Robust Optimal Control with Adjustable Uncertainty Sets [1]

Xiaojing Zhang[a], Maryam Kamgarpour[a], Angelos Georghiou[a], Paul Goulart[b], John Lygeros[a]

[a]*Automatic Control Laboratory, ETH Zurich, Switzerland*
[b]*Department of Engineering Science, Oxford University, United Kingdom*

Abstract

In this paper, we develop a unified framework for studying constrained robust optimal control problems with *adjustable uncertainty sets.* In contrast to standard constrained robust optimal control problems with known uncertainty sets, we treat the uncertainty sets in our problems as additional decision variables. In particular, given a finite prediction horizon and a metric for adjusting the uncertainty sets, we address the question of determining the optimal size and shape of the uncertainty sets, while simultaneously ensuring the existence of a control policy that will keep the system within its constraints for all possible disturbance realizations inside the adjusted uncertainty set. Since our problem subsumes the classical constrained robust optimal control design problem, it is computationally intractable in general. We demonstrate that by restricting the families of admissible uncertainty sets and control policies, the problem can be formulated as a tractable convex optimization problem. We show that our framework captures several families of (convex) uncertainty sets of practical interest, and illustrate our approach on a demand response problem of providing control reserves for a power system.

Keywords: Robust Optimal Control, Adjustable Uncertainty Sets, Affine Policies, Robust Optimization

1. Introduction

Robust finite-horizon optimal control of constrained linear systems subject to additive uncertainty has been studied extensively in the literature, both in the control [1, 2, 3] and operations research community [4, 5, 6]. Apart from issues such as stability and recursive feasibility that arise in the context of Model Predictive Control, significant amount of research is concerned with the approximation and efficient computation of the optimal control policies associated with such problems [7, 8, 9].

---

[1]This manuscript is the preprint of a paper submitted to Automatica and is subject to Elsevier copyright. Elsevier maintains the sole rights of distribution or publication of the work in all forms and media. If accepted, the copy of record will be available at http://www.journals.elsevier.com/automatica/.

✩E-mail addresses: `xiaozhan@control.ee.ethz.ch` (X. Zhang), `mkargam@control.ee.ethz.ch` (M. Kamgarpour), `angelosg@control.ee.ethz.ch` (A. Georghiou), `paul.goulart@eng.ox.ac.uk` (P. Goulart), `jlygeros@control.ee.ethz.ch` (J. Lygeros).

Commonly, robust control problems of constrained systems over a finite horizon deal with uncertainty sets that are known a priori. In this paper, we add another layer of complexity to these problems by allowing the uncertainty sets to be decision variables of our problems, and refer to such problems as constrained robust optimal control problems with *adjustable uncertainty sets.* For example, if the uncertainty sets are interpreted as a system's resilience against disturbance, then our framework can be used in a robustness analysis setup for determining the limits of robustness of a given system. The goal then is to determine the optimal size and shape of the uncertainty sets which maximize a given metric, while ensuring the existence of a control policy that will keep the system within its constraints. Unfortunately, such problems are computationally intractable in general, since they subsume the standard robust optimal control problem with fixed uncertainty set. The aim of this paper is to propose a systematic method for finding approximate solutions in a computationally efficient way.

Our work is motivated by reserve provision problems, where the adjustable uncertainty set is interpreted as a *reserve capacity*, which a system can offer to third parties, and for which it receives (financial) reward. In this case, the maximum reserve capacity can be computed by maximizing the size of the uncertainty set. Moreover, the reserve capacity is to be chosen such that for every admissible reserve demand, i.e. for every realization within the reserve capacity set, our system is indeed able to provide this reserve without violating its constraints. Reserve provision problems of this kind were first formulated and studied in [10] where it was shown that for uncertainty sets described by norm balls, the problems can be reformulated as tractable convex optimization problems. Another problem that admits the interpretation as a robust control problem with adjustable uncertainty set is *robust input tracking* [11, 12], where the aim is to determine the largest set of inputs that can be tracked by a system without violating its constraints. Reserve provision and input tracking problems have recently received increased attention in demand response applications of control reserves for electrical power grids [13, 14, 15].

The purpose of this paper is two-fold: First, we generalize the work of [10, 12] by considering a larger class of adjustable uncertainty sets based on techniques of conic convex optimization. Second, we provide a unified framework for studying reserve provision, input tracking and robustness analysis problems under the umbrella of constrained robust optimal control with adjustable uncertainty sets. The main contributions of this paper with respect to the existing literature can be summarized as follows:

- We show that if (i) the uncertainty sets are restricted to those that can be expressed as affine transformations of properly selected *primitive* convex sets, and (ii) the control policies are restricted to be affine with respect to the elements in these primitive sets, then the problems admit convex reformulations that can be solved efficiently, and whose size grows polynomially in the decision parameters. In particular, we extend the results of [10, 12] in two ways: First, we show that any convex set can be used as a primitive set, allowing us to target a much larger class of uncertainty sets. Second, by

allowing the primitive sets to be defined on higher dimensional spaces than those of the uncertainty sets, we are able to design more flexible uncertainty sets.

- We identify families of uncertainty sets of practical interest, including norm-balls, ellipsoids and hyper-rectangles, and show that they can be adjusted efficiently. Extending the work of [10, 12], we also show that by choosing the primitive set as the simplex, our framework enables us to efficiently optimize over compact polytopes with a predefined number of vertices. Furthermore, we prove that if the primitive sets are polytopes (e.g. the simplex), then our policy approximation gives rise to continuous piece-wise affine controllers.

- We study a reserve provision problem that arises in power systems, and show how it can be formulated as a robust optimal control problem with an adjustable uncertainty set. The problem is addressed using the developed tools, and we show that it can be formulated as a linear optimization problem of modest size that can be solved efficiently within 0.3 seconds, making it also practically applicable. In additional to results available in the literature [12, 13], we verify that our method produces the optimal solution for this problem for prediction horizons of up to eight time steps. For larger horizons, we show that the (relative) suboptimality gap is bounded by 0.71%.

This paper is organized as follows: Section 2 introduces the general problem setup. Section 3 focuses on the problem of adjusting the uncertainty sets, while in Section 4, we return to the original problem and restrict the family of control policies to obtain tractable instances thereof. Section 5 illustrates our approach on a demand response problem, while Section 6 demonstrates the usefulness of allowing the uncertainty sets to be projections of primitive convex sets. Finally, Section 7 concludes the paper. The Appendix contains auxiliary results needed to prove the main results of the paper.

Notation

For given matrices $(A_1, \ldots, A_n)$, we define $A := \mathrm{diag}(A_1, \ldots, A_n)$ as the block-diagonal matrix with elements $(A_1, \ldots, A_n)$ on its diagonal. A_{ij} denotes the (i, j)-th element of the matrix A, while $A_{.j}$ denotes the j-th column of A. Given a cone $\mathcal{K} \subseteq \mathbb{R}^l$ and two vectors $a, b \in \mathbb{R}^l$, $a \preceq_{\mathcal{K}} b$ implies $(b - a) \in \mathcal{K}$. For a matrix $B \in \mathbb{R}^{m \times l}$, $B \succeq_{\mathcal{K}} 0$ denotes row-wise inclusion in $\mathcal{K}$. For a symmetric matrix $C \in \mathbb{R}^{n \times n}$, $C \succeq 0$ denotes positive semi-definiteness of C. Given vectors $(v_1, \ldots, v_m)$, $v_i \in \mathbb{R}^l$, we denote their convex hull as $\mathrm{conv}(v_1, \ldots, v_m)$. Moreover, $[v_1, \ldots, v_m] := [v_1^\top \; \ldots \; v_m^\top]^\top \in \mathbb{R}^{lm}$ denotes their vector concatenation.

2. Problem Formulation

In this section, we formulate the robust optimal control problem with adjustable uncertainty set. We consider uncertain linear systems of the form

$$x_{k+1} = Ax_k + Bu_k + Ew_k, \tag{1}$$

where $x_k \in \mathbb{R}^{n_x}$ is the state at time step k given an initial state $x_0 \in \mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$ is the control input and $w_k \in \mathbb{W}_k \subseteq \mathbb{R}^{n_w}$ is an uncertain disturbance. We consider compact polytopic state and input constraints

$$
\begin{aligned}
x_k &\in \mathbb{X} := \{x \in \mathbb{R}^{n_x} : F_x x \leq f_x\}, \ k = 1, \dots, N, \\
u_k &\in \mathbb{U} := \{u \in \mathbb{R}^{n_u} : F_u u \leq f_u\}, \ k = 0, \dots, N-1,
\end{aligned}
\tag{2}
$$

where $F_x \in \mathbb{R}^{n_f \times n_x}$, $f_x \in \mathbb{R}^{n_f}$, $F_u \in \mathbb{R}^{n_g \times n_u}$, $f_u \in \mathbb{R}^{n_g}$, and n_f (n_g) is the number of state (input) constraints. Given a planning horizon N, we denote by $\phi_k(\mathbf{u}, \mathbf{w})$ the predicted state after k time steps resulting from the input sequence $\mathbf{u} := [u_0, \dots, u_{N-1}] \in \mathbb{R}^{N n_u}$ and disturbance sequence $\mathbf{w} := [w_0, \dots, w_{N-1}] \in \mathbb{R}^{N n_w}$.

In contrast to classical robust control problem formulations, we assume that the uncertainty set $\mathbb{W}_k$ is not fixed and needs to be adjusted according to some objective function $\varrho : \mathcal{P}(\mathbb{R}^{n_w}) \to \mathbb{R}$, where $\mathcal{P}(\mathbb{R}^{n_w})$ denotes the power set of $\mathbb{R}^{n_w}$. For example, we may think of $\varrho(\mathbb{W}_k)$ as the volume of $\mathbb{W}_k$, although depending on the application, it can represent other qualities such as the diameter or circumference of $\mathbb{W}_k$. Our objective is to maximize $\varrho(\mathbb{W}_k)$, while simultaneously minimizing some operating cost and ensuring satisfaction of input and state constraints. Hence, the cost to be minimized is given by

$$
\max_{\mathbf{w} \in \mathcal{W}} \{J(\mathbf{u}, \mathbf{w})\} - \lambda \sum_{k=0}^{N-1} \varrho(\mathbb{W}_k),
\tag{3}
$$

where $J(\mathbf{u}, \mathbf{w}) := \ell_f(\phi_N(\mathbf{u}, \mathbf{w})) + \sum_{k=0}^{N-1} \ell(\phi_{k+1}(\mathbf{u}, \mathbf{w}), u_k)$ is some "nominal" cost function with $\ell : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}$ and $\ell_f : \mathbb{R}^{n_x} \to \mathbb{R}$ linear, and $\lambda \geq 0$ is a user-defined weighting factor. Note that convex quadratic cost can also be incorporated in our framework by taking the certainty-equivalent cost $J(\mathbf{u}, \bar{\mathbf{w}})$ instead of the min-max cost in (3), where $\bar{\mathbf{w}}$ is some fixed (or expected) uncertainty. Due to the presence of the uncertainties $\mathbf{w}$, we consider the design of a causal disturbance feedback policy $\boldsymbol{\pi}(\cdot) := [\pi_0(\cdot), \dots, \pi_{N-1}(\cdot)]$, with each $\pi_k : \mathbb{W}_0 \times \dots \times \mathbb{W}_k \to \mathbb{R}^{n_u}$, such that the control input at each time step is given by $u_k = \pi_k(w_0, \dots, w_k)$[2]. Combining (1)–(3), we express the optimal control problem compactly as

$$
\begin{aligned}
\min \quad & \max_{\mathbf{w} \in \mathcal{W}} \{\mathbf{c}^\top \boldsymbol{\pi}(\mathbf{w})\} - \lambda \varrho(\mathcal{W}) \\
\text{s.t.} \quad & \boldsymbol{\pi}(\cdot) \in \mathcal{C}, \ \mathcal{W} \in \mathcal{P}(\mathbb{R}^{N n_w}), \\
& \mathbf{C}\boldsymbol{\pi}(\mathbf{w}) + \mathbf{D}\mathbf{w} \leq \mathbf{d}, \quad \forall \mathbf{w} \in \mathcal{W},
\end{aligned}
\tag{4}
$$

with decision variables $(\boldsymbol{\pi}(\cdot), \mathcal{W})$, and $\mathcal{W} := \mathbb{W}_0 \times \dots \times \mathbb{W}_{N-1}$, $\varrho(\mathcal{W}) := \sum_{k=0}^{N-1} \varrho(\mathbb{W}_k)$, and $\mathcal{C} := \{[\pi_0(\cdot), \dots, \pi_{N-1}(\cdot)] : \pi_k : \mathbb{W}_0 \times \dots \times \mathbb{W}_k \to \mathbb{R}^{n_u}, \ k = 0, \dots, N-1\}$ is the space of all causal control policies. $\mathbf{c} \in \mathbb{R}^{N n_u}$, $\mathbf{C} \in \mathbb{R}^{N(n_f + n_g) \times N n_u}$, $\mathbf{D} \in \mathbb{R}^{N(n_f + n_g) \times N n_w}$, and $\mathbf{d} \in \mathbb{R}^{N(n_f + n_g)}$ are matrices constructed from the problem data, see e.g. [16] for an example of such a construction. We call (4) a *robust optimal control problem with*

---

[2]Strictly causal policies can be incorporated by restricting $\pi_k(\cdot)$ to depend on $(w_0, \dots, w_{k-1})$ only. For simplicity, this paper considers causal policies. However, all subsequent results apply to strictly causal policies with minor modifications.

adjustable uncertainty set, since the uncertainty set $\mathcal{W}$ is a decision variable. Note that if $\mathcal{W}$ is given and fixed, then the above problem is a standard robust optimal control problem with additive uncertainty, see e.g. [16, 17] and the references therein.

