A Hierarchy of Near-Optimal Policies for Multistage Adaptive Optimization

Dimitris Bertsimas, Dan Andrei Iancu, *Student Member, IEEE*, and Pablo A. Parrilo, *Senior Member, IEEE*

***Abstract*—In this paper, we propose a new tractable framework for dealing with linear dynamical systems affected by uncertainty, applicable to multistage robust optimization and stochastic programming. We introduce a hierarchy of near-optimal polynomial disturbance-feedback control policies, and show how these can be computed by solving a single semidefinite programming problem. The approach yields a hierarchy parameterized by a single variable (the degree of the polynomial policies), which controls the trade-off between the optimality gap and the computational requirements. We evaluate our framework in the context of three classical applications—two in inventory management, and one in robust regulation of an active suspension system—in which very strong numerical performance is exhibited, at relatively modest computational expense.**

***Index Terms*—Constrained control, optimization, robust adaptive control, semidefinite programming, sums-of-squares, uncertain systems.**

I. Introduction

MULTISTAGE optimization problems under uncertainty are prevalent in numerous fields of engineering, economics, finance, and have elicited interest on both a theoretical and a practical level from diverse research communities. Among the most established methodologies for dealing with such problems are dynamic programming (DP) [1], stochastic programming [2], robust control [3], [4], and, more recently, robust optimization (see [5]–[8] and references therein).

In the current paper, we consider discrete-time, linear dynamical systems of the form

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

evolving over a finite planning horizon, $k = 0, \dots, T-1$. 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 known. The system is affected by unknown,[1] additive disturbances, $\boldsymbol{w}(k)$, which are assumed to lie in a given compact, basic semialgebraic set

$$\mathcal{W}_k \stackrel{\text{def}}{=} \{\boldsymbol{w}(k) \in \mathbb{R}^{n_w} : g_j(\boldsymbol{w}(k)) \geq 0,\ j \in \mathcal{J}_k\} \tag{2}$$

where $g_j \in \mathbb{R}[w]$ are multivariate polynomials depending on the vector of uncertainties at time k, $\boldsymbol{w}(k)$, and $\mathcal{J}_k$ is a finite index set. We note that this formulation captures many uncertainty sets of interest, such as polytopes (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 III-C for details).

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

$$\begin{aligned} E_x(k)\,\boldsymbol{x}(k) + E_u(k)\,\boldsymbol{u}(k) &\leq \boldsymbol{f}(k), \qquad k = 0, \dots, T-1 \\ E_x(T)\,\boldsymbol{x}(T) &\leq \boldsymbol{f}(T) \end{aligned} \tag{3}$$

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 piecewise affine and convex in the states and controls

$$\begin{aligned} &h\left(k, \boldsymbol{x}(k), \boldsymbol{u}(k)\right) \\ &\quad = \max_{i \in \mathcal{I}_k} \left[\, c_0(k,i) + \boldsymbol{c}_x(k,i)^T \boldsymbol{x}(k) + \boldsymbol{c}_u(k,i)^T \boldsymbol{u}(k) \,\right] \end{aligned} \tag{4}$$

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

$$\begin{aligned} J = &h\left(0, \boldsymbol{x}(0), \boldsymbol{u}(0)\right) \\ &+ \max_{\boldsymbol{w}(0)} \Big[\, h\left(1, \boldsymbol{x}(1), \boldsymbol{u}(1)\right) + \cdots \\ &\qquad + \max_{\boldsymbol{w}(T-2)} \big[\, h\left(T-1, \boldsymbol{x}(T-1), \boldsymbol{u}(T-1)\right) \\ &\qquad\qquad + \max_{\boldsymbol{w}(T-1)} h\left(T, \boldsymbol{x}(T)\right) \big] \dots \Big]. \end{aligned}$$

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

[1]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 $\boldsymbol{w}(k)$.

Manuscript received June 26, 2009; revised April 29, 2010; accepted January 26, 2011. Date of publication August 18, 2011; date of current version December 07, 2011. This work was supported in part by the National Science Foundation under Grants DMI-0556106 and EFRI-0735905, by the Air Force Office of Scientific Research under Grant FA9550-06-1-0303, and the the Department of Mathematical Sciences, IBM T. J. Watson Research Center. Recommended by Associate Editor C.-H. Chen.

D. Bertsimas is with the Faculty of the Sloan School of Management and the Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: dbertsim@mit.edu).

D. A. Iancu is with the Faculty of the Graduate School of Business, Stanford University, Stanford, CA 94305 USA (e-mail: iancu_dan@gsb.stanford.edu).

P. A. Parrilo is with the Faculty of Electrical Engineering and Computer Science and the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: parrilo@mit.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2011.2162878

to compute optimal policies, $\boldsymbol{u}^\star(k, \boldsymbol{x}(k))$, and optimal value functions, $J^\star(k, \boldsymbol{x}(k))$ (see [1] for details). 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 multiparametric programming [9]–[11]), or resort to approximations, such as approximate DP [12], [13], stochastic approximation [14], simulation-base optimization [15], [16], 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 suboptimality gaps resulting from specific classes of policies (the interested reader can refer to the books [1], [12], and [13] for a thorough review).

An alternative approach, originally proposed in the stochastic programming community (see [2], [17], and references therein), is to consider control policies that are parametrized directly in the sequence of observed uncertainties, and typically referred to as *recourse decision rules*. 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 piecewise quadratic, convex cost, the authors in [17] show that piecewise 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 [18] and [19]).

Disturbance-feedback parameterizations have recently been used by researchers in robust control and robust optimization (see [5]–[7], [20]–[26], 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 convex optimization problems, which vary from linear and quadratic (e.g., [6], [21]), to second-order conic and semidefinite programs (e.g., [6], [20], [24], and [25]). Some of the first steps toward analyzing the properties of disturbance-affine policies were taken in [6] and [21], where it was shown that, under suitable conditions, the resulting parametrization has certain desirable system theoretic properties (stability and robust invariance), and 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 open-loop and prestabilizing control policies.

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. The only result in a constrained, robust setting that the authors are aware of is our recent paper [8], in which it is shown that, in the case of one-dimensional systems, with independent state and control constraints $\left(L_k \le u_k \le U_k, L_k^x \le x_k \le U_k^x\right)$, linear control costs and *any* convex state costs, disturbance-affine policies *are*, in fact, optimal, and can be found efficiently. As a downside, the same paper presents simple examples of multidimensional systems where affine policies are suboptimal.

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 at the end of [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 [27]–[29] 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 [30], 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 [26], 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 [31] 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 multistage dynamical systems, involving complicated constraints on both states and controls.

The goal of our current paper is to introduce a new framework for modeling and (approximately) solving such multistage dynamical problems. 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 III-C for a discussion). Our main contributions are summarized below:

- 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 [32]. Our approach is advantageous

[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 [1] and references therein.

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

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 one in robust control (regulation of an active suspension system), 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 paper is organized as follows. Section II presents the mathematical formulation of the problem, briefly discusses relevant solution techniques in the literature, and introduces our framework. Section III, which is the main body of the paper, 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 III-C also elaborates on immediate extensions of the framework to more general multistage decision problems. Section V translates three classical problems into our framework, and Section VI presents our computational results, exhibiting the strong performance of polynomial policies. Section VII concludes the paper and suggests directions of future research.

Notation

Throughout the rest of the paper, we denote scalar quantities by lowercase, nonbold 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 jth 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 $\boldsymbol{w}_{[k]} \stackrel{\text{def}}{=} (\boldsymbol{w}_0, \ldots, \boldsymbol{w}_{k-1})$ 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, $\boldsymbol{w}_{[0]} \equiv []$, i.e., an empty vector. 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

$$\mathcal{B}_d(\boldsymbol{x}) \stackrel{\text{def}}{=} \left(1, x_1, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_1 x_n, \ldots, x_n^d\right) \quad (5)$$

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

$$p(\boldsymbol{x}) = p(x_1, \ldots, x_n) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} p_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} = \boldsymbol{p}^T \mathcal{B}_d(\boldsymbol{x}) \quad (6)$$

where $\boldsymbol{x}^{\boldsymbol{\alpha}} \stackrel{\text{def}}{=} x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$, 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 (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.

