Adaptive Robust Optimization with Applications in Inventory and Revenue Management

by

Dan Andrei Iancu

B.S., Yale University (2004)
S.M., Harvard University (2006)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2010

Author . . . .
Sloan School of Management
June 8, 2010

Certified by . . . .
Dimitris Bertsimas
Boeing Leaders for Global Operations Professor
Thesis Supervisor

Certified by . . . .
Pablo A. Parrilo
Professor
Thesis Supervisor

Accepted by . . . .
Patrick Jaillet
Co-director, Operations Research Center

Adaptive Robust Optimization with Applications in Inventory and Revenue Management

by

Dan Andrei Iancu

Submitted to the Sloan School of Management
on June 8, 2010, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Operations Research

Abstract

In this thesis, we examine a recent paradigm for solving dynamic optimization problems under uncertainty, whereby one considers decisions that depend directly on the sequence of observed disturbances. The resulting policies, called *recourse decision rules*, originated in Stochastic Programming, and have been widely adopted in recent works in Robust Control and Robust Optimization; the specific subclass of affine policies has been found to be tractable and to deliver excellent empirical performance in several relevant models and applications.

In the first chapter of the thesis, using ideas from polyhedral geometry, we *prove* that disturbance-affine policies are optimal in the context of a one-dimensional, constrained dynamical system. Our approach leads to policies that can be computed by solving a single linear program, and which bear an interesting decomposition property, which we explore in connection with a classical inventory management problem. The result also underscores a fundamental distinction between robust and stochastic models for dynamic optimization, with the former resulting in qualitatively simpler problems than the latter.

In the second chapter, we introduce a hierarchy of polynomial policies that are also directly parameterized in the observed uncertainties, and that can be efficiently computed using semidefinite optimization methods. The hierarchy is asymptotically optimal and guaranteed to improve over affine policies for a large class of relevant problems. To test our framework, we consider two problem instances arising in inventory management, for which we find that quadratic policies considerably improve over affine ones, while cubic policies essentially close the optimality gap.

In the final chapter, we examine the problem of dynamically pricing inventories in multiple items, in order to maximize revenues. For a linear demand function, we propose a distributionally robust uncertainty model, argue how it can be constructed from limited historical data, and show how pricing policies depending on the observed model misspecifications can be computed by solving second-order conic or semidefinite optimization problems. We calibrate and test our model using both synthetic data, as well as real data from a large US retailer. Extensive Monte-Carlo simulations show that adaptive robust policies considerably improve over open-loop formulations, and are competitive with popular heuristics in the literature.

Thesis Supervisor: Dimitris Bertsimas

Title: Boeing Leaders for Global Operations Professor

Thesis Supervisor: Pablo A. Parrilo

Title: Professor

Acknowledgments

My deepest gratitude goes to my two advisers and mentors, Dimitris Bertsimas and Pablo Parrilo. It is hard to imagine this thesis without their insightful questions, shrewd suggestions, and perpetual patience and words of encouragement in tough times. Thank you, Dimitris and Pablo, for being a beacon of energy and inspiration, for guiding my steps well beyond the academic life, and for believing in me! I owe both of you a debt of gratitude larger than I can express here.

I was also extremely fortunate to have Georgia Perakis and Retsef Levi as my two other committee members. Apart from providing me with invaluable research feedback, they have been a continuous source of support and advice. Thank you, Georgia, for allowing me to be your teaching assistant (twice!), and for making the job-search process infinitely more tolerable with your incredible kindness and warm thoughts! Thank you, Retsef, for teaching me everything I know about inventory management, and for your candid comments about research and academia in general.

I am also indebted to Rama Ramakrishnan for his support in obtaining the data set used in the last chapter of the dissertation, and for sharing his extensive industry experience and insights.

My academic life was also shaped by numerous wonderful teachers - in particular, I would like to acknowledge Aharon Ben-Tal and Andreas Schulz, for lecturing three of the best classes in optimization that I have every taken, and Paul Flondor, Ana Niţă, Liliana Mihăilă, and Mihaela Popa, for nurturing my early interests in mathematics and physics.

During my four years at MIT, the Operations Research Center has undoubtedly been a "home away from home". The co-directors - Cindy Barnhart and Dimitris Bertsimas - have been great managers, considerate and mindful of all issues pertaining to student life, while the administrative staff - Paulette Mosley, Laura Rose and Andrew Carvalho - have always made sure the center runs smoothly, and the students have everything they need to focus on work. However, by far, the best thing about the ORC has been the wonderful group of colleagues and friends, with whom I have shared many hours of work and fun, alike - Thibault Leguen, Diana Michalek, Martin Quinteros, Jason Acimovic, Jonathan Kluberg, Adrian Becker, Alex Rikun, Gareth Williams, David Goldberg, and Andy Sun. I was also lucky to forge strong friendships with Ilan and Ruben Lobel, Premal Shah, Margrét Bjarnadóttir, Doug Fearing, Théo Weber, Parikshit Shah and Amir-Ali Ahmadi. A special debt of gratitude is due to my two best friends and colleagues in the ORC, Nikos Trichakis and Kostas Bimpikis - thank you both for being a great support in dire times, and joyful company in good times!

Life in the Cambridge and Boston area would certainly have been radically different, had it not been for the Romanian students at Harvard and MIT. Throughout the six years that we have spent together, they have become some of my closest and dearest friends, and simply mentioning their names does not do justice for all the wonderful moments we spent together. Thank you, Emma Voinescu, Cristi Jitianu, Emanuel Stoica, Florin Moroşan, Florin Constantin, Ştefan Horneţ, Alex Sălcianu, Daniel Nedelcu and Cristi Boboilă!

Last, but most certainly not least, I would like to thank my parents, my grandparents, and my girlfriend, Laura. Their tireless and unconditional love, support, and encouragement has kept me going through the toughest of times, and I owe them more than I could ever express here. To them I dedicate this thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Multi-stage optimization problems under uncertainty have been prevalent in numerous fields of science and engineering, and have elicited interest from diverse research communities, on both a theoretical and a practical level. Several solution approaches have been proposed throughout the years, with various degrees of generality, tractability, and performance guarantees. Some of the most successful ones include exact and approximate dynamic programming, stochastic programming, sampling-based methods, and, more recently, robust and adaptive optimization, which is the main focus of the present thesis.

The key underlying philosophy behind the robust optimization approach is that, in many practical situations, a complete stochastic description of the uncertainty may not be available, and one may only have information with less detailed structure, such as bounds on the magnitude of the uncertain quantities or rough algebraic relations linking multiple unknown parameters. In such cases, one may be able to describe the unknowns by specifying a set in which any realization should lie, the so-called *uncertainty set*. The goal of the decision maker is then to ensure that the constraints in the problem remain feasible for *any* possible realization, while optimizing an objective that corresponds to the worst possible outcome.

In its original form, proposed by Soyster [136] and Falk [64] in the early 1970s, robust optimization was mostly concerned with linear programming problems in which the data was inexact. The former paper considered cases where the column vectors of the constraint matrix (interpreted as the consumption of some finite resource) and the right-hand side vector (the resource availability), were only known to belong to closed, convex sets, and the goal was to find an allocation, given by the decision variables, which would remain feasible for any realization of the consumption and availability. The latter paper dealt with an uncertain objective, with coefficients only known to lie in a convex set - as such, the goal was to find a feasible solution which would optimize the worst-case outcome for the objective.

Interestingly enough, following these early contributions, the approach remained unnoticed in the operations research literature, until the late 1990s. The sequence of papers by Ben-Tal and Nemirovski [10, 18, 11, 12], Ben-Tal et al. [13], El-Ghaoui and Lebret [62], El-Ghaoui et al. [61], and then Bertsimas and Sim [30, 31], Bertsimas et al. [34] and Goldfarb and Iyengar [75] considerably generalized the earlier framework, by extending it to other classes of convex optimization problems (quadratic, conic and semidefinite programs), as well as more complex descriptions of the uncertainty sets (intersections of ellipsoids, cardinality-constrained uncertainty sets, etc). Throughout the papers, the key emphases were on

1. Tractability - under what circumstances can a nominal problem with uncertain data be formulated as a tractable (finite dimensional, convex) optimization problem, and what is the complexity of solving this resulting *robust counterpart*. As it turns out, many interesting classes of nominal optimization problems result in robust counterparts within the same (or related) complexity classes, which allows the use of fast, interior point methods developed for convex optimization (Nesterov and Nemirovski [109]).

2. Degree of conservativeness and probabilistic guarantees - Robust Optimization constructs solutions that are feasible for *any* realization of the unknown parameters within the uncertainty set, and optimizes worst-case outcomes. In many realistic situations, particularly cases where the uncertainties are really stochastic, these prescriptions might lead to overly pessimistic solutions, which simultaneously guard against violations in constraints and low-quality objec-

tives. In this case, as the papers above show, one can still use the framework of Robust Optimization to construct uncertainty sets so that, when solving the (deterministic) robust problem, one obtains solutions which are, with high probability, feasible for the original (stochastic) problem. In such formulations, the structure and size of the uncertainty sets are directly related to the desired probabilistic guarantees, and several systematic ways for trading off between conservativeness and probability of constraint violation exist.

Most of these early contributions were focused on robustification of mathematical programs in static settings. That is, the decision process typically involved a single stage/period, and all the decisions were to be taken at the same time, *before* the uncertainty was revealed. Recognizing that this was a modelling limitation which was not adequate in many realistic settings, a sequence of later papers (Ben-Tal et al. [14, 15, 17]) considered several extensions of the base model. Ben-Tal et al. [14] introduced a setting in which a subset of the decision variables in a linear program could be decided *after* the uncertainty wass revealed, hence resulting in adjustable policies, or *decision rules*. The paper showed that allowing arbitrary adjustable rules typically results in intractable problems, and then proceeded to consider the special class of *affine* rules, i.e., decisions that depend affinely on model disturbances. Under the assumption of fixed recourse, the paper showed that such affine policies remain tractable for several interesting classes of uncertainty sets. For cases without fixed recourse, the paper suggested several approximation techniques, using tools derived from linear systems and control theory. In Ben-Tal et al. [15, 17], the same approach was extended to multi-period linear dynamical systems affected by uncertainty, and tractable exact or approximate reformulations were presented, which allow the computation of affine decision rules.

A related stream of work, focused mostly on applications of robust optimization in different areas of operations management, also considered multi-period models. Formulations have been proposed for several variations of inventory management problems (e.g., Ben-Tal et al. [16], Bertsimas and Thiele [32], Bienstock and Özbay [40]), for dynamic pricing and network revenue management (e.g., Perakis and Roels

[114], Adida and Perakis [1], Thiele [139], Thiele [140]), or for portfolio optimization (Goldfarb and Iyengar [74], Tütüncü and Koenig [142], Ceria and Stubbs [47], Pinar and Tütüncü [116], Bertsimas and Pachamanova [27]). For more references, and a more comprehensive overview, we refer the reader to the recent review paper Bertsimas et al. [35] and the book Ben-Tal et al. [19].

In the context of multi-period decision making, we should note that a parallel stream of work, focusing on similar notions of robustness, has also existed for several decades in the field of dynamical systems and control. The early thesis Witsenhausen [145] and the paper Witsenhausen [146] first formulated problems of state estimation with a set-based membership description of the uncertainty, and the thesis Bertsekas [25] and paper Bertsekas and Rhodes [22] considered the problem of deciding under what conditions the state of a dynamical system affected by uncertainties is guaranteed to lie in specific ellipsoidal or polyhedral tubes (the latter two references showed that, under some conditions, control policies that are linear in the states are sufficient for such a task). The literature on robust control received a tremendous speed-up in the 1990s, with contributions from numerous groups (e.g., Doyle et al. [56], Fan et al. [65]), resulting in two published books on the topic (Zhou and Doyle [148], Dullerud and Paganini [57]. Typically, in most of this literature, the main objective was to design control laws that ensured the dynamical system remained stable under uncertainty, and the focus was on coming up with computationally efficient procedures for synthesizing such controllers. Several (more recent) papers, particularly in the field of model predictive control, have also considered multi-period formulations with different objectives, and have shown how specific classes of policies (typically, open-loop or affine) can be computed efficiently (e.g., Löfberg [99], Kerrigan and Maciejowski [87], Bemporad et al. [9], Goulart and Kerrigan [76], Kerrigan and Maciejowski [88], Bertsimas and Brown [26], Skaf and Boyd [133]).

A unifying theme in both the operations research and robust control literature mentioned above has been that, whenever one deals with multi-period decision problems affected by uncertainty, one always faces the unpleasant conundrum of choosing between *optimality* and *tractability*. If one insists on finding optimal decision policies, then one typically resorts to a formulation via Dynamic Programming (DP) (Bertsekas [21]). More precisely, with a properly defined notion of the state of the dynamical system, one tries to use the Bellman recursions in order to find optimal decision policies and optimal value functions that depend on the underlying state. While DP is a powerful theoretical tool for the characterization of optimal decision policies, it is plagued by the well-known *curse of dimensionality*, in that the complexity of the underlying recursive equations grows quickly with the size of the state-space, rendering the approach ill suited to the computation of actual policy parameters. Therefore, in practice, one would typically either solve the recursions numerically (e.g., by multi-parametric programming Bemporad et al. [7, 8, 9]), or resort to approximations of the value functions, by approximate DP techniques (Bertsekas and Tsitsiklis [23], Powell [120]), sampling (Calafiore and Campi [45], Calafiore and Campi [46]), or other methods.

Instead of considering policies in the states, one could equivalently look for decisions that are directly parametrized in the sequence of observed uncertainties. The resulting policies, usually called *recourse decision rules*, were originally proposed in the Stochastic Programming community (see Birge and Louveaux [41], Garstka and Wets [70] and references therein), and have been widely adopted in recent works in robust control and robust optimization, typically under the names of *disturbance-feedback parametrizations* or *adjustable robust counterparts*. While allowing general decision rules is just as intractable as solving the DP formulation (Ben-Tal et al. [14], Nemirovski and Shapiro [107], Dyer and Stougie [60]), searching for specific functional forms, such as the *affine* class, can often be done by solving convex optimization problems, which vary from linear and quadratic (e.g. Ben-Tal et al. [15], Kerrigan and Maciejowski [88]), to second-order conic and semidefinite programs (e.g. Löfberg [99], Ben-Tal et al. [15], Skaf and Boyd [133]).

Contributing to the popularity of the affine decision rules was also their empirical success, reported in a variety of applications (Ben-Tal et al. [16], Mani et al. [102], Adida and Perakis [1], Lobel and Perakis [97], Babonneau et al. [6]). Ben-Tal et al. [16] performed simulations in the context of a supply chain contracts problem, and found that in only two out of three hundred instances were the affine policies suboptimal (in fact, Chapter 14 of the recent book Ben-Tal et al. [19] contains a slight modification of the model in Ben-Tal et al. [16], for which the authors find that in *all* tested instances, the affine class is optimal!). By comparing (computationally) with appropriate dual formulations, the recent paper Kuhn et al. [90] also found that affine policies were always optimal.

While convenient from a tractability standpoint, the restriction to the affine case could potentially result in large optimality gaps, and it is rarely obvious apriori when that is the case - in the words of Ben-Tal et al. [19], *"in general, [...], we have no idea of how much we lose in terms of optimality when passing from general decision rules to the affine rules. At present, we are not aware of any theoretical tools for evaluating such a loss."*

While proving optimality for affine policies in non-trivial multi-stage problems would certainly be interesting, one might also take a different approach - namely, considering other classes of *tractable* policies, which are guaranteed to improve over the affine case. Along this train of thought, recent works have considered parameterizations that are affine in a new set of variables, derived by lifting the original uncertainties into a higher dimensional space. For example, the authors in Chen and Zhang [50], Chen et al. [52], Sim and Goh [131] suggest using so-called *segregated linear decision rules*, which are affine parameterizations in the positive and negative parts of the original uncertainties. Such policies provide more flexibility, and their computation (for two-stage decision problems in a robust setting) requires roughly the same complexity as that needed for a set of affine policies in the original variables. Another example following similar ideas is Chatterjee et al. [49], where the authors consider arbitrary functional forms of the disturbances, and show how, for specific types of p-norm constraints on the controls, the problems of finding the coefficients of the parameterizations can be relaxed into convex optimization problems. A similar approach is taken in Skaf and Boyd [134], where the authors also consider arbitrary functional forms for the policies, and show how, for a problem with convex state-control constraints and convex costs, such policies can be found by convex optimization, combined with Monte-Carlo sampling (to enforce constraint satisfaction). Chapter 14 of the recent book Ben-Tal et al. [19] also contains a thorough review of several other classes of such adjustable rules, and a discussion of cases when sophisticated rules can actually improve over the affine ones. The main drawback of some of the above approaches is that the *right* choice of functional form for the decision rules is rarely obvious, and there is no systematic way to influence the trade-off between the performance of the resulting policies and the computational complexity required to obtain them, rendering the frameworks ill-suited for general multi-stage dynamical systems, involving complicated constraints on both states and controls.

With the above issues in mind, we now arrive at the point of discussing the main questions addressed in the present thesis, and the ensuing results. Our main contributions can be summarized as follows:

- In Chapter 2, we consider a similar problem to that in Ben-Tal et al. [19], namely a one-dimensional, linear dynamical system evolving over a finite horizon, with box constraints on states and controls, affected by bounded uncertainty, and under an objective consisting of linear control penalties and any convex state penalties. For this model, we *prove* that disturbance-affine policies are optimal. Furthermore, we show that a certain (affine) relaxation of the state costs is also possible, without any loss of optimality, which gives rise to very efficient algorithms for computing the optimal affine policies when the state costs are piece-wise affine. Our theoretical constructions are tight, and the proof of the theorem itself is atypical, consisting of a forward induction and making use of polyhedral geometry to construct the optimal affine policies. Thus, we gain insight into the structure and properties of these policies, which we explore in connection with a classical inventory management problem.

  We remark that two concepts are central to our constructions. First, considering policies over an enlarged state space (i.e., the history of all disturbances) is *essential*, in the sense that affine state-feedback controllers depending only on the current state are, in general, *suboptimal* for the problems we consider.

Second, the construction makes full use of the fact that the problem objective is of mini-max type, which allows the decision maker the freedom of computing policies that are not optimal in every state of the system evolution (but rather, only in states that could result in worst-case outcomes). This underscores a fundamental distinction between robust and stochastic models for decision making under uncertainty, and it suggests that utilizing the framework of Dynamic Programming to solve multi-period robust problems might be an unnecessary overkill, since simpler (not necessarily "Bellman optimal") policies might be sufficient to achieve the optimal worst-case outcome.

- In Chapter 3, we consider a multi-dimensional system, under more general state-control constraints and piece-wise affine, convex state-control costs. For such problems, we introduce a natural extension of the aforementioned affine decision rules, by considering control policies that depend *polynomially* on the observed disturbances. For a fixed polynomial degree d, we develop a convex reformulation of the constraints and objective of the problem, using Sums-Of-Squares (SOS) techniques. In the resulting framework, polynomial policies of degree d can be computed by solving a single semidefinite programming problem (SDP). Our approach is advantageous from a modelling perspective, since it places little burden on the end user (the only choice is the polynomial degree d), while at the same time providing a lever for directly controlling the trade-off between performance and computation (higher d translates into policies with better objectives, obtained at the cost of solving larger SDPs).

  To test our polynomial framework, we consider two classical problems arising in inventory management (single echelon with cumulative order constraints, and serial supply chain with lead-times), and compare the performance of affine, quadratic and cubic control policies. The results obtained are very encouraging - in particular, for all problem instances considered, quadratic policies considerably improve over affine policies (typically by a factor of 2 or 3), while cubic policies essentially close the optimality gap (the relative gap in *all simulations* is less than 1%, with a median gap of less than 0.01%).

- Finally, Chapter 4 considers a classical problem arising in operations management, namely that of dynamically adjusting the prices of inventories in order to maximize the revenues obtained from customers. For the multi-product case, under a linear demand function, we propose a distributionally robust model for the uncertainties, and argue how it can be constructed from limited historical data. We then consider polynomial pricing policies parameterized directly in the observed model mis-specifications, and show how these can be computed by solving second-order conic or semidefinite programming problems.

  In order to test our framework, we consider both simulated data, as well as real data from a large US retailer. We discuss issues related to the calibration of our model, and present extensive Monte-Carlo simulations, which show that adjustable robust policies improve considerably over open-loop robust formulations, and are competitive with popular heuristics in the literature.

Chapter 2

Optimality of Disturbance-Affine Policies

2.1 Introduction

We begin our treatment by examining the following multi-period problem:

Problem 1. *Consider a one-dimensional, discrete-time, linear dynamical system,*

$$x_{k+1} = \alpha_k \cdot x_k + \beta_k \cdot u_k + \gamma_k \cdot w_k \,, \tag{2.1}$$

where $\alpha_k, \beta_k, \gamma_k \neq 0$ *are known scalars, and the initial state* $x_1 \in \mathbb{R}$ *is specified. The random disturbances* w_k *are unknown, but bounded,*

$$w_k \in \mathcal{W}_k \stackrel{\text{def}}{=} [\underline{w}_k, \overline{w}_k]. \tag{2.2}$$

We would like to find a sequence of robust controllers $\{u_k\}$*, obeying upper and lower bound constraints,*

$$u_k \in [L_k, U_k] \,, \tag{2.3}$$

($L_k, U_k \in \mathbb{R}$ *are known and fixed), and minimizing the following cost function over a*

finite horizon $1, \ldots, T$*,*

$$J = c_1\, u_1 + \max_{w_1} \Big[h_1(x_2) + c_2\, u_2 + \max_{w_2} \Big[h_2(x_3) + \ldots + \max_{w_{T-1}} \big[c_T\, u_T + \max_{w_T} \, h_T(x_{T+1}) \, \big] \ldots \Big] \Big], \quad (2.4)$$

where the functions $h_k : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ *are extended-real and convex, and* $c_k \geq 0$ *are fixed and known.*

The problem corresponds to a situation in which, at every time step k, the decision maker has to compute a control action u_k, in such a way that certain constraints (2.3) are obeyed, and a cost penalizing both the state $(h_k(x_{k+1}))$ and the control $(c_k \cdot u_k)$ is minimized. The uncertainty, w_k, always acts so as to maximize the costs, hence the problem solved by the decision maker corresponds to a worst-case scenario (a minimization of the maximum possible cost). An example of such a problem, which we use extensively in the current paper, is the following:

Example 1. *Consider a retailer selling a single product over a planning horizon* $1, \ldots, T$*. The demands* w_k *from customers are only known to be bounded, and the retailer can replenish her inventory* x_k *by placing capacitated orders* u_k*, at the beginning of each period, for a cost of* c_k *per unit of product. After the demand* w_k *is realized, the retailer incurs holding costs* $H_k \cdot \max\{0, x_k + u_k - w_k\}$ *for all the amounts of supply stored on her premises, and penalties* $B_k \cdot \max\{w_k - x_k - u_k, 0\}$*, for any demand that is backlogged.*

Other examples of Problem 1 are the norm-1/∞ and norm-2 control, i.e., $h_k(x) = r_k\, |x|$ or $h_k(x) = r_k\, x^2$, all of which have been studied extensively in the control literature in the unconstrained case (see Zhou and Doyle [148] and Dullerud and Paganini [57]).

The solution to Problem 1 could be obtained using a "classical" Dynamic Programming (DP) formulation (Bertsekas [21]), in which the optimal policies $u_k^\star(x_k)$ and the optimal value functions $J_k^\star(x_k)$ are computed backwards in time, starting at the end of the planning horizon, $k = T$. The resulting policies are *piecewise affine* in the states x_k, and have properties that are well known and documented in the literature (e.g., for the inventory model above, they exactly correspond to the base-stock ordering policies of Scarf [129] and Kasugai and Kasegai [86]). We remark that the *piecewise* structure is essential, i.e., control policies that are only *affine* in the states x_k are, in general, suboptimal.

As detailed in the introduction, our goal is to study the performance of a new class of policies, where instead of regarding the controllers u_k as functions of the state x_k, one seeks *disturbance-feedback policies*, i.e., policies that are directly parameterizations in the observed disturbances:

$$u_k : \mathcal{W}_1 \times \mathcal{W}_2 \times \cdots \times \mathcal{W}_{k-1} \to \mathbb{R}. \tag{2.5}$$

One such example (of particular interest) is the *disturbance-affine* class, i.e., policies of the form (2.5) which are also affine. In this new framework, we require that constraint (2.3) should be robustly feasible, i.e.,

$$u_k(\boldsymbol{w}) \in [L_k, U_k], \qquad \forall\, \boldsymbol{w} \in \mathcal{W}_1 \times \cdots \times \mathcal{W}_{k-1}. \tag{2.6}$$

Note that if we insisted on this category of parameterizations, then we would have to consider a new state for the system, $\boldsymbol{X}_k$, which would include at least all the past-observed disturbances, as well as possibly other information (e.g., the previous controls $\{u_t\}_{1\le t<k}$, the previous states $\{x_t\}_{1\le t<k}$, or some combination thereof). Compared with the original, compact state formulation, x_k, the new state $\boldsymbol{X}_k$ would become much larger, and solving the DP with state variable $\boldsymbol{X}_k$ would produce exactly the same optimal objective function value. Therefore, one should rightfully ask what the benefit for introducing such a complicated state might be.

The hope is that, by considering policies over a larger state, *simpler functional forms* might be sufficient for optimality, for instance, *affine* policies. These have a very compact representation, since only the coefficients of the parameterization are needed, and, for certain classes of convex costs $h_k(\cdot)$, there may be efficient procedures available for computing them.

This approach is also not new in the literature. It has been originally advocated in the context of stochastic programming (see Charnes et al. [48], Garstka and Wets [70], and references therein), where such policies are known as *decision rules*. More recently, the idea has received renewed interest in robust optimization (Ben-Tal et al. [14]), and has been extended to linear systems theory (Ben-Tal et al. [15, 17]), with notable contributions from researchers in robust model predictive control and receding horizon control (see Löfberg [99], Bemporad et al. [9], Kerrigan and Maciejowski [88], Skaf and Boyd [133], and references therein). In all the papers, which usually deal with the more general case of multi-dimensional linear systems, the authors typically restrict attention, for purposes of tractability, to the class of *disturbance-affine* policies, and show how the corresponding policy parameters can be found by solving specific types of optimization problems, which vary from linear and quadratic programs (Ben-Tal et al. [15], Kerrigan and Maciejowski [87, 88]) to conic and semi-definite (Löfberg [99], Ben-Tal et al. [15]), or even multi-parametric, linear or quadratic programs (Bemporad et al. [9]). The tractability and empirical success of disturbance-affine policies in the robust framework have lead to their reexamination in stochastic settings, with several recent papers (Nemirovski and Shapiro [107], Chen et al. [52], Kuhn et al. [90]) providing tractable methods for determining the best parameters of the policies, in the context of both single-stage and multi-stage linear stochastic programming problems.

The first steps towards analyzing the properties of such parameterizations were made in Kerrigan and Maciejowski [88], where the authors show that, under suitable conditions, the resulting affine parameterization has certain desirable system theoretic properties (stability and robust invariance). Other notable contributions were Goulart and Kerrigan [76] and Ben-Tal et al. [15], who prove that the class of affine disturbance feedback policies is equivalent to the class of affine state feedback policies with memory of prior states, thus subsuming the well known classes of open-loop and pre-stabilizing control policies. In terms of characterizing the optimal objective obtained by using affine parameterizations, most research efforts thus far focus on providing tractable dual formulations, which allow a computation of lower or upper bounds to the problems, and hence an assessment of the degree of sub-optimality (see Kuhn et al. [90] for details). Empirically, several authors have observed that affine policies deliver excellent performance, with Ben-Tal et al. [16] and Kuhn et al. [90] reporting many instances in which they are actually optimal. However, to the best of our knowledge, apart from these advances, there has been very little progress in *proving* results about the quality of the objective function value resulting from the use of such parameterizations.

Our main result, summarized in Theorem 1 of Section 2.3, is that, for Problem 1 stated above, disturbance-affine policies of the form (2.5) *are optimal*. Furthermore, we prove that a certain (affine) relaxation of the state costs is also possible, without any loss of optimality, which gives rise to very efficient algorithms for computing the optimal affine policies when the state costs $h_k(\cdot)$ are piece-wise affine. To the best of our knowledge, this is the first result of its kind, and it is surprising, particularly since similar policies, i.e., decision rules, are known to be severely suboptimal for stochastic problems (see, e.g., Garstka and Wets [70], and our discussion in Section 2.4.5). The result provides intuition and motivation for the widespread advocation of such policies in both theory and applications. Our theoretical constructions are tight, i.e., if the conditions in Problem 1 are slightly perturbed, then simple counterexamples for Theorem 1 can be found (see Section 2.4.5). The proof of the theorem itself is atypical, consisting of a forward induction and making use of polyhedral geometry to construct the optimal affine policies. Thus, we gain insight into the structure and properties of these policies, which we explore in connection with the inventory management problem in Example 1.

We remark that two concepts are central to our constructions. First, considering policies over an enlarged state space (here, the history of all disturbances) is *essential*, in the sense that affine state-feedback controllers depending only on the current state x_k (e.g., $u_k(x_k) = \ell_k x_k + \ell_{k,0}$) are, in general, *suboptimal* for the problems we consider. Second, the construction makes full use of the fact that the problem objective is of mini-max type, which allows the decision maker the freedom of computing policies that are not optimal in every state of the system evolution (but rather, only in states that could result in worst-case outcomes). This is a fundamental distinction between robust and stochastic models for decision making under uncertainty, and it suggests that utilizing the framework of Dynamic Programming to solve multi-period robust problems might be an unnecessary overkill, since simpler (not necessarily "Bellman optimal") policies might be sufficient to achieve the optimal worst-case outcome.

The chapter is organized as follows. Section 2.2 presents an overview of the Dynamic Programming formulation in state variable x_k, extracting the optimal policies $u_k^\star(x_k)$ and optimal value functions $J_k^\star(x_k)$, as well as some of their properties. Section 2.3 contains our main result, and briefly discusses some immediate extensions and computational implications. In Section 2.4, we introduce the constructive proof for building the affine control policies and the affine cost relaxations, and present counterexamples that prevent a generalization of the results. Section 2.5 concludes the chapter, by discussing our results in connection with the classical inventory management problem of Example 1.

2.1.1 Notation.

Throughout the rest of the chapter, the subscripts k and t are used to denote time-dependency, and vector quantities are distinguished by bold-faced symbols, with optimal quantities having a $\star$ superscript, e.g., $J_k^\star$. Also, $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ stands for the set of extended reals.

Since we seek policies parameterized directly in the uncertainties, we introduce $\boldsymbol{w}_{[k]} \stackrel{\text{def}}{=} (w_1, \ldots, w_{k-1})$ to denote the history of known disturbances in period k, and $\mathcal{H}_k \stackrel{\text{def}}{=} \mathcal{W}_1 \times \cdots \times \mathcal{W}_{k-1}$ to denote the corresponding uncertainty set (a hypercube in $\mathbb{R}^{k-1}$). A function q_k that depends affinely on variables $w_1, \ldots, w_{k-1}$ is denoted by $q_k(\boldsymbol{w}_{[k]}) \stackrel{\text{def}}{=} q_{k,0} + \boldsymbol{q}_k' \boldsymbol{w}_{[k]}$, where $\boldsymbol{q}_k$ is the vector of coefficients, and $'$ denotes the usual transpose.

2.2 Dynamic Programming Solution.

As mentioned in the introduction, the solution to Problem 1 can be obtained using a "classical" DP formulation (see, e.g., Bertsekas [21]), in which the state is taken to be x_k, and the optimal policies $u_k^\star(x_k)$ and optimal value functions $J_k^\star(x_k)$ are computed starting at the end of the planning horizon, $k = T$, and moving backwards in time. In this section, we briefly outline the DP solution for our problem, and state some of the key properties that are used throughout the rest of the paper. For completeness, a full proof of the results is included in Section A.1 of the Appendix.

In order to simplify the notation, we remark that, since the constraints on the controls u_k and the bounds on the disturbances w_k are time-varying, and independent for different time-periods, we can restrict attention, without loss of generality[1], to a system with $\alpha_k = \beta_k = \gamma_k = 1$. With this simplification, the problem that we would like to solve is the following:

$$
\begin{aligned}
&\min_{u_1}\Big[c_1\,u_1 + \max_{w_1}\Big[h_1(x_2) + \dots + \min_{u_k}\Big[c_k\,u_k + \max_{w_k}\Big[h_k(x_{k+1}) + \dots \\
&\qquad\qquad + \min_{u_T}\Big[c_T\,u_T + \max_{w_T}\; h_T(x_{T+1})\Big]\dots\Big]\Big]
\end{aligned}
$$

$$
\begin{aligned}
&\text{s.t.} \quad x_{k+1} = x_k + u_k + w_k \\
(DP) \qquad & L_k \le u_k \le U_k \qquad\qquad \forall\, k \in \{1,2,\dots,T\} \\
& w_k \in \mathcal{W}_k = [\underline{w}_k, \overline{w}_k].
\end{aligned}
$$

The corresponding Bellman recursion for (DP) can then be written as follows:

$$
J_k^\star(x_k) \overset{\text{def}}{=} \min_{L_k \le u_k \le U_k}\Big[c_k\,u_k + \max_{w_k \in \mathcal{W}_k}\Big[\,h_k(x_k + u_k + w_k) + J_{k+1}^\star\,(x_k + u_k + w_k)\Big]\Big]\,,
$$

[1] Such a system can always be obtained by the linear change of variables $\tilde{x}_k = \frac{x_k}{\prod_{i=1}^{k-1}\alpha_i}$, and by suitably scaling the bounds $L_k, U_k, \underline{w}_k, \overline{w}_k$.

where $J^\star_{T+1}(x_{T+1}) \equiv 0$. By defining:

$$y_k \stackrel{\text{def}}{=} x_k + u_k \tag{2.7a}$$

$$g_k(y_k) \stackrel{\text{def}}{=} \max_{w_k \in \mathcal{W}_k} \left[\; h_k(y_k + w_k) + J^\star_{k+1}\left(y_k + w_k\right) \; \right], \tag{2.7b}$$

we obtain the following solution to the Bellman recursion (see Section A.1 in the Appendix for the derivation):

$$u^\star_k(x_k) = \begin{cases} U_k, & \text{if } x_k < y^\star_k - U_k \\ -x_k + y^\star_k, & \text{otherwise} \\ L_k, & \text{if } x_k > y^\star_k - L_k \end{cases} \tag{2.8}$$

$$J^\star_k(x_k) = c_k \cdot u^\star_k(x_k) + g_k\big(x_k + u^\star_k(x_k)\big) = \begin{cases} c_k \cdot U_k + g_k(x_k + U_k), & \text{if } x_k < y^\star_k - U_k \\ c_k \cdot (y^\star_k - x_k) + g_k(y^\star), & \text{otherwise} \\ c_k \cdot L_k + g_k(x_k + L_k), & \text{if } x_k > y^\star_k - L_k \end{cases}, \tag{2.9}$$

where $y^\star_k$ represents the minimizer[2] of the convex function $c_k \cdot y + g_k(y)$ (for the inventory Example 1, $y^\star_k$ is the basestock level in period k, i.e., the inventory position just after ordering, and before seeing the demand). A typical example of the optimal control law and the optimal value function is shown in Figure 2-1.

The main properties of the solution relevant for our later treatment are listed below:

(P1) The optimal control law $u^\star_k(x_k)$ is piecewise affine, continuous and non-increasing.

(P2) The optimal value function, $J^\star_k(x_k)$, and the function $g_k(y_k)$ are convex.

(P3) The difference in the values of the optimal control law at two distinct arguments

[2] For simplicity of exposition, we work under the assumption that the minimizer is unique. The results can be extended to the case of multiple minimizers.

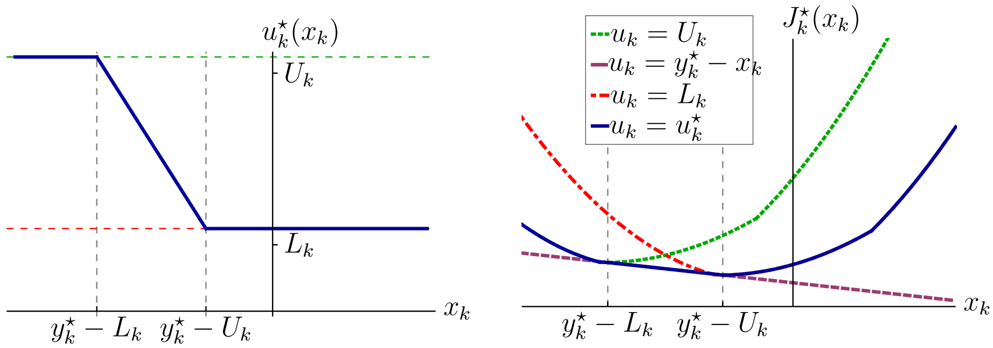


Figure 2-1: Optimal control law $u_k^\star(x_k)$ and optimal value function $J_k^\star(x_k)$ at time k.

$s \le t$ always satisfies: $0 \le u_k^\star(s) - u_k^\star(t) \le t - s$. Equivalently, $x_k + u_k^\star(x_k)$ is non-decreasing as a function of x_k.

2.3 Optimality of Affine Policies in the History of Disturbances.

In this section, we introduce our main contribution, namely a proof that policies that are affine in the disturbances $\boldsymbol{w}_{[k]}$ are, in fact, optimal for problem (DP). Using the same notation as in Section 2.2, and with $J_1^\star(x_1)$ denoting the optimal overall value, we can summarize our main result in the following theorem:

Theorem 1 (Optimality of disturbance-affine policies)**.** *Affine disturbance-feedback policies are optimal for Problem 1 stated in the introduction. More precisely, for every time step $k = 1, \ldots, T$, the following quantities exist:*

an affine control policy, $$q_k(\boldsymbol{w}_{[k]}) \stackrel{\text{def}}{=} q_{k,0} + \boldsymbol{q}_k' \boldsymbol{w}_{[k]}, \tag{2.10a}$$

an affine running cost, $$z_k(\boldsymbol{w}_{[k+1]}) \stackrel{\text{def}}{=} z_{k,0} + \boldsymbol{z}_k' \boldsymbol{w}_{[k+1]}, \tag{2.10b}$$

such that the following properties are obeyed:

$$L_k \leq q_k(\boldsymbol{w}_{[k]}) \leq U_k, \qquad \forall\, \boldsymbol{w}_{[k]} \in \mathcal{H}_k, \tag{2.11a}$$

$$z_k(\boldsymbol{w}_{[k+1]}) \geq h_k\left(x_1 + \sum_{t=1}^{k}\Big(q_t(\boldsymbol{w}_{[t]}) + w_t\Big)\right), \qquad \forall\, \boldsymbol{w}_{[k+1]} \in \mathcal{H}_{k+1}, \tag{2.11b}$$

$$J_1^{\star}(x_1) = \max_{\boldsymbol{w}_{[k+1]} \in \mathcal{H}_{k+1}} \left[\sum_{t=1}^{k}\Big(c_t \cdot q_t(\boldsymbol{w}_{[t]}) + z_t(\boldsymbol{w}_{[t+1]})\Big) + J_{k+1}^{\star}\left(x_1 + \sum_{t=1}^{k}\Big(q_t(\boldsymbol{w}_{[t]}) + w_t\Big)\right)\right]. \tag{2.11c}$$

Let us interpret the main statements in the theorem. Equation (2.11a) confirms the existence of an affine policy $q_k(\boldsymbol{w}_{[k]})$ that is robustly feasible, i.e., that obeys the control constraints, no matter what the realization of the disturbances may be. Equation (2.11b) states the existence of an affine cost $z_k(\boldsymbol{w}_{[k+1]})$ that is always larger than the convex state cost $h_k(x_{k+1})$ incurred when the affine policies $\{q_t(\cdot)\}_{1\leq t\leq k}$ are used. Equation (2.11c) guarantees that, despite using the (suboptimal) affine control law $q_k(\cdot)$, and incurring a (potentially larger) affine stage cost $z_k(\cdot)$, the overall objective function value $J_1^{\star}(x_1)$ is, in fact, not increased. This translates in the following two main results:

- *Existential result.* Affine policies $q_k(\boldsymbol{w}_{[k]})$ are, in fact, optimal for Problem 1.

- *Computational result.* When the convex costs $h_k(x_{k+1})$ are piecewise affine, the optimal affine policies $\big\{q_k(\boldsymbol{w}_{[k]})\big\}_{1\leq k\leq T}$ can be computed by solving a Linear Programming problem.

To see why the second implication would hold, suppose that $h_k(x_{k+1})$ is the maximum of m_k affine functions, $h_k(x_{k+1}) = \max\big(p_k^i \cdot x_{k+1} + p_{k,0}^i\big),\ i \in \{1,\ldots,m_k\}$. Then the optimal affine policies $q_k(\boldsymbol{w}_{[k]})$ can be obtained by solving the following optimization problem (see Ben-Tal et al. [16]):

$$
(AARC)\quad
\begin{aligned}
\min_{J;\{q_{k,t}\};\{z_{k,t}\}} \quad & J \\
\text{s.t.} \quad & \forall\, \boldsymbol{w} \in \mathcal{H}_{T+1}, \quad \forall\, k \in \{1,\dots,T\}: \\
& J \geq \sum_{k=1}^{T} \left[c_k \cdot q_{k,0} + z_{k,0} + \sum_{t=1}^{k-1} (c_t \cdot q_{k,t} + z_{k,t}) \cdot w_t + z_{k,k} \cdot w_k \right], \\
& z_{k,0} + \sum_{t=1}^{k} z_{k,t} \cdot w_t \geq p_k^i \cdot \left[x_1 + \sum_{t=1}^{k} \left(q_{t,0} + \sum_{\tau=1}^{t-1} q_{t,\tau} \cdot w_\tau + w_t \right) \right] + p_{k,0}^i\,, \\
& \qquad\qquad \forall\, i \in \{1,\dots,m_k\}, \\
& L_k \leq q_{k,0} + \sum_{t=1}^{k-1} q_{k,t} \cdot w_t \leq U_k.
\end{aligned}
\tag{2.12}
$$

Although Problem $(AARC)$ is still a semi-infinite LP (due to the requirement of robust constraint feasibility, $\forall\,\boldsymbol{w}$), since all the constraints are inequalities that are bi-affine in the decision variables and the uncertain quantities, a very compact reformulation of the problem is available. In particular, with a typical constraint in $(AARC)$ written as

$$
\lambda_0(\boldsymbol{x}) + \sum_{t=1}^{T} \lambda_t(\boldsymbol{x}) \cdot w_t \leq 0, \qquad \forall\, \boldsymbol{w} \in \mathcal{H}_{T+1}\,,
$$

where $\lambda_i(\boldsymbol{x})$ are affine functions of the decision variables $\boldsymbol{x}$, it can be shown (see Ben-Tal and Nemirovski [12], Ben-Tal et al. [14] for details) that the previous condition is equivalent to:

$$
\begin{cases}
\lambda_0(\boldsymbol{x}) + \sum_{t=1}^{T} \left(\lambda_t(\boldsymbol{x}) \cdot \frac{\underline{w}_t + \overline{w}_t}{2} + \frac{\overline{w}_t - \underline{w}_t}{2} \cdot \xi_t \right) \leq 0 \\
-\xi_t \leq \lambda_t(\boldsymbol{x}) \leq \xi_t, \quad t = 1,\dots,T\,,
\end{cases}
\tag{2.13}
$$

which are linear constraints in the decision variables $\boldsymbol{x}, \boldsymbol{\xi}$. Therefore, $(AARC)$ can be reformulated as a Linear Program, with $O\left(T^2 \max_k m_k\right)$ variables and $O\left(T^2 \max_k m_k\right)$

constraints, which can be solved very efficiently using commercially available software.

We conclude our observations by making one last remark related to an immediate extension of the results. Note that in the statement of Problem 1, there was no mention about constraints on the states x_k of the dynamical system. In particular, one may want to incorporate lower or upper bounds on the states, as well,

$$L_k^x \leq x_k \leq U_k^x. \tag{2.14}$$

We claim that, in case the mathematical problem including such constraints remains feasible[3], then affine policies are, again, optimal. The reason is that such constraints can always be simulated in our current framework, by adding suitable convex barriers to the stage costs $h_k(x_{k+1})$. In particular, by considering the modified, convex stage costs

$$\tilde{h}_k(x_{k+1}) \stackrel{\text{def}}{=} h_k(x_{k+1}) + \mathbf{1}_{[L_{k+1}^x, U_{k+1}^x]}(x_{k+1}),$$

where $\mathbf{1}_S(x) \stackrel{\text{def}}{=} \big\{0, \text{ if } x \in S; \infty, \text{ otherwise}\big\}$, it can be easily seen that the original problem, with convex stage costs $h_k(\cdot)$ and state constraints (2.14), is equivalent to a problem with the modified stage costs $\tilde{h}_k(\cdot)$ and no state constraints. And, since affine policies are optimal for the latter problem, the result is immediate. Therefore, our decision to exclude such constraints from the original formulation was made only for sake of brevity and conciseness of the proofs, but without loss of generality.

2.4 Proof of Main Theorem.

The current section contains the proof of Theorem 1. Before presenting the details, we first give some intuition behind the strategy of the proof, and introduce the organization of the material.

Unlike most Dynamic Programming proofs, which utilize backward induction on

[3]Such constraints may lead to infeasible problems. For example, $T = 1, x_1 = 0, u_1 \in [0, 1], w_1 \in [0, 1], x_2 \in [5, 10]$.

the time-periods, we proceed with a *forward* induction. Section 2.4.1 presents a test of the first step of the induction, and then introduces a detailed analysis of the consequences of the induction hypothesis.

We then separate the completion of the induction step into two parts. In the first part, discussed in Section 2.4.2, by exploiting the structure provided by the forward induction hypothesis, and making critical use of the properties of the optimal control law $u_k^\star(x_k)$ and optimal value function $J_k^\star(x_k)$ (the DP solutions), we introduce a candidate affine policy $q_k(\boldsymbol{w}_{[k]})$. In Section 2.4.2, we then prove that this policy is robustly feasible, and preserves the min-max value of the overall problem, $J_1^\star(x_1)$, when used in conjunction with the original, convex state costs, $h_k(x_{k+1})$.

Similarly, for the second part of the inductive step (Section 2.4.3), by re-analyzing the feasible sets of the optimization problems resulting after the use of the (newly computed) affine policy $q_k(\boldsymbol{w}_{[k]})$, we determine a candidate affine cost $z_k(\boldsymbol{w}_{[k+1]})$, which we prove to be always larger than the original convex state costs, $h_k(x_{k+1})$. However, despite this fact, in Section 2.4.3 we also show that when this affine cost is incurred, the overall min-max value $J_1^\star(x_1)$ remains unchanged, which completes the proof of the inductive step.

Section 2.4.4 concludes the proof of Theorem 1, and outlines several counterexamples that prevent an immediate extension of the result to more general cases.

2.4.1 Induction Hypothesis.

As mentioned before, the proof of the theorem utilizes a *forward* induction on the time-step k. We begin by verifying the induction at $k = 1$.

Using the same notation as in Section 2.2, by taking the affine control to be $q_1 \stackrel{\text{def}}{=} u_1^\star(x_1)$, we immediately get that q_1, which is simply a constant, is robustly feasible, so (2.11a) is obeyed. Furthermore, since $u_1^\star(x_1)$ is optimal, we can write the overall optimal objective value as:

$$\begin{aligned}J_1^\star(x_1) &= \min_{u_1 \in [L_1, U_1]} \left[c_1 \cdot u_1 + g_1(x_1 + u_1) \right] = c_1 \cdot q_1 + g_1\left(x_1 + q_1\right) \\ &= \big(\text{by (2.7b) and convexity of } h_1, J_2^\star\big) \\ &= c_1 \cdot q_1 + \max\big\{ \left(h_1 + J_2^\star\right)\left(x_1 + q_1 + \underline{w}_1\right), \left(h_1 + J_2^\star\right)\left(x_1 + q_1 + \overline{w}_1\right)\big\}. \end{aligned} \tag{2.15}$$

Next, we introduce the affine cost $z_1(w_1) \stackrel{\text{def}}{=} z_{1,0} + z_{1,1} \cdot w_1$, where we constrain the coefficients $z_{1,i}$ to satisfy the following two linear equations:

$$z_{1,0} + z_{1,1} \cdot w_1 = h_1(x_1 + q_1 + w_1), \forall\, w_1 \in \big\{\underline{w}_1, \overline{w}_1\big\}.$$

Note that for fixed x_1 and q_1, the function $z_1(w_1)$ is nothing but a linear interpolation of the mapping $w_1 \mapsto h_1(x_1 + q_1 + w_1)$, matching the value at points $\{\underline{w}_1, \overline{w}_1\}$. Since h_1 is convex, the linear interpolation defined above clearly dominates it, so condition (2.11b) is readily satisfied. Furthermore, by (2.15), $J_1^\star(x_1)$ is achieved for $w_1 \in \{\underline{w}_1, \overline{w}_1\}$, so condition (2.11c) is also obeyed.

Having checked the induction at time $k = 1$, let us now assume that the statements of Theorem 1 are true for times $t = 1, \ldots, k$. Equation (2.11c) written for stage k then yields:

$$\begin{aligned}J_1^\star(x_1) &= \max_{\boldsymbol{w}_{[k+1]} \in \mathcal{H}_{k+1}} \left[\sum_{t=1}^{k} \Big(c_t \cdot q_t(\boldsymbol{w}_{[t]}) + z_t\left(\boldsymbol{w}_{[t+1]}\right)\Big) + J_{k+1}^\star\Big(x_1 + \sum_{t=1}^{k} \big(q_t(\boldsymbol{w}_{[t]}) + w_t\big)\Big) \right] \\ &= \max_{(\theta_1, \theta_2) \in \Theta} \Big[\theta_1 + J_{k+1}^\star(\theta_2) \Big], \text{ where}\end{aligned} \tag{2.16}$$

$$\begin{aligned}\Theta \stackrel{\text{def}}{=} \Big\{ (\theta_1, \theta_2) \in \mathbb{R}^2 : \theta_1 &\stackrel{\text{def}}{=} \sum_{t=1}^{k} \big(c_t \cdot q_t(\boldsymbol{w}_{[t]}) + z_t\left(\boldsymbol{w}_{[t+1]}\right)\big), \\ \theta_2 &\stackrel{\text{def}}{=} x_1 + \sum_{t=1}^{k} \big(q_t(\boldsymbol{w}_{[t]}) + w_t\big),\ \boldsymbol{w}_{[k+1]} \in \mathcal{H}_{k+1} \Big\}.\end{aligned} \tag{2.17}$$

Since $\{q_t\}_{1 \le t \le k}$ and $\{z_t\}_{1 \le t \le k}$ are affine functions, this implies that, although the uncertainties $\boldsymbol{w}_{[k+1]} = (w_1, \ldots, w_k)$ lie in a set with 2^k vertices (the hyperrectangle $\mathcal{H}_{k+1}$), they are only able to affect the objective J_{mM} through two affine combinations (θ_1 summarizing all the past stage costs, and θ_2 representing the next state, x_{k+1}), taking values in the set Θ. Such a polyhedron, arising as a 2-dimensional affine projection of a k-dimensional hyperrectangle, is called a *zonogon* (see Figure 2-2 for an example). It belongs to a larger class of polytopes, known as *zonotopes*, whose combinatorial structure and properties are well documented in the discrete and computational geometry literature. The interested reader is referred to Chapter 7 of Ziegler [149] for a very nice and accessible introduction.

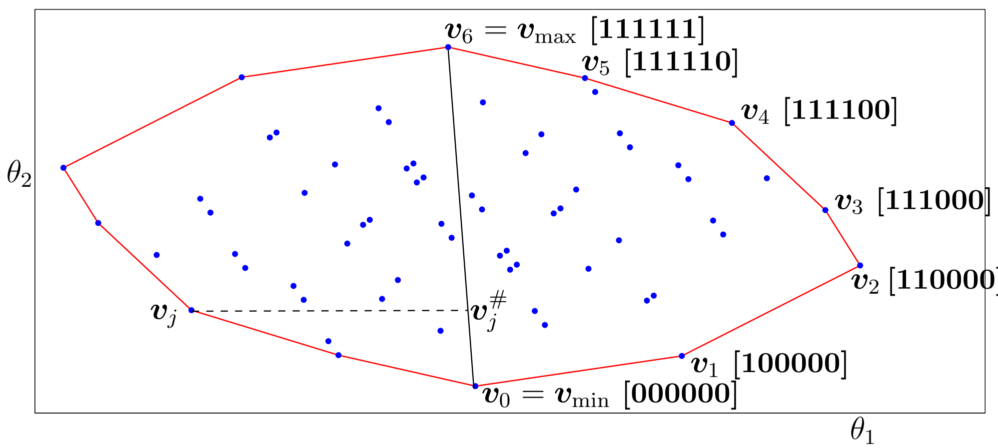


Figure 2-2: Zonogon obtained from projecting a hypercube in $\mathbb{R}^6$.

The main properties of a zonogon that we are interested in are summarized in Lemma 13, found in the Appendix. In particular, the set Θ is centrally symmetric, and has at most $2k$ vertices (see Figure 2-2 for an example). Furthermore, by numbering the vertices of Θ in counter-clockwise fashion, starting at

$$\boldsymbol{v}_0 \equiv \boldsymbol{v}_{\min} \stackrel{\text{def}}{=} \arg\max\left\{\theta_1 \,:\, \boldsymbol{\theta} \in \arg\min\{\theta_2' \,:\, \boldsymbol{\theta}' \in \Theta\}\right\}, \tag{2.18}$$

we establish the following result concerning the points of Θ that are relevant in our problem:

Lemma 1. *The maximum value in* (2.16) *is achieved for some* $(\theta_1, \theta_2) \in \{\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$.

Proof. The optimization problem described in (2.16) and (2.17) is a maximization of a convex function over a convex set. Therefore (see Section 32 of Rockafellar [126]), the maximum is achieved at the extreme points of the set Θ, namely on the set $\{\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{2p-1}, \boldsymbol{v}_{2p} \equiv \boldsymbol{v}_0\}$, where $2p$ is the number of vertices of Θ. Letting $\boldsymbol{O}$ denote the center of Θ, by part (iii) of Lemma 13 in the Appendix, we have that the vertex symmetrically opposed to $\boldsymbol{v}_{\min}$, namely $\boldsymbol{v}_{\max} \stackrel{\text{def}}{=} 2\boldsymbol{O} - \boldsymbol{v}_{\min}$, satisfies $\boldsymbol{v}_{\max} = \boldsymbol{v}_p$.

Consider any vertex $\boldsymbol{v}_j$ with $j \in \{p+1, \ldots, 2p-1\}$. From the definition of $\boldsymbol{v}_{\min}, \boldsymbol{v}_{\max}$, for any such vertex, there exists a point $\boldsymbol{v}_j^{\#} \in [\boldsymbol{v}_{\min}, \boldsymbol{v}_{\max}]$, with the same θ_2-coordinate as $\boldsymbol{v}_j$, but with a θ_1-coordinate larger than $\boldsymbol{v}_j$ (refer to Figure 2-2). Since such a point will have an objective in problem (2.16) at least as large as $\boldsymbol{v}_j$, and $\boldsymbol{v}_j^{\#} \in [\boldsymbol{v}_0, \boldsymbol{v}_p]$, we can immediately conclude that the maximum of problem (2.16) is achieved on the set $\{\boldsymbol{v}_0, \ldots, \boldsymbol{v}_p\}$. Since $2p \leq 2k$ (see part (ii) of Lemma 13), we immediately arrive at the conclusion of the lemma. □

Since the argument presented in the lemma is recurring throughout several of our proofs and constructions, we end this subsection by introducing two useful definitions, and generalizing the previous result.

Consider the system of coordinates (θ_1, θ_2) in $\mathbb{R}^2$, and let $\mathcal{S} \subset \mathbb{R}^2$ denote an arbitrary, finite set of points and $\mathcal{P}$ denote any (possibly non-convex) polygon such that its set of vertices is exactly $\mathcal{S}$. With $\boldsymbol{y}_{\min} \stackrel{\text{def}}{=} \arg\max\big\{\theta_1 : \boldsymbol{\theta} \in \arg\min\{\theta_2' : \boldsymbol{\theta}' \in \mathcal{P}\}\big\}$ and $\boldsymbol{y}_{\max} \stackrel{\text{def}}{=} \arg\max\big\{\theta_1 : \boldsymbol{\theta} \in \arg\max\{\theta_2' : \boldsymbol{\theta}' \in \mathcal{P}\}\big\}$, by numbering the vertices of the convex hull of $\mathcal{S}$ in a counter-clockwise fashion, starting at $\boldsymbol{y}_0 \stackrel{\text{def}}{=} \boldsymbol{y}_{\min}$, and with $\boldsymbol{y}_m = \boldsymbol{y}_{\max}$, we define the *right side* of $\mathcal{P}$ and the *zonogon hull* of $\mathcal{S}$ as follows:

Definition 1. *The* ***right side*** *of an arbitrary polygon $\mathcal{P}$ is:*

$$\text{r-side}\,(\mathcal{P}) \stackrel{\text{def}}{=} \{\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}. \tag{2.19}$$

Definition 2. *The* ***zonogon hull*** *of a set of points $\mathcal{S}$ is:*

$$\text{z-hull}\,(\mathcal{S}) \stackrel{\text{def}}{=} \left\{\boldsymbol{y} \in \mathbb{R}^2 \;:\; \boldsymbol{y} = \boldsymbol{y}_0 + \sum_{i=1}^{m} w_i \cdot (\boldsymbol{y}_i - \boldsymbol{y}_{i-1}) \;,\; 0 \leq w_i \leq 1\right\}. \tag{2.20}$$

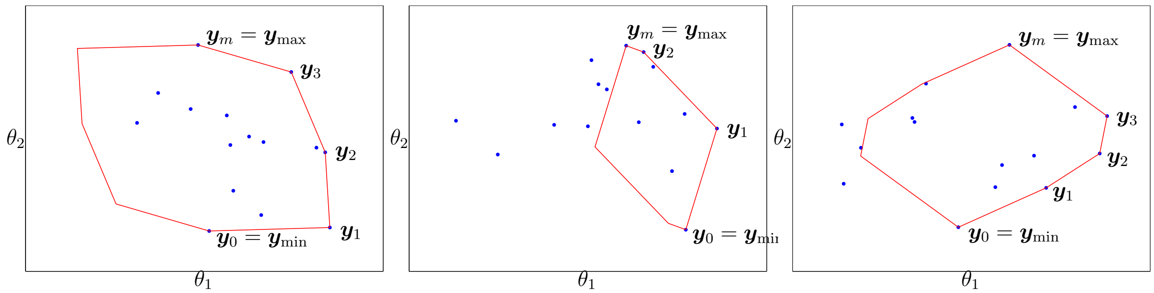


Figure 2-3: Examples of zonogon hulls for different sets $\mathcal{S} \in \mathbb{R}^2$.

Intuitively, r-side($\mathcal{P}$) represents exactly what the names hints at, i.e., the vertices found on the right side of $\mathcal{P}$. An equivalent definition using more familiar operators would be

$$\text{r-side}(\mathcal{P}) \equiv \text{ext}\Big(\text{cone}\big(\left[\begin{smallmatrix}-1\\0\end{smallmatrix}\right]\big) + \text{conv}\,(\mathcal{P})\Big),$$

where cone($\cdot$) and conv($\cdot$) represent the conic and convex hull, respectively, and ext($\cdot$) denotes the set of extreme points.

Using Definition 3 in Section A.2 of the Appendix, one can see that the *zonogon hull* of a set $\mathcal{S}$ is simply a zonogon that has exactly the same vertices on the right side as the convex hull of $\mathcal{S}$, i.e., $\text{r-side}\,(\text{z-hull}\,(\mathcal{S})) = \text{r-side}\,(\text{conv}\,(\mathcal{S}))$. Some examples of zonogon hulls are shown in Figure 2-3 (note that the initial points in $\mathcal{S}$ do not necessarily fall inside the zonogon hull, and, as such, there is no general inclusion relation between the zonogon hull and the convex hull). The reason for introducing this object is that it allows for the following immediate generalization of Lemma 1:

Corollary 1. *If $\mathcal{P}$ is any polygon in $\mathbb{R}^2$ (coordinates $(\theta_1, \theta_2) \equiv \boldsymbol{\theta}$) with a finite set $\mathcal{S}$ of vertices, and $f(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \theta_1 + g(\theta_2)$, where $g : \mathbb{R} \to \bar{\mathbb{R}}$ is any convex function, then the following chain of equalities holds:*

$$\begin{aligned}\max_{\boldsymbol{\theta}\in\mathcal{P}} f(\boldsymbol{\theta}) = \max_{\boldsymbol{\theta}\in\text{conv}(\mathcal{P})} f(\boldsymbol{\theta}) &= \max_{\boldsymbol{\theta}\in\mathcal{S}} f(\boldsymbol{\theta}) = \max_{\boldsymbol{\theta}\in\text{r-side}(\mathcal{P})} f(\boldsymbol{\theta}) \\ &= \max_{\boldsymbol{\theta}\in\text{z-hull}(\mathcal{S})} f(\boldsymbol{\theta}) = \max_{\boldsymbol{\theta}\in\text{r-side}(\text{z-hull}(\mathcal{S}))} f(\boldsymbol{\theta}).\end{aligned}$$

Proof. The proof is identical to that of Lemma 1, and is omitted for brevity. ◻

Using this result, whenever we are faced with a maximization of a convex function $\theta_1 + g(\theta_2)$, we can switch between different feasible sets, without affecting the overall optimal value of the optimization problem.

In the context of Lemma 1, the above result allows us to restrict attention from a potentially large set of relevant points (the 2^k vertices of the hyperrectangle $\mathcal{H}_{k+1}$), to the $k+1$ vertices found on the right side of the zonogon Θ, which also gives insight into why the construction of an affine controller $q_{k+1}(\boldsymbol{w}_{[k+1]})$ with $k+1$ degrees of freedom, yielding the same overall objective function value J_{mM}, might actually be possible.

In the remaining part of Section 2.4.1, we further narrow down this set of relevant points, by using the structure and properties of the optimal control law $u^\star_{k+1}(x_{k+1})$ and optimal value function $J^\star_{k+1}(x_{k+1})$, derived in Section 2.2. Before proceeding, however, we first reduce the notational clutter by introducing several simplifications and assumptions.

Simplified Notation and Assumptions.