Problem (4) is intractable because (i) the optimization of the uncertainty set $\mathcal{W}$ is performed over arbitrary subsets of $\mathbb{R}^{Nn_w}$; (ii) the constraints must be satisfied robustly for every uncertainty realization; and (iii) the optimization of the policies is performed over the space of causal functions $\mathcal{C}$. The main goal of this paper is to present approximations of instances of (4) that can be solved efficiently using tools from convex optimization. In Section 3, we focus on the optimization of the uncertainty sets without the presence of the control inputs, and discuss several families of uncertainty sets that can be handled efficiently. We then return to the original problem in Section 4 and show that by applying affine policy approximation, the resulting optimization problem is convex, and has a number of decision variables and constraints that grows polynomially in the problem parameters.

3. Uncertainty Set Approximation via Primitive Sets

We consider initially the following simplified problem

$$
\begin{aligned}
\max \quad & \varrho(\mathcal{W}) \\
\text{s.t.} \quad & \mathcal{W} \in \mathcal{P}(\mathbb{R}^{Nn_w}), \\
& \mathbf{Dw} \le \mathbf{d}, \quad \forall \mathbf{w} \in \mathcal{W},
\end{aligned}
\tag{5}
$$

where $\mathcal{W} = \mathbb{W}_0 \times \ldots \times \mathbb{W}_{N-1}$ is the decision variable. Problem (5) can be used, for example, to perform robustness analysis on a closed-loop or autonomous systems, where the control policy $\boldsymbol{\pi}(\cdot)$ in (4) is fixed. The objective would be to compute the largest set $\mathcal{W}$, according to metric $\varrho(\cdot)$, which the system can tolerate without violating its constraints. Note that Problem (5) becomes intractable when $\mathcal{W}$ is a polytope with a fixed number of vertices and $\varrho(\cdot)$ is the volume [18]. To gain computational tractability, we restrict $\mathcal{W}$ to be the affine transformation of a *primitive set*, and choose $\varrho(\cdot)$ from a pre-specified family of concave functions.

Definition 1 (Primitive Set)**.** Given a convex cone $\mathcal{K} \subseteq \mathbb{R}^l$, we call a compact set $\mathbb{S} \subset \mathbb{R}^{n_s}$, $n_s \ge n_w$, a primitive set if it has a non-empty relative interior and can be represented as

$$
\mathbb{S} := \{s \in \mathbb{R}^{n_s} : Gs \preceq_{\mathcal{K}} g\},
\tag{6}
$$

for some $G \in \mathbb{R}^{l \times n_s}$ and $g \in \mathbb{R}^l$. □

Note that *any* compact convex set is a primitive set, since every convex set admits a conic representation of the form (6) [19, p.15]. From now on, we *restrict* the uncertainty sets $\mathbb{W}_k$ to be of the form

$$
\mathbb{W}_k = Y_k \mathbb{S}_k + y_k := \{Y_k s + y_k : s \in \mathbb{S}_k\} \in \mathcal{P}(\mathbb{R}^{n_w}),
$$

$$
(Y_k, y_k) \in \mathbb{Y}_k \subseteq (\mathbb{R}^{n_w \times n_s} \times \mathbb{R}^{n_w}),
$$

where $\mathbb{S}_k$ is a primitive set. For simplicity, we assume that for all stages $k = 0, \dots, N-1$, $\mathbb{S}_k$ is described by the same cone $\mathcal{K} \subseteq \mathbb{R}^l$ and the same matrices G and g. However, all subsequent results can be easily extended to cases when the primitive sets are described by time-varying cones $\mathcal{K}_k \subset \mathbb{R}^{l_k}$ and matrices G_k and g_k. We refer to (Y_k, y_k) as *shaping parameters* and assume they take values in a convex set $\mathbb{Y}_k$. Note that if $n_w < n_s$, then $\mathbb{W}_k$ can be seen as the projection of $\mathbb{S}_k$ onto $\mathbb{R}^{n_w}$. Therefore, intuitively speaking, the matrix Y_k can be used to scale, rotate and project the set $\mathbb{S}_k$, while the vector y_k can be used to translate it. By defining $\mathcal{S} := \mathbb{S}_0 \times \dots \times \mathbb{S}_{N-1}$ and

$$\mathcal{Y} := \left\{ (\mathbf{Y}, \mathbf{y}) \in \mathbb{R}^{Nn_w \times Nn_s} \times \mathbb{R}^{Nn_w} : \exists \{(Y_k, y_k) \in \mathbb{Y}_k\}_{k=0}^{N-1}, \right.$$
$$\left. \mathbf{Y} = \mathrm{diag}(Y_0, \dots, Y_{N-1}), \; \mathbf{y} = [y_0, \dots, y_{N-1}] \right\},$$

we can compactly express the restriction on $\mathcal{W}$ as

$$\mathcal{W} = \mathbf{Y}\mathcal{S} + \mathbf{y} \in \mathcal{P}(\mathbb{R}^{Nn_w}), \quad (\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}. \tag{7}$$

Applying restriction (7) to problem (5) yields the following problem with decision variables $(\mathcal{W}, \mathbf{Y}, \mathbf{y})$:

$$\begin{aligned} \max \quad & \varrho(\mathcal{W}) \\ \text{s.t.} \quad & \mathcal{W} = \mathbf{Y}\mathcal{S} + \mathbf{y}, \; (\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}, \\ & \mathbf{D}\mathbf{w} \leq \mathbf{d}, \quad \forall \mathbf{w} \in \mathcal{W}. \end{aligned} \tag{8a}$$

Using standard duality arguments, e.g. [20, Theorem 3.1], to reformulate the semi-infinite constraint leads to a bilinear and hence non-convex optimization problem. To circumvent this difficulty, we eliminate the variables $\mathbf{w}$ and $\mathcal{W}$ in (8a) using (7). This leads to the following optimization problem with decision-independent uncertainty set $\mathcal{S}$:

$$\begin{aligned} \max \quad & \varrho(\mathbf{Y}\mathcal{S} + \mathbf{y}) \\ \text{s.t.} \quad & (\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}, \\ & \mathbf{D}(\mathbf{Y}\mathbf{s} + \mathbf{y}) \leq \mathbf{d}, \quad \forall \mathbf{s} \in \mathcal{S}, \end{aligned} \tag{8b}$$

with decision variables $(\mathbf{Y}, \mathbf{y})$. The following proposition shows that problems (8a) and (8b) are in fact equivalent.

Proposition 1. *Problems* (8a) *and* (8b) *are equivalent in the following sense: both problems have the same optimal value, and there exists a mapping between feasible solutions in both problems.*

Proof. The equivalence follows immediately from statement (i) in Lemma 1 in the Appendix, and by substituting $\mathcal{W}$ with $\mathbf{Y}\mathcal{S} + \mathbf{y}$. $\qquad \square$

Since $\mathcal{S}$ is convex, the semi-infinite constraint in (8b) can be reformulated using techniques based on strong duality of convex optimization, see e.g. [20, Theorem 3.1]. This leads to the following finite-

dimensional optimization problem:

$$\max \quad \varrho(\mathbf{Y}\mathcal{S} + \mathbf{y})$$
$$\text{s.t.} \quad (\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}, \ \mathbf{\Lambda} \in \mathbb{R}^{N(n_f+n_g)\times Nl}, \ \mathbf{\Lambda} \succeq_{\bar{\mathcal{K}}^\star} 0, \tag{9}$$
$$\mathbf{D}\mathbf{y} + \mathbf{\Lambda}\mathbf{g} \leq \mathbf{d}, \ \mathbf{\Lambda}\mathbf{G} = \mathbf{D}\mathbf{Y},$$

with decision variables $(\mathbf{Y}, \mathbf{y}, \mathbf{\Lambda})$, where $\mathbf{G} := \mathrm{diag}(G, \ldots, G) \in \mathbb{R}^{Nl \times Nn_s}$ $\mathbf{g} := [g, \ldots, g] \in \mathbb{R}^{Nl}$, $\bar{\mathcal{K}} := \mathcal{K} \times \ldots \times \mathcal{K} \subset \mathbb{R}^{Nl}$ and $\bar{\mathcal{K}}^\star$ is its dual cone. Problem (9) has convex constraints, and constitutes an inner approximation of the infinite-dimensional problem (5) due to restriction (7). Note that both the number of decision variables and the number of constraints in problem (9) grow polynomially with respect to the problem parameters $(N, n_x, n_s, n_w, n_f, n_g, l)$. If in addition the objective function $\varrho(\cdot)$ is concave in the shaping variables $(\mathbf{Y}, \mathbf{y})$, then problem (9) is a tractable convex optimization problem.

In the following, we discuss choices of the function $\varrho(\cdot)$ that will render problem (9) convex, and show that a large family of uncertainty sets of practical interest is captured by the affine transformation (7).

3.1. Volume Maximization

A natural objective in problem (5) is to maximize the volume of the sets $\mathbb{W}_k$. Indeed, for the special case of $n_s = n_w$, it is well-known that $\mathrm{vol}(\mathbb{W}_k) = \sqrt{\det(Y_k^\top Y_k)} \, \mathrm{vol}(\mathbb{S}_k)$ [21, Section 8.3.1], where $\mathrm{vol}(\mathbb{W}_k)$ and $\mathrm{vol}(\mathbb{S}_k)$ are the Lebesgue measures of the sets $\mathbb{W}_k$ and $\mathbb{S}_k$, respectively. If $\mathbb{Y}_k$ is chosen such that Y_k is constrained to be positive-definite, i.e. $\mathbb{Y}_k := \{(Y_k, y_k) \in (\mathbb{R}^{n_w \times n_w} \times \mathbb{R}^{n_w}) : Y_k \succ 0\}$, then maximizing $\varrho(Y_k\mathbb{S}_k + y_k) = \log\det(Y_k)$ maximizes $\mathrm{vol}(\mathbb{W}_k)$, see [21, Section 3.1.5] for more details[3]. Since $\log\det(\cdot)$ is a concave function, problem (9) is a convex optimization problem that can be solved efficiently [22]. For the case $n_s > n_w$, maximizing the volume of $\mathbb{W}_k$ is generally a non-convex optimization problem, and a convex, albeit heuristic, objective function should be chosen to keep the optimization problem tractable. Examples of such cost functions are given in Section 3.2 below.

3.2. Specific Families of Uncertainty Sets

In practical applications, uncertainty sets $\mathbb{W}_k$ are typically constrained to have a specific geometric form, such as ellipsoidal or rectangular. This requirement can be easily incorporated in problem (5) by including an additional constraint of the form $\mathbb{W}_k \in \Omega_k \subset \mathcal{P}(\mathbb{R}^{n_w})$, where Ω_k is the family of admissible uncertainty sets. In the following, we study families of uncertainty sets that often arise in robust control problems, focusing on norm balls, ellipsoids, hyper-rectangles and polytopes. We show that these representations can be obtained by appropriately selecting the primitive set $\mathbb{S}_k$ and constraint set $\mathbb{Y}_k$. To simplify notation, we omit the time indices k in the subsequent discussion.

---

[3]Similarly, the volume of $\mathcal{W}$ is maximized by choosing $\varrho(\mathbf{Y}\mathcal{S} + \mathbf{y}) = \log\det(\mathbf{Y})$, subject to $n_w = n_s$ and $Y_k \succ 0$.

Ball Uncertainty Set

p-norm ball uncertainty sets, with $p \in [1, \infty]$, take the form $\mathbb{W} = \{w \in \mathbb{R}^{n_w} : \|w - y\|_p \leq r\}$, with parameters $r \geq 0$ and $y \in \mathbb{R}^{n_w}$. In the spirit of [10], we choose the primitive set $\mathbb{S} = \{s \in \mathbb{R}^{n_w} : \|s\|_p \leq 1\}$, where $\mathcal{K}$ in (6) is the p-order cone, and restrict $\mathbb{Y} = \{(Y, y) \in (\mathbb{R}^{n_w \times n_w} \times \mathbb{R}^{n_w}) : \exists r \in \mathbb{R}_+, \, Y = rI\}$, where $I \in \mathbb{R}^{n_w \times n_w}$ is the identity matrix. This allows us to express p-norm ball uncertainty sets as $\mathbb{W} = Y\mathbb{S} + y$, subject to $(Y, y) \in \mathbb{Y}$. A natural choice for the objective function is $\varrho(\mathbb{W}) = r$, which is proportional to $\log \det(Y) = n_w r$ and thus maximizes the volume of $\mathbb{W}$. Since $\mathcal{K}^\star$ is the q-order cone, where $1/p + 1/q = 1$, and the set $\mathbb{Y}$ is described by linear constraints, problem (9) is a conic convex optimization problem that can be solved efficiently [23].

