knowledge that this set is contained in a known set $\mathcal{B}$. This set is usually chosen in such a way that one can compute all the moments of the measure $\mu_{\mathcal{B}}$ in a closed form.

However, the problem above obviously requires the knowledge of the measure μ whose support $\mathcal{K}$ we are trying to determine. To be able to solve this problem by using only knowledge of moments consider a bounding set $\mathcal{B}$ defined by a set of polynomial inequalities; i.e.,

$$\mathcal{B} = \{x \in \mathbb{R}^n : g_j(x) \geq 0, j = 1, ..., l\}$$

where g_j, $j = 1, 2, \ldots, l$ are given polynomials. As before, let $\mu_{\mathcal{B}}$ be the Lebesgue measure supported in $\mathcal{B}$ with α-th moment $y_{\mathcal{B}_\alpha}$. Moreover, let the (infinite) vector $\mathbf{y}$ be the vector containing all the moments of the measure μ. Then, define the following optimization problem (which has an infinite number of constraints).

Problem

$$\mathbf{P_3^*} := \min_{p_\alpha, \sigma_j} \sum_{\alpha=0}^{d} p_\alpha y_{\mathcal{B}_\alpha} \tag{5.4}$$

$$\text{s.t.} \quad \mathcal{P}_d(x) = \sum_{\|\alpha\|_1 \leq d} p_\alpha x^\alpha \tag{5.4a}$$

$$\mathcal{P}_d(x) \geq 0, \text{for all } x \in \mathcal{B} \tag{5.4b}$$

$$M_\infty((\mathcal{P}_d(x) - 1)\mathbf{y}) \succcurlyeq 0 \tag{5.4c}$$

The problem above is a first step towards an implementable version of Problem 5.3.1. The objective function is the same in both, just represented as a function of the moments of $\mu_{\mathcal{B}}$ in Problem 5.3.1. Constraint (5.4b) enforces $\mathcal{P}_d$ to be positive on the set $\mathcal{B}$. Finally, based on Lemma 5 constraint (5.4c) implies that moment sequence $\mathbf{y}$ has a representing measure μ supported on the set $\{x \in \mathbb{R}^n : \mathcal{P}_d(x) \geq 1\}$. Hence, $\mathcal{P}_d$ is larger than one in the support set of μ.

Since one cannot solve the problem above, we start be proposing the following relaxation.

Problem

$$\mathbf{P_4^*} := \min_{p_\alpha, \sigma_j} \sum_{\alpha=0}^{d} p_\alpha y_{\mathcal{B}_\alpha} \tag{5.5}$$

$$\text{s.t.} \quad \mathcal{P}_d(x) = \sum_{\|\alpha\|_1 \leq d} p_\alpha x^\alpha \tag{5.5a}$$

$$\mathcal{P}_d(x) \geq 0, \text{for all } x \in \mathcal{B} \tag{5.5b}$$

$$M_r((\mathcal{P}_d(x) - 1)\mathbf{y}) \succcurlyeq 0 \tag{5.5c}$$

where, $r \geq 1$ is relaxation order. In other words, we truncate the infinite moment localization matrix.

Theorem 24. *Polynomial $\mathcal{P}_d^*(x) = \sum_{\|\alpha\|_1 \leq d} p_\alpha^* x^\alpha$ constructed by p_α^*, the optimal solution of the Problem 5.3.1, converges to the indicator function of support of measure with known moments $\boldsymbol{y}$.*

Sketch of proof: The feasibility set of Problem 5.3.1 contains the feasibility set of Problem 5.3.1 and converges to it as $r \to \infty$. Hence, optimal value of Problem 5.3.1 converges to the optimal value of Problem 5.3.1.

The truncation of the moment localization matrix provides an approximation of the constraint $\mathcal{P}_d(x) \geq 1$ for all $x \in \mathcal{K}$. Although, if r is "large" one has acceptable estimates of the support set, for "low" values of r this can lead to estimates of the support set that are less accurate than desirable.

To be able to solve the problem above, in this work we use following SOS relaxation.

Problem

$$\mathbf{P_5^*} := \min_{p_\alpha, \sigma_j} \sum_{\alpha=0}^{d} p_\alpha y_{\mathcal{B}_\alpha} \tag{5.6}$$

$$\text{s.t.} \quad \mathcal{P}_d(x) = \sum_{\|\alpha\|_1 \leq d} p_\alpha x^\alpha \tag{5.6a}$$

$$\mathcal{P}_d(x) = \sigma_0(x) + \sum_{j=1}^{l} \sigma_j(x) g_j(x) \tag{5.6b}$$

$$\sigma_j \in \Sigma^2[x]; j = 0, 1, \ldots, l \tag{5.6c}$$

$$deg(\sigma_0) \leq d_{sos}; \tag{5.6d}$$

$$deg(\sigma_j g_j) \leq d_{sos}; j = 1, 2, \ldots, l \tag{5.6e}$$

$$M_r((\mathcal{P}_d(x) - 1)\mathbf{y}) \succcurlyeq 0 \tag{5.6f}$$

where, d_{sos} is SOS relaxation order. One can see that constraint (5.6b) enforces $\mathcal{P}_d$ to be positive on the set $\mathcal{B}$. Furthermore, as $d_{sos} \to \infty$, standard results on SOS relaxations can be used to show that one converges to the solution of Problem 5.3.1. One should note that the problem above can be formulated as a standard SDP; i.e., minimization of a linear function subject to Linear Matrix Inequalities (LMIs).

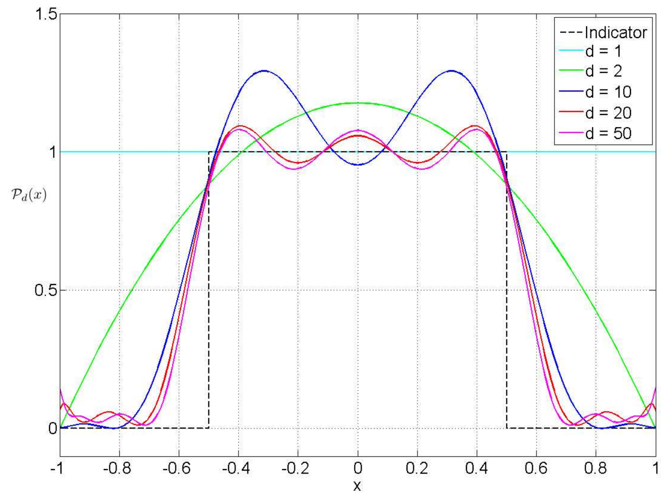

Figure 5.1: Result of SDP in (5.6) For Example 1

Example 1: Let, $\mathbf{y}$ be a moment sequence of uniform probability measure μ supported on $[-0.5, 0.5]$. The α-th moment of uniform distribution $U[a, b]$ is $y_\alpha = \frac{b^{\alpha+1}-a^{\alpha+1}}{(b-a)(\alpha+1)}$. For this example, we take $\mathcal{B} = [-1, 1]$, and use the moments up to order $2d$. To solve the SDP (5.6), Yalmip with Sedumi SDP solver is used. The obtained results are depicted in Fig 5.1. One can see as d, the order of polynomial, increases $\mathcal{P}_d(x)$ converges to indicator function of support of uniform measure. Hence, the semialgebraic set $\mathcal{K}_d = \{x \in \mathbb{R} : \mathcal{P}_d(x) \geq 1\}$ provides better approximations of the support as one increases d. However, as one can see in Fig 5.1, $\mathcal{P}_d$ can be below one in a significant subset of the support of μ.

5.3.2 An Heuristic for Improved Performance

To minimize the measure of the subset of the support of the measure μ where $\mathcal{P}_d$ is below one, we propose to maximize the values of $\mathcal{P}_d(x)$ inside the support of the measure while still trying to bring its values as low as possible everywhere else in $\mathcal{B}$. This results in following modified SDP.

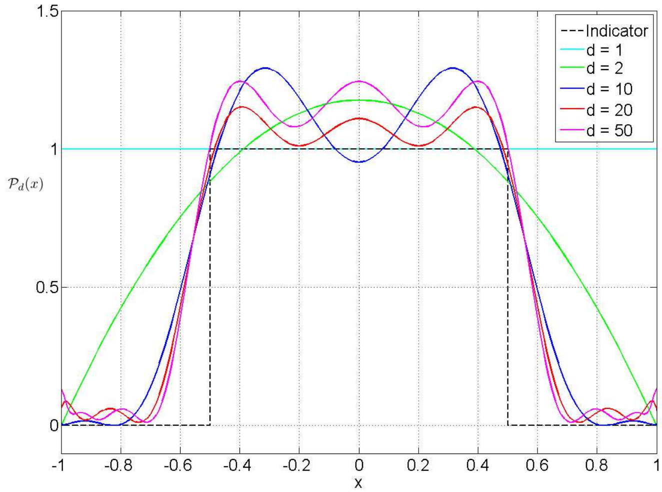

Figure 5.2: Result of SDP in (5.7) For Example 1

Problem

$$\mathbf{P_6^*} := \min_{p_\alpha, \sigma_j} \sum_{\alpha=0}^{d} p_\alpha y_{\mathcal{B}_\alpha} - \omega_h h \tag{5.7}$$

$$\text{s.t.} \quad \mathcal{P}_d(x) = \sum_{\|\alpha\|_1 \le d} p_\alpha x^\alpha \tag{5.7a}$$

$$\mathcal{P}_d(x) = \sigma_0(x) + \sum_{j=1}^{l} \sigma_j(x) g_j(x) \tag{5.7b}$$

$$\sigma_j \in \Sigma^2[x]; j = 0, 1, \dots, l \tag{5.7c}$$

$$deg(\sigma_0) \le d_{sos}; \tag{5.7d}$$

$$deg(\sigma_j g_j) \le d_{sos}; j = 1, 2, \dots, l \tag{5.7e}$$

$$M_r((\mathcal{P}_d(x) - h)\mathbf{y}) \succcurlyeq 0 \tag{5.7f}$$

$$1 \le h \le 1 + \Delta h \tag{5.7g}$$

where, ω_h and Δh are positive design parameters.

To show the effectiveness of the modified SDP, we again consider the uniform measure in Example 6.6.1. Fig 5.2 shows the results obtained by solving the modified SDP with parameters $\omega_h = 1.2$ and $\Delta h = 0.2$. As it is seen, on obtains a substantial improvement in the estimate of the support set.

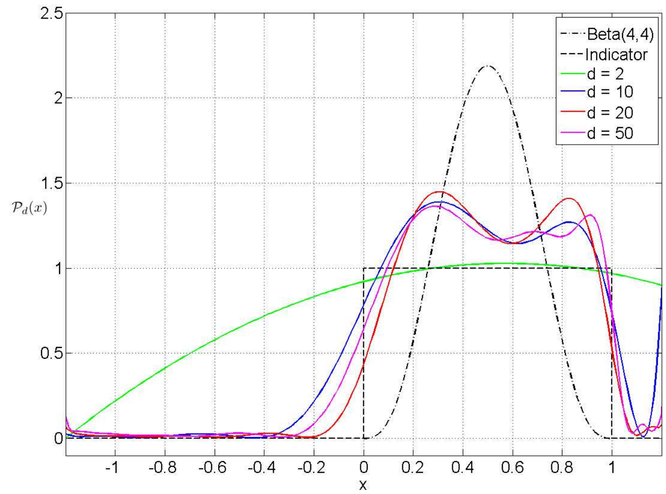

Figure 5.3: Result of SDP in (5.7) For Example 2

Example 2 In this example, we consider a Beta$(4,4)$ probability measure on $[0,1]$. The α-th moment of Beta distribution Beta(a,b) over [0,1] is $y_\alpha = \frac{a+k-1}{(a+b+\alpha-1)} y_{\alpha-1}$ and $y_0 = 1$. We assume that set $\mathcal{B} = [-1.2, 1.2]$, and use the moments up to order $2d$. The obtained results by solving SDP (5.7) with parameters $\omega_h = 1.2$ and $\Delta h = 0.2$ are depicted in Fig 5.3.

This is a more difficult problem than previous ones since, in terms of probability, there is a "smooth transition" from the interior to the exterior of the support set. Nevertheless, if one uses enough moments, one can get a very good approximation of the support.

Example 3 Here, we consider a 2-dimensional example where one wants to approximate the support of a uniform probability measure on $[-0.5, 0.5]^2$. The results obtained by solving SDP (5.7) with parameters $d = 14$, $\omega_h = 1.2$, and $\Delta h = 0.2$ are depicted in Fig 5.4.

5.3.3 Support Reconstruction for Uniform Measures

In this section, we present a modification of our approach aimed specifically at uniform distributions. In the development to follow, we rely on a result in [48] which provides criteria under which polynomials vanish on the boundary of support of the uniform measure of interest. We now elaborate on this.

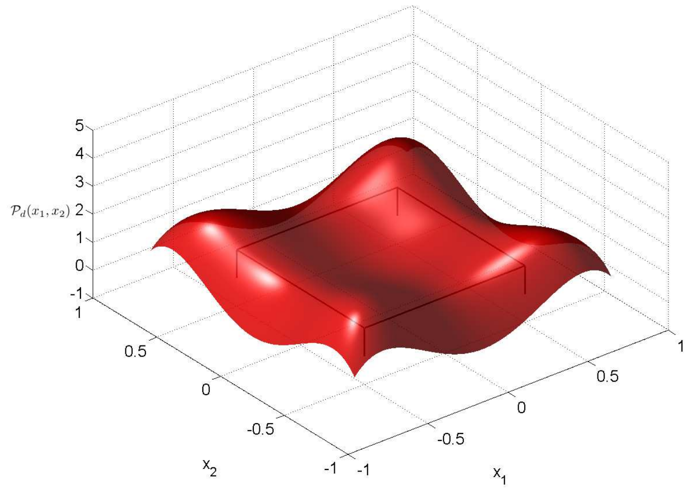

Figure 5.4: Result of SDP in (5.7) For Example 3

Define

$$\bar{M}_r(\mathbf{y})(i,j) = \frac{n+|i|+|j|}{n+|i|} y_{\alpha^{(i)}+\alpha^{(j)}}, \quad 1 \le i,j \le S_{n,r}, \tag{5.8}$$

where $\mathbf{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$ are the moments of the uniform distribution of interest. The results in [48] show that a polynomial $\mathcal{P}(x)$ whose vector of coefficients $\mathbf{p}$ is the eigenvector associated with zero eigenvalue of the matrix $\bar{M}_r$, vanishes on the boundary of support of measure. More precisely, under some technical conditions,

$$\bar{M}_r(\mathbf{y})\mathbf{p} = 0 \Rightarrow \mathcal{P}(x) = 0 \text{ for all } x \in \partial\mathcal{K} \tag{5.9}$$

where, $\partial\mathcal{K}$ denotes the boundary of support set $\mathcal{K}$. However, without any additional constraints, this polynomial can also be zero in the interior of $\mathcal{K}$ and, hence, it might not provide a good estimate of the support.

Nevertheless, one can take advantage of this property and modify our approach as follows.

Problem

$$\mathbf{P}_7^* := \min_{p_\alpha,\sigma_j} \sum_{\alpha=0}^{d} p_\alpha y_{\mathcal{B}_\alpha} - \omega_h h + \omega_M \|\bar{M}_d(\mathbf{y})(\mathbf{p}-1)\|_2 \tag{5.10}$$

$$\text{s.t.} \quad \mathcal{P}_d(x) = \sum_{\|\alpha\|_1 \le d} p_\alpha x^\alpha \tag{5.10a}$$

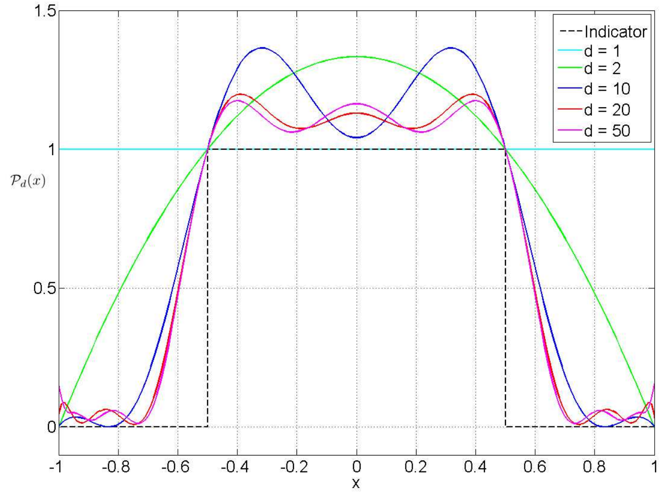

Figure 5.5: Result of SDP in (5.10) For Example 1

$$\mathcal{P}_d(x) = \sigma_0(x) + \sum_{j=1}^{l} \sigma_j(x) g_j(x) \tag{5.10b}$$

$$\sigma_j \in \Sigma^2[x]; j = 0, 1, \ldots, l \tag{5.10c}$$

$$deg(\sigma_0) \leq d_{sos}; \tag{5.10d}$$

$$deg(\sigma_j g_j) \leq d_{sos}; j = 1, 2, \ldots, l \tag{5.10e}$$

$$M_r((\mathcal{P}_d(x) - h)\mathbf{y}) \succcurlyeq 0 \tag{5.10f}$$

$$1 \leq h \leq 1 + \Delta h \tag{5.10g}$$

where, ω_M, ω_h and Δh are positive design parameters, $\mathbf{p} = \{p_\alpha\}_{\alpha \in \mathbb{N}^n}$ denotes the vector of polynomial coefficients and $\|.\|_2$ denotes the l_2 norm.

In fact in (5.10), we aim at "pushing" the coefficients of the polynomial $(\mathcal{P}(x) - 1)$ as close as possible to the null space of $\bar{M}_d$ by minimizing the term $\|\bar{M}_d(\mathbf{p} - 1)\|_2$. In this case obtained polynomial $\mathcal{P}_d(x)$ becomes close to one at the boundary of support while we still aim at having $\mathcal{P}_d$ larger than one inside the support.

To show the effectiveness of proposed method, we reconstruct the support for the measure of Example 1 by solving the SDP (5.10) with parameter $\omega_M = 10$. The obtained result are depicted in Fig 5.5, where semialgebraic set $\mathcal{K}_d = \{x \in \mathbb{R} : \mathcal{P}(x)_d \geq 1\}$ for any polynomial order $d \geq 2$ exactly reconstructs the support of measure.

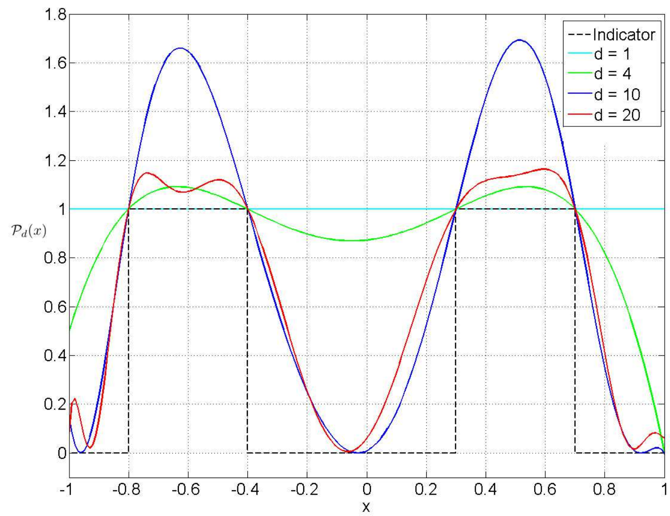

Figure 5.6: Result of SDP in (5.10) For Example 4

Example 4 To further show the effectiveness of our approach, we now consider a uniform distribution with disconnected support. More precisely, we aim at estimating the support of a uniform probability measure over the union of the sets $[-0.8, -0.4]$ and $[0.3, 0.7]$. We assume that $\mathcal{B} = [-1, 1]$ and use moments up to order $2d$. The results obtained by solving SDP (5.10) with parameters $\omega_h = 1.2$, $\omega_M = 10$ and $\Delta h = 0.2$ are depicted in Fig 5.6, where one can see that the semialgebraic set $\mathcal{K}_d = \{x \in \mathbb{R} : \mathcal{P}_d(x) \geq 1\}$ for $d \geq 4$ exactly reconstructs the support of measure.

5.4 Conclusion

In this chapter, we present a novel approach to the problem of uncertainty propagation and reconstruction of support of measures from their moments. A sequence of semidefinite relaxations is provided whose solution converges to the support of the measure of interest. Examples are provided that show that one does obtain a good approximation of support using only a finite number of moments. Further research effort is now being put on developing methods for support reconstruction for specific classes of measures which have provable performance.

Chapter 6

Constrained Volume Optimization Problem

In this chapter, we generalize the chance optimization problems and introduce *constrained volume optimization* where enables us to obtain convex formulation for challenging problems in systems and control [73]. We show that many different problems can be cast as a particular cases of this framework. In constrained volume optimization, we aim at maximizing the volume of a semialgebraic set under some semialgebraic constraints. Building on the theory of measures and moments, a sequence of semidefinite programs are provided, whose sequence of optimal values is shown to converge to the optimal value of the original problem. We show that different problems in the area of systems and control that are known to be nonconvex can be reformulated as special cases of this framework. Particularly, in this work, we address the problems of probabilistic control of uncertain systems as well as inner approximation of region of attraction and invariant sets of polynomial systems. Numerical examples are presented to illustrate the computational performance of the proposed approach.

6.1 Introduction

The purpose of the proposed approach is to develop convex tractable relaxations for different problems in the area of systems and control that are known to be "*hard*". We introduce the so-called constrained volume optimization and show that many challenging problems can be cast as a particular type of this framework. More precisely, we aim at maximizing the volume of a semialgebraic set under some semialgebraic constraints; i.e., let $\mathcal{S}_1(a)$ and $\mathcal{S}_2(a)$ be semialgebraic sets described by set of polynomial inequalities as follows

$$\mathcal{S}_1(a) := \{x \in \chi : \ \mathcal{P}_{1j}(x,a) \geq 0, j = 1, \ldots, o_1\} \tag{6.1}$$

$$\mathcal{S}_2(a) := \{x \in \chi : \ \mathcal{P}_{2j}(x,a) \geq 0, j = 1, \ldots, o_2\} \tag{6.2}$$

where a denotes the vector of design parameters. The objective is to find a parameter vector a such that maximizes the volume of the set $\mathcal{S}_1(a)$ under the constraint $\mathcal{S}_1(a) \subseteq \mathcal{S}_2(a)$. More precisely, we aim at solving the following problem

$$\mathbf{P}^*_{\mathbf{vol}} := \sup_{a \in \mathcal{A}} \ \mathrm{vol}_{\mu_x} \mathcal{S}_1(a), \tag{6.3}$$

$$\text{s.t.} \quad \mathcal{S}_1(a) \subseteq \mathcal{S}_2(a) \tag{6.3a}$$

where, $\mathrm{vol}_{\mu_x}\mathcal{S}_1(a) = \int_{\mathcal{S}_1(a)} d\mu_x$ is the volume of the set $\mathcal{S}_1(a)$ with respect to a given measure μ_x. Many well-known problems can be formulated as a constrained volume optimization problem. As an example, consider the problem of finding the maximal region of attraction (ROA) set for dynamical systems. For a given polynomial system $\dot{x} = f(x)$, maximal ROA set is the largest set of all initial states whose trajectories converge to the origin. This set can be approximated by level sets of a polynomial Lyapunov function $V(x)$. The level set of Lyapunov function $\{x \in \mathbb{R}^n : 0 \leq V(x) \leq 1\}$ is ROA set if it is contained in the region described by $\{x \in \mathbb{R}^n : \dot{V}(x) \leq \epsilon \|x\|_2^2\}$. By characterizing $V(x)$ with a finite order polynomial with unknown coefficients vector a and defining $\mathcal{S}_1(a) := \{x \in \mathbb{R}^n : \ 0 \leq V(x,a) \leq 1\}$ and $\mathcal{S}_2(a) := \left\{x \in \mathbb{R}^n : \ \dot{V}(x,a) \leq \epsilon \|x\|_2^2\right\}$, the problem of finding maximal ROA set can be reformulated as a constrained volume optimization problem.

More details are provided in Section 6.3 where we reformulate different problems in system and control area as constrained volume optimization problems. More precisely, we address the problems of probabilistic control of uncertain systems, inner approximation of region of attraction set and invariant set of polynomial systems, and we also introduce generalized sum of squares problems. The defined constrained volume optimization problem in this work is in general non-convex optimization problem. In this work, relying on measures and moments theory as well as sum of squares theory, we provide a sequence of convex relaxations whose solution converge to the solution of the original problem.

The outline of the chapter is as follows. In Section 6.2, we precisely define the constrained volume optimization problem with respect to semialgebraic constraints. In Section 6.3, some well-known nonconvex problems in system and control area are reformulated as constrained volume optimization problems. In Sections 6.4, we propose equivalent convex problem and sequences of SDP relaxations to the original problem and show that the sequence of optimal solutions to SDP relaxations converges to the solutions of the original problems. In Section 6.5, the problems dual to the convex problems given in Section 6.4 are provided. In Section 6.6, some numerical results are presented to illustrate the numerical performance of the proposed approach, and finally, conclusion is stated in Section 6.7.

6.2 Problem Statement

In this work, we consider *constrained volume optimization problems* defined as follows: Let (χ, Σ_x, μ_x) be a given measure space with Σ_x denoting the Borel σ-algebra of $\chi \subset \mathbb{R}^n$ and μ_x denoting a finite nonnegative Borel measure on Σ_x. Consider semialgebraic sets $\mathcal{S}_1 : \mathbb{R}^n \to \Sigma_x$ and $\mathcal{S}_2 : \mathbb{R}^n \to \Sigma_x$ as follows

$$\mathcal{S}_1(a) := \{x \in \chi : \ \mathcal{P}_{1j}(x, a) \geq 0, j = 1, \ldots, o_1\} \tag{6.4}$$

$$\mathcal{S}_2(a) := \{x \in \chi : \ \mathcal{P}_{2j}(x, a) \geq 0, j = 1, \ldots, o_2\} \tag{6.5}$$

where $\mathcal{P}_{1j} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $j = 1, 2, \ldots, o_1$, and $\mathcal{P}_{2j} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $j = 1, 2, \ldots, o_2$ are given polynomials. We focus on the following problem.

$$\mathbf{P}^*_{\mathbf{vol}} := \sup_{a \in \mathcal{A}} \ \mathrm{vol}_{\mu_x} \mathcal{S}_1(a), \tag{6.6}$$

$$\text{s.t.} \quad \mathcal{S}_1(a) \subseteq \mathcal{S}_2(a) \tag{6.6a}$$

where, $\text{vol}_{\mu_x}\mathcal{S}_1(a) = \int_{\mathcal{S}_1(a)} d\mu_x$ is the volume of the set $\mathcal{S}_1(a)$ with respect to given finite Borel measure μ_x. In the problem (6.6), we are looking for $a \in \mathcal{A} \subset \mathbb{R}^m$, the parameters of sets $\mathcal{S}_1$ and $\mathcal{S}_2$, such that volume of the set $\mathcal{S}_1(a)$ becomes maximum while it is contained in the set $\mathcal{S}_2(a)$.

6.3 Applications in Systems and Control

In this section we focus on some well-known challenging problems in the area of system and control which are, in general, nonconvex and computationally hard. As an first step in the development of convex relaxations of these problems, we reformulate them as a constrained volume optimization problem.

6.3.1 Region of Attraction

Consider a continuous-time system of the form

$$\dot{x} = f(x) \tag{6.7}$$

where, $f : \mathbb{R}^n \to \mathbb{R}^n$ is a polynomial function, $x \in \chi \subset \mathbb{R}^n$ are system states and χ is compact that contains the origin. Let the origin $x = 0$ be an asymptotically stable equilibrium point for the system. The region of attraction (ROA) set $\mathcal{R}_x \subseteq \chi$ is defined as largest set of all initial states whose trajectories converge to the origin.