For the remaining part of the chapter, we seek a simplified notation as much as possible, in order to clarify the key ideas. To start, we omit the time subscript $k+1$ whenever possible, so that we write $\boldsymbol{w}_{[k+1]} \equiv \boldsymbol{w}$, $q_{k+1}(\cdot) \equiv q(\cdot)$, $J^\star_{k+1}(\cdot) \equiv J^\star(\cdot)$, $g_{k+1}(\cdot) \equiv g(\cdot)$. The affine functions $\theta_{1,2}(\boldsymbol{w}_{[k+1]})$ and $q_{k+1}(\boldsymbol{w}_{[k+1]})$ are written:

$$\theta_1(\boldsymbol{w}) \stackrel{\text{def}}{=} a_0 + \boldsymbol{a}'\,\boldsymbol{w}; \quad \theta_2(\boldsymbol{w}) \stackrel{\text{def}}{=} b_0 + \boldsymbol{b}'\,\boldsymbol{w}; \quad q(\boldsymbol{w}) \stackrel{\text{def}}{=} q_0 + \boldsymbol{q}'\,\boldsymbol{w}\ , \tag{2.21}$$

where $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^k$ are the *generators* of the zonogon Θ. Since θ_2 is nothing but the state x_{k+1}, instead of referring to $J^\star_{k+1}(x_{k+1})$ and $u^\star_{k+1}(x_{k+1})$, we use $J^\star(\theta_2)$ and $u^\star(\theta_2)$.

Since our exposition relies heavily on sets given by maps $\boldsymbol{\gamma} : \mathbb{R}^k \mapsto \mathbb{R}^2$ ($k \geq 2$), in order to reduce the number of symbols, we denote the resulting coordinates in $\mathbb{R}^2$ by γ_1, γ_2, and use the following overloaded notation:

- $\gamma_i[\boldsymbol{v}]$ denotes the γ_i-coordinate of the point $\boldsymbol{v} \in \mathbb{R}^2$,

- $\gamma_i(\boldsymbol{w})$ is the value assigned by the i-th component of the map $\boldsymbol{\gamma}$ to $\boldsymbol{w} \in \mathbb{R}^k$ (equivalently, $\gamma_i(\boldsymbol{w}) \equiv \gamma_i[\boldsymbol{\gamma}(\boldsymbol{w})]$).

The different use of parentheses should remove any ambiguity from the notation (particularly in the case $k = 2$). For the same (γ_1, γ_2) coordinate system, we use $\text{cotan}\big(\boldsymbol{M},\, \boldsymbol{N}\big)$ to denote the cotangent of the angle formed by an oriented line segment $[\boldsymbol{M}, \boldsymbol{N}] \in \mathbb{R}^2$ with the γ_1-axis,

$$\text{cotan}\big(\boldsymbol{M},\, \boldsymbol{N}\big) \overset{\text{def}}{=} \frac{\gamma_1[\boldsymbol{N}] - \gamma_1[\boldsymbol{M}]}{\gamma_2[\boldsymbol{N}] - \gamma_2[\boldsymbol{M}]}. \tag{2.22}$$

Also, to avoid writing multiple functional compositions, since most quantities of interest depend solely on the state x_{k+1} (which is the same as θ_2), we use the following shorthand notation for any point $\boldsymbol{v} \in \mathbb{R}^2$, with corresponding θ_2-coordinate given by $\theta_2[\boldsymbol{v}]$:

$$u^\star\big(\theta_2[\boldsymbol{v}]\big) \equiv u^\star(\boldsymbol{v}); \quad J^\star\big(\theta_2[\boldsymbol{v}]\big) \equiv J^\star(\boldsymbol{v}); \quad g\big(\theta_2[\boldsymbol{v}] + u^\star(\theta_2[\boldsymbol{v}])\big) \equiv g(\boldsymbol{v}).$$

We use the same counter-clockwise numbering of the vertices of Θ as introduced earlier in Section 2.4.1,

$$\boldsymbol{v}_0 \overset{\text{def}}{=} \boldsymbol{v}_{\min}, \ldots, \boldsymbol{v}_p \overset{\text{def}}{=} \boldsymbol{v}_{\max}, \ldots, \boldsymbol{v}_{2p} = \boldsymbol{v}_{\min} \ , \tag{2.23}$$

where $2p$ is the number of vertices of Θ, and we also make the following simplifying assumptions:

Assumption 1. *The uncertainty vector at time $k+1$, $\boldsymbol{w}_{[k+1]} = (w_1, \ldots, w_k)$, belongs to the unit hypercube of $\mathbb{R}^k$, i.e., $\mathcal{H}_{k+1} = \mathcal{W}_1 \times \cdots \times \mathcal{W}_k \equiv [0,1]^k$.*

Assumption 2. *The zonogon Θ has a maximal number of vertices, i.e., $p = k$.*

Assumption 3. *The vertex of the hypercube projecting to $\boldsymbol{v}_i$, $i \in \{0, \ldots, k\}$, is exactly $[1, 1, \ldots, 1, 0, \ldots, 0]$, i.e., 1 in the first i components and 0 thereafter (see Figure 2-2).*

These assumptions are made only to facilitate the exposition, and result in no loss of generality. To see this, note that the conditions of Assumption 1 can always be achieved by adequate translation and scaling of the generators $\boldsymbol{a}$ and $\boldsymbol{b}$ (refer to Section A.2 of the Appendix for more details), and Assumption 3 can be satisfied by renumbering and possibly reflecting[4] the coordinates of the hyperrectangle, i.e., the disturbances $w_1, \ldots, w_k$. As for Assumption 2, we argue that an extension of our construction to the degenerate case $p < k$ is immediate (one could also remove the degeneracy by applying an infinitesimal perturbation to the generators $\boldsymbol{a}$ or $\boldsymbol{b}$, with infinitesimal cost implications).

Further Analysis of the Induction Hypothesis.

In the simplified notation, equation (2.16) can now be rewritten, using (2.9) to express $J^\star(\cdot)$ as a function of $u^\star(\cdot)$ and $g(\cdot)$, as follows:

$$(OPT) \quad J_{mM} = \max_{(\gamma_1,\gamma_2)\in\Gamma^\star} \Big[\gamma_1 + g(\gamma_2) \Big], \tag{2.24a}$$

$$\Gamma^\star \stackrel{\text{def}}{=} \Big\{ (\gamma_1^\star, \gamma_2^\star) \; : \; \gamma_1^\star \stackrel{\text{def}}{=} \theta_1 + c\cdot u^\star(\theta_2), \quad \gamma_2^\star \stackrel{\text{def}}{=} \theta_2 + u^\star(\theta_2), \quad (\theta_1,\theta_2)\in\Theta \Big\}. \tag{2.24b}$$

In this form, (OPT) represents the optimization problem solved by the uncertainties $\boldsymbol{w} \in \mathcal{H}$ when the optimal policy, $u^\star(\cdot)$, is used at time $k+1$. The significance of $\gamma_{1,2}^\star$ in the context of the original problem is straightforward: $\gamma_1^\star$ stands for the cumulative past stage costs, plus the current-stage control cost $c \cdot u^\star$, while $\gamma_2^\star$, which is the same variable as y_{k+1}, is the sum of the state and the control (in the inventory Example 1, it would represent the inventory position just after ordering, before seeing the demand).

Note that we have $\Gamma^\star \equiv \boldsymbol{\gamma}^\star(\Theta)$, where a characterization for the map $\boldsymbol{\gamma}^\star$ can be

[4] Reflection would represent a transformation $w_i \mapsto 1 - w_i$. As we show in a later result (Lemma 4 of Section 2.4.2), reflection is actually not needed, but this is not obvious at this point.

obtained by replacing the optimal policy, given by (2.8), in equation (2.24b):

$$\boldsymbol{\gamma}^\star : \mathbb{R}^2 \to \mathbb{R}^2,\ \boldsymbol{\gamma}^\star(\boldsymbol{\theta}) \equiv \big(\gamma_1^\star(\boldsymbol{\theta}),\, \gamma_2^\star(\boldsymbol{\theta})\big) = \begin{cases} (\theta_1 + c\cdot U,\ \theta_2 + U)\ , & \text{if } \theta_2 < y^\star - U \\ (\theta_1 - c\cdot \theta_2 + c\cdot y^\star,\ y^\star)\ , & \text{otherwise} \\ (\theta_1 + c\cdot L,\ \theta_2 + L)\ , & \text{if } \theta_2 > y^\star - L \end{cases} \tag{2.25}$$

The following is a compact characterization for the maximizers in problem (OPT) from (2.24a):

Lemma 2. *The maximum in problem* (OPT) *over* $\Gamma^\star$ *is reached on the right side of:*

$$\Delta_{\Gamma^\star} \stackrel{\text{def}}{=} \operatorname{conv}\left(\{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_k^\star\}\right)\ , \tag{2.26}$$

where:

$$\boldsymbol{y}_i^\star \stackrel{\text{def}}{=} \boldsymbol{\gamma}^\star(\boldsymbol{v}_i) = \big(\theta_1[\boldsymbol{v}_i] + c\cdot u^\star(\boldsymbol{v}_i),\, \theta_2[\boldsymbol{v}_i] + u^\star(\boldsymbol{v}_i)\big), \quad i \in \{0, \ldots, k\}. \tag{2.27}$$

Proof. By Lemma 1, the maximum in (2.16) is reached at one of the vertices $\boldsymbol{v}_0$, $\boldsymbol{v}_1$, $\ldots$, $\boldsymbol{v}_k$ of the zonogon Θ. Since this problem is equivalent to problem (OPT) in (2.24b), written over $\Gamma^\star$, we can immediately conclude that the maximum of the latter problem is reached at the points $\{\boldsymbol{y}_i^\star\}_{1\le i\le k}$ given by (2.27). Furthermore, since $g(\cdot)$ is convex (see Property (**P2**) of the optimal DP solution, in Section 2.2), we can apply Corollary 1, and replace the points $\boldsymbol{y}_i^\star$ with the right side of their convex hull, $\text{r-side}\,(\Delta_{\Gamma^\star})$, without changing the result of the optimization problem, which completes the proof. ☐

Since this result is central to our future construction and proof, we spend the remaining part of the subsection discussing some of the properties of the main object of interest, the set, $\text{r-side}(\Delta_{\Gamma^\star})$. To understand the geometry of the set $\Delta_{\Gamma^\star}$, and its connection with the optimal control law, note that the mapping $\boldsymbol{\gamma}^\star$ from Θ to $\Gamma^\star$ discriminates points $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \Theta$ depending on their position relative to the horizontal band

$$\mathcal{B}_{LU} \stackrel{\text{def}}{=} \left\{ \ (\theta_1, \theta_2) \in \mathbb{R}^2 \ : \ \theta_2 \in [y^\star - U, y^\star - L] \ \right\}. \tag{2.28}$$

In terms of the original problem, the band $\mathcal{B}_{LU}$ represents the portion of the state space x_{k+1} (i.e., θ_2) in which the optimal control policy $u^\star$ is unconstrained by the bounds L, U. More precisely, points below $\mathcal{B}_{LU}$ and points above $\mathcal{B}_{LU}$ correspond to state-space regions where the upper-bound, U, and the lower bound, L, are active, respectively.

With respect to the geometry of $\Gamma^\star$, we can use (2.25) and the definition of $\boldsymbol{v}_0, \dots, \boldsymbol{v}_k$ to distinguish a total of four distinct cases. The first three, shown in Figure 2-4, are very easy to analyze:

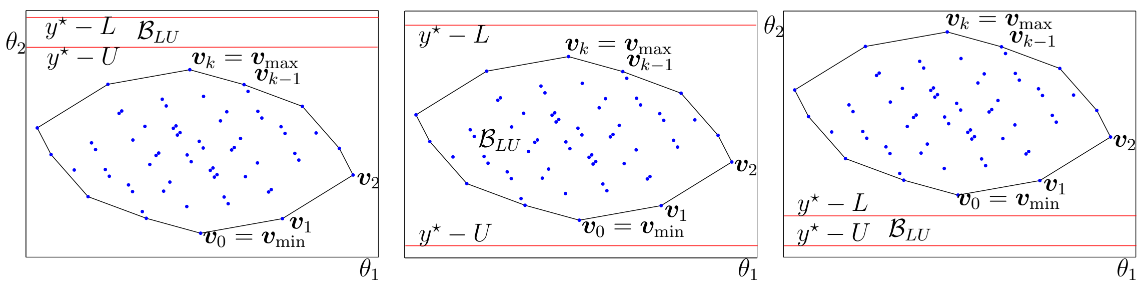


Figure 2-4: Trivial cases, when zonogon Θ lies entirely **[C1]** below, **[C2]** inside, or **[C3]** above the band $\mathcal{B}_{LU}$.

[C1] If the entire zonogon Θ falls below the band $\mathcal{B}_{LU}$, i.e., $\theta_2\left[\boldsymbol{v}_k\right] < y^\star - U$, then $\Gamma^\star$ is simply a translation of Θ, by $(c \cdot U, U)$, so that $\text{r-side}\left(\Delta_{\Gamma^\star}\right) = \{\boldsymbol{y}_0^\star, \boldsymbol{y}_1^\star, \dots, \boldsymbol{y}_k^\star\}$.

[C2] If Θ lies inside the band $\mathcal{B}_{LU}$, i.e., $y^\star - U \leq \theta_2\left[\boldsymbol{v}_0\right] \leq \theta_2\left[\boldsymbol{v}_k\right] \leq y^\star - L$, then all the points in $\Gamma^\star$ will have $\gamma_2^\star = y^\star$, so $\Gamma^\star$ will be a line segment, and $|\text{r-side}\left(\Delta_{\Gamma^\star}\right)| = 1$.

[C3] If the entire zonogon Θ falls above the band $\mathcal{B}_{LU}$, i.e., $\theta_2\left[\boldsymbol{v}_0\right] > y^\star - L$, then $\gamma^\star$ is again a translation of Θ, by $(c \cdot L, L)$, so, again $\text{r-side}\left(\Delta_{\Gamma^\star}\right) = \{\boldsymbol{y}_0^\star, \boldsymbol{y}_1^\star, \dots, \boldsymbol{y}_k^\star\}$.

The remaining case, **[C4]**, is when Θ intersects the horizontal band $\mathcal{B}_{LU}$ in a nontrivial fashion. We can separate this situation in the three sub-cases shown in

Figure 2-5, depending on the position of the vertex $\boldsymbol{v}_t \in$ r-side(Θ), where the index t

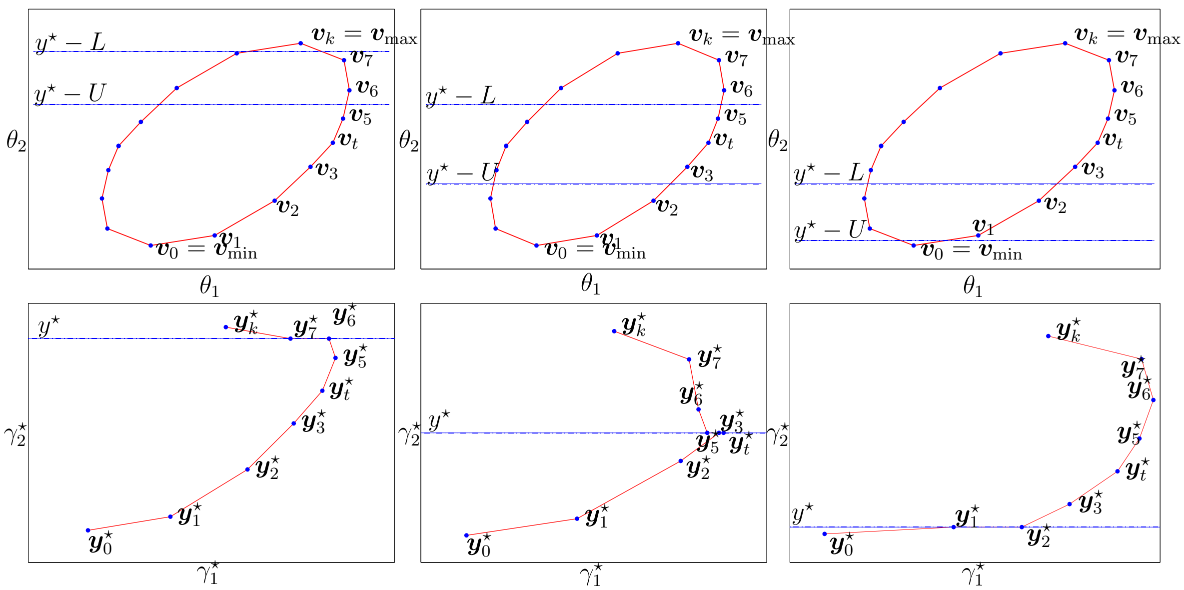


Figure 2-5: Case **[C4]**. Original zonogon Θ (first row) and the set $\Gamma^\star$ (second row) when $\boldsymbol{v}_t$ falls (a) under, (b) inside or (c) above the band $\mathcal{B}_{LU}$.

relates the per-unit control cost, c, with the geometrical properties of the zonogon:

$$t \stackrel{\text{def}}{=} \begin{cases} 0 \ , & \text{if } \frac{a_1}{b_1} \leq c \\ \max\Big\{ i \in \{1, \ldots, k\} : \frac{a_i}{b_i} > c \Big\} \ , & \text{otherwise} \ . \end{cases} \tag{2.29}$$

We remark that the definition of t is consistent, since, by the simplifying Assumption 3, the generators $\boldsymbol{a}, \boldsymbol{b}$ of the zonogon Θ always satisfy:

$$\begin{cases} \frac{a_1}{b_1} > \frac{a_2}{b_2} > \cdots > \frac{a_k}{b_k} \\ b_1, b_2, \ldots, b_k \geq 0. \end{cases} \tag{2.30}$$

An equivalent characterization of $\boldsymbol{v}_t$ can be obtained as the result of an optimization problem,

$$\boldsymbol{v}_t \equiv \arg\min \Big\{ \theta_2 : \boldsymbol{\theta} \in \arg\max \big\{ \theta_1' - c \cdot \theta_2' \ : \ \boldsymbol{\theta}' \in \Theta \big\} \Big\}.$$

The following lemma summarizes all the relevant geometrical properties corresponding to this case:

Lemma 3. *When the zonogon* Θ *has a non-trivial intersection with the band* $\mathcal{B}_{LU}$ *(case* **[C4]***), the convex polygon* $\Delta_{\Gamma^\star}$ *and the set of points on its right side,* r-side$(\Delta_{\Gamma^\star})$*, verify the following properties:*

1. r-side$(\Delta_{\Gamma^\star})$ *is the union of two sequences of consecutive vertices (one starting at* $\boldsymbol{y}_0^\star$*, and one ending at* $\boldsymbol{y}_k^\star$*), and possibly an additional vertex,* $\boldsymbol{y}_t^\star$*:*

$$\text{r-side}(\Delta_{\Gamma^\star}) = \{\boldsymbol{y}_0^\star, \boldsymbol{y}_1^\star, \ldots, \boldsymbol{y}_s^\star\} \cup \{\boldsymbol{y}_t^\star\} \cup \left\{\boldsymbol{y}_r^\star, \boldsymbol{y}_{r+1}^\star \ldots, \boldsymbol{y}_k^\star\right\},$$

   for some $s \leq r \in \{0, \ldots, k\}$*.*

2. *With* $\text{cotan}(\cdot,\,\cdot)$ *given by* (2.22) *applied to the* $(\gamma_1^\star, \gamma_2^\star)$ *coordinates, we have that:*

$$\begin{cases} \text{cotan}\left(\boldsymbol{y}_s^\star,\, \boldsymbol{y}_{\min(t,r)}^\star\right) \geq \frac{a_{s+1}}{b_{s+1}}, & \text{whenever } t > s \\ \text{cotan}\left(\boldsymbol{y}_{\max(t,s)}^\star,\, \boldsymbol{y}_r^\star\right) \leq \frac{a_r}{b_r}, & \text{whenever } t < r. \end{cases} \tag{2.31}$$

While the proof of the lemma is slightly technical (which is why we have decided to leave it for Section A.2.1 of the Appendix), its implications are more straightforward. In conjuction with Lemma 2, it provides a compact characterization of the points $\boldsymbol{y}_i^\star \in \Gamma^\star$ which are potential maximizers of problem (OPT) in (2.24a), which immediately narrows the set of relevant points $\boldsymbol{v}_i \in \Theta$ in optimization problem (2.16), and, implicitly, the set of disturbances $\boldsymbol{w} \in \mathcal{H}_{k+1}$ that can achieve the overall min-max cost.

2.4.2 Construction of the Affine Control Law.

Having analyzed the consequences that result from using the induction hypothesis of Theorem 1, we now return to the task of completing the inductive proof, which amounts to constructing an affine control law $q_{k+1}(\boldsymbol{w}_{[k+1]})$ and an affine cost $z_{k+1}(\boldsymbol{w}_{[k+2]})$ that verify conditions (2.11a), (2.11b), and (2.11c) in Theorem 1. We separate this task into two parts. In the current section, we exhibit an affine control law $q_{k+1}(\boldsymbol{w}_{[k+1]})$ that is robustly feasible, i.e., satisfies constraint (2.11a), and that leaves the overall min-max cost $J_1^\star(x_1)$ unchanged, when used at time $k+1$ in conjunction with the original convex state cost, $h_{k+1}(x_{k+2})$. The second part of the induction, i.e., the construction of the affine costs $z_{k+1}(\boldsymbol{w}_{[k+2]})$, is left for Section 2.4.3.

In the simplified notation introduced earlier, the problem we would like to solve is to find an affine control law $q(\boldsymbol{w})$ such that:

$$J_1^\star(x_1) = \max_{\boldsymbol{w}\in\mathcal{H}_{k+1}} \Big[\, \theta_1(\boldsymbol{w}) + c\cdot q(\boldsymbol{w}) + g\big(\theta_2(\boldsymbol{w}) + q(\boldsymbol{w})\big) \Big]$$

$$L \le q(\boldsymbol{w}) \le U\ , \quad \forall\, \boldsymbol{w} \in \mathcal{H}_{k+1}.$$

The maximization represents the problem solved by the disturbances, when the affine controller, $q(\boldsymbol{w})$, is used instead of the optimal controller, $u^\star(\theta_2)$. As such, the first equation amounts to ensuring that the overall objective function remains unchanged, and the inequalities are a restatement of the robust feasibility condition. The system can be immediately rewritten as

$$(AFF) \quad J_1^\star(x_1) = \max_{(\gamma_1,\gamma_2)\in\Gamma} \Big[\, \gamma_1 + g\left(\gamma_2\right) \Big] \tag{2.32a}$$

$$L \le q(\boldsymbol{w}) \le U\ , \quad \forall\, \boldsymbol{w} \in \mathcal{H}_{k+1} \tag{2.32b}$$

where

$$\Gamma \stackrel{\text{def}}{=} \Big\{ \ (\gamma_1,\gamma_2) \ : \ \gamma_1 \stackrel{\text{def}}{=} \theta_1(\boldsymbol{w}) + c\cdot q(\boldsymbol{w}), \quad \gamma_2 \stackrel{\text{def}}{=} \theta_2(\boldsymbol{w}) + q(\boldsymbol{w}), \quad \boldsymbol{w} \in \mathcal{H}_{k+1} \Big\}. \tag{2.33}$$

With this reformulation, all our decision variables, i.e., the affine coefficients of $q(\boldsymbol{w})$, have been moved to the feasible set Γ of the maximization problem (AFF) in (2.32a). Note that, with an affine controller $q(\boldsymbol{w}) = q_0 + \boldsymbol{q}'\,\boldsymbol{w}$, and $\theta_{1,2}$ affine in $\boldsymbol{w}$, the feasible set Γ will represent a new zonogon in $\mathbb{R}^2$, with generators given by $\boldsymbol{a} + c\cdot\boldsymbol{q}$ and $\boldsymbol{b} + \boldsymbol{q}$. Furthermore, since the function g is convex, the optimization problem (AFF) over Γ is of the exact same nature as that in (2.16), defined over the zonogon Θ. Thus, in perfect analogy with our discussion in Section 2.4.1 (Lemma 1 and Corollary 1), we can conclude that the maximum in (AFF) must occur at a vertex of Γ found in r-side(Γ).

In a different sense, note that optimization problem (AFF) is also very similar to problem (OPT) in (2.24b), which was the problem solved by the uncertainties $\boldsymbol{w}$ when the optimal control law, $u^\star(\theta_2)$, was used at time $k+1$. Since the optimal value of the latter problem is exactly equal to the overall min-max value, $J_1^\star(x_1)$, we interpret the equation in (2.32a) as comparing the optimal values in the two optimization problems, (AFF) and (OPT).

As such, note that the same convex objective function, $\gamma_1 + g(\gamma_2)$, is maximized in both problems, but over different feasible sets, $\Gamma^\star$ for (OPT) and Γ for (AFF), respectively. From Lemma 2 in Section 2.4.1, the maximum of problem (OPT) is reached on the set r-side($\Delta_{\Gamma^\star}$), where $\Delta_{\Gamma^\star} = \text{conv}\left(\{\boldsymbol{y}_0^\star, \boldsymbol{y}_1^\star, \ldots, \boldsymbol{y}_k^\star\}\right)$. From the discussion in the previous paragraph, the maximum in problem (AFF) occurs on r-side(Γ). Therefore, in order to compare the two results of the maximization problems, we must relate the sets r-side($\Delta_{\Gamma^\star}$) and r-side(Γ).

In this context, we introduce the central idea behind the construction of the affine control law, $q(\boldsymbol{w})$. Recalling the concept of a *zonogon hull* introduced in Definition 2, we argue that, if the affine coefficients of the controller, $q_0, \boldsymbol{q}$, were computed in such a way that the zonogon Γ actually corresponded to the zonogon hull of the set $\{\boldsymbol{y}_0^\star, \boldsymbol{y}_1^\star, \ldots, \boldsymbol{y}_k^\star\}$, then, by using the result in Corollary 1, we could immediately conclude that the optimal values in (OPT) and (AFF) are the same.

To this end, we introduce the following procedure for computing the affine control law $q(\boldsymbol{w})$:

Algorithm 1 Compute affine controller $q(\boldsymbol{w})$

Require: $\theta_1(\boldsymbol{w}), \theta_2(\boldsymbol{w}), g(\cdot), u^\star(\cdot)$

if (Θ falls below $\mathcal{B}_{LU}$) **or** ($\Theta \subseteq \mathcal{B}_{LU}$) **or** ($\Theta$ falls above $\mathcal{B}_{LU}$) **then**

  Return $q(\boldsymbol{w}) = u^\star(\theta_2(\boldsymbol{w}))$.

else

  Apply the mapping (2.25) to obtain the points $\boldsymbol{y}_i^\star,\ i \in \{0, \ldots, k\}$.

  Compute the set $\Delta_{\Gamma^\star} = \text{conv}\left(\{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_k^\star\}\right)$.

  Let $\text{r-side}(\Delta_{\Gamma^\star}) = \{\boldsymbol{y}_0^\star, \boldsymbol{y}_1^\star, \ldots, \boldsymbol{y}_s^\star\} \cup \{\boldsymbol{y}_t^\star\} \cup \{\boldsymbol{y}_r^\star, \ldots, \boldsymbol{y}_k^\star\}$.

  Solve the following system for $q_0, \ldots, q_k$ and K_U, K_L:

$$(S) \begin{cases} q_0 + \cdots + q_i = u^\star(\boldsymbol{v}_i), & \forall\, \boldsymbol{y}_i^\star \in \text{r-side}(\Delta_{\Gamma^\star}) \quad \text{(matching)} \\ \dfrac{a_i + c \cdot q_i}{b_i + q_i} = K_U, & \forall\, i \in \{s+1, \ldots, \min(t, r)\} \quad \text{(alignment below } t\text{)} \\ \dfrac{a_i + c \cdot q_i}{b_i + q_i} = K_L, & \forall\, i \in \{\max(t, s)+1, \ldots, r\} \quad \text{(alignment above } t\text{)} \end{cases} \tag{2.34}$$

  Return $q(\boldsymbol{w}) = q_0 + \sum_{i=1}^{k} q_i w_i$.

end if

Before proving that the construction is well-defined and produces the expected result, we first give some intuition for the constraints in system (2.34). In order to have the zonogon Γ be the same as the zonogon hull of $\{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_k^\star\}$, we must ensure that the vertices on the right side of Γ exactly correspond to the points on the right side of $\Delta_{\Gamma^\star} = \text{conv}\left(\{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_k^\star\}\right)$.This is achieved in two stages. First, we ensure that vertices $\boldsymbol{w}_i$ of the hypercube $\mathcal{H}_{k+1}$ that are mapped by the optimal control law $u^\star(\cdot)$ into points $\boldsymbol{v}_i^\star \in \text{r-side}(\Delta_{\Gamma^\star})$ (through the succession of mappings $\boldsymbol{w}_i \overset{(2.17)}{\mapsto} \boldsymbol{v}_i \in \text{r-side}(\Theta) \overset{(2.27)}{\mapsto} \boldsymbol{y}_i^\star \in \text{r-side}(\Delta_{\Gamma^\star})$), will be mapped by the affine control law, $q(\boldsymbol{w}_i)$, into the same point $\boldsymbol{y}_i^\star$ (through the mappings $\boldsymbol{w}_i \overset{(2.17)}{\mapsto} \boldsymbol{v}_i \in \text{r-side}(\Theta) \overset{(2.33)}{\mapsto} \boldsymbol{y}_i^\star \in \text{r-side}(\Delta_{\Gamma^\star})$). This is done in the first set of constraints, by *matching* the value of the optimal control law at any such points. Second, we ensure that any such matched points $\boldsymbol{y}_i^\star$ actually correspond to the vertices on the right side of the zonogon Γ. This is done in the second and third set of constraints in (2.34), by computing the affine coefficients q_j in such a way that the resulting segments in the generators of the zonogon Γ, namely $\binom{a_j+c\cdot q_j}{b_j+q_j}$, are all *aligned*, i.e., have the same cotangent, given by the K_U, K_L variables. Geometrically, this exactly corresponds to the situation shown in Figure 2-6 below.

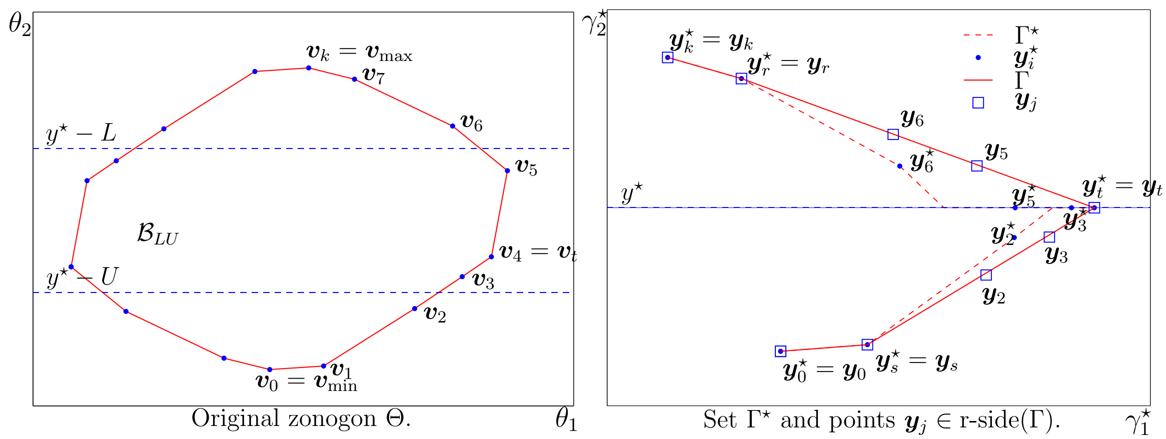


Figure 2-6: Outcomes from the matching and alignment performed in Algorithm 1.

We remark that the above algorithm does not explicitly require that the control $q(\boldsymbol{w})$ be robustly feasible, i.e., condition (2.32b). However, this condition turns out to hold as a direct result of the way matching and alignment are performed in Algorithm 1.

Affine Controller Preserves Overall Objective and Is Robust.

In this section, we prove that the affine control law $q(\boldsymbol{w})$ produced by Algorithm 1 satisfies the requirements of (2.32a), i.e., it is robustly feasible, and it preserves the overall objective function $J_1^\star(x_1)$, when used in conjunction with the original convex state costs, $h(\cdot)$. With the exception of Corollary 1, all the key results that we are using are contained in Section 2.4.1 (Lemmas 2 and 3). Therefore, we preserve the same notation and case discussion as initially introduced there.

First consider the condition on line 1 of Algorithm 1, and note that this corresponds to the three trivial cases **[C1]**, **[C2]** and **[C3]** of Section 2.4.1. In particular, since $\theta_2 \equiv x_{k+1}$, we can use (2.8) to conclude that in these cases, the optimal control law $u^\star(\cdot)$ is actually affine:

[C1] If Θ falls below the band $\mathcal{B}_{LU}$, then the upper bound constraint on the control at time k is always active, i.e., $u^\star\big(\theta_2(\boldsymbol{w})\big) = U, \ \forall\, \boldsymbol{w} \in \mathcal{H}_{k+1}$.

[C2] If $\Theta \subseteq \mathcal{B}_{LU}$, then the constraints on the control at time k are never active, i.e., $u^\star\big(\theta_2(\boldsymbol{w})\big) = y^\star - \theta_2(\boldsymbol{w})$, hence affine in $\boldsymbol{w}$, since θ_2 is affine in $\boldsymbol{w}$, by (2.21).

[C3] If Θ falls above the band $\mathcal{B}_{LU}$, then the lower bound constraint on the control is always active, i.e., $u^\star\big(\theta_2(\boldsymbol{w})\big) = L, \ \forall\, \boldsymbol{w} \in \mathcal{H}_{k+1}$.

Therefore, with the assignment in line 2 of Algorithm 1, we obtain an affine control law that is always feasible and also optimal.

When none of the trivial cases holds, we are in case **[C4]** of Section 2.4.1. Therefore, we can invoke the results from Lemma 3 to argue that the right side of the set $\Delta_{\Gamma^\star}$ is exactly the set on line 7 of the algorithm, i.e., r-side$(\Delta_{\Gamma^\star}) = \{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_s^\star\} \cup \{\boldsymbol{y}_t^\star\} \cup \{\boldsymbol{y}_r^\star, \ldots, \boldsymbol{y}_k^\star\}$. In this setting, we can now formulate the first claim about system (2.34) and its solution:

Lemma 4. *System* (2.34) *is always feasible, and the solution satisfies:*

1. $-b_i \le q_i \le 0, \ \forall\, i \in \{1, \ldots, k\}$.
2. $L \le q(\boldsymbol{w}) \le U, \ \forall\, \boldsymbol{w} \in \mathcal{H}_{k+1}$.

Proof. Note first that system (2.34) has exactly $k+3$ unknowns, two for the cotangents K_U, K_L, and one for each coefficient $q_i, 0 \le i \le k$. Also, since $|\text{r-side}(\Delta_{\Gamma^\star})| \le |\text{ext}(\Delta_{\Gamma^\star})| \le k+1$, and there are exactly $|\text{r-side}(\Delta_{\Gamma^\star})|$ matching constraints, and $k + 3 - |\text{r-side}(\Delta_{\Gamma^\star})|$ alignment constraints, it can be immediately seen that the system is always feasible.

Consider any q_i with $i \in \{1, \ldots, s\} \cup \{r+1, \ldots, k\}$. From the matching conditions, we have that $q_i = u^\star(\boldsymbol{v}_i) - u^\star(\boldsymbol{v}_{i-1})$. By Property (**P3**) from Section 2.2, the difference in the values of the optimal control law $u^\star(\cdot)$ satisfies:

$$\begin{aligned} u^\star(\boldsymbol{v}_i) - u^\star(\boldsymbol{v}_{i-1}) &\stackrel{\text{def}}{=} u^\star(\theta_2[\boldsymbol{v}_i]) - u^\star(\theta_2[\boldsymbol{v}_{i-1}]) \\ \big(\text{by } (\mathbf{P3})\big) \quad &= -f \cdot (\theta_2[\boldsymbol{v}_i] - \theta_2[\boldsymbol{v}_{i-1}]) \\ &\stackrel{(2.21)}{=} -f \cdot b_i, \qquad \text{where } f \in [0,1]. \end{aligned}$$

Since, by (2.30), $b_j \geq 0, \forall\, j \in \{1, \ldots, k\}$, we immediately obtain $-b_i \leq q_i \leq 0$, for $i \in \{1, \ldots, s\} \cup \{r+1, \ldots, k\}$.

Now consider any index $i \in \{s+1, \ldots, t \wedge r\}$, where $t \wedge r \equiv \min(t, r)$. From the conditions in system (2.34) for alignment below t, we have $q_i = \frac{a_i - K_U \cdot b_i}{K_U - c}$. By summing up all such relations, we obtain:

$$\begin{aligned}
\sum_{i=s+1}^{t\wedge r} q_i &= \frac{\sum_{i=s+1}^{t\wedge r} a_i - K_U \cdot \sum_{i=s+1}^{t\wedge r} b_i}{K_U - c} \quad \Leftrightarrow \qquad \text{(using the matching)} \\
u^\star(\boldsymbol{v}_{t\wedge r}) - u^\star(\boldsymbol{v}_s) &= \frac{\sum_{i=s+1}^{t\wedge r} a_i - K_U \cdot \sum_{i=s+1}^{t\wedge r} b_i}{K_U - c} \quad \Leftrightarrow \\
K_U &= \frac{\sum_{i=s+1}^{t\wedge r} a_i + c \cdot (u^\star(\boldsymbol{v}_{t\wedge r}) - u^\star(\boldsymbol{v}_s))}{\sum_{i=s+1}^{t\wedge r} b_i + u^\star(\boldsymbol{v}_{t\wedge r}) - u^\star(\boldsymbol{v}_s)} \\
&= \frac{\left[\sum_{i=0}^{t\wedge r} a_i + c \cdot u^\star(\boldsymbol{v}_{t\wedge r})\right] - \left[\sum_{i=0}^{s} a_i + c \cdot u^\star(\boldsymbol{v}_s)\right]}{\left[\sum_{i=0}^{t\wedge r} b_i + u^\star(\boldsymbol{v}_{t\wedge r})\right] - \left[\sum_{i=0}^{s} b_i + u^\star(\boldsymbol{v}_s)\right]} \\
&\overset{(2.27)}{=} \frac{\gamma_1^\star[\boldsymbol{y}_{t\wedge r}^\star] - \gamma_1^\star[\boldsymbol{y}_s^\star]}{\gamma_2^\star[\boldsymbol{y}_{t\wedge r}^\star] - \gamma_2^\star[\boldsymbol{y}_s^\star]} \overset{(2.22)}{=} \operatorname{cotan}\bigl(\boldsymbol{y}_s^\star,\, \boldsymbol{y}_{t\wedge r}^\star\bigr).
\end{aligned}$$

In the first step, we have used the fact that both $\boldsymbol{v}_s^\star$ and $\boldsymbol{v}_{\min(t,r)}^\star$ are matched, hence the intermediate coefficients q_i must sum to exactly the difference of the values of $u^\star(\cdot)$ at $\boldsymbol{v}_{\min(t,r)}$ and $\boldsymbol{v}_s$ respectively. In this context, we can see that K_U is simply the cotangent of the angle formed by the segment $[\boldsymbol{y}_s^\star, \boldsymbol{y}_{\min(t,r)}^\star]$ with the horizontal (i.e., $\gamma_1^\star$) axis. In this case, we can immediately recall result (2.31) from Lemma 3, to argue that $K_U \geq \frac{a_{s+1}}{b_{s+1}}$. Combining with (2.29) and (2.30), we obtain:

$$K_U \geq \frac{a_{s+1}}{b_{s+1}} \overset{(2.30)}{\geq \cdots \geq} \frac{a_{\min(t,r)}}{b_{\min(t,r)}} \geq \frac{a_t}{b_t} \overset{(2.29)}{>} c.$$

Therefore, we immediately have that for any $i \in \{s+1, \ldots, \min(t, r)\}$,

$$\begin{cases} a_i - K_U \cdot b_i \leq 0 \\ K_U - c > 0 \end{cases} \Rightarrow q_i = \frac{a_i - K_U \cdot b_i}{K_U - c} \leq 0\,, \quad \begin{cases} a_i - c \cdot b_i > 0 \\ q_i + b_i = \dfrac{a_i - c \cdot b_i}{K_U - c} \end{cases} \Rightarrow q_i + b_i \geq 0.$$

The argument for indices $i \in \{\max(t, s)+1, \ldots, r\}$ proceeds in exactly the same fashion, by recognizing that K_L defined in the algorithm is the same as $\operatorname{cotan}\bigl(\boldsymbol{y}_{\max(t,s)}^\star,\, \boldsymbol{y}_r^\star\bigr)$, and then applying (2.31) to argue that $K_L < \frac{a_r}{b_r} \leq \frac{a_{\max(t,s)+1}}{b_{\max(t,s)+1}} \leq \frac{a_{t+1}}{b_{t+1}} \leq c$. This will allow us to use the same reasoning as above, completing the proof of part (i) of the claim.

To prove part (ii), consider any $\boldsymbol{w} \in \mathcal{H}_{k+1} \stackrel{\text{def}}{=} [0,1]^k$. Using part (i), we obtain:

$$q(\boldsymbol{w}) \stackrel{\text{def}}{=} q_0 + \sum_{i=1}^{k} q_i \cdot w_i \ \leq \ (\text{since } w_i \in [0,1], q_i \leq 0) \ \leq q_0 \stackrel{(**)}{=} u^\star(\boldsymbol{v}_0) \leq U \ ,$$
$$q(\boldsymbol{w}) \geq q_0 + \sum_{i=1}^{k} q_i \cdot 1 \stackrel{(**)}{=} u^\star(\boldsymbol{v}_k) \geq L.$$

Note that in step $(**)$, we have critically used the result from Lemma 3 that, when $\Theta \nsubseteq \mathcal{B}_{LU}$, the points $\boldsymbol{v}_0^\star, \boldsymbol{v}_k^\star$ are always among the points on the right side of $\Delta_{\Gamma^\star}$, and, therefore, we always have the equations $q_0 = u^\star(\boldsymbol{v}_0), q_0 + \sum_{i=1}^{k} q_i = u^\star(\boldsymbol{v}_k)$ among the matching equations of system (2.34). For the last arguments, we have simply used the fact that the optimal control law, $u^\star(\cdot)$, is always feasible, hence $L \leq u^\star(\cdot) \leq U$. □

This completes our first goal, namely proving that the affine controller $q(\boldsymbol{w})$ is always robustly feasible. To complete the construction, we introduce the following final result:

Lemma 5. *The affine control law $q(\boldsymbol{w})$ computed in Algorithm 1 verifies equation* (2.32a)*.*

Proof. From (2.33), the affine controller $q(\boldsymbol{w})$ induces the generators $\boldsymbol{a}+c\cdot\boldsymbol{q}$ and $\boldsymbol{b}+\boldsymbol{q}$ for the zonogon Γ. This implies that Γ will be the Minkowski sum of the following segments in $\mathbb{R}^2$:

$$\begin{bmatrix} a_1+c\cdot q_1 \\ b_1+q_1 \end{bmatrix}, \ldots, \begin{bmatrix} a_s+c\cdot q_s \\ b_s+q_s \end{bmatrix}, \begin{bmatrix} K_U\cdot(b_{s+1}+q_{s+1}) \\ b_{s+1}+q_{s+1} \end{bmatrix}, \ldots, \begin{bmatrix} K_U\cdot\big(b_{\min(t,r)}+q_{\min(t,r)}\big) \\ b_{\min(t,r)}+q_{\min(t,r)} \end{bmatrix},$$
$$\begin{bmatrix} K_L\cdot\big(b_{\max(t,s)+1}+q_{\max(t,s)+1}\big) \\ b_{\max(t,s)+1}+q_{\max(t,s)+1} \end{bmatrix} \ldots, \begin{bmatrix} K_L\cdot(b_r+q_r) \\ b_r+q_r \end{bmatrix}, \begin{bmatrix} a_{r+1}+c\cdot q_{r+1} \\ b_{r+1}+q_{r+1} \end{bmatrix}, \ldots, \begin{bmatrix} a_k+c\cdot q_k \\ b_k+q_k \end{bmatrix}. \quad (2.35)$$

From Lemma 4, we have that $q_i + b_i \geq 0, \forall\, i \in \{1,\ldots,k\}$. Therefore, if we consider the points in $\mathbb{R}^2$:

$$\boldsymbol{y}_i = \left(\sum_{j=0}^{i} (a_j + c \cdot q_j), \quad \sum_{j=0}^{i} (b_j + q_j) \right), \quad \forall\, i \in \{0, \ldots, k\},$$

we can make the following simple observations:

- For any vertex $\boldsymbol{v}_i \in \Theta$, $i \in \{0, \ldots, k\}$, that is matched, i.e., $\boldsymbol{y}_i^\star \in \text{r-side}(\Delta_{\Gamma^\star})$, if we let $\boldsymbol{w}_i$ represent the unique[5] vertex of the hypercube $\mathcal{H}_k$ projecting onto $\boldsymbol{v}_i$, i.e., $\boldsymbol{v}_i = (\theta_1(\boldsymbol{w}_i), \theta_2(\boldsymbol{w}_i))$, then we have:

$$\boldsymbol{y}_i \overset{(2.33)}{=} \big(\gamma_1(\boldsymbol{w}_i),\, \gamma_2(\boldsymbol{w}_i)\big) \overset{(2.34)}{=} \big(\gamma_1^\star(\boldsymbol{v}_i),\, \gamma_2^\star(\boldsymbol{v}_i)\big) \overset{(2.27)}{=} \boldsymbol{y}_i^\star.$$

  The first equality follows from the definition of the mapping that characterizes the zonogon Γ. The second equality follows from the fact that for any matched vertex $\boldsymbol{v}_i$, the coordinates in $\Gamma^\star$ and Γ are exactly the same, and the last equality is simply the definition of the point $\boldsymbol{y}_i^\star$.

- For any vertex $\boldsymbol{v}_i \in \Theta$, $i \in \{0, \ldots, k\}$, that is *not* matched, we have:

$$\begin{aligned} \boldsymbol{y}_i &\in [\boldsymbol{y}_s, \boldsymbol{y}_{\min(t,r)}], \quad \forall\, i \in \{s+1, \ldots, \min(t,r)-1\} \\ \boldsymbol{y}_i &\in [\boldsymbol{y}_{\max(t,s)}, \boldsymbol{y}_r], \quad \forall\, i \in \{\max(t,s)+1, \ldots, r-1\}. \end{aligned}$$

  This can be seen directly from (2.35), since the segments in $\mathbb{R}^2$ given by $[\boldsymbol{y}_s, \boldsymbol{y}_{s+1}]$, $\ldots$, $[\boldsymbol{y}_{\min(t,r)-1}, \boldsymbol{y}_{\min(t,r)}]$ are always aligned (with common cotangent, given by K_U), and, similarly, the segments $[\boldsymbol{y}_{\max(t,s)}, \boldsymbol{y}_{\max(t,s)+1}], \ldots, [\boldsymbol{y}_{r-1}, \boldsymbol{y}_r]$ are also aligned (with common cotangent K_L).

This exactly corresponds to the situation shown earlier in Figure 2-6. By combining the two observations, it can be seen that the points $\big\{\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_s, \boldsymbol{y}_{\max(t,s)},$

[5]This vertex is unique due to our standing Assumption 2 that the number of vertices in Θ is $2k$ (also see part (iv) of Lemma 13 in the Appendix).

$\boldsymbol{y}_{\min(t,r)}, \boldsymbol{y}_r, \ldots, \boldsymbol{y}_k\}$ will satisfy the following properties:

$$\begin{aligned} \boldsymbol{y}_i &= \boldsymbol{y}_i^\star, \ \forall\, \boldsymbol{y}_i^\star \in \text{r-side}(\Delta_{\Gamma^\star}) \ , \\ \text{cotan}\big(\boldsymbol{y}_0,\, \boldsymbol{y}_1\big) &\geq \text{cotan}\big(\boldsymbol{y}_1,\, \boldsymbol{y}_2\big) \geq \cdots \geq \text{cotan}\big(\boldsymbol{y}_{s-1},\, \boldsymbol{y}_s\big) \geq \text{cotan}\big(\boldsymbol{y}_s,\, \boldsymbol{y}_{\min(t,r)}\big) \geq \\ &\geq \text{cotan}\big(\boldsymbol{y}_{\max(t,s)},\, \boldsymbol{y}_r\big) \geq \text{cotan}\big(\boldsymbol{y}_r,\, \boldsymbol{y}_{r+1}\big) \geq \cdots \geq \text{cotan}\big(\boldsymbol{y}_{k-1},\, \boldsymbol{y}_k\big), \end{aligned}$$

where the second relation follows simply because the points $\boldsymbol{y}_i^\star \in \text{r-side}(\Delta_{\Gamma^\star})$ are extreme points on the right side of a convex hull, and thus satisfy the same string of inequalities. This immediately implies that this set of $\boldsymbol{y}_i$ exactly represent the right side of the zonogon Γ, which, in turn, implies that $\Gamma \equiv \text{z-hull}\big(\{\boldsymbol{y}_0^\star, \boldsymbol{y}_1^\star, \ldots, \boldsymbol{y}_s^\star, \boldsymbol{y}_{\max(t,s)}^\star, \boldsymbol{y}_{\min(t,r)}^\star, \boldsymbol{y}_r^\star, \boldsymbol{y}_{r+1}^\star, \ldots, \boldsymbol{y}_k^\star\}\big)$. But then, by Corollary 1, the maximum value of problem (OPT) in (2.24b) is equal to the maximum value of problem (AFF) in (2.32a), and, since the former is always J_{mM}, so is that latter. □

This concludes the construction of the affine control law $q(\boldsymbol{w})$. We have shown that the policy computed by Algorithm 1 satisfies the conditions (2.32b) and (2.32a), i.e., is robustly feasible (by Lemma 4) and, when used in conjunction with the original convex state costs, preserves the overall optimal min-max value $J_1^\star(x_1)$ (Lemma 5).

2.4.3 Construction of the Affine State Cost.

Note that we have essentially completed the first part of the induction step. For the second part, we would still need to show how an affine stage cost can be computed, such that constraints (2.11b) and (2.11c) are satisfied. We return temporarily to the notation containing time indices, so as to put the current state of the proof into perspective.

In solving problem (AFF) of (2.32a), we have shown that there exists an affine $q_{k+1}(\boldsymbol{w}_{[k+1]})$ such that:

$$\begin{aligned} J_1^\star(x_1) \ &= \max_{\boldsymbol{w}_{[k+1]} \in \mathcal{H}_{k+1}} \Big[\theta_1(\boldsymbol{w}_{[k+1]}) + c_{k+1}\, q_{k+1}(\boldsymbol{w}_{[k+1]}) + g_{k+1}\big(\theta_2(\boldsymbol{w}_{[k+1]}) + q_{k+1}(\boldsymbol{w}_{[k+1]})\big)\Big] \\ &\overset{(2.33)}{=} \max_{\boldsymbol{w}_{[k+1]} \in \mathcal{H}_{k+1}} \Big[\, \gamma_1(\boldsymbol{w}_{[k+1]}) + g_{k+1}\big(\gamma_2(\boldsymbol{w}_{[k+1]})\big) \,\Big]. \end{aligned}$$

Using the definition of $g_{k+1}(\cdot)$ from (2.7b), we can write the above (only retaining the second term) as:

$$\begin{aligned} J_1^\star(x_1) &= \max_{\boldsymbol{w}_{[k+1]} \in \mathcal{H}_k} \Big[\gamma_1(\boldsymbol{w}_{[k+1]}) + \max_{w_{k+1} \in \mathcal{W}_{k+1}} \Big[h_{k+2}\big(\gamma_2(\boldsymbol{w}_{[k+1]}) + w_{k+2}\big) \\ &\qquad\qquad + J_{k+2}^\star\big(\gamma_2(\boldsymbol{w}_{[k+1]}) + w_{k+2}\big) \Big] \Big] \\ &\stackrel{\text{def}}{=} \max_{\boldsymbol{w}_{[k+2]} \in \mathcal{H}_{k+2}} \Big[\tilde{\gamma}_1(\boldsymbol{w}_{[k+2]}) + h_{k+2}\big(\tilde{\gamma}_2(\boldsymbol{w}_{[k+2]})\big) + J_{k+2}^\star\big(\tilde{\gamma}_2(\boldsymbol{w}_{[k+2]})\big) \Big], \end{aligned}$$

where $\tilde{\gamma}_1(\boldsymbol{w}_{[k+2]}) \stackrel{\text{def}}{=} \gamma_1(\boldsymbol{w}_{[k+1]})$, and $\tilde{\gamma}_2(\boldsymbol{w}_{[k+2]}) \stackrel{\text{def}}{=} \gamma_2(\boldsymbol{w}_{[k+1]}) + w_{k+2}$. In terms of physical interpretation, $\tilde{\gamma}_1$ has the same significance as γ_1, i.e., the cumulative past costs (including the control cost at time $k+1$, $c \cdot q_{k+1}$), while $\tilde{\gamma}_2$ represents the state at time $k+2$, i.e., x_{k+2}.

Geometrically, is is easy to note that

$$\tilde{\Gamma} \stackrel{\text{def}}{=} \Big\{ \big(\tilde{\gamma}_1(\boldsymbol{w}_{[k+2]}),\, \tilde{\gamma}_2(\boldsymbol{w}_{[k+2]})\big) \;:\; \boldsymbol{w}_{[k+2]} \in \mathcal{H}_{k+2} \Big\} \tag{2.36}$$

represents yet another zonogon, obtained by projecting a hyperrectangle $\mathcal{H}_{k+2} \subset \mathbb{R}^{k+1}$ into $\mathbb{R}^2$. It has a particular shape relative to the zonogon $\Gamma = (\gamma_1, \gamma_2)$, since the generators of $\tilde{\Gamma}$ are simply obtained by appending a 0 and a 1, respectively, to the generators of Γ, which implies that $\tilde{\Gamma}$ is the convex hull of two translated copies of Γ, where the translation occurs on the $\tilde{\gamma}_2$ axis. As it turns out, this fact will bear little importance for the discussion to follow, so we include it here only for completeness.

In this context, the problem we would like to solve is to replace the convex function $h_{k+2}\big(\tilde{\gamma}_2(\boldsymbol{w}_{[k+2]})\big)$ with an affine function $z_{k+2}(\boldsymbol{w}_{[k+2]})$, such that the analogues of conditions (2.11b) and (2.11c) are obeyed:

$$\begin{aligned} z_{k+2}(\boldsymbol{w}_{[k+2]}) \geq h_{k+2}\big(\tilde{\gamma}_2(\boldsymbol{w}_{[k+2]})\big), &\qquad \forall\, \boldsymbol{w}_{[k+2]} \in \mathcal{H}_{k+2}, \\ J_1^\star(x_1) = \max_{\boldsymbol{w}_{[k+2]} \in \mathcal{H}_{k+2}} \Big[\tilde{\gamma}_1(\boldsymbol{w}_{[k+2]}) + z_{k+2}(\boldsymbol{w}_{[k+2]}) &+ J_{k+2}^\star\big(\tilde{\gamma}_2(\boldsymbol{w}_{[k+2]})\big) \Big]. \end{aligned}$$

We can now switch back to the simplified notation, where the time subscript $k+2$ is removed. Furthermore, to preserve as much of the familiar notation from Section 2.4.1, we denote the generators of zonogon $\tilde{\Gamma}$ by $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^{k+1}$, and the coefficients of $z(\boldsymbol{w})$ by $z_0, \boldsymbol{z}$, so that we have:

$$\tilde{\gamma}_1(\boldsymbol{w}) = a_0 + \boldsymbol{a}'\,\boldsymbol{w}, \qquad \tilde{\gamma}_2(\boldsymbol{w}) = b_0 + \boldsymbol{b}'\,\boldsymbol{w}, \qquad z(\boldsymbol{w}) = z_0 + \boldsymbol{z}'\,\boldsymbol{w}. \tag{2.37}$$

In perfect analogy to our discussion in Section 2.4.1, we can introduce:

$$\begin{aligned}
&\boldsymbol{v}_{\min} \stackrel{\text{def}}{=} \arg\max\big\{\tilde{\gamma}_1 : \tilde{\boldsymbol{\gamma}} \in \arg\min\{\xi_2' : \boldsymbol{\xi}' \in \tilde{\Gamma}\}\big\};\\
&\boldsymbol{v}_{\max} \stackrel{\text{def}}{=} 2\boldsymbol{O} - \boldsymbol{v}_{\min} \ \Big(\boldsymbol{O} \text{ is the center of } \tilde{\Gamma}\Big)\\
&\boldsymbol{v}_0 \stackrel{\text{def}}{=} \boldsymbol{v}_{\min}, \ldots, \boldsymbol{v}_{2p_1} = \boldsymbol{v}_{\min} \ \big(\text{counter-clockwise numbering of } \tilde{\Gamma}\text{'s vertices}\big).
\end{aligned} \tag{2.38}$$

Without loss of generality, we work, again, under Assumptions 1, 2, and 3, i.e., we analyze the case when $\mathcal{H}_{k+2} = [0,1]^{k+1}$, $p_1 = k+1$ (the zonogon $\tilde{\Gamma}$ has a maximal number of vertices), and $\boldsymbol{v}_i = [1, 1, \ldots, 1, 0, \ldots, 0]$ (ones in the first i positions). We also use the same overloaded notation when referring to the map $\tilde{\boldsymbol{\gamma}} : \mathbb{R}^{k+1} \to \mathbb{R}^2$ (i.e., $\tilde{\gamma}_{1,2}(\boldsymbol{w})$ denote the value assigned by the map to a point $\boldsymbol{w} \in \mathcal{H}_{k+2}$, while $\tilde{\gamma}_{1,2}[\boldsymbol{v}_i]$ are the $\tilde{\gamma}_{1,2}$ coordinates of a point $\boldsymbol{v}_i \in \mathbb{R}^2$), and we write $h(\boldsymbol{v}_i)$ and $J^\star(\boldsymbol{v}_i)$ instead of $h(\tilde{\gamma}_2[\boldsymbol{v}_i])$ and $J^\star(\tilde{\gamma}_2[\boldsymbol{v}_i])$, respectively.

With the simplified notation, the goal is to find $z(\boldsymbol{w})$ such that:

$$z(\boldsymbol{w}) \geq h\big(\tilde{\gamma}_2(\boldsymbol{w})\big), \qquad \forall\, \boldsymbol{w} \in \mathcal{H}_{k+1} \tag{2.39a}$$

$$\max_{(\tilde{\gamma}_1, \tilde{\gamma}_2) \in \tilde{\Gamma}} \Big[\tilde{\gamma}_1 + h(\tilde{\gamma}_2) + J^\star(\tilde{\gamma}_2)\Big] = \max_{\boldsymbol{w} \in \mathcal{H}_{k+1}} \Big[\tilde{\gamma}_1(\boldsymbol{w}) + z(\boldsymbol{w}) + J^\star\big(\tilde{\gamma}_2(\boldsymbol{w})\big)\Big] \tag{2.39b}$$

In (2.39b), the maximization on the left corresponds to the problem solved by the uncertainties, $\boldsymbol{w}$, when the original convex state cost, $h(\tilde{\gamma}_2)$, is incurred. As such, the result of the maximization is always exactly equal to $J_1^\star(x_1)$, the overall min-max value. The maximization on the right corresponds to the problem solved by the uncertainties when the affine cost, $z(\boldsymbol{w})$, is incurred instead of the convex cost. Requiring that the two optimal values be equal thus amounts to preserving the overall min-max value.

Since h and $J^\star$ are convex (see Property (**P2**) in Section 2.2), we can immediately use Lemma 1 to conclude that the optimal value in the left maximization problem in (2.39b) is reached at one of the vertices $\boldsymbol{v}_0, \ldots, \boldsymbol{v}_{k+1}$ found in r-side$(\tilde{\Gamma})$. Therefore, by introducing the points:

$$\boldsymbol{y}_i^\star \stackrel{\text{def}}{=} \big(\tilde{\gamma}_1[\boldsymbol{v}_i] + h(\boldsymbol{v}_i), \; \tilde{\gamma}_2[\boldsymbol{v}_i] \big), \forall\, i \in \{0, \ldots, k+1\}, \tag{2.40}$$

we can immediately conclude the following result:

Lemma 6. *The maximum in problem:*

$$(OPT) \qquad \max_{(\pi_1,\pi_2)\in\Pi^\star} \Big[\pi_1 + J^\star(\pi_2) \Big], \tag{2.41a}$$

$$\Pi^\star \stackrel{\text{def}}{=} \Big\{ (\pi_1^\star, \pi_2^\star) \in \mathbb{R}^2 : \pi_1^\star \stackrel{\text{def}}{=} \tilde{\gamma}_1 + h(\tilde{\gamma}_2), \quad \pi_2^\star \stackrel{\text{def}}{=} \tilde{\gamma}_2, \quad (\tilde{\gamma}_1, \tilde{\gamma}_2) \in \tilde{\Gamma} \Big\}, \tag{2.41b}$$

is reached on the right side of:

$$\Delta_{\Pi^\star} \stackrel{\text{def}}{=} \text{conv}\left(\{ \boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_{k+1}^\star \} \right). \tag{2.42}$$

Proof. The result is analogous to Lemma 2, and the proof is a rehashing of similar ideas. In particular, first note that problem (OPT) is a rewriting of the left maximization in (2.39b). Therefore, since the maximum of the latter problem is reached at the vertices $\boldsymbol{v}_i$, $i \in \{0, \ldots, k+1\}$, of zonogon $\tilde{\Gamma}$, by the definition (2.40) of the points $\boldsymbol{y}_i^\star$, we can conclude that the maximum in problem (OPT) must be reached on the set $\{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_{k+1}^\star\}$. Noting that the function maximized in (OPT) is convex, this set of points can be replaced with its convex hull, $\Delta_{\Pi^\star}$, without affecting the result. Furthermore, since $J^\star$ is convex, by applying the results in Corollary 1, and replacing the set by the right-side of its convex hull, r-side$(\Delta_{\Pi^\star})$, the optimal value remains unchanged. □

The significance of the new variables $\pi_{1,2}^\star$ is as follows. $\pi_1^\star$ represents the cumulative past stage costs, plus the true (*i.e.*, ideal) convex cost as stage $k+1$, while $\pi_2^\star$, just like $\tilde{\gamma}_2$, stands for the state at the next time-step, x_{k+2}.