II. Problem Description

Using the notation mentioned in the introduction, our goal is to find nonanticipatory 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

$$\begin{aligned} & \min_{\boldsymbol{u}_0} \Big[h_0(\boldsymbol{x}_0, \boldsymbol{u}_0) + \max_{\boldsymbol{w}_0} \min_{\boldsymbol{u}_1} \Big[h_1(\boldsymbol{x}_1, \boldsymbol{u}_1) + \cdots + \\ (P) \quad & + \min_{\boldsymbol{u}_{T-1}} \big[h_{T-1}(\boldsymbol{x}_{T-1}, \boldsymbol{u}_{T-1}) + \max_{\boldsymbol{w}_{T-1}} h_T(\boldsymbol{x}_T) \big] \ldots \Big] \Big] \end{aligned} \quad (7a)$$

$$\text{s.t. } \boldsymbol{x}_{k+1} = A_k \boldsymbol{x}_k + B_k \boldsymbol{u}_k + \boldsymbol{w}_k, \; \forall k \in \{0, \ldots, T-1\} \quad (7b)$$

$$E_x(k) \boldsymbol{x}_k + E_u(k) \boldsymbol{u}_k \leq \boldsymbol{f}_k, \; \forall k \in \{0, \ldots, T-1\} \quad (7c)$$

$$E_x(T) \boldsymbol{x}_T \leq \boldsymbol{f}_T. \quad (7d)$$

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, denoted by $\mathcal{F}_k$, resulting in control policies $\boldsymbol{u}_k : \mathcal{F}_k \to \mathbb{R}^{n_u}$.

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) [1], to compute optimal policies, $\boldsymbol{u}^\star_k(\boldsymbol{x}_k)$, and optimal value functions, $J^\star_k(\boldsymbol{x}_k)$. 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 [12], [13], [16].

In this paper, we take a slightly different approach, and consider instead policies parametrized directly in the observed uncertainties

$$\boldsymbol{u}_k : \mathcal{W}_0 \times \mathcal{W}_1 \times \cdots \times \mathcal{W}_{k-1} \to \mathbb{R}^{n_u}. \quad (8)$$

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

In this context, the decisions that must be taken are the parameters defining the specific functional form sought for $\boldsymbol{u}_k$. One such example of disturbance-feedback policies, often considered in the literature, 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 (7b) used to express the dependency of states $\boldsymbol{x}_k$ on past uncertainties, the state-control constraints (7c), (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 [2], [5]–[7], [17], [20]–[25]. The typical restriction to the subclass of affine policies, done for purposes of tractability, almost invariably results in loss of performance [19], with the gap being sometimes very large.

To illustrate this effect, we introduce the following simple example,[5] motivated by a similar case in [27].

1) Example 1: Consider a two-stage problem, where $\boldsymbol{w} \in \mathcal{W}$ is the uncertainty, with $\mathcal{W} = \{\boldsymbol{w} \in \mathbb{R}^N : \|\boldsymbol{w}\|_2 \le 1\}$, $x \in \mathbb{R}$ 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:

$$\begin{aligned} \underset{x, \boldsymbol{y}(\boldsymbol{w})}{\text{minimize}} \quad & x \\ \text{such that} \quad & x \ge \sum_{i=1}^N y_i, \quad \forall \boldsymbol{w} \in \mathcal{W}, \\ & y_i \ge w_i^2, \quad \forall \boldsymbol{w} \in \mathcal{W}. \end{aligned} \tag{9}$$

It can be easily shown (see Lemma 1 in Section A) that the optimal objective in Problem (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 paper 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 $\mathcal{F}_k$

$$\boldsymbol{w}_{[k]} \overset{\text{def}}{=} (\boldsymbol{w}_0, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{k-1}) \in \mathbb{R}^{k\cdot n_w} \tag{10}$$

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

$$\boldsymbol{u}_k(\boldsymbol{w}_{[k]}) = L_k \, \mathcal{B}_d(\boldsymbol{w}_{[k]}) \tag{11}$$

where $\mathcal{B}_d(\boldsymbol{w}_{[k]})$ is the canonical basis of $\mathcal{P}_d[\boldsymbol{w}_{[k]}]$, given by (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.

[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.

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 [33], 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 nonconvex 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 [34]–[36].

III. Polynomial Policies and Convex Reformulations Using Sums-of-Squares

Under polynomial policies of the form (11), one can use the dynamical (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 jth row of $E_x(k)$ and $E_u(k)$, respectively, a typical state-control constraint (7c) can be written

$$\begin{aligned} &\boldsymbol{e}_x(k,j)^T \boldsymbol{x}_k + \boldsymbol{e}_u(k,j)^T \boldsymbol{u}_k \le \boldsymbol{f}_j(k) \\ &\qquad \Leftrightarrow p_{j,k}^{\text{con}}(\boldsymbol{w}_{[k]}) \ge 0, \; \forall \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]} \end{aligned}$$

where

$$p_{j,k}^{\text{con}}(\boldsymbol{w}_{[k]}) \overset{\text{def}}{=} \boldsymbol{f}_j(k) - \boldsymbol{e}_x(k,j)^T \boldsymbol{x}_k - \boldsymbol{e}_u(k,j)^T \boldsymbol{u}_k.$$

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 $p_{j,k}^{\text{con}}(\boldsymbol{w}_{[k]}), j = 1, \ldots, r_k$, are nonnegative on the domain $\mathcal{W}_{[k]}$.

Similarly, the expression (4) for the stage cost at time k can be written as

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

i.e., the cost h_k is a piecewise 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:

$$\begin{aligned} &\min_{L_0}\Big[\max_{i\in\mathcal{I}_1} p_i^{\text{cost}}(\boldsymbol{w}_{[0]}) + \max_{\boldsymbol{w}_0}\min_{L_1}\Big[\max_{i\in\mathcal{I}_2} p_i^{\text{cost}}(\boldsymbol{w}_{[1]}) + \cdots \\ &\qquad + \max_{\boldsymbol{w}_{T-2}}\min_{L_{T-1}}\big[\max_{i\in\mathcal{I}_{T-1}} p_i^{\text{cost}}(\boldsymbol{w}_{[T-1]}) \\ (P_{\text{POP}}) &\qquad + \max_{\boldsymbol{w}_{T-1}}\max_{i\in\mathcal{I}_T} p_i^{\text{cost}}(\boldsymbol{w}_{[T]})\big]\ldots\Big]\Big] \end{aligned} \tag{12a}$$

$$\text{s.t.} \quad p_{j,k}^{\text{con}}(\boldsymbol{w}_{[k]}) \ge 0, \; \forall k = 0, \ldots, T, \; \forall j = 1, \ldots, r_k, \; \forall \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]}. \tag{12b}$$

In this formulation, the decision variables are the coefficients $\{L_t\}_{0\le t\le T-1}$, and (12b) summarize all the state-control constraints. We emphasize that the expression of the polynomial controls(11) and the dynamical system (7b) should not be interpreted as real constraints in the problem (rather, they are only used to derive the dependency of the polynomials $p_i^{\mathrm{cost}}(\boldsymbol{w}_{[k]})$ and $p_{j,k}^{\mathrm{con}}(\boldsymbol{w}_{[k]})$ on $\{L_t\}_{0\leq t\leq k-1}$ and $\boldsymbol{w}_{[k]}$).

A. Reformulating the Constraints

As mentioned in the previous section, under polynomial control policies, a typical state-control constraint (12b) in program (P_{POP}) can now be written as

$$p(\boldsymbol{\xi}) \geq 0, \quad \forall \boldsymbol{\xi} \in \mathcal{W}_{[k]} \tag{13}$$

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, $0 \leq t \leq k-1$. It is easy to see that constraint (13) can be rewritten equivalently as

$$\begin{aligned} &p(\boldsymbol{\xi}) \geq 0, \ \forall \boldsymbol{\xi} \in \mathcal{W}_{[k]} \\ &\mathcal{W}_{[k]} \stackrel{\mathrm{def}}{=} \left\{\boldsymbol{\xi} \in \mathbb{R}^{k\cdot n_w} : g_j(\boldsymbol{\xi}) \geq 0, \ j = 1, \ldots, m\right\} \end{aligned} \tag{14}$$

where $\{g_j\}_{1\leq j\leq m}$ 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 (2). In this form, (14) falls in the general class of constraints that require testing polynomial nonnegativity 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 (14) to hold is

$$p = \sigma_0 + \sum_{j=1}^{m} \sigma_j \, g_j \tag{15}$$

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 nonnegativity 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 [34]–[36] for details), which, for a fixed precision, can be done in polynomial time, by interior point methods [32].

