Imperial College London
Department of Computing

Decision Under Uncertainty: Problems in Control Theory, Robust Optimization and Game Theory

Michael J. Hadjiyiannis

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Computing of Imperial College and
the Diploma of Imperial College, July 2014

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Abstract

Decision making under uncertainty is a widely-studied area spanning a number of fields such as computational optimization, control theory, utility theory and game theory. A typical problem of decision making under uncertainty requires the design of an optimal decision rule, control policy, or behavioural function, that takes into account all available information regarding the uncertain parameters and returns the decision that is most suitable for the given objective. A popular requirement is to determine robust decisions that maintain certain desired characteristics despite the presence of uncertainty.

In this thesis, we study three distinct problems that involve the design of robust decisions under different types of uncertainty. We investigate a dynamic multi-stage control problem with stochastic exogenous uncertainty, a dynamic two-stage robust optimization problem with epistemic exogenous uncertainty, and finally a game theoretic problem with both stochastic endogenous and epistemic exogenous uncertainty. Specifically,

a) we develop an efficient algorithm that bounds the performance loss of affine policies operating in discrete-time, finite-horizon, stochastic systems with expected quadratic costs and mixed linear state and input constraints. Finding the optimal control policy for such problems is generally computationally intractable, but suboptimal policies can be computed by restricting the class of admissible policies to be affine on the observation. Our algorithm provides an estimate of the loss of optimality due to the use of such affine policies, and it is based on a novel dualization technique, where the dual variables are restricted to have an affine structure;

b) we develop an efficient algorithm to bound the suboptimality of linear decision rules in two-stage dynamic linear robust optimization problems, where they have been shown to suffer a worst-case performance loss of the order $\Omega(\sqrt{m})$ for problems with m linear constraints. Our algorithm is based on a scenario selection technique, where the original problem is evaluated only over a finite subset of the possible uncertain parameters. This set is constructed from the Lagrange multipliers associated with the computation of the linear adaptive decision rules. The resulting instance-wise bounds outperform known bounds, including the aforementioned worst-case bound, in the vast majority of problem instances;

c) we develop an algorithm that enumerates all behavioural functions that are at equilibrium in a game where players face epistemic uncertainty regarding their opponent's utility functions. Traditionally,

these games are solved as complete-information games where players are assumed to be risk-neutral, with a utility function that is positively affine in the monetary payoffs. We demonstrate that this assumption imposes severe limitations on the problem structure, and we propose that these games should be formulated as incomplete private information games where each player may have any increasing or increasing concave utility function. If the players are ambiguity-averse, then under these assumptions, they play either a pure strategy, a max-min strategy, or a convex combination of the two. By utilizing this result, we develop an algorithm that can enumerate all equilibria of the game.

Acknowledgements

First and foremost, I would like to extend my gratitudes to my supervisor Dr. Daniel Kuhn. As my undergraduate project supervisor, his brilliance and passion were a great influence on me choosing the path of the Ph.D., whilst his support,guidance and patience, along with his always-open door-policy, have been instrumental in successfully completing it. I am also grateful to have had as my second supervisor one of my esteemed friends, Dr. Wolfram Wiesemann. Even though our relationship pre-dates his professional academic career, I consider my self privileged to be his first Ph.D. student, now that he is a lecturer. During the first two years of my Ph.D., I was fortunate to work with an additional supervisor, Dr Paul Goulart. Dr. Goulart has been an invaluable source of advice and support. I would also like to express my sincere gratitude to Professor BerÃ§ Rustem for his constant support and encouragement as well as his willingness to offer a fresh perspective throughout the course of this work.

I am grateful to my friends and colleagues at Imperial College, Ryan Duy Luong, Dimitra Bampou, Phebe Vayanos, Eva Kalyvianaki, Panos Parpas, Michalis Kapsos, Paula Rocha, Polyxeni Kleniati, and many others who made my time as a graduate student extremely enjoyable. A special thanks to Angelos Gheorgiou, Iakovos Kakouris, Loizos Markides and Vladimir Roitch for all the lunch-time discussions and shared chickens. I also sincerely thank my flatmates and friends Stelios and Philip, my ski and write-up partner Juan Jerez and all my friends in London and Cyprus for their constant encouragement and support.

Last but not least, I want to thank my parents Elias and Angela, my sisters Eirini and Marina and my brother Yianni for their loving support and guidance through all these years, and for always believing in me. Without them I would not be where I am today.

Declaration

The work presented in this thesis is the result of the research that I carried out as part of my Ph.D. degree at Imperial College London, between October 2009 and September 2013. This work is my own and contains no materials previously published or written by another person, except where due acknowledgement is made. Any contribution made to my research by others with whom I have worked with is explicitly acknowledged in this thesis. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

Contents

1 Introduction **1**

2 Background Theory **7**

2.1 Stochastic Optimization 7

2.2 Robust Optimization 10

2.3 Expected Utility Theory 13

2.4 Game Theory 17

3 Affine Feedback Controllers in Stochastic Control Problems **21**

3.1 Motivation 22

3.2 Problem Statement and Assumptions 24

3.2.1 Controller Information Structure 28

3.3 Upper Approximations of $\mathcal{P}$ and Primal Affine Control Policies 32

3.4 Lower Approximations of $\mathcal{P}$ and Dual Affine Control Policies 38

3.5 Geometric Properties 49

3.6 Numerical Examples 52

3.6.1 Perfect State Measurements 52

3.6.2 Imperfect State Measurements . . . 53

3.7 Conclusion . . . 54

4 Linear Decision Rules in Two-stage Robust Optimization Problems 57

4.1 Motivation . . . 58

4.2 Problem Statement and Assumptions . . . 60

4.3 Upper Approximations of $\mathcal{P}$ and Primal Linear Decision Rules . . . 60

4.3.1 Dual variables and the robust counterpart $\tilde{\mathcal{U}}$ . . . 61

4.4 Lower Approximations of $\mathcal{P}$ . . . 63

4.4.1 Dual Linear Decision Rules Approximation . . . 63

4.4.2 Scenario Based Approximation . . . 66

4.5 Numerical Results . . . 73

4.5.1 Temporal Networks Example . . . 73

4.5.2 Randomly Generated Instances . . . 74

4.6 Conclusion . . . 75

5 Ambiguous Utilities in Two-Player Games 77

5.1 Motivation . . . 80

5.2 Problem Statement and Assumptions . . . 83

5.2.1 Ambiguity Aversion and Robust Equilibria . . . 86

5.3 Properties of Robust Equilibria . . . 88

5.3.1 Discontinuity in Best-response Strategies . . . 88

5.3.2 Difference Between Infinite and Finite Type Games . . . 91

5.4 Structural Reformulation 101

5.4.1 The Type Sets Θ^i and the Utility Matrix Sets $\mathcal{U}^i$ 103

5.4.2 The Extreme Strategies Equilibrium Condition 107

5.5 The Best-response Functions 110

5.5.1 The Best-response Strategy-set 112

5.5.2 The Best-response Extremes-set 118

5.6 Equilibria Enumeration 125

5.7 Numerical Example 129

5.8 Conclusion 131

6 Conclusions and Future Work **133**

A Numerical Calculations of Chapter 3 **137**

A.1 Perfect State Measurements Example 137

A.1.1 The System Matrices 138

A.1.2 The Noise-process Matrices 139

A.1.3 Calculating the Regions X_ℓ and X_u 140

A.2 Imperfect State Measurements Example 142

B Numerical Calculations of Chapter 4 **145**

B.1 Randomly Generated Problems 145

Bibliography **150**

List of Tables

4.1 Bound comparison for temporal network example. 74

4.2 Bound comparison for polyhedral uncertainty sets. 75

4.3 Bound comparison for spherical uncertainty sets. 75

5.1 The strategies that each player plays at their extreme types $\text{ext}\{\mathcal{U}^i\}$, $i = 1, 2$, at both equilibrium behavioural functions $\hat{b^i}(\cdot)$, $i = 1, 2$, and $b^i(\cdot)$, $i = 1, 2$. 98

5.2 The set $\mathcal{E}^i_{P^i}\left[b^{-i}\right]$ for partitions A^i, C^i, D^i, F^i, expressed through conditions on $b^i = (\underline{b^i}, \overline{b^i})$ for both cases where i is a coordination (COOR) and an anti-coordination (A-COOR) player. 128

5.3 The set $\mathcal{E}^i_{P^i}\left[b^{-i}\right]$ for partitions B^i, E^i expressed via conditions on $b^i = (\underline{b^i}, \overline{b^i})$ for both cases where i is a coordination (COOR) and an anti-coordination (A-COOR) player. The conditions $\underline{b} = \underline{m^i}\left[b^{-i}\right]$ and $\overline{b} = \overline{m^i}\left[b^{-i}\right]$ can be expressed through two-dimensional linear complementarity problems. 129

A.1 The optimal values of the upper and lower bounds of the perfect state measurement example, for each of the horizon lengths $T = 2, \ldots, 10$. 138

A.2 The optimal values of the upper and lower bounds of the imperfect state measurement example, for each of the horizon lengths $T = 2, \ldots, 20$. 143

B.1 The optimal values for the problem instance described in Appendix B.1. 150

List of Figures

3.1 Upper and lower bounds for the perfect state measurements example. 52

3.2 Sets of initial states X_ℓ and X_u for horizon $T = 8$, for the perfect state measurements example. 53

3.3 The upper and lower bounds for the imperfect state measurements example. The two bounds $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ coincide meaning that for this problem instance affine policies are optimal. 54

5.1 All the equilibria of the odd-even game, that can be reached given all possible combinations of increasing and concave utility functions for the two players. The x-axis shows the probability that row assigns to her first action (o) and the y-axis the probability that column assigns to her first action (o). 83

5.2 The nominal utility function (shown as the dashed line) for the risky-safe game. The two solid lines represent the outer envelope of all possible utility functions of each player, given the ambiguity about the utility they assign to £3. 89

5.3 The utility functions of Player 1. The three functions indicated with solid lines in Figures 5.3a and 5.3b correspond to the three extreme points of range $\left(r^1\right)$. 95

5.4 Player 1's and 2's sets $\mathcal{U}^i$, $i = 1, 2$ respectively. The three extreme points (labelled clockwise) correspond to the three extreme utility functions of each player. 96

5.5 The best-response extremes-set $\mathcal{E}^i\left[b^{-i}\right] \subseteq \mathcal{L}$ of a coordination player. The triangle represents the two-dimensional domain of the parameter $b^{-i} = (\underline{b^{-i}}, \overline{b^{-i}}) \in \mathcal{L}$, whilst the labels of the partitions $A^i, \ldots, F^i$, represent the unique values that the correspond to $\mathcal{E}^i\left[b^{-i}\right]$ when b^{-i} is in each partition. 121

5.6 The best-response extremes-set $\mathcal{E}^i[\cdot] \subseteq \mathcal{L}$ of an anti-coordination player. 122

Chapter 1

Introduction

Decision makers often face the task to take decisions whose consequences extend well into the future and cannot be accurately predicted. Unforeseeable events such as variations in stock prices, changes in consumer taste and changes in competitor strategies introduce significant uncertainty regarding the effectiveness of any decision. Such situations arise naturally whenever decision makers operate in environments where they cannot control or measure all the parameters affecting their choices. For example, some parameters may be random, and their realizations are only observed after the decision has been implemented, such as the consumer demand for a product or the travel time on a specific route. Similarly, some parameters may be controlled by competing agents, such as the level of a competitor's bid in an auction or the price of a rival product.

Whether it is the control of a steel mill, the design of investment portfolios, the expansion of production capacity or the pricing of a product, decision makers should account for all the uncertain factors when choosing their strategies. Failing to account for uncertainty often leads to decisions that have undesirable or even disastrous outcomes, such as monetary losses or complete breakdowns of manufacturing processes. As a result, a number of fields are devoted to the study of decision making under uncertainty, including computational optimization, control theory, decision theory and game theory. In this thesis, we investigate three separate but closely related problems of decision making from these various fields, each presenting a unique set of challenges. The differences in each problem arise mainly due to

a) the *source of the uncertainty*, which may be *endogenous* or *exogenous* to the decision making process,

and

b) the *available distributional information* about the uncertain parameters, resulting in either *stochastic* uncertainty (when such information is available) or *epistemic* uncertainty (when no such information is available).

Specifically, we examine two problems with exogenous stochastic and exogenous epistemic uncertainty respectively, as well as one game-theoretic problem that exhibits a combination of endogenous stochastic and exogenous epistemic uncertainty. Each problem is presented in its own, self-contained chapter. The chapters can be summarized as follows.

Chapter 3. In Chapter 3 we investigate the problem of operating a time-variant, linear, discrete-time system over a finite time-horizon under stochastic exogenous uncertainty. More precisely, we study a multi-stage problem where a controller observes the current state of the system (*e.g.* the temperature of a chemical reaction) and then decides on a control action (*e.g.* increase the amount of a catalyser) in order to drive the system to a target state (*e.g.* optimal reaction speed), whilst minimizing the system's expected operation costs (*e.g.* the amount of catalyser used) and remaining within the system's safe operating parameters (*e.g.* temperature tolerances). The exogenous uncertainty affects both the state evolution (*e.g.* through inconsistencies in the catalyser composition) as well as the state measurement (*e.g.* through inaccuracies of the temperature sensor). In our case, the aim of the decision maker is to operate the system so as to minimize an expected quadratic cost function, subject to linear, joint state and input constraints. This class of problems is at the cutting edge of linear, discrete-time stochastic control and has manifold applications in areas as diverse as energy, medicine, engineering, aeronautics and finance. Unfortunately, such problems have been shown to be computationally intractable in general and as a result research has focused on designing suboptimal controls that sacrifice performance in exchange of tractability. A number of different control schemes have been proposed with varying degrees of complexity and resource requirements. We examine one such suboptimal control scheme, where the control policy is restricted to be affine in the observation. We provide a problem-specific measure of the sub-optimality of these controls, by developing a progressive bound on the best possible performance that any control policy, including the optimal one, can achieve. Our approach follows the general method of [60] and is based on a novel dualization technique where the original problem is dualized and the dual variables are then restricted to the space of affine functions. Our methodology

results in a computationally tractable algorithm that calculates a progressive estimate of the best possible performance that can be achieved by the optimal policy given a specific problem. This estimate can be used to benchmark the performance of different approximation schemes, and thus decide whether there is scope to improve a given solution by adopting a more complex control. The contents of this chapter are published in

1. M. J. Hadjiyiannis, P. J. Goulart and D. Kuhn. An efficient method to estimate the suboptimality of affine controllers. *IEEE Transactions on Automatic Control*, 56(12):2841-2853, 2011,

and they have been presented in the UKACC International Conference on Control 2010.

Chapter 4 In Chapter 4 we investigate a dynamic two-stage decision problem with exogenous epistemic uncertainty, linear worst-case objective and linear robust constraints. This is an example of a dynamic robust optimization problem in reactive environments, involving a *design* decision and an *adaptive* decision. The decision process has the following sequence. The decision maker first chooses the design decision (*e.g.* the generation capacity of a power plan), then she observes the realization of the uncertain parameter (*e.g.* the annual demand for electricity) and finally reacts to that observation by selecting an appropriate adaptive decision (*e.g.* adding additional generators to expand capacity). The decision maker faces epistemic uncertainty, where she has no information regarding the probability distribution governing the uncertain parameters. Instead, the decision maker takes a worst-case approach where she assumes that the worst possible realization will occur for any decision she takes (*e.g.* the demand will be small if she decides to build a large plan and vice versa if she decides to build a small plan). The objective of the decision maker is then to choose the decision with the best possible performance in lieu of the worst-case scenario materializing. At the same time, the decision maker needs to guarantee that she satisfies a number of constraints for all possible realizations of the uncertain parameters (*e.g.* always meeting a percentage of the demand). These problems are generally intractable and instead, as in the case of stochastic uncertainty, research is focused on restricting the adaptive decision to be linear in the observation. Whilst this approximation leads to efficiently computable solutions, it has been shown to suffer a loss of optimality of the order $\Omega(\sqrt{m})$ in the worst-case, where m is the number of linear constraints in the problem [17]. Even though this a priori worst-case performance bound is occasionally tight, it is too pessimistic for the vast majority of problem instances. In this chapter, we develop an posteriori (*i.e.* problem-specific)

bound on the performance loss of linear adaptive decisions. To this end, we first demonstrate that the methodology developed in Chapter 3 is unsuitable for deriving such bounds for problems where no distributional information about the uncertain parameters is available. We then propose a different methodology based on a scenario selection technique in which the original problem is evaluated on a finite subset of the uncertain parameter realizations, which are chosen by inspecting the Lagrange multipliers associated with the computed linear adaptive decisions. The contents of this chapter are published in

1. M. J. Hadjiyiannis, P. J. Goulart and D. Kuhn. A scenario approach for estimating the suboptimality of linear decision rules in two-stage robust optimization problems. In *50th IEEE Conference on Decision and Control and European Control Conference*, 7386-7391, 2011.

Chapter 5 In Chapter 5 we examine simultaneous, non-cooperative two-player two-action games with incomplete information, where players have limited information regarding their opponent's utility function. This is an example of a game-theoretic problem that exhibits both stochastic endogenous and epistemic exogenous uncertainty. A simultaneous non-cooperative game is one where players choose their actions at the same time, and the outcome of the game is revealed after all the decisions have been made. Endogenous uncertainty arises because a player's decision (*e.g.* to enter a specific market) affects the other players (*e.g.* the other market participants), who must take precautionary steps (*e.g.* aggressive price discounting or bolstering of their marketing campaigns) in order to protect their interests (*e.g.* their market capitalization). All the players need to anticipate all of their opponents' possible precautions, as these will have a significant bearing on the suitability of their decisions. Game theory describes how each player can account for her opponents' actions, by assigning meaningful probabilities to them and then adjusting her decision according to those probabilities. In our setting, the players engage in competitive games where they know their own as well as their opponents' potential rewards. However, players face exogenous uncertainty in that they do not know their opponents' exact risk preferences, but only that they are risk-averse or that they have non-satiated preferences. This is an example of an incomplete information game with private knowledge. To the best of our knowledge, no algorithm has been published that can solve such games, even for the two-player two-action case. In Chapter 5, we develop a solution methodology for these games based on the robust equilibrium concept [1], which we demonstrate to be a suitable concept for the specific problem. This chapter is the content of the following working paper.

1. M. J. Hadjiyiannis, W. Wiesemann and D. Kuhn. Ambiguous utilities in two-player games. Working paper, *Department of Computing, Imperial College London*,

Chapter 2

Background Theory

2.1 Stochastic Optimization

Stochastic optimization is a computational paradigm that is typically used to formulate and solve decision problems involving stochastic exogenous uncertainty [21, 57, 83]. In its simplest form, a stochastic optimization problem takes the form of

$$\begin{aligned} \min \quad & \mathbb{E}\left[f\left(u,\xi\right)\right] \\ \text{s.t.} \quad & u \in \mathbb{R}^n \\ & g^i\left(u,\xi\right) \leq 0 \quad \mathbb{P}\text{-a.s.} \quad i = 1,\ldots,M, \end{aligned}$$

where $u \in \mathbb{R}^n$ is the decision vector, $\xi \in \mathbb{R}^k$ is the uncertain parameter that is realized according to the known probability distribution $\mathbb{P}$ and $\mathbb{E}\left[\cdot\right]$ is the expectation operator with respect to that distribution. The function $f : \mathbb{R}^n \times \mathbb{R}^k \mapsto \mathbb{R}$ is referred to as the objective function, which in this case might represent a measure of cost or risk that we wish to minimize in expectation. Since we assume full knowledge about the distribution $\mathbb{P}$ of the uncertain parameter ξ, we can potentially compute the multi-dimensional integral in order to evaluate the expectation in the objective. The functions $g^i : \mathbb{R}^n \times \mathbb{R}^k \mapsto \mathbb{R}$, $i = 1,\ldots,M$, are referred to as constraint functions and they represent a set of conditions that any acceptable decision vector u needs to satisfy *almost surely*. The latter implies that the constraints should hold for all possible realizations of the uncertain parameter ξ in the support of its probability distribution $\mathbb{P}$, except possibly for some sets with zero measure. Therefore, if $\mathbb{P}$ is a

continuous probability distribution, it will have an infinite support which will imply that the problem has an infinite number of constraints.

Stochastic optimization offers the advantage that it can model *dynamic* environments where the uncertain parameters are revealed in a sequential manner as time progresses, allowing for decisions to be taken dynamically as these parameters become observed. In such cases, the optimization problem is given as

$$\begin{aligned} \min \quad & \mathbb{E}\left[\sum_{t=1}^{T} f_t\left(u_t\left(\xi^t\right), \xi^t\right)\right] \\ \text{s.t.} \quad & u_t \in \mathcal{L}^2_{kt,n},\ t = 1, \ldots, T \\ & g_t^i\left(u_t\left(\xi^t\right), \xi^t\right) \leq 0 \quad \mathbb{P}\text{-a.s.}, \quad i = 1, \ldots, M_t, \quad t = 1, \ldots, T. \end{aligned}$$

This dynamic decision problem with stochastic exogenous uncertainty has the following structure. At time t, the decision maker observes a realization of the uncertain parameter $\xi_t \in \mathbb{R}^k$. Then, based on the history of observations up to that time period t, denoted by $\xi^t = (\xi_1, \ldots, \xi_t) \in \mathbb{R}^{kt}$, she chooses a decision $u_t\left(\xi^t\right) \in \mathbb{R}^n$ in order to minimize the objective function. This sequence of alternating observations and decisions occurs for all time periods $t = 1, \ldots, T$, where T is the finite *time-horizon* of the problem. In order to account for the arrival of new information at each time period $t = 1, \ldots, T$, the decision vectors $u_t \in \mathcal{L}^2_{kt,n}$ are no longer finite vectors in $\mathbb{R}^n$, but instead they are functions of the history ξ^t of all the uncertain parameters that have been observed up to that point. Furthermore, the functions $u_t(\cdot)$ must be square-integrable so as to ensure that the expectation in the objective is well-defined. Finally, care is taken so that each function $u_t(\cdot)$ depends only on those parameters $\xi_1, \ldots, \xi_t$ that have been observed up to that period t, and does not depend on any of the parameters $\xi_{t+1}, \ldots, \xi_T$ that have not yet been observed. This feature reflects the *non-anticipative* nature of the dynamic decision problem and ensures its causality.

Stochastic optimization also allows the modelling of problems with *state dynamics*. These problems are prevalent in control theory, where the decision maker aims to design a *control policy* that drives a dynamical system towards a desired state whilst simultaneously satisfying a number of constraints and minimizing a specific objective. In order to accommodate for this feature, a state variable is introduced

along with a function describing the state dynamics. Such problems are expressed as

$$
\begin{aligned}
\min \quad & \mathbb{E}\left[\sum_{t=1}^{T} f_t\left(u_t\left(x^t\right), x_t\left(\xi^{t-1}\right), \xi^t\right)\right] \\
\text{s.t.} \quad & u_t \in \mathcal{L}^2_{mt,n},\ t = 1, \ldots, T \\
& x_t \in \mathcal{L}^2_{k(t-1),m},\ t = 2, \ldots, T+1 \\
& \left.\begin{aligned}
& x_{t+1}\left(\xi^t\right) = h_t\left(x_t\left(\xi^{t-1}\right), u_t\left(x^t\right), \xi^t\right) \quad \mathbb{P}\text{-a.s.} \\
& g_t^i\left(u_t\left(x^t\right), x_t\left(\xi^{t-1}\right), \xi^t\right) \leq 0 \quad \mathbb{P}\text{-a.s.} \quad i = 1, \ldots, M_t
\end{aligned}\right\} \quad t = 1, \ldots, T \\
& g_{T+1}^i\left(x_{T+1}\left(\xi^T\right)\right) \leq 0 \quad \mathbb{P}\text{-a.s.} \quad i = 1, \ldots, M_{T+1}.
\end{aligned}
$$

Here, the system starts at a known state $x_1\left(\xi^0\right) = x_1 \in \mathbb{R}^n$. At each time period $t = 1, \ldots, T$, the decision maker observes the realized state $x_t\left(\xi^{t-1}\right) \in \mathbb{R}^m$, and decides on an a control input $u_t\left(x^t\right) \in \mathbb{R}^n$ based on the history of the observed states up to that time period, given by $x^t = \left(x_1\left(\xi^1\right), \ldots, x_t\left(\xi^{t-1}\right)\right) \in \mathbb{R}^{mt}$. Once the input is chosen, the state evolves according to the state dynamical functions $h_t(\cdot)$. There are a few but important differences when compared to the problem without state dynamics. The objective function and the constraints depend jointly on the states and the inputs, whilst the terminal state $x_{T+1}\left(\xi^T\right)$ needs to satisfy a set of *terminal constraints*. The vector of the uncertain parameters $\xi_t \in \mathbb{R}^k$, or *noise*, now affects how the current state $x_t\left(\xi^{t-1}\right)$ will evolve to the next state $x_{t+1}\left(\xi^t\right)$, and is realized after the current state $x_t\left(\xi^{t-1}\right)$ has been observed and the control input $u_t\left(x^t\right)$ has been chosen. Thus, the state variable $x_t \in \mathcal{L}^2_{k(t-1),m}$ at time period $t = 2, \ldots, T+1$, depends on the history of the realized noise $\xi^{t-1} = \xi_1, \ldots, \xi_{t-1} \in \mathbb{R}^{k(t-1)}$ up to the previous time period. Furthermore, the decision maker now does not observe the uncertain parameters $\xi_t \in \mathbb{R}^k$ directly, but can only indirectly observe them through their influence on the state $x_{t+1}\left(\xi^t\right)$. Thus, the control functions (control policies) $u_t \in \mathcal{L}^2_{mt,n}$, $t = 1, \ldots, T$, are now functions of the history of the observed states x^t instead of the history of the realized uncertain parameters ξ^t. One can extend these problems further by introducing a measurement variable, that is, a potentially noisy projection of the state. The decision maker then no longer observes the state directly, but only through its noisy measurement, and thus the control policies at each time period are now functions of the state measurement.

In Chapter 3, we investigate such dynamical control problems with quadratic state and input objective, linear joint state and input constraints, linear state dynamics and linear measurements, both of which are affected by bounded uncertainty. Computing an optimal control policy for such systems is

computationally intractable in general [82]. Methods for solving such problems typically rely on some variation of robust dynamic programming [15], multi-parametric programming techniques [37, 76] or vertex enumeration methods [81], and they typically require the solution of optimization problems that grow exponentially with the size of the problem data. In this thesis, we investigate another solution mechanism based on an approximation of the original problem, where the control policies are restricted to be affine functions of the measurements. This approach has been studied by a number of authors, and it is a common theme in the literature of control theory and stochastic optimization [11, 21, 28, 29, 40, 61, 63, 67, 82, 86].

2.2 Robust Optimization

Robust optimization is a powerful modelling paradigm for decision making under uncertainty [9, 10, 11, 12]. It is tailored to decision problems in which the distribution of the uncertain parameters is unknown except for its support. By definition, the support represents the range of all possible parameter realizations and is commonly referred to as the *uncertainty set*. Robust optimization models are designed to find the best decision in view of the worst-case realization of the uncertain parameters within their uncertainty set.

Classical *static* robust optimization models involve only *design decisions*, which are of here-and-now type and must be selected before any of the uncertain parameters are observed. A typical static robust optimization problem has the form of

$$\begin{array}{ll} \min & \max\limits_{\xi \in \Xi} f(u, \xi) \\ \text{s.t.} & u \in \mathbb{R}^n \\ & g^i(u, \xi) \leq 0 \quad \forall \xi \in \Xi \quad i = 1, \ldots, M, \end{array}$$

where $u \in \mathbb{R}^n$ is the design decision, $\xi \in \Xi$ is the vector of uncertain parameters and $\Xi \subseteq \mathbb{R}^k$ is the support of these uncertain parameters. The goal is to choose a decision that minimizes the objective function under the assumption that the realized uncertain parameter will be one that maximizes that objective. Robust optimization problems are usually reformulated using an epigraph variable. That

is, by introducing a scalar $t \in \mathbb{R}$, the problem can be expressed as

$$\begin{array}{ll} \min & \rho \\ \text{s.t.} & u \in \mathbb{R}^n,\, \rho \in \mathbb{R} \\ & \left.\begin{array}{l} f(u,\xi) \le \rho \\ g^i(u,\xi) \le 0 \quad i = 1,\ldots,M \end{array}\right\} \quad \forall \xi \in \Xi. \end{array} \tag{2.1}$$

The constraint $f(u,\xi) \le \rho,\, \forall \xi \in \Xi$ guarantees that the max-function $\max_{\xi \in \Xi} f(u,\xi) \le \rho$, and since there is no other constraint with regards to variable ρ, the minimization over ρ ensures that this constraint is satisfied at equality. The constraints in the optimization problem are enforced robustly, *i.e.* for all possible realizations of $\xi \in \Xi$. When the uncertainty set Ξ is infinite, the problem (2.1) represents a semi-infinite optimization problem, as it has finitely many variables but infinitely many constraints (*i.e.* for each of the elements in Ξ), which render the problem intractable [12, 73] even in the case where the constraint functions are linear. Fortunately, depending on the support Ξ and the nature of the constraints of the problem, one can derive an equivalent finite-dimensional reformulation of problem (2.1), which is referred to as the *robust counterpart* [9, 12, 13]. This is achieved by utilizing a known result from duality theory, based on the *dual cone* of the set Ξ.

A cone is defined as a set that is closed under multiplications with non-negative scalars. Let

$$\mathcal{K} := \operatorname{cone}(\Xi) = \left\{ z \in \mathbb{R}^k \;:\; z = \lambda\xi,\, \lambda \ge 0,\, \xi \in \Xi \right\}$$

be the cone generated by the set Ξ. The dual cone is then defined as follows.

Definition 2.2.1 (Dual cone)**.** *[22, §2.6.1] For any cone $\mathcal{K} \subseteq \mathbb{R}^k$, its dual cone is given by*

$$\mathcal{K}^* = \left\{ \sigma \in \mathbb{R}^k \;:\; \sigma^\top z \ge 0,\, \forall z \in \mathcal{K} \right\}.$$

The dual cone $\mathcal{K}$ is always convex even if the primal cone $\mathcal{K}$ is not.

The definition of the dual cone is the key ingredient in reformulating the infinite number of constraints in problem (2.1) into finitely many equivalent constraints. Indeed, for any linear function of ξ

$$g(u,\xi) := u^\top A\xi + \mu^\top \xi = \left(u^\top A + \mu^\top\right)\xi,$$

where $A \in \mathbb{R}^{n\times k}$ and $\mu \in \mathbb{R}^k$, we have by linear homogeneity that

$$\left(u^\top A + \mu^\top\right)\xi \leq 0 \quad \forall \xi \in \Xi \quad \Longleftrightarrow \quad \left(u^\top A + \mu^\top\right) z \leq 0 \quad \forall z \in \mathcal{K}.$$

Then, by definition of the dual cone, the following equivalence also follows:

$$-\left(u^\top A + \mu^\top\right) z \geq 0 \quad \forall z \in \mathcal{K} \quad \Longleftrightarrow \quad -\left(u^\top A + \mu^\top\right)^\top \in \mathcal{K}^*.$$

If Ξ has a benign structure, *i.e.* it can be expressed as an intersection of half-spaces and ellipsoids, then the latter constraint takes the form of a finite, tractable cone or linear constraint [10, 12]. We will use this result throughout the thesis and investigate it in more detail in Lemmas 3.3.1 and 4.3.1. Consequently, if problem (2.1) involves an objective and constraint functions that are linear in ξ, whilst the support Ξ has a benign structure allowing for the efficient calculation of its dual cone $\mathcal{K}^*$, then the above result can be used to reformulate problem (2.1) as a tractable optimization problem.

Recently, this result has also been used in *dynamic* robust optimization problems [8, 9, 10, 11, 12, 22]. Such problems incorporate additional *adaptive decisions*, which are of wait-and-see type and can be selected after the uncertain parameters have been revealed. A typical dynamic robust optimization problem takes the form

$$\begin{aligned} \min \quad & c^\top u_1 + \max_{\xi\in\Xi} d^\top u_2(\xi) \\ \text{s.t.} \quad & u_0 \in \mathbb{R}^{n_1},\ u_1 \in \mathcal{L}_{k,n_2} \\ & g^i(u_1, u_2, \xi) \leq 0 \quad \forall \xi \in \Xi \quad i = 1, \ldots, M, \end{aligned}$$

where $u_1 \in \mathbb{R}^{n_1}$ is the first-stage design decision that does not depend on the uncertain parameters, and $u_2 \in \mathcal{L}_{k,n_2}$ is the second-stage adaptive decision, also known as a *decision rule*, which is a function from the uncertainty set Ξ to $\mathbb{R}^{n_2}$. By restricting the adaptive decisions to be linear in ξ, *e.g.* $u_2(\xi) = U\xi$ for some $U \in \mathbb{R}^{n_2\times k}$, we ensure that the objective function is also linear in ξ. Then, if the constraint functions $g^i(\cdot)$ are also linear in ξ when we make the substitution $u_2(\xi) = U\xi$, we can use the aforementioned result to solve such problems in polynomial time. These problems are the subject of investigation in Chapter 4.

2.3 Expected Utility Theory

Expected utility (EU) theory is a decision theory describing how decision makers value different risky propositions using the available information, in order to choose the one they prefer the most. These propositions are expressed as *lotteries*, which are essentially probability distributions over a known, finite set of potential outcomes. Expected utility theory traces its routes back to Daniel Bernoulli, who introduced the notion of a utility function as a means to resolve the St. Petersburg Paradox [14].

The St. Petersburg paradox, first stated by Nicolas Bernoulli in 1713, is a lottery game highlighting the discrepancy between the value that people are willing to assign to a specific lottery, compared to that lottery's expected reward. The St. Petersburg paradox describes a game of chance for a single player in which a fair coin is tossed at each stage. The pot starts with a dollar and is doubled every time a head appears. The first time a tail appears, the game ends and the player wins whatever is in the pot. Thus, at the first toss, the player has a 50%-chance of winning one dollar and 50%-chance of proceeding to the next round where the pot is doubled and the game is repeated. The expected value of this game (where X denotes the lottery representing the game) is then given by

$$\begin{aligned}\mathbb{E}[X] = &\ \tfrac{1}{2}1 + \tfrac{1}{2}\left(\tfrac{1}{2}2 + \tfrac{1}{2}\left(\tfrac{1}{2}4 + \tfrac{1}{2}(\dots)\right)\right)\\ = &\ \tfrac{1}{2}1 + \tfrac{1}{4}2 + \tfrac{1}{8}4 + \tfrac{1}{16}8\dots\\ = &\ \textstyle\sum_{i=1}^{\infty}\frac{1}{2} = \infty.\end{aligned}$$

Any person that takes into consideration only the expected value of a lottery in her decisions, would prefer to participate in that game than to receive any certain sum of money, regardless of how generous it is, a choice that presumably no rational individual would be actually willing to take.

Bernoulli's resolution of the paradox in 1738 (republished in [14]) came with the explicit introduction of a *utility function*, an *expected utility hypothesis*, and the presumption of *diminishing marginal utility of money*. In Daniel Bernoulli's own words: "The determination of the value of an item must not be based on the price, but rather on the utility it yields[...]. There is no doubt that a gain of one thousand ducats is more significant to the pauper than to a rich man though both gain the same amount". Then, according to the expected utility hypothesis, people value any risky proposition based just on its expected utility, not its expected monetary reward. Bernoulli went as far as to suggest the logarithmic function as a potential utility function, where a monetary reward c results in a utility

of $u(c) = \log(c)$. The logarithmic utility function, also known as log-utility, has the advantage that it models the concept of diminishing marginal utility of money, where higher rewards do not result in a proportional increase in utility. A gambler with a log-utility would value the game in the St. Petersburg paradox as

$$\begin{aligned}\mathbb{E}\left[u\left(X\right)\right] = &\ \tfrac{1}{2}u(1) + \tfrac{1}{2}\left(\tfrac{1}{2}u(2) + \tfrac{1}{2}\left(\tfrac{1}{2}u(4) + \tfrac{1}{2}\left(\ldots\right)\right)\right) \\ = &\ \textstyle\sum_{i=1}^{\infty} \frac{1}{2^i} u\left(2^{i-1}\right) = \sum_{i=1}^{\infty} \frac{1}{2^i} \log\left(2^{i-1}\right) < \infty.\end{aligned}$$

Whilst Bernoulli set the foundations of EU theory, it was von-Neumann and Morgenstern that transformed Bernoulli's EU hypothesis into a rigorous theory. In [87], von-Neumann and Morgenstern proposed a set of four axioms that every rational individual should adhere to when presented with a choice between any given lotteries L, M and N. These four axioms, called the VNM-axioms, represent a set of axiomatic statements which describe rational behaviour and are as follows:

- Axiom 1 (Completeness): Either lottery L is preferred, in which case $M \prec L$, lottery M is preferred, in which case $L \prec M$, or there is no preference between the two, in which case $L \approx M$.
- Axiom 2 (Transitivity): If $L \prec M$ and $M \prec N$ then $L \prec N$.
- Axiom 3 (Continuity): If $L \prec M \prec N$, then there exists a probability $p \in (0, 1)$ such that $pL + (1-p)N \approx M$.
- Axiom 4 (Independence): If $L \prec M$, then for any N and $p \in (0, 1]$, $pL+(1-p)N \prec pM+(1-p)N$.

Whilst the four axioms are fairly unassuming, their consequence is the central pillar of EU theory, as stated in the following theorem.

Theorem 2.3.1 (Expected Utility Theorem)**.** *[87] For any agent whose preferences satisfy the VNM-axioms, there exists a function $u\left(\cdot\right)$ assigning to any outcome c, in the lotteries L and M, a real number $u(c)$, such that*

$$L \prec M \iff \mathbb{E}\left[u\left(L\right)\right] < \mathbb{E}\left[u(M)\right],$$

where $\mathbb{E}\left[u\left(L\right)\right]$, $\mathbb{E}\left[u\left(M\right)\right]$ denote the expected values of $u\left(\cdot\right)$, given the probabilities associated with each outcome in L and M respectively. Conversely, any agent acting to maximize the expectation of a utility function $u\left(\cdot\right)$ will obey the four VNM-axioms.

According to the EU theorem, any rational agent adhering to the four VNM-axioms of rationality will, when having to choose among different propositions, choose the one that maximizes her personal utility function in expectation.

One of the major criticisms of EU theory is its postulation of the absoluteness of probabilities, where rational agents use the same probabilities to calculate their expected utility, regardless of the difference in information that is available to each one of them. Any difference in their decisions is only due to the difference in their utility functions, not their personal beliefs or perceptions. In order to overcome this limitation, Savage introduced the subjective expected utility (SEU) theory [79], where he combined the personal utility function with a personal Bayesian probability distribution. Savage then showed that if an agent is VNM-rational with a personal utility function $u(\cdot)$ and a personal belief $\mathbb{P}_L(c_i)$ regarding the probability of each outcome c_i, $i = 1, \ldots, n$ in the lottery L materializing, then the agent assigns a subjective expected utility to the lottery

$$\mathbb{E}_{\mathbb{P}_L}[u(L)] = \sum_{i=1}^{n} \mathbb{P}_L(L = c_i)u(c_i).$$

Thus, two agents with the same utility function may value a lottery differently because they have different beliefs about the probabilities of each outcome in that lottery materializing.

Subjective expected utility theory is one of the most popular decision theories, used widely in game theory and economics. However, the theory has significant shortcomings, as demonstrated by the Ellsberg [35] and the Allais [2] paradoxes. These paradoxes have led to the development of additional theories, such as the max-min expected utility theory [43], the Croquet expected utility [42], and other non-expected utility theories [85]. Researches have also proposed abandoning utility theory altogether, with prospect theory being the main candidate to replace it [55]. For a review of alternative utility theories we refer the reader to [26, 55, 85].

2.4 Game Theory

Game theory is a branch of decision science that investigates decision making in environments involving multiple agents with distinct motives and priorities. The central theme of game theory is that each player's reward in a particular game does not depend only on that player's decision, but on the decisions

made by all the players of the game.

A typical simultaneous, one-shot, two-player two-action game is given in strategic form using the matrix notation as follows:

	a_1^2	a_2^2
	A	C
a_1^1	a	b
	B	C
a_2^1	c	d

where player 1 determines the row of the matrix where the outcome of the game resides, whilst player 2 determines the column. Hence, each player i must choose an action from the set $\mathcal{A}^i := \left\{a_1^i, a_2^i\right\}$, $i = 1, 2$. When both players choose their first action a_1^i, then player 1 obtains a payoff of a and player 2 a payoff of A.

In a complete information setting, all the parameters of the game, such as each player's action set and payoffs, are *common knowledge*. However, in a *non-cooperative* environment the players are unable to pre-commit to a specific action or communicate their intentions truthfully to the opponents. As a result, the players may need to choose their actions according to a probability distribution, referred to as a *mixed strategy*. Player i's mixed strategy then resides in a probability simplex Δ^i containing all probability distributions over her action set $\mathcal{A}^i$. For a mixed strategy $s^i \in \Delta^i$, the j-th element of s^i denotes the probability with which player i chooses her j-th action. When the players play according to mixed strategies $s^i \in \Delta^i$, $i \in \mathcal{N}$, we can calculate the expected payoff that they receive by using their individual payoff matrices

$$P^1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad P^2 = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Then, each player i receives an expected payoff of

$$\pi\left[s^i;\, P^i, s^{-i}\right] := {s^i}^\top P^i s^{-i},$$

where $-i$ denotes the opponent of i.

Game theory studies how each player decides on her mixed strategy in non-cooperative games, and

the most prevalent solution concept is due to Nash [72], who in turn builds on von Neumann's and Morgenstern's EU theory [87]. According to EU theory, a decision maker that adheres to the VNM-axioms of rationality always chooses the actions that maximize her expected utility. Thus, Nash postulated that if a player i is rational and knows that her opponent chooses a strategy $s^{-i} \in \Delta^{-i}$, then player i must choose an action only from the set of strategies

$$\mathcal{S}^i\left[s^{-i}\right] := \underset{s^i \in \Delta^i}{\arg\max}\ \pi\left[s^i;\, P^i, s^{-i}\right],$$

that maximize i's expected utility. Nash then defined an *equilibrium* as a pair of such strategies, one for each player, where no rational player has an incentive to unilaterally deviate from her corresponding strategy. That is, no player can increase her expected utility by unilaterally playing a strategy other than her respective equilibrium strategy. Hence, in a two-player game, an equilibrium is a pair of strategies $s^i \in \Delta$, $i = 1, 2$, satisfying

$$s^i \in \mathcal{S}^i\left[s^{-i}\right], \quad i = 1, 2.$$

In his famous proof [72], Nash showed that any complete information game with finitely many players and finite action sets will always admit such an equilibrium in mixed strategies.

Incomplete Information Games

Incomplete information games are games where players do not know with certainty all the parameters of the game's structure. Instead, some of the parameters, such as the players' available actions or payoffs, are either random and realized only after the players choose their actions, or they are only known to some of the players but not to everyone. Incomplete information games were initially proposed by von Neumann and Morgenstern in 1944 [87], but they were not solved until 1967 when Harsanyi introduced Bayesian games and the concept of the Bayes equilibrium [49].

Bayesian games are a class of incomplete private information games, where private information refers to knowledge that some players have that is not available to the others, such as their personal utility function, available actions, economic constraints or physical resources. Harsanyi used the notion of *types*, where each player's type $\theta^i \in \Theta^i$ is an encoding of all the private information available to that

player i. He showed that all forms of private information can be represented as private information about each player's payoff matrix. Thus, a player's type $\theta^i \in \Theta^i$ is used to determine her payoff matrix, using her payoff function $r^i\left(\theta^i\right) := P^i$. Incomplete information is then modelled by having players know their own type but be ignorant of their opponents' exact types. Instead, players share a *common prior* $\mathbb{P}$, that is, a common knowledge joint probability distribution over all the players' sets $\underset{i}{\times} \Theta^i$ of possible types. Particularly, any Bayesian incomplete information game with private knowledge can be reformulated so that each player's set of potential actions $\mathcal{A}^i$, set of potential types Θ^i, payoff functions $r^i(\cdot)$ and common prior $\mathbb{P}$ over the sets of types, are common knowledge. Each player i knows her own type $\theta^i \in \Theta^i$ but does not know her opponent's type, only that it belongs in the type set Θ^{-i} and that it is realized according to the common prior $\mathbb{P}$.

Harsanyi observed that a utility-maximizing player would play a different strategy for each of her types $\theta^i \in \Theta^i$, according to her type-dependent payoff matrix $r^i\left(\theta^i\right)$. Thus, each player now plays according to a *behavioural function*

$$b^i : \Theta^i \mapsto \Delta^i.$$

As opposed to complete information games, players now face an opponent that does not have a uniquely defined strategy. Instead, the opponent's strategy depends on her type, which the player does not know with certainty. An equilibrium must thus be defined in terms of behavioural functions, not individual strategies.

Harsanyi's major breakthrough was to show that players can derive subjective but meaningful probabilities regarding their opponent's unknown strategy. In order to do so, the players use the common prior along with their type to derive the conditional probabilities over the opponent's type space. They then use this conditional probability to calculate their subjective beliefs as to how a rational opponent will play. Then, according to SEU theory, if the players adhere to the VNM-axioms of rationality [79], they will choose strategies that maximize their subjective expected utility given these beliefs. Formally, if player i is of type $\theta^i \in \Theta^i$ and given her opponent's behavioural function $b^{-i}(\cdot)$ along with the common prior $\mathbb{P}$, then player i's best response set is given by:

$$\mathcal{S}^i\left[\theta^i\,;\, b^{-i}(\cdot)\right] := \underset{s^i \in \Delta}{\arg\max}\ \mathbb{E}_{\mathbb{P}}\left[\pi\left[s^i;\, r^i(\theta^i),\, b^{-i}\left(\theta^{-i}\right)\right]|\theta^i\right].$$

Thus, similar to the complete information case, an equilibrium is defined as a set of behavioural

functions, one for each player, where no rational player has an incentive to unilaterally deviate from the strategy prescribed by her behavioural function for each of her possible types. Thus, for a two-player game, an equilibrium is given as a pair of behavioural functions $b^i(\cdot)$, $i = 1, 2$, satisfying

$$b^i(\theta^i) \in \mathcal{S}^i\left[\theta^i\,;\, b^{-i}(\cdot)\right] \quad \forall \theta^i \in \Theta^i, \quad i = 1, 2.$$

We investigate such games in Chapter 5. For more details with regards to game theory and incomplete information games, we refer the reader to [49, 66, 71, 75].

