

A Tractable Approach for designing Piecewise Affine Policies in Dynamic Robust Optimization

Aharon Ben-Tal · Omar El Housni · Vineet Goyal

Received: date / Accepted: date

Abstract We consider the problem of designing piecewise affine policies for two-stage adjustable robust linear optimization problems under right hand side uncertainty. It is well known that a piecewise affine policy is optimal although the number of pieces can be exponentially large. A significant challenge in designing a practical piecewise affine policy is constructing good pieces of the uncertainty set. Here we address this challenge by introducing a new framework in which the uncertainty set is “approximated” by a “dominating” simplex. The corresponding policy is then based on the map from the uncertainty set to the simplex. Although our piecewise affine policy has exponentially many pieces, it can be computed efficiently by solving a compact linear program. Furthermore, the performance of our policy is significantly better than the affine policy for many important uncertainty sets both theoretically and numerically. For instance, for hypersphere uncertainty set, our piecewise affine policy can be computed by an LP and gives a $O(m^{1/4})$ -approximation whereas the affine policy requires us to solve a second order cone program and has a worst-case performance bound of $O(\sqrt{m})$. To the best of our knowledge, this is the first tractable approach for designing piecewise affine policies with significantly improved theoretical performance guarantees.

Aharon Ben-Tal
Industrial Engineering and Management, Technion - Israel Institute of Technology and
CentER, Tilburg University, Netherlands
E-mail: abental@ie.technion.ac.il

Omar El Housni, Vineet Goyal
Industrial Engineering and Operations Research, Columbia University
E-mail: {oe2148, vg2277}@columbia.edu

1 Introduction

Addressing uncertainty in problem parameters in an optimization problem is a fundamental challenge in most real world problems where decisions often need to be made in the face of uncertainty. Stochastic and robust optimization are two approaches that have been studied extensively to handle uncertainty. In a stochastic optimization framework, uncertainty is modeled using a probability distribution and the goal is to optimize an expected objective [15]. We refer the reader to Kall and Wallace [21], Prekopa [22], Shapiro [23], Shapiro et al. [24] for a detailed discussion on stochastic optimization. While it is a reasonable approach in certain settings, it is intractable in general and suffers from the “curse of dimensionality”. Moreover, in many applications, we may not have sufficient historical data to estimate a joint probability distribution over the uncertain parameters.

Robust optimization is another paradigm where we consider an adversarial model of uncertainty using an uncertainty set and the goal is to optimize over the worst-case realization from the uncertainty set. This approach was first introduced by Soyster [25] and has been extensively studied in recent past. We refer the reader to Ben-Tal and Nemirovski [3–5], El Ghaoui and Lebret [16], Bertsimas and Sim [13,14], Goldfarb and Iyengar [19], Bertsimas et al. [7] and Ben-Tal et al. [1] for a detailed discussion of robust optimization. Robust optimization leads to a tractable approach where an optimal static solution can be computed efficiently for a large class of problems. Moreover, designing an uncertainty set is significantly less challenging than estimating a joint probability distribution for high-dimensional uncertainty. However, computing an optimal adjustable (or dynamic) solution for a multi-stage problem is generally hard even in the robust optimization framework.

In this paper, we consider two-stage adjustable robust linear optimization problems with covering constraints and uncertain right-hand-side. In particular, we consider the following model $\Pi_{\text{AR}}(\mathcal{U})$:

$$\begin{aligned} z_{\text{AR}}(\mathcal{U}) = \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} + \max_{\mathbf{h} \in \mathcal{U}} \min_{\mathbf{y}(\mathbf{h})} \mathbf{d}^T \mathbf{y}(\mathbf{h}) \\ \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y}(\mathbf{h}) \geq \mathbf{h} \quad \forall \mathbf{h} \in \mathcal{U} \\ \mathbf{x} \in \mathbb{R}_+^{n_1} \\ \mathbf{y}(\mathbf{h}) \in \mathbb{R}_+^{n_2}, \end{aligned} \tag{1.1}$$

where $\mathbf{A} \in \mathbb{R}_+^{m \times n_1}$, $\mathbf{c} \in \mathbb{R}_+^{n_1}$, $\mathbf{d} \in \mathbb{R}_+^{n_2}$, $\mathbf{B} \in \mathbb{R}^{m \times n_2}$. The right-hand-side \mathbf{h} belongs to the uncertainty set $\mathcal{U} \subseteq \mathbb{R}_+^m$. The goal in this problem is to select the first-stage decision \mathbf{x} , and the second-stage recourse decision, $\mathbf{y}(\mathbf{h})$, as a function of the uncertain right hand side realization, \mathbf{h} such that the worst-case cost over all realizations of $\mathbf{h} \in \mathcal{U}$ is minimized. We assume without loss of generality that $n_1 = n_2 = n$ and the uncertainty set \mathcal{U} satisfies the following assumption.

Assumption 1. $\mathcal{U} \subseteq [0, 1]^m$ is convex, full-dimensional with $\mathbf{e}_i \in \mathcal{U}$ for all $i = 1, \dots, m$, and *down-monotone*, i.e., $\mathbf{h} \in \mathcal{U}$ and $\mathbf{0} \leq \mathbf{h}' \leq \mathbf{h}$ implies that $\mathbf{h}' \in \mathcal{U}$.