For the system in (6.7), assume there exist a function $V(x)$ such that

$$V(0) = 0, V(x) > 0 \text{ for x} \neq 0 \tag{6.8}$$

The level set defined as $\mathcal{R} = \{x \in \chi : V(x) \leq r\}$ is an inner approximation of ROA if $\dot{V}(x) < 0$ for all $x \in \mathcal{R}$ and $\dot{V}(x) = 0$ for $x = 0$ [103]. In this case function $V(x)$ is a Lyapunov function for the system in (6.7). We assume that polynomial system in (6.7) admits a polynomial Lyapunov function (see [104] for discussion on existence of a polynomial Lyapunov function). Hence, we can describe it as a finite order polynomial $V(x) = \sum_{\|i\|_1 \leq d} a_i x^i \in \mathbb{R}_d[s]$, where $a \in \mathcal{A} \subset \mathbb{R}^{S_{n,d}}$ is a vector of unknown coefficients, ([105, 106]). The following optimization problem can be used to find $V(x)$ and corresponding inner approximation of maximal ROA. For a given system in (6.7) and given compact sets χ and $\mathcal{A}$, solve

$$\mathbf{P}^*_{\mathbf{ROA}} := \max_{a\in\mathcal{A}} \mathrm{vol}_{\mu_x}(\mathcal{R}) \tag{6.9}$$

$$\text{s.t.} \quad V(x) = \sum_{\|i\|_1\le d} a_i x^i,\ V(0) = 0 \tag{6.9a}$$

$$V(x) > 0,\ \text{for all } x \neq 0 \tag{6.9b}$$

$$\mathcal{R} = \{x \in \chi :\ 0 \le V(x) \le r\} \tag{6.9c}$$

$$\mathcal{R} \subseteq \left\{x \in \chi : \dot{V}(x) \le -\epsilon_r \|x\|_2^2\right\} \tag{6.9d}$$

where, $\mathrm{vol}_{\mu_x}(\mathcal{R}) = \int_{\mathcal{R}} d\mu_x$ is the volume of the set $\mathcal{R}$ with respect to Lebesgue measure μ_x supported on χ, $\|.\|_2$ is l_2 norm and $r > 0$ and $\epsilon_r > 0$ are known constants. By solving problem in (6.9), we are in fact looking for a Lyapunov function $V(x)$ among the space of polynomial functions of order at most d. By defining the sets $\mathcal{S}_1(a) = \{x \in \chi :\ 0 \le V(x,a) \le r\}$ and $\mathcal{S}_2(a) = \{x \in \chi : \dot{V}(x,a) \le -\epsilon_r\|x\|_2^2\}$, problem in (6.9) can be restated as volume optimization problem in (6.6). With the same reasoning, one can extend the problem in (6.9) for discrete-time systems $x_{k+1} = f(x_k)$ by replacing the derivative of Lyapunov function $\dot{V}(x)$ with the difference Lyapunov function $\triangle V(x) = V(x_{k+1}) - V(x_k)$.

6.3.2 Maximal Invariant Set

Consider a discrete time system

$$x_{k+1} = f(x_k) \tag{6.10}$$

where, $f : \mathbb{R}^n \to \mathbb{R}^n$ is a polynomial function and $x_k \in \chi_{ext} \subset \mathbb{R}^n$ are system states. Given a compact set $\chi \subset \chi_{ext}$, the set $\mathcal{V} \subset \chi$ is *robustly invariant* if

$$f(x) \in \mathcal{V},\ \text{for all } x \in \mathcal{V} \tag{6.11}$$

Hence, *maximal invariant set* is the maximal set of all initial states whose trajectories remains inside the set. The following statement holds true for robustly invariant sets.

Consider the set $\mathcal{V} = \{x \in \chi : \ \mathcal{P}(x) \geq 0\}$, for bounded above function $\mathcal{P}(x)$. The set $\mathcal{V}$ is an robustly invariant set for dynamical system above if $\mathcal{P}(f(x)) \geqslant 0$ for all $x \in \mathcal{V}$.

In order to find a polynomial approximation of maximal robustly invariant set for dynamical system (6.10), we characterize the function $\mathcal{P}(x)$ with finite order polynomial as $\mathcal{P}(x) = \sum_{\|i\|_1 \leq d} a_i x^i \in \mathbb{R}_d[x]$, where $a \in \mathcal{A} \subset \mathbb{R}^{S_{n,d}}$ is vector of unknown coefficients. Then, we consider following optimization problem to obtain unknown coefficients.

Assume that the given compact set χ can be described as $\chi = \{x \in \mathbb{R}^n : \mathcal{P}_\chi(x) \geq 0\} \subset \mathbb{R}^n$, for some polynomial $\mathcal{P}_\chi(x)$. Then for a given system in (6.10) and given compact sets $\chi \subset \mathbb{R}^n$ and $\mathcal{A} \subset \mathbb{R}^{S_{n,d}}$, solve

$$\mathbf{P}^*_{\mathbf{INV}} := \max_{a \in \mathcal{A}} \mathrm{vol}_{\mu_x}(\mathcal{V}) \tag{6.12}$$

$$\text{s.t.} \quad \mathcal{P}(x) = \sum_{\|i\|_1 \leq d} a_i x^i, \tag{6.12a}$$

$$\mathcal{V} = \{x \in \chi : \ \mathcal{P}(x) \geq 0\} \tag{6.12b}$$

$$\mathcal{V} \subset \{x \in \chi : \mathcal{P}(f(x)) \geq 0\} \tag{6.12c}$$

where, $\mathrm{vol}_{\mu_x}(\mathcal{V}) = \int_{\mathcal{V}} d\mu_x$ is the volume of the set $\mathcal{V}$ with respect to Lebesgue measure μ_x supported on χ_{ext}.

By defining the sets $\mathcal{S}_1(a) = \{x : \mathcal{P}(x, a) \geq 0, \ \mathcal{P}_\chi(x) \geq 0\}$ and $\mathcal{S}_2(a) = \{x : \mathcal{P}(f(x), a) \geq 0, \ \mathcal{P}_\chi(f(x)) \geq 0\}$, problem (6.12) can be restated as volume optimization problem (6.6).

6.3.3 Generalized Sum of Squares Problem

Using SOS representation Lemma 7, we can find a polynomial that is strictly positive on the given semialgebraic set. Many different problems in system and control can be reformulated as SOS representation problem which results in semidefinite programming problems.

In this work, we generalize the notion of SOS and introduce a new class of SOS problems where enable us to find a strictly positive polynomial on some unknown semialgebraic sets. More precisely, we define the *Generalized Sum of Squares* (GSOS) problem as follow.

Generalized Sum of Squares: Consider polynomial $\mathcal{P}(x,a) \in \mathbb{R}_d[x]$ and semi-algebraic set $\mathcal{S}_1(a) := \{x \in \chi : \ \mathcal{P}_{1j}(x,a) \geq 0, j = 1,\ldots,o_1\}$ where $a \in \mathcal{A} \subset \mathbb{R}^m$ are unknown parameters. We aim at finding parameters a such that polynomial $\mathcal{P}(x,a)$ is strictly positive on the set $\mathcal{S}_1(a)$.

This problem can be restated as a volume optimization problem in (6.6) by defining the set $\mathcal{S}_2(a) := \{x \in \chi : \ \mathcal{P}(x,a) \geq 0\}$. As an example of GSOS, we could consider the problem of finding polynomial Lyapunov function as in section 6.3.1 where we are looking for a polynomial $\epsilon||x||_2^2 - \dot{V}(x)$ to be strictly positive on the set $\{x \in \chi : \ 0 \leq V(x) \leq r\}$, where $\epsilon < 0$ and $r > 0$.

6.4 Equivalent Convex Problem on Measures and Moments

In this section, we first provide an equivalent infinite linear program (LP) on finite nonnegative Borel measures to solve the constrained volume optimization problem in (6.6). Then, we provide a finite dimensional semidefinite program (SDP) in moment space. Consider the sets of volume optimization problem $\mathcal{S}_1$ and $\mathcal{S}_2$ defined in (6.4) and (6.5). Given polynomials $\mathcal{P}_{1j} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ for $j = 1,2,\ldots,o_1$, and polynomials $\mathcal{P}_{2j} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ for $j = 1,2,\ldots,o_2$, we define following sets as

$$\mathcal{K}_1 := \{(x,a) : \mathcal{P}_{1j}(x,a) \geq 0, j = 1,\ldots,o_1\} \tag{6.13}$$

$$\mathcal{K}_2 := \{(x,a) : \mathcal{P}_{2j}(x,a) \geq 0, j = 1,\ldots,o_2\} \tag{6.14}$$

Assumption 3. *Assume that $\mathcal{K}_1$ and $\mathcal{K}_2$ satisfy Putinar's property represented in Chapter 2. This implies that sets $\mathcal{K}_1$ and $\mathcal{K}_2$ are compact; Hence the projections of the sets $\mathcal{K}_1$ and $\mathcal{K}_2$ onto x-coordinates and onto a-coordinates are also compact. Also, we assume without loss of generality that $x \in \chi := [-1,1]^n$ and $a \in \mathcal{A} := [-1,1]^m$ and the set $(\chi \times \mathcal{A}) \setminus \mathcal{K}_1 = \{(x,a) \in \chi \times \mathcal{A} : (x,a) \notin \mathcal{K}_1\}$ has a nonempty interior.*

6.4.1 Linear Program on Measures

As an intermediate step in the development of finite convex relaxations of the original problem in (6.6), a related infinite dimensional convex problem on measures is provided as follows. Let μ_x be the given measure supported on χ defined in constrained volume problem (6.6). The sets $\mathcal{K}_1$ and $\mathcal{K}_2$ are defined as (6.13) and (6.14) and let $\overline{\mathcal{K}_1}$ be the complement of the set $\mathcal{K}_1$. Consider the following problem on measures

$$\mathbf{P}^*_{\mathbf{measure}} := \sup_{\mu,\mu_a} \int d\mu, \tag{6.15}$$

$$\text{s.t. } \mu \preccurlyeq \mu_a \times \mu_x, \tag{6.15a}$$

$$\mu_a \text{ is a probability measure, } \mu_a \in \mathcal{M}_+(\mathcal{A}), \tag{6.15b}$$

$$\mu_a \times \mu_x \in \mathcal{M}_+(\overline{\mathcal{K}_1} \cup \mathcal{K}_2), \quad \mu \in \mathcal{M}_+(\mathcal{K}_1). \tag{6.15c}$$

In the problem (6.15), we are looking for measures μ and μ_a supported on $\mathcal{K}_1$ and $\mathcal{A}$, such that μ is bounded by product measure $\mu_a \times \mu_x$ supported on $\overline{\mathcal{K}_1} \cup \mathcal{K}_2$. The following theorem, shows the equivalency of the problem in (6.15) and the original volume problem in (6.6).

Theorem 25. *The optimization problem in* (6.6) *is equivalent to the infinite LP in* (6.15) *in the following sense:*

i) The optimal values are the same, i.e., $\mathbf{P}^*_{\mathbf{vol}} = \mathbf{P}^*_{\mathbf{measure}}$.

ii) If an optimal solution to (6.15) *exists, call it* μ_a^*, *then any* $a^* \in supp(\mu_a^*)$ *is an optimal solution to* (6.6).

iii) If an optimal solution to (6.6) *exists, call it* a^*, *then* $\mu_a = \delta_{a^*}$, *Dirac measure at* a^*, *and* $\mu = \delta_{a^*} \times \mu_x$ *is an optimal solution to* (6.15).

Proof. See Appendix A. □

Problem (6.15), requires information of the set $\mathcal{K}_1$ and its complement $\overline{\mathcal{K}_1}$. From a numerical implementation point of view, this results on an ill conditioned problem. To solve problem (6.15), we first need to obtain finite relaxations that provide an outer approximation of the sets of problem (6.15). The outer approximation of the sets $\mathcal{K}_1$ and $\overline{\mathcal{K}_1}$ intersect and thus poor performance of the solution is observed. To solve this, we modify the provided problem on measures as follows.

We first aim at finding the approximation of the constraint of the original problem (6.6). The set that approximates the constraint of volume problem includes all design parameters $a \in \mathcal{A}$ for which the set $\mathcal{S}_1(a)$ is a subset of the set $\mathcal{S}_2(a)$. Next, we look for parameter a inside the obtained set that maximizes the volume of the set $\mathcal{S}_1(a)$.

Let $\mathcal{A}_{\mathcal{F}}$ be the set of all parameters $a \in \mathcal{A}$ for which the set $\mathcal{S}_1(a)$ is a subset of the set $\mathcal{S}_2(a)$, i.e.

$$\mathcal{A}_{\mathcal{F}} := \{a \in \mathcal{A} : \mathcal{S}_1(a) \subseteq \mathcal{S}_2(a)\} \tag{6.16}$$

To obtain the approximation of the set $\mathcal{A}_{\mathcal{F}}$, consider the following infinite LP on continuous functions:

$$\mathbf{P}^*_{\mathcal{A}_\mathbf{f}} := \inf_{f \in \mathcal{C}(a)} \int_{\mathcal{A}} f(a) d\mu_{\mathcal{A}}, \tag{6.17}$$

$$\text{s.t.} \quad f(a) \geq 1 \text{ on } \mathcal{K}_1 \cap \overline{\mathcal{K}_2}, \tag{6.17a}$$

$$f(a) \geq 0 \text{ on } \chi \times \mathcal{A}. \tag{6.17b}$$

where, $f \in \mathcal{C}(a)$ and $\overline{\mathcal{K}_2}$ is the complement of the set $\mathcal{K}_2$.

Then, following Theorem holds true.

Theorem 26. *Let $\mu_{\mathcal{A}}$ be the Lebesgue measure of the set $\mathcal{A}$. Also, let $\mathcal{I}_{\overline{\mathcal{A}_{\mathcal{F}}}}$ be the indicator function of the set $\overline{\mathcal{A}_{\mathcal{F}}} := \{a \in \mathcal{A} : \mathcal{S}_1(a) \not\subseteq \mathcal{S}_2(a)\}$; i.e., $\mathcal{I}_{\overline{\mathcal{A}_{\mathcal{F}}}}(a) = 1$ if $a \in \overline{\mathcal{A}_{\mathcal{F}}}$ and 0 otherwise. Then*

i) There is a sequence of continuous functions $f_i(a)$ to Problem (6.17) that converges to the $\mathcal{I}_{\overline{\mathcal{A}_{\mathcal{F}}}}$ in L_1-norm sense, i.e., $lim_{i\to\infty} \int_{\mathcal{A}} |f_i(a) - \mathcal{I}_{\overline{\mathcal{A}_{\mathcal{F}}}}(a)| da = 0$.

ii) The set $\mathcal{A}_{f_i} = \{a \in \mathcal{A} : \ f_i(a) < 1\}$ converges to the set $\mathcal{A}_{\mathcal{F}}$, i.e., $lim_{i\to\infty} \mu_{\mathcal{A}}(\mathcal{A}_{\mathcal{F}} - \mathcal{A}_{f_i}) = 0$, and $\mathcal{A}_{f_i} \subseteq \mathcal{A}_{\mathcal{F}}$.

Proof. See Appendix B. □

Now, to obtain an approximate of the solution of the original volume problem (6.6), consider infinite LP on measures as follows. Let, μ_x be the given measure supported on χ as in problem (6.6) and $\mathcal{K}_1$ be the set as in (6.13). Then, consider the following problem

$$\mathbf{P}^*_{\mathbf{f_i}} := \sup_{\mu,\mu_a} \int d\mu, \tag{6.18}$$

$$\text{s.t.} \quad \mu \preccurlyeq \mu_a \times \mu_x, \tag{6.18a}$$

$$\mu_a \text{ is a probability measure}, \tag{6.18b}$$

$$\mathcal{A}_{f_i} = \{a \in \mathcal{A} : \; f_i(a) < 1\}, \tag{6.18c}$$

$$\mu_a \in \mathcal{M}_+(\mathcal{A}_{f_i}), \quad \mu \in \mathcal{M}_+(\mathcal{K}_1). \tag{6.18d}$$

Now, following theorem establishes the equivalence of volume optimization problem in (6.6) and infinite LP in (6.18).

Theorem 27. *Let,* $(\mu_a^*(f_i), \mu^*(f_i), \mathbf{P}^*_{\mathbf{f_i}})$ *be an optimal solution and value of the LP in* (6.18) *for the obtained function* f_i *and the set* $\mathcal{A}_{f_i}$ *solving Problem* (6.17). *Also, assume that volume problem* (6.6) *has a unique optimal solution and value* $(a^*, \mathbf{P}^*_{\mathbf{vol}})$. *As the set* $\mathcal{A}_{f_i} = \{a \in \mathcal{A} : \; f_i(a) < 1\}$ *defined in Theorem* (26) *converges to the set* $\mathcal{A}_{\mathcal{F}}$, *we have the following results:*

i) The optimal value $\mathbf{P}^*_{\mathbf{f_i}}$ *converges to the* $\mathbf{P}^*_{\mathbf{vol}}$.

ii) Measures $\mu_a^*(f_i)$ *and* $\mu^*(f_i)$ *converge to* $\mu_a = \delta_{a^*}$, *Dirac measure at* a^*, *and* $\mu = \delta_{a^*} \times \mu_x$, *respectively.*

iii) Any point in the support of the measure $\mu_a^*(f_i)$, *i.e.,* $a_i \in supp(\mu_a^*(f_i))$, *converges to the* a^*.

Proof. See Appendix C. □

In the next section, we provide the tractable finite relaxation to infinite LP in (6.18) and (6.17).

6.4.2 Finite Semidefinite Programming on Moments

In this section, we provide a finite dimensional SDP whose feasible region is defined over real sequences. We show that the corresponding sequence of optimal solutions arbitrarily approximate the optimal solution of (6.18). Unlike problem (6.18) in which we are looking for measures, in the provided SDP formulation, we aim at finding moment sequences corresponding to measures that are optimal to (6.18).

For this purpose, we first need to obtain the semialgebraic approximation of the set $\mathcal{A}_{\mathcal{F}}$ in (6.16), the set of all parameters $a \in \mathcal{A}$ for which the set $\mathcal{S}_1(a)$ is a subset of the set $\mathcal{S}_2(a)$. In the previous section, the continuous function f and the infinite LP in (6.17) are used to obtain an inner approximation of the set $\mathcal{A}_{\mathcal{F}}$. Here, we use polynomial $\mathcal{P}^d_{\mathcal{A}}(a) \in \mathbb{R}_d[a]$ and finite finite SDP problem below

$$\mathbf{P}^*_{\mathcal{A}_\mathbf{d}} := \min_{\mathcal{P}^d_{\mathcal{A}}(a) \in \mathbb{R}_d[a]} \int_{\mathcal{A}} \mathcal{P}^d_{\mathcal{A}}(a) d\mu_{\mathcal{A}}, \tag{6.19}$$

$$\text{s.t. } \mathcal{P}^d_{\mathcal{A}}(a) - 1 \in \mathcal{QM}_i \left(\{\mathcal{P}_{1j}\}_{j=1}^{o_1}, -\mathcal{P}_{2i} \right), i = 1, \ldots, o_2 \tag{6.19a}$$

$$\mathcal{P}^d_{\mathcal{A}}(a) \in \mathcal{QM} \left(\{(1 - x_i^2)\}_{i=1}^n, \{(1 - a_i^2)\}_{i=1}^m \right). \tag{6.19b}$$

where, $\mu_{\mathcal{A}}$ is the Lebesgue measure over the set $\mathcal{A}$ and d is the order of polynomial $\mathcal{P}^d_{\mathcal{A}}(a)$. $\mathcal{QM}_i$ and $\mathcal{QM}$ as defined in (2.1) are the quadratic modules generated by polynomials of set $\mathcal{K}_1 \cap \overline{\mathcal{K}_2} = \{(x, a) : \cup_{i=1}^{o_2} \{-\mathcal{P}_{2i} > 0, \mathcal{P}_{1j} \geq 0, j = 1, \ldots, o_1\}\}$, and polynomials of hyper cube $\chi \times \mathcal{A}$, respectively. According to the Lemma 7, constraints (6.19a) and (6.19b) imply that polynomials $(\mathcal{P}^d_{\mathcal{A}}(a) - 1)$ and $\mathcal{P}^d_{\mathcal{A}}(a)$ are positive on the sets $\mathcal{K}_1 \cap \overline{\mathcal{K}_2}$ and hyper cube $\chi \times \mathcal{A}$, respectively. Problem in (6.19) is a SDP, where objective function is a weighted summation of coefficients of polynomial $\mathcal{P}^d_{\mathcal{A}}(a)$ with respect to the moments of Lebesgue measure $\mu_{\mathcal{A}}$ and constraints are convex linear matrix inequalities in terms of coefficients of polynomial $\mathcal{P}^d_{\mathcal{A}}(a)$.

The following theorem hold true for the problems (6.19).

Theorem 28. *Let $\mathcal{P}^d_{\mathcal{A}}(a)$ be an optimal solution of SDP* (6.19) *and consider the set:*

$$\mathcal{A}_d = \left\{ a \in \mathcal{A} : \ \mathcal{P}^d_{\mathcal{A}}(a) < 1 \right\} \tag{6.20}$$

Also, let $\mathcal{I}_{\overline{\mathcal{A}_{\mathcal{F}}}}$ be the indicator function of the set $\overline{\mathcal{A}_{\mathcal{F}}} := \{a \in \mathcal{A} : \mathcal{S}_1(a) \not\subseteq \mathcal{S}_2(a)\}$. Then,

i) The sequence of optimal solutions to the finite SDP in (6.19) *converges to the $\mathcal{I}_{\overline{\mathcal{A}_{\mathcal{F}}}}$ in L_1-norm sense as $d \to \infty$, i.e., $lim_{d\to\infty} \int_{\mathcal{A}} |\mathcal{P}^d_{\mathcal{A}}(a) - \mathcal{I}_{\overline{\mathcal{A}_{\mathcal{F}}}}(a)| da = 0$.*

ii) The set $\mathcal{A}_d$ converges to the set $\mathcal{A}_{\mathcal{F}}$ in (6.16)*, i.e., $lim_{d\to\infty} \mu_{\mathcal{A}}(\mathcal{A}_{\mathcal{F}} - \mathcal{A}_d) = 0$, and $\mathcal{A}_d \subseteq \mathcal{A}_{\mathcal{F}}$.*

Proof. See Appendix D. □

Given that the indicator function $\mathcal{I}_{\overline{\mathcal{A}_{\mathcal{F}}}}$ can be approximated by sequence of polynomials of increasing order $\mathcal{P}^d_{\mathcal{A}}$ as in Theorems 28, we restrict the continuous function f_i in problem (6.18) to be polynomials. Hence, we can approximate the optimal solution of infinite LP (6.18) on measures with finite dimensional SDP on moments. In order to have tractable approximations to the infinite dimensional LP in (6.18), we consider the SDP (6.21), known as Lasserre's hierarchy [68], where $\mathbf{y_x} := \{y_{x_\beta}\}_{\beta\in\mathbb{N}^n_{2r}}$ and $\mathbf{y_a} = \{y_{a_\alpha}\}_{\alpha\in\mathbb{N}^m_{2r}}$ are the truncated moment sequence of measures μ_x and μ_a.

$$\mathbf{P^*_r} := \sup_{\mathbf{y}\in\mathbb{R}^{S_{n+m,2r}},\ \mathbf{y_a}\in\mathbb{R}^{S_{m,2r}}} (\mathbf{y})_{\mathbf{0}}, \tag{6.21}$$

$$\text{s.t.}\quad M_r(\mathbf{y}) \succcurlyeq 0,\ M_{r-r_j}(\mathbf{y};\mathcal{P}_{1j}) \succcurlyeq 0,\ j=1,\dots,o_1, \tag{6.21a}$$

$$(\mathbf{y_a})_{\mathbf{0}} = 1, \tag{6.21b}$$

$$M_r(\mathbf{y_a}) \succcurlyeq 0,\ M_{r-r_a}(\mathbf{y_a}; 1-\mathcal{P}^d_{\mathcal{A}}(a)) \succcurlyeq 0,\ M_{r-1}(\mathbf{y_a}; 1-a_i^2) \succcurlyeq 0, i=1,...,m \tag{6.21c}$$

$$M_r(\mathbf{y_a}\times\mathbf{y_x}-\mathbf{y}) \succcurlyeq 0. \tag{6.21d}$$

In (6.21), $(\mathbf{y})_{\mathbf{0}}$ is the first element of the truncated moment sequence of measure μ, $r\in\mathbb{Z}_+$ is relaxation order of matrices, d_j is the degree of polynomial $\mathcal{P}_{1j}$ in the set $\mathcal{S}_1$, $r_j := \left\lceil \frac{d_j}{2} \right\rceil$ for all $1\le j\le o_1$. Sequence $\mathbf{y_a}\times\mathbf{y_x} = \bar{\mathbf{y}}$ is truncated moment sequence of measure $\mu_a\times\mu_x$ such that $(\bar{\mathbf{y}})_\theta = (\mathbf{y_a})_\alpha(\mathbf{y_x})_\beta$ for all $\theta=(\alpha,\beta)\in\mathbb{N}^{n+m}_{2r}$. Matrices $M_r(\mathbf{y})$, $M_r(\mathbf{y_a})$, and $M_r(\mathbf{y_a}\times\mathbf{y_x}-\mathbf{y})$ are moment matrices constructed by moment sequences $\mathbf{y}$, $\mathbf{y_a}$, and $\mathbf{y_a}\times\mathbf{y_x}-\mathbf{y}$, respectively. Also, $M_{r-r_j}(\mathbf{y};\mathcal{P}_{1j}), j=1,\dots,o_1$ and $\{\ M_{r-1}(\mathbf{y_a}; 1-a_i^2)\}_{i=1}^m$ are localization matrices constructed by polynomials of the set $\mathcal{K}_1$ and hyper cube $\chi\times\mathcal{A}$, respectively. Finally, $M_{r-r_a}(\mathbf{y_a}; 1-\mathcal{P}^d_{\mathcal{A}}(a))$ is localization matrix constructed by polynomial of the set $\mathcal{A}_d$ in (6.20), i.e., $(1-\mathcal{P}^d_{\mathcal{A}}(a))$, where $\mathcal{P}^d_{\mathcal{A}}(a)$ is an optimal solution of SDP (6.19).

Remark 6.4.1. *To be able to work with closed sets $\mathcal{A}_d$ in* (6.20) *and $\mathcal{K}_1\cap\overline{\mathcal{K}_2}$ which are used in constraints* (6.21c) *and* (6.19a)*, we use positive small $\epsilon_{\mathcal{A}},\epsilon_{\mathcal{K}}\to 0$ and also to satisfy the Putinar's property, we add the polynomial $\sqrt{m}^2-\|a\|^2\ge 0$, i.e.,*

$$\mathcal{A}_d = \left\{a\in\mathcal{A}:\ \mathcal{P}^d_{\mathcal{A}}(a)\le 1-\epsilon_{\mathcal{A}}, \sqrt{m}^2-\|a\|^2\ge 0\right\}$$

and

$$\mathcal{K}_1 \cap \overline{\mathcal{K}_2} = \{(x,a) : \cup_{i=1}^{o_2}\{-\mathcal{P}_{2i} \geq \epsilon_{\mathcal{K}}, \mathcal{P}_{1j} \geq 0, j = 1, \ldots, o_1\}\}$$

.

6.4.3 Illustrative Example

In this section, we present a simple example of constrained volume optimization problem in (6.6) and show how the proposed finite SDP in (6.21) effectively works. For illustrative purposes, the provided example is low dimensional and consists of sets described by polynomials in $x \in \chi \subset \mathbb{R}$ and parameter $a \in \mathcal{A} \subset \mathbb{R}$. We consider the volume optimization problem in (6.6) with following sets

$$\mathcal{S}_1(a) := \left\{x \in \chi :\ 0.25 - a^2 - x^2 \geq 0\right\} \tag{6.22}$$

$$\mathcal{S}_2(a) := \left\{x \in \chi :\ 0.09 - a^2 - 0.8a - x^2 \geq 0\right\} \tag{6.23}$$

where, $\chi = [-1, 1]$ and given measure μ_x is the Lebesgue measure supported on $\mathcal{A} = [-1, 1]$.