Chapter 3

Affine Feedback Controllers in Stochastic Control Problems

In this chapter, we consider robust feedback control of time-varying, linear discrete-time systems operating over a finite horizon. For such systems, we consider the problem of designing robust causal controllers that minimize the expected value of a convex quadratic cost function, subject to mixed linear state and input constraints. Determination of an optimal control policy for such problems is generally computationally intractable, but suboptimal policies can be computed by restricting the class of admissible policies to be affine on the observation. By using a suitable re-parameterization and robust optimization techniques, these approximations can be solved efficiently as convex optimization problems. We investigate the loss of optimality due to the use of such affine policies. Using duality arguments and by imposing an affine structure on the dual variables, we provide an efficient method to estimate a lower bound on the value of the optimal cost function for any causal policy, by solving a cone program whose size is a polynomial function of the problem data. This lower bound can then be used to quantify the loss of optimality incurred by the affine policy.

3.1 Motivation

We are interested in characterizing the degree of suboptimality of affine feedback policies for linear discrete-time systems with mixed state and input constraints, bounded disturbances and an expected

value cost.

The problem of computing an optimal control policy for such systems, either in a minimax or expected value sense, is computationally intractable in general, and current solution methods involve optimization problems that typically grow exponentially with the size of the problem data [15, 37, 76, 81]. As a result, significant research effort has focused on methods for finding *suboptimal* control policies that can be computed via solution of a tractable optimization problem.

A common approach is to restrict the class of control policies considered to those based on perturbations to some fixed stabilizing linear controller [61, 67]. More generally, one can compute a controller based on affine disturbance or measurement-error feedback, a technique suggested by a number of authors [11, 63, 86]. In the affine feedback case, characterization of the set of constraint admissible policies is achieved by following the general approach proposed in [11] for robust optimization problems with linear decision rules, leading to a computationally tractable optimization problem.

An attractive feature of such affine parameterizations is that they can be shown to be equivalent (in the state feedback case) to parameterizations of control policies as affine functions of prior states [48, 84], or (in the output feedback case) as affine functions of prior measurements [9, 46]. The idea underpinning these equivalence results is akin to that of the well-known Youla parameterization (or Q-parameterization) in linear systems [91], and relies on a similar nonlinear transformation to produce a convex set of constraint admissible policies over which one can optimize.

Further refinements to the basic idea of affine uncertainty feedback policies have also been proposed, *e.g.* policies employing "segregated" disturbance feedback based on some partitioning of the uncertainty set [41, 88]. However, a fundamental difficulty with all of the present proposals is that they generally do not provide any estimate of the degree of suboptimality introduced by restricting the optimal control problem to the particular class of control policies proposed. One notable exception to this is in the case of affine disturbance feedback policies for SISO systems, for which such policies can be shown to be optimal for minimax problems in very limited circumstances [18].

In this chapter we provide a general method for estimating the degree of suboptimality of affine feedback controllers when minimizing the expected value of a quadratic cost, for systems with polyhedral state and input constraints and uncertainties whose support is characterized by conic constraints. Our approach follows the general method of [60], and is predicated on a dualization of the original optimal

control problem followed by a restriction of the dual variables to those parameterized by a linear decision rule. This results in a tractable optimization problem that provides a *lower* bound on the finite horizon cost achievable, and is a natural counterpart to the *upper* bound on the achievable cost that is found when restricting the class of control policies to those in affine feedback form. The gap between these bounds then serves as an estimate of the worst-case suboptimality of controllers based on affine feedback policies.

The ability to compute the suboptimality gap in an efficient manner provides a valuable insight to anyone wishing to design a control system. If the gap is small, then affine policies are near-optimal and there is little room for improvement. On the contrary, if the gap is large one may consider investing more time and effort to improve on the affine policies, *e.g.* by using deflected or segregated policies as in [41, 45, 88]. One attractive property of the proposed approach is that due to the symmetry between the upper and lower bound, one may improve both bounds and thus decrease the optimality gap by employing similar techniques on both the primal as well as the dual controllers (*e.g.* segregation of the uncertainty space [41]).

The chapter is organized as follows: Section 3.2 describes the problem of interest and details a number of standing assumptions. Section 3.3 outlines the restriction of control policies to those in affine form, and shows how such policies can be calculated via a tractable conic optimization problem. Section 3.4 describes a method for estimating the degree of suboptimality of affine feedback policies, based on a novel approximation to the dual of the optimal control problem of interest. We show how this method allows us to compute lower bounds on the optimal cost via solution of a tractable conic optimization problem. Section 3.5 summarizes a number of general observations about the relationship between the lower- and upper-bounding problems discussed in the previous sections. Section 3.6 presents numerical examples illustrating the efficacy of the proposed method, with some conclusions drawn in Section 3.7.

Chapter Notation All random vectors appearing in this chapter are defined on an abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $\mathbb{E}(\cdot)$ denotes the expectation operator with respect to $\mathbb{P}$. Random vectors will be represented in boldface, while their realizations will be denoted by the same symbols in normal face. For notational convenience, we denote by $\mathcal{L}_n^2 := \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$ the space of all $\mathcal{F}$-measurable, square-integrable random vectors valued in $\mathbb{R}^n$. Given a stochastic process $\boldsymbol{z} := (\boldsymbol{z}_0, \ldots, \boldsymbol{z}_T)$, we will denote as $\boldsymbol{z}^t := (\boldsymbol{z}_0, \ldots, \boldsymbol{z}_t)$ the history of the $\boldsymbol{z}$ process up to time t.

The identity matrix in $\mathbb{R}^n$ is denoted $I^{(n)}$ — the superscript will be omitted when the dimension is clear from the context. For matrices $(A, B) \in \mathbb{R}^{m\times n}$, $\ker(A)$ denotes the kernel or null space of A, $A^\dagger$ its pseudo-inverse, and $A \geq B$ denotes component-wise inequality. For every $C \in \mathbb{R}^{n\times n}$ $\mathrm{tr}(C)$ denotes the trace of C. We denote the p-norm in $\mathbb{R}^n$ as $\|\cdot\|_p$ for any $p \in [1, \infty]$. The p-order cone in $\mathbb{R}^{n+1}$ is denoted $\mathcal{K}_p := \{(t, x) \in \mathbb{R}^{n+1} : \|x\|_p \leq t\}$, and its dual cone as $\mathcal{K}_p^* = \mathcal{K}_q$, where $1/p + 1/q = 1$. For any $(y, z) \in \mathbb{R}^{n+1}$ the relation $y \succeq_{\mathcal{K}_p} z$ implies that $(y - z) \in \mathcal{K}_p$. For a matrix $D \in \mathbb{R}^{(n+1)\times(n+1)}$, $D \succeq_{\mathcal{K}_p} 0$ denotes column-wise inclusion in $\mathcal{K}_p$. The vector e_0 denotes the unit vector $e_0 = (1, 0, \ldots, 0)$.

3.2 Problem Statement and Assumptions

In this chapter we are concerned with linear time-varying systems with known state dynamics

$$\boldsymbol{x}_{t+1} = A_t\boldsymbol{x}_t + B_t\boldsymbol{u}_t + C_t\boldsymbol{\xi}_t \quad \text{for} \quad t = 1, \ldots, T-1, \tag{3.1a}$$

where $\boldsymbol{x}_t \in \mathcal{L}^2_{n_x}$ denotes the state of the system with $\boldsymbol{x}_1 = x_1$ known, $\boldsymbol{u}_t \in \mathcal{L}^2_{n_u}$ denotes the control input, and $\boldsymbol{\xi}_t \in \mathcal{L}^2_{n_\xi}$ the process noise. The state of the system is partially observable in the sense that the measured output $\boldsymbol{y}_t \in \mathcal{L}^2_{n_y}$ at time t depends linearly on the state and the noise, *i.e.*

$$\boldsymbol{y}_t = D_t\boldsymbol{x}_t + E_t\boldsymbol{\xi}_t \quad \text{for} \quad t = 1, \ldots, T-1. \tag{3.1b}$$

The system equations (3.1) can be rewritten compactly in matrix notation as

$$\boldsymbol{x} = B\boldsymbol{u} + C\boldsymbol{\xi} \quad \text{and} \quad \boldsymbol{y} = D\boldsymbol{x} + E\boldsymbol{\xi}\,, \tag{3.2}$$

with

$$\begin{aligned}
\boldsymbol{x} &:= (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T) \in \mathcal{L}^2_{N_x}, & N_x &:= n_x T \\
\boldsymbol{u} &:= (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{T-1}) \in \mathcal{L}^2_{N_u}, & N_u &:= n_u(T-1) \\
\boldsymbol{\xi} &:= (\boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_{T-1}) \in \mathcal{L}^2_{N_\xi}, & N_\xi &:= 1 + n_\xi(T-1) \\
\boldsymbol{y} &:= (\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{T-1}) \in \mathcal{L}^2_{N_y}, & N_y &:= 1 + n_y(T-1),
\end{aligned}$$

where $\boldsymbol{\xi}_0$ and $\boldsymbol{y}_0$ represent degenerate random variables that are almost surely equal to 1. The matrices $B \in \mathbb{R}^{N_x\times N_u}$, $C \in \mathbb{R}^{N_x\times N_\xi}$, $D \in \mathbb{R}^{N_y\times N_x}$ and $E \in \mathbb{R}^{N_y\times N_\xi}$ in (3.2) are easily constructed from the

problem data in (3.1) and are defined as

$$B := \begin{bmatrix} 0 & & & & \\ \mathcal{A}_2^2 B_1 & 0 & & & \\ \mathcal{A}_2^3 B_1 & \mathcal{A}_3^3 B_2 & 0 & & \\ \vdots & & & \ddots & \\ \vdots & & & & 0 \\ \mathcal{A}_2^T B_1 & \mathcal{A}_3^T B_2 & \cdots & \cdots & \mathcal{A}_T^T B_{T-1} \end{bmatrix}, \quad C := \begin{bmatrix} \mathcal{A}_1^1 x_1 & & & & \\ \mathcal{A}_1^2 x_1 & \mathcal{A}_2^2 C_1 & & & \\ \mathcal{A}_1^3 x_1 & \mathcal{A}_2^3 C_1 & \mathcal{A}_3^3 C_2 & & \\ \vdots & & & \ddots & \\ \mathcal{A}_1^T x_1 & \mathcal{A}_2^T C_1 & \mathcal{A}_3^T C_2 & \cdots & \mathcal{A}_T^T C_{T-1} \end{bmatrix},$$

$$D := \begin{bmatrix} 0 & & & \\ D_1 & 0 & & \\ & \ddots & \ddots & \\ & & D_{T-1} & 0 \end{bmatrix} \quad \text{and} \quad E := \begin{bmatrix} 1 & & & \\ & E_1 & & \\ & & \ddots & \\ & & & E_{T-1} \end{bmatrix},$$

where $\mathcal{A}_s^t := A_{t-1} A_{t-2} \cdots A_s$ for $s < t$, $\mathcal{A}_t^t := I^{(n_x)}$, and x_1 is the given value of the initial state $\boldsymbol{x}_1$. We call $\boldsymbol{x}$ the state process, $\boldsymbol{u}$ the control policy or control process, $\boldsymbol{\xi}$ the noise or disturbance process and $\boldsymbol{y}$ the measurement or observation process. We include the degenerate random variables $(\boldsymbol{\xi}_0, \boldsymbol{y}_0)$ as a conceptual device that will simplify significantly later mathematical developments in the chapter.

Remark 3.2.1. *We assume without loss of generality that the initial state $\boldsymbol{x}_1$ is known and given by x_1. If that is not the case, we can set $\boldsymbol{x}_1$ to an arbitrary value and identify $\boldsymbol{x}_2$ with the initial state. We can then assign to $\boldsymbol{x}_2$ a prescribed distribution by tying it to $\boldsymbol{\xi}_1$ through a suitable choice of the dynamic system matrices for $t = 1$.*

We consider only physically implementable *causal* control policies, which rely solely on information available by observing the measurement process $\boldsymbol{y}$. To this end, we let $\mathcal{N}$ be the linear space of all *causal* control policies

$$\mathcal{N} := \times_{t=1}^{T-1} \mathcal{L}^2(\Omega, \mathcal{F}_t^y, \mathbb{P}; \mathbb{R}^{n_u}) \subseteq \mathcal{L}_{N_u}^2,$$

where $\mathcal{F}_t^y := \sigma(\boldsymbol{y}_0, \ldots, \boldsymbol{y}_t)$ is the σ-algebra generated by the history of the observation process up to

time t. We are interested in stochastic optimal control problems of the following type:

$$\begin{array}{ll} \inf & \mathbb{E}\left[\boldsymbol{u}^\top J_u \boldsymbol{u} + \boldsymbol{x}^\top J_x \boldsymbol{x}\right] \\ \text{s.t.} & \boldsymbol{u} \in \mathcal{N},\ \boldsymbol{s} \in \mathcal{L}^2_{N_s} \\ & \left.\begin{array}{l} \boldsymbol{x} = B\boldsymbol{u} + C\boldsymbol{\xi} \\ F_u \boldsymbol{u} + F_x \boldsymbol{x} + F_s \boldsymbol{s} = h \\ \boldsymbol{s} \geq 0 \end{array}\right\} \ \mathbb{P}\text{-a.s.} \end{array} \tag{$\mathcal{P}$}$$

Our aim is to find a causal control policy $\boldsymbol{u} \in \mathcal{N}$ that minimizes the expectation of the quadratic cost $\boldsymbol{u}^\top J_u \boldsymbol{u} + \boldsymbol{x}^\top J_x \boldsymbol{x}$. The requirement that $\boldsymbol{u} \in \mathcal{N}$ is equivalent to that of each $\boldsymbol{u}_t$ being a function of the observation history $\boldsymbol{y}^t = (\boldsymbol{y}_0, \ldots, \boldsymbol{y}_t)$ available at that time. The control policy is selected subject to linear joint state and control constraints of the form $F_u\boldsymbol{u}+F_x\boldsymbol{x}+F_s\boldsymbol{s} = h$, where $h \in \mathbb{R}^{N_c}$, $F_u \in \mathbb{R}^{N_c \times N_u}$, $F_x \in \mathbb{R}^{N_c \times N_x}$, $F_s \in \mathbb{R}^{N_c \times N_s}$, where N_c is the number of state and control constraints. The random vector $\boldsymbol{s} \in \mathcal{L}^2_{N_s}$ can be interpreted as a vector of N_s slack variables, restricted to be non-negative. Note that any row of F_s that contains only zeros corresponds to an equality constraint involving only $\boldsymbol{u}$ and $\boldsymbol{x}$. All other constraints include slack variables and thus correspond to inequality constraints for $\boldsymbol{u}$ and $\boldsymbol{x}$. It is assumed without loss of generality that the matrix F_s has full column rank. Finally, we require the dynamic system equation $\boldsymbol{x} = B\boldsymbol{u} + C\boldsymbol{\xi}$ to hold. Recall that this constraint includes the requirement that the initial state $\boldsymbol{x}_1 = x_1$.

Some conditions need to be imposed for problem $\mathcal{P}$ to be well-defined. We assume that the matrices J_u and J_x are positive semidefinite to ensure convexity of the objective function. Additionally, we assume that the support of the noise process $\boldsymbol{\xi}$ is non-empty, compact and representable as

$$\Xi := \left\{\xi \in \mathbb{R}^{N_\xi} \colon W_i \xi \succeq_{\mathcal{K}_p} 0,\ i = 1 \ldots l,\ e_0^\top \xi = 1\right\} \tag{3.3}$$

for some matrices $W_i \in \mathbb{R}^{N_\xi \times N_\xi}$. The last constraint in (3.3) ensures that $\xi_0 = 1$ for all $\xi \in \Xi$, which is consistent with our previous assumption that $\boldsymbol{\xi}_0 = 1$ almost surely. We assume that the second-order moment matrix associated with the noise process $\boldsymbol{\xi}$ is known and defined as

$$M := \mathbb{E}\left(\boldsymbol{\xi}\boldsymbol{\xi}^\top\right). \tag{3.4}$$

We further assume that the support Ξ spans the state space $\mathbb{R}^{N_\xi}$ of the noise process. This is equivalent

to assuming that there is a $\xi \in \Xi$ that satisfies the strict inequalities $W_i\xi \succ_{\mathcal{K}_p} 0$ for all $i = 1 \dots l$. Note that this last assumption is non-restrictive and can often be enforced by reducing the dimension of the noise process. This assumption also ensures that M is invertible and consequently that $M \succ 0$.

The cases of greatest general interest are $p \in \{1, 2, \infty\}$. In particular, it is easy to show that any subset of the hyperplane $\{\xi \in \mathbb{R}^{N_\xi} : \xi_0 = 1\}$ that results from a finite intersection of arbitrary ellipsoids and half spaces is representable in the form (3.3) for $p = 2$, so that the class of disturbances we consider includes those with polyhedral support.

3.2.1 Controller Information Structure

The information available to the causal controllers $\boldsymbol{u} \in \mathcal{N}$ at each time instance t can be interpreted as the σ-algebra $\mathcal{F}_t^y$, *i.e.* the information available by the history of the observation process $\boldsymbol{y}$ up to time t. For any causal policy $\boldsymbol{u} \in \mathcal{N}$, $\boldsymbol{u}_t$ must be $\mathcal{F}_t^y$-measurable for each $t = 1, \dots, T-1$. Since the observation process $\boldsymbol{y}$ depends on the control policy $\boldsymbol{u}$, the filtration $\mathbb{F}^y := \{\mathcal{F}_t^y\}_{t=0}^{T-1}$ seems to depend on $\boldsymbol{u}$ as well. In order to show that $\mathbb{F}^y$ is in fact independent of $\boldsymbol{u}$, we consider the purified observation process

$$\boldsymbol{\eta} := (\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_{T-1}) \in \mathcal{L}_{N_y}^2$$

with $\boldsymbol{\eta} = (DC + E)\boldsymbol{\xi} = G\boldsymbol{\xi}$, where $G := (DC + E) \in \mathbb{R}^{N_y \times N_\xi}$. As described in [8], the purified observation $\boldsymbol{\eta}_t$ can be interpreted as the difference between the actual observation $\boldsymbol{y}_t$ and the observation that would have resulted at time t from a completely noise-free system obeying the same control policy. Note that $\boldsymbol{y}$ is linear in the control policy $\boldsymbol{u}$ and the purified observation process $\boldsymbol{\eta}$, *i.e.* $\boldsymbol{y} = DB\boldsymbol{u} + \boldsymbol{\eta}$.

Now let $\mathcal{F}_t := \sigma(\boldsymbol{\eta}_0, \dots, \boldsymbol{\eta}_t)$ be the σ-algebra generated by the history of the purified observation process up to time t. By construction, $\boldsymbol{\eta}$ and the filtration $\mathbb{F} := \{\mathcal{F}_t\}_{t=0}^{T-1}$ are independent of the control strategy $\boldsymbol{u}$.

Proposition 3.2.1. *The filtrations $\mathbb{F}^y$ and $\mathbb{F}$ are identical.*

Proof. The statement is proved by induction. For any causal policy $\boldsymbol{u} \in \mathcal{N}$, $\boldsymbol{u}_t$ must be $\mathcal{F}_t^y$-measurable for each $t = 1, \dots, T-1$. This is equivalent to asserting the existence of Borel-measurable functions $\varphi_t : \mathbb{R}^{1+n_y t} \to \mathbb{R}^{n_u}$ such that $\boldsymbol{u}_t = \varphi_t(\boldsymbol{y}_0, \dots, \boldsymbol{y}_t)$, see e.g. [5, Theorem 6.4.2 (c)]. It is clear that $\mathcal{F}_0^y = \mathcal{F}_0$ since $\boldsymbol{y}_0 = \boldsymbol{\eta}_0$, which is 1 almost surely. Assume now that $\mathcal{F}_s^y = \mathcal{F}_s$ for all $0 \leq s < t$ where

$0 < t \leq T-1$. Thus, there are Borel-measurable functions ψ_s and χ_s such that $\boldsymbol{y}_s = \psi_s(\boldsymbol{\eta}_0, \ldots, \boldsymbol{\eta}_s)$ and $\boldsymbol{\eta}_s = \chi_s(\boldsymbol{y}_0, \ldots, \boldsymbol{y}_s)$ for all $0 \leq s < t$. Moreover, by the causality of the control policy $\boldsymbol{u}$ we have

$$\begin{aligned} \boldsymbol{y}_t &= \sum\nolimits_{s=1}^{t-1} D_t A_{s+1}^t B_s \boldsymbol{u}_s + \sum\nolimits_{s=0}^{t-1} D_t A_{s+1}^t C_s \boldsymbol{\xi}_s + E_t \boldsymbol{\xi}_t \\ &= \sum\nolimits_{s=1}^{t-1} D_t A_{s+1}^t B_s \, \varphi_s(\psi_0(\boldsymbol{\eta}_0), \ldots, \psi_s(\boldsymbol{\eta}_0, \ldots, \boldsymbol{\eta}_s)) + \boldsymbol{\eta}_t \\ &=: \psi_t(\boldsymbol{\eta}_0, \ldots, \boldsymbol{\eta}_t), \end{aligned} \tag{3.5}$$

and from (3.5) we infer that

$$\begin{aligned} \boldsymbol{\eta}_t &= \boldsymbol{y}_t - \sum\nolimits_{s=1}^{t-1} D_t A_{s+1}^t B_s \, \varphi_s(\boldsymbol{y}_0 \ldots, \boldsymbol{y}_s) \\ &=: \chi_t(\boldsymbol{y}_0, \ldots, \boldsymbol{y}_t). \end{aligned} \tag{3.6}$$

The relation (3.5) implies that

$$\mathcal{F}_t^y = \sigma(\psi_0(\boldsymbol{\eta}_0), \ldots, \psi_t(\boldsymbol{\eta}_0, \ldots, \boldsymbol{\eta}_t)) \subseteq \sigma(\boldsymbol{\eta}_0, \ldots, \boldsymbol{\eta}_t) = \mathcal{F}_t, \tag{3.7a}$$

and the relation (3.6) implies that

$$\mathcal{F}_t = \sigma(\chi_0(\boldsymbol{y}_0), \ldots, \chi_t(\boldsymbol{y}_0, \ldots, \boldsymbol{y}_t)) \subseteq \sigma(\boldsymbol{y}_0, \ldots, \boldsymbol{y}_t) = \mathcal{F}_t^y. \tag{3.7b}$$

The inclusions (3.7a) and (3.7b) yield $\mathcal{F}_t^y = \mathcal{F}_t$, and thus the claim follows. □

Proposition 3.2.1 shows that the information structure generated by the observation process is not decision-dependent. Furthermore, it shows that the observation process $\boldsymbol{y}$ and the purified observation process $\boldsymbol{\eta}$ convey the same amount of information. This in turn implies that the set of all control policies that are adapted to $\boldsymbol{y}$ is equivalent to the set of all policies that are adapted to $\boldsymbol{\eta}$. As a result, we can re-define the set of all implementable causal control policies $\mathcal{N}$ in terms of the σ-algebras $\{\mathcal{F}_t\}_{t=0}^{T-1}$.

In the remainder of the chapter we will provide tractable methods for calculating upper and lower approximations to the problem $\mathcal{P}$ using affine decision rules. Although no further assumptions are required for the upper bound to be presented in Section 3.3, some further assumption relating to the purified observation process $\boldsymbol{\eta}$ will be required for the lower bounds presented in Section 3.4:

A1 (Noise Process) *The expectation of the noise process* $\boldsymbol{\xi}$ *conditioned on the purified observation process* $\boldsymbol{\eta}$ *is linear in* $\boldsymbol{\eta}$, i.e. *there exists a matrix* $L \in \mathbb{R}^{N_\xi \times N_y}$ *so that*

$$\mathbb{E}\left[\boldsymbol{\xi} \,|\boldsymbol{\eta}\right] = L\,\boldsymbol{\eta} \;\mathbb{P}\text{-}a.s. \tag{3.8}$$

Furthermore, the conditional expectation of $\boldsymbol{\eta}$ *given its history* $\boldsymbol{\eta}^t := (\boldsymbol{\eta}_0, \ldots, \boldsymbol{\eta}_t)$ *up to time* t, *is linear in* $\boldsymbol{\eta}^t$, i.e. *there exist matrices* $H_t \in \mathbb{R}^{N_y \times (1+tn_y)}$ *so that*

$$\mathbb{E}\left[\boldsymbol{\eta}|\,\boldsymbol{\eta}^t\right] = H_t \boldsymbol{\eta}^t \;\mathbb{P}\text{-}a.s. \text{ for all } t = 1, \ldots, T-1. \tag{3.9}$$

The assumptions of **A1** appear to be restrictive. However, they are satisfiable in two special cases of considerable practical and theoretical interest:

Proposition 3.2.2 (Elliptically Contoured Distributions)**.** *The conditions of* ***A1*** *are satisfied if the distribution of the noise process* $\boldsymbol{\xi}$ *follows an elliptically contoured distribution.*

Proof. To demonstrate that (3.8) holds, we can use [24, Theorem 1] to show that the joint random vector $(\boldsymbol{\xi}, \boldsymbol{\eta})$, as the image of $\boldsymbol{\xi}$ under the linear transformation $\left(I^{(N_\xi)}, G^\top\right)^\top$, also follows an elliptically contoured distribution. This result allows us to invoke [24, Cor. 5] which guarantees that the conditional expectation of $\boldsymbol{\xi}$ given $\boldsymbol{\eta}$ is almost surely linear in $\boldsymbol{\eta}$. A similar argument can be used to establish (3.9). □

The preceding result confirms that **A1** is very broadly applicable, since the class of elliptically contoured distributions covers a wide range of distributions including multivariate normal, logistic, symmetric stable, generalized-hyperbolic and t-distributions as well as their truncated versions [30].

In the case of perfect state information, a weaker set of conditions are required to ensure satisfaction of **A1**:

Proposition 3.2.3 (Perfect State Information)**.** *The conditions of* ***A1*** *are satisfied if the following conditions all hold:*

i) The system state can be measured perfectly, i.e. $\boldsymbol{y}_t = \boldsymbol{x}_t$ *for all* $t = 1, \ldots, T-1$.

ii) There exist matrices $N_t \in \mathbb{R}^{N_\xi \times (1+tn_\xi)}$ *such that*

$$\mathbb{E}\left[\boldsymbol{\xi}\middle|\,\boldsymbol{\xi}^t\right] = N_t \boldsymbol{\xi}^t \ \mathbb{P}\text{-a.s. for all } \ t = 1, \ldots, T. \tag{3.10}$$

iii) The matrices C_t *have full column rank for all* $t = 1, \ldots, T-1$.

Proof. In the perfect information case under consideration, the system equations (3.1) simplify because $n_y = n_x$ while $D_t = I$ and $E_t = 0$ for all $t = 1, \ldots, T-1$. The matrix G then simplifies to

$$G = \begin{bmatrix} 1 & & & & & \\ x_1 & & & & & \\ \mathcal{A}_1^2 x_1 & C_1 & & & & \\ \mathcal{A}_1^3 x_1 & \mathcal{A}_2^3 C_1 & C_2 & & & \\ \vdots & & & \ddots & & \\ \mathcal{A}_1^T x_1 & \mathcal{A}_2^T C_1 & \mathcal{A}_3^T C_2 & \cdots & C_{T-2} & 0 \end{bmatrix},$$

where the last n_ξ columns contain only zeros. We can now define $\hat{G}$ as the sub-matrix of G which is obtained by removing the last n_ξ columns of G. By construction, we have $\boldsymbol{\eta} = G\boldsymbol{\xi} = \hat{G}\boldsymbol{\xi}^{T-2}$ since $\boldsymbol{\eta}$ does not depend on $\boldsymbol{\xi}_{T-1}$ when $E_{T-1} = 0$. Note that $\hat{G}$ has full column rank, and thus the purified observation process $\boldsymbol{\eta}$ and the truncated noise process $\boldsymbol{\xi}^{T-2}$ generate the same σ-algebra. This implies that

$$\begin{aligned} \mathbb{E}\left[\boldsymbol{\xi}|\boldsymbol{\eta}\right] = \ & \mathbb{E}[\boldsymbol{\xi}|\boldsymbol{\xi}^{T-2}] && = N_{T-2}\,\boldsymbol{\xi}^{T-2} \\ = \ & N_{T-2}\,\hat{G}^\dagger\,\hat{G}\,\boldsymbol{\xi}^{T-2} && = N_{T-2}\,\hat{G}^\dagger\,\boldsymbol{\eta}. \end{aligned}$$

A similar argument can be used to establish (3.9). □

Remark 3.2.2. *The conditions required in Proposition 3.2.3 are very mild. Condition ii) is trivially satisfied if, for example, the noise vectors* $\boldsymbol{\xi}_t$ *are serially independent. Condition iii) is a common assumption for control problems of the type* $\mathcal{P}$ *with perfect state measurements, and can usually be enforced by reducing the dimension of the noise process.*

3.3 Upper Approximations of $\mathcal{P}$ and Primal Affine Control Policies

The optimization problem $\mathcal{P}$ generally involves an infinite number of decision variables and constraints. Without a suitable approximation, this problem is not amenable to numerical solution. In this section we impose a restriction on the feasible set to achieve computational tractability. Any reduction of the feasible set corresponds to a conservative approximation of the original problem and thus results in an upper bound on the true optimal value of $\mathcal{P}$. Here, we calculate such an upper bound by restricting the class of causal control policies $\mathcal{N}$ to the subclass of all control policies that are affine with respect to the observations. The affine observation-feedback policies can be written as

$$\boldsymbol{u}_t = u_{t,0} + \sum_{s=1}^{t} U_{t,s}\boldsymbol{y}_s \quad \text{for} \quad t = 1, \ldots, T-1,$$

for some matrices $U_{t,s} \in \mathbb{R}^{n_u \times n_y}$. In compact notation, we have $\boldsymbol{u} = U\boldsymbol{y}$ almost surely (since $\boldsymbol{y}_0 = 1$ almost surely), where

$$U := \begin{bmatrix} u_{1,0} & U_{1,1} & & & \\ u_{2,0} & U_{2,1} & U_{2,2} & & \\ \vdots & & & \ddots & \\ u_{T-1,0} & U_{T-1,1} & U_{T-1,2} & \cdots & U_{T-1,T-1} \end{bmatrix}. \tag{3.11}$$

We denote by $\mathcal{U}$ the linear space of all block lower triangular matrices of the form (3.11).

Proposition 3.3.1. *An upper bound to the problem $\mathcal{P}$ can be found by solving the following optimization problem, where the control policy $\boldsymbol{u}$ is parameterized as an affine function of the outputs $\boldsymbol{y}$:*

$$\begin{array}{ll} \inf & \mathbb{E}\left[\boldsymbol{u}^\top J_u \boldsymbol{u} + \boldsymbol{x}^\top J_x \boldsymbol{x}\right] \\ s.t. & U \in \mathcal{U},\ \boldsymbol{s} \in \mathcal{L}^2_{N_s} \\ & \left.\begin{array}{l} \boldsymbol{x} = B\boldsymbol{u} + C\boldsymbol{\xi},\ \boldsymbol{u} = U\boldsymbol{y} \\ F_u\boldsymbol{u} + F_x\boldsymbol{x} + F_s\boldsymbol{s} = h \\ \boldsymbol{s} \geq 0 \end{array}\right\} \ \mathbb{P}\text{-}a.s. \end{array}$$

The above problem is itself equivalent to the following optimization problem, where the control policy $\boldsymbol{u}$

is parameterized as an affine function of the purified outputs $\boldsymbol{\eta} = G\boldsymbol{\xi}$*:*

$$\begin{array}{ll} \inf & \mathbb{E}\left[\boldsymbol{u}^\top J_u \boldsymbol{u} + \boldsymbol{x}^\top J_x \boldsymbol{x}\right] \\ s.t. & Q \in \mathcal{U},\ \boldsymbol{s} \in \mathcal{L}^2_{N_s} \\ & \left.\begin{array}{l} \boldsymbol{x} = B\boldsymbol{u} + C\boldsymbol{\xi},\ \boldsymbol{u} = QG\boldsymbol{\xi} \\ F_u \boldsymbol{u} + F_x \boldsymbol{x} + F_s \boldsymbol{s} = h \\ \boldsymbol{s} \geq 0 \end{array}\right\} \ \mathbb{P}\text{-}a.s. \end{array} \qquad (\mathcal{P}_u)$$

Note that the two equivalent problem formulations in Proposition 3.3.1 differ only in their representation of the control policy $\boldsymbol{u}$. In the first case, the control input is parameterized as a causal affine function of prior measurements $\boldsymbol{y}$, while in the second case the input is parameterized as a causal affine function of the purified outputs $\boldsymbol{\eta} = G\boldsymbol{\xi}$. In both cases, the set of feasible control policies is a subset of the set of all possible control policies $\mathcal{N}$.

Before proceeding to the proof of Proposition 3.3.1, we note that problem $\mathcal{P}_u$ remains seemingly intractable due to the presence of functional decision variables and constraints. We therefore require an additional result that provides a method for solving the problem $\mathcal{P}_u$ as a tractable convex optimization problem.

Proposition 3.3.2. *A solution to problem* $\mathcal{P}_u$ *can be found by solving the following equivalent convex optimization problem:*

$$\begin{array}{ll} \inf & \operatorname{tr}\left(G^\top Q^\top (J_u + B^\top J_x B) QGM + 2C^\top J_x BQGM + C^\top J_x CM\right) \\ s.t. & Q \in \mathcal{U},\ S \in \mathbb{R}^{N_s \times N_\xi},\ \Lambda_i \in \mathbb{R}^{N_\xi \times N_s},\ \mu \in \mathbb{R}^{N_s} \\ & (F_u + F_x B)QG + F_x C + F_s S - he_0^\top = 0 \\ & S = \mu e_0^\top + \sum_{i=1}^{l} \Lambda_i^\top W_i \\ & \Lambda_i \succeq_{\mathcal{K}_p^*} 0,\ i = 1 \ldots l,\ \mu \geq 0, \end{array} \qquad (\tilde{\mathcal{P}}_u)$$

The problem $\tilde{\mathcal{P}}_u$ *is a tractable conic optimization problem and can be solved in polynomial time for any* $p \in [1, +\infty]$.

The problem $\tilde{\mathcal{P}}_u$ is of interest because it provides a tractable (though suboptimal) solution to $\mathcal{P}$ by restricting the class of control policies over which the optimization is performed. However, there are at present no sophisticated methods available for estimating the degree of suboptimality introduced

by imposing such a restriction. In Section 3.4, we develop a method for calculating a *lower* bound on $\mathcal{P}$ by formulating a problem dual to $\mathcal{P}$ and restricting the dual variables to affine form to ensure tractability. This will enable us to bound the suboptimality in $\tilde{\mathcal{P}}_u$ in Section 3.5.

Remark 3.3.1. *The second part of Proposition 3.3.1 provides a method for optimizing over affine output feedback policies $\boldsymbol{u} = U\boldsymbol{y}$ by re-parameterizing the problem into one of optimizing over a different but equivalent class of parametric policies, for which $\boldsymbol{u}$ and $\boldsymbol{x}$ become affine functions of the parameters [9, 46, 84]. The resulting optimization problem can then be solved using Proposition 3.3.2. The idea underpinning this re-parameterization follows a similar approach to the classical Youla- or Q-parameterization procedure [91]. The benefit of this approach is that it avoids direct optimization over the parameter $U \in \mathcal{U}$, which is impractical since both $\boldsymbol{u}$ and $\boldsymbol{x}$ become nonlinear (rational) functions of the coefficients of U. In order to facilitate comparison with the lower bounding methods of Section 3.4, we sketch the proofs using our notation.*

Corollary 3.3.1. *A trivial upper bound to $\mathcal{P}$ can be found by considering only stationary (*i.e. *open-loop) policies:*

$$\begin{array}{ll} \inf & u^\top (J_u + B^\top J_x B)u + 2u^\top B^\top J_x C\mathbb{E}\left[\boldsymbol{\xi}\right] + \operatorname{tr}\left(C^\top J_x C M\right) \\ s.t. & u \in \mathbb{R}^{N_u},\, S \in \mathbb{R}^{N_s \times N_\xi},\, \Lambda_i \in \mathbb{R}^{N_\xi \times N_s},\, \mu \in \mathbb{R}^{N_s} \\ & (F_u + F_x B)ue_0^\top + F_x C + F_s S - he_0^\top = 0 \\ & S = \mu e_0^\top + \sum_{i=1}^{l} \Lambda_i^\top W_i \\ & \Lambda_i \succeq_{\mathcal{K}_p^*} 0,\ i = 1 \dots l \\ & \mu \geq 0. \end{array} \qquad (\mathcal{P}_{uu})$$

where $u \in \mathbb{R}^{N_u}$ is the vector of the stationary controls. It is easy to verify that

$$\mathcal{P} \leq \mathcal{P}_u \leq \mathcal{P}_{uu}.$$

Problem $\mathcal{P}_{uu}$ is a control optimization problem over open-loop robust controllers. Note that $\mathcal{P}_{uu}$ is a restriction of $\mathcal{P}_u$. One can derive $\mathcal{P}_{uu}$ from $\mathcal{P}_u$ by restricting the control policies to be stationary instead of affine on $\boldsymbol{y}$. As a result $\mathcal{P}_u$ is a tighter upper bound to $\mathcal{P}$ than $\mathcal{P}_{uu}$.

In the remainder of this section we supply proofs for Propositions 3.3.1 and 3.3.2.

Proof of Proposition 3.3.1

Proof of the first part of Proposition 3.3.1 is straightforward, since any control law in the affine form $\boldsymbol{u} = U\boldsymbol{y}$ is causal due to the lower triangular structure imposed by the constraint $U \in \mathcal{U}$, hence $\boldsymbol{u} \in \mathcal{N}$.

To prove the second part, we show that any control law in the affine output feedback form $\boldsymbol{u} = U\boldsymbol{y}$ for some $U \in \mathcal{U}$ can be matched exactly by a control law in the purified output feedback form $\boldsymbol{u} = Q\boldsymbol{\eta} = QG\boldsymbol{\xi}$ for some $Q \in \mathcal{U}$, and vice-versa. To prove the first case, assume that $\boldsymbol{u} = U\boldsymbol{y}$ and solve for the control input $\boldsymbol{u}$ in terms of the purified outputs $\boldsymbol{\eta}$, yielding

$$\boldsymbol{u} = (I - UDB)^{-1}U\boldsymbol{\eta}.$$

We then set $Q = (I - UDB)^{-1}U$, noting that the required matrix inverse exists since UDB is strictly lower triangular, which also ensures that $Q \in \mathcal{U}$. Proof of the second case is provided by a similar argument. □

Proof of Proposition 3.3.2

Both $\boldsymbol{u}$ and $\boldsymbol{x}$ are linear in $\boldsymbol{\xi}$ in the problem $\mathcal{P}_u$. By pre-multiplying the joint state and control constraint of $\mathcal{P}_u$ with the left inverse of F_s (which exists since F_s is assumed to have full column rank), it can be seen that the vector of slack variables $\boldsymbol{s}$ appearing in $\mathcal{P}_u$ is also linear in $\boldsymbol{\xi}$, so that $\boldsymbol{s} = S\boldsymbol{\xi}$ for some matrix $S \in \mathbb{R}^{N_s \times N_\xi}$. Noting also that $h = he_0^\top \boldsymbol{\xi}$ almost surely, we can rewrite the problem $\mathcal{P}_u$ as

$$\begin{array}{lll} \inf & \mathbb{E}\left[\boldsymbol{u}^\top J_u \boldsymbol{u} + \boldsymbol{x}^\top J_x \boldsymbol{x}\right] & \\ \text{s.t.} & Q \in \mathcal{U},\, S \in \mathbb{R}^{N_s \times N_\xi} & \\ & \left.\begin{array}{l} \boldsymbol{x} = B\boldsymbol{u} + C\boldsymbol{\xi},\ \boldsymbol{u} = QG\boldsymbol{\xi} \\ \left((F_u + F_x B)\,QG + F_x C + F_s S - he_0^\top\right)\boldsymbol{\xi} = 0 \\ S\boldsymbol{\xi} \geq 0 \end{array}\right\} & \mathbb{P}\text{-a.s.} \end{array} \tag{3.12}$$

Observe that due to the continuity of the constraint functions in $\boldsymbol{\xi}$, the almost sure constraints in (3.12) must hold for all $\xi \in \Xi$. We can therefore replace the almost sure constraints with semi-infinite

constraints to obtain the equivalent semi-infinite problem

$$\begin{array}{ll} \inf & \operatorname{tr}\left(G^\top Q^\top (J_u + B^\top J_x B) QGM + 2C^\top J_x BQGM + C^\top J_x CM\right) \\ \text{s.t.} & Q \in \mathcal{U},\, S \in \mathbb{R}^{N_s \times N_\xi} \\ & \left.\begin{array}{l} (F_u + F_x B)\, QG + F_x C + F_s S - h e_0^\top)\xi = 0 \\ S\xi \geq 0 \end{array}\right\} \quad \forall \xi \in \Xi, \end{array} \tag{3.13}$$

where we have also eliminated $\boldsymbol{u}$ and $\boldsymbol{x}$ and rewritten the objective function in terms of the second-order moment matrix $M = \mathbb{E}\left[\boldsymbol{\xi}\boldsymbol{\xi}^\top\right]$.

The equality constraints in (3.13) must hold for all $\xi \in \Xi$, meaning that the linear hull of Ξ belongs to the null space of the linear operator $(F_u + F_x B)QG + F_s S + F_x C - h e_0^\top$. Since Ξ is assumed to span the entire space $\mathbb{R}^{N_\xi}$, these semi-infinite equality constraints are equivalent to the matrix equality $(F_u + F_x B)QG + F_s S + F_x C - h e_0^\top = 0$. Next, we simplify the semi-infinite inequality constraints by using techniques that are commonly used in robust optimization [10, 12, 13]. The following lemma, which is a special case of [13, Theorem 3.1], describes the key enabling mechanism:

Lemma 3.3.1. *For any $z \in \mathbb{R}^{N_\xi}$, the following statements are equivalent:*

i) $z^\top \xi \geq 0$ *for all* $\xi \in \Xi$

ii) There exist vectors $\lambda_i \in \mathbb{R}^{N_\xi}$ *for* $i \in \{1 \ldots l\}$ *and a scalar* $\mu \geq 0$ *such that*

$$\begin{aligned} z^\top &= \mu e_0^\top + \sum_{i=1}^{l} \lambda_i^\top W_i \\ \lambda_i &\succeq_{\mathcal{K}_p^*} 0\,,\ i = 1 \ldots l. \end{aligned}$$

Proof. The proposition follows from a strong duality argument. Statement (i) is equivalent to

$$\begin{array}{rll} 0 \leq & \min & z^\top \xi \\ & \text{s.t.} & \xi \in \mathbb{R}^{N_\xi} \\ & & W_i\, \xi \succeq_{\mathcal{K}_p} 0,\, i = 1 \ldots l \\ & & e_0^\top \xi = 1. \end{array} \tag{3.14}$$

Due to our assumptions on the support Ξ, the feasible region of the minimization problem (3.14) has

nonempty relative interior. By strong conic duality, (3.14) therefore is equivalent to

$$\begin{aligned} 0 \leq \quad & \max \quad \mu \\ & \text{s.t.} \quad \mu \in \mathbb{R},\ \lambda_i \in \mathbb{R}^N_\xi \\ & \qquad z^\top = \textstyle\sum_{i=1}^l \lambda_i^\top W_i + \mu e_0^\top \\ & \qquad \lambda_i \succeq_{\mathcal{K}_p^*} 0,\ \ i = 1 \ldots l, \end{aligned}$$

which is manifestly equivalent to statement (ii). Thus the claim follows. □

Let $z_\kappa^\top$ denote the κ-th row of the matrix S in problem (3.13). We can then use Lemma 3.3.1 to replace the semi-infinite constraint $z_\kappa^\top \xi \geq 0 \, \forall \xi \in \Xi$ by finitely many linear constraints involving the new decision variables $\mu_\kappa \in \mathbb{R}$ and $\lambda_{\kappa,i} \in \mathbb{R}^{N_\xi}$ for $i = 1, \ldots, l$. The constraint $S\xi \geq 0 \, \forall \xi \in \Xi$ can thus be reformulated in terms of $\mu = (\mu_1, \ldots, \mu_{N_s})^\top \in \mathbb{R}^{N_s}$ and $\Lambda_i = (\lambda_{1,i} \ \cdots \ \lambda_{N_s,i}) \in \mathbb{R}^{N_\xi \times N_s}$, $i = 1, \ldots, l$ to yield the optimization problem $\tilde{\mathcal{P}}_u$. Solvability of $\tilde{\mathcal{P}}_u$ in polynomial time for any $p \in [1, +\infty]$ is ensured by [90]. □

3.4 Lower Approximations of $\mathcal{P}$ and Dual Affine Control Policies

The use of primal affine control policies leads to a conservative approximation for the original problem $\mathcal{P}$, whose computational complexity scales gracefully with the horizon length and the dimension of the system (3.1). It also yields an implementable control policy which is feasible, but typically suboptimal, in $\mathcal{P}$. Our main goal is to estimate the loss of optimality due to the use of affine policies. To this end, we now establish a lower bound on $\mathcal{P}$ by extending the dual linear decision rule techniques that were recently introduced in the context of stochastic programming; see [60]. The loss of optimality incurred by using affine controllers can then be bounded by the difference between the upper and lower bounds on the optimization problem $\mathcal{P}$. In order to derive the lower bound, we first reformulate $\mathcal{P}$ as a min-max problem by introducing a dual control policy $\boldsymbol{\nu} \in \mathcal{L}^2_{N_c}$ and by moving the joint state and

control constraints to the objective function, to obtain

$$\begin{array}{lll} \inf & \sup\limits_{\boldsymbol{\nu}\in\mathcal{L}^2_{N_c}} & \mathbb{E}\left[\boldsymbol{u}^\top J_u \boldsymbol{u} + \boldsymbol{x}^\top J_x \boldsymbol{x} \right. \\ & & \qquad \left. + \boldsymbol{\nu}^\top \left[F_u \boldsymbol{u} + F_x \boldsymbol{x} + F_s \boldsymbol{s} - h\right]\right] \\ \text{s.t.} & \boldsymbol{u} \in \mathcal{N},\, \boldsymbol{s} \in \mathcal{L}^2_{N_s} & \\ & \left.\begin{array}{l} \boldsymbol{x} = B\boldsymbol{u} + C\boldsymbol{\xi} \\ \boldsymbol{s} \geq 0 \end{array}\right\} & \mathbb{P}\text{-a.s.} \end{array} \qquad (\mathcal{P}')$$

The inner maximization over the dual control policies in $\mathcal{P}'$ ensures that any violation of the state and control constraints on a set of strictly positive probability incurs an infinite penalty. The two problems $\mathcal{P}$ and $\mathcal{P}'$ are equivalent. In order to derive a tractable lower bounding approximation for $\mathcal{P}$, we will adopt the same approach as in the primal problem in Section 3.3 and restrict the space of all dual control policies to those that depend linearly on the noise process, and are thus representable as $\boldsymbol{\nu} = Y\boldsymbol{\xi}$ for some $Y \in \mathbb{R}^{N_c \times N_\xi}$. We will then show that the resulting approximation problem can be solved by solving an equivalent tractable conic optimization problem.