At first sight, condition (15) might seem overly restrictive. However, it is motivated by recent powerful results in real algebraic geometry [37], [38], which, under mild conditions[6] on the functions g_j, state that *any* polynomial that is strictly positive on a compact semi-algebraic set $\mathcal{W}_{[k]}$ *must* admit a representation of the form (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 (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.

[6]These are readily satisfied when g_j are affine, or can be satisfied by simply appending a redundant constraint that bounds the 2-norm of the vector $\boldsymbol{\xi}$

B. Reformulating the Objective

Recall from our discussion in the beginning of Section III 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

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

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

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

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 (7a), we seek to solve

$$\begin{aligned} \min \quad & J \\ \text{s.t.} \quad & J \geq \sum_{k=0}^{T} \tilde{h}_k(\boldsymbol{w}_{[k]}), \ \forall \boldsymbol{w}_{[T]} \in \boldsymbol{w}_{[T]} & \text{(16a)} \\ & \tilde{h}_k(\boldsymbol{w}_{[k]}) \geq p_i^{\mathrm{cost}}(\boldsymbol{w}_{[k]}), \ \forall \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]}, \ \forall i \in \mathcal{I}_k. & \text{(16b)} \end{aligned}$$

The advantage of this approach is that constraints (16a) and (16b) are now of the exact same nature as (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, in general, over-bound the optimal cost of Problem (P), due to several reasons.

1) We are replacing the (true) piecewise polynomial cost h_k with an *upper bound* given by the polynomial cost $\tilde{h}_k$. Therefore, the optimal value J of problem (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 (16a), (16b), and (12b), are enforced using SOS polynomials with fixed degree (see the discussion in Section III-A), and this is sufficient, but not necessary.

However, despite these multiple layers of approximation, our numerical experiments (Section VI), suggest that most of the above considerations are second-order effects when compared with the fact that polynomial policies of the form (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:

1) Consider polynomial control policies in the disturbances, $\boldsymbol{u}_k(\boldsymbol{w}_{[k]}) = L_k\,\mathcal{B}_d(\boldsymbol{w}_{[k]})$.
2) Express all the states $\boldsymbol{x}_k$ according to (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 $\sum_k \tilde{h}_k$.
5) Replace a typical constraint $p(\boldsymbol{w}_{[k]}) \ge 0$, $\forall \boldsymbol{w}_{[k]} \in \{\boldsymbol{\xi} : g_j(\boldsymbol{\xi}) \ge 0,\ j=1,\ldots,m\}$ (for either state-control or costs) with the requirements:

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

6) Solve the resulting SDP to obtain the coefficients L_k.

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

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

Above, $\hat{d} \overset{\text{def}}{=} \max\Big(d, \max_j\big(\deg(g_j)\big)\Big)$, i.e., the largest of 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 $d\ge 2$).

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 $\binom{q\cdot n_w+\lceil\hat{d}/2\rceil}{\lceil\hat{d}/2\rceil}$, and only $\mathcal{O}\big(\sum_k|\mathcal{J}_k|\big)$ SDP constraints of size $\binom{T\cdot n_w+\lceil\hat{d}/2\rceil}{\lceil\hat{d}/2\rceil}$.

C. 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 $\mathcal{W}_{[k]}$, i.e.,

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

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 (13) and the cost reformulation constraints (16a), (16b).

2) While we restrict the exposition to uncertainties that are only affecting the system dynamics additively, i.e., by means of (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 paper, 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 $\mathcal{W}_0\times\cdots\times\mathcal{W}_{T-1}$, and known probability distribution function $\mathbb{F}$. The goal

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

is to find $\boldsymbol{u}_0, \ldots, \boldsymbol{u}_{T-1}$ so as to obey the state-control constraints (3) almost surely, and to minimize the expected costs,

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

In this case, since our framework already enforces almost sure (robust) constraint satisfaction, the only potential modifications would be in the 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* $\mathbb{F}$

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

In this case, if partial information (such as the moments of the distribution up to degree d) is available, then the framework in (a) could 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 introduced in our framework can actually become tight. More precisely, the paper [39] 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 $\sup_{\mathbb{F}\in\mathcal{F}}\mathbb{E}_{\boldsymbol{w}_{[k]}}\,h_k(\boldsymbol{x}_k,\boldsymbol{u}_k)$ (where h_k are piecewise 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 constrained by (16b). In other words, introducing a (potentially dominating) polynomial stage cost $\tilde{h}_k$ does not increase the optimal value of the problem under the distributionally-robust framework.

In general, if more information about the measures in the set $\mathcal{F}$ is available, such as uni-modality, symmetry, directional deviations ([40]), then one should be able to obtain better bounds on the stage costs h_k, by employing appropriate Tchebycheff-type inequalities (see [39], [41], [42], and the recent papers [28], [29], and [43], 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 paper to the original worst-case formulation.

IV. 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 $\cap$-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 [31]. 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).

A. Affine and Quadratic Policies for $\cap$-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 (this is one of the most general classes of uncertainty sets treated in the robust optimization literature, see, e.g., [31]).

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 piecewise 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}):

$$\min_{L_k,\boldsymbol{z}_k,z_{k,0},J}\quad J \tag{17a}$$

$$(P_{\mathrm{AFF}})\qquad J \geq \sum_{k=0}^{T}\big(\boldsymbol{z}_k^T\boldsymbol{w}_{[k]} + z_{k,0}\big) \tag{17b}$$

$$\begin{aligned}\boldsymbol{z}_k^T\mathcal{B}_1(\boldsymbol{w}_{[k]}) &\geq c_0(k,i) + \boldsymbol{c}_x(k,i)^T\boldsymbol{x}_k(\boldsymbol{w}_{[k]}) \\ &\quad + \boldsymbol{c}_u(k,i)^T\boldsymbol{u}_k(\boldsymbol{w}_{[k]}), \\ &\quad \forall\,\boldsymbol{w}_{[k]}\in\mathcal{W}_{[k]},\ \forall\,i\in\mathcal{I}_k,\ \forall\,k\in\{0,\ldots,T-1\}\end{aligned} \tag{17c}$$

$$\begin{aligned}\boldsymbol{z}_T^T\mathcal{B}_1(\boldsymbol{w}_{[T]}) &\geq c_0(T,i) + \boldsymbol{c}_x(T,i)^T\boldsymbol{x}_T(\boldsymbol{w}_{[T]}), \\ &\quad \forall\,\boldsymbol{w}_{[T]}\in\mathcal{W}_{[T]},\ \forall\,i\in\mathcal{I}_T\end{aligned} \tag{17d}$$

$$\begin{aligned}\big(\,\boldsymbol{x}_{k+1}(\boldsymbol{w}_{[k+1]}) &= A_k\,\boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + B_k\,\boldsymbol{u}_k(\boldsymbol{w}_{[k]}) + \boldsymbol{w}(k),\ \big) \\ &\quad \forall\,k\in\{0,\ldots,T-1\}\end{aligned} \tag{17e}$$

$$\begin{aligned}\boldsymbol{f}_k &\geq E_x(k)\,\boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + E_u(k)\,\boldsymbol{u}_k(\boldsymbol{w}_{[k]}), \\ &\quad \forall\,\boldsymbol{w}_{[k]}\in\mathcal{W}_{[k]}, \\ &\quad \forall\,k\in\{0,\ldots,T-1\}\end{aligned} \tag{17f}$$

[8]We are grateful to one of the anonymous referees for pointing out reference [31], which was not at our disposal at the time of conducting the research.

[9]This is the same approach as that taken in [31]; when the stage costs h_k are already affine in $\boldsymbol{x}_k, \boldsymbol{u}_k$, the step is obviously not necessary.

$$\boldsymbol{f}_T \geq E_x(T)\,\boldsymbol{x}_T(\boldsymbol{w}_{[T]}), \quad \forall\, \boldsymbol{w}_{[T]} \in \mathcal{W}_{[T]}. \tag{17g}$$

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 (17e) 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 (17c), (17d), (17f), and (17g). 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 [31] can be used to compute the optimal decisions in (P_{AFF}).

Letting $J^\star_{\mathrm{AFF}}$ 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 1: 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:

$$J^\star_{\mathrm{AFF}} \geq J^\star_{d=1} \geq J^\star_{d=2} \geq \cdots$$

Proof: See Section VIII-C of the Appendix. ∎

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 (17c) and (17d), 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 [31] for details). If we let $J^\star_{\mathrm{QUAD}}$ denote the optimal value resulting from this method, a proof paralleling that of Theorem 1 can be used to show that $J^\star_{\mathrm{QUAD}} \geq J^\star_{d=2}$, 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.