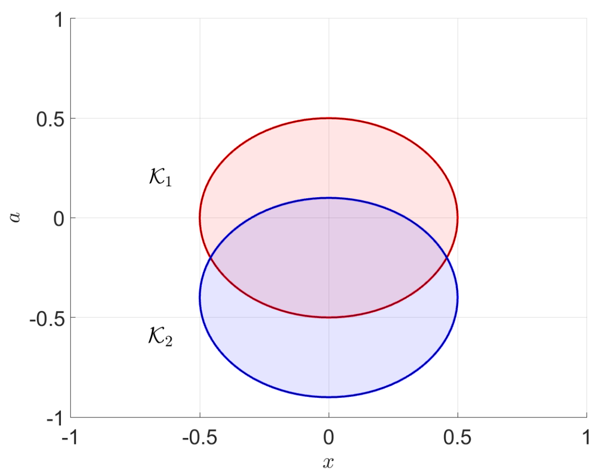

Figure 6.1: Sets $\mathcal{K}_1$ and $\mathcal{K}_2$

Based on given sets $\mathcal{S}_1$ and $\mathcal{S}_2$, we define the sets $\mathcal{K}_1 := \left\{(x,a) : 0.25 - a^2 - x^2 \geq 0\right\}$ and $\mathcal{K}_2 := \left\{(x,a) : 0.09 - a^2 - 0.8a - x^2 \geq 0\right\}$. Figure 6.1 displays the sets $\mathcal{K}_1$ and $\mathcal{K}_2$.

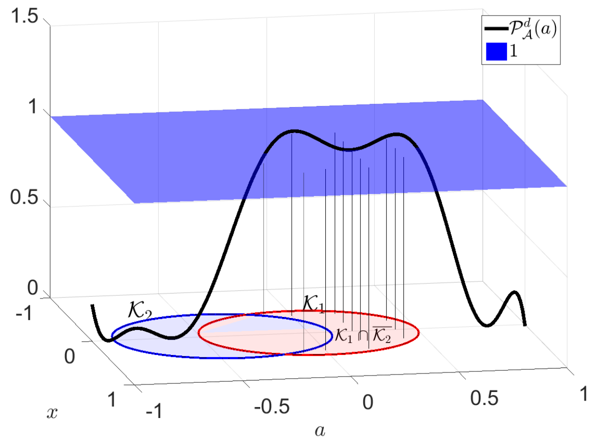

Figure 6.2: Polynomial $\mathcal{P}_{\mathcal{A}}^d(a)$ obtained by SDP (6.19) for $d = 7$

To obtain an approximate solution of constrained volume optimization problem, we solve finite SDPs in (6.21) and (6.19). First, we solve the SDP in (6.19) to obtain the polynomial $\mathcal{P}_{\mathcal{A}}^d(a)$. To this, we use Yalmip which is a MATLAB-based toolbox for polynomial and SOS optimization [96]. Figure 6.2 displays polynomial $\mathcal{P}_{\mathcal{A}}^d(a)$ obtained by SDP (6.19) for polynomial order $d = 7$. As constraints of SDP (6.19), $\mathcal{P}_{\mathcal{A}}^d(a)$ is greater than 1 on $\mathcal{K}_1 \cap \overline{\mathcal{K}_2}$ and is positive on $\chi \times \mathcal{A} = [-1, 1]^2$. Hence, based on Theorem 28 the set $\mathcal{A}_d = \left\{a \in \mathbb{R} : \ \mathcal{P}_{\mathcal{A}}^7(a) \leq 1 - \epsilon_{\mathcal{A}}, 1 - a^2 \geq 0, \epsilon_{\mathcal{A}} = 0.05\right\}$ is an inner approximation of the set $\mathcal{A}_{\mathcal{F}}$, the set of all parameter $a \in \mathcal{A}$ that set $\mathcal{S}_1$ is subset of $\mathcal{S}_2$. Clearly, based on the Figure 6.1, $\mathcal{A}_{\mathcal{F}} = (-1 \leq a \leq -0.2) \cup (0.5 \leq a \leq 1)$, where $\mathcal{S}_1(a) \subseteq \mathcal{S}_2(a)$ for $a \in [-0.5, -0.2]$, $\mathcal{S}_1(a) = \varnothing \subseteq \mathcal{S}_2(a)$ for $a \in [-0.9, -0.5)$, and $\mathcal{S}_1(a) = \mathcal{S}_2(a) = \varnothing$ for $a \in [-1, -0.9) \cup (0.5, 1]$. Figure 6.3 displays polynomial $\mathcal{P}_{\mathcal{A}}^d(a)$ obtained by SDP (6.19) for different polynomial orders $d = 2, 4, 6, 7$ and also the sets $\mathcal{A}_d$.

We take $\mathcal{P}_{\mathcal{A}}^7(a)$ and solve SDP in (6.21). Based on moments of Lebesgue measure μ_x on $\mathcal{A} = [-1, 1]$ as $(y_x)_\alpha = \frac{1}{(\alpha+1)}(1^{\alpha+1} - (-1)^{\alpha+1})$, we construct the matrices in constraints of SDP (6.21) in terms of unknown moment vectors $\mathbf{y} \in \mathbb{R}^{S_{2,2r}}$ and $\mathbf{y_a} \in \mathbb{R}^{S_{1,2r}}$. Since the order of maximum degree of polynomials in $\mathcal{S}_1$ and $\mathcal{A}_d$ is 7, the minimum relaxation order for SDP (6.21) is $r = 4$, which requires the moments up to order 8. The SDP in (6.21) with $r = 6$ is solved using GloptiPoly. Based on obtained solution for moment vectors, we approximate the solution to volume problem a with $y_{a1} = -0.2050$ and estimate the optimal volume $\mathbf{P}_{\mathbf{vol}}^*$ with $\mathbf{P_r} = y_{00} = 1.239$. Clearly,

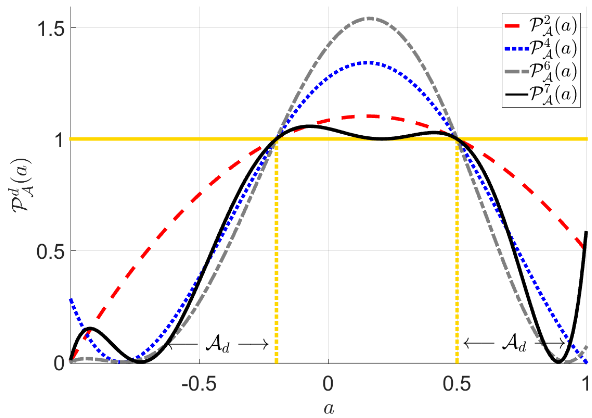

Figure 6.3: Polynomial $\mathcal{P}^d_{\mathcal{A}}(a)$ obtained by SDP (6.19) for $d = 2, 4, 6, 7$

for obtained $a = -0.2050$, the set $\mathcal{S}_1 = \{x : 0.4561^2 - x^2 \geq 0\}$ is a subset of $\mathcal{S}_2 = \{x : 0.4604^2 - x^2 \geq 0\}$. Based on Figure 6.1, the true solution for the volume optimization problem is $a^* = -0.2$ with volume $\mathbf{P}^*_{\mathbf{vol}} = 0.9165$. To obtain better estimates of the optimum volume, one needs to increase the relaxation order r. Also, Section 3.2.3, where we provided some methods to improve the estimated volume of the semialgebraic sets in the similar setup.

6.5 Dual Convex Problem on Function Space

In this section, we provide an infinite LP on continuous functions which is dual to the infinite LP on measure in (6.18). The provided dual problem gives a new insight on solving the volume problem. Also, from computational efficiency perspective we can take advantage of polynomial convex optimization techniques. For instance, to handle large scale SDPs, one can employ DSOS optimization technique where relies on linear and second order cone programming ([107, 108]).

To obtain a dual problem to the infinite LP in (6.18), let $\mathcal{C}(\chi \times \mathcal{A})$ be the Banach space of continuous functions on $\chi \times \mathcal{A}$. Then, Lagrangian dual of (6.18) is:

$$\mathbf{P}^*_{\mathbf{Dual}} := \inf_{\beta \in \mathbb{R}, \mathcal{W} \in \mathcal{C}(\chi \times \mathcal{A})} \beta, \tag{6.24}$$

$$\text{s.t.} \quad \mathcal{W}(x, a) \geq 1 \quad \text{on} \quad \mathcal{K}_1, \tag{6.24a}$$

$$\beta - \int_{\chi} \mathcal{W}(x, a) d\mu_x \geq 0 \quad \text{on} \quad \mathcal{A}_{f_i}, \tag{6.24b}$$

$$\mathcal{W}(x,a) \geq 0,\ \ \beta \geq 0. \tag{6.24c}$$

where, $\mathcal{K}_1$ is defined as (6.13), μ_x is a given Borel measure, and $\mathcal{A}_{f_i}$ is a set defined in (6.18c). We can interpret the obtained dual problem as follow. If we assume that a is given, then the optimal solution for $\mathcal{W}(x,a)$ is the indicator function of the set $\mathcal{K}_1$ and the optimal value $\mathbf{P}^*_{\mathbf{Dual}}$ is the volume of the set $\mathcal{K}_1$, i.e., $\mathbf{P}^*_{\mathbf{Dual}} = \beta = \int_\chi \mathcal{W}(x,a)d\mu_x$. Otherwise, $\int_\chi \mathcal{W}(x,a)d\mu_x$ is an upper bound for the volume of the set $\mathcal{K}_1$.

The following theorem establish the equivalence of problems in (6.18) and (6.24).

Theorem 29. *There is no duality gap between the infinite LP on measure in* (6.18) *and infinite LP on continuous function in* (6.24) *in the sense that the optimal values are the same, i.e.,* $\mathbf{P}^*_{\mathbf{f_i}} = \mathbf{P}^*_{\mathbf{Dual}}$

Proof. See Appendix E. □

To be able to obatin a tractable relaxation of infinite LP in (6.24), we use polynomial approximation of continuous function $\mathcal{W}$ and use SOS relaxation to satisfy the nonnegativity constraints, where results in following finite SDP on polynomials:

$$\mathbf{P}^*_{\mathbf{d}} := \min_{\beta\in\mathbb{R},\mathcal{P}^d_{\mathcal{W}}\in\mathbb{R}_d[x,a]} \beta, \tag{6.25}$$

$$\text{s.t.}\quad \mathcal{P}^d_{\mathcal{W}}(x,a) - 1 \in \mathcal{QM}\left(\{\mathcal{P}_{1j}\}_{j=1}^{o_1}\right), \tag{6.25a}$$

$$\beta - \int_\chi \mathcal{P}^d_{\mathcal{W}}(x,a)d\mu_x \in \mathcal{QM}\left(\{1-\mathcal{P}^d_{\mathcal{A}}(a)\}, \{(1-a_i^2)\}_{i=1}^m\right), \tag{6.25b}$$

$$\mathcal{P}^d_{\mathcal{W}}(x,a) \geq 0,\ \ \beta \geq 0. \tag{6.25c}$$

where, $\mathcal{P}^d_{\mathcal{W}}(x,a) \in \mathbb{R}_d[x,a]$, μ_x is a given finite Borel measure and $\mathcal{QM}$ defined in (2.1) is quadratic module generated by polynomials. According to the Lemma 7, constraints (6.25a) and (6.25b) imply that polynomials $\mathcal{P}^d_{\mathcal{W}}(x,a) - 1$ and $\beta - \int_\chi \mathcal{P}^d_{\mathcal{W}}(x,a)d\mu_x$ are positive on the sets $\mathcal{K}_1$ in (6.13) and $\mathcal{A}_d = \{a \in \mathcal{A} : 1-\mathcal{P}^d_{\mathcal{A}}(a) > 0\}$ in (6.20), respectively, where $\mathcal{P}^d_{\mathcal{A}}(a)$ is an optimal solution of SDP (6.19). Problem in (6.25) is a SDP, where objective function is a linear and constraints are convex linear matrix inequalities in terms of coefficients of polynomial $\mathcal{P}^d_{\mathcal{W}}$. To be able to work with closed set $\mathcal{A}_d$, see the Remark 6.4.1.

The following theorem establish the equivalence of problems in (6.21) and (6.25).

Theorem 30. *There is no duality gap between the finite SDP on moments in* (6.21) *and finite SDP on polynomials in* (6.25) *in the sense that the optimal values are the same, i.e.,* $\mathbf{P}_{\mathbf{r}}^{*} = \mathbf{P}_{\mathbf{d}}^{*}$.

Proof. See Appendix F. □

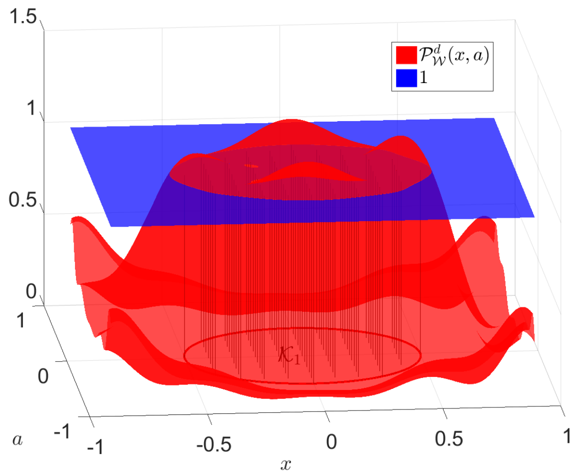

Figure 6.4: Polynomial $\mathcal{P}_{\mathcal{W}}^{d}(x,a)$ obtained by SDP (6.25) for $d = 12$

Remark 6.5.1. *In low dimensional problems, we can replace the global positivity condition in* (6.25c) *with local constraint as* $\{\mathcal{P}_{\mathcal{W}}^{d}(x,a) \geq 0 \quad on \quad \chi \times \mathcal{A}\}$ *to improve the obtained results.*

6.5.1 Illustrative Example

Consider the simple example provided in section 6.4.3. Here, to obtain an approximate solution, we solve the dual problem provided in finite SDP (6.25). We take $\mathcal{P}_{\mathcal{A}}^{7}(a)$ obtained by solving (6.19) and solve SDP in (6.25) for polynomial order $d = 12$ by Yalmip. Figure 6.4 displays obtained $\mathcal{P}_{\mathcal{W}}^{d}(x,a)$ which is greater than 1 on the set $\mathcal{K}_1$ and is positive on $\chi \times \mathcal{A} = [-1,1]^2$ as in constraint (6.25a). Figure 6.5 displays obtained β and also $\int_{\chi} \mathcal{P}_{\mathcal{W}}^{d}(x,a) d\mu_x$. As in constraint (6.25b) β is greater than $\int_{\chi} \mathcal{P}_{\mathcal{W}}^{d}(x,a) d\mu_x$ on the set $\mathcal{A}_d = \{a \in \mathcal{A} : \mathcal{P}_{\mathcal{A}}^{7}(a) < 1\}$. Based on obtained β and $\mathcal{P}_{\mathcal{W}}^{12}(x,a)$, we approximate the solution to the volume optimization problem with $a = -0.2050$ that maximizes polynomial $\int_{\chi} \mathcal{P}_{\mathcal{W}}^{d}(x,a) d\mu_x$ on the $\mathcal{A}_d$ and estimate the optimal volume $\mathbf{P}_{\mathbf{vol}}^{*}$ with $\mathbf{P}_{\mathbf{d}} = \beta = 1.239$. Based on the Theorem 30, the obtained solution by solving dual SDP in (6.25) matches the solution obtained by SDP in (6.21).

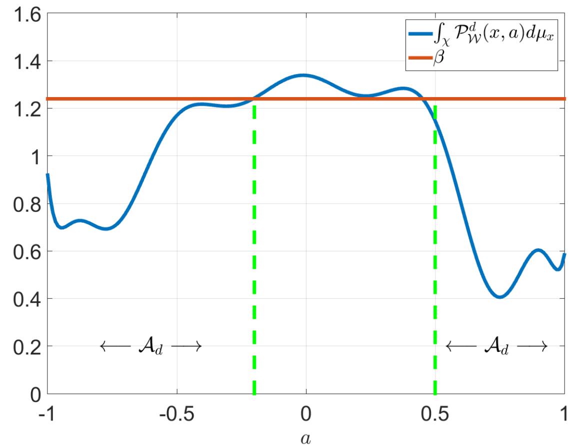

Figure 6.5: β and $\int_\chi \mathcal{P}^d_{\mathcal{W}}(x,a)d\mu_x$ obtained by SDP (6.25) for $d = 12$

6.6 Implementation and Numerical Results

In this section, numerical examples are presented that illustrate the performance of proposed method. The presented example are problem of inner approximation of ROA set defined in section 6.3.1.

6.6.1 Example 1: ROA set of system

In this example, we address the problem of approximating ROA set defined in 6.3.1. Consider the following locally stable nonlinear system.

$$\begin{aligned} \dot{x}_1 &= -\, x_2 \\ \dot{x}_2 &= x_1 + (4x_1^2 - 1)x_2 \end{aligned} \tag{6.26}$$

where, states of the system $x \in \chi = [-1,1]^2$. To approximate the ROA set of the system in the unit box, the Lyapunov function is described as

$$V(x) = 3\|x\|_2^2 + 3a_1x_1x_2 + 3a_2x_1^3x_2 + 3a_3x_1x_2^3 \tag{6.27}$$

where $a = [a_1, a_2, a_3] \in \mathcal{A} = [-1, 1]^3$ is the vector of unknown coefficients. The equivalent constrained volume optimization problem is stated as (6.9). To obtain an approximate solution, we solve finite SDPs in (6.21) and (6.19). First, we solve the SDP in (6.19) to obtain the polynomial $\mathcal{P}_{\mathcal{A}}^d(a)$ for $d = 10$. The polynomials describing the sets $\mathcal{K}_1$ and $\mathcal{K}_2$ are:

$$\mathcal{P}_{11} = V(x,a), \quad \mathcal{P}_{12} = 1 - V(x,a) \tag{6.28}$$

$$\mathcal{P}_{21} = -\epsilon_r \|x\|_2^2 - \frac{\partial V(x,a)}{\partial x_1}\dot{x}_1 - \frac{\partial V(x,a)}{\partial x_1}\dot{x}_2 \tag{6.29}$$

We set ϵ_r to 0.001 and $\epsilon_{\mathcal{K}}$ and $\epsilon_{\mathcal{A}}$ as in Remark 6.4.1 to 0.1 and 0.02, respectively. Figure 6.6 shows the obtained set $\left\{(a_1, a_2, a_3) : \; \mathcal{P}_{\mathcal{A}}^{10}(a) \leq 1 - \epsilon_{\mathcal{A}}\right\}$. Based on Theorem 28, the set $\mathcal{A}_d = \left\{(a_1, a_2, a_3) : \; \mathcal{P}_{\mathcal{A}}^{10}(a) \leq 1 - \epsilon_{\mathcal{A}}, \{1 - a_i^2 \geq 0\}_{i=1}^3\right\}$ is an inner approximation of the set of all coefficients (a_1, a_2, a_3) for which the set $\{x \in \chi : \; 0 \leq V(x,a) \leq 1\}$ is subset of the set $\{x \in \chi : \dot{V}(x,a) \leq -\epsilon_r \|x\|_2^2\}$.

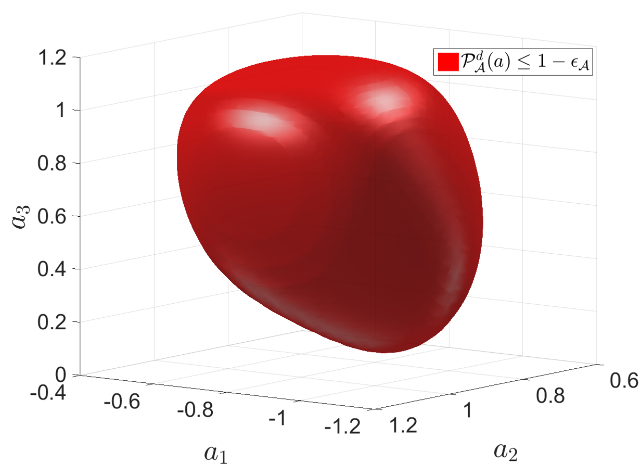

Figure 6.6: The set $\left\{(a_1, a_2, a_3) : \; \mathcal{P}_{\mathcal{A}}^d(a) \leq 1 - \epsilon_{\mathcal{A}}\right\}$ obtained by SDP (6.19) for $d = 10$ and $\epsilon_{\mathcal{A}} = 0.02$

We take $\mathcal{A}_{10}$ and solve SDP in (6.21). Based on moments of Lebesgue measure μ_x on $\chi = [-1, 1]^2$ we construct the matrices in constraints of SDP (6.21) in terms of unknown moment vectors $\mathbf{y} \in \mathbb{R}^{S_{5,2r}}$ and $\mathbf{y_a} \in \mathbb{R}^{S_{3,2r}}$. The SDP in (6.21) with $r = 7$ is solved using GloptiPoly. Based on obtained solution for moment vectors, we approximate the (a_1, a_2, a_3) with the first order moments of vector y_a

as $(y_{a_{100}}, y_{a_{010}}, y_{a_{001}}) = (-0.999362, 0.853458, 0.132566)$. Figure 6.7 shows the sets $\mathcal{S}_1(a) = \{x \in \chi : \; 0 \leq V(x,a) \leq 1\}$ and $\mathcal{S}_2(a) = \{x \in \chi : \dot{V}(x,a) \leq -\epsilon_r \|x\|_2^2\}$ for obtained coefficients a. For obtained a, the set $\mathcal{S}_1(a)$ is subset of the set $\mathcal{S}_2(a)$; hence, is an inner approximation of the ROA set.

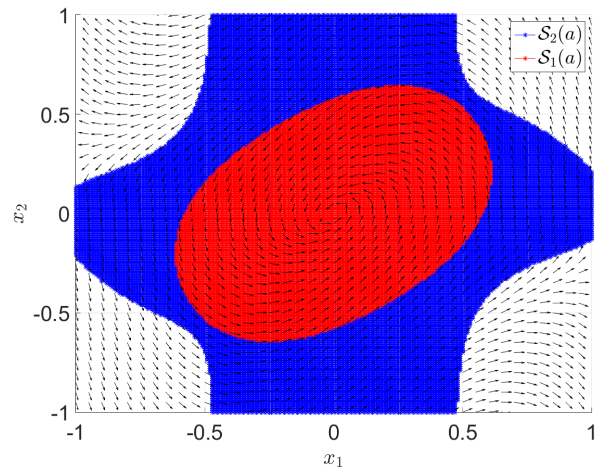

Figure 6.7: The sets $\mathcal{S}_1(a) = \{x \in \chi : \; 0 \leq V(x,a) \leq 1\}$ and $\mathcal{S}_2(a) = \{x \in \chi : \dot{V}(x,a) \leq -\epsilon_r \|x\|_2^2\}$ for obtained a

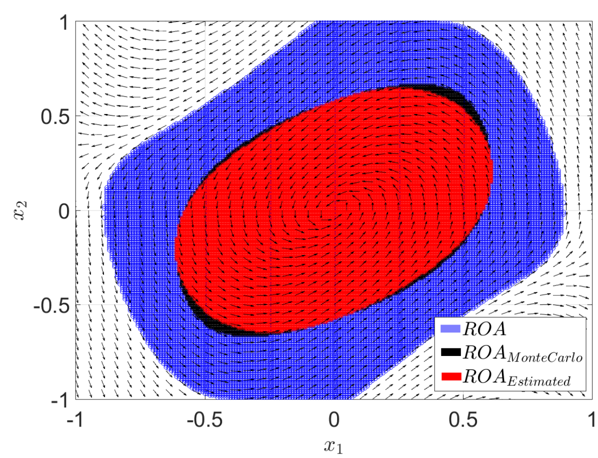

Figure 6.8: The true and estimated ROA sets

To test the accuracy of the obtained results, we used Monte Carlo simulation. The obtained result for coefficients of provided Lyapunov function by Monte Carlo method are $(a_1^*, a_2^*, a_3^*) = (-0.6, 0, -0.9)$. Figure 6.8 depicts the true ROA set for the system inside the unite box as well as obtained ROA using Monte Carlo method and convex approach provided in this work.

6.7 Conclusion

In this chapter, constrained volume optimization problems are introduced, where one aims at maximizing the volume of a set defined by polynomial inequalities such that it is contained in other semialgebraic set. We showed that many nonconvex problems in system and control can be reformulated as constrained volume optimization problems. To be able to obtain a equivalent convex problem, the results from theory of measure and moments as well as duality theory are used. Sequence of semidefinite relaxations is provided whose sequence of optimal values is shown to converge to the optimal value of the original problem. Numerical examples are provided that show that one can obtain reasonable approximations to the optimal solution.

6.8 Appendix A: Proof of Theorem 25

Let $\mathcal{A}_{\mathcal{F}}$ be the set of all parameters $a \in \mathcal{A}$ for which the set $\mathcal{S}_1(a)$ is a subset of the set $\mathcal{S}_2(a)$.Then, consider the following problem over the measures μ_a

$$\mathbf{P}_{\mu_\mathbf{a}} := \sup_{\mu_a \in \mathcal{M}_+(\mathcal{A}_{\mathcal{F}})} \left\{ \int_{\mathcal{A}} \mathrm{vol}_{\mu_x}(\mathcal{S}_1(a))\ d\mu_a : \mu_a(\mathcal{A}_{\mathcal{F}}) = 1 \right\} \tag{6.30}$$

We first want to show that $\mathbf{P}^*_{\mathbf{vol}} = \mathbf{P}_{\mu_\mathbf{a}}$. Let μ_a be a feasible solution to (6.30). Since, $\mathrm{vol}_{\mu_x}(\mathcal{S}_1(a)) \leq \mathbf{P}^*_{\mathbf{vol}}$ for all $a \in \mathcal{A}$, we have $\int_{\mathcal{A}} \mathrm{vol}_{\mu_x}(\mathcal{S}_1(a))\ d\mu_a \leq \mathbf{P}^*_{\mathbf{vol}}$. Thus, $\mathbf{P}_{\mu_\mathbf{a}} \leq \mathbf{P}^*_{\mathbf{vol}}$. Conversely, let $a \in \mathcal{A}$ be a feasible solution to the problem in (6.6); hence, a belongs to the set $\mathcal{A}_{\mathcal{F}}$. Let δ_a denotes the Dirac measure at a. The objective value of a in (6.6) is equal to $\mathrm{vol}_{\mu_x}(\mathcal{S}_1(a))$. Moreover, $\mu_a = \delta_a$ is a feasible solution to the problem in (6.30) with objective value equal to $\mathrm{vol}_{\mu_x}(\mathcal{S}_1(a))$. This implies that $\mathbf{P}^*_{\mathbf{vol}} \leq \mathbf{P}_{\mu_\mathbf{a}}$. Hence, $\mathbf{P}^*_{\mathbf{vol}} = \mathbf{P}_{\mu_\mathbf{a}}$, and (6.30) can be rewritten as

$$\mathbf{P}^*_{\mathbf{vol}} = \sup_{\mu_a \in \mathcal{M}_+(\mathcal{A}_{\mathcal{F}})} \left\{ \int_{\mathcal{A}} \int_{\mathcal{S}_1(a)} d\mu_x d\mu_a :\ \mu_a(\mathcal{A}_{\mathcal{F}}) = 1 \right\} \tag{6.31}$$

$$= \sup_{\mu_a \in \mathcal{M}_+(\mathcal{A}_\mathcal{F})} \left\{ \int_{\mathcal{K}_1} d\mu_a \mu_x : \ \mu_a(\mathcal{A}_\mathcal{F}) = 1 \right\} \tag{6.32}$$

and using the Lemma 6, we obtain

$$\mathbf{P}^*_{\mathbf{vol}} = \sup_{\mu_a \in \mathcal{M}_+(\mathcal{A}_\mathcal{F}), \mu \in \mathcal{M}_+(\mathcal{K}_1)} \int d\mu \tag{6.33}$$

$$\text{s.t.} \quad \mu \preccurlyeq \mu_a \times \mu_x, \ \mu_a(\mathcal{A}_\mathcal{F}) = 1. \tag{6.33a}$$

For a given $a \in \mathcal{A}$, the set $\mathcal{S}_1(a)$ is subset of the set $\mathcal{S}_2(a)$, if the set $\{x \in \chi : \ (x, a) \in \mathcal{K}_1 \cap \overline{\mathcal{K}_2}\}$ is an empty set. Hence, for any measure $\mu_a \in \mathcal{M}_+(\mathcal{A}_\mathcal{F})$, we have $\mu_a \times \mu_x \in \mathcal{M}_+(\overline{\mathcal{K}_1} \cup \mathcal{K}_2)$ considering that μ_x is supported on χ. Therefore, in problem (6.33) we can look for μ_a that is supported on $\mathcal{A}$ such that measure $\mu_a \times \mu_x$ is supported on $\overline{\mathcal{K}_1} \cup \mathcal{K}_2$. This results in the following problem:

$$\mathbf{P}^*_{\mathbf{vol}} = \sup_{\mu_a \in \mathcal{M}_+(\mathcal{A}), \mu \in \mathcal{M}_+(\mathcal{K}_1)} \int d\mu \tag{6.34}$$

$$\text{s.t.} \quad \mu \preccurlyeq \mu_a \times \mu_x, \ \mu_a(\mathcal{A}) = 1, \tag{6.34a}$$

$$\mu_a \times \mu_x \in \mathcal{M}_+(\overline{\mathcal{K}_1} \cup \mathcal{K}_2). \tag{6.34b}$$

Therefore, $\mathbf{P}^*_{\mathbf{vol}} = \mathbf{P}^*_{\mathbf{measure}}$.