Ellipsoidal Uncertainty Set

Ellipsoidal uncertainty sets take the form $\mathbb{W} = \{w \in \mathbb{R}^{n_w} : (w - y)^\top \Sigma^{-1}(w - y) \leq 1\}$, with parameters $\Sigma = \Sigma^\top \succ 0$ and $y \in \mathbb{R}^{n_w}$. By choosing the primitive set $\mathbb{S} = \{s \in \mathbb{R}^{n_w} : \|s\|_2 \leq 1\}$, where $\mathcal{K}$ in (6) is the second-order cone, and setting $\mathbb{Y} = \{(Y, y) \in (\mathbb{R}^{n_w \times n_w} \times \mathbb{R}^{n_w}) : Y = Y^\top \succ 0\}$, we can express the ellipsoid as $\mathbb{W} = Y\mathbb{S} + y$ by identifying Y with $\Sigma^{1/2}$, see e.g. [21, Section 2.2.2]. The volume of $\mathbb{W}$ is maximized by choosing $\varrho(Y\mathbb{S} + y) = \log \det(Y)$. Since $\mathcal{K}^\star = \mathcal{K}$ and $\mathbb{Y}$ is described by linear matrix inequalities, problem (9) can be solved efficiently [22].

Axis-aligned Hyper-Rectangular Uncertainty Sets

Axis-aligned rectangular uncertainty sets take the general form $\mathbb{W} = \{w \in \mathbb{R}^{n_w} : -\gamma \leq w - y \leq \gamma\}$, with parameters $\gamma \in \mathbb{R}_+^{n_w}$ and $y \in \mathbb{R}^{n_w}$. By choosing $\mathbb{S} = \{s \in \mathbb{R}^{n_w} : \|s\|_\infty \leq 1\}$, where $\mathcal{K}$ in (6) is the ∞-order cone, and $\mathbb{Y} = \{(Y, y) \in (\mathbb{R}^{n_w \times n_w} \times \mathbb{R}^{n_w}) : \exists \gamma \in \mathbb{R}_+^{n_w}, \, Y = \text{diag}(\gamma)\}$, we can write $\mathbb{W} = Y\mathbb{S} + y$ subject to $(Y, y) \in \mathbb{Y}$. The volume of $\mathbb{W}$ is maximized with $\varrho(Y\mathbb{S} + y) = \log \det(Y) = \sum_{i=1}^{n_w} \log Y_{ii}$. For given $c_i \in \mathbb{R}$, $i \in \{1, \ldots, n_w\}$, another choice of the objective could be $\varrho(Y\mathbb{S} + y) = \sum_{i=1}^{n_w} c_i Y_{ii}$, which maximizes the (weighted) circumference of the rectangle $\mathbb{W}$. Since the dual cone $\mathcal{K}^\star$ is the first-order cone and $\mathbb{Y}$ is described by linear constraints, problem (9) is a convex optimization problem for either choice of the objective function.

Polyhedral Uncertainty Set

Compact polytopes with $m \geq n_w$ vertices take the general form $\mathbb{W} = \text{conv}(v^{(1)}, \ldots, v^{(m)})$, with parameters $(v^{(1)}, \ldots, v^{(m)})$. By choosing the primitive set as the simplex $\mathbb{S} = \text{conv}(e_1, \ldots, e_m)$, where $e_i \in \mathbb{R}^m$ is the i-th unit vector and $\mathcal{K} = \mathbb{R}_+^m$ in (6), and $\mathbb{Y} = \{(Y, y) \in (\mathbb{R}^{n_w \times m} \times \mathbb{R}^{n_w}) : y = 0\}$, the polytope can be expressed as $\mathbb{W} = Y\mathbb{S}$, where $Y_{\cdot j}$ corresponds to $v^{(j)}$ for all $j = 1, \ldots, m$. Unfortunately, maximizing the volume of a generic polytope with m vertices is computationally intractable [18]. Instead, one can, for example, use the linear objective function $\varrho(Y\mathbb{S}) = \sum_{j=1}^m c_j^\top Y_{\cdot j}$, which places the individual vertices of $\mathbb{W}$ in user-specified directions $c_j \in \mathbb{R}^{n_w}$, $j = 1, \ldots, m$ ("pushing"). With all constraints linear, the resulting

problem is a linear optimization problem. An alternative cost function is $\varrho(Y\mathbb{S}) = -\sum_{j=1}^{m} \|d_j - Y_j\|_2^2$, which places the vertices of $\mathbb{W}$ close to chosen points $d_j \in \mathbb{R}^{n_w}$, $j = 1, \ldots, m$ ("pulling"), resulting in a convex quadratic optimization problem.

4. Policy Approximation via Affine Decision Rules

To provide a tractable approximation for optimizing over policies, we first restrict the uncertainty set $\mathcal{W}$ to admit the affine representation (7), giving rise to the following infinite-dimensional problem:

$$
\begin{aligned}
\min \quad & \max_{\mathbf{w} \in \mathcal{W}} \left\{ \mathbf{c}^\top \boldsymbol{\pi}(\mathbf{w}) \right\} - \lambda \varrho(\mathcal{W}) \\
\text{s.t.} \quad & \boldsymbol{\pi}(\cdot) \in \mathcal{C}, \ \mathcal{W} = \mathbf{Y}\mathcal{S} + \mathbf{y}, \ (\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}, \\
& \mathbf{C}\boldsymbol{\pi}(\mathbf{w}) + \mathbf{D}\mathbf{w} \leq \mathbf{d}, \qquad \forall \mathbf{w} \in \mathcal{W},
\end{aligned}
\tag{10a}
$$

with decision variables $(\boldsymbol{\pi}(\cdot), \mathcal{W}, \mathbf{Y}, \mathbf{y})$. Similar to problem (8a), problem (10a) has a uncertainty set parametrized by the decision variables $(\mathbf{Y}, \mathbf{y})$. Therefore, we consider the following reformulation of problem (10a), parametrized in terms of $\mathbf{s}$, and with decision-independent uncertainty set:

$$
\begin{aligned}
\min \quad & \max_{\mathbf{s} \in \mathcal{S}} \left\{ \mathbf{c}^\top \widetilde{\boldsymbol{\pi}}(\mathbf{s}) \right\} - \lambda \varrho(\mathbf{Y}\mathcal{S} + \mathbf{y}) \\
\text{s.t.} \quad & \widetilde{\boldsymbol{\pi}}(\cdot) \in \widetilde{\mathcal{C}}, \ (\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}, \\
& \mathbf{C}\widetilde{\boldsymbol{\pi}}(\mathbf{s}) + \mathbf{D}(\mathbf{Y}\mathbf{s} + \mathbf{y}) \leq \mathbf{d}, \quad \forall \mathbf{s} \in \mathcal{S},
\end{aligned}
\tag{10b}
$$

with decision variables $(\widetilde{\boldsymbol{\pi}}(\cdot), \mathbf{Y}, \mathbf{y})$, and where $\widetilde{\mathcal{C}} := \{[\widetilde{\pi}_0(\cdot), \ldots, \widetilde{\pi}_{N-1}(\cdot)] : \widetilde{\pi}_k : \mathbb{S}_0 \times \cdots \times \mathbb{S}_k \to \mathbb{R}^{n_u}, \ k = 0, \ldots, N-1\}$. The set $\widetilde{\mathcal{C}}$ ensures that $\widetilde{\boldsymbol{\pi}}(\cdot)$ are causal policies, depending only on the first k elements of $\mathbf{s}$. The following proposition shows that problems (10a) and (10b) are equivalent.

Proposition 2. *Problems* (10a) *and* (10b) *are equivalent in the following sense: both problems have the same optimal value, and there exists a (not necessarily unique) mapping between feasible solutions in both problems.*

Proof. The proof is provided in the Appendix. $\qquad\square$

Proposition 2 allows us to focus on problem (10b), which belongs to the class of multistage adaptive optimization problems, involving a continuum of decision variables and inequality constraints [24]. Following [4], we restrict the space of admissible policies to exhibit the following affine structure

$$
\widetilde{\pi}_k(s_0, \ldots, s_k) := p_k + \sum_{j=0}^{k} P_{k,j} s_j, \quad k = 0, \ldots, N-1,
\tag{11}
$$

for some $P_{k,j} \in \mathbb{R}^{n_u \times n_s}$ and $p_k \in \mathbb{R}^{n_u}$. The concatenated policy is expressed by $\widetilde{\boldsymbol{\pi}}(\mathbf{s}) = \mathbf{P}\mathbf{s} + \mathbf{p}$, where $\mathbf{P} \in \mathbb{R}^{Nn_u \times Nn_s}$ is a lower-block-triangular matrix defined by the elements $P_{k,j}$ and $\mathbf{p} := [p_0, \ldots, p_{N-1}] \in \mathbb{R}^{Nn_u}$.

This restriction yields the following semi-infinite problem,

$$
\begin{aligned}
\min \quad & \max_{\mathbf{s} \in \mathcal{S}} \left\{ \mathbf{c}^\top (\mathbf{Ps} + \mathbf{p}) \right\} - \lambda \varrho(\mathbf{Y}\mathcal{S} + \mathbf{y}) \\
\text{s.t.} \quad & \mathbf{P} \in \mathbb{R}^{Nn_u \times Nn_s}, \ \mathbf{p} \in \mathbb{R}^{Nn_u}, \ (\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}, \\
& \mathbf{C}(\mathbf{Ps} + \mathbf{p}) + \mathbf{D}(\mathbf{Ys} + \mathbf{y}) \le \mathbf{d}, \quad \forall \mathbf{s} \in \mathcal{S},
\end{aligned}
\tag{12}
$$

with a finite number of decision variables $(\mathbf{P}, \mathbf{p}, \mathbf{Y}, \mathbf{y})$ and an infinite number of constraints. Since $\mathcal{S}$ is convex, one can employ the duality argument of [20, Theorem 3.1] to reformulate problem (12) into the following finite-dimensional convex optimization problem:

$$
\begin{aligned}
\min \quad & \tau - \lambda \varrho(\mathbf{Y}\mathcal{S} + \mathbf{y}) \\
\text{s.t.} \quad & \tau \in \mathbb{R}, \ \mathbf{P} \in \mathbb{R}^{Nn_u \times Nn_s}, \ \mathbf{p} \in \mathbb{R}^{Nn_u}, \ (\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}, \\
& \boldsymbol{\mu} \in \mathbb{R}^{Nl}, \ \boldsymbol{\mu} \succeq_{\bar{\mathcal{K}}^\star} 0, \ \boldsymbol{\Lambda} \in \mathbb{R}^{N(n_f + n_g) \times Nl}, \ \boldsymbol{\Lambda} \succeq_{\bar{\mathcal{K}}^\star} 0, \\
& \mathbf{c}^\top \mathbf{p} + \boldsymbol{\mu}^\top \mathbf{g} \le \tau, \ \mathbf{G}^\top \boldsymbol{\mu} = \mathbf{P}^\top \mathbf{c}, \\
& \mathbf{Cp} + \mathbf{Dy} + \boldsymbol{\Lambda}\mathbf{g} \le \mathbf{d}, \ \boldsymbol{\Lambda}\mathbf{G} = \mathbf{CP} + \mathbf{DY},
\end{aligned}
\tag{13}
$$

with decision variables $(\tau, \mathbf{P}, \mathbf{p}, \mathbf{Y}, \mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\Lambda})$. Note that the size of problem (13) grows polynomially in the parameters $(N, n_x, n_u, n_s, n_w, n_f, n_g, l)$. Furthermore, if $\varrho(\cdot)$ is chosen as discussed in Sections 3.1 and 3.2, then problem (13) is a convex optimization problem that can be solved efficiently; in particular, the affine policy does not alter the problem class.

The following theorem summarizes the relationship between the original problem (4), and its approximation (13), obtained through the restrictions (7) and (11):

Theorem 1. *Problem* (13) *constitutes an inner approximation of the original problem* (4) *in the following sense: every feasible solution of* (13) *is feasible in* (4), *and an upper bound to problem* (4) *can be found by solving problem* (13).

4.1. Relationship between the policies $\boldsymbol{\pi}(\cdot)$ and $\widetilde{\boldsymbol{\pi}}(\cdot)$

By restricting our attention to affine policies $\widetilde{\boldsymbol{\pi}}(\cdot)$ in (11), we were able to obtain a tractable optimization problem in (13). By Proposition 2, these affine policies are equivalent to some, possibly non-linear, policy $\boldsymbol{\pi}(\cdot) \in \mathcal{C}$. In the following, we show that if $\mathcal{S}$ is a polytope, then affine policies $\widetilde{\boldsymbol{\pi}}(\cdot)$ correspond to continuous piece-wise affine policies $\boldsymbol{\pi}(\cdot)$. To see this, let $\widetilde{\mathcal{C}}_{\text{aff}} \subset \widetilde{\mathcal{C}}$ be the class of causal affine policies in problem (10b) defined as in (11); $\mathcal{C}_{\text{aff}} \subset \mathcal{C}$ be the class of causal affine policies in problem (10a); and $\mathcal{C}_{\text{pwa}} \subset \mathcal{C}$ be the class of causal, piece-wise affine, continuous policies in problem (10a).