B. 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 III 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 (2) are actually affine. It can be shown (see Theorem 2 in [11]) that piecewise 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, $\boldsymbol{w}_{[T]} \in \mathrm{ext}(\mathcal{W}_0)\times\cdots\times\mathrm{ext}(\mathcal{W}_{T-1})$. 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 problem (see [6], [10], and [11] for a proof):

$$\min_{\boldsymbol{u}_k(\boldsymbol{w}_{[k]}),z_k(\boldsymbol{w}_{[k]}),J} \quad J \tag{18a}$$

$$J \geq \sum_{k=0}^{T} z_k(\boldsymbol{w}_{[k]}) \tag{18b}$$

$$(P)_{\mathrm{ext}} \quad z_k(\boldsymbol{w}_{[k]}) \geq h_k\big(\boldsymbol{x}_k(\boldsymbol{w}_{[k]}), \boldsymbol{u}_k(\boldsymbol{w}_{[k]})\big), \quad k=0,\ldots,T-1 \tag{18c}$$

$$z_T(\boldsymbol{w}_{[T]}) \geq h_T\left(\boldsymbol{x}_T(\boldsymbol{w}_{[T]})\right) \tag{18d}$$

$$\boldsymbol{x}_{k+1}(\boldsymbol{w}_{[k+1]}) = A_k\,\boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + B_k\,\boldsymbol{u}_k(\boldsymbol{w}_{[k]}) + \boldsymbol{w}(k), \quad k=0,\ldots,T-1 \tag{18e}$$

$$\boldsymbol{f}_k \geq E_x(k)\,\boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + E_u(k)\,\boldsymbol{u}_k(\boldsymbol{w}_{[k]}), \quad k=0,\ldots,T-1 \tag{18f}$$

$$\boldsymbol{f}_T \geq E_x(T)\,\boldsymbol{x}_T(\boldsymbol{w}_{[T]}). \tag{18g}$$

In this formulation, *nonanticipatory* $\boldsymbol{u}_k(\boldsymbol{w}_{[k]})$ control values and corresponding states $\boldsymbol{x}_k(\boldsymbol{w}_{[k]})$ are computed for every vertex of the disturbance set, i.e., for every $\boldsymbol{w}_{[k]} \in \mathrm{ext}(\mathcal{W}_0)\times\cdots\times\mathrm{ext}(\mathcal{W}_{k-1})$, $k=0,\ldots,T-1$. 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 (18c), (18d) can be immediately rewritten in linear form, since the functions $h_k(\boldsymbol{x},\boldsymbol{u})$, $h_T(\boldsymbol{x})$ are piecewise 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 *nonanticipativity* 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 $(P)_{\mathrm{ext}}$ 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, $\mathcal{O}\big(\prod_{k=0}^{T-1}|\mathrm{ext}(\mathcal{W}_k)|\big)$. Therefore, solving $(P)_{\mathrm{ext}}$ is relevant only for small horizons, but is very useful for benchmarking purposes, since it provides the optimal value of the original problem.

[10] One could also immediately extend the result of [17] to argue that disturbance-feedback policies $\boldsymbol{u}_k(\boldsymbol{w}_{[k]})$ are also optimal.

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

We conclude this section by examining a particular example when the uncertainty sets take an even simpler form, and polynomial policies (11) are provably optimal. More precisely, we consider the case of scalar uncertainties $(n_w = 1)$, and

$$\boldsymbol{w}(k) \in \mathcal{W}(k) \stackrel{\text{def}}{=} [\underline{w}_k, \bar{w}_k] \subset \mathbb{R},\ \forall\, k = 0, \ldots, T-1 \quad (19)$$

known in the literature as *box uncertainty* [23], [44]. Under this model, any partial uncertain sequence $\boldsymbol{w}_{[k]} \stackrel{\text{def}}{=} (w_0, \ldots, w_{k-1})$ will be a k-dimensional vector, lying inside the hypercube $\mathcal{W}_{[k]} \stackrel{\text{def}}{=} \mathcal{W}_0 \times \cdots \times \mathcal{W}_{k-1} \subset \mathbb{R}^k$.

Introducing the subclass of *multiaffine* policies[12] of degree d, given by

$$u_j(k, \boldsymbol{w}_{[k]}) = \sum_{\boldsymbol{\alpha} \in \{0,1\}^k} \ell_{\boldsymbol{\alpha}} (\boldsymbol{w}_{[k]})^{\boldsymbol{\alpha}},\ \text{where} \sum_{i=1}^{k} \alpha_i \le d \quad (20)$$

one can show (see Theorem 2 in the Appendix) that multiaffine 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 [31] show one can seek *separable* polynomial policies of the form

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

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 exact (refer to Lemma 14.3.4 in [31]). 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 [31]). However, as we will see in our numerical results (to be presented in Section VI), polynomials policies that are *not* separable, i.e., are of the general form (11), can and do improve over the affine case.

V. Examples From Inventory Management

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

A. Single Echelon With Cumulative Order Constraints

This first example was originally discussed in a robust framework by [45], in the context of a more general model for the problem of negotiating flexible contracts between a retailer and a supplier in the presence of uncertain orders from customers. We describe a simplified version of the problem, which is sufficient to illustrate the benefit of our approach, and refer the interested reader to [45] for more details.

The setting is the following: consider a single-product, single-echelon, multiperiod 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, $L_k \le u_k \le U_k$, as well as on the cumulative orders places in consecutive periods, $\hat{L}_k \le \sum_{t=0}^{k} u_t \le \hat{U}_k$. 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 $\mathcal{W}_k = \{w_k \in \mathbb{R} : |w_k - \bar{d}_k| \le \rho \cdot \bar{d}_k\}$, 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

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

and the costs

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

We remark that the cumulative order constraints, $\hat{L}_k \le \sum_{t=0}^{k} u_t \le \hat{U}_k$, are needed here, since otherwise, the resulting (one-dimensional) system would fit the theoretical results from [8], which would imply that polynomial policies of the form (11) and polynomial stage costs of the form (16b) are already optimal for degree $d = 1$ (affine). Therefore, testing for higher order polynomial policies would not add any benefit.

B. 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

[12]Note that these are simply polynomial policies of the form (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 $\alpha_i \in \{0,1\}$.

a planning horizon of T periods by a centralized decision maker. The jth 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 jth 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 nonnegative, and we assume that the last echelon, J, can replenish inventory from a supplier with infinite capacity.

Following a standard requirement in inventory theory [46], 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, $w(k)$.

The dynamics of the linear system can then be formulated as

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

with the following constraints on the states and controls

$$\begin{aligned} & u_j(k) \ge 0,\ j = 1,\ldots,J,\ k = 0,\ldots,T-1, \\ &\qquad \text{(nonnegative shipments)} \\ & x_j(k) \ge u_{j-1}(k), j = 2,\ldots,J,\ k = 0,\ldots,T-1, \\ &\qquad \text{(downstream order} \le \text{upstream inventory)} \end{aligned}$$

and the costs

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

[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.

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).

C. Active Suspension System

Our third example, originally suggested in [47] and also appearing in [11], consists of the problem of robustly regulating to the origin the following active suspension system:

$$\begin{aligned} \boldsymbol{x}(k+1) &= \begin{bmatrix} 0.809 & 0.009 & 0 & 0 \\ -36.93 & 0.80 & 0 & 0 \\ 0.191 & -0.009 & 1 & 0.01 \\ 0 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{x}(k) \\ &\quad + \begin{bmatrix} 0.0005 \\ 0.0935 \\ -0.005 \\ -0.0100 \end{bmatrix} u(k) + \begin{bmatrix} -0.009 \\ 0.191 \\ -0.0006 \\ 0 \end{bmatrix} w(k). \end{aligned}$$

We follow the formulation in [11], and consider $T = 4$, with a cost function

$$\|P\,\boldsymbol{x}(T+1)\|_\infty + \sum_{k=1}^{T} \big(\|Q\,\boldsymbol{x}(k)\|_\infty + |R\,u(k)|\big)$$

where $P = Q = \operatorname{diag}\{5000,\ 0.1,\ 400,\ 0.1\}$, $R = 1.8$, input constraints $-5 \le u(k) \le 5$, state constraints

$$\begin{bmatrix} -0.02 \\ -\infty \\ -0.05 \\ -\infty \end{bmatrix} \le \boldsymbol{x}(k) \le \begin{bmatrix} 0.02 \\ +\infty \\ 0.05 \\ +\infty \end{bmatrix}$$

and a disturbance set $-0.4 \le w(k) \le 0.4$, $\forall\, k \in \{1,\ldots,T\}$.