Continuing the analogy with Section 2.4.2, the right optimization in (2.39b) can be rewritten as

$$
(AFF) \quad \begin{aligned} & \max_{(\pi_1,\pi_2)\in\Pi} \left[\pi_1 + J^\star(\pi_2) \right] \\ & \Pi \stackrel{\text{def}}{=} \left\{ (\pi_1, \pi_2) \ : \ \pi_1(\boldsymbol{w}) \stackrel{\text{def}}{=} \tilde{\gamma}_1(\boldsymbol{w}) + z(\boldsymbol{w}), \quad \pi_2(\boldsymbol{w}) \stackrel{\text{def}}{=} \tilde{\gamma}_2(\boldsymbol{w}), \ \boldsymbol{w} \in \mathcal{H}_{k+2} \right\}. \end{aligned} \tag{2.43}
$$

In order to examine the maximum in problem (AFF), we remark that its feasible set, $\Pi \subset \mathbb{R}^2$, also represents a zonogon, with generators given by $\boldsymbol{a}+\boldsymbol{z}$ and $\boldsymbol{b}$, respectively. Therefore, by Lemma 1, the maximum of problem (AFF) is reached at one of the vertices on r-side(Π).

Using the same key idea from the construction of the affine control law, we now argue that, if the coefficients of the affine cost, z_i, were computed in such a way that Π represented the *zonogon hull* of the set of points $\{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_{k+1}^\star\}$, then (by Corollary 1), the maximum value of problem (AFF) would be the same as the maximum value of problem (OPT).

To this end, we introduce the following procedure for computing the affine cost $z(\boldsymbol{w})$:

Algorithm 2 Compute affine stage cost $z(\boldsymbol{w})$

Require: $\tilde{\gamma}_1(\boldsymbol{w}), \tilde{\gamma}_2(\boldsymbol{w}), h(\cdot), J^\star(\cdot)$.

1: Apply the mapping (2.40) to obtain $\boldsymbol{v}_i^\star$, $\forall\, i \in \{0, \ldots, k+1\}$.

2: Compute the set $\Delta_{\Pi^\star} = \text{conv}\big(\{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_{k+1}^\star\}\big)$.

3: Let r-side$(\Delta_{\Pi^\star}) \stackrel{\text{def}}{=} \{\boldsymbol{y}_{s(1)}^\star, \ldots, \boldsymbol{y}_{s(n)}^\star\}$, where $s(1) \le s(2) \le \cdots \le s(n) \in \{0, \ldots, k+1\}$ are the sorted indices of points on the right side of $\Delta_{\Pi^\star}$.

4: Solve the following system for z_j, $(j \in \{0, \ldots, k+1\})$, and $K_{s(i)}$, $(i \in \{2, \ldots, n\})$:

$$
\begin{cases} z_0 + z_1 + \cdots + z_{s(i)} = h\left(\boldsymbol{v}_{s(i)}\right), & \forall\, \boldsymbol{y}_{s(i)}^\star \in \text{r-side}(\Delta_{\Pi^\star}) \quad \text{(matching)} \\ \dfrac{z_j + a_j}{b_j} = K_{s(i)}, & \forall\, j \in \{s(i-1)+1, \ldots, s(i)\}, \\ & \forall\, i \in \{2, \ldots, n\} \qquad \text{(alignment)} \end{cases} \tag{2.44}
$$

5: Return $z(\boldsymbol{w}) = z_0 + \sum_{i=1}^{k+1} z_i \cdot w_i$.

To visualize how the algorithm is working, an extended example is included in Figure 2-7.

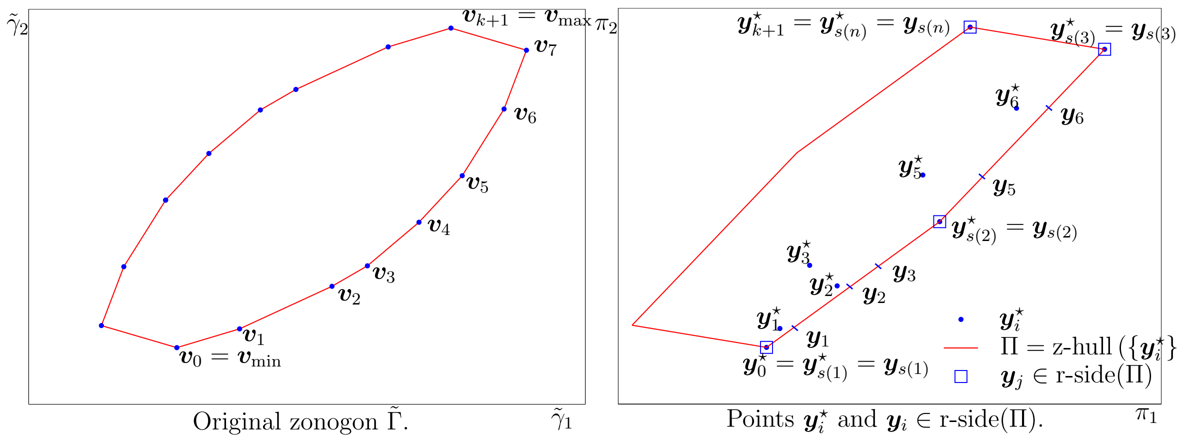


Figure 2-7: Matching and alignment performed in Algorithm 2.

The intuition behind the construction is the same as that presented in Section 2.4.2. In particular, the *matching* constraints in system (2.44) ensure that for any vertex $\boldsymbol{w}$ of the hypercube $\mathcal{H}_{k+2}$ that corresponds to a potential maximizer in problem (OPT) $\big($through $\boldsymbol{w} \in \mathcal{H}_{k+2} \overset{(2.37)}{\mapsto} \boldsymbol{v}_i \in \tilde{\Gamma} \overset{(2.40)}{\mapsto} \boldsymbol{y}^\star_i \in \text{r-side}(\Delta_{\Pi^\star})\big)$, the value of the affine cost $z(\boldsymbol{w})$ is equal to the value of the initial convex cost, $h(\boldsymbol{v}_i)$, implying that the value in problem (AFF) of (2.43) at $\big(\pi_1(\boldsymbol{w}), \pi_2(\boldsymbol{w})\big)$ is equal to the value in problem (OPT) of (2.41a) at $\boldsymbol{y}^\star_i$. The *alignment* constraints in system (2.44) ensure that any such *matched* points, $\big(\pi_1(\boldsymbol{w}), \pi_2(\boldsymbol{w})\big)$, actually correspond to the vertices on the right side of the zonogon Π, which implies that, as desired, $\Pi \equiv \text{z-hull}\big(\{\boldsymbol{y}^\star_0, \ldots, \boldsymbol{y}^\star_{k+1}\}\big)$.

We conclude our preliminary remarks by noting that, similar to the affine construction, system (2.44) does not directly impose the robust domination constraint (2.39a). However, as we will soon argue, this result is a byproduct of the way the matching and alignment are performed in Algorithm 2.

Affine Cost $z(\cdot)$ Dominates Convex Cost $h(\cdot)$ and Preserves Overall Objective.

In this section, we prove that the affine cost $z(\boldsymbol{w})$ computed in Algorithm 2 not only robustly dominates the original convex cost (2.39a), but also preserves the overall min-max value (2.39b). The following lemma summarizes the first main result:

Lemma 7. *System* (2.44) *is always feasible, and the solution* $z(\boldsymbol{w})$ *always satisfies equation* (2.39b)*.*

Proof. We first note that $s(1) = 0$ and $s(n) = k+1$, i.e., $\boldsymbol{y}_0^\star$, $\boldsymbol{y}_{k+1}^\star \in \text{r-side}(\Delta_{\Pi^\star})$. To see why that is the case, note that, by (2.38), $\boldsymbol{v}_0$ will always have the smallest $\tilde{\gamma}_2$ coordinate in the zonogon $\tilde{\Gamma}$. Since the transformation (2.40) yielding $\boldsymbol{y}_i^\star$ leaves the second coordinate unchanged, it is always true that

$$\boldsymbol{y}_0^\star = \arg\max\Big\{\pi_1 : \boldsymbol{\pi} \in \arg\min\big\{\pi_2' : \boldsymbol{\pi}' \in \{\boldsymbol{y}_i^\star,\, i \in \{0,\ldots,k+1\}\}\big\}\Big\},$$

which immediately implies that $\boldsymbol{y}_0^\star \in \text{r-side}(\Delta_{\Pi^\star})$. The proof for $\boldsymbol{y}_{k+1}^\star$ follows in an identical matter, since $\boldsymbol{v}_{k+1}$ has the largest $\tilde{\gamma}_2$ coordinate in $\tilde{\Gamma}$.

It can then be checked that the following choice of z_i always satisfies system (2.44):

$$\begin{aligned}
z_0 &= h(\boldsymbol{v}_0); \quad z_j = K_{s(i)} \cdot b_j - a_j, \quad \forall\, j \in \{s(i-1)+1,\ldots,s(i)\}, \quad \forall\, i \in \{2,\ldots,n\},\\
K_{s(i)} &= \frac{z_{s(i-1)+1} + \cdots + z_{s(i)} + a_{s(i-1)+1} + \cdots + a_{s(i)}}{b_{s(i-1)+1} + \cdots + b_{s(i)}}\\
&= \frac{h(v_{s(i)}) - h(v_{s(i-1)}) + a_{s(i-1)+1} + \cdots + a_{s(i)}}{b_{s(i-1)+1} + \cdots + b_{s(i)}}.
\end{aligned}$$

The proof of the second part of the lemma is analogous to that of Lemma 5. To start, consider the feasible set of problem (AFF) in (2.43), namely the zonogon Π, and note that, from (2.37), its generators are given by $\boldsymbol{a}+\boldsymbol{z}$ and $\boldsymbol{b}$,

$$\begin{bmatrix} \boldsymbol{a}+\boldsymbol{z} \\ \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} a_1+z_1 & \ldots & a_{s(i)}+z_{s(i)} & a_{s(i)+1}+z_{s(i)+1} & \ldots & a_{k+1}+z_{k+1} \\ b_1 & \ldots & b_{s(1)} & b_{s(1)+1} & \ldots & b_{k+1} \end{bmatrix}. \tag{2.45}$$

By introducing the following points in $\mathbb{R}^2$,

$$\boldsymbol{y}_i = \left(\sum_{j=0}^{i} (a_j + z_j),\ \sum_{j=0}^{i} b_j \right),$$

we have the following simple claims:

- For any $\boldsymbol{v}_i \in$ r-side$(\tilde{\Gamma})$ that is *matched*, i.e., $\boldsymbol{y}_i^\star \in$ r-side$(\Delta_{\Pi^\star})$, with $\boldsymbol{w}_i = [1, 1, \ldots, 1, 0, \ldots, 0]$ denoting the unique[6] vertex of $\mathcal{H}_{k+2}$ satisfying $\big(\tilde{\gamma}_1(\boldsymbol{w}_i), \tilde{\gamma}_2(\boldsymbol{w}_i)\big) = \boldsymbol{v}_i$, we have

$$\boldsymbol{y}_i \overset{(2.43)}{=} \big(\tilde{\gamma}_1(\boldsymbol{w}_i) + z(\boldsymbol{w}_i),\, \tilde{\gamma}_2(\boldsymbol{w}_i)\big) \overset{(2.44)}{=} \big(\tilde{\gamma}_1[\boldsymbol{v}_i] + h(\boldsymbol{v}_i),\, \tilde{\gamma}_2[\boldsymbol{v}_i]\big) \overset{(2.40)}{=} \boldsymbol{y}_i^\star.$$

  The first equality follows from the definition of the zonogon Π, the second follows because any $\boldsymbol{y}_i^\star \in$ r-side$(\Delta_{\Pi^\star})$ is *matched* in system (2.44), and the third equality represents the definition of the points $\boldsymbol{y}_i^\star$.

- For any vertex $\boldsymbol{v}_j \in$ r-side$(\tilde{\Gamma})$, which is *not* matched, i.e., $\boldsymbol{y}_j^\star \notin$ r-side$(\Delta_{\Pi^\star})$, and $s(i) < j < s(i{+}1)$ for some i, we have $\boldsymbol{y}_j \in [\boldsymbol{y}_{s(i)}, \boldsymbol{y}_{s(i+1)}]$. This can be seen by using the *alignment* conditions in system (2.44), in conjunction with (2.45), since the segments in $\mathbb{R}^2$ given by $\big[\boldsymbol{y}_{s(i)}, \boldsymbol{y}_{s(i)+1}\big], \big[\boldsymbol{y}_{s(i)+1}, \boldsymbol{y}_{s(i)+2}\big], \ldots, \big[\boldsymbol{y}_{s(i+1)-1}, \boldsymbol{y}_{s(i+1)}\big]$ are always parallel, with common cotangent given by $K_{s(i+1)}$.

For a geometric interpretation, the reader is referred back to Figure 2-7. Corroborating these results with the fact that $\big\{\boldsymbol{y}_{s(1)}^\star, \ldots, \boldsymbol{y}_{s(n)}^\star\big\} =$ r-side$(\Delta_{\Pi^\star})$ always satisfy:

$$\text{cotan}\big(\boldsymbol{y}_{s(1)}^\star,\, \boldsymbol{y}_{s(2)}^\star\big) \geq \text{cotan}\big(\boldsymbol{y}_{s(2)}^\star,\, \boldsymbol{y}_{s(3)}^\star\big) \geq \cdots \geq \text{cotan}\big(\boldsymbol{y}_{s(n-1)}^\star,\, \boldsymbol{y}_{s(n)}^\star\big), \qquad (2.46)$$

we immediately obtain that the points $\big\{\boldsymbol{y}_{s(1)}, \boldsymbol{y}_{s(2)}, \ldots, \boldsymbol{y}_{s(n)}\big\}$ exactly represent the right side of the zonogon Π, which, in turn, implies that $\Pi \equiv$ z-hull $\big(\big\{\boldsymbol{y}_0^\star, \boldsymbol{y}_1^\star, \ldots, \boldsymbol{y}_{k+1}^\star\big\}\big)$. But then, by Corollary 1, the maximum value of problem (OPT) in (2.41a) is equal to the maximum value of problem (AFF) in (2.43), and, since the former is always $J_1^\star(x_1)$, so is that latter. □

In order to complete the second step of the induction, we must only show that

[6] We are working under Assumption 2, which implies uniqueness of the vertex.

the robust domination constraint (2.39a) is also obeyed:

$$z(\boldsymbol{w}) \geq h\big(\tilde{\gamma}_2(\boldsymbol{w})\big) \;\; \Leftrightarrow$$
$$z_0 + z_1 \cdot w_1 + \cdots + z_{k+1} \cdot w_{k+1} \geq h\left(b_0 + b_1 \cdot w_1 + \cdots + b_{k+1} \cdot w_{k+1}\right), \;\; \forall\, \boldsymbol{w} \in \mathcal{H}_{k+1}.$$

The following lemma takes us very close to the desired result:

Lemma 8. *The coefficients for the affine cost $z(\boldsymbol{w})$ computed in Algorithm 2 always satisfy the following property:*

$$h\big(b_0 + b_{j(1)} + \cdots + b_{j(m)}\big) \leq z_0 + z_{j(1)} + \cdots + z_{j(m)},$$
$$\forall\, j(1), \ldots, j(m) \in \{1, \ldots, k+1\}, \forall\, m \in \{1, \ldots, k+1\}.$$

Proof. Before proceeding with the proof, we first list several properties related to the construction of the affine cost. We claim that, upon termination, Algorithm 2 always produces a solution to the following system:

$$\begin{cases} z_0 & = h\left(\boldsymbol{v}_{s(1)}\right) \\ z_0 + z_1 + \cdots + z_{s(2)} & = h\left(\boldsymbol{v}_{s(2)}\right) \\ \quad\vdots & \quad\vdots \\ z_0 + z_1 + \cdots + z_{s(n)} & = h\left(\boldsymbol{v}_{s(n)}\right) \\ \frac{z_1 + a_1}{b_1} & = \cdots = \frac{z_{s(2)} + a_{s(2)}}{b_{s(2)}} = K_{s(2)} \\ \quad\vdots & \quad\quad\vdots \\ \frac{z_{s(n-1)+1} + a_{s(n-1)+1}}{b_{s(n-1)+1}} & = \cdots = \frac{z_{s(n)} + a_{s(n)}}{b_{s(n)}} = K_{s(n)} \end{cases} \tag{2.47}$$

$$K_{s(2)} \geq \cdots \geq K_{s(n)} \tag{2.48}$$

$$
\begin{cases}
\frac{h(\boldsymbol{v}_j)-h(\boldsymbol{v}_0)+a_1+\cdots+a_j}{b_1+\cdots+b_j} \le K_{s(2)} \le \frac{h\left(\boldsymbol{v}_{s(2)}\right)-h(\boldsymbol{v}_j)+a_{j+1}+\cdots+a_{s(1)}}{b_{j+1}+\cdots+b_{s(1)}}, \\
\qquad \forall\, j \in \{1,\ldots,s(2)-1\} \\
\qquad \vdots \qquad \vdots \\
\frac{h(\boldsymbol{v}_j)-h\left(\boldsymbol{v}_{s(n-1)}\right)+a_{s(n-1)+1}+\cdots+a_j}{b_{s(n-1)+1}+\cdots+b_j} \le K_{s(n)} \le \frac{h\left(\boldsymbol{v}_{s(n)}\right)-h(\boldsymbol{v}_j)+a_{j+1}+\cdots+a_{s(n)}}{b_{j+1}+\cdots+b_{s(n)}}, \\
\qquad \forall\, j \in \{s(n-1)+1,\ldots,s(n)-1\}\ .
\end{cases}
\tag{2.49}
$$

Let us explain the significance of all the equations. (2.47) is simply a rewriting of the original system (2.44), which states that at any vertex $\boldsymbol{v}_{s(i)}$, the value of the affine function should exactly match the value assigned by the convex function $h(\cdot)$, and the coefficients z_i between any two matched vertices should be such that the resulting segments, $[z_j + a_j, b_j]$, are aligned (i.e., the angles they form with the π_1 axis have the same cotangent, specified by $K_{(\cdot)}$ variables). We note that we have explicitly used the fact that $s(1) = 0$, which we have shown in the first paragraph of the proof of Lemma 7.

Equation (2.48) is a simple restatement of (2.46), that the cotangents on the right side of a convex hull must be decreasing.

Equation (2.49) is a direct consequence of the fact that $\{\boldsymbol{y}^\star_{s(1)}, \boldsymbol{y}^\star_{s(2)}, \ldots, \boldsymbol{y}^\star_{s(n)}\}$ represent r-side$(\Delta_{\Pi^\star})$. To see why that is, consider an arbitrary $j \in \{s(i)+1,\ldots,s(i+1)-1\}$. Since $\boldsymbol{y}^\star_j \notin$ r-side$(\Delta_{\Pi^\star})$, we have:

$$
\begin{aligned}
&\operatorname{cotan}\left(\boldsymbol{y}^\star_{s(i)},\, \boldsymbol{y}^\star_j\right) \le \operatorname{cotan}\left(\boldsymbol{y}^\star_j,\, \boldsymbol{y}^\star_{s(i+1)}\right) \overset{(2.37),(2.40)}{\Leftrightarrow} \\
&\frac{a_{s(i)+1}+\cdots+a_j+h\left(\boldsymbol{v}_j\right)-h\left(\boldsymbol{v}_{s(i)}\right)}{b_{s(i)+1}+\cdots+b_j} \le \frac{a_{j+1}+\cdots+a_{s(i+1)}+h\left(\boldsymbol{v}_{s(i+1)}\right)-h\left(\boldsymbol{v}_j\right)}{b_{j+1}+\cdots+b_{s(i+1)}} \Leftrightarrow \\
&\frac{a_{s(i)+1}+\cdots+a_j+h\left(\boldsymbol{v}_j\right)-h\left(\boldsymbol{v}_{s(i)}\right)}{b_{s(i)+1}+\cdots+b_j} \le K_{s(i+1)} \\
&\qquad \le \frac{a_{j+1}+\cdots+a_{s(i+1)}+h\left(\boldsymbol{v}_{s(i+1)}\right)-h\left(\boldsymbol{v}_j\right)}{b_{j+1}+\cdots+b_{s(i+1)}},
\end{aligned}
$$

where, in the last step, we have used the mediant inequality[7] and the fact that, from (2.47), $K_{s(i+1)} = \operatorname{cotan}\left(\boldsymbol{y}^\star_{s(i)},\, \boldsymbol{y}^\star_{s(i+1)}\right) = \frac{a_{s(i)+1}+\cdots+a_{s(i+1)}+h\left(\boldsymbol{v}_{s(i+1)}\right)-h\left(\boldsymbol{v}_{s(i)}\right)}{b_{s(i)+1}+\cdots+b_{s(i+1)}}$ (refer

[7]If $b, d > 0$ and $\frac{a}{b} \le \frac{c}{d}$, then $\frac{a}{b} \le \frac{a+c}{b+d} \le \frac{c}{d}$.

back to Figure 2-7 for a geometrical interpretation).

With these observations, we now prove the claim of the lemma. The strategy of the proof will be to use induction on the size of the subsets, m. First, we show the property for any subset of indices $j(1), \dots, j(m) \in \{s(1) = 0, \dots, s(2)\}$, and then extend it to $j(1), \dots, j(m) \in \{s(i)+1, \dots, s(i+1)\}$ for any i, and then to any subset of $\{1, \dots, k+1\}$.

The following implications of the conditions (2.47), (2.48) and (2.49), are stated here for convenience, since they are used throughout the rest of the proof:

$$h\left(\boldsymbol{v}_{s(1)}\right) = h(\boldsymbol{v}_0) = z_0; \quad h(\boldsymbol{v}_{s(2)}) = z_0 + z_1 + \dots + z_{s(2)}. \tag{2.50}$$

$$h(\boldsymbol{v}_j) - h(\boldsymbol{v}_0) \le z_1 + \dots + z_j, \qquad \forall\, j \in \{1, \dots, s(2)-1\}. \tag{2.51}$$

$$\frac{z_1}{b_1} \le \dots \le \frac{z_j}{b_j} \le \dots \le \frac{z_{s(2)}}{b_{s(2)}}, \qquad \forall\, j \in \{1, \dots, s(2)-1\}. \tag{2.52}$$

Their proofs are straightforward. (2.50) follows directly from system (2.47), and:

$$\frac{h(\boldsymbol{v}_j) - h(\boldsymbol{v}_0) + a_1 + \dots + a_j}{b_1 + \dots + b_j} \overset{(2.49)}{\le} K_{s(2)} \overset{(2.47)}{=} \frac{z_1 + \dots + z_j + a_1 + \dots + a_j}{b_1 + \dots + b_j} \Rightarrow (2.51).$$

$$\begin{cases} (2.47): \ \frac{a_1+z_1}{b_1} = \dots = \frac{a_j+z_j}{b_j} = \dots = \frac{a_{s(2)}+z_{s(2)}}{b_{s(2)}} \\ \Pi \text{ zonogon } \Rightarrow \frac{a_1}{b_1} > \dots > \frac{a_j}{b_j} > \dots > \frac{a_{s(2)}}{b_{s(2)}} \end{cases} \Rightarrow \ (2.52).$$

We can now proceed with the proof, by checking the induction for $m = 1$. We would like to show that $h\left(b_0 + b_j\right) \le z_0 + z_j,\ \forall\, j \in \{1, \dots, s(2)\}$. Writing $b_0 + b_j$ as $b_0 + b_j = (1-\lambda) \cdot b_0 + \lambda \cdot (b_0 + \dots + b_j),$ with $\lambda = b_j/(b_1 + \dots + b_j)$, we obtain:

$$\begin{aligned} h(b_0 + b_j) &\le (1-\lambda) \cdot h(b_0) + \lambda \cdot \underbrace{h(b_0 + \dots + b_j)}_{\equiv h(\boldsymbol{v}_j)} \\ &= h(\boldsymbol{v}_0) + \frac{b_j}{b_1 + \dots + b_j} \left[\, h(\boldsymbol{v}_j) - h(\boldsymbol{v}_0) \,\right] \le \ \text{(by (2.50) or (2.51))} \\ &\le z_0 + \frac{b_j}{b_1 + \dots + b_j} \left(z_1 + \dots + z_j\right) \ \le \ \text{(by (2.52) and mediant inequality)} \\ &\le z_0 + z_j. \end{aligned}$$

Assume the property is true for any subsets of size m. Consider a subset $j(1), \ldots,$ $j(m)$, $j(m+1)$, and, without loss of generality, let $j(m+1)$ be the largest index. With the convex combination:

$$\begin{aligned} b^{\star} &\overset{\text{def}}{=} b_0 + b_{j(1)} + \cdots + b_{j(m)} + b_{j(m+1)} \\ &= (1-\lambda)\cdot(b_0 + b_{j(1)} + \cdots + b_{j(m)}) + \lambda\cdot(b_0 + b_1 + \cdots + b_{j(m+1)-1} + b_{j(m+1)}), \end{aligned}$$

where $\lambda = \dfrac{b_{j(m+1)}}{(b_1 + b_2 + \cdots + b_{j(m+1)}) - (b_{j(1)} + b_{j(2)} + \cdots + b_{j(m)})}$,

we obtain:

$$\begin{aligned} h(b^{\star}) &\leq (1-\lambda)\cdot h(b_0 + b_{j(1)} + \cdots + b_{j(m)}) + \lambda \cdot h\left(\boldsymbol{v}_{i(m+1)}\right) \\ &\leq \big(\text{by induction hypothesis and (2.50), (2.51)}\big) \\ &\leq (1-\lambda)\cdot(z_0 + z_{j(1)} + \cdots + z_{j(m)}) + \lambda\cdot\big(z_0 + z_1 + \cdots + z_{i(m+1)}\big) \\ &= z_0 + z_{j(1)} + \cdots + z_{j(m)} + \frac{b_{j(m+1)}}{(b_1 + b_2 + \cdots + b_{j(m+1)}) - (b_{j(1)} + b_{j(2)} + \cdots + b_{j(m)})}\cdot \\ &\qquad \cdot\big[(z_1 + z_2 + \cdots + z_{j(m+1)}) - (z_{j(1)} + z_{j(2)} + \cdots + z_{j(m)})\big] \\ &\leq \big(\text{by (2.52) and the mediant inequality}\big) \\ &\leq z_0 + z_{j(1)} + \cdots + z_{j(m)} + z_{j(m+1)}. \end{aligned}$$

We claim that the exact same procedure can be repeated for a subset of indices from $\{s(i)+1, \ldots, s(i+1)\}$, for any index $i \in \{1, \ldots, n-1\}$. We would simply be using the adequate inequality from (2.49), and the statements equivalent to (2.50), (2.51) and (2.52). The following results are immediate:

$$\begin{aligned} &h\left((b_0 + b_1 + \cdots + b_{s(i)}) + b_{j(1)} + \cdots + b_{j(m)}\right) \\ &\qquad \leq \left(z_0 + z_1 + \cdots + z_{s(i)}\right) + z_{j(1)} + \cdots + z_{j(m)}, \forall\, i \in \{1, \ldots, n\}, \\ &\qquad\qquad \forall\, j(1), \ldots, j(m) \in \{s(i)+1, \ldots, s(i+1)\}. \end{aligned} \tag{2.53}$$

Note that instead of the term b_0 for the argument of $h(\cdot)$, we would use the complete sum $b_0 + b_1 + \cdots + b_{s(i)}$, and, similarly, instead of z_0 we would have the complete sum

$z_0 + z_1 + \cdots + z_{s(i)}$. With these results, we can make use of the increasing increments property of convex functions,

$$\frac{h(x_1+\Delta)-h(x_1)}{\Delta} \le \frac{h(x_2+\Delta)-h(x_2)}{\Delta}, \quad \forall\, \Delta > 0, x_1 \le x_2 \;,$$

to obtain the following result:

$$\begin{aligned}
&h\Big(b_0 + \underbrace{b_{j(1)} + \cdots + b_{j(m)}}_{j(\cdot)\in\{1,\dots,s(2)\}} + \underbrace{b_{i(1)} + \cdots + b_{i(l)}}_{i(\cdot)\in\{s(2)+1,\dots,s(3)\}}\Big) - h\left(b_0 + b_{j(1)} + \cdots + b_{j(m)}\right) \;\le \\
&\quad \le h\Big(b_0 + \underbrace{b_1 + \cdots + b_{s(2)}}_{\text{all indices in } \{1,\dots,s(2)\}} + b_{i(1)} + \cdots + b_{i(l)}\Big) - \underbrace{h\left(b_0 + b_1 + \cdots + b_{s(2)}\right)}_{\overset{\text{def}}{=} h(\boldsymbol{v}_{s(2)})} \\
&\quad \overset{(2.50),(2.53)}{\le} \big(z_0 + z_1 + \cdots + z_{s(2)}\big) + z_{i(1)} + \cdots + z_{i(l)} - \big(z_0 + z_1 + \cdots + z_{s(2)}\big) \\
&\quad = z_{i(1)} + \cdots + z_{i(l)},
\end{aligned}$$

which would imply

$$\begin{aligned}
h\big(b_0 + b_{j(1)} + \cdots + b_{j(m)} + b_{i(1)} + \cdots + b_{i(l)}\big) \\
&\le h\left(b_0 + b_{j(1)} + \cdots + b_{j(m)}\right) + z_{i(1)} + \cdots + z_{i(l)} \\
&\overset{(2.53)}{\le} z_0 + z_{j(1)} + \cdots + z_{j(m)} + z_{i(1)} + \cdots + z_{i(l)}.
\end{aligned}$$

We showed the property for indices drawn only from the first two intervals, $\{s(1)+1,\dots,s(2)\}$ and $\{s(2)+1,\dots,s(3)\}$, but it should be clear how the argument can be immediately extended to any collection of indices, drawn from any intervals. We omit the details for brevity, and conclude that the claim of the lemma is true. $\square$

We are now ready for the last major result:

Lemma 9. *The affine cost $z(\boldsymbol{w})$ computed by Algorithm 2 always dominates the convex cost $h\big(\tilde{\gamma}_2(\boldsymbol{w})\big)$:*

$$h\left(b_0 + \sum_{i=1}^{k+1} b_i \cdot w_i\right) \le z_0 + \sum_{i=1}^{k+1} z_i \cdot w_i, \quad \forall\, \boldsymbol{w} \in \mathcal{H}_{k+1} = [0,1]^{k+1}.$$

Proof. Note first that the function $f(\boldsymbol{w}) \stackrel{\text{def}}{=} h\left(b_0 + \sum_{i=1}^{k+1} b_i \cdot w_i\right) - (z_0 + \sum_{i=1}^{k+1} z_i \cdot w_i)$ is a convex function of $\boldsymbol{w}$. Furthermore, the result of Lemma 8 can be immediately rewritten as:

$$h\left(b_0 + \sum_{i=1}^{k+1} b_i \cdot w_i\right) \leq z_0 + \sum_{i=1}^{k+1} z_i \cdot w_i, \, \forall\, \boldsymbol{w} \in \{0,1\}^{k+1} \;\Leftrightarrow\; f(\boldsymbol{w}) \leq 0, \, \forall\, \boldsymbol{w} \in \{0,1\}^{k+1}.$$

Since the maximum of a convex function on a polytope occurs on the extreme points of the polytope, and $\text{ext}(\mathcal{H}_{k+1}) = \{0,1\}^{k+1}$, we immediately have that: $\max_{\boldsymbol{w}\in\mathcal{H}_{k+1}} f(\boldsymbol{w}) = \max_{\boldsymbol{w}\in\{0,1\}^{k+1}} f(\boldsymbol{w}) \leq 0$, which completes the proof of the lemma. $\square$

We can now conclude the proof of correctness in the construction of the affine stage cost, $z(\boldsymbol{w})$. With Lemma 9, we have that the affine cost always dominates the convex cost $h(\cdot)$, thus condition (2.39a) is obeyed. Furthermore, from Lemma 7, the overall min-max cost remains unchanged even when incurring the affine stage cost, $z(\boldsymbol{w})$, hence condition (2.39b) is also true. This completes the construction of the affine cost, and hence also the full step of the induction hypothesis.

2.4.4 Proof of Main Theorem.

To finalize the current section, we summarize the steps that have lead us to the result, thereby proving the main Theorem 1.

Theorem 1. In Section 2.4.1, we have verified the induction hypothesis at time $k = 1$. With the induction hypothesis assumed true for times $t = 1, \ldots, k$, we have listed the initial consequences in Lemma 1 and Corollary 1 of Section 2.4.1. By exploring the structure of the optimal control law, $u^\star_{k+1}(x_{k+1})$, and the optimal value function, $J^\star_{k+1}(x_{k+1})$, in Section 2.4.1, we have finalized the analysis of the induction hypothesis, and summarized our findings in Lemmas 2 and 3.

Section 2.4.2 then introduced the main construction of the affine control law, $q_{k+1}(\boldsymbol{w}_{[k+1]})$, which was shown to be robustly feasible (Lemma 4). Furthermore, in Lemma 5, we have shown that, when used in conjuction with the original convex state costs, $h_{k+1}(x_{k+2})$, this affine control preserves the min-max value of the overall problem.

In Section 2.4.3, we have also introduced an affine stage cost, $z_{k+1}(\boldsymbol{w}_{[k+1]})$, which, if incurred at time $k+1$, will always preserve the overall min-max value (Lemma 7), despite being always larger than the original convex cost, $h_{k+1}(x_{k+2})$ (Lemma 9). $\square$

2.4.5 Counterexamples for potential extensions.

On first sight, one might be tempted to believe that the results in Theorem 1 could be immediately extended to more general problems. In particular, one could be tempted to ask one of the following natural questions:

1. Would both results of Theorem 1 (i.e., existence of affine control laws *and* existence of affine stage costs) hold for a problem which also included linear constraints coupling the controls u_t across different time-steps? (see Ben-Tal et al. [16] for a situation when this might be of interest)

2. Would both results of Theorem 1 hold for multi-dimensional linear systems? (i.e., problems where $x_k \in \mathbb{R}^d$, $\forall\, k$, with $d \geq 2$)

3. Are affine policies in the disturbances optimal for the two problems above?

4. Are affine policies also optimal for stochastic versions of this problem, e.g., for the case where w_k is uniformly distributed in $\mathcal{W}_k = [\underline{w}_k, \overline{w}_k]$, and the goal is to minimize expected costs?

In the rest of the current section, we argue that all of the above questions can be answered negatively. To address the first three, we use the following simple counterexample:

Example 2 (Suboptimality of affine policies and affine cost relaxations)**.**

$$T = 4, c_k = 1,\ h_k(x_{k+1}) = \max\{18.5 \cdot x_{k+1},\ -24 \cdot x_{k+1}\},\ L_k = 0, U_k = \infty,\ 1 \le k \le 4,$$
$$w_1 \in [-7, 0], w_2 \in [-11, 0], w_3 \in [-8, 0], w_4 \in [-44, 0]\ ,$$
$$\sum_{i=1}^{k} u_i \le 10 \cdot k\ , \forall\, k \in \{1, \ldots, 4\}.$$

The first two rows describe a one-dimensional problem that fits the conditions of Problem 1 in Section 2.1. The third row corresponds to a coupling constraint for controls at different times, so that the problem fits question (i) above. Furthermore, since the state in such a problem consists of two variables (one for x_k and one for $\sum_{i=1}^{k} u_k$), the example also fits question (ii) above.

The optimal min-max value for Example 2 above can be found by solving an optimization problem (see Ben-Tal et al. [16]), in which non-anticipatory decisions are computed at all the extreme points of the uncertainty set, i.e., for $\{\underline{w}_1, \overline{w}_1\} \times \{\underline{w}_2, \overline{w}_2\} \times \{\underline{w}_3, \overline{w}_3\} \times \{\underline{w}_4, \overline{w}_4\}$. The resulting model, which is a large linear program, can be solved to optimality, resulting in a corresponding value of approximately 838.493 for Example 2.

To compute the optimal min-max objective obtained by using affine policies $q_k(\boldsymbol{w}_{[k]})$ and incurring affine costs $z_k(\boldsymbol{w}_{[k+1]})$, one can amend the model $(AARC)$ from Section 2.3 by including constraints for the cumulative controls (see Ben-Tal et al. [16] for details), and then using (2.13) to rewrite the resulting model as a linear program. The optimal value of this program for Example 2 was approximately 876.057, resulting in a gap of 4.4%, and thus providing a negative answer to questions (i) and (ii).

To investigate question (iii), we remark that the smallest objective achievable by using affine policies of the type $q_k(\boldsymbol{w}_{[k]})$ can be found by solving another linear optimization problem, having as decision variables the affine coefficients $\{q_{k,t}\}_{0 \le t < k \le T}$, as well as (non-anticipatory) stage cost variables $z_k^{\boldsymbol{w}}$ for every time step $k \in \{1, \ldots, T\}$ and every extreme point $\boldsymbol{w}$ of the uncertainty set. Solving the resulting linear program for Example 2 gave an optimal value of 873.248, so strictly larger than the (true) optimum (838.493), and strictly smaller than the optimal value of the model utilizing both affine control policies *and* affine stage costs (876.057).

Thus, with question (iii) also answered negatively, we conclude that policies that are affine in the disturbances, $q_k(\boldsymbol{w}_{[k]})$, are in general *suboptimal* for problems with cumulative control constraints or multiple dimensions, and that replacing the convex state costs $h_k(x_{k+1})$ by (larger) affine costs $z_k(\boldsymbol{w}_{[k+1]})$ would, in general, result in even *further* deterioration of the objective.

As for question (iv), the following simple example suggests that affine rules are, in general, suboptimal, and that the gap can be arbitrarily large:

Example 3 (Suboptimality of affine policies in stochastic problems)**.**

$$\begin{aligned} J = & \mathbb{E}_{w_1}\left[\min_{u_2(w_1)} (u_2 - w_1)^2\right] \\ \text{s.t. } & 0 \le u_2 \le \frac{1}{K}, \\ & w_1 \sim \text{Uniform}[0,1], \\ & K \in (1,3), \text{ fixed and known.} \end{aligned}$$

From the convexity of the problem, it is easy to see that the optimal policy is

$$u_2^\star(w_1) = \begin{cases} w_1, & \text{if } 0 \le w_1 \le \frac{1}{K} \\ \frac{1}{K}, & \text{otherwise,} \end{cases}$$

which results in an objective $J^\star = \frac{(K-1)^3}{3K^3}$. It can also be easily shown that the optimal objective achievable under affine rules (that satisfy the constraint almost surely) is $J^{\text{AFF}} = \frac{(K-1)^2}{4K^2}$, for $u_2^{\text{AFF}}(w_1) = \frac{3-K}{2K}\, w_1 + \frac{K-1}{2K}$. In particular, note that the relative optimality gap, $\frac{J^{\text{AFF}} - J^\star}{J^\star} = \frac{4-K}{4(K-1)}$, can be made arbitrarily large, by taking $K \searrow 1$.

2.5 An application in inventory management.

In this section, we would like to explore our results in connection with the classical inventory problem mentioned in Example 1. This example was originally considered by Ben-Tal et al. [16], in the context of a more general model: a retailer-supplier with flexible commitment contracts problem. We first describe the problem in detail, and then draw a connection with our results.

The setting is the following: consider a single-product, single-echelon, multi-period supply chain, in which inventories are managed periodically over a planning horizon of T periods. The unknown demands w_t from customers arrive at the (unique) echelon, henceforth referred to as the *retailer*, and are satisfied from the on-hand inventory, denoted by x_t at the beginning of period t. The retailer can replenish the inventory by placing orders u_t, at the beginning of each period t, for a cost of c_t per unit of product. These orders are immediately available, i.e., there is no lead-time in the system, but there are capacities on how much the retailer can order: $L_t \leq u_t \leq U_t$. After the demand w_t is realized, the retailer incurs holding costs $H_t \cdot \max\{0, x_t + u_t - w_t\}$ for all the amounts of supply stored on her premises, as well as penalties $B_t \cdot \max\{w_t - x_t - u_t, 0\}$, for any demand that is backlogged.

In the spirit of robust optimization, we assume that the only information available about the demand at time t is that it resides within a certain inverval centered around a *nominal* (or mean) demand $\bar{d}_t$, which results in the uncertainty set $\mathcal{W}_t = \{\, |w_t - \bar{d}_t| \leq \rho \cdot \bar{d}_t \,\}$, where $\rho \in [0, 1]$ can be interpreted as an *uncertainty level.* As such, if we take the objective function to be minimized as the cost resulting in the worst-case scenario, we immediately obtain an instance of our original Problem 1, with $\alpha_t = \beta_t = 1, \gamma_t = -1$, and the convex state costs $h_t(\cdot)$ denoting the Newsvendor costs, $h_t(x_{t+1}) = H_t \cdot \max\{x_t + u_t - w_t, 0\} + B_t \cdot \max\{w_t - x_t - u_t, 0\}$.

Therefore, the results in Theorem 1 are immediately applicable to conclude that no loss of optimality is incurred when we restrict attention to affine order quantities q_t that depend on the history of available demands at time t, $q_t(\boldsymbol{w}_{[t]}) = q_{t,0} + \sum_{\tau=1}^{t-1} q_{t,\tau} \cdot w_\tau$, and when we replace the Newsvendor costs $h_t(x_{t+1})$ by some (potentially larger) affine costs $z_t(\boldsymbol{w}_{[t+1]})$. The main advantage is that, with these substitutions, the problem of finding the optimal affine policies becomes an LP (see the discussion in Section 2.3 and Ben-Tal et al. [16] for more details).

The more interesting connection with our results comes if we recall the construction in Algorithm 1. In particular, we have the following simple claim:

Proposition 1. *If the affine orders $q_t(\boldsymbol{w}_{[t]})$ computed in Algorithm 1 are implemented at every time step t, and we let: $x_k(\boldsymbol{w}_{[k]}) = x_1 + \sum_{t=1}^{k-1} \left(q_t(\boldsymbol{w}_{[t]}) - w_t \right) \stackrel{\text{def}}{=} x_{t,0} + \sum_{t=1}^{k-1} x_{k,t} \cdot w_t$ denote the affine dependency of the inventory x_k on the history of demands, $\boldsymbol{w}_{[k]}$, then:*

1. *If a certain demand w_t is fully satisfied by time $k \geq t+1$, i.e., $x_{k,t} = 0$, then all the (affine) orders q_τ placed after time k will not depend on w_t.*

2. *Every demand w_t is at most satisfied by the future orders q_k, $k \geq t+1$, and the coefficient $q_{k,t}$ represents what fraction of the demand w_t is satisfied by the order q_k.*

Proof. To prove the first claim, recall that, in our notation from Section 2.4.1, $x_k \equiv \theta_2 = b_0 + \sum_{t=1}^{k-1} b_t \cdot w_t$. Applying part (i) of Lemma 4 in the current setting[8], we have that $0 \leq q_{k,t} \leq -x_{k,t}$. Therefore, if $x_{k,t} = 0$, then $q_{k,t} = 0$, which implies that $x_{k+1,t} = 0$. By induction, we immediately get that $q_{\tau,t} = 0, \forall \tau \in \{k, \ldots, T\}$.

To prove the second part, note that any given demand, w_t, initially has an affine coefficient of -1 in the state x_{t+1}, i.e., $x_{t+1,t} = -1$. By part (i) of Lemma 4, $0 \leq q_{t+1,t} \leq -x_{t+1,t} = 1$, so that $q_{t+1,t}$ represents a fraction of the demand w_t satisfied by the order q_{t+1}. Furthermore, $x_{t+2,t} = x_{t+1,t} + q_{t+1,t} \in [-1, 0]$, so, by induction, we immediately have that $q_{k,t} \in [0,1], \forall k \geq t+1$, and $\sum_{k=t+1}^{T} q_{k,t} \leq 1$. □

In view of this result, if we think of $\{q_k\}_{k \geq t+1}$ as future orders that are partially satisfying the demand w_t, then every future order quantity $q_k(\boldsymbol{w}_{[k]})$ satisfies exactly a fraction of the demand w_t (since the coefficient for w_t in q_k is always in $[0,1]$), and

[8] The signs of the inequalities are changed because every disturbance, w_t, is entering the system dynamics with a coefficient -1, instead of $+1$, as was the case in the discussion from Section 2.4.1.

every demand is at most satisfied by the sequence of orders following after it appears. This interpretation bears some similarity with the unit decomposition approach of Muharremoglu and Tsitsiklis [105], where every unit of supply can be interpreted as satisfying a particular unit of the demand. Here, we are accounting for fractions of the total demand, as being satisfied by future order quantities.

2.5.1 Capacity Commitment and Negotiation.

Our theoretical result can also be employed in solving an interesting capacity commitment problem. In particular, we introduce the following modification of our original problem:

Problem 2. *Consider an identical setup as Problem 1, i.e., a dynamical system described by* (2.1)*, with scalar uncertainties given by* (2.2) *and control constraints described by* (2.3)*, but assume that the bounds on the controls,* L_k, U_k*, are not fixed, but part of the decision process. In particular,* $\boldsymbol{L} \stackrel{\text{def}}{=} (L_1, \ldots, L_T) \in \mathbb{R}^T$ *and* $\boldsymbol{U} \stackrel{\text{def}}{=} (U_1, \ldots, U_T) \in \mathbb{R}^T$ *must be decided at time* $k = 1$*, before observing any disturbances.*

The goal is to find a sequence of constrained controllers $\{u_k\}_{1 \le k \le T}$*, minimizing the following cost function over a finite horizon* $1, \ldots, T$*,*

$$\tilde{J} = J + \mathcal{F}(\boldsymbol{U}) - \mathcal{R}(\boldsymbol{L}), \tag{2.54}$$

where J *is the original cost given in* (2.4)*, and* $\mathcal{F} : \mathbb{R}^T \to \bar{\mathbb{R}}$ *is an extended-real, convex function, while* $\mathcal{R} : \mathbb{R}^T \to \mathbb{R}$ *is a concave function.*

An example of such a problem, which arises naturally in the context of the inventory example discussed earlier, is in negotiating supply contracts. In particular, since U_k represents an upper bound on the replenishment order quantity u_k that can be obtained in every period, the function $\mathcal{F}$ can be interpreted as a *cost of flexibility*, which the retailer must pay the supplier (at the beginning of the horizon) for having additional capacity available. Similarly, since L_k are commitments to ordering specific amounts in every period k, the function $\mathcal{R}$ can be interpreted as a *rebate for* *commitment*, which the retailer obtains from the supplier. The convexity restriction on $\mathcal{F}$ can arise naturally in practice - for instance, when the production of additional units requires installing technologies with increasing marginal cost Zipkin [150], or overtime costs paid to employees. Similarly, the concavity assumption on $\mathcal{R}$ can be seen as an effect of economies of scale (in the rebate payments of the supplier).

Under this setup, we have the following simple result concerning the problem that the retailer has to solve.

Lemma 10. *Assuming that an oracle providing subgradients for the functions $\mathcal{F}$ and $\mathcal{R}$ is available, the computation of the optimal capacities $\boldsymbol{U}$, commitments $\boldsymbol{L}$ and replenishment policies $\{u_k\}_{1\leq k\leq T}$ can be done by solving a subgradient optimization problem. Furthermore, if $\mathcal{F}$ and $\mathcal{R}$ are also piecewise affine, then the retailer only needs to solve a single linear program.*

Proof. Consider a fixed choice of $\boldsymbol{L}, \boldsymbol{U}$. By the result in Theorem 1, the retailer must solve the linear program $(AARC)$ in (2.12) to determine the optimal affine ordering policies. In this LP, $\boldsymbol{L}$ and $\boldsymbol{U}$ appear as right-hand side vectors; therefore, letting $J^\star(\boldsymbol{L}, \boldsymbol{U})$ denote the optimal value of $(AARC)$ as a function of $\boldsymbol{L}, \boldsymbol{U}$, it can be argued by standard results in linear programming duality (see Chapter 5 of Bertsimas and Tsitsiklis [33]) that:

- $J^\star$ is piece-wise affine and convex
- The optimal dual variables corresponding to the constraints involving $\boldsymbol{L}$ and $\boldsymbol{U}$ represent a valid subgradient for $J^\star$.

Therefore, at any fixed $\boldsymbol{L}, \boldsymbol{U}$, the retailer has access to subgradients for the functions $\mathcal{F}(\boldsymbol{U}), \mathcal{R}(\boldsymbol{L})$ and $J^\star(\boldsymbol{L}, \boldsymbol{U})$. Since the objective is always convex, standard nonlinear programming algorithms based on subgradient methods can be used to solve the resulting problem (refer to Bertsekas [20] for a detailed discussion).

Now suppose the functions $\mathcal{F}, \mathcal{R}$ are also piecewise affine, i.e., $\mathcal{F}(\boldsymbol{U}) = \max_{i\in\mathcal{I}} \boldsymbol{f}_i' \boldsymbol{U}$ and $\mathcal{R}(\boldsymbol{L}) = \min_{j\in\mathcal{J}} \boldsymbol{r}_j' \boldsymbol{L}$, where $\mathcal{I}$ and $\mathcal{J}$ are finite index sets, and $\boldsymbol{f}_i, \boldsymbol{r}_j \in \mathbb{R}^T$, $\forall\, i$, $\forall\, j$. Then the retailer can consider a slight modification of problem $(AARC)$, where $\boldsymbol{L}$

and $\boldsymbol{U}$ are decision variables, and the objective is to minimize $J + J_F - J_R$, where J is constrained just as in (2.12), while J_F, J_R are constrained by:

$$J_F \geq \boldsymbol{f}_i' \boldsymbol{U}, \; \forall \, i \in \mathcal{I},$$
$$J_R \leq \boldsymbol{r}_j' \boldsymbol{L}, \; \forall \, j \in \mathcal{J}.$$

In can be easily seen that the resulting problem is an LP, and has the same optimal value as the problem with cost $\mathcal{F}$ and rebate $\mathcal{R}$. □

Chapter 3

A Hierarchy of Near-Optimal Polynomial Policies in the Disturbances

3.1 Introduction

In Chapter 2, we studied a particular instance of multi-stage dynamical systems, where the class of disturbance-affine policies was provably optimal. While insightful from a theoretical viewpoint, the model suffered from several limitations, including the one-dimensional dynamics, the independent (box) state-control constraints, the linear control cost, and the simple structure of the uncertainty sets (box). In the present chapter, we seek to relax several of these modelling pitfalls.

To make things concrete, we consider discrete-time, linear dynamical systems of the form

$$\boldsymbol{x}(k+1) = A(k)\,\boldsymbol{x}(k) + B(k)\,\boldsymbol{u}(k) + \boldsymbol{w}(k), \tag{3.1}$$

evolving over a finite planning horizon, $k = 0, \ldots, T-1$. The variables $\boldsymbol{x}(k) \in \mathbb{R}^n$ represent the state, and the controls $\boldsymbol{u}(k) \in \mathbb{R}^{n_u}$ denote actions taken by the decision maker. $A(k)$ and $B(k)$ are matrices of appropriate dimensions, describing the evolution of the system, and the initial state, $\boldsymbol{x}(0)$, is assumed fixed. The system is affected by unknown[1], additive disturbances, $\boldsymbol{w}(k)$, which are assumed to lie in a given compact, basic semialgebraic set,

$$\mathcal{W}_k \stackrel{\text{def}}{=} \{\boldsymbol{w}(k) \in \mathbb{R}^{n_w} : g_j(\boldsymbol{w}(k)) \geq 0,\, j \in \mathcal{J}_k\}\,, \tag{3.2}$$

where $g_j \in \mathbb{R}[\boldsymbol{w}]$ are multivariate polynomials depending on the vector of uncertainties at time k, $\boldsymbol{w}(k)$, and $\mathcal{J}_k$ is a finite index set. For simplicity, we omit pre-multiplying $\boldsymbol{w}(k)$ by a matrix $C(k)$ in (3.1), since such an evolution could be recast in the current formulation by defining a new uncertainty, $\tilde{\boldsymbol{w}}(k) = C(k)\boldsymbol{w}(k)$, evolving in a suitably adjusted set $\tilde{\mathcal{W}}_k$.

We note that this formulation captures many uncertainty sets of interest in the robust optimization literature (see Ben-Tal et al. [19]), such as polytopic (all g_j affine), p-norms, ellipsoids, and intersections thereof. For now, we restrict our description to uncertainties that are *additive* and *independent across time*, but our framework can also be extended to cases where the uncertainties are *multiplicative* (e.g., affecting the system matrices), and also dependent across time (please refer to Section 3.3.3 for details).

We assume that the dynamic evolution of the system is constrained by a set of linear inequalities,

$$\begin{cases} E_x(k)\,\boldsymbol{x}(k) + E_u(k)\,\boldsymbol{u}(k) \leq \boldsymbol{f}(k),\;\; k = 0,\ldots,T-1, \\ E_x(T)\,\boldsymbol{x}(T) \leq \boldsymbol{f}(T), \end{cases} \tag{3.3}$$

where $E_x(k) \in \mathbb{R}^{r_k \cdot n}, E_u(k) \in \mathbb{R}^{r_k \cdot n_u}, \boldsymbol{f}(k) \in \mathbb{R}^{r_k}$ for the respective k, and the system incurs penalties that are piece-wise affine and convex in the states and controls,

$$h\left(k, \boldsymbol{x}(k), \boldsymbol{u}(k)\right) = \max_{i \in \mathcal{I}_k} \left[\, c_0(k,i) + \boldsymbol{c}_x(k,i)^T \boldsymbol{x}(k) + \boldsymbol{c}_u(k,i)^T \boldsymbol{u}(k) \,\right], \tag{3.4}$$

[1]Just as in Chapter 2, we use the convention that the disturbance $\boldsymbol{w}(k)$ is revealed in period k *after* the control action $\boldsymbol{u}(k)$ is taken, so that $\boldsymbol{u}(k+1)$ is the first decision allowed to depend on $\boldsymbol{w}(k)$.

where $\mathcal{I}_k$ is a finite index set, and $c_0(k,i) \in \mathbb{R}, c_x(k,i) \in \mathbb{R}^n, c_u(k,i) \in \mathbb{R}^{n_u}$ are pre-specified cost parameters. The goal is to find non-anticipatory control policies $\boldsymbol{u}(0), \boldsymbol{u}(1), \ldots, \boldsymbol{u}(T-1)$ that minimize the cost incurred by the system in the worst-case scenario,

$$J = h\left(0, \boldsymbol{x}(0), \boldsymbol{u}(0)\right) + \max_{\boldsymbol{w}(0)} \Big[h\left(1, \boldsymbol{x}(1), \boldsymbol{u}(1)\right) + \ldots \\ + \max_{\boldsymbol{w}(T-2)} \big[h\left(T-1, \boldsymbol{x}(T-1), \boldsymbol{u}(T-1)\right) + \max_{\boldsymbol{w}(T-1)} h\left(T, \boldsymbol{x}(T)\right) \big] \ldots \Big].$$

Examples of such systems naturally arise in many different contexts. One particular instance, in the area of operations management, is the problem of deciding optimal replenishment orders in multi-echelon networks. There, $\boldsymbol{x}(k)$ denotes the vector of all inventories (of potentially different items) stored at various echelons in the supply chain, as well as the replenishment orders that are in the pipeline (i.e., en-route between the echelons), $\boldsymbol{u}_k$ denotes the new replenishment orders placed at the beginning of period k, and $\boldsymbol{w}_k$ denotes exogenous demand from customers. The cost functions represent combinations of holding, backlogging, and inventory reordering costs. The interested reader is referred to the books Zipkin [150], Simchi-Levi et al. [132] and Porteus [119] for more examples and details.

With the *state* of the dynamical system at time k given by $\boldsymbol{x}(k)$, one can resort to the Bellman optimality principle of DP Bertsekas [21] to compute optimal policies, $\boldsymbol{u}^\star(k, \boldsymbol{x}(k))$, and optimal value functions, $J^\star(k, \boldsymbol{x}(k))$. Although DP is a powerful technique as to the theoretical characterization of the optimal policies, it is plagued by the well-known *curse of dimensionality*, in that the complexity of the underlying recursive equations grows quickly with the size of the state-space, rendering the approach ill suited to the computation of actual policy parameters. Therefore, in practice, one would typically solve the recursions numerically (e.g., by multi-parametric programming Bemporad et al. [7, 8, 9]), or resort to approximations, such as approximate DP Bertsekas and Tsitsiklis [23], Powell [120], stochastic approximation Asmussen and Glynn [3], simulation based optimization (Glasserman and Tayur [73], Marbach and Tsitsiklis [103]), and others. Some of the approximations also come with performance guarantees in terms of the objective value in the problem, and many ongoing research efforts are placed on characterizing the sub-optimality gaps resulting from specific classes of policies (the interested reader can refer to the books Bertsekas [21], Bertsekas and Tsitsiklis [23] and Powell [120] for a thorough review).

An alternative approach, which we have already encountered in Chapter 2, is to consider control policies that are parametrized directly in the sequence of observed uncertainties. For the case of linear constraints on the controls, with uncertainties regarded as random variables having bounded support and known distributions, and the goal of minimizing an expected piece-wise quadratic, convex cost, the authors in Garstka and Wets [70] show that piece-wise affine decision rules are optimal, but pessimistically conclude that computing the actual parameterization is usually an "impossible task" (for a precise quantification of that statement, see Dyer and Stougie [60] and Nemirovski and Shapiro [107]).

As briefly discussed in Chapter 2, such disturbance-feedback parameterizations have gained a lot of attention from researchers in robust control and robust optimization (see Löfberg [99], Kerrigan and Maciejowski [87, 88], Goulart and Kerrigan [76], Ben-Tal et al. [14, 15, 17], Skaf and Boyd [133, 134], and references therein). In most of the papers, the authors restrict attention to the case of *affine policies*, and show how reformulations can be done that allow the computation of the policy parameters by solving specific convex optimization problems.

However, with the exception of a few classical cases, such as linear quadratic Gaussian or linear exponential quadratic Gaussian[2], characterizing the performance of affine policies in terms of objective function value is typically very hard. Chapter 2 presented a proof for a one-dimensional case, and also introduced simple examples of multi-dimensional systems where affine policies are (very) sub-optimal.

In fact, in most applications, the restriction to the affine case is done for purposes of tractability, and almost invariably results in loss of performance (see the remarks

[2]These refer to problems that are unconstrained, with Gaussian disturbances, and the goal of minimizing expected costs that are quadratic or exponential of a quadratic, respectively. For these, the optimal policies are affine in the states - see Bertsekas [21] and references therein.

at the end of Nemirovski and Shapiro [107] and in Chapter 14 of Ben-Tal et al. [19]), with the optimality gap being sometimes very large. In an attempt to address this problem, recent work has considered parameterizations that are affine in a new set of variables, derived by lifting the original uncertainties into a higher dimensional space. For example, the authors in Chen and Zhang [50], Chen et al. [52], Sim and Goh [131] suggest using so-called *segregated linear decision rules*, which are affine parameterizations in the positive and negative parts of the original uncertainties. Such policies provide more flexibility, and their computation (for two-stage decision problems in a robust setting) requires roughly the same complexity as that needed for a set of affine policies in the original variables. Another example following similar ideas is Chatterjee et al. [49], where the authors consider arbitrary functional forms of the disturbances, and show how, for specific types of p-norm constraints on the controls, the problems of finding the coefficients of the parameterizations can be relaxed into convex optimization problems. A similar approach is taken in Skaf and Boyd [134], where the authors also consider arbitrary functional forms for the policies, and show how, for a problem with convex state-control constraints and convex costs, such policies can be found by convex optimization, combined with Monte-Carlo sampling (to enforce constraint satisfaction). Chapter 14 of the recent book Ben-Tal et al. [19] also contains a thorough review of several other classes of such adjustable rules, and a discussion of cases when sophisticated rules can actually improve over the affine ones.

The main drawback of some of the above approaches is that the *right* choice of functional form for the decision rules is rarely obvious, and there is no systematic way to influence the trade-off between the performance of the resulting policies and the computational complexity required to obtain them, rendering the frameworks ill-suited for general multi-stage dynamical systems, involving complicated constraints on both states and controls.

The goal of the current chapter is to introduce a new framework for modeling and (approximately) solving such multi-stage dynamical problems. In keeping with the philosophy introduced in our earlier work, we examine the performance of *disturbance-feedback* policies, i.e., policies which are directly parameterized in the sequence of observed uncertainties. While we restrict attention mainly to the robust, mini-max objective setting, our ideas can be extended to deal with stochastic problems, in which the uncertainties are random variables with known, bounded support and distribution that is either fully or partially known[3] (see Section 3.3.3 for a discussion, and Chapter 4 for a more elaborate example). Our main contributions are summarized as follows:

- We introduce a natural extension of the aforementioned affine decision rules, by considering control policies that depend *polynomially* on the observed disturbances. For a fixed polynomial degree d, we develop a convex reformulation of the constraints and objective of the problem, using Sums-Of-Squares (SOS) techniques. In the resulting framework, polynomial policies of degree d can be computed by solving a single semidefinite programming problem (SDP), which, for a fixed precision, can be done in polynomial time (Vandenberghe and Boyd [143]). Our approach is advantageous from a modelling perspective, since it places little burden on the end user (the only choice is the polynomial degree d), while at the same time providing a lever for directly controlling the trade-off between performance and computation (higher d translates into policies with better objectives, obtained at the cost of solving larger SDPs).

- To test our polynomial framework, we consider two classical problems arising in inventory management (single echelon with cumulative order constraints, and serial supply chain with lead-times), and compare the performance of affine, quadratic and cubic control policies. The results obtained are very encouraging - in particular, for all problem instances considered, quadratic policies considerably improve over affine policies (typically by a factor of 2 or 3), while cubic policies essentially close the optimality gap (the relative gap in *all simulations* is less than 1%, with a median gap of less than 0.01%).

The chapter is organized as follows. Section 3.2 presents the mathematical formulation of the problem, briefly discusses relevant solution techniques in the literature,

[3] In the latter case, the cost would correspond to the worst-case distribution consistent with the partial information.

and introduces our framework. Section 3.3, which is the main body of the chapter, first shows how to formulate and solve the problem of searching for the optimal polynomial policy of fixed degree, and then discusses the specific case of polytopic uncertainties. Section 3.3.3 also elaborates on immediate extensions of the framework to more general multi-stage decision problems. Section 3.5 translates two classical problems from inventory management into our framework, and Section 3.6 presents our computational results, exhibiting the strong performance of polynomial policies.