The above assumption is without loss of generality since we can appropriately scale the uncertainty set and consider a down-monotone completion without affecting the two-stage problem (1.1).

We would like to note that the objective coefficients \mathbf{c} , \mathbf{d} , the first-stage constraint matrix \mathbf{A} , and the decision variables $\mathbf{x}, \mathbf{y}(\mathbf{h})$ are all non-negative. This is restrictive as compared to general two-stage linear programs but the above model still captures many important applications including set cover, facility location and network design problems under uncertain demand. Here the right hand side, \mathbf{h} models the uncertain demand and the covering constraints capture the requirement of satisfying the uncertain demand.

The worst case scenario of problem (1.1) occurs on extreme points of \mathcal{U} . Therefore, given an explicit list of the extreme points of the uncertainty set \mathcal{U} , the adjustable robust optimization problem (1.1) can be solved efficiently by including the second-stage decisions and the covering constraints only for the extreme points of \mathcal{U} . However, in general the adjustable robust optimization problem (1.1) is intractable; for example, when the number of extreme points is large or due to other structural complexities of \mathcal{U} . In fact, Feige et al. [18] show that $\Pi_{\text{AR}}(\mathcal{U})$ (1.1) is hard to approximate within any factor that is better than $\Omega(\log m)$. This motivates us to consider approximations for the problem. Static robust and affinely adjustable solution approximations have been studied in the literature for this problem. In a static robust solution, we compute a single optimal solution (\mathbf{x}, \mathbf{y}) that is feasible for all realizations of the uncertain right hand side. Bertsimas et al. [11] relate the performance of static solution to the symmetry of the uncertainty set and show that it provides a good approximation to the adjustable problem if the uncertainty is close to being centrally symmetric. However, the performance of static solutions can be arbitrarily large for a general convex uncertainty set with the worst case performance being $\Omega(m)$. El Housni and Goyal [17] consider piecewise static policies for two-stage adjustable robust problem with uncertain constraint coefficients. These are a generalization of static policies where we divide the uncertainty set into several pieces and specify a static solution for each piece. However, they show that, in general, there is no piecewise static policy with a polynomial number of pieces that has a significantly better performance than an optimal static policy.

Ben-Tal et al. [2] introduce an affine adjustable solution (also known as affine policy) to approximate adjustable robust problems. An affine policy restricts the second-stage decisions, $\mathbf{y}(\mathbf{h})$ to being an affine function of the uncertain right-hand-side \mathbf{h} , i.e., $\mathbf{y}(\mathbf{h}) = \mathbf{P}\mathbf{h} + \mathbf{q}$ for some $\mathbf{P} \in \mathbb{R}^{n \times m}$ and $\mathbf{q} \in \mathbb{R}^m$ are decision variables. An optimal affine policy can be computed efficiently for a large class of problems and has a strong empirical performance. Bertsimas et al. [12] and Iancu et al. [20] show that affine policies are optimal for a class of multistage problems where there is a single parameter uncertain

in each period. Bertsimas and Goyal [10] show that affine policies are optimal if the uncertainty set \mathcal{U} is a simplex and give a worst case bound of $O(\sqrt{m})$ on the performance of affine policy for general uncertainty sets for the adjustable robust problem (1.1). Moreover, they show that this bound is tight for an uncertainty set quite analogous to the intersection of the unit ℓ_2 -norm ball and the non-negative orthant, i.e.,

$$\mathcal{U} = \{\mathbf{h} \in \mathbb{R}_+^m \mid \|\mathbf{h}\|_2 \leq 1, \mathbf{h} \geq \mathbf{0}\}. \quad (1.2)$$

Bertsimas and Bidkhori [6] provide improved approximation bounds for affine policies for $\Pi_{\text{AR}}(\mathcal{U})$ that depend on the geometric properties of the uncertainty set.

1.1 Our Contributions.

In this paper, we present a piecewise affine policy for the adjustable problem (1.1) where we consider pieces $\mathcal{U}_i, i \in [k]$ of the uncertainty set \mathcal{U} such that $\mathcal{U}_i \subseteq \mathcal{U}$ and \mathcal{U} is covered by the union of all pieces. For each \mathcal{U}_i , we have an affine solution $\mathbf{y}_i(\mathbf{h})$ for $\mathbf{h} \in \mathcal{U}_i$. The piecewise affine policy is significantly more general than static and affine policies and is known to be optimal for the adjustable robust problem (1.1) if the uncertainty set is a polytope. However, the number of pieces can be exponentially large. Furthermore, finding the optimal pieces is intractable in general. In fact, Bertsimas and Caramanis [8] show that it is NP-hard to construct the optimal pieces for piecewise policies with only two pieces for two-stage robust linear programs. In a recent paper, Bertsimas and Dunning [9] give a MIP based algorithm to adaptively partition the uncertainty set. However, there are no theoretical guarantees on the performance or the number of partitions. Our main contributions in this paper are the following.