6.9 Appendix B: Proof of Theorem 26

We want to obtain the set $\mathcal{A}_\mathcal{F}$ in (6.16), set of all possible decision parameters $a \in \mathcal{A}$ for which $\mathcal{S}_1(a) \subseteq \mathcal{S}_2(a)$. The idea is to approximate the indicator function of the set $\mathcal{A}_\mathcal{F}$; i.e., $\mathcal{I}_{\mathcal{A}_\mathcal{F}}(a) = 1$ if $a \in \mathcal{A}_\mathcal{F}$ and 0 otherwise, with continuous functions ([67], Section 3.2). There exist a sequence of functions $f_i \in \mathcal{C}[a]$ that converges from above to the indicator function of set $\mathcal{A}_\mathcal{F}$ as $i \to \infty$ ([80], Theorem A6.6, Urysohns Lemma A4.2), which results in an outer approximation of the set $\mathcal{A}_\mathcal{F}$ as $\mathcal{A}_{f_i} = \{a \in \mathcal{A} : \ f_i(a) \geq 1\} \supset \mathcal{A}_\mathcal{F}$.

To avoid the outer approximation and obtain the inner approximation of the set $\mathcal{A}_\mathcal{F}$ instead, we obtain the outer approximation of the complement set $\overline{\mathcal{A}_\mathcal{F}}$ by approximating its indicator function. Hence, if $f \in \mathcal{C}$ approximates the indicator function of $\overline{\mathcal{A}_\mathcal{F}}$ from above, the set $\mathcal{A}_f = \{a \in \mathcal{A} : \ f(a) < 1\}$ is an inner approximation of the set $\mathcal{A}_{\mathcal{F}}$.

To find such function $f(a)$, we use the LP in (6.17). For a given $a \in \mathcal{A}$, the set $\mathcal{S}_1(a)$ is subset of the set $\mathcal{S}_2(a)$, if the set $\left\{x \in \chi : \ (x,a) \in \mathcal{K}_1 \cap \overline{\mathcal{K}_2}\right\}$ is an empty set. Hence, the set $\mathcal{A}_{\mathcal{F}}$ can be describe as $\mathcal{A}_{\mathcal{F}} = \{a \in \mathcal{A} : \ \nexists x \in \chi \text{ s.t. } (x,a) \in \mathcal{K}_1 \cap \overline{\mathcal{K}_2}\}$. As a result, the complement set $\overline{\mathcal{A}_{\mathcal{F}}}$ read as

$$\overline{\mathcal{A}_{\mathcal{F}}} = \{a \in \mathcal{A} : \ \exists x \in \chi \text{ s.t. } (x,a) \in \mathcal{K}_1 \cap \overline{\mathcal{K}_2}\} \tag{6.35}$$

Therefore, to approximate the indicator function of $\overline{\mathcal{A}_{\mathcal{F}}}$ in (6.35), the continuous function $f(a)$ should be greater 1 over the set $\left\{(x,a) \in \mathcal{K}_1 \cap \overline{\mathcal{K}_2}\right\}$ and 0 otherwise, as in (6.17a) and (6.17b), the constrains of the LP. By minimizing the L_1-norm of $f(a)$ as in the objective function of (6.17), we converge to the indicator function of the set $\overline{\mathcal{A}_{\mathcal{F}}}$ from the above. Hence, $\mathcal{A}_{f_i}$ can arbitrarily approximate the set $\mathcal{A}_{\mathcal{F}}$ in (6.16) and (i) and (ii) hold true.

6.10 Appendix C: Proof of Theorem 27

First, consider the following problem over the measures supported in the set $\mathcal{A}_{\mathcal{F}}$ in (6.16).

$$\mathbf{P}_{\mu_\mathbf{a}} := \sup_{\mu_a \in \mathcal{M}_+(\mathcal{A}_{\mathcal{F}})} \left\{ \int_{\mathcal{A}} \mathrm{vol}_{\mu_x}(\mathcal{S}_1(a)) \; d\mu_a : \ \mu_a(\mathcal{A}_{\mathcal{F}}) = 1 \right\} \tag{6.36}$$

As in Appendix in 6.8, we can show that $\mathbf{P}^*_{\mathbf{vol}} = \mathbf{P}_{\mu_\mathbf{a}}$ and Eq. (6.37) is true.

$$\mathbf{P}^*_{\mathbf{vol}} = \sup_{\mu_a \in \mathcal{M}_+(\mathcal{A}_{\mathcal{F}}), \mu \in \mathcal{M}_+(\mathcal{K}_1)} \int d\mu \tag{6.37}$$

$$\text{s.t.} \quad \mu \preccurlyeq \mu_a \times \mu_x, \ \ \mu_a(\mathcal{A}_{\mathcal{F}}) = 1. \tag{6.37a}$$

Now, we replace the set $\mathcal{A}_{\mathcal{F}}$ in problem (6.37) with the set $\mathcal{A}_{f_i}$, where results in Problem (6.18). Hence, as the set $\mathcal{A}_{f_i} = \{a \in \mathcal{A} : \ f_i(a) < 1\}$ defined in Theorem 26 converges to the set $\mathcal{A}_{\mathcal{F}}$, the optimal value $\mathbf{P}^*_{\mathbf{f_i}}$ converges to the $\mathbf{P}^*_{\mathbf{vol}}$ and measures $\mu^*_a(f_i)$ and $\mu^*(f_i)$ converge to $\mu_a = \delta_{a^*}$, Dirac measure at a^*, and $\mu = \delta_{a^*} \times \mu_x$, respectively. Also, since $\mathcal{A}_{f_i}$ is an inner approximation of the set $\mathcal{A}_{\mathcal{F}}$, any point a_i in the support of measure $\mu^*_a(f_i)$ is also contained in the set $\mathcal{A}_{\mathcal{F}}$; Hence is an optimal solution to (6.6) and converges to the a^*.

6.11 Appendix D: Proof of Theorem 28

To drive a finite convex relaxation of the infinite LP problem in (6.17), we use finite order polynomial $\mathcal{P}_{\mathcal{A}}^d(a)$ to approximate the continuous function $f(a)$ and SOS relaxations to satisfy the constraints of the problem in (6.17), ([67] Section 3.3, [102]). Based on Stone-Weierstrass Theorem [80], every continuous function can be uniformly approximated as closely as desired by a polynomial. Then, to make such polynomial $\mathcal{P}_{\mathcal{A}}^d$ to satisfy the constraints of the problem in (6.17), SOS relaxations are used as in Lemma 7. Constraints (6.19a) implies that $\mathcal{P}_{\mathcal{A}}^d(a) - 1$ belongs to the quadratic module generated by polynomials of set $\mathcal{K}_1 \cap \overline{\mathcal{K}_2} = \{(x,a) : \cup_{i=1}^{o_2}\{-\mathcal{P}_{2i} > 0, \mathcal{P}_{1j} \geq 0, j = 1, \ldots, o_1\}\}$; hence, $\mathcal{P}_{\mathcal{A}}^d(a) - 1$ is nonnegative on the set $\mathcal{K}_1 \cap \overline{\mathcal{K}_2}$. Also, constraint (6.19b) implies that $\mathcal{P}_{\mathcal{A}}^d(a)$ belong to the quadratic module generated by polynomials describing the hyper cube $[-1,1]^n \times [-1,1]^m$, so $\mathcal{P}_{\mathcal{A}}^d(a)$ is nonnegative over the set $\chi \times \mathcal{A} = [-1,1]^n \times [-1,1]^m$. Therefore, by minimizing the L_1-norm of the $\mathcal{P}_{\mathcal{A}}^d$ similar to (6.17), and $d \to \infty$, $\mathcal{P}_{\mathcal{A}}^d$ converges to the indicator function of the set $\overline{\mathcal{A}_{\mathcal{F}}}$. Therefore, the set $\mathcal{A}_d = \left\{a \in \mathcal{A} : \ \mathcal{P}_{\mathcal{A}}^d(a) < 1\right\}$ converges to the set $\mathcal{A}_{\mathcal{F}}$ in (6.16) as in the Theorem 26.

6.12 Appendix E: Proof of Theorem 29

The LP in (6.18) can be rewritten as

$$\mathbf{P_1^*} := \ \sup\langle\gamma, c\rangle \tag{6.38}$$

$$\text{s.t.} \quad A^*\gamma = b \tag{6.38a}$$

$$\gamma \in \mathcal{M}_+(\mathcal{K}_1) \times \mathcal{M}_+(\mathcal{A}_{f_i}). \tag{6.38b}$$

where, $\gamma := (\mu, \mu_a) \in \mathcal{M}_+(\mathcal{K}_1) \times \mathcal{M}_+(\mathcal{A}_{f_i})$ is the variable vector, and $c := (1,0) \in \mathcal{C}_+(\mathcal{K}_1) \times \mathcal{C}_+(\mathcal{A}_{f_i})$, so objective function is $\langle\gamma, c\rangle = \int d\mu$. Also, $A^* : \mathcal{M}_+(\mathcal{K}_1) \times \mathcal{M}_+(\mathcal{A}_{f_i}) \to \mathcal{M}_+(\chi \times \mathcal{A}) \times \mathbb{R}_+$ is the linear operator that is defined by $A^*\gamma := (\mu - \mu_a \times \mu_x, \int_{\mathcal{A}} d\mu_a)$ and $b := (0,1) \in \mathcal{M}_+(\chi \times \mathcal{A}) \times \mathbb{R}_+$,([49], Theorem 2, [79, 82]). The problem in (6.38) is infinite LP defined in cone of nonnegative measures. The cone of nonnegative continuous functions are dual to cone of nonnegative measures. Based on standard results on LP the dual problem of (6.38) reads as

$$\mathbf{P}_{\mathbf{2}}^{*} := \inf \langle b, z \rangle \tag{6.39}$$

$$\text{s.t.} \quad Az - c \in \mathcal{C}_{+}(\mathcal{K}_1) \times \mathcal{C}_{+}(\mathcal{A}_{f_i}) \tag{6.39a}$$

where, $z := (\mathcal{W}(x,a), \beta) \in \mathcal{C}_{+}(\chi \times \mathcal{A}) \times \mathbb{R}_{+}$ is the variable vector, so the objective function is $\langle b, z \rangle = \beta$. The linear operator $A : \mathcal{C}_{+}(\chi \times \mathcal{A}) \times \mathbb{R}_{+} \to \mathcal{C}_{+}(\mathcal{K}_1) \times \mathcal{C}_{+}(\mathcal{A}_{f_i})$ satisfies adjoint relation$\langle A^{*}\gamma, z \rangle = \langle \gamma, Az \rangle$; hence, is defined by $Az := (\mathcal{W}(x,a), \beta - \int_{\mathcal{A}} \mathcal{W}(x,a) d\mu_x)$. As a result, the dual problem (6.39) is equal to the problem (6.24).

If problem in (6.38) is consistent with finite value and the set

$$D := \{(A^{*}\gamma, \langle \gamma, c \rangle) : \gamma \in \mathcal{M}_{+}(\mathcal{K}_1) \times \mathcal{M}_{+}(\mathcal{A}_{f_i})\}$$

is closed, then there is no duality gap between (6.38) and (6.39). The support of measures in (6.38) are compact. Also, the measure μ is constrained by the measure $\mu_a \times \mu_x$ in which, measure μ_a is probability measure; i.e., $\mu_a(\mathcal{A}_{f_i}) = 1$, and μ_x is finite Borel measure defined on compact set χ. Hence, $\mathbf{P}_{\mathbf{1}}^{*} = \sup \int d\mu < \infty$. Also, the feasible set of (6.38) is nonempty for instance $(\delta_a \times \mu_x, \delta_a)$ for $a \in \mathcal{A}_{f_i}$ is a feasible solution; therefor $0 \leq \mathbf{P}_{\mathbf{1}}^{*} = \sup \int d\mu < \infty$. Using sequential Banach$-$Alaoglu theorem [80] and weak-$\star$ continuity of the A^{*}, there exist an accumulation point of $\gamma_k = (\mu_k, \mu_{ak})$ in the weak-$\star$ topology of nonnegative measures such that $\lim_{k \to \infty} ((A^{*}\gamma_k, \langle \gamma_k, c \rangle)) \in D$; hence, D is closed, ([49], Theorem 2).

6.13 Appendix F: Proof of Theorem 30

For simplicity, we denote the polynomials $\left(\{1 - \mathcal{P}_{\mathcal{A}}^{d}\}, \{(1 - a_i^2)\}_{i=1}^{m}\right)$ that construct the set $\mathcal{A}_d$ by $\{\mathcal{P}_{\mathcal{A}j}(a)\}_{j=1}^{o_a}$. Matrices of the problem (6.21) can be rewritten as follow, ([67]). $M_r(\mathbf{y}) = \sum_{\alpha} A_{\alpha} y_{\alpha}$ and $M_{r-r_j}(\mathbf{y}; \mathcal{P}_{1j}) = \sum_{\alpha} B_{\alpha}^{j} y_{\alpha}$. Also, $M_r(\mathbf{y_a}) = \sum_{\alpha} D_{\alpha} y_{a_{\alpha}}$, $M_{r-r_a}(\mathbf{y_a}; \mathcal{P}_{\mathcal{A}j}(a)) = \sum_{\alpha} E_{\alpha}^{j} y_{a_{\alpha}}$, and $M_r(\mathbf{y_a} \times \mathbf{y_x} - \mathbf{y}) = \sum_{\alpha} F_{\alpha} y_{a_{\alpha}} - \sum_{\alpha} A_{\alpha} y_{\alpha}$ for appropriate real symmetric matrices $(A_{\alpha}, \{B_{\alpha}^{j}\}_{j=1}^{o_1}, D_{\alpha}, \{E_{\alpha}^{j}\}_{j=1}^{o_a}, F_{\alpha})$ and $0 \leq |\alpha| \leq 2r$. Let, $\gamma = (\mathbf{y} \in \mathbb{R}^{S_{n+m,2r}},\ \mathbf{y_a} \in \mathbb{R}^{S_{m,2r}})$. Then problem in (6.21) can be rewritten as a standard form as follow:

$$\mathbf{P}_{\mathbf{r}}^* := \sup_{\gamma} b^T\gamma, \tag{6.40}$$

$$\text{s.t.} \quad C_1 + \sum_{\alpha} \hat{A}_\alpha \gamma_\alpha \succcurlyeq 0, \tag{6.40a}$$

$$C_2^j + \sum_{\alpha} \hat{B}_\alpha^j \gamma_\alpha \succcurlyeq 0, \quad j = 1, \ldots, o_1 \tag{6.40b}$$

$$C_3 - \sum_{\alpha} \hat{C}_\alpha \gamma_\alpha \succcurlyeq 0, \tag{6.40c}$$

$$C_4 + \sum_{\alpha} \hat{D}_\alpha \gamma_\alpha \succcurlyeq 0, \tag{6.40d}$$

$$C_5^j + \sum_{\alpha} \hat{E}_\alpha^j \gamma_\alpha \succcurlyeq 0, \quad j = 1, \ldots, o_a \tag{6.40e}$$

$$C_6 + \sum_{\alpha} \hat{F}_\alpha \gamma_\alpha \succcurlyeq 0, \tag{6.40f}$$

where, $b = (1, \mathbf{0}) \in \mathbb{R}^{S_{n+m,2r}+S_{m,2r}}$, $(C_1, C_2, C_4, C_5, C_6)$ are zero matrices, $(\hat{A}_\alpha, \{\hat{B}_\alpha^j\}_{j=1}^{o_1}$ $, \hat{D}_\alpha, \{\hat{E}_\alpha^j\}_{j=1}^{o_a}, \hat{F}_\alpha)$ are real symmetric matrices, $C_3 = 1$, and $\hat{C}^T = (\mathbf{0} \in \mathbb{R}^{S_{n+m,2r}}, 1, \mathbf{0} \in \mathbb{R}^{S_{m,2r}-1}) \in \mathbb{R}^{S_{n+m,2r}+S_{m,2r}}$. Based on standard results on duality of SDP, the dual problem to (6.40) reads as

$$\mathbf{P}_{\mathbf{d}}^* := \inf_{\{X^j\}_{j=0}^{o_1}, \{Y^j\}_{j=0}^{o_a}, Z, \beta} \langle C_1, X^0\rangle + \sum_{j=1}^{o_1} \left\langle C_2^j, X^j \right\rangle + \langle C_3, \beta\rangle + \langle C_4, Y^0\rangle + \sum_{j=1}^{o_a} \left\langle C_5^j, Y^j \right\rangle + \langle C_6, Z\rangle \tag{6.41}$$

$$\text{s.t.} \quad \beta - \langle A_\alpha, X^0\rangle - \sum_{j}^{o_1} \langle B_\alpha^j, X^j\rangle - \langle D_\alpha, Y^0\rangle - \sum_{j}^{o_a} \langle E_\alpha^j, Y^j\rangle - \langle F_\alpha, Z\rangle = b_\alpha, \quad \alpha = 0, \tag{6.41a}$$

$$- \langle A_\alpha, X^0\rangle - \sum_{j}^{o_1} \langle B_\alpha^j, X^j\rangle - \langle D_\alpha, Y^0\rangle - \sum_{j}^{o_a} \langle E_\alpha^j, Y^j\rangle - \langle F_\alpha, Z\rangle = b_\alpha, \quad 0 < |\alpha| \leq 2r, \tag{6.41b}$$

$$X^0, \{X^j\}_{j=1}^{o_1}, Y^0, \{Y^j\}_{j=1}^{o_a}, Z, \beta \succcurlyeq 0 \tag{6.41c}$$

where, $\langle X, Y\rangle = \text{trace}(XY)$. This problem is equal to the problem in (6.25). Based on the defined matrices and vectors, the cost function of (6.41) is equal to β. Also, let $\mathcal{B}_d$ denote the vector comprised of the monomial basis of $\mathbb{R}_{\mathrm{d}}[a, x]$. We can represent the polynomials of (6.25) as $\mathcal{P}_{\mathcal{W}}^d(x, a) = \mathcal{B}_d^T X^0 \mathcal{B}_d$, $\mathcal{QM}\left(\{\mathcal{P}_{1j}\}_{j=1}^{o_1}\right) = \sum_j^{o_1} \mathcal{B}_d^T X^j \mathcal{B}_d$, $\int \mathcal{P}_{\mathcal{W}}^d(x, a) d\mu_x = \mathcal{B}_d^T Y^0 \mathcal{B}_d$, $\mathcal{QM}\left(\{\mathcal{P}_{\mathcal{A}j}\}_{j=1}^{o_a}\right) = \sum_j^{o_a} \mathcal{B}_d^T Y^j \mathcal{B}_d$, and $\hat{\mathcal{P}}_{\mathcal{W}}^d(x, a)$

$= \mathcal{B}_d^T Z \mathcal{B}_d$. Then constraints (6.41a) and (6.41b) are conditions for α-th coefficient of polynomial $\mathcal{P}_{\mathcal{W}}^d(x,a)$ so that as constraints (6.25a) and (6.25b), $\mathcal{P}_{\mathcal{W}}^d(x,a) - 1 \in \mathcal{QM}\left(\{\mathcal{P}_{1j}\}_{j=1}^{o_1}\right)$, $\beta - \hat{\mathcal{P}}_{\mathcal{W}}^d(x,a) \in \mathcal{QM}\left(\{\mathcal{P}_{\mathcal{A}j}\}_{j=1}^{o_a}\right)$, and $\hat{\mathcal{P}}_{\mathcal{W}}^d(x,a) = \int_{\chi} \mathcal{P}_{\mathcal{W}}^d(x,a) d\mu_x$ are satisfied.

Based om Slater's sufficient condition,if the feasible set of strictly positive matrices in constraint of primal SDP is nonempty, then there is no duality gap. Consider SDP in (6.21). Let μ_a uniform measure on $\mathcal{A}_d$ and $\mu = \mu_a \times \mu_x$. Since set $\mathcal{K}_1$ and $\mathcal{A}_d$ have a nonempty interior, then $M_r(\mathbf{y}) \succ 0$, $M_{r-r_j}(\mathbf{y}; \mathcal{P}_{1j}) \succ 0, j = 1, \ldots, o_1$, $M_r(\mathbf{y_a}) \succ 0$, and $M_{r-r_a}(\mathbf{y_a}; \mathcal{P}_{aj}) \succ 0, j = 1, \ldots, o_a$. Based on Assumption 3, $\chi \times \mathcal{A} \setminus \mathcal{K}_1$ has nonempty interior; hence $M_r(\mathbf{y_a} \times \mathbf{y_x} - \mathbf{y}) \succ 0$. Therefore, Slater's condition holds, (see [67] for similar setup).

Chapter 7

Sparse Data Reconstruction in Sensory Networks

In this work, a novel approach to reconstruct a noisy sparse n-dimensional data is proposed. The main idea is to complete the data with least possible complexity. The complexity is defined as the number of exponential signals that can describe the data. We show that the number of exponential signals that can describe a data set corresponds to the rank of block Hankel matrix constructed from given data. In this context, the problem of data reconstruction can be reformulated as matrix completion and rank minimization problems where the nuclear norm is used as a convex relaxation of matrix rank. To be able to deal with large scale data, a first-order augmented Lagrangian algorithm is implemented for solving the resulting optimization problem. To illustrate the performance of the proposed approach, the results obtained by applying the method to practical data set is presented.

7.1 Introduction

This work addresses the problem of reconstructing a sparse and noisy data where we aim at finding the missing part of given data. This problem arises in different areas such as sensor networks, where the sensors do not completely cover the area of interest; Hence, the sampled data are usually inadequate. Moreover, reconstruction of corrupted image or videos can can be reformulated as a special case of this problem. Motivated by low-rank structure methods, we propose a novel approach to reconstruct a noisy sparse data with least possible complexity. In this work, to obtain the complete data, we look for minimum rank block Hankel matrix associated with given sparse and noisy data. We show that minimizing the rank of constructed block Hankel matrix is equal to minimizing the number of exponential signals that describes the data. The proposed method, with out making any assumption on the structure of data and data loss pattern, could reconstruct the complete data. To be able to deal with large scale data, a first-order augmented Lagrangian algorithm is implemented for solving the resulting optimization problem.

7.2 Problem Formulation

The problem of reconstructing a sparse and noisy n-dimensional data is formulate as follow. Consider a sensor network where sensors are scattered in a n-dimensional space $\mathbb{R}^{l_n \times \ldots \times l_2 \times l_1}$. We assume that the space is discretized in a uniform way where sensors are placed in a specific set of nodes denoted by $(k_n^*, ..., k_2^*, k_1^*)$ in the space, i.e., $k_i^* \in \mathbb{R}^{l_i^*},\ l_i^* \subset l_i,\ i = 1, ..., n$.

Given a sparse n-dimensional array of measurement $\hat{Y} \in \mathbb{R}^{l_n \times \ldots \times l_2 \times l_1}$ with noisy sensory measurement $\hat{Y}_{k_n^*, \ldots, k_2^*, k_1^*}$, i.e., measurement for the node $(k_n^*, ..., k_2^*, k_1^*)$, the objective is to denoise the measurement data and estimate the data for the missing measurements, with the least complex extension of the given measurements. As mentioned before, we define the notion of complexity as the number of exponential signals describing the data.

More precisely, we assume that any n-dimensional data array $Y \in \mathbb{R}^{l_n \times \ldots \times l_2 \times l_1}$ can be approximated by weighted sum of complex exponential signals of the form below

$$Y_{k_n, \ldots, k_2, k_1} \approx \sum_{i=1}^{N} a_i z_{n_i}^{k_n} \ldots z_{2_i}^{k_2} z_{1_i}^{k_1}, \quad k_j = 0, ..., l_j - 1, j = 1, ..., n \tag{7.1}$$

where, a_i is the *i-th* complex amplitude, and $(z_{1_i}, z_{2_i}, ..., , z_{n_i}) = (e^{j\omega_{1_i}}, e^{j\omega_{2_i}}, ..., e^{j\omega_{n_i}})$ defines the *i-th* n-dimensional complex frequency $(\omega_{1_i}, \omega_{2_i}, ..., \omega_{n_i})$, and N is the number of the n-dimensional distinguish complex frequencies. Hence, the problem completing the data can be defined as follows

$$\mathbf{P_1^*} := \min_{Y\in\mathbb{R}^{l_n\times\ldots\times l_2\times l_1}, \{a_i,(z_{1_i},z_{2_i},\ldots,z_{n_i})\}_{i=1}^N} N \tag{7.2}$$

$$\text{s.t.} \quad \|Y_{k_n^*,\ldots,k_2^*,k_1^*} - \hat{Y}_{k_n^*,\ldots,k_2^*,k_1^*}\|_2 \leq \epsilon, \tag{7.2a}$$

$$Y_{k_n,\ldots,k_2,k_1} = \sum_{i=1}^{N} a_i z_{n_i}^{k_n} \ldots z_{2_i}^{k_2} z_{1_i}^{k_1}, \quad k_j = 0,\ldots,l_j-1, j=1,\ldots,n \tag{7.2b}$$

where, $Y \in \mathbb{R}^{l_n\times\ldots\times l_2\times l_1}$ is constructed noiseless complete n-dimensional data, N is the number of extensional signals that describes the data Y as in (7.1), $\hat{Y} \in \mathbb{R}^{l_n\times\ldots\times l_2\times l_1}$ is given sparse noisy n-dimensional sensory data, $(k_n^*,\ldots,k_2^*,k_1^*)$ are indexes of known parts of measurement data $\hat{Y}$, $\|.\|_2$ is l_2 norm, and $\epsilon > 0$. In chapter, we provide the convex equivalent formulation of the Problem 1. For this purpose, we use Hankel matrix notion which is defined in the next section.

7.3 Hankel Matrix

To provide a convex equivalent problem for Problem 1 in (7.22), we first define Hankel matrix as follows.

Hankel Matrix: Given a vector $y = \{y_k \in \mathbb{R}, k=0,\ldots,l-1\}$, *Hankel* matrix $\mathcal{H}(y)$ is defined as

$$\mathcal{H}(y) = \begin{bmatrix} y(0) & y(1) & \ldots & y(l-M) \\ y(1) & y(2) & \ldots & y(l-M+1) \\ . & . & \ldots & . \\ . & . & \ldots & . \\ y(l-M) & y(l-M+1) & \ldots & y(l-1) \end{bmatrix} \tag{7.3}$$

where, $M \leq l$ is window size.