Proposition 3. *Let $\mathcal{S} := \mathbb{S}_0 \times \ldots \times \mathbb{S}_{N-1}$. Then, the following hold:*

(i) *For every $(\boldsymbol{\pi}(\cdot), \mathcal{W}, (\mathbf{Y}, \mathbf{y})) \in \mathcal{C}_{\text{aff}} \times \mathcal{P}(\mathbb{R}^{Nn_w}) \times \mathcal{Y}$ feasible in problem* (10a), *there exists a $(\widetilde{\boldsymbol{\pi}}(\cdot), (\mathbf{Y}, \mathbf{y})) \in \widetilde{\mathcal{C}}_{\text{aff}} \times \mathcal{Y}$ feasible in problem* (10b) *that achieves the same objective value.*

(ii) *If* **Y** *is invertible, then for every* $(\widetilde{\boldsymbol{\pi}}(\cdot), (\mathbf{Y}, \mathbf{y})) \in \tilde{\mathcal{C}}_{\mathrm{aff}} \times \mathcal{Y}$ *feasible in problem* (10b), *there exists a* $(\boldsymbol{\pi}(\cdot), \mathcal{W}, (\mathbf{Y}, \mathbf{y})) \in \mathcal{C}_{\mathrm{aff}} \times \mathcal{P}(\mathbb{R}^{Nn_w}) \times \mathcal{Y}$ *feasible in problem* (10a) *that achieves the same objective.*

(iii) *If* $\mathcal{S}$ *is a polytope, then for every* $(\widetilde{\boldsymbol{\pi}}(\cdot), (\mathbf{Y}, \mathbf{y})) \in \widetilde{\mathcal{C}}_{\mathrm{aff}} \times \mathcal{Y}$ *feasible in problem* (10b), *there exists a* $(\boldsymbol{\pi}(\cdot), \mathcal{W}, (\mathbf{Y}, \mathbf{y})) \in \mathcal{C}_{\mathrm{pwa}} \times \mathcal{P}(\mathbb{R}^{Nn_w}) \times \mathcal{Y}$ *feasible in problem* (10a) *that achieves the same objective value.*

Proof. The proof is provided in the Appendix. $\qquad\square$

From Proposition 3 (*i*) it follows that applying affine policies to problem (10b) produces at least as good results as applying affine policies to (10a). From part (*ii*) it follows that if **Y** is invertible (e.g. $\mathbf{Y} \succ 0$), then imposing $\widetilde{\boldsymbol{\pi}}(\cdot) \in \widetilde{\mathcal{C}}_{\mathrm{aff}}$ in problem (10b) is equivalent to imposing $\boldsymbol{\pi} \in \mathcal{C}_{\mathrm{aff}}$ in problem (10a), but has the advantage that a convex optimization problem is solved instead of a bilinear optimization problem. Finally, it follows from the proof of Proposition 3 that if $(\mathbf{P}^\star, \mathbf{p}^\star, \mathbf{Y}^\star, \mathbf{y}^\star)$ are the optimizers of problem (13), then a causal policy $\boldsymbol{\pi}(\cdot) \in \mathcal{C}$ can be obtained as

$$\boldsymbol{\pi}(\mathbf{w}) := \mathbf{P}^\star \left[(\mathbf{Y}^\star)^{-1}(\mathbf{w} - \mathbf{y}^\star) \right] + \mathbf{p}^\star \tag{14}$$

if $\mathbf{Y}^\star$ is invertible, and otherwise as

$$\boldsymbol{\pi}(\mathbf{w}) := \mathbf{P}^\star L(\mathbf{w}) + \mathbf{p}^\star, \tag{15}$$

where $L : \mathcal{W} \to \mathcal{S}$ is a so-called *lifting operator* (see Appendix). From a practical point of view, (14) and (15) also provide a way for computing the control input if a disturbance realization $\mathbf{w} \in \mathcal{W}$ is given or measured.

We close this section by pointing out that instead of approximating $\tilde{\boldsymbol{\pi}}(\cdot)$ with an affine-policy, it can be approximated using more complex, e.g. piece-wise affine, policies. Similar to the case of $\tilde{\boldsymbol{\pi}}(\cdot) \in \tilde{\mathcal{C}}_{\mathrm{aff}}$, it can be shown that if $\tilde{\boldsymbol{\pi}}(\cdot) \in \tilde{\mathcal{C}}_{\mathrm{pwa}}$ and $\mathcal{S}$ is a polytope, then $\boldsymbol{\pi}(\cdot)$ will be piece-wise affine. The interested reader is referred to [25, 26] for tractable methods of designing piece-wise affine policies $\tilde{\boldsymbol{\pi}}(\cdot)$ using convex optimization techniques.

5. Example 1: Frequency Reserve Provision

In a power grid, the supply and demand of electricity must be balanced at all times to maintain frequency stability. This balance is achieved by the grid operator by procuring so-called frequency reserves that are dispatched whenever supply does not meet demand [27]. Recently, there has been increasing interest in so-called demand response schemes, where flexible loads such as office buildings are used to provide such reserves [28]. Buildings participating in the demand response program are asked in real-time by the grid operator to adjust their electricity consumption within a given range (the so-called *reserve capacity*). They are rewarded financially for this service through a payment that typically depends both on the size of the

offered reserve capacity and on the actual reserves provided. In the following, we investigate the amount of frequency reserves commercial buildings can offer to the power grid.

5.1. Problem Formulation

Following [29, 30], we assume that the building dynamics are of the form $x_{k+1} = Ax_k + B(u_k + \Delta u_k) + Ev_k$, where the states represent the temperatures in the room, walls, floor, and ceiling. The affine term $v_k \in \mathbb{R}^{n_v}$ represents uncontrolled inputs acting on the building, such as ambient temperature, solar radiation, and internal gains (number of building occupants, etc.). For simplicity, we assume that v_k can be predicted perfectly, and refer to [29, 31] for methods of estimating these. The control inputs consist of the building's heating, ventilation and air condition (HVAC) system, which we model through two input vectors u_k and Δu_k. We interpret the (strictly causal) input u_k as the nominal control input which we apply if no reserve is demanded at time k. The (causal) input Δu_k models the necessary input change at time k to follow the actual *reserve demand* $w_k \in \mathbb{W}_k$, where $\mathbb{W}_k \subseteq \mathbb{R}$ is the reserve capacity. Note that Δu_k is required to depend on w_k since it is used in real-time to follow the reserve demand w_k, which is a regulation signal sent by the grid operator. Following the literature [30, 31], we model the electricity consumption of the HVAC system through a linear function of the form $u_k \mapsto \eta^\top u_k$, so that $e_k := \eta^\top u_k$ is the nominal electricity consumption, and $\Delta e_k := \eta^\top \Delta u_k$ is the change in electricity consumption if the input is changed from u_k to $u_k + \Delta u_k$. In practice, Δu_k is chosen such that it matches the reserve demand, satisfying $w_k = \eta^\top \Delta u_k$ for all $k = 0, \ldots, N-1$.

The objective function is composed of two competing quatities: the maximization of reserve capacity, and the minimization of electricity cost. The electricity cost takes the form $\sum_{k=0}^{N-1} c_k \eta^\top u_k$, where $c_k \in \mathbb{R}_+$ is the electricity price at time k. Following the current norm in Switzerland [32], we consider symmetric reserve capacity $\mathbb{W}_k = [-Y_k, Y_k]$, where $Y_k \in \mathbb{R}_+$. We assume that offering reserve capacity $\mathbb{W}_k$ of size Y_k is compensated with a reward of the form λY_k. For simplicity, the reward associated with the actual reserves provided $\mathbf{w} := [w_0, \ldots, w_{N-1}]$ is not considered. Since the reserve demand $\mathbf{w}$ is uncertain at the time of planning, we introduce the policies $\boldsymbol{\pi}(\cdot) := [\pi_0(\cdot), \ldots, \pi_{N-1}(\cdot)] \in \mathcal{C}_s$ and $\boldsymbol{\kappa}(\cdot) := [\kappa_0(\cdot), \ldots, \kappa_{N-1}(\cdot)] \in \mathcal{C}$ for $[u_0, \ldots, u_{N-1}]$ and $[\Delta u_0, \ldots, \Delta u_{N-1}]$, respectively, where $\mathcal{C}_s$ is the class of all strictly causal policies. Note that the inputs $u_k = \pi_k(w_0, \ldots, w_{k-1})$ are strictly causal because in practice they are decided and applied independent of the current realization w_k. In contrast, the inputs $\Delta u_k = \kappa_k(w_0, \ldots, w_k)$ are only causal since they must satisfy the (instantaneous) reserve request $w_k = \eta^\top \Delta u_k$ for all $w_k \in \mathbb{W}_k$.

The frequency reserve provision problem can now be formulated as the following robust optimal control

problem with adjustable uncertainty set:

$$
\begin{aligned}
\min \quad & \max_{\mathbf{w} \in \mathcal{W}} \left\{ \mathbf{c}^\top \mathbf{e}(\mathbf{w}) \right\} - \lambda \, \varrho(\mathcal{W}) \\
\text{s.t.} \quad & \boldsymbol{\pi}(\cdot) \in \mathcal{C}_s, \ \boldsymbol{\kappa}(\cdot) \in \mathcal{C}, \ \mathcal{W} = \times_{k=0}^{N-1} [-Y_k, Y_k], \ \mathbf{Y} \geq 0 \\
& \left. \begin{aligned}
& \mathbf{x}(\mathbf{w}) = \mathbf{A} x_0 + \mathbf{B}(\boldsymbol{\pi}(\mathbf{w}) + \boldsymbol{\kappa}(\mathbf{w})) + \mathbf{E} \mathbf{v}, \\
& \mathbf{F} \mathbf{x}(\mathbf{w}) \leq \mathbf{f}, \ \mathbf{G}(\boldsymbol{\pi}(\mathbf{w}) + \boldsymbol{\kappa}(\mathbf{w})) \leq \mathbf{g}, \\
& \mathbf{e}(\mathbf{w}) = \boldsymbol{\eta}^\top \boldsymbol{\pi}(\mathbf{w}), \ \mathbf{w} = \boldsymbol{\eta}^\top \boldsymbol{\kappa}(\mathbf{w}),
\end{aligned} \right\} \ \forall \mathbf{w} \in \mathcal{W},
\end{aligned}
\tag{16}
$$

with decision variables $(\boldsymbol{\pi}(\cdot), \boldsymbol{\kappa}(\cdot), \mathcal{W}, \mathbf{Y})$, and $\mathbf{c} := [c_0, \ldots, c_{N-1}]$, $\mathbf{Y}$ is the collection of all $\{Y_k\}_{k=0}^{N-1}$, $\varrho(\mathcal{W}) := \sum_{k=0}^{N-1} Y_k$, $\mathbf{v} := [v_0, \ldots, v_{N-1}]$, $x_0 \in \mathbb{R}^{n_x}$ is the initial state, and $\boldsymbol{\eta}^\top := \mathrm{diag}(\eta^\top, \ldots, \eta^\top)$. The matrices $(\mathbf{A}, \mathbf{B}, \mathbf{E}, \mathbf{F}, \mathbf{f}, \mathbf{G}, \mathbf{g})$ can easily be constructed from the problem data. The inequality constraint $\mathbf{F} \mathbf{x}(\mathbf{w}) \leq \mathbf{f}$, ensures that comfort constraints are satisfied at all times. Similarly, the inequality constraint $\mathbf{G}(\boldsymbol{\pi}(\mathbf{w}) + \boldsymbol{\kappa}(\mathbf{w})) \leq \mathbf{g}$, ensures that the constraints on the HVAC systems are always satisfied. The first equality constraint describes the state dynamics, while constraint $\mathbf{e}(\mathbf{w}) = \boldsymbol{\eta}^\top \boldsymbol{\pi}(\mathbf{w})$ models the (nominal) power consumption. Finally, the constraint $\mathbf{w} = \boldsymbol{\eta}^\top \boldsymbol{\kappa}(\mathbf{w})$ ensures that for every reserve demand sequence $\mathbf{w}$, there exists a feasible input change $\boldsymbol{\kappa}(\mathbf{w})$ that matches the reserve demand.

5.2. Simulation Results

We consider a one-room building model with one-hour discretization and horizon length $N = 24$, based on the reduced building model presented in [29, Section 4.5]. The states are $x_k = [x_{k,1}, x_{k,2}, x_{k,3}] \in \mathbb{R}^3$, where $x_{k,1}$ is the room temperature, $x_{k,2}$ the inside wall temperature, and $x_{k,3}$ the outside wall temperature. We consider a deterministic sequence for the uncontrolled inputs $v_k = [v_{k,1}, v_{k,2}, v_{k,3}] \in \mathbb{R}^3$, taken from [31]. In particular, $v_{k,1}$ represents the outside air temperature, $v_{k,2}$ the solar radiation, and $v_{k,3}$ the number of occupants in the building. We consider four inputs $u_k = [u_{k,1}, u_{k,2}, u_{k,3}, u_{k,4}] \in \mathbb{R}^4$, representing actuators commonly found in Swiss office buildings: a radiator heater, a cooled ceiling, a floor heating system, and mechanical ventilation [29, 30, 31]. Comfort constraints are given as $21°C \leq x_{k,1} \leq 25°C$ for all $k = 0, \ldots, N-1$, while the inputs are upper- and lower-bounded[4]. Moreover, we consider two electricity prices $\mathbf{c}^{\mathrm{wd}}$ and $\mathbf{c}^{\mathrm{we}}$, where the former exhibits a typical weekday price pattern and the latter a typical weekend price pattern. The prices are depicted in Figure 1 (top).

Problem (16) is an instance of (4), and can be (inner) approximated based on the ideas discussed in Sections 3 and 4. First, following Section 3.2, we express $\mathcal{W} = \mathbf{Y} \mathcal{S}$ with $\mathbf{Y} := \mathrm{diag}(Y_0, \ldots, Y_{N-1})$ and $\mathcal{S} := \{ \mathbf{s} \in \mathbb{R}^N : -\mathbf{1} \leq \mathbf{s} \leq \mathbf{1} \}$, where $\mathbf{1} \in \mathbb{R}^N$ is the all-one vector. Second, we restrict the input policies to admit an affine structure with respect to the primitive set $\mathcal{S}$ as in (11).