VI. Numerical Experiments

In this section, we present numerical simulations testing the performance of polynomial policies in each of the three problems mentioned in Section V.

All our computations were done in a MATLAB environment, on the MIT Operations Research Center computational machine (3-GHz Intel Dual Core Xeon Processor, with 8 GB of RAM memory, running Ubuntu Linux). The optimization problems were formulated using YALMIP [48], and the resulting SDPs were solved with SDPT3 [49].

A. 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 perform the following.

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 following.
    - The optimal cost OPT, by solving the exponential Linear Program $(P)_{\text{ext}}$.

TABLE I
RELATIVE GAPS (IN PERCENT) 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 II
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

- 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 III.

  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.

Tables I and II 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, Fig. 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. Fig. 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 Fig. 2(b), shows promise for scalability—for horizon $T = 10$, the median and average solver times are below 15 s.

  We remark that the computational times could be substantially reduced by exploiting the structure of the polynomial optimization problems (e.g., [50]), and by utilizing more suitable techniques for solving smooth large-scale SDPs (see, e.g., [51] and the references therein). Such techniques are immediately applicable to our setting, and could provide a large speed-up over general-purpose algorithms (such as the interior point methods implemented in SDPT3), hence allowing much larger and more complicated instances to be solved.

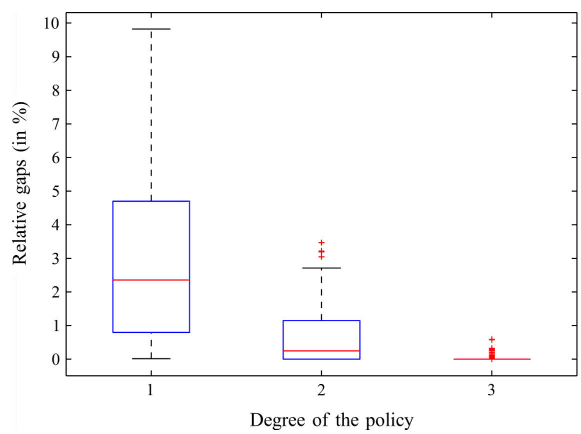


Fig. 1. Box plots comparing the performance of different polynomial policies for horizon $T = 6$.

B. 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 Tables III and IV, 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 Fig. 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 Fig. 3 (b), which corresponds to quadratic policies).

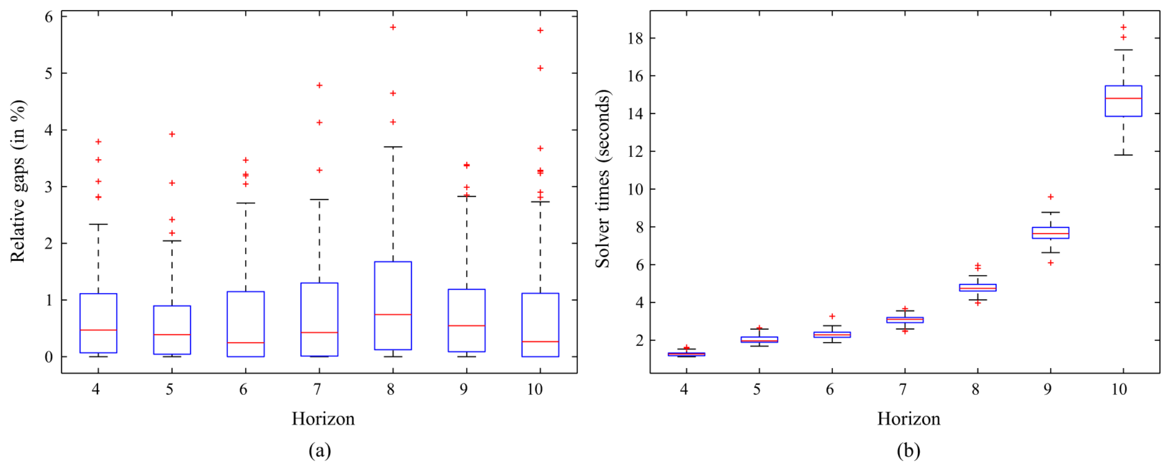


Fig. 2. Performance of quadratic policies for Example 1. (a) illustrates the weak dependency of the improvement on the problem size (measured in terms of the horizon T), while (b) compares the solver times required for different problem sizes.

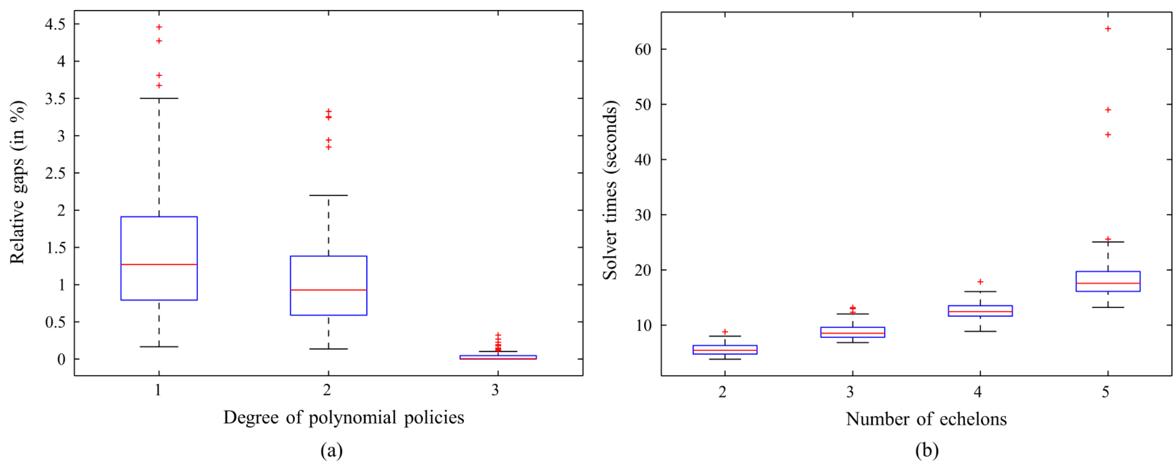


Fig. 3. Performance of polynomial policies for Example 2. (a) compares the three policies for problems with $J = 3$ echelons, and (b) shows the solver times needed to compute quadratic policies for different problem sizes.

TABLE III
RELATIVE GAPS (IN PERCENT) 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 IV
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

C. Third Example

For the third example, we fix the initial state to $\boldsymbol{x}(1) = (-0.01, -1.0, -0.03, -0.5)$, and compute polynomial control policies with degree $d \in \{1, \ldots, 3\}$. The worst case cost for $d = 1$ is approximately 422.28, while the worst-case cost under $d = 2$ and $d = 3$ is 367.12 (in this case, cubic polynomials do not offer considerable improvement over quadratic).

To understand the actions taken under the different policies, we simulate 10000 uncertainty sequences, and we compute the relevant controls and states. Fig. 4 depicts box-plots of the state vector $\boldsymbol{x}(k)$ under the three different schemes, with (a), (b), and (c) denoting polynomials with $d = 1$, $d = 2$ and $d = 3$, respectively. As can be seen from the figures, all three control schemes are able to drive the state vector close to the origin, and the effect is particularly obvious for the first component, $x_1(k)$, which is