Block Hankel Matrix: For a given array consisting of l Hankel matrices $\mathcal{H}_i, i =$

$0, ..., l-1$, *Block Hankel* matrix is defined as [109]

$$\mathcal{H}(\{\mathcal{H}_i\}_{i=0}^{l-1}) = \begin{bmatrix} \mathcal{H}_0 & H_1 & ... & \mathcal{H}_{l-M} \\ \mathcal{H}_1 & H_2 & ... & \mathcal{H}_{l-M+1} \\ . & . & ... & . \\ . & . & ... & . \\ \mathcal{H}_{l-M} & \mathcal{H}_{l-M+1} & ... & H_{l-1} \end{bmatrix} \tag{7.4}$$

where, $M \leq l$ is window size.

n-Dimensional Block Hankel: Based on definition of *Hankel* and *Block Hankel* matrices, we define n-dimensional *block Hankel* matrix for n-dimensional data $Y \in \mathbb{R}^{l_n \times ... \times l_2 \times l_1}$. For a given n-dimensional data array $Y_{k_n,...,k_2,k_1}$, *n-Dimensional Block Hankel* matrix $\mathcal{H}^{nD}$ is defined as

$$\mathcal{H}^{nD} = \begin{bmatrix} \mathcal{H}_0^{(n-1)D} & \mathcal{H}_1^{(n-1)D} & ... & \mathcal{H}_{l_n-M_n}^{(n-1)D} \\ \mathcal{H}_1^{(n-1)D} & \mathcal{H}_2^{(n-1)D} & ... & \mathcal{H}_{l_n-M_n+1}^{(n-1)D} \\ . & . & ... & . \\ . & . & ... & . \\ \mathcal{H}_{l_n-M_n-1}^{(n-1)D} & \mathcal{H}_{l_n-M_n+1}^{(n-1)D} & ... & \mathcal{H}_{l_n-1}^{(n-1)D} \end{bmatrix} \tag{7.5}$$

where, M_n is window size, $\mathcal{H}_i^{(n-1)D}$ is *(n-1)-dimensional Block Hankel* matrix of (n-1)-dimensional data $Y_{k_n=i,...,k_2,k_1}$ and $i = 0, ..., l_n - 1$.

For example the 2D, 3D and 4D block Hankel matrices are as follow:

2D Block Hankel Matrix: For a given 2D data array $Y \in \mathbb{R}^{3\times3}$, and window sizes $M_i = 2,\ i = 1, 2$, 2D *Block Hankel* matrix $\mathcal{H}^{2D}$ is defined as

$$\mathcal{H}^{2D} = \begin{bmatrix} \mathcal{H}_0^{1D} & \mathcal{H}_1^{1D} \\ \mathcal{H}_1^{1D} & \mathcal{H}_2^{1D} \end{bmatrix} \tag{7.6}$$

where, $\mathcal{H}_0^{1D}, \mathcal{H}_1^{1D}$, and $\mathcal{H}_2^{1D}$ are 1D Hankel matrices of 1D data Y_{0,k_1}, Y_{1,k_1}, and Y_{2,k_1}, respectively as below

$$\mathcal{H}_0^{1D} = \begin{bmatrix} Y_{0,0} & Y_{0,1} \\ Y_{0,1} & Y_{0,2} \end{bmatrix}, \mathcal{H}_1^{1D} = \begin{bmatrix} Y_{1,0} & Y_{1,1} \\ Y_{1,1} & Y_{1,2} \end{bmatrix}, \mathcal{H}_2^{1D} = \begin{bmatrix} Y_{2,0} & Y_{2,1} \\ Y_{2,1} & Y_{2,2} \end{bmatrix}$$

3D Block Hankel Matrix: For a given 3D array $Y \in \mathbb{R}^{3\times3\times3}$, and window sizes $M_i = 2,\ i = 1, ..., 3$, 3D *Block Hankel* matrix $\mathcal{H}^{3D}$ is defined as

$$\mathcal{H}^{3D} = \begin{bmatrix} \mathcal{H}_0^{2D} & \mathcal{H}_1^{2D} \\ \mathcal{H}_1^{2D} & \mathcal{H}_2^{2D} \end{bmatrix} \tag{7.7}$$

where, $\mathcal{H}_0^{2D}, \mathcal{H}_1^{2D}$, and $\mathcal{H}_2^{2D}$ are 2D Hankel matrices of 2D data $Y_{0,k_2,k_1}, Y_{1,k_2,k_1}$, and Y_{2,k_2,k_1}, respectively. Matrix $\mathcal{H}_0^{2D}$ reads as

$$\mathcal{H}_0^{2D} = \begin{bmatrix} \mathcal{H}_0^{1D} & \mathcal{H}_1^{1D} \\ \mathcal{H}_1^{1D} & \mathcal{H}_2^{1D} \end{bmatrix}$$

$$\mathcal{H}_0^{1D} = \begin{bmatrix} Y_{0,0,0} & Y_{0,0,1} \\ Y_{0,0,1} & Y_{0,0,2} \end{bmatrix}, \mathcal{H}_1^{1D} = \begin{bmatrix} Y_{0,1,0} & Y_{0,1,1} \\ Y_{0,1,1} & Y_{0,1,2} \end{bmatrix}, \mathcal{H}_2^{1D} = \begin{bmatrix} Y_{0,2,0} & Y_{0,2,1} \\ Y_{0,2,1} & Y_{0,2,2} \end{bmatrix}$$

where, $\mathcal{H}_0^{1D}, \mathcal{H}_1^{1D}$, and $\mathcal{H}_2^{1D}$ are 1D Hankel matrices of 1D data $Y_{0,0,k_1}, Y_{0,1,k_1}$, and $Y_{0,2,k_1}$, respectively. Matrix $\mathcal{H}_1^{2D}$ reads as

$$\mathcal{H}_1^{2D} = \begin{bmatrix} \mathcal{H}_0^{1D} & \mathcal{H}_1^{1D} \\ \mathcal{H}_1^{1D} & \mathcal{H}_2^{1D} \end{bmatrix}$$

$$\mathcal{H}_0^{1D} = \begin{bmatrix} Y_{1,0,0} & Y_{1,0,1} \\ Y_{1,0,1} & Y_{1,0,2} \end{bmatrix}, \mathcal{H}_1^{1D} = \begin{bmatrix} Y_{1,1,0} & Y_{1,1,1} \\ Y_{1,1,1} & Y_{1,1,2} \end{bmatrix}, \mathcal{H}_2^{1D} = \begin{bmatrix} Y_{1,2,0} & Y_{1,2,1} \\ Y_{1,2,1} & Y_{1,2,2} \end{bmatrix}$$

where, $\mathcal{H}_0^{1D}, \mathcal{H}_1^{1D}$, and $\mathcal{H}_2^{1D}$ are 1D Hankel matrices of 1D data $Y_{1,0,k_1}, Y_{1,1,k_1}$, and $Y_{1,2,k_1}$, respectively. Matrix $\mathcal{H}_2^{2D}$ reads as

$$\mathcal{H}_2^{2D} = \begin{bmatrix} \mathcal{H}_0^{1D} & \mathcal{H}_1^{1D} \\ \mathcal{H}_1^{1D} & \mathcal{H}_2^{1D} \end{bmatrix}$$

$$\mathcal{H}_0^{1D} = \begin{bmatrix} Y_{2,0,0} & Y_{2,0,1} \\ Y_{2,0,1} & Y_{2,0,2} \end{bmatrix}, \mathcal{H}_1^{1D} = \begin{bmatrix} Y_{2,1,0} & Y_{2,1,1} \\ Y_{2,1,1} & Y_{2,1,2} \end{bmatrix}, \mathcal{H}_2^{1D} = \begin{bmatrix} Y_{2,2,0} & Y_{2,2,1} \\ Y_{2,2,1} & Y_{2,2,2} \end{bmatrix}$$

where, $\mathcal{H}_0^{1D}, \mathcal{H}_1^{1D}$, and $\mathcal{H}_2^{1D}$ are 1D Hankel matrices of 1D data $Y_{2,0,k_1}, Y_{2,1,k_1}$, and

$Y_{2,2,k_1}$, respectively. Therefore,

$$
\mathcal{H}^{3D} = \left[\begin{array}{cc|cc|cc|cc}
Y_{0,0,0} & Y_{0,0,1} & Y_{0,1,0} & Y_{0,1,1} & Y_{1,0,0} & Y_{1,0,1} & Y_{1,1,0} & Y_{1,1,1} \\
Y_{0,0,1} & Y_{0,0,2} & Y_{0,1,1} & Y_{0,1,2} & Y_{1,0,1} & Y_{1,0,2} & Y_{1,1,1} & Y_{1,1,2} \\
\hdashline
Y_{0,1,0} & Y_{0,1,1} & Y_{0,2,0} & Y_{0,2,1} & Y_{1,1,0} & Y_{1,1,1} & Y_{1,2,0} & Y_{1,2,1} \\
Y_{0,1,1} & Y_{0,1,2} & Y_{0,2,1} & Y_{0,2,2} & Y_{1,1,1} & Y_{1,1,2} & Y_{1,2,1} & Y_{1,2,2} \\
\hdashline
Y_{1,0,0} & Y_{1,0,1} & Y_{1,1,0} & Y_{1,1,1} & Y_{2,0,0} & Y_{2,0,1} & Y_{2,1,0} & Y_{2,1,1} \\
Y_{1,0,1} & Y_{1,0,2} & Y_{1,1,1} & Y_{1,1,2} & Y_{2,0,1} & Y_{2,0,2} & Y_{2,1,1} & Y_{2,1,2} \\
\hdashline
Y_{1,1,0} & Y_{1,1,1} & Y_{1,2,0} & Y_{1,2,1} & Y_{2,1,0} & Y_{2,1,1} & Y_{2,2,0} & Y_{2,2,1} \\
Y_{1,1,1} & Y_{1,1,2} & Y_{1,2,1} & Y_{1,2,2} & Y_{2,1,1} & Y_{2,1,2} & Y_{2,2,1} & Y_{2,2,2}
\end{array}\right]
$$

4D Hankel Matrix: For a given 4D array $Y \in R^{3\times3\times3\times3}$, and window sizes $M_i = 2,\ i = 1, ..., 4$, 4D *Block Hankel* matrix $\mathcal{H}^{4D}$is defined as

$$
\mathcal{H}^{4D} = \begin{bmatrix} \mathcal{H}_0^{3D} & \mathcal{H}_1^{3D} \\ \mathcal{H}_1^{3D} & \mathcal{H}_2^{3D} \end{bmatrix} \tag{7.8}
$$

where, $\mathcal{H}_0^{3D}, \mathcal{H}_1^{3D}$, and $\mathcal{H}_2^{3D}$ are 3D Hankel matrices of 3D data $Y_{0,k_3,k_2,k_1}, Y_{1,k_3,k_2,k_1}$, and Y_{2,k_3,k_2,k_1}, respectively. Based on the example of 3D block Hankel matrix, the matrices $\mathcal{H}_0^{3D}, \mathcal{H}_1^{3D}$, and $\mathcal{H}_2^{3D}$ read as

$$
\mathcal{H}_0^{3D} = \left[\begin{array}{cc|cc|cc|cc}
Y_{0,0,0,0} & Y_{0,0,0,1} & Y_{0,0,1,0} & Y_{0,0,1,1} & Y_{0,1,0,0} & Y_{0,1,0,1} & Y_{0,1,1,0} & Y_{0,1,1,1} \\
Y_{0,0,0,1} & Y_{0,0,0,2} & Y_{0,0,1,1} & Y_{0,0,1,2} & Y_{0,1,0,1} & Y_{0,1,0,2} & Y_{0,1,1,1} & Y_{0,1,1,2} \\
\hdashline
Y_{0,0,1,0} & Y_{0,0,1,1} & Y_{0,0,2,0} & Y_{0,0,2,1} & Y_{0,1,1,0} & Y_{0,1,1,1} & Y_{0,1,2,0} & Y_{0,1,2,1} \\
Y_{0,0,1,1} & Y_{0,0,1,2} & Y_{0,0,2,1} & Y_{0,0,2,2} & Y_{0,1,1,1} & Y_{0,1,1,2} & Y_{0,1,2,1} & Y_{0,1,2,2} \\
\hdashline
Y_{0,1,0,0} & Y_{0,1,0,1} & Y_{0,1,1,0} & Y_{0,1,1,1} & Y_{0,2,0,0} & Y_{0,2,0,1} & Y_{0,2,1,0} & Y_{0,2,1,1} \\
Y_{0,1,0,1} & Y_{0,1,0,2} & Y_{0,1,1,1} & Y_{0,1,1,2} & Y_{0,2,0,1} & Y_{0,2,0,2} & Y_{0,2,1,1} & Y_{0,2,1,2} \\
\hdashline
Y_{0,1,1,0} & Y_{0,1,1,1} & Y_{0,1,2,0} & Y_{0,1,2,1} & Y_{0,2,1,0} & Y_{0,2,1,1} & Y_{0,2,2,0} & Y_{0,2,2,1} \\
Y_{0,1,1,1} & Y_{0,1,1,2} & Y_{0,1,2,1} & Y_{0,1,2,2} & Y_{0,2,1,1} & Y_{0,2,1,2} & Y_{0,2,2,1} & Y_{0,2,2,2}
\end{array}\right]
$$

$$\mathcal{H}_1^{3D} = \left[\begin{array}{cc|cc|cc|cc}
Y_{1,0,0,0} & Y_{1,0,0,1} & Y_{1,0,1,0} & Y_{1,0,1,1} & Y_{1,1,0,0} & Y_{1,1,0,1} & Y_{1,1,1,0} & Y_{1,1,1,1} \\
Y_{1,0,0,1} & Y_{1,0,0,2} & Y_{1,0,1,1} & Y_{1,0,1,2} & Y_{1,1,0,1} & Y_{1,1,0,2} & Y_{1,1,1,1} & Y_{1,1,1,2} \\
\hline
Y_{1,0,1,0} & Y_{1,0,1,1} & Y_{1,0,2,0} & Y_{1,0,2,1} & Y_{1,1,1,0} & Y_{1,1,1,1} & Y_{1,1,2,0} & Y_{1,1,2,1} \\
Y_{1,0,1,1} & Y_{1,0,1,2} & Y_{1,0,2,1} & Y_{1,0,2,2} & Y_{1,1,1,1} & Y_{1,1,1,2} & Y_{1,1,2,1} & Y_{1,1,2,2} \\
\hline
Y_{1,1,0,0} & Y_{1,1,0,1} & Y_{1,1,1,0} & Y_{1,1,1,1} & Y_{1,2,0,0} & Y_{1,2,0,1} & Y_{1,2,1,0} & Y_{1,2,1,1} \\
Y_{1,1,0,1} & Y_{1,1,0,2} & Y_{1,1,1,1} & Y_{1,1,1,2} & Y_{1,2,0,1} & Y_{1,2,0,2} & Y_{1,2,1,1} & Y_{1,2,1,2} \\
\hline
Y_{1,1,1,0} & Y_{1,1,1,1} & Y_{1,1,2,0} & Y_{1,1,2,1} & Y_{1,2,1,0} & Y_{1,2,1,1} & Y_{1,2,2,0} & Y_{1,2,2,1} \\
Y_{1,1,1,1} & Y_{1,1,1,2} & Y_{1,1,2,1} & Y_{1,1,2,2} & Y_{1,2,1,1} & Y_{1,2,1,2} & Y_{1,2,2,1} & Y_{1,2,2,2}
\end{array}\right]$$

$$\mathcal{H}_2^{3D} = \left[\begin{array}{cc|cc|cc|cc}
Y_{2,0,0,0} & Y_{2,0,0,1} & Y_{2,0,1,0} & Y_{2,0,1,1} & Y_{2,1,0,0} & Y_{2,1,0,1} & Y_{2,1,1,0} & Y_{2,1,1,1} \\
Y_{2,0,0,1} & Y_{2,0,0,2} & Y_{2,0,1,1} & Y_{2,0,1,2} & Y_{2,1,0,1} & Y_{2,1,0,2} & Y_{2,1,1,1} & Y_{2,1,1,2} \\
\hline
Y_{2,0,1,0} & Y_{2,0,1,1} & Y_{2,0,2,0} & Y_{2,0,2,1} & Y_{2,1,1,0} & Y_{2,1,1,1} & Y_{2,1,2,0} & Y_{2,1,2,1} \\
Y_{2,0,1,1} & Y_{2,0,1,2} & Y_{2,0,2,1} & Y_{2,0,2,2} & Y_{2,1,1,1} & Y_{2,1,1,2} & Y_{2,1,2,1} & Y_{2,1,2,2} \\
\hline
Y_{2,1,0,0} & Y_{2,1,0,1} & Y_{2,1,1,0} & Y_{2,1,1,1} & Y_{2,2,0,0} & Y_{2,2,0,1} & Y_{2,2,1,0} & Y_{2,2,1,1} \\
Y_{2,1,0,1} & Y_{2,1,0,2} & Y_{2,1,1,1} & Y_{2,1,1,2} & Y_{2,2,0,1} & Y_{2,2,0,2} & Y_{2,2,1,1} & Y_{2,2,1,2} \\
\hline
Y_{2,1,1,0} & Y_{2,1,1,1} & Y_{2,1,2,0} & Y_{2,1,2,1} & Y_{2,2,1,0} & Y_{2,2,1,1} & Y_{2,2,2,0} & Y_{2,2,2,1} \\
Y_{2,1,1,1} & Y_{2,1,1,2} & Y_{2,1,2,1} & Y_{2,1,2,2} & Y_{2,2,1,1} & Y_{2,2,1,2} & Y_{2,2,2,1} & Y_{2,2,2,2}
\end{array}\right]$$

7.3.1 n-Dimensional Block Hankel Matrix Decomposition

In this section, we generalize the 2D block hankel matrix decomposition shown in [109, 110] and extend it to n-dimensional block Hankel matrix. Consider *n-Dimensional Block Hankel* matrix $\mathcal{H}^{nD}$ as (7.5) constructed with n-dimensional data array $Y \in R^{l_n \times \ldots \times l_2 \times l_1}$.

Theorem 31. *n-dimensional* Block Hankel *matrix $\mathcal{H}^{nD}$ can be decomposed as*

$$\mathcal{H}^{nD} = E_{n_L} A E_{n_R} \tag{7.9}$$

where, A, E_L^n, and E_R^n are the matrices of the form

$$A = \begin{bmatrix} a_1 & 0 & .. & 0 \\ 0 & a_2 & .. & 0 \\ 0 & 0 & .. & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & ... & a_N \end{bmatrix} \; Z_n = \begin{bmatrix} z_{n_1} & 0 & .. & 0 \\ 0 & z_{n_2} & .. & 0 \\ 0 & 0 & .. & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & .. & z_{n_N} \end{bmatrix} \tag{7.10}$$

$$E_{n_L} = \begin{bmatrix} E_{n-1_L} Z_n^0 \\ E_{n-1_L} Z_n^1 \\ \vdots \\ E_{n-1_L} Z_n^{M_n - 1} \end{bmatrix}, E_{n_R} = \begin{bmatrix} Z_n^0 E_{n-1_R} & Z_n^1 E_{n-1_R} & ... & Z_n^{l_n - M_n} E_{n-1_R} \end{bmatrix} \tag{7.11}$$

$$E_{1_L} = \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \\ z_{1_1}^2 & z_{1_2}^2 & .. & z_{1_N}^2 \\ \vdots & \vdots & & \vdots \\ z_{1_1}^{M_1-1} & z_{1_2}^{M_1-1} & .. & z_{1_N}^{M_1-1} \end{bmatrix} E_{1_R} = \begin{bmatrix} 1 & z_{1_1} & z_{1_1}^2 & .. & z_{1_1}^{l_1-M_1} \\ 1 & z_{1_2} & z_{1_2}^2 & .. & z_{1_2}^{l_1-M_1} \\ \vdots & \vdots & \vdots & & \\ 1 & z_{1_N} & z_{1_N}^2 & .. & z_{1_N}^{l_1-M_1} \end{bmatrix} \tag{7.12}$$

(in fact, $E_{n_R} = E_{n_L}^T$ where M_i is replaced by $l_i - M_i + 1$)

Proof. See Appendix A. □

Consider the following examples.

Example 1 Consider 2D data array $Y \in R^{3\times 3}$ as

$$Y_{k_2,k_1} = \sum_{i=1}^{N=2} a_i z_{2_i}^{k_2} z_{1_i}^{k_1}, \quad k_1, k_2 = 0, 1, 2$$

Hankel matrix $\mathcal{H}^{2D}$ with $M_1 = M_2 = 2$, takes the form of

$$\mathcal{H}^{2D} = \left[\begin{array}{cc|cc} Y_{0,0} & Y_{0,1} & Y_{1,0} & Y_{1,1} \\ Y_{0,1} & Y_{0,2} & Y_{1,1} & Y_{1,2} \\ ---- & ---- & ---- & ---- \\ Y_{1,0} & Y_{1,1} & Y_{2,0} & Y_{2,1} \\ Y_{1,1} & Y_{1,2} & Y_{2,1} & Y_{2,2} \end{array}\right]$$

$$= \left[\begin{array}{cc|cc} a_1 z_{2_1}^0 z_{1_1}^0 + a_2 z_{2_2}^0 z_{1_2}^0 & a_1 z_{2_1}^0 z_{1_1}^1 + a_2 z_{2_2}^0 z_{1_2}^1 & a_1 z_{2_1}^1 z_{1_1}^0 + a_2 z_{2_2}^1 z_{1_2}^0 & a_1 z_{2_1}^1 z_{1_1}^1 + a_2 z_{2_2}^1 z_{1_2}^1 \\ a_1 z_{2_1}^0 z_{1_1}^1 + a_2 z_{2_2}^0 z_{1_2}^1 & a_1 z_{2_1}^0 z_{1_1}^2 + a_2 z_{2_2}^0 z_{1_2}^2 & a_1 z_{2_1}^1 z_{1_1}^1 + a_2 z_{2_2}^1 z_{1_2}^1 & a_1 z_{2_1}^1 z_{1_1}^2 + a_2 z_{2_2}^1 z_{1_2}^2 \\ ---- & ---- & ---- & ---- \\ a_1 z_{2_1}^1 z_{1_1}^0 + a_2 z_{2_2}^1 z_{1_2}^0 & a_1 z_{2_1}^1 z_{1_1}^1 + a_2 z_{2_2}^1 z_{1_2}^1 & a_1 z_{2_1}^2 z_{1_1}^0 + a_2 z_{2_2}^2 z_{1_2}^0 & a_1 z_{2_1}^2 z_{1_1}^1 + a_2 z_{2_2}^2 z_{1_2}^1 \\ a_1 z_{2_1}^1 z_{1_1}^1 + a_2 z_{2_2}^1 z_{1_2}^1 & a_1 z_{2_1}^1 z_{1_1}^2 + a_2 z_{2_2}^1 z_{1_2}^2 & a_1 z_{2_1}^2 z_{1_1}^1 + a_2 z_{2_2}^2 z_{1_2}^1 & a_1 z_{2_1}^2 z_{1_1}^2 + a_2 z_{2_2}^2 z_{1_2}^2 \end{array}\right]$$

$$= \left[\begin{array}{c|c} \begin{bmatrix} 1 & 1 \\ z_{1_1} & z_{1_2} \end{bmatrix} \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} \begin{bmatrix} z_{2_1}^0 & 0 \\ 0 & z_{2_2}^0 \end{bmatrix} \begin{bmatrix} 1 & z_{1_1} \\ 1 & z_{1_2} \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ z_{1_1} & z_{1_2} \end{bmatrix} \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} \begin{bmatrix} z_{2_1}^1 & 0 \\ 0 & z_{2_2}^1 \end{bmatrix} \begin{bmatrix} 1 & z_{1_1} \\ 1 & z_{1_2} \end{bmatrix} \\ ---- & ---- \\ \begin{bmatrix} 1 & 1 \\ z_{1_1} & z_{1_2} \end{bmatrix} \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} \begin{bmatrix} z_{2_1}^1 & 0 \\ 0 & z_{2_2}^1 \end{bmatrix} \begin{bmatrix} 1 & z_{1_1} \\ 1 & z_{1_2} \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ z_{1_1} & z_{1_2} \end{bmatrix} \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} \begin{bmatrix} z_{2_1}^2 & 0 \\ 0 & z_{2_2}^2 \end{bmatrix} \begin{bmatrix} 1 & z_{1_1} \\ 1 & z_{1_2} \end{bmatrix} \end{array}\right]$$

$$= \left[\begin{array}{c} \begin{bmatrix} 1 & 1 \\ z_{1_1} & z_{1_2} \end{bmatrix} \begin{bmatrix} z_{2_1}^0 & 0 \\ 0 & z_{2_2}^0 \end{bmatrix} \\ ---- \\ \begin{bmatrix} 1 & 1 \\ z_{1_1} & z_{1_2} \end{bmatrix} \begin{bmatrix} z_{2_1}^1 & 0 \\ 0 & z_{2_2}^1 \end{bmatrix} \end{array}\right] \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} \left[\begin{array}{c|c} \begin{bmatrix} z_{2_1}^0 & 0 \\ 0 & z_{2_2}^0 \end{bmatrix} \begin{bmatrix} 1 & z_{1_1} \\ 1 & z_{1_2} \end{bmatrix} & \begin{bmatrix} z_{2_1}^1 & 0 \\ 0 & z_{2_2}^1 \end{bmatrix} \begin{bmatrix} 1 & z_{1_1} \\ 1 & z_{1_2} \end{bmatrix} \end{array}\right]$$

$$= E_{2_L} A E_{2_R}$$

where, E_{2_R}, E_{2_L}, A and Z_2 are as in (7.10).

7.3.2 Row Permutation of Block Hankel Matrix

In this section, we generalize the row permutation of 2D block hankel matrix shown in [109] and extend it to n-dimensional block Hankel matrix. Consider n-dimensional *Block Hankel* matrix $\mathcal{H}^{nD}$ constructed by n-dimensional data array $Y \in \mathbb{R}^{l_n \times \ldots \times l_1 \times l_1}$

as $\mathcal{H}^{nD} = E_{n_L} A E_{n_R}$. The matrices E_{n_L} and E_{n_R} contains all frequencies $\{z_{1_i}, i = 1, ..., N\}, \{z_{2_i}, i = 1, ..., N\}, ..., \{z_{n_i}, i = 1, ..., N\}$. Using row permutation one can change the position of frequencies in one dimension with one another. More precisely, there exist row permutation matrix P_j^r such that the position of frequencies in rth dimension $\{z_{r_i}, i = 1, ..., N\}$ in the matrix E_{n_L} (or E_{n_R}) is same as the position of the frequencies in jth dimension$\{z_{j_i}, i = 1, ..., N\}$ in the shuffled matrix $\hat{E}_{n_L} = P_j^r E_{n_L}$ (or $\hat{E}_{n_R} = P_j^r E_{n_R}$) and also the position of frequencies in jth dimension $\{z_{j_i}, i = 1, ..., N\}$ in the matrix E_{n_L} (or E_{n_R}) is same as frequencies in rth dimension$\{z_{r_i}, i = 1, ..., N\}$ in the shuffled matrix $\hat{E}_{n_L}$ (or $\hat{E}_{n_R}$). In this case, matrices of the form (Vandermond)

$$V_i = \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{i_1} & z_{i_2} & .. & z_{i_N} \\ z_{i_1}^2 & z_{i_2}^2 & .. & z_{i_N}^2 \\ \vdots & \vdots & & \vdots \\ z_{i_1}^{M_i-1} & z_{i_2}^{M_i-1} & .. & z_{i_N}^{M_i-1} \end{bmatrix}, i = 1, ..., n$$

are sub-matrices of E_{n_L} and its row shuffled matrices $\hat{E}_{n_L}$.

Consider the following Examples, due to the similarity of E_{n_L} and E_{n_R}, we just consider the E_{n_L} in the examples.