3.1.1 Notation

Throughout the rest of the chapter, we denote scalar quantities by lowercase, non-bold face symbols (e.g. $x \in \mathbb{R}, k \in \mathbb{N}$), vector quantities by lowercase, boldface symbols (e.g. $\boldsymbol{x} \in \mathbb{R}^n, n > 1$), and matrices by uppercase symbols (e.g. $A \in \mathbb{R}^{n \cdot n}, n > 1$). Also, in order to avoid transposing vectors several times, we use the *comma* operator (,) to denote vertical vector concatenation, e.g. with $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\boldsymbol{y} = (y_1, \ldots, y_m) \in \mathbb{R}^m$, we write $(\boldsymbol{x}, \boldsymbol{y}) \stackrel{\text{def}}{=} (x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^{m+n}$.

We refer to quantities specific to time-period k by either including the index in parenthesis, e.g. $\boldsymbol{x}(k)$, $J^\star(k, \boldsymbol{x}(k))$, or by using an appropriate subscript, e.g. $\boldsymbol{x}_k$, $J^\star_k(\boldsymbol{x}_k)$. When referring to the j-th component of a vector at time k, we always use the parenthesis notation for time, and subscript for j, e.g., $x_j(k)$.

Since we seek policies parameterized directly in the uncertainties, we introduce $\boldsymbol{w}_{[k]} \stackrel{\text{def}}{=} (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{k-1})$ to denote the history of known disturbances at the beginning of period k, and $\mathcal{W}_{[k]} \stackrel{\text{def}}{=} \mathcal{W}_1 \times \cdots \times \mathcal{W}_{k-1}$ to denote the corresponding uncertainty set. By convention, $\boldsymbol{w}_{[0]} \equiv \{\emptyset\}$.

With $\boldsymbol{x} = (x_1, \ldots, x_n)$, we denote by $\mathbb{R}[\boldsymbol{x}]$ the ring of polynomials in variables $x_1, \ldots, x_n$, and by $\mathcal{P}_d[\boldsymbol{x}]$ the $\mathbb{R}$-vector space of polynomials in $x_1, \ldots, x_n$, with degree at most d. We also let

$$\mathcal{B}_d(\boldsymbol{x}) \stackrel{\text{def}}{=} \left(1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_1x_n, x_2^2, x_2x_3 \ldots, x_n^d\right) \tag{3.5}$$

be the canonical basis of $\mathcal{P}_d[\boldsymbol{x}]$, and $s(d) \stackrel{\text{def}}{=} \binom{n+d}{d}$ be its dimension. Any polynomial

$p \in \mathcal{P}_d[\boldsymbol{x}]$ is written as a finite linear combination of monomials,

$$p(\boldsymbol{x}) = p(x_1, \ldots, x_n) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} p_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} = \boldsymbol{p}^T \mathcal{B}_d(\boldsymbol{x}), \tag{3.6}$$

where $\boldsymbol{x}^{\boldsymbol{\alpha}} \overset{\text{def}}{=} x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$, and the sum is taken over all n-tuples $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}^n$ satisfying $\sum_{i=1}^n \alpha_i \leq d$. In the expression above, $\boldsymbol{p} = (p_{\boldsymbol{\alpha}}) \in \mathbb{R}^{s(r)}$ is the vector of coefficients of $p(\boldsymbol{x})$ in the basis (3.5). In situations where the coefficients $p_{\boldsymbol{\alpha}}$ of a polynomial are decision variables, in order to avoid confusions, we refer to $\boldsymbol{x}$ as the *indeterminate* (similarly, we refer to $p(\boldsymbol{x})$ as a polynomial in indeterminate $\boldsymbol{x}$). By convention, we take $p(\emptyset) \equiv p_{0,0,\ldots,0}$, i.e., a polynomial without indeterminate is simply a constant.

For a polynomial $p \in \mathbb{R}[\boldsymbol{x}]$, we use $\deg(p)$ to denote the largest degree of a monomial present in p.

3.2 Problem Description

Using the notation mentioned in the introduction, our goal is to find non-anticipatory control policies $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{T-1}$ that minimize the cost incurred by the system in the worst-case scenario. In other words, we seek to solve the problem:

$$\min_{\boldsymbol{u}_0} \Big[h_0(\boldsymbol{x}_0, \boldsymbol{u}_0) + \max_{\boldsymbol{w}_0} \min_{\boldsymbol{u}_1} \Big[h_1(\boldsymbol{x}_1, \boldsymbol{u}_1) + \cdots + \\ + \min_{\boldsymbol{u}_{T-1}} \big[h_{T-1}(\boldsymbol{x}_{T-1}, \boldsymbol{u}_{T-1}) + \max_{\boldsymbol{w}_{T-1}} h_T(\boldsymbol{x}_T) \big] \ldots \Big] \Big] \tag{3.7a}$$

$$(P) \qquad \text{s.t.} \quad \boldsymbol{x}_{k+1} = A_k \boldsymbol{x}_k + B_k \boldsymbol{u}_k + \boldsymbol{w}_k, \quad \forall\, k \in \{0, \ldots, T-1\}, \tag{3.7b}$$

$$E_x(k)\, \boldsymbol{x}_k + E_u(k)\, \boldsymbol{u}_k \leq \boldsymbol{f}_k, \qquad \forall\, k \in \{0, \ldots, T-1\}, \tag{3.7c}$$

$$E_x(T)\, \boldsymbol{x}_T \leq \boldsymbol{f}_T. \tag{3.7d}$$

As already mentioned, the control actions $\boldsymbol{u}_k$ do not have to be decided entirely at time period $k = 0$, i.e., (P) does not have to be solved as an open-loop problem.

Rather, $\boldsymbol{u}_k$ is allowed to depend on the information set available[4] at time k, resulting in control policies $\boldsymbol{u}_k : \mathcal{F}_k \to \mathbb{R}^{n_u}$, where $\mathcal{F}_k$ consists of past states, controls and disturbances, $\mathcal{F}_k = \{\boldsymbol{x}_t\}_{0 \le t \le k} \cup \{\boldsymbol{u}_t\}_{0 \le t < k} \cup \{\boldsymbol{w}_t\}_{0 \le t < k}$.

While $\mathcal{F}_k$ is a large (expanding with k) set, the state $\boldsymbol{x}_k$ represents sufficient information for taking optimal decisions at time k. Thus, with control policies depending on the states, one can resort to the Bellman optimality principle of Dynamic Programming (DP) (Bertsekas [21]), to compute optimal policies, $\boldsymbol{u}_k^\star(\boldsymbol{x}_k)$, and optimal value functions, $J_k^\star(\boldsymbol{x}_k)$. As suggested in the introduction, the approach is limited due to the *curse of dimensionality*, so that, in practice, one typically resorts to approximate schemes for computing suboptimal, state-dependent policies (Bertsekas and Tsitsiklis [23], Powell [120], Marbach and Tsitsiklis [103]).

In keeping with the approach introduced in Chapter 2, we take a slightly different view, and consider instead policies parametrized directly in the observed uncertainties,

$$\boldsymbol{u}_k : \mathcal{W}_0 \times \mathcal{W}_1 \times \cdots \times \mathcal{W}_{k-1} \to \mathbb{R}^{n_u}. \tag{3.8}$$

In this context, the decisions that must be taken are the parameters defining the specific functional form sought for $\boldsymbol{u}_k$. A particular example of disturbance-feedback policies, which we have already encountered in Chapter 2, is the *affine* case, i.e., $\boldsymbol{u}_k = L_k \cdot (1, \boldsymbol{w}_0, \ldots, \boldsymbol{w}_{k-1})$, where the decision variables are the coefficients of the matrices $L_k \in \mathbb{R}^{n_u \times (1 + k \times n_w)}$, $k = 0, \ldots, T-1$.

In this framework, with (3.7b) used to express the dependency of states $\boldsymbol{x}_k$ on past uncertainties, the state-control constraints (3.7c), (3.7d) at time k can be written as functions of the parametric decisions $L_0, \ldots, L_k$ and the uncertainties $\boldsymbol{w}_0, \ldots, \boldsymbol{w}_{k-1}$, and one typically requires these constraints to be obeyed *robustly*, i.e., for any possible realization of the uncertainties.

As already mentioned, this approach has been explored before in the literature, in both the stochastic and robust frameworks (Birge and Louveaux [41], Garstka and Wets [70], Löfberg [99], Kerrigan and Maciejowski [87, 88], Goulart and Kerrigan

[4]More formally, the decision process u_k is adapted to the filtration generated by past values of the disturbances and controls.

[76], Ben-Tal et al. [14, 15, 17], Bertsimas and Brown [26], Skaf and Boyd [133]). The typical restriction to the sub-class of affine policies, done for purposes of tractability, almost invariably results in loss of performance Nemirovski and Shapiro [107], with the gap being sometimes very large. To illustrate this effect, we introduce the following simple example[5], motivated by a similar case in Chen and Zhang [50]:

Example 4. *Consider a two-stage problem, where* $\boldsymbol{w} \in \mathcal{W}$ *is the uncertainty, with* $\mathcal{W} = \big\{\boldsymbol{w} \in \mathbb{R}^N : \|\boldsymbol{w}\|_2 \leq 1\big\}$*,* $x \in \mathbb{R}$ *is a first-stage decision (taken before* $\boldsymbol{w}$ *is revealed), and* $\boldsymbol{y} \in \mathbb{R}^N$ *is a second-stage decision (allowed to depend on* $\boldsymbol{w}$*). We would like to solve the following optimization:*

$$
\begin{aligned}
\underset{x,\boldsymbol{y}(\boldsymbol{w})}{\text{minimize}} \quad & x \\
\text{such that} \quad & x \geq \sum_{i=1}^{N} y_i, \quad \forall\, \boldsymbol{w} \in \mathcal{W}, \\
& y_i \geq w_i^2, \qquad \forall\, \boldsymbol{w} \in \mathcal{W}.
\end{aligned}
\tag{3.9}
$$

It can be easily shown (see Lemma 14 in Appendix B.1) that the optimal objective in Problem (3.9) is 1, corresponding to $y_i(\boldsymbol{w}) = w_i^2$, while the best objective achievable under *affine* policies $\boldsymbol{y}(\boldsymbol{w})$ is N, for $y_i(\boldsymbol{w}) = 1,\ \forall\, i$. In particular, this simple example shows that the optimality gap resulting from the use of affine policies can be made arbitrarily large (as the problem size increases).

Motivated by these facts, in the current chapter, we explore the performance of a more general class of disturbance-feedback control laws, namely policies that are *polynomial* in past-observed uncertainties. More precisely, for a specified degree d, and with $\boldsymbol{w}_{[k]}$ denoting the vector of all disturbances in $\mathcal{F}_k$,

$$
\boldsymbol{w}_{[k]} \stackrel{\text{def}}{=} (\,\boldsymbol{w}_0, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{k-1}\,) \in \mathbb{R}^{k \cdot n_w}, \tag{3.10}
$$

we consider a control law at time k in which every component is a polynomial of

[5] We note that this example can be easily cast as an instance of Problem (P). We opt for the simpler notation to keep the ideas clear.

degree at most d in variables $\boldsymbol{w}_{[k]}$, i.e., $u_j(k, \boldsymbol{w}_{[k]}) \in \mathcal{P}_d[\boldsymbol{w}_{[k]}]$, and thus:

$$\boldsymbol{u}_k(\boldsymbol{w}_{[k]}) = L_k \, \mathcal{B}_d(\boldsymbol{w}_{[k]}), \tag{3.11}$$

where $\mathcal{B}_d(\boldsymbol{w}_{[k]})$ is the canonical basis of $\mathcal{P}_d[\boldsymbol{w}_{[k]}]$, given by (3.5). The new decision variables become the matrices of coefficients $L_k \in \mathbb{R}^{n_u \cdot s(d)}$, $k = 0, \ldots, T-1$, where $s(d) = \binom{k \cdot n_w + d}{d}$ is the dimension of $\mathcal{P}_d[\boldsymbol{w}_{[k]}]$. Therefore, with a fixed degree d, the number of decision variables remains polynomially bounded in the size of the problem input, T, n_u, n_w.

This class of policies constitutes a natural extension of the disturbance-affine control laws, i.e., the case $d = 1$. Furthermore, with sufficiently large degree, one can expect the performance of the polynomial policies to become near-optimal - recall that, by the Stone-Weierstrass Theorem (Rudin [127]), any continuous function on a compact set can be approximated as closely as desired by polynomial functions. The main drawback of the approach is that searching over arbitrary polynomial policies typically results in non-convex optimization problems. To address this issue, in the next section, we develop a tractable, convex reformulation of the problem based on Sum-Of-Squares (SOS) techniques (Parrilo [112, 113], Lasserre [94]).

3.3 Polynomial Policies and Convex Reformulations Using Sums-Of-Squares

Under polynomial policies of the form (3.11), one can use the dynamical equation (3.7b) to express every component of the state at time k, $x_j(k)$, as a polynomial in indeterminate $\boldsymbol{w}_{[k]}$, whose coefficients are linear combinations of the entries in $\{L_t\}_{0 \le t \le k-1}$. As such, with $\boldsymbol{e}_x(k,j)^T$ and $\boldsymbol{e}_u(k,j)^T$ denoting the j-th row of $E_x(k)$ and $E_u(k)$, respectively, a typical state-control constraint (3.7c) can be written

$$\boldsymbol{e}_x(k,j)^T \boldsymbol{x}_k + \boldsymbol{e}_u(k,j)^T \boldsymbol{u}_k \le \boldsymbol{f}_j(k) \quad \Leftrightarrow$$
$$p^{\text{con}}_{j,k}(\boldsymbol{w}_{[k]}) \stackrel{\text{def}}{=} \boldsymbol{f}_j(k) - \boldsymbol{e}_x(k,j)^T \boldsymbol{x}_k - \boldsymbol{e}_u(k,j)^T \boldsymbol{u}_k \ge 0, \, \forall \, \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]}.$$

In particular, feasibility of the state-control constraints at time k is equivalent to ensuring that the coefficients $\{L_t\}_{0\le t\le k-1}$ are such that the polynomials $p^{\text{con}}_{j,k}(\boldsymbol{w}_{[k]}), j = 1,\ldots,r_k$, are non-negative on the domain $\mathcal{W}_{[k]}$.

Similarly, the expression (3.4) for the stage cost at time k can be written as

$$h_k(\boldsymbol{x}_k,\boldsymbol{u}_k) = \max_{i\in\mathcal{I}_k} p^{\text{cost}}_i(\boldsymbol{w}_{[k]}),$$
$$p^{\text{cost}}_i(\boldsymbol{w}_{[k]}) \stackrel{\text{def}}{=} c_0(k,i) + \boldsymbol{c}_x(k,i)^T\boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + \boldsymbol{c}_u(k,i)^T\boldsymbol{u}_k(\boldsymbol{w}_{[k]}),$$

i.e., the cost h_k is a piece-wise polynomial function of the past-observed disturbances $\boldsymbol{w}_{[k]}$. Therefore, under polynomial control policies, we can rewrite the original Problem (P) as the following polynomial optimization problem:

$$(P_{\text{POP}}) \quad \min_{L_0}\Bigg[\max_{i\in\mathcal{I}_1} p^{\text{cost}}_i(\boldsymbol{w}_{[0]}) + \max_{\boldsymbol{w}_0}\min_{L_1}\Big[\max_{i\in\mathcal{I}_2} p^{\text{cost}}_i(\boldsymbol{w}_{[1]}) + \ldots + \max_{\boldsymbol{w}_{T-2}}\min_{L_{T-1}}\big[\max_{i\in\mathcal{I}_{T-1}} p^{\text{cost}}_i(\boldsymbol{w}_{[T-1]}) + \max_{\boldsymbol{w}_{T-1}}\max_{i\in\mathcal{I}_T} p^{\text{cost}}_i(\boldsymbol{w}_{[T]})\big]\ldots\Big]\Bigg] \quad (3.12a)$$

$$\text{s.t.} \quad p^{\text{con}}_{j,k}(\boldsymbol{w}_{[k]}) \ge 0, \qquad \forall\, k=0,\ldots,T,\ \forall\, j=1,\ldots,r_k,\ \forall\, \boldsymbol{w}_{[k]}\in\mathcal{W}_{[k]}. \quad (3.12b)$$

In this formulation, the decision variables are the coefficients $\{L_t\}_{0\le t\le T-1}$, and (3.12b) summarize all the state-control constraints. We emphasize that the expression of the polynomial controls (3.11) and the dynamical system equation (3.7b) should not be interpreted as real constraints in the problem (rather, they are only used to derive the dependency of the polynomials $p^{\text{cost}}_i(\boldsymbol{w}_{[k]})$ and $p^{\text{con}}_{j,k}(\boldsymbol{w}_{[k]})$ on $\{L_t\}_{0\le t\le k-1}$ and $\boldsymbol{w}_{[k]}$).

3.3.1 Reformulating the Constraints

As mentioned in the previous section, under polynomial control policies, a typical state-control constraint (3.12b) in program (P_{POP}) can now be written as:

$$p(\boldsymbol{\xi}) \ge 0, \quad \forall\,\boldsymbol{\xi}\in\mathcal{W}_{[k]}, \qquad (3.13)$$

where $\boldsymbol{\xi} \equiv \boldsymbol{w}_{[k]} \in \mathbb{R}^{k \cdot n_w}$ is the history of disturbances, and $p(\boldsymbol{\xi})$ is a polynomial in variables $\xi_1, \xi_2, \ldots, \xi_{k \cdot n_w}$ with degree at most d,

$$p(\boldsymbol{\xi}) = \boldsymbol{p}^T \mathcal{B}_d(\boldsymbol{\xi}),$$

whose coefficients p_i are affine combinations of the decision variables L_t, $0 \le t \le k-1$. It is easy to see that constraint (3.13) can be rewritten equivalently as

$$p(\boldsymbol{\xi}) \ge 0, \qquad \forall\, \boldsymbol{\xi} \in \mathcal{W}_{[k]} \stackrel{\text{def}}{=} \left\{ \boldsymbol{\xi} \in \mathbb{R}^{k \cdot n_w} : g_j(\boldsymbol{\xi}) \ge 0,\ j = 1, \ldots, m \right\}, \tag{3.14}$$

where $\{g_j\}_{1 \le j \le m}$ are all the polynomial functions describing the compact basic semi-algebraic set $\mathcal{W}_{[k]} \equiv \mathcal{W}_0 \times \cdots \times \mathcal{W}_{k-1}$, immediately derived from (3.2). In this form, (3.14) falls in the general class of constraints that require testing polynomial non-negativity on a basic closed, semi-algebraic set, i.e., a set given by a finite number of polynomial equalities and inequalities. To this end, note that a *sufficient* condition for (3.14) to hold is:

$$p = \sigma_0 + \sum_{j=1}^{m} \sigma_j\, g_j, \tag{3.15}$$

where $\sigma_j \in \mathbb{R}[\boldsymbol{\xi}]$, $j = 0, \ldots, m$, are polynomials in the variables $\boldsymbol{\xi}$ which are furthermore *sums of squares* (SOS). This condition translates testing the non-negativity of p on the set $\mathcal{W}_{[k]}$ into a system of linear equality constraints on the coefficients of p and σ_j, $j = 0, \ldots, m$, and a test whether σ_j are SOS. The main reason why this is valuable is because testing whether a polynomial of fixed degree is SOS is equivalent to solving a semidefinite programming problem (SDP) (refer to Parrilo [112, 113], Lasserre [94] for details), which, for a fixed precision, can be done in polynomial time, by interior point methods (Vandenberghe and Boyd [143]).

On first sight, condition (3.15) might seem overly restrictive. However, it is motivated by recent powerful results in real algebraic geometry (Putinar [121], Jacobi and Prestel [84]), which, under mild conditions[6] on the functions g_j, state that *any*

[6]These are readily satisfied when g_j are affine, or can be satisfied by simply appending a redundant

polynomial that is strictly positive on a compact semi-algebraic set $\mathcal{W}_{[k]}$ *must* admit a representation of the form (3.15), where the degrees of the σ_j polynomials are not a priori bounded. In our framework, in order to obtain a tractable formulation, we furthermore restrict these degrees so that the total degree of every product $\sigma_j \, g_j$ is at most $\max\Big(d, \max_j\big(\deg(g_j)\big)\Big)$, the maximum between the degree of the control policies (3.11) under consideration and the largest degree of the polynomials g_j giving the uncertainty sets. While this requirement is more restrictive, and could, in principle, result in conservative parameter choices, it avoids ad-hoc modeling decisions and has the advantage of keeping a single parameter that is adjustable to the user (the degree d), which directly controls the trade-off between the size of the resulting SDP formulation and the quality of the overall solution. Furthermore, in our numerical simulations, we find that this choice performs very well in practice, and never results in infeasible conditions.

3.3.2 Reformulating the Objective

Recall from our discussion in the beginning of Section 3.3 that, under polynomial control policies, a typical stage cost becomes a piecewise polynomial function of past uncertainties, i.e., a maximum of several polynomials. A natural way to bring such a cost into the framework presented before is to introduce, for every stage $k = 0, \ldots, T$, a polynomial function of past uncertainties, and require it to be an upper-bound on the true (piecewise polynomial) cost.

More precisely, and to fix ideas, consider the stage cost at time k, which, from our earlier discussion, can be written as

$$\begin{aligned} h_k(\boldsymbol{x}_k, \boldsymbol{u}_k) &= \max_{i \in \mathcal{I}_k} p_i^{\text{cost}}(\boldsymbol{w}_{[k]}), \\ p_i^{\text{cost}}(\boldsymbol{w}_{[k]}) &= c_0(k,i) + \boldsymbol{c}_x(k,i)^T \boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + \boldsymbol{c}_u(k,i)^T \boldsymbol{u}_k(\boldsymbol{w}_{[k]}), \, \forall \, i \in \mathcal{I}_k. \end{aligned}$$

In this context, we introduce a modified stage cost $\tilde{h}_k \in \mathcal{P}_d[\boldsymbol{w}_{[k]}]$, which we con-

constraint that bounds the 2-norm of the vector $\boldsymbol{\xi}$.

strain to satisfy

$$\tilde{h}_k(\boldsymbol{w}_{[k]}) \geq p_i^{\text{cost}}(\boldsymbol{w}_{[k]}), \qquad \forall\, \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]},\ \forall\, i \in \mathcal{I}_k,$$

and we replace the overall cost for Problem (P_{POP}) with the sum of the modified stage costs. In other words, instead of minimizing the objective (3.7a), we seek to solve:

$$\begin{aligned}
&\min J \\
\text{s.t.} \qquad & J \geq \sum_{k=0}^{T} \tilde{h}_k(\boldsymbol{w}_{[k]}), \quad \forall\, \boldsymbol{w}_{[T]} \in \boldsymbol{w}_{[T]}, && (3.16\text{a}) \\
&\tilde{h}_k(\boldsymbol{w}_{[k]}) \geq p_i^{\text{cost}}(\boldsymbol{w}_{[k]}), \quad \forall\, \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]},\ \forall\, i \in \mathcal{I}_k. && (3.16\text{b})
\end{aligned}$$

The advantage of this approach is that, now, constraints (3.16a) and (3.16b) are of the exact same nature as (3.13), and thus fit into the SOS framework developed earlier. As a result, we can use the same semidefinite programming approach to enforce them, while preserving the tractability of the formulation and the trade-off between performance and computation delivered by the degree d. The main drawback is that the cost J may conceivably, in general, over-bound the optimal cost of Problem (P), due to several reasons:

1. We are replacing the (true) piece-wise polynomial cost h_k with an *upper bound* given by the polynomial cost $\tilde{h}_k$. Therefore, the optimal value J of problem (3.16a) may, in general, be larger than the true cost corresponding to the respective polynomial policies, i.e., the cost of problem (P_{POP}).

2. All the constraints in the model, namely (3.16a), (3.16b), and the state-control constraints (3.12b), are enforced using SOS polynomials with fixed degree (see the discussion in Section 3.3.1), and this is sufficient, but not necessary.

However, despite these multiple layers of approximation, our numerical experiments, presented in Section 3.6, suggest that most of the above considerations are second-order effects when compared with the fact that polynomial policies of the form (3.11), are themselves, in general, suboptimal. In fact, our results suggest that with a modest polynomial degree (3, and sometimes even 2), one can close most of the optimality gap between the SDP formulation and the optimal value of Problem (P).

To summarize, our framework can be presented as the sequence of steps below:

Algorithm 3 Framework for computing polynomial policies of degree d

1: Consider polynomial control policies in the disturbances, $\boldsymbol{u}_k(\boldsymbol{w}_{[k]}) = L_k\,\mathcal{B}_d\big(\boldsymbol{w}_{[k]}\big)$.

2: Express all the states $\boldsymbol{x}_k$ according to equation (3.7b). Each component of a typical state $\boldsymbol{x}_k$ becomes a polynomial in indeterminate $\boldsymbol{w}_{[k]}$, with coefficients given by linear combinations of $\{L_t\}_{0\le t\le k-1}$.

3: Replace a typical stage cost $h_k(\boldsymbol{x}_k, \boldsymbol{u}_k) = \max_{i\in\mathcal{I}_k} p_i^{\text{cost}}(\boldsymbol{w}_{[k]})$ with a modified stage cost $\tilde{h}_k \in \mathcal{P}_d[\boldsymbol{w}_{[k]}]$, constrained to satisfy $\tilde{h}_k(\boldsymbol{w}_{[k]}) \ge p_i^{\text{cost}}(\boldsymbol{w}_{[k]}), \forall\, \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]}, \forall\, i \in \mathcal{I}_k$.

4: Replace the overall cost with the sum of the modified stage costs.

5: Replace a typical constraint $p(\boldsymbol{w}_{[k]}) \ge 0,\ \forall\, \boldsymbol{w}_{[k]} \in \big\{\boldsymbol{\xi} : g_j(\boldsymbol{\xi}) \ge 0,\ j = 1, \dots, m\big\}$ (for either state-control or costs) with the requirements:

$$
\begin{aligned}
& p = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j && \big(\text{linear constraints on coefficients}\big) \\
& \sigma_j \text{ SOS},\ j = 0, \dots, m. && \big(m+1 \text{ SDP constraints}\big) \\
& \quad \deg(\sigma_j\, g_j) \le \max\Big(d, \max_j\big(\deg(g_j)\big)\Big), \\
& \quad \deg(\sigma_0) = \max_j\big(\deg(\sigma_j\, g_j)\big).
\end{aligned}
$$

6: Solve the resulting SDP to obtain the coefficients L_k of the policies.

The size of the overall formulation is controlled by the following parameters:

- There are $\mathcal{O}\Big(T^2 \cdot \max_k(r_k + |\mathcal{I}_k|) \cdot (\max_k |\mathcal{J}_k|) \cdot \binom{T\cdot n_w + \hat{d}}{\hat{d}}\Big)$ linear constraints
- There are $\mathcal{O}\big(T^2 \cdot \max_k(r_k + |\mathcal{I}_k|) \cdot (\max_k |\mathcal{J}_k|)\big)$ SDP constraints, each of size at most $\binom{T\cdot n_w + \lceil \frac{\hat{d}}{2} \rceil}{\lceil \frac{\hat{d}}{2} \rceil}$
- There are $\mathcal{O}\Big(T \cdot \big[n_u + T \cdot \max_k(r_k + |\mathcal{I}_k|) \cdot (\max_k |\mathcal{J}_k|)\big] \binom{T\cdot n_w + \hat{d}}{\hat{d}}\Big)$ variables

Above, $\hat{d} \stackrel{\text{def}}{=} \max\Big(d, \max_j\big(\deg(g_j)\big)\Big)$, i.e., the largest between d and the degree of any polynomial g_j defining the uncertainty sets. Since, for all practical purposes, most uncertainty sets considered in the literature are polyhedral or quadratic, the main parameter that controls the complexity is d (for $d \geq 2$).

As the main computational bottleneck comes from the SDP constraints, we note that their size and number could be substantially reduced by requiring the control policies to only depend on a partial history of the uncertainties, e.g., by considering $\boldsymbol{u}_k : \mathcal{W}_{k-q} \times \mathcal{W}_{k-q+1} \times \cdots \times \mathcal{W}_{k-1}$, for some fixed $q > 0$, and by restricting $\boldsymbol{x}_k$ in a similar fashion. In this case, there would be $\mathcal{O}\big(T \cdot q \cdot \max_k(r_k + |\mathcal{I}_k|) \cdot (\max_k |\mathcal{J}_k|)\big)$ SDP constraints, each of size at most $\binom{q \cdot n_w + \lceil \frac{\hat{d}}{2} \rceil}{\lceil \frac{\hat{d}}{2} \rceil}$, and only $\mathcal{O}\big(\sum_k |\mathcal{J}_k|\big)$ SDP constraints of size $\binom{T \cdot n_w + \lceil \frac{\hat{d}}{2} \rceil}{\lceil \frac{\hat{d}}{2} \rceil}$.

3.3.3 Extensions

For completeness, we conclude our discussion by briefly mentioning several modelling extensions that can be readily captured in our framework:

1. Although we only consider uncertainties that are "independent" across time, i.e., the history $\boldsymbol{w}_{[k]}$ always belongs to the cartesian product $\mathcal{W}_0 \times \cdots \times \mathcal{W}_{k-1}$, our approach could be immediately extended to situations in which the uncertainty sets characterize partial sequences. As an example, instead of $\mathcal{W}_k$, we could specify a semi-algebraic description for the history $\mathcal{W}_{[k]}$,

$$(\boldsymbol{w}_0, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{k-1}) \in \mathcal{W}_{[k]} = \big\{\boldsymbol{\xi} \in \mathbb{R}^{k \times n_w} \;:\; g_j(\boldsymbol{\xi}) \geq 0, \forall\, j \in \tilde{\mathcal{J}}_k\big\},$$

   which could be particularly useful in situations where the uncertainties are generated by processes that are dependent across time. The only modification would be to use the new specification for the set $\mathcal{W}_{[k]}$ in the typical state-control constraints (3.13) and the cost reformulation constraints (3.16a), (3.16b).

2. While we restrict the exposition to uncertainties that are only affecting the system dynamics additively, i.e., by means of equation (3.1), the framework

can be extended to situations where the system and constraint matrices, $A(k)$, $B(k)$, $E_x(k)$, $E_u(k)$, $\boldsymbol{f}(k)$ or the cost parameters, $\boldsymbol{c}_x(k,i)$ or $\boldsymbol{c}_u(k,i)$ are also affected by uncertainty. These situations are of utmost practical interest, in both the inventory examples that we consider in the current chapter, but also in other realistic dynamical systems. As an example, suppose that the matrix $A(k)$ is affinely dependent on uncertainties $\boldsymbol{\zeta}_k \in \mathcal{Z}_k \subset \mathbb{R}^{n_\zeta}$,

$$A(k) = A_0(k) + \sum_{i=1}^{n_\zeta} \zeta_i(k) A_i(k),$$

where $A_i(k) \in \mathbb{R}^{n \times n}, \forall\, i \in \{0, \ldots, n_\zeta\}$ are deterministic matrices, and $\mathcal{Z}_k$ are closed, basic semi-algebraic sets. Then, provided that the uncertainties $\boldsymbol{w}_k$ and $\boldsymbol{\zeta}_k$ are both observable in every period[7], our framework can be immediately extended to decision policies that depend on the histories of both sources of uncertainty, i.e., $\boldsymbol{u}_k(\boldsymbol{w}_0, \ldots, \boldsymbol{w}_{k-1}, \boldsymbol{\zeta}_0, \ldots, \boldsymbol{\zeta}_{k-1})$.

3. Note that, instead of considering uncertainties as lying in given sets, and adopting a min-max (worst-case) objective, we could accommodate the following modelling assumptions:

   (a) The uncertainties are random variables, with bounded support given by the set $\mathcal{W}_0 \times \mathcal{W}_1 \times \ldots \mathcal{W}_{T-1}$, and known probability distribution function $\mathbb{F}$. The goal is to find $\boldsymbol{u}_0, \ldots, \boldsymbol{u}_{T-1}$ so as to obey the state-control constraints (3.3) almost surely, and to minimize the expected costs,

$$\begin{aligned}\min_{\boldsymbol{u}_0} \Big[h_0(\boldsymbol{x}_0, \boldsymbol{u}_0) + \mathbb{E}_{\boldsymbol{w}_0 \sim \mathbb{F}} \min_{\boldsymbol{u}_1} \big[h_1(\boldsymbol{x}_1, \boldsymbol{u}_1) + \ldots \\ + \min_{\boldsymbol{u}_{T-1}} \big[h_{T-1}(\boldsymbol{x}_{T-1}, \boldsymbol{u}_{T-1}) + \mathbb{E}_{\boldsymbol{w}_{T-1} \sim \mathbb{F}} h_T(\boldsymbol{x}_T) \big] \ldots \big] \Big]. \end{aligned} \quad (3.17)$$

   In this case, since our framework already enforces almost sure (robust) constraint satisfaction, the only potential modifications would be in the

[7] When only the states $\boldsymbol{x}_k$ are observable, then one might not be able to simultaneously discriminate and measure both uncertainties.

reformulation of the objective. Since the distribution of the uncertainties is assumed known, and the support is bounded, the moments exist and can be computed up to any fixed degree d. Therefore, we could preserve the reformulation of state-control constraints and stage-costs in our framework (i.e., Steps 2 and 4), but then proceed to minimize the *expected* sum of the polynomial costs $\tilde{h}_k$ (note that the expected value of a polynomial function of uncertainties can be immediately obtained as a linear function of the moments).

(b) The uncertainties are random variables, with the same bounded support as above, but unknown distribution function $\mathbb{F}$, belonging to a given set of distributions, $\mathcal{F}$. The goal is to find control policies obeying the constraints almost surely, and minimizing the expected costs corresponding to the *worst-case distribution* $\mathbb{F}$,

$$\min_{\boldsymbol{u}_0}\Bigg[h_0\left(\boldsymbol{x}_0, \boldsymbol{u}_0\right) + \sup_{\mathbb{F}\in\mathcal{F}} \mathbb{E}_{\boldsymbol{w}_0} \min_{\boldsymbol{u}_1}\Big[h_1\left(\boldsymbol{x}_1, \boldsymbol{u}_1\right) + \cdots + \min_{\boldsymbol{u}_{T-1}}\Big[h_{T-1}\left(\boldsymbol{x}_{T-1}, \boldsymbol{u}_{T-1}\right) + \sup_{\mathbb{F}\in\mathcal{F}} \mathbb{E}_{\boldsymbol{w}_{T-1}} h_T\left(\boldsymbol{x}_T\right)\Big] \ldots\Big]\Bigg]. \tag{3.18}$$

In this case, if partial information (such as the moments of the distribution up to degree d) is available, then the framework in (a) can be applied. Otherwise, if the only information available about $\mathbb{F}$ were the support, then our framework could be applied without modification, but the solution obtained would exactly correspond to the min-max approach, and hence be quite conservative.

We note that, under moment information, some of the seemingly "ad-hoc" substitutions that we performed in our framework can actually become tight. More precisely, the recent paper Zuluaga and Pena [151] argues that, when the set of measures $\mathcal{F}$ is characterized by a compact support and fixed moments up to degree d, then the optimal value in the worst-case expected cost problem $\sup_{\mathbb{F}\in\mathcal{F}} \mathbb{E}_{\boldsymbol{w}_{[k]}} h_k\left(\boldsymbol{x}_k, \boldsymbol{u}_k\right)$ (where h_k are piece-wise polynomial functions) *exactly* corresponds to the cost $\sup_{\mathbb{F}\in\mathcal{F}} \mathbb{E}_{\boldsymbol{w}_{[k]}} \tilde{h}_k(\boldsymbol{w}_{[k]})$, where $\tilde{h}_k$ are exactly given by the constraints (3.16b). In other words, introducing a single modified polynomial stage cost of the form does not increase the optimal value of the problem under the distributionally-robust framework.

In general, under the distributionally robust framework, if more information about the measures in the set $\mathcal{F}$ is available, such as uni-modality, symmetry, directional deviations (Chen et al. [51]), then one should be able to obtain better bounds on the stage costs h_k, by employing appropriate Tchebycheff-type inequalities (Bertsimas and Popescu [29], Popescu [117], Zuluaga and Pena [151]). The interested reader to the recent papers Popescu [118], Natarajan et al. [106], Chen et al. [52], Sim and Goh [131], which take similar approaches in related contexts.

While these extensions are certainly worthy of attention, we do not pursue them here, and restrict our discussion in the remainder of the chapter to the original worst-case formulation. For a more elaborate discussion of the distributionally-robust framework (in a slightly different setting), we refer the interested reader to Chapter 4 of the thesis.

3.4 Other Methodologies for Computing Decision Rules or Exact Values

Our goal in the current section is to discuss the relation between our polynomial hierarchy and several other established methodologies in the literature[8] for computing *affine* or *quadratic* decision rules. More precisely, for the case of ∩-ellipsoidal uncertainty sets, we show that our framework delivers policies of degree 1 or 2 with performance at least as good as that obtained by applying the methods in Ben-Tal et al. [19]. In the second part of the section, we discuss the particular case of polytopic uncertainty sets, where exact values for Problem (P) can be found (which are very useful for benchmarking purposes).

[8] We are grateful to one of the anonymous referees for pointing out reference Ben-Tal et al. [19], which was not at our disposal at the time of conducting the research.

3.4.1 Affine and Quadratic Policies for ∩-Ellipsoidal Uncertainty Sets

Let us consider the specific case when the uncertainty sets $\mathcal{W}_k$ are given by the intersection of finitely many convex quadratic forms, and have nonempty interior - one of the most general classes of uncertainty sets treated in the robust optimization literature (see, e.g., Ben-Tal et al. [19]).

We first focus attention on *affine* disturbance-feedback policies, i.e., $\boldsymbol{u}_k(\boldsymbol{w}_{[k]}) = L_k\, \mathcal{B}_1(\boldsymbol{w}_{[k]})$, and perform the same substitution of a piece-wise affine stage cost with an affine cost that over-bounds it[9]. Finding the optimal affine policies then requires solving the following instance of Problem (P_{POP}):

$$
(P_{\text{AFF}}) \quad
\begin{aligned}
&\min_{L_k, \boldsymbol{z}_k, z_{k,0}, J} \quad J && \text{(3.19a)}\\
&J \geq \sum_{k=0}^{T} \big(\boldsymbol{z}_k^T \boldsymbol{w}_{[k]} + z_{k,0} \big), && \text{(3.19b)}\\
&\boldsymbol{z}_k^T \mathcal{B}_1(\boldsymbol{w}_{[k]}) \geq c_0(k,i) + \boldsymbol{c}_x(k,i)^T \boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + \boldsymbol{c}_u(k,i)^T \boldsymbol{u}_k(\boldsymbol{w}_{[k]}), && \text{(3.19c)}\\
&\qquad \forall\, \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]},\ \forall\, i \in \mathcal{I}_k,\ \forall\, k \in \{0, \dots, T-1\}, &&\\
&\boldsymbol{z}_T^T \mathcal{B}_1(\boldsymbol{w}_{[T]}) \geq c_0(T,i) + \boldsymbol{c}_x(T,i)^T \boldsymbol{x}_T(\boldsymbol{w}_{[T]}), && \text{(3.19d)}\\
&\qquad \forall\, \boldsymbol{w}_{[T]} \in \mathcal{W}_{[T]},\ \forall\, i \in \mathcal{I}_T, &&\\
&\big(\ \boldsymbol{x}_{k+1}(\boldsymbol{w}_{[k+1]}) = A_k\, \boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + B_k\, \boldsymbol{u}_k(\boldsymbol{w}_{[k]}) + \boldsymbol{w}(k),\ \big) && \text{(3.19e)}\\
&\qquad \forall\, k \in \{0, \dots, T-1\}, &&\\
&\boldsymbol{f}_k \geq E_x(k)\, \boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + E_u(k)\, \boldsymbol{u}_k(\boldsymbol{w}_{[k]}), && \text{(3.19f)}\\
&\qquad \forall\, \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]},\ \forall k \in \{0, \dots, T-1\}, &&\\
&\boldsymbol{f}_T \geq E_x(T)\, \boldsymbol{x}_T(\boldsymbol{w}_{[T]}),\ \forall\, \boldsymbol{w}_{[T]} \in \mathcal{W}_{[T]}. && \text{(3.19g)}
\end{aligned}
$$

In this formulation, the decision variables are $\{L_k\}_{0 \leq k \leq T-1}, \{\boldsymbol{z}_k\}_{0 \leq k \leq T}$ and J, and equation (3.19e) should be interpreted as giving the dependency of $\boldsymbol{x}_k$ on $\boldsymbol{w}_{[k]}$ and the decision variables, which can then be used in the constraints (3.19c), (3.19d), (3.19f),

[9] This is the same approach as that taken in Ben-Tal et al. [19]; when the stage costs h_k are already affine in $\boldsymbol{x}_k, \boldsymbol{u}_k$, the step is obviously not necessary

and (3.19g). Note that, in the above optimization problem, all the constraints are bi-affine functions of the uncertainties and the decision variables, and thus, since the uncertainty sets $\mathcal{W}_{[k]}$ have tractable conic representations, the techniques in Ben-Tal et al. [19] can be used to compute the optimal decisions in (P_{AFF}).

Letting $J^{\star}_{\text{AFF}}$ denote the optimal value in (P_{AFF}), and with $J^{\star}_{d=r}$ representing the optimal value obtained from our polynomial hierarchy (with SOS constraints) for degree $d = r$, we have the following result.

Theorem 2. *If the uncertainty sets $\mathcal{W}_k$ are given by the intersection of finitely many convex quadratic forms, and have nonempty interior, then the objective functions obtained from the polynomial hierarchy satisfy the following relation*

$$J^{\star}_{AFF} \geq J^{\star}_{d=1} \geq J^{\star}_{d=2} \geq \dots$$

Proof. First, note that the hierarchy can only improve when the polynomial degree d is increased (this is because any feasible solutions for a particular degree d remain feasible for degree $d+1$). Therefore, we only need to prove the first inequality.

Consider any feasible solution to Problem (P_{AFF}) under disturbance-affine policies, i.e., any choice of matrices $\{L_k\}_{0 \leq k \leq T-1}$, coefficients $\{\boldsymbol{z}_k\}_{0 \leq k \leq T}$ and cost J, such that all constraints in (P_{AFF}) are satisfied.

Note that a typical constraint in Problem (P_{AFF}) becomes

$$f(\boldsymbol{w}_{[k]}) \geq 0, \, \forall \, \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]},$$

where f is a degree 1 polynomial in indeterminate $\boldsymbol{w}_{[k]}$, with coefficients that are affine functions of the decision variables. By the assumption in the statement of the theorem, the sets $\mathcal{W}_k$ are convex, with nonempty interior, $\forall \, k \in \{0, \dots, T-1\}$, which implies that $\mathcal{W}_{[k]} = \mathcal{W}_0 \times \dots \times \mathcal{W}_{k-1}$ is also convex, with non-empty interior.

Therefore, the typical constraint above can be written as

$$f(\boldsymbol{w}_{[k]}) \geq 0, \, \forall \, \boldsymbol{w}_{[k]} \in \left\{ \boldsymbol{\xi} \in \mathbb{R}^{k \times n_w} \, : \, g_j(\boldsymbol{\xi}) \geq 0, \, j \in \mathcal{J} \right\},$$

where $\mathcal{J}$ is a finite index set, and $g_j(\cdot)$ are convex. By the nonlinear Farkas Lemma (see, e.g., Proposition 3.5.4 in Bertsekas et al. [24]), there must exist multipliers $0 \leq \lambda_j \in \mathbb{R}, \forall\, j \in \mathcal{J}$, such that

$$f(\boldsymbol{w}_{[k]}) \geq \sum_{j \in \mathcal{J}} \lambda_j g_j(\boldsymbol{w}_{[k]}).$$

But then, recall that our SOS framework required the existence of polynomials $\sigma_j(\boldsymbol{w}_{[k]}), j \in \{0\} \cup \mathcal{J}$, such that

$$f(\boldsymbol{w}_{[k]}) = \sigma_0(\boldsymbol{w}_{[k]}) + \sum_{j \in \mathcal{J}} \sigma_j(\boldsymbol{w}_{[k]})\, g_j(\boldsymbol{w}_{[k]}).$$

By choosing $\sigma_j(\boldsymbol{w}_{[k]}) \equiv \lambda_j, \forall\, j \in \mathcal{J}$, and $\sigma_0(\boldsymbol{w}_{[k]}) = f(\boldsymbol{w}_{[k]}) - \sum_{j \in \mathcal{J}} \lambda_j g_j(\boldsymbol{w}_{[k]})$, we can immediately see that:

- $\forall\, j \neq 0, \sigma_j$ are SOS (they are positive constants)
- Since g_j are quadratic, and f is affine, σ_0 is a quadratic polynomial which is non-negative, for any $\boldsymbol{w}_{[k]}$. Therefore, since any such polynomial can be represented as a sum-of-squares (see Parrilo [113], Lasserre [94]), we also have that σ_0 is SOS.

By these two observations, we can conclude that the particular choice $L_k, \boldsymbol{z}_k, J$ will also remain feasible in our SOS framework applied to degree $d = 1$, and, hence, $J^{\star}_{\text{AFF}} \geq J^{\star}_{d=1}$. □

The above result suggests that the performance of our polynomial hierarchy can never be worse than that of the best affine policies.

For the same case of $\mathcal{W}_k$ given by intersection of convex quadratic forms, a popular technique introduced by Ben-Tal and Nemirovski in the robust optimization literature, and based on using the approximate S-Lemma, could be used for computing *quadratic* decision rules. More precisely, the resulting problem (P_{QUAD}) can be obtained from (P_{AFF}) by using $\boldsymbol{u}_k(\boldsymbol{x}_k) = L_k \cdot \mathcal{B}_2(\boldsymbol{w}_{[k]})$, and by replacing $\boldsymbol{z}_k^T \mathcal{B}_2(\boldsymbol{w}_{[k]})$ and $\boldsymbol{z}_T^T \mathcal{B}_2(\boldsymbol{w}_{[T]})$ in (3.19c) and (3.19d), respectively. Since all the constraints become quadratic polynomials in indeterminates $\boldsymbol{w}_{[k]}$, one can use the Approximate S-Lemma to enforce the resulting constraints (See Chapter 14 in Ben-Tal et al. [19] for details). If we let $J^{\star}_{\text{QUAD}}$ denote the optimal value resulting from this method, a proof paralleling that of Theorem 2 can be used to show that $J^{\star}_{\text{QUAD}} \geq J^{\star}_{d=2}$, i.e., the performance of the polynomial hierarchy for $d \geq 2$ cannot be worse than that delivered by the S-Lemma method.

In view of these results, one can think of the polynomial framework as a generalization of two classical methods in the literature, with the caveat that (for degree $d \geq 3$), the resulting SOS problems that need to be solved can be more computationally challenging.

3.4.2 Determining the Optimal Value for Polytopic Uncertainties

Here, we briefly discuss a specific class of Problems (P), for which the *exact* optimal value can be computed by solving a (large) mathematical program. This is particularly useful for benchmarking purposes, since it allows a precise assessment of the polynomial framework's performance (note that the approach presented in Section 3.3 is applicable to the general problem, described in the introduction).

Consider the particular case of polytopic uncertainty sets, i.e., when all the polynomial functions g_j in (3.2) are actually affine. It can be shown (see Theorem 2 in Bemporad et al. [9]) that piece-wise affine state-feedback policies[10] $\boldsymbol{u}_k(\boldsymbol{x}_k)$ are optimal for the resulting Problem (P), and that the sequence of uncertainties that achieves the min-max value is an extreme point of the uncertainty set, that is, $\boldsymbol{w}_{[T]} \in \text{ext}(\mathcal{W}_0) \times \cdots \times \text{ext}(\mathcal{W}_{T-1})$. As an immediate corollary of this result, the optimal value for Problem (P), as well as the optimal decision at time $k = 0$ for a fixed initial state $\boldsymbol{x}_0$, $\boldsymbol{u}_0^{\star}(\boldsymbol{x}_0)$, can be computed by solving the following optimization

[10] One could also immediately extend the result of Garstka and Wets [70] to argue that disturbance-feedback policies $\boldsymbol{u}_k(\boldsymbol{w}_{[k]})$ are also optimal.

problem (see Ben-Tal et al. [15], Bemporad et al. [8, 9] for a proof):

$$
(P)_{\text{ext}} \quad
\begin{aligned}
\min_{\boldsymbol{u}_k(\boldsymbol{w}_{[k]}),z_k(\boldsymbol{w}_{[k]}),J} \quad & J && \text{(3.20a)}\\
\text{s.t.} \quad & \forall\, t \in \{0,\dots,T-1\},\ \forall\, \boldsymbol{w}_{[k]} \in \text{ext}(\mathcal{W}_0)\times\cdots\times\text{ext}(\mathcal{W}_{k-1}), &&\\
& J \geq \sum_{k=0}^{T} z_k(\boldsymbol{w}_{[k]}), && \text{(3.20b)}\\
& z_k(\boldsymbol{w}_{[k]}) \geq h_k\big(\boldsymbol{x}_k(\boldsymbol{w}_{[k]}),\boldsymbol{u}_k(\boldsymbol{w}_{[k]})\big), && \text{(3.20c)}\\
& z_T(\boldsymbol{w}_{[T]}) \geq h_T\left(\boldsymbol{x}_T(\boldsymbol{w}_{[T]})\right), && \text{(3.20d)}\\
& \boldsymbol{x}_{k+1}(\boldsymbol{w}_{[k+1]}) = A_k\,\boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + B_k\,\boldsymbol{u}_k(\boldsymbol{w}_{[k]}) + \boldsymbol{w}(k), && \text{(3.20e)}\\
& \boldsymbol{f}_k \geq E_x(k)\,\boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + E_u(k)\,\boldsymbol{u}_k(\boldsymbol{w}_{[k]}), && \text{(3.20f)}\\
& \boldsymbol{f}_T \geq E_x(T)\,\boldsymbol{x}_T(\boldsymbol{w}_{[T]}). && \text{(3.20g)}
\end{aligned}
$$

In this formulation, *non-anticipatory* control values $\boldsymbol{u}_k(\boldsymbol{w}_{[k]})$ and corresponding states $\boldsymbol{x}_k(\boldsymbol{w}_{[k]})$ are computed for every vertex of the disturbance set, i.e., for every $\boldsymbol{w}_{[k]} \in \text{ext}(\mathcal{W}_0)\times\cdots\times\text{ext}(\mathcal{W}_{k-1})$, $k = 0,\dots,T-1$. The variables $z_k(\boldsymbol{w}_{[k]})$ are used to model the stage cost at time k, in scenario $\boldsymbol{w}_{[k]}$. Note that constraints (3.20c), (3.20d) can be immediately rewritten in linear form, since the functions $h_k(\boldsymbol{x},\boldsymbol{u})$, $h_T(\boldsymbol{x})$ are piece-wise affine and convex in their arguments.

We emphasize that the formulation does not seek to compute an actual *policy* $\boldsymbol{u}_k^\star(\boldsymbol{x}_k)$, but rather the values that this policy would take (and the associated states and costs), when the uncertainty realizations are restricted to extreme points of the uncertainty set. As such, the variables $\boldsymbol{u}_k(\boldsymbol{w}_{[k]}), \boldsymbol{x}_k(\boldsymbol{w}_{[k]})$ and $z_k(\boldsymbol{w}_{[k]})$ must also be forced to satisfy a *non-anticipativity* constraint[11], which is implicitly taken into account when only allowing them to depend on the portion of the extreme sequence available at time k, i.e., $\boldsymbol{w}_{[k]}$. Due to this coupling constraint, Problem $(P)_{\text{ext}}$ results in a Linear Program which is doubly-exponential in the horizon T, with the number of variables and the number of constraints both proportional to the number of extreme sequences in the uncertainty set, $\mathcal{O}\big(\prod_{k=0}^{T-1}|\text{ext}(\mathcal{W}_k)|\big)$. Therefore, solving $(P)_{\text{ext}}$ is

[11] In our current notation, non-anticipativity is equivalent to requiring that, for any two sequences $(\boldsymbol{w}_0,\dots,\boldsymbol{w}_{T-1})$ and $(\hat{\boldsymbol{w}}_0,\dots,\hat{\boldsymbol{w}}_{T-1})$ satisfying $\boldsymbol{w}_t = \hat{\boldsymbol{w}}_t, \forall\, t \in \{0,\dots,k-1\}$, we have $\boldsymbol{u}_t(\boldsymbol{w}_{[t]}) = \boldsymbol{u}_t(\hat{\boldsymbol{w}}_{[t]}), \forall\, t \in \{0,\dots,k\}$.

relevant only for small horizons, but is very useful for benchmarking purposes, since it provides the optimal value of the original problem.

We conclude this section by examining a particular example when the uncertainty sets take an even simpler form, and polynomial policies (3.11) are *provably* optimal. More precisely, we consider the case of scalar uncertainties ($n_w = 1$), and

$$\boldsymbol{w}(k) \in \mathcal{W}(k) \stackrel{\text{def}}{=} [\underline{w}_k, \overline{w}_k] \subset \mathbb{R}, \ \forall\, k = 0, \ldots, T-1, \tag{3.21}$$

which corresponds to the exact case of one-dimensional *box uncertainty* which we considered in Chapter 2. Under this model, any partial uncertain sequence $\boldsymbol{w}_{[k]}$ will be a k-dimensional vector, lying inside the hypercube $\mathcal{W}_{[k]} \subset \mathbb{R}^k$.

Introducing the subclass of *multi-affine* policies[12] of degree d, given by

$$u_j(k, \boldsymbol{w}_{[k]}) = \sum_{\boldsymbol{\alpha} \in \{0,1\}^k} \ell_{\boldsymbol{\alpha}} \, (\boldsymbol{w}_{[k]})^{\boldsymbol{\alpha}}, \text{ where } \sum_{i=1}^{k} \alpha_i \leq d, \tag{3.22}$$

one can show (see Theorem 3 in Appendix B) that multi-affine policies of degree $T-1$ are, in fact, optimal for Problem (P). While this theoretical result is of minor practical importance (due to the large degree needed for the policies, which translates into prohibitive computation), it provides motivation for restricting attention to polynomials of smaller degree, as a midway solution that preserves tractability, while delivering high quality objective values.

For completeness, we remark that, for the case of box-uncertainty, the authors in Ben-Tal et al. [19] show one can seek *separable* polynomial policies of the form

$$u_j(k, \boldsymbol{w}_{[k]}) = \sum_{i=1}^{k} p_i(w_i), \ \forall\, j \in \{1, \ldots, n_u\}, \ \forall\, k \in \{0, \ldots, T-1\},$$

where $p_i \in \mathcal{P}_d[x]$ are univariate polynomials in indeterminate x. The advantage of this approach is that the reformulation of a typical state-control constraint would be

[12] Note that these are simply polynomial policies of the form (3.11), involving only square-free monomials, i.e., every monomial, $\boldsymbol{w}_{[k]}^{\boldsymbol{\alpha}} \stackrel{\text{def}}{=} \prod_{i=0}^{k-1} w_i^{\alpha_i}$, satisfies the condition $\alpha_i \in \{0, 1\}$.

exact (refer to Lemma 14.3.4 in Ben-Tal et al. [19]). The main pitfall, however, is that, for the case of box-uncertainty, such a rule would never improve over purely affine rules, i.e., where all the polynomials p_i have degree 1 (refer to Lemma 14.3.6 in Ben-Tal et al. [19]). However, as we will see in our numerical results (to be presented in Section 3.6), polynomials policies that are *not* separable, i.e., are of the general form (3.11), can and do improve over the affine case.

3.5 Examples from Inventory Management

To test the performance of our proposed policies, we consider two problems arising in inventory management.

3.5.1 Single Echelon with Cumulative Order Constraints

Our first example corresponds to a slight generalization of the instance we considered in Chapter 2, namely the problem of negotiating flexible contracts between a retailer and a supplier in the presence of uncertain orders from customers, originally discussed in a robust framework by Ben-Tal et al. [16]. We describe the version of the problem here, and refer the interested reader to Ben-Tal et al. [16] for more details.

The setting is the following: consider a single-product, single-echelon, multi-period supply chain, in which inventories are managed periodically over a planning horizon of T periods. The unknown demands w_k from customers arrive at the (unique) echelon, henceforth referred to as the *retailer*, and are satisfied from the on-hand inventory, denoted by x_k at the beginning of period k. The retailer can replenish the inventory by placing orders u_k, at the beginning of each period k, for a cost of c_k per unit of product. These orders are immediately available, i.e., there is no lead-time in the system, but there are capacities on the order size in every period, $L_k \leq u_k \leq U_k$, as well as on the cumulative orders places in consecutive periods, $\hat{L}_k \leq \sum_{t=0}^{k} u_t \leq \hat{U}_k$. After the demand w_k is realized, the retailer incurs holding costs $H_{k+1} \cdot \max\{0, x_k + u_k - w_k\}$ for all the amounts of supply stored on her premises, as well as penalties $B_{k+1} \cdot \max\{w_k - x_k - u_k, 0\}$, for any demand that is backlogged.

In the spirit of robust optimization, we assume that the only information available about the demand at time k is that it resides within an interval centered around a *nominal* (mean) demand $\bar{d}_k$, which results in the uncertainty set $\mathcal{W}_k = \{w_k \in \mathbb{R} : \left|w_k - \bar{d}_k\right| \leq \rho \cdot \bar{d}_k\}$, where $\rho \in [0,1]$ can be interpreted as an *uncertainty level*.

With the objective function to be minimized as the cost resulting in the worst-case scenario, we immediately obtain an instance of our original Problem (P), i.e., a linear system with $n = 2$ states and $n_u = 1$ control, where $x_1(k)$ represents the on-hand inventory at the beginning of time k, and $x_2(k)$ denotes the total amount of orders placed in prior times, $x_2(k) = \sum_{t=0}^{k-1} u(t)$. The dynamics are specified by

$$
\begin{aligned}
x_1(k+1) &= x_1(k) + u(k) - w(k), \\
x_2(k+1) &= x_2(k) + u(k),
\end{aligned}
$$

with the constraints

$$
\begin{aligned}
L_k &\leq u(k) \leq U_k, \\
\hat{L}_k &\leq x_2(k) + u(k) \leq \hat{U}_k,
\end{aligned}
$$

and the costs

$$
\begin{aligned}
h_k(\boldsymbol{x}_k, u_k) &= \max\left\{c_k\, u_k + [H_k, 0]^T \boldsymbol{x}_k,\ c_k\, u_k + [-B_k,\, 0]^T \boldsymbol{x}_k\right\}, \\
h_T(\boldsymbol{x}_T) &= \max\left\{[H_T,\, 0]^T \boldsymbol{x}_T,\ [-B_T,\, 0]^T \boldsymbol{x}_T\right\}.
\end{aligned}
$$

We remark that the cumulative order constraints, $\hat{L}_k \leq \sum_{t=0}^{k} u_t \leq \hat{U}_k$, are needed here, since otherwise, the resulting (one-dimensional) system would fit the theoretical results from Bertsimas et al. [37], which would imply that polynomial policies of the form (3.11) and polynomial stage costs of the form (3.16b) are already optimal for degree $d = 1$ (affine). Therefore, testing for higher order polynomial policies would not add any benefit.

3.5.2 Serial Supply Chain

As a second problem, we consider a serial supply chain, in which there are J echelons, numbered $1, \ldots, J$, managed over a planning horizon of T periods by a centralized decision maker. The j-th echelon can hold inventory on its premises, for a per-unit cost of $H_j(k)$ in time period k. In every period, echelon 1 faces the unknown, external demands $w(k)$, which it must satisfy from the on-hand inventory. Unmet demands can be backlogged, incurring a particular per-unit cost, $B_1(k)$. The j-th echelon can replenish its on-hand inventory by placing orders with the immediate echelon in the upstream, $j+1$, for a per-unit cost of $c_j(k)$. For simplicity, we assume the orders are received with zero lead-time, and are only constrained to be non-negative, and we assume that the last echelon, J, can replenish inventory from a supplier with infinite capacity.

Following a standard requirement in inventory theory (Zipkin [150]), we maintain that, under centralized control, orders placed by echelon j at the beginning of period k cannot be backlogged at echelon $j+1$, and thus must always be sufficiently small to be satisfiable from on-hand inventory at the beginning[13] of period k at echelon $j+1$. As such, instead of referring to orders placed by echelon j to the upstream echelon $j+1$, we will refer to physical shipments from $j+1$ to j, in every period.

This problem can be immediately translated into the linear systems framework mentioned before, by introducing the following states, controls, and uncertainties:

- Let $x_j(k)$ denote the local inventory at stage j, at the beginning of period k.
- Let $u_j(k)$ denote the shipment sent in period k from echelon $j+1$ to echelon j.
- Let the unknown external demands arriving at echelon 1 represent the uncertainties, $w(k)$.

[13] This implies that the order placed by echelon j in period k (to the upstream echelon, $j+1$) cannot be used to satisfy the order in period k from the downstream echelon, $j-1$. Technically, this corresponds to an effective lead time of 1 period, and a more appropriate model would redefine the state vector accordingly. We have opted to keep our current formulation for simplicity.

The dynamics of the linear system can then be formulated as

$$
\begin{aligned}
&x_1(k+1) = x_1(k) + u_1(k) - w(k), && k = 0, \dots, T-1, \\
&x_j(k+1) = x_j(k) + u_j(k) - u_{j-1}(k), && j = 2, \dots, J,\ k = 0, \dots, T-1,
\end{aligned}
$$

with the following constraints on the states and controls

$$
\begin{aligned}
&u_j(k) \geq 0, && j = 1, \dots, J,\ k = 0, \dots, T-1, && \text{(non-negative shipments)} \\
&x_j(k) \geq u_{j-1}(k), && j = 2, \dots, J,\ k = 0, \dots, T-1, && \text{(downstream order} \\
&&&&& \leq \text{upstream inventory)}
\end{aligned}
$$

and the costs

$$
\begin{aligned}
h_1\big(k, x_1(k), u_1(k)\big) &= c_1(k) u_1(k) + \max\big\{H_1(k)\, x_1(k),\ -B_1(k)\, x_1(k)\big\},\ k = 0, \dots, T-1 \\
h_1\big(T, x_1(T)\big) &= \max\big\{H_1(T)\, x_1(T),\ -B_1(T)\, x_1(T)\big\}, \\
h_j\big(k, x_j(k), u_j(k)\big) &= c_j(k)\, u_j(k) + H_j(k)\, x_j(k),\ k = 0, \dots, T-1 \\
h_j\big(T, x_j(T)\big) &= H_j(T)\, x_j(T).
\end{aligned}
$$

With the same model of uncertainty as before, $\mathcal{W}_k = \big[\bar{d}_k(1-\rho),\, \bar{d}_k(1+\rho)\big]$, for some known mean demand $\bar{d}_k$ and uncertainty level $\rho \in [0,1]$, and the goal to decide shipment quantities $u_j(k)$ so as to minimize the cost in the worst-case scenario, we obtain a different example of Problem (P).