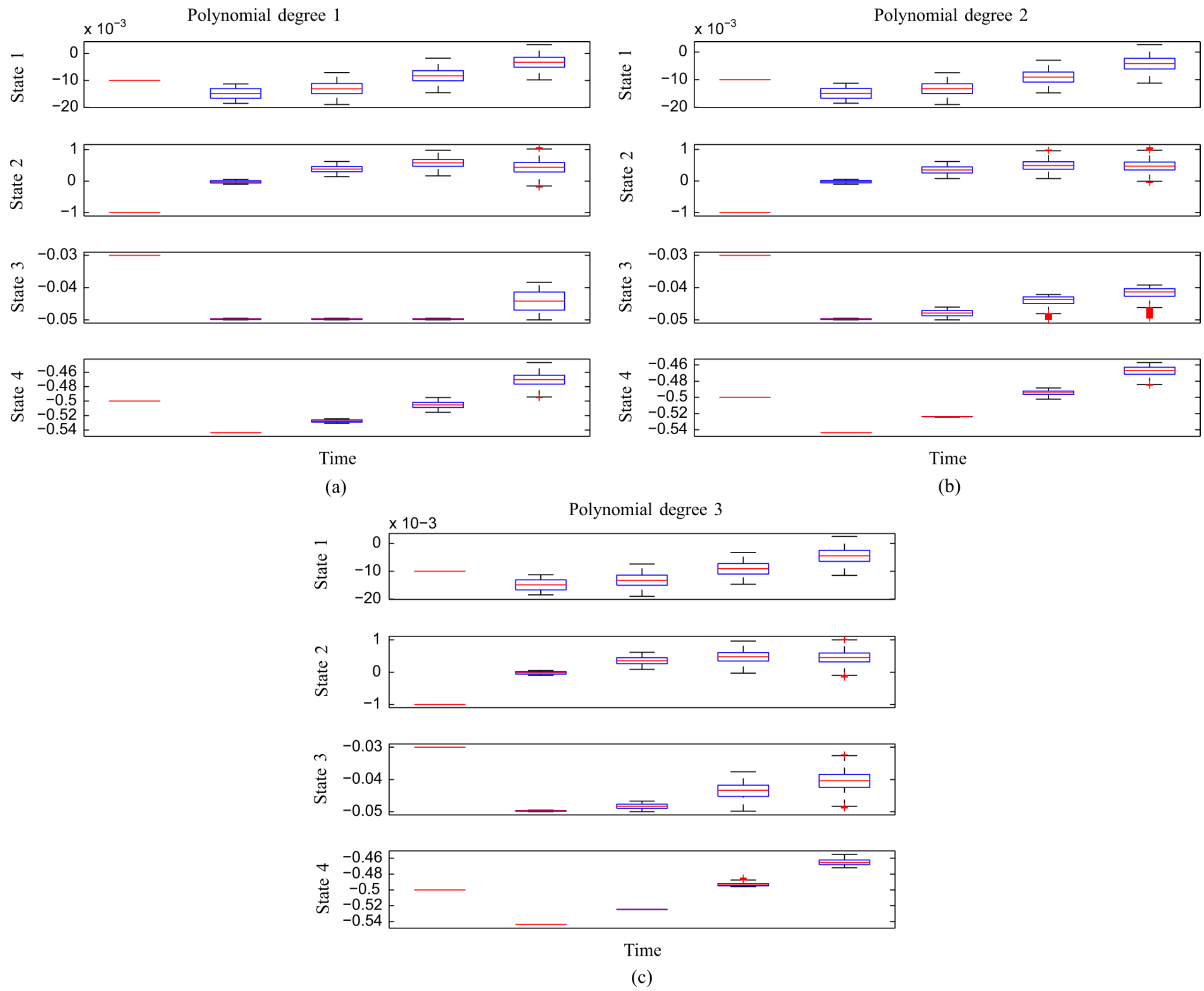


Fig. 4. Performance of polynomial policies in Example 3. A total of 10000 uncertainty sequences are generated, and (a), (b), and (c) record box-plots of the corresponding states under policies of degree 1, 2 and 3, respectively.

the main driver of the cost. However, quadratic and cubic policies typically generate less variability of the state values than affine policies, and also incur a considerably smaller worst-case cost (a reduction of around 13%).

VII. Conclusion

In this paper, we have presented a new method for dealing with multistage decision problems affected by uncertainty, applicable to robust optimization and stochastic programming. Our approach consists of constructing a hierarchy of suboptimal polynomial policies, parameterized directly in the observed uncertainties. The problem of computing such an optimal polynomial policy can be reformulated as an SDP, which can be solved efficiently with interior point methods. Furthermore, the approach allows modelling flexibility, in that the degree of the polynomial policies explicitly controls the trade-off between the quality of the approximation and the computational requirements. To test the quality of the policies, we have considered two applications in inventory management and one in the robust control of an active suspension system. For all our examples, quadratic policies (requiring modest computational requirements) were able to substantially reduce the optimality gap, and cubic policies (under more computational requirements) were always within 1% of optimal. Given that our tests were run using publicly-available, general-purpose SDP solvers, we believe that, with the advent of more powerful (commercial) packages for interior point methods, as well as dedicated algorithms for solving SOS problems, our method should have applicability to large scale, real-world optimization problems.

Appendix

Suboptimality of Affine Policies:

Lemma 1: Consider Problem (9), written below for convenience. Recall that x is a (first-stage) nonadjustable decision, while $\boldsymbol{y}$ is a second-stage adjustable *policy* (allowed to depend on $\boldsymbol{w}$).

$$\begin{aligned} &\underset{x,\boldsymbol{y}(\boldsymbol{w})}{\text{minimize}} && x \\ &\text{such that} && x \geq \sum_{i=1}^{N} y_i, \quad \forall \boldsymbol{w} \in \mathcal{W} = \left\{\boldsymbol{w} \in \mathbb{R}^N : \|\boldsymbol{w}\| \leq 1\right\} && \text{(21a)} \\ & && y_i \geq w_i^2, \quad \forall \boldsymbol{w} \in \mathcal{W}. && \text{(21b)} \end{aligned}$$

The optimal value in the problem is 1, corresponding to policies $y_i(\boldsymbol{w}) = w_i^2,\ i = 1, \ldots, N$. 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 (21b), 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, $y_i^{\mathrm{AFF}}(\boldsymbol{w}) = \beta_i + \boldsymbol{\alpha}_i^T \boldsymbol{w},\ i = 1, \ldots, N$. With $\boldsymbol{e}_1$ denoting the first unit vector (1 in the first component, 0 otherwise), for any $i = 1, \ldots, N$, we have

$$\left.\begin{array}{l} \boldsymbol{w} = \boldsymbol{e}_1 \in \mathcal{W} \Rightarrow \beta_i + \alpha_i(1) \geq 1 \\ \boldsymbol{w} = -\boldsymbol{e}_1 \in \mathcal{W} \Rightarrow \beta_i - \alpha_i(1) \geq 1 \end{array}\right\} \Rightarrow \beta_i \geq 1.$$

This implies that $x^{\mathrm{AFF}} \geq \sum_{i=1}^N y_i^{\mathrm{AFF}}(\boldsymbol{w}) \geq N + \sum_{i=1}^N \boldsymbol{\alpha}_i^T \boldsymbol{w}$. In particular, with $\boldsymbol{w} = \boldsymbol{0} \in \mathcal{W}$, we have $x^{\mathrm{AFF}} \geq N$. The optimal choice, in this case, will be to set $\boldsymbol{\alpha}_i = \boldsymbol{0}$, resulting in $x^{\mathrm{AFF}} = N$.

Optimality of MultiAffine Policies:

Theorem 2: Multiaffine policies of the form (20), 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 multiaffine policy u_j of the form (20) 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 $(P)_{\mathrm{ext}}$ from Section IV-B. Let us denote the optimal decisions obtained from solving problem $(P)_{\mathrm{ext}}$ by $\boldsymbol{u}_k^{\mathrm{ext}}(\boldsymbol{w}_{[k]}), \boldsymbol{x}_k^{\mathrm{ext}}(\boldsymbol{w}_{[k]})$, respectively. Note that, at time k, there are at most 2^k such distinct values $\boldsymbol{u}_k^{\mathrm{ext}}(\boldsymbol{w}_{[k]})$, and, correspondingly, at most 2^k values $\boldsymbol{x}_k^{\mathrm{ext}}(\boldsymbol{w}_{[k]})$, due to the nonanticipativity condition and the fact that the extreme uncertainty sequences at time k, $\boldsymbol{w}_{[k]} \in \mathrm{ext}(\mathcal{W}_{[k]}) = \mathrm{ext}(\mathcal{W}_0) \times \cdots \times \mathrm{ext}(\mathcal{W}_{k-1})$, 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 $\boldsymbol{u}_{T-1}^{\mathrm{ext}}(\boldsymbol{w}_{[T-1]})$ computed.

Consider now a multiaffine policy of the form (20), of degree $T-1$, implemented at time $T-1$. By Observation 1, the number of coefficients in the jth 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

$$\boldsymbol{u}_{T-1}(\boldsymbol{w}_{[T-1]}) = \boldsymbol{u}_{T-1}^{\mathrm{ext}}(\boldsymbol{w}_{[T-1]}),\ \forall\, \boldsymbol{w}_{[T-1]} \in \mathrm{ext}(\mathcal{W}_{[T-1]}) \tag{22}$$

i.e., the value of the multiaffine policy exactly matches the 2^{T-1} optimal decisions computed in $(P)_{\mathrm{ext}}$, 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 multiaffine policies of degree at most[14] $T - 1$ that match the values $\boldsymbol{u}_k^{\mathrm{ext}}(\boldsymbol{w}_{[k]})$ at the extreme points of $\mathcal{W}_{[k]}$.