Example: Consider the E_{2_L} in 2D block Hankel matrix $\mathcal{H}^{2D}$ as

$$E_{2_L} = \begin{bmatrix} E_{1_L} Z_2^0 \\ ---- \\ E_{1_L} Z_2^1 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \end{bmatrix} Z_2^0 \\ ---- \\ \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \end{bmatrix} Z_2^1 \end{bmatrix} \tag{7.13}$$

By expanding the matrices, E_{2_L} takes the form

$$E_{2_L} = \begin{bmatrix} \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \end{bmatrix} \\ ---- \\ \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \end{bmatrix} \begin{bmatrix} z_{2_1} & 0 & \dots & 0 \\ 0 & z_{2_2} & \dots & 0 \\ 0 & 0 & \dots & z_{2_N} \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_{1_1} & z_{1_2} & \dots & z_{1_N} \end{bmatrix} \\ ---- \\ \begin{bmatrix} z_{2_1} & z_{2_2} & \dots & z_{2_N} \\ z_{1_1} z_{2_1} & z_{1_2} z_{2_2} & \dots & z_{1_N} z_{2_N} \end{bmatrix} \end{bmatrix}$$

By row permutation of matrix E_{2_L}, one obtains $\hat{E}_{2_L}$ as

$$\hat{E}_{2_L} = \begin{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_{2_1} & z_{2_2} & \dots & z_{2_N} \end{bmatrix} \\ ---- \\ \begin{bmatrix} z_{1_1} & z_{1_2} & \dots & z_{1_N} \\ z_{1_1} z_{2_1} & z_{1_2} z_{2_2} & \dots & z_{1_N} z_{2_N} \end{bmatrix} \end{bmatrix} =$$

$$\begin{bmatrix} \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{2_1} & z_{2_2} & .. & z_{2_N} \end{bmatrix} \\ ---- \\ \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{2_1} & z_{2_2} & .. & z_{2_N} \end{bmatrix} \begin{bmatrix} z_{1_1} & 0 & \dots & 0 \\ 0 & z_{1_2} & \dots & 0 \\ 0 & 0 & \dots & z_{1_N} \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{2_1} & z_{2_2} & .. & z_{2_N} \end{bmatrix} Z_1^0 \\ ---- \\ \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{2_1} & z_{2_2} & .. & z_{2_N} \end{bmatrix} Z_1^1 \end{bmatrix} \tag{7.14}$$

The position of frequencies of first dimension $\{z_{1_i}, i = 1, ..., N\}$ in the matrix E_{2_L} (7.13) is same as the position of frequencies of second dimension $\{z_{2_i}, i = 1, ..., N\}$ in $\hat{E}_{2_L}$ (7.14) and the position of z_{2_i} in E_{2_L} is same as the position of z_{1_i} in $\hat{E}_{2_L}$ [109]. Also, matrices of the form

$$V_1 = \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \end{bmatrix}, V_2 = \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{2_1} & z_{2_2} & .. & z_{2_N} \end{bmatrix}$$

are sub-matrices of E_{2_L} and its row shuffled matrices $\hat{E}_{2_L}$, respectively.

Example: Consider the E_{3_L} in 3D block Hankel matrix $\mathcal{H}^{3D}$ as

$$E_{3_L} = \begin{bmatrix} E_{2_L} Z_3^0 \\ ---- \\ E_{2_L} Z_3^1 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \end{bmatrix} Z_2^0 \\ ---- \\ \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \end{bmatrix} Z_2^1 \end{bmatrix} Z_3^0 \\ ------------ \\ \begin{bmatrix} \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \end{bmatrix} Z_2^0 \\ ---- \\ \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \end{bmatrix} Z_2^1 \end{bmatrix} Z_3^1 \end{bmatrix} \tag{7.15}$$

By expanding the matrices, E_{3_L} takes the form

$$E_{3_L} = \begin{bmatrix} \begin{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_{1_1} & z_{1_2} & \dots & z_{1_N} \end{bmatrix} \\ -\,-\,-\,- \\ \begin{bmatrix} z_{2_1} & z_{2_2} & \dots & z_{2_N} \\ z_{1_1}z_{2_1} & z_{1_2}z_{2_2} & \dots & z_{1_N}z_{2_N} \end{bmatrix} \end{bmatrix} Z_3^0 \\ -\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,- \\ \begin{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_{1_1} & z_{1_2} & \dots & z_{1_N} \end{bmatrix} \\ -\,-\,-\,- \\ \begin{bmatrix} z_{2_1} & z_{2_2} & \dots & z_{2_N} \\ z_{1_1}z_{2_1} & z_{1_2}z_{2_2} & \dots & z_{1_N}z_{2_N} \end{bmatrix} \end{bmatrix} Z_3^1 \end{bmatrix} =$$

$$\begin{bmatrix} \begin{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_{1_1} & z_{1_2} & \dots & z_{1_N} \end{bmatrix} \\ -\,-\,-\,- \\ \begin{bmatrix} z_{2_1} & z_{2_2} & \dots & z_{2_N} \\ z_{1_1}z_{2_1} & z_{1_2}z_{2_2} & \dots & z_{1_N}z_{2_N} \end{bmatrix} \end{bmatrix} \\ -\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,- \\ \begin{bmatrix} \begin{bmatrix} z_{3_1} & z_{3_2} & \dots & z_{3_N} \\ z_{1_1}z_{3_1} & z_{1_2}z_{3_2} & \dots & z_{1_N}z_{3_N} \end{bmatrix} \\ -\,-\,-\,- \\ \begin{bmatrix} z_{2_1}z_{3_1} & z_{2_2}z_{3_2} & \dots & z_{2_N}z_{3_N} \\ z_{1_1}z_{2_1}z_{3_1} & z_{1_2}z_{2_2}z_{3_2} & \dots & z_{1_N}z_{2_N}z_{3_N} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

By row permutation of matrix E_{3_L}, one can change the position of frequencies in the first, second and third dimension with one another. For example, we aim at changing the position of frequencies in the first and third dimension with one another by following row permutation of matrix E_{3_L}.

$$\hat{E}_{3_L} = \begin{bmatrix} \begin{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_{3_1} & z_{3_2} & \dots & z_{3_N} \end{bmatrix} \\ ---- \\ \begin{bmatrix} z_{2_1} & z_{2_2} & \dots & z_{2_N} \\ z_{3_1}z_{2_1} & z_{3_2}z_{2_2} & \dots & z_{3_N}z_{2_N} \end{bmatrix} \end{bmatrix} \\ ------------ \\ \begin{bmatrix} \begin{bmatrix} z_{1_1} & z_{1_2} & \dots & z_{1_N} \\ z_{1_1}z_{3_1} & z_{1_2}z_{3_2} & \dots & z_{1_N}z_{3_N} \end{bmatrix} \\ ---- \\ \begin{bmatrix} z_{2_1}z_{1_1} & z_{2_2}z_{1_2} & \dots & z_{2_N}z_{1_N} \\ z_{1_1}z_{2_1}z_{3_1} & z_{1_2}z_{2_2}z_{3_2} & \dots & z_{1_N}z_{2_N}z_{3_N} \end{bmatrix} \end{bmatrix} \end{bmatrix} =$$

$$\begin{bmatrix} \begin{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_{3_1} & z_{3_2} & \dots & z_{3_N} \end{bmatrix} \\ ---- \\ \begin{bmatrix} z_{2_1} & z_{2_2} & \dots & z_{2_N} \\ z_{3_1}z_{2_1} & z_{3_2}z_{2_2} & \dots & z_{3_N}z_{2_N} \end{bmatrix} \end{bmatrix} Z_1^0 \\ ------------ \\ \begin{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_{3_1} & z_{3_2} & \dots & z_{3_N} \end{bmatrix} \\ ---- \\ \begin{bmatrix} z_{2_1} & z_{2_2} & \dots & z_{2_N} \\ z_{3_1}z_{2_1} & z_{3_2}z_{2_2} & \dots & z_{3_N}z_{2_N} \end{bmatrix} \end{bmatrix} Z_1^1 \end{bmatrix}$$

$$
= \begin{bmatrix}
\begin{bmatrix}
\begin{bmatrix} 1 & 1 & .. & 1 \\ z_{3_1} & z_{3_2} & .. & z_{3_N} \end{bmatrix} Z_2^0 \\
- - - - \\
\begin{bmatrix} 1 & 1 & .. & 1 \\ z_{3_1} & z_{3_2} & .. & z_{3_N} \end{bmatrix} Z_2^1
\end{bmatrix} Z_1^0 \\
- - - - - - - - - - - - \\
\begin{bmatrix}
\begin{bmatrix} 1 & 1 & .. & 1 \\ z_{3_1} & z_{3_2} & .. & z_{3_N} \end{bmatrix} Z_2^0 \\
- - - - \\
\begin{bmatrix} 1 & 1 & .. & 1 \\ z_{3_1} & z_{3_2} & .. & z_{3_N} \end{bmatrix} Z_2^1
\end{bmatrix} Z_1^1
\end{bmatrix} \tag{7.16}
$$

As it is seen, the position of frequencies of first dimension $\{z_{1_i}, i = 1, ..., N\}$ in the matrix E_{3_L} (7.15) is same as the position of frequencies of third dimension $\{z_{3_i}, i = 1, ..., N\}$ in $\hat{E}_{3_L}$ (7.16) and the position of z_{3_i} in E_{3_L} is same as position of z_{1_i} in $\hat{E}_{3_L}$. Also, matrices of the form

$$
V_1 = \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \end{bmatrix}, V_3 = \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{3_1} & z_{3_2} & .. & z_{3_N} \end{bmatrix}
$$

are sub-matrices of E_{3_L} and its row shuffled matrices $\hat{E}_{3_L}$, respectively. In similar way, by changing the position of the second dimension frequencies $\{z_{2_i}, i = 1, ..., N\}$ with first dimension frequencies $\{z_{1_i}, i = 1, ..., N\}$, it can be shown that the matrix

$$
V_2 = \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{2_1} & z_{2_2} & .. & z_{2_N} \end{bmatrix}
$$

is sub-matrix of resulted row shuffled matrix $\hat{E}_{3_L}$.

7.3.3 Rank of Block Hankel Matrix

In [109], it is shown that rank of 2D hankel matrix is equivalent to the number exponential signals that describes the data. Here, we extend the results to n-dimensional Hankel matrix. Consider *n-D Block Hankel* matrix $\mathcal{H}^{nD}$ constructed from n-dimensional data array $Y \in \mathbb{R}^{l_n \times \ldots \times l_2 \times l_1}$ as $\mathcal{H}^{nD} = E_{n_L} A E_{n_R}$. We assume that given n-dimensional data can be written in terms of exponential signal as in (7.1).

Theorem 32. *Rank of $n-$dimensional Hankel matrix $\mathcal{H}^{nD}$ is equal to the number of exponential signals that describes the data if*

$$N \leq M_i \leq l_i - N + 1$$

where, N is the number of exponential signals that describes the data, M_i and l_i are window size of $n-$dimensional Hankel matrix and number of measurement for i-th dimension.

Proof. See Appendix B. □

7.4 Equivalent Problem and Convex Relaxation

As an intermediate step in the development of convex relaxation of the original problem, a equivalent problem is provided. This is achieved by solving the following problem:

Problem 2: Solve

$$min_{Y \in \mathbb{R}^{l_n \times \ldots \times l_2 \times l_1}} \{ \;\; Rank(\; \mathcal{H}^{nD}(Y) \;) : \|Y_{k_n^*, \ldots, k_2^*, k_1^*} - \hat{Y}_{k_n^*, \ldots, k_2^*, k_1^*}\|_2 \leq \epsilon \} \tag{7.17}$$

where, $Y \in \mathbb{R}^{l_n \times \ldots \times l_2 \times l_1}$ is constructed noiseless complete n-dimensional data, $\mathcal{H}^{nD}(Y)$ is a n dimensional block Hankel matrix of the form (7.5), $\hat{Y} \in \mathbb{R}^{l_n \times \ldots \times l_2 \times l_1}$ is given sparse noisy n-dimensional sensory data, $(k_1^*, k_2^*, \ldots, k_n^*)$ are indexes of known parts of measurement data $\hat{Y}$, and $\epsilon > 0$.

Theorem 33. *Problem 2 is equivalent to Problem 1.*

Proof. In Theorem 32, we showed that rank of n-dimensional Hankel matrix is equivalent to the number of exponentiation describing the data. Hence, minimizing the rank of n-dimensional Hankel matrix in Problem 2 is equivalent to minimizing the number of exponential signals in Problem 1. □

Equivalent Problem 2 involves nonconvex problem of rank minimization. To avoid this, we use nuclear norm of matrix as a convex relaxation of rank of matrix. Therefore, convex relaxation of Problem 2 is as follow:

Problem 3: Solve

$$min_{Y \in \mathbb{R}^{l_n \times \ldots \times l_2 \times l_1}} \{ \|\mathcal{H}(Y)\|_* : \|Y_{k_n^*,\ldots,k_2^*,k_1^*} - \hat{Y}_{k_n^*,\ldots,k_2^*,k_1^*}\|_2 \leq \epsilon \} \tag{7.18}$$

where $\|.\|_*$ stands for nuclear norm of a matrix.

7.5 Implementation and Numerical Results

In the previous sections, we showed that sparse and noisy data reconstruction problem can be reformulated as nuclear norm minimization problem. To be able to deal with large scale data, one need to implement efficient and fast convex optimization algorithm. Recently, first-order augmented Lagrangian algorithm has been proposed to deal with large semidefinite programs. We adapt this algorithm to solve resulting convex optimization problem as (7.18).

To be able to use convex optimization methods to solve nuclear norm minimization Problem 3 which contains linear structured Hankel matrices, one needs to reformulated the Problem 3 as follows.

Problem 4: Solve

$$min_{Y,H,y} \{ \|H\|_* : \mathcal{H}(Y) - H = 0, \mathcal{A}(Y) - b - s = 0, s \in \mathcal{Q} \} \tag{7.19}$$

where, $Y \in R^{l_1 \times l_2 \ldots \times l_n}$ is complete measurement matrix and $\mathcal{H}(Y)$ is associated block Hankel matrix, H is a matrix with appropriate dimension, $\mathcal{A}$ is a linear operator such that $\mathcal{A}(Y) = Y_{k_n^*,\ldots,k_2^*,k_1^*}$, b is vector of known sensory measurement as $\hat{Y}_{k_n^*,\ldots,k_2^*,k_1^*}$, s is a slack variable and $\mathcal{Q}$ is a closed convex set of the form $\mathcal{Q} = \{s : \|s\| \leq \epsilon\}$, and$\|.\|_*$ stands for nuclear norm.

Table 7.1: First-Order Augmented Lagrangian Optimization Algorithm

1: $s^{(0)} \leftarrow \mathcal{A}(Y^{(0)}) - b$, $\eta \leftarrow \|H^{(0)}\|_*$, $\theta_1^{(1)} \leftarrow 0$, $\theta_2^{(1)} \leftarrow 0$, $k \leftarrow 0$
2: While (FALC STOP is false) do
3: $k \leftarrow k+1$
4: $f^{(k)}(H,Y,s) := \frac{1}{2}\|\mathcal{H}(Y) - H - \lambda^{(k)}\theta_1^{(k)}\|$
$+\frac{1}{2}\|\mathcal{A}(Y) - b - s - \lambda^{(k)}\theta_2^{(k)}\|$
4: $\eta_1^{(k)} \leftarrow \eta + \frac{\lambda^{(k)}}{2}(\|\theta_1^{(k)}\|_2^2 + \|\theta_2^{(k)}\|_2^2)$
5: $H_1^{(0)} \leftarrow H^{(0)}$, $H_2^{(1)} \leftarrow H^{(0)}$, $Y_1^{(0)} \leftarrow Y^{(0)}$, $Y_2^{(1)} \leftarrow Y^{(0)}$
$s_1^{(0)} \leftarrow s^{(0)}$, $s_2^{(1)} \leftarrow s^{(0)}$, $t^{(1)} = 1, l = 0$
7: While (APG STOP is false) do
8: $l \leftarrow l+1$
9: $[H_1^{(1)}, Y_1^{(1)}, s_1^{(1)}] \leftarrow argmin$

$$\left\{ \begin{array}{l} \lambda^{(k)}\|H\|_* + \begin{bmatrix} \nabla_H f^{(k)}(H_2^{(l)}, Y_2^{(l)}, s_2^{(l)}) \\ \nabla_Y f^{(k)}(H_2^{(l)}, Y_2^{(l)}, s_2^{(l)}) \\ \nabla_s f^{(k)}(H_2^{(l)}, Y_2^{(l)}, s_2^{(l)}) \end{bmatrix}^T \begin{bmatrix} H - H_2^{(l)} \\ Y - Y_2^{(l)} \\ s - s_2^{(l)} \end{bmatrix} \\ + \frac{L_Y}{2}\|Y - Y_2^{(l)}\| + \frac{L_H}{2}\|H - H_2^{(l)}\| + \frac{L_s}{2}\|s - s_2^{(l)}\| : s \in \mathcal{Q} \end{array} \right\}$$

10: $t^{(l+1)} \leftarrow (1 + \sqrt{1 + 4(t^{(l)})^2})/2$
11: $H_2^{(l+1)} \leftarrow H_1^{(l)} + (\frac{t^{(l)} - 1}{t^{(l+1)}})(H_1^{(l)} - H_1^{(l-1)})$
$Y_2^{(l+1)} \leftarrow Y_1^{(l)} + (\frac{t^{(l)} - 1}{t^{(l+1)}})(Y_1^{(l)} - Y_1^{(l-1)})$
$s_2^{(l+1)} \leftarrow s_1^{(l)} + (\frac{t^{(l)} - 1}{t^{(l+1)}})(s_1^{(l)} - s_1^{(l-1)})$
12: end APG while
13: $H^{(k)} \leftarrow H_1^{(l)}, Y^{(k)} \leftarrow Y_1^{(l)}, s^{(k)} \leftarrow s_1^{(l)}$
14: $\theta_1^{(k+1)} \leftarrow \theta_1^{(k)} - \frac{\mathcal{H}(Y^{(k)}) - H^{(k)}}{\lambda^{(k)}}$
$\theta_2^{(k+1)} \leftarrow \theta_2^{(k)} - \frac{\mathcal{A}(Y^{(k)}) - b - s^{(k)}}{\lambda^{(k)}}$
15: end FALC while

7.5.1 First-Order Augmented Lagrangian Optimization Algorithm

First-order augmented Lagrangian algorithm (FALC)is shown in Table 7.1. FALC derives the solution of the nuclear norm minimization problem of the form of Problem

4, by inexactly solving a sequence of sub problems of the form:

$$min_{Y,H,s\in\mathcal{Q}}\{\lambda\|H\|_* + \frac{1}{2}\|\mathcal{H}(Y) - H - \lambda\theta_1\|_2^2 + \frac{1}{2}\|\mathcal{A}(Y) - b - s - \lambda\theta_2\|_2^2\} \tag{7.20}$$

where, λ and θ_i are penalty parameter and Lagrangian dual variables, respectively. These sub problems are solved using an Accelerated Proximal Gradient (APG) algorithm, where in each update it solves the problem of the form:

$$min_{Y,H,s\in\mathcal{Q}} \left\{ \begin{array}{c} \lambda\|H\|_* + \begin{bmatrix} \nabla_H f(\tilde{H},\tilde{Y},\tilde{s}) \\ \nabla_Y f(\tilde{H},\tilde{Y},\tilde{s}) \\ \nabla_s f(\tilde{H},\tilde{Y},\tilde{s}) \end{bmatrix}^T \begin{bmatrix} H-\tilde{H} \\ Y-\tilde{Y} \\ s-\tilde{s} \end{bmatrix} \\ +\frac{L}{2}\|Y-\tilde{Y}\|_2^2 + \frac{L}{2}\|H-\tilde{H}\|_F^2 + \frac{L}{2}\|s-\tilde{s}\|_2^2 \end{array} \right\} \tag{7.21}$$

where, $f(H,Y,s) = \frac{1}{2}\|\mathcal{H}(Y) - H - \lambda\theta_1\|_2^2 + \frac{1}{2}\|\mathcal{A}(Y) - b - s - \lambda\theta_2\|_2^2$ for a given $(\tilde{H},\tilde{Y},\tilde{s})$, and L is Lipschitz constant. If $\sigma_{max}(.)$ denotes the largest singular value of a matrix, then $L = \sigma_{max}^2(\mathcal{H}) + \sigma_{max}^2(\mathcal{A})$. The gradients can be obtained as $\nabla_H f = -(\mathcal{H}(Y) - H - \lambda\theta_1)$, $\nabla_Y f = -\mathcal{H}^*(\mathcal{H}(Y) - H - \lambda\theta_1) - \mathcal{A}^*(\mathcal{A}(Y) - b - s - \lambda\theta_2)$, $\nabla_s f = -\mathcal{A}^*(\mathcal{A}(Y) - b - s - \lambda\theta_2)$. The problem of the form (7.21) is separable in H, Y, and s variables and reduces to *constrained shrinkage* problem in H and Y, and *Euclidean projection* problem onto $\mathcal{Q}$ in s.

7.5.2 Numerical Examples

We now consider the problem of data loss for different types of data. To complete the given noisy and lost data we solve, convex problem in (7.18) using the first older algorith provided in section 7.5.

Example 1: 1-dimensional data Consider given signal in Fig 7.1. The aim is to complete the sampled sparse signal. The obtained signal by solving the convex optimization problem provided in section 7.5 is shown in Fig 7.2. Also, Fig 7.2 shows the nonzero singular values of Hankel matrix of obtained data where corresponds to rank of matrix.

Example 2: 2-dimensional data Consider given sparse 2-dimensional signal in Fig 7.3. The aim is to complete the sparse signal using provided convex optimization method. The obtained signal and The singular values of Hankel matrix constructed by obtained signal is shown in Fig 7.4.

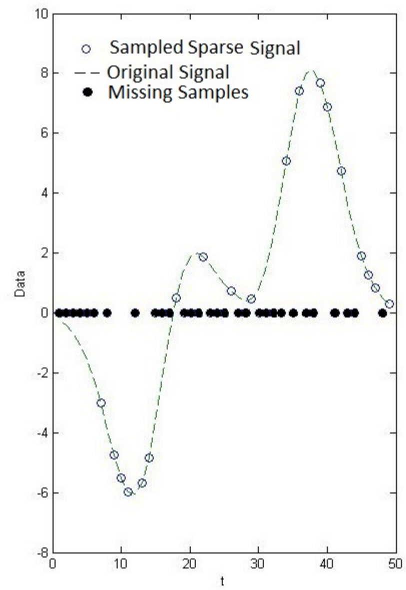

Figure 7.1: Original and sparse signal of Example 1

Example 3: Corrupted Image Consider corrupted image in Fig 7.5. The aim is to reconstruct the corrupted signal using provided convex optimization method. The singular values of Hankel matrix constructed by sparse signal is shown in Fig 7.6. The obtained signal is shown in Fig 7.5. As shown in Fig 7.6 the nuclear rank of Hankel matrix of reconstructed signal is decreased.

Example 4: Lost and Noisy Sensor Data We apply proposed method on four set of sensory data which contains light and temperature information of ocean and a indoor place. The indoor experiment contains 49 nodes placed in a room which each node reports the temperature and light data over the time. The measured data for 149 sampling time, forms two 49×149 measurement matrices for light and temperature as in Fig 7.7 and 7.8. The Ocean experiment contains 10 nodes deployed in the sea which each node reports temperature, and light data over the time. The obtained data for 42 sampling time, constructs two 10×42 measurement matrices for light and temperature as in Fig 7.7 and 7.8. To show the performance of the proposed method, first, each measurement matrix is corrupted by randomly dropping of the elements; then, the proposed method is used to find the missing parts of each measurement matrix. In order to measure the error of reconstructed data, errors on the missing data are used to define a *Error Ratio* (ER) notion as

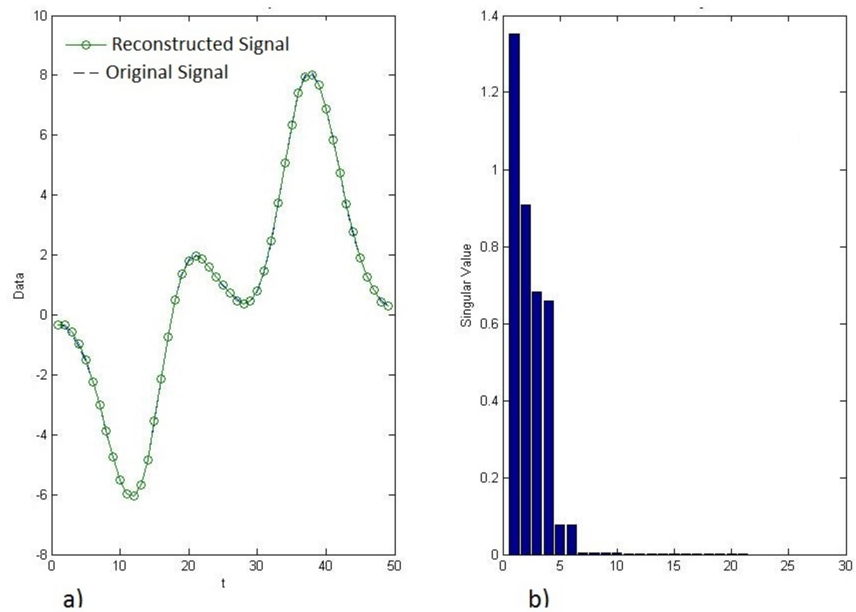

Figure 7.2: Results of Example 1: a) Reconstructed signal, b) Singular values of Hankel matrix of reconstructed sparse signal

$ER = \dfrac{\sqrt{\sum_{(k_1,k_2)\neq(k_1^*,k_2^*)}(Y(k_1,k_2)-\hat{Y}(k_1,k_2))^2}}{\sqrt{\sum_{(k_1,k_2)\neq(k_1^*,k_2^*)}(Y(k_1,k_2))^2}}$ is used. The results for 50% loss problem for indoor and ocean light temperature are shown in Fig 7.9 to 7.12. It shows singular values of block Hankel matrix constructed by corrupted and reconstructed data, initial error data and obtained error data, the Error ratio and time at each FALC iteration.

7.6 Conclusion

In this work, we presented a novel approach for solving the problem of reconstructing spars and noisy data. The proposed method reformulates the problem as a minimum rank problem and completes the data with least possible complexity, where the complexity is defined as the number of exponential signals that could describe the data. Provided method allows us to benefit the space-time features of data and correlation between the sensory nodes in the data. To solve the resulting convex optimization problem first order augmented Lagrangian optimization algorithm is implemented where enables us to deal with large scale data. Numerical examples show effective

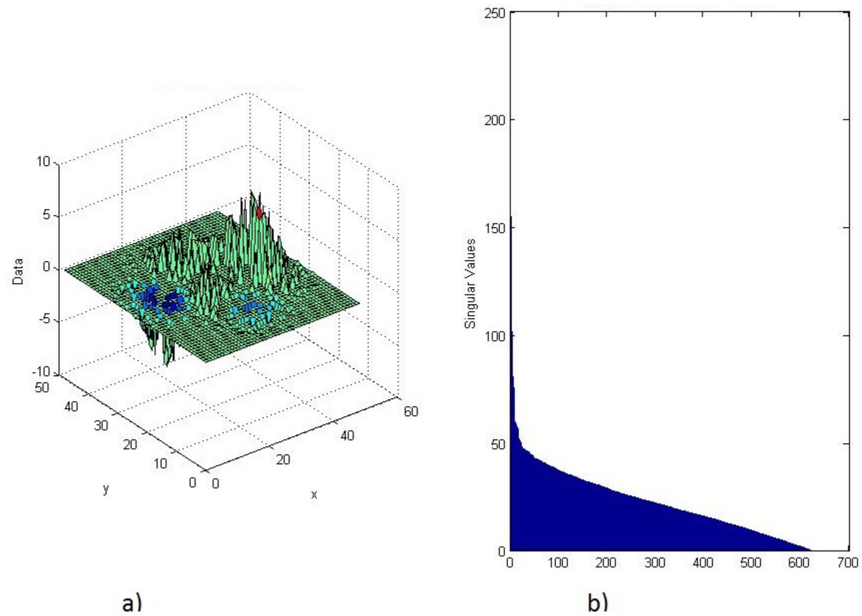

Figure 7.3: Example 2: a) Original and sparse signal, b) Singular values of Hankel matrix constructed by sparse signal

reconstruction even in face of massive missing values in the data.