New Framework for Piecewise affine policy. We present a new framework to efficiently construct a “good” piecewise affine policy for the adjustable robust problem. As we mention earlier, one of the significant challenges in designing a piecewise affine policy arises from constructing good pieces of the uncertainty set. We consider a new approach where instead of directly finding an explicit partition of \mathcal{U} , we approximate \mathcal{U} with a “simple” set $\hat{\mathcal{U}}$ satisfying the following two properties:

1. the adjustable robust problem (1.1) over $\hat{\mathcal{U}}$ can be solved efficiently,
2. $\hat{\mathcal{U}}$ “dominates” \mathcal{U} , i.e., for any $\mathbf{h} \in \mathcal{U}$, there exists $\hat{\mathbf{h}} \in \hat{\mathcal{U}}$ such that $\mathbf{h} \leq \hat{\mathbf{h}}$.

The domination property of $\hat{\mathcal{U}}$ preserves the feasibility of the adjustable robust problem if we consider the uncertainty set $\hat{\mathcal{U}}$ instead of \mathcal{U} . We consider $\hat{\mathcal{U}}$ to be a simplex dominating \mathcal{U} in our framework. Therefore, the adjustable robust problem (1.1) over \mathcal{U} can be solved efficiently since $\hat{\mathcal{U}}$ only has $m + 1$ extreme points. The piecewise affine policy is constructed from an optimal adjustable robust solution over $\hat{\mathcal{U}}$ using a piecewise affine map from $\mathbf{h} \in \mathcal{U}$

to a point $\hat{\mathbf{h}}$ that dominates \mathbf{h} . We show that the performance of our policy is significantly better than the affine policy for many important uncertainty sets both theoretically and numerically. We elaborate on the two ingredients of designing our piecewise affine policy below, namely, constructing $\hat{\mathcal{U}}$ and constructing the piecewise map below.

- a) **Constructing dominating uncertainty set.** Our framework is based on choosing an appropriate *dominating simplex* $\hat{\mathcal{U}}$ based on the geometric structure of \mathcal{U} . We consider $\hat{\mathcal{U}}$ to be a simplex of the following form

$$\hat{\mathcal{U}} = \beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v}),$$

where $\beta > 0$ and $\mathbf{v} \in \mathcal{U}$ are chosen appropriately such that $\hat{\mathcal{U}}$ dominates \mathcal{U} . Solving the adjustable robust problem over $\hat{\mathcal{U}}$ gives a feasible solution for the adjustable robust problem over \mathcal{U} due to the domination property. Moreover, the optimal adjustable solution over $\hat{\mathcal{U}}$ gives a β -approximation for the adjustable robust problem over \mathcal{U} since $\hat{\mathcal{U}} = \beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v}) \subseteq \beta \cdot \mathcal{U}$. The approximation bound, β , is related to a geometric *scaling factor* that represents the Banach-Mazur distance between \mathcal{U} and $\hat{\mathcal{U}}$. We would like to note that $\hat{\mathcal{U}}$ does not necessarily contain \mathcal{U} .

We also give an algorithm to construct the dominating set $\hat{\mathcal{U}}$ for a general uncertainty set \mathcal{U} . However, the algorithm requires us to solve an MIP which is computationally much harder than the case of permutation invariant sets but can be computed efficiently in practice. Moreover, the construction of the dominating set $\hat{\mathcal{U}}$ is independent from the parameters of the adjustable problem and depends only on the uncertainty set, \mathcal{U} . Therefore, it can be pre-computed offline and used to construct the piecewise affine policy for the adjustable problem efficiently.

- b) **Piecewise affine mapping.** We consider the following piecewise affine mapping that maps any $\mathbf{h} \in \mathcal{U}$ to a dominating point $\hat{\mathbf{h}}$ such that $\mathbf{h} \leq \hat{\mathbf{h}}$.

$$\hat{\mathbf{h}}(\mathbf{h}) = \beta \mathbf{v} + (\mathbf{h} - \beta \mathbf{v})^+.$$

We show that for any $\mathbf{h} \in \mathcal{U}$, $\hat{\mathbf{h}}(\mathbf{h})$ is contained in the down-monotone completion of $2 \cdot \hat{\mathcal{U}}$. Our piecewise affine policy is based on the above piecewise affine mapping and gives a 2β -approximation for the adjustable robust problem over \mathcal{U} . In our piecewise affine policy, $\beta \mathbf{v}$ is covered by the static component and $(\mathbf{h} - \beta \mathbf{v})^+$ is covered by the piecewise linear component of our policy. This is quite analogous to *threshold policies* that are widely used in dynamic optimization. We would like to note that $\hat{\mathbf{h}}$ does not necessarily belong to $\hat{\mathcal{U}}$ but is contained in the down-monotone completion of $2 \cdot \hat{\mathcal{U}}$ and therefore, we get an approximation factor of 2β instead of β . We can construct a set-dependent piecewise affine map between \mathcal{U} and $\hat{\mathcal{U}}$ that allows us to construct a piecewise affine policy with a performance bound of β .