[14]In fact, multiaffine policies of degree k would be sufficient at time k

With such multiaffine control policies, it is easy to see that the states $\boldsymbol{x}_k$ become multiaffine functions of $\boldsymbol{w}_{[k]}$. Furthermore, we have $\boldsymbol{x}_k(\boldsymbol{w}_{[k]}) = \boldsymbol{x}_k^{\mathrm{ext}}(\boldsymbol{w}_{[k]}),\ \forall\, \boldsymbol{w}_{[k]} \in \mathrm{ext}(\boldsymbol{w}_{[k]})$. A typical state-control constraint (7c) written at time k amounts to ensuring that

$$\begin{aligned} \boldsymbol{e}_x(k,j)^T \boldsymbol{x}_k(\boldsymbol{w}_{[k]}) + \boldsymbol{e}_u(k,j)^T \boldsymbol{u}_k(\boldsymbol{w}_{[k]}) \\ -\boldsymbol{f}_j(k) \leq 0, \quad \forall\, \boldsymbol{w}_{[k]} \in \mathcal{W}_{[k]} \end{aligned}$$

where $\boldsymbol{e}_x(k,j)^T, \boldsymbol{e}_u(k,j)^T$ denote the jth row of $E_x(k)$ and $E_u(k)$, respectively. Note that the left-hand side of this expression is also a multiaffine function of the variables $\boldsymbol{w}_{[k]}$. Since, by our observation, the maximum of multiaffine functions is reached at the vertices of the feasible set, i.e., $\boldsymbol{w}_{[k]} \in \mathrm{ext}(\mathcal{W}_{[k]})$, and, by (22), we have that for any such vertex, $u_k(\boldsymbol{w}_{[k]}) = \boldsymbol{u}_k^{\mathrm{ext}}(\boldsymbol{w}_{[k]}), \boldsymbol{x}_k(\boldsymbol{w}_{[k]}) = \boldsymbol{x}_k^{\mathrm{ext}}(\boldsymbol{w}_{[k]})$, we immediately conclude that the constraint above is satisfied, since $\boldsymbol{u}_k^{\mathrm{ext}}(\boldsymbol{w}_{[k]}), \boldsymbol{x}_k^{\mathrm{ext}}(\boldsymbol{w}_{[k]})$ are certainly feasible.

A similar argument can be invoked for constraint (7d), and also to show that the maximum of the objective function is reached on the set of vertices $\mathrm{ext}(\mathcal{W}_{[T]})$, and, since the values of the multiaffine policies exactly correspond to the optimal decisions in program $(P)_{\mathrm{ext}}$, optimality is preserved. ■

Comparison With Other Methodologies:

Theorem: 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:

$$J_{\mathrm{AFF}}^\star \geq J_{d=1}^\star \geq J_{d=2}^\star \geq \cdots$$

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, $\forall\, k \in \{0, \ldots, T-1\}$, which implies that $\mathcal{W}_{[k]} = \mathcal{W}_0 \times \cdots \times \mathcal{W}_{k-1}$ is also convex, with nonempty 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 [52]), 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 nonnegative, for any $\boldsymbol{w}_{[k]}$. Therefore, since any such polynomial can be represented as a sum-of-squares (see [35] and [36]), we also have that σ_0 is SOS.

By these two observations, we can conclude that the particular choice L_k, z_k, J remains feasible in our SOS framework applied to degree $d = 1$, and, hence, $J^\star_{\mathrm{AFF}} \geq J^\star_{d=1}$. ∎

ACKNOWLEDGMENT

The authors would like to thank the five referees and the associate editor for the careful reading of the paper and the thoughtful comments and suggestions, which have considerably improved the quality of the paper.

REFERENCES

[1] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 2001.

[2] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. New York: Springer, 2000.

[3] K. Zhou and J. C. Doyle, *Essentials of Robust Control*. : Prentice Hall, 1998.

[4] G. E. Dullerud and F. Paganini, *A Course in Robust Control Theory*. New York: Springer, 2005.

[5] E. Kerrigan and J. Maciejowski, "On robust optimization and the optimal control of constrained linear systems with bounded state disturbances," in *Proc. 2003 Eur. Control Conf.*, Cambridge, UK, Sep. 2003.

[6] A. Ben-Tal, S. Boyd, and A. Nemirovski, *Control of uncertainty-affected discrete time linear systems via convex programming*. 2005, submitted for publication.

[7] A. Ben-Tal, S. Boyd, and A. Nemirovski, "Extending scope of robust optimization: Comprehensive robust counterparts of uncertain problems," *Math. Program.*, vol. 107, no. 1, pp. 63–89, 2006.

[8] D. Bertsimas, D. A. Iancu, and P. A. Parrilo, "Optimality of affine policies in multistage robust optimization," *Math. Oper. Res.*, vol. 35, no. 2, pp. 363–394, 2010.

[9] A. Bemporad, F. Borrelli, and M. Morari, "Optimal controllers for hybrid systems: Stability and piecewise linear explicit form," in *Proc. 39th IEEE Conf. Decision and Control*, 2000, vol. 2, pp. 1810–1815.

[10] A. Bemporad, F. Borrelli, and M. Morari, "Model predictive control based on linear programming—the explicit solution," *IEEE Trans. Autom. Control*, vol. 47, no. 12, pp. 1974–1985, Dec. 2002.

[11] A. Bemporad, F. Borrelli, and M. Morari, "Min-max control of constrained uncertain discrete-time linear systems," *IEEE Trans. Autom. Control*, vol. 48, no. 9, pp. 1600–1606, Sep. 2003.

[12] D. P. Bertsekas and J. N. Tsitsiklis, *Neurodynamic Programming*. Belmont, MA: Athena Scientific, 1996.

[13] W. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality, ser. Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley-Interscience, 2007.

[14] S. Asmussen and P. Glynn, *Stochastic Simulation: Algorithms and Analysis*. New York: Springer, 2007.

[15] P. Glasserman and S. Tayur, "Sensitivity analysis for base-stock levels in multiechelon production-inventory systems," *Manage. Sci.*, vol. 41, no. 2, pp. 263–281, 1995.

[16] P. Marbach and J. N. Tsitsiklis, "Simulation-based optimization of markov reward processes," *IEEE Trans. Autom. Control*, vol. 46, no. 2, pp. 191–209, Feb. 2001.

[17] S. J. Garstka and R. J.-B. Wets, "On decision rules in stochastic programming," *Math. Program.*, vol. 7, no. 1, pp. 117–143, 1974.

[18] M. Dyer and L. Stougie, "Computational complexity of stochastic programming problems," *Math. Program.*, vol. 106, no. 3, pp. 423–432, 2006.

[19] A. Nemirovski and A. Shapiro, "On Complexity of Stochastic Programming Problems," in *Continuous Optimization: Current Trends and Applications*, ser. Applied Optimization. New York: Springer, 2005, vol. 99, pp. 111–146.

[20] J. Löfberg, "Approximations of closed-loop minimax MPC," in *Proc. 42nd IEEE Conf. Decision and Control*, Dec. 2003, vol. 2, pp. 1438–1442.

[21] E. Kerrigan and J. Maciejowski, "Properties of a new parameterization for the control of constrained systems with disturbances," in *Proc. 2004 Amer. Control Conf.*, Jun.–Jul. 2004, vol. 5, pp. 4669–4674 [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01384049

[22] P. J. Goulart and E. Kerrigan, "Relationships between affine feedback policies for robust control with constraints," in *Proc. 16th IFAC World Congress on Autom. Control*, Jul. 2005.

[23] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, "Adjustable robust solutions of uncertain linear programs," *Math. Program.*, vol. 99, no. 2, pp. 351–376, 2004.

[24] D. Bertsimas and D. B. Brown, "Constrained stochastic LQC: A tractable approach," *IEEE Trans. Autom. Control*, vol. 52, no. 10, pp. 1826–1841, Oct. 2007.

[25] J. Skaf and S. Boyd, "Design of affine controllers via convex optimization," *IEEE Trans. Autom. Control*, vol. 55, no. 11, pp. 2476–2487, Nov. 2010.