The first of these results forms a natural lower-bounding counterpart to Proposition 3.3.1:

Proposition 3.4.1. *A lower bound to the problem $\mathcal{P}$ (equivalently $\mathcal{P}'$) can be found by solving the following optimization problem, where the dual policy variable $\boldsymbol{\nu}$ is parameterized as an affine function of the disturbance $\boldsymbol{\xi}$:*

$$\begin{array}{lll} \inf & \sup\limits_{Y\in\mathbb{R}^{N_c\times N_\xi}} & \mathbb{E}\left[\boldsymbol{u}^\top J_u \boldsymbol{u} + \boldsymbol{x}^\top J_x \boldsymbol{x} \right. \\ & & \qquad \left. + \boldsymbol{\nu}^\top \left[F_u \boldsymbol{u} + F_x \boldsymbol{x} + F_s \boldsymbol{s} - h\right]\right] \\ \textit{s.t.} & \boldsymbol{u} \in \mathcal{N},\, \boldsymbol{s} \in \mathcal{L}^2_{N_s} & \\ & \left.\begin{array}{l} \boldsymbol{x} = B\boldsymbol{u} + C\boldsymbol{\xi} \\ \boldsymbol{\nu} = Y\boldsymbol{\xi},\ \boldsymbol{s} \geq 0 \end{array}\right\} & \mathbb{P}\text{-}a.s. \end{array}$$

The above min-max problem is itself equivalent to the following minimization problem, whose solution

provides a lower bound to the problem $\mathcal{P}$:

$$\left.\begin{array}{ll} \inf & \mathbb{E}\left[\boldsymbol{u}^\top J_u \boldsymbol{u} + \boldsymbol{x}^\top J_x \boldsymbol{x}\right] \\ s.t. & \boldsymbol{u} \in \mathcal{N},\ \boldsymbol{s} \in \mathcal{L}^2_{N_s} \\ & \boldsymbol{x} = B\boldsymbol{u} + C\boldsymbol{\xi} \\ & \mathbb{E}\left[(F_u \boldsymbol{u} + F_x \boldsymbol{x} + F_s \boldsymbol{s} - h)\,\boldsymbol{\xi}^\top\right] = 0 \\ & \boldsymbol{s} \geq 0 \end{array}\right\} \quad \mathbb{P}\text{-}a.s. \tag{$\mathcal{P}_\ell$}$$

Before proceeding to the proof of Proposition 3.4.1, we note that although $\mathcal{P}_\ell$ has only finitely many state and control constraints, it remains seemingly intractable since it involves functional decision variables and functional inequality constraints. We therefore require a method for solving the problem $\mathcal{P}_\ell$ as a tractable convex optimization problem. The following result provides a solution to this problem and forms a natural lower-bounding counterpart to Proposition 3.3.2:

Proposition 3.4.2. *If **A1** holds, then a lower-bounding solution to problem $\mathcal{P}_\ell$ can be found by solving the following convex optimization problem:*

$$\begin{array}{ll} \inf & \operatorname{tr}\left(G^\top Q^\top (J_u + B^\top J_x B) Q G M + 2C^\top J_x B Q G M + C^\top J_x C M\right) \\ s.t. & Q \in \mathcal{U},\ S \in \mathbb{R}^{N_s \times N_\xi} \\ & (F_u + F_x B) Q G + F_s S + F_x C - h e_0^\top = 0 \\ & W_i M S^\top \succeq_{\mathcal{K}_p} 0,\ i = 1 \ldots l \\ & e_0^\top M S^\top \geq 0. \end{array} \tag{$\tilde{\mathcal{P}}_\ell$}$$

The problem $\tilde{\mathcal{P}}_\ell$ is a tractable conic optimization problem and can be solved in polynomial time for any $p \in [1, +\infty]$.

Remark 3.4.1. *The problem $\mathcal{P}_\ell$ can be shown to be equivalent to the finite-dimensional problem $\tilde{\mathcal{P}}_\ell$ whenever $\mathcal{P}_\ell$ contains a strictly feasible point. Proof of this claim is based on the observation that the constraints in $\tilde{\mathcal{P}}_\ell$ represent the closure of a set of constraints in a problem equivalent to $\mathcal{P}_\ell$ (and which is constructed in the proof). We omit details of this argument for the sake of brevity.*

The problem $\tilde{\mathcal{P}}_\ell$ provides a tractable (though suboptimal) method for finding a lower bound to $\mathcal{P}$, and therefore allows us to bound the degree of approximation in the affine policy optimization problem $\tilde{\mathcal{P}}_u$.

Furthermore, there is an attractive structural similarity to the constraints in $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ – we elaborate on these points in Section 3.5.

Corollary 3.4.1. *A trivial lower bound to $\mathcal{P}$ can be found by considering a certainty equivalence problem:*

$$\begin{aligned} \inf \quad & \overline{u}^\top J_u \overline{u} + \overline{x}^\top J_x \overline{x} \\ s.t. \quad & \overline{u} \in \mathbb{R}^{N_u},\, \overline{s} \in \mathbb{R}^{N_s} \\ & \overline{x} = B\overline{u} + C\mathbb{E}\left[\boldsymbol{\xi}\right] \\ & F_u \overline{u} + F_x \overline{x} + F_s \overline{s} - h = 0 \\ & \overline{s} \geq 0 \end{aligned} \qquad . \tag{$\mathcal{P}_{\ell\ell}$}$$

It is easy to verify that $\mathcal{P} \geq \mathcal{P}_\ell \geq \mathcal{P}_{\ell\ell}$.

Problem $\mathcal{P}_{\ell\ell}$ is a certainty equivalence problem where only the mean of the uncertainty is considered. Note that the feasible region of $\mathcal{P}_{\ell\ell}$ is a subset of the feasible region of $\mathcal{P}_\ell$. One can retrieve the feasible region of $\mathcal{P}_{\ell\ell}$ by restricting the dual policy variable $\boldsymbol{\nu} \in \mathcal{L}^2_{N_c}$ in $\mathcal{P}'$ to be stationary instead of affine on $\boldsymbol{\xi}$. Furthermore, Jensen's Inequality ensures that the objective function in $\mathcal{P}_{\ell\ell}$ is a lower bound to the objective function in $\mathcal{P}_\ell$. As a result $\mathcal{P}_\ell$ is a tighter lower bound to $\mathcal{P}$ than $\mathcal{P}_{\ell\ell}$.

In the remainder of this section we supply proofs for Propositions 3.4.1 and 3.4.2.

Proof of Proposition 3.4.1

Proof of the first part of Proposition 3.4.1 is straightforward, since any dual variable in the affine form $\boldsymbol{\nu} = Y\boldsymbol{\xi}$ also satisfies $\boldsymbol{\nu} \in \mathcal{L}^2_{N_c}$. Hence the inner maximization over Y is a lower bound on the inner maximization that appears in the problem $\mathcal{P}'$.

Proof of the second part follows from the first since

$$\begin{aligned} & \sup_Y \mathbb{E}\left[\boldsymbol{\xi}^\top Y^\top \left[F_u \boldsymbol{u} + F_x \boldsymbol{x} + F_s \boldsymbol{s} - h\right]\right] \\ =& \sup_Y \mathbb{E}\left[\mathrm{tr}\left\{\left[F_u \boldsymbol{u} + F_x \boldsymbol{x} + F_s \boldsymbol{s} - h\right]\boldsymbol{\xi}^\top Y^\top\right\}\right] \\ =& \begin{cases} 0 & \text{if } \left[F_u \boldsymbol{u} + F_x \boldsymbol{x} + F_s \boldsymbol{s} - h\right]\boldsymbol{\xi}^\top = 0 \ \mathbb{P}\text{-a.s.} \\ \infty & \text{otherwise.} \end{cases} \end{aligned}$$

Problem $\mathcal{P}_\ell$ is clearly a relaxation of $\mathcal{P}$ since any $(\boldsymbol{u}, \boldsymbol{s})$ satisfying the almost sure state and control constraints in $\mathcal{P}$ will therefore also satisfy the expectation constraint in $\mathcal{P}_\ell$. □

Proof of Proposition 3.4.2

We first eliminate the state process $\boldsymbol{x}$ in $\mathcal{P}_\ell$ by substituting the dynamic system constraint into both the objective function and the expectation constraint in $\mathcal{P}_\ell$. Recalling the definition of the second-order moment matrix M, the approximate problem $\mathcal{P}_\ell$ can then be rewritten as

$$\begin{array}{ll} \inf & \mathbb{E}\left[\boldsymbol{u}^\top (J_u + B^\top J_x B)\boldsymbol{u}\right] + 2\,\mathrm{tr}\left(C^\top J_x B \mathbb{E}\left[\boldsymbol{u}\boldsymbol{\xi}^\top\right]\right) + \mathrm{tr}\left(C^\top J_x C M\right) \\ \text{s.t} & \boldsymbol{u} \in \mathcal{N},\ \boldsymbol{s} \in \mathcal{L}^2_{N_s} \\ & (F_u + F_x B)\mathbb{E}\left[\boldsymbol{u}\boldsymbol{\xi}^\top\right] + F_s \mathbb{E}\left[\boldsymbol{s}\boldsymbol{\xi}^\top\right] + (F_x C - h e_0^\top) M = 0 \\ & \boldsymbol{s} \geq 0 \quad \mathbb{P}\text{-a.s.} \end{array} \tag{3.15}$$

The optimization problem (3.15) contains a variety of terms composed of the functional decision variables $(\boldsymbol{u}, \boldsymbol{s})$. In the remainder of the proof, we introduce a number of technical lemmas that will enable us to eliminate each of these terms in turn, replacing them with finite-dimensional variables.

The first result will allow us to eliminate terms in the form $\mathbb{E}[\boldsymbol{u}\boldsymbol{\xi}^\top]$ and $\mathbb{E}[\boldsymbol{s}\,\boldsymbol{\xi}^\top]$:

Lemma 3.4.1. *For every $\boldsymbol{s} \in \mathcal{L}^2_{N_s}$ there exists a matrix $S \in \mathbb{R}^{N_s \times N_\xi}$ that satisfies*

$$SM = \mathbb{E}[\boldsymbol{s}\,\boldsymbol{\xi}^\top]. \tag{3.16}$$

*Likewise, if **A1** holds, then for every $\boldsymbol{u} \in \mathcal{N}$ there exists a block lower triangular matrix $Q \in \mathcal{U}$ that satisfies*

$$QGM = \mathbb{E}[\boldsymbol{u}\boldsymbol{\xi}^\top]. \tag{3.17}$$

Proof. Proof of the first statement is immediate since M is invertible and we impose no special structure (such as block lower triangularity) on the matrix S.

Proof of the second statement is subdivided into three parts. Before beginning the proof, we introduce

a sequence of truncation operators

$$P_t : \mathbb{R}^{N_y} \to \mathbb{R}^{1+tn_y},\ \eta \mapsto \eta^t = (\eta_0, \ldots, \eta_t),\ t = 1, \ldots, T-1. \tag{3.18}$$

It is easy to verify that $P_t\boldsymbol{\eta} = P_t G\boldsymbol{\xi} = G_t\boldsymbol{\xi}$ where $G_t \in \mathbb{R}^{(1+tn_y)\times N_\xi}$ is the sub-matrix of G which is obtained by removing the last $((T-1)-t)n_y$ rows of G.

Part 1 We first show that for every $\boldsymbol{u} \in \mathcal{N}$ there exists a matrix $Q_t \in \mathbb{R}^{n_u \times (1+n_y t)}$ that satisfies

$$Q_t G_t M G_t^\top = \mathbb{E}\left[\boldsymbol{u}_t\ (G_t\boldsymbol{\xi})^\top\right]. \tag{3.19}$$

Since the matrix M is positive definite by assumption, the null space relationship $\ker(G_t M G_t^\top) = \ker([G_t M G_t^\top]^\dagger) = \ker(G_t^\top)$ holds. We set

$$Q_t = E\left[\boldsymbol{u}_t\ (G_t\boldsymbol{\xi})^\top\right](G_t M G_t^\top)^\dagger.$$

Then $Q_t v = 0$ for all $v \in \ker(G_t^\top)$, and $Q_t G_t M G_t^\top v = \mathbb{E}\left[\boldsymbol{u}_t\ (G_t\boldsymbol{\xi})^\top\right] v$ for all v in the orthogonal complement of $\ker(G_t^\top)$, which guarantees (3.19).

Part 2 Next we show that for every $\boldsymbol{u} \in \mathcal{N}$ there exists a $Q \in \mathcal{U}$ so that

$$QGMG = \mathbb{E}\left[\boldsymbol{u}\boldsymbol{\eta}^\top\right]. \tag{3.20}$$

By assumption, there exists a matrix $H_t \in \mathbb{R}^{N_y \times (1+tn_y)}$ so that $\mathbb{E}\left[\boldsymbol{\eta}|\,\boldsymbol{\eta}^t\right] = H_t P_t\,\boldsymbol{\eta} = H_t G_t\,\boldsymbol{\xi}$ $\mathbb{P}$-a.s. Therefore, we find

$$\begin{aligned}
\mathbb{E}[\boldsymbol{u}_t\,\boldsymbol{\eta}^\top] &= \mathbb{E}[\boldsymbol{u}_t\,\mathbb{E}[\boldsymbol{\eta}|\,\boldsymbol{\eta}^t]^\top] &&= \mathbb{E}[\boldsymbol{u}_t\,(G_t\,\boldsymbol{\xi})^\top]\,H_t^\top \\
&= Q_t\,G_t\,M\,G_t^\top\,H_t^\top &&= Q_t\,\mathbb{E}[G_t\,\boldsymbol{\xi}\,(H_t\,G_t\,\boldsymbol{\xi})^\top] \\
&= Q_t\,\mathbb{E}[\boldsymbol{\eta}^t\,\mathbb{E}[\boldsymbol{\eta}|\,\boldsymbol{\eta}^t]^\top] &&= Q_t\,\mathbb{E}[\boldsymbol{\eta}^t\,\boldsymbol{\eta}^\top] \\
&= Q_t\,G_t\,\mathbb{E}[\boldsymbol{\xi}\,\boldsymbol{\xi}^\top]G^\top &&= Q_t\,P_t\,G\,M\,G^\top,
\end{aligned} \tag{3.21}$$

where the first line follows from the law of iterated conditional expectations and the $\mathcal{F}_t$ -measurability of $\boldsymbol{u}_t$. The second line holds due to (3.19) and the definition of M, and the third line follows from

the definition of H_t and the law of iterated conditional expectations. The last line follows from the definition of G_t and M. Finally, we define

$$Q := \begin{pmatrix} Q_0 P_0 \\ Q_1 P_1 \\ \vdots \\ Q_{T-1} P_{T-1} \end{pmatrix} \in \mathcal{U}.$$

By construction, Q satisfies (3.20).

Part 3 To conclude the proof, we use the relation (3.20) in conjunction with assumption **A1** that $\mathbb{E}[\boldsymbol{\xi}|\boldsymbol{\eta}] = L\boldsymbol{\eta}$ $\mathbb{P}$-a.s. We have

$$\begin{aligned} \mathbb{E}\left[\boldsymbol{u}\,\boldsymbol{\xi}^\top\right] &= \mathbb{E}\left[\boldsymbol{u}\,\mathbb{E}\left(\boldsymbol{\xi}|\boldsymbol{\eta}\right)^\top\right] \\ &= \mathbb{E}\left[\boldsymbol{u}\,\boldsymbol{\eta}^\top\right] L^\top \\ &= QGMG^\top L^\top = QGM, \end{aligned} \tag{3.22}$$

where that last equality follows from

$$\begin{aligned} GMG^\top L^\top &= G\mathbb{E}\left[\boldsymbol{\xi}\,\boldsymbol{\xi}^\top\right] G^\top L^\top = \mathbb{E}\left[\boldsymbol{\eta}\,\boldsymbol{\eta}^\top\right] L^\top \\ &= \mathbb{E}\left[\boldsymbol{\eta}\,\mathbb{E}\left(\boldsymbol{\xi}|\boldsymbol{\eta}\right)^\top\right] = \mathbb{E}\left[\boldsymbol{\eta}\,\boldsymbol{\xi}^\top\right] \\ &= GM. \end{aligned} \tag{3.23}$$

This observation completes the proof of the second statement. □

Remark 3.4.2. *For any $Q \in \mathcal{U}$, the affine controller $\boldsymbol{u} = Q\boldsymbol{\eta} \in \mathcal{N}$ satisfies* (3.17) *since*

$$\mathbb{E}[Q\boldsymbol{\eta}\,\boldsymbol{\xi}^\top] = QG\mathbb{E}[\boldsymbol{\xi}\boldsymbol{\xi}^\top] = QGM.$$

The second part of Lemma 3.4.1 investigates the converse situation in which some $\boldsymbol{u} \in \mathcal{N}$ (not necessarily affine in $\boldsymbol{\eta}$) is given, and we seek some $Q \in \mathcal{U}$ satisfying (3.17).

The results of Lemma 3.4.1 allow us to add new decision variables $Q \in \mathcal{U}$ and $S \in \mathbb{R}^{N_s \times N_\xi}$ to problem (3.15) and append (3.16) and (3.17) as additional constraints without constraining the choice $\boldsymbol{u} \in \mathcal{N}$

and $\boldsymbol{s} \in \mathcal{L}^2_{N_s}$ in (3.15). Therefore (3.15) is equivalent to the optimization problem

$$\begin{array}{ll} \inf & \mathbb{E}\left[\boldsymbol{u}(J_u + B^\top J_x B)\boldsymbol{u}^\top\right] + 2\,\mathrm{tr}\left(C^\top J_x B Q G M\right) + \mathrm{tr}(C^\top J_x C M) \\ \text{s.t.} & \boldsymbol{u} \in \mathcal{N},\, Q \in \mathcal{U},\, S \in \mathbb{R}^{N_s \times N_\xi} \\ & (F_u + F_x B)QGM + F_s SM + (F_x C - h e_0^\top)M = 0 \\ & QGM = \mathbb{E}\left[\boldsymbol{u}\,\boldsymbol{\xi}^\top\right] \\ & \exists \boldsymbol{s} \in \mathcal{L}^2_{N_s} : SM = \mathbb{E}\left[\boldsymbol{s}\,\boldsymbol{\xi}^\top\right],\ \boldsymbol{s} \geq 0 \quad \mathbb{P}\text{-a.s.} \end{array} \tag{3.24}$$

We next introduce a technical Lemma that will enable us to eliminate the functional decision variable $\boldsymbol{u}$ in (3.24) by explicitly minimizing the part of the objective function containing this term:

Lemma 3.4.2. *Consider the convex optimization problem*

$$\begin{array}{ll} \min & \mathbb{E}\left[\boldsymbol{u}^\top J \boldsymbol{u}\right] \\ \textit{s.t.} & \boldsymbol{u} \in \mathcal{N} \\ & QGM = \mathbb{E}\left[\boldsymbol{u}\,\boldsymbol{\xi}^\top\right] \end{array} \tag{3.25}$$

where J is positive semi-definite and $Q \in \mathcal{U}$ is fixed. A minimizer for this problem is $\boldsymbol{u}^* = QG\boldsymbol{\xi}$.

Proof. We first consider the relaxed problem

$$\min_{\boldsymbol{u} \in \mathcal{L}^2_{N_u}} \left\{ \mathbb{E}\left[\boldsymbol{u}^\top J \boldsymbol{u}\right] \,:\, QGM = \mathbb{E}\left[\boldsymbol{u}\,\boldsymbol{\xi}^\top\right] \right\}, \tag{3.26}$$

where $\boldsymbol{u}$ is not restricted to be in $\mathcal{N}$. We will show that this problem has a solution in $\mathcal{N}$, which is therefore also a solution for (3.25). Consider the Lagrangian of the relaxed problem (3.26),

$$L(\boldsymbol{u}, \Lambda) = \mathbb{E}\left[\boldsymbol{u}^\top J \boldsymbol{u}\right] + \mathrm{tr}\left(\Lambda^\top \left(QGM - \mathbb{E}\left[\boldsymbol{u}\,\boldsymbol{\xi}^\top\right]\right)\right), \tag{3.27}$$

where $\Lambda \in \mathbb{R}^{N_u \times N_\xi}$ is the matrix of Lagrange multipliers associated with the equality constraints. A set of first-order optimality conditions is found by setting the Gâteaux differential of the Lagrangian with respect to $\boldsymbol{u}$ to zero for all descent directions $\boldsymbol{h} \in L^2_{N_u}$.

$$\begin{array}{rl} \delta L(\boldsymbol{u}, \Lambda; \boldsymbol{h}) = & \mathbb{E}\left[\boldsymbol{h}^\top (2\,J\,\boldsymbol{u} - \Lambda\,\boldsymbol{\xi})\right] = 0 \quad \forall \boldsymbol{h} \in L^2_{N_u} \\ \iff & 2J\boldsymbol{u} - \Lambda\boldsymbol{\xi} = 0 \quad \mathbb{P}\text{-a.s.} \end{array} \tag{3.28}$$

It can be verified that $\Lambda^* = 2JQG$ and $\boldsymbol{u}^* = QG\boldsymbol{\xi} = Q\boldsymbol{\eta}$ satisfy both the optimality conditions (3.28) and the constraints of problem (3.26). Thus $\boldsymbol{u}^*$ constitutes a valid but not necessarily unique solution of the convex problem (3.26). Furthermore, since Q is an element of $\mathcal{U}$, the affine policy $\boldsymbol{u}^*$ is non-anticipative, that is $\boldsymbol{u}^* \in \mathcal{N}$. As a result, $\boldsymbol{u}^*$ is also an optimal solution for problem (3.25), which is more restrictive than (3.26). Thus the claim follows. □

We can apply Lemma 3.4.2 with $J := (J_u + B^\top J_x B)$ to (3.24) and replace $\boldsymbol{u}$ by $QG\boldsymbol{\xi}$. This yields the following optimization problem:

$$\begin{array}{ll} \inf & \operatorname{tr}\left(G^\top Q^\top (J_u + B^\top J_x B) QGM + 2C^\top J_x BQGM + C^\top J_x CM\right) \\ \text{s.t.} & Q \in \mathcal{U},\ S \in \mathbb{R}^{N_s \times N_\xi} \\ & (F_u + F_x B)QG + F_s S + F_x C - he_0^\top = 0 \\ & \exists \boldsymbol{s} \in \mathcal{L}^2_{N_s} : SM = \mathbb{E}\left[\boldsymbol{s}\,\boldsymbol{\xi}^\top\right],\ \boldsymbol{s} \geq 0 \quad \mathbb{P}\text{-a.s.} \end{array} \tag{3.29}$$

Note also that we removed the second-order moment matrix M from the state and control constraint by post-multiplying the corresponding constraint in (3.24) with M^{-1}.

The last constraint in (3.29) requires the solution of N_s moment problems, *i.e.* for a given $S \in \mathbb{R}^{N_s \times N_\xi}$ we must assert the existence of N_s non-negative Borel measures whose vectors of zero- and first-order moments coincide with the rows of SM and which have square integrable densities with respect to $\mathbb{P}$. The following lemma provides a means for dealing with this constraint by replacing it with a set of conic constraints that enclose its feasible region:

Lemma 3.4.3. *Define the cones $\mathcal{C} \subseteq \mathbb{R}^{N_\xi}$ and $\mathcal{C}_{\boldsymbol{\xi}} \subseteq \mathbb{R}^{N_\xi}$ as*

$$\begin{aligned} \mathcal{C}_{\boldsymbol{\xi}} &:= \left\{ z \in \mathbb{R}^{N_\xi} \ : \ \exists \boldsymbol{s} \in \mathcal{L}^2_1 \text{ with } z = \mathbb{E}\left[\boldsymbol{s}\boldsymbol{\xi}\right], \boldsymbol{s} \geq 0\ \mathbb{P}\text{-}a.s \right\} \\ \mathcal{C} &:= \left\{ z \in \mathbb{R}^{N_\xi} \ : \ W_i\, z \succeq_{\mathcal{K}_p} 0, i = 1 \ldots l, e_0^\top z \geq 0 \right\} \\ &= \operatorname{cone}(\Xi). \end{aligned}$$

Then $\emptyset \neq \operatorname{int}\mathcal{C} \subseteq \mathcal{C}_{\boldsymbol{\xi}} \subseteq \mathcal{C}$.

Proof. First, observe that $\emptyset \neq \operatorname{int}\mathcal{C}$. This follows from the assumption that Ξ has a non-empty relative interior and spans $\mathbb{R}^{N_\xi}$. In the remainder of the proof we show that $\operatorname{int}\mathcal{C} \subseteq \mathcal{C}_{\boldsymbol{\xi}} \subseteq \mathcal{C}$. We will show this using the methodologies used in [60, Proposition 3].

Let $\mathcal{M}^+$ be the set of all non-negative finite measures on $(\Xi, \mathcal{B}(\Xi))$ with finite second moments, and let $\mathcal{M}^+_{\boldsymbol{\xi}}$ be the subset of all measures in $\mathcal{M}^+$ that have a square-integrable density with respect to the distribution $\mathbb{P}_{\boldsymbol{\xi}}$ of $\boldsymbol{\xi}$. Define two convex cones in $\mathbb{R}^{N_\xi}$ as

$$\tilde{\mathcal{C}} := \left\{ \int_\Xi \xi \, \mu(\mathrm{d}\xi) : \mu \in \mathcal{M}^+ \right\} \text{ and}$$
$$\tilde{\mathcal{C}}_{\boldsymbol{\xi}} := \left\{ \int_\Xi \xi \, \mu(\mathrm{d}\xi) : \mu \in \mathcal{M}^+_{\boldsymbol{\xi}} \right\}.$$

By construction $\tilde{\mathcal{C}}_{\boldsymbol{\xi}} \subseteq \tilde{\mathcal{C}}$. Since the density of any $\mu \in \mathcal{M}^+_{\boldsymbol{\xi}}$ with respect to $\mathbb{P}_{\boldsymbol{\xi}}$ can be identified with a function in $\mathcal{L}^2_1$, it is clear that $\tilde{\mathcal{C}}_{\boldsymbol{\xi}} = \mathcal{C}_{\boldsymbol{\xi}}$. Furthermore, as $\mathcal{M}^+_{\boldsymbol{\xi}}$ is weakly dense in $\mathcal{M}^+$ (since Ξ constitutes the support of $\mathbb{P}_{\boldsymbol{\xi}}$), and the identity mapping $\xi \mapsto \xi$ is continuous, it follows that $\tilde{\mathcal{C}}_{\boldsymbol{\xi}}$ is dense in $\tilde{\mathcal{C}}$. Keeping in mind that $\tilde{\mathcal{C}}_{\boldsymbol{\xi}}$ is also convex, the above findings imply

$$\operatorname{int} \tilde{\mathcal{C}} \subseteq \tilde{\mathcal{C}}_{\boldsymbol{\xi}} = \mathcal{C}_{\boldsymbol{\xi}} \subseteq \tilde{\mathcal{C}}\,. \tag{3.30}$$

We now proceed to show that $\tilde{\mathcal{C}} = \mathcal{C}$, which will complete the proof. For any $z \in \mathcal{C}$ there exists a $z' \in \Xi$ and a scaling factor $\lambda \geq 0$ such that $z = \lambda z'$. Next, we define $\mu_z := \lambda \delta_{z'}$ where $\delta_{z'}$ is the Dirac measure which concentrates unit mass at the point z'. It is easy to verify that $\mu_z \in \mathcal{M}^+$. We then have

$$z = \int_\Xi \xi \, \mu_z(\mathrm{d}\xi) \in \tilde{\mathcal{C}},$$

yielding the relation $\mathcal{C} \subseteq \tilde{\mathcal{C}}$. In order to derive the opposite relation, we select any $z \in \tilde{\mathcal{C}}$. By the definition of $\tilde{\mathcal{C}}$ there exists a $\mu \in \mathcal{M}^+$ so that $z = \int_\Xi \xi \, \mu(\mathrm{d}\xi)$. We set $\lambda := \mu(\Xi)$. If $\lambda = 0$, then $z = 0 \in \mathcal{C}$, since μ is a non-negative measure. From now on we thus assume that $\lambda > 0$ and set $z' := \frac{1}{\lambda} \int_\Xi \xi \, \mu(\mathrm{d}\xi)$. Observe that μ/λ constitutes a probability measure. Then, it is easy to verify that $z' \in \Xi$, since z' is the mean value of a probability distribution supported on the convex set Ξ. Since $\lambda > 0$, we have $z = \lambda z' \in \mathcal{C}$. As the choice of $z \in \tilde{\mathcal{C}}$ was arbitrary, we have $\tilde{\mathcal{C}} \subseteq \mathcal{C}$, and thus we conclude that $\tilde{\mathcal{C}} = \mathcal{C}$. Substituting this result into (3.30) completes the proof. □

Using the definition of $\mathcal{C}_{\boldsymbol{\xi}}$ in Lemma 3.4.3, we can rewrite (3.29) in the form

$$\begin{array}{ll} \inf & \mathrm{tr}\left(G^\top Q^\top (J_u + B^\top J_x B) QGM + 2C^\top J_x BQGM + C^\top J_x CM\right) \\ \text{s.t.} & Q \in \mathcal{U},\ S \in \mathbb{R}^{N_s \times N_\xi} \\ & (F_u + F_x B)QG + F_s S + F_x C - he_0^\top = 0 \\ & (MS^\top)_i \in C_{\boldsymbol{\xi}},\ \forall i \in \{1, \ldots, N_s\}, \end{array} \tag{3.31}$$

where $(MS^\top)_i$ denotes the i^{th} column of $MS^\top$. The problem (3.31) is therefore equivalent to $\mathcal{P}_\ell$. A lower bound to (3.31) can then be found by taking the closure of its constraints via application of Lemma 3.4.3, replacing $\mathcal{C}_{\boldsymbol{\xi}}$ with $\mathcal{C}$, which results immediately in the problem $\tilde{\mathcal{P}}_\ell$. It can be shown that the application of this closure operation does not affect the optimal value except for pathological cases of little practical interest (c.f. Remark 3.4.1), so that in most cases $\tilde{\mathcal{P}}_\ell$ is actually equivalent to (3.31). Note that $\mathcal{C}$ is the cone generated by Ξ.

Solvability of $\tilde{\mathcal{P}}_\ell$ in polynomial time for any $p \in [1, +\infty]$ is ensured by [90]. □

3.5 Geometric Properties

In this section we make a number of general observations about the relationship between the lower- and upper-bounding problems $\mathcal{P}_\ell$ and $\mathcal{P}_u$. Where necessary for clarity, we will make the dependence of the problem $\mathcal{P}$ on the initial state x_1 explicit by denoting it $\mathcal{P}(x_1)$, and its optimal value $\mathcal{P}^*(x_1)$. We adopt a similar notation in relation to the problems $\mathcal{P}_\ell$ and $\mathcal{P}_u$ and their tractable reformulations $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$. The first result is a natural consequence of Propositions 3.3.2 and 3.4.2 and is the central result in the chapter:

Theorem 3.5.1. *The optimal values of the problem $\mathcal{P}$ and the tractable problems $\tilde{\mathcal{P}}_u$ and $\tilde{\mathcal{P}}_\ell$ satisfy*

$$\tilde{\mathcal{P}}_\ell^*(x_1) \leq \mathcal{P}^*(x_1) \leq \tilde{\mathcal{P}}_u^*(x_1) \quad \forall x_1 \in \mathbb{R}^{n_x}.$$

The above result is significant because it allows one to bound the degree of suboptimality incurred when employing affine decision rules to solve approximately the intractable optimization problem $\mathcal{P}(x_1)$.

Further insight is possible by comparing the sets of feasible decision variables in the finite dimensional

upper- and lower-bounding problems $\tilde{\mathcal{P}}_u(x_1)$ and $\tilde{\mathcal{P}}_\ell(x_1)$. Define the sets of feasible decision variables for these problems as

$$\Pi_u(x_1) := \left\{ (Q,S) \ : \ \begin{array}{l} \tilde{\mathcal{P}}_u(x_1) \text{ is feasible for some} \\ \mu \geq 0,\ \Lambda_i \succeq_{\mathcal{K}_p^*} 0,\ i \in \{1 \dots l\} \end{array} \right\}$$
$$\Pi_\ell(x_1) := \left\{ (Q,S) \ : \ \tilde{\mathcal{P}}_\ell(x_1) \text{ is feasible} \right\},$$

and define the sets of initial states x_1 for which a feasible policy can be found for these problems as

$$X_u := \{x_1 \in \mathbb{R}^{n_x} \ : \ \Pi_u(x_1) \neq \emptyset\}$$
$$X_\ell := \{x_1 \in \mathbb{R}^{n_x} \ : \ \Pi_\ell(x_1) \neq \emptyset\}.$$

The next result guarantees that these sets are nested:

Proposition 3.5.1. *The policy sets $\Pi_u(x_1)$ and $\Pi_\ell(x_1)$ satisfy*

$$\Pi_u(x_1) \subseteq \Pi_\ell(x_1) \quad \forall x_1 \in \mathbb{R}^{n_x},$$

and the initial state sets X_u and X_ℓ satisfy $X_u \subseteq X_\ell$.

Proof. To prove the first part, assume that $(Q,S) \in \Pi_u(x_1)$ is specified, so that $S = \mu e_0^\top + \sum_{i=1}^l \Lambda_i^\top W_i$ for some $\Lambda_i \succeq_{\mathcal{K}_p^*} 0$, $i = 1, \dots, l$, and $\mu \geq 0$. Then

$$\begin{aligned} W_i M S^\top = \ & W_i M (e_0 \mu^\top + \textstyle\sum_{i=1}^l W_i^\top \Lambda_i) \\ = \ & \mathbb{E}\left[W_i \boldsymbol{\xi}\right] \mu^\top + \textstyle\sum_{i=1}^l W_i M W_i^\top \Lambda_i \\ = \ & \mathbb{E}\left[W_i \boldsymbol{\xi}\right] \mu^\top + \textstyle\sum_{i=1}^l \mathbb{E}\left[(W_i \boldsymbol{\xi})(\Lambda_i^\top W_i \boldsymbol{\xi})^\top\right] \\ & \succeq_{\mathcal{K}_p} 0, \end{aligned}$$

since $W_i \xi \succeq_{\mathcal{K}_p} 0$ and $\Lambda_i^\top W_i \xi \geq 0\ \forall \xi \in \Xi$, $i = 1, \dots, l$. This ensures that the conic inequality in $\tilde{\mathcal{P}}_\ell(x_1)$ is satisfied. Furthermore,

$$\begin{aligned} e_0^\top M S^\top \ &= \ e_0^\top M (e_0 \mu^\top + \textstyle\sum_{i=1}^l W_i^\top \Lambda_i) \\ &= \ \mu^\top + \textstyle\sum_{i=1}^l e_0^\top M W_i^\top \Lambda_i \\ &= \ \mu^\top + \textstyle\sum_{i=1}^l \mathbb{E}\left[(\Lambda_i^\top W_i \boldsymbol{\xi})^\top\right] \geq 0. \end{aligned}$$

This ensures that the final inequality constraint in $\tilde{\mathcal{P}}_\ell(x_1)$ is satisfied. Since all other constraints in $\tilde{\mathcal{P}}_\ell(x_1)$ and $\tilde{\mathcal{P}}_\ell(x_1)$ are identical, $(Q, S) \in \Pi_\ell(x_1)$.

The second part of Proposition 3.5.1 follows immediately from the first part. □

Remark 3.5.1. *An important application for solutions to robust finite horizon problems such as $\mathcal{P}$ is in receding horizon control (RHC). If one equips problem $\mathcal{P}$ (alternatively, $\tilde{\mathcal{P}}_u$) with appropriate terminal conditions on the constraints and objective function, then a RHC law synthesized from repeated solutions to $\mathcal{P}$ can be shown to endow the resulting closed-loop system with desirable stability and invariance properties; c.f. [68, §4],[47, 88].*

*It is difficult in general to assess the degradation of performance (if any) of such a controller if one substitutes receding horizon implementation of solutions to problem $\mathcal{P}$ with solutions to its sub-optimal approximation $\tilde{\mathcal{P}}_u$. In particular, there is no obvious method for directly inferring stability properties (*e.g. *input-to-state gain) from the value function $\mathcal{P}^*$, which typically plays the role of a Lyapunov function in RHC. On the other hand, if one defines $X := \{x \ : \ \mathcal{P}(x) \text{ is feasible}\}$, then clearly $X_u \subseteq X \subseteq X_\ell$. Given appropriate terminal conditions, the set X (alternatively, X_u) is the region of attraction of such a RHC controller, and the set difference $X_u \backslash X_\ell$ provides an estimate of the conservatism of a RHC synthesized from $\tilde{\mathcal{P}}_u$ with respect to region of attraction.*

Finally, we show that the upper and lower bounds calculable via $\tilde{\mathcal{P}}_u$ and $\tilde{\mathcal{P}}_\ell$ coincide for problems with equality constraints only, which demonstrates that our approximation methods are not unnecessarily conservative in this case:

Corollary 3.5.1. *If $F_s = 0$ then the optimal values achieved in the upper- and lower-bounding problems $\tilde{\mathcal{P}}_u$ and $\tilde{\mathcal{P}}_\ell$ coincide,* i.e.

$$\tilde{\mathcal{P}}_\ell^*(x_1) = \mathcal{P}^*(x_1) = \tilde{\mathcal{P}}_u^*(x_1) \ .$$

Proof. If $F_s = 0$, then both of the problems $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ are constrained only by the causality condition $Q \in \mathcal{U}$ and the equality $(F_u + F_x B)QG + F_x C - he_0^\top = 0$. This ensures that $\Pi_u(x_1) = \Pi_l(x_1)$, and consequently $\tilde{\mathcal{P}}_\ell^*(x_1) = \tilde{\mathcal{P}}_u^*(x_1)$. The result then follows from Proposition 3.3.2, 3.4.2 and Theorem 3.5.1. □

Remark 3.5.2. *The above result ensures that affine policies computed using both the primal and dual methods of Section 3.3 and 3.4 are optimal for control problems of the type $\mathcal{P}$ if $\boldsymbol{u}$ and $\boldsymbol{x}$ are restricted*

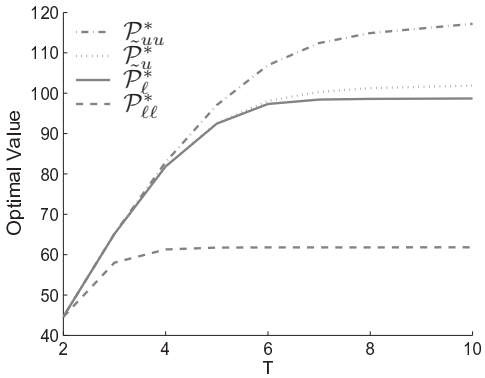


Figure 3.1: Upper and lower bounds for the perfect state measurements example.

only to a subspace. In the particular case that (J_x, J_u) *are block diagonal and* $(F_x, F_u, F_s) = 0$*, both problems reduce to the standard Linear Quadratic Regulator problem, for which linear feedback policies are well-known to be optimal [3].*

3.6 Numerical Examples

3.6.1 Perfect State Measurements

We consider the following time-invariant, discrete-time linear system, which allows for perfect state measurements:

$$\boldsymbol{x}_{t+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}_t + \begin{bmatrix} 0.2 \\ 1 \end{bmatrix} \boldsymbol{u}_t + \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} \boldsymbol{\xi}_t, \ \boldsymbol{y}_t = \boldsymbol{x}_t,$$

where $t = 1, \ldots, T-1$. The initial state of the system is set to $\boldsymbol{x}_1 = (4.7, 0)^\top$. We assume that the $\boldsymbol{\xi}_t$ are independent and uniformly distributed on $[0, 2]$ for $t = 1, \ldots, T$, while $\boldsymbol{\xi}_0 = 1$ $\mathbb{P}$-a.s. Our objective is to minimize $\mathbb{E}\left[\boldsymbol{u}^\top \boldsymbol{u} + \boldsymbol{x}^\top \boldsymbol{x}\right]$ subject to

$$\left.\begin{array}{l} (-5,-5)^\top \leq \boldsymbol{x}_t \leq (5,5)^\top, \ t = 1, \ldots, T \\ (1,1)\, \boldsymbol{x}_t \leq 5, \ t = 1, \ldots, T \\ (1,-1)\, \boldsymbol{x}_t \leq 5, \ t = 1, \ldots, T \\ |\boldsymbol{u}_t| \leq 1, \ t = 1, \ldots, T-1 \end{array}\right\} \ \mathbb{P}\text{-a.s.}$$

We solve the approximate problems $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ as well as the trivial bound problems $\mathcal{P}_{\ell\ell}$ and $\mathcal{P}_{uu}$ for

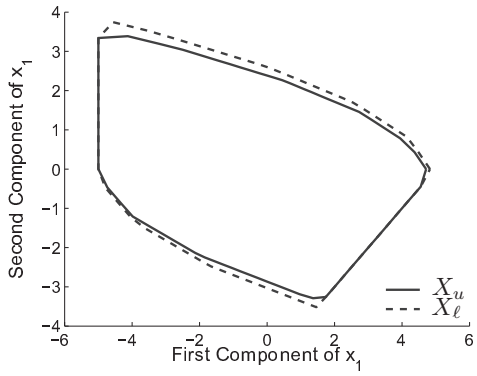


Figure 3.2: Sets of initial states X_ℓ and X_u for horizon $T = 8$, for the perfect state measurements example.

$T = 2, \ldots, 10$. Figure 3.1 shows results for the optimal value achieved by solving these problems for the given horizon lengths. For this particular example, the upper and lower bounds $\tilde{\mathcal{P}}_u$ and $\tilde{\mathcal{P}}_\ell$ are close to each other but diverge slowly. The bounds do not coincide suggesting that affine policies may not be optimal, especially as the time horizon increases beyond 4. The trivial bounds $\mathcal{P}_{\ell\ell}$ and $\mathcal{P}_{uu}$ fail to detect the near-optimality of affine policies for time horizons less than 4. Figure 3.2 shows the two regions X_ℓ and X_u of initial states x_1, for which a feasible policy can be found for $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ respectively, for a time horizon of $T = 8$. Observe that $X_u \subseteq X_\ell$.

3.6.2 Imperfect State Measurements

We consider a similar system but with imperfect state information and measurement errors. The system equations are:

$$\boldsymbol{x}_{t+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}_t + \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \boldsymbol{u}_t + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \boldsymbol{\xi}_t$$

$$\boldsymbol{y}_t = \begin{bmatrix} 1 & 1 \end{bmatrix} \boldsymbol{x}_t + \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \boldsymbol{\xi}_t$$

for all $t = 1 \ldots T-1$. The initial state is set to $\boldsymbol{x}_1 = (0, 1)^\top$ and the support Ξ is defined as

$$\Xi := \left\{ \xi \in \mathbb{R}^{N_\xi} : \ \xi \succeq_{\mathcal{K}_2} 0, \, e_0^\top \xi = 1 \right\}.$$

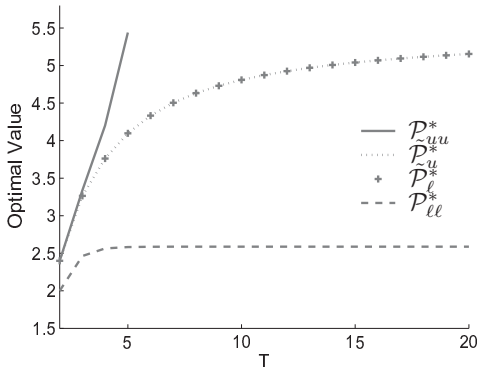


Figure 3.3: The upper and lower bounds for the imperfect state measurements example. The two bounds $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ coincide meaning that for this problem instance affine policies are optimal.

The process noise $\boldsymbol{\xi}$ is uniformly distributed on Ξ. Our objective is to minimize $\mathbb{E}\left[\boldsymbol{u}^\top \boldsymbol{u} + \boldsymbol{x}^\top \boldsymbol{x}\right]$ subject to

$$\left.\begin{array}{ll} (-5,-5)^\top \leq \boldsymbol{x}_t \leq (5,5)^\top, & t = 1,\ldots,T \\ |\boldsymbol{u}_t| \leq 2, & t = 1,\ldots,T-1 \end{array}\right\} \quad \mathbb{P}\text{-a.s.}$$

We solve the approximate problems $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ and the trivial bound problems $\mathcal{P}_{\ell\ell}$ and $\mathcal{P}_{uu}$ for $T = 2,\ldots,20$. The results are shown in Figure 3.3. For this example, the bounds $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ coincide, meaning there is no optimality gap and thus affine policies are optimal in this instance. This is of course not true in general, but as the proposed bounds are problem-specific, they can attest distinct instances of problems where affine controllers are optimal. Problem $\mathcal{P}_{uu}$ becomes infeasible for $T \geq 6$, rendering the trivial upper bound meaningless for those horizons.