Given the dominating set, $\hat{\mathcal{U}}$, our piecewise affine policy can be computed efficiently; in fact, it can be computed even more efficiently than an affine solution over \mathcal{U} in many cases because the adjustable problem over $\hat{\mathcal{U}}$ is a simple LP with only $m + 1$ constraints while the affine problem over \mathcal{U} is a general convex program for general convex uncertainty sets.

Performance bounds for Permutation invariant sets. We consider the class of permutation invariant sets including norm-balls, intersection of norm-balls and budget of uncertainty sets. This is an important family of uncertainty sets that are widely used in practice. We show that we can efficiently construct the dominating set $\hat{\mathcal{U}}$ and compute the scaling factor β for any permutation invariant set \mathcal{U} . In particular, we give an efficiently computable closed form expression for computing β and $\mathbf{v} \in \mathcal{U}$ to construct $\hat{\mathcal{U}}$. Therefore we can efficiently construct our piecewise affine solution with a performance bound of $2\beta(\mathcal{U}, \hat{\mathcal{U}})$.

Using this framework, we provide approximation bounds for our piecewise affine policy that are significantly better than the performance bound of an optimal affine policy in [6] for many permutation invariant uncertainty sets. For instance, we show that our policy gives a $O(m^{1/4})$ approximation for the two-stage adjustable robust problem (1.1) with hypersphere uncertainty set as in (1.2), while the affine policy has an approximation bound of $O(\sqrt{m})$ [6]. More generally, the performance bound for our policy for the p -norm ball is $O(m^{\frac{p-1}{p^2}})$ as opposed to $O(m^{\frac{1}{p}})$ given by an affine policy in [6]. Table 1 summarizes the comparison between the performance bounds for our policy as compared to the bounds for affine policy in [6]. We also present computational experiments and observe that our policy outperforms affine policy both in terms of objective and computation time on the family test of instances considered. (We would like to note that in [6], in Tables 1 and 2, there is a typo in the performance bound for affine policies for p -norm balls. According to Theorem 3 in [6], the bound should be

$$\frac{m^{\frac{p-1}{p}} + m}{m^{\frac{p-1}{p}} + m^{\frac{1}{p}}} = O\left(m^{\frac{1}{p}}\right),$$

instead of $\frac{m^{\frac{p-1}{p}} + m}{m^{\frac{1}{p}} + m}$ as mentioned in Table 2 in [6].

General uncertainty sets. We give an algorithm to construct the dominating set $\hat{\mathcal{U}}$ and a piecewise affine policy for general uncertainty set \mathcal{U} . The algorithm requires us to solve an MIP which is computationally much harder than the case of permutation invariant sets but can be computed efficiently in practice. Moreover, the construction of the dominating set $\hat{\mathcal{U}}$ is independent from the parameters of the adjustable problem and depends only on the uncertainty set, \mathcal{U} . Therefore, it can be pre-computed offline and used to construct the piecewise affine policy for the adjustable robust problem efficiently.

We show that our policy gives a $O(\sqrt{m})$ -approximation for general uncertainty sets which is same as the worst-case performance bound for affine

No.	Uncertainty set	Bounds in [6]	Our Bounds
1	$\{\mathbf{h} \in \mathbb{R}_+^m \mid \ \mathbf{h}\ _2 \leq 1\}$	$O(\sqrt{m})$	$O\left(m^{\frac{1}{4}}\right)$
2	$\{\mathbf{h} \in \mathbb{R}_+^m \mid \mathbf{h}^T \Sigma \mathbf{h} \leq 1\}$	—	$O\left(m^{\frac{2}{3}}\right)$
3	$\{\mathbf{h} \in \mathbb{R}_+^m \mid \ \mathbf{h}\ _p \leq 1\}$	$O\left(m^{\frac{1}{p}}\right)$	$O\left(m^{\frac{p-1}{p^2}}\right)$
4	$\{\mathbf{h} \in \mathbb{R}_+^m \mid \ \mathbf{h}\ _p \leq 1, \ \mathbf{h}\ _q \leq r\}$	$O\left(r^{-1}m^{\frac{1}{q}}\right)$	$O\left(\min\left(r^{\frac{1-p}{p}}, m^{\frac{p-1}{pq}}, r^{\frac{1}{q}}m^{\frac{q-1}{q^2}}\right)\right)$
5	$\{\mathbf{h} \in \mathbb{R}_+^m \mid \sum_{i=1}^m h_i \leq k, \forall i \in [m] h_i \leq 1\}$	$O\left(\frac{k^2+m}{k^2+m}\right)$	$O\left(\min\left(k, \frac{m}{k}\right)\right)$

Table 1 Comparison with performance bounds for affine policies in Bertsimas and Bikhori [6]. The ellipsoid in Example 2 is assumed to be permutation invariant set. There is no specialized bound for this Ellipsoid in [6]. For intersection of norm-balls (Example 4 in the table), we assume $m^{\frac{1}{q} - \frac{1}{p}} \geq r \geq 1$.

policy. We also show that the bound of $O(\sqrt{m})$ is tight. In particular, we show that for the budget uncertainty set

$$\mathcal{U} = \left\{ \mathbf{h} \in \mathbb{R}_+^m \mid \sum_{i=1}^m h_i = \sqrt{m}, 0 \leq h_i \leq 1 \quad \forall i \in [m] \right\},$$

the performance bound of our piecewise affine policy is $\Theta(\sqrt{m})$. Furthermore, we show that the bound of $\Theta(\sqrt{m})$ holds even if we consider dominating sets with a polynomial number of extreme points that are significantly more general than a simplex. While this example shows that the worst-case performance of our policy is the same as the worst-case performance of the affine policy, we would like to emphasize that our policy gives a significantly better approximation than affine policies for many important uncertainty sets as discussed above.