---

[4]Often, no temperature constraints are imposed in office buildings during nights. Although not considered here, this feature can easily be incorporated in our problem setup.
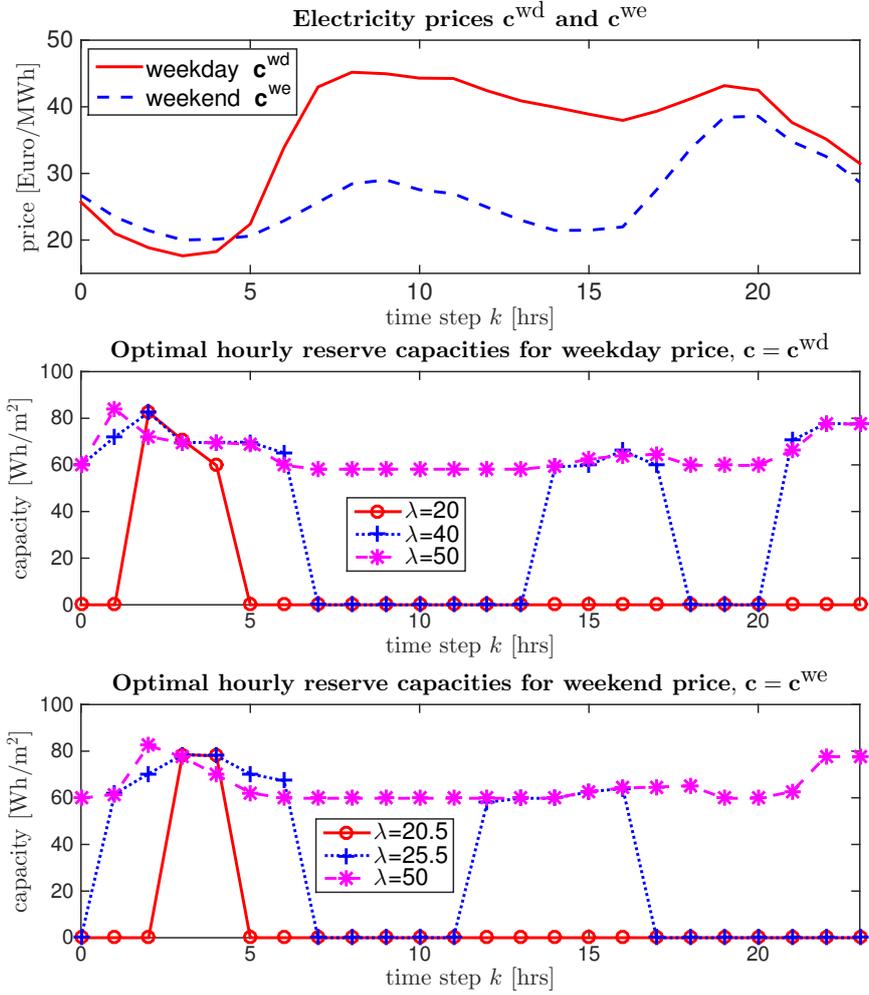
Figure 1: *(Top)* Hourly day-ahead spot market prices for weekday (solid red line) and weekend (dashed blue line). The weekday prices are averaged over the workdays 14th–18th September 2015, and the weekend prices averaged over 19th–20th September 2015. Data were obtained from `www.epexspot.com`. *(Center)* Optimal reserve capacities $Y_k^\star$ for weekday prices and reward factors $\lambda \in \{20, 40, 50\}$. *(Bottom)* Optimal reserve capacities $Y_k^\star$ for weekend prices and reward factors $\lambda \in \{20.5, 25.5, 50\}$.

In our first numerical study, we examine the effect of electricity price and reward factor on the reserve capacity provided by the building. In particular, we solve problem (16) using the weekday prices with $\lambda \in \{20, 40, 50\}$, and using the weekend prices with $\lambda \in \{20.5, 25.5, 50\}$. The results are reported in Figure 1 (center and bottom), which show the optimal reserve capacities $[Y_0^\star, \ldots, Y_{N-1}^\star]$ as a function of time. As we can see from the figures, for a given reward factor λ, the building provides reserve capacities only at times during which the reward factor is higher than the electricity price, i.e. whenever $\lambda > c_k$. This behavior can be explained as follows: For a building to provide a certain reserve capacity Y_k, it must increase its nominal
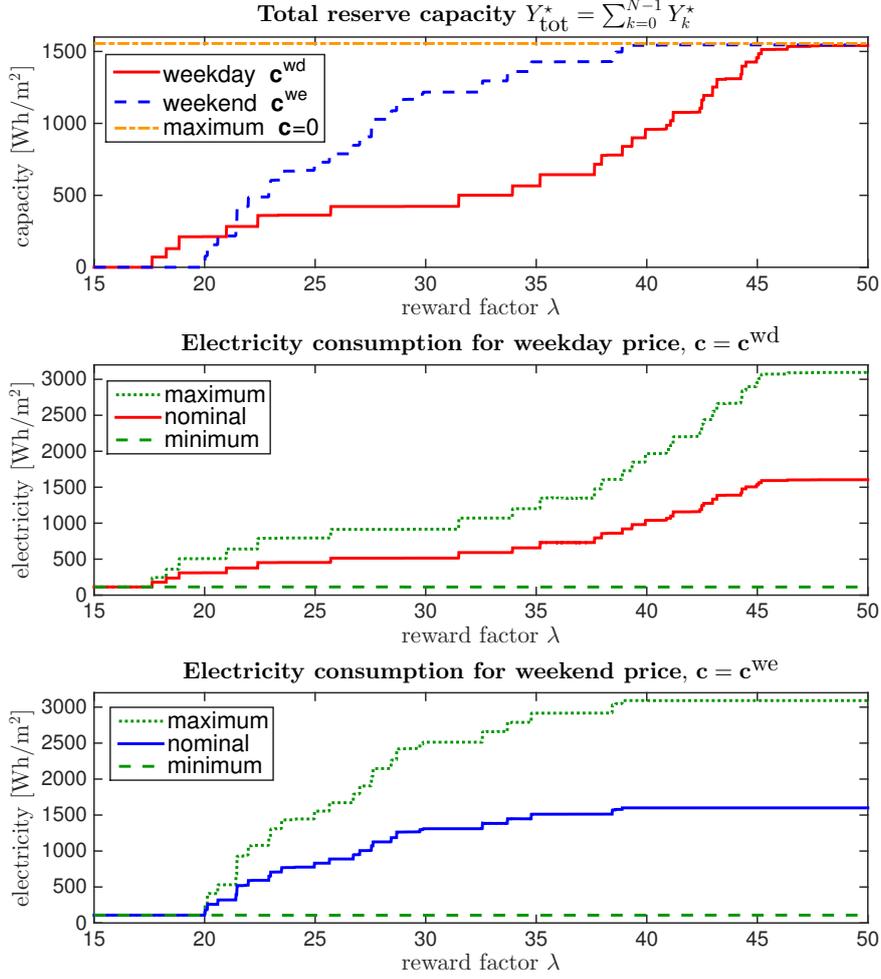
14

Figure 2: *(Top)* Total reserve capacity $Y^\star = \sum_{k=0}^{N-1} Y_k^\star$ for weekday (solid red line) and weekend (dashed blue line). The orange dash-dotted line is the maximum possible offered reserve capacity, obtained by setting $\mathbf{c} = 0$. *(Center)* Nominal electricity consumption (solid red line), minimal electricity consumption (dashed green line), and maximum electricity consumption (dotted green line) for weekday price. *(Bottom)* Same as center, but with weekend price.

electricity consumption by at least Y_k, or else it would not be able to reduce its electricity consumption by Y_k. However, since an increase of electricity consumption will incur an additional cost of $c_k Y_k$, the building will only provide reserve capacity of size Y_k if the reward for doing so outweighs the additional electricity cost, i.e. whenever $\lambda Y_k > c_k Y_k$. Otherwise, if $\lambda Y_k < c_k Y_k$, the building will choose not to provide reserves, resulting in the jumps in the reserve capacity profile shown in Figure 1 (center and bottom). If non-symmetric reserve capacities were allowed, then the building would provide (positive) reserves $\mathbb{W}_k = [0, \bar{Y}_k]$ when $\lambda < c_k$. During these times, the building is able to increase its electricity consumption by $w_k = \eta^\top \Delta u_k$, without the

15

need to increase its nominal electricity consumption $e_k = \eta^\top u_k$, since $w_k \geq 0$. Finally, we point out that, due to the dynamics of the buildings and the comfort constraints, the level of reserves provided can change as λ takes on different values. For example, we see from Figure 1 that the reserve profile with $\lambda = 50$ is not always higher than that with $\lambda = 40$ (center figure) or $\lambda = 25.5$ (bottom figure), even if the total reserve capacity turns out to be higher, as we see later on.

In our second numerical study, we study the effect of the reward factor λ on the total reserve capacity $Y_{\text{tot}}^\star := \sum_{k=1}^{N-1} Y_k^\star$ offered by the system. To this end, we solve problem (16) with weekday and weekend prices for reward factors $\lambda \in [15, 50]$. The results are depicted in Figure 2 (top), which compares the total reserves for $\mathbf{c} = \mathbf{c}^{\text{wd}}$ (solid red line), $\mathbf{c} = \mathbf{c}^{\text{we}}$ (dashed blue line), and the maximum possible capacity obtained when $\mathbf{c} = \mathbf{0}$ (dash-dotted line). Figure 2 (top) allows us to make the following two observations: First, for most values of λ, the building provides more reserves during the weekend than during weekdays. This is intuitive because the electricity price during the weekend is typically lower than during weekday, compare with Figure 1 (top). Second, when λ is small, more reserves are provided during weekdays, which can be explained by the fact that the minimum electricity price is lower on weekdays than during the weekend, incentivizing the building to provide reserves even for small values of λ. From a practical point of view, Figure 2 (top) can be used to determine the optimal level of reserves that should be offered for a given reward λ. Alternatively, in case of conditional bidding (i.e. the amount of reserve bid in the auction is a function of the reward), Figure 2 (top) can be used to determine the optimal tradeoff between the amount of reserves and the reward factor.

In our final numerical study, we investigate the effect of the reward factor λ on the building's electricity consumption. To this end, we define the nominal, maximal and minimal electricity consumption as $\eta^\top (\pi^\star(\mathbf{0}) + \kappa^\star(\mathbf{0}))$, $\max_{\mathbf{w} \in \mathcal{W}^\star} \{\eta^\top (\pi^\star(\mathbf{w}) + \kappa^\star(\mathbf{w}))\}$, and $\min_{\mathbf{w} \in \mathcal{W}^\star} \{\eta^\top (\pi^\star(\mathbf{w}) + \kappa^\star(\mathbf{w}))\}$, respectively, where $\pi^\star(\cdot)$ and $\kappa^\star(\cdot)$ are the optimal affine policies computed from problem (16), and $\mathcal{W}^\star$ is the associated optimal reserve capacity. The results are depicted in Figure 2 (center and bottom). It can be seen that as the building provides more reserves for increasing λ, it must also increase its nominal electricity consumption. Indeed, as can be verified, the nominal electricity consumption increases in the exact same fashion as the total reserve capacity depicted in Figure 2 (top). Similar to what we have seen before, such behavior is to be expected since to provide a certain level of reserves, the building must increase its nominal electricity consumption by the same amount. Moreover, the interval between the maximal and minimal electricity consumption can be interpreted as the "slack" of the building, within which the building can adjust its electricity consumption without violating its constraints.

We finish this section with three concluding remarks: (i) Our framework also allows for non-symmetric reserve capacities. For example, non-symmetric reserve capacities of the form $[-\underline{Y}_k, \overline{Y}_k]$ can be obtained by replacing $\mathcal{W} = \mathbf{Y}\mathcal{S}$ in (16) with $\mathcal{W} = \mathbf{Y}\mathcal{S} + \mathbf{y}$ and $-\mathbf{Y} \leq \text{diag}(\mathbf{y}) \leq \mathbf{Y}$. Similarly, positive reserve capacities are obtained by choosing the primitive set $\mathcal{S}$ in (16) as $\mathcal{S} = [0, 1]^N$, or as $\mathcal{S} = [-1, 0]^N$ if negative reserves

are desired. (*ii*) It is well-known that restricting the input policies to be affine can be conservative, and that measuring this loss of optimality exactly is as challenging as determining the optimal policy. Nevertheless, for small horizons, the solution to problem (16) can be computed exactly by enumerating all corner points of $\mathcal{W}$, and then solving the corresponding optimization problem [33]. Unfortunately, this approach is only applicable if N is small, since the problem size grows exponentially in N [33]. Nevertheless, if $N \leq 8$ in (16), by using this enumeration approach, we verified that affine policies are indeed optimal for all $\lambda \in [15, 50]$. For longer horizons $9 \leq N \leq 24$, a bound on the loss of optimality can be obtained based on the method introduced in [34]. Using this approach, we have verified that the relative gap between the exact solution and the affine policies is consistently below 0.71%, for all $\lambda \in [15, 50]$. This demonstrates that the affine policy approximation achieves a high degree of precision for this example, giving us hope that affine policies are indeed optimal in this case. (*iii*) We highlight that under the affine policy restriction, problem (16) with $N = 24$ can be written as a linear optimization problem with roughly $13\,000$ decision variables and $18\,000$ constraints. For fixed λ, each problem was solved within 0.3 seconds using the optimization solver Mosek on a standard laptop equipped with 16 GB RAM and a 2.6 GHz quad-core Intel i7 processor, demonstrating that our proposed solution method is indeed computationally tractable.