3.7 Conclusion

We investigate discrete time, time-varying linear systems with input and measurement uncertainty. Finding causal controllers that minimize an expected cost function is computationally intractable in the presence of state constraints. Therefore, we restrict attention to affine controllers over which one can optimize in polynomial time by solving a tractable conic optimization problem. The optimal value of this problem constitutes an upper bound on the minimal cost achievable. Estimating the degree of suboptimality of the best *affine* controller has been a long-standing open problem. By linearizing the dual variables in the original control problem, we devise a relaxed problem whose optimal value

provides a lower bound on the minimal cost achievable by any *causal* controller. We argue that this problem is again equivalent to a tractable conic optimization problem. The difference between the two bounds constitutes an a posteriori measure for the degree of suboptimality of the best affine controller. This error estimate depends on the structure of the underlying control problem and can be calculated in polynomial time.

Chapter 4

Linear Decision Rules in Two-stage Robust Optimization Problems

In this chapter, we study two-stage robust optimization problems with linear robust constraints and a min-max linear objective. Two-stage robust optimization problems are computationally intractable in general. Tractable upper bounding approximations can be obtained by requiring the adaptive decisions to be representable as linear decision rules. In this chapter we investigate families of tractable lower bounding approximations, which serve to estimate the degree of suboptimality of the best linear decision rule. These approximations are obtained either by solving a dual version of the robust optimization problem in linear decision rules or by discretizing the problem's uncertainty set. The quality of the resulting lower bounds depends on the distribution assigned to the uncertain parameters or the choice of the discretization points within the uncertainty set, respectively. We prove that identifying the best possible lower bound is generally intractable in both cases and propose an efficient procedure to construct suboptimal lower bounds. We also demonstrate numerically that the resulting instance-wise bounds outperform known worst-case bounds in the vast majority of problem instances.

4.1 Motivation

Robust optimization problems involving adaptive decisions are known to be computationally intractable [11, 17] in general. To reduce this complexity, methods for restricting the space of adaptive decision rules have

been suggested, *e.g.* by restricting the decision rules to those with linear [11], piecewise linear [29, 45] or polynomial structure [19]. Because of their desirable scalability properties, linear decision rules enjoy the widest popularity. Indeed, the best linear decision rule for a given robust optimization problem can typically be computed by solving a tractable linear or second-order cone program.

Unfortunately, linear decision rules may be severely suboptimal or even infeasible in the original optimization problem. In order to assess the appropriateness of using linear decision rules, one thus needs to estimate their loss of optimality. We distinguish two complementary approaches for this purpose: *a priori* methods evaluate the worst-case approximation ratio of linear decision rules over a whole class of problems, while *a posteriori* methods estimate the approximation error for each problem instance individually.

We first review the existing a priori methods. Linear decision rules have been shown to be optimal for two-stage min-max problems with simplicial uncertainty sets and for certain one-dimensional robust control problems [18]. However, instances satisfying these idealized conditions are rarely encountered in practice. Under restrictive non-negativity conditions on the problem data, it has also been shown that the worst-case approximation ratio of linear decision rules in two-stage min-max problems with m linear constraints is of the order $\Omega(\sqrt{m})$ [17]. Even though it is occasionally tight, this worst-case performance bound is too pessimistic for the vast majority of problem instances.

Thus, there is considerable merit in developing good a posteriori methods. Instance-wise upper and lower bounds on multi-stage *stochastic* programs can be obtained by solving both the original problem and its dual in linear decision rules; see Chapter 3 of this thesis, see also [60]. The gap between the bounds provides an a posteriori measure for the suboptimality of linear decision rules. This approach is directly applicable to robust optimization problems. However, the dual (lower) bound becomes non-unique in the robust case as it depends on the probability distribution of the uncertain parameters (which can be chosen freely from amongst distributions with appropriate support).

In this chapter we demonstrate that the quality of the dual linear decision rule bound for robust problems is highly sensitive to the choice of the distribution governing the uncertain parameters. We show that the best (maximum) lower bound is achieved by a *discrete* distribution. However, we further show that finding this distribution is as hard as solving the original (intractable) problem. A different class of tractable lower bounds is obtained by replacing the original uncertainty set with a finite scenario set. Again, finding the best scenario set of a prescribed cardinality is generically hard. We

therefore propose an efficient method in which scenario sets are taken from the Lagrange multipliers associated with the primal linear decision rule problem. Next, we establish a link between the dual linear decision rule method and the scenario approach. This analysis allows us to identify new problem classes for which linear decision rules are optimal. Extensive numerical experiments demonstrate that our scenario-based lower bound consistently outperforms the known a priori bounds as well as the dual linear decision rule bound associated with naive distributional choices.

This chapter is structured as follows. Section 4.2 introduces the problem formulation and reviews an approximate solution method based on linear decision rules. Section 4.4.1 and 4.4.2 describe bounding techniques for estimating the suboptimality of linear decision rules by using dual linear decision rules and a discretization of the uncertainty set, respectively. The performance of the resulting bounds is analysed in Section 4.5 on a set of randomly generated test problems, and some conclusions are drawn in Section 4.6.

Chapter Notation The optimal value of any optimization problem $\mathcal{P}$ is denoted by $\mathcal{P}^*$. For matrices $S, T \in \mathbb{R}^{k+1\times l}$ and a proper cone $\mathcal{K} \subseteq \mathbb{R}^{k+1}$, the relation $S \succeq_{\mathcal{K}} T$ ($S \preceq_{\mathcal{K}} T$) indicates that the columns of $S - T$ ($T - S$) are included in the cone $\mathcal{K}$. The dual cone of $\mathcal{K}$ is denoted $\mathcal{K}^*$. Moreover, the kernel of a matrix A is denoted $\ker(A)$. Finally, we use $A^{(i)}$ to denote the i-th row of a matrix A, and $A_{(i)}$ to denote its i-th column.

4.2 Problem Statement and Assumptions

We study linear two-stage robust optimization problems with affine right hand side uncertainty and a min-max objective. Such problems can be represented as

$$\begin{array}{ll} \inf & c^\top x \; + \; \sup\limits_{\xi\in\Xi} \; d^\top y(\xi) \\ \text{s.t.} & Ax \; + \; By(\xi) \leq C\xi \quad \forall \xi \in \Xi, \end{array} \tag{$\mathcal{P}$}$$

where $\Xi \subseteq \mathbb{R}^{k+1}$ represents the uncertainty set. The first-stage decision $x \in \mathbb{R}^{n_1}$ is a rigid design decision, and the second-stage decision y is a fully adaptive decision rule, that is, a continuous function from $\mathbb{R}^{k+1}$ to $\mathbb{R}^{n_2}$. We denote by m the number of constraints. Without much loss of generality, we assume that $\mathcal{P}^*$ is finite and that the minimum is attained by an optimal decision (x^*, y^*). Moreover,

we make the following assumption about Ξ.

A2 (Convex uncertainty set) *The set Ξ is the intersection of proper cone $\mathcal{K} \subseteq \mathbb{R}^{k+1}$ with the hyperplane $\left\{\xi \in \mathbb{R}^{k+1} \ : \ e_0^\top \xi = 1\right\}$, where e_0 denotes the first canonical basis vector in $\mathbb{R}^{k+1}$:*

$$\Xi := \left\{\xi \in \mathbb{R}^{k+1} \ : \ \xi \in \mathcal{K},\ e_0^\top \xi = 1\right\}.$$

We assume that Ξ is non-empty and bounded.

Remark 4.2.1. *Introducing a degenerate uncertain variable ξ_0 that is equal to 1 on Ξ allows us to express any affine function of the non-degenerate uncertain parameters $(\xi_1, \ldots, \xi_m)$ on Ξ in a compact way as a linear function of $\xi = (\xi_0, \ldots, \xi_m)$.*

4.3 Upper Approximations of $\mathcal{P}$ and Primal Linear Decision Rules

Problem $\mathcal{P}$ is computationally intractable [11, Theorem 2.2], involving an infinite number of constraints and decision variables. Therefore, finding a suboptimal solution necessitates a trade-off between accuracy and tractability, usually in the form of a restriction on the structure of the adaptive decision rules in $\mathcal{P}$. In this chapter, we investigate such an approximation, where the second-stage adaptive decision rule y is restricted to be a linear function of ξ, *i.e.*, $y(\xi) = Y\xi$ for some matrix $Y \in \mathbb{R}^{n_2 \times (k+1)}$. The problem of identifying an optimal linear decision rule can be represented as

$$\begin{array}{ll} \inf & c^\top x \ + \ \sup\limits_{\xi \in \Xi} \ d^\top Y \xi \\ \text{s.t.} & Ax \ + \ BY\xi \leq C\xi \quad \forall \xi \in \Xi, \end{array} \tag{$\mathcal{U}$}$$

where the second-stage decision is now encoded by the matrix Y instead of the continuous function y. Problem $\mathcal{U}$ constitutes a linear robust optimization problem with semi-infinite constraints. By using

robust optimization techniques [11], it can be reformulated as a conic optimization problem of the form

$$\begin{array}{ll} \inf & c^\top x + t \\ \text{s.t.} & \left(d^\top Y - te_0^\top\right)^\top \preceq_{\mathcal{K}^*} 0 \\ & \left(Axe_0^\top + BY - C\right)^\top \preceq_{\mathcal{K}^*} 0. \end{array} \qquad (\tilde{\mathcal{U}})$$

Problem $\mathcal{U}$ and its equivalent conic reformulation $\tilde{\mathcal{U}}$ are derived via a restriction of the set of admissible second-stage decisions in $\mathcal{P}$. In general, an optimal fully adaptive decision rule y^* is continuous piecewise linear in ξ, with a possibly exponential number of pieces [7]. Restricting y to be linear in ξ introduces an optimality gap between $\mathcal{P}$ and $\mathcal{U}$ with $\mathcal{P}^* \leq \mathcal{U}^* = \tilde{\mathcal{U}}^*$. The loss of performance due to the use of linear decision rules varies greatly according to the problem data. Identifying a bound on this optimality gap requires the derivation of a lower bound for $\mathcal{P}^*$ tailored to the specific problem instance. In the remainder of this chapter we demonstrate how one can derive such problem-specific lower bounds in an efficient manner.

4.3.1 Dual variables and the robust counterpart $\tilde{\mathcal{U}}$

Before we discuss the first lower bound, we comment on the relationship between the semi-infinite problem $\mathcal{U}$ and its conic equivalent $\tilde{\mathcal{U}}$ and demonstrate how one can derive $\tilde{\mathcal{U}}$ from $\mathcal{U}$ using two distinct approaches. The first approach relies on the following lemma, which captures the essence of robust optimization (see Lemma 3.3.1 for detailed proof):

Lemma 4.3.1 ([13, Theorem 3.1]). *For any $\sigma \in \mathbb{R}^{k+1}$ we have*

$$\sigma^\top \xi \geq 0 \quad \forall \xi \in \Xi \iff \sigma \in \mathcal{K}^*,$$

where $\mathcal{K}$ is the cone generated by Ξ, and $\mathcal{K}^*$ its dual cone.

Problem $\tilde{\mathcal{U}}$ can be derived from $\mathcal{U}$ via application of Lemma 4.3.1, which has the effect of replacing the semi-infinite linear inequality constraints in $\mathcal{U}$ with a set of finite dimensional constraints that define the dual cone $\mathcal{K}^*$.

We suggest now an alternative derivation based on Lagrangian functions, which provides a different insight into the relationship between $\mathcal{U}$ and $\tilde{\mathcal{U}}$. Let $\mathcal{M}_+^d$ be the space of all d-dimensional non-negative

Borel measures on Ξ. Let f be any continuous function from $\mathbb{R}^{k+1}$ to $\mathbb{R}^d$. Then we have

$$\sup_{\psi \in \mathcal{M}_+^d} \int_\Xi f(\xi)^\top \psi(\mathrm{d}\xi) = \begin{cases} \infty & \text{if } \exists \xi \in \Xi : \ f(\xi) > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

Let $q \in \mathcal{M}_+^m$ and $z \in \mathcal{M}_+$ be m-dimensional and one-dimensional non-negative Borel measures on Ξ, respectively. By using (4.1), we can derive the Lagrangian of $\mathcal{U}$ using q and z as the dual variables for the semi-infinite inequality constraints, *i.e.*,

$$\begin{aligned} L_{\mathcal{U}}(x, Y, t; q, z) = \ & c^\top x + t + \left(d^\top Y - t e_0^\top\right) \textstyle\int_\Xi \xi z(\mathrm{d}\xi) \\ & + \operatorname{tr}\left[(A x e_0^\top + BY - C) \textstyle\int_\Xi \xi q^\top(\mathrm{d}\xi)\right]. \end{aligned}$$

The emergence of Borel measures as dual variables in the Lagrangian is a consequence of the semi-infinite nature of the constraints in $\mathcal{P}$. However, these dual Borel measures only affect the Lagrangian through their first moments. Consider therefore the variable substitution $\Lambda = \int_\Xi \xi q^\top(\mathrm{d}\xi)$ and $\lambda = \int_\Xi \xi z(\mathrm{d}\xi)$, where the new variables Λ and λ are the first moments of the Borel measures q and z, respectively. The non-negativity constraints $q \in \mathcal{M}_+^m$ and $z \in \mathcal{M}_+$ translate to the new variables as $(\lambda, \Lambda) \succeq_{\mathcal{K}} 0$. Substituting these new variables into $L_{\mathcal{U}}$ yields:

$$\begin{aligned} L_{\mathcal{U}} \ & = c^\top x + t + \left(d^\top Y - t e_0^\top\right) \lambda + \operatorname{tr}\left[(A x e_0^\top + BY - C)\Lambda\right] \\ & = L_{\tilde{\mathcal{U}}}(x, Y, t; \lambda, \Lambda), \end{aligned} \tag{4.2}$$

which is precisely the Lagrangian of $\tilde{\mathcal{U}}$, where Λ and λ are the dual variables of the conic inequality constraints.

The preceding line of argument provides an alternate perspective on how the conic problem $\tilde{\mathcal{U}}$ relates to the semi-infinite problem $\mathcal{U}$. The established interpretation dictates that problem $\tilde{\mathcal{U}}$ is derived from its semi-infinite equivalent via a manipulation of the constraints of $\mathcal{U}$ that utilizes convex duality arguments, through the mechanism of Lemma 4.3.1. This alternative derivation can be interpreted as a simple variable substitution in the Lagrangians, where the dual Borel measures in $L_{\mathcal{U}}$ are substituted by their respective moments to obtain the Lagrangian $L_{\tilde{\mathcal{U}}}$.

4.4 Lower Approximations of $\mathcal{P}$

We are interested in characterising the suboptimality of linear decision rules in two-stage robust problems. To this end, we need to derive a lower bound on the performance of any decision rule (including the global optimal one) for the given two-stage robust optimization problem. In this chapter, we describe two methodologies for deriving such a bound.

4.4.1 Dual Linear Decision Rules Approximation

The problem of deriving a lower bound for two-stage decision problems was addressed successfully in the stochastic programming setting [60], where a lower bound was proposed for linear problems similar to $\mathcal{P}$, but with expectation objective and a prescribed probability distribution for ξ. The bound derived in [60] employed a dualization of the original decision rule problem, followed by a restriction of the dual variables to be linear in ξ. The resulting semi-infinite problem is a relaxation of $\mathcal{P}$, where the infinitely many linear constraints, one for each uncertainty realization, are multiplied by their corresponding ξ and aggregated into finitely many expectation constraints.

Let $\mathcal{C}$ be the set of all probability distributions supported on Ξ. For any given $\mathbb{P} \in \mathcal{C}$, we can follow the approach of [60] (which was also utilized in Chapter 3), to construct a lower bound for the min-max problem $\mathcal{P}$ as

$$\begin{array}{ll} \inf & c^\top x + t \\ \text{s.t.} & \mathbb{E}_{\mathbb{P}}\left[\left(d^\top y(\xi) - t\right)\xi^\top\right] \leq 0 \\ & \mathbb{E}_{\mathbb{P}}\left[\left(Ax + By(\xi) - C\xi\right)\xi^\top\right] \leq 0. \end{array} \qquad (\mathcal{L}(\mathbb{P}))$$

Even though $\mathcal{L}(\mathbb{P})$ involves functional decision variables and is therefore seemingly intractable, it has been shown that under a mild strict feasibility condition, $\mathcal{L}(\mathbb{P})$ can be reformulated as a conic optimization problem [60]. One can immediately verify that $\mathcal{L}(\mathbb{P})$ provides a lower bound on $\mathcal{P}$ for any possible distribution $\mathbb{P} \in \mathcal{C}$:

Proposition 4.4.1. *For any probability distribution $\mathbb{P} \in \mathcal{C}$, we have $\mathcal{L}^*(\mathbb{P}) \leq \mathcal{P}^* \leq \mathcal{U}^*$.*

Any solution that is feasible in $\mathcal{P}$ will of course also satisfy the less restrictive expectation constraints in $\mathcal{L}(\mathbb{P})$, regardless of the probability distribution used. However, the choice of a distribution $\mathbb{P}$ has a central role in the performance of the bound $\mathcal{L}^*(\mathbb{P})$.

Constraints in the original problem $\mathcal{P}$ corresponding to uncertainties with higher probability mass under $\mathbb{P}$ are more likely to be satisfied by the optimal solution of $\mathcal{L}(\mathbb{P})$. In the absence of any information about the distribution $\mathbb{P}$, as is the case with $\mathcal{P}$, we are free to select any distribution from $\mathcal{C}$ to be used in $\mathcal{L}(\mathbb{P})$.

In the extreme case, one can adopt a Dirac distribution from $\mathcal{C}$ that concentrates probability mass on a single uncertainty realization $\xi \in \Xi$. A solution to $\mathcal{L}(\mathbb{P})$ will then only need to satisfy the constraints corresponding to that single uncertainty realization, *i.e.*, the problem becomes deterministic. In the event that such constraints are not binding, the optimal value of $\mathcal{L}(\mathbb{P})$ is $-\infty$, and the lower bound is trivial.

Consequently, if one wishes to obtain a useful lower bound from $\mathcal{L}(\mathbb{P})$, then the distribution $\mathbb{P}$ must be carefully chosen, ideally from the set of worst-case distributions

$$\mathcal{C}_w := \arg\max_{\mathbb{P}\in\mathcal{C}} \mathcal{L}(\mathbb{P}),$$

which can be shown to be non-empty. Selecting any distribution $\mathbb{P}_w \in \mathcal{C}_w$ ensures that $\mathcal{L}^*(\mathbb{P}_w)$ constitutes the best possible lower bound for $\mathcal{P}$ among all dual linear decision rule bounds. Unfortunately, finding an element $\mathbb{P}_w \in \mathcal{C}_w$ and evaluating $\mathcal{L}^*(\mathbb{P}_w)$ is no less difficult than computing the optimal value of $\mathcal{P}$:

Theorem 4.4.1. *The set $\mathcal{C}_w$ is non-empty, and for any worst-case distribution $\mathbb{P}_w \in \mathcal{C}_w$, the lower bounding problem $\mathcal{L}(\mathbb{P}_w)$ satisfies*

$$\mathcal{L}^*(\mathbb{P}_w) = \mathcal{P}^*.$$

Thus, computing the best possible dual linear decision rule bound is as hard as computing the optimal value of $\mathcal{P}$.

Proof. First define the optimal second-stage cost of $\mathcal{P}$, which depends parametrically on the first-stage decision x and the uncertainty realization ξ, *i.e.*,

$$\begin{aligned} Q(x,\xi) \;:=\; \min_{y} \quad & d^\top y \\ \text{s.t.} \quad & Ax + By \leq C\xi. \end{aligned} \tag{4.3}$$

The optimal value function Q is known to be convex and continuous on its effective domain. Problem

$\mathcal{P}$ can now be expressed as the following min-max problem:

$$\mathcal{P} = \inf \left[c^\top x + \sup_{\xi \in \Xi} Q(x, \xi) \right].$$

Let x^* be an optimal first-stage decision and Ξ_w be the set of worst-case scenarios for the corresponding optimal second-stage parametric cost,

$$\Xi_w := \arg\max_{\xi \in \Xi} Q(x^*, \xi). \tag{4.4}$$

Note that Ξ_w is non-empty since Q is continuous on its effective domain and Ξ is compact. For any scenario $\xi_w \in \Xi_w$ and any $x \in \mathbb{R}^{n_1}$, we have $Q(x, \xi_w) \geq Q(x^*, \xi_w)$. Now let δ_w be any Dirac distribution concentrating mass on a single element $\xi_w \in \Xi_w$. Then, we find

$$\begin{aligned} \mathcal{L}^*(\delta_w) &= \inf_{\mathrm{x}} \ \mathrm{c}^\top \mathrm{x} + \mathrm{Q}(\mathrm{x}, \xi_{\mathrm{w}}) \\ &= c^\top x^* + Q(x^*, \xi_w) \\ &= \mathcal{P}^*. \end{aligned}$$

Thus, $\mathcal{C}_w$ is non-empty, and any $\mathbb{P}_w \in \mathcal{C}_w$ results in an exact lower bound $\mathcal{L}^*(\mathbb{P}_w) = \mathcal{P}^*$. Since the problem $\mathcal{L}(\delta_w)$ is actually finite and therefore a tractable convex cone program, finding a Dirac distribution in $\delta_w \in \mathcal{C}_w$ has the same complexity as finding the optimal value of $\mathcal{P}$. □

Theorem 4.4.1 has two important implications. First, finding the best dual linear decision rule bound is actually as hard as deriving the optimal value of the original problem $\mathcal{P}$. Using any lower bound $\mathcal{L}^*(\mathbb{P})$ associated with a distribution $\mathbb{P} \notin \mathcal{C}_w$ will provide only a conservative estimate of the suboptimality of linear decision rules.

Second, the proof of Theorem 4.4.1 illustrates that there exist specific uncertainty realizations, namely the elements of Ξ_w, from which Dirac distributions $\mathbb{P}_w$ can be constructed that provide tight lower bounds $\mathcal{L}^*(\mathbb{P}_w) = \mathcal{P}^*$.

As a result, a different method of bounding $\mathcal{P}$ can be designed, which is inspired by the existence of the set Ξ_w and relies on a discretization of the uncertainty set.

4.4.2 Scenario Based Approximation

We propose a procedure in which we identify a discrete subset $Z \subset \Xi$ and solve $\mathcal{P}(Z)$, a variant of the problem $\mathcal{P}$ in which the uncertainty set Ξ is replaced by Z. This results in a solution that is robust only with respect to a finite subset of all scenarios $\xi \in \Xi$, namely the elements of Z. Problem $\mathcal{P}(Z)$ is thus also finite, whose solution provides a lower bound on $\mathcal{P}^*$.

Theorem 4.4.2. *For any finite subset $Z \subset \Xi$, problem $\mathcal{P}(Z)$ provides a lower bound on $\mathcal{P}$ that can be obtained via the solution of a finite linear program. Furthermore, for any distribution $\mathbb{P}_z \in \mathcal{C}$ supported on Z, $\mathcal{P}(Z)$ provides a tighter lower bound on $\mathcal{P}$ than $\mathcal{L}(\mathbb{P}_z)$, that is,*

$$\mathcal{L}^*(\mathbb{P}_z) \leq \mathcal{P}^*(Z) \leq \mathcal{P}^*.$$

Proof. The proof of the first statement follows immediately because $Z \subset \Xi$. Thus, any solution that is feasible in $\mathcal{P}$ will also be feasible in $\mathcal{P}(Z)$, resulting in the inequality $\mathcal{P}^*(Z) \leq \mathcal{P}^*$. The second part of the proof follows by the definition of $\mathcal{L}^*(\mathbb{P}_z)$, which is designed to be a lower bound on the problem $\mathcal{P}(Z)$. □

The quality of this lower bound depends on the choice of Z, in a manner equivalent to the dependence of $\mathcal{L}(\mathbb{P})$ on a selection of the distribution $\mathbb{P}$. The next result follows immediately from the preceding discussion and is a generalization of Theorem 4.4.1.

Theorem 4.4.3. *There exists a non-empty set $\Xi_w \subseteq \Xi$, so that for any $Z \subseteq \Xi$ with $Z \bigcap \Xi_w \neq \emptyset$, we have*

$$\mathcal{P}^*(Z) = \mathcal{P}^*.$$

We note that Theorem 4.4.3 is reminiscent of Theorem 4.4.1. While Theorem 4.4.1 relates to the existence of worst-case discrete distributions that maximize the dual linear decision rule bound $\mathcal{L}^*(\mathbb{P})$, Theorem 4.4.3 guarantees the existence of finite worst-case scenario sets Z that maximize the discretization bound $\mathcal{P}^*(Z)$. Of course, finding a worst-case scenario set, or equivalently an element of Ξ_w is at least as hard as computing $\mathcal{P}^*$ and thus is itself intractable. Thus, one must resort to other methods for choosing the set Z.

A naïve approach is to let Z be the set of all extreme points of Ξ. It is known that Ξ_w contains at least one of these extreme points [17], and therefore such Z is guaranteed to yield a tight bound.

Proposition 4.4.2. *Let $Z_v := ext(\Xi)$ be the set of extreme points of Ξ. Then $\Xi_w \bigcap Z_v \neq \emptyset$ and $\mathcal{P}^*(Z_v) = \mathcal{P}^*$.*

If Ξ is strictly convex, then Z_v has infinite cardinality, and $\mathcal{P}(Z_v)$ remains an infinite problem. In the particular case where Z_v is finite, *e.g.* for polyhedral uncertainty sets, it is in principle possible to enumerate the vertices of Ξ. However, $\mathcal{P}(Z_v)$ is still generally intractable.

Remark 4.4.1. *It is also possible to generate the discrete set Z by drawing finitely many samples from any distribution $\mathbb{P} \in \mathcal{C}$ supported on Ξ. The resulting scenario problem $\mathcal{P}(Z)$ yields a stochastic lower bound for $\mathcal{P}$ whose quality depends on the number of samples and the choice of $\mathbb{P}$. Scenario problems of this type have been investigated in [16, 23].*

The Critical Set Δ

Theorem 4.4.3 guarantees that pruning the uncertainty set from Ξ to any non-empty subset of Ξ_w does not alter the optimal value of $\mathcal{P}$, even though such a reduction amounts to a relaxation of the constraints in $\mathcal{P}$. However, finding such a subset is intractable.

We therefore turn to the binding uncertainty realizations in the linear decision rule problem $\mathcal{U}$ as a proxy, *i.e.*, we propose to use a finite scenario set Δ, which is derived from $\mathcal{U}$ in a similar way as Ξ_w is derived from $\mathcal{P}$. Such a set can be constructed efficiently and enjoys properties with respect to the linear decision rule problem $\mathcal{U}$ similar to those of Ξ_w with respect to the original problem $\mathcal{P}$.

Theorem 4.4.4. *There exists a non-empty set $\mathbb{D} \subseteq 2^{\Xi}$, where each $\Delta \in \mathbb{D}$ is a finite subset of Ξ with $|\Delta| \leq m+1$ and satisfies*

$$\mathcal{U}^*(\Delta) = \mathcal{U}^*.$$

Here, $\mathcal{U}(Z)$ denotes a variant of problem $\mathcal{U}$ in which Ξ is replaced by Z. One particular set $\Delta \in \mathbb{D}$ can be constructed efficiently via the solution of the dual of the cone program $\tilde{\mathcal{U}}$.

Proof. We first derive explicit an formulation of problem $\mathcal{U}(\Delta)$. Let $[\Delta]$ be the matrix whose columns are the elements of the critical set Δ. Recall that $\Delta \subset \Xi$ can have at most $m+1$ elements. Thus, the

matrix $[\Delta]$ has at most $m+1$ columns of dimension $k+1$. We can write $\mathcal{U}(\Delta)$ as

$$\begin{array}{ll} \inf & c^\top x + t \\ \text{s.t.} & d^\top Y\,[\Delta] - te_0^\top\,[\Delta] \leq 0 \\ & Axe_0^\top\,[\Delta] + BY\,[\Delta] - C\,[\Delta] \leq 0. \end{array} \qquad (\mathcal{U}(\Delta))$$

We construct the set Δ in a particular manner, such that any solution (x^*, Y^*, t^*) of $\tilde{\mathcal{U}}$ will also be optimal in $\mathcal{U}(\Delta)$, yielding the relation $\mathcal{U}^*(\Delta) = \tilde{\mathcal{U}}^* = \mathcal{U}^*$.

KKT conditions for $\tilde{\mathcal{U}}$ Consider the Lagrangian function of $\tilde{\mathcal{U}}$ described in (4.2), where $\lambda \in \mathbb{R}^{k+1}$ and $\Lambda \in \mathbb{R}^{(k+1)\times m}$ are the Lagrange multipliers for the conic inequality constraints of $\tilde{\mathcal{U}}$, with $(\lambda, \Lambda) \succeq_{\mathcal{K}} 0$. From the Lagrangian, we derive the optimality conditions

$$\begin{array}{rl} \text{Primal Feasibility:} & \begin{cases} (d^\top Y - te_0^\top)^\top \preceq_{\mathcal{K}^*} 0 \\ (Axe_0^\top + BY - C)^\top \preceq_{\mathcal{K}^*} 0 \end{cases} \\ \text{Dual Feasibility:} & \begin{cases} 1 - e_0^\top \lambda = 0, \\ \Lambda B + \lambda d^\top = 0 \\ e_0^\top \Lambda A + c^\top = 0 \\ (\lambda, \Lambda) \succeq_{\mathcal{K}} 0 \end{cases} \\ \text{Complementarity:} & \begin{cases} \left(d^\top Y - te_0^\top\right)\lambda = 0 \\ \operatorname{tr}(\left(Axe_0^\top + BY - C\right)\Lambda) = 0. \end{cases} \end{array} \qquad (4.5)$$

Let $(x^*, Y^*, t^*; \lambda^*, \Lambda^*)$ be a KKT point of $\tilde{\mathcal{U}}$. We can obtain (λ^*, Λ^*) by solving the dual of $\tilde{\mathcal{U}}$. From the KKT conditions we have that $(\lambda^*, \Lambda^*) \succeq_{\mathcal{K}} 0$ and $\lambda^* \in \Xi$. For each column $\Lambda^*_{(i)}$, $i = 1, \ldots, m$, there are two possibilities:

1. $\Lambda^*_{(i)} \neq 0$. As $\Lambda^*_{(i)} \neq 0$ and $\Lambda^*_{(i)} \in K$, then $e_0^\top \Lambda^*_{(i)} > 0$ (since Ξ is compact). Thus, there exists a positive scalar s_i so that $\xi_i = s_i \Lambda^*_{(i)} \in \Xi$. Note that ξ_i makes the i-th constraint binding, *i.e.*, $A^{(i)}x^* + B^{(i)}Y^*\xi_i = C^{(i)}\xi_i$.

2. $\Lambda^*_{(i)} = 0$. As $\Lambda^*_{(i)}$ is vanishing, the i-th constraint does not bind the optimal LDR and can be omitted from (4.2) without affecting the remaining KKT point.

Furthermore, $\xi_0 = \lambda^* \in \Xi$ is a worst-case realization for the optimal LDR, meaning that $d^\top Y^* \xi_0 = t^*$. We can now construct the critical set as $\Delta = \{\xi_i : i \in I\} \bigcup \{\xi_0\}$, where I is the index set of the nonzero columns of Λ^*.

KKT conditions for $\mathcal{U}(\Delta)$ Let $v \in \mathbb{R}^{m+1}$ and $V \in \mathbb{R}^{(m+1)\times m}$ be the dual multipliers for the linear inequality constraints of $\mathcal{U}(\Delta)$. The KKT conditions of $\mathcal{U}(\Delta)$ are:

$$\begin{aligned}
&\text{Primal Feasibility:} && \begin{cases} d^\top Y\,[\Delta] - t e_0^\top\,[\Delta] \leq 0 \\ A x e_0^\top\,[\Delta] + BY\,[\Delta] - C\,[\Delta] \leq 0 \end{cases} \\
&\text{Dual Feasibility:} && \begin{cases} 1 - e_0^\top\,[\Delta]\, v = 0 \\ [\Delta]\, VB + [\Delta]\, v d^\top = 0 \\ e_0^\top\,[\Delta]\, VA + c^\top = 0 \\ (v, V) \geq 0 \end{cases} \\
&\text{Complementarity:} && \begin{cases} \left(d^\top Y - t e_0^\top\right)[\Delta]\, v = 0 \\ \operatorname{tr}\left(\left(A x e_0^\top + BY - C\right)[\Delta]\, V\right) = 0. \end{cases}
\end{aligned} \tag{4.6}$$

When Δ is defined as before, it is easy to show that there exists $(v^*, V^*) \geq 0$, so that $[\Delta]\, v^* = \lambda^*$ and $[\Delta]\, V^* = \Lambda^*$. As a result, the dual feasibility and complementarity conditions in (4.6) become identical to those in (4.5). Furthermore, any solution of $\tilde{\mathcal{U}}$ will satisfy the primal feasibility conditions in (4.6) since $\mathcal{U}(\Delta)$ is a relaxation of $\tilde{\mathcal{U}}$. Thus, the optimal solution (x^*, Y^*, t^*) of $\tilde{\mathcal{U}}$ is also optimal in $\mathcal{U}(\Delta)$. □

We therefore propose to use $\mathcal{P}^*(\Delta)$ for some $\Delta \in \mathbb{D}$ as a lower bound on $\mathcal{P}^*$. The motivation for this choice is predicated on the reasonable expectation that there may exist at least one worst-case uncertainty realization for $\mathcal{P}$ that is also a worst-case realization for $\mathcal{U}$, *i.e.*, that there exists $\Delta \in \mathbb{D}$ such that $\Delta \bigcap \Xi_w \neq \emptyset$.

In cases where the above assumption holds and $\Delta \in \mathbb{D}$ is chosen correctly, $\mathcal{P}^*(\Delta)$ will be equal to the true optimal value of $\mathcal{P}$, and the optimality gap $\mathcal{U}^* - \mathcal{P}^*(\Delta)$ will be an exact characterization of the suboptimality of linear decision rules. On the other hand, our approach can return a conservative optimality gap for one of two reasons:

i) There does not exist any $\Delta \in \mathbb{D}$ with $\Delta \bigcap \Xi_w \neq \emptyset$. Such a situation is illustrated in Example

4.5.1.

ii) There does exist some $\Delta \in \mathbb{D}$ with $\Delta \bigcap \Xi_w \neq \emptyset$, but $|\mathbb{D}| \geq 2$ and there is no mechanism of selecting the right one. In the numerical examples of 4.5.2, we encountered several problem instances for which $\mathcal{P}^*(\Delta) < \mathcal{U}^*$ when in fact $\mathcal{P}^* = \mathcal{U}^*$.

Despite this limitation, we have found the approach to perform well in practice and to reliably identify problem instances for which linear decision rules are optimal.

Furthermore, properties of a given set $\Delta \in \mathbb{D}$ can in certain circumstances be used to identify immediately that linear decision rules are optimal without computing a lower bound.

Theorem 4.4.5 (Optimality of linear decision rules)**.** *If there exists a $\Delta \in \mathbb{D}$ whose elements are linearly independent, then linear decision rules are optimal for the associated instance of $\mathcal{P}$, and the optimality gap derived from such a Δ is zero,* i.e.*,*

$$\mathcal{P}^*(\Delta) = \mathcal{P}^* = \mathcal{U}^*(\Delta) = \mathcal{U}^*.$$

Proof. The rightmost equality follows from Theorem 4.4.4. The remainder of the proof relies on the fact that linear decision rules are optimal for simplicial uncertainty sets [17]. If there exists a Δ with linearly independent elements, then taking the convex hull of Δ will produce a (maybe degenerate) simplex. Let Δ_c be the convex hull of Δ. As Δ contains all the extreme points of Δ_c, we have $\mathcal{U}^*(\Delta_c) = \mathcal{U}^*(\Delta)$ and similarly $\mathcal{P}^*(\Delta_c) = \mathcal{P}^*(\Delta)$. Furthermore, since linear decision rules are optimal for Δ_c, it follows that $\mathcal{U}^*(\Delta_c) = \mathcal{P}^*(\Delta_c)$. □

Remark 4.4.2. *Theorems 4.4.4 and 4.4.5 together provide a convenient method for establishing the optimality of linear decision rules for a specific instance of $\mathcal{P}$ without explicitly calculating a lower bound. One need only obtain the set Δ via the solution of the dual of $\tilde{\mathcal{U}}$ and check whether its elements are linearly independent.*

Corollary 4.4.1. *Let $\mathbb{P}_\Delta$ be any probability distribution supported on Δ. Under the assumptions of Theorem 4.4.5, we find*

$$\mathcal{L}^*(\mathbb{P}_\Delta) = \mathcal{P}^*(\Delta) = \mathcal{P}^* = \mathcal{U}^*(\Delta).$$

Thus, the lower bound $\mathcal{L}^*(\mathbb{P}_\Delta)$ also certifies the optimality of linear decision rules in these situations.

Efficient Identification of Worst-Case Scenarios

Despite the general intractability result of Theorem 4.4.1, it is sometimes possible to compute an element of Ξ_w in an efficient manner.

Theorem 4.4.6. *Suppose that* $\ker(B^\top) \subseteq \ker(C^\top)$. *Then,*

$$\Xi_w = \arg\max_{\xi \in \Xi} \ -\mu^\top C\xi, \tag{4.7}$$

where μ *is any second-stage dual feasible variable for problem* $\mathcal{P}$ *satisfying the dual feasibility conditions* $\mu \geq 0$ *and* $B^\top \mu = -d$.

Proof. Consider the optimal second-stage cost $Q(x, \xi)$ defined in (4.3). By linear programming duality, we obtain

$$\begin{aligned} Q(x,\xi) \ = \ \mu^\top(Ax - C\xi) + \max_{p} \ & p^\top(Ax - C\xi) \\ \text{s.t.} \quad & B^\top p = 0 \\ & p \geq -\mu. \end{aligned}$$

Whenever the kernel condition holds, the maximization term in the above expression is independent of ξ. Recalling the definition (4.4) of Ξ_w we have

$$\Xi_w = \arg\max_{\xi \in \Xi} \ Q(x^*, \xi) = \arg\max_{\xi \in \Xi} \ \mu^\top C\xi,$$

and thus an element of Ξ_w can be found by solving a convex optimization problem. □

Similar results can be obtained whereby alternative assumptions to those in Theorem 4.4.6 ensure that solutions to (4.7) are contained in Ξ_w, and consequently that $\mathcal{P}$ is efficiently solvable. For example one can replace the assumption that $\ker(B^\top) \subseteq \ker(C^\top)$ with the assumption that both $A = 0$ and $C^\top \ker(B^\top) \subseteq \mathcal{K}^*$; in this case the maximization term vanishes for all possible x and $\xi \in \Xi$, and thus Theorem 4.4.6 still holds. Such a situation occurs in Example 4.5.1.

4.5 Numerical Results

Throughout this section, we compute lower bounds on $\mathcal{P}$ by solving $\mathcal{P}(Z)$ where: $Z = \Delta$ is the set defined in Theorem 4.4.4; $Z = \{\xi_m\}$ is a single scenario set, where ξ_m is chosen via a solution of (4.7); and $Z = Z_v$ is the set of vertices of the support.[1] For completeness, we also provide a set of dual linear decision rule bounds $\mathcal{L}^*(\mathbb{P})$, parametrized by three different uncertainty distributions: $\mathbb{P} = \mathbb{P}_\Xi$ is the uniform distribution on the support Ξ; $\mathbb{P} = \mathbb{P}_\Delta$ is the uniform distribution on the set Δ; and $\mathbb{P} = \mathbb{P}_{Z_v}$ is the uniform distribution on the set of vertices Z_v of the support. The estimated optimality gaps provided by each bound are then calculated as the difference between $\tilde{\mathcal{U}}^*$ and the different lower bounds described above, whilst the percentage gaps correspond to that difference divided by the respective lower bounds.

4.5.1 Temporal Networks Example

We investigate a specific instance of a temporal network problem described in [89], for which linear decision rules are known to be suboptimal.

$$\begin{array}{ll} \inf & \sup\limits_{\xi\in\Xi}\ y_2(\xi) \\ \text{s.t.} & \left.\begin{array}{l} y_1(\xi) \geq \max(\xi_1, 1-\xi_1) \\ y_2(\xi) \geq y_1(\xi) + \max(\xi_2, 1-\xi_2) \end{array}\right\} \forall \xi \in \Xi \end{array} \tag{4.8}$$

There are two fully adaptive one-dimensional decisions y_1 and y_2 but no design decisions. The uncertainty set is given by $\Xi := \left\{\xi \in \mathbb{R}^3 \ : \ \xi_1 = 1, (\xi_1 - \frac{1}{2})^2 + (\xi_2 - \frac{1}{2})^2 \leq (\frac{1}{2})^2\right\}$. The optimal value of (4.8) can be calculated analytically [89]. Furthermore, one can find a worst-case scenario in Ξ_w by solving (4.7). The results for the various lower bounds are reported in Table 4.1.

4.5.2 Randomly Generated Instances

We assess the performance of the different lower bounds based on a sample of randomly generated instances of $\mathcal{P}$ with fixed dimensions $k = 16$, $m = 16$, $n_1 = 3$ and $n_2 = 5$. The instances are generated

[1] Recall from Proposition 4.4.2 that $\mathcal{P}^*(Z_v) = \mathcal{P}^*$, so that $\mathcal{U}^* - \mathcal{P}^*(Z_v)$ measures the true degree of suboptimality of linear decision rules.

Problem	Opt. Value	Gap	% Gap
$\tilde{\mathcal{U}}$	2.00	-	-
$\mathcal{P}$	1.71	0.29	16%
$\mathcal{P}(\{\xi_m\})$	1.71	0.29	16%
$\mathcal{P}(\Delta)$	1.50	0.50	33%
$\mathcal{L}(\mathbb{P}_\Delta)$	1.40	0.60	43%
$\mathcal{L}(\mathbb{P}_\Xi)$	1.25	0.75	60%

Table 4.1: Bound comparison for temporal network example.

to adhere to the assumptions described in [17], where it has been shown that $\mathcal{U}^*/\mathcal{P}^*$ is bounded below by $\Omega(\sqrt{m}) = 4$ in the worst case. This means that linear decision rules can incur a percentage gap of up to 300%. The problem instances are generated as follows:

- A constant symmetric uncertainty set Ξ is defined as $\Xi := \left\{(1,\xi)^\top \in \mathbb{R}^{k+1} \;:\; \|\xi\|_p \leq 1\right\}$ for $p \in \{1, 2, \infty\}$.
- The elements of the matrices A and B are uniformly distributed on the interval $[-5, 5]$.
- The matrix C is randomly generated so that each row of C is in the dual-cone $\mathcal{K}^*$. This guarantees that $C\xi \geq 0$ for all $\xi \in \Xi$.
- A dual vector $\mu \in \mathbb{R}^m$ is randomly generated according to a uniform distribution on $[0, 1]^m$. The vector μ is used to generate the cost vectors $c = -A^\top\mu$ and $d = -B^\top\mu$, ensuring that $c \geq 0$ and $d \geq 0$. The non-negativity restrictions are necessitated by the assumptions in [17]. The existence of a vector μ satisfying the above is required by the postulated dual feasibility of the original problem $\mathcal{P}$.

We use this procedure to generate 1000 instances of $\mathcal{P}$ for each $p \in \{1, 2, \infty\}$. The statistics for instances with box ($p = \infty$) and diamond ($p = 1$) uncertainty sets are reported in Table 4.2. In these cases, the vertices of Ξ can be enumerated. Thus, we can count how often any particular bound coincides with the true optimal value of $\mathcal{P}$ (tight instances), and how often the bound detects the optimality of linear decision rules among all instances for which linear decision rules are optimal (opt. detection). The results for the spherical uncertainty set ($p = 2$) are reported in Table 4.3.

Problem	Average % Gap	% Tight Inst.	% Opt. Detections
$\mathcal{P}(Z_v)$	7.05	100	100
$\mathcal{P}(\Delta)$	10.27	49.75	99.6
$\mathcal{P}(\{\xi_w\})$	27.21	2.15	4.317
$\mathcal{L}(\mathbb{P}_{Z_v})$	39.39	0	0
$\mathcal{L}(\mathbb{P}_{\Delta})$	22.77	3.95	8.032
$\mathcal{L}(\mathbb{P}_{\Xi})$	45.61	0	0

Table 4.2: Bound comparison for polyhedral uncertainty sets.

Problem:	$\mathcal{P}(\Delta)$	$\mathcal{P}(\{\xi_m\})$	$\mathcal{L}(\mathbb{P}_{\Delta})$	$\mathcal{L}(\mathbb{P}_{\Xi})$
% Gap:	18.76	34.43	37.17	47.73

Table 4.3: Bound comparison for spherical uncertainty sets.

4.6 Conclusion

We study the performance of linear decision rules in two-stage robust optimization problems with a linear min-max objective and linear robust constraints. Calculating the optimal decision for these problems is typically intractable and a popular measure to overcome this intractability is to restrict the decision rule to be linear in the uncertainty. However, for a problem with m constraints, linear decision rules have been shown to suffer a loss of performance of the order $\Omega(m)$ in the worst case. This worst-case a priori bound is too pessimistic for the majority of problem instances. For example, linear decision rules have been shown to be optimal in problem instances with simplicial uncertainty sets. In this chapter, we develop an algorithm that provides a problem-specific bound on the performance loss of linear decision rules. The bound is based on a scenario selection technique, where we use the Lagrange multipliers associated with the computation of the linear decision rule to construct a finite subset of the uncertain parameters' realizations. We then calculate the performance of the optimal decision rule when the uncertainty set is restricted to this finite subset, which in turn provides a progressive estimate on the performance of the optimal decision in the original problem. This estimate can then be used to bound the performance loss of the linear decision rule. We demonstrate that this methodology outperforms existing bounds in numerical instances, and can be used to derive an alternative proof for the optimality of decision rules in cases where the uncertainty is simplicial.