7.7 Appendix A: Proof of Theorem 31

Consider 1D data array $Y \in \mathbb{R}^{l_1}$ as $Y_{k_1} = \sum_{i=1}^{N} a_i z_{1_i}^{k_1}, \quad k_1 = 0, ..., l_1$. Based on (7.5) and (7.3), Hankel matrix $\mathcal{H}^{1D}$ reads as

$$\mathcal{H}^{1D} = \begin{bmatrix} Y_0 & Y_1 & ... & Y_{l_1-M_1} \\ Y_1 & Y_2 & ... & Y_{l_1-M_1+1} \\ . & . & ... & . \\ . & . & ... & . \\ Y_{l_1-M_1} & Y_{l_1-M_1+1} & ... & Y_{l_1-1} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{N} a_i z_{1_i}^{0} & \sum_{i=1}^{N} a_i z_{1_i}^{1} & ... & \sum_{i=1}^{N} a_i z_{1_i}^{l_1-M_1} \\ \sum_{i=1}^{N} a_i z_{1_i}^{1} & \sum_{i=1}^{N} a_i z_{1_i}^{2} & ... & \sum_{i=1}^{N} a_i z_{1_i}^{l_1-M_1+1} \\ . & . & ... & . \\ . & . & ... & . \\ \sum_{i=1}^{N} a_i z_{1_i}^{l_1-M_1} & \sum_{i=1}^{N} a_i z_{1_i}^{l_1-M_1+1} & ... & \sum_{i=1}^{N} a_i z_{1_i}^{l_1-1} \end{bmatrix}$$

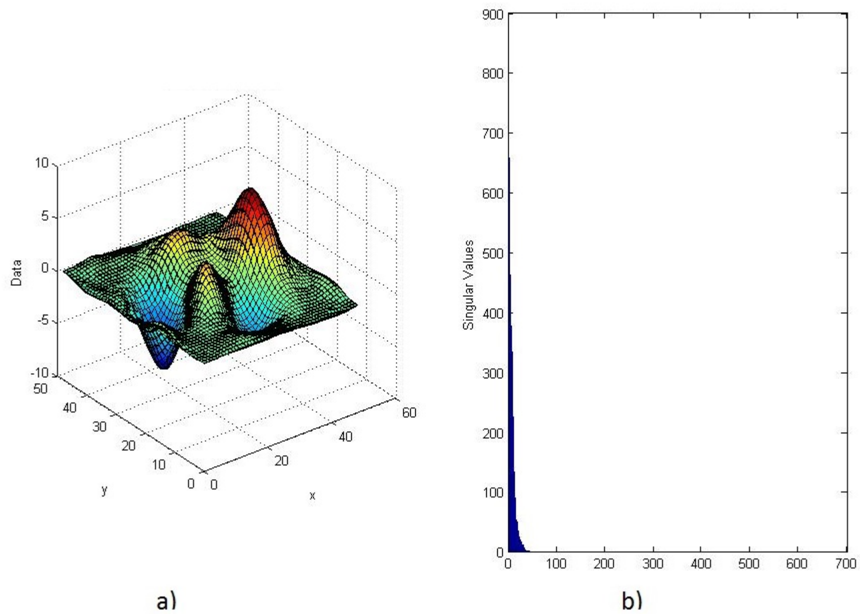

Figure 7.4: Example 2: a) Reconstructed signal, b) Singular values of Hankel matrix of reconstructed sparse signal

$$\begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \\ z_{1_1}^2 & z_{1_2}^2 & .. & z_{1_N}^2 \\ \vdots & \vdots & & \vdots \\ z_{1_1}^{M_1-1} & z_{1_2}^{M_1-1} & .. & z_{1_N}^{M_1-1} \end{bmatrix} \begin{bmatrix} a_1 & 0 & .. & 0 \\ 0 & a_2 & .. & 0 \\ 0 & 0 & .. & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & ... & a_N \end{bmatrix} \begin{bmatrix} 1 & z_{1_1} & z_{1_1}^2 & .. & z_{1_1}^{l_1-M_1} \\ 1 & z_{1_2} & z_{1_2}^2 & .. & z_{1_2}^{l_1-M_1} \\ \vdots & \vdots & \vdots & & \\ 1 & z_{1_N} & z_{1_N}^2 & .. & z_{1_N}^{l_1-M_1} \end{bmatrix} = E_{1_L} A E_{1_R} \tag{7.22}$$

Hence, (7.9) holds true for 1D Hankel matrix. Now, we show that (7.9) holds true for 2D Hankel matrix. Based on (7.5) , for a given 2D array $Y \in \mathbb{R}^{l_2 \times l_1}$ 2D Block Hankel matrix $\mathcal{H}^{2D}$ reads as

$$\mathcal{H}^{2D} = \begin{bmatrix} \mathcal{H}_0^{(1)D} & \mathcal{H}_1^{(1)D} & ... & \mathcal{H}_{l_2-M_2}^{(1)D} \\ \mathcal{H}_1^{(1)D} & \mathcal{H}_2^{(1)D} & ... & \mathcal{H}_{l_2-M_2+1}^{(1)D} \\ . & . & ... & . \\ . & . & ... & . \\ \mathcal{H}_{l_2-M_2}^{(1)D} & \mathcal{H}_{l_2-M_2+1}^{(1)D} & ... & \mathcal{H}_{l_2-1}^{(1)D} \end{bmatrix}$$

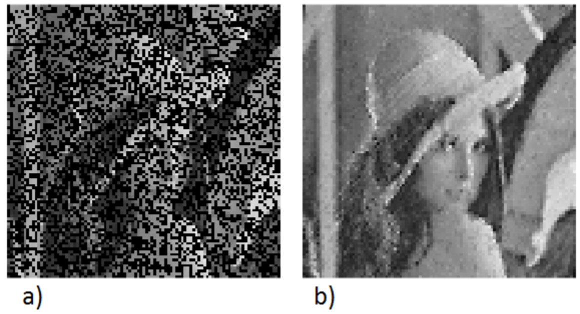

Figure 7.5: a) corrupted image, b) reconstructed image

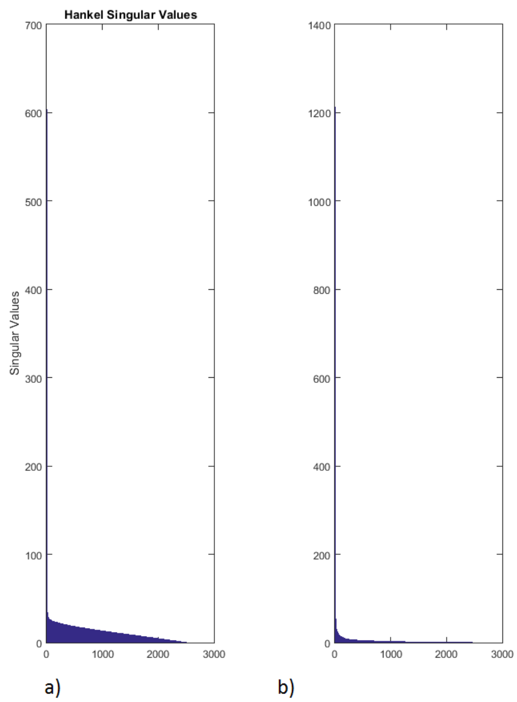

Figure 7.6: Singular values of Hankel matrix of a) corrupted and b)reconstructed image

Based on (7.22), 2D Block Hankel matrix $\mathcal{H}^{2D}$ can be rewritten as

$$\mathcal{H}^{2D} = \begin{bmatrix} E_{1_L} Z_2^{0} A E_{1_R} & E_{1_L} Z_2^{1} A E_{1_R} & \dots & E_{1_L} Z_2^{l_2 - M_2} A E_{1_R} \\ E_{1_L} Z_2^{1} A E_{1_R} & E_{1_L} Z_2^{2} A E_{1_R} & \dots & E_{1_L} Z_2^{l_2 - M_2 + 1} A E_{1_R} \\ . & . & \dots & . \\ . & . & \dots & . \\ E_{1_L} Z_2^{l_2 - M_2} A E_{1_R} & E_{1_L} Z_2^{l_2 - M_2 + 1} A E_{1_R} & \dots & E_{1_L} Z_2^{l_2 - 1} A E_{1_R} \end{bmatrix}$$

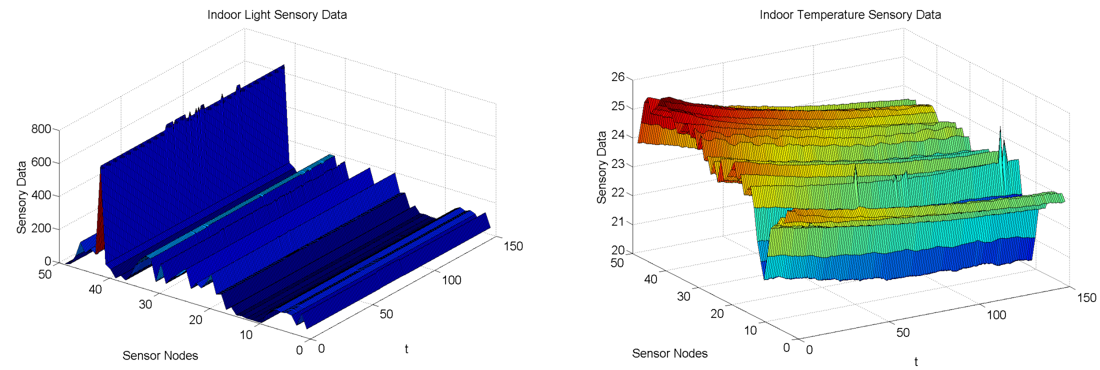

Figure 7.7: Indoor light and temperature sensory data

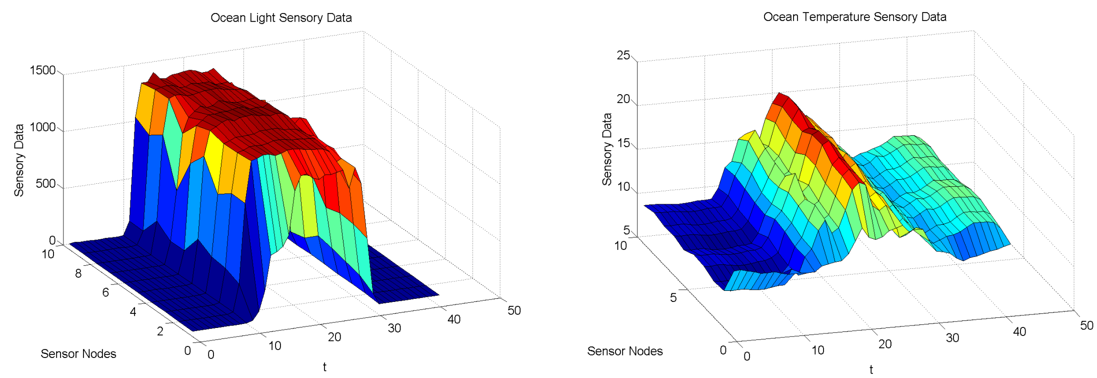

Figure 7.8: Ocean light and temperature sensory data

$$= \begin{bmatrix} E_{1_L} Z_2^0 \\ E_{1_L} Z_2^1 \\ \vdots \\ E_{1_L} Z_2^{M_2 - 1} \end{bmatrix} A \begin{bmatrix} Z_2^0 E_{1_R} & Z_2^1 E_{1_R} & \dots & Z_2^{l_2 - M_2} E_{1_R} \end{bmatrix} = E_{2_L} A E_{2_R}$$

where, Z_2, E_{1_L}, E_{1_R}, and A is as (7.10). Hence, (7.9) holds true for 2D data.

Now, assume that (7.9) holds true for (n-1)D data. We want to show that (7.9) is valid for n-D data. Based on (7.4), For a given n-dimensional data array $Y \in \mathbb{R}^{l_n \times \ldots \times l_2 \times l_1}$, *n-Dimensional Block Hankel* matrix $\mathcal{H}^{nD}$ reads as

$$\mathcal{H}^{nD} = \begin{bmatrix} \mathcal{H}_0^{(n-1)D} & \mathcal{H}_1^{(n-1)D} & ... & \mathcal{H}_{l_n-M_n}^{(n-1)D} \\ \mathcal{H}_1^{(n-1)D} & \mathcal{H}_2^{(n-1)D} & ... & \mathcal{H}_{l_n-M_n+1}^{(n-1)D} \\ . & . & ... & . \\ . & . & ... & . \\ \mathcal{H}_{l_n-M_n}^{(n-1)D} & \mathcal{H}_{l_n-M_n+1}^{(n-1)D} & ... & \mathcal{H}_{l_n-1}^{(n-1)D} \end{bmatrix} \tag{7.23}$$

Based on the assumption that (7.9) hold true for (n-1)-D data, $\mathcal{H}^{nD}$ can be rewritten as

$$\mathcal{H}^{nD} = \begin{bmatrix} E_{n-1_L} Z_n^0 A Z_n^0 E_{n-1_R} & E_{n-1_L} Z_n^0 A Z_n^1 E_{n-1_R} & ... & E_{n-1_L} Z_n^0 A Z_n^{l_n-M_n} E_{n-1_R} \\ E_{n-1_L} Z_n^0 A Z_n^1 E_{n-1_R} & E_{n-1_L} Z_n^0 A Z_n^2 E_{n-1_R} & ... & E_{n-1_L} Z_n^0 A Z_n^{l_n-M_n+1} E_{n-1_R} \\ . & . & ... & . \\ . & . & ... & . \\ E_{n-1_L} Z_n^0 A Z_n^{l_n-M_n} E_{n-1_R} & E_{n-1_L} Z_n^0 A Z_n^{l_n-M_n+1} E_{n-1_R} & ... & E_{n-1_L} Z_n^0 A Z_n^{l_n-1} E_{n-1_R} \end{bmatrix}$$

$$= \begin{bmatrix} E_{n-1_L} Z_n^0 \\ E_{n-1_L} Z_n^1 \\ \vdots \\ E_{n-1_L} Z_n^{M_n-1} \end{bmatrix} A \begin{bmatrix} Z_n^0 E_{n-1_R} & Z_n^1 E_{n-1_R} & ... & Z_n^{l_n-M_n} E_{n-1_R} \end{bmatrix} = E_{n_L} A E_{n_R}$$

where, Z_n, E_{n_L}, E_{n_R}, and A is as (7.10). Hence, (7.9) holds true for n-D data as well.

7.8 Appendix B: Proof of Theorem 32

From the structure of $n-$dimensional Hankel matrix $\mathcal{H}^{nD}$, we know that Rank($\mathcal{H}^{nD}$) $= N$, iff Rank(E_{n_L}) $=$ Rank(E_{n_R}) $= N$, where N is the number of exponential signals. Now, we need to find the conditions on the free parameters M_i (window parameter of ith dimension) under which Rank(E_{n_L}) $=$ Rank(E_{n_R}) $= N$. Since the structure of E_{n_L} and E_{n_R} are similar, only E_{n_L} is considered.

We show that Rank(E_{n_L})$= N$ iff

$$N \leq M_i \leq l_i - N + 1 \tag{7.24}$$

where M_i and l_i are window size parameter and number of data samples in the ith dimension, respectively.

First, it is clear form the structure of $\mathcal{H}^{nD}$ as (7.9) that

$$Rank(\mathcal{H}^{nD}) \leq N \tag{7.25}$$

Now, consider the E_{n_L} matrix, as it is shown in the section 7.3.2 by row permutation one can change the position of frequencies. Hence, the position of frequencies in the higher dimensions (2, 3,...,n) can be changed with the frequencies of first dimension. Also the frequencies of the first dimension appear in the form of E_{1_L} as (7.12) in the matrix of E_{n_L}. Therefore, matrices of the form (Vandermond)

$$V_i = \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{i_1} & z_{i_2} & .. & z_{i_N} \\ z_{i_1}^2 & z_{i_2}^2 & .. & z_{i_N}^2 \\ \vdots & \vdots & & \vdots \\ z_{i_1}^{M_i-1} & z_{i_2}^{M_i-1} & .. & z_{i_N}^{M_i-1} \end{bmatrix} \qquad i = 1, 2, ..., n$$

are sub-matrices of E_{n_L} and its row shuffled matrices as section 7.3.2. Hence, [109]

$$Rank(E_{n_L}) \geq Rank \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} = Rank \begin{bmatrix} \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{1_1} & z_{1_2} & .. & z_{1_N} \\ z_{1_1}^2 & z_{1_2}^2 & .. & z_{M_N}^2 \\ \vdots & \vdots & & \vdots \\ z_{1_1}^{M_1-1} & z_{n_2}^{M_1-1} & .. & z_{n_N}^{M_1-1} \end{bmatrix} \\ \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{2_1} & z_{2_2} & .. & z_{2_N} \\ z_{2_1}^2 & z_{2_2}^2 & .. & z_{2_N}^2 \\ \vdots & \vdots & & \vdots \\ z_{2_1}^{M_2-1} & z_{2_2}^{M_2-1} & .. & z_{2_N}^{M_2-1} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} 1 & 1 & .. & 1 \\ z_{n_1} & z_{n_2} & .. & z_{n_N} \\ z_{n_1}^2 & z_{n_2}^2 & .. & z_{n_N}^2 \\ \vdots & \vdots & & \vdots \\ z_{n_1}^{M_n-1} & z_{n_2}^{M_n-1} & .. & z_{i_N}^{M_n-1} \end{bmatrix} \end{bmatrix} \tag{7.26}$$

Since, $\{(z_{1_i}, z_{2_i}, ..., z_{n_i}), i = 1, 2, ...N\}$ are distinct, the N columns of right hand side matrix in (7.26) are linearly independent provided $M_i \geq N, i = 1, ..., n$ (so that E_{i_L} each have no less than N row). Hence, the sufficient condition for E_{n_L} to be of the full rank N is

$$M_i \geq N, \ i = 1, ..., n. \tag{7.27}$$

The necessary condition for E_{n_L} to be of the full rank N is that the number of rows of E_{n_L} is no less than N. Hence:

$$\prod_{i=1}^{n} M_i \geq N. \tag{7.28}$$

Due to the similarity between E_{n_L} and E_{n_R}, it can be similarly shown that the necessary and sufficient conditions of $Rank(E_{n_R}) = N$ are

$$L_i - M_i + 1 \geq N, \ i = 1, ..., n. \tag{7.29}$$

$$\prod_{i=1}^{n}(L_i - M_i + 1) \geq N \tag{7.30}$$

Hence, combining the (7.25), (7.27) and (7.29), the sufficient condition (7.24) is proven.