Outline. In Section 2, we present the new framework for approximating the two-stage adjustable robust problem (1.1) via dominating uncertainty sets and constructing piecewise affine policies. In Section 3 we provide improved approximation bounds for (1.1) for permutation invariant sets. We present the case of general uncertainty sets in Section 4. In Section 5, we present a family of lower-bound instances where our piecewise affine policy has the worst performance bound and finally in Section 6, we present a computational study to test our policy and compare it to an affine policy over \mathcal{U} .

2 Our framework for piecewise affine policies

We present a piecewise affine policy to approximate the two-stage adjustable robust problem (1.1). Our policy is based on approximating the uncertainty set \mathcal{U} with a simple set $\hat{\mathcal{U}}$ such that the adjustable problem (1.1) can be efficiently solved over $\hat{\mathcal{U}}$. In particular, we select $\hat{\mathcal{U}}$ such that it *dominates* \mathcal{U} and it is *close* to \mathcal{U} . We make these notions precise with the following definitions.

Definition 1 (Domination) Given an uncertainty set $\mathcal{U} \subseteq \mathbb{R}_+^m$, $\hat{\mathcal{U}} \subseteq \mathbb{R}_+^m$ dominates \mathcal{U} if for all $\mathbf{h} \in \mathcal{U}$, there exists $\hat{\mathbf{h}} \in \hat{\mathcal{U}}$ such that $\hat{\mathbf{h}} \geq \mathbf{h}$.

Definition 2 (Scaling factor) Given a full-dimensional uncertainty set $\mathcal{U} \subseteq \mathbb{R}_+^m$ and $\hat{\mathcal{U}} \subseteq \mathbb{R}_+^m$ that dominates \mathcal{U} . We define the scaling factor $\beta(\mathcal{U}, \hat{\mathcal{U}})$ as following

$$\beta(\mathcal{U}, \hat{\mathcal{U}}) = \min \left\{ \beta > 0 \mid \hat{\mathcal{U}} \subseteq \beta \cdot \mathcal{U} \right\}.$$

For the sake of simplicity, we denote the scaling factor $\beta(\mathcal{U}, \hat{\mathcal{U}})$ by β in the rest of this paper. Note that this scaling factor always exists since \mathcal{U} is full-dimensional. Moreover, it is greater than one because $\hat{\mathcal{U}}$ dominates \mathcal{U} . We would like to emphasize that the dominating set $\hat{\mathcal{U}}$ does not necessarily contain \mathcal{U} . We illustrate this in the following example.

Example. Consider the uncertainty set \mathcal{U} defined in (1.2) which is the intersection of the unit ℓ_2 -norm ball and the non-negative orthant. We show later in this paper (Proposition 1) that the simplex $\hat{\mathcal{U}}$ dominates \mathcal{U} where

$$\hat{\mathcal{U}} = m^{\frac{1}{4}} \cdot \text{conv} \left(e_1, \dots, e_m, \frac{1}{\sqrt{m}} e \right). \quad (2.1)$$

In Figures 1 and 2, we demonstrate the sets \mathcal{U} and $\hat{\mathcal{U}}$ for $m = 3$. Note that $\hat{\mathcal{U}}$ does not contain \mathcal{U} but only dominates \mathcal{U} . This is an important property in our framework.

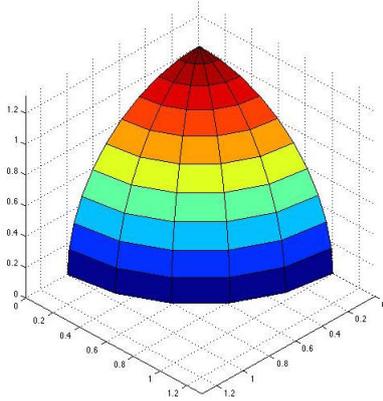


Fig. 1 The uncertainty set (1.2)

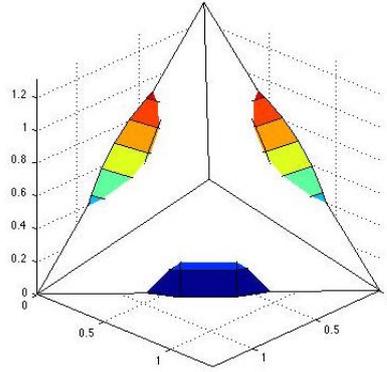


Fig. 2 The dominating set $\hat{\mathcal{U}}$ (2.1)

The following theorem shows that solving the adjustable problem over the set $\hat{\mathcal{U}}$ gives a β -approximation to the two-stage adjustable robust problem (1.1).