Chapter 5

Ambiguous Utilities in Two-Player Games

Game theory is a fundamental tool in the study of decision making under uncertainty. It is used to model situations where the payoff of each agent depends not only on her own action, but also on the actions of the other players. In a complete information setting, all the parameters of the game, such as each player's possible actions and payoffs, are common knowledge. Uncertainty is then only introduced by the non-cooperative nature of the game, whereby players are unable to communicate and pre-commit to specific actions, and therefore play mixed (randomized) strategies. In this situation, every player is uncertain about her opponents' strategies, and thus is uncertain about her own payoff. In his famous proof [72], Nash showed that if all players are maximizing their expected utility, then they can predict the outcome of the game as an equilibrium in mixed strategies, from which none of the players has an incentive to unilaterally deviate. Consequently, game theory and the Nash equilibrium concept is built on expected utility theory, where the players' payoffs are assumed to be expressed in terms of utilities. Unfortunately, in most real-life situations, the players' utilities are either unavailable or hard to measure precisely, and instead the payoffs are specified in monetary units. The traditional approach in these scenarios is to treat the monetary payoffs as utility values, which in turn implies that the utility-maximizing players are risk-neutral. In this chapter, we make an explicit distinction between utility values, which capture the players' risk preferences, and their monetary payoffs.

We consider simultaneous, one shot, two-player, two-action games where the players know about their own and their opponent's available actions and monetary payoffs, as well as their own utility function, but only posses limited information about their opponent's utility function. For example, players

may only know that they are facing an opponent having non-satiated preferences, meaning that the opponent can have any utility function that is increasing in the payoffs; or they may know that the opponent is both greedy and risk-averse, meaning that she can have any utility function that is concave as well as increasing in the payoffs. Such games arise in a number of social and economic situations. Examples include agent-based network systems, where computer agents compete for bandwidth or other shared resources. The agents' risk attributes depend on their quality-of-service requirements, with video streaming agents being more risk-averse than those that require bandwidth in order to send an email. Another example is on-line poker where the interactions between players are both impersonal and anonymous, depriving the players of any cues, initially at least, according to which they can infer their opponents' risk attributes. Finally, by considering games where players may have any increasing utility function, we can model situations in which the games are described in terms of ordinal preference relations, where each player is known to prefer one outcome of the game over another, but no specific (monetary or utility) value is associated with each of those outcomes. This is true in a number of social and behavioural experiments where the payoffs are not necessarily expressed in terms of monetary rewards.

The resulting models are games of incomplete private information, where each player assumes that the opponent can have any utility function from a specific infinite set. We show that these games can be reformulated as games of incomplete private information with no common prior, where each player knows that the opponent's utility payoffs belong to a polyhedral set, which we can construct from the opponent's set of potential utility functions and her common knowledge monetary payoffs. To the best of our knowledge, there is no algorithm that can solve such games, even for the two-player two-action case. In this chapter, we develop a solution algorithm for games with this kind of infinite private information. We build on the robust equilibrium concept proposed by Aghassi and Bertsimas [1], which we demonstrate to be a suitable concept for our games. We then exploit the specific properties of our information structure and develop a methodology to find all equilibria of such games. Our ability to do so is a direct consequence of the structural characterization of each player's optimal strategies: in the presence of incomplete information about the opponent's utility function, a player will either, depending on her utility function:

1. play a pure strategy,

2. play a prudent strategy, that is, the max-min strategy that does not depend on the opponent's

behaviour,

3. or be indifferent between the above two, in which case she plays a convex combination of the two as dictated by the equilibrium.

The equilibria, which are expressed as a pair of sets containing all strategies that each player may play depending on her utility function, have a very specific structure. As a result, we can use an exhaustive approach to characterize all equilibria of the game for different sets of payoffs and utility functions.

This chapter is organized as follows. In Section 5.1 we motivate our approach by illustrating the shortcomings of treating monetary payoffs as utility values. In Section 5.2 we introduce our game's information structure where utility values are separated from monetary payoffs and we demonstrate the natural applicability of the robust equilibrium concept for such games. In Section 5.3 we derive some numerical properties of the equilibria that arise in our games, highlighting the inability of current techniques to solve this kind of games. In Section 5.4 we express the problem of finding equilibria for these games in a form that is amenable to an algorithmic solution method. In Section 5.5 we state our main results which provide the enabling mechanism for solving the problem and in Section 5.6 we use these results to design a solution algorithm that is able to enumerate all the equilibria of our game. In 5.7 we demonstrate the algorithm through numerical examples and some conclusions are drawn in Section 5.8.

5.1 Motivation

Consider the following constant-sum game. Two players decide to split £5 by playing an odds-even game. To this end, they simultaneously present their fist with either the index or both middle and index finger extended, indicating one (o) or two (t). The sum of fingers determines the winner, with an odd sum favouring the row player and an even sum favouring the column player. However, the game rewards the players more if they manage to get their desired outcome using two (t) fingers instead of

one (o). The strategic form of the game has the following payoff matrix:

	(o)	(t)
(o)	3 2	2 3
(t)	1 4	4 1

Traditionally, this game is modelled as a complete information game and solved using Nash's equilibrium concept. The game has a unique Nash equilibrium in mixed strategies given by $\{0.75, 0.5\}$, where the row player chooses to play (o) with probability 0.75 and (t) with probability 0.25, whilst the column player randomizes between the two actions with equal probability. This solution is reasonable as long as both players' aim is to maximize their expected monetary wealth, in the presence of the uncertainty solely generated by their opponent's mixed strategy. This assumption can be broken down into two separate statements:

a) the players are von Neumann-Morgenstern rational (VNM-rational),

b) the players are risk-neutral.

VNM-rationality is the brainchild of von Neumann and Morgenstern [87], and it is the main driving force behind the Nash equilibrium. VNM-rationality describes a specific way in which decision makers should behave in the presence of uncertainty, and the description is very compelling: a list of four, arguably reasonable axioms, called the VNM-axioms. Whilst the VNM-axioms are fairly unassuming, their consequences are far reaching. In fact, a VNM-rational agent is one that maximizes the expected value of a particular utility function. By assuming that players are VNM-rational, we accept that they will always play strategies that maximize their expected utilities. In this way, we can define an equilibrium as a pair of strategies that are consistent with both players' VNM-rational behaviours, in that each strategy maximizes the corresponding player's expected utility given the opponent's corresponding strategy. Given a pair of equilibrium strategies, a VNM-rational player does not have an incentive to unilaterally deviate from her equilibrium strategy. However, in our odds-even game, we are only given the monetary payoffs that the players receive at each outcome, not how much utility they associate with these payoffs. In the absence of any information regarding the players' utility func-

tions, a sensible thing is to assume that the players are risk-neutral, with a utility function that is an affine positive function in the monetary payoffs. Under that assumption, any strategy that maximizes their expected monetary wealth also maximizes their expected utility, and thus we can compute an equilibrium for VNM-rational players as in the case above.

Whilst the two assumptions appear to be equally necessary compromises for solving the game, they are actually very different. VNM-rationality is a normative assumption, that is, an axiomatic statement about how things ought to be, or in this case, how decision makers ought to behave. It is a statement we need to make in order to be able to model and study situations involving decision making under uncertainty, such as the odds-even game discussed here. On the other hand, risk-neutrality is a positive statement, that is, a potentially false assumption which we are forced to make because we have no information regarding each player's utility function. The difference is that while VNM-rationality is essential as such for the computation of a Nash equilibrium, risk-neutrality is not. We can alter our assumption regarding the players' risk attributes (and hence their utility functions), and we will still be able to compute a Nash equilibrium. In fact, risk-neutrality is a controversial assumption. It is what causes the St. Petersburg paradox, whose resolution by Bernoulli in 1738 led to the development of the expected utility (EU) theory [14], and it is atypical of the observed behaviour of humans, who are generally considered to be risk-averse [4, 14, 20, 44, 51, 77]. If we were to revisit the odds-even game with the modified assumption that the players are risk-averse with a logarithmic utility function instead of being risk-neutral, the equilibrium would change to $\{0.77, 0.61\}$. The change happens exactly because the players are VNM-rational. Even though the monetary payoffs remain the same, the utilities associated with each outcome are different to the risk-neutral case, and thus the utility-maximizing players adjust their strategies accordingly.

The logarithmic utility function, proposed by Bernoulli himself, is one of the oldest and most commonly used utility functions, and it has been shown to follow naturally from elementary assumptions on an individual's preference relation [64]. Nonetheless, a number of equally acceptable utility functions can be considered, with both theoretical and empirical evidence to support them. Examples include families of parametric functions with desirable risk properties, such as the exponential utilities which exhibit constant absolute risk aversion, or the isoelastic power utilities which model constant relative risk aversion [4, 38, 77]. Note however, that these functions require calibration, and there is no specific set of parameters that can be a-priori considered correct or representative of human behaviour [6, 31, 33].

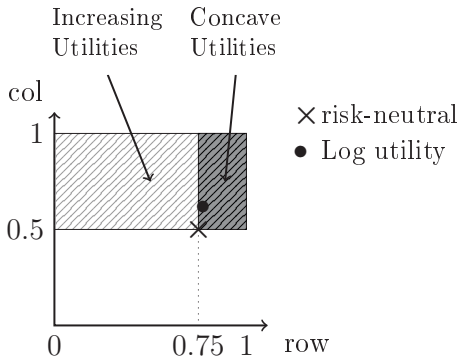


Figure 5.1: All the equilibria of the odd-even game, that can be reached given all possible combinations of increasing and concave utility functions for the two players. The x-axis shows the probability that row assigns to her first action (o) and the y-axis the probability that column assigns to her first action (o).

In the absence of empirical unanimity, there is no theoretically compelling argument why an individual would have any one utility function instead of another. The requirement of VNM-rationality merely implies the existence of utility functions that are increasing in the monetary payoffs and that are unique up to positive affine transformations. If we also accept the presumption of the diminishing marginal utility of money, then we can at most restrict our focus to risk-averse utility functions that are concave in the monetary payoffs. However, our assumptions regarding the players' utility functions have a direct impact on the equilibrium point we compute, as we demonstrated in the odds-even game when we considered the logarithmic utility. Different assumptions for the players' utility functions lead to different equilibria. Figure 5.1 shows the equilibria of our odds-evens game that can be reached simply by varying these assumptions. The outer area shows all the equilibria that are generated by all the possible combinations of increasing utility functions for the two players, while the inner shaded area shows the equilibria resulting from all possible combinations of concave utility functions.

All of the equilibria shown in Figure 5.1 are plausible because they differ only in those assumptions that were not part of the original specification of the game and for which there is no particular way to reason with. It is thus sensible to design games where players face uncertainty regarding their opponent's utility functions, even when the monetary payoffs are common knowledge. In this chapter, we examine such games of incomplete information, where players know each other's monetary payoffs, they know that they face a greedy or risk-averse opponent, but they have no other information regarding their opponent's exact utility function. We propose that players should treat this uncertainty as ambiguity and take a worst-case view when deciding their optimal strategies, as described in [1]. In the next

section, we show that this approach arises naturally when players are ambiguity-averse, and has the advantage of leading to equilibria that solve the original problem without the need for additional, unadulterated assumptions regarding the opponent's risk attributes and utility functions.

5.2 Problem Statement and Assumptions

We consider two-player two-action games. We denote by $P^i \in \mathbb{R}^{2\times 2}$, $i = 1, 2$, the monetary payoff matrix for the i-th player, where player i's strategy determines the row of her monetary payoff matrix from which her payoff is chosen, whilst the opponent's strategy determines the column. We use $-i$ to denote the opponent of the i-th player. We use

$$\Delta := \left\{ s \in \mathbb{R}^2 \ : \ s \geq 0,\, s_1 + s_2 = 1 \right\}$$

to denote the two-dimensional simplex where all mixed strategies of both players reside. If the players play strategies $s^i \in \Delta$, $i = 1, 2$, then player i receives an expected monetary payoff given by

$$\pi(s^i;\, P^i, s^{-i}) := {s^i}^\top P^i s^{-i}.$$

We make a separation between the monetary payoffs earned by players playing the game and the perceived value that each player assigns to each payoff. Whilst the players receive a common knowledge monetary payoff depending on the game's outcome, each player values the earnings differently, according to her own private utility function. The information structure of the game has the following properties:

- the monetary payoff matrices P^i, $i = 1, 2$, are common knowledge,
- each player has a type $\theta^i \in \Theta^i$ which she is fully aware of. A player's type determines the player's personal utility function. Thus, a type $\theta^i \in \Theta^i$ corresponds to a specific utility function $u^i_{\theta^i} : \mathbb{R} \mapsto \mathbb{R}$,
- a player i doesn't know her opponent's type, but she knows it belongs in a set Θ^{-i}. If player i

knows that she is facing a risk-averse opponent, then she considers a type set

$$\Theta^{-i} = \left\{\theta \ : \ u_\theta^{-i}(\cdot) \text{ any strictly increasing concave function}\right\},$$

and if she knows she is facing a greedy opponent, then she considers the type set

$$\Theta^{-i} = \left\{\theta \ : \ u_\theta^{-i}(\cdot) \text{ any strictly increasing function}\right\}.$$

Each player chooses a strategy depending on her type. Thus each player $i = 1, 2$, plays according to her behavioural function

$$b^i : \Theta^i \mapsto \Delta.$$

Given that the players play according to the behavioural functions $b^i(\cdot)$, $i = 1, 2$, and they are of types θ^i, $i = 1, 2$, then player i receives an expected utility payoff of:

$$\pi\Big[b^i(\theta^i);\, r^i\left(\theta^i\right), b^{-i}(\theta^{-i})\Big],$$

where

$$\begin{aligned} r^i &: \quad \Theta^i \mapsto \mathbb{R}^{2\times 2}, \\ r^i\left(\theta^i\right) &:= \quad u^i_{\theta^i}(P^i), \end{aligned}$$

is the component-wise application of player i's corresponding utility function $u^i_{\theta^i}(\cdot)$ to the elements of her monetary payoff matrix P^i.

This problem is an example of an incomplete private information game, where players do not know with certainty all the parameters of the game's mathematical structure. Instead, each player's personal utility function, and hence each player's utility associated with each of the game's outcomes, is only known to that player and not to the opponent.

Traditionally, incomplete private information games are modelled as Bayesian games and solved using the concept of Bayes equilibrium, introduced by Harsanyi [49] in 1967. However, Bayesian games build on Savage's subjective expected utility (SEU) theory [79] and as such depend on the availability of a unique common prior from which players can derive subjective probabilities regarding the uncertainty they face. This assumption has attracted some criticism [70], and researchers have subsequently relaxed it, most importantly with the introduction of universal type spaces large enough to accommodate

players with higher order beliefs and uncertainty regarding the common prior [69]. However, these approaches still rely on the availability of some higher-order distributional information about the uncertain parameters which may not always be available, as is the case with the games we investigate here.

In this chapter, we do not wish to introduce any (positive) assumptions regarding the availability of a common prior or any other common knowledge distributional information about the players' types. Instead, we treat this uncertainty as ambiguity and assume that players are ambiguity-averse and rational as described in Gilboa's and Schmeider's axiomatic max-min expected utility (MEU) theory [43]. Hence, we adopt the worst-case response introduced in the context of robust games [1], where each player's aim is to maximize her worst-case expected utility in the presence of the ambiguity generated by the opponent's unknown utility function. This allows us to solve the game without the need to introduce foreign parameters, such as a common prior or other distributional information, in the problem data. To keep this chapter self contained, we use the next subsection to introduce the concept of ambiguity aversion, its manifestation in MEU theory and its relation with the robust equilibrium concept.

5.2.1 Ambiguity Aversion and Robust Equilibria

Ambiguity refers to uncertainty for which players have no distributional information, as opposed to risk, which refers to uncertainty for which distributional information is available [59]. Both complete information games and Bayesian games only deal with risk. Players can use their private and common knowledge, along with the common knowledge of some form of rationality, either from EU theory or SEU theory, in order to assign objective or otherwise meaningful probabilities to their opponent's choice of actions. In our setting, players face both risk and ambiguity. Risk arises in the traditional game-theoretic sense, from the uncertainty the players face about the opponent's specific strategy, assuming that they know their opponent's exact type. Ambiguity, on the other hand, arises from the epistemic exogenous uncertainty that players face about their opponent's exact type, as this uncertainty is not accompanied by any distributional information, neither first-order nor higher-order. The need to distinguish between these two forms of uncertainty was first noticed in the economic literature by Knight [59]. It was later solidified by Ellsberg's thought experiments [35], which demonstrated that decision makers treat the two concepts differently, in a manner that violates the rationality axioms of

both the EU and SEU theory.

The famous Ellsberg paradox showed that in the presence of ambiguity, decision makers tend to adopt a worst-case approach and attempt to mitigate their losses by assuming the worst possible outcome. This max-min behaviour has since been supported by extensive evidence, both experimentally [25, 78] and theoretically [42, 43, 80]. In their MEU theory [43], Gilboa and Schmeider propose a set of axioms which encapsulate ambiguity aversion as a preference of known risks over unknown risks. Then, by modelling ambiguity as a set of multiple priors, they show that if decision makers behave according to these axioms, they will choose the actions that maximize their worst-case expected utility over that set of priors. Consider the setting where player i is given the opponent's behavioural function $b^{-i}(\cdot)$. For each of her types $\theta^{-i} \in \Theta^{-i}$, the opponent plays a mixed strategy according to her behavioural function $b^{-i}\left(\theta^{-i}\right)$. Thus, from player i's perspective, every possible evaluation of that function represents a prior. Since player i lacks any distributional information over the opponent's type and thus over the priors, if i is rational as per Gilboa's and Schmeider's axioms, she will take a worst-case approach over all of the possible priors, or equivalently, over all of the opponent's possible types. Thus, if players are ambiguity-averse and rational as per Gilboa and Schmeider, then player i, given her opponent's behavioural function $b^{-i}(\cdot)$ and for a specific type $\theta^i \in \Theta^i$, she will choose a strategy from the set (denoted the best-response strategy-set):

$$\begin{aligned} \mathcal{S}^i\left[\theta^i\,;\, b^{-i}(\cdot)\right] :=& \quad \arg\max_{s^i \in \Delta} \quad \min_{\theta^{-i} \in \Theta^{-i}} \quad \pi\left[s^i;\, r^i\left(\theta^i\right),\, b^{-i}\left(\theta^{-i}\right)\right] \\ \equiv& \quad \arg\max_{s^i \in \Delta} \quad \min_{s^{-i} \in \text{range}(b^{-i})} \quad \pi\left[s^i;\, r^i\left(\theta^i\right),\, s^{-i}\right], \end{aligned} \tag{BS}$$

where $\text{range}\left(b^{-i}\right)$ denotes the range of the opponent's behavioural function $b^{-i}(\cdot)$, which can also be perceived as the set of multiple priors. Then, similar to the Bayesian case, an equilibrium is defined as a set of behavioural functions, one for each player, so that, each player's behavioural function returns a strategy that maximizes the player's worst-case expected utility for each of her types, given the opponent's corresponding behavioural function. Formally, an equilibrium is given as a pair of behavioural functions $b^i(\cdot)$, $i = 1, 2$, satisfying

$$b^i(\theta^i) \in \mathcal{S}^i\left[\theta^i\,;\, b^{-i}(\cdot)\right] \quad \forall \theta^i \in \Theta^i, \quad i = 1, 2. \tag{EQ}$$

5.3 Properties of Robust Equilibria

5.3.1 Discontinuity in Best-response Strategies

Even the slightest ambiguity about each player's utility functions can have a dramatic effect on the game's nature and on the equilibrium strategies. Strategies that are at equilibrium in a complete information game become suboptimal as soon as we introduce ambiguity in the utility functions of each player. This is demonstrated in the following example.

Example 5.3.1 (Risky-safe game: effects of ambiguity on the equilibrium strategies)**.** *Consider a game with the following monetary payoff matrix:*

	(r)	*(s)*
	0	3
(r)	10	0
	10	3
(s)	3	3

*Each player has a risky strategy (r), yielding payoffs of either £*0 *or £*10*, and a safe strategy (s) which guarantees a payoff of £*3*. The game however is not symmetric, because the row player, denoted as Player* 1*, obtains her maximum payoff of £*10 *when both her and her opponent choose the risky strategy (r), whilst the column player, denoted as Player* 2*, obtains her maximum payoff when she chooses her risky strategy (r) whilst her opponent chooses the safe strategy (s). Consider the case where both players have a common knowledge concave utility function* $u(\cdot)$*, where* $u(0) = 0$*,* $u(3) = 5$ *and* $u(10) = 10$ *(shown as the dashed line in Figure 5.2). This is a complete information game which has a unique Nash equilibrium given by* $\{0.5, 0.5\}$*, where both players randomize between their two actions with equal probability. On expectation, both players receive an expected utility of* 5*. As a result, neither player has a preference of whether to play the game as the row or the column player. Now, consider the case where players don't know their opponent's exact utility function. Instead, they know that the opponent has any utility function where, as before,* $u(0) = 0$ *and* $u(10) = 10$*, but where* $u(3)$ *can now be any value in the set* $5 + [-\epsilon, +\epsilon]$ *for some* $\epsilon > 0$ *(shown as the set within the solid lines in Figure 5.2). Such uncertainty can be generated in instances where the players know the opponent's risk premiums up to a specific accuracy.*

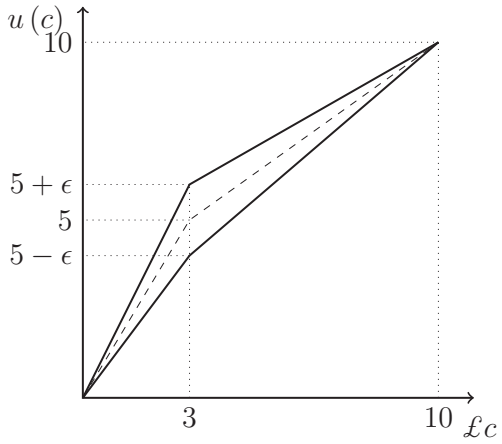


Figure 5.2: The nominal utility function (shown as the dashed line) for the risky-safe game. The two solid lines represent the outer envelope of all possible utility functions of each player, given the ambiguity about the utility they assign to £3.

A player knows her own utility function but not that of the opponent. A player's type contains information only about her utility function, and specifically the utility she assigns to her safe monetary payoff of £3. It is thus easy to verify that there is a one-to-one relationship between each player's type set Θ^i *and the interval* $[-\epsilon, +\epsilon]$*. Hence, for* $\theta^1, \theta^2 \in [-\epsilon, +\epsilon]$*, the corresponding utility payoff matrices for the two players are given by*

$$r^1\left(\theta^1\right) = \begin{bmatrix} 10 & 0 \\ 5+\theta^1 & 5+\theta^1 \end{bmatrix} \quad \textit{and} \quad r^2\left(\theta^2\right) = \begin{bmatrix} 0 & 10 \\ 5+\theta^2 & 5+\theta^2 \end{bmatrix}.$$

Remark 5.3.1. *The original equilibrium is no longer an equilibrium of this game. To see this, consider the case where Player* 1 *has a type* $\theta^1 > 0$*. Then, if Player* 2 *plays her complete information equilibrium strategy of* $\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$*, Player* 1*'s best-response will be her safe action (s) instead of her own equilibrium strategy of* $\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$*, and if Player* 1*'s action is (s) then Player* 2 *should respond with (r) etc.*

We can compute the robust equilibrium using the techniques of this chapter, and we obtain a pair of behavioural functions that represent the unique equilibrium of this game. Player 1*'s behavioural function at the equilibrium is given by:*

$$b^1\left(\theta^1\right) = \begin{cases} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv (s), & \theta^1 \in (-\epsilon, +\epsilon], \\ \begin{pmatrix} 0.5 - \frac{\epsilon}{10} \\ 0.5 + \frac{\epsilon}{10} \end{pmatrix}, & \theta^1 = -\epsilon, \end{cases}$$

whilst Player 2*'s behavioural function is given by*

$$b^2\left(\theta^2\right) = \begin{cases} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv (r), & \theta^2 \in [-\epsilon, +\epsilon), \\ \begin{pmatrix} 0.5 - \frac{\epsilon}{10} \\ 0.5 + \frac{\epsilon}{10} \end{pmatrix}, & \theta^2 = +\epsilon. \end{cases}$$

Even without knowledge of the computation procedure we can still verify that these behavioural functions are at equilibrium, by checking that they satisfy the equilibrium condition (EQ) *when the best-response strategy-set is as defined in* (BS)*. Consider Player* 1*'s worst-case utility for each of her actions, given Player* 2*'s behavioural function* $b^2(\cdot)$ *described above. When Player* 1 *has a type* $\theta^1 \in [-\epsilon, +\epsilon]$*, then for each of her actions she will receive a worst-case utility of (using row-wise minimization)*

$$\begin{aligned} \min_{\theta^2 \in [-\epsilon,+\epsilon]} \pi\left[:; r^1\left(\theta^1\right), b^2\left(\theta^2\right)\right] = & \quad \min\left\{ \begin{bmatrix} 10 & 0 \\ 5+\theta^1 & 5+\theta^1 \end{bmatrix} s^2 \;:\; s^2 \in \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.5 - \frac{\epsilon}{10} \\ 0.5 + \frac{\epsilon}{10} \end{pmatrix} \right\} \right\} \\ = & \begin{cases} 5 - \epsilon, & \text{if row plays } \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ 5 + \theta^1, & \text{if row plays } \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{cases} \end{aligned}$$

It is easy to see that Player 1*'s best-response is the safe strategy (s), (or equivalently* $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$*) for all her types* $\theta^1 \in (-\epsilon, +\epsilon]$*. When Player* 1 *has a type of* $\theta^1 = -\epsilon$*, then she is indifferent between the two actions, and thus the strategy* $b^1(-\epsilon) = \begin{pmatrix} 0.5 - \frac{\epsilon}{10} \\ 0.5 + \frac{\epsilon}{10} \end{pmatrix}$ *is a valid best-response. Hence, given Player* 2*'s behavioural function, Player* 1*'s behavioural function belongs in the best-response strategy-set defined in* (BS) *for all of her possible types. A symmetric argument can be made for Player* 2*'s behavioural function. Thus, the equilibrium condition* (EQ) *holds.*

The effects of the uncertainty are dramatic. Whilst in the complete information game both players choose to randomize between their two actions with equal probability, the introduction of ambiguity forces the players to play one of their pure strategies in almost all of their types, apart from one. The results are not that surprising. An ambiguity-averse player will prefer a known risk over an unknown one. Thus, it is not unreasonable for players to prefer the safe action whose expected reward is not affected by ambiguity, instead of choosing the risky action where the expected reward fluctuates depending on the opponent's actual type. This behaviour is reflected in Player 1*'s behavioural function. Player* 1*, having no distributional information about her opponent's type, chooses her safe strategy (s)*

in almost all of her types except from the one where she is least risk-averse (corresponding to $\theta^1 = -\epsilon$). When Player 1 is least risk-averse, she is willing to randomize between the two actions. One would expect that Player 2 will exhibit a similar behaviour. However, Player 2 has the advantage that she receives her maximum payoff when the opponent plays the safe strategy. Thus, Player 2 prefers her risky strategy in almost all of her types even in the presence of ambiguity, apart from the one where she has the most risk-averse utility function (corresponding to $\theta^2 = \epsilon$). When Player 2 is most risk-averse, she is willing to assign some probability mass to the safe action (s) and not just play the pure risky action, and thus randomizes between the two.

5.3.2 Difference Between Infinite and Finite Type Games

Games with worst-case response sets have been suggested in a number of different settings. The oldest example is von Neumann's and Morgenstern's "max-min" formulation of players' behaviour in non-cooperative complete information games [87]. This approach leads to the celebrated Nash equilibria in the case of zero-sum and constant-sum games. However, the majority of worst-case responses have been used in order to model situations where players entertain ambiguity about various aspects of the game. For example, the worst-case approach was used to define equilibria in complete information games where each player has a certain degree of doubt regarding her opponents' ability or willingness to implement a specific strategy [34, 58, 65]. Similarly, the worst-case approach was used to define equilibria where players face ambiguity because the opponents have multiple best-responses to choose from [62]. The worst-case approach was used in incomplete information games as well. For instance, it was used in incomplete information games without private information, where players have epistemic uncertainty about their own payoff functions as well as their opponent's ability to implement a best-response strategy [74]. In incomplete information games with private information, the worst-case approach was used in a Bayesian setting where players face ambiguity because they have multiple common priors instead of just one [56]. Finally, for games with private information and no common prior like the ones we study here, the worst-case approach has been proposed in the context of robust games [1] where the best-response strategy-set is defined as in (BS), as well as in the context of mechanism design where the best-response strategy-set is defined using a minimax regret criterion [54]. This chapter builds on the results of [1], where it has been shown that the best-response strategy-sets defined in (BS) admit an equilibrium in behavioural functions, even in the case where the type sets Θ^i, $i = 1, 2$, are infinite

[1, Theorem 8].

The existing methodologies for solving incomplete information games with private information and no common prior utilize various optimization techniques. In [1], the problem of finding robust equilibria is reformulated as a linear complementarity problem. Similarly, in [54] the problem is reformulated as a semi-infinite mixed integer linear problem. Whilst these problems are in general intractable [52], the current state of the art solvers can readily handle two-player two-action games. However, both methods are restricted to games where players have finite type sets, as opposed to the games with infinite type sets that we study in this chapter. Fortunately, the worst-case approach allows for some complexity reduction of the original problem, where the majority of the types in the type sets Θ^i, $i = 1, 2$, can be ignored. This is in contrast to the Bayesian approach, where each player needs to account for all of her opponents' possible types according to the common prior. To this end, we make the following observation:

Observation 5.3.1. *A player i can construct her best-response strategy-set $\mathcal{S}^i[\cdot]$ defined in* (BS) *without having a full description of her opponent's behavioural function $b^{-i}(\cdot)$. Instead, player i only needs to know the extreme strategies of her opponent's behavioural function, given by* $\operatorname{ext}\operatorname{range}\left(b^{-i}\right)$.

Proof. This follows immediately from the linearity of the expected utility operator $\pi(\cdot)$:

$$\mathcal{S}^i\left[\theta^i\,;\, b^{-i}(\cdot)\right] \equiv \underset{s^i\in\Delta}{\arg\max} \min_{s^{-i}\in \operatorname{ext}\operatorname{range}(b^{-i})} \pi\left[s^i;\, r^i\left(\theta^i\right),\, s^{-i}\right], \quad \forall \theta^i \in \Theta^i.$$

□

Consequently, the players only need to know the opponent's extreme types, that is, the types that correspond to the extreme strategies of the opponent's behavioural function, given by the set[1]

$$\tilde{\Theta}^i := \left\{\tilde{\theta}^i \in \Theta^i \;:\; b^i\left(\tilde{\theta}^i\right) \in \operatorname{ext}\operatorname{range}\left(b^i\right)\right\}.$$

The following relationship then trivially holds:

$$\mathcal{S}^i\left[\theta^i\,;\, b^{-i}(\cdot)\right] \equiv \underset{s^i\in\Delta}{\arg\max} \min_{\tilde{\theta}^{-i}\in\tilde{\Theta}^{-i}} \pi\left[s^i;\, r^i\left(\theta^i\right),\, b^{-i}\left(\tilde{\theta}^{-i}\right)\right], \quad \forall \theta^i \in \Theta^i,$$

[1] This set may contain types that correspond to the same extreme strategy. For our purposes, it is sufficient to consider any set containing exactly one type for each extreme strategy. That is, we can consider any set $\tilde{\Theta}^i$ of cardinality equal to that of the set $\operatorname{ext}\operatorname{range}\left(b^i\right)$ so that $\forall \tilde{s} \in \operatorname{ext}\operatorname{range}\left(b^i\right), \exists \tilde{\theta} \in \tilde{\Theta}^i : b^i\left(\tilde{\theta}\right) = \tilde{s}$.

which implies that any two behavioural functions $b^i(\cdot)$, $i = 1, 2$, that are at equilibrium for the game with the full type sets Θ^i, $i = 1, 2$, are also at equilibrium for the game with the extreme type sets $\tilde{\Theta}^i$. In the case of two-action games where the behavioural functions map to two-dimensional simplices, there are at most two extreme strategies, thus there only need to be two types in each of the extreme type sets $\tilde{\Theta}^i$. As a result, there is at least one equilibrium for the game with these finite types sets $\tilde{\Theta}^i$ that is also an equilibrium for the game with the infinite types sets Θ^i. The conjecture is that one may be able to obtain this equilibrium by solving the finite game using the existing methods.

Finding the extreme type sets $\tilde{\Theta}^i$ is not trivial. In fact, it is as hard as solving the original problem. A naive approach would consider the types that correspond to the extremes of the range of the utility function $r^i(\cdot)$, instead of those of the behavioural function $b^i(\cdot)$. That is, instead of solving games where the type sets are $\tilde{\Theta}^i$, we solve games for the type sets $\hat{\Theta}^i$ where

$$\hat{\Theta}^i := \left\{ \hat{\theta}^i \in \Theta^i \ : \ r^i\left(\hat{\theta}^i\right) \in \operatorname{ext}\operatorname{range}\left(r^i\right) \right\}. \tag{5.1}$$

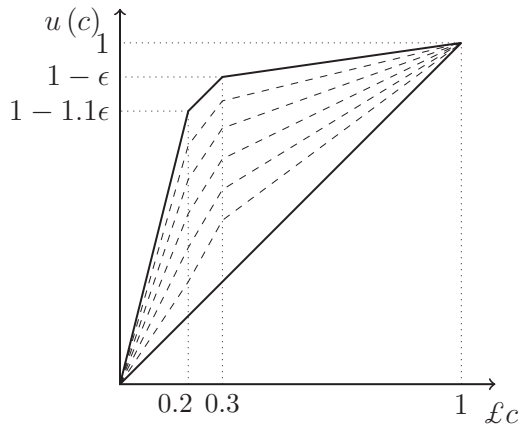
Given the specific nature of our uncertainty, we will show that these sets are easy to construct and will always be finite. Unfortunately, equilibria derived using these finite sets do not necessarily correspond to equilibria for the original game with the infinite type sets. This result is demonstrated in Example 5.3.2.

Example 5.3.2 (Differences in equilibria of the game with infinite type sets Θ^i from the equilibria of the game with the finite type sets $\hat{\Theta}^i$)**.** *Consider the following game given in strategic form with monetary payoff matrix:*

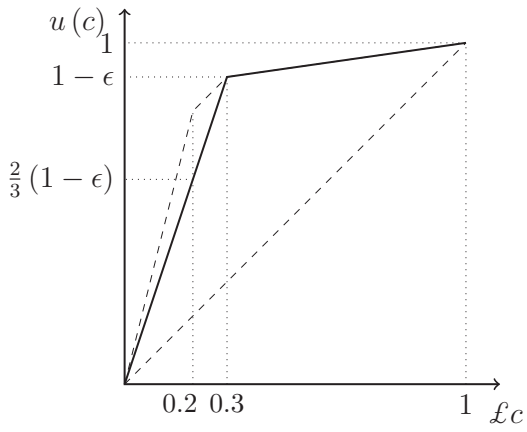
	(l)	*(r)*
	0.1	0.2
(t)	1	0.2
	1	0
(b)	0	0.3

Both players are uncertain about their opponent's utility function, but they know their opponent is greedy and risk-averse. Thus, both players solve the game where the opponent can have any ***strictly increasing***[2] *and* ***concave*** *utility function. We denote the row player as Player* 1 *and the column*

[2]For implementation reasons, a strictly increasing function is modelled as any function $u : \mathbb{R} \mapsto \mathbb{R}$ with a slope greater or equal to some solver parameter $\epsilon > 0$.



(a) The risk-neutral and the most risk-averse utility for Player 1. The two solid lines represent the outer envelope of all possible concave utility functions.



(b) The utility function that corresponds to the third extreme point in the range of the payoff function $r^1(\cdot)$, labelled as θ_3^1 in Figure 5.4a.

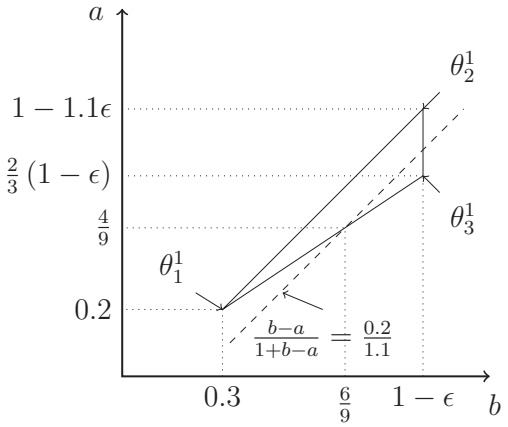
Figure 5.3: The utility functions of Player 1. The three functions indicated with solid lines in Figures 5.3a and 5.3b correspond to the three extreme points of range (r^1).

player as Player 2. *Let* Θ^i, $i = 1, 2$, *denote the players' infinite type sets encoding all increasing concave utility functions. By the equivalence of utility functions up to positive affine transformations, we only need to consider the set of all increasing concave utility functions with* $u(0) = 0$ *and* $u(1) = 1$ *(c.f. Proposition 5.4.1). The set of all utility functions for Player* 1 *is illustrated in Figure 5.3a, where all concave increasing utility functions with* $u(0) = 0$ *and* $u(1) = 1$ *lie within the two solid lines in the figure.*

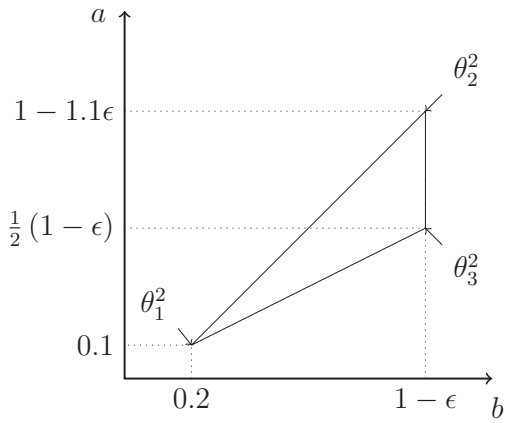
*In our setting, a player's type contains information only about the same player's utility function, and specifically, the utility values that she assigns at each of her monetary payoffs. Since the players have constant utility values at their extreme payoffs £*0 *and £*1, *their types* $\theta^i \in \Theta^i$, $i = 1, 2$, *contain information only about the utility values that the players assign to their intermediate payoffs, which for Player* 1 *are £*0.2 *and £*0.3. *Consequently, we can replace the type sets* Θ^i, $i = 1, 2$, *with the two-dimensional sets* $\mathcal{U}^i$, $i = 1, 2$, *where* $\mathcal{U}^1$ *is defined for Player* 1 *as*

$$\mathcal{U}^1 := \left\{ (a, b) \in \mathbb{R}^2 \;:\; \begin{array}{c} a = u(0.2),\, b = u(0.3), \\ u(\cdot) \text{ any concave increasing utility function} \\ \text{with } u(0) = 0,\, u(1) = 1 \end{array} \right\},$$

and $\mathcal{U}^2$ *is defined similarly for Player* 2. *It is easy to show that the sets* $\mathcal{U}^i$, $i = 1, 2$, *have the form illustrated in Figures 5.4a and 5.4b (c.f. Proposition 5.4.2).*



(a) Player 1's set $\mathcal{U}^1$ where $a = u(0.2)$ and $b = u(0.3)$.



(b) Player 2's set $\mathcal{U}^2$ where $a = u(0.1)$ and $b = u(0.2)$.

Figure 5.4: Player 1's and 2's sets $\mathcal{U}^i$, $i = 1, 2$ respectively. The three extreme points (labelled clockwise) correspond to the three extreme utility functions of each player.

We can write the players' utility matrices given by $r^i(\cdot)$, $i = 1, 2$, *as functions of the polyhedral sets* $\mathcal{U}^i$, $i = 1, 2$. *Thus, for* $\left(a^i, b^i\right) \in \mathcal{U}^i$, $i = 1, 2$, *the utility matrices are given as*

$$r^1\left(\left(a^1, b^1\right)\right) = \begin{bmatrix} 1 & a^1 \\ 0 & b^1 \end{bmatrix} \quad \textit{and} \quad r^2\left(\left(a^2, b^2\right)\right) = \begin{bmatrix} a^2 & 1 \\ b^2 & 0 \end{bmatrix}.$$

As a result, the extreme types in $\hat{\Theta}^i$ *defined in* (5.1) *correspond to the extreme points of* $\mathcal{U}^i$. *According to the discussion earlier, one would hope that the equilibrium for the game where the players consider only the three extreme points in the sets* $\mathcal{U}^i$, $i = 1, 2$, *is an equilibrium for the game where the players consider the entire sets* $\mathcal{U}^i$, $i = 1, 2$. *However, this is not the case. In order to illustrate this, we calculate the equilibrium behavioural functions for both the game with the infinite uncertainty sets (given by* $\mathcal{U}^i$, $i = 1, 2$*) as well as the finite sets (given by* $\text{ext}\{\mathcal{U}^i\}$, $i = 1, 2$*). We denote with* $b^i(\cdot)$, $i = 1, 2$, *the equilibrium functions of the game with the infinite sets, and with* $\hat{b}^i(\cdot)$, $i = 1, 2$, *the equilibrium functions of the game with the finite sets.*

The equilibrium for the infinite game is calculated using the techniques of this chapter. The behavioural functions $b^i(\cdot)$, $i = 1, 2$, *are given as functions of the sets* $\mathcal{U}^i$, $i = 1, 2$. *For Player* 1, *for any* $(a, b) \in \mathcal{U}^1$,

the equilibrium behavioural function is any function satisfying (c.f. Theorem 5.5.1):

$$b^1\left((a,b)\right) = \begin{cases} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, & \text{when } \frac{b-a}{1+b-a} < \frac{0.2}{1.1}, \\ \begin{pmatrix} x \\ 1-x \end{pmatrix},\ x \in \left[\frac{b}{1+b-a}, 1\right], & \text{when } \frac{b-a}{1+b-a} = \frac{0.2}{1.1}, \\ \frac{1}{1+b-a}\begin{pmatrix} b \\ 1-a \end{pmatrix}, & \text{otherwise.} \end{cases}$$

Note that the strategy $\frac{1}{1+b-a}\begin{pmatrix} b \\ 1-a \end{pmatrix}$ *is in-fact the max-min strategy when Player* 1 *is of type* $(a,b) \in \mathcal{U}^1$. *That is, the strategy* $\frac{1}{1+b-a}\begin{pmatrix} b \\ 1-a \end{pmatrix}$ *is the one that guarantees the highest possible payoff (equal to b) for Player* 1 *regardless of what the opponent does (c.f. Proposition 5.5.2). Thus, according to the equilibrium behavioural function* $b^1(\cdot)$, *Player* 1 *plays either the pure strategy (t), the max-min strategy* $\frac{1}{1+b-a}\begin{pmatrix} b \\ 1-a \end{pmatrix}$, *or is indifferent between any convex combination of the two. One can immediately verify that the high extreme of the range of* $b^i(\cdot)$ *(with respect to its first dimension) is* $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. *We can calculate the corresponding low extreme of the range by finding the lowest value (with respect to the first dimension) of the max-min strategy* $\frac{1}{1+b-a}\begin{pmatrix} b \\ 1-a \end{pmatrix}$ *for any* $(a,b) \in \mathcal{U}^1$ *for which Player* 1 *plays that strategy. That is, we need to find the smallest value of* $\frac{b}{1+b-a}$, *for any* $(a,b) \in \mathcal{U}^1$ *with* $\frac{b-a}{1+b-a} \geq \frac{0.2}{1.1}$. *Since the max-min strategy is quasilinear on* (a,b), *it obtains its lowest value at one of the extremes of it's feasible region. It is easy to verify that this minimum materializes at* $(a,b) = \left(\frac{6}{9}, \frac{4}{9}\right)$ *(see Figure 5.4a), resulting in a low extreme strategy of* $\frac{1}{1.1}\begin{pmatrix} 0.6 \\ 0.5 \end{pmatrix}$. *Thus the range of any behavioural function of the above form is*

$$\text{ext range}\left(b^1\right) = \left\{\frac{1}{1.1}\begin{pmatrix} 0.6 \\ 0.5 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}.$$

Similarly, Player 2's equilibrium behavioural function is uniquely defined to be the max-min strategy for all of her types:

$$b^2\left((a,b)\right) = \frac{1}{1+b-a}\begin{pmatrix} b \\ 1-a \end{pmatrix}, \quad \forall\,(a,b) \in \mathcal{U}^2.$$

Using a similar argument as above, we can find Player 2*'s range extremes by evaluating this strategy at each of her extreme types (shown as the second row in Table 5.1b). We thus obtain*

$$\text{ext range}\left(b^2\right) = \left\{\frac{1}{1.1}\begin{pmatrix} 0.2 \\ 0.9 \end{pmatrix}, \frac{1}{1+0.1\epsilon}\begin{pmatrix} 1-\epsilon \\ 1.1\epsilon \end{pmatrix}\right\}.$$

	Equilibrium strategies for Player 1 at her three extreme types in $\text{ext}\{\mathcal{U}^1\}$		
	θ_1^1	θ_2^1	θ_3^1
Finite $\hat{b}^1(\cdot)$	$\frac{1}{3-\epsilon}\begin{pmatrix} 2 \\ 1-\epsilon \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\frac{1}{4-\epsilon}\begin{pmatrix} 3-3\epsilon \\ 1+2\epsilon \end{pmatrix}$
Infinite $b^1(\cdot)$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\frac{1}{4-\epsilon}\begin{pmatrix} 3-3\epsilon \\ 1+2\epsilon \end{pmatrix}$

(a) Player 1's strategies at the extreme type sets $\text{ext}\{\mathcal{U}^1\}$. The labels of the extreme type sets correspond to the labelled points in Figure 5.4a.