3.6 Numerical Experiments

In this section, we present numerical simulations testing the performance of polynomial policies in each of the two problems mentioned in Section 3.5. In order to examine the dependency of our results on the size of the problem, we proceed in the following fashion.

3.6.1 First Example

For the first model (single echelon with cumulative order constraints), we vary the horizon of the problem from $T = 4$ to $T = 10$, and for every value of T, we:

1. Create 100 problem instances, by randomly generating the cost parameters and the constraints, in which the performance of polynomial policies of degree 1 (affine) is suboptimal.

2. For every such instance, we compute:

    - The optimal cost OPT, by solving the exponential Linear Program $(P)_{\text{ext}}$.
    - The optimal cost $\bar{P}_d$ obtained with polynomial policies of degree $d = 1, 2,$ and 3, respectively, by solving the corresponding associated SDP formulations, as introduced in Section 3.3.

    We also record the relative optimality gap corresponding to each polynomial policy, defined as $(\bar{P}_d - OPT)/OPT$, and the solver time.

3. We compute statistics over the 100 different instances (recording the mean, standard deviation, min, max and median) for the optimality gaps and solver times corresponding to all three polynomial parameterizations.

Table 3.1 and Table 3.2 record these statistics for relative gaps and solver times, respectively. The following conclusions can be drawn from the results:

- Policies of higher degree decrease the performance gap considerably. In particular, while affine policies yield an average gap between 2.8% and 3.7% (with a median gap between 2% and 2.7%), quadratic policies reduce both average and median gaps by a factor of 3, and cubic policies essentially close the optimality gap (all gaps are smaller than 1%, with a median gap smaller than 0.01%). To better see this, Figure 3-1 illustrates the box-plots corresponding to the three policies for a typical case (here, $T = 6$).

- The reductions in the relative gaps are not very sensitive to the horizon, T. Figure 3-2(a) illustrates this effect for the case of quadratic policies, and similar plots can be drawn for the affine and cubic cases.

- The computational time grows polynomially with the horizon size. While computations for cubic policies are rather expensive, the quadratic case, shown in Figure 3-2(b), shows promise for scalability - for horizon $T = 10$, the median and average solver times are below 15 seconds.

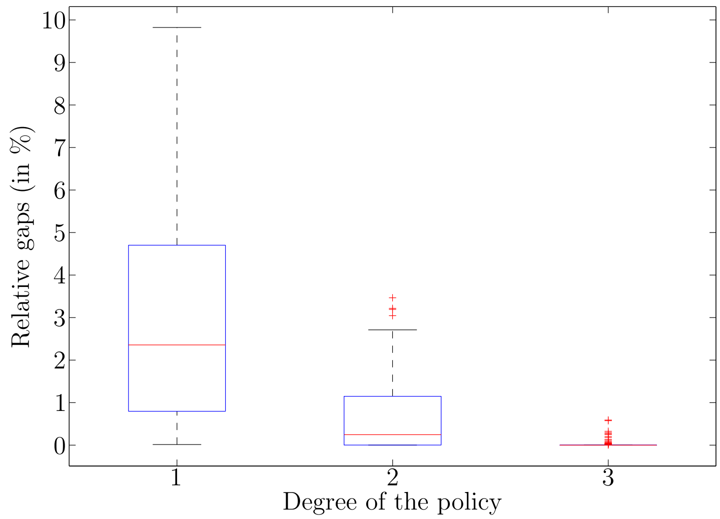
3.6.2 Second Example

For the second model (serial supply chain), we fix the problem horizon to $T = 7$, and vary the number of echelons from $J = 2$ to $J = 5$. For every resulting size, we go through the same steps 1-3 as outlined above, and record the same statistics, displayed in Table 3.3 and Table 3.4, respectively. Essentially the same observations as before hold. Namely, policies of higher degree result in strict improvements of the objective function, with cubic policies always resulting in gaps smaller than 1% (see Figure 3-3(a) for a typical case). Also, increasing the problem size (here, this corresponds to the number of echelons, J) does not affect the reductions in gaps, and the computational requirements do not increase drastically (see Figure 3-3(b), which corresponds to quadratic policies).

All our computations were done in a MATLAB® environment, on the MIT Operations Research Center computational machine (3 GHz Intel® Dual Core Xeon® 5050 Processor, with 8GB of RAM memory, running Ubuntu Linux). The optimization problems were formulated using YALMIP (Löfberg [100]), and the resulting SDPs were solved with SDPT3 (Toh et al. [141]).

We remark that the computational times could be substantially reduced by exploiting the structure of the polynomial optimization problems (e.g., Nie [110]), and by utilizing more suitable techniques for solving smooth large-scale SDPs (see, e.g., Lan et al. [92] and the references therein). Such techniques are immediately applicable to our setting, and could provide a large speed-up over general-purpose algorithms

Figure 3-1: Box plots comparing the performance of different polynomial policies for horizon $T = 6$



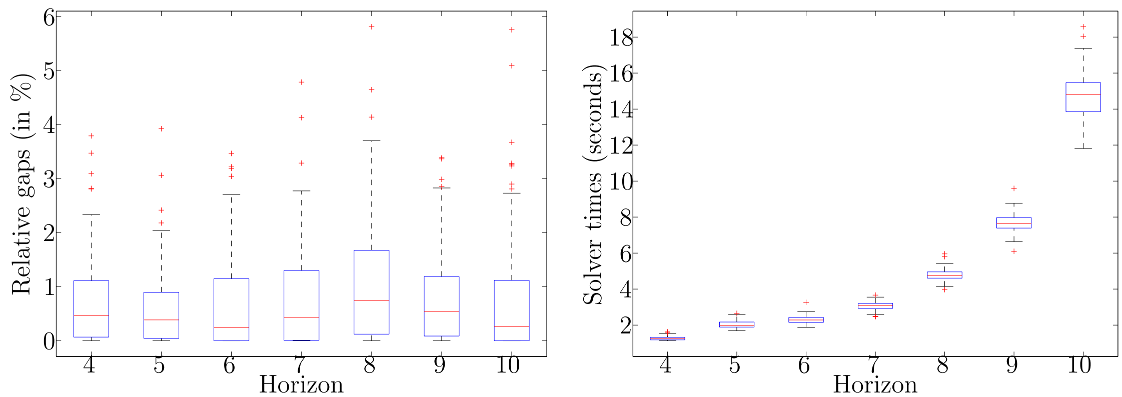
(such as the interior point methods implemented in SDPT3), hence allowing much larger and more complicated instances to be solved.

Table 3.1: Relative gaps (in %) for polynomial policies in Example 1

	Degree $d = 1$					Degree $d = 2$					Degree $d = 3$				
T	**avg**	**std**	**mdn**	**min**	**max**	**avg**	**std**	**mdn**	**min**	**max**	**avg**	**std**	**mdn**	**min**	**max**
4	2.84	2.41	2.18	0.02	9.76	0.75	0.85	0.47	0.00	3.79	0.03	0.12	0.00	0.00	0.91
5	2.82	2.29	2.52	0.04	11.22	0.62	0.71	0.39	0.00	3.92	0.02	0.09	0.00	0.00	0.56
6	3.09	2.63	2.36	0.01	9.82	0.69	0.89	0.25	0.00	3.47	0.03	0.10	0.00	0.00	0.59
7	3.25	2.95	2.58	0.13	15.00	0.83	0.99	0.43	0.00	4.79	0.06	0.17	0.00	0.00	0.93
8	3.66	3.29	2.69	0.03	18.36	1.06	1.17	0.74	0.00	5.81	0.10	0.17	0.00	0.00	0.99
9	2.93	2.78	2.12	0.05	11.56	0.80	0.86	0.55	0.00	3.39	0.07	0.13	0.00	0.00	0.61
10	3.44	3.60	2.09	0.00	18.20	0.76	1.16	0.26	0.00	5.76	0.05	0.12	0.00	0.00	0.74

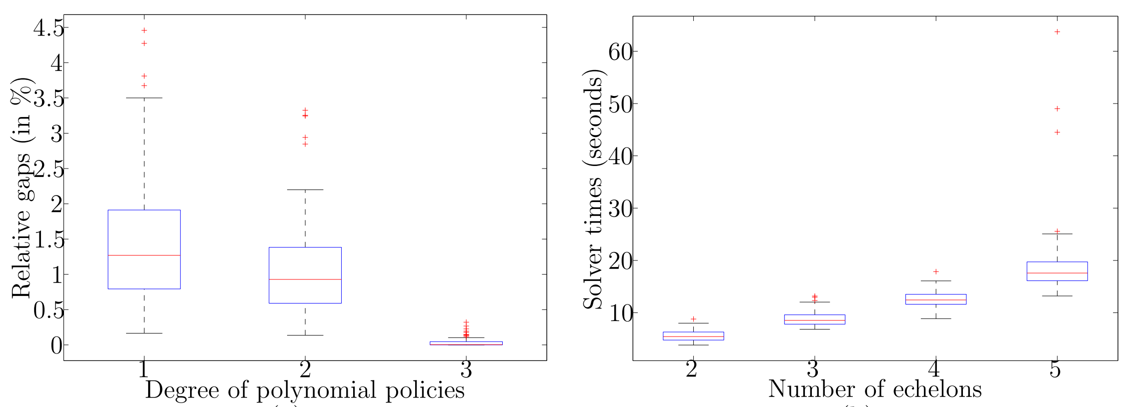
Table 3.2: Solver times (in seconds) for polynomial policies in Example 1

	Degree $d = 1$					Degree $d = 2$					Degree $d = 3$				
T	**avg**	**std**	**mdn**	**min**	**max**	**avg**	**std**	**mdn**	**min**	**max**	**avg**	**std**	**mdn**	**min**	**max**
4	0.47	0.05	0.46	0.38	0.63	1.27	0.10	1.27	1.13	1.62	3.33	0.21	3.24	3.01	4.03
5	0.58	0.06	0.58	0.46	0.75	2.03	0.20	1.97	1.69	2.65	7.51	0.91	7.27	6.58	12.08
6	0.73	0.11	0.72	0.62	1.50	2.29	0.22	2.28	1.87	3.26	18.96	2.54	18.25	16.07	31.86
7	0.88	0.08	0.87	0.72	1.07	3.08	0.23	3.10	2.47	3.67	48.83	5.63	47.99	40.65	74.09
8	1.13	0.12	1.11	0.94	1.92	4.79	0.32	4.75	3.97	5.96	157.73	20.67	153.91	126.15	217.80
9	1.53	0.17	1.51	1.27	2.66	7.65	0.51	7.65	6.10	9.59	420.75	60.10	411.09	334.71	760.13
10	1.31	0.15	1.30	1.07	2.19	14.77	1.24	14.80	11.81	18.57	1846.94	600.89	1640.10	1313.18	4547.09



(a) (b)

Figure 3-2: Performance of quadratic policies for Example 1 - (a) illustrates the weak dependency of the improvement on the problem size (measured in terms of the horizon T), while (b) compares the solver times required for different problem sizes.



(a) (b)

Figure 3-3: Performance of polynomial policies for Example 2. (a) compares the three policies for problems with $J = 3$ echelons, and (b) shows the solver times needed to compute quadratic policies for different problem sizes.

Table 3.3: Relative gaps (in %) for polynomial policies in Example 2

	Degree $d = 1$					Degree $d = 2$					Degree $d = 3$				
J	**avg**	**std**	**mdn**	**min**	**max**	**avg**	**std**	**mdn**	**min**	**max**	**avg**	**std**	**mdn**	**min**	**max**
2	1.87	1.48	1.47	0.00	8.27	1.38	1.16	1.11	0.00	6.48	0.06	0.14	0.01	0.00	0.96
3	1.47	0.89	1.27	0.16	4.46	1.08	0.68	0.93	0.14	3.33	0.04	0.06	0.00	0.00	0.32
4	1.14	2.46	0.70	0.05	24.63	0.67	0.53	0.53	0.01	2.10	0.04	0.07	0.00	0.00	0.38
5	0.35	0.37	0.21	0.03	1.85	0.27	0.32	0.15	0.00	1.59	0.02	0.03	0.00	0.00	0.15

Table 3.4: Solver times (in seconds) for polynomial policies Example 2

	Degree $d = 1$					Degree $d = 2$					Degree $d = 3$				
J	**avg**	**std**	**mdn**	**min**	**max**	**avg**	**std**	**mdn**	**min**	**max**	**avg**	**std**	**mdn**	**min**	**max**
2	1.22	0.20	1.18	0.86	2.35	5.58	1.05	5.44	3.82	8.79	81.64	14.02	80.88	52.55	116.56
3	1.72	0.26	1.70	1.21	3.09	8.84	1.40	8.53	6.83	13.19	115.08	20.91	109.96	77.29	183.84
4	1.57	0.22	1.55	1.20	2.85	12.59	1.63	12.44	8.86	17.86	160.05	19.34	159.29	82.11	207.56
5	2.59	1.46	1.97	1.51	8.18	18.97	6.59	17.59	13.21	63.71	250.43	109.96	227.56	144.54	952.37

Chapter 4

Polynomial Policies for Multi-Item Dynamic Pricing

4.1 Introduction

In the final chapter of the thesis, we examine a different variation of a multi-period decision problem under uncertainty, arising in the field of revenue management (RM). More precisely, we consider a setting in which a single firm (a monopolist) is selling a set of nonperishable products to an incoming stream of (non-strategic) customers, and is seeking a pricing policy that would maximize its revenue over a finite selling season.

Variations of this problem have received attention from numerous research groups in the dynamic pricing and RM community (the interested reader is referred to the books Talluri and van Ryzin [138] and Phillips [115], and the review papers Bitran and Mondschein [42] and Elmaghraby and Keskinocak [63] for an in-depth overview of the field). Most models typically assume that the unknown quantities affecting the system can be characterized completely through a probability distribution function. While this paradigm is certainly justified in stationary environments with abundant historical information, it does not fit several interesting situations arising naturally in RM, such as the introduction of new products or the changing dynamics of existing markets (e.g., due to the entry of a new competitor or some unforeseen event strongly affecting supply and/or demand).

In recognition of this shortfall, several recent papers have considered formulations and models that avoid specifying complete distributional information for the unknown parameters. One such approach is to consider robust optimization formulations, based on either maximizing the minimum possible revenue (Thiele [139], Adida and Perakis [1], Lim and Shanthikumar [95], Thiele [140], Lobel and Perakis [97]), minimizing the worst-case regret (Perakis and Roels [114], Lobel and Perakis [97]), or maximizing the competitive ratio (Lan et al. [93]). An alternative approach, which has been popular in several recent papers (Lobo and Boyd [98], Bertsimas and Perakis [28], Aviv and Pazgal [4, 5], Lin [96], Araman and Caldentey [2], Kachani et al. [85], Cope [54], Besbes and Zeevi [38], Farias and Van Roy [67], Besbes and Zeevi [39], Cooper et al. [53]) is to attempt to *learn* the unknown parameters in the model from realized sales, by performing suitable updates (Bayesian or otherwise). Yet another approach is to simply resort to completely non-parametric formulations, which make direct use of data, and are hence inherently "distribution-free" (see, e.g., Kleinberg and Leighton [89], Rusmevichientong et al. [128]).

The formulation we pursue in the current chapter is mostly in line with the first set of approaches above. Namely, we consider a setting where the demand model belongs to a known parametric class, and the goal is to estimate the correct parameters, while computing pricing policies so as to maximize revenues. Moreover, we focus on a situation where the estimation and the optimization stages are segregated, in that one first uses historical data to estimate the model, and then proceeds to solve the ensuing optimization problem, based on the estimated model. However, instead of believing that the constructed model is correct, we take the pragmatic view of *knowingly accepting* that it is most likely incorrect, and thus focus on robust(ified) formulations, which account for potential mis-specification. As a final ingredient in our approach, we also recognize that static decisions (i.e., open-loop controls) intrinsically miss several key dynamic features of the problem, and hence, we focus on formulations that allow the computation of *adjustable* policies. Our contributions in the current chapter are as follows:

- We consider a multi-period, multi-item dynamic pricing problem under a linear demand function, with additive uncertainties. For such a model, we propose distributionally robust formulations, in which the uncertainties are characterized by support and limited moment information, and argue how the ensuing models can be constructed and calibrated from limited historical data.

- For the resulting dynamic optimization model, we consider policies that depend polynomially on the observed model disturbances, and show how the policy parameters can be computed by solving tractable optimization problems (second-order conic or semidefinite programs).

- We present extensive computational results, based on both simulated and real data from a large US retailer. These show that robust policies with minimal degree of adjustability (e.g., affine policies) already improve considerably over open-loop robust policies, and are competitive with popular heuristics in the literature.

The chapter is organized as follows. Section 4.2 introduces the complete model description, and briefly discusses relevant results in the RM literature. Section 4.3 introduces the polynomial policies, presents tractable reformulations based on semidefinite relaxations and sums-of-squares techniques, and discusses alternative heuristic methods for solving the problem. Section 4.4 discusses possible model extensions. Section 4.5 introduces our data-set, and elaborates on several aspects related to estimation using the real data, and Section 4.6 presents the numerical results obtained using both simulated and real data.

4.1.1 Notation

Similar to Talluri and van Ryzin [138], we denote the price vector at time t by $\boldsymbol{p}_t \in \mathbb{R}^n_+$. For a given price vector $\boldsymbol{p}_t$, we let $\boldsymbol{d}_t(\boldsymbol{p}_t) \in \mathbb{R}^n_+$ denote the deterministic part of the demand function at time t, and $D_t\big(\boldsymbol{d}_t(\boldsymbol{p}_t), \boldsymbol{\varepsilon}_t\big)$ the unknown (realized) demand, which also depends on an unknown component $\boldsymbol{\varepsilon}$. For a vector $\boldsymbol{x}_t \in \mathbb{R}^n$, we use $x_{i,t}$ to denote the i-th component. We also let $\mathbf{1}$ denote the vector of all ones.

Throughout the chapter, since much of the exposition is centered on polynomial policies, an identical notation to that introduced in Chapter 3 will be in place. In particular, since we work extensively with quantities which depend on the entire history of available information at a given time t, we define, for any time-varying vector quantity $\{\boldsymbol{x}_t \in \mathbb{R}^n\}_{t=1,\dots,T}$, the following stacked vector $\boldsymbol{x}_{[t]} \stackrel{\text{def}}{=} (\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_{t-1}) \in \mathbb{R}^{n\times(t-1)}$, which represents measurements available at the beginning of period t. Similarly, if $\boldsymbol{x}_t \in \mathcal{X}_t, \forall\, t$, we define $\mathcal{X}_{[t]} \stackrel{\text{def}}{=} \mathcal{X}_1 \times \cdots \times \mathcal{X}_{t-1}$ as the cartesian-product support of the quantity $\boldsymbol{x}_{[t]}$.

For a vector $\boldsymbol{x} \in \mathbb{R}^n$, we use $\text{diag}(\boldsymbol{x})$ to denote the $n \times n$ matrix that has $\boldsymbol{x}$ on the main diagonal and zeros everywhere else.

4.2 Model Description

We consider a setting in which a single firm (a monopolist) is selling a set of n nonperishable products, denoted by $i \in \mathcal{I} \stackrel{\text{def}}{=} \{1, \dots, n\}$, over a finite planning horizon, $t \in \mathcal{T} \stackrel{\text{def}}{=} \{1, \dots, T\}$. The initial inventory in each product (i.e., the capacity) is denoted by C_i. In every period, the firm is selecting the prices for all products, $\boldsymbol{p}_t \stackrel{\text{def}}{=} (p_{1,t}, \dots, p_{n,t})$, subject to certain constraints, $\boldsymbol{p}_{[t+1]} \stackrel{\text{def}}{=} (\boldsymbol{p}_1, \dots, \boldsymbol{p}_t) \in \Omega^p_{t+1}$, where the set $\Omega^p_{t+1} \subseteq \mathbb{R}^{n\times t}$ is assumed to be polyhedral. Such constraints could include price non-negativity, as well as *mark-down* (i.e., $\boldsymbol{p}_t \leq \boldsymbol{p}_{t-1}$) or *mark-up* (i.e., $\boldsymbol{p}_t \geq \boldsymbol{p}_{t-1}$) constraints.

After setting the prices, the firm observes the resulting customer demand, $\boldsymbol{D}_t \stackrel{\text{def}}{=} (D_{1,t}, \dots, D_{n,t})$, which is influenced by the prices, as well as by unknown external factors $\boldsymbol{\varepsilon}_t$. We assume that the customers are non-strategic, and also that backlogging of demand is possible, at no cost to the firm, but that any remaining backlog *must* be satisfied by placing a constrained order $\boldsymbol{u} = (u_1, \dots, u_n)$, in period T, at a cost of $r_i \in \bar{\mathbb{R}}^+$ per unit of item i. We assume that the order can be decided in the last period (i.e., there is no requirement for a pre-commitment), but the order is constrained, $\boldsymbol{u} \in \Omega^u \subseteq \mathbb{R}^+$, where Ω^u is a polyhedral set (containing, e.g., non-negativity or capacity constraints).

The problem that the firm would like to solve is to find a sequence of prices and a final-period order so as to maximize its revenue (net of reordering cost) collected from an unknown stream of customers.

Clearly, in order to complete the description of the model, we must further specify two key ingredients:

1. The functional form for the demand $D_{i,t}$ - in particular, how the price vector $\boldsymbol{p}_t$ influences the customer demand for the different items $i \in \mathcal{I}$.

2. The exact way in which the firm is quantifying its preference over uncertain outcomes - recall that the demand depends on a set of unknown factors $\boldsymbol{\varepsilon}_t$, hence a recipe must be prescribed for measuring all the uncertain quantities (revenue stream, realized sales, etc.)

In the next sections, we discuss in detail each of these two aspects. Since the choices involved with the former will influence our modelling decisions related to the latter, we begin by describing the demand models.

4.2.1 Demand Model

While several choices of demand models are possible (see Chapter 7 of Talluri and van Ryzin [138] for more details and examples), we restrict attention to one of the most popular options in the RM literature, namely the linear demand model under additive noise (for extensions to other relevant demand models, we refer the reader to Section 4.4). This model is characterized by:

$$\boldsymbol{d}_t(\boldsymbol{p_t}) = \boldsymbol{b}_t + A_t\, \boldsymbol{p}_t, \tag{4.1a}$$

$$\boldsymbol{D}_t(\boldsymbol{d}_t, \boldsymbol{\varepsilon}_t) = \boldsymbol{d}_t + \boldsymbol{\varepsilon}_t, \tag{4.1b}$$

where the terms have the following significance:

- $\boldsymbol{p}_t$ is the price vector at the beginning of period t

- $\boldsymbol{d}_t$ is the *planned demand*, i.e., the deterministic component of the demand function, dependent only on the price vector $\boldsymbol{p}_t$

- $\boldsymbol{b}_t \in \mathbb{R}^n$ represents a base demand in period t

- $A_t \in \mathbb{R}^{n \times n}$ represents a matrix of price sensitivity coefficients in period t

- $\boldsymbol{D}_t$ is the realized demand in period t

- $\boldsymbol{\varepsilon}_t$ is an exogenous noise.

In particular, the functional form $\boldsymbol{d}_t$ is *linear* in the prices, and the noise affects the demand in an *additive* fashion. This model is quite popular due to its simplicity, and the ease of estimation from data. As such, it has been used extensively in both theoretical, as well as experimental studies (see Talluri and van Ryzin [138] and Mas-Colell et al. [104]). Standard assumptions on the matrices A_t include the following:

Assumption 4. *The diagonal coefficients of A_t are non-positive, i.e., $a_{ii} \leq 0, \forall\, t \in \mathcal{T}$.*

This assumption is a fundamental law in economics, and reflects the fact that decreasing the price of a given product makes it more attractive to the customers. Items not satisfying this requirement (known as Veblen or Giffen goods Mas-Colell et al. [104]) are usually ignored in the revenue management literature.

Assumption 5. *The matrices A_t are strictly row-diagonally dominant, i.e., $|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \forall\, i \in \mathcal{I}$.*

The latter fact states that the demand for a product i is more sensitive to changes in its own price, rather than simultaneous changes in the prices of other products. Alternatively, one sometimes requires that

Assumption 6. *The matrices A_t are strictly column diagonally-dominant, i.e., $|a_{jj}| > \sum_{i \neq j} |a_{ij}|, \forall\, j \in \mathcal{I}$.*

This would reflect that changes in the price of one product impacts the demand of that product more than the total demand of other products combined. All of the above assumptions have well grounded economic justifications (Mas-Colell et al. [104]), and have been widely adopted in the operations management literature. In can be seen by standard facts in linear algebra (see Theorem 5.6.17 in Horn and Johnson [79] and Chapter 2 in Horn and Johnson [80]) that the first assumption corroborated with any of the latter two ensure that:

1. The matrices A_t are invertible. This is convenient, since it allows inverting the price-demand relation to obtain a specific price $\boldsymbol{p}_t$ that would generate a particular demand $\boldsymbol{d}_t$. In this sense, we can equivalently think of the decisions as being the demands $\boldsymbol{d}_t$, rather than the prices.

2. The eigenvalues of A_t all have non-positive real parts. Since the A_t matrices are also usually taken to be symmetric, this latter fact has the direct implication that the revenue function, $r(\boldsymbol{p}_t) \stackrel{\text{def}}{=} \boldsymbol{p}_t' \, \boldsymbol{d}_t(\boldsymbol{p}_t)$, is concave in the prices, which ensures the existence of a unique revenue-maximizing price (see Chapter 7 of Talluri and van Ryzin [138]).

Despite these attactive theoretical properties, the model does suffer from several pitfalls. On a theoretical level, it requires bounding the range of feasible prices in order to ensure the demand is non-negative (e.g., in a single product case, we would need $p_t \leq -a_t/b_t$). This also implies that the model violates another typical requirement in the OM literature, namely that the range of the revenue function $r(\boldsymbol{p}_t)$ span the entire positive half-line (refer to Section 7.3 of Talluri and van Ryzin [138] for more details). For recent work that provides a natural extension of the linear demand model which does not suffer from some of these shortcomings, we refer the reader to Farahat and Perakis [66].

On a practical level, several empirical studies (e.g., Smith and Achabal [135]), as well as several patent filings (Woo et al. [147], Boyd et al. [44]) have found the model to under-perform other functional forms, such as exponential or power sensitivity. However, despite these shortcomings, due to widespread use of the model in both theory and practice (see, e.g., Heching et al. [78], Bertsimas and Perakis [28], Maglaras and Meissner [101], Adida and Perakis [1], Thiele [140] and references therein), we have chosen it as the main object of study in the current chapter.

In terms of the estimation requirements, since the functional form of the demand is linear (4.1a), and the noise affects the model in an additive fashion (4.1b), one can use ordinary least-squares regression techniques (OLS) (Greene [77]) to estimate the parameters of the model. More precisely, with dependent variables $y_{it} = D_{it}, \forall\, i \in \mathcal{I}, \forall\, t \in \{2, \ldots, T\}$, and with independent variables $x_{it} = \{p_{it}, \delta_t\}$ (where δ_t is an indicator for period t, with $t \in \{1, \ldots, T-1\}$), one can compute estimates $\hat{A}_t, \hat{\boldsymbol{b}}_t$, as well as associated confidence intervals.

The key underlying assumptions supporting the use of OLS techniques (see Chapter 2 of Greene [77]) are the standard Gauss-Markov requirements, namely

(i) The linearity of the functional form (i.e., equation (4.1a), in our case)

(ii) The full rank assumption on the data matrix X (consisting of the $p_{i,t}, \delta_t$ variables)

(iii) Exogeneity of the independent variables, i.e., $\mathbb{E}\left[\varepsilon_{i,t} \,|\, x_{jt}\right] = 0,\ \forall\, i, j \in \mathcal{I},\ \forall\, t \in \mathcal{T}$. In words, the expected value of the disturbance corresponding to a particular observation should not be a function of the independent variables x_{it} corresponding to any observation (including the current one).

(iv) Homoscedasticity and nonautocorrelation, i.e., the disturbances ε_{it} should have the same finite variance and be uncorrelated across i and t. More precisely, $\text{var}[\varepsilon_{it} \,|\, X] = 0, \forall\, i \in \mathcal{I},\ \forall\, t \in \mathcal{T}$, $\text{cov}[\varepsilon_{it}\, \varepsilon_{j\tau} \,|\, X] = 0,\ \forall\, i, j \in \mathcal{I},\ \forall\, t, \tau \in \mathcal{T}$ with $(it) \neq (j\tau)$.

(v) Normality, i.e., that the disturbances ε_{it} follow a Gaussian distribution.

Since, in reality, several of these assumptions are violated, procedures have been designed to test for mis-specifications, and several extensions of the regression techniques are available for more general cases (see Greene [77] for a complete account and more references). In Section 4.6.3, we revisit some of these issues in the specific context of estimating the linear demand model (4.1a), and we also discuss several aspects related to our own data-set.

4.2.2 Model Uncertainties and Preferences Over Uncertain Outcomes

From our earlier discussion, it is evident that there are several potential sources of disturbances affecting our model. In particular, apart from the noise ε_{it} reflected in (4.1b), one might also be introducing errors through the estimation procedure itself. For instance, it is likely that the true form of the demand function is not linear, and that the disturbances ε_{it} affecting the model are both heteroscedastic and autocorrelated (the latter is a particularly common phenomenon when dealing with panel data Greene [77], such as bulk transactional data from an RM system), resulting in potentially systematic mis-specifications of the model.

At the same time, in many practical settings (including the one we face with our data-set), records are affected by mistakes, as well as scarce (e.g., few seasons available, preventing an adequate estimation of the non-stationary factors). Such issues not only affect the quality of the regressions, but also prevent one from performing adequate tests for violations of the standard assumptions (Greene [77]) or constructing adequate distributions for stochastic quantities.

In particular, we maintain that a much more sensible requirement would be to estimate the support and moments of stochastic quantities, instead of complete distributions. Bearing the above issues in mind, we model the uncertain quantities using a *distributionally robust* framework. More precisely, we assume that $\{\boldsymbol{\varepsilon}_t\}_{t\in\mathcal{T}}$ represents a stochastic process defined on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, but that the measure $\mathbb{P}$ is not completely specified. Rather, the only information available is that the measure $\mathbb{P}$ belongs to the class of all measures $\mathcal{P}$ which are characterized by the following partial information:

- All measures in $\mathcal{P}$ are supported on the set $\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_T$, where $\mathcal{E}_t \subseteq \mathbb{R}^n$

is a closed, basic semialgebraic set, i.e.,

$$\mathcal{E}_t \overset{\text{def}}{=} \{\boldsymbol{\varepsilon}_t \in \mathbb{R}^n : g_j(\boldsymbol{\varepsilon}_t) \geq 0,\, j \in \mathcal{J}_t\}\,, \tag{4.2}$$

where $g_j \in \mathbb{R}[\boldsymbol{\varepsilon}_t]$ are multivariate polynomials depending on the disturbances at time t, and $\mathcal{J}_t$ is a finite index set.

- All measures in $\mathcal{P}$ have a given set of moments, up to a specified degree.

4.2.3 Complete Formulation

We now return to the original problem formulation, which we describe in detail. The goal of the firm is to choose a sequence of pricing policies $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_T$ and a last-period order $\boldsymbol{u}$, so as to maximize the worst-case net expected revenue, that is

$$\max_{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_T,\boldsymbol{u}} \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\boldsymbol{\varepsilon}_{[T+1]}\sim\mathbb{P}} \left[\sum_{t=1}^{T} \boldsymbol{p}_t' \,\boldsymbol{D}_t(\boldsymbol{p}_t, \boldsymbol{\varepsilon}_t) - \boldsymbol{r}'\,\boldsymbol{u} \right] \tag{4.3a}$$

$$(P) \quad \text{such that} \quad \sum_{t=1}^{T} \boldsymbol{D}_t(\boldsymbol{p}_t, \boldsymbol{\varepsilon}_t) \leq \boldsymbol{C} + \boldsymbol{u} \tag{4.3b}$$

$$\boldsymbol{p}_{[t]} \in \Omega_t^p, \qquad \forall\, t \in \{2, \ldots, T+1\}. \tag{4.3c}$$

$$\boldsymbol{u} \in \Omega^u. \tag{4.3d}$$

In the above formulation, the inner (minimization) corresponds to the problem solved by nature, which chooses the worst possible measure in the set $\mathcal{P}$ for the uncertain quantities $\boldsymbol{\varepsilon}_{[T+1]} \overset{\text{def}}{=} (\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_T)$. The outer maximization corresponds to the problem the firm is seeking to solve, namely choosing the prices so as to maximize its expected revenue $\sum_{t=1}^{T} \boldsymbol{p}_t' \,\boldsymbol{D}_t(\boldsymbol{p}_t, \boldsymbol{\varepsilon}_t)$, net of the reordering cost $\boldsymbol{r}'\,\boldsymbol{u}$. The constraint (4.3b) reflects that fact that sales should not exceed available capacity, while (4.3c) and (4.3d) capture the constraints that the planned prices and the planned order quantity should obey, respectively.

In analogy to our exposition in Chapter 3, we make the following two additional remarks:

- The price chosen in period t, $\boldsymbol{p}_t$, is allowed to depend on any information that is available at the beginning of period t. Put formally, the firm is free to seek pricing policies $\boldsymbol{p}_t$ that are adapted to the filtration induced by $\boldsymbol{\varepsilon}_{[t]}$.

- We ask that any constraints involving stochastic quantities should be obeyed almost surely, i.e., for any possible realization of the uncertain quantities.

As stated, our model falls in the large class of mini-max stochastic programming formulations, pioneered by Žáčová [144] and Dupačová [58, 59]. Such models have seen renewed interest in several recent papers (Popescu [118], Natarajan et al. [106], Delage and Ye [55], Sim and Goh [131], Bertsimas et al. [36]), in which tractable reformulations and computational aspects are discussed, typically in the context of two-stage problems, with more general objectives (concave utility functions). However, in the level of generality considered in Problem (P), these models are typically severely computationally intractable, hence the usual approach is to look for approximate solutions, most often by restricting attention to specific classes of policies (Shapiro et al. [130]).

We remark that, when the only information about the measures in the set $\mathcal{P}$ is the support, the distributionally robust model above becomes equivalent to the robust optimization models, which we have extensively discussed in Chapter 2 and Chapter 3. Similar models have been considered recently in the RM literature, and are gaining increased attention due to their advantageous computational properties. One of the initial papers to make use of such formulations is Adida and Perakis [1], which considers a firm pricing several products that utilize a common production capacity, and in which both ordering and pricing decisions are possible in every period. For a linear demand model without cross-item price effects (i.e., a diagonal A_t), the authors compare different robust formulations (affine adjustable and open loop) with closed-loop (dynamic programming) solutions, and conclude that the robust models perform well, while remaining tractable. Thiele [139] and Thiele [140] also considers robust models under open-loop (i.e., non-adjustable) policies, discusses managerial insights of the robust formulation for the single-item case, and presents computational results for the multi-product case under linear demand models. Perakis and Roels [114] also considers robust (maximin or minimax-regret) formulations in network RM problems, and find that open-loop minimax-regret controls perform very well on average, despite their worst-case focus, and outperform traditional controls when demand is censored. The recent paper Lobel and Perakis [97] employs sampling-based techniques in the context of multi-period network revenue management, and computes affinely-adjustable policies which deliver excellent empirical performance when compared with heuristic policies.

4.3 Polynomial Policies and Tractable Robust Reformulations

As a natural follow-up to the approach introduced in Chapter 3, we consider policies for both pricing and reordering that are adjustable in the sequence of observed model disturbances, i.e.,

$$\begin{aligned} \boldsymbol{p}_t &: \mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_{t-1} \to \mathbb{R}^n, \forall\, t \in \mathcal{T}, \\ \boldsymbol{u} &: \mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_T \to \mathbb{R}^n. \end{aligned}$$

Furthermore, we restrict attention to policies in which every component is a *polynomial* function of the history. That is, for a fixed degree d, we seek $\boldsymbol{p}_{it}(\boldsymbol{\varepsilon}_{[t]}) \in \mathcal{P}_d[\boldsymbol{\varepsilon}_{[t]}]$ and $u_i(\boldsymbol{\varepsilon}_{[T+1]}) \in \mathcal{P}_d[\boldsymbol{\varepsilon}_{[T+1]}]$, $\forall\, t \in \mathcal{T}$, $\forall\, i \in \mathcal{I}$, i.e.,

$$\begin{aligned} \boldsymbol{p}_t(\boldsymbol{\varepsilon}_{[t]}) &= L_t\, \mathcal{B}_d(\boldsymbol{\varepsilon}_{[t]}), \\ \boldsymbol{u}(\boldsymbol{\varepsilon}_{[t]}) &= U\, \mathcal{B}_d(\boldsymbol{\varepsilon}_{[T+1]}), \end{aligned} \tag{4.4}$$

where $\mathcal{B}_d(\boldsymbol{\varepsilon}_{[t]})$ is the canonical basis of $\mathcal{P}_d[\boldsymbol{\varepsilon}_{[k]}]$. The new decision variables become the matrices of coefficients $L_t \in \mathbb{R}^{n \cdot s(t,d)}$, $t = 1, \ldots, T$, and $U \in \mathbb{R}^{n \cdot s(T,d)}$, where $s(t,d) = \binom{t \cdot n + d}{d}$ is the dimension of polynomial ring in $t \cdot n$ variables.

We note that, under the assumption that the demand functions are invertible (i.e., the sensitivity matrices A_t in (4.1a) are non-singular), we could equivalently look for polynomial disturbance-feedback *demand policies* $d_{it}(\boldsymbol{\varepsilon}_{[t]}) \in \mathcal{P}_d[\boldsymbol{\varepsilon}_{[t]}]$. For a brief discussion of cases when this approach might be advantageous, please refer to Section 4.4.

Just as with the approach in Chapter 2, note that, under polynomial policies of the form (4.4), the original Problem (P) becomes non-convex in the decision variables. Hence, the current section is devoted to showing how to formulate tractable convex optimization problems that allow the computation of the optimal policy parameters.

Let us first consider the capacity constraints (4.3b) in Problem (P). Note that, with polynomial pricing policies of the form (4.4), the realized demand function $\boldsymbol{D}_t(\boldsymbol{p}_t, \boldsymbol{\varepsilon}_t)$ can be written as

$$\boldsymbol{D}_t(\boldsymbol{p}_t, \boldsymbol{\varepsilon}_t) = A_t L_t \boldsymbol{\xi}_t + \boldsymbol{b}_t + \boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\xi}_t \equiv \mathcal{B}_d(\boldsymbol{\varepsilon}_{[t]})$ denotes all the monomials in indeterminates $\boldsymbol{\varepsilon}_{[t]}$ of degree $\leq d$. In particular, it is a polynomial function of the history $\boldsymbol{\varepsilon}_{[t+1]}$, which implies that the capacity constraints (4.3b) (written in vector form) become:

$$\boldsymbol{f}(\boldsymbol{\xi}_t) \stackrel{\text{def}}{=} \boldsymbol{C} + U\boldsymbol{\xi}_{T+1} - \sum_{t=1}^{T} \big(A_t L_t \boldsymbol{\xi}_t + \boldsymbol{b}_t + \boldsymbol{\varepsilon}_t \big) \geq 0, \quad \forall\, \boldsymbol{\varepsilon}_{[T+1]} \in \mathcal{E}_{[T+1]}.$$

Above, $\boldsymbol{C}, A_t$ and $\boldsymbol{b}_t$ are data, $U, \{L_t\}_{t \in \mathcal{T}}$ are decision variables, and $\boldsymbol{\xi}_t$ are monomials of uncertain quantities. In particular, every constraint is a polynomial with coefficients that are affine combinations of the decision variables, and with indeterminates $\boldsymbol{\varepsilon}_{[T]}$. Since the goal is to test whether the polynomial is non-negative on the set $\mathcal{E}_{[T+1]}$, and by (4.2), the latter set is simply a basic, closed semialgebraic set, we can immediately see that these constraints fall in the same category as the state-control constraints of Chapter 3. In particular, if we denote by $\tilde{g}_j$ the polynomials[1] generating

[1]These can be directly obtained from (4.2).

the set $\mathcal{E}_{[T+1]}$, i.e.,

$$\mathcal{E}_{[T+1]} \equiv \big\{ \tilde{g}_j(\boldsymbol{\varepsilon}_{[T]}) \geq 0,\ j = 1, \ldots, m \big\},$$

a sufficient condition for the capacity constraints to hold is

$$f_i = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j,\ \forall\, i \in \mathcal{I},$$

where the polynomials σ_j are SOS. As such, all the remarks in Chapter 3 pertaining to the relation between the choice of degree for σ_j, the size of the resulting SDP formulation, and the degree of conservativeness also apply here, as well (the reader is referred to Section 3.3 for details).

In an analogous fashion, the constraints on the prices (4.3c) and on the order quantity (4.3d) can also be written as polynomial functions, with coefficients depending on the decision variables $\{L_t\}_{t\in\mathcal{T}}$, and with indeterminates $\boldsymbol{\varepsilon}_{[T+1]}$. Thus, the SOS framework can be applied here, as well, to derive a safe convex reformulation of the constraints.

4.3.1 Reformulating the Objective

We now focus attention on the objective in Problem (P), which can be written concisely as

$$\begin{aligned} &\max_{\{L_t\},U} \ \inf_{\mathbb{P}\in\mathcal{P}} \ \mathbb{E}_{\boldsymbol{\varepsilon}_{[T+1]}\sim\mathbb{P}} \big[J(L_1, \ldots, L_T, U, \boldsymbol{\varepsilon}_{[T+1]}) \big], \\ &J(L_1, \ldots, L_T, U, \boldsymbol{\varepsilon}_{[T+1]}) \stackrel{\text{def}}{=} \sum_t \boldsymbol{\xi}_t' L_t' \big(A_t L_t \boldsymbol{\xi}_t + \boldsymbol{b}_t + \boldsymbol{\varepsilon}_t \big) - \boldsymbol{r}'\, U \boldsymbol{\xi}_{T+1}, \end{aligned} \tag{4.5}$$

where we use the same shorthand notation introduced earlier, $\boldsymbol{\xi}_t \equiv \mathcal{B}_d(\boldsymbol{\varepsilon}_{[t]})$. At this point, we segregate our discussion into two parts. The first considers the case when the only information available about the set $\mathcal{P}$ is the support, while the second extends the discussion to a situation when moments are also available.

Known Support Information

The following proposition characterizes the first case:

Proposition 2 (Remark 23 in Shapiro et al. [130])**.** *If the only information available on the measures in the set $\mathcal{P}$ is the support $\mathcal{E}_{[T+1]} = \mathcal{E}_1 \times \cdots \times \mathcal{E}_T$, then problem* (4.5) *is equivalent to the following (deterministic) problem:*

$$\max_{\{L_t\},U} \min_{\boldsymbol{\varepsilon}_{[T+1]} \in \mathcal{E}_{[T+1]}} J(L_1, \ldots, L_T, U, \boldsymbol{\varepsilon}_{[T+1]}).$$

Proof. First note that, since $\boldsymbol{\varepsilon}_{[T+1]}$ is a compact set and $J(L_1, \ldots, L_T, U, \boldsymbol{\varepsilon}_{[T+1]})$ is a polynomial in $\boldsymbol{\varepsilon}_{[T+1]}$, then for any choice of the decision variables L_t, U, we have

$$\inf_{\boldsymbol{\varepsilon}_{[T+1]} \in \mathcal{E}_{[T+1]}} J(L_1, \ldots, L_T, U, \boldsymbol{\varepsilon}_{[T+1]}) = \min_{\boldsymbol{\varepsilon}_{[T+1]} \in \mathcal{E}_{[T+1]}} J(L_1, \ldots, L_T, U, \boldsymbol{\varepsilon}_{[T+1]}) \stackrel{\text{def}}{=} \bar{J},$$

i.e., the infimum is achieved. The proof is now immediate, since any measure that assigns non-zero probability to a set $\tilde{\mathcal{E}} \subset \mathcal{E}_{[T+1]}$ not achieving $\bar{J}$ is dominated by a singleton measure that assigns all mass to (one of) the points in $\arg\min_{\boldsymbol{\varepsilon}_{[T+1]} \in \mathcal{E}_{[T+1]}} J(L_1, \ldots, L_T, U, \boldsymbol{\varepsilon}_{[T+1]})$. ☐

This proposition allows us to formulate the following simple claim:

Lemma 11. *Under Assumptions 4 and 5, when the only information about the set $\mathcal{P}$ is the support $\mathcal{E}_{[T+1]}$, solving Problem* (4.5) *is equivalent to providing an efficient test for the condition:*

$$Q(L_1, \ldots, L_t, U, \boldsymbol{\varepsilon}_{[T+1]}) \succeq 0 \,, \forall\, \boldsymbol{\varepsilon}_{[T+1]} \in \mathcal{E}_{[T+1]}, \tag{4.6}$$

where

$$Q(L_1, \ldots, L_t, U, \boldsymbol{\varepsilon}_{[T+1]}) \stackrel{\text{def}}{=}$$

$$\begin{pmatrix} \sum_t \boldsymbol{\xi}_t' L_t' (\boldsymbol{b}_t + \boldsymbol{\varepsilon}_t) - \boldsymbol{r}' U \boldsymbol{\xi}_{T+1} - J & \boldsymbol{\xi}_1' L_1' & \boldsymbol{\xi}_2' L_2' & \ldots & \boldsymbol{\xi}_T' L_T' \\ L_1 \boldsymbol{\xi}_1 & -A_1^{-1} & 0 & \ldots & 0 \\ L_2 \boldsymbol{\xi}_2 & 0 & -A_2^{-1} & \ldots & 0 \\ \vdots & \vdots & \ldots & \ddots & \vdots \\ L_T \boldsymbol{\xi}_T & 0 & 0 & \ldots & -A_T^{-1} \end{pmatrix}. \quad (4.7)$$

Proof. By Proposition 2, solving Problem (4.5) is equivalent to solving the following optimization:

$$\begin{aligned} \max_{\{L_t\}, U, J} \quad & J \\ \text{s.t.} \quad & J \leq \sum_t \boldsymbol{\xi}_t' L_t' \big(A_t L_t \boldsymbol{\xi}_t + \boldsymbol{b}_t + \boldsymbol{\varepsilon}_t \big) - \boldsymbol{r}' U \boldsymbol{\xi}_{T+1}, \ \forall \boldsymbol{\xi}_{T+1} \in \mathcal{E}_{T+1}. \end{aligned}$$

Under Assumptions 4 and 5, the matrices A_t are negative definite (see Horn and Johnson [80]). Therefore, with $-A_t \succ 0$, the second constraint above is equivalent to

$$\sum_t \boldsymbol{\xi}_t' L_t' \big(A_t L_t \boldsymbol{\xi}_t + \boldsymbol{b}_t + \boldsymbol{\varepsilon}_t \big) - \boldsymbol{r}' U \boldsymbol{\xi}_{T+1} - J \geq 0, \ \forall \boldsymbol{\xi}_{T+1} \in \mathcal{E}_{T+1} \quad \Leftrightarrow$$

(by Shur complement)

$$\begin{pmatrix} \sum_t \boldsymbol{\xi}_t' L_t' (\boldsymbol{b}_t + \boldsymbol{\varepsilon}_t) - \boldsymbol{r}' U \boldsymbol{\xi}_{T+1} - J & \boldsymbol{\xi}_1' L_1' & \boldsymbol{\xi}_2' L_2' & \ldots & \boldsymbol{\xi}_T' L_T' \\ L_1 \boldsymbol{\xi}_1 & -A_1^{-1} & 0 & \ldots & 0 \\ L_2 \boldsymbol{\xi}_2 & 0 & -A_2^{-1} & \ldots & 0 \\ \vdots & \vdots & \ldots & \ddots & \vdots \\ L_T \boldsymbol{\xi}_T & 0 & 0 & \ldots & -A_T^{-1} \end{pmatrix} \succeq 0 \, , \forall \boldsymbol{\varepsilon}_{[T+1]} \in \mathcal{E}_{[T+1]}.$$

□

Note that, in condition (4.6), all the entries in the matrix Q are polynomials in the indeterminates $\boldsymbol{\varepsilon}_{[T+1]}$, with coefficients depending affinely on the decision variables

$\{L_t\}_{t\in\mathcal{T}}$, U. As such, condition (4.6) requires testing when a polynomial matrix is positive semidefinite over a basic compact semialgebraic set.

While such conditions are, in general, NP-hard, several recent papers (see Nie [111] and references therein) have provided characterizations for cases of interest when polynomial-time tests are available. Unfortunately, in the general setting that we consider here, the conditions of Nie [111] do not apply, but a *sufficient* condition for testing is available. In particular, note that (4.6) is equivalent to:

$$\underbrace{[y_0, \boldsymbol{y}'] \, Q(L_1, \ldots, L_t, U, \boldsymbol{\varepsilon}_{[T+1]}) \begin{bmatrix} y_0 \\ \boldsymbol{y} \end{bmatrix}}_{\stackrel{\text{def}}{=} q(L_1, \ldots, L_t, U, \boldsymbol{\varepsilon}_{[T+1]}, y_0, \boldsymbol{y})} \geq 0, \ \forall \, (y_0, \boldsymbol{y}) \in \mathbb{R}^{1+n \cdot T}, \ \forall \, \boldsymbol{\varepsilon}_{[T+1]} \in \mathcal{E}_{[T+1]}. \quad (4.8)$$

In the last inequality, $q(\cdot)$ is a polynomial in indeterminates $y_0, \boldsymbol{y}, \boldsymbol{\varepsilon}_{[T+1]}$, with coefficients that are affine functions of the decision variables $\{L_t\}_{t\in\mathcal{T}}$, U. Hence, the condition requires testing non-negativity of a polynomial over a set that is the cartesian product of the Euclidean space $\mathbb{R}^{1+n\cdot T}$ and a basic semialgebraic set $\mathcal{E}_{[T+1]}$. A sufficient condition for the latter is simply:

$$q = \sigma_0(y_0, \boldsymbol{y}, \boldsymbol{\varepsilon}_{[T+1]}) + \sum_{j=1}^{m} \sigma_j(y_0, \boldsymbol{y}, \boldsymbol{\varepsilon}_{[T+1]}) \, g_j(\boldsymbol{\varepsilon}_{[T+1]}), \quad \sigma_j \text{ s.o.s.}, \quad (4.9)$$

where g_j are all the polynomial constraints giving the set $\boldsymbol{\varepsilon}_{[T+1]}$, and the σ_j polynomials are all sums-of-squares. As such, this condition directly fits into the SDP framework that we introduced earlier, resulting in the following algorithm for solving the overall pricing problem:

Algorithm 4 Computing pricing and ordering policies of degree d under support information

1: Consider polynomial pricing and ordering policies in the disturbances, $\boldsymbol{p}_t(\boldsymbol{\varepsilon}_{[t]}) = L_t\,\mathcal{B}_d\big(\boldsymbol{\varepsilon}_{[t]}\big)$, $\boldsymbol{u}(\boldsymbol{\varepsilon}_{[T+1]}) = U\,\mathcal{B}_d(\boldsymbol{\varepsilon}_{[T+1]})$.

2: Express the planned and realized demands according to (4.1a). Each component of a price, planned or realized demand becomes a polynomial in indeterminate $\boldsymbol{\varepsilon}_{[t]}$, with coefficients given by linear combinations of $\{L_t\}_{1\le t\le t-1}$.

3: Express the revenue polynomial $q(\cdot)$ according to (4.8), and replace constraint (4.8) with the tractable constraint (4.9).

4: Replace a typical constraint $f(\boldsymbol{\varepsilon}_{[t]}) \ge 0,\ \forall\, \boldsymbol{\varepsilon}_{[t]} \in \mathcal{E}_{[t]} \overset{\text{def}}{=} \big\{\boldsymbol{\varepsilon} : g_j(\boldsymbol{\varepsilon}) \ge 0,\, j = 1,\dots,m\big\}$ (for capacity, price, order quantity or revenue) with the requirements:

$$
\begin{aligned}
& p = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j && \big(\text{linear constraints on coefficients}\big)\\
& \sigma_j \text{ SOS},\ j = 0,\dots,m. && \big(m+1 \text{ SDP constraints}\big)\\
& \quad \deg(\sigma_j\, g_j) \le \max\Big(d, \max_j\big(\deg(g_j)\big)\Big), &&\\
& \quad \deg(\sigma_0) = \max_j\big(\deg(\sigma_j\, g_j)\big). &&
\end{aligned}
$$

5: Solve the resulting SDP to obtain the coefficients L_t, U of the policies.

The main observations made in Section 3.3.2 of Chapter 3 with respect to the size of the overall formulation apply here, as well. However, the size of the SDP constraints here is potentially even larger than that of Chapter 3, due to the introduction of the additional variables $\boldsymbol{y}, y_0$. Natural choices for reducing the size would be to consider pricing policies that (a) do not depend on errors from all the items (e.g., the price $p_{i,t}$ could be restricted to depend on disturbances $\varepsilon_{i,1},\dots,\varepsilon_{i,t-1}$), or (b) do not depend on the entire history (e.g., $\boldsymbol{p}_t$ could be restricted to depend on $\boldsymbol{\varepsilon}_{t-W}, \boldsymbol{\varepsilon}_{t-W+1},\dots,\boldsymbol{\varepsilon}_{t-1}$, where W is the size of a rolling window).

Known Support and Moment Information

We now consider the second case, namely when both the support $\boldsymbol{\varepsilon}_{[T+1]}$, as well as moment information is available for the set $\mathcal{P}$. In particular, we make the following assumption about the number of moments:

Assumption 7. *All the measures $\mathbb{P}$ in the set $\mathcal{P}$ have specified moments at least up to degree $2d$. In particular, for any $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_T)$, with $\alpha_i \in \mathbb{N}^n$ and $\mathbf{1}'\,\boldsymbol{\alpha} \leq 2d$, we have*

$$\mathbb{E}_{\boldsymbol{\varepsilon}_{[T+1]}\sim\mathbb{P}}\left[\prod_{t=1}^{T} \boldsymbol{\varepsilon}_t^{\boldsymbol{\alpha}_t}\right] = \mu_{\boldsymbol{\alpha}}, \quad \forall\, \mathbb{P} \in \mathcal{P},$$

where $\mu_{\boldsymbol{\alpha}} \in \mathbb{R}$ are given values which constitute a valid set of moments.

We note that, in general, testing membership in the set $\mathcal{P}$ when both support and moments are specified is NP-hard - even with mean and covariance information available, a particular instance of this problem requires testing whether a matrix is copositive, which is known to be co-NP-complete (see Quist et al. [122]). To avoid this issue, we follow the same pragmatic approach as Popescu [118], and explicitly assume that the moments $\mu_{\boldsymbol{\alpha}}$ are specified so that the set $\mathcal{P}$ is nonempty. As we will later see in Section 4.2.2, when one is free to construct the set of measures $\mathcal{P}$ from available data samples, this can always be ensured.

We now return to examine Problem (4.5). Note that the term under the expectation operator, i.e.,

$$\sum_t \boldsymbol{\xi}_t' L_t' \big(A_t L_t \boldsymbol{\xi}_t + \boldsymbol{b}_t + \boldsymbol{\varepsilon}_t\big) - \boldsymbol{r}'\, U \boldsymbol{\xi}_{T+1},$$

is a polynomial in indeterminates $\boldsymbol{\varepsilon}_{[T+1]}$ of degree $\leq 2d$ (recall that $\boldsymbol{\xi}_t \stackrel{\text{def}}{=} \mathcal{B}_d(\boldsymbol{\varepsilon}_{[t]})$ represents all monomials of degree $\leq d$). Therefore, with Assumption 7 in place, the expectation operator in (4.5) simply resumes to replacing any term of the form $\prod_{t=1}^{T} \boldsymbol{\varepsilon}_t^{\boldsymbol{\alpha}_t}$ with the corresponding $\mu_{\boldsymbol{\alpha}}$. Therefore, the expression becomes independent of the measure $\mathbb{P}$, and hence the infimum operator in (4.5) has no effect. Furthermore, it can be immediately seen that, if the set $\mu_{\boldsymbol{\alpha}}$ constitutes a valid set of moments, then the new objective,

$$\sum_t \mathrm{Tr}\big(A_t L_t \mathbb{E}[\boldsymbol{\xi}_t \boldsymbol{\xi}_t'] L_t'\big) + \sum_t \boldsymbol{b}_t' \, L_t \, \mathbb{E}[\boldsymbol{\xi}_t] + \mathrm{Tr}\big(\boldsymbol{\varepsilon}_t \boldsymbol{\xi}_t' L_t'\big) - \boldsymbol{r}' \, U \, \mathbb{E}[\boldsymbol{\xi}_{T+1}],$$

is a concave quadratic function of the decision variables $\{L_t\}_{t \in \mathcal{T}}$ and U - a function which can be very efficiently optimized.

We can now see that the scheme under support and moment information only entails a trivial modification of Algorithm 1. In particular, instead of Step 3, we simply replace all the moments $\mathbb{E}_{\mathbb{P}}\left[\prod_{t=1}^{T} \boldsymbol{\varepsilon}_t^{\boldsymbol{\alpha}_t}\right]$ by the values μ_α, and then solve the resulting SDP formulation to obtain the desired policy parameters. The main advantage of including moment information is that one can preserve a potentially simpler structure for the robust counterpart. In particular, note that if the supports $\mathcal{E}_t$ are polytopic or ellipsoidal, and we restrict attention to degrees $d \leq 1$ (i.e., non-adjustable or affine adjustable), the resulting robust counterpart is a second-order conic optimization problem (see Ben-Tal et al. [19]), which can be solved very efficiently even for large sizes using state-of-the-art solvers such as CPLEX (ILOG [82]).

4.3.2 Other Methods for Solving the Problem

In order to test the performance of our policies, we also consider several alternative methods for solving the original Problem (P), which we briefly discuss in the current section. First note that, under our setting where the uncertainties are specified by support (and moment) information, solving the problem exactly by Dynamic Programming would be prohibitive, not only due to the number of items, but also since the state space at time t would have to involve the entire sequence of realized uncertainties, $\boldsymbol{\varepsilon}_{[t]}$. Therefore, we discuss a set of heuristic policies against which we benchmark the performance of our robustified polynomial policies, under specific choices of degree d.

Certainty Equivalent

The first approximate method that we consider is the Certainty Equivalent (CE) heuristic, also known as Model Predictive Control (MPC) (Garcia et al. [69], Bertsekas [21], Bemporad et al. [8]). This procedure replaces the uncertain quantities corresponding to any future periods by a "sufficient statistic" (usually, the conditional mean), solves the resulting deterministic problem to obtain optimal open-loop decisions, and then proceeds to implement these decisions for the first (or first couple of) time periods. The heuristic is usually implemented in a rolling-horizon fashion, by resolving at successive periods.

Under the additive uncertainty model of Section 4.2.1, a typical CE step (at time k) would involve solving the following problem:

$$
\begin{aligned}
& \max_{\boldsymbol{u},\boldsymbol{p}_k,\dots,\boldsymbol{p}_T} \sum_{t=k}^{T} \boldsymbol{p}_t' \left(A_t \boldsymbol{p}_t + \boldsymbol{b}_t + \bar{\boldsymbol{\varepsilon}}_t\right) - \boldsymbol{r}'\,\boldsymbol{u} \\
(CE) \quad & \text{s.t.} \quad \sum_{t=k}^{T} (A_t\, \boldsymbol{p}_t + \boldsymbol{b}_t + \bar{\boldsymbol{\varepsilon}}_t) \le \boldsymbol{C} + \boldsymbol{u} - \sum_{t=1}^{k-1} (A_t\, \bar{\boldsymbol{p}}_t + \boldsymbol{b}_t + \boldsymbol{\varepsilon}_t) \\
& \qquad (\bar{\boldsymbol{p}}_1, \dots, \bar{\boldsymbol{p}}_{k-1}, \boldsymbol{p}_k, \dots, \boldsymbol{p}_T) \in \Omega_t^p,\ \forall\, t \in \{k, \dots, T\} \\
& \qquad \boldsymbol{u} \in \Omega^u.
\end{aligned}
\tag{4.10}
$$

Here, $\bar{\boldsymbol{p}}_1, \dots, \bar{\boldsymbol{p}}_{k-1}$ are prices that have been implemented in past periods, and $\boldsymbol{\varepsilon}_1$, $\dots$, $\boldsymbol{\varepsilon}_{k-1}$ are realized (known) values, so that these quantities act as data for the optimization problem. The decision variables are the open-loop controls $\boldsymbol{u}$, $\{\boldsymbol{p}_t\}_{t=k,\dots,T}$, while $\{\bar{\boldsymbol{\varepsilon}}_t\}_{t=k,\dots,T}$ represents the sufficient statistic of $\boldsymbol{\varepsilon}_t$ which we referred to earlier. In particular, if moments are available, then the conditional mean can be directly used. If only support information is provided, then a good substitute would be to replace every $\boldsymbol{\varepsilon}_t$ with a point that is "central" in $\mathcal{E}_t$ (e.g., since most sensible supports for $\boldsymbol{\varepsilon}_t$ should contain the point 0, one could simply take $\bar{\boldsymbol{\varepsilon}}_t = 0$).

We note that, since the sets Ω_t^p and Ω^u are polyhedral, and the objective is concave, the above problem is a Quadratic Program (QP) of fairly small size, and can be solved efficiently using commercially available software, such as CPLEX (ILOG [82]), SDPT3

(Toh et al. [141]) or SeDuMi (Sturm et al. [137]).

Sample Average Approximation

A second heuristic that we consider is a variation of the *Sample Average Approximation* (SAA) (Shapiro et al. [130], Birge and Louveaux [41]). Here, we assume that N sample-path realizations are available for the stochastic process of disturbances, i.e., we have $\boldsymbol{\varepsilon}^{(i)}_{[T+1]}, i = 1, \ldots, N$. In practice, these could either be obtained from historical data, or from a simulation engine.