6. Example 2: Robustness Analysis

We next apply our framework to a robustness analysis problem, where the goal is to determine the largest uncertainty set a system can tolerate without violating its constraints. We consider a system of the form (1) with parameters $n_x = 2$, $n_u = 1$, $n_w = 2$, and system matrices $A = I$, $B = [1, 0.7]$ and $E = -I$, where $I \in \mathbb{R}^{2 \times 2}$ is the identity matrix. To simplify the subsequent analysis, we set $N = 1$, $x_0 = \bar{x} = 0$, and $u_0 = \pi_0(w_0) \in \mathcal{C}$. Hence, the robustness analysis problem can be written as a robust optimal control problem with adjustable uncertainty set:

$$
\begin{aligned}
\max \quad & \mathrm{vol}(\mathbb{W}) \\
\text{s.t.} \quad & \pi(\cdot) \in \mathcal{C},\ \mathbb{W} \in \mathcal{P}(\mathbb{R}^2),\ \ \bar{x} = 0, \\
& \left.\begin{aligned} & x(w) = A\bar{x} + B\pi(w) - w, \\ & x(w) \in \mathcal{X},\ \pi(w) \in \mathcal{U}, \end{aligned}\right\} \forall w \in \mathbb{W},
\end{aligned}
\tag{17}
$$

with optimization variables $(\mathbb{W}, \pi(\cdot))$, where, for simplicity, we have omitted the time indices in our one-step problem. The state and input constraints are defined as $\mathcal{X} := \{[x_1, x_2] \in \mathbb{R}^2 : [-10, -10] \leq [x_1, x_2] \leq [10, 10],\ -15 \leq -x_1 + x_2 \leq 15,\ -15 \leq x_1 + x_2 \leq 15\}$ and $\mathcal{U} := \{u \in \mathbb{R} : -5 \leq u \leq 5\}$. Problem (17) is an instance of (4), and can thus be dealt with by first expressing the uncertainty set as $\mathbb{W} = Y\mathbb{S} + y$, where $\mathbb{S}$ is a primitive set, and then restricting the input policy as $\tilde{\pi}(\cdot) \in \tilde{\mathcal{C}}_{\mathrm{aff}}$. Finally, a suitable concave objective function $\varrho(\cdot)$ is chosen that approximates the objective function $\mathrm{vol}(\cdot)$.
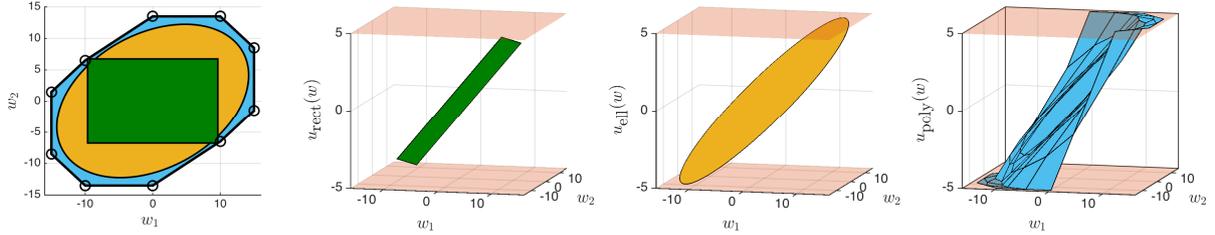
Figure 3: *(Left)* The optimal uncertainty sets $\mathbb{W}^\star_{\mathrm{rect}}$ (green rectangle), $\mathbb{W}^\star_{\mathrm{ell}}$ (yellow ellipse) and $\mathbb{W}^\star_{\mathrm{poly}}$ (blue polytope). *(Center left)* Affine policy $\pi_{\mathrm{rect}}(\cdot)$ for rectangular uncertainty set recovered from $\tilde{\pi}^\star(\cdot) \in \tilde{\mathcal{C}}_{\mathrm{aff}}$. *(Center right)* Affine policy $\pi_{\mathrm{ell}}(\cdot)$ for ellipsoidal uncertainty set recovered from $\tilde{\pi}^\star(\cdot) \in \tilde{\mathcal{C}}_{\mathrm{aff}}$. *(Right)* Piecewise affine policy $\pi_{\mathrm{poly}}(\cdot)$ for polytopic uncertainty set recovered from $\tilde{\pi}^\star(\cdot) \in \tilde{\mathcal{C}}_{\mathrm{aff}}$.

In our first study, we examine how the choice of the uncertainty set family affects the solution quality. The following three families of uncertainty sets, all discussed in Section 3.2, are considered: (i) rectangular uncertainty sets $\mathbb{W}_{\mathrm{rect}} \in \Omega_{\mathrm{rect}}$; (ii) ellipsoidal uncertainty sets $\mathbb{W}_{\mathrm{ell}} \in \Omega_{\mathrm{ell}}$; and (iii) polytopic uncertainty sets $\mathbb{W}_{\mathrm{poly}} \in \Omega_{\mathrm{poly}}$ with 30 vertices. Following Section 3.1, we use $\varrho_{\mathrm{rect}}(Y\mathbb{S}+y) = \varrho_{\mathrm{ell}}(Y\mathbb{S}+y) = \log\det(Y)$ to maximize $\mathrm{vol}(\mathbb{W}_{\mathrm{rect}})$ and $\mathrm{vol}(\mathbb{W}_{\mathrm{ell}})$, respectively. Since we cannot directly maximize the volume of the polytope $\mathbb{W}_{\mathrm{poly}}$, we use the quadratic cost function $\varrho_{\mathrm{poly}}(Y\mathbb{S}) = -\sum_{j=1}^{n_s} \|d_j^\top - Y_{\cdot j}\|_2^2$ described in Section 3.2 to approximate $\mathrm{vol}(\mathbb{W}_{\mathrm{poly}})$, where $n_s = 30$, and d_j are uniformly distributed on a centered circle of radius forty, i.e. $d_j := [40\cos(2\pi(j-1)/n_s),\ 40\sin(2\pi(j-1)/n_s)]$. Figure 3 (left) depicts the optimal uncertainty sets $\mathbb{W}^\star_{\mathrm{rect}}$, $\mathbb{W}^\star_{\mathrm{ell}}$, and $\mathbb{W}^\star_{\mathrm{poly}}$, while Table 1 lists their volumes. For this example, we see that $\mathrm{vol}(\mathbb{W}^\star_{\mathrm{rect}})$ is much smaller than $\mathrm{vol}(\mathbb{W}^\star_{\mathrm{ell}})$ and $\mathrm{vol}(\mathbb{W}^\star_{\mathrm{poly}})$, because Ω_{rect} restricts the sets to be axis-aligned, whereas both Ω_{ell} and Ω_{poly} allow for rotated and skewed sets. Moreover, it is interesting to observe that even though $\varrho_{\mathrm{poly}}(\cdot)$ does not maximize the volume of $\mathbb{W}_{\mathrm{poly}}$ explicitely, it yet yields the largest uncertainty. For this example, it can be verified by taking all combinations of the vertices of the sets $\mathcal{X}$ and $\mathcal{U}$ that the polytopic approximation method returns the optimal solution to problem (17), i.e. $\mathbb{W}^\star = \mathbb{W}^\star_{\mathrm{poly}}$. Observe now from Figure 3 (left) that while $\mathbb{W}^\star_{\mathrm{rect}}$ corresponds to the largest volume rectangle contained in $\mathbb{W}^\star$, $\mathbb{W}^\star_{\mathrm{ell}}$ is *not*

Table 1: Optimal value of different approximation schemes.

Set Family	vol($\mathbb{W}$)	Approx. Quality
(i) $\mathbb{W}^\star_{\mathrm{rect}} \in \Omega_{\mathrm{rect}}$	260.4	42.0%
(ii) $\mathbb{W}^\star_{\mathrm{ell}} \in \Omega_{\mathrm{ell}}$	514.4	83.0%
(iii) $\mathbb{W}^\star_{\mathrm{poly}} \in \Omega_{\mathrm{poly}}$	620.2	100%

the maximum volume ellipse contained in $\mathbb{W}_{\mathrm{poly}}$. As we will see in the following paragraph, this is due to the requirement that $\tilde{\pi}_{\mathrm{ell}}(\cdot)$ is affine.

In the second part of this study, we examine the structure of the policy $\pi(\cdot)$ for the different uncertainty sets. Figure 3 (center left – right) depicts the policies $\pi(\cdot)$ for all three cases, where we recovered $\pi_{\mathrm{rect}}(\cdot)$ and $\pi_{\mathrm{ell}}(\cdot)$ using (14), and $\pi_{\mathrm{poly}}(\cdot)$ using (15) with $L(\cdot)$ in (A.1). We see from Figure 3 (center right) that the reason $\mathbb{W}_{\mathrm{ell}}^{\star}$ is not the maximum volume ellipse is due to the restriction $\pi_{\mathrm{ell}}(\cdot) \in \mathcal{C}_{\mathrm{aff}}$. Indeed, we observe that even for $\mathbb{W}_{\mathrm{ell}} = \mathbb{W}_{\mathrm{ell}}^{\star}$, the associated policy $\pi_{\mathrm{ell}}(\cdot)$ hits the input constraints, and that if $\mathbb{W}_{\mathrm{ell}}$ were expanded in the $45°$-direction, then the input constraints would be violated. This drawback, in principle, can be alleviated by allowing $\pi_{\mathrm{ell}}(\cdot)$ to be non-linear, using e.g. the techniques in [25, 26]. Note that this issue is not present in the polytopic approximation, since the policy $\pi_{\mathrm{poly}}(\cdot)$ is piece-wise affine by construction, c.f. Proposition 3 (iii). Indeed, a careful inspection of $\pi_{\mathrm{poly}}(\cdot)$ reveals that it is the piece-wise affine nature that allows $\mathbb{W}_{\mathrm{poly}}^{\star}$ to obtain its complex shape with vertices that are not captured by either $\mathbb{W}_{\mathrm{rect}}^{\star}$ or $\mathbb{W}_{\mathrm{poly}}^{\star}$, and which gives rise to the uncertainty set with biggest volume.

7. Conclusion

This paper presents a unified framework for studying robust optimal control problems with adjustable uncertainty sets. We have shown that these problems can be approximated as a convex optimization problem if (i) the uncertainty sets are representable as affine transformations of primitive sets, and (ii) the control policies are restricted to be affine with respect to the primitive sets. Under these restrictions, the resulting convex optimization problem grows polynomially in the size of the problem data and the description of the primitive set, resulting in a tractable optimization problem. The applicability and importance of such problems was demonstrated through a reserve capacity problem arising in power systems. A numerical example confirms that these problems can be solved efficiently, and that optimal bidding and reserve strategies can be derived.

Future work will focus on reserve provision problems, addressing stability and recursive feasibility when the reserve capacities are bid in a receding horizon fashion. Moreover, we plan on investigating the optimality of affine policies in the reserve provision context, as well as the impact external uncertainties have on the provided reserves. Finally, the presented framework could be extended to cases when only partial state observations are available.

Appendix A. Definition, Lemma and Proofs

In this section, we provide the proofs for Propositions 2 and 3. To do this, we first need to introduce the *lifting operator*, and an auxiliary result given in Lemma 1. We will use similar theoretical tools as those developed in [25, 26].

Definition 2 (Lifting operator). Given $(\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}$ and $\mathcal{S} = \mathbb{S}_0 \times \ldots \times \mathbb{S}_{N-1}$, where each $\mathbb{S}_k$ is a primitive set, let $\mathcal{W} := \mathbf{Y}\mathcal{S} + \mathbf{y}$. We call a mapping $L : \mathcal{W} \to \mathcal{S}$ a *lifting operator* if it satisfies the following properties:

(P1) $L(\cdot) = [L_0(\cdot), \ldots, L_{N-1}(\cdot)]$, where, for all $k = 0, \ldots, N-1$, $L_k : \mathbb{W}_k \to \mathbb{S}_k$;

(P2) $L_k(w) \in \mathbb{S}_k$, for all $w \in \mathbb{W}_k$;

(P3) $\mathbf{Y}L(\mathbf{w}) + \mathbf{y} = \mathbf{w}$, for all $\mathbf{w} \in \mathcal{W}$.

We refer to $L(\cdot)$ as the lifting operator since typically $\dim(\mathcal{S}) \geq \dim(\mathcal{W})$, i.e. $n_s \geq n_w$. Properties (P1) and (P2) ensure that $L(\cdot)$ preserves causality since at time k, since $[L_0(\cdot), \ldots, L_k(\cdot)]$ contains information up to and including stage k only. Notice that, although $L(\cdot)$ is defined through $(\mathbf{Y}, \mathbf{y})$ and $\mathcal{S}$, there can be multiple choices of $L(\cdot)$ that satisfy (P1)–(P3). The following example demonstrates the existence of such a lifting for given $(\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}$, $\mathcal{S}$, and $\mathcal{W} := \mathbf{Y}\mathcal{S} + \mathbf{y}$, while highlighting some of its properties.