	Equilibrium strategies for Player 2 at her three extreme types in $\text{ext}\{\mathcal{U}^2\}$		
	θ_1^2	θ_2^2	θ_3^2
Finite $\hat{b}^2(\cdot)$	$\frac{1}{1.1}\begin{pmatrix} 0.2 \\ 0.9 \end{pmatrix}$	$\frac{1}{1+0.1\epsilon}\begin{pmatrix} 1-\epsilon \\ 1.1\epsilon \end{pmatrix}$	$\frac{1}{1.1}\begin{pmatrix} 0.1 \\ 1 \end{pmatrix}$
Infinite $b^2(\cdot)$	$\frac{1}{1.1}\begin{pmatrix} 0.2 \\ 0.9 \end{pmatrix}$	$\frac{1}{1+0.1\epsilon}\begin{pmatrix} 1-\epsilon \\ 1.1\epsilon \end{pmatrix}$	$\frac{1}{3-\epsilon}\begin{pmatrix} 2-2\epsilon \\ 1+\epsilon \end{pmatrix}$

(b) Player 2's strategies at the extreme type sets $\text{ext}\{\mathcal{U}^2\}$. The labels of the extreme type sets correspond to the labelled points in Figure 5.4b.

Table 5.1: The strategies that each player plays at their extreme types $\text{ext}\{\mathcal{U}^i\}$, $i = 1, 2$, at both equilibrium behavioural functions $\hat{b}^i(\cdot)$, $i = 1, 2$, and $b^i(\cdot)$, $i = 1, 2$.

For the finite type case, we can calculate the equilibrium behavioural functions

$$\hat{b}^i : \text{ext}\{\mathcal{U}^i\} \mapsto \Delta,\ i = 1, 2,$$

as well as verify their uniqueness using the techniques described in [1]. Each behavioural function $\hat{b}^i(\cdot)$, $i = 1, 2$, is represented as a set of three strategies, corresponding to the three types in $\text{ext}\{\mathcal{U}^i\}$. Thus, the equilibrium behavioural functions of the finite game $\hat{b}^i(\cdot)$ are illustrated in the first row in Tables 5.1a and 5.1b for players 1 and 2 respectively. The labels of the extreme types in Tables 5.1a and 5.1b correspond to the labelled points in Figures 5.4a and 5.4b. For comparison, in the second row of Tables 5.1a and 5.1b, we show the (uniquely defined) strategies that the players play at each extreme type, under the equilibrium behavioural functions of the the infinite game $b^i(\cdot)$.

One can immediately verify that for each player i, the two behavioural functions $b^i(\cdot)$ and $\hat{b}^i(\cdot)$ map to different strategies in at least one of the extreme types. However, the differences are more significant than that, and we highlight them in the following two remarks:

Remark 5.3.2. *The equilibrium behavioural functions $b^i(\cdot)$, $i = 1, 2$, of the game with infinite sets $\mathcal{U}^i$, $i = 1, 2$, are not at equilibrium for the game with the finite sets $\text{ext}\{\mathcal{U}^i\}$, $i = 1, 2$.*

Proof. We will prove this by considering Player 1's equilibrium behavioural function $b^1(\cdot)$ only at

Player 1's extreme types, given by $\text{ext}\{\mathcal{U}^1\}$ (we use the labels of Figures 5.4a and Figures 5.4b to identify the extreme types). By taking Player 1's behavioural function only at the types in $\text{ext}\{\mathcal{U}^1\}$, we will proceed to calculate Player 2's best-response at her extreme type labelled as θ_3^2. We will then show that this best-response is unique and it is different than what Player 2 plays at that type under her equilibrium behavioural function $b^2(\cdot)$. This way, we demonstrate that $b^2(\cdot)$ is not a best-response to $b^1(\cdot)$, when we only consider the types in $\text{ext}\{\mathcal{U}^1\}$, $i = 1, 2$.

For the extreme types $\theta^1 \in \text{ext}\{\mathcal{U}^1\}$, $b^1(\cdot)$ has the following form:

$$b^1\left(\theta^1\right) = \begin{cases} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, & \text{if } \theta^1 \text{ is type } \theta_1^1 \text{ or } \theta_2^1, \\ \frac{1}{4-\epsilon}\begin{pmatrix} 3-3\epsilon \\ 1+2\epsilon \end{pmatrix}, & \text{if } \theta^1 \text{ is type } \theta_3^1. \end{cases}$$

Now, consider Player 2's utility payoff matrix at her type θ_3^2, given by $r^2\left(\theta_3^2\right) = \begin{bmatrix} \frac{1}{2}(1-\epsilon) & 1 \\ 1-\epsilon & 0 \end{bmatrix}$. We calculate Player 2's worst-case utility for each of her possible actions, given by (row-wise minimization)

$$\begin{aligned} \min_{\theta^1 \in \text{ext}\{\mathcal{U}^1\}} \pi\left[:; r^2\left(\theta_3^2\right), b^1\left(\theta^1\right)\right] = & \min\left\{ \begin{bmatrix} \frac{1}{2}(1-\epsilon) & 1 \\ 1-\epsilon & 0 \end{bmatrix} s^1 \ : \ s^1 \in \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{3-3\epsilon}{4-\epsilon} \\ \frac{1+2\epsilon}{4-\epsilon} \end{pmatrix} \right\} \right\} \\ = & \begin{cases} \frac{1+2\epsilon+\frac{3}{2}(1-\epsilon)^2}{4-\epsilon}, & \text{if Player 2 plays } \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \frac{3(1-\epsilon)^2}{4-\epsilon}, & \text{if Player 2 plays } \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{cases} \end{aligned}$$

As $\epsilon \to 0$, Player 2's best-response is $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, which is different than the strategy prescribed by the equilibrium behavioural function $b^2\left(\theta_3^2\right) = \frac{1}{3-\epsilon}\begin{pmatrix} 2-2\epsilon \\ 1+\epsilon \end{pmatrix} \underset{\epsilon \to 0}{\Longrightarrow} \frac{1}{3}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$. □

Remark 5.3.2 highlights the unsuitability of using the equilibrium behavioural functions of the infinite game to derive equilibrium behavioural functions for the finite game. However, we are more interested in the opposite case, that is, using the equilibrium behavioural functions of the finite game to derive equilibria for the infinite game. We now show that this is not possible either. According to Observation 5.3.1, player i only needs access to the opponent's range extremes, given by $\text{ext}\,\text{range}\left(b^{-i}\right)$*, in order to calculate her best-response strategy-set* $\mathcal{S}^i[\cdot]$ *over her entire type set* $\mathcal{U}^i$*. Thus, one would hope to use the range extremes provided by the equilibrium behavioural functions* $\hat{b}^i(\cdot)$*,* $i = 1, 2$*, of the finite game,*

in order to derive equilibrium behavioural functions for the infinite game. This is not the case as we will demonstrate in the following remark:

Remark 5.3.3. *The equilibrium behavioural functions $\hat{b}^i(\cdot)$, $i=1,2$, of the finite game with type sets $\text{ext}\{\mathcal{U}^i\}$, $i=1,2$, cannot be used to calculate an equilibrium for the games with the infinite sets $\mathcal{U}^i$.*

Proof. From Tables 5.1a and 5.1b, we can verify that the range extremes of the equilibrium behavioural functions $\hat{b}^i(\cdot)$, $i=1,2$, are

$$\text{ext range}\left(\hat{b}^1\right) = \left\{ \frac{1}{3-\epsilon}\begin{pmatrix} 2 \\ 1-\epsilon \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

and

$$\text{ext range}\left(\hat{b}^2\right) = \left\{ \frac{1}{1.1}\begin{pmatrix} 0.1 \\ 1 \end{pmatrix}, \frac{1}{1+(\delta-1)\epsilon}\begin{pmatrix} 1-\epsilon \\ \delta\epsilon \end{pmatrix} \right\}.$$

Observation 5.3.1 states that we can calculate correctly the best-response strategy-sets $\mathcal{S}^i[\cdot]$, for any type in $\mathcal{U}^i$, by using the range extremes of the opponent's behavioural function. For Player 1, for any $(a,b) \in \mathcal{U}^1$, given $\text{ext range}\left(\hat{b}^2\right)$, the best-response strategy-set is (c.f. Theorem 5.5.1):

$$\mathcal{S}^1\left((a,b)\,;\hat{b}^2(\cdot)\right) = \begin{cases} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}, & \text{when } \frac{b-a}{1+b-a} < \frac{0.1}{1.1}, \\ \left\{ \begin{pmatrix} x \\ 1-x \end{pmatrix} \;:\; x \in \left[\frac{b}{1+b-a}, 1\right] \right\}, & \text{when } \frac{b-a}{1+b-a} = \frac{0.1}{1.1}, \\ \left\{ \frac{1}{1+b-a}\begin{pmatrix} b \\ 1-a \end{pmatrix} \right\}, & \text{otherwise.} \end{cases}$$

Consider a function $\bar{b}^1 : \mathcal{U}^1 \mapsto \Delta$ that satisfies $\bar{b}^1(\theta) \in \mathcal{S}^1\left[\theta; \hat{b}^2(\cdot)\right]$ $\forall \theta \in \mathcal{U}^1$. We can immediately verify that any function of this form is inherently different to the behavioural function $b^1(\cdot)$ we calculated earlier, which we argue is the only form of equilibrium that exists in this game. Nevertheless, consider such a function $\bar{b}^1(\cdot)$. Using the same reasoning as before, it is easy to show that

$$\text{ext range}\left(\bar{b}^1\right) = \left\{ \frac{1}{1.1}\begin{pmatrix} 0.3 \\ 0.8 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \neq \text{ext range}\left(\hat{b}^1\right).$$

Thus, for Player 1, any function that is a best-response given Player 2's $\text{ext range}\left(\hat{b}^2\right)$, will have range extremes that are different to $\text{ext range}\left(\hat{b}^1\right)$. In fact, by using these new range extremes to calculate Player 2's best-response strategy-sets in a similar manner, we obtain a behavioural function for Player 2

with range extremes that are different to ext range $\left(\hat{b}^2\right)$. Consequently, neither of the two equilibrium behavioural functions $\hat{b}^i(\cdot)$, $i = 1, 2$, of the finite game, report the correct range extremes in order to calculate the equilibrium behavioural functions of the infinite game. □

To conclude, whilst one may be able to obtain an equilibrium of the infinite game via a solution of a finite game using the extreme types in $\tilde{\Theta}^i$, the same is not true if we use the extreme types in $\hat{\Theta}^i$. To the best of our knowledge, there are no techniques for deriving the set $\tilde{\Theta}^i$, nor for solving such games with infinite type sets. As a result, a new algorithm is developed in this chapter that can provide a solution to the problem we investigate, which we will demonstrate in the subsequent sections.

5.4 Structural Reformulation

In the previous sections, we introduced our problem and highlighted our inability to solve it using existing techniques. As a result, in the remainder of this chapter we will develop a new algorithm that can solve such games where players face infinite ambiguity regarding their opponent's utility function. We begin this endeavour by reformulating the problem so that it is amenable to a solution method. There are two difficulties that arise when trying to solve this game:

a) the type sets Θ^i, $i = 1, 2$, reside in infinite-dimensional functional spaces,

b) the equilibrium condition (EQ) involves the infinite-dimensional functionals $b^i(\cdot)$, $i = 1, 2$.

In the next two subsections, we will address both these issues separately. However, before we proceed, we state two results that trivially hold because we only consider strictly increasing utility functions:

Lemma 5.4.1 (Dominant strategies in the monetary payoff matrix P^i)**.** *Any (weakly) dominant strategy in the monetary payoff matrix P^i will remain (weakly) dominant in the utility payoff matrix $u_{\theta^i}\left(P^i\right)$ for any strictly increasing utility function $u_{\theta^i}(\cdot)$. If there is a weakly dominant strategy in P^i for some player $i \in \{1, 2\}$, then the equilibria remain the same whether we consider the complete information game with payoff matrices P^i, $i = 1, 2$, or the incomplete information game where players may have any strictly increasing (or strictly increasing concave) utility function.*

Proof. For any strictly increasing utility function $u_{\theta^i}(\cdot)$, $\theta^i \in \Theta^i$, the ordering of the elements of the utility matrix $u_{\theta^i}\left(P^i\right)$ is the same as the ordering of the elements of P^i, and hence any (weakly) dominant strategy in the monetary payoff matrix P^i will remain (weakly) dominant in the utility payoff matrix $u_{\theta^i}\left(P^i\right)$. For a two-player two-action game, let player i have a weakly dominant strategy. Player i will always choose that dominant strategy, unless the opponent always plays the pure strategy that makes i indifferent on whether to choose the dominant or the dominated strategy. Thus, for any equilibrium, at least one of the players must play the same pure strategy for all of her types. The proof then follows as in Lemma 5.4.2. □

Lemma 5.4.2 (Pure equilibria in the complete information game P^i, $i = 1, 2$)**.** *Any pure equilibrium in the complete information game with payoff matrices P^i, $i = 1, 2$, will remain an equilibrium in the incomplete information game where players may have any strictly increasing (or strictly increasing concave) utility function.*

Proof. For any strictly increasing utility function $u_{\theta^i}(\cdot)$, $\theta^i \in \Theta^i$, the ordering of the elements in $u_{\theta^i}\left(P^i\right)$ is the same as the ordering of the elements of P^i. Consider two pure strategies e^i that are at equilibrium in the complete information game. Then the following holds: $e^i \in \arg\max_{s \in \Delta} \pi\left[s;\, P^i, e^{-i}\right]$, $i = 1, 2$. This means that strategy e^i selects the largest element in the column of P^i determined by the opponent's pure strategy e^{-i}. Consequently, for any strictly increasing utility function $u_{\theta^i}(\cdot)$, the following also holds: $e^i \in \arg\max_{s \in \Delta} \pi\left[s;\, u_{\theta^i}\left(P^i\right), e^{-i}\right]$, $i = 1, 2$. Thus, assuming that the opponent plays the pure strategy e^{-i}, player i's best-response strategy-set will always include the pure strategy e^i for any of her types, and hence she can choose a behavioural function $b^i\left(\theta^i\right) = e^i$ for all $\theta^i \in \Theta^i$. By a symmetrical argument, the opponent can choose a behavioural function $b^{-i}\left(\theta^{-i}\right) = e^{-i}$ for all $\theta^{-i} \in \Theta^{-i}$, and hence the two pure strategies are also at equilibrium for the game with incomplete information with type sets Θ^i, $i = 1, 2$. □

According to Lemma 5.4.1, games where the monetary payoff matrix P^i has a weakly dominant strategy have the same equilibria whether we treat them as complete information games with payoff matrices P^i, $i = 1, 2$, or incomplete information games where the players have the type sets Θ^i we consider in this chapter. We thus narrow our focus on games where there are no weakly dominant strategies in P^i. To this end, we make the following assumption:

A3 (No weakly dominant strategies.) *There are no weakly dominant strategies in P^i, $i = 1, 2$. That is, for $P^i := \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, one of the two sets of inequalities must hold:*

- *$A > C$ and $D > B$, in which case player i is a coordination player, since she receives greater rewards when choosing the same actions as her opponent, or,*
- *$A < C$ and $D < B$, in which case player i is an anti-coordination player.*

5.4.1 The Type Sets Θ^i and the Utility Matrix Sets $\mathcal{U}^i$

We now proceed to examine in more detail the type sets Θ^i. The type $\theta^i \in \Theta^i$ of each player i is used to define her utility function $u_{\theta^i}(\cdot)$ and we assume that no other private information exists. To this end, we state the following proposition:

Proposition 5.4.1 (Type set Θ^i)**.** *Consider a game where a player is known to be risk-averse with non-satiated preferences. The smallest type set containing all of the player's possible types is given by*

$$\Theta^i := \left\{ \theta^i \ : \begin{array}{l} u_{\theta^i} : \mathbb{R} \mapsto \mathbb{R}, \\ u_{\theta^i}\left(\min\left\{P^i\right\}\right) = 0,\ u_{\theta^i}\left(\max\left\{P^i\right\}\right) = 1, \\ u_{\theta^i}(\cdot) \textit{ strictly increasing}, u_{\theta^i}(\cdot) \ \textit{ concave} \end{array} \right\},$$

where $\min\left\{P^i\right\}$ and $\max\left\{P^i\right\}$ denote the smallest and largest element of the payoff matrix P^i. If the player is not known to be risk-averse but is only known to be greedy, then we can omit the concavity requirement.

Proof. It is well known that individuals with non-satiated preferences have a utility function that is strictly increasing in the monetary payoffs. Furthermore, if the individuals are risk-averse, then their utility function is also concave in the monetary payoffs. We will now show that the requirements that all utility functions $u_{\theta^i}(\cdot)$ assign 0 and 1 to the smallest and largest monetary payoffs in P^i, ensure that we only consider uniquely defined increasing utility functions.

We are interested in utility functions that represent the utility that a rational agent assigns to a specific wealth level. Such functions are unique up to positive affine transformations [87]. Consequently, any

pair of utility functions $u(\cdot), w(\cdot)$, satisfying

$$u(c) = \frac{1}{\lambda} w(c) - \frac{t}{\lambda},$$

are equivalent for any $t \in \mathbb{R}$ and any $\lambda > 0$ and since the following set-equivalence holds

$$\arg\max_{s \in \Delta} \min_{s^{-i} \in \text{range}(b^{-i})} \pi\Big[s;\, u(P^i), s^{-i}\Big] \;\equiv\; \arg\max_{s \in \Delta} \min_{s^{-i} \in \text{range}(b^{-i})} \pi\Big[s;\, w(P^i), s^{-i}\Big],$$

players with either of the two utility functions will be indistinguishable in the sense that they will be choosing strategies from the same best-response strategy-set. Now let $w(\cdot)$ be a strictly increasing concave utility function corresponding to a risk-averse player, and assume that $w(\cdot)$ violates either $w\left(min\left\{P^i\right\}\right) = 0$ or $w\left(max\left\{P^i\right\}\right) = 1$. Let $t = w\left(min\left\{P^i\right\}\right)$ and $\lambda = w\left(max\left\{P^i\right\}\right) - w\left(min\left\{P^i\right\}\right)$. Since $w(\cdot)$ is strictly increasing and assuming P^i has more than one distinct entries, then $\lambda > 0$. Thus, $w(\cdot)$ is equivalent to its positive affine transformation $u(\cdot)$ defined above. As a result, any risk-averse player with a utility function $w(\cdot)$ can be represented with the utility function $u(\cdot)$. Since $u(\cdot)$ is strictly increasing concave with $u\left(\min\left\{P^i\right\}\right) = 0$ and $u\left(\max\left\{P^i\right\}\right) = 1$, there is already a type in Θ^i corresponding to $u(\cdot)$, and thus $w(\cdot)$. □

The set Θ^i described in Proposition 5.4.1 includes all unique strictly increasing concave utility functions that a risk-averse player may have. However, for the purposes of solving the game, the only relevant information is given by the utility values that the player assigns to the monetary payoffs in their monetary payoff matrices P^i, $i = 1, 2$. As a result, instead of considering games where player i's private information (and hence type) is encoded in the set Θ^i defined in Proposition 5.4.1, we can consider games where player i's private information belongs in the utility matrix set

$$\mathcal{U}^i = \left\{ u_{\theta^i}(P^i) \in [0,1]^{2\times 2} \;\; : \;\; \theta^i \in \Theta^i \right\}, \tag{5.2}$$

where $\mathcal{U}^i$ is a projection of the set Θ^i along the values of P^i. Thus, the behavioural functions $b^i(\cdot)$ can be expressed as functions of the utility matrices instead of the types

$$b^i : \mathcal{U}^i \mapsto \Delta.$$

Similarly, the best-response strategy-sets $\mathcal{S}^i$ can be expressed as

$$\mathcal{S}^i\left[U^i; b^{-i}(\cdot)\right] := \quad \arg\max_{s\in\Delta} \min_{s^{-i}\in \text{range}(b^{-i})} \pi\Big[s;\, U^i, s^{-i}\Big] \tag{BS}$$

and the equilibrium condition (EQ) as

$$b^i(U^i) \quad \in \quad \mathcal{S}^i\left[U^i; b^{-i}(\cdot)\right] \quad \forall U^i \in \mathcal{U}^i, \quad i = 1, 2. \tag{EQ}$$

Whilst the type sets Θ^i are difficult to construct, the utility matrix sets have a polyhedral representation. To this end, we state the following proposition:

Proposition 5.4.2 (Polyhedral representation of $\mathcal{U}^i$)**.** *Let player i be a risk-averse player with a type set Θ^i defined in Proposition 5.4.1, and let the utility matrix set $\mathcal{U}^i$ be the projection of Θ^i along P^i, as defined in* (5.2)*. The utility matrix set $\mathcal{U}^i$ has a polyhedral representation:*

$$\mathcal{U}^i \equiv \left\{ U^i \in [0,1]^{2\times 2} \;:\; \begin{array}{l} C_{eq}\left(P^i\right)\text{vec}\left(U^i\right) = (1,0)^\top, \\ C_{in}(P^i)\,\text{vec}\left(U^i\right) < 0, \\ C_{con}(P^i)\,\text{vec}\left(U^i\right) \leq 0 \end{array} \right\}.$$

where $\text{vec}\left(U^i\right)$ is a vector representation of the matrix U^i, whilst $C_{eq}(P^i)$, $C_{in}(P^i)$ and $C_{con}(P^i)$ are constraint matrices (of appropriate size) that depend on P^i.

Proof. We start with the equality constraints. According to Proposition 5.4.1, all utility functions in the set Θ^i map the smallest and the largest elements of P^i to the values 0 and 1 respectively. It is thus easy to verify that any matrix U^i in the set $\mathcal{U}^i$ defined in (5.2), must satisfy the equality condition $C_{eq}\left(P^i\right)\text{vec}\left(U^i\right) = (1,0)^\top$ for an appropriate projection matrix $C_{eq}\left(P^i\right)$. We now examine the strict-inequality constraints. These constraints are equivalent to the condition that any matrix $U \in \mathcal{U}^i$ is a strictly increasing transformation of the payoff matrix P^i. This follows by the definition of a strictly increasing function $u\,(\cdot)$, where

$$x > y \iff u(x) > u(y).$$

For any two elements of the matrix P^i indexed by j and k, namely P^i_j and P^i_k, and for any $U^i \in \mathcal{U}^i$, we have $P^i_j < P^i_k \iff U^i_j < U^i_k$. We need to check that this relationship is true for all possible pairs

of elements in the matrices. Since there are finitely many elements in P^i, there are finitely many pairs to check. Thus, there is a matrix $C_{in}(P^i)$ so that $C_{in}(P^i)\,\mathrm{vec}\left(P^i\right) < 0$ expresses all possible binary strict inequalities that hold among the different pairs of elements in P^i. Iff P^i satisfies these strict inequalities, then any $U^i \in \mathcal{U}^i$ must also satisfy these inequalities. Finally, we examine the inequality constraints, which ensure that the matrices in $\mathcal{U}^i$ correspond to concave utility functions on P^i. The argument is very similar to the one above. A function $u(\cdot)$ is concave iff it satisfies

$$u(tx + (1-t)y) \leq tu(x) + (1-t)u(y),$$

for any $t \in [0,1]$ and for any x, y. We need to check that this relationship holds at the elements of P^i. Let $P^i_j < P^i_k < P^i_z$ be three distinct elements of P^i, with $P^i_k = t_k\, P^i_j + (1-t_k)P^i_z$ for some specific $t_k \in (0,1)$. Then, for any $U^i \in \mathcal{U}^i$, we have

$$U^i_k \leq t_k U^i_j + (1-t_k)U^i_z.$$

This relationship needs to hold for all possible distinct triplets in P^i and for the appropriate values of t. The rest of the argument parallels the reasoning in the strictly increasing case. □

5.4.2 The Extreme Strategies Equilibrium Condition

We now derive a finite equivalent of the equilibrium condition (EQ). Condition (EQ) states that two behavioural functions $b^i(\cdot)$, $i = 1, 2$, are at equilibrium iff each functional $b^i(\cdot)$ maps to strategies in the best-response strategy-set $\mathcal{S}^i\left[U^i; b^{-i}(\cdot)\right]$ defined in (BS) for every $U^i \in \mathcal{U}^i$. In this section, we will reformulate this condition on the two infinite-dimensional functionals $b^i(\cdot)$, $i = 1, 2$, as an equivalent condition on two pairs of strategies (one pair for each player) in the strategy simplex Δ.

According to Observation 5.3.1, one can construct the best-response strategy-set $\mathcal{S}^i\left[U^i; b^{-i}(\cdot)\right]$ from the extreme strategies in the range of $b^{-i}(\cdot)$, denoted by $\mathrm{ext\,range}\left(b^{-i}\right)$. Since $b^i(\cdot)$, $i = 1, 2$, maps to strategies in the two-dimensional probability simplex Δ which constitutes a line segment in $\mathbb{R}^2$, there are only two extremes in the sets $\mathrm{ext\,range}\left(b^i\right)$, $i = 1, 2$. Let $\underline{b^i}$, $\overline{b^i} \in \Delta$ represent the two strategies in $\mathrm{ext\,range}\left(b^i\right)$, for $i = 1, 2$. We then denote with $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right]$ the best-response strategy-set $\mathcal{S}^i\left[U^i; b^{-i}(\cdot)\right]$ but parametrized in the two extreme strategies $\underline{b^{-i}}$, $\overline{b^{-i}}$ instead of the behavioural

function $b^{-i}(\cdot)$.

The fact that we can express the best-response strategy-sets in terms of the extreme strategies instead of the behavioural functions allows us to reformulate the equilibrium condition (EQ) for the functionals $b^{-i}(\cdot)$, $i = 1, 2$, as an equivalent condition for the extreme strategies $\underline{b^i}$, $\overline{b^i} \in \Delta$, $i = 1, 2$. To this end, we state the following theorem:

Theorem 5.4.1 (Extreme strategies equilibrium)**.** *Let $b^i(\cdot)$, $i = 1, 2$, be a pair of behavioural functions and let $\left\{\underline{b^i}, \overline{b^i}\right\} = \operatorname{ext}\operatorname{range}\left(b^i\right)$, $i = 1, 2$, be the pair of their two extreme strategies respectively. The following statements are equivalent:*

a) the behavioural functions $b^i(\cdot)$, $i = 1, 2$, satisfy the equilibrium condition (EQ)*,*

b) the two pairs of extreme strategies $\underline{b^i}$, $\overline{b^i}$, $i = 1, 2$, satisfy:

$$\left.\begin{array}{rl} \forall U^i \in \mathcal{U}^i: & \mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] \bigcap [\,\underline{b^i}, \overline{b^i}\,] \neq \emptyset, \\ \exists U_1^i, U_2^i \in \mathcal{U}^i: & U_1^i \neq U_2^i, \\ & \underline{b^i} \in \mathcal{S}^i\left[U_1^i; \underline{b^{-i}}, \overline{b^{-i}}\right], \\ & \overline{b^i} \in \mathcal{S}^i\left[U_2^i; \underline{b^{-i}}, \overline{b^{-i}}\right], \end{array}\right\} i = 1, 2.$$

For any two pairs of strategies that satisfy the above condition, one can derive two behavioural functions that satisfy the equilibrium condition (EQ)*, and vice versa.*

Proof. We will first show how one can use the two pairs of strategies $\underline{b^i}$, $\overline{b^i} \in \Delta$, $i = 1, 2$, satisfying the conditions in Theorem 5.4.1, to construct two behavioural functions $b^i(\cdot)$, $i = 1, 2$, satisfying the equilibrium condition (EQ). Let $\underline{b^i}$, $\overline{b^i} \in \Delta$, U_1^i, $U_2^i \in \mathcal{U}^i$, $i = 1, 2$, denote the strategies and utility matrices for which the conditions in the theorem are satisfied. Consider two behavioural functions $z^i : \mathcal{U}^i \mapsto \Delta$, $i = 1, 2$, defined as:

$$z^i\left(U^i\right) := \begin{cases} \underline{b^i} \text{ when } U^i = U_1^i, \\ \overline{b^i} \text{ when } U^i = U_2^i, \\ \text{ any element in } \mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] \bigcap [\,\underline{b^i}, \overline{b^i}\,] \text{ otherwise.} \end{cases} \tag{5.3}$$

For both $i = 1, 2$, the following statements are true:

a) the behavioural functions $z^i(\cdot)$ are well defined,

b) the behavioural functions $z^i(\cdot)$ have extreme strategies equal to $\underline{b^i}$, $\overline{b^i}$. That is, $\operatorname{ext\,range}\left(z^i\right) = \left\{\underline{b^i}, \overline{b^i}\right\}$,

c) for every $U \in \mathcal{U}^i$, the condition $z^i\left(U^i\right) \in \mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right]$ holds.

Thus, it is easy to see that the two behavioural functions $z^i(\cdot)$, $i = 1, 2$, satisfy the equilibrium condition (EQ). We will now proceed to show the converse implication. Let two behavioural functions $b^i(\cdot)$, $i = 1, 2$, satisfy the equilibrium condition (EQ), and let the two pairs of strategies $\underline{b^i}$, $\overline{b^i}$, $i = 1, 2$ be the extreme strategies in the range of $b^i(\cdot)$. That is, $\left\{\underline{b^i}, \overline{b^i}\right\} = \operatorname{ext\,range}\left(b^i\right)$. Then

$$\begin{aligned} \left\{\underline{b^i}, \overline{b^i}\right\} = \operatorname{ext\,range}\left(b^i\right) \quad &\Longleftrightarrow \quad \operatorname{range}\left(b^i\right) \subseteq \left[\underline{b^i}, \overline{b^i}\right] \\ &\Longleftrightarrow \quad \forall U^i \in \mathcal{U}^i \,:\, b^i\left(U^i\right) \in \left[\underline{b^i}, \overline{b^i}\right]. \end{aligned}$$

Furthermore, by definition of (EQ), we have $\forall U^i \in \mathcal{U}^i \,:\, b^i\left(U^i\right) \in \mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right]$. Combining the two, we derive the requirement $\forall U^i \in \mathcal{U}^i \,:\, \mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] \bigcap [\,\underline{b^i}, \overline{b^i}\,] \neq \emptyset$. The remaining two conditions follow immediately because $\underline{b^i}$, $\overline{b^i}$ are in the range of $b^i(\cdot)$. □

The equilibrium reformulation of Theorem 5.4.1 allows for the representation of an equilibrium not as a pair of infinite-dimensional behavioural functions, but as a pair of two strategies in the two-dimensional simplex. This reformulation is not restricted to the problem where players face uncertainty about the opponent's utility functions, as encapsulated in the sets $\mathcal{U}^i$, but it applies to any two-player two-action problems where the best-response strategy-set is defined as in (BS). However, the condition in Theorem 5.4.1 still involves an infinite number of constraints, namely one for each $U^i \in \mathcal{U}^i$ for each player. In the remainder of this chapter, we will exploit the specific properties of the matrix sets $\mathcal{U}^i$ along with specific properties of the best-response strategy-sets $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right]$, to reformulate these infinite constraints in a manner amenable to a solution, and hence derive all extreme strategies that satisfy the conditions of Theorem 5.4.1. This in turn will enable us to construct all behavioural functions that satisfy the equilibrium condition (EQ).

5.5 The Best-response Functions

The main insight of Theorem 5.4.1 is that any equilibrium of the game can be expressed as a condition on the extreme strategies of the behavioural functions, instead of the behavioural functions themselves. Thus, we have reduced the infinite-dimensional description of an equilibrium to that of a finite number of points in Δ, namely the two pairs of strategies denoted by $\underline{b^i}$, $\overline{b^i} \in \Delta$, $i = 1, 2$. For any two such pairs satisfying the conditions in Theorem 5.4.1, we can construct two behavioural functions in an ad-hoc manner that will satisfy the equilibrium condition (EQ). We will now devise a technique to find such strategies with the help of a new best-response function. In the traditional game theoretic framework, where the equilibrium is expressed as a pair of strategies (or behavioural functions), player i's best-response function returns the set containing all of player i's strategies (or behavioural functions) that unilaterally satisfy the equilibrium conditions in view of fixed strategies (or behavioural functions) of the opponent. In our case, the equilibrium is expressed as a pair of extreme strategies, thus player i's best-response function takes the opponent's extreme strategies $\underline{b^{-i}}$, $\overline{b^{-i}} \in \Delta$ as input and returns the set containing all pairs of player i's extreme strategies $\underline{b^i}$, $\overline{b^i} \in \Delta$, that unilaterally satisfy the conditions in Theorem 5.4.1 for that specific $\underline{b^{-i}}$, $\overline{b^{-i}}$. Thus, a best-response function for player i can be defined as a set-valued mapping $\mathcal{E}^i : \Delta^2 \rightrightarrows \Delta^2$ (denoted as the best-response extremes-set)

$$\mathcal{E}^i\left[\underline{b^{-i}}, \overline{b^{-i}}\right] := \left\{ \left(\underline{b^i}, \overline{b^i}\right) \in \Delta^2 \ : \ \begin{array}{l} \left(\underline{b^i}, \overline{b^i}\right) \text{ satisfies the conditions} \\ \text{for } i \text{ in Theorem 5.4.1, given } \underline{b^{-i}}, \overline{b^{-i}} \end{array} \right\}. \tag{BE}$$

An equilibrium can then be found by finding a pair of extreme strategies $b^i = \left(\underline{b^i}, \overline{b^i}\right) \in \Delta^2$, $i = 1, 2$, where

$$b^i \in \mathcal{E}^i\left[b^{-i}\right], \ i = 1, 2.$$

In this section, we will derive an expression for the best-response extremes-set defined in (BE), which we will subsequently use to calculate the equilibria of the game. However, before we proceed with our analysis, we need to introduce a one-dimensional representation of the behavioural functions, best-response strategy-sets and extreme strategies that will be used in the remainder of this chapter. The behavioural functions $b^i(\cdot)$ map utility matrices to the two-dimensional simplex Δ. Whilst the functions $b^i(\cdot)$ have a two-dimensional range, they only have one degree of freedom. To this end, let $b^i_1(\cdot) \in [0, 1]$ be the first element of $b^i(\cdot)$, corresponding to the probability that player i assigns to her

first action under $b^i(\cdot)$. For any $U^i \in \mathcal{U}^i$, we can define $b^i(\cdot)$ using $b_1^i(\cdot)$ as

$$b^i(U^i) = \begin{pmatrix} b_1^i(U^i) \\ 1 - b_1^i(U^i) \end{pmatrix}.$$

To this end, we will use $\underline{b_1^i}, \overline{b_1^i} \in [0,1]$ to denote the extreme strategies in $\operatorname{ext}\operatorname{range}\left(b_1^i\right)$, where

$$\underline{b_1^i} := \min_{U^i \in \mathcal{U}^i} b_1^i(U^i), \quad \overline{b_1^i} := \max_{U^i \in \mathcal{U}^i} b_1^i(U^i), \tag{5.4}$$

correspond to the lowest and highest probabilities that player i assigns to her first action under the behavioural function $b^i(\cdot)$. Let

$$\mathcal{L} := \left\{(\underline{b}, \overline{b}) \in [0,1]^2 \ : \ \underline{b} \leq \overline{b}\right\}$$

be a triangular subset of $\mathbb{R}^2$. Then any pair of the one-dimensional extreme strategies defined in (5.4) can be described as a single point in $\mathcal{L}$.

Finally, let a generic utility matrix $U^i \in \mathcal{U}^i$ for player i be represented as:

$$U^i = \begin{pmatrix} \text{a} & \text{b} \\ \text{c} & \text{d} \end{pmatrix} \in \mathcal{U}^i.$$

By carrying out the matrix multiplications explicitly, we can re-formulate the best-response strategy-set $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right]$ as a sub-interval of $[0,1]$:

$$\mathcal{S}_1^i\left[U^i; \underline{b_1^{-i}}, \overline{b_1^{-i}}\right] := \operatorname*{arg\,max}_{x \in [0,1]} \ \min_{y \in \{\underline{b_1^{-i}}, \overline{b_1^{-i}}\}} \ x\left[ay + b(1-y)\right] + (1-x)\left[cy + d(1-y)\right], \tag{BR}$$

where $\left(\underline{b_1^{-i}}, \overline{b_1^{-i}}\right) \in \mathcal{L}$ are the opponent's one-dimensional extreme strategies defined in (5.4). For the remainder of this chapter, any reference to the best-response strategy-set $\mathcal{S}^i[\cdot]$, the best-response extremes-set $\mathcal{E}^i[\cdot]$ and the behavioural functions $b^i(\cdot)$, will be to their one-dimensional representations respectively.

We are now in a position to present the two main results of this chapter, which are the enabling mechanisms that allow us to solve the problem. In the next subsection, we will derive an analytical expression of the best-response strategy-set $\mathcal{S}^i[\cdot]$. Then, in the last part of this section, we will use that expression to derive a formulation of the best-response extremes-set $\mathcal{E}^i[\cdot]$, which will allow for the

efficient computation of all pairs of extreme strategies $\left(b^1, b^2\right) \in \mathcal{L}^2$ satisfying $b^i \in \mathcal{E}^i\left[b^{-i}\right]$, $i = 1, 2$, and thus the efficient computation of equilibria in behavioural functions of the games discussed in this chapter.

5.5.1 The Best-response Strategy-set

In order to find the pairs $\left(b^1, b^2\right) \in \mathcal{L}^2$ satisfying $b^i \in \mathcal{E}^i\left[b^{-i}\right]$, $i = 1, 2$ and solve the game, it is crucial to understand how the best-response strategy-sets $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right]$ change with the opponent's extreme strategies $\left(\underline{b^{-i}}, \overline{b^{-i}}\right) \in \mathcal{L}$, for any $U^i \in \mathcal{U}^i$. In this subsection, we will derive an analytical expression for $\mathcal{S}^i[\cdot]$ highlighting our main result: depending on her utility matrix and the opponent's extreme strategies, a player either plays a pure strategy, the max-min strategy, or is indifferent between any convex combination of the two.

Using the convention $U^i = \begin{pmatrix} \text{a} & \text{b} \\ \text{c} & \text{d} \end{pmatrix} \in \mathcal{U}^i$, we define two behavioural functions that will play an important role in the remainder of the chapter. We set

$$\begin{aligned} g^i &: \mathcal{U}^i \mapsto [0,1], \\ m^i &: \mathcal{U}^i \mapsto [0,1], \end{aligned}$$

where

$$g^i\left(U^i\right) := \frac{d-b}{a-c+d-b},$$

and

$$m^i\left(U^i\right) := \begin{cases} 0, & \frac{d-c}{a-c+d-b} < 0, \\ 1, & \frac{d-c}{a-c+d-b} > 1, \\ \frac{d-c}{a-c+d-b}, & \text{otherwise.} \end{cases}$$

Lemma 5.5.1. *If Assumption **A3** holds, the functions $g^i(\cdot)$ and $m^i(\cdot)$ are well-defined.*

Proof. We will show that the denominator $a-b+d-c$ does not vanish for any utility matrix $U^i \in \mathcal{U}^i$. By definition of the utility matrix set $\mathcal{U}^i$, all matrices $U^i \in \mathcal{U}^i$ preserve the element-wise ordering of the monetary payoff matrix P^i. Then, by Assumption **A3**, $a-c$ and $d-b$ have the same sign and are never 0 for all the utility matrices $U^i \in \mathcal{U}^i$, and thus $a-c+d-c \neq 0$. It remains to be shown that

the function $g^i(\cdot)$ always maps to a number in $[0,1]$. It is easy to see that

$$g^i\left(U^i\right) = \frac{1}{1+\frac{a-c}{d-b}}.$$

The claim then follows from the observation that $\frac{a-c}{d-b} > 0$. □

Proposition 5.5.1 (Indifference strategy)**.** *The function $g^i\left(U^i\right)$ represents the level of the opponent's strategy for which player i is indifferent as to what to play, when she has a utility matrix $U^i \in \mathcal{U}^i$. That is, when the opponent's strategy $s^{-i} \in [0,1]$ is equal to $g^i\left(U^i\right)$, then*

$$\forall s^i \in [0,1] \,:\, \pi\left[s^i; U^i, s^{-i}\right] = \frac{ad-bc}{a-c+d-b}$$

and thus player i has no preference as to which strategy $s^i \in [0,1]$ to play.

Proof. The proof follows immediately by carrying out the multiplications. For any $s^i \in [0,1]$ we have:

$$\begin{aligned}
\pi\left[s^i; U^i, s^{-i}\right] = & \begin{pmatrix} s^i \\ 1-s^i \end{pmatrix}^\top U^i \begin{pmatrix} s^{-i} \\ 1-s^{-i} \end{pmatrix} = \begin{pmatrix} s^i \\ 1-s^i \end{pmatrix}^\top U^i \begin{pmatrix} g^i\left(U^i\right) \\ 1-g^i\left(U^i\right) \end{pmatrix} \\
= & \tfrac{ab-dc}{a-c+d-b} \begin{pmatrix} s^i \\ 1-s^i \end{pmatrix}^\top \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \tfrac{ab-dc}{a-c+d-b}.
\end{aligned}$$

□

Proposition 5.5.2 (Max-min strategy)**.** *Under assumption **A3**, the function $m^i\left(U^i\right)$ represents player i's unique max-min strategy, when she has a utility matrix $U^i \in \mathcal{U}^i$. That is, when the opponent can choose any strategy $s^{-i} \in [0,1]$, then*

$$\left\{m^i\left(U^i\right)\right\} = \arg\max_{s^i \in [0,1]} \min_{s^{-i} \in [0,1]} \pi\left[s^i; U^i, s^{-i}\right].$$

Proof. The proof follows immediately, because s^i is the maximizer of the point-wise minimization of the two linear functions (found by setting s^{-i} at the extremes of it's feasible region)

$$\pi\left[s^i; U^i, 0\right] = s^i(b-d) + d, \text{ and } \pi\left[s^i; U^i, 1\right] = s^i(a-c) + c.$$

Thus, $\frac{d-c}{a-c+d-b}$ is the value of s^i at the intersection of these lines. It is easy to verify that if that intersection lies within the feasible region $[0,1]$ of s^i, then it is the unique maximizer. If the intersection

lies on the left of the feasible region, that is $\frac{d-c}{a-c+d-b} < 0$ then $s^i = 0$ is the maximizer whilst if it's on the right, that is $\frac{d-c}{a-c+d-b} > 1$, then $s^i = 1$ is the maximizer. □

We now state our main theorem:

Theorem 5.5.1 (Best-response strategy-set)**.** *When player i is a coordination player as described in Assumption **A3**, the best-response strategy-set $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] \subseteq [0,1]$ is*

$$\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \begin{cases} \{1\}, & \text{if } g^i\left(U^i\right) < \underline{b^{-i}}, \\ \left[m^i\left(U^i\right), 1\right], & \text{if } g^i\left(U^i\right) = \underline{b^{-i}}, \\ \{m^i\left(U^i\right)\}, & \text{if } \underline{b^{-i}} < g^i\left(U^i\right) < \overline{b^{-i}}, \\ \left[0, m^i\left(U^i\right)\right], & \text{if } g^i\left(U^i\right) = \overline{b^{-i}}, \\ \{0\}, & \text{if } \overline{b^{-i}} < g^i\left(U^i\right). \end{cases}$$

*When player i is an anti-coordination player as described in Assumption **A3**, the best-response strategy-set $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] \subseteq [0,1]$ is*

$$\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \begin{cases} \{0\}, & \text{if } g^i\left(U^i\right) < \underline{b^{-i}}, \\ \left[0, m^i\left(U^i\right)\right], & \text{if } g^i\left(U^i\right) = \underline{b^{-i}}, \\ \{m^i\left(U^i\right)\}, & \text{if } \underline{b^{-i}} < g^i\left(U^i\right) < \overline{b^{-i}}, \\ \left[m^i\left(U^i\right), 1\right], & \text{if } g^i\left(U^i\right) = \overline{b^{-i}}, \\ \{1\}, & \text{if } \overline{b^{-i}} < g^i\left(U^i\right). \end{cases}$$

Before we proceed with the proof of Theorem 5.5.1, we summarize the theorem in the following remark which helps clarify the relationship between the best-response strategy-set $\mathcal{S}^i[\cdot]$ and its parameters:

Remark 5.5.1. *Depending on the opponent's extreme strategies $\left(\underline{b^{-i}}, \overline{b^{-i}}\right) \in \mathcal{L}$ and her indifference strategy $g^i\left(U^i\right) \in [0,1]$ for the specific utility matrix $U^i \in \mathcal{U}^i$, player i's best-response strategy set consists of*

a) the pure strategy (given as an element of $\{0,1\}$),

b) the max-min strategy $m^i\left(U^i\right)$ for that utility matrix, or,

c) all convex combinations of a pure and the max-min strategy.