In the SAA method, Problem (4.10) above is replaced with the following optimization:

$$
\begin{aligned}
& \max_{\boldsymbol{u}, \boldsymbol{p}_k, \ldots, \boldsymbol{p}_T} \sum_{t=k}^{T} \frac{1}{N} \boldsymbol{p}_t' \left(A_t \boldsymbol{p}_t + \boldsymbol{b}_t + \boldsymbol{\varepsilon}_t^{(i)} \right) - \boldsymbol{r}' \boldsymbol{u} \\
(SAA) \quad & \text{s.t.} \quad \sum_{t=k}^{T} (A_t \, \boldsymbol{p}_t + \boldsymbol{b}_t + \boldsymbol{\varepsilon}_t^{(i)}) \leq \boldsymbol{C} + \boldsymbol{u} - \sum_{t=1}^{k-1} (A_t \, \bar{\boldsymbol{p}}_t + \boldsymbol{b}_t + \boldsymbol{\varepsilon}_t^{(i)}) \\
& \quad (\bar{\boldsymbol{p}}_1, \ldots, \bar{\boldsymbol{p}}_{k-1}, \boldsymbol{p}_k, \ldots, \boldsymbol{p}_T) \in \Omega_t^p, \; \forall \, t \in \{k, \ldots, T\} \\
& \quad \boldsymbol{u} \in \Omega^u.
\end{aligned}
\tag{4.11}
$$

Note that, here, we are essentially using an empirical distribution measure to estimate the true measure of the stochastic quantities. If the latter measure were actually unique (i.e., the set $\mathbb{P}$ contained a singleton), then, under mild technical conditions, one could expect the objective in Problem (4.11) to converge (uniformly) to the true objective of the problem, as N gets large (see Chapter 5 of Shapiro et al. [130] for details). Certain estimates for the size of N are also available, which guarantee that the solution to the SAA approximation is feasible, with high probability, for the original problem (see Calafiore and Campi [45], Calafiore and Campi [46], Nemirovski and Shapiro [108] for the case of known distribution, and Iyengar and Erdoğan [83] for a distributionally robust setting, similar to the one we consider here). The advantage of the SAA approach is that one could also embed adjustability, by allowing decisions to depend in a parametric fashion on the realized uncertainties (we refer the interested reader to the recent paper Lobel and Perakis [97] for more details). Here, we simply consider the non-adjustable SAA described in (4.11), and allow resolving (in a similar fashion as for the CE heuristic), at particular points in time.

Perfect Hindsight

The perfect hindsight heuristic, as the name suggests, is a sample-path optimization which has the entire realization $\boldsymbol{\varepsilon}_{[T+1]}$ available (the optimization to be solved looks exactly like the one in (4.10), except that $\bar{\boldsymbol{\varepsilon}}_t$ is replaced with the realized $\boldsymbol{\varepsilon}_t$). This is clearly not an implementable policy, but it provides an upper-bound for the achievable revenue, against which we can compare the different heuristics.

While several other computational approaches are also possible, for instance, based on one- or two-step look-ahead policies (Bertsekas [21]) or by Approximate Dynamic Programming (Bertsekas and Tsitsiklis [23]), we have decided to restrict attention to a subset, and leave a more comprehensive comparison for future research.

4.4 Extensions

In this section, we introduce several relevant extensions of the models presented thus far. In particular, we discuss multiplicative disturbances, disturbances affecting the sensitivity matrices A_t, and also potential generalizations to log-linear (or exponential) demand functions.

4.4.1 Multiplicative Disturbances

Note that the linear demand model we presented in Section 4.2.1 was affected by additive disturbances, i.e., via (4.1b). The pitfall of this approach is that, for large, negative disturbances $\boldsymbol{\varepsilon}_t$, one can obtain negative sales. While, in some applications, this may be suitable (e.g., to capture the effect of returns of merchandise), it is often undesirable, and avoided in models (see the comments in Section 7.3.4.1 of Talluri and van Ryzin [138]). Therefore, we would like to briefly discuss the case of *multiplicative* uncertainty, i.e., when the realized demand depends on the planned demand by

$$\boldsymbol{D}_t(\boldsymbol{d}_t, \boldsymbol{\zeta}_t) = \text{diag}(\boldsymbol{\zeta}_t)\, \boldsymbol{d}_t.$$

Under this model, the usual assumption in the literature is that ζ_{it} are non-negative random variables, with mean 1, $\forall\, i \in \mathcal{I}$, $\forall\, t \in \mathcal{T}$. For simplicity, we focus on the case where $\text{diag}(\boldsymbol{\zeta}_t) = \varepsilon_t^2 \boldsymbol{I}$ (i.e., the same multiplicative factor affects all demands), but several of our ideas can be immediately extended to the case of distinct disturbances. Here, we model the quantities ε_t as before. In particular, we assume that $\boldsymbol{\varepsilon}_{[T+1]}$ is distributed according to an unknown probability measure $\mathbb{P}$, belonging to a set $\mathcal{P}$ characterized by a known support of type (4.2) (restricted to be in the non-negative orthant), and (possibly) having known moments up to degree $2d$.

Under this new setting, we can also consider polynomial pricing policies of the form $\boldsymbol{p}_t = L_t\, \boldsymbol{\xi}_t$, where $\boldsymbol{\xi}_t \equiv \mathcal{B}_d(\boldsymbol{\varepsilon}_{[t]})$. The following remarks outline the similarities and changes from our previous discussion for additive uncertainty:

- Every capacity, pricing and order quantity constraint still represents a polynomial inequality, where the polynomial is in indeterminates $\boldsymbol{\varepsilon}_{[t]}$, and with coefficients affinely depending on $\{L_t\}_{t \in \mathcal{T}}, U$. Thus, they can be processed exactly as described in the prior section, using the SOS framework.

- The objective can be written as

$$\max_{\{L_t\}, U} \; \inf_{\mathbb{P} \in \mathcal{P}} \; \mathbb{E}_{\boldsymbol{\varepsilon}_{[T+1]} \sim \mathbb{P}} \big[J(L_1, \dots, L_T, U, \boldsymbol{\varepsilon}_{[T+1]}) \big],$$

$$J(L_1, \dots, L_T, U, \boldsymbol{\varepsilon}_{[T+1]}) \overset{\text{def}}{=} \sum_t \boldsymbol{\xi}_t' L_t' \varepsilon_t^2 (A_t L_t \boldsymbol{\xi}_t + \boldsymbol{b}_t) - \boldsymbol{r}'\, U \boldsymbol{\xi}_{T+1}$$

  As such, we can discuss the same two cases encountered earlier.

- When the only information about the measure is the support, then a similar result to 2 holds, and, under Assumptions 4 and 5, the Shur Complement Lemma can be invoked to obtain a condition such as

$$\begin{pmatrix} \sum_t \varepsilon_t^2 \boldsymbol{\xi}_t' L_t' \boldsymbol{b}_t - \boldsymbol{r}' U \boldsymbol{\xi}_{T+1} - J & \varepsilon_1 \boldsymbol{\xi}_1' L_1' & \varepsilon_2 \boldsymbol{\xi}_2' L_2' & \dots & \varepsilon_T \boldsymbol{\xi}_T' L_T' \\ \varepsilon_1 L_1 \boldsymbol{\xi}_1 & -A_1^{-1} & 0 & \dots & 0 \\ \varepsilon_2 L_2 \boldsymbol{\xi}_2 & 0 & -A_2^{-1} & \dots & 0 \\ \vdots & \vdots & \dots & \ddots & \vdots \\ \varepsilon_T L_T \boldsymbol{\xi}_T & 0 & 0 & \dots & -A_T^{-1} \end{pmatrix} \succeq 0,$$

$$\forall\, \boldsymbol{\varepsilon}_{[T+1]} \in \mathcal{E}_{[T+1]}.$$

In this form, we can again rewrite the condition as in (4.8), (4.9), with the only difference being the slightly larger degree of the resulting polynomial $q(\cdot)$ of (4.8).

- When moment information is also available, we can simply apply the same procedure as before, and replace all the monomials in $\boldsymbol{\varepsilon}_{[t]}$ with the respective moments. It is easy to see that the resulting expression for the objective remains concave in the variables L_t, U, and, therefore, the exact same approach as before is immediately applicable.

We note that the model above could also be interpreted as corresponding to a case when there are disturbances ε_t^2 affecting the sensitivity matrices A_t. Combining such a model with our earlier one, on additive disturbances, and under the additional assumption that one can *simultaneously observe*[2] both sources of uncertainty, one could then use the same SOS framework to look for adjustable polynomial policies.

4.4.2 Exponential (or Log-Linear) Demand Model with Multiplicative Noise

One of the major arguments against the demand model (4.1a), which we have examined thus far, is that the linear functional dependency has often been found to

[2]Note that, even for a single item with demand given by $D_t(d_t, \varepsilon_t, \zeta_t) = \zeta_t\, d_t + \varepsilon_t$, where ε_t and ζ_t are additive and multiplicative disturbances, respectively, if one only observes the realized demand D_t, then one might not be able to simultaneously estimate ε_t and ζ_t.

deliver poor performance in practice. A different form, which has been quite popular in econometric studies, and that has also received a lot of attention in the RM literature (see Rakesh and Steinberg [123], Gallego and van Ryzin [68], Smith and Achabal [135], and Talluri and van Ryzin [138] for more details) is the exponential (or log-linear) model with multiplicative uncertainty. That is,

$$\log \boldsymbol{d}_t(\boldsymbol{p}_t) = \boldsymbol{b}_t + A_t\, \boldsymbol{p}_t,$$
$$\log \boldsymbol{D}_t(\boldsymbol{d}_t, \boldsymbol{\zeta}_t) = \log \boldsymbol{d}_t + \boldsymbol{\zeta}_t,$$

where the $\log(\cdot)$ operator is interpreted component-wise, and the parameters have the same significance as in Section 4.4.2. We note that referring to this as a *multiplicative* model is in keeping with the fact that the realized demand for item i is given by $D_{it} = d_{it}e^{\zeta_{it}}$, hence one could equivalently consider as disturbances $\varepsilon_{it} \equiv e^{\zeta_{it}}$, obtaining a typical instance of the multiplicative models in Talluri and van Ryzin [138]. A main advantage of this model is that (i) the demand function is non-negative for any (non-negative) value of the price, and (ii) the model is well suited for estimation by OLS regression techniques, provided the sales are sufficiently frequent[3].

With respect to restrictions on the model parameters, one typically requires the same Assumptions 4 and 5 (or 6) to argue that the matrices A_t are non-singular, so that an inverse demand function always exists, and corresponding prices can be computed for any given demand vector $\boldsymbol{d}_t$. This is the approach we take here, as well. In particular, letting our decisions be the *demand policies* $\boldsymbol{d}_t$, we can rewrite the earlier equations as

$$\boldsymbol{p}_t(\boldsymbol{d}_t) = \tilde{A}_t \log \boldsymbol{d}_t + \tilde{\boldsymbol{b}}_t \tag{4.13a}$$
$$\boldsymbol{D}_t(\boldsymbol{d}_t, \boldsymbol{\zeta}_t) = \operatorname{diag}(\boldsymbol{\varepsilon}_t)\, \boldsymbol{d}_t, \tag{4.13b}$$

where $\tilde{A}_t = A_t^{-1}$ and $\tilde{\boldsymbol{b}}_t = -A_t^{-1}\, \boldsymbol{b}_t$.

We focus our remaining discussion on the case of a *single item*, with *time-invariant*

[3] Note that, in case there are records with 0 sales/demand, one has to deal with the quantity $\log(0)$.

sensitivity, and discuss the limitations of the approach. As mentioned, we look for demand policies that depend polynomially on the observed uncertainties, i.e.,

$$d_t = \boldsymbol{\ell}_t' \, \boldsymbol{\xi}_t$$

where $\boldsymbol{\xi}_t \overset{\text{def}}{=} \mathcal{B}_d(\boldsymbol{\varepsilon}_{[t]})$. The decision variables are now the vectors $\boldsymbol{\ell}_t \in \mathbb{R}^{\binom{n\cdot(t-1)+d}{d}\times 1}$. The modifications/similarities from our earlier approach are as follows:

- The capacity constraint, as well as any constraints on the order quantity u or on a demand sequence $\boldsymbol{d}_{[t]}$, resume to testing polynomial non-negativity, where the coefficients of the polynomial are affine in the decision variables $\{\boldsymbol{\ell}_t\}_{t\in\mathcal{T}}, \boldsymbol{u}$. Thus, any such constraint can be directly enforced using the SOS framework.

- Under Assumption 4, a price lowerbound would translate to $p_t \geq \Gamma \;\Leftrightarrow\; d_t \leq \exp\left(\frac{\Gamma - b_t}{-a_t}\right)$, which can also be immediately enforced in the SOS framework. Similarly, price upper-bounds or price monotonicity can also be re-written equivalently as affine constraints on the demands, and hence can be accommodated.

  However, we remark that incorporating arbitrary affine constraints on the price sequence $p_{[t]}$ is not possible. More precisely, since any such constraint $\sum_t \alpha_t \, p_t \geq \beta$ is equivalent to $\prod_t d_t^{\alpha_t \, a_t} \geq e^{\beta - \sum_t \alpha_t \, b_t}$, arbitrary coefficients α_t lead to non-linear constraints in the d_t polynomials, hence are outside the scope of our approach.

- For the objective, note that a typical stage revenue can be written as

  $$(d_t \, \varepsilon_t) \, p_t(d_t) = (d_t \, \varepsilon_t)(\tilde{a}_t \, \log d_t + \tilde{b}_t).$$

  The term potentially presenting problems is $\tilde{a}_t \, \varepsilon_t \, d_t \, \log d_t$. Since $\varepsilon_t \geq 0$, and $\tilde{a}_t \leq 0$, this is always a concave function of d_t, and, as such, we can introduce a piece-wise affine, concave under-estimator for it. More precisely, consider a finite number of pieces $\{\alpha_k, \beta_k\}, k \in \mathcal{I}_t$, such that

$$\min_{k \in \mathcal{I}_t} \big(\alpha_k x + \beta_k \big) \leq \tilde{a}_t \, x \, \log x, \ \forall \, x \in (0, +\infty).$$

The number of pieces, $|\mathcal{I}_t|$, as well as the slopes and intercepts, α_k, β_k, can be chosen (offline) so as to achieve a good trade-off between maximum revenue loss and computational burden. Once the under-estimators are fixed, we can introduce a new polynomial stage revenue, $\mathcal{C}_t(\boldsymbol{\varepsilon}_{[t+1]}) \stackrel{\text{def}}{=} \boldsymbol{c}_t' \, \boldsymbol{\xi}_{t+1}$, constrained to satisfy

$$\mathcal{C}_t(\boldsymbol{\varepsilon}_{[t+1]}) \leq \varepsilon_t \tilde{b}_t \, \boldsymbol{l}' \boldsymbol{\xi}_t + \varepsilon_t \, \alpha_k \boldsymbol{l}' \boldsymbol{\xi}_t + \varepsilon_t \beta_k, \, \forall \, \boldsymbol{\varepsilon}_{[t+1]} \in \mathcal{E}_{[t+1]}, \ \forall \, k \in \mathcal{I}_t.$$

Such constraints can be directly enforced within the SOS framework. The corresponding overall objective would then be to maximize $\sum_t \mathcal{C}_t - r \, u(\boldsymbol{\varepsilon}_{[T+1]})$. Since this term is also a polynomial in indeterminates $\boldsymbol{\varepsilon}_{[T+1]}$, with coefficients that are affine in the variables $\boldsymbol{c}_t, u$, they can directly be accommodated for the case of known support or known moments.

The approach as presented can also be extended to the case of multiple products sharing a common capacity (e.g., Adida and Perakis [1]), as long as there are no price-interaction terms (i.e., the matrices A_t are diagonal). For the case of non-diagonal A_t, note that the revenue would involve the function $f : \mathbb{R}^n \to \mathbb{R}$, $f(\boldsymbol{d}) = \boldsymbol{d}' A_t^{-1} \log(\boldsymbol{d})$. The complication is that, even when A_t satisfies Assumptions 4, 5 and/or 6, it may be that $f(\boldsymbol{d})$ is not concave in $\boldsymbol{d}$. In this situation, finding under-estimators as we did above might be considerably more challenging. However, if one can compute, by some other techniques, a concave, piece-wise (or quadratic) underestimator for the function $f(\boldsymbol{d})$, then the SOS framework as described is immediately applicable to this setting.

4.5 Calibrating the Models from Real Data

In the current section, we briefly discuss our data-set, and describe the techniques we used for calibrating the models directly from data.

4.5.1 Data Set

Our original set consisted of one season of sales (30 weeks) from a large US retailer in the fashion industry. After appropriate cleaning, the data contained a total of 102 different stock keeping units (SKU), corresponding to one division of the retailer. The organizational structure (a sub-part of which is depicted in Figure 4-1), consisted of 6 different departments, with each department segregated into subclasses, and each subclass containing a specific number of different SKUs - refer to Table 4.1 below for a breakdown of the SKUs into the higher organizational units[4].

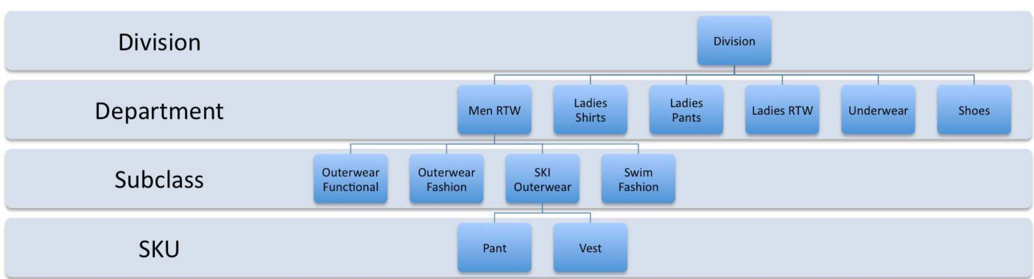


Figure 4-1: Organizational chart for the division.

Department #	**1**	**2**	**3**	**4**	**5**	**6**
Subclasses	2	7	5	3	2	3
Total SKUs	3	38	38	12	8	3

Table 4.1: Size and composition of each department and subclass.

For each SKU, the following fields were available:

- A brief description (containing the name of the SKU), and a unique SKU id

[4]The original names of the units have been masked for privacy, but the numbers correspond to the actual data.

- The production cost of the SKU (in $)
- The full price of the SKU (in $)
- The ticket price charged in each week (in $)
- The average sell price in each week (in $)
- The number of items sold in each week
- The inventory at the end of each week
- The number of units received in each week.

Before proceeding, we make the following remarks with respect to the various fields.

1. The *ticket price* for each SKU corresponded to the price displayed on the sticker at the beginning of each week. This price was typically discounted during the selling season, with most SKUs having between 3 and 7 markdowns, and the average size of a markdown being 27% (see Figure 4-2 for a histogram). Typically, in all dynamic pricing problems, this would be the variable that one would be optimizing over, i.e., the $\boldsymbol{p}_t$ variables in Problem (P).

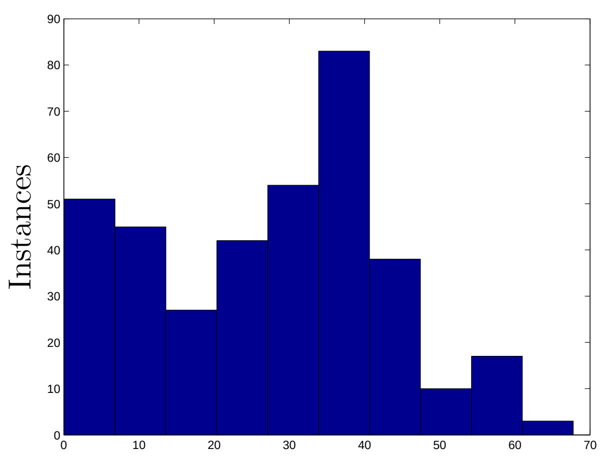


Figure 4-2: Histogram of the discounts in the division.

However, note that, in the data, the *ticket price* is actually different from the *average sell price*, which is the actual price received in any given week. In fact, a boxplot of the data (see Figure 4-3) revealed that the latter price can be considerably lower than the former, particularly in certain periods of the year (during the major selling season, and then also towards the end of the season).

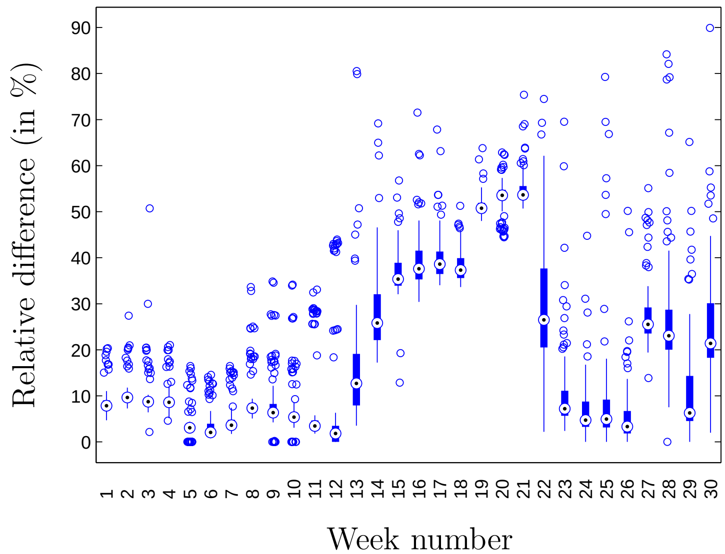


Figure 4-3: Boxplot of the relative difference between ticket price and average selling price.

The main reason for the discrepancy seems to be related to additional coupons for discounts, which the retailer is sending directly to its customers or employees Ramakrishnan [125]. The literature on dynamic pricing with promotions and coupons is certainly abundant (see, e.g., Chapter 9 of Talluri and van Ryzin [138] for examples of such models), but most work assumes that the coupons/promotions are endogenous decisions, rather than exogenous (unknown) factors. Since our data-set contained no information whatsoever about these coupons, apart from the observed effect on the prices, we have decided to ignore this issue in our ensuing model, and simply treat the ticket price as the relevant decision in each week.

We note that, in practice, this might not be the best possible choice, since the effect is certainly relevant. An alternative might be to represent the actual prices received in each period as random, e.g., $P_{it} = p_{it}\zeta_{it}$, where P_{it} is the received price for the i-th SKU, p_{it} is the planned price, and ζ_{it} is a multiplicative uncertainty. One could then construct a description for the disturbances $\boldsymbol{\zeta}_t$, and consider policies in both $\boldsymbol{\zeta}_t$ and $\boldsymbol{\varepsilon}_t$. However, such an approach was outside the scope of the current work, so we decided to leave it for future consideration.

2. Note that there is a field entitled *number of units received in each week*. This relates to the fact that, for several SKUs in our data-set, there were items *received* during the selling season. This does not refer to items returned from customers (in our data, the latter would reflect in the sales units for the respective week), but rather to additional units sent from a central store/warehouse to the outlets. As evidenced in the boxplot of Figure 4-4, most of the receipts occurred during the first 7 weeks of the selling season, and some were quite sizeable relative to the initial inventory in the respective SKU.

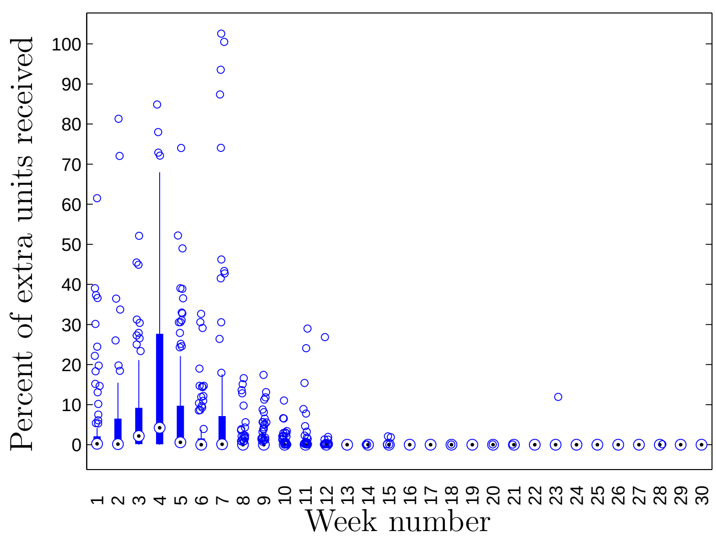


Figure 4-4: Percent of (extra) units received in each week, relative to initial inventory.

Since our data-set provided no additional information with respect to the potential sizes and times of such receipts, we, again, decided to ignore this factor. In our handling of the data, we simply added all such receipts to the initial inventory and operated under the premise that the initial capacity $\boldsymbol{C}$ was larger.

It is important to note that, even under the original inventory (i.e., ignoring the effect of receipts completely), none of the SKUs ran out of items by week 12, so that we were not accidentally ignoring instances of lost sales by adding the receipts in this way - see Figure 4-5 for boxplots of the (normalized) original inventory for all the SKUs, as well as the transformed one (by adding the receipts). Section 4.5.2 further elaborates on this issue.

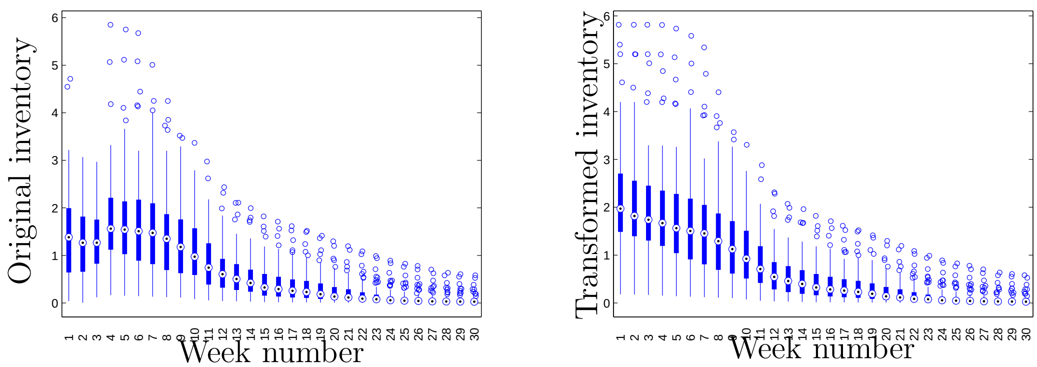


Figure 4-5: Original inventory (left) and transformed inventory, by adding receipts (right). Y-axis is normalized due to privacy reasons.

4.5.2 Demand Model Estimation and Calibration

We now discuss some aspects related to the estimation of the models using our specific data-set. We begin by focusing on the linear demand model from Section 4.2.1. Recall that the functional dependency introduced there was given by (4.1a), (4.1b), which we paste below, for convenience:

$$\boldsymbol{D}_t(\boldsymbol{p_t}, \boldsymbol{\varepsilon}_t) = \boldsymbol{b}_t + A_t \, \boldsymbol{p}_t + \boldsymbol{\varepsilon}_t.$$

While the model is certainly a simplification of reality, since it ignores several salient features (such as the effect of inventory on sales Smith and Achabal [135], the effect of promotions and coupons Woo et al. [147], Boyd et al. [44], the strategic customer behavior Talluri and van Ryzin [138], etc.), it remains very popular in the academic literature, and also in practice. One of the main attractive features of the model is the ease of estimation from data - more precisely, with unconstrained demand as the dependent variable, and price as an independent variable, one could utilize regression techniques to estimate the sensitivity matrices A_t and the market-size factors $\boldsymbol{b}_t$.

In practice, however, several issues can arise. Firstly, it is easy to see that the number of parameters to be estimated can quickly become very large, since it is proportional with both the number of items and the horizon. In particular, in case only a few selling seasons are available (in our data-set, we only have one!), estimating independent b_{it} for each item is practically infeasible. Therefore, what is often done in practice is to aggregate data from multiple items together, and/or to ignore some of the time dependencies. For instance, a popular choice (Talluri and van Ryzin [138], Ramakrishnan [124]) is to assume that the items in different organizational units are independent, that the price sensitivity matrix is time-invariant, i.e., $A_t = A, \forall\, t \in \mathcal{T}$, and that the $\boldsymbol{b}_t$ component can be separated into a base demand $\boldsymbol{b} \in \mathbb{R}^n$, which is time-invariant, and a seasonal factor $\boldsymbol{s}_t \in \mathbb{R}^n$, often assumed to be the same for all items in a particular organizational group. For instance, if all the items $i \in \mathcal{S}$ were taken to have the same seasonality, and be independent of items in $\mathcal{I} \setminus \mathcal{S}$, then the functional equation for the demand of items in $\mathcal{S}$ would become

$$\boldsymbol{D}_t(\boldsymbol{p_t}, \boldsymbol{\varepsilon}_t) = \boldsymbol{b} + A\,\boldsymbol{p}_t + \mathbf{1}\,s_t + \boldsymbol{\varepsilon}_t,\ \forall\, t \in \mathcal{T}, \tag{4.14}$$

where $A \in \mathbb{R}^{|\mathcal{S}|\times|\mathcal{S}|}$, $\boldsymbol{b} \in \mathbb{R}^{|\mathcal{S}|}$, and $s_t \in \mathbb{R}$ would represent an additive seasonal factor corresponding to period t. The aggregation of the items can be performed either by using sensible business rules Ramakrishnan [124], Talluri and van Ryzin [138], or by using other statistical techniques, such as clustering, classification and regression trees or time-series analysis (see, e.g., Kumar and Patel [91], Ghysels et al. [72] or the books Greene [77] and Box et al. [43]).

Due to these considerations, we decided to also make the following simplifications in our model:

1. *We assume that SKUs in different subclasses are independent.*

2. *We assume that all the SKUs inside a given subclass have the same seasonality*

factor s_t, but different market sizes, b_i.

3. *We assume that the demand-sensitivity matrices are time-invariant, i.e., $A_t = A, \forall\, t \in \mathcal{T}$.*

4. *We assume that each item's demand only depends on its own price and the average price of the other items inside the same subclass. Furthermore, we assume that the effects are the same across all the SKUs in a particular subclass.* More precisely, we take:

$$D_{it} = b_i + a\, p_{it} + a_{-} \sum_{j \in \mathcal{S} \setminus \{i\}} p_{jt} + s_t + \varepsilon_{it}, \tag{4.15}$$

   where a represents the effect of SKU i's own price, while a_{-} denotes the effect from the prices of all the other items j inside the same subclass $\mathcal{S}$.

These assumptions are made more out of necessity (i.e., to enable an adequate estimation), rather than out of solid economic or business considerations. In reality, even items inside the same subclass can be quite "different" in terms of seasonality patterns, and one can expect both substitutability, as well as complementarity effects to exist across subclasses[5]. Such effects could be captured with a significantly larger data-set, consisting of several selling seasons involving the same items, but were outside the scope of our data.

The second remark we would like to make is that some of the requirements in our model description (most importantly, Assumptions 4 and 5) might not hold if the parameters are estimated by running an OLS regression. One immediate correction for this would be to run a constrained regression, in which the parameters are forced, via inequality constraints, to obey the properties mentioned in our discussion in Section 4.4.2. This approach does not present any computational difficulties (one would have to solve a constrained quadratic program), but has the main pitfall of invalidating most of the standard statistical analysis in linear regression (e.g., inferences

[5]For an example of the former, imagine an item in fashion outerwear is discounted, hence one prefers to buy that rather than a functional outerwear item; for the latter, suppose a shirt is discounted, inducing the purchase of a matching pant, from a different subclass

based on t- or F-statistics are no longer possible under inequality constrained linear regression Geweke [71], so one must resort to other techniques, such as bootstrapping, for testing statistical significance). Our regression results, presented in Section 4.5.4, frequently encountered this problem, thus requiring a pragmatic choice that traded off between (a) the convenient theoretical properties of OLS regression and (b) the consistency of the model parameters with standard microeconomic theory.

Our third (and final) remark is related to the fact that our data-set contained *sales*, rather than direct *demand* information. The distinction becomes relevant when one might be dealing with a censoring effect, whereby, once on-hand inventory becomes 0, one observes a truncated demand function. There are standard tools in regression modelling for dealing with such situations (e.g., tobit regression Greene [77], the expectation-maximization algorithm, Gibbs sampling or the Kaplan-Meier estimator Talluri and van Ryzin [138]). However, in our data-set, the vast majority of SKUs still had remaining inventory after the end of the sales period, thus the number of records that could have suffered from censoring effects was very small. Therefore, we decided to ignore this issue in our regression estimation procedures.

4.5.3 Estimating the Model for the Uncertainties

With the above simplifications in place, one can perform panel regressions within each subclass $\mathcal{S}$ to obtain estimates $\hat{\boldsymbol{b}}$, $\hat{s}_t$ and $\hat{A}$ for the demand model corresponding to all the items $i \in \mathcal{S}$. One last component of our model must still be described, namely the construction of the support (and moment) information for the random terms $\boldsymbol{\varepsilon}_t$.

Note that, as a result of performing the OLS (or constrained) regression, one also obtains sample paths of the disturbances ε_{it} by means of the regression residuals. In particular, we have

$$e_{it} \stackrel{\text{def}}{=} D_{it} - \Big(\hat{b}_i + \hat{a}\, p_{it} + \hat{a}_{-} \sum_{j \in \mathcal{S} \setminus \{i\}} p_{jt} + \hat{s}_t\Big),\ \forall\, i \in \mathcal{S}, \forall\, t \in \mathcal{T}.$$

Based on these residuals, we propose the following simple scheme for constructing the supports and moments of the stochastic terms ε_{it}:

- Construct the support using a box model, i.e., take $\varepsilon_{it} \in [l_{it}, u_{it}], \forall\, i \in \mathcal{S}, \forall\, t \in \mathcal{T}$, where the bounds l_{it} and u_{it} are given by quantiles of the empirical distribution of the residuals e_{it}. A very similar model was recently considered by Perakis and Roels [114], in the context of network RM. The recommended choices there are the twenty-fifth and seventy-fifth percentiles of the empirical distribution, since they are less sensitive to censored data, and make the results more robust to the actual shape of the distribution or the location of the mode. In our models, we have also attempted using other variations, based on quantiles or widths controlled by standard deviations, but we generally found that the rule in Perakis and Roels [114] works quite well, and is less sensitive to the underlying (true) model of the disturbance terms.

  We note that many different approaches for constructing these supports are possible. Another option could be to additionally use the confidence intervals for the coefficients b_i and s_t, which (especially for highly variable periods), might better incorporate the original data. However, we decided to not pursue these further in our current model.

- Due to the scarcity of our data-set, estimating arbitrary moments is clearly not feasible without additional assumptions about the error terms ε_{it}. In particular, there are two natural assumptions that one could make: (a) that the disturbances ε_{it} are independent across the items, but correlated across time, or (b) that the disturbances are independent across time, but correlated across the items. For our analysis, we chose to make the following standing assumption about the error terms:

  Assumption 8. *The stochastic error terms ε_{it} are independent across the items $i \in \mathcal{S}$.*

  This simplification then allows us to estimate the raw (i.e., non-central) moments up to a pre-specified degree $2d$, by using the sample moments,

$$\mathbb{E}\Big[\prod_{t \in \tilde{T}} \varepsilon_{i,t}\Big] = \frac{1}{|\mathcal{S}|} \sum_{j=1}^{|\mathcal{S}|} \prod_{t \in \tilde{T}} e_{j,t}, \quad \forall\, i \in \mathcal{I},\, \forall\, \tilde{T} \subseteq \mathcal{T} \text{ s.t. } |\tilde{T}| \leq 2d.$$

For cases when the estimated mean did not lie in the support of the quantities (not very frequent), we opted to replace the estimated mean with the estimated median, which never suffered from this issue.

Assumption 8, which might appear as a gross oversimplification, is motivated by our belief that, in our data-set, most of the variability and poor(er) prediction came from residuals that are strongly correlated in time and heteroscedastic (as evidenced by the results in Section 4.5.4). As such, while cross-sectional (i.e., cross-item) correlations might indeed exist, we chose to ignore them for the remainder of the analysis.

Before proceeding to present our numerical results, we would like to make one last clarification with regards to the motivation behind our approach, and some of the choices involved. We recognize that, under the belief/assumption that the residuals in a regression model are correlated and/or heterscedastic, one can take the following approach:

(a) Test for such a phenomenon. There are well established procedures, for both heteroscedasticity (White, Goldfeld-Quandt or Breusch-Pagan tests - see Greene [77] for details), as well as auto-correlation (Box-Pierce, Durbin-Watson, etc.)

(b) If the phenomena are identified, one can attempt to adjust the regression model to correct for them. For instance, one could estimate a covariance matrix for the errors terms, and run a Generalized Least Squares (GLS) regression (see Chapter 13 of Greene [77] for details). Or, if one finds auto-regressive conditional heteroscedasticity, one can use powerful tools in time-series (ARCH, GARCH) to amend the initial model.

In our regressions, we have actually attempted some of the above procedures, as well as non-linear regressions which accounted for potential AR(p) disturbances (see page 257 of Greene [77] for a theoretical description). However, even in the corrected models, we still found evidence of the phenomena, most likely due to the other model mis-specifications (e.g., the shape of the demand functional form itself, the fact that SKUs inside the same subclass do not have identical seasonalities, etc.). In this context, we took the pragmatic approach of (a) accepting the fact that the models are most likely mis-specified, and (b) looking for robustified, adjustable policies, which partially allow one to correct for such problems.

4.5.4 Regression Results

With the above simplifications in place, we began our tests by running individual panel regressions (Greene [77]) for several subclasses. We restrict our descriptions below to one of the larger subclasses, namely subclass 2 of department 1, with 21 SKUs, but similar observations apply to some of the smaller ones.

The results for an unconstrained regression in department 2, subclass 1, are presented below. In particular, the regression had an $R^2 = 0.51$, an adjusted $R^2 = 0.50$, the two price coefficients,

$$\hat{a} = -95.384 \qquad\qquad \hat{a}_{-} = 13.930$$

were both significant at the 95% confidence level, and 8 (out of 29) seasonality terms $\hat{s}_t$ were found to be significant. Summaries for the values of the coefficients $\hat{b}_i$ and the seasonality factors are shown in Figure 4-6.

In particular, it can be seen that the results suffer from two of the caveats mentioned in Section 4.5.2, namely that several of the b_i terms are not positive, and the A matrix resulting from $\hat{a}$ and $\hat{a}_{-}$ is not diagonally dominant (in fact, it is not even negative semi-definite).

Furthermore, three different heteroscedasticity tests with respect to both the price variables and the time variables (Breusch-Pagan-Koenker, White and modified White) delivered p-values in the range of 10^{-9}, leading to a rejection of the hypothesis that

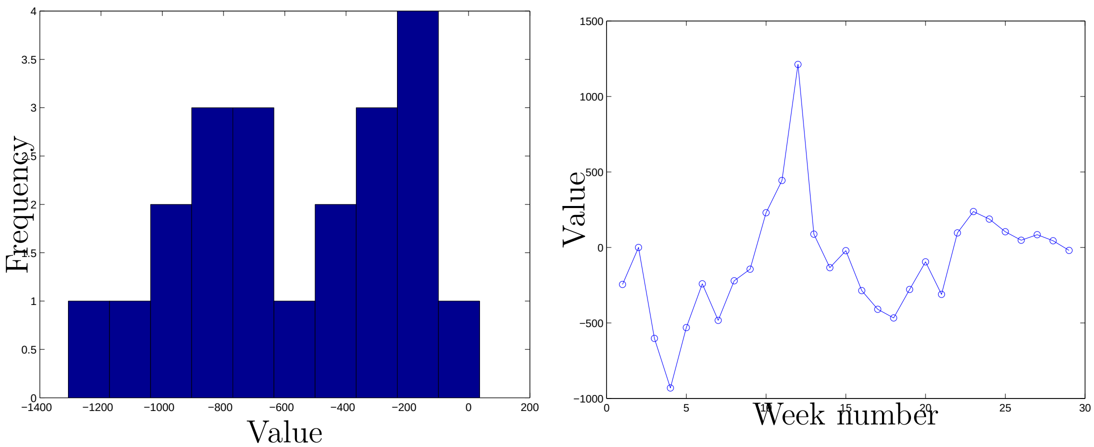


Figure 4-6: Results using OLS regression. Histogram of the values $\hat{b}_i$ (left) and plot of the values $\hat{s}_t$ (right)

the residuals are homoscedastic. The Durbin-Watson test for autocorrelation also produced a p-value of 10^{-214}, confirming our suspicion of autocorrelation. Very similar results were obtained for the other subclasses mentioned above - in fact, in *all* the cases, the hypotheses for homoscedasticity and non-autocorrelation were rejected at levels of confidence $\geq 99.99\%$.

As already mentioned, although we attempted several techniques to correct the regression model by accounting for these undesirable effects, in most instances, the problems persisted in the new regressions, as well. Furthermore, the issues related to the matrix A not being negative semidefinite and the coefficients b_i being negative also persisted throughout.

Therefore, we have taken the pragmatic decision of giving up the OLS regression, and running, instead, a version of constrained regression, where the structure given by Assumptions 4 and 5 was pre-imposed on the regression. The resulting price-sensitivity coefficients (for department 2, subclass 1), are

$$\hat{a} = -104.115 \qquad \hat{a}_{-} = 5.199,$$

and the coefficients $\hat{b}_i$ and seasonality terms $\hat{s}_t$ are represented in Figure 4-7 below.

We note that we have also attempted a version of regression where the $\hat{b}_i$ were

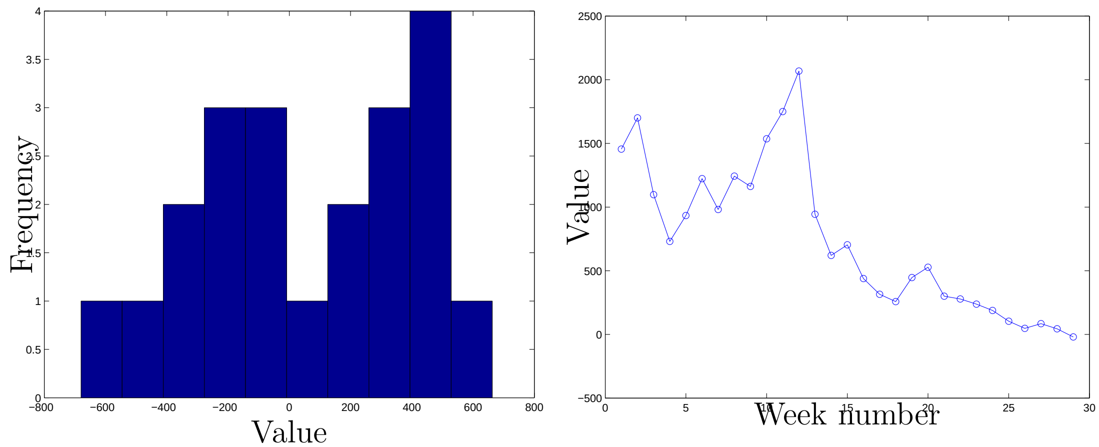


Figure 4-7: Results using constrained regression. Histogram of the values $\hat{b}_i$ (left) and plot of the values $\hat{s}_t$ (right)

also constrained to be non-negative. This did not result in significant changes in the price sensitivity coefficients, but rather a readjustment of the seasonal factors s_t to accommodate for the new requirement. Since all the coefficients s_t, as well as the b_i factors, were essentially computed relative to a baseline (the last period additive sales, the indicator of which was removed from the regression[6]), it appeared as though constraining $\hat{b}_i$ would not add much.

A similar process was run for the other subclasses mentioned above, as well as for several smaller subclasses. We remark that, in all the results, the coefficients $\hat{a}_-$ were always positive (suggesting substitutability effects in the data), and the regression constraining only $\hat{a}$ and $\hat{a}_-$ already returned positive $\hat{b}_i$'s (hence the issue mentioned in the above paragraph might have been specific to the subclass under consideration there). We also attempted the following modifications/extensions:

- Building models that performed data aggregations at higher levels (e.g., imposing the same seasonality for all items in a given division, but allowing individual price sensitivity coefficients at the subclass level).

- Using robust regression techniques Huber and Ronchetti [81] to correct for some of the outliers in the data. We tested several different weighting schemes (An-

[6] By removing one indicator from the regression, one is automatically introducing a bias. A different procedure, suggested in Greene [77], is to run a regression where the indicators are all constrained to sum up to 1. While this might remove some of the bias, it was outside the scope of our present work.

drews, bi-square, Cauchy, Welsch, Talwar, Welsch), and found that, while there was, occasionally, improvement in the number of significant coefficients, the quality of the overall prediction was not necessarily better than that obtained using the regular (OLS-based) methods.

Since the results were rather mixed, and not necessarily better than our baseline model, we decided to keep the initial choice of subclass-level aggregation, with constrained regression for the A matrix.

Results for the Uncertainty Models

Once the regressions were run, we used the residuals to construct the support and moments of the uncertain quantities ε_{it}, as described in Section 4.2.2. A typical boxplot of the residuals from the regression (here, again, department 2, subclass 1) is shown in Figure 4-8.

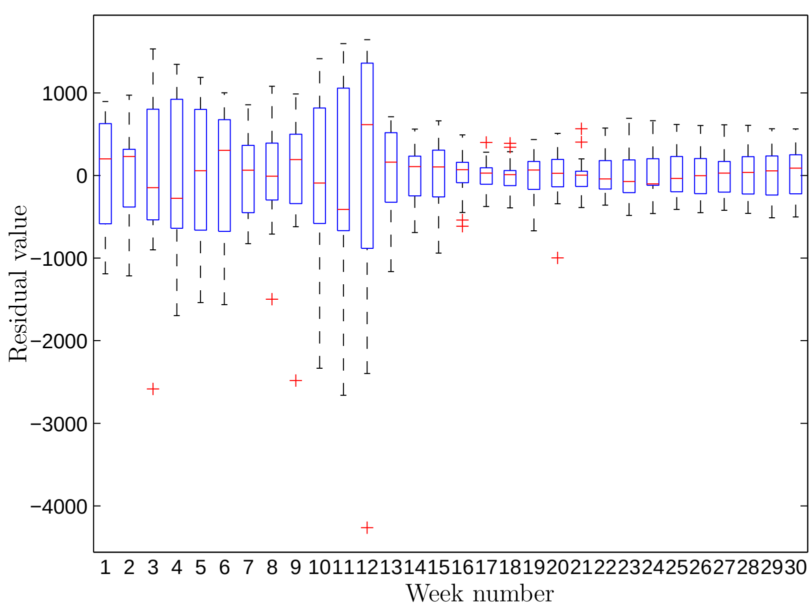


Figure 4-8: Residuals from the constrained regression in Subclass 1 of Department 2.

It can be seen even directly from the figure that the residuals are exhibiting heteroscedasticity (with considerably larger variability in the first half of the selling season), as well as strong autocorrelation (the sample autocorrelation matrix revealed a succession of clusters of strong negative correlation, followed by clusters of strong positive correlation). Therefore, a typical model for the residuals would involve a second-moment matrix with large (in absolute value) entries, of both positive and negative signs.

4.6 Testing the Polynomial Pricing Policies for the Linear Demand Model

Ideally, one would like to test the (combined) results of the estimation and optimization in an out-of-sample fashion. Unfortunately, due to the limited data available, and also the nature of the dynamic pricing problem (with pricing decisions influencing the observed demand), such a test is quite difficult to achieve. With this motivation in mind, we decided to test our policies on both the real data, as well as simulated data, which we artificially generated. The current section describes the exact procedures used throughout, and discusses the numerical results.

4.6.1 Testing with Simulated Data

As a first step in testing our algorithm, we constructed our own data-generating process, which produced historical records based on which the model would be estimated and policies would be computed. The advantage of this procedure is that it allowed us to test the performance of the scheme under the *true* demand model.

In order to better understand the interplay between the estimation and optimization engines, as well as to isolate the impact of particular parameters on the results, we began our tests by considering a case with *no price interactions*, i.e., when the A matrix in (4.14) is diagonal. More precisely, we proceeded in the following fashion:

Algorithm 5 Testing the policies of degree d with simulated data

1: For a collection $\mathcal{S}$ of n items, $\mathcal{S} = \{1, \ldots, n\}$, fix a set of *nominal* values for the parameters of the demand model (4.15). More precisely, take $\bar{\boldsymbol{b}} = \bar{b} \cdot \mathbf{1} \in \mathbb{R}^n$, $\bar{A} = \bar{a}\, I \in \mathbb{R}^{n \times n}$ (with $\bar{a} \leq 0$), a seasonal pattern $\bar{\boldsymbol{s}}_t = \bar{s}_t \cdot \mathbf{1} \in \mathbb{R}^n, \forall\, t \in \mathcal{T}$, and a stochastic model for $\{\boldsymbol{\varepsilon}_t\}_{t \in \mathcal{T}}$, given by a collection of nominal parameters $\bar{\Sigma}$.

2: Fix a particular pricing sequence for every item $i \in \mathcal{S}$.

3: Set the *true* model parameters to $\bar{\boldsymbol{b}}$, $\bar{A}$, $\bar{\boldsymbol{s}}$, $\bar{\Sigma}$.

4: **for** several values of a particular parameter η **do**

5: Generate "historical" records for each SKU using the *true* model and the pricing sequences.

6: **for** every SKU $i \in \mathcal{S}$ **do**

7: Using the data for all items $j \in \mathcal{S}$, $j \neq i$, construct a linear demand model of type (4.15), with the assumptions discussed in Section 4.5.2, and the additional simplification that $\bar{a}_{-} = 0$ (i.e., no interaction effects between the items).

8: Using the residuals from the regression model, estimate the support and moments of the disturbances $\varepsilon_{j,t}$, as discussed in Section 4.2.2.

9: Using the constructed model for the demand function and error terms, compute policies of degree 0 and 1 for item i. Here, the constraints in the sets Ω_t^p are price and demand non-negativity and price mark-down, while the only constraint in Ω^u is non-negativity.

10: Compare the performance (realized revenue) by Monte-Carlo simulation. More precisely,

(a) Generate noise terms according to different distributions, which may or may not obey the model constructed in Step 8 (i.e., in terms of support and moments).

(b) Compare the revenue under polynomial policies with the revenue achieved by the heuristics of Section 4.3.2.

11: **end for**

12: **end for**

We note that some of the steps in the above procedure have been left ambiguous: the specification of the noise model, the exact choice of the distributions for performing Monte-Carlo simulation, and the choice of parameter η to vary in Step 4. While many options are possible, we decided for the following:

- For the "true" noise model, we generate the noise for any item[7] according to an AR(1) process, i.e., $\varepsilon_{t+1} = \rho\,\varepsilon_t + u_t$, $\forall\, t \in \mathcal{T}$, where the terms u_t are i.i.d. random variables, and $|\rho| < 1$ determines the level of correlation. For u_t, we consider several possibilities: Gaussian (with mean $p\,L_t + (1-p)\,H_t$, and standard deviation σ_t), truncated Gaussian (with mean and standard deviation as before, and truncated in the interval $[L_t,\, H_t]$), mixture of Gaussians (two Gaussians, each with standard deviation σ_t, with means L_t and H_t, respectively, and with the former occurring with probability p), uniform (in the interval $[L_t,\, H_t]$). As such, the collection of parameters describing the noise model is $\Sigma \stackrel{\text{def}}{=} \{\rho, \sigma_t, L_t, H_t, p\}$.

- For the Monte-Carlo step, we either use the original model to generate "true" noise terms, or we fit a Gaussian or mixture of Gaussians (so that the moments are matched), or a uniform distribution (so that the range information is matched).

- For the parameter η in Step 4, we choose σ_t (the standard deviation of the residuals), ρ (the auto-correlation of the residuals), p (which controls the mean of the residuals) and a (the price sensitivity coefficient).

Throughout all the tests, the nominal values of the parameters that we used were $\bar{\sigma}_t = \sigma = 1.0$, $\bar{\rho} = 0.0$, $\bar{b} = 20$, $\bar{a} = -1.0$, $L_t = L = -1.0$, $H_t = H = 1.0$.

The results are presented in a sequence of tables and figures in Appendix C. Every case (corresponding to a particular parameter varying) is accompanied by two tables, a collection of boxplots, and a collection of histograms. We explain their significance for the first case, where the coefficient that varies is σ_t, and the meaning for the remaining ones is analogous.

[7] Recall that we are operating under the standing Assumption 8, hence we can drop the index i.

The first table and collection of boxplots always pertain to relative gaps from the perfect hindsight solution. For example, Table C.1 records statistics (average, standard deviation, minimum, maximum and median), while the accompanying Figure C-1 shows box-plots for the same relative gaps.

The second table and the collection of histograms pertain to performance gaps computed relative to the highest-degree polynomial policy (here, $d = 1$). As an example, Table C.2 records the same statistics mentioned above, while Figure C-2 then presents a histogram of these relative gaps.

The acronyms pertaining to the heuristics are as follows:

- **ALY** - *As Last Year* - that is, simply use the same price sequences as the historical ones.

- **CESO** - *Certainty Equivalent Solved Once* - this is the Certainty Equivalent procedure described in Section 4.3.2, solved only once (at the beginning of the horizon).

- **CEST** - *Certainty Equivalent Same Times* - the Certainty Equivalent procedure of Section 4.3.2, but with resolving (at the same set of times when the prices were discounted in the previous year).

- **SAA** - *Sample Average Approximation* - the procedure described in Section 4.3.2, solved only once (at the beginning of the horizon).

From the simulations, we can draw the following conclusions:

- The heuristic "As Last Year" performs very poorly, which is certainly justified, since the periods and sizes of the discounts in the historical sequence were chosen randomly (this heuristic has more meaning when applied to real data, since the historical choices in that context are most likely based on sensible reasons).

- Adjustability results in increased performance for robust policies. In particular, policies with $d = 1$ improve quite systematically over policies with $d = 0$ (i.e., robust, non-adjustable), both in terms of worst-case expected revenue, as well as in Monte-Carlo simulations on various distributions. This is particularly evident in the histograms of Figures C-2, C-4, C-6 and C-8, which clearly outline the improvements that one obtains by introducing minimal adjustability (i.e., degree 1).

- The heuristics CESO and CEST deliver comparable performance, and are, in many cases, quite close to the robust policies. In fact, these heuristics often outperform robust non-adjustable policies ($d = 0$), but are typically inferior to the adjustable robust ones, as evidenced by both the average and standard deviation of the optimality gaps (also refer to the same set of figures mentioned in the previous paragraph, and note that the histograms tend to have thicker left-tails, indicating under-performance). The most notable cases when the performance gaps increase (i.e., adjustable robust policies are even better) are cases where the standard deviation of the residuals, σ_t, is reasonably large (see Tables C.1 and C.2). This observation is certainly in line with our expectation that adjustable robust policies should guard against highly heteroscedastic residuals.

- Many of the heuristics are very close to the PH solution. This is mostly due to the choice in parameters, and - as we shall see in the next set of experiments - there are certainly interesting cases where the typical gaps from PH can be much larger.

4.6.2 Multi-Product Tests with Simulated Data

For the second category of tests, we considered several items (here, $n = 3$), and a price-sensitivity matrix A that was diagonally dominant and with equal off-diagonal terms (i.e., the demand equation given by (4.15)). Since our goal was more to test the quality of the optimization engine, we decided to make the following changes to the procedure described in Section 4.6.1:

- Instead of generating historical sales data, and then estimating the models, we proceeded to directly construct a system model (i.e., matrix A, vector $\boldsymbol{b}$, seasonalities s_t, etc.).

- We directly generated historical samples for the disturbance sequences ε_t.
- We no longer imposed a markdown constraint on the prices.

An instance of such a simulation is reported in Table C.9 and Figures C-9 and C-10. Here, the *true* distribution used for generating the disturbance terms was uniform, with 0-mean and a reasonably large support, and the values of ε_t in different periods were strongly *negatively* correlated. The testing distribution was chosen to be either the true one (i.e, uniform), or a Gaussian or mixture of Gaussians, matching the first two moments of the generated sample.

Several interesting observations emerged from our tests:

- The CESO and CEST heuristics can actually have noticeably different performance. In particular, while it is easy to think of instances when the latter improves over the former (i.e., resolving the problem increases the objective), we chose this particular example to show that the reverse case can actually hold, as well[8].
- Adjustable robust policies deliver very good performance, while open-loop formulations are considerably worse (note the average gap of 24% under all the testing distributions). The SAA and CESO heuristics also deliver very good performance, and are quite close to the affine policies (average gaps of $1 - 2\%$). As with our simulations for the single-item case, these gaps tend to become more pronounced when using distributions with larger variance or wider supports.
- Removing the markdown constraint resulted in more instances with larger optimality gaps from the PH solution, as well as larger gaps between the heuristics

[8]The main reason for the behavior here, which became obvious once the pricing sequences from the two heuristics were examined, is the following: since the residuals ε_t in successive periods are strongly, negatively correlated, when the CE is resolved in a particular period (e.g., an odd period), it can respond to a large residual in the preceding period, and adjust prices disproportionately in the wrong direction (since it cannot anticipate the fact that the residual in the succeeding period will have an opposite sign).

and the adjustable robust policies. The reason is intuitively clear, as not having a cap on the prices is more valuable (in relative terms) for an adjustable policy, than it is for open-loop formulations.

4.6.3 Real Data

A similar behavior was observed when testing with the real data. As an example, Table C.10 records the relative gaps from policies of degree $d = 0$ (open-loop), obtained for data in Department 2, Subclass 1. In this case, it can be noticed that adjustable policies with $d = 1$ and the CESO, CEST and SAA policies deliver comparable results, better than open-loop robust policies, and the ALY heuristic.

Chapter 5

Conclusions and Future Research

In this dissertation, we have discussed several theoretical and computational aspects related to disturbance-feedback policies in multi-period robust optimization, and have explored several potential applications to problems in inventory and revenue management.

In Chapter 2, we introduced a novel theoretical result concerning the optimality of *affine* disturbance-feedback policies, in the context of a one-dimensional, constrained, multi-period dynamical system. Our proof technique strongly utilized the connections between the geometrical properties of the feasible sets (zonogons), and the objective functions being optimized, in order to prune the set of relevant points and derive properties that the optimal policies for the problem should obey. We have also shown an interesting implication of our theoretical results in the context of a classical problem in inventory management, consisting of a single (risk-averse) retailer replenishing inventory in the face of unknown demand.

Chapter 3 then proceeded to introduce an extension of the affine policies to multi-dimensional linear dynamical systems, by considering a hierarchy of polynomial disturbance-feedback policies, parametrized by the degree d. We showed how the problem of computing such policies can be reformulated as a semi-definite program, and hence solved efficiently by interior point methods. To test the quality of the policies, we considered two applications in inventory management, and noted that quadratic policies (requiring modest computational requirements) were able to substantially reduce the optimality gap, while cubic policies (under more computational requirements) were always within 1% of the optimal solution.

Finally, Chapter 4 considered a different version of a multi-period dynamical system, arising in the context of dynamic pricing applications in revenue management.For the multi-product case, under a linear demand function, we proposed a distributionally robust model for the uncertainties, and argued how it can be constructed from limited historical data. We then considered polynomial pricing policies parameterized directly in the observed model mis-specifications, and showed how these can be computed by solving second-order conic or semidefinite programming problems. Extensive simulation results on both real and synthetic data allowed us to conclude that considering adjustable policies (versus open-loop formulations) considerably improves the quality of the objective, and yields pricing policies that are competitive with some of the popular heuristics in the literature.

On a theoretical level, one immediate direction of future research would be to explore potential generalizations of the optimality result in Chapter 2 to non-trivial multi-dimensional systems. It would also be worthwhile to get a better understanding of the connections between the matching performed in Algorithm 1 and Algorithm 2 and the properties of convex (or supermodular) functions, as well as to explore extensions of the approach to handle different cost functions. Another potential development would be to use our analysis tools to quantify the performance of affine or polynomial policies even in problems where they are known to be suboptimal. This could potentially lead to fast approximation algorithms, with solid theoretical foundations.

In a different sense, our research thus far suggests that multi-stage, worst-case oriented decision making, results, in a fundamental sense, in "simpler" optimization problems than stochastic decision-making (recall that, even in the simple example of Chapter 2, disturbance-affine policies are severely suboptimal for the latter problem). However, this type of "freedom" cannot be explored if one uses Dynamic Programming formulations to solve the resulting problems! Thus, new theoretical tools have to be developed, which are capable of exploiting this very property when computing optimal actions at every stage. This may yield very interesting results in terms of the structure and properties of the solution (in particular, it may well be that optimal policies in robust decision making have far simpler structure than their stochastic counterparts...)

On a more practical level, it would be interesting to explore connections between the robust optimization formulations that we have seen in this thesis and several problems arising in risk management and risk-adjusted decision making. In particular, several recent developments in the literature on coherent risk measures, combined with some of the techniques that we developed for multi-stage, worst-case oriented decision making, might provide novel ways of modeling and solving large-scale risk-adjusted decision problems, with very interesting applications in a variety of fields, from operations to financial engineering. In this direction, a key development would be to better understand (a) how to translate particular business goals into risk-adjusted objectives, and (b) how to construct uncertainty sets that correspond to the respective objectives, and that remain tractable for multi-period, adjustable optimization.

Appendix A

Appendix for Chapter 2

A.1 Dynamic Programming Solution.