Theorem 1 Consider an uncertainty set \mathcal{U} that verifies Assumption 1 and $\hat{\mathcal{U}} \subseteq \mathbb{R}_+^m$ dominates $\mathcal{U} \subseteq \mathbb{R}_+^m$. Let β be the scaling factor of $(\mathcal{U}, \hat{\mathcal{U}})$. Moreover, let $z_{\text{AR}}(\mathcal{U})$ and $z_{\text{AR}}(\hat{\mathcal{U}})$ be the optimal values for (1.1) on \mathcal{U} and $\hat{\mathcal{U}}$, respectively. Then,

$$z_{\text{AR}}(\mathcal{U}) \leq z_{\text{AR}}(\hat{\mathcal{U}}) \leq \beta \cdot z_{\text{AR}}(\mathcal{U}).$$

The proof of Theorem 1 is presented in Appendix A.

2.1 Choice of $\hat{\mathcal{U}}$

Theorem 1 provides a new framework for approximating the two-stage adjustable robust problem $\Pi_{\text{AR}}(\mathcal{U})$ (1.1). Note that we require that $\hat{\mathcal{U}}$ to be such that it dominates \mathcal{U} and $\Pi_{\text{AR}}(\hat{\mathcal{U}})$ can be efficiently solved over $\hat{\mathcal{U}}$. In fact, the latter is satisfied if the number of extreme points of $\hat{\mathcal{U}}$ is small and explicitly given (typically polynomial of m). In our framework, we choose the dominating set to be a simplex of the following form

$$\hat{\mathcal{U}} = \beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v}), \quad (2.2)$$

for some $\mathbf{v} \in \mathcal{U}$. The coefficient β and $\mathbf{v} \in \mathcal{U}$ are chosen such that $\hat{\mathcal{U}}$ dominates \mathcal{U} . For a given $\hat{\mathcal{U}}$ (i.e., β and $\mathbf{v} \in \mathcal{U}$), the adjustable robust problem, $\Pi_{\text{AR}}(\hat{\mathcal{U}})$ (1.1) can be solved efficiently as it can be reduced to the following LP:

$$\begin{aligned} z_{\text{AR}}(\hat{\mathcal{U}}) = \min \quad & \mathbf{c}^T \mathbf{x} + z \\ & z \geq \mathbf{d}^T \mathbf{y}_i, \quad \forall i \in [m+1] \\ & \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y}_i \geq \beta \mathbf{e}_i, \quad \forall i \in [m] \\ & \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y}_{m+1} \geq \beta \mathbf{v} \\ & \mathbf{x} \in \mathbb{R}_+^n, \quad \mathbf{y}_i \in \mathbb{R}_+^n, \quad \forall i \in [m+1]. \end{aligned}$$

2.2 Mapping to dominating points

Consider the following piecewise affine mapping for any $\mathbf{h} \in \mathcal{U}$:

$$\forall \mathbf{h} \in \mathcal{U}, \quad \hat{\mathbf{h}}(\mathbf{h}) = \beta \mathbf{v} + (\mathbf{h} - \beta \mathbf{v})_+. \quad (2.3)$$

We show that this maps any $\mathbf{h} \in \mathcal{U}$ to a dominating point contained in the down-monotone completion of $2 \cdot \hat{\mathcal{U}}$. First, let us introduce the following structural result.

Lemma 1 (Structural Result) *Consider an uncertainty set \mathcal{U} that verifies Assumption 1. Consider β and $\mathbf{v} \in \mathcal{U}$ such that $\hat{\mathcal{U}} = \beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v})$ dominates \mathcal{U} . Then,*

$$\frac{1}{\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ \leq 1, \quad \forall \mathbf{h} \in \mathcal{U}. \quad (2.4)$$

Proof Consider $\mathbf{h} \in \mathcal{U}$. Since $\hat{\mathcal{U}}$ dominates \mathcal{U} , there exists $\alpha_1, \alpha_2, \dots, \alpha_{m+1} \geq 0$ with $\alpha_1 + \dots + \alpha_{m+1} = 1$ such that

$$h_i \leq \beta (\alpha_i + \alpha_{m+1} v_i), \quad \forall i = 1, \dots, m. \quad (2.5)$$

Let

$$I(\mathbf{h}) = \left\{ i \in [m] \mid h_i - \beta v_i \geq 0 \right\}.$$

Then,

$$\begin{aligned} \sum_{i=1}^m (h_i - \beta v_i)^+ &= \sum_{i \in I(\mathbf{h})} h_i - \beta \sum_{i \in I(\mathbf{h})} v_i \\ &\leq \sum_{i \in I(\mathbf{h})} \beta (\alpha_i + \alpha_{m+1} v_i) - \beta \sum_{i \in I(\mathbf{h})} v_i \\ &= \beta \sum_{i \in I(\mathbf{h})} \alpha_i + (\alpha_{m+1} - 1) \beta \sum_{i \in I(\mathbf{h})} v_i \\ &\leq \beta, \end{aligned}$$

where the first inequality follows from (2.5) and the last inequality holds because $\alpha_{m+1} - 1 \leq 0$, $v_i \geq 0$, $\beta \geq 0$ and $\sum_{i \in I(\mathbf{h})} \alpha_i \leq 1$. We conclude that

$$\frac{1}{\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ \leq 1.$$

□

The following lemma shows that the mapping in (2.3) maps any $\mathbf{h} \in \mathcal{U}$ to a dominating point that belongs to the down-monotone completion of $2 \cdot \hat{\mathcal{U}}$.