[26] J. Skaf and S. Boyd, "Nonlinear Q-design for convex stochastic control," *IEEE Trans. Autom. Control*, vol. 54, no. 10, pp. 2426–2430, Oct. 2009.

[27] X. Chen and Y. Zhang, "Uncertain linear programs: Extended affinely adjustable robust counterparts," *Oper. Res.*, vol. 57, no. 6, pp. 1469–1482, Nov./Dec. 2009.

[28] X. Chen, M. Sim, P. Sun, and J. Zhang, "A linear decision-based approximation approach to stochastic programming," *Oper. Res.*, vol. 56, no. 2, pp. 344–357, 2008.

[29] M. Sim and J. Goh, "Distributionally robust optimization and its tractable approximations," *Oper. Res.*, Apr. 2009.

[30] D. Chatterjee, P. Hokayem, and J. Lygeros, Stochastic Model Predictive Control With Bounded Control Inputs: A Vector Space Approach Mar. 2009 [Online]. Available: http://arxiv.org/abs/0903.5444

[31] A. Ben-Tal, L. El-Ghaoui, and A. Nemirovski, *Robust Optimization,*, ser. Princeton Series in Applied Mathematics. Princeton, NJ: Princeton Univ. Press, 2009.

[32] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, 1996.

[33] W. Rudin, *Principles of Mathematical Analysis*. New York: Mc-Graw Hill Science/Engineering/Math, 1976.

[34] P. A. Parrilo, "Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization," Ph.D. dissertation, California Inst. Technol., Pasadena, CA, 2000.

[35] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Math. Programm. Series B*, vol. 96, no. 2, pp. 293–320, May 2003.

[36] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM J. Optim.*, vol. 11, pp. 796–817, 2001.

[37] M. Putinar, "Positive polynomials on compact semi-algebraic sets," *Indiana Univ. Math. J.*, vol. 42, pp. 969–984, 2003.

[38] T. Jacobi and A. Prestel, "Distinguished representations of strictly positive polynomials," *J. Reine Angewandte Math. (Crelles Journal)*, vol. 2001, no. 532, pp. 223–235, 2001.

[39] L. F. Zuluaga and J. F. Pena, "A conic programming approach to generalized Tchebycheff inequalities," *Math. Oper. Res.*, vol. 30, no. 2, pp. 369–388, 2005.

[40] X. Chen, M. Sim, and P. Sun, "A robust optimization perspective on stochastic programming," *Oper. Res.*, vol. 55, no. 6, pp. 1058–1071, 2007.

[41] D. Bertsimas and I. Popescu, "Optimal inequalities in probability theory: A convex optimization approach," *SIAM J. Optim.*, vol. 15, no. 3, pp. 780–804, 2005.

[42] I. Popescu, "A semidefinite programming approach to optimal-moment bounds for convex classes of distributions," *Math. Oper. Res.*, vol. 30, no. 3, pp. 632–657, 2005.

[43] K. Natarajan, M. Sim, and J. Uichanco, "Tractable robust expected utility and risk models for portfolio optimization," *Math. Finance*, Dec. 2008.

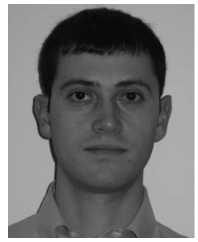
[44] A. Ben-Tal and A. Nemirovski, "Robust optimization—methodology and applications," *Math. Program.*, vol. 92, no. 3, pp. 453–480, 2002.

[45] A. Ben-Tal, B. Golany, A. Nemirovski, and J.-P. Vial, “Retailer-supplier flexible commitments contracts: A robust optimization approach,” *Manufact. Service Oper. Manage.*, vol. 7, no. 3, pp. 248–271, 2005.
[46] P. Zipkin*, Foundations of Inventory Management*. New York: McGraw Hill, 2000.
[47] D. Hrovat, “Survey of advanced suspension developments and related optimal control applications,” *Automatica*, vol. 33, no. 10, pp. 1781–1817, 1997.
[48] J. Löfberg, “YALMIP: A toolbox for modeling and optimization in MATLAB,” in *Proc. CACSD Conf.*, Taipei, Taiwan, 2004 [Online]. Available: http://control.ee.ethz.ch/~joloef/yalmip.php
[49] K. Toh, M. Todd, and R. Tütüncü, SDPT3 a MATLAB Software Package for Semidefinite Programming 1999, pp. 545–581 [Online]. Available: http://www.math.nus.edu.sg/~mattohkc/sdpt3.html
[50] J. Nie, Regularization Methods for Sum of Squares Relaxations in Large Scale Polynomial Optimization Sep. 2009 [Online]. Available: http://arxiv.org/abs/0909.3551
[51] G. Lan, Z. Lu, and R. D. C. Monteiro, “Primal-dual first-order methods with $O(1/\epsilon)$ iteration-complexity for cone programming,” *Math. Program.*, 2009.
[52] D. P. Bertsekas, A. Nedic, and A. Ozdaglar*, Convex Analysis and Optimization*. Belmont, MA: Athena Scientific, 2003.

Dimitris Bertsimas received the B.S. degree in electrical engineering and computer science at the National Technical University of Athens, Athens, Greece in 1985, the M.S. degree in operations research at the Massachusetts Institute of Technology (MIT), Cambridge, in 1987, and the Ph.D. degree in applied mathematics and operations research at MIT in 1988.

He is currently the Boeing Professor of Operations Research and the Codirector of the Operations Research Center at MIT and has been with the MIT faculty since 1988. His research interests include optimization, stochastic systems, data mining, and their application. In recent years he has worked in robust optimization, health care, and finance. He has coauthored more than 125 scientific papers and he has coauthored the following books: *Introduction to Linear Optimization* (with J. Tsitsiklis, Athena Scientific and Dynamic Ideas, 2008), *Data, Models and Decisions* (with R. Freund, Dynamic Ideas, 2004), and *Optimization over Integers* (with R. Weismantel, Dynamic Ideas, 2005). He is currently a Department Editor in *Optimization for Management Science* and former Area Editor for *Operations Research in Financial Engineering*. He has supervised 42 doctoral students and he is currently supervising ten others.

Dr. Bertsimas is a member of the National Academy of Engineering, and he has received numerous research awards including the Farkas Prize (2008), the Erlang Prize (1996), the SIAM Prize in Optimization (1996), the Bodossaki Prize (1998), and the Presidential Young Investigator Award (1991–1996).



Dan Andrei Iancu (S’09) received the B.S. degree in electrical engineering and computer science from Yale University, New Haven, CT, in 2004, the S.M. degree in engineering sciences from Harvard University, Cambridge, MA, in 2006, and the Ph.D. degree in operations research from the Massachusetts Institute of Technology, Cambridge, in 2010.

He is currently an Assistant Professor of operations, information and technology at the Graduate School of Business, Stanford University, Stanford, CA. Prior to joining Stanford, he spent one year as a Goldstine Fellow in the Mathematical Sciences Department, IBM T. J. Watson Research Center. His research interests include dynamic optimization under uncertainty, with an emphasis on robust optimization, connections with risk and utility theory, and applications in operations management and financial engineering.

Dr. Iancu was the recipient of the Best Student Paper Prize of the INFORMS Optimization Society (2009), and is a member of INFORMS and SIAM.

Pablo A. Parrilo (S’92–A’95–SM’07) received the electronics engineering undergraduate degree from the University of Buenos Aires, Buenos Aires, Argentina, in 1995 and the Ph.D. degree in control and dynamical systems from the California Institute of Technology, Pasadena, in 2000.

He has held short-term visiting appointments at the University of California—Santa Barbara, Lund Institute of Technology, and University of California—Berkeley. From 2001 to 2004, he was Assistant Professor of analysis and control systems at the Swiss Federal Institute of Technology (ETH Zurich). He is currently a Professor in the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, where he holds the Finmeccanica Career Development chair. He is affiliated with the Laboratory for Information and Decision Systems (LIDS) and the Operations Research Center (ORC). His research interests include optimization methods for engineering applications, control and identification of uncertain complex systems, robustness analysis and synthesis, and the development and application of computational tools based on convex optimization and algorithmic algebra to practically relevant engineering problems.

Dr. Parrilo received the 2005 Donald P. Eckman Award of the American Automatic Control Council, and the SIAM Activity Group on Control and Systems Theory (SIAG/CST) Prize. He is currently on the Board of Directors of the Foundations of Computational Mathematics (FoCM) Society and on the editorial board of the MPS/SIAM Book Series on Optimization.