Example 1. *Given $(\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}$, $\mathcal{S} = \mathbb{S}_0 \times \ldots \times \mathbb{S}_{N-1}$ and $\mathcal{W} := \mathbf{Y}\mathcal{S} + \mathbf{y}$, let us define $L(\cdot) := [L_0(\cdot), \ldots, L_{N-1}(\cdot)]$ through its components*

$$L_k(w) := \arg\min_{z \in \mathbb{S}_k}\{\|z\|_2^2 : Y_k z + y_k = w\}, \tag{A.1}$$

for all $w \in \mathbb{W}_k$. It is easy to verify that this lifting operator satisfies conditions (P1)–(P3). Moreover, if $n_s = n_w$ and $\mathbf{Y}$ is a full rank matrix, then $L(\cdot)$ is the bijective operator $L(\mathbf{w}) = \mathbf{Y}^{-1}(\mathbf{w} - \mathbf{y})$. However, if $\mathbf{Y}$ is rank deficient, then the mapping from $\mathcal{W}$ to $\mathcal{S}$ is not unique, and therefore, operator (A.1) chooses the mapping that minimizes the 2-norm in the $\mathcal{S}$ space. Finally, note that the above optimization problem defining $L(\cdot)$ is always feasible by construction, and that alternative liftings can be defined using different (also non-linear) objective functions.

The following lemma establishes the relation between constraint functions in the $\mathcal{W}$-space and in the $\mathcal{S}$-space.

Lemma 1. *Given $(\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}$, $\mathcal{S} = \mathbb{S}_0 \times \ldots \times \mathbb{S}_{N-1}$, where each $\mathbb{S}_k$ is a primitive set, let $\mathcal{W} := \mathbf{Y}\mathcal{S} + \mathbf{y}$ and $L(\cdot)$ be a lifting operator. For any two functions $g : \mathbb{R}^{Nn_w} \to \mathbb{R}$ and $f : \mathbb{R}^{Nn_s} \to \mathbb{R}$, it holds:*

(i) $g(\mathbf{w}) \leq 0, \ \forall \mathbf{w} \in \mathcal{W} \iff g(\mathbf{Y}\mathbf{s} + \mathbf{y}) \leq 0, \ \forall \mathbf{s} \in \mathcal{S}$;

(ii) $f(\mathbf{s}) \leq 0, \ \forall \mathbf{s} \in \mathcal{S} \implies f(L(\mathbf{w})) \leq 0, \ \forall \mathbf{w} \in \mathcal{W}$.

Proof. We prove " $\Longrightarrow$ " of (i) by contradiction. Assume that $g(\mathbf{w}) \leq 0$ for all $\mathbf{w} \in \mathcal{W}$, but there exists $\mathbf{s} \in \mathcal{S}$ such that $g(\mathbf{Ys} + \mathbf{y}) > 0$. Since $\mathbf{Ys} + \mathbf{y} \in \mathcal{W}$ for all $\mathbf{s} \in \mathcal{S}$ by definition of $\mathcal{W}$, this leads to a contradiction. The " $\Longleftarrow$ " part of (i) is clear, since $\mathcal{W} = \mathbf{Y}\mathcal{S} + \mathbf{y}$ by definition. Statement (ii) follows immediately from property (P2) in the definition of the lifting operator. $\qquad\square$

We now have the necessary tools to prove Propositions 2 and 3.

Proof of Proposition 2. To simplify notation, we first rewrite problem (10a) using the epigraph formulation

$$
\begin{aligned}
\min \quad & \tau - \lambda \varrho(\mathcal{W}) \\
\text{s.t.} \quad & \tau \in \mathbb{R}, \ \boldsymbol{\pi}(\cdot) \in \mathcal{C}, \ \mathcal{W} = \mathbf{Y}\mathcal{S} + \mathbf{y}, \ (\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}, \\
& \left.\begin{aligned} \mathbf{c}^\top \boldsymbol{\pi}(\mathbf{w}) &\leq \tau, \\ \mathbf{C}\boldsymbol{\pi}(\mathbf{w}) + \mathbf{Dw} &\leq \mathbf{d}, \end{aligned}\right\} \forall \mathbf{w} \in \mathcal{W},
\end{aligned}
\tag{A.2}
$$

with decision variables $(\tau, \mathcal{W}, \boldsymbol{\pi}(\cdot), \mathbf{Y}, \mathbf{y})$, and problem (10b) as

$$
\begin{aligned}
\min \quad & \tau - \lambda \varrho(\mathbf{Y}\mathcal{S} + \mathbf{y}) \\
\text{s.t.} \quad & \tau \in \mathbb{R}, \ \widetilde{\boldsymbol{\pi}}(\cdot) \in \widetilde{\mathcal{C}}, \ (\mathbf{Y}, \mathbf{y}) \in \mathcal{Y}, \\
& \left.\begin{aligned} \mathbf{c}^\top \widetilde{\boldsymbol{\pi}}(\mathbf{s}) &\leq \tau, \\ \mathbf{C}\widetilde{\boldsymbol{\pi}}(\mathbf{s}) + \mathbf{D}(\mathbf{Ys} + \mathbf{y}) &\leq \mathbf{d}, \end{aligned}\right\} \forall \mathbf{s} \in \mathcal{S},
\end{aligned}
\tag{A.3}
$$

with decision variables $(\tau, \widetilde{\boldsymbol{\pi}}(\cdot), \mathbf{Y}, \mathbf{y})$.

We now show that any feasible solution in (A.2) corresponds to a feasible solution in (A.3) with the same objective value, and vice versa. Suppose that $(\tau, \mathcal{W}, \boldsymbol{\pi}(\cdot), \mathbf{Y}, \mathbf{y})$ is a feasible solution to (A.2). We define $\widetilde{\boldsymbol{\pi}}(\cdot)$ such that $\widetilde{\boldsymbol{\pi}}(\mathbf{s}) := \boldsymbol{\pi}(\mathbf{Ys} + \mathbf{y})$, for all $\mathbf{s} \in \mathcal{S}$. Note that due to the block-diagonal form of $\mathbf{Y}$, the constructed policy $\widetilde{\boldsymbol{\pi}}(\cdot)$ is causal, satisfying the constraint $\widetilde{\boldsymbol{\pi}}(\cdot) \in \widetilde{\mathcal{C}}$. Since $(\tau, \mathcal{W}, \boldsymbol{\pi}(\cdot), \mathbf{Y}, \mathbf{y})$ is feasible in (A.2), it implies that

$$
\begin{aligned}
& \left.\begin{aligned} \mathbf{C}\boldsymbol{\pi}(\mathbf{w}) + \mathbf{Dw} &\leq \mathbf{d}, \\ \mathbf{c}^\top \boldsymbol{\pi}(\mathbf{w}) &\leq \tau, \end{aligned}\right\} \forall \mathbf{w} \in \mathcal{W}, \\
\Longleftrightarrow \quad & \left.\begin{aligned} \mathbf{C}\boldsymbol{\pi}(\mathbf{Ys} + \mathbf{y}) + \mathbf{D}(\mathbf{Ys} + \mathbf{y}) &\leq \mathbf{d}, \\ \mathbf{c}^\top \boldsymbol{\pi}(\mathbf{Ys} + \mathbf{y}) &\leq \tau, \end{aligned}\right\} \forall \mathbf{s} \in \mathcal{S}, \\
\Longleftrightarrow \quad & \left.\begin{aligned} \mathbf{C}\widetilde{\boldsymbol{\pi}}(\mathbf{s}) + \mathbf{D}(\mathbf{Ys} + \mathbf{y}) &\leq \mathbf{d}, \\ \mathbf{c}^\top \widetilde{\boldsymbol{\pi}}(\mathbf{s}) &\leq \tau, \end{aligned}\right\} \forall \mathbf{s} \in \mathcal{S},
\end{aligned}
$$

where the first equivalence is due to statement (i) in Lemma 1, and the second equivalence by definition of $\widetilde{\boldsymbol{\pi}}(\cdot)$. Hence, the quadruple $(\tau, \widetilde{\boldsymbol{\pi}}(\cdot), \mathbf{Y}, \mathbf{y})$ is feasible in (A.3). Moreover, since $\mathcal{W} = \mathbf{Y}\mathcal{S} + \mathbf{y}$, problems (A.2) and (A.3) share the same objective function. Hence, both problems achieve the same objective value.

Conversely, suppose that $(\tau, \widetilde{\boldsymbol{\pi}}(\cdot), \mathbf{Y}, \mathbf{y})$ is feasible in (A.3), and let $\mathcal{W} := \mathbf{Y}\mathcal{S} + \mathbf{y}$. We now define $\boldsymbol{\pi}(\cdot)$ through $\boldsymbol{\pi}(\mathbf{w}) := \widetilde{\boldsymbol{\pi}}(L(\mathbf{w}))$, for all $\mathbf{w} \in \mathcal{W}$, where $L(\cdot)$ is a lifting operator satisfying Definition 2, for example

the one given in (A.1). By construction of $L(\cdot)$, $\boldsymbol{\pi}(\cdot)$ remains causal, satisfying $\boldsymbol{\pi}(\cdot) \in \mathcal{C}$. Since $\widetilde{\boldsymbol{\pi}}(\cdot)$ is feasible in (A.3), we have that

$$
\left.\begin{array}{l}
\mathbf{C}\widetilde{\boldsymbol{\pi}}(\mathbf{s}) + \mathbf{D}(\mathbf{Ys} + \mathbf{y}) \le \mathbf{d}, \\
\mathbf{c}^\top \widetilde{\boldsymbol{\pi}}(\mathbf{s}) \le \tau,
\end{array}\right\} \forall \mathbf{s} \in \mathcal{S},
$$

$$
\implies \left.\begin{array}{l}
\mathbf{C}\widetilde{\boldsymbol{\pi}}(L(\mathbf{w})) + \mathbf{D}(\mathbf{Y}L(\mathbf{w}) + \mathbf{y}) \le \mathbf{d}, \\
\mathbf{c}^\top \widetilde{\boldsymbol{\pi}}(L(\mathbf{w})) \le \tau,
\end{array}\right\} \forall \mathbf{w} \in \mathcal{W},
$$

$$
\iff \left.\begin{array}{l}
\mathbf{C}\boldsymbol{\pi}(\mathbf{w}) + \mathbf{D}\mathbf{w} \le \mathbf{d}, \\
\mathbf{c}^\top \boldsymbol{\pi}(\mathbf{w}) \le \tau,
\end{array}\right\} \forall \mathbf{w} \in \mathcal{W},
$$

where the second line follows from statement (ii) of Lemma 1, and the third line follows from property (P2) in the definition of the lifting operator and the definition of $\boldsymbol{\pi}(\cdot)$. Hence, $(\tau, \mathcal{W}, \boldsymbol{\pi}(\cdot), \mathbf{Y}, \mathbf{y})$ is feasible in (A.2). Since both problems share the same objective function due to $\mathcal{W} = \mathbf{Y}\mathcal{S} + \mathbf{y}$, both problems achieve the same objective value. $\qquad\square$

Proof of Proposition 3. To prove (i), let $(\tau, \mathcal{W}, \boldsymbol{\pi}(\cdot), \mathbf{Y}, \mathbf{y})$ be a feasible solution to (A.2) with $\boldsymbol{\pi}(\cdot) \in \mathcal{C}_{\mathrm{aff}}$. Since $\boldsymbol{\pi}(\cdot) \in \mathcal{C}_{\mathrm{aff}}$, it can be written as $\boldsymbol{\pi}(\mathbf{w}) = \mathbf{Q}\mathbf{w} + \mathbf{q}$ for some $\mathbf{Q} \in \mathbb{R}^{Nn_u \times Nn_w}$ lower block-triangular, and $\mathbf{q} \in \mathbb{R}^{Nn_u}$. Now, let us consider the candidate solution $(\tau, \widetilde{\boldsymbol{\pi}}(\cdot), \mathbf{Y}, \mathbf{y})$ for (A.3), with $\widetilde{\boldsymbol{\pi}}(\mathbf{s}) := (\mathbf{Q}\mathbf{Y})\mathbf{s} + (\mathbf{Q}\mathbf{y} + \mathbf{q})$. Since $\mathbf{Y}$ is block-diagonal, the matrix $(\mathbf{Q}\mathbf{Y})$ is also lower block-triangular, thus ensuring the causality of $\widetilde{\boldsymbol{\pi}}(\cdot)$. Therefore, $\widetilde{\boldsymbol{\pi}}(\mathbf{s}) := (\mathbf{Q}\mathbf{Y})\mathbf{s} + (\mathbf{Q}\mathbf{y} + \mathbf{q}) \in \mathcal{C}_{\mathrm{aff}}$. Moreover, since $\boldsymbol{\pi}(\cdot)$ is feasible in (A.2), we have that

$$
\left.\begin{array}{l}
\mathbf{C}[\mathbf{Q}\mathbf{w} + \mathbf{q}] + \mathbf{D}\mathbf{w} \le \mathbf{d}, \\
\mathbf{c}^\top [\mathbf{Q}\mathbf{w} + \mathbf{q}] \le \tau,
\end{array}\right\} \forall \mathbf{w} \in \mathcal{W},
$$

$$
\iff \left.\begin{array}{l}
\mathbf{C}[\mathbf{Q}(\mathbf{Ys} + \mathbf{y}) + \mathbf{q}] + \mathbf{D}(\mathbf{Ys} + \mathbf{y}) \le \mathbf{d}, \\
\mathbf{c}^\top [\mathbf{Q}(\mathbf{Ys} + \mathbf{y}) + \mathbf{q}] \le \tau,
\end{array}\right\} \forall \mathbf{s} \in \mathcal{S},
$$

$$
\iff \left.\begin{array}{l}
\mathbf{C}\widetilde{\boldsymbol{\pi}}(\mathbf{s}) + \mathbf{D}(\mathbf{Ys} + \mathbf{y}) \le \mathbf{d}, \\
\mathbf{c}^\top \widetilde{\boldsymbol{\pi}}(\mathbf{s}) \le \tau,
\end{array}\right\} \forall \mathbf{s} \in \mathcal{S},
$$

where the first equivalence is due to statement (i) in Lemma 1, and the second equivalence follows from the definition of $\widetilde{\boldsymbol{\pi}}(\cdot)$. Hence, $(\tau, \widetilde{\boldsymbol{\pi}}(\cdot), \mathbf{Y}, \mathbf{y})$ is a feasible solution for (A.3), achieving the same objective value.