Lemma 2 *For all $\mathbf{h} \in \mathcal{U}$, $\hat{\mathbf{h}}(\mathbf{h})$ as defined in (2.3) is a dominating point that belongs to the down-monotone completion of $2 \cdot \hat{\mathcal{U}}$.*

Proof It is clear that $\hat{\mathbf{h}}(\mathbf{h})$ dominates \mathbf{h} because $\hat{\mathbf{h}}(\mathbf{h}) \geq \beta \mathbf{v} + (\mathbf{h} - \beta \mathbf{v}) = \mathbf{h}$. Moreover, for all $\mathbf{h} \in \mathcal{U}$, we have

$$\hat{\mathbf{h}}(\mathbf{h}) = \beta \mathbf{v} + \frac{1}{\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ \beta \mathbf{e}_i.$$

From Lemma 1,

$$1 + \frac{1}{\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ \leq 2.$$

Therefore, $\hat{\mathbf{h}}(\mathbf{h})$ belongs to the down-monotone completion of $2 \cdot \hat{\mathcal{U}}$. □

2.3 Piecewise affine policy

We construct a piecewise affine policy over \mathcal{U} from the optimal solution of $\Pi_{\text{AR}}(\hat{\mathcal{U}})$ based on the piecewise affine mapping in (2.3). Let $\hat{\mathbf{x}}, \hat{\mathbf{y}}(\hat{\mathbf{h}})$ for $\hat{\mathbf{h}} \in \hat{\mathcal{U}}$ be an optimal solution of $\Pi_{\text{AR}}(\hat{\mathcal{U}})$. Since $\hat{\mathcal{U}}$ is a simplex, we can compute this efficiently. We consider the following piecewise affine policy for $\Pi_{\text{AR}}(\mathcal{U})$ (1.1).

Piecewise affine policy.

$$\begin{aligned} \mathbf{x} &= 2\hat{\mathbf{x}} \\ \mathbf{y}(\mathbf{h}) &= \frac{1}{\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ \hat{\mathbf{y}}(\beta \mathbf{e}_i) + \hat{\mathbf{y}}(\beta \mathbf{v}), \quad \forall \mathbf{h} \in \mathcal{U}. \end{aligned} \quad (2.6)$$

The following theorem shows that the above piecewise affine policy gives a 2β -approximation for $\Pi_{\text{AR}}(\mathcal{U})$ (1.1).

Theorem 2 *Consider an uncertainty set \mathcal{U} that verifies Assumption 1 and $\hat{\mathcal{U}} = \beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v})$ be a dominating set where $\mathbf{v} \in \mathcal{U}$. The piecewise affine solution in (2.6) is feasible and gives a 2β -approximation for the adjustable robust problem, $\Pi_{\text{AR}}(\mathcal{U})$ (1.1).*

Proof First, we show that the policy (2.6) is feasible. We have,

$$\begin{aligned} \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y}(\mathbf{h}) &= 2\mathbf{A}\hat{\mathbf{x}} + \mathbf{B} \left(\frac{1}{\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ \hat{\mathbf{y}}(\beta \mathbf{e}_i) + \hat{\mathbf{y}}(\beta \mathbf{v}) \right) \\ &= (\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}(\beta \mathbf{v})) + \mathbf{A}\hat{\mathbf{x}} + \frac{1}{\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ \mathbf{B}\hat{\mathbf{y}}(\beta \mathbf{e}_i) \\ &\geq (\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}(\beta \mathbf{v})) + \frac{1}{\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ (\mathbf{B}\hat{\mathbf{y}}(\beta \mathbf{e}_i) + \mathbf{A}\hat{\mathbf{x}}) \\ &\geq \beta \mathbf{v} + \sum_{i=1}^m (h_i - \beta v_i)^+ \mathbf{e}_i \\ &\geq \beta \mathbf{v} + \sum_{i=1}^m (h_i - \beta v_i) \mathbf{e}_i = \mathbf{h}, \end{aligned}$$

where the first inequality follows from Lemma 1 and the non-negativity of $\hat{\mathbf{x}}$ and \mathbf{A} . The second inequality follows from the feasibility of $\hat{\mathbf{x}}, \hat{\mathbf{y}}(\hat{\mathbf{h}})$.