Proof of Theorem 5.5.1. We will prove the theorem only for a coordination player. A symmetrical argument can be used to derive the best-response strategy-set for an anti-coordination player. However, for the sake of brevity, we omit the latter proof. We now show that the best-response strategy-set $\mathcal{S}^i[\cdot]$ for a coordination player is as shown in the theorem. Recall that a coordination player has a payoff matrix $P^i = \begin{pmatrix} \text{A} & \text{B} \\ \text{C} & \text{D} \end{pmatrix}$ with $A > C$ and $D > B$. For any utility matrix $U^i = \begin{pmatrix} \text{a} & \text{b} \\ \text{c} & \text{d} \end{pmatrix} \in \mathcal{U}^i$, let $\hat{m}^i\left(U^i\right) = \frac{d-c}{a-c+d-b} - m^i\left(U^i\right)$ denote the distance between the fraction $\frac{d-c}{a-c+d-b}$ and the interval $[0,1]$. Note that

$$\begin{aligned} \hat{m}^i\left(U^i\right) = 0 \quad &\text{when} \quad m^i\left(U^i\right) \in [0,1], \\ \hat{m}^i\left(U^i\right) \geq 0 \quad &\text{when} \quad m^i\left(U^i\right) = 1, \text{ and} \\ \hat{m}^i\left(U^i\right) \leq 0 \quad &\text{when} \quad m^i\left(U^i\right) = 0. \end{aligned}$$

Since player i is assumed to be a coordination player as described in Assumption **A3**, and since all matrices $U^i \in \mathcal{U}^i$ preserve the element-wise ordering of the monetary payoff matrix P^i, we have $a > c$ and $d > b$. Consequently $a - c + d - b > 0$ for all $U^i \in \mathcal{U}^i$, and thus, we find

$$\begin{aligned} \mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] &= \underset{x\in[0,1]}{\arg\max} \quad \underset{y\in\{\underline{b^{-i}}, \overline{b^{-i}}\}}{\min} \quad x\left[ay + b(1-y)\right] + (1-x)\left[cy + d(1-y)\right] \\ &= \underset{x\in[0,1]}{\arg\max} \quad \underset{y\in\{\underline{b^{-i}}, \overline{b^{-i}}\}}{\min} \quad x\left((a-c+d-b)y + b - d\right) + (c-d)y + d \\ &= \underset{x\in[0,1]}{\arg\max} \quad \underset{y\in\{\underline{b^{-i}}, \overline{b^{-i}}\}}{\min} \quad x\left(y - \tfrac{d-b}{a-c+d-b}\right) - \tfrac{d-c}{a-c+d-b}y + \tfrac{d}{a-c+d-b} \\ &= \underset{x\in[0,1]}{\arg\max} \quad \underset{y\in\{\underline{b^{-i}}, \overline{b^{-i}}\}}{\min} \quad x\left(y - g^i\left(U^i\right)\right) - \left(m^i\left(U^i\right) + \hat{m}^i\left(U^i\right)\right) y, \end{aligned}$$

where the second to last line follows by dividing all terms by the positive constant $a-c+d-b$, whilst the last line follows by removing the constant term from the objective function and using the definitions of $g^i\left(U^i\right)$ and $m^i\left(U^i\right) + \hat{m}^i\left(U^i\right)$. From the last line of the equivalence, we can immediately verify that if $y - g^i\left(U^i\right) > 0$ for all $y \in \left\{\underline{b^{-i}}, \overline{b^{-i}}\right\}$, then $x = 1$ is the unique maximizer. Since $\overline{b^{-i}} \geq \underline{b^{-i}}$ we have

$$\forall y \in \left\{\underline{b^{-i}}, \overline{b^{-i}}\right\} : y - g^i\left(U^i\right) > 0 \iff \underline{b^{-i}} > g^i\left(U^i\right),$$

and thus we have proven the case where $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \{1\}$ if $g^i\left(U^i\right) < \underline{b^{-i}}$. Similar arguments can be used to establish that $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \{0\}$ if $\overline{b^{-i}} < g^i\left(U^i\right)$.

We will now deal with the case $\underline{b^{-i}} < g^i\left(U^i\right) < \overline{b^{-i}}$, which will result in the best-response strategy-set

$\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \left\{m^i\left(U^i\right)\right\}$. Apply the variable substitution $x = m^i\left(U^i\right) + \alpha$, where $\alpha \in [0,1] - m^i\left(U^i\right)$. Note that $m^i(U^i) \in [0,1]$ for all $U^i \in \mathcal{U}^i$, and thus 0 is always in the set $[0,1] - m^i\left(U^i\right)$ for any $U^i \in \mathcal{U}^i$. Then we have

$$\begin{aligned}\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] &= \underset{x\in[0,1]}{\arg\max} \quad \min_{y\in\{\underline{b^{-i}}, \overline{b^{-i}}\}} \quad x\left(y - g^i\left(U^i\right)\right) - \left(m^i\left(U^i\right) - \hat{m}^i\left(U^i\right)\right) y \\ &= \underset{\alpha\in[0,1]-m^i(U^i)}{\arg\max} \quad \min_{y\in\{\underline{b^{-i}}, \overline{b^{-i}}\}} \quad \alpha\left(y - g^i\left(U^i\right)\right) - \hat{m}^i\left(U^i\right) y.\end{aligned}$$

For $\underline{b^{-i}} < g^i\left(U^i\right) < \overline{b^{-i}}$, we can distinguish the following three cases:

- if $\hat{m}^i\left(U^i\right) = 0$, then $\min_{y\in\{\underline{b^{-i}}, \overline{b^{-i}}\}} \alpha\left(y - g^i\left(U^i\right)\right) \leq 0$ for all $\alpha \in \mathbb{R}$, and thus $\alpha = 0$ is the unique maximizer. This corresponds to the optimal strategy $x = m^i\left(U^i\right)$.

- If $\hat{m}^i\left(U^i\right) > 0$, we can conclude that $m^i\left(U^i\right) = 1$, and thus $\alpha \in [-1, 0]$. Since both $\alpha \leq 0$ and $-\hat{m}^i\left(U^i\right) < 0$, the inner minimum is attained at $y = \overline{b^{-i}}$, and since $\overline{b^{-i}} > g^i\left(U^i\right)$, we have $\alpha\left(\overline{b^{-i}} - g^i\left(U^i\right)\right) + \hat{m}^i\left(U^i\right)\overline{b^{-i}} \leq \hat{m}^i\left(U^i\right)\overline{b^{-i}}$. Therefore, $\alpha = 0$ is optimal in the outer maximization problem. This corresponds to the optimal strategy $x = m^i\left(U^i\right)$.

- If $\hat{m}^i\left(U^i\right) < 0$, we can conclude that $m^i\left(U^i\right) = 0$, and thus $\alpha \in [0, 1]$. Since both $\alpha \geq 0$ and $-\hat{m}^i\left(U^i\right) > 0$, the inner minimum is attained at $y = \underline{b^{-i}}$, and since $\underline{b^{-i}} < g^i\left(U^i\right)$, we have $\alpha\left(\underline{b^{-i}} - g^i\left(U^i\right)\right) + \hat{m}^i\left(U^i\right)\underline{b^{-i}} \leq \hat{m}^i\left(U^i\right)\underline{b^{-i}}$. Therefore, $\alpha = 0$ is optimal in the outer maximization problem. This corresponds to the optimal strategy $x = m^i\left(U^i\right)$.

Finally, we investigate the case $\underline{b^{-i}} = g^i\left(U^i\right)$, which will result in the best-response strategy-set $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \left[m^i\left(U^i\right), 1\right]$. We have $\underline{b^{-i}} = g^i\left(U^i\right) < \overline{b^{-i}}$. As before, there are three cases:

- if $\hat{m}^i\left(U^i\right) = 0$, then $\min_{y\in\{g^i(U^i), \overline{b^{-i}}\}} \alpha\left(y - g^i\left(U^i\right)\right) \leq 0$ for any $\alpha \in \mathbb{R}$. Thus $\alpha \in \left[0, 1 - m^i\left(U^i\right)\right]$ is optimal in the outer maximization problem. This set corresponds to a best-response strategy-set that contains all strategies in $\left[m^i\left(U^i\right), 1\right]$.

- If $\hat{m}^i\left(U^i\right) > 0$, then the argument is as in the case where $\underline{b^{-i}} < g^i\left(U^i\right) < \overline{b^{-i}}$, where the inner minimum is attained at $y = \overline{b^{-i}}$ and the unique maximizer is $\alpha = 0$. This corresponds to the unique strategy $x = m^i\left(U^i\right)$. However, $\hat{m}^i\left(U^i\right) > 0$ when $m^i\left(U^i\right) = 1$. Thus, the set $\left[m^i\left(U^i\right), 1\right]$ contains only the strategy $m^i\left(U^i\right)$, or equivalently the pure strategy 1.

- If $\hat{m}^i\left(U^i\right) < 0$, then the argument is as in the case where $\underline{b^{-i}} < g^i\left(U^i\right) < \overline{b^{-i}}$, where the inner minimum is attained at $y = \underline{b^{-i}} = g^i\left(U^i\right)$. Thus, the optimization variable α vanishes from the outer maximization problem and hence player i is indifferent as to what to play. Consequently, the optimal strategy for player i is any strategy in $[0,1]$. Since $\hat{m}^i\left(U^i\right) < 0$ when $m^i\left(U^i\right) = 0$, the set $\left[m^i\left(U^i\right), 1\right]$ contains all the strategies and thus represents the best-response strategy-set.

Thus, for $\underline{b^{-i}} = g^i\left(U^i\right)$ the optimal best-response strategy-set is given by $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \left[m^i\left(U^i\right), 1\right]$. Similar arguments can be used to prove that $\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \left[0, m^i\left(U^i\right)\right]$ when $\overline{b^{-i}} = g^i\left(U^i\right)$. □

5.5.2 The Best-response Extremes-set

In the last part of this section, we will use the properties of the best-response strategies-set described in Theorem 5.5.1 to construct (a one-dimensional representation of) the parametric best-response extremes-set $\mathcal{E}^i[\cdot]$ as defined in (BE). That is, we will construct the best-response extremes-sets $\mathcal{E}^i\left[\left(\underline{b^{-i}}, \overline{b^{-i}}\right)\right] \subseteq \mathcal{L}$ that map the opponent's extreme strategies $\left(\underline{b^{-i}}, \overline{b^{-i}}\right) \in \mathcal{L}$ to all extreme strategies $\left(\underline{b^i}, \overline{b^i}\right) \in \mathcal{L}$ that unilaterally satisfy the conditions in Theorem 5.4.1 for player i.

By Theorem 5.5.1, a player either plays a pure strategy, the max-min strategy or is indifferent between any convex combination of the two. For every utility matrix $U^i \in \mathcal{U}^i$, player i's best-response depends on the relationship between the opponent's extreme strategies $\underline{b^{-i}}$, $\overline{b^{-i}}$ and the strategy that makes player i indifferent for that U^i, given by $g^i\left(U^i\right)$.

To this end, given the opponent's extreme strategies $\underline{b^{-i}}$, $\overline{b^{-i}}$, we can partition player i's utility matrix set $\mathcal{U}^i$ into three open subsets and their boundaries, where:

1. $\mathcal{U}^i_b\left[\underline{b^{-i}}\right] := \left\{U^i \in \mathcal{U}^i \ : \ g^i\left(U^i\right) < \underline{b^{-i}}\right\}$, which we denote as the bottom set,
2. $\mathcal{U}^i_m\left[\underline{b^{-i}}, \overline{b^{-i}}\right] := \left\{U^i \in \mathcal{U}^i \ : \ \underline{b^{-i}} < g^i\left(U^i\right) < \overline{b^{-i}}\right\}$, which we denote as the middle set, and
3. $\mathcal{U}^i_t\left[\overline{b^{-i}}\right] := \left\{U^i \in \mathcal{U}^i \ : \ g^i\left(U^i\right) > \overline{b^{-i}}\right\}$, which we denote as the top set.

Then from Theorem 5.5.1, we derive the following two Corollaries:

Corollary 5.5.1. *When player i is a coordination player, then*

$$\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \begin{cases} \{1\}, & \forall U^i \in \mathcal{U}_b^i\left[\underline{b^{-i}}\right], \\ \{m^i\left(U^i\right)\}, & \forall U^i \in \mathcal{U}_m^i\left[\underline{b^{-i}}, \overline{b^{-i}}\right], \\ \{0\}, & \forall U^i \in \mathcal{U}_t^i\left[\overline{b^{-i}}\right]. \end{cases}$$

Furthermore, whenever a utility matrix $U^i \in \mathcal{U}^i$ is at the boundary of two partitions, then the best-response strategy-set contains all convex combinations of the two strategies corresponding to each partition:

$$\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \begin{cases} \left[m^i\left(U^i\right), 1\right], & \forall U^i \in \partial\mathcal{U}_b^i\left[\underline{b^{-i}}\right] \bigcap \partial\mathcal{U}_m^i\left[\underline{b^{-i}}, \overline{b^{-i}}\right], \\ \left[0, m^i\left(U^i\right)\right], & \forall U^i \in \partial\mathcal{U}_t^i\left[\overline{b^{-i}}\right] \bigcap \partial\mathcal{U}_m^i\left[\underline{b^{-i}}, \overline{b^{-i}}\right]. \end{cases}$$

Corollary 5.5.2. *When player i is an anti-coordination player, then*

$$\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \begin{cases} \{0\}, & \forall U^i \in \mathcal{U}_b^i\left[\underline{b^{-i}}\right], \\ \{m^i\left(U^i\right)\}, & \forall U^i \in \mathcal{U}_m^i\left[\underline{b^{-i}}, \overline{b^{-i}}\right], \\ \{1\}, & \forall U^i \in \mathcal{U}_t^i\left[\overline{b^{-i}}\right]. \end{cases}$$

Furthermore, whenever a utility matrix $U^i \in \mathcal{U}^i$ is at the boundary of two partitions, then the best-response strategy-set contains all convex combinations of the two strategies corresponding to each partition:

$$\mathcal{S}^i\left[U^i; \underline{b^{-i}}, \overline{b^{-i}}\right] = \begin{cases} \left[0, m^i\left(U^i\right)\right], & \forall U^i \in \partial\mathcal{U}_b^i\left[\underline{b^{-i}}\right] \bigcap \partial\mathcal{U}_m^i\left[\underline{b^{-i}}, \overline{b^{-i}}\right], \\ \left[m^i\left(U^i\right), 1\right], & \forall U^i \in \partial\mathcal{U}_t^i\left[\overline{b^{-i}}\right] \bigcap \partial\mathcal{U}_m^i\left[\underline{b^{-i}}, \overline{b^{-i}}\right]. \end{cases}$$

Let player i be a coordination player. Then, from Corollary 5.5.1, we know that, given the opponent's extreme strategies $\left(\underline{b^{-i}}, \overline{b^{-i}}\right) \in \mathcal{L}$, if the bottom set $\mathcal{U}_b^i\left[\underline{b^{-i}}\right]$ is non-empty, then there is a utility matrix in $\mathcal{U}^i$ where player i's best-response strategy is unique and equal to 1. In that case, player i's high extreme strategy $\overline{b^i}$ is equal to 1. Similarly, if the top set $\mathcal{U}_t^i\left[\overline{b^{-i}}\right]$ is non-empty, then player i's low extreme strategy $\underline{b^i}$ is equal to 0.

On the other hand, if all of player i's utility matrices $U^i \in \mathcal{U}^i$ satisfy $\underline{b^{-i}} < g^i\left(U^i\right) < \overline{b^{-i}}$, that is $\mathcal{U}^i = \mathcal{U}_m^i\left[\underline{b^{-i}}, \overline{b^{-i}}\right]$ where all matrices belong in the middle set, then player i must play her max-min strategy $m^i\left(U^i\right)$ for every matrix $U^i \in \mathcal{U}^i$. In that case, player i's extreme strategies are her lowest

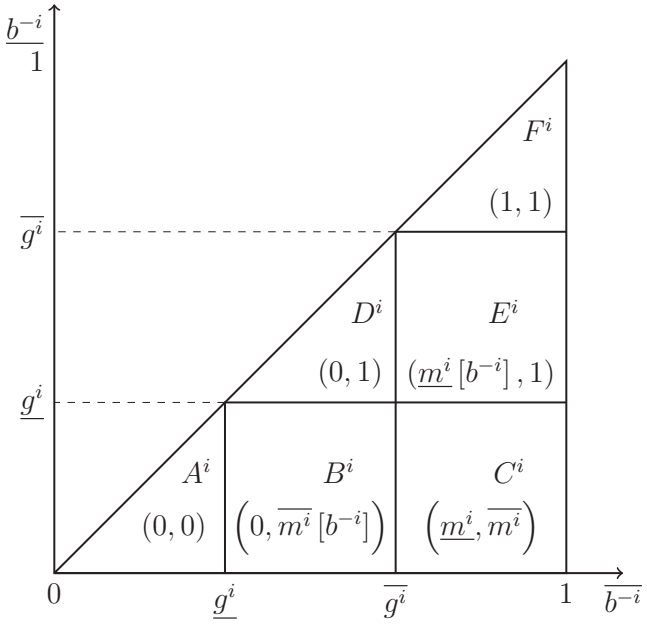


Figure 5.5: The best-response extremes-set $\mathcal{E}^i[b^{-i}] \subseteq \mathcal{L}$ of a coordination player. The triangle represents the two-dimensional domain of the parameter $b^{-i} = (\underline{b^{-i}}, \overline{b^{-i}}) \in \mathcal{L}$, whilst the labels of the partitions $A^i, \ldots, F^i$, represent the unique values that the correspond to $\mathcal{E}^i[b^{-i}]$ when b^{-i} is in each partition.

and highest max-min strategies, given by the value functions

$$\underline{m^i}\left[\left(\underline{b^{-i}}, \overline{b^{-i}}\right)\right] := \inf_{U^i \in \mathcal{U}^i_m\left[\left(\underline{b^{-i}}, \overline{b^{-i}}\right)\right]} m^i\left(U^i\right),$$

and

$$\overline{m^i}\left[\underline{b^{-i}}, \overline{b^{-i}}\right] := \sup_{U^i \in \mathcal{U}^i_m\left[\underline{b^{-i}}, \overline{b^{-i}}\right]} m^i\left(U^i\right),$$

or equivalently, since $\mathcal{U}^i_m\left[\underline{b^{-i}}, \overline{b^{-i}}\right] = \mathcal{U}^i$, by the static optimization problems

$$\underline{m^i} := \inf_{U^i \in \mathcal{U}^i} m^i\left(U^i\right) \quad \text{and} \quad \overline{m^i} := \sup_{U^i \in \mathcal{U}^i} m^i\left(U^i\right).$$

Consequently, we can derive player i's extreme strategies by examining how the opponent's extreme strategies $\left(\underline{b^{-i}}, \overline{b^{-i}}\right) \in \mathcal{L}$ partition player i's utility matrix set $\mathcal{U}^i$ into its three subsets. To this end, let

$$\underline{g^i} := \inf_{U^i \in \mathcal{U}^i} g^i\left(U^i\right) \quad \text{and} \quad \overline{g^i} := \sup_{U^i \in \mathcal{U}^i} g^i\left(U^i\right) \tag{5.5}$$

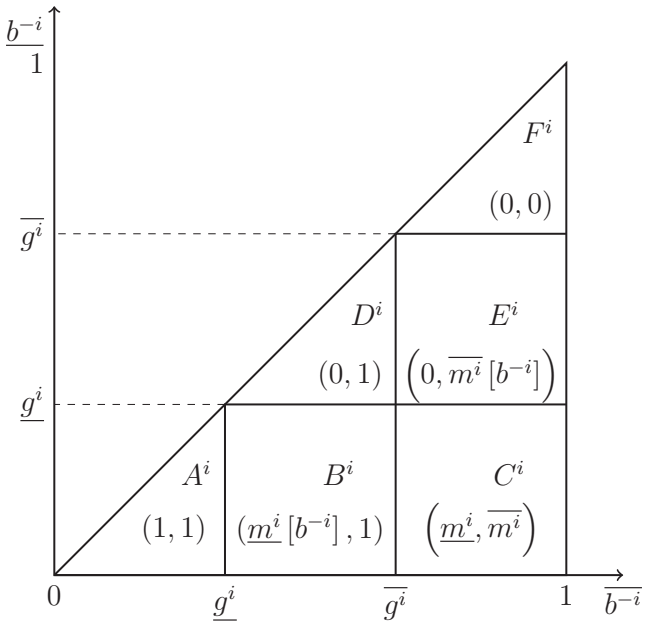


Figure 5.6: The best-response extremes-set $\mathcal{E}^i[\cdot] \subseteq \mathcal{L}$ of an anti-coordination player.

denote player i's lowest and highest indifference strategies $g^i(\cdot)$ given her utility matrix set $\mathcal{U}^i$. Then it is easy to see that when $\underline{b^{-i}} > \underline{g^i}$, the bottom set $\mathcal{U}_b^i\left[\underline{b^{-i}}\right]$ is non-empty, and thus player i's extreme strategy is $\overline{b^i} = 1$. Similarly, if $\overline{b^{-i}} < \overline{g^i}$, then the top set $\mathcal{U}_t^i\left[\overline{b^{-i}}\right]$ is non-empty, and thus $\underline{b^i} = 0$. By expanding this analysis to encompass all possible ordering scenarios between the opponent's extreme strategies $\underline{b^{-i}}$, $\overline{b^{-i}}$ and player i's extreme indifference strategies $\underline{g^i}$, $\overline{g^i}$, we can derive an analytical expression for player i's best-response extremes-set $\mathcal{E}^i[\cdot]$.

Proposition 5.5.3 (Best-response extremes-set)**.** *Depending on the opponent's extreme strategies* $b^{-i} = \left(\underline{b^{-i}}, \overline{b^{-i}}\right) \in \mathcal{L}$ *and her lowest and highest indifference strategies* $\underline{g^i}, \overline{g^i} \in [0,1]$ *for her utility matrix set* $\mathcal{U}^i$*, player* i*'s low extreme strategy* $\underline{b^i}$ *is*

a) a pure strategy from within $\{0,1\}$*,*

b) the lowest max-min strategy $\underline{m^i}\left[b^{-i}\right]$*, or,*

c) any convex combination of the two.

Similarly, player i*'s high extreme strategy* $\overline{b^i}$ *is a pure strategy, the highest max-min strategy* $\overline{m^i}\left[b^{-i}\right]$*, or, any convex combination of the two.*

The best-response extremes-set $\mathcal{E}^i\left[b^{-i}\right] \subseteq \mathcal{L}$ *is as illustrated in Figure 5.5, when player* i *is a coordination player, and as illustrated in Figure 5.6, when she is an anti-coordination player. The axes*

represent the two-dimensional domain of the parameter $b^{-i} = \left(\underline{b^{-i}}, \overline{b^{-i}}\right) \in \mathcal{L}$. *In each case, the domain* $\mathcal{L}$ *of the parametric set* $\mathcal{E}^i[\cdot]$ *is partitioned into the 6 regions* $A^i, \ldots, F^i$. *The label in each of the partitions* $A^i, \ldots, F^i$, *represents the value of the unique element belonging in the best-response extremes-set* $\mathcal{E}^i\left[b^{-i}\right]$, *when* b^{-i} *is inside that partition. When* b^{-i} *is at the boundaries of two or more partitions, the best-response extremes-set is equal to the smallest interval containing all the values of* $\mathcal{E}^i\left[b^{-i}\right]$ *when* b^{-i} *is in each of the partitions involved. For example, for* $b^{-i} \in \mathcal{L}$ *at the boundary of partitions* A^i *and* B^i *in Figure 5.5, we have* $\mathcal{E}^i\left[b^{-i}\right] = \left[(0,0), \left(0, \overline{m^i}\left[b^{-i}\right]\right)\right]$.

Proof. We only prove the proposition for a coordination player (see Figure 5.5), and only for the first dimension of the set $\mathcal{E}^i[\cdot]$, which corresponds to player i's low extreme strategy $\underline{b^i}$. Symmetrical arguments can be used to prove the case for the second dimension, corresponding to player i's high extreme strategy $\overline{b^i}$, as well as for the case of an anti-coordination player.

Consider a pair of the opponent's extreme strategies $b^{-i} = \left(\underline{b^{-i}}, \overline{b^{-i}}\right) \in \mathcal{L}$. If $\overline{b^{-i}} < \overline{g^i}$, implying that b^{-i} lies in one of the partitions A^i, B^i, or D^i, there exists a utility matrix $U^i \in \mathcal{U}^i$ with $\overline{b^{-i}} < g^i\left(U^i\right)$. Thus, the top set $\mathcal{U}_t^i\left[\overline{b^{-i}}\right]$ is non-empty, and player i's low extreme strategy is 0. On the other hand, if $\overline{b^{-i}} \geq \underline{b^{-i}} > \overline{g^i}$, meaning that b^{-i} lies inside partition F^i, then $\underline{b^{-i}} > g^i\left(U^i\right)$ for all $U^i \in \mathcal{U}^i$. Thus, all matrices $U^i \in \mathcal{U}^i$ belong in the bottom set $\mathcal{U}_b^i\left[\underline{b^{-i}}\right]$. As a result player i's low (and high) extreme strategy is 1.

We now consider the case where $\underline{b^{-i}} < \overline{g^i}$ and $\overline{b^{-i}} > \overline{g^i}$, which corresponds to partitions C^i and E^i. As $\overline{b^{-i}} > \overline{g^i}$, there is no utility matrix $U^i \in \mathcal{U}^i$ with $\overline{b^{-i}} < g^i\left(U^i\right)$. Thus, there is no utility matrix in the top set $\mathcal{U}_t^i\left[\overline{b^{-i}}\right]$, and the pure strategy 0 is never a best-response. On the other hand, as $\underline{b^{-i}} < \overline{g^i}$, there are utility matrices that satisfy $\underline{b^{-i}} < g^i\left(U^i\right) < \overline{b^{-i}}$, meaning that the middle set $\mathcal{U}_m^i\left[\underline{b^{-i}}, \overline{b^{-i}}\right]$ is non-empty. Hence, the lowest max-min strategy that player i plays, and hence player i's low extreme strategy, is given by the value function $\underline{m^i}\left[b^{-i}\right]$. For any $b^{-i} = \left(\underline{b^{-i}}, \overline{b^{-i}}\right)$ in partition C^i, we have $\underline{b^{-i}} < \underline{g^i}$ and $\overline{b^{-i}} > \overline{g^i}$, thus all matrices $U^i \in \mathcal{U}^i$ satisfy $\underline{b^{-i}} < g^i\left(U^i\right) < \overline{b^{-i}}$. Hence $\mathcal{U}_m^i\left[\underline{b^{-i}}, \overline{b^{-i}}\right] = \mathcal{U}^i$ and the value function $\underline{m^i}\left[b^{-i}\right]$ is constant and equal to $\underline{m^i}$.

Finally, we investigate the case where b^{-i} lies at the boundary of two or more partitions. If b^{-i} lies on the boundary of partitions B^i and C^i, then $\overline{b^{-i}} = \overline{g^i}$ and $\underline{b^{-i}} \leq \underline{g^i}$. Thus $\underline{b^{-i}} \leq g^i\left(U^i\right) \leq \overline{b^{-i}}$ for all $U^i \in \mathcal{U}^i$, and the following statements are true: $\underline{m^i}\left[b^{-i}\right] = \underline{m^i}$ and $\overline{m^i}\left[b^{-i}\right] = \overline{m^i}$. Furthermore, the middle set $\mathcal{U}_m^i\left[b^{-i}\right]$ is non-empty, and thus player i must play the strategy $\underline{m^i}$ for some matrix in

$\mathcal{U}_m^i\left[b^{-i}\right]$. As a result, player i's low extreme strategy is at most $\underline{m^i}$. However, since $\overline{b^{-i}} = \overline{g^i}$, there is no utility matrix with $g^i\left(U^i\right) > \overline{b^{-i}}$. Thus, there is no utility matrix in the top set $\mathcal{U}_t^i\left[\overline{b^{-i}}\right]$ where player i must play the pure strategy 0. On the other hand, there is at least one matrix at the boundary of the top set and the middle set, that is, there exists a matrix $U^i \in \mathcal{U}^i$ satisfying $g^i\left(U^i\right) = \overline{b^{-i}}$. For this matrix, player i can choose any strategy from the interval $\left[0, m^i\left(U^i\right)\right]$, where by definition $m^i\left(U^i\right) \geq \underline{m^i}$. Recalling that $\underline{b^i} \leq \underline{m^i}$, we conclude that player i's extreme strategy can be any strategy in the interval $\left[0, \underline{m^i}\right]$. Similar arguments hold for the boundary of the partitions E^i and F^i, where the low extreme strategy is any element in $\left[\underline{m^i}\left[b^{-i}\right], 1\right]$. □

Lemma 5.5.2 (Linear reformulation of $\underline{m^i}\left[b^{-i}\right]$ and $\overline{m^i}\left[b^{-i}\right]$)**.** *Define $m, g, d \in \mathbb{R}^4$ so that for any $U^i \in \mathcal{U}^i$*

$$m^i\left(U^i\right) = \frac{m^\top \operatorname{vec}\left(U^i\right)}{d^\top \operatorname{vec}\left(U^i\right)} \text{ and } g^i\left(U^i\right) = \frac{g^\top \operatorname{vec}\left(U^i\right)}{d^\top \operatorname{vec}\left(U^i\right)}.$$

where $\operatorname{vec}\left(U^i\right)$ is a vector expression of the matrix U^i. When i is a coordination player, then for any $b^{-i} = \left(\underline{b^{-i}}, \overline{b^{-i}}\right) \in \mathcal{L}$, we have

$$\begin{aligned} \underline{m^i}\left[b^{-i}\right] = \quad \inf \quad & m^\top v \\ s.t. \quad & v \in \mathbb{R}^4, \tau \in \mathbb{R} \\ & \underline{b^{-i}} \leq g^\top v \leq \overline{b^{-i}} \\ & d^\top \tau = 1,\ \tau > 0 \\ & v \in \operatorname{cone}\left(\operatorname{vec}\left(\mathcal{U}^i\right)\right), \end{aligned} \tag{5.6}$$

where $\operatorname{vec}\left(\mathcal{U}^i\right)$ is a vector representation of the set $\mathcal{U}^i$, so that $U^i \in \mathcal{U}^i \iff \operatorname{vec}\left(U^i\right) \in \operatorname{vec}\left(\mathcal{U}^i\right)$, and $\operatorname{cone}\left(\operatorname{vec}\left(\mathcal{U}^i\right)\right)$ is the cone generated by that set. The value function $\overline{m^i}\left[b^{-i}\right]$ is identical except that the infimum in the objective is replaced by a supremum. For an anti-coordination, the problems are again identical except that the last two constraints are $\tau < 0$ and $-v \in \operatorname{cone}\left(\mathcal{U}^i\right)$, respectively.

Proof. Apply the Charnes-Cooper transformation [27], where $v = \frac{\operatorname{vec}\left(U^i\right)}{d^\top \operatorname{vec}\left(U^i\right)}$, and $\tau = \frac{1}{d^\top \operatorname{vec}\left(U^i\right)}$. Let $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{U}^i$. Then $d^\top \operatorname{vec}\left(U^i\right) = a - c + d - b$. Thus, when i is a coordination player we have $d^\top \operatorname{vec}\left(U^i\right) > 0$ for all $U^i \in \mathcal{U}^i$, and thus $\tau \geq 0$. Furthermore, as $\operatorname{vec}\left(U^i\right) = \frac{v}{\tau}$, we find $\operatorname{vec}\left(U^i\right) \in \mathcal{U}^i \iff v \in \operatorname{cone}\mathcal{U}^i$. □

Remark 5.5.2 (Efficient computation of $\mathcal{E}^i[\cdot]$)**.** *Given the opponent's extreme strategies $b^{-i} \in \mathcal{L}$, and*

player i's polyhedral utility matrix set $\mathcal{U}^i$ described in Proposition 5.4.2, computing the best-response extremes-set $\mathcal{E}^i\left[b^{-i}\right]$ requires the solution of the following optimization problems:

- *the static optimization problems $\underline{g^i}$ and $\overline{g^i}$ and*
- *the parametric problems involved in the value functions $\underline{m^i}\left[b^{-i}\right]$ and $\overline{m^i}\left[b^{-i}\right]$ or their static counterparts $\underline{m^i}$ and $\overline{m^i}$.*

The above problems are linear-fractional optimization problems with polyhedral constraints and, in the case of the parametric problems $\underline{m^i}\left[b^{-i}\right]$ and $\overline{m^i}\left[b^{-i}\right]$, with additional linear-fractional constraints. Consequently, all these problems can be reformulated as linear programs using the Charnes-Cooper transformation [27] as demonstrated in Lemma 5.5.2, and thus the best-response extremes-strategy set $\mathcal{E}^i[\cdot]$ can be computed in efficiently.

Remark 5.5.3 (Monotonicity of $\mathcal{E}^i[\cdot]$)**.** *The best-response extremes-set $\mathcal{E}^i[b{-}i]$ is non-decreasing (non-increasing) on its parameter $b^{-i} \in \mathcal{L}$ with respect to component-wise inequality, whenever player i is a coordination (anti-coordination) player. That is, for any $b_1^{-i}, b_2^{-i} \in \mathcal{L}$ with $b_1^{-i} \geq b_2^{-i}$, we have*

$$\forall b_1^i \in \mathcal{E}^i\left[b_1^{-i}\right], \forall b_2^i \in \mathcal{E}^i\left[b_2^{-i}\right] : b_1^i \geq (\leq)\, b_2^i.$$

According to Proposition 5.5.3, for a $b^{-i} \in \mathcal{L}$, the best-response extremes-set $\mathcal{E}^i\left[b^{-i}\right]$ is singled-valued in most of its domain with a finite piece-wise representation, where most of the pieces are constant. Player i's extreme strategies either involve a pure strategy, an extreme max-min strategy $\underline{m^i}\left[b^{-i}\right]$ or $\overline{m^i}\left[b^{-i}\right]$, or all convex combinations of the two, depending on which partition $A^i, \ldots, F^i$, of the domain of $\mathcal{E}^i\left[b^{-i}\right]$, that contains the parameter b^{-i}. Since any equilibrium $\left(b^1, b^2\right) \in \mathcal{L}^2$ satisfies $b^i \in \mathcal{E}^i\left[b^{-i}\right]$, $i = 1, 2$, there are only a finite number of forms that an equilibrium may adopt. In the next section, we utilize this result to develop a full enumeration algorithm for finding all equilibria of the game.

5.6 Equilibria Enumeration

An equilibrium in extreme strategies is any pair of two-dimensional elements $b^1, b^2 \in \mathbb{R}^2$ satisfying

$$\left.\begin{array}{l} b^i \in \mathcal{L} \\ b^i \in \mathcal{E}^i\left[b^{-i}\right] \end{array}\right\} i = 1, 2. \tag{$\mathcal{EQ}$}$$

A naive approach for finding equilibria is to calculate a fixed point of the product function

$$\begin{aligned} f &: \mathcal{L}^2 \rightrightarrows \mathcal{L}^2 \\ f\left(b^1, b^2\right) &:= \mathcal{E}^1\left(b^2\right) \times \mathcal{E}^2\left(b^1\right), \end{aligned}$$

where, for any fixed point we have

$$\left(b^1, b^2\right) \in f\left(b^1, b^2\right) \iff b^1 \in \mathcal{E}^1\left[b^2\right] \text{ and } b^2 \in \mathcal{E}^2\left[b^1\right],$$

as required by the equilibrium conditions. Unfortunately, the product function $f(\cdot)$ is not a contraction and nor is it necessarily monotone (c.f. Remark 5.5.3). Consequently, finding a fixed-point of $f(\cdot)$ and hence an equilibrium of the game cannot be guaranteed with an iterative process involving $f(\cdot)$.

Remark 5.6.1 (Equilibrium existence proof)**.** *According to Berge's maximum theorem, the best-response extremes-set $\mathcal{E}^i\left[b^{-i}\right]$ is upper hemi-continuous. Thus the product function $f(\cdot)$ is also upper hemi-continuous. Hence one can use Kakutani's fixed-point theorem to prove the existence of a fixed-point in the product function $f(\cdot)$ and equivalently the existence of an equilibrium in behavioural functions. This reasoning provides an alternate existence-proof to the one presented in [1].*

It seems impossible to express ($\mathcal{EQ}$) as a feasibility problem that we can solve computationally, due to the complicated structure of the best-response extremes-set $\mathcal{E}^i[\cdot]$ over the domain $\mathcal{L}$. However, the structural characterization of $\mathcal{E}^i[\cdot]$ described in Proposition 5.5.3, allows for the formulation of feasibility problems involving only specific partitions of the domain, within which the best-response extremes-set $\mathcal{E}^i[\cdot]$, $i = 1, 2$ have a simple description. That is, instead of solving the feasibility problem ($\mathcal{EQ}$) in order to obtain all equilibria $\left(b^1, b^2\right) \in \mathcal{L}^2$, we select two subsets $\mathcal{P}^i \subset \mathcal{L}$, $i = 1, 2$ and solve the problem

$$\left.\begin{array}{l} b^{-i} \in \mathcal{P}^i \\ b^i \in \mathcal{E}^i\left[b^{-i}\right] \end{array}\right\} i = 1, 2 \tag{5.7}$$

in order to retrieve all equilibria satisfying $\left(b^1, b^i\right) \in \mathcal{P}^2 \times \mathcal{P}^1$. The subsets $\mathcal{P}^i \subset \mathcal{L}$, $i = 1, 2$, must be chosen so that for any $b^{-i} \in \mathcal{P}^i$, the best-response extremes-set $\mathcal{E}^i\left[b^{-i}\right]$ has a simple description, for which the feasibility problem (5.7) can be solved. To this end, define $\mathcal{E}^i_{P^i}\left[b^{-i}\right]$ so that

$$\forall b^{-i} \in \mathcal{P}^i \,:\, b^i \in \mathcal{E}^{-i}\left[b^i\right] \iff b^i \in \mathcal{E}^i_{\mathcal{P}^i}\left[b^{-i}\right].$$

Corollary 5.6.1. *Let $\mathcal{P}^i$ be one of the partitions A^i, C^i, D^i or F^i, excluding their boundaries with other partitions, or let $\mathcal{P}^i$ be one of those boundaries. The set $\mathcal{E}^i_{\mathcal{P}^i}\left[b^{-i}\right]$ is illustrated via conditions on $b^i = \left(\underline{b^i}, \overline{b^i}\right)$ in the third and fourth column of Table 5.2 for a coordination (COOR) and an anti-coordination (A-COOR) player i, respectively. For any such $\mathcal{P}^i$, the set $\mathcal{E}^i_{\mathcal{P}^i}\left[b^{-i}\right]$ involves only linear constraints.*

Lemma 5.6.1. *Let $b^{-i} \in \mathcal{L}$, and $\underline{b}, \overline{b} \in [0, 1]$. The conditions $\underline{b} = \underline{m^i}\left[b^{-i}\right]$ and $\overline{b} = \overline{m^i}\left[b^{-i}\right]$ can be expressed as two-dimensional linear complementarity problems that do not include the parameter b^{-i} in the complementarity conditions.*

Proof. The optimization problems involved in $\underline{m^i}\left[b^{-i}\right]$ and $\overline{m^i}\left[b^{-i}\right]$ can be reformulated as linear optimization problems as demonstrated in Lemma 5.5.2, where the parameter b^{-i} appears as the right-hand-side of the linear constraints . By expressing the problem as a solution of the KKT conditions, we obtain a linear complementarity problem. Since the parameter b^{-i} is the right-hand-side of a linear constraint, it does not appear in the complementarity condition. □

Corollary 5.6.2. *Let $\mathcal{P}^i$ be one of the partitions B^i, E^i, excluding their boundaries with other partitions, or let $\mathcal{P}^i$ be one of those boundaries. The set $\mathcal{E}^i_{\mathcal{P}^i}\left[b^{-i}\right]$ is illustrated via conditions on $b^i = \left(\underline{b^i}, \overline{b^i}\right)$ in the third and fourth column of Table 5.3 for a coordination (COOR) and an anti-coordination (A-COOR) player i, respectively. For any such $\mathcal{P}^i$, the set $\mathcal{E}^i_{\mathcal{P}^i}\left[b^{-i}\right]$ involves only linear constraints.*

Consequently, all equilibria of the game can be found by solving the feasibility problem

$$\left.\begin{array}{l} b^{-i} \in \mathcal{P}^i \\ b^i \in \mathcal{E}^i_{\mathcal{P}^i}\left[b^{-i}\right] \end{array}\right\} i = 1, 2, \qquad (\mathcal{E}Q_P)$$

where $\mathcal{P}^i$ and $\mathcal{P}^{-i}$ take the values shown in Tables 5.2 and 5.3. The constraints of problem $(\mathcal{E}Q_P)$ are

Partition $\mathcal{P}^i$		$\mathcal{E}^i_{P^i}[b^{-i}]$ COOR	$\mathcal{E}^i_{P^i}[b^{-i}]$ A-COOR
A^i	$[0,\underline{g^i}) \times [0,\underline{g^i})$	$\underline{b^i} = 0$ $\overline{b^i} = 0$	$\underline{b^i} = 1$ $\overline{b^i} = 1$
C^i	$[0,\underline{g^i}) \times \left(\overline{g^i},1\right]$	$\underline{b^i} = \underline{m^i}$ $\overline{b^i} = \overline{m^i}$	$\underline{b^i} = \underline{m^i}$ $\overline{b^i} = \overline{m^i}$
D^i	$\left(\underline{g^i},\overline{g^i}\right) \times \left(\underline{g^i},\overline{g^i}\right)$	$\underline{b^i} = 0$ $\overline{b^i} = 1$	$\underline{b^i} = 0$ $\overline{b^i} = 1$
F^i	$\left(\overline{g^i},1\right] \times \left(\overline{g^i},1\right]$	$\underline{b^i} = 1$ $\overline{b^i} = 1$	$\underline{b^i} = 0$ $\overline{b^i} = 0$
$A^i \bigcap B^i$	$[0,\underline{g^i}] \times \{\underline{g^i}\}$	$\underline{b^i} = 0$ $0 \leq \overline{b^i} \leq \overline{m^i}\left[(0,\underline{g^i})\right]$	$\underline{m^i}\left[(0,\underline{g^i})\right] \leq \underline{b^i} \leq 1$ $\overline{b^i} = 1$
$B^i \bigcap C^i$	$[0,\underline{g^i}] \times \left\{\overline{g^i}\right\}$	$0 \leq \underline{b^i} \leq \underline{m^i}$ $\overline{b^i} = \overline{m^i}$	$\underline{b^i} = \underline{m^i}$ $\overline{m^i} \leq \overline{b^i} \leq 1$
$C^i \bigcap E^i$	$\{\underline{g^i}\} \times \left[\overline{g^i},1\right]$	$\underline{b^i} = \underline{m^i}$ $\overline{m^i} \leq \overline{b^i} \leq 1$	$0 \leq \underline{b^i} \leq \underline{m^i}$ $\overline{b^i} = \overline{m^i}$
$E^i \bigcap F^i$	$\left\{\overline{g^i}\right\} \times \left[\overline{g^i},1\right]$	$\underline{m^i}\left[\left(\overline{g^i},1\right)\right] \leq \underline{b^i} \leq 1$ $\overline{b^i} = 1$	$\underline{b^i} = 0$ $0 \leq \overline{b^i} \leq \overline{m^i}\left[\left(\overline{g^i},1\right)\right]$
$A^i \bigcap D^i$	$\{(\underline{g^i},\underline{g^i})\}$	$\underline{b^i} = 0$ $0 \leq \overline{b^i} \leq 1$	$0 \leq \underline{b^i} \leq 1$ $\overline{b^i} = 1$
$F^i \bigcap D^i$	$\left\{\left(\overline{g^i},\overline{g^i}\right)\right\}$	$0 \leq \underline{b^i} \leq 1$ $\overline{b^i} = 1$	$\underline{b^i} = 0$ $0 \leq \overline{b^i} \leq 1$
$C^i \bigcap D^i$	$\left\{\left(\underline{g^i},\overline{g^i}\right)\right\}$	$0 \leq \underline{b^i} = \underline{m^i}$ $\overline{m^i} \leq \overline{b^i} \leq 1$	$0 \leq \underline{b^i} = \underline{m^i}$ $\overline{m^i} \leq \overline{b^i} \leq 1$

Table 5.2: The set $\mathcal{E}^i_{P^i}[b^{-i}]$ for partitions A^i, C^i, D^i, F^i, expressed through conditions on $b^i = (\underline{b^i},\overline{b^i})$ for both cases where i is a coordination (COOR) and an anti-coordination (A-COOR) player.

Partition $\mathcal{P}^i$		$\mathcal{E}^i_{P^i}$ COOR	A-COOR
B^i	$[0, \underline{g^i}) \times \left(\underline{g^i}, \overline{g^i}\right)$	$\underline{b^i} = 0$ $b^i = \overline{m^i}\left[\overline{b^{-i}}\right]$	$\underline{b^i} = \underline{m^i}\,[b^{-i}]$ $\overline{b^i} = 1$
E^i	$\left(\underline{g^i}, \overline{g^i}\right) \times \left(\overline{g^i}, 1\right]$	$\underline{b^i} = \underline{m^i}\,[b^{-i}]$ $\overline{b^i} = 1$	$\underline{b^i} = 0$ $\overline{b^i} = \overline{m^i}\,[b^{-i}]$
$B^i \bigcap D^i$	$\{\underline{g^i}\} \times \left(\underline{g^i}, \overline{g^i}\right)$	$\underline{b^i} = 0$ $\overline{b} \leq \overline{b^i} \leq 1$ $\overline{b} = \overline{m^i}\left[\overline{b^{-i}}\right]$	$0 \leq \underline{b^i} \leq \underline{b}$ $\overline{b^i} = 1$ $b = \underline{m^i}\,[b^{-i}]$
$E^i \bigcap D^i$	$\left(\underline{g^i}, \overline{g^i}\right) \times \left\{\overline{g^i}\right\}$	$0 \leq \underline{b^i} \leq \underline{b}$ $\overline{b^i} = 1$ $\underline{b} = \underline{m^i}\,[b^{-i}]$	$\underline{b^i} = 0$ $\overline{b} \leq \overline{b^i} \leq 1$ $\overline{b} = \overline{m^i}\left[\overline{b^{-i}}\right]$

Table 5.3: The set $\mathcal{E}^i_{P^i}[b^{-i}]$ for partitions B^i, E^i expressed via conditions on $b^i = (\underline{b^i}, \overline{b^i})$ for both cases where i is a coordination (COOR) and an anti-coordination (A-COOR) player. The conditions $\underline{b} = \underline{m^i}[b^{-i}]$ and $\overline{b} = \overline{m^i}\,[b^{-i}]$ can be expressed through two-dimensional linear complementarity problems.

1. linear constraints or

2. linear complementarity constraints.

In both cases, all feasible points can be enumerate efficiently [52, 53].

5.7 Numerical Example

Consider a two-player two-action game with the following characteristics:

1. the game is symmetric, both players are anti-coordination players with the same payoff matrix

$$P^i = \begin{array}{|c|c|}\hline 0.33 & 0.67 \\ \hline 1 & 0 \\ \hline \end{array}, \; i = 1, 2,$$

2. both players can have any utility function that is increasing and concave. Thus, the utility

matrix set (c.f. Proposition 5.4.2), for both $i = 1, 2$ is

$$\mathcal{U}^i = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in [0,1]^{2\times 2} \ : \ \begin{array}{c} d = 0,\, c = 1 \\ d \leq_\epsilon a \leq_\epsilon b \leq_\epsilon c \\ c + a - 2b \leq 0 \\ b + d - 2a \leq 0 \end{array} \right\},$$

where $a \leq_\epsilon b \iff a \leq b - \epsilon$ is the solver implementation of the strict inequality constraint $a < b$, for a parameter $\epsilon = 10^{-5}$.