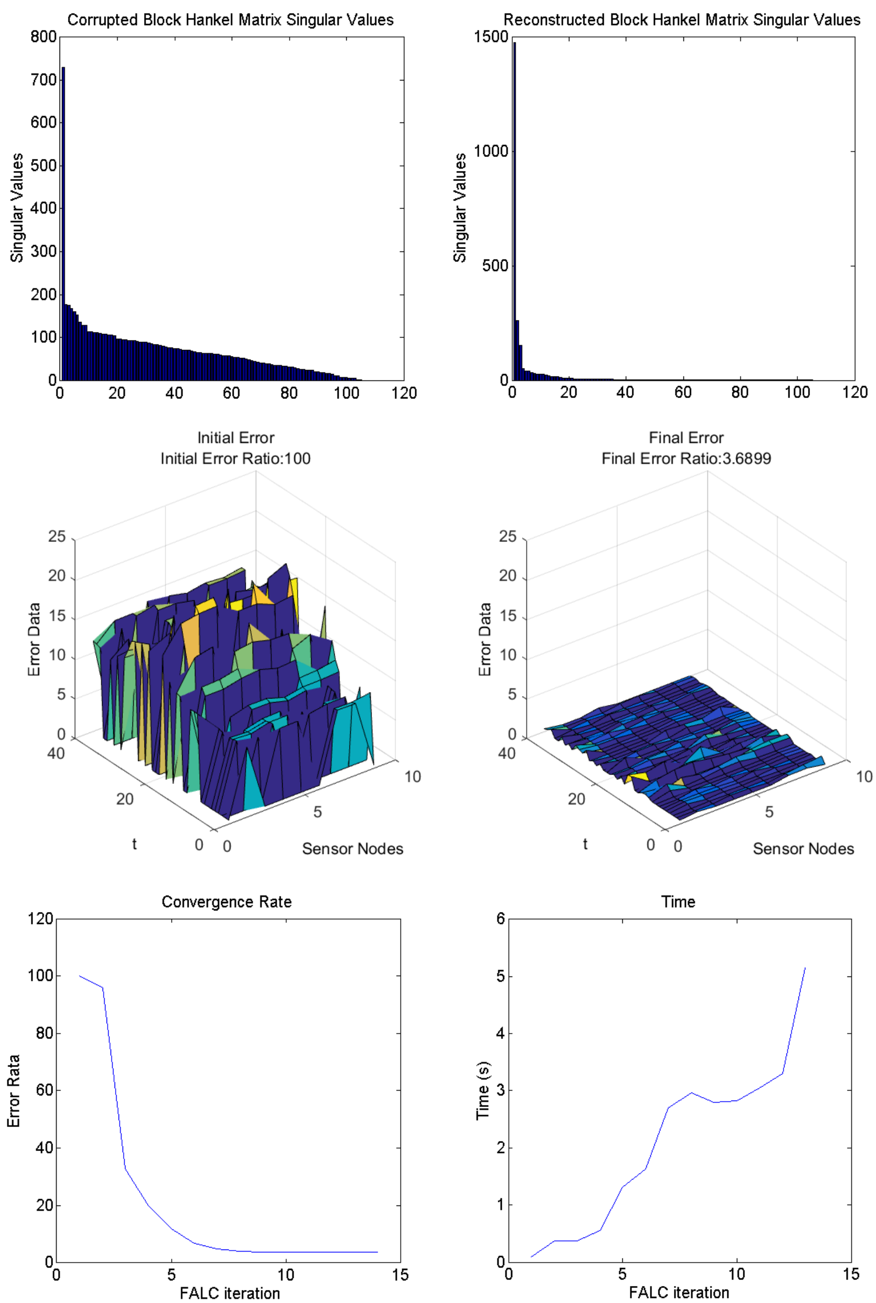

Figure 7.9: 50% sparse Ocean temperature data reconstruction

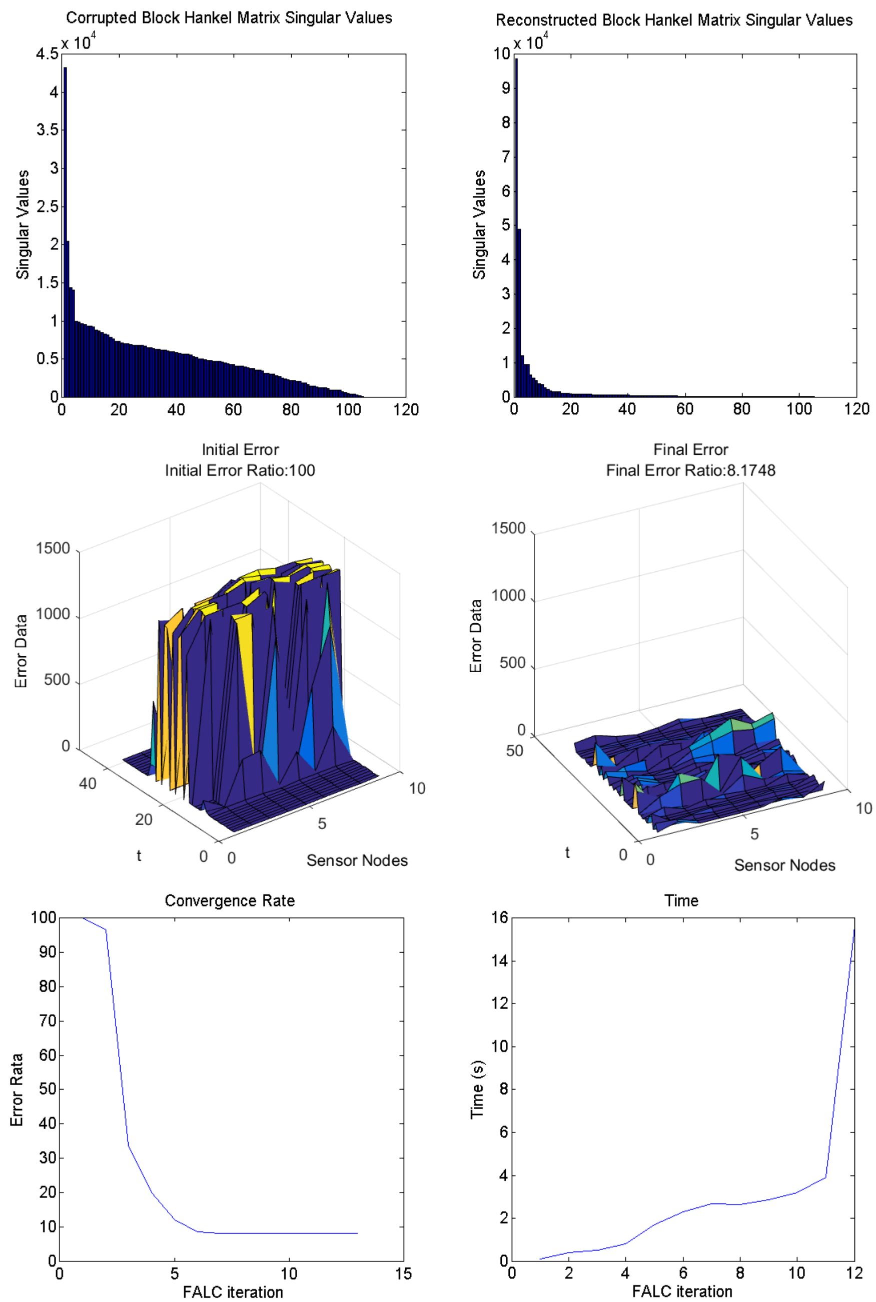

Figure 7.10: 50% sparse Ocean light data reconstruction

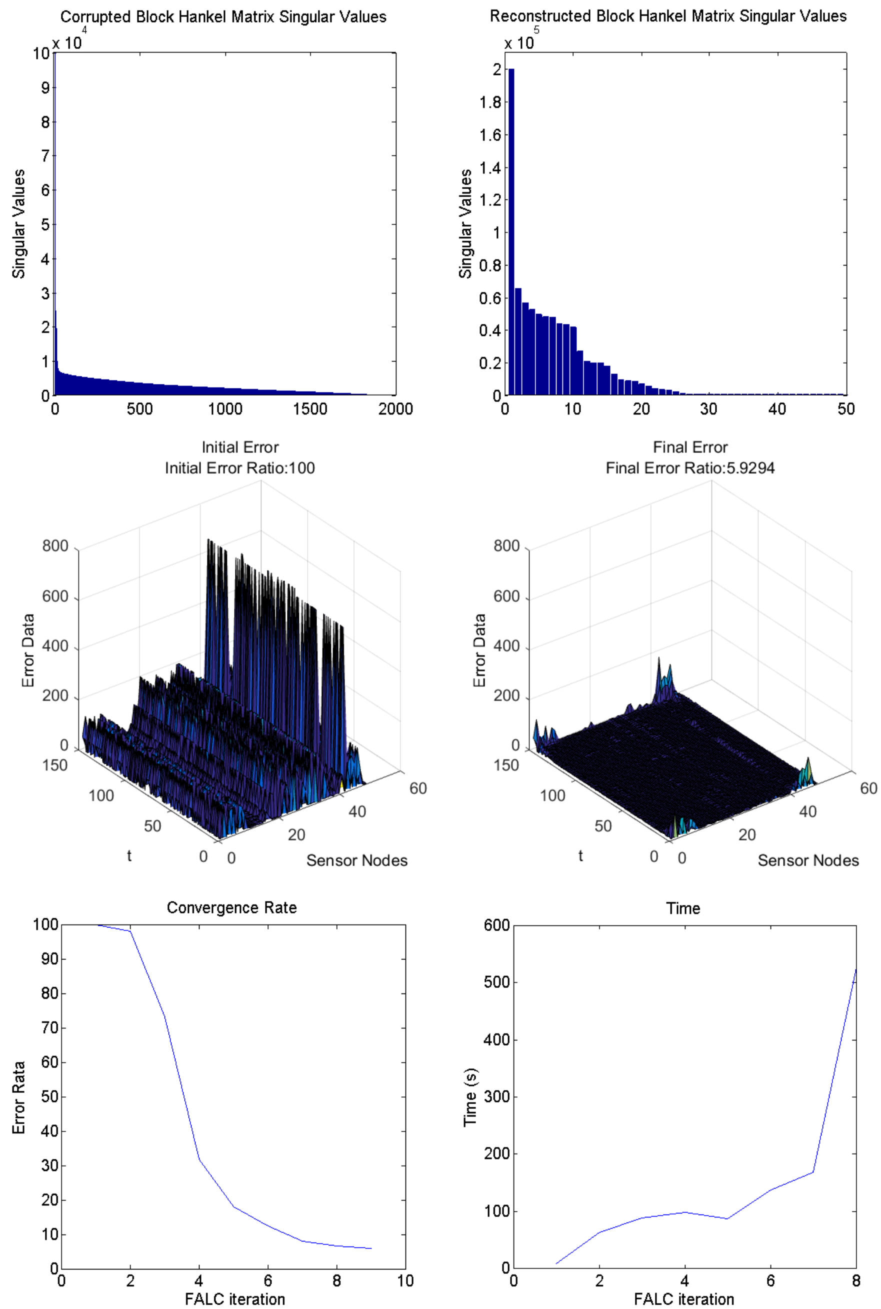

Figure 7.11: 50% sparse indoor light data reconstruction

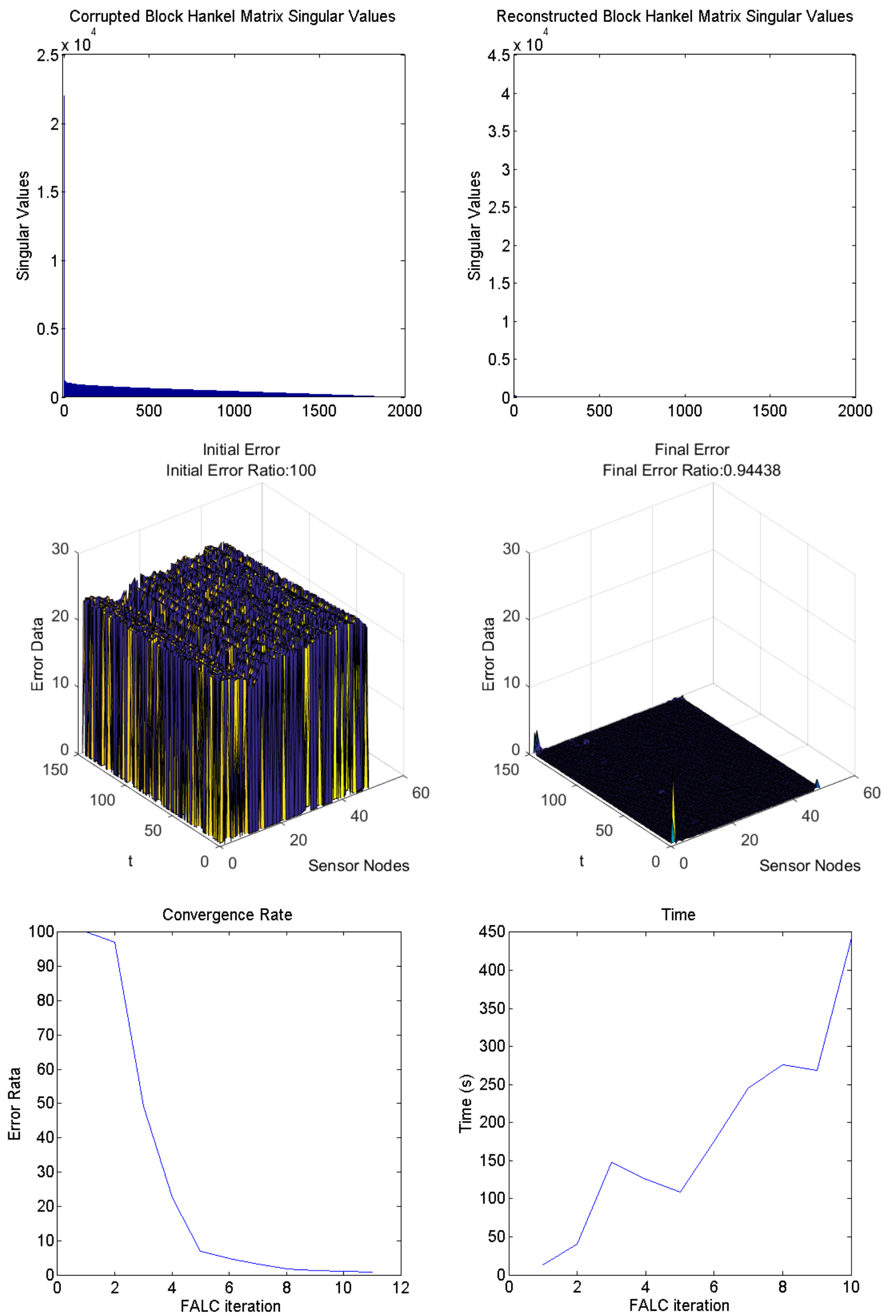

Figure 7.12: 50% sparse indoor Temperature data reconstruction

Chapter 8

Conclusion and Discussion

In this thesis, "chance optimization" problems are introduced, where one aims at maximizing the probability of a set defined by polynomial inequalities. These problems are, in general, nonconvex and computationally hard. A sequence of semidefinite relaxations is provided whose sequence of optimal values is shown to converge to the optimal value of the original problem. We also presented a novel approach based on chance optimization results to the chance constrained controller design when the objective is to reach a given target set with high probability. Moreover, we provided a novel method to the problem of uncertainty propagation and reconstruction of support of measures from their moments. In this thesis, constrained volume optimization problems are introduced, where one aims at maximizing the volume of a set defined by polynomial inequalities such that it is contained in other semialgebraic set. We showed that many nonconvex problems in system and control can be reformulated as constrained volume optimization problems. To be able to obtain a equivalent convex problem, the results from theory of measure and moments as well as duality theory are used. In addition, we presented a novel approach for solving the problem of reconstructing spars and noisy data. The proposed method reformulates the problem as a minimum rank problem and completes the data with least possible complexity, where the complexity is defined as the order of linear differential equations describing the signal data or equivalently the number of exponential signals that could describe the data.

To solve the semidefinite programs of increasing size obtained by relaxing the original chance optimization problem, a first-order augmented Lagrangian algorithm is implemented which enables us to solve much larger size semidefinite programs that interior point methods can deal with. Numerical examples are provided that show that one can obtains reasonable approximations to the optimal solution and the corresponding optimal probability even for lower order relaxations. In terms of future work, by exploiting algebraic structures i.e., sparsity, symmetry, we can reduce the complexity of obtained semidefinite programs. Also, to improve the performance, we can develop the results for specific classes of measures, polynomials, and systems.

Bibliography

[1] Calafiore, G. and M. Campi (2005) "Uncertain convex programs: randomized solutions and confidence levels," *Mathematical Programming*, **102**(1), pp. 25–46.

[2] ——— (2006) "The Scenario Approach to Robust Control Design," *IEEE Transactions on Automatic Control*, **51**(5), pp. 742–753.

[3] Nemirovski, A. and H. M. Stewart (2012) "On Safe Tractable Approximations of Chance Constraints," *Eur. J. Oper. Res.*, (219), pp. 707–718.

[4] Nemirovski, A. and A. Shapiro (2005) "Scenario Approximations of Chance Constraints," *Probabilistic and Randomized Methods for Design under Uncertainty, Springer*, pp. 3–47.

[5] Tempo, R., G. Calafiore, and F. Dabbene (2013) *Randomized Algorithms for Analysis and Control of Uncertain Systems*, Communications and Control Engineering, Springer London, London.

[6] Ben-Tal, A. and A. Nemirovski (2000) "Robust solutions of Linear Programming problems contaminated with uncertain data," *Mathematical Programming*, **88**(3), pp. 411–424.

[7] Ben, A., T. Laurent, E. Ghaoui, and A. Nemirovski (1998) "Robust Semidefinite Programming *," in *Handbook on Semidefinite Programming,New York: Kluwer.*

[8] Ben-Tal, A., A. Nemirovski, and C. Roos (2006) "Robust solutions of uncertain quadratic and conic-quadratic problems," *SIAM J. Optim.,*, **13**(2), pp. 535–560.

[9] Ben-Tal, A., A. Goryashko, E. Guslitzer, and A. Nemirovski (2004) "Adjustable robust solutions of uncertain linear programs," *Math. Program., Ser. A*, **99**, pp. 351–376.

[10] A. Ben-tal, A. N. (1999) "Robust solutions of uncertain linear programs," *Operations Research Letters*, **25**(1), pp. 1–13.

[11] Bertsimas, D. and M. Sim (2004) "The Price of Robustness," *Operations Research*, **52**(1), pp. 35–53.

[12] Miller, B. L. and H. M. Wagner (1965) "Chance Constrained Programming with Joint Constraints," *Operations Research*, **13**(6).

[13] Nemirovski, A. (2003) "On tractable approximations of randomly perturbed convex constraints," in *42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475)*, vol. 3, IEEE, pp. 2419–2422.

[14] Nemirovski, A. and A. Shapiro (2007) "Convex Approximations of Chance Constrained Programs," *SIAM Journal on Optimization*, **17**(4), pp. 969–996.

[15] Pintr, J. (1989) "Deterministic approximations of probability inequalities," *ZOR Zeitschrift fr Operations Research Methods and Models of Operations Research*, **33**(4), pp. 219–239.

[16] Dabbene, F., C. Feng, and C. M. Lagoa (2009) "Robust and chance-constrained optimization under polynomial uncertainty," in *2009 American Control Conference*, IEEE, pp. 379–384.

[17] Feng, C. and C. Lagoa (2015) "Distributional Robustness Analysis for Nonlinear Uncertainty Structures," *IEEE Transactions on Automatic Control*, **61**(7), pp. 1–13.

[18] Feng, C., F. Dabbene, and C. M. Lagoa (2011) "A Kinship Function Approach to Robust and Probabilistic Optimization Under Polynomial Uncertainty," *IEEE Transactions on Automatic Control*, **56**(7), pp. 1509–1523.

[19] Lagoa, C., F. Dabbene, and R. Tempo (2008) "Hard Bounds on the Probability of Performance With Application to Circuit Analysis," *IEEE Transactions on Circuits and Systems I: Regular Papers*, **55**(10), pp. 3178–3187.

[20] Calafiore, G. C. and L. E. Ghaoui (2006) "On Distributionally Robust Chance-Constrained Linear Programs," *Journal of Optimization Theory and Applications*, **130**(1), pp. 1–22.

[21] Chen, W., M. Sim, J. Sun, and C.-P. Teo (2010) "From CVaR to Uncertainty Set: Implications in Joint Chance-Constrained Optimization," *Operations Research*, **58**(2), pp. 470–485.

[22] Cheung, S.-S., A. M.-C. So, and K. Wang (2012) "Linear Matrix Inequalities with Stochastically Dependent Perturbations and Applications to Chance-Constrained Semidefinite Optimization," *SIAM J. Optim.*,, **22**(4), pp. 1394–1430.

[23] Xu, H., C. Caramanis, and S. Mannor (2012) "Optimization Under Probabilistic Envelope Constraints," *Operations Research*, **60**(3), pp. 682–699.

[24] Zymler, S., D. Kuhn, and B. Rustem (2013) "Distributionally robust joint chance constraints with second-order moment information," *Mathematical Programming*, **137**(1-2), pp. 167–198.

[25] Ben-Tal, A., D. Bertsimas, and D. B. Brown (2010) "A Soft Robust Model for Optimization Under Ambiguity," *OPERATIONS RESEARCH*, **58**(2), pp. 1220–1234.

[26] Ben-Tal, A., S. Boyd, and A. Nemirovski (2006) "Extending Scope of Robust Optimization: Comprehensive Robust Counterparts of Uncertain Problems," *Mathematical Programming*, **107**(1-2), pp. 63–89.

[27] Lagoa, C. M., X. Li, and M. Sznaier (2005) "Probabilistically Constrained Lnear Programs and Risk-Adjusted Controller Design," **15**(3), pp. 938–951.

[28] Prékopa, A. (1995) *Stochastic Programming*, Springer Netherlands, Dordrecht.

[29] Calafiore, G. C. and F. Dabbene (2007) "Probabilistic Robust Control," in *2007 American Control Conference*, IEEE, pp. 147–158.

[30] Wang, Q. and R. F. Stengel "Probabilistic Control of Nonlinear Uncertain Systems," in *Probabilistic and Randomized Methods for Design under Uncertainty, Springer*, pp. 381–414.

[31] Calafiore, G., F. Dabbene, and R. Tempo (2000) "Randomized algorithms for probabilistic robustness with real and complex structured uncertainty," *IEEE Transactions on Automatic Control*, **45**(12), pp. 2218–2235.

[32] Vitus, M. P. and C. J. Tomlin (2011) "On feedback design and risk allocation in chance constrained control," in *IEEE Conference on Decision and Control and European Control Conference*, IEEE, pp. 734–739.

[33] Lyons, D., J. Calliess, and U. D. Hanebeck (2012) "Chance constrained model predictive control for multi-agent systems with coupling constraints," in *2012 American Control Conference (ACC)*, IEEE, pp. 1223–1230.

[34] Ono, M., L. Blackmore, and B. C. Williams (2010) "Chance Constrained Finite Horizon Optimal Control with Nonconvex Constraints," in *Proceedings of the 2010 American Control Conference*, pp. 1145–1152.

[35] Oldewurtel, F., C. N. Jones, and M. Morari (2008) "A tractable approximation of chance constrained stochastic MPC based on affine disturbance feedback," in *2008 47th IEEE Conference on Decision and Control*, IEEE, pp. 4731–4736.

[36] Blackmore, L. and M. Ono (2009) "Convex Chance Constrained Predictive Control without Sampling," in *AIAA Guidance, Navigation, and Control Conference, Guidance, Navigation, and Control Conferences*.

[37] Blackmore, L., M. Ono, A. Bektassov, and B. C. Williams (2010) "A Probabilistic Particle-Control Approximation of Chance-Constrained Stochastic Predictive Control," *IEEE Transactions on Robotics*, **26**(3), pp. 502–517.

[38] Streif, S., M. Kögel, T. Bäthge, and R. Findeisen (2014) "Robust Nonlinear Model Predictive Control with Constraint Satisfaction: A Relaxation-based Approach," in *FAC World Congress*, vol. 47, pp. 11073–11079.

[39] Cannon, M., J. Buerger, B. Kouvaritakis, and S. Rakovic (2011) "Robust Tubes in Nonlinear Model Predictive Control," *IEEE Transactions on Automatic Control*, **56**(8), pp. 1942–1947.

[40] Jasour, A. and M. Farrokhi (2014) "Adaptive neuro-predictive control for redundant robot manipulators in presence of static and dynamic obstacles: A Lyapunov-based approach," *International Journal of Adaptive Control and Signal Processing*, **28**(3-5), pp. 386–411.

[41] ——— (2010) "Fuzzy improved adaptive neuro-NMPC for online path tracking and obstacle avoidance of redundant robotic manipulators," *International Journal of Automation and Control*, **4**(2).

[42] ——— (2009) "Adaptive Neuro-NMPC Control of Redundant Manipulators for Path Tracking and Obstacle Avoidance," in *European Control Conference (ECC)*, IEEE, Budapest, Hungary, pp. 2181–2186.

[43] ——— (2009) "Path tracking and obstacle avoidance for redundant robotic arms using fuzzy NMPC," in *American Control Conference (ACC)*, IEEE, St. Louis, Missouri, pp. 1353–1358.

[44] ——— (2010) "Control of redundant manipulators in non-stationary environments using neural networks and model predictive control," *IFAC Workshop on Intelligent Control Systems*.

[45] Blackmore, L., H. Hui Li, and B. Williams (2006) "A probabilistic approach to optimal robust path planning with obstacles," in *2006 American Control Conference*, IEEE.

[46] Gravin, N., J. Lasserre, D. V. Pasechnik, and S. Robins (2012) "The Inverse Moment Problem for Convex Polytopes," *Discrete & Computational Geometry*, **48**(3), pp. 596–621, `arXiv:1106.5723v2`.

[47] Gustafsson, B., C. He, P. Milanfar, and M. Putinar (2000) "Reconstructing planar domains from their moments," *Inverse Problems*, **16**, pp. 1053–1070.

[48] Putinar, M. and J.-B. Lasserre (2015) "Algebraicexponential Data Recovery from Moments," *Discrete & Computational Geometry*, **54**(4), pp. 993–1012.

[49] Henrion, D. and M. Korda (2014) "Convex computation of the region of attraction of polynomial control systems," *IEEE Transactions on Automatic Control*, `arXiv:1208.1751v2`.

[50] Korda, M., D. Henrion, and C. N. Jones (2013) "Convex computation of the maximum controlled invariant set for discrete-time polynomial control systems?" in *Proceedings of the IEEE Conference on Decision and Control*, `1303.6469`.

[51] Korda, M., D. Henrion, and N. Jones, Colin (2013) *Inner Approximations of the Region of Attraction for Polynomial Dynamical Systems*, vol. 9.

[52] Majumdar, A., A. A. Ahmadi, and R. Tedrake (2013) "Control Design along Trajectories with Sums of Squares Programming," in *IEEE International Conference on Robotics and Automation (ICRA)*.

[53] ——— (2014) "Control and verification of high-dimensional systems with DSOS and SDSOS programming," in *53rd IEEE Conference on Decision and Control*, IEEE, pp. 394–401.

[54] Khashayar Kotobi, C. T. S. B., P.B. Mainwaring (2015) "Data throughput enhancement using data mining informed cognitive radio," *Electronics*.

[55] S.G. Biln, C. T., K. Kotobi (2014) "Data mining informed cognitive radio networks," *New England Workshop on Software Defined Radio,*.

[56] Cover, T. and P. Hart (1967) "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, **13**(1), pp. 21–27.

[57] Kong, L., D. Jiang, and M.-Y. Wu (2010) "Optimizing the Spatio-temporal Distribution of Cyber-Physical Systems for Environment Abstraction," in *2010 IEEE 30th International Conference on Distributed Computing Systems*, IEEE, pp. 179–188.

[58] Zhu, H., Y. Zhu, M. Li, and L. M. Ni (2009) "SEER: Metropolitan-Scale Traffic Perception Based on Lossy Sensory Data," in *IEEE INFOCOM 2009 - The 28th Conference on Computer Communications*, IEEE, pp. 217–225.

[59] Donoho, D. (2006) "Compressed sensing," *IEEE Transactions on Information Theory*, **52**(4), pp. 1289–1306.

[60] Kutyniok, G. (2012) "Theory and Applications of Compressed Sensing," *GAMM-Mitteilungen*, **36**(1), pp. 79–101, `1203.3815`.

[61] H.Huang and A. Makur (2011) "Optimized Measurement Matrix for Compressive Sensing," in *The 9th International Conference on Sampling Theory and Applications*.

[62] Kong, L., M. Xia, X.-Y. Liu, M.-Y. Wu, and X. Liu (2013) "Data loss and reconstruction in sensor networks," in *2013 Proceedings IEEE INFOCOM*, IEEE, pp. 1654–1662.

[63] Sznaier, M. and O. Camps (2005) "A Hankel Operator Approach to Texture Modelling and Inpainting," in *International Workshop on Texture Analysis and Synthesis*, pp. 125–130.

[64] Ding, T., M. Sznaier, and O. Camps (2006) "Robust Identification of 2-D Periodic Systems with Applications to Texture Synthesis and Classification," in *Proceedings of the 45th IEEE Conference on Decision and Control*, IEEE, pp. 3678–3683.

[65] Jasour, A. and C. Lagoa (2012) "Semidefinite relaxations of chance constrained algebraic problems," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, IEEE, Maui, Hawaii, pp. 2527–2532.

[66] Jasour, A., N. S. Aybat, and C. M. Lagoa (2015) "Semidefinite Programming For Chance Constrained Optimization Over Semialgebraic Sets," *SIAM Journal on Optimization*, **25**(3), pp. 1411–1440.

[67] Henrion, D., J. B. Lasserre, and C. Savorgnan (2009) "Approximate Volume and Integration for Basic Semialgebraic Sets," *SIAM Review*, **51**(4), pp. 722–743.

[68] Lasserre, J. B. (2011) "Global Optimization with Polynomials and the Problem of Moments," *SIAM Journal on Optimization*, **11**, pp. 796–817.

[69] ——— (2010) *Moments, positive polynomials and their applications*, Imperial College Press.

[70] Jasour, A. and C. Lagoa (2013) "Convex relaxations of a probabilistically robust control design problem," in *52nd IEEE Conference on Decision and Control (CDC)*, IEEE, Florence, Italy, pp. 1892–1897.

[71] ——— (2016) "Convex Chance Constrained Model Predictive Control," *55th IEEE Conference on Decision and Control (CDC)*.

[72] ——— (2014) "Reconstruction of support of a measure from its moments," *53rd IEEE Conference on Decision and Control (CDC)*.

[73] ——— (2016) "Convex constrained semialgebraic volume optimization: Application in systems and control," *arXiv:1701.08910*.

[74] Laurent, M. (2008) "Sums of Squares, Moment Matrices and Optimization Over Polynomials," chap. Emerging A, Springer New York, pp. 157–270.

[75] Parrilo, P. A. (2003) "Semidefinite programming relaxations for semialgebraic problems," *Mathematical Programming*, **96**(2), pp. 293–320.

[76] Putinar, M. (1993) "Positive polynomials on compact semi-algebraic sets," *Indiana University Mathematics Journal*, **42**(3), pp. 969–984.

[77] Laraki, R. and J. B. Lasserre (2012) "Semidefinite programming for min-max problems and games," *Mathematical Programming*, **131**(1-2), pp. 305–332.

[78] Lasserre, J. B. (2007) "A semidefinite programming approach to the generalized problem of moments," *Mathematical Programming*, **112**(1), pp. 65–92.

[79] Anderson, E. J. and P. Nash (1987) *Linear programming in infinite-dimensional spaces: theory and applications*, John Wiley & Sons.

[80] Ash, R. B. (1972) *Real Analysis and Probability*.

[81] Royden, H. L. (2010) *Real analysis*, 4th ed., Prentice Hall.

[82] A. Barvinok (2002) *A course in convexity*, American Mathematical Society, Providence.

[83] M. Trnovska (2005) "Strong Duality Conditions in Semidefinite Programming," *Journal of Electrical Engineering*, **56**(12).

[84] Todd, M. (2001) "Semidefinite Optimization," *Acta Numerica*, **10**, pp. 515–560.

[85] Sturm, J. F. (1999) "Using SeDuMi 1.02, A MATLAB toolbox for optimization over symmetric cones, Optim," *Methods Softw.*, **11**, pp. 625–653.

[86] Battles, Z., L. N. Trefethen, and S. J. C. Sci "An Extension of Matlab to Continuous Functions and Operators," *SIAM J. Sci. Comput.*, **25**(5), pp. 1743–1770.

[87] Fazel, M., T. K. Pong, D. Sun, and P. Tseng (2013) "Hankel Matrix Rank Minimization with Applications to System Identification and Realization," *SIAM Journal on Matrix Analysis and Applications*, **34**(3), pp. 946–977.

[88] Fazel, M., H. Hindi, and S. Boyd (2003) "Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices," in *Proceedings of the 2003 American Control Conference, 2003.*, vol. 3, IEEE, pp. 2156–2162.

[89] Aybat, N. S. and G. Iyengar (2013) "An Augmented Lagrangian Method for Conic Convex Programming," *http://arxiv.org/abs/1302.6322v1*.

[90] Beck, A. and M. Teboulle (2009) "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *SIAM Journal on Imaging Sciences*, **2**(1), pp. 183–202.

[91] Nesterov, Y. (2005) "Smooth minimization of non-smooth functions," *Mathematical Programming*, **103**(1), pp. 127–152.

[92] ——— (2004) *Introductory Lectures on Convex Optimization*, vol. 87 of *Applied Optimization*, Springer US, Boston, MA.

[93] P. Tseng (2008) *On Accelerated Proximal Gradient Methods for Convex-Concave Optimization,, Tech. rep.*

[94] Aybat, N. S. and G. Iyengar (2014) "A unified approach for minimizing composite norms," *Math. Program., Ser. A*, **144**, pp. 181–226.

[95] Josz, C. and D. Henrion (2016) "Strong duality in Lasserre's hierarchy for polynomial optimization," **10**(1), pp. 3–10.

[96] Efberg, J. (2004) "YALMIP : A toolbox for modeling and optimization in MATLAB," *IEEE International Symposium on Computer Aided Control Systems Design Taipei*.

[97] Henrion, D., J.-B. Lasserre, and J. Löfberg (2009) "GloptiPoly 3: moments, optimization and semidefinite programming," *Optimization Methods and Software*, **24**(4-5), pp. 761–779.

[98] Recht, B., M. Fazel, and P. A. Parrilo (2010) "Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization," *SIAM Rev., ,*, **52**(3), pp. 471–501.

[99] Wang, Y. and S. Boyd (2008) "Fast Model Predictive Control Using Online Optimization," in *17th IFAC World Congress*, pp. 6974–6979.

[100] Lasserre, J. B. (2011) "Min-max and robust polynomial optimization," *Journal of Global Optimization*, **51**(1), pp. 1–10.

[101] Lasserre, J. B. and T. P. Thanh (2012) "A joint + marginal heuristic for 0/1 programs," *Journal of Global Optimization*, **54**(4), pp. 729–744.

[102] Dabbene, F., D. Henrion, and C. Lagoa (2015) "Simple Approximations of Semialgebraic Sets and their Applications to Control," *arXiv:1509.04200*, `arXiv:1509.04200v1`.

[103] Khalil, H. K. (2002) "Nonlinear systems, 3rd," *New Jewsey, Prentice Hall*, **9**.

[104] Ali Ahmadi, A., M. Krstic, and P. A. Parrilo (2011) "A Globally Asymptotically Stable Polynomial Vector Field with no Polynomial Lyapunov Function," in *IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*.

[105] Papachristodoulou, A. and S. Prajna (2002) "On the construction of Lyapunov functions using the sum of squares decomposition," in *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, vol. 3, IEEE, pp. 3482–3487.

[106] Majumdar, A., A. A. Ahmadi, and R. Tedrake (2013) "Control Design along Trajectories with Sums of Squares Programming," in *IEEE International Conference on Robotics and Automation (ICRA)*.

[107] Ahmadi, A. A. and A. Majumdar (2014) "DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization," in *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, IEEE, pp. 1–5.

[108] ——— (2016) "Some applications of polynomial optimization in operations research and real-time decision making," *Optimization Letters*, **10**(4), pp. 709–729.

[109] Hua, Y. (1992) "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Transactions on Signal Processing*, **40**(9), pp. 2267–2280.

[110] Chen, Y. and Y. Chi (2014) "Robust Spectral Compressed Sensing via Structured Matrix Completion," *IEEE Transactions on Information Theory*, **60**(10), pp. 6576–6601.

Vita

Ashkan M. Jasour

http://jasour.mit.edu/ashkan, jasour@mit.edu

Education:

- Postdoctoral Associate, Computer Science and Artificial Intelligence Laboratory (CSAIL), 2016 –2018,
  Massachusetts Institute of Technology (MIT),
  Model-based Embedded and Robotic Systems Group (MERS).

- PhD, Electrical Engineering and Computer Science, 2010 –2016,
  The Pennsylvania State University (PSU),
  Robust Machine Intelligence and Control Lab,
  Adviser: Professor Constantino Lagoa,
  Co-Adviser: Professor Necdet Serhat Aybat,
  Thesis Title: Convex Approximation of Chance Constrained Problems: Application in Systems and Control. Defense Date: Aug 26 2016.

- PhD Minor, Mathematics,
  The Pennsylvania State University.

- MS, Electrical Engineering,
  Iran University of Science and Technology (IUST),
  Intelligent Systems Research Lab,
  Advisor: Professor Mohammad Farrokhi,
  Thesis Title: Path Tracking and Obstacle Avoidance for Redundant Robotic Manipulators in Stationary and Non-Stationary Environments Using Adaptive Nonlinear Model Predictive Control.

- BS, Electrical and Computer Engineering,
  Tabriz University, Iran,
  Advisor: Professor Mohammad Taghi Vakili,
  Thesis Title: Fuzzy Controller Design Using Evolutionary Algorithms.

Experience:

- ANSYS, Inc., System Modeling and Testing Internship, 2014.

- Research and Teaching Assistant, The Pennsylvania State University.

- Research and Teaching Assistant, Iran University of Science and Technology.

- Soccer Robots Team, Tabriz University.