This section contains a detailed proof for the solution of the Dynamic Programming formulation, initially introduced in Section 2.2. Recall that the problem we would like to solve is the following:

$$\min_{u_1}\Bigg[c_1\,u_1+\max_{w_1}\Big[h_1(x_2)+\cdots+\min_{u_k}\Big[c_k\,u_k+\max_{w_k}\Big[h_k(x_{k+1})+\dots$$

$$+\min_{u_T}\Big[c_T\,u_T+\max_{w_T}\;h_T(x_{T+1})\Big]\dots\Big]\Bigg]$$

$$\begin{aligned}
&\text{s.t. } x_{k+1}=x_k+u_k+w_k\\
(DP)\qquad &\qquad L_k\le u_k\le U_k \qquad\qquad \forall\,k\in\{1,2,\dots,T\}\\
&\qquad w_k\in\mathcal{W}_k=[\underline{w}_k,\overline{w}_k],
\end{aligned}$$

which gives rise to the corresponding Bellman recursion:

$$J_k^\star(x_k)\stackrel{\text{def}}{=}\min_{L_k\le u_k\le U_k}\Bigg[c_k\,u_k+\max_{w_k\in\mathcal{W}_k}\Big[\;h_k(x_k+u_k+w_k)+J_{k+1}^\star\left(x_k+u_k+w_k\right)\Big]\Bigg].$$

According to our definition of running cost and cost-to-go, the cost at $T+1$ is $J^\star_{T+1} = 0$, which yields the following Bellman recursion at time T:

$$J^\star_T(x_T) \stackrel{\text{def}}{=} \min_{L_T \le u_T \le U_T} \Big[c_T \cdot u_T + \max_{w_T \in \mathcal{W}_T} h_T(x_T + u_T + w_T) \Big].$$

First consider the inner (maximization) problem. Letting $y_T \stackrel{\text{def}}{=} x_T + u_T$, we obtain:

$$\begin{aligned} g_T(y_T) &\stackrel{\text{def}}{=} \max_{w_T \in [\underline{w}_T, \overline{w}_T]} h_T(x_T + u_T + w_T) \\ \text{(since } h_T(\cdot) \text{ convex)} \quad &= \max \left\{ h_T\left(y_T + \underline{w}_T\right), h_T\left(y_T + \overline{w}_T\right) \right\}. \end{aligned} \tag{A.1}$$

Note that g_T is the maximum of two convex functions of y_T, hence it is also convex (see [126]). The outer (minimization) problem at time T becomes:

$$\begin{aligned} J^\star_T(x_T) &= \min_{L_T \le u_T(\cdot) \le U_T} c_T \cdot u_T + g_T(x_T + u_T) \\ &= -c_T \cdot x_T + \min_{L_T \le u_T(\cdot) \le U_T} \left[\, c_T \cdot (x_T + u_T) + g_T(x_T + u_T) \, \right] \end{aligned}$$

For any x_T, $c_T \cdot (x_T + u_T) + g_T(x_T + u_T)$ is a convex function of its argument $y_T = x_T + u_T$. As such, by defining $y^\star_T$ to be the minimizer[1] of the convex function $c_T \cdot y + g_T(y)$, we obtain that the optimal controller and optimal value function at time T will be:

$$u^\star_T(x_T) = \begin{cases} U_T, & \text{if } x_T < y^\star_T - U_T \\ -x_T + y^\star_T, & \text{otherwise} \\ L_T, & \text{if } x_T > y^\star_T - L_T \end{cases} \tag{A.2}$$

$$J^\star_T(x_T) = \begin{cases} c_T \cdot U_T + g_T(x_T + U_T), & \text{if } x_T < y^\star_T - U_T \\ c_T \cdot (y^\star_T - x_T) + g_T(y^\star_T), & \text{otherwise} \\ c_T \cdot L_T + g_T(x_T + L_T), & \text{if } x_T > y^\star_T - L_T. \end{cases} \tag{A.3}$$

The following properties are immediately obvious:

[1] We assume, again, that the minimizer is unique. The results can be extended to a compact set of minimizers, $[\underline{y}_T, \overline{y}_T]$.

1. $u_T^\star(x_T)$ is piecewise affine (with at most 3 pieces), continuous, monotonically decreasing in x_T.

2. $J_T^\star(x_T)$ is convex, since it represents a partial minimization of a convex function with respect to one of the variables (see Proposition 2.3.6 in [24]).

The results can be immediately extended by induction on k:

Lemma 12. *The optimal control policy $u_k^\star(x_k)$ is piecewise affine, with at most 3 pieces, continuous, and monotonically decreasing in x_k. The optimal objective function $J_k^\star(x_k)$ is convex in x_t.*

Proof. The induction is checked at $k = T$. Assume the property is true at $k+1$. Letting $y_k \stackrel{\text{def}}{=} x_k + u_k$, the Bellman recursion at k becomes:

$$\begin{aligned} J_k^\star(x_k) &\stackrel{\text{def}}{=} \min_{L_k \le u_k \le U_k} \Big[\ c_k \cdot u_k + g_k\left(x_k + u_k\right)\ \Big] \\ g_k\left(y_k\right) &\stackrel{\text{def}}{=} \max_{w_k \in \mathcal{W}_k} \Big[\ h_k(y_k + w_k) + J_{k+1}^\star\left(y_k + w_k\right)\ \Big]. \end{aligned}$$

Consider first the maximization problem. Since h_k is convex, and (by the induction hypothesis) $J_{k+1}^\star$ is also convex, the maximum will be reached on the boundary of $\mathcal{W}_k = [\underline{w}_k, \overline{w}_k]$,

$$g_k(y_k) = \max_{w_k \in \{\underline{w}_k, \overline{w}_k\}} \big[\, h_k(y_k + w_k) + J_{k+1}^\star\left(y_k + w_k\right)\big], \tag{A.4}$$

and $g_k(y_k)$ will be also be convex. The minimization problem becomes:

$$\begin{aligned} J_k^\star(x_k) &= \min_{L_k \le u_k \le U_k} \ [\ c_k \cdot u_k + g_k\left(x_k + u_k\right)\ ] \\ &= -c_k \cdot x_k + \min_{L_k \le u_k \le U_k} \ [\ c_k \cdot (x_k + u_k) + g_k\left(x_k + u_k\right)\ ] \end{aligned} \tag{A.5}$$

Defining, as before, $y_k^\star$ as the minimizer of $c_k \cdot y + g_k(y)$, we get:

$$u_k^\star(x_k) = \begin{cases} U_k, & \text{if } x_k < y_k^\star - U_k \\ -x_k + y_k^\star, & \text{otherwise} \\ L_k, & \text{if } x_k > y_k^\star - L_k \end{cases} \tag{A.6}$$

$$J_k^\star(x_k) = \begin{cases} c_k \cdot U_k + g_k(x_k + U_k), & \text{if } x_k < y_k^\star - U_k \\ c_k \cdot (y_k^\star - x_k) + g_k(y_k^\star), & \text{otherwise} \\ c_k \cdot L_k + g_k(x_k + L_k), & \text{if } x_k > y_k^\star - L_k. \end{cases} \tag{A.7}$$

In particular, $u_k^\star$ will be piecewise affine with 3 pieces, continuous, monotonically decreasing, and $J_k^\star$ will be convex (as the partial minimization of a convex function with respect to one of the variables). A typical example of the optimal control law and the optimal value function is shown in Figure 2-1 of Section 2.2. □

A.2 Zonotopes and Zonogons.

In this section of the Appendix, we would like to outline several useful properties of the main geometrical objects of interest in our exposition, namely *zonotopes*. The presentation here parallels that in Chapter 7 of [149], to which the interested reader is referred for a much more comprehensive treatment.

Zonotopes are special polytopes that can be viewed in various ways: as projections of hypercubes, as Minkowski sums of line segments, and as sets of bounded linear combinations of vector configurations. Each description gives a different insight into the combinatorics of zonotopes, and there exist some very interesting results that unify the different descriptions under a common theory. For our purposes, it will be sufficient to understand zonotopes under the first two descriptions. In particular, letting $\mathcal{H}_k$ denote the k-dimensional hypercube, $\mathcal{H}_k = \{\boldsymbol{w} \in \mathbb{R}^k : 0 \le w_i \le 1, \forall\, i\}$, we can introduce the following definition:

Definition 3 (7.13 in [149])**.** *A* ***zonotope*** *is the image of a hypercube under an affine* *projection, that is, a d-polytope $Z \subseteq \mathbb{R}^d$ of the form*

$$Z = Z(V) := V \cdot \mathcal{H}_k + \boldsymbol{z} = \{V\boldsymbol{w} + \boldsymbol{z} \ : \ \boldsymbol{w} \in \mathcal{H}_k\}$$
$$= \{\boldsymbol{x} \in \mathbb{R}^d \ : \ \boldsymbol{x} = \boldsymbol{z} + \sum_{i=1}^{k} w_i \boldsymbol{v}_i, \ 0 \leq w_i \leq 1\}$$

for some matrix (vector configuration) $V = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) \in \mathbb{R}^{d \times k}$ and some $\boldsymbol{z} \in \mathbb{R}^d$.

The rows of the matrix V are often referred to as the *generators* defining the zonotope. An equivalent description of the zonotope can be obtained by recalling that every k-cube $\mathcal{H}_k$ is a product of line segments $\mathcal{H}_k = \mathcal{H}_1 \times \cdots \times \mathcal{H}_1$. Since for a linear operator π we always have: $\pi(\mathcal{H}_1 \times \cdots \times \mathcal{H}_1) = \pi(\mathcal{H}_1) + \cdots + \pi(\mathcal{H}_1)$, by considering an affine map given by $\pi(\boldsymbol{w}) = V\boldsymbol{w} + \boldsymbol{z}$, it is easy to see that every zonotope is the Minkowski sum of a set of line segments:

$$Z(V) = [0, \boldsymbol{v}_1] + \cdots + [0, \boldsymbol{v}_k] + \boldsymbol{z}.$$

For completeness, we remark that there is no loss of generality in regarding a zonotope as a projection from the unit hypercube $\mathcal{H}_k$, since any projection from an arbitrary hyperrectangle in $\mathbb{R}^k$ can be seen as a projection from the unit hypercube in $\mathbb{R}^k$. To see this, consider an arbitrary hyperrectangle in $\mathbb{R}^k$:

$$\mathcal{W}_k = [\underline{w}_1, \overline{w}_1] \times [\underline{w}_2, \overline{w}_2] \times \cdots \times [\underline{w}_k, \overline{w}_k],$$

and note that, with $V \in \mathbb{R}^{d \times k}$, and $\boldsymbol{a}' \in \mathbb{R}^k$ denoting the j-th row of V, the j-th component of $Z(V) \stackrel{\text{def}}{=} V \cdot \mathcal{W}_k + \boldsymbol{z}$ can be written:

$$Z(V)_j \stackrel{\text{def}}{=} z_j + \sum_{i=1}^{k} (a_i \cdot w_i) = \left(z_j + \sum_{i=1}^{k} a_i \cdot \underline{w}_i \right) + \sum_{i=1}^{k} a_i \cdot (\overline{w}_i - \underline{w}_i) \cdot y_i,$$

where $y_i \in [0, 1]$, $\forall\, 1 \leq i \leq k$.

An example of a subclass of zonotopes are the *zonogons*, which are all centrally symmetric, two-dimensional $2p$-gons, arising as the projection of p-cubes to the plane.

An example is shown in Figure 2-2 of Section 2.4.1. These are the main objects of interest in our treatment, and the following lemma summarizes their most important properties:

Lemma 13. *Let* $\mathcal{H}_k = [0,1]^k$ *be a* k*-dimensional hypercube,* $k \geq 2$*. For fixed* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^k$ *and* $a_0, b_0 \in \mathbb{R}$*, consider the affine transformation* $\pi : \mathbb{R}^k \to \mathbb{R}^2$*,* $\pi(\boldsymbol{w}) = \begin{bmatrix} \boldsymbol{a}' \\ \boldsymbol{b}' \end{bmatrix} \cdot \boldsymbol{w} + \begin{bmatrix} a_0 \\ b_0 \end{bmatrix}$ *and the zonogon* $\Theta \subset \mathbb{R}^2$*:*

$$\Theta = \pi(\mathcal{H}_k) \stackrel{\text{def}}{=} \left\{ \boldsymbol{\theta} \in \mathbb{R}^2 : \exists\, \boldsymbol{w} \in \mathcal{H}_k \ s.t.\ \boldsymbol{\theta} = \pi(\boldsymbol{w}) \right\}.$$

If we let $\mathcal{V}_\Theta$ *denote the set of vertices of* Θ*, then the following properties are true:*

1. $\exists\, \boldsymbol{O} \in \Theta$ *such that* Θ *is symmetric around* $\boldsymbol{O}$: $\forall\, \boldsymbol{x} \in \Theta \Rightarrow 2\boldsymbol{O} - \boldsymbol{x} \in \Theta$.

2. $|\mathcal{V}_\Theta| = 2p \leq 2k$ *vertices. Also,* $p < k$ *if and only if* $\exists\, i \neq j \in \{1, \ldots, k\}$ *such that* $\operatorname{rank}\left(\begin{bmatrix} a_i & a_j \\ b_i & b_j \end{bmatrix} \right) < 2$.

3. *If we number the vertices of* $\mathcal{V}_\Theta$ *in cyclic order:*

$$\mathcal{V}_\Theta = (\boldsymbol{v}_0, \ldots, \boldsymbol{v}_i, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}_{2p-1}) \quad (\boldsymbol{v}_{2p+i} \stackrel{\text{def}}{=} \boldsymbol{v}_{(2p+i) \mod (2p)})$$

then $2\boldsymbol{O} - \boldsymbol{v}_i = \boldsymbol{v}_{i+p}$*, and we have the following representation for* Θ *as a Minkowski sum of line segments:*

$$\begin{aligned} \Theta &= \boldsymbol{O} + \left[-\frac{\boldsymbol{v}_1 - \boldsymbol{v}_0}{2}, \frac{\boldsymbol{v}_1 - \boldsymbol{v}_0}{2} \right] + \cdots + \left[-\frac{\boldsymbol{v}_p - \boldsymbol{v}_{p-1}}{2}, \frac{\boldsymbol{v}_p - \boldsymbol{v}_{p-1}}{2} \right] \\ &\stackrel{\text{def}}{=} \boldsymbol{O} + \sum_{i=1}^{p} \lambda_i \cdot \frac{\boldsymbol{v}_i - \boldsymbol{v}_{i-1}}{2}, \quad -1 \leq \lambda_i \leq 1. \end{aligned}$$

4. *If* $\exists\, \boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathcal{H}_k$ *such that* $\boldsymbol{v}_1 \stackrel{\text{def}}{=} \pi(\boldsymbol{w}_1) = \boldsymbol{v}_2 \stackrel{\text{def}}{=} \pi(\boldsymbol{w}_2)$ *and* $\boldsymbol{v}_{1,2} \in \mathcal{V}_\Theta$*, then* $\exists\, j \in \{1, \ldots, k\}$ *such that* $a_j = b_j = 0$.

5. *With the same numbering from (iii) and* $k = p$*, for any* $i \in \{0, \ldots, 2p-1\}$*, the vertices of the hypercube that are projecting to* $\boldsymbol{v}_i$ *and* $\boldsymbol{v}_{i+1}$*, respectively, are adjacent, i.e., they only differ in exactly one component.*

Proof. We will omit a complete proof of the lemma, and will instead simply suggest the main ideas needed for checking the validity of the statements.

For part (i), it is easy to argue that the center of the hypercube, $\boldsymbol{O}_{\mathcal{H}} = [1/2, 1/2, \ldots, 1/2]'$, will always project into the center of the zonogon, i.e., $\boldsymbol{O} = \pi(\boldsymbol{O}_{\mathcal{H}})$. This implies that any zonogon will be centrally symmetric, and will therefore have an even number of vertices.

Part (ii) can be shown by induction on the dimension k of the hypercube, $\mathcal{H}_k$. For instance, to prove the first claim, note that the projection of a polytope is simply the convex hull of the projections of the vertices, and therefore projecting a hypercube of dimension k simply amounts to projecting two hypercubes of dimension $k-1$, one for $w_k = 0$ and another for $w_k = 1$, and then taking the convex hull of the two resulting polytopes. It is easy to see that these two polytopes in $\mathbb{R}^2$ are themselves zonogons, and are translated copies of each other (by an amount $[a_k, b_k]'$). Therefore, by the induction hypothesis, they have at most $2(k-1)$ vertices, and taking their convex hull introduces at most two new vertices, for a total of at most $2(k-1)+2 = 2k$ vertices. The second claim can be proved in a similar fashion.

One way to prove part (iii) is also by induction on p, by taking any pair of opposite (i.e., parallel, of the same length) edges and showing that they correspond to a Minkowski summand of the zonogon.

Part (iv) also follows by induction. Using the same argument as for part (ii), note that the only ways to have two distinct vertices of the hypercube $\mathcal{H}_k$ (of dimension k) project onto the same vertex of the zonogon Θ is to either have this situation happen for one of the two $k-1$ dimensional hypercubes (in which case the induction hypothesis would complete the proof), or to have zero translation between the two zonogons, which could only happen if $a_k = b_k = 0$.

Part (v) follows by using parts (iii) and (iv) and the definition of a zonogon as the Minkowski sum of line segments. In particular, since the difference between two consecutive vertices of the zonogon, $\boldsymbol{v}_i, \boldsymbol{v}_{i+1}$, for the case $k = p$, is always given by a single column of the projection matrix (i.e., $[a_j, b_j]'$, for some j), then the unique vertices of $\mathcal{H}_k$ that were projecting onto $\boldsymbol{v}_i$ and $\boldsymbol{v}_{i+1}$, respectively, must be incidence vectors that differ in exactly one component, i.e., are adjacent on the hypercube $\mathcal{H}_k$.

□

A.2.1 Technical Lemmas.

This section of the Appendix contains a detailed proof for the technical Lemma 3 introduced in Section 2.4.1, which we include below, for convenience.

LEMMA 3. *When the zonogon Θ has a non-trivial intersection with the band $\mathcal{B}_{LU}$ (case* [**C4**]*), the convex polygon $\Delta_{\Gamma^\star}$ and the set of points on its right side,* r-side$(\Delta_{\Gamma^\star})$*, satisfy the following properties:*

1. r-side$(\Delta_{\Gamma^\star})$ *is the union of two sequences of consecutive vertices (one starting at $\boldsymbol{y}_0^\star$, and one ending at $\boldsymbol{y}_k^\star$), and possibly an additional vertex, $\boldsymbol{y}_t^\star$:*

$$\text{r-side}(\Delta_{\Gamma^\star}) = \{\boldsymbol{y}_0^\star, \boldsymbol{y}_1^\star, \dots, \boldsymbol{y}_s^\star\} \cup \{\boldsymbol{y}_t^\star\} \cup \left\{\boldsymbol{y}_r^\star, \boldsymbol{y}_{r+1}^\star \dots, \boldsymbol{y}_k^\star\right\},$$

   for some $s \leq r \in \{0, \dots, k\}$.

2. *With* $\text{cotan}\big(\cdot,\, \cdot\big)$ *given by* (2.22) *applied to the* $(\gamma_1^\star, \gamma_2^\star)$ *coordinates, we have that:*

$$\begin{cases} \text{cotan}\big(\boldsymbol{y}_s^\star,\, \boldsymbol{y}_{\min(t,r)}^\star\big) \geq \frac{a_{s+1}}{b_{s+1}}, & \text{whenever } t > s \\ \text{cotan}\big(\boldsymbol{y}_{\max(t,s)}^\star,\, \boldsymbol{y}_r^\star\big) \leq \frac{a_r}{b_r}, & \text{whenever } t < r. \end{cases}$$

Lemma 3. In the following exposition, we use the same notation as introduced in Section 2.4.1. Recall that case [**C4**] on which the lemma is focused corresponds to a nontrivial intersection of the zonotope Θ with the horizontal band $\mathcal{B}_{LU}$ defined in (2.28). As suggested in Figure 2-5 of Section 2.4.1, this case can be separated into three subcases, depending on the position of the vertex $\boldsymbol{v}_t$ relative to the band $\mathcal{B}_{LU}$, where the index t is defined in (2.29). Since the proof of all three cases is essentially identical, we will focus on the more "complicated" situation, namely when $\boldsymbol{v}_t \in \mathcal{B}_{LU}$. The corresponding arguments for the other two cases should be straightforward.

First, recall that $\Delta_{\Gamma^\star}$ is given by (2.26), i.e., $\Delta_{\Gamma^\star} = \text{conv}\left(\{\boldsymbol{y}_0^\star, \dots, \boldsymbol{y}_k^\star\}\right)$, where the points $\boldsymbol{y}_i^\star$ are given by (2.27), which results from applying mapping (2.25) to $\boldsymbol{v}_i \in \Theta$. From Definition 1 of the *right side*, it can be seen that the points of interest to us, namely r-side$(\Delta_{\Gamma^\star})$, will be a maximal subset $\{\boldsymbol{y}_{i(1)}^\star, \boldsymbol{y}_{i(2)}^\star, \ldots, \boldsymbol{y}_{i(m)}^\star\} \subseteq \{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_k^\star\}$, satisfying:

$$\begin{cases} \boldsymbol{y}_{i(1)}^\star = \arg\max\Big\{\gamma_1 : \boldsymbol{\gamma} \in \arg\min\big\{\gamma_2' : \boldsymbol{\gamma}_2' \in \{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_k^\star\}\big\}\Big\} \\ \boldsymbol{y}_{i(m)}^\star = \arg\max\Big\{\gamma_1 : \boldsymbol{\gamma} \in \arg\max\big\{\gamma_2' : \boldsymbol{\gamma}_2' \in \{\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_k^\star\}\big\}\Big\} \\ \text{cotan}\big(\boldsymbol{y}_{i(1)}^\star,\, \boldsymbol{y}_{i(2)}^\star\big) > \text{cotan}\big(\boldsymbol{y}_{i(2)}^\star,\, \boldsymbol{y}_{i(3)}^\star\big) > \cdots > \text{cotan}\big(\boldsymbol{y}_{i(m-1)}^\star,\, \boldsymbol{y}_{i(m)}^\star\big). \end{cases} \tag{A.8}$$

For the analysis, we find it useful to define the following two indices:

$$\begin{aligned} \hat{s} &\stackrel{\text{def}}{=} \min\big\{i \in \{0, \ldots, k\} : \theta_2(\boldsymbol{v}_i) \geq y^\star - U\big\}, \\ \hat{r} &\stackrel{\text{def}}{=} \max\big\{i \in \{0, \ldots, k\} : \theta_2(\boldsymbol{v}_i) \leq y^\star - L\big\}. \end{aligned} \tag{A.9}$$

In particular, $\hat{s}$ is the index of the first vertex of r-side(Θ) falling inside $\mathcal{B}_{LU}$, and $\hat{r}$ is the index of the last vertex of r-side(Θ) falling inside $\mathcal{B}_{LU}$. Since we are in the situation when $\boldsymbol{v}_t \in \mathcal{B}_{LU}$, it can be seen that $0 \leq \hat{s} \leq t \leq \hat{r} \leq k$, and thus, from (2.29) (the definition of t) and (2.30) (typical conditions for the right side of a zonogon):

$$\frac{a_1}{b_1} > \cdots > \frac{a_{\hat{s}}}{b_{\hat{s}}} > \cdots > \frac{a_t}{b_t} > c \geq \frac{a_{t+1}}{b_{t+1}} > \cdots > \frac{a_{\hat{r}}}{b_{\hat{r}}} > \cdots > \frac{a_k}{b_k}. \tag{A.10}$$

With this new notation, we proceed to prove the first result in the claim. First, consider all the vertices $\boldsymbol{v}_i \in$ r-side(Θ) falling strictly below the band $\mathcal{B}_{LU}$, i.e., satisfying $\theta_2[\boldsymbol{v}_i] < y^\star - U$. From the definition of $\hat{s}$, (A.9), these are exactly $\boldsymbol{v}_0, \ldots, \boldsymbol{v}_{\hat{s}-1}$, and mapping (2.25) applied to them will yield $\boldsymbol{y}_i^\star = (\ \theta_1[\boldsymbol{v}_i] + c \cdot U,\ \theta_2[\boldsymbol{v}_i] + U\ )$. In other words, any such points will simply be translated by $(c \cdot U, U)$. Similarly, any points $\boldsymbol{v}_i \in$ r-side(Θ) falling strictly above the band $\mathcal{B}_{LU}$, i.e., $\theta_2[\boldsymbol{v}_i] > y^\star - L$, will be translated by $(c \cdot L, L)$, so that we have:

$$\begin{aligned} \boldsymbol{y}_i^\star &= \boldsymbol{v}_i + (c \cdot U, U), \quad i \in \{0, \ldots, \hat{s} - 1\}, \\ \boldsymbol{y}_i^\star &= \boldsymbol{v}_i + (c \cdot L, L), \quad i \in \{\hat{r} + 1, \ldots, k\}, \end{aligned} \tag{A.11}$$

which immediately implies, since $\boldsymbol{v}_i \in$ r-side(Θ), that:

$$\begin{cases} \operatorname{cotan}\big(\boldsymbol{y}_0^\star,\, \boldsymbol{y}_1^\star\big) > \operatorname{cotan}\big(\boldsymbol{y}_1^\star,\, \boldsymbol{y}_2^\star\big) > \cdots > \operatorname{cotan}\big(\boldsymbol{y}_{\hat{s}-2}^\star,\, \boldsymbol{y}_{\hat{s}-1}^\star\big), \\ \operatorname{cotan}\big(\boldsymbol{y}_{\hat{r}+1}^\star,\, \boldsymbol{y}_{\hat{r}+2}^\star\big) > \operatorname{cotan}\big(\boldsymbol{y}_{\hat{r}+2}^\star,\, \boldsymbol{y}_{\hat{r}+3}^\star\big) > \cdots > \operatorname{cotan}\big(\boldsymbol{y}_{k-1}^\star,\, \boldsymbol{y}_k^\star\big). \end{cases} \tag{A.12}$$

For any vertices inside $\mathcal{B}_{LU}$, i.e., $\boldsymbol{v}_i \in$ r-side$(\Theta) \cap \mathcal{B}_{LU}$, mapping (2.25) will yield:

$$\boldsymbol{y}_i^\star = (\ \theta_1[\boldsymbol{v}_i] - c \cdot \theta_2[\boldsymbol{v}_i] + c \cdot y^\star,\ y^\star\ ), \quad i \in \{\hat{s}, \ldots, t, \ldots, \hat{r}\}, \tag{A.13}$$

that is, they will be mapped into points with the same $\gamma_2^\star$ coordinates. Furthermore, using (2.21), it can be seen that $\boldsymbol{y}_t^\star$ will have the largest $\gamma_1^\star$ coordinate among all such $\boldsymbol{y}_i^\star$:

$$\begin{aligned} \gamma_1^\star[\boldsymbol{y}_t^\star] - \gamma_1^\star[\boldsymbol{y}_i^\star] &\overset{\text{def}}{=} \theta_1[\boldsymbol{v}_t] - \theta_1[\boldsymbol{v}_i] - c \cdot (\theta_2[\boldsymbol{v}_t] - \theta_2[\boldsymbol{v}_i]) \\ &\overset{(2.21)}{=} \begin{cases} \sum_{j=i+1}^{t} a_j - c \cdot \sum_{j=i+1}^{t} b_j \overset{(A.10)}{\geq} 0, & \text{if } \hat{s} \leq i < t \\ -\sum_{j=t+1}^{i} a_j + c \cdot \sum_{j=t+1}^{i} b_j \overset{(A.10)}{\geq} 0, & \text{if } t < i \leq \hat{r}. \end{cases} \end{aligned} \tag{A.14}$$

Furthermore, since the mapping (2.25) yielding $\gamma_2^\star$ is only a function of θ_2, and is monotonic non-decreasing (strictly monotonic increasing outside the band $\mathcal{B}_{LU}$), vertices $\boldsymbol{v}_0, \ldots, \boldsymbol{v}_k \in$ r-side(Θ) will be mapped into points $\boldsymbol{y}_0^\star, \ldots, \boldsymbol{y}_k^\star \in \gamma^\star$ with non-decreasing $\gamma_2^\star$ coordinates:

$$\begin{aligned} &\gamma_2^\star[\boldsymbol{y}_0^\star] < \gamma_2^\star[\boldsymbol{y}_1^\star] < \cdots < \gamma_2^\star[\boldsymbol{y}_{\hat{s}-1}^\star] < \\ &\qquad < y^\star = \gamma_2^\star[\boldsymbol{y}_{\hat{s}}^\star] = \cdots = \gamma_2^\star[\boldsymbol{y}_t^\star] = \cdots = \gamma_2^\star[\boldsymbol{y}_{\hat{r}}^\star] < \gamma_2^\star[\boldsymbol{y}_{\hat{r}+1}^\star] < \cdots < \gamma_2^\star[\boldsymbol{y}_k^\star]. \end{aligned}$$

Therefore, combining this fact with (A.12) and (A.14), we can conclude that the points $\boldsymbol{y}_i^\star$ satisfying conditions (A.8) are none other than:

$$\text{r-side}(\Delta_{\Gamma^\star}) = \left\{\boldsymbol{y}_0^\star, \boldsymbol{y}_1^\star, \ldots, \boldsymbol{y}_s^\star, \boldsymbol{y}_t^\star, \boldsymbol{y}_r^\star, \boldsymbol{y}_{r+1}^\star, \boldsymbol{y}_k^\star\right\},$$

where the indices s and r are given as:

$$
\begin{aligned}
s &\stackrel{\text{def}}{=} \begin{cases} \max\left\{i \in \{1,\ldots,\hat{s}-1\} \;:\; \operatorname{cotan}\big(\boldsymbol{y}^\star_{i-1},\, \boldsymbol{y}^\star_i\big) > \operatorname{cotan}\big(\boldsymbol{y}^\star_i,\, \boldsymbol{y}^\star_t\big)\right\} \\ 0, \text{ if the above condition is never true,} \end{cases} \\
r &\stackrel{\text{def}}{=} \begin{cases} \min\left\{i \in \{\hat{r}+1,\ldots,k-1\} \;:\; \operatorname{cotan}\big(\boldsymbol{y}^\star_t,\, \boldsymbol{y}^\star_i\big) > \operatorname{cotan}\big(\boldsymbol{y}^\star_i,\, \boldsymbol{y}^\star_{i+1}\big)\right\} \\ k, \text{ if the above condition is never true.} \end{cases}
\end{aligned} \tag{A.15}
$$

This completes the proof of part (i) of the Lemma. We remark that, for the cases when $\boldsymbol{v}_t$ falls strictly below $\mathcal{B}_{LU}$ or strictly above $\mathcal{B}_{LU}$, one can repeat the exact same reasoning, and immediately argue that the same result would hold.

In order to prove the first claim in part (ii), we first recall that, from (A.15), if $s < \hat{s} - 1$, we must have:

$$\operatorname{cotan}\big(\boldsymbol{y}^\star_s,\, \boldsymbol{y}^\star_{s+1}\big) \le \operatorname{cotan}\big(\boldsymbol{y}^\star_{s+1},\, \boldsymbol{y}^\star_t\big),$$

since otherwise, we would have taken $s+1$ instead of s in (A.15). But this immediately implies that:

$$\operatorname{cotan}\big(\boldsymbol{y}^\star_s,\, \boldsymbol{y}^\star_{s+1}\big) \le \operatorname{cotan}\big(\boldsymbol{y}^\star_{s+1},\, \boldsymbol{y}^\star_t\big) \;\stackrel{(2.22)}{\Leftrightarrow}\; \frac{\gamma^\star_1[\boldsymbol{y}^\star_{s+1}] - \gamma^\star_1[\boldsymbol{y}^\star_s]}{\gamma^\star_2[\boldsymbol{y}^\star_{s+1}] - \gamma^\star_2[\boldsymbol{y}^\star_s]} \le \frac{\gamma^\star_1[\boldsymbol{y}^\star_t] - \gamma^\star_1[\boldsymbol{y}^\star_{s+1}]}{\gamma^\star_2[\boldsymbol{y}^\star_t] - \gamma^\star_1[\boldsymbol{y}^\star_{s+1}]},$$

which, by the mediant inequality, then implies

$$\frac{\gamma^\star_1[\boldsymbol{y}^\star_{s+1}] - \gamma^\star_1[\boldsymbol{y}^\star_s]}{\gamma^\star_2[\boldsymbol{y}^\star_{s+1}] - \gamma^\star_2[\boldsymbol{y}^\star_s]} \le \frac{\gamma^\star_1[\boldsymbol{y}^\star_t] - \gamma^\star_1[\boldsymbol{y}^\star_s]}{\gamma^\star_2[\boldsymbol{y}^\star_t] - \gamma^\star_1[\boldsymbol{y}^\star_s]} \;\stackrel{(A.11)}{\Leftrightarrow}\; \frac{a_{s+1}}{b_{s+1}} \le \operatorname{cotan}\big(\boldsymbol{y}^\star_s,\, \boldsymbol{y}^\star_t\big),$$

which is exactly the first claim in part (ii). Thus, the only case to discuss is $s = \hat{s}-1$. Since $s \ge 0$, it must be that, in this case, there are vertices $\boldsymbol{v}_i \in$ r-side(Θ) falling strictly below the band $\mathcal{B}_{LU}$. Therefore, we can introduce the following point in Θ:

$$M \stackrel{\text{def}}{=} \arg\max\Big\{\, \theta_1 : (\theta_1, \theta_2) \in \Theta,\, \theta_2 = y^\star - U \,\Big\} \tag{A.16}$$

Referring back to Figure 2-6 in Section 2.4.2, it can be seen that M represents the point with smallest θ_2 coordinate in $\mathcal{B}_{LU} \cap$ r-side(Θ), and $M \in [\boldsymbol{v}_{\hat{s}-1}, \boldsymbol{v}_{\hat{s}}]$. If we let $(\theta_1[M], \theta_2[M])$ denote the coordinates of M, then by applying mapping (2.25) to M, the coordinates of the point $\tilde{M} \in \gamma^\star$ are:

$$\tilde{M} = (\, \theta_1[M] + c \cdot U,\ \theta_2[M] + U \,) = (\, \theta_1[M] + c \cdot U,\ y^\star \,)\,. \tag{A.17}$$

Furthermore, a similar argument with (A.14) can be invoked to show that $\gamma_1^\star[\tilde{M}] \leq \gamma_1^\star[\boldsymbol{y}_t^\star]$. With $s = \hat{s} - 1$, we then have:

$$\begin{aligned}
\operatorname{cotan}\big(\boldsymbol{y}_s^\star,\, \boldsymbol{y}_t^\star\big) &\overset{(2.22)}{=} \frac{\gamma_1^\star[\boldsymbol{y}_t^\star] - \gamma_1^\star[\boldsymbol{y}_{\hat{s}-1}^\star]}{\gamma_2^\star[\boldsymbol{y}_t^\star] - \gamma_2^\star[\boldsymbol{y}_{\hat{s}-1}^\star]} \geq \quad (\text{since } \gamma_2^\star[\boldsymbol{y}_t^\star] = \gamma_2^\star[\tilde{M}] = y^\star > \gamma_2^\star[\boldsymbol{y}_{\hat{s}-1}^\star]) \\
&\geq \frac{\gamma_1^\star[\tilde{M}] - \gamma_1^\star[\boldsymbol{y}_{\hat{s}-1}^\star]}{\gamma_2^\star[\tilde{M}] - \gamma_2^\star[\boldsymbol{y}_{\hat{s}-1}^\star]} \\
&\overset{(A.11),(A.17)}{=} \frac{\theta_1[M] - \theta_1[\boldsymbol{v}_{\hat{s}-1}]}{\theta_2[M] - \theta_2[\boldsymbol{v}_{\hat{s}-1}]} = \quad (\text{since } M \in [\boldsymbol{v}_{\hat{s}-1}, \boldsymbol{v}_{\hat{s}}]) \\
&= \frac{a_{s+1}}{b_{s+1}}\,,
\end{aligned}$$

which completes the proof of the first claim in part (ii).

The proof of the second claim in (ii) proceeds in an analogous fashion, by first examining the trivial case $r > \hat{r}+1$ in (A.15), and then introducing $N \overset{\text{def}}{=} \arg\max\big\{\, \theta_1 : (\theta_1, \theta_2) \in \Theta,\ \theta_2 = y^\star - L \,\big\}$ for the case $r = \hat{r} + 1$. □

Appendix B

Appendix for Chapter 3

B.1 Suboptimality of Affine Policies

Lemma 14. *Consider Problem (3.9), written below for convenience. Recall that x is a (first-stage) non-adjustable decision, while $\boldsymbol{y}$ is a second-stage adjustable* policy *(allowed to depend on $\boldsymbol{w}$).*

$$\begin{aligned}
&\underset{x,\boldsymbol{y}(\boldsymbol{w})}{\textit{minimize}} && x \\
&\textit{such that} && x \geq \sum_{i=1}^{N} y_i, \quad \forall\, \boldsymbol{w} \in \mathcal{W} = \big\{(w_1,\dots,w_N) \in \mathbb{R}^N \;:\; \|w\|_2 \leq 1\big\}, && \text{(B.1a)} \\
& && y_i \geq w_i^2, \qquad \forall\, \boldsymbol{w} \in \mathcal{W}. && \text{(B.1b)}
\end{aligned}$$

The optimal value in the problem is 1, corresponding to policies $y_i(\boldsymbol{w}) = w_i^2$, $i = 1,\dots,N$. Furthermore, the optimal achievable objective under affine *policies $\boldsymbol{y}(\boldsymbol{w})$ is N.*

Proof. Note that for any feasible $x, \boldsymbol{y}$, we have $x \geq \sum_{i=1}^{N} y_i \geq \sum_{i=1}^{N} w_i^2$, for any $\boldsymbol{w} \in \mathcal{W}$. Therefore, with $\sum_{i=1}^{N} w_i^2 = 1$, we must have $x \geq 1$. Also note that $y_i^\star(\boldsymbol{w}) = w_i^2$ is robustly feasible for constraint (B.1b), and results in an objective $x^\star = \max_{\boldsymbol{w}\in\mathcal{W}} \sum_{i=1}^{N} w_i^2 = 1$, which equals the lower bound, and is hence optimal.

Consider an affine policy in the second stage, $y_i^{\mathrm{AFF}}(\boldsymbol{w}) = \beta_i + \boldsymbol{\alpha}_i^T \boldsymbol{w}, i = 1,\dots,N$.

With $\boldsymbol{e}_1$ denoting the first unit vector (1 in the first component, 0 otherwise), for any $i = 1, \ldots, N$, we have:

$$\left.\begin{array}{r} \boldsymbol{w} = \boldsymbol{e}_1 \in \mathcal{W} \;\Rightarrow\; \beta_i + \alpha_i(1) \geq 1 \\ \boldsymbol{w} = -\boldsymbol{e}_1 \in \mathcal{W} \;\Rightarrow\; \beta_i - \alpha_i(1) \geq 1 \end{array}\right\} \;\Rightarrow\; \beta_i \geq 1.$$

This implies that $x^{\text{AFF}} \geq \sum_{i=1}^{N} y_i^{\text{AFF}}(\boldsymbol{w}) \geq N + \sum_{i=1}^{N} \boldsymbol{\alpha}_i^T \boldsymbol{w}$. In particular, with $\boldsymbol{w} = \boldsymbol{0} \in \mathcal{W}$, we have $x^{\text{AFF}} \geq N$. The optimal choice, in this case, will be to set $\boldsymbol{\alpha}_i = \boldsymbol{0}$, resulting in $x^{\text{AFF}} = N$. □

B.2 Optimality of Multi-affine Policies

Theorem 3. *Multi-affine policies of the form* (3.22)*, with degree at most* $d = T - 1$*, are optimal for problem* (P).

Proof. The following trivial observation will be useful in our analysis:

Observation 1. *A multi-affine policy* u_j *of the form* (3.22) *is an* affine *function of a given variable* w_i*, when all the other variables* w_l*,* $l \neq i$*, are fixed. Also, with* u_j *of degree at most* d*, the number of coefficients* $\ell_{\boldsymbol{\alpha}}$ *is* $\binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{d}$.

Recall that the optimal value in Problem (P) is that same as the optimal value in Problem $(P)_{\text{ext}}$ from Section 3.4.2. Let us denote the optimal decisions obtained from solving problem $(P)_{\text{ext}}$ by $\boldsymbol{u}_k^{\text{ext}}(\boldsymbol{w}_{[k]}), \boldsymbol{x}_k^{\text{ext}}(\boldsymbol{w}_{[k]})$, respectively. Note that, at time k, there are at most 2^k such distinct values $\boldsymbol{u}_k^{\text{ext}}(\boldsymbol{w}_{[k]})$, and, correspondingly, at most 2^k values $\boldsymbol{x}_k^{\text{ext}}(\boldsymbol{w}_{[k]})$, due to the non-anticipativity condition and the fact that the extreme uncertainty sequences at time k, $\boldsymbol{w}_{[k]} \in \text{ext}(\mathcal{W}_{[k]}) = \text{ext}(\mathcal{W}_0) \times \cdots \times \text{ext}(\mathcal{W}_{k-1})$, are simply the vertices of the hypercube $\mathcal{W}_{[k]} \subset \mathbb{R}^k$. In particular, at the last time when decisions are taken, $k = T - 1$, there are at most 2^{T-1} distinct optimal values $\boldsymbol{u}_{T-1}^{\text{ext}}(\boldsymbol{w}_{[T-1]})$ computed.

Consider now a multi-affine policy of the form (3.22), of degree $T-1$, implemented at time $T - 1$. By Observation 1, the number of coefficients in the j-th component of such a policy is exactly $\binom{T-1}{0} + \binom{T-1}{1} + \cdots + \binom{T-1}{T-1} = 2^{T-1}$, by Newton's binomial formula. Therefore, the total $n_u \cdot 2^{T-1}$ coefficients for $\boldsymbol{u}_{T-1}$ could be computed so that

$$\boldsymbol{u}_{T-1}(\boldsymbol{w}_{[T-1]}) = \boldsymbol{u}_{T-1}^{\text{ext}}(\boldsymbol{w}_{[T-1]}), \ \forall\ \boldsymbol{w}_{[T-1]} \in \text{ext}(\mathcal{W}_{[T-1]}), \tag{B.2}$$

that is, the value of the multi-affine policy exactly matches the 2^{T-1} optimal decisions computed in $(P)_{\text{ext}}$, at the 2^{T-1} vertices of $\mathcal{W}_{[T-1]}$. The same process can be conducted for times $k = T-2, \ldots, 1, 0$, to obtain multi-affine policies of degree at most[1] $T-1$ that match the values $\boldsymbol{u}_k^{\text{ext}}(\boldsymbol{w}_{[k]})$ at the extreme points of $\mathcal{W}_{[k]}$.

With such multi-affine control policies, it is easy to see that the states $\boldsymbol{x}_k$ become multi-affine functions of $\boldsymbol{w}_{[k]}$. Furthermore, we have $\boldsymbol{x}_k(\boldsymbol{w}_{[k]}) = \boldsymbol{x}_k^{\text{ext}}(\boldsymbol{w}_{[k]}), \forall\, \boldsymbol{w}_{[k]} \in \text{ext}(\boldsymbol{w}_{[k]})$. A typical state-control constraint (3.7c) written at time k amounts to ensuring that

$$\boldsymbol{e}_x(k,j)^T \boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + \boldsymbol{e}_u(k,j)^T \boldsymbol{u}_k(\boldsymbol{w}_{[k]}) - \boldsymbol{f}_j(k) \le 0,$$
$$\forall\, \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]},$$

where $\boldsymbol{e}_x(k,j)^T, \boldsymbol{e}_u(k,j)^T$ denote the j-th row of $E_x(k)$ and $E_u(k)$, respectively. Note that the left-hand side of this expression is also a multi-affine function of the variables $\boldsymbol{w}_{[k]}$. Since, by our observation, the maximum of multi-affine functions is reached at the vertices of the feasible set, i.e., $\boldsymbol{w}_{[k]} \in \text{ext}(\mathcal{W}_{[k]})$, and, by (B.2), we have that for any such vertex, $u_k(\boldsymbol{w}_{[k]}) = \boldsymbol{u}_k^{\text{ext}}(\boldsymbol{w}_{[k]}), \boldsymbol{x}_k(\boldsymbol{w}_{[k]}) = \boldsymbol{x}_k^{\text{ext}}(\boldsymbol{w}_{[k]})$, we immediately conclude that the constraint above is satisfied, since $\boldsymbol{u}_k^{\text{ext}}(\boldsymbol{w}_{[k]}), \boldsymbol{x}_k^{\text{ext}}(\boldsymbol{w}_{[k]})$ are certainly feasible.

A similar argument can be invoked for constraint (3.7d), and also to show that the maximum of the objective function is reached on the set of vertices $\text{ext}(\mathcal{W}_{[T]})$, and, since the values of the multi-affine policies exactly correspond to the optimal decisions in program $(P)_{\text{ext}}$, optimality is preserved. □

[1] In fact, multi-affine policies of degree k would be sufficient at time k

Appendix C

Appendix for Chapter 4

		ALY	CESO	CEST	SAA	d=0	d=1
$\sigma_t = 1.0$	**avg**	-34.15	-2.60	-2.28	-2.60	-2.01	-1.78
	std	13.68	4.14	3.92	4.14	3.27	3.23
	mdn	-33.83	-0.39	-0.33	-0.39	-0.45	-0.38
	min	-69.73	-25.09	-24.62	-25.09	-23.93	-24.28
	max	-5.24	-0.00	-0.00	-0.00	-0.00	-0.00
$\sigma_t = 2.0$	**avg**	-37.34	-6.21	-5.47	-6.21	-5.22	-4.24
	std	20.65	8.15	7.55	8.15	7.11	6.03
	mdn	-34.56	-2.09	-1.69	-2.08	-2.00	-1.50
	min	-98.29	-44.95	-43.26	-44.95	-58.32	-34.95
	max	-3.20	-0.01	-0.01	-0.01	-0.01	-0.01
$\sigma_t = 3.0$	**avg**	-42.52	-9.50	-8.28	-19.33	-9.01	-6.73
	std	25.33	11.41	10.20	12.47	11.85	8.68
	mdn	-38.12	-4.35	-3.68	-14.60	-4.70	-3.31
	min	-121.11	-77.22	-54.28	-89.10	-107.06	-65.24
	max	-2.83	-0.02	-0.02	-4.75	-0.01	-0.01

Table C.1: Relative gaps (in %) from perfect hindsight. Here, the noise terms u_t are Gaussian, and the standard deviation σ_t varies. Testing distribution is the *true* one.

	ALY	CESO	CEST	SAA	PH	d=0
avg	-34.94	-2.18	-1.35	-2.18	4.92	-1.04
std	20.37	3.87	3.78	3.87	7.80	3.90
mdn	-32.53	-0.25	0.03	-0.25	1.52	-0.01
min	-97.72	-21.94	-12.78	-21.94	0.01	-40.90
max	7.24	18.26	18.26	18.25	53.72	18.26

Table C.2: Relative gaps (in %) from polynomial policies with $d = 1$. Here, the noise terms u_t are Gaussian with $\sigma_t = 2.0$. Testing distribution is the *true* one.

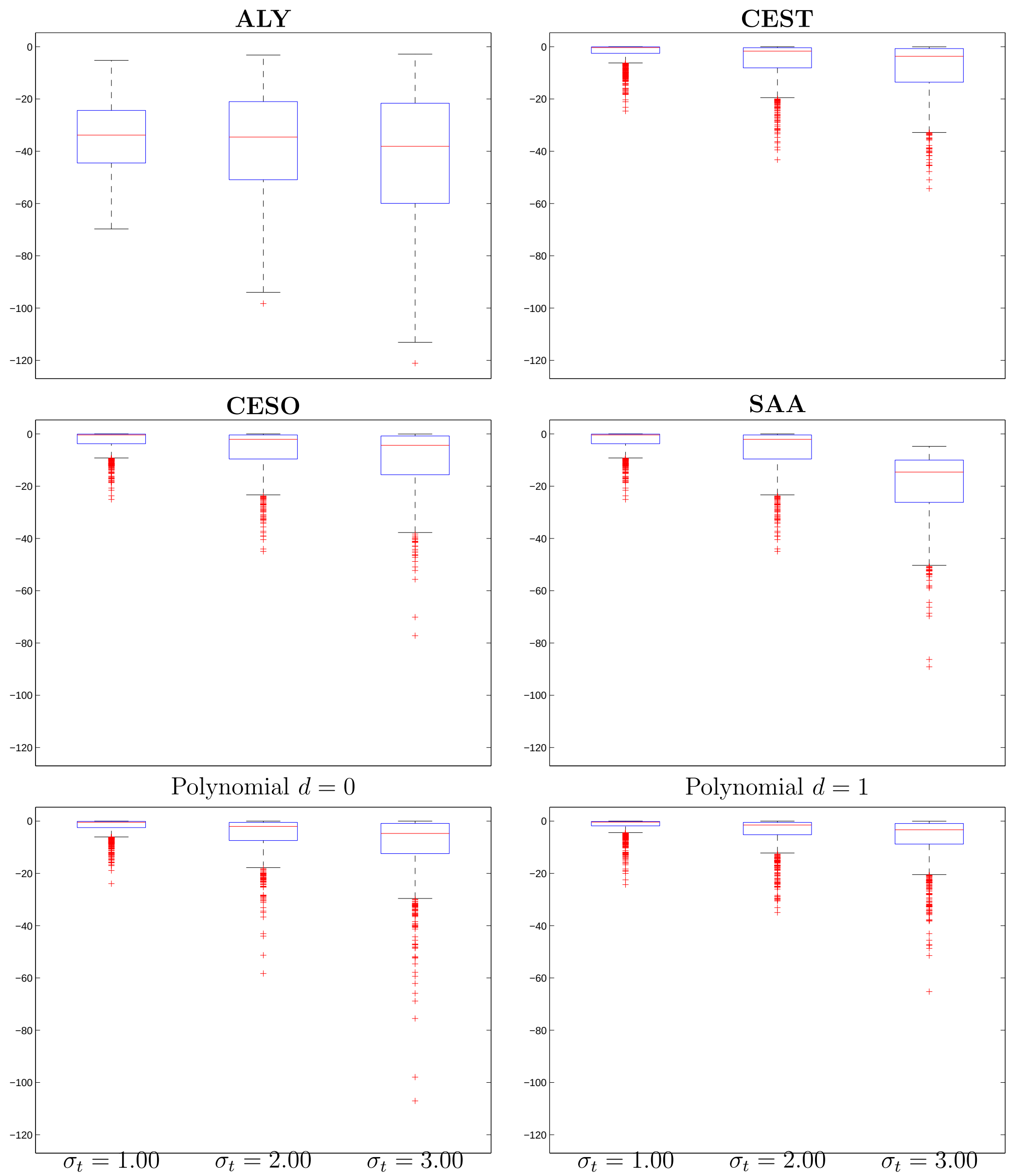


Figure C-1: Boxplots for the relative gaps (in %) from the perfect hindsight solution. Here, the noise terms u_t are Gaussian, and the standard deviation σ_t varies. Testing distribution is the *true* one.

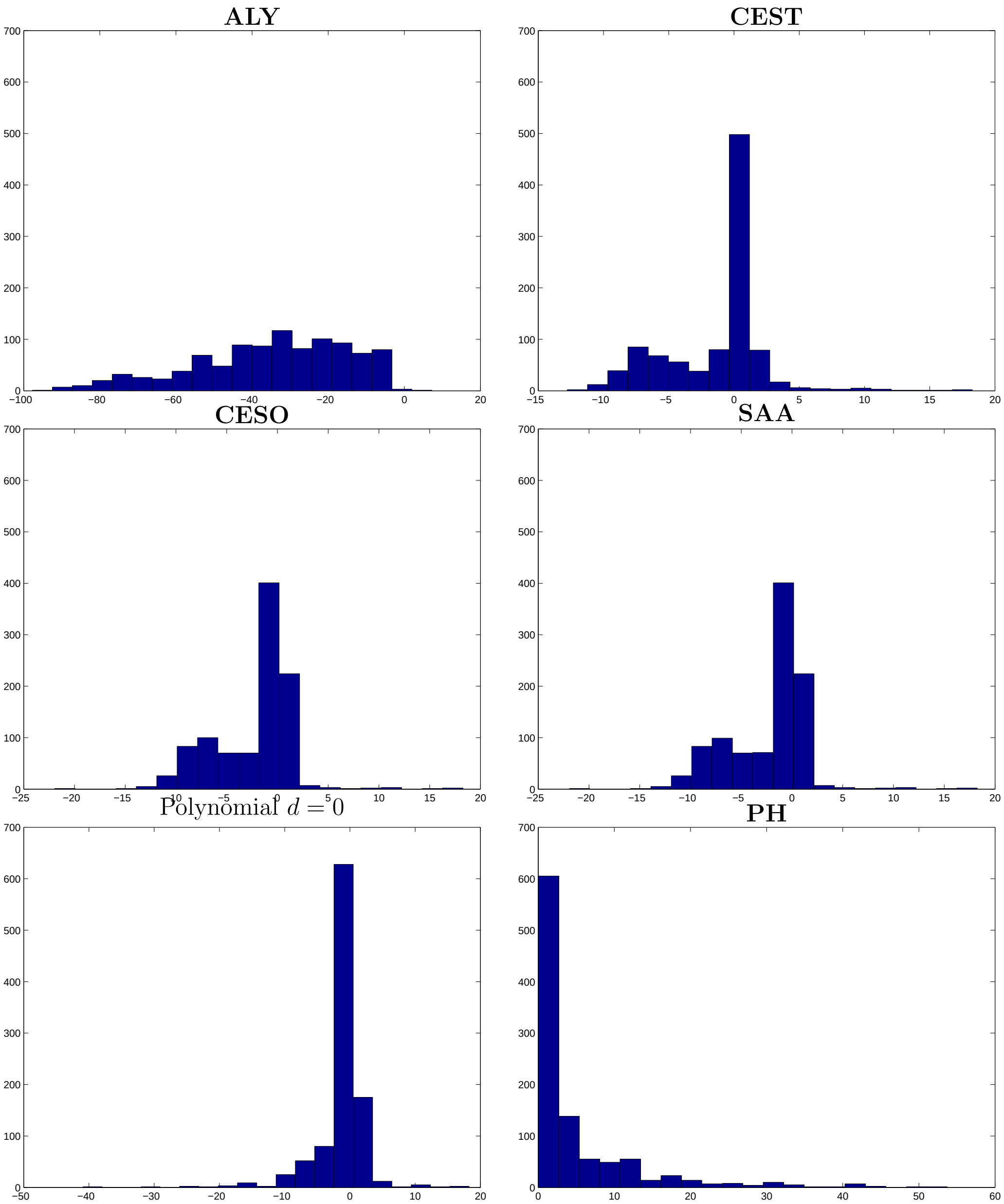


Figure C-2: Histograms of relative gaps (in %) from polynomial policies with $d = 1$. Here, the noise terms u_t are Gaussian, with $\sigma_t = 2.0$. Testing distribution is the *true* one.

		ALY	CESO	CEST	SAA	d=0	d=1
$\rho = 1.0$	**avg**	-33.89	-2.80	-2.52	-2.80	-2.37	-2.25
	std	15.06	4.50	4.38	4.50	3.87	4.04
	mdn	-32.44	-0.42	-0.32	-0.42	-0.43	-0.40
	min	-74.17	-29.59	-29.35	-29.59	-28.85	-31.83
	max	-5.58	-0.00	-0.00	-0.00	-0.00	-0.00
$\rho = 0.6$	**avg**	-34.66	-2.44	-2.16	-2.44	-1.77	-1.59
	std	12.92	3.95	3.72	3.95	2.94	2.76
	mdn	-34.98	-0.33	-0.29	-0.33	-0.36	-0.39
	min	-67.46	-24.57	-24.04	-24.57	-19.02	-19.02
	max	-4.32	-0.00	-0.00	-0.00	-0.00	-0.00
$\rho = 0.2$	**avg**	-35.32	-1.59	-1.49	-1.59	-1.21	-0.92
	std	9.94	2.67	2.48	2.67	1.98	1.65
	mdn	-36.13	-0.18	-0.19	-0.18	-0.22	-0.26
	min	-59.44	-19.06	-18.43	-19.05	-12.78	-14.69
	max	-4.33	-0.00	-0.00	-0.00	-0.00	-0.01

Table C.3: Relative gaps (in %) from perfect hindsight. Here, the noise terms u_t are Gaussian, and the correlation ρ varies, so as to make the disturbances in different time-periods less correlated. Testing distribution is the *true* one.

	ALY	CESO	CEST	SAA	PH	d=0
avg	-33.69	-0.89	-0.59	-0.89	1.70	-0.17
std	12.58	2.18	1.97	2.18	3.11	1.81
mdn	-33.54	0.00	0.03	0.00	0.39	0.01
min	-61.12	-8.76	-6.21	-8.76	0.00	-13.23
max	-4.14	7.31	7.31	7.31	23.48	7.31

Table C.4: Relative gaps (in %) from polynomial policies with $d = 1$. Here, the noise terms u_t are Gaussian with $\rho = 0.6$. Testing distribution is the *true* one.

		ALY	CESO	CEST	SAA	d=0	d=1
$a = -1.0$	**avg**	-33.04	-3.28	-2.91	-3.28	-2.68	-1.97
	std	10.37	4.35	4.17	4.35	3.69	3.04
	mdn	-33.29	-1.19	-0.81	-1.19	-1.21	-0.73
	min	-59.88	-24.47	-24.47	-24.47	-24.07	-19.48
	max	-8.68	-0.00	-0.00	-0.00	-0.00	-0.00
$a = -0.8$	**avg**	-31.30	-3.23	-2.82	-3.23	-3.34	-1.33
	std	12.07	4.42	4.23	4.42	4.75	1.94
	mdn	-32.25	-1.07	-0.75	-1.07	-1.24	-0.60
	min	-66.46	-23.31	-23.32	-23.31	-36.55	-15.07
	max	-8.21	-0.00	-0.00	-0.00	-0.00	-0.01
$a = -0.6$	**avg**	-31.84	-2.48	-2.17	-21.82	-5.35	-1.39
	std	15.42	3.85	3.62	5.70	7.84	1.74
	mdn	-32.60	-0.30	-0.29	-20.19	-1.34	-0.57
	min	-73.79	-30.02	-30.02	-50.81	-38.88	-18.17
	max	-6.71	-0.00	-0.00	-11.77	0.00	0.00

Table C.5: Relative gaps (in %) from perfect hindsight. Here, the noise terms u_t are Gaussian, and the value a in the price sensitivity matrix varies. Testing distribution is the *true* one.

	ALY	CESO	CEST	SAA	PH	d=0
avg	-30.40	-1.95	-1.52	-1.95	1.39	-2.06
std	12.03	3.47	3.63	3.47	2.13	3.83
mdn	-31.20	0.02	0.09	0.02	0.60	-0.46
min	-62.30	-13.82	-13.82	-13.83	0.01	-27.79
max	-8.16	4.60	4.60	4.60	17.74	4.60

Table C.6: Relative gaps (in %) from polynomial policies with $d = 1$. Here, the price sensitivity coefficient is $a = -0.8$. Testing distribution is the *true* one.

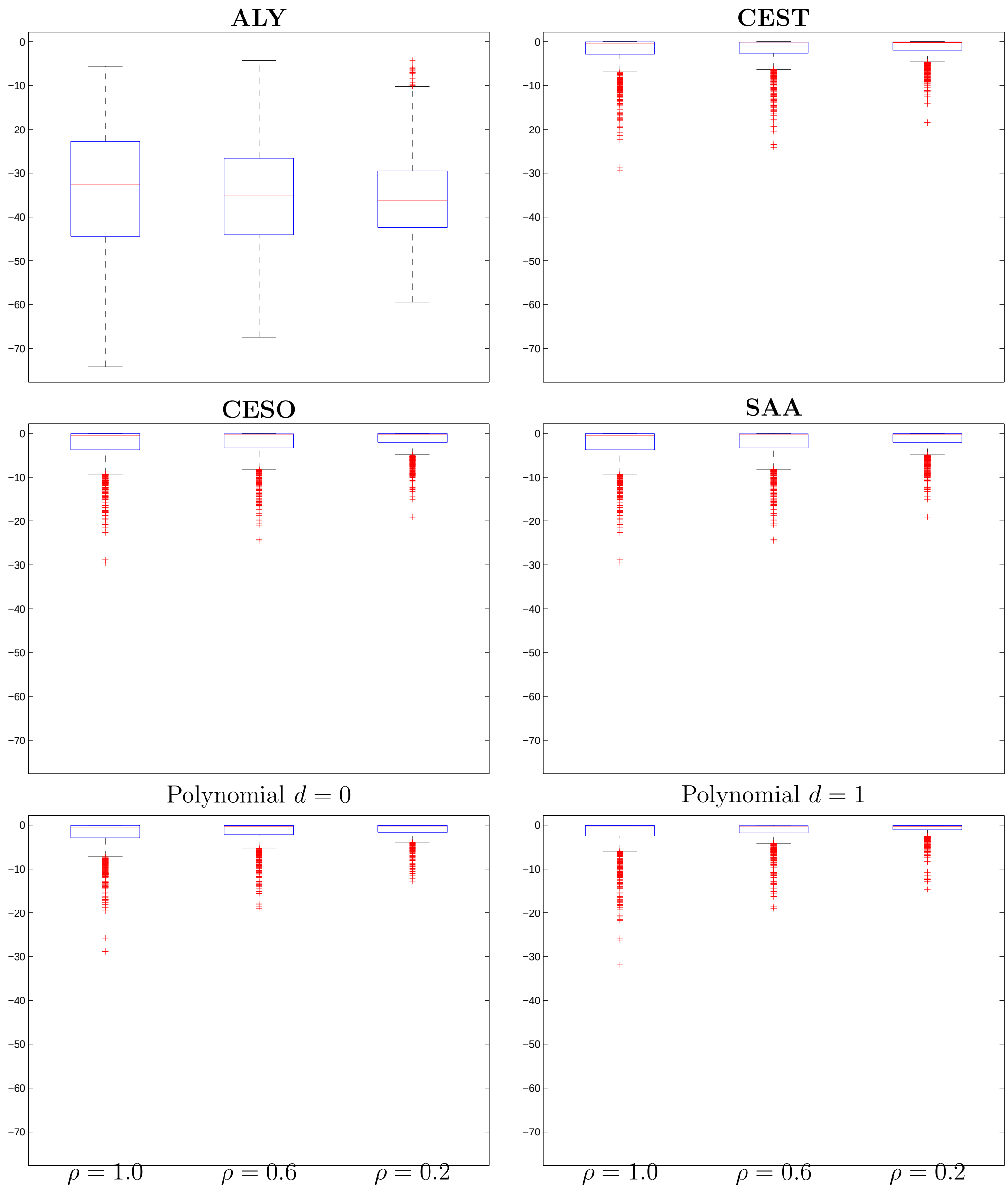


Figure C-3: Boxplots for the relative gaps (in %) from the perfect hindsight solution. Here, the noise terms u_t are Gaussian, and the correlation ρ varies, so as to make the disturbances in different time-periods less correlated. Testing distribution is the *true* one.