To prove (ii), let $(\tau, \widetilde{\boldsymbol{\pi}}(\cdot), \mathbf{Y}, \mathbf{y})$ be a feasible solution to (A.3) with $\widetilde{\boldsymbol{\pi}}(\cdot) \in \mathcal{C}_{\mathrm{aff}}$ and $\mathbf{Y}$ invertible, and define $\mathcal{W} := \mathbf{Y}\mathcal{S} + \mathbf{y}$. Since $\widetilde{\boldsymbol{\pi}}(\cdot) \in \mathcal{C}_{\mathrm{aff}}$, it can be written as $\widetilde{\boldsymbol{\pi}}(\mathbf{s}) = \mathbf{P}\mathbf{s} + \mathbf{p}$ for some $\mathbf{P} \in \mathbb{R}^{Nn_u \times Nn_s}$ lower block-triangular, and $\mathbf{p} \in \mathbb{R}^{Nn_u}$. Consider now the candidate solution $(\tau, \mathcal{W}, \boldsymbol{\pi}(\cdot), \mathbf{Y}, \mathbf{y})$ for (A.2), with $\boldsymbol{\pi}(\mathbf{w}) := \mathbf{P}\left[\mathbf{Y}^{-1}(\mathbf{w} - \mathbf{y})\right] + \mathbf{p}$, for all $\mathbf{w} \in \mathcal{W}$. It is easy to see that $\boldsymbol{\pi}(\cdot)$ is an affine policy. Moreover, since $\mathbf{Y}$ is block-diagonal by construction, see (7), $\boldsymbol{\pi}(\cdot)$ is also causal, hence satisfying $\boldsymbol{\pi}(\cdot) \in \mathcal{C}_{\mathrm{aff}}$. Finally, since

$\widetilde{\boldsymbol{\pi}}(\cdot)$ is feasible in (A.3), we have that

$$
\left.\begin{aligned}
&\mathbf{C}[\mathbf{Ps} + \mathbf{p}] + \mathbf{D}(\mathbf{Ys} + \mathbf{y}) \leq \mathbf{d}, \\
&\mathbf{c}^\top[\mathbf{Ps} + \mathbf{p}] \leq \tau,
\end{aligned}\right\} \forall \mathbf{s} \in \mathcal{S},
$$

$$
\iff \left.\begin{aligned}
&\mathbf{C}[\mathbf{PY}^{-1}(\mathbf{w} - \mathbf{y}) + \mathbf{p}] + \mathbf{Dw} \leq \mathbf{d}, \\
&\mathbf{c}^\top[\mathbf{PY}^{-1}(\mathbf{w} - \mathbf{y}) + \mathbf{p}] \leq \tau,
\end{aligned}\right\} \forall \mathbf{w} \in \mathcal{W},
$$

$$
\iff \left.\begin{aligned}
&\mathbf{C}\boldsymbol{\pi}(\mathbf{w}) + \mathbf{Dw} \leq \mathbf{d}, \\
&\mathbf{c}^\top \boldsymbol{\pi}(\mathbf{w}) \leq \tau,
\end{aligned}\right\} \forall \mathbf{w} \in \mathcal{W},
$$

where the first equivalence follows from statement (i) of Lemma 1 and the fact that $\mathbf{w} = \mathbf{Ys} + \mathbf{y}$ is equal to $\mathbf{s} = \mathbf{Y}^{-1}(\mathbf{w} - \mathbf{y})$, and the second equivalence follows from the definition of $\boldsymbol{\pi}(\cdot)$. Hence, $(\tau, \mathcal{W}, \boldsymbol{\pi}(\cdot), \mathbf{Y}, \mathbf{y})$ is a feasible solution for (A.2), and achieves the same objective value.

To prove (iii), let $(\tau, \widetilde{\boldsymbol{\pi}}(\cdot), \mathbf{Y}, \mathbf{y})$ be a feasible solution to (A.3) with $\widetilde{\boldsymbol{\pi}}(\cdot) \in \mathcal{C}_{\mathrm{aff}}$, and define $\mathcal{W} := \mathbf{Y}\mathcal{S} + \mathbf{y}$. Since $\widetilde{\boldsymbol{\pi}}(\cdot) \in \mathcal{C}_{\mathrm{aff}}$, it can be written as $\widetilde{\boldsymbol{\pi}}(\mathbf{s}) = \mathbf{Ps} + \mathbf{p}$ for some $\mathbf{P} \in \mathbb{R}^{Nn_u \times Nn_s}$ lower block-triangular, and $\mathbf{p} \in \mathbb{R}^{Nn_u}$. Let $L(\cdot) = [L_0(\cdot), \dots, L_{N-1}(\cdot)]$ be the lifting operator with its elements $L_k(\cdot)$ defined as in (A.1), and consider the candidate solution $(\tau, \mathcal{W}, \boldsymbol{\pi}(\cdot), \mathbf{Y}, \mathbf{y})$ for (A.2), with $\boldsymbol{\pi}(\mathbf{w}) := \mathbf{P}L(\mathbf{w}) + \mathbf{p}$, for all $\mathbf{w} \in \mathcal{W}$. Under the assumption of $\mathcal{S}$ being a polytope, $L(\cdot)$ is a continuous piece-wise affine function defined through the solution of a parametric quadratic optimization problem [35, Theorem 7.7]. Moreover, since $L(\cdot)$ satisfies properties (P1) and (P2) in Definition 2, it preserves causality, and $\boldsymbol{\pi}(\mathbf{w}) := \mathbf{P}L(\mathbf{w}) + \mathbf{p} \in \mathcal{C}_{\mathrm{pwa}}$ is a causal policy. Finally, since $\widetilde{\boldsymbol{\pi}}(\cdot)$ is feasible in (A.3), we have that

$$
\left.\begin{aligned}
&\mathbf{C}(\mathbf{Ps} + \mathbf{p}) + \mathbf{D}(\mathbf{Ys} + \mathbf{y}) \leq \mathbf{d}, \\
&\mathbf{c}^\top(\mathbf{Ps} + \mathbf{p}) \leq \tau,
\end{aligned}\right\} \forall \mathbf{s} \in \mathcal{S},
$$

$$
\implies \left.\begin{aligned}
&\mathbf{C}(\mathbf{P}L(\mathbf{w}) + \mathbf{p}) + \mathbf{D}(\mathbf{Y}L(\mathbf{w}) + \mathbf{y}) \leq \mathbf{d}, \\
&\mathbf{c}^\top(\mathbf{P}L(\mathbf{w}) + \mathbf{p}) \leq \tau,
\end{aligned}\right\} \forall \mathbf{w} \in \mathcal{W},
$$

$$
\iff \left.\begin{aligned}
&\mathbf{C}(\mathbf{P}L(\mathbf{w}) + \mathbf{p}) + \mathbf{Dw} \leq \mathbf{d}, \\
&\mathbf{c}^\top(\mathbf{P}L(\mathbf{w}) + \mathbf{p}) \leq \tau,
\end{aligned}\right\} \forall \mathbf{w} \in \mathcal{W},
$$

where the second line follows from statement (ii) of Lemma 1, and the third line follows from property (P3). Hence, $(\tau, \mathcal{W}, \boldsymbol{\pi}(\cdot), \mathbf{Y}, \mathbf{y})$ is a feasible solution for (A.2), and achieves the same objective value. $\qquad \square$

References

[1] A. Bemporad and M. Morari. Robust model predictive control: A survey. In *Robustness in identification and control*, volume 245 of *Lecture Notes in Control and Information Sciences*, pages 207–226. Springer London, 1999.

[2] J. Löfberg. Approximations of closed-loop minimax MPC. In *IEEE Conference on Decision and Control*, Dec 2003.

[3] J. Skaf and S. P. Boyd. Design of affine controllers via convex optimization. *IEEE Transactions on Automatic Control*, 55(11):2476–2487, 2010.

[4] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.

[5] D. Bertsimas and A. Thiele. A robust optimization approach to inventory theory. *Operations Research*, 54(1):150–168, 2006.

[6] K. Postek and D. den Hertog. Multi-stage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. Available online at: `http://www.optimization-online.org/DB_FILE/2014/10/4575.pdf`, 2014.

[7] D. Q. Mayne, J.B. Rawlings, C.V. Rao, and P.O.M. Scokaert. Constrained model predictive control: stability and optimality. *Automatica*, 36:789–814, 2000.

[8] C. Wang, C.-J. Ong, and M. Sim. Model Predictive Control Using Segregated Disturbance Feedback. *IEEE Transactions on Automatic Control*, 55(4):831–840, April 2010.

[9] E. F. Camacho and C. Bordons. *Model Predictive Control*. Springer Verlag, 2013.

[10] X. Zhang, M. Kamgarpour, P.J. Goulart, and J. Lygeros. Selling Robustness Margins: A Framework for Optimizing Reserve Capacities for Uncertain Systems. In *IEEE Conference on Decision and Control*, Los Angeles, California, USA, 2014.

[11] E. Vrettos, F. Oldewurtel, F. Zhu, and G. Andersson. Robust provision of frequency reserves by office building aggregations. In *IFAC World Congress*, Cape Town, South Africa, 2014.

[12] T. Gorecki, A. Bitlislioglu, G. Stathopoulos, and C. Jones. Guaranteeing Input Tracking for Constrained Systems: Theory and Application to Demand Response. In *American Control Conference*, 2015.

[13] E. Vrettos, F. Oldewurtel, and G. Andersson. Robust energy-constrained frequency reserves from aggregations of commercial buildings. *IEEE Transactions on Power Systems*, 2016.

[14] W. Wu, J. Chen, B. Zhang, and H. Su. A Robust Wind Power Optimization Method for Look-Ahead Power Dispatch. *IEEE Transactions on Sustainable Energy*, 5(2):507–515, April 2014.

[15] X. Zhang, E. Vrettos, M. Kamgarpour, G. Andersson, and J. Lygeros. Stochastic frequency reserve provision by chance-constrained control of commercial buildings. In *European Control Conference*, Linz, Austria, 2015.

[16] P. Goulart, E. Kerrigan, and J. Maciejowski. Optimization over state feedback policies for robust control with constraints. *Automatica*, 42(4):523–533, 2006.

[17] J. Löfberg. *Minimax approaches to robust model predictive control*. PhD thesis, Linköping University, 2003.

[18] M. E. Dyer and A. M. Frieze. On the complexity of computing the volume of a polyhedron. *SIAM Journal on Computing*, 17(5):967–974, 1988.

[19] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[20] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25(1):1–13, 1999.

[21] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[22] L. Vandenberghe, S. Boyd, and S-P. Wu. Determinant maximization with linear matrix inequality constraints. *SIAM Journal on Matrix Analysis and Applications*, 19(2):499–533, 1998.

[23] G. Xue and Y. Ye. An efficient algorithm for minimizing a sum of p-norms. *SIAM Journal on Optimization*, 10(2):551–579, 2000.

[24] D. Bertsimas and V. Goyal. On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical Programming*, 134(2):491–531, 2012.

[25] A. Georghiou, W. Wiesemann, and D. Kuhn. Generalized decision rule approximations for stochastic programming via liftings. *Mathematical Programming*, 152(1-2):301–338, 2015.

[26] D. Bertsimas and A. Georghiou. Binary decision rules in multistage adaptive mixed-integer optimization. *Mathematical Programming (conditionally accepted)*, 2016.

[27] Y. Rebours, D. Kirschen, M. Trotignon, and S. Rossignol. A survey of frequency and voltage control ancillary services – Part I: Technical features. *IEEE Transactions on Power Systems*, 22(1), 2007.

[28] D. Callaway and I. Hiskens. Achieving controllability of electric loads. *Proceedings of the IEEE*, 99(1):184–199, 2011.

[29] F. Oldewurtel. *Stochastic Model Predictive Control for Energy Efficient Building Climate Control*. PhD thesis, ETH Zurich, August 2011.

[30] F. Oldewurtel, A. Parisio, C. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and M. Morari. Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, 45:15–27, 2012.

[31] X. Zhang, G. Schildbach, D. Sturzenegger, and M. Morari. Scenario-based MPC for Energy-Efficient Building Climate Control under Weather and Occupancy Uncertainty. In *European Control Conference*, Zurich, Switzerland, 2013.

[32] Swissgrid Ltd. Basic principles of ancillary service products. Available online at: `https://www.swissgrid.ch/dam/swissgrid/experts/ancillary_services/Dokumente/D151019_AS-Products_V9R1_en.pdf`, 2015.

[33] P. O. M. Scokaert and D. Q. Mayne. Min-max feedback model predictive control for constrained linear systems. *IEEE Transactions on Automatic Control*, 43(8):1136–1142, 1998.

[34] D. Bertsimas and A. Georghiou. Design of Near Optimal Decision Rules in Multistage Adaptive Mixed-Integer Optimization. *Operations Research*, 63(3):610–627, 2015.

[35] F. Borrelli, A. Bemprad, and M. Morari. *Predictive Control for linear and hybrid systems*. Available online at: `http://www.mpc.berkeley.edu/mpc-course-material`.