From the set $\mathcal{U}^i$, we calculate the following quantities via solutions of linear problems as described in Lemma 5.5.2:

$$\begin{array}{ccc} \underline{g^i} = 0.5 & \text{and} & \overline{g^i} = 1, \\ \underline{m^i} = 0.67 & \text{and} & \overline{m^i} = 1, \\ \underline{m^i}\left[(0, \underline{g^i})\right] = 0.75 & \text{and} & \overline{m^i}\left[\left(\overline{g^i}, 1\right)\right] = 1. \end{array}$$

By solving the feasibility problem $(\mathcal{E}Q_P)$ for the partitions in Tables 5.2 and 5.3, we obtain the following three types of extreme strategies that satisfy the equilibrium conditions in Theorem 5.4.1, for both $i = 1, 2$ (since the game is symmetric):

a) $b^i = \left(0, \overline{b^i}\right)$ and $b^{-i} = (1, 1)$, where $\overline{b^i} \in [0, 0.5]$,

b) $b^i = (0.5, 1)$ and $b^{-i} = (0.5, 1)$.

c) $b^i = (0, 1)$ and $b^{-i} = (0.67, 1)$,

The first pair of extreme strategies, $b^i = \left(0, \overline{b^i}\right)$ and $b^{-i} = (1, 1)$, where $\overline{b^i} \in [0, 0.5]$, for some $i \in \{1, 2\}$, indicates that whenever one of the players plays the pure strategy 1 for all of her types, *i.e.* $b^{-i}\left(U^{-i}\right) = 1$ for all $U^{-i} \in \mathcal{U}^{-i}$, then there is an infinite set of extreme strategies for player i that will satisfy the conditions in Theorem 5.4.1. For the given extreme strategies, when both players are risk-neutral, *i.e.* $U^i = P^i$, $i = 1, 2$, then player i plays the pure strategy 0 and $-i$ the pure strategy 1. These strategies form a pure equilibrium if we treat this game as a complete information game. According to those strategies, the players receive a materialized utility of 1 and 0.67 respectively.

The second pair of equilibrium extreme strategies $\overline{b^i} \in [0, 0.5]$, $i = 1, 2$, results in equilibrium behavioural functions where, when both players are risk-neutral, they both play the mixed strategy 0.5.

When both players choose that strategy, they end up playing the unique mixed strategy equilibrium of the complete information game, where both players receive a payoff of 0.5.

Finally, the last pair of extreme strategies, $b^i = (0, 1)$ and $b^{-i} = (0.67, 1)$, for one of $i \in \{1, 2\}$, corresponds to an equilibrium where player i plays from the entire simplex Δ, and thus player $-i$ plays the max-min strategy for all of her types, *i.e.* player $-i$ has an equilibrium behavioural function $b^{-i}\left(U^{-i}\right) = m^i\left(U^{-i}\right)$ for all $U^{-i} \in \mathcal{U}^i$. When both players are risk-neutral, then player i plays the pure strategy 0 whilst player $-i$ plays her max-min strategy for that type, which, when $U^{-i} = P^{-i}$, is equal to $m^i\left(U^{-i}\right) = 0.75$. These strategies do not correspond to an equilibrium in the complete information game. When both players play according to those strategies, then player i receives a utility of 0.75 and player $-i$ a utility of 0.67.

5.8 Conclusion

In this chapter, we investigate games where we explicitly distinguish monetary rewards from utility values. We demonstrate that treating games with known monetary payoffs as complete information games, where those payoffs coincide with the players' utilities, leads to equilibrium strategies that do not necessarily maximize a player's corresponding expected utility when that player is not risk-neutral. Thus, even when the monetary payoffs of the game are known, players should still assume that they are engaging in an incomplete information game where they do not know their opponents' risk preferences and hence the utilities that they attach to their monetary rewards. The resulting games are expressed as incomplete private information games with no common prior, where each player can have any utility matrix from a specific utility matrix set. We show that these utility matrix sets have a polyhedral representation whenever we consider games where players are known to have non-satiated preferences or to be risk-averse. If the players are ambiguity-averse, they choose their behavioural functions so as to maximize their worst-case expected utility. We show that in this case, players either play a pure strategy, the max-min strategy or some convex combination of the two, depending on the opponent's behavioural function and their own utility matrix. By expressing the equilibrium condition in behavioural functions as an equivalent condition in extreme strategies, we develop an algorithm that provides all equilibria of the game via the solution of a finite number of linear feasibility and linear complementarity problems.

Chapter 6

Conclusions and Future Work

In this thesis we investigated three separate yet related problems from stochastic control theory, dynamic two-stage robust optimization and game theory. All three problems involve optimal decision making in the presence of uncertainty, and the main differences among them can be attributed to the source of the uncertainty and the distributional information we have about the uncertain parameters.

Our main contributions can be summarised as follows:

- We developed an efficient algorithm that bounds the performance loss of affine policies operating in discrete-time, finite-horizon, stochastic systems with expected quadratic costs and mixed linear state and input constraints. Designing an optimal control policy for such problems can lead to computational intractabilities, and a common resolution is to approximate the optimal solution by a sub-optimal but efficiently computable affine controller. We developed a methodology that provides a conservative estimation of the performance penalty incurred by utilizing these suboptimal controllers. The methodology is based on a dualization of the original problem and the imposition of a linear structure on the dual variables. The resulting lower bound is a conic optimization problem that can be solved efficiently.

- We designed an efficient algorithm to estimate the suboptimality of linear decision rules in two-stage robust optimization problems, where they have been shown to suffer a worst-case performance loss of the order $\Omega(\sqrt{m})$ for problems with m linear constraints. Our algorithm is based on a scenario selection technique, where the original problem is solved for only a finite subset

of the uncertain parameters. This subset is constructed according to the Lagrange multipliers associated with the computation of the linear adaptive decision rules. This algorithm does not rely on the existence of probabilistic information with regards to the uncertain parameters, and has outperformed known bounds, including the aforementioned dual bound, in the vast majority of problem instances with epistemic uncertainty.

- We developed an algorithm that enumerates all behavioural functions that are at equilibrium in a game where players face epistemic uncertainty regarding their opponent's utility functions. Traditionally, these games are solved as complete-information games where players are assumed to be risk-neutral, with a utility function that is positively affine in the monetary payoffs. This assumption imposes severe limitations on the problem structure, whilst our approach allows for these games to be formulated and solved as incomplete private information games where each player may have any increasing or increasing concave utility function. Our resulting algorithm can enumerate all equilibria of such games.

During our research, we have identified several interesting avenues for future work. Firstly, the bound developed in Chapter 3 relies on the restriction of the dual decision rules to be affine in structure. An extension would be to consider dual decision rules that are polynomial or piecewise-linear in structure. Whilst these extensions would still result in lower bounds to the original problem, they will be less conservative due to the higher flexibility of the dual rules when compared to the affine case. Additional extensions would be to consider $H - \infty$ control problems, where the objective function considers the worst-case cost incurred over the time. A second topic for future research is the development of a scenario-based lower bound for multi-stage stochastic optimization problems. Whilst our bound in Chapter 4 has been successful in two-stage problems, extending it to the multi-stage case is non-trivial, due to the additional non-anticipativity constraints. Further extensions could adapt the bound to problems involving segregated (piecewise) linear decision rules [41, 88]. Finally, an interesting but challenging problem for future research will be to extend our robust game methodology developed in Chapter 5 to the case of multi-action or even multi-player games. Unfortunately, the problem of finding a Nash equilibrium in two-player games is PPAD-complete even for normal form games with complete information [32]. With three or more players, the problem becomes FIXP-complete even for the zero-sum case [36]. However, one could potentially explore approximation schemes that have worked well with other classes of games, such as smoothing techniques [50] or other other methodologies that lead

to ϵ-equilibria [39].

Appendix A

Numerical Calculations of Chapter 3

A.1 Perfect State Measurements Example

In Section 3.6, we consider a time-invariant, discrete-time linear system which allows for perfect state measurements. The system has the following dynamics:

$$\boldsymbol{x}_{t+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}_t + \begin{bmatrix} 0.2 \\ 1 \end{bmatrix} \boldsymbol{u}_t + \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} \boldsymbol{\xi}_t, \ \boldsymbol{y}_t = \boldsymbol{x}_t,$$

where $t = 1, \ldots, T-1$. The initial state of the system is set to $\boldsymbol{x}_1 = (4.7, 0)^\top$. We assume that the $\boldsymbol{\xi}_t$ are independent and uniformly distributed on $[0, 2]$ for $t = 1, \ldots, T$, while $\boldsymbol{\xi}_0 = 1$ $\mathbb{P}$-a.s. Our objective is to minimize $\mathbb{E}\left[\boldsymbol{u}^\top \boldsymbol{u} + \boldsymbol{x}^\top \boldsymbol{x}\right]$ subject to

$$\left.\begin{array}{l} (-5,-5)^\top \leq \boldsymbol{x}_t \leq (5,5)^\top, \ t = 1, \ldots, T \\ (1,1)\, \boldsymbol{x}_t \leq 5, \ t = 1, \ldots, T \\ (1,-1)\, \boldsymbol{x}_t \leq 5, \ t = 1, \ldots, T \\ |\boldsymbol{u}_t| \leq 1, \ t = 1, \ldots, T-1 \end{array}\right\} \ \mathbb{P}\text{-a.s.}$$

We solve the approximate problems $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ as well as the trivial bound problems $\mathcal{P}_{\ell\ell}$ and $\mathcal{P}_{uu}$ a total of nine times, for each of the time horizons lengths $T = 2, \ldots, 10$. The resulting optimal values are illustrated in Figure 3.1, but are also presented here in Table A.1.

T	$\mathcal{P}^*_{ll}$	$\tilde{\mathcal{P}}^*_{l}$	$\tilde{\mathcal{P}}^*_{u}$	$\mathcal{P}^*_{uu}$
2	44.5892	44.6385	44.6385	44.6385
3	58.0287	65.1370	65.1495	65.1711
4	61.2732	81.8446	81.8534	82.9130
5	61.7444	92.4568	92.5867	97.0683
6	61.7942	97.3216	98.0970	106.9390
7	61.7958	98.4240	100.3132	112.4232
8	61.7999	98.6015	101.2557	114.9087
9	61.8107	98.6465	101.5933	116.0412
10	61.8254	98.6822	101.8700	117.1779

Table A.1: The optimal values of the upper and lower bounds of the perfect state measurement example, for each of the horizon lengths $T = 2, \ldots, 10$.

Additionally, in Figure 3.2, we illustrate the two regions X_ℓ and X_u of initial states x_1, for which a feasible policy can be found for $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ respectively, when the time horizon is $T = 8$. In this section, we will describe in detail the specific optimization problems that are solved in order to derive the demonstrated results.

A.1.1 The System Matrices

From the problem description one can readily verify that the state-dynamics matrices are

$$A_t := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B_t := \begin{bmatrix} 0.2 \\ 1 \end{bmatrix} \quad \text{and } C_t := \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$$

for $t = 1, \ldots, T-1$. Furthermore, as the system allows for perfect state measurements, we have $y_t = x_t$ for all $t = 1, \ldots, T-1$, which corresponds to the measurement matrices

$$D_t := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and } E_t := \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

for $t = 1, \ldots, T-1$.

We can now construct the compact-notation matrices B, C, D, E shown in equation (3.2). For the specific case where the time horizon length is $T = 4$, and given that the initial state $x_1 = (4.7, 0)^\top$,

the respective matrices are

$$B := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.2 & 0 & 0 \\ 1.0 & 0 & 0 \\ 1.2 & 0.2 & 0 \\ 1.0 & 1.0 & 0 \\ 2.2 & 1.2 & 0.2 \\ 1.0 & 1.0 & 1.0 \end{bmatrix}, \quad C := \begin{bmatrix} 4.7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 4.7 & 0.1 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 4.7 & 0.2 & 0.1 & 0 \\ 0 & 0.1 & 0.1 & 0 \\ 4.7 & 0.3 & 0.2 & 0.1 \\ 0 & 0.1 & 0.1 & 0.1 \end{bmatrix},$$

$$D := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad E := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

A.1.2 The Noise-process Matrices

We assume that the uncertain parameters $\boldsymbol{\xi}_t \in \mathbb{R}$ are independent and uniformly distributed on $[0, 2]$ for $t = 1, \ldots, T$, while $\boldsymbol{\xi}_0 = 1$ $\mathbb{P}$-a.s. Thus, for the case where the time horizon length is $T = 4$, the support Ξ can be defined as

$$\Xi := \left\{ \left(1, \xi^\top\right)^\top \in \mathbb{R}^5 \ : \ \xi \in \mathbb{R}^4, 0 \leq \xi \leq 2 \right\}$$

One can readily verify that the support Ξ can be defined in a manner consistent to the definition (3.3), as

$$\Xi = \left\{ \xi \in \mathbb{R}^5 \ : \ W_i \xi \succeq_{\mathcal{K}_p} 0, \ i = 1, 2, \ e_0^\top \xi = 1 \right\}$$

where

$$W_1 := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad W_2 := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 2 & 0 & -1 & 0 \\ 2 & 0 & 0 & -1 \end{bmatrix},$$

and p can be any norm from $\{1, \infty\}$.

Under the assumption that each of the ξ_t, $t = 1, \ldots, T-1$ is independent and continuously uniformly distributed on $[0, 2]$, the moment matrix (for $T = 4$) is derived analytically to be

$$M := \mathbb{E}\left[\xi\xi^\top\right] = \begin{bmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 1.0000 & 1.3333 & 1.0000 & 1.0000 \\ 1.0000 & 1.0000 & 1.3333 & 1.0000 \\ 1.0000 & 1.0000 & 1.0000 & 1.3333 \end{bmatrix}.$$

The first column and first row follows because $\xi_0 = 1$ almost surely. The remaining columns follow from the fact that for a multidimensional random variable $\xi \in \mathbb{R}^5$, the covariance matrix $\Sigma \in \mathbb{R}^{5\times 5}$ satisfies

$$\Sigma = \mathbb{E}\left[\xi\xi^\top\right] - \mathbb{E}\left[\xi\right]\mathbb{E}\left[\xi\right]^\top,$$

whilst for each ξ_t, $t = 1, \ldots, T-1$, $\mathbb{E}\left[\xi_t\right] = 1$, $\sigma\left(\xi_t, \xi_t\right) = \frac{1}{3}$, and the covariance between any pair of them is 0.

A.1.3 Calculating the Regions X_ℓ and X_u

In Figure 3.2, we illustrate the two regions X_ℓ and X_u of initial states x_1, for which a feasible policy can be found for $\tilde{\mathcal{P}}_\ell$ and $\tilde{\mathcal{P}}_u$ respectively, when the time horizon is $T = 8$. We calculate these regions by computing their vertices. This is done by repeatedly solving a pair of bi-linear feasibility problems, where the initial state x_1 is maximized along a specific direction, subject to constraints that ensure that $\tilde{\mathcal{P}}_\ell(x_1)$ and $\tilde{\mathcal{P}}_u(x_1)$ remain feasible.

Recall from Section 3.2, that in the Problem $\mathcal{P}$, only the matrix C depends on the initial state x_1, which is assumed to be known (c.f. Remark 3.2.1). Let $C\left(x_1\right)$ denote the matrix C, where the initial state x_1 is passed as a parameter. Then $G(x_1) = \left(DC\left(x_1\right) + E\right)$ is the matrix G when x_1 is passed as

a parameter. For given initial state x_1, Problem $\tilde{\mathcal{P}}_u(x_1)$ is feasible iff the constraints

$$\begin{aligned}
&(F_u + F_x B)QG\,(x_1) + F_x C\,(x_1) + F_s S - he_0^\top = 0\\
&S = \mu e_0^\top + \textstyle\sum_{i=1}^{l} \Lambda_i^\top W_i\\
&\Lambda_i \succeq_{\mathcal{K}_p^*} 0,\ i = 1 \dots l,\ \mu \geq 0
\end{aligned}$$

can be satisfied for some $Q \in \mathcal{U}$, $S \in \mathbb{R}^{N_s \times N_\xi}$, $\Lambda_i \in \mathbb{R}^{N_\xi \times N_s}$, $i = 1 \dots l$ and $\mu \in \mathbb{R}^{N_s}$.

We can express x_1 using the polar coordinate system. For a two-dimensional $x_1 \in \mathbb{R}^2$ we have

$$x_1 = \alpha \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix}$$

for some $\alpha \in \mathbb{R}$ and some radian $\theta \in [0, 2\pi]$. Thus, in order to calculate the extremes of region X_u along a specific direction θ, we solve the optimization problem

$$\begin{aligned}
\max_{\alpha,Q,S,\Lambda_i}\quad & \alpha\\
\text{s.t.}\quad & \alpha \in \mathbb{R},\ x_1 = \alpha \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix}\\
& Q \in \mathcal{U},\ S \in \mathbb{R}^{N_s \times N_\xi},\ \Lambda_i \in \mathbb{R}^{N_\xi \times N_s},\ \mu \in \mathbb{R}^{N_s}\\
& (F_u + F_x B)QG\,(x_1) + F_x C\,(x_1) + F_s S - he_0^\top = 0\\
& S = \mu e_0^\top + \textstyle\sum_{i=1}^{l} \Lambda_i^\top W_i\\
& \Lambda_i \succeq_{\mathcal{K}_p^*} 0,\ i = 1 \dots l,\ \mu \geq 0.
\end{aligned}$$

We can then find an approximation of the entire region X_u by solving the problem repeatedly for $\theta = \kappa\pi$, where $\kappa = 0.01, 0.02, \dots, 1.99, 2.00$. Observe that the bi-linearities appear only in the first term of the joint state and control constraint $(F_u + F_x B)QG\,(x_1)$, where we optimize over both Q and x_1. However, due to its low dimensions, the problem can be solved to optimality using currently available global optimization tools.

A similar process is followed in order to calculate the edges of region X_ℓ. The optimization problem

that we solve is

$$\begin{aligned}
\max_{\alpha,Q,S} \quad & \alpha \\
\text{s.t.} \quad & \alpha \in \mathbb{R},\ x_1 = \alpha \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix} \\
& Q \in \mathcal{U},\ S \in \mathbb{R}^{N_s \times N_\xi} \\
& (F_u + F_x B) Q G(x_1) + F_x C(x_1) + F_s S - h e_0^\top = 0 \\
& W_i M S^\top \succeq_{\mathcal{K}_p} 0,\ i = 1 \dots l \\
& e_0^\top M S^\top \geq 0.
\end{aligned}$$

for, as before, $\theta = \kappa\pi$, $\kappa = 0.01, 0.02, \dots, 1.99, 2.00$.

A.2 Imperfect State Measurements Example

For the imperfect state measurement example described in Section 3.6, the problems $\tilde{\mathcal{P}}_\ell$, $\tilde{\mathcal{P}}_u$ and $\mathcal{P}_{\ell\ell}$ and $\mathcal{P}_{uu}$ where solved a total of 19 times, for each of the time horizon lengths $T = 2, \dots, 20$. The results are illustrated in Figure 3.3, but we also show them here in Table A.2.

In this example, the uncertain parameters are uniformly distributed in a circle with center at the origin and radius equal to one. The moment matrix can be analytically calculated because for each $\xi_t \in \mathbb{R}^3$, $t = 1, \dots, T-1$, $\mathbb{E}[\xi_t] = 0$, $\sigma(\xi_t, \xi_t) = \frac{1}{3T-1}$ (using the formula $\sigma = \frac{r^2}{d+2}$, where r is the radius of the support and d is the dimension of the uncertain parameters, which is equal to $N_\xi - 1 = 3(T-1)$, i.e. we exclude the dummy variable $\xi_0 = 1$ $\mathbb{P}$-a.s.). Finally, as the uncertain parameters ξ_t, $t = 1, \dots, T-1$ are independent, the covariance between them is 0. Thus, for a time horizon length equal to $T = 3$, the moment matrix is

$$M := \mathbb{E}\left[\xi\xi^\top\right] = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.125 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.125 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.125 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.125 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.125 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.125
\end{bmatrix}.$$

T	$\mathcal{P}^*_{uu}$	$\tilde{\mathcal{P}}^*_u$	$\tilde{\mathcal{P}}^*_l$	$\mathcal{P}^*_{ll}$
2	2.0000	2.4000	2.4000	2.4000
3	2.4631	3.2640	3.2640	3.3381
4	2.5643	3.7612	3.7612	4.2007
5	2.5828	4.0972	4.0972	5.4400
6	2.5865	4.3335	4.3335	Inf
7	2.5873	4.5044	4.5044	Inf
8	2.5874	4.6321	4.6321	Inf
9	2.5875	4.7307	4.7307	Inf
10	2.5875	4.8090	4.8090	Inf
11	2.5875	4.8726	4.8726	Inf
12	2.5875	4.9253	4.9253	Inf
13	2.5875	4.9697	4.9697	Inf
14	2.5875	5.0076	5.0076	Inf
15	2.5875	5.0404	5.0404	Inf
16	2.5875	5.0689	5.0689	Inf
17	2.5875	5.0940	5.0940	Inf
18	2.5875	5.1163	5.1163	Inf
19	2.5875	5.1362	5.1362	Inf
20	2.5875	5.1541	5.1541	Inf

Table A.2: The optimal values of the upper and lower bounds of the imperfect state measurement example, for each of the horizon lengths $T = 2, \ldots, 20$.

Appendix B

Numerical Calculations of Chapter 4

B.1 Randomly Generated Problems

In Section 4.5.2 we assess the performance of a number of different lower bounds based on a sample of randomly generated problem instances. The instances are generated to adhere to the assumptions described in [17], where it has been shown that linear decision rules incur a worst-case performance loss by a factor of 4. We will now illustrate step-by-step how such problems are generated and how the performance metrics in Tables 4.2 and 4.3 are calculated.

In Section 4.5.2 the problem dimensions for each randomly generated instance are set to $k = 16$, $m = 16$, $n_1 = 3$ and $n_2 = 5$. For the sake of readability, here we consider a problem with dimensions $k = 4$, $m = 10$, $n_1 = n_2 = 3$. The problem is generated according to the following steps:

- The uncertainty set Ξ is defined as $\Xi := \left\{(1, \xi)^\top \in \mathbb{R}^{k+1} \ : \ \|\xi\|_p \leq 1\right\}$, where for our instance $p = 1$. This corresponds to a diamond-shaped uncertainty set with vertices (presented as the

columns of the matrix):

$$[Z_v] := \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}.$$

- The elements of the matrices $A \in \mathbb{R}^{m\times n_1}$ and $B \in \mathbb{R}^{m\times n_2}$ are uniformly sampled from the interval $[-5, 5]$. In this problem instance, the sample matrices are

$$A := \begin{bmatrix} -1.68 & -1.86 & -4.24 \\ 0.19 & -0.66 & 2.23 \\ -2.33 & 2.08 & 2.48 \\ -4.54 & -2.50 & 3.22 \\ -2.46 & 1.26 & 2.25 \\ -2.14 & -3.61 & -1.88 \\ 1.92 & -4.07 & -2.27 \\ -3.91 & 2.95 & 3.57 \\ -1.04 & -1.96 & 1.58 \\ -1.19 & 3.95 & -3.53 \end{bmatrix} \quad \text{and } B := \begin{bmatrix} -1.44 & -2.14 & -2.49 \\ 3.72 & -3.10 & -1.87 \\ 3.71 & -3.56 & -4.80 \\ 3.21 & -0.11 & 4.11 \\ 2.49 & -3.76 & -0.80 \\ -0.31 & -1.09 & 1.54 \\ 3.12 & 2.68 & 3.07 \\ -4.57 & -2.68 & -3.62 \\ -4.92 & 3.52 & -2.83 \\ -2.65 & -0.21 & 0.26 \end{bmatrix}.$$

- The matrix $C \in \mathbb{R}^{m\times(k+1)}$ is randomly generated so that each row of C is in the dual cone of Ξ, namely $C^\top \succeq_{\mathcal{K}^*} 0$, which guarantees that $C\xi \geq 0$ for all $\xi \in \Xi$. This is achieved as follows:

  1. We generate m random vectors $\lambda_i \in \mathbb{R}^k$, $i = 1, \ldots, m$, drawn from the interval $[-5, +5]$ according to the uniform distribution.
  2. For each random vector $\lambda_i \in \mathbb{R}^k$, $i = 1, \ldots, m$, we set $\tau_i = \|\lambda_i\|_{p^*} + \epsilon_i$, where p^* is the dual norm of p (thus for $p = 1$ we have $p^* = \infty$), and ϵ_i is a random scalar uniformly distributed on the interval $[0, 10]$.
  3. The matrix C is then defined as

$$C := \begin{bmatrix} \tau_1 & \lambda_1^\top \\ \vdots & \vdots \\ \tau_m & \lambda_m^\top \end{bmatrix}.$$

Note that each row of C is $\left(\tau_i, \lambda_i^\top\right) \in \mathbb{R}^{k+1}$, $i = 1, \ldots, m$, with $\|\lambda_i\|_{p^*} \leq \tau_i$. Thus, by definition $C^\top \succeq_{\mathcal{K}^*} 0$.

For our problem the sample matrix C is

$$C := \begin{bmatrix} 9.06 & -1.99 & 1.32 & -0.73 & -3.10 \\ 5.87 & -1.23 & -3.78 & -0.20 & -1.74 \\ 6.14 & 1.30 & -1.60 & -2.53 & -4.87 \\ 8.97 & 1.30 & -0.87 & -0.58 & 3.74 \\ 10.01 & 1.11 & -2.40 & -0.25 & -4.75 \\ 11.35 & -1.71 & 3.56 & 3.56 & -0.44 \\ 5.09 & -1.52 & 0.51 & -4.74 & 3.54 \\ 4.30 & -3.73 & -1.79 & 1.53 & 0.37 \\ 10.72 & 3.85 & 2.81 & -2.69 & 0.21 \\ 5.93 & -0.05 & 2.72 & 1.52 & -0.04 \end{bmatrix}.$$

- The objective function vectors $c \in \mathbb{R}^{n_1}$ and $d \in \mathbb{R}^{n_2}$ are generated according to the mappings $c = -A^\top \mu$ and $d = -B^\top \mu$. The vector $\mu \in \mathbb{R}^m$ is repeatedly sampled from $[0, 1]^m$ until both $c \geq 0$ and $d \geq 0$ are satisfied. The non-negativity restrictions for c and d are necessitated by the assumptions in [17]. The existence of a vector μ satisfying the above is required by the postulated dual feasibility of the original problem $\mathcal{P}$. In our problem instance, $\mu := (3.71, 1.75, 0.66, 1.85, 2.87, 3.91, 4.84, 4.35, 2.91, 2.86)^\top$, hence

$$c := \begin{pmatrix} 45.40 \\ 23.11 \\ 6.06 \end{pmatrix} \text{ and } d := \begin{pmatrix} 11.19 \\ 20.03 \\ 12.69 \end{pmatrix}.$$

We obtain an upper bound on $\mathcal{P}$ by solving the primal linear decision rule problem $\tilde{\mathcal{U}}$

$$\begin{array}{ll} \inf\limits_{x,Y} & c^\top x + t \\ \text{s.t.} & x \in \mathbb{R}^{n_1},\, Y \in \mathbb{R}^{n_2 \times (k+1)} \\ & \left(d^\top Y - t e_0^\top\right)^\top \preceq_{\mathcal{K}^*} 0 \\ & \left(A x e_0^\top + BY - C\right)^\top \preceq_{\mathcal{K}^*} 0 \end{array} \qquad (\tilde{\mathcal{U}})$$

resulting in an optimal value $\tilde{\mathcal{U}}^* = -111.46$.

For the lower bounds, we solve six separate problems. Three of these problems are instances of the scenario-based lower bound $\mathcal{P}(Z)$, for three different finite uncertainty sets $Z \subset \Xi$. Specifically, when the finite set $Z \subset \Xi$ has l elements (denoted by $\xi_i \in Z$, $i = 1, \ldots, l$), we solve the finite optimization problem

$$\begin{array}{ll} \inf\limits_{x, y_i} & c^\top x + t \\ \text{s.t.} & x \in \mathbb{R}^{n_1},\, y_i \in \mathbb{R}^{n_2},\, i = 1, \ldots, l \\ & \left.\begin{array}{l} d^\top y_i \leq t \\ Ax + By_i \leq C\xi_i \end{array}\right\} i = 1, \ldots, l \end{array} . \qquad (\mathcal{P}(Z))$$

for the three different sets:

a) $Z = \Delta$, where $|\Delta| \leq m + 1$ is the critical set of dual variables derived from the problem $\tilde{\mathcal{U}}$, as described in Section 4.4.2. For our instance, the set Δ is (presented in matrix form)

$$[\Delta] := \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.55 & 0.38 & 0 & 0.28 & 0 & 0 & 0.37 & -0.37 & 0 \\ 0 & -0.10 & 0 & 0.19 & -1 & 0 & 0.20 & 07 & -1 \\ -0.45 & 0 & 0 & -0.07 & 0 & 1 & 0 & 0.63 & 0 \\ 0 & 0.52 & -1 & -0.06 & 0 & 0 & -0.43 & -0 & 0 \end{bmatrix}.$$

b) $Z = \{\xi_m\}$, where ξ_m is a solution of Problem (4.7). In this instance, $\xi_m = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

c) $Z = Z_v$ the set of vertices of the support.

For completeness, we also provide a set of dual linear decision rule bounds $\mathcal{L}^*(\mathbb{P})$, parametrized by three different uncertainty distributions. Specifically, we solve the optimization problem (described in

more detail in Chapter 3)

$$\begin{array}{ll} \inf\limits_{x,Y} & c^\top x + t \\ \text{s.t.} & x \in \mathbb{R}^{n_1},\, Y \in \mathbb{R}^{n_2 \times (k+1)} \\ & \left(d^\top Y - t e_0^\top\right) \mathbb{E}_{\mathbb{P}}\left[\xi\xi^\top\right] 0 \\ & \left(A x e_0^\top + BY - C\right) \mathbb{E}_{\mathbb{P}}\left[\xi\xi^\top\right] \leq 0. \end{array} \tag{$\mathcal{L}(\mathbb{P})$}$$

for the following distributions:

a) $\mathbb{P} = \mathbb{P}_\Xi$ the uniform distribution on the support Ξ. The corresponding second-order moment is

$$\mathbb{E}_{\mathbb{P}_\Xi}\left[\xi\xi^\top\right] := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.0667 & 0 & 0 & 0 \\ 0 & 0 & 0.0667 & 0 & 0 \\ 0 & 0 & 0 & 0.0667 & 0 \\ 0 & 0 & 0 & 0 & 0.0667 \end{bmatrix},$$

b) $\mathbb{P} = \mathbb{P}_\Delta$ the uniform distribution on the set Δ, where

$$\mathbb{E}_{\mathbb{P}_\Delta}\left[\xi\xi^\top\right] := \begin{bmatrix} 1 & 0.1340 & -0.1893 & 0.1230 & -0.1106 \\ 0.1340 & 0.0883 & 0.0101 & -0.0556 & 0.0018 \\ -0.1893 & 0.0101 & 0.2319 & -0.0015 & -0.0172 \\ 0.1230 & -0.0556 & -0.0015 & 0.1785 & 0.0007 \\ -0.1106 & 0.0018 & -0.0172 & 0.0007 & 0.1628 \end{bmatrix}, \text{ and}$$

c) $\mathbb{P} = \mathbb{P}_{Z_v}$ the uniform distribution on the set of vertices Z_v of the support, where

$$\mathbb{E}_{\mathbb{P}_{Z_v}}\left[\xi\xi^\top\right] := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 & 0 \\ 0 & 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0 & 0.25 & 0 \\ 0 & 0 & 0 & 0 & 0.25 \end{bmatrix}$$

The optimal values of the six lower bounds and the upper bound are shown in Table B.1. The gap is defined as the difference between each lower bound with the upper bound, whilst the relative gap (%

Problem:	$\tilde{\mathcal{U}}$	$\mathcal{P}(Z_v)$	$\mathcal{P}(\Delta)$	$\mathcal{P}(\{\xi_w\})$	$\mathcal{L}(\mathbb{P}_{Z_v})$	$\mathcal{L}(\mathbb{P}_{\Delta})$	$\mathcal{L}(\mathbb{P}_{\Xi})$
Opt. Value	-111.46	-113.98	-119.24	-116.87	-123.43	-127.29	-138.04
Gap	-	2.52	7.78	5.41	11.97	15.83	26.58
%Gap	-	2.2%	6.5%	4.6%	9.7%	12%	19%

Table B.1: The optimal values for the problem instance described in Appendix B.1.

Gap) is that difference divided by its respective lower bound. As expected, $\mathcal{P}^*(Z_v)$ is the tightest lower bound, since (from Proposition 4.4.2) $\mathcal{P}^*(Z_v) = \mathcal{P}^*$. However, problem $\mathcal{P}(Z_v)$ relies on the enumeration of the vertices of the support, which is usually not feasible in practice. In this problem instance, one can immediately verify that all scenario based lower bounds outperform the dual linear decision rule bounds.

Bibliography

[1] Aghassi, M., and Bertsimas, D. Robust game theory. *Mathematical Programming 107*, 1 (2006), 231–273.

[2] Allais, M. Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine. *Econometrica 21*, 4 (1953), 503–546.

[3] Anderson, B. D. O., and Moore, J. B. *Optimal Control: Linear Quadratic Methods*. Prentice Hall, 1990.

[4] Arrow, K. J. *Essays in the Theory of Risk-bearing*, vol. 1. Markham Publishing Company, Chicago, 1971.

[5] Ash, R. B. *Real Analysis and Probability*. Academic Press, New York, 1972.

[6] Barsky, R., Juster, T., Kimball, M., and Shapiro, M. D. Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics 112*, 2 (1997), 537–579.

[7] Bemporad, A., Borrelli, F., and Morari, M. Min-max control of constrained uncertain discrete-time linear systems. *IEEE Transactions on Automatic Control 48*, 9 (2003), 1600– 606.

[8] Ben-Tal, A., Boyd, S., and Nemirovski, A. Control of uncertainty affected discrete time linear systems via convex programming. Tech. rep., Minerva Optimization Center, Technion, 2005.

[9] Ben-Tal, A., Boyd, S., and Nemirovski, A. Extending scope of robust optimization: Comprehensive robust counterparts of uncertain problems. *Mathematical Programming 107*, 1–2 (2006), 63–89.

[10] Ben-Tal, A., El Ghaoui, L., and Nemirovski, A. *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press, 2009.

[11] Ben-Tal, A., Goryashko, A., Guslitzer, E., and Nemirovski, A. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming 99*, 2 (2004), 351–376.

[12] Ben-Tal, A., and Nemirovski, A. Robust convex optimization. *Mathematics Of Operations Research 23*, 4 (1998), 769–805.

[13] Ben-Tal, A., and Nemirovski, A. Robust solutions of uncertain linear programs. *Operation Research Letters 25*, 1 (1999), 1–13.

[14] Bernoulli, D. Exposition of a new theory on the measurement of risk. *Econometrica 22*, 1 (1954), 23–36. Originally published in 1738; translated by Dr. Louise Sommer.

[15] Bertsekas, D., and Rhodes, I. On the minimax reachability of target sets and target tubes. *Automatica 7*, 2 (1971), 233–247.

[16] Bertsimas, D., and Caramanis, C. Adaptability via sampling. In *46th IEEE Conference on Decision and Control* (2007), pp. 4717–4722.

[17] Bertsimas, D., and Goyal, V. On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical Programming 134*, 2 (2012), 491–531.

[18] Bertsimas, D., Iancu, D., and Parrilo, P. Optimality of affine policies in multi-stage robust optimization. *Mathematics Of Operations Research 35*, 2 (2010), 363–394.

[19] Bertsimas, D., Iancu, D., and Parrilo, P. A hierarchy of near-optimal policies for multi-stage adaptive optimization. *IEEE Transactions on Automatic Control 56*, 12 (2011), 2809–2824.

[20] Binswanger, H. P. Attitudes toward risk: Experimental measurement in rural India. *American Journal of Agricultural Economics 62*, 3 (1980), 395–407.

[21] Birge, J., and Louveaux, F. V. *Introduction to Stochastic Programming*. Springer, 1997.

[22] Boyd, S., and Vandenberghe, L. *Convex Optimization*. Cambridge University Press, 2004.

[23] Calafiore, G., and Campi, M. C. Uncertain convex programs: Randomized solutions and confidence levels. *Mathematical Programming 102*, 1 (2005), 25–46.

[24] Cambanis, S., Huang, S., and Simons, G. On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis 11*, 3 (1981), 368 – 385.

[25] Camerer, C. Individual decision making. In *Handbook of Experimental Economics*, J. Kagel and A. Roth, Eds. Princeton University Press, 1995, pp. 587–704.

[26] Camerer, C., and Weber, M. Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty 5*, 4 (1992), 325–370.

[27] Charnes, A., and Cooper, W. W. Programming with linear fractional functionals. *Naval Research Logistics Quarterly 9*, 3-4 (1962), 181–186.

[28] Chen, X., Sim, M., and Sun, P. A robust optimization perspective on stochastic programming. *Operations Research 55*, 6 (2007), 1058–1071.

[29] Chen, X., Sim, M., Sun, P., and Zhang, J. A linear decision-based approximation approach to stochastic programming. *Operations Research 56*, 2 (2008), 344 –357.

[30] Chmielewski, M. Elliptically symmetric distributions: A review and bibliography. *International Statistical Review 49*, 1 (1981), 67–74.

[31] Choi, S., Fisman, R., Gale, D., and Kariv, S. Consistency and heterogeneity of individual behavior under uncertainty. *The American Economic Review 97*, 5 (2007), 1921–1938.

[32] Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. H. The complexity of computing a nash equilibrium. *SIAM Journal on Computing 39*, 1 (2008), 195–259.

[33] Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey. *IZA Discussion Paper No. 17300* (2005). Available at SSRN.

[34] Dow, J., and Werlang, S. Nash equilibrium under Knightian uncertainty: Breaking down backward induction. *Journal of Economic Theory 64*, 2 (1994), 305–324.

[35] Ellsberg, D. Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics 75*, 4 (1961), 643–669.

[36] Etessami, K., and Yannakakis, M. On the complexity of nash equilibria and other fixed points. In *48th Annual IEEE Symposium on Foundations of Computer Science(FOCS-07)* (2007), pp. 113–123.

[37] Faísca, N. P., Kouramas, K. I., Saraiva, P. M., Rustem, B., and Pistikopoulos, E. N. A multi-parametric programming approach for constrained dynamic programming problems. *Optimization Letters 2*, 2 (2008), 267–280.

[38] Friend, I., and Blume, M. E. The demand for risky assets. *The American Economic Review 65*, 5 (1975), 900–922.

[39] Ganzfried, S., and Sandholm, T. Computing equilibria in multiplayer stochastic games of imperfect information. In *21st International Jont Conference on Artifical Intelligence* (2009), pp. 140–146.

[40] Garstka, S. J., and Wets, R. J. On decision rules in stochastic programming. *Mathematical Programming 7*, 1 (1974), 117–143.

[41] Georghiou, A., Wiesemann, W., and Kuhn, D. Generalized decision rule approximations for stochastic programming via liftings. Tech. rep., Imperial College London, 2010. Available on Optimization Online.

[42] Gilboa, I. Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics 16*, 1 (1987), 65–88.

[43] Gilboa, I., and Schmeidler, D. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*, 2 (1989), 141–153.

[44] Goeree, J. K., Holt, C., and Palfrey, T. Risk averse behavior in generalized matching pennies games. *Games and Economic Behaviour 45*, 1 (2003), 97–113.

[45] Goh, J., and Sim, M. Distributionally robust optimization and its tractable approximations. *Operations Research 58*, 4-Part-1 (2010), 902–917.

[46] Goulart, P. J., and Kerrigan, E. C. Output feedback receding horizon control of constrained systems. *International Journal of Control 80*, 1 (2007), 8–20.

[47] Goulart, P. J., and Kerrigan, E. C. Input-to-state stability of robust receding horizon control with an expected value cost. *Automatica 44*, 4 (2008), 1171–1174.

[48] Goulart, P. J., Kerrigan, E. C., and Maciejowski, J. Optimization over state feedback policies for robust control with constraints. *Automatica 42*, 4 (2006), 523 – 533.

[49] Harsanyi, J. C. Games with incomplete information played by "Bayesian" players, Part I–III. *Management Science 14*, 3, 5, 7 (1967, 1968), 159–182, 320–334, 486–502.

[50] Hoda, S., Gilpin, A., Pena, J., and Sandholm, T. Smoothing techniques for computing nash equilibria of sequential games. *Mathematics of Operations Research 35*, 2 (2010), 494–512.

[51] Holt, C. A., and Laury, S. K. Risk aversion and incentive effects. *The American Economic Review 92*, 5 (2002), 1644–1655.

[52] Horst, R., Pardalos, P., and Romeijn, H. E. *Handbook of Global Optimization*, vol. 2. Springer, 2002.

[53] Horst, R., Pardalos, P., and Thoai, N. V. *Introduction to Global Optimization*. Springer, 2000.

[54] Hyafil, N., and Boutilier, C. Regret minimizing equilibria and mechanisms for games with strict type uncertainty. In *20th Conference on Uncertainty in Artificial Intelligence* (2004), pp. 268–277.

[55] Kahneman, D., and Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica 47*, 2 (1979), 263–291.

[56] Kajii, A., and Ui, T. Incomplete information games with multiple priors. *Japanese Economic Review 56*, 3 (2005), 332–351.

[57] Kall, P., and Wallace, S. W. *Stochastic Programming*. Wiley, 1994.

[58] Klibano, P. Uncertainty, decision, and normal form games, 1993. Manuscript, MIT.

[59] Knight, F. *Risk, Uncertainty and Profit*. Houghton Mifflin, Boston, 1921.

[60] Kuhn, D., Wiesemann, W., and Georghiou, A. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming 130*, 1 (2011), 177–209.

[61] Lee, Y., and Kouvaritakis, B. Robust receding horizon predictive control for systems with uncertain dynamics and input saturation. *Automatica 36*, 10 (2000), 1497–1504.

[62] Lo, K. C. Equilibrium in beliefs under uncertainty. *Journal of Economic Theory 71*, 2 (1996), 443–484.

[63] Löfberg, J. *Minimax Approaches to Robust Model Predictive Control*. PhD thesis, Linköping University, Apr. 2003.

[64] Luenberger, D. G. A preference foundation for log mean-variance criteria in portfolio choice problems. *Journal of Economic Dynamics and Control 17*, 5 (1993), 887–906.

[65] Marinacci, M. Ambiguous games. *Games and Economic Behaviour 31*, 2 (2000), 191–219.

[66] Maschler, M., Zamir, S., and Solan, E. *Game Theory*. Cambridge University Press, 2007.

[67] Mayne, D., Seron, M., and Raković, S. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica 41*, 2 (2005), 219–24.

[68] Mayne, D. Q., Rawlings, J. B., Rao, C. V., and Scokaert, P. O. M. Constrained model predictive control: Stability and optimality. *Automatica 36*, 6 (2000), 789–814.

[69] Mertens, J.-F., and Zamir, S. Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory 14*, 1 (1985), 1–29.

[70] Morris, S. The common prior assumption in economic theory. *Economics and Philosophy 11*, 2 (1995), 227–227.

[71] Myerson, R. B. *Game Theory: Analysis of Conflict*. Harvard University Press, 2013.

[72] Nash, J. F. Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences of the United States of America 36*, 1 (1950), 48–49.

[73] Nemirovskii, A. Several NP-hard problems arising in robust stability analysis. *Mathematics of Control, Signals and Systems 6*, 2 (1993), 99–105.

[74] Nishimura, R., Hayashi, S., and Fukushima, M. Robust Nash equilibria in N-person non-cooperative games: Uniqueness and reformulation. *Pacific Journal of Optimization 5*, 2 (2009), 237–259.

[75] Osborne, M. J. *A Course in Game Theory*. The MIT Press, Cambridge, 1994.

[76] Pistikopoulos, E., Georgiadis, M., and Dua, V. *Multi-Parametric Model-Based Control: Theory and Applications*. Wiley-VCH, 2007.

[77] Pratt, J. W. Risk aversion in the small and in the large. *Econometrica 32*, 1/2 (1964), 122–136.

[78] Pulford, B. D., and Colman, A. M. Ambiguous games: Evidence for strategic ambiguity aversion. *The Quarterly Journal of Experimental Psychology 60*, 8 (2007), 1083–1100.

[79] Savage, L. J. *The Foundations of Statistics*. Wiley, New York, 1954.

[80] Schmeidler, D. Subjective probability and expected utility without additivity. *Econometrica 57*, 3 (1989), 571–587.

[81] Scokaert, P., and Mayne, D. Min-max feedback model predictive control for constrained linear systems. *IEEE Transactions on Automatic Control 43*, 8 (1998), 1136–1142.

[82] Shapiro, A., and Nemirovski, A. On complexity of stochastic programming problems. In *Continuous optimization*. Springer, 2005, pp. 111–146.

[83] Shapiro, A., and Ruszczyński, A. P. *Stochastic Programming*. Elsevier, 2003.

[84] Skaf, J., and Boyd, S. Design of affine controllers via convex optimization. *IEEE Transactions on Automatic Control 55*, 11 (2010), 2476–2487.

[85] Starmer, C. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature 38*, 2 (2000), 332–382.

[86] van Hessem, D., and Bosgra, O. A conic reformulation of model predictive control including bounded and stochastic disturbances under state and input constraints. In *41st IEEE Conference on Decision and Control* (2002), vol. 4, pp. 4643–4648.

[87] von Neumann, J., and Morgenstern, O. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[88] Wang, C., Ong, C. J., and Sim, M. Model predictive control using segregated disturbance feedback. *IEEE Transactions on Automatic Control 55*, 4 (2010), 831–840.

[89] Wiesemann, W., Kuhn, D., and Rustem, B. Robust resource allocations in temporal networks. *Mathematical Programming 135*, 1-2 (2012), 437–471.

[90] Xue, G., and Ye, Y. An efficient algorithm for minimizing a sum of p-norms. *SIAM Journal on Optimization 10*, 2 (1999), 551–579.

[91] Youla, D., Jabr, H., and Bongiorno, J. Modern Wiener-Hopf design of optimal controllers: Part II. *IEEE Transactions on Automatic Control AC-21*, 3 (1976), 319–338.