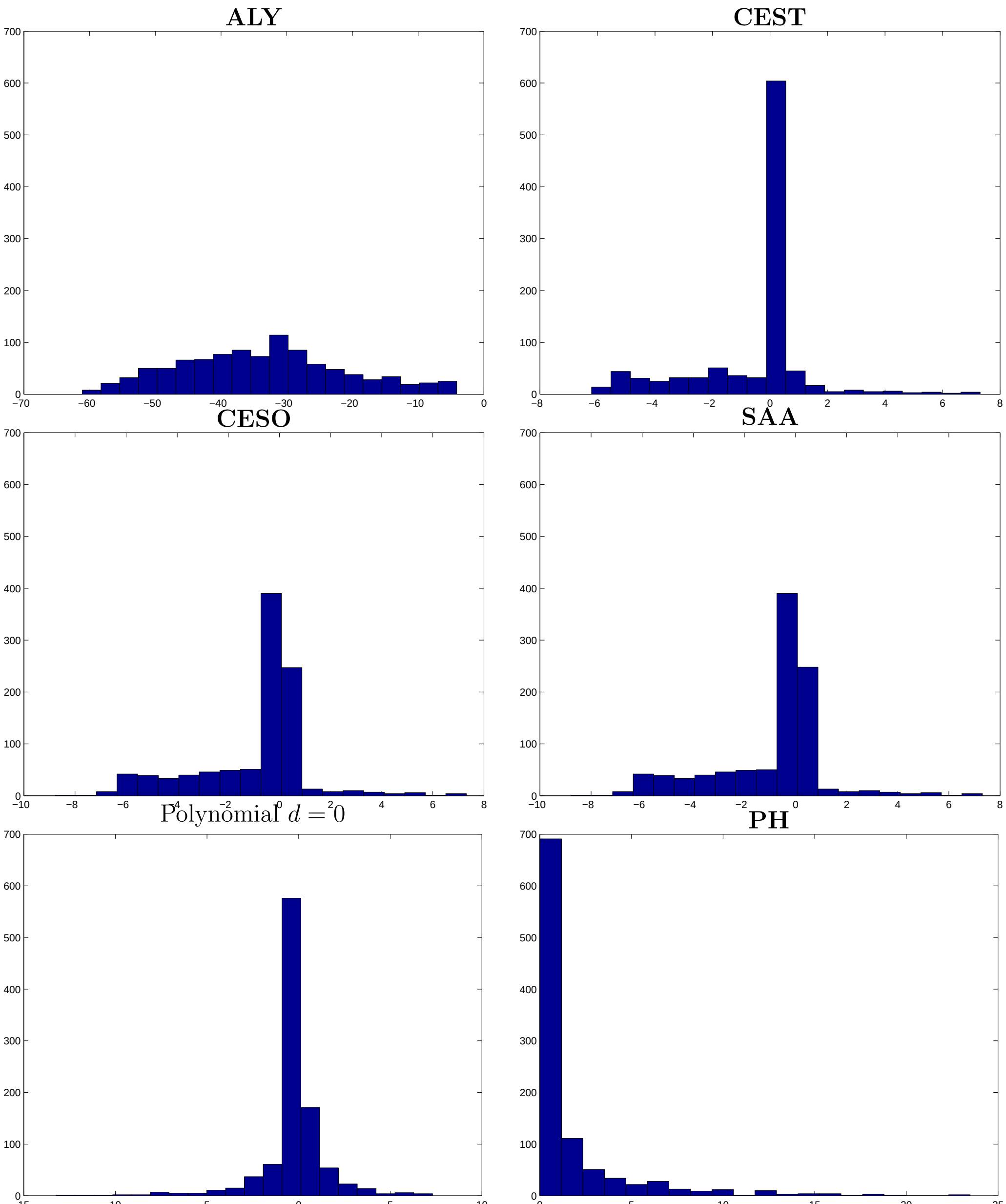


Figure C-4: Histograms of relative gaps (in %) from polynomial policies with $d = 1$. Here, the noise terms u_t are Gaussian, with $\rho = 0.6$. Testing distribution is the *true* one.

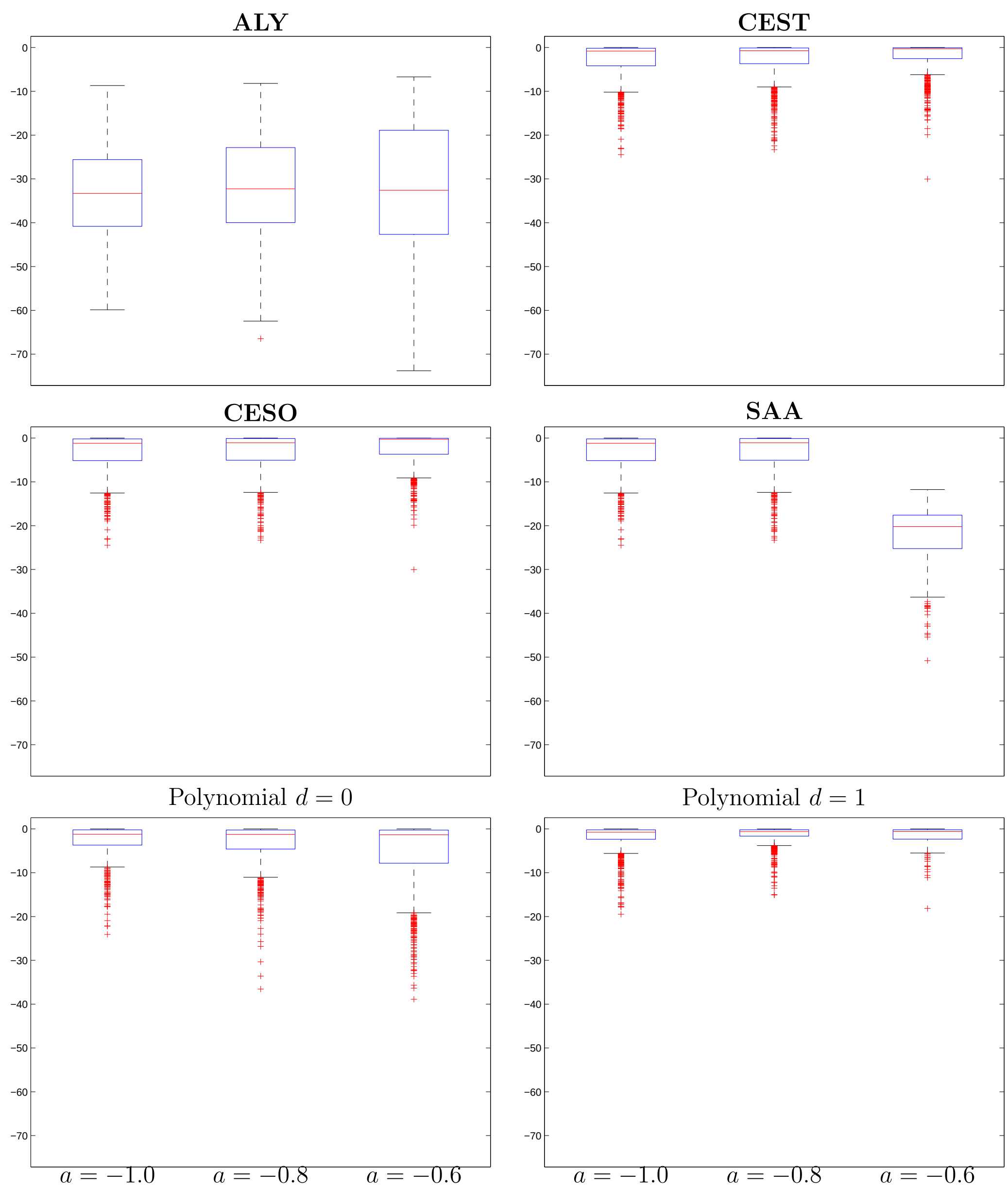


Figure C-5: Boxplots of relative gaps (in %) from perfect hindsight. Here, the noise terms u_t are Gaussian, and the value a in the price sensitivity matrix varies. Testing distribution is the *true* one.

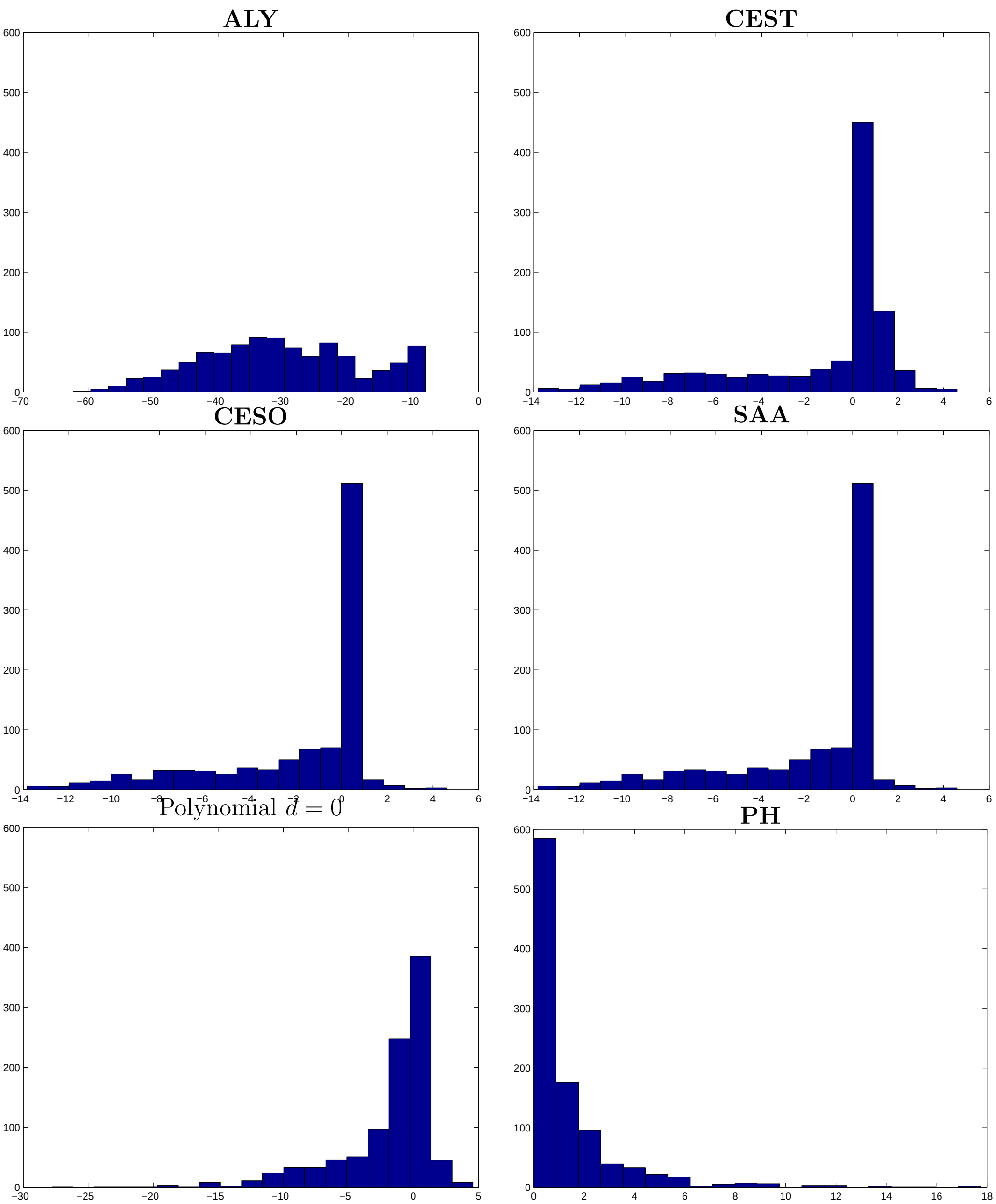


Figure C-6: Histogram of relative gaps (in %) from polynomial policies with $d = 1$. Here, the price sensitivity coefficient is $a = -0.8$. Testing distribution is the *true* one.

		ALY	CESO	CEST	SAA	d=0	d=1
$\bar{p} = 0.2$	**avg**	-39.90	-4.41	-4.21	-4.41	-3.28	-3.84
	std	13.12	5.92	5.78	5.92	4.70	5.43
	mdn	-39.91	-0.70	-0.51	-0.70	-0.50	-0.81
	min	-73.15	-28.73	-28.26	-28.73	-26.41	-37.43
	max	-5.38	-0.00	-0.00	-0.00	-0.00	-0.00
$\bar{p} = 0.5$	**avg**	-34.38	-2.68	-2.39	-2.68	-2.02	-1.87
	std	14.05	4.31	4.10	4.31	3.36	3.28
	mdn	-34.48	-0.39	-0.33	-0.39	-0.45	-0.43
	min	-73.61	-26.64	-26.23	-26.64	-21.71	-21.70
	max	-5.40	-0.00	-0.00	-0.00	-0.00	-0.00
$\bar{p} = 0.8$	**avg**	-29.80	-2.18	-1.55	-2.18	-2.41	-1.43
	std	13.11	3.40	2.79	3.40	3.49	2.42
	mdn	-29.31	-0.76	-0.56	-0.76	-0.97	-0.66
	min	-66.90	-28.00	-27.63	-28.00	-23.52	-29.52
	max	-5.20	-0.00	-0.00	-0.01	-0.00	-0.00

Table C.7: Relative gaps (in %) from perfect hindsight. Here, the coefficient $\bar{p}$ varies, changing the mean of the perturbations. Testing distribution is the *true* one.

	ALY	CESO	CEST	SAA	PH	d=0
avg	-33.25	-0.84	-0.54	-0.84	2.04	-0.14
std	13.62	2.16	1.96	2.16	3.79	1.76
mdn	-32.54	-0.00	0.02	-0.00	0.44	0.01
min	-66.75	-9.75	-5.78	-9.75	0.00	-13.94
max	-5.29	8.54	9.11	8.53	27.72	8.54

Table C.8: Relative gaps (in %) from polynomial policies with $d = 1$. Here, the coefficient $\bar{p} = 0.5$, corresponding to 0-mean perturbations u_t. Testing distribution is the *true* one.

Testing distribution		ALY	CESO	CEST	SAA	PH	d=0
True	**avg**	-73.12	-2.28	-12.47	-1.74	5.48	-24.92
	std	3.81	3.49	2.89	3.45	3.34	2.96
	mdn	-72.79	-2.65	-12.47	-2.07	4.64	-25.62
	min	-85.84	-10.81	-19.17	-10.38	0.93	-29.63
	max	-61.43	13.67	-1.67	13.74	25.01	-8.94
Gaussian	**avg**	-72.94	-1.98	-12.14	-1.46	5.46	-24.67
	std	3.66	3.82	3.05	3.64	3.58	3.17
	mdn	-72.41	-1.80	-12.14	-1.40	4.53	-25.40
	min	-87.55	-11.86	-22.14	-11.37	1.20	-30.13
	max	-62.66	15.51	3.62	15.52	29.19	-4.92
Gauss mix	**avg**	-73.26	-1.67	-11.98	-1.15	5.85	-24.45
	std	3.57	3.75	3.25	3.68	5.60	3.45
	mdn	-72.70	-1.64	-11.98	-1.09	4.72	-25.29
	min	-88.61	-11.46	-19.91	-11.25	0.83	-29.93
	max	-63.43	16.63	18.24	19.56	100.64	5.20

Table C.9: Relative gaps (in %) from polynomial policies with $d = 1$. Here, the noise terms are uniform and strongly negatively correlated, and the testing distributions are uniform, Gaussian or mixture of Gaussians.

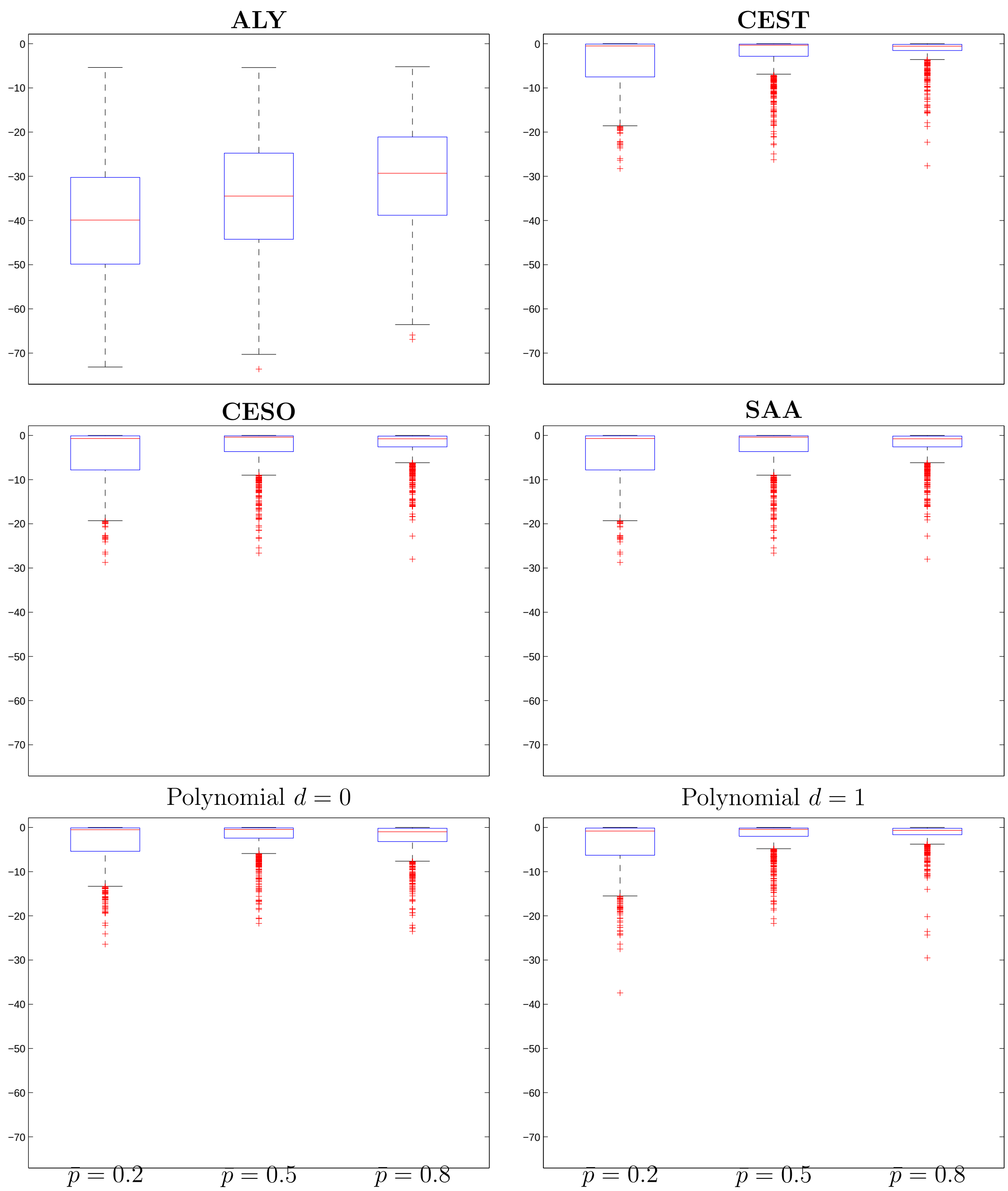


Figure C-7: Boxplots of relative gaps (in %) from perfect hindsight. Here, the coefficient $\bar{p}$ varies, changing the mean of the perturbations. Testing distribution is the *true* one.

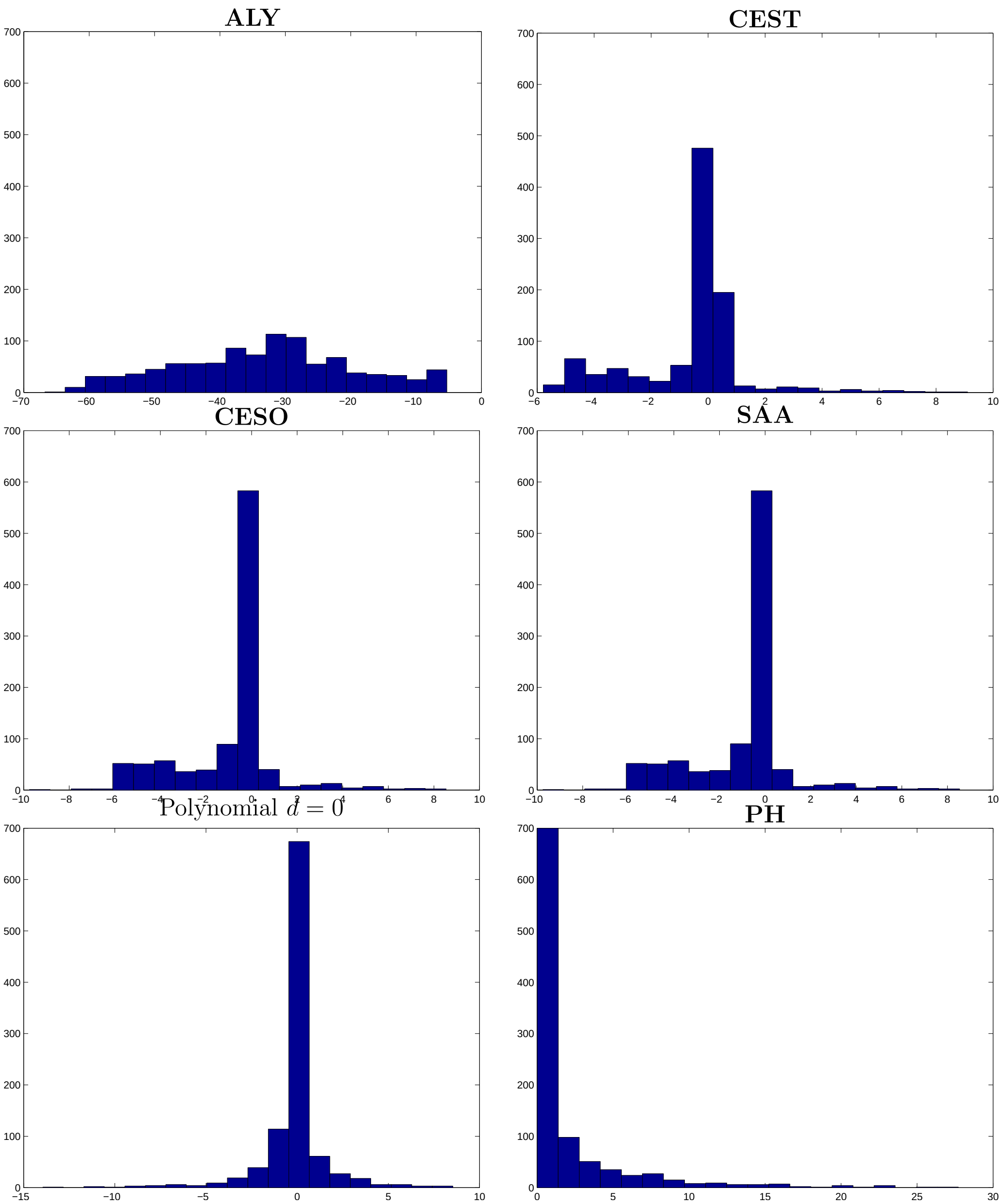


Figure C-8: Histograms of relative gaps (in %) from polynomial policies with $d = 1$. Here, the coefficient $\bar{p} = 0.5$, corresponding to 0-mean perturbations u_t. Testing distribution is the *true* one.

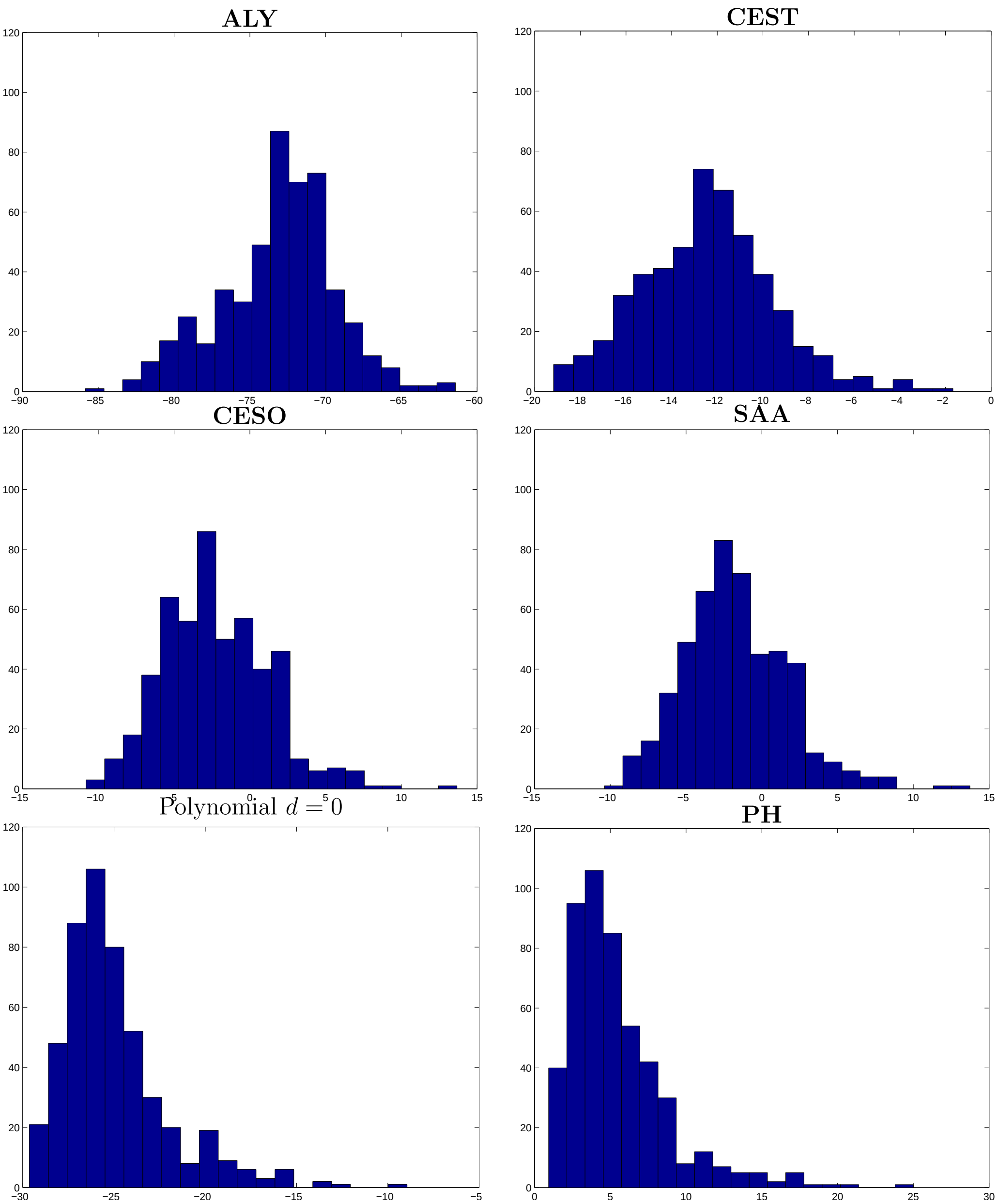


Figure C-9: Histograms of relative gaps (in %) from polynomial policies with $d = 1$. Here, the noise terms are uniform and strongly, negatively correlated, and the testing distribution is the *true* one.

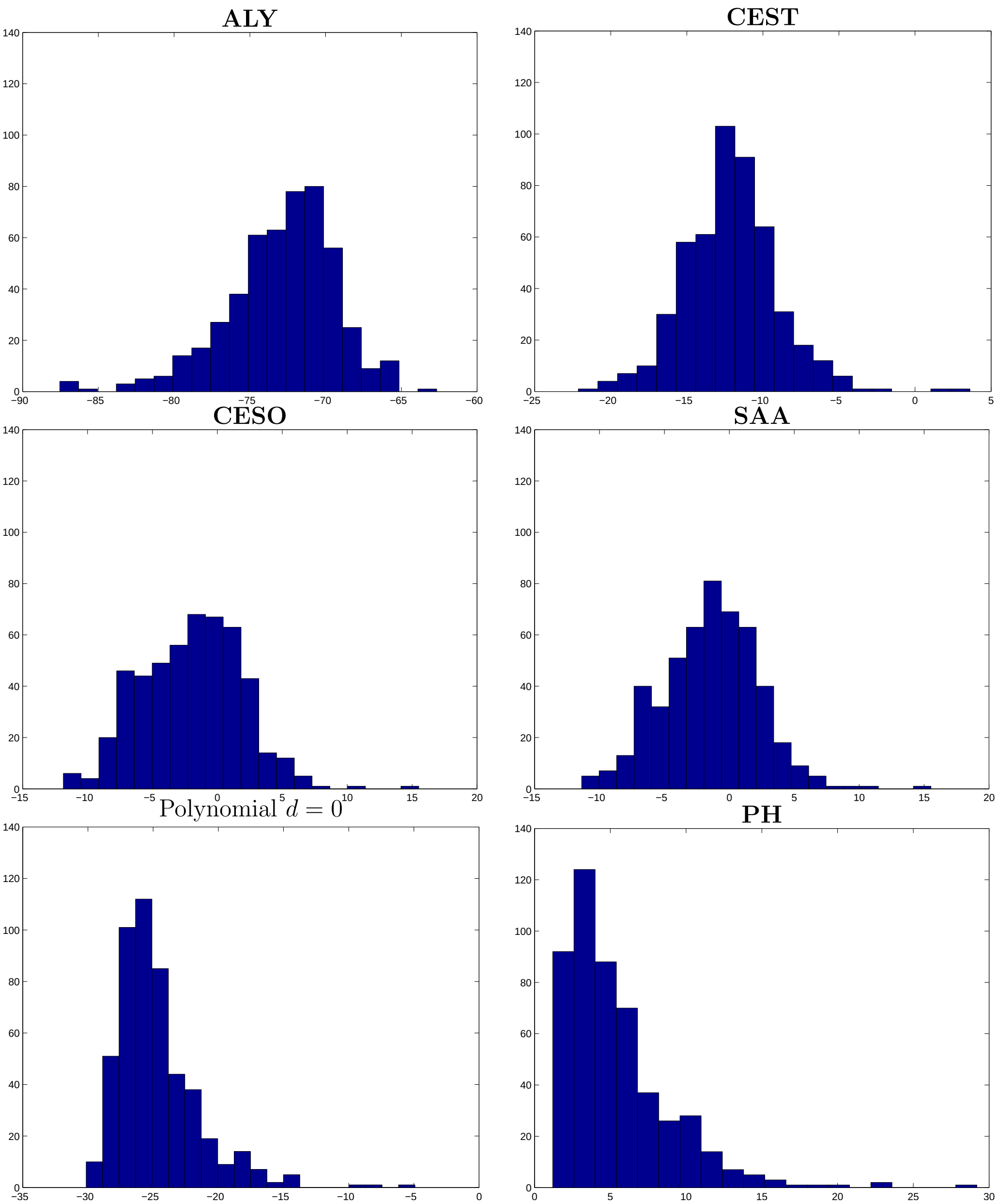


Figure C-10: Histograms of relative gaps (in %) from polynomial policies with $d = 1$. Here, the noise terms are uniform and strongly, negatively correlated, and the testing distribution is *gaussian*.

Testing distribution		ALY	CESO	CEST	SAA	PH	d=1
Uniform	**avg**	-22.29	0.43	0.43	0.41	3.60	0.43
	std	3.15	0.44	0.44	0.48	0.55	0.46
	cv	-0.14	1.02	1.02	1.16	0.15	1.06
	mdn	-21.67	0.48	0.48	0.41	3.59	0.47
	min	-30.88	-0.74	-0.74	-0.78	2.21	-0.70
	max	-15.35	1.60	1.60	1.45	5.19	1.58
Gaussian	**avg**	-26.37	2.41	2.41	2.57	6.17	2.39
	std	3.24	3.12	3.12	3.25	4.10	3.13
	cv	-0.12	1.30	1.30	1.27	0.66	1.31
	mdn	-26.91	1.53	1.53	1.54	5.01	1.51
	min	-32.50	-0.81	-0.81	-0.67	1.55	-0.83
	max	-11.00	22.24	22.24	22.92	30.88	22.25
Gauss mix	**avg**	-26.00	2.94	2.94	3.09	6.77	2.92
	std	3.47	3.62	3.62	3.77	4.72	3.63
	cv	-0.13	1.23	1.23	1.22	0.70	1.24
	mdn	-26.52	1.98	1.98	2.07	5.80	1.95
	min	-31.56	-0.64	-0.64	-0.75	1.65	-0.70
	max	-10.65	23.38	23.38	24.00	32.23	23.36

Table C.10: Test using real data (department 2, subclass 1). Table records relative gaps (in %) from polynomial policies with $d = 0$. Here, the noise terms are uniform and strongly negatively correlated, and the testing distributions are uniform, Gaussian or mixture of Gaussians.

Bibliography

[1] Elodie Adida and Georgia Perakis. Dynamic pricing and inventory control: Robust vs. stochastic uncertainty models. unpublished., January 2007. URL `http://tigger.uic.edu/~elodie/adjustable.pdf`.

[2] Victor Araman and Rene Caldentey. Dynamic pricing for non-perishable products with demand learning. Submitted to Operations Research, 2005.

[3] Soren Asmussen and Peter Glynn. *Stochastic Simulation: Algorithms and Analysis*. Springer, 2007. ISBN 978-0387306797.

[4] Y. Aviv and A. Pazgal. Pricing of short life-cycle products through active learning. Working paper, Washington University in St. Louis, St. Louis, MO, 2002.

[5] Yossi Aviv and Amit Pazgal. A partially observed markov decision process for dynamic pricing. *Manage. Sci.*, 51(9):1400–1416, 2005. ISSN 0025-1909. doi: http://dx.doi.org/10.1287/mnsc.1050.0393.

[6] F. Babonneau, J.-P. Vial, and R. Apparigliato. *Uncertainty and Environmental Decision Making*, volume 138 of *International Series in Operations Research & Management Science*, chapter Robust Optimization for Environmental and Energy Planning, pages 79–126. Springer US, 2010.

[7] A. Bemporad, F. Borrelli, and M. Morari. Optimal Controllers for Hybrid Systems: Stability and Piecewise Linear Explicit Form. In *Decision and Control, 2000. Proceedings of the 39th IEEE Conference on*, volume 2, pages 1810–1815 vol.2, 2000. doi: 10.1109/CDC.2000.912125.

[8] A. Bemporad, F. Borrelli, and M. Morari. Model Predictive Control Based on Linear Programming - The Explicit Solution. *IEEE Transactions on Automatic Control*, 47(12):1974–1985, 2002.

[9] A. Bemporad, F. Borrelli, and M. Morari. Min-Max Control of Constrained Uncertain Discrete-Time Linear Systems. *IEEE Transactions on Automatic Control*, 48(9):1600–1606, 2003. ISSN 0018-9286.

[10] A. Ben-Tal and A. Nemirovski. Robust Convex Optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998. doi: 10.1287/moor.23.4.769. URL `http://mor.journal.informs.org/cgi/content/abstract/23/4/769`.

[11] A. Ben-Tal and A. Nemirovski. Robust Solutions of Uncertain Linear Programs. *Operations Research Letters*, 25:1–13, 1999.

[12] A. Ben-Tal and A. Nemirovski. Robust optimization - methodology and applications. *Math. Program.*, 92(3):453–480, 2002. ISSN 0025-5610. doi: http://dx.doi.org/10.1007/s101070100286.

[13] A. Ben-Tal, A. Nemirovski, and C. Roos. Robust Solutions of Uncertain Quadratic and Conic-Quadratic Problems. *SIAM Journal on Optimization*, 13(2):535–560, 2002. ISSN 1052-6234. doi: http://dx.doi.org/10.1137/S1052623401392354.

[14] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Math. Program.*, 99(2):351–376, 2004. ISSN 0025-5610. doi: http://dx.doi.org/10.1007/s10107-003-0454-y.

[15] A. Ben-Tal, S. Boyd, and A. Nemirovski. Control of Uncertainty-Affected Discrete Time Linear Systems via Convex Programming. Working paper, 2005.

[16] A. Ben-Tal, B. Golany, A. Nemirovski, and J.-P. Vial. Retailer-Supplier Flexible Commitments Contracts: A Robust Optimization Approach. *Manufacturing & Service Operations Management*, 7(3):248–271, 2005. ISSN 1526-5498. doi: http://dx.doi.org/10.1287/msom.1050.0081.

[17] A. Ben-Tal, S. Boyd, and A. Nemirovski. Extending Scope of Robust Optimization: Comprehensive Robust Counterparts of Uncertain Problems. *Mathematical Programming*, 107(1):63–89, 2006. ISSN 0025-5610. doi: http://dx.doi.org/10.1007/s10107-005-0679-z.

[18] Aharon Ben-Tal and Arkadi Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming, Series A*, 88(3):411–424, September 2000.

[19] Aharon Ben-Tal, Laurent El-Ghaoui, and Arkadi Nemirovski. *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press, 2009.

[20] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 1999. ISBN 1886529000.

[21] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, MA, 2001. ISBN 1886529272.

[22] D. P. Bertsekas and I.B. Rhodes. On the minmax reachability of target tubes. *Automatica*, 7:233–247, 1971. doi: http://dx.doi.org/10.1007/s10107-005-0679-z.

[23] D. P. Bertsekas and J. N. Tsitsiklis. *Neurodynamic Programming*. Athena Scientific, Belmont, MA, 1996.

[24] D. P. Bertsekas, A. Nedic, and A. Ozdaglar. *Convex Analysis and Optimization*. Athena Scientific, 2003. ISBN 1886529450.

[25] Dimitri Bertsekas. *Control of Uncertain Systems with a Set-Membership Description of the Uncertainty*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1971.

[26] Dimitris Bertsimas and David B. Brown. Constrained Stochastic LQC: A Tractable Approach. *IEEE Transactions on Automatic Control*, 52(10):1826–1841, Oct. 2007.

[27] Dimitris Bertsimas and Dessislava Pachamanova. Robust multiperiod portfolio management in the presence of transaction costs. *Computers and Operations Research*, 35(1):3–17, 2008.

[28] Dimitris Bertsimas and Georgia Perakis. Dynamic pricing: A learning approach. *Mathematical and Computational Models for Congestion Charging*, 101:45–79, 2006. URL `http://dx.doi.org/10.1007/0-387-29645-X_3`.

[29] Dimitris Bertsimas and Ioana Popescu. Optimal Inequalities in Probability Theory: A Convex Optimization Approach. *SIAM Journal on Optimization*, 15(3):780–804, 2005. doi: 10.1137/S1052623401399903. URL `http://link.aip.org/link/?SJE/15/780/1`.

[30] Dimitris Bertsimas and Melvyn Sim. Robust Discrete Optimization and Network Flows. *Mathematical Programming*, 98(1-3):49–71, 2003.

[31] Dimitris Bertsimas and Melvyn Sim. The Price of Robustness. *Operations Research*, 52(1):35–53, 2004. URL `http://or.journal.informs.org/cgi/content/abstract/52/1/35`.

[32] Dimitris Bertsimas and Aurelie Thiele. A Robust Optimization Approach to Inventory Theory. *Operations Research*, 54(1):150–168, 2006. URL `http://or.journal.informs.org/cgi/content/abstract/54/1/150`.

[33] Dimitris Bertsimas and John N. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, Belmont, MA, 1997.

[34] Dimitris Bertsimas, Dessislava Pachamanova, and Melvyn Sim. Robust linear optimization under general norms. *Operations Research Letters*, 32(6):510 – 516, 2004.

[35] Dimitris Bertsimas, Constantine Caramanis, and David B. Brown. Theory and Applications of Robust Optimization. Submitted for publication., 2008.

[36] Dimitris Bertsimas, Xuan Vinh Doan, Karthik Natarajan, and Chung-Piaw Teo. Models for minimax stochastic linear optimization problems with risk aversion. *Mathematics of Operations Research*, 2010. doi: 10.1287/moor.1100.0445.

[37] Dimitris Bertsimas, Dan A. Iancu, and Pablo A. Parrilo. Optimality of Affine Policies in Multistage Robust Optimization. *MATHEMATICS OF OPERATIONS RESEARCH*, 35(2):363–394, 2010. doi: 10.1287/moor.1100.0444. URL http://mor.journal.informs.org/cgi/content/abstract/35/2/363.

[38] Omar Besbes and Assaf Zeevi. On the minimax complexity of pricing in a changing environment. Submitted for publication, 2008. URL http://opimweb.wharton.upenn.edu/documents/research/Minimax_cpl_ch_pt_ob_az_posted.pdf.

[39] Omar Besbes and Assaf Zeevi. Dynamic Pricing Without Knowing the Demand Function: Risk Bounds and Near-Optimal Algorithms. *OPERATIONS RESEARCH*, 57(6):1407–1420, 2009. doi: 10.1287/opre.1080.0640. URL http://or.journal.informs.org/cgi/content/abstract/57/6/1407.

[40] Daniel Bienstock and Nuri Özbay. Computing Robust Basestock Levels. *Discrete Optimization*, 5(2):389 – 414, 2008. ISSN 1572-5286. doi: DOI:10.1016/j.disopt.2006.12.002. URL http://www.sciencedirect.com/science/article/B7GWV-4R1FSG6-4/2/46b3823dc847bd0c7904b1cb6f9db078. In Memory of George B. Dantzig.

[41] John R. Birge and Francois Louveaux. *Introduction to Stochastic Programming*. Springer, 2000. ISBN 0387982175.

[42] Gabriel R. Bitran and Susana V. Mondschein. Pricing perishable products: An application to the retail industry. MIT Working Paper #3592-93, July 2003. URL http://dspace.mit.edu/handle/1721.1/2482.

[43] George Box, Gwilym M. Jenkins, and Gregory Reinsel. *Time Series Analysis: Forecasting and Control*. Prentice Hall, 3 edition, 1994.

[44] Dean Weldon Boyd, Prashandt Balepur Narayan, Henry Frederick Schwarz, Philip David Reginalt Apps, Ravishankar Venkata Nandiwada, Brian Lawrence Monteiro, and Thomas Edward Guardino. Promotion Pricing System and Method. United States Patent US 7,072,848 B2, Manugistics, Inc., Rockville, MA USA, July 2006.

[45] Giuseppe Calafiore and Marco C. Campi. Uncertain Convex Programs: Randomized Solutions and Confidence Levels. *Mathematical Programming*, 102(1): 25–46, January 2005.

[46] Giuseppe Calafiore and Marco C. Campi. The Scenario Approach to Robust Control Design. *IEEE Transactions on Automatic Control*, 51(5):742–753, May 2006.

[47] Sebastián Ceria and Robert A. Stubbs. Incorporating estimation errors into portfolio selection: Robust portfolio construction. *Journal of Asset Management*, 7(2):109–127, July 2006.

[48] A. Charnes, W. W. Cooper, and G. H. Symonds. Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil. *Management Science*, 4(3):235–263, April 1958.

[49] Debasish Chatterjee, Peter Hokayem, and John Lygeros. Stochastic model predictive control with bounded control inputs: a vector space approach. Submitted for publication., March 2009. URL `http://arxiv.org/abs/0903.5444`.

[50] Xin Chen and Yuhan Zhang. Uncertain Linear Programs: Extended Affinely Adjustable Robust Counterparts. *Operations Research*, page opre.1080.0605, 2009. doi: 10.1287/opre.1080.0605. URL `http://or.journal.informs.org/cgi/content/abstract/opre.1080.0605v1`.

[51] Xin Chen, Melvyn Sim, and Peng Sun. A robust optimization perspective on stochastic programming. *Operations Research*, 55(6):1058–1071, 2007. doi: 10.1287/opre.1070.0441.

[52] Xin Chen, Melvyn Sim, Peng Sun, and Jiawei Zhang. A Linear Decision-Based Approximation Approach to Stochastic Programming. *Operations Research*, 56 (2):344–357, 2008. doi: 10.1287/opre.1070.0457.

[53] William L. Cooper, Tito Homem-de Mello, and Anton J. Kleywegt. Learning and pricing with models that do not explicitly incorporate competition. working paper, November 2009.

[54] Eric Cope. Bayesian strategies for dynamic pricing in e-commerce. *Naval Research Logistics*, 54(3):265–281, 2007. doi: 10.1002/nav.20204.

[55] Erick Delage and Yinyu Ye. Distributionally Robust Optimization Under Moment Uncertainty with Application to Data-Driven Problems. *OPERATIONS RESEARCH*, page opre.1090.0741, 2010. doi: 10.1287/opre.1090.0741. URL `http://or.journal.informs.org/cgi/content/abstract/opre.1090.0741v1`.

[56] John Doyle, Keith Glover, Pramod Khargonekar, and Bruce Francis. State-space solutions to standard h-2 and h-infinity control problems. *IEEE Transactions on Automatic Control*, 34(8):831–847, 1989.

[57] G. E. Dullerud and F. Paganini. *A Course in Robust Control Theory*. Springer, 2005. ISBN 0387989455.

[58] J. Dupačová. *Minimax stochastic programs with nonseparable penalties*, volume 1, chapter Lecture Notes in Control and Information Sciences, pages 157–163. Springer Verlag, Berlin, 1980.

[59] J. Dupačová. The minimax approach to stochastic programming and an illustrative application. *Stochastics*, 20:73–88, 1987.

[60] Martin Dyer and Leen Stougie. Computational complexity of stochastic programming problems. *Mathematical Programming*, 106(3):423–432, 2006. ISSN 0025-5610. doi: http://dx.doi.org/10.1007/s10107-005-0597-0.

[61] L. El-Ghaoui, F. Oustry, and H. Lebret. Robust Solutions to Uncertain Semidefinite Programs. *SIAM J. Optimization*, 9(1):33–52, 1998.

[62] Laurent El-Ghaoui and Hervé Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18(4):1035–1064, October 1997.

[63] Wedad Elmaghraby and Pinar Keskinocak. Dynamic Pricing in the Presence of Inventory Considerations: Research Overview, Current Practices, and Future Directions. *MANAGEMENT SCIENCE*, 49(10):1287–1309, 2003. doi: 10.1287/mnsc.49.10.1287.17315. URL `http://mansci.journal.informs.org/cgi/content/abstract/49/10/1287`.

[64] James E. Falk. Exact solutions of inexact linear programs. *Operations Research*, 24(4):783–787, 1976.

[65] M.K.H. Fan, A. Tits, and J. Doyle. Robustness in the Presence of Mixed Parametric Uncertainty and Unmodeled Dynamics. *IEEE Transactions on Automatic Control*, 36(1):25–38, 1991. ISSN 0018-9286.

[66] Amr Farahat and Georgia Perakis. Price competition among multiproduct firms. working paper, June 2008.

[67] Vivek Farias and Benjamin Van Roy. Dynamic pricing with a prior on market response. Operations Research - forthcoming, 2009.

[68] Guillermo Gallego and Garrett van Ryzin. Optimal Dynamic Pricing of Inventories with Stochastic Demand over Finite Horizons. *Management Science*, 40(8): 999–1020, 1994. doi: 10.1287/mnsc.40.8.999. URL `http://mansci.journal.informs.org/cgi/content/abstract/40/8/999`.

[69] C. E. Garcia, D. M. Prett, and M. Morari. Model Predictive Control: Theory and Practice—A Survey. *Automatica*, 25(3):335–348, 1989. ISSN 0005-1098. doi: http://dx.doi.org/10.1016/0005-1098(89)90002-2.

[70] Stanley J. Garstka and Roger J.-B. Wets. On Decision Rules in Stochastic Programming. *Mathematical Programming*, 7(1):117–143, 1974. ISSN 0025-5610. doi: 10.1007/BF01585511.

[71] John Geweke. Exact inference in the inequality constrained normal linear regression model. *Journal of Applied Econometrics*, 1(2):127–141, 1986. ISSN 08837252. URL `http://www.jstor.org/stable/2096611`.

[72] Eric Ghysels, Denise R. Osborn, and Paulo M.M. Rodrigues. Handbook of economic forecasting. In C.W.J. Granger G. Elliott and A. Timmermann, editors, *Forecasting Seasonal Time Series*, volume 1 of *Handbook of Economic Forecasting*, chapter 13, pages 659 – 711. Elsevier, 2006. doi: DOI:10.1016/S1574-0706(05)01013-X. URL http://www.sciencedirect.com/science/article/B7P5J-4JSMTWJ-J/2/8ebd1c9d33e50f751ca151daaeaeb45a.

[73] Paul Glasserman and Sridhar Tayur. Sensitivity Analysis for Base-Stock Levels in Multiechelon Production-Inventory Systems. *Management Science*, 41(2): 263–281, 1995. doi: 10.1287/mnsc.41.2.263. URL http://mansci.journal.informs.org/cgi/content/abstract/41/2/263.

[74] D. Goldfarb and G. Iyengar. Robust Portfolio Selection Problems. *Mathematics of Operations Research*, 28(1):1–38, 2003. doi: 10.1287/moor.28.1.1.14260. URL http://mor.journal.informs.org/cgi/content/abstract/28/1/1.

[75] Donald Goldfarb and Garud Iyengar. Robust Convex Quadratically Constrained Programs. *Mathematical Programming*, 97:495–515, 2002.

[76] P. J. Goulart and E.C. Kerrigan. Relationships Between Affine Feedback Policies for Robust Control with Constraints. In *Proceedings of the 16th IFAC World Congress on Automatic Control*, July 2005.

[77] William Greene. *Econometric Analysis*. Prentice Hall, 5th edition edition, 2002.

[78] Aliza Heching, Guillermo Gallego, and Garrett van Ryzin. Mark-down pricing: An empirical analysis of policies and revenue potential at one apparel retailer. *Journal of Revenue & Pricing Management*, 1(2):139–160, July 2002.

[79] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 19 edition, 2005.

[80] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 8 edition, 2007.

[81] Peter Huber and Elvezio Ronchetti. *Robust Statistics*. Wiley, 2 edition, February 2009.

[82] IBM ILOG. IBM ILOG CPLEX, 2010.

[83] Garud Iyengar and Emre Erdoğan. Ambiguous Chance Constrained Problems and Robust Optimization. *Mathematical Programming*, 107(1-2):37–61, June 2006.

[84] T. Jacobi and A. Prestel. Distinguished Representations of Strictly Positive Polynomials. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2001(532):223–235, 2001. doi: 10.1515/crll.2001.023. URL http://www.reference-global.com/doi/abs/10.1515/crll.2001.023.

[85] Soulaymane Kachani, Georgia Perakis, and Carine Simon. *Modeling the Transient Nature of Dynamic Pricing with Demand Learning in a Competitive Environment*, volume 102 of *International Series in Operations Research & Management Science*, chapter 11. Springer US, June 2007.

[86] Hiroshi Kasugai and Tadami Kasegai. Characteristics of Dynamic Maximin Ordering Policy. *Journal of the Operations Research Society of Japan*, 3(1): 11–26, 1960.

[87] E.C. Kerrigan and J.M. Maciejowski. On robust optimization and the optimal control of constrained linear systems with bounded state disturbances. In *Proceedings of the 2003 European Control Conference*, Cambridge, UK, September 2003.

[88] E.C. Kerrigan and J.M. Maciejowski. Properties of a New Parameterization for the Control of Constrained Systems with Disturbances. *Proceedings of the 2004 American Control Conference*, 5:4669–4674 vol.5, June-July 2004. ISSN 0743-1619. URL `http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01384049`.

[89] Robert Kleinberg and Tom Leighton. The Value of Knowing a Demand Curve: Bounds on Regret for Online Posted-Price Auctions. In *FOCS '03: Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, page 594. IEEE Computer Society, 2003.

[90] Daniel Kuhn, Wolfram Wiesemann, and Angelos Georghiou. Primal and Dual Linear Decision Rules in Stochastic and Robust Optimization. *Mathematical Programming*, pages 1–33, December 2009. doi: 10.1007/s10107-009-0331-4.

[91] Mahesh Kumar and Nitin R. Patel. Using clustering to improve sales forecasts in retail merchandising. *Annals of Operations Research*, 174(1):33–46, February 2010.

[92] Guangui Lan, Zhaosong Lu, and Renato D. C. Monteiro. Primal-dual First-order Methods with O($1/\epsilon$) Iteration-complexity for Cone Programming. *Mathematical Programming*, 2009. doi: 10.1007/s10107-008-0261-6.

[93] Yingjie Lan, Huina Gao, Michael O. Ball, and Itir Karaesmen. Revenue Management with Limited Demand Information. *MANAGEMENT SCIENCE*, 54 (9):1594–1609, 2008. doi: 10.1287/mnsc.1080.0859. URL `http://mansci.journal.informs.org/cgi/content/abstract/54/9/1594`.

[94] J. B. Lasserre. Global Optimization with Polynomials and the Problem of Moments. *SIAM Journal on Optimization*, 11:796–817, 2001.

[95] Andrew E. B. Lim and J. George Shanthikumar. Relative Entropy, Exponential Utility, and Robust Dynamic Pricing. *OPERATIONS RESEARCH*, 55(2):198–214, 2007. doi: 10.1287/opre.1070.0385. URL `http://or.journal.informs.org/cgi/content/abstract/55/2/198`.

[96] Kyle Y. Lin. Dynamic pricing with real-time demand learning. *European Journal of Operational Research*, 174(1):522–538, 2006. ISSN 0377-2217. doi: DOI: 10.1016/j.ejor.2005.01.041. URL `http://www.sciencedirect.com/science/article/B6VCT-4FTXXMN-3/2/de8a104b502dd441b48c442a6e22d0f9`.

[97] Ruben Lobel and Georgia Perakis. Dynamic pricing through sampling based optimization. working paper, 2010.

[98] M. Lobo and Stephen Boyd. Pricing and Learning with Uncertain Demand. working paper, 2003.

[99] J. Löfberg. Approximations of closed-loop minimax MPC. *Proceedings of the 42nd IEEE Conference on Decision and Control*, 2:1438–1442 Vol.2, Dec. 2003. ISSN 0191-2216. doi: 10.1109/CDC.2003.1272813.

[100] J. Löfberg. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. URL `http://control.ee.ethz.ch/~joloef/yalmip.php`.

[101] Constantinos Maglaras and Joern Meissner. Dynamic pricing strategies for multiproduct revenue management problems. *Manufacturing & Service Operations Management*, 8(2):136–148, 2006. ISSN 1526-5498. doi: http://dx.doi.org/10.1287/msom.1060.0105.

[102] Murari Mani, Ashish K. Sing, and Michael Orshansky. Joint Design-time and Post-silicon Minimization of Parametric Yield Loss Using Adjustable Robust Optimization. In *ICCAD '06: Proceedings of the 2006 IEEE/ACM international conference on Computer-aided design*, pages 19–26, New York, NY, USA, 2006. ACM. ISBN 1-59593-389-1. doi: http://doi.acm.org/10.1145/1233501.1233507.

[103] Peter Marbach and John N. Tsitsiklis. Simulation-based optimization of Markov reward processes. *IEEE Transactions on Automatic Control*, 46(2):191–209, January 2001. doi: 10.1109/9.905687.

[104] Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, June 1995.

[105] Alp Muharremoglu and John N. Tsitsiklis. A Single-Unit Decomposition Approach to Multiechelon Inventory Systems. *Operations Research*, 56(5):1089–1103, 2008. doi: 10.1287/opre.1080.0620. URL `http://or.journal.informs.org/cgi/content/abstract/56/5/1089`.

[106] Karthik Natarajan, Melvyn Sim, and Joline Uichanco. Tractable robust expected utility and risk models for portfolio optimization. to appear in Mathematical Finance, December 2008.

[107] Arkadi Nemirovski and Alexander Shapiro. *Continuous Optimization: Current Trends and Applications*, volume 99 of *Applied Optimization*, chapter On Complexity of Stochastic Programming Problems, pages 111–146. Springer US, 2005. doi: 10.1007/b137941.

[108] Arkadi Nemirovski and Alexander Shapiro. Convex approximations of chance constrained programs. *SIAM J. on Optimization*, 17(4):969–996, 2006. ISSN 1052-6234. doi: http://dx.doi.org/10.1137/050622328.

[109] Yurii Nesterov and Arkadi Nemirovski. *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics, 1994.

[110] Jiawang Nie. Regularization Methods for Sum of Squares Relaxations in Large Scale Polynomial Optimization. Submitted for publication., September 2009. URL `http://arxiv.org/abs/0909.3551`.

[111] Jiawang Nie. Polynomial matrix inequality and semidefinite representation. working paper., May 2010. URL `http://arxiv.org/abs/0908.0364v1`.

[112] Pablo A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, 2000.

[113] Pablo A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming Series B*, 96(2):293–320, May 2003. doi: http://dx.doi.org/10.1007/s10107-003-0387-5.

[114] Georgia Perakis and Guillaume Roels. Robust Controls for Network Revenue Management. *MANUFACTURING SERVICE OPERATIONS MANAGEMENT*, 12(1):56–76, 2010. doi: 10.1287/msom.1080.0252. URL `http://msom.journal.informs.org/cgi/content/abstract/12/1/56`.

[115] Robert Phillips. *Pricing and Revenue Management*. Stanford Business Books, 1 edition, August 2005.

[116] Mustafa Pinar and Reha Tütüncü. Robust profit opportunities in risky financial portfolios. *Operations Research Letters*, 33(4):331 – 340, 2005. ISSN 0167-6377. doi: DOI:10.1016/j.orl.2004.08.005.

[117] Ioana Popescu. A Semidefinite Programming Approach to Optimal-Moment Bounds for Convex Classes of Distributions. *MATHEMATICS OF OPERATIONS RESEARCH*, 30(3):632–657, 2005. doi: 10.1287/moor.1040.0137. URL `http://mor.journal.informs.org/cgi/content/abstract/30/3/632`.

[118] Ioana Popescu. Robust Mean-Covariance Solutions for Stochastic Optimization. *OPERATIONS RESEARCH*, 55(1):98–112, 2007. doi: 10.1287/opre.1060.0353. URL `http://or.journal.informs.org/cgi/content/abstract/55/1/98`.

[119] Evan L. Porteus. *Foundations of Stochastic Inventory Management*. Stanford Business Books, 2002.

[120] Warren Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Wiley Series in Probability and Statistics. Wiley-Interscience, 2007.

[121] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42:969–984, 2003.

[122] A.J. Quist, E. De Clerk, C. Roos, and T. Terlaky. Copositive relaxation for general quadratic programming. *Optimization Methods and Software*, 9(1):185–208, 1998.

[123] Arvind Rajan Rakesh and Richard Steinberg. Dynamic Pricing and Ordering Decisions by a Monopolist. *Management Science*, 38(2):240–262, 1992. doi: 10.1287/mnsc.38.2.240. URL `http://mansci.journal.informs.org/cgi/content/abstract/38/2/240`.

[124] Rama Ramakrishnan. Markdown Management. personal communication., 2009.

[125] Rama Ramakrishnan. Personal communication., 2009-2010.

[126] T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970. ISBN 0691080690.

[127] Walter Rudin. *Principles of Mathematical Analysis*. Mc-Graw Hill Science/Engineering/Math, 1976. ISBN 007054235X.

[128] Paat Rusmevichientong, Benjamin Van Roy, and Peter W. Glynn. A Nonparametric Approach to Multiproduct Pricing. *OPERATIONS RESEARCH*, 54(1):82–98, 2006. doi: 10.1287/opre.1050.0252. URL `http://or.journal.informs.org/cgi/content/abstract/54/1/82`.

[129] H. Scarf. *A Min-Max Solution to an Inventory Problem*, chapter 2, pages 201–209. Studies in the Mathematical Theory of Inventory and Production. Stanford University Press, Stanford, CA, USA, 1958.

[130] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on Stochastic Programming*. MPS / SIAM Series on Optimization. SIAM, 2009.

[131] Melvyn Sim and Joel Goh. Distributionally Robust Optimization and its Tractable Approximations. Accepted in Operations Research, April 2009.

[132] David Simchi-Levi, Xin Chen, and Julien Bramel. *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management*. Springer Series in Operations Research and Financial Engineering. Springer, 2 edition, 2004.

[133] J. Skaf and S. Boyd. Design of Affine Controllers Via Convex Optimization. Submitted to *IEEE Transactions on Automatic Control*, 2008. URL http://www.stanford.edu/~boyd/papers/affine_contr.html.

[134] J. Skaf and S. Boyd. Nonlinear Q-Design for Convex Stochastic Control. Submitted to *IEEE Transactions on Automatic Control*, 2008. URL http://www.stanford.edu/~boyd/papers/nonlin_Q_param.html.

[135] Stephen A. Smith and Dale D. Achabal. Clearance pricing and inventory policies for retail chains. *Management Science*, 44(3):285–300, 1998. ISSN 00251909. URL http://www.jstor.org/stable/2634668.

[136] A. L. Soyster. Technical Note–Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming. *Operations Research*, 21(5):1154–1157, 1973. doi: 10.1287/opre.21.5.1154. URL http://or.journal.informs.org/cgi/content/abstract/21/5/1154.

[137] J. Sturm, P. Imre, and T. Terlaky. Sedumi 1.3, 2006. URL http://sedumi.ie.lehigh.edu/.

[138] Kalyan T. Talluri and Garrett J. van Ryzin. *The Theory and Practice of Revenue Management*. Springer Science+Business Media, Inc., 233 Spring Street, New York, NY 10013, USA, paperback edition, 2005.

[139] Aurélie Thiele. Single-product Pricing via Robust Optimization. Submitted to Operations Research, January 2006.

[140] Aurélie Thiele. Multi-product pricing via robust optimisation. *Journal of Revenue & Pricing Management*, 8(1):67 – 80, 2009. ISSN 14766930.

[141] K.C. Toh, M.J. Todd, and R. Tütüncü. SDPT3 a MATLAB Software Package for Semidefinite Programming, 1999. URL http://www.math.nus.edu.sg/~mattohkc/sdpt3.html.

[142] Reha Tütüncü and M. Koenig. Robust asset allocation. *Annals of Operations Research*, 132(1-4):157–187, November 2004.

[143] Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996. ISSN 0036-1445. doi: http://dx.doi.org/10.1137/1038003.

[144] J. Žáčková. On Minimax Solution of Stochastic Linear Programming Problems. *Časopis pro Pěstování Matematiky*, 91(4):423–430, 1966.

[145] Hans S. Witsenhausen. *Minimax Control of Uncertain Systems*. PhD thesis, Massachusetts Institute of Technology, Cambridge MA, USA, 1966.

[146] Hans S. Witsenhausen. Sets of possible states of linear systems given perturbed observations. *IEEE Transactions on Automatic Control*, AC(13), 1968.

[147] Jonathan Woo, Michael Levy, and John Bible. Inventory and Price Decision Support. United States Patent US 6,910,017 B1, ProfitLogic, Inc., Cambridge MA, USA, June 2005.

[148] K. Zhou and J. C. Doyle. *Essentials of Robust Control*. Prentice Hall, 1998. ISBN 0135258332.

[149] G. Ziegler. *Lectures on Polytopes*. Springer, 2nd edition, 2003. ISBN 0387404090.

[150] P. Zipkin. *Foundations of Inventory Management*. McGraw Hill, 2000. ISBN 0256113793.

[151] Luis F. Zuluaga and Javier F. Pena. A Conic Programming Approach to Generalized Tchebycheff Inequalities. *MATHEMATICS OF OPERATIONS RESEARCH*, 30(2):369–388, 2005. doi: 10.1287/moor.1040.0124. URL `http://mor.journal.informs.org/cgi/content/abstract/30/2/369`.