To compute the performance of (2.6), we have for any $\mathbf{h} \in \mathcal{U}$,

$$\begin{aligned} \mathbf{c}^T \mathbf{x} + \mathbf{d}^T \mathbf{y}(\mathbf{h}) &= 2 \left(\mathbf{c}^T \hat{\mathbf{x}} + \mathbf{d}^T \left(\frac{1}{2\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ \hat{\mathbf{y}}(\beta \mathbf{e}_i) + \frac{1}{2} \hat{\mathbf{y}}(\beta \mathbf{v}) \right) \right) \\ &\leq 2 \left(\mathbf{c}^T \hat{\mathbf{x}} + \max_{\hat{\mathbf{h}} \in \hat{\mathcal{U}}} \mathbf{d}^T \hat{\mathbf{y}}(\hat{\mathbf{h}}) \left(\frac{1}{2\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ + \frac{1}{2} \right) \right) \\ &\leq 2 \left(\mathbf{c}^T \hat{\mathbf{x}} + \max_{\hat{\mathbf{h}} \in \hat{\mathcal{U}}} \mathbf{d}^T \hat{\mathbf{y}}(\hat{\mathbf{h}}) \right) \\ &= 2 \cdot z_{\text{AR}}(\hat{\mathcal{U}}), \end{aligned}$$

where the second last inequality follows from Lemma 1. From Theorem 1, $z_{\text{AR}}(\hat{\mathcal{U}}) \leq \beta \cdot z_{\text{AR}}(\mathcal{U})$. Therefore, the cost of the piecewise affine policy for any $\mathbf{h} \in \mathcal{U}$

$$\mathbf{c}^T \mathbf{x} + \mathbf{d}^T \mathbf{y}(\mathbf{h}) \leq 2\beta \cdot z_{\text{AR}}(\mathcal{U}),$$

which implies that the piecewise affine solution (2.6) gives 2β -approximation for the adjustable robust problem, $\Pi_{\text{AR}}(\mathcal{U})$ (1.1). \square

The above proof shows that it is sufficient to find β and $\mathbf{v} \in \mathcal{U}$ satisfying (2.4) in Lemma 1 to construct a piecewise affine policy that gives a 2β -approximation for $\Pi_{\text{AR}}(\mathcal{U})$ (1.1). Moreover, for any $\beta, \mathbf{v} \in \mathcal{U}$ satisfying (2.4), $2\beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v})$ dominates \mathcal{U} in this case. In particular, we have the following corollary.

Corollary 1 *Consider an uncertainty set \mathcal{U} satisfying Assumption 1. Consider any β and $\mathbf{v} \in \mathcal{U}$ satisfying (2.4). Then, the piecewise affine solution in (2.6) gives a 2β -approximation for the adjustable robust problem, $\Pi_{\text{AR}}(\mathcal{U})$ (1.1). Moreover, $2\beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v})$ dominates \mathcal{U} .*

Proof The first part of the corollary follows directly from the proof of Theorem 2. Let us prove that $2\beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v})$ dominates \mathcal{U} . Denote $s \geq 0$ such that

$$\frac{1}{\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ + s = 1.$$

For any $\mathbf{h} \in \mathcal{U}$, let

$$\hat{\mathbf{h}} = \sum_{i=1}^m (h_i - \beta v_i)^+ \mathbf{e}_i + \beta(1+s)\mathbf{v}.$$

Then for all $i = 1, \dots, m$,

$$\begin{aligned} \hat{h}_i &= (h_i - \beta v_i)^+ + \beta(1+s)v_i \\ &\geq (h_i - \beta v_i)^+ + \beta v_i \geq h_i. \end{aligned}$$

Therefore, $\hat{\mathbf{h}}$ dominates \mathbf{h} . Moreover,

$$\hat{\mathbf{h}} = 2\beta \left(\sum_{i=1}^m \frac{(h_i - \beta v_i)^+}{2\beta} \mathbf{e}_i + \frac{1+s}{2} \mathbf{v} \right) \in 2\beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v}).$$

We conclude that $2\beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v})$ dominates \mathcal{U} . \square

We would like to note that our piecewise affine policy is not necessarily an optimal piecewise policy. However, for a large class of uncertainty sets, we show that our policy is significantly better than affine policy and can even be computed more efficiently than an affine policy.

3 Performance Bounds for Permutation Invariant Sets

In this section, we consider the class of permutation invariant sets including norm-balls, intersection of norm-balls and budget of uncertainty sets and present the performance bounds for our policy. These are widely used sets in theory and practice and have nice symmetry properties.

Definition 3 (Permutation Invariant Sets) \mathcal{U} is a permutation invariant set if $\mathbf{x} \in \mathcal{U}$ implies that for any permutation τ of $\{1, 2, \dots, m\}$, $\mathbf{x}^\tau \in \mathcal{U}$ where $x_i^\tau = x_{\tau(i)}$.

We first introduce some structural properties of permutation invariant sets. Consider any permutation invariant set, \mathcal{U} satisfying Assumption 1. For all $k = 1, \dots, m$, let

$$\gamma(k) = \frac{1}{k} \cdot \max \left\{ \sum_{i=1}^k h_i \mid \mathbf{h} \in \mathcal{U} \right\}. \quad (3.1)$$

The coefficients, $\gamma(k)$ for all $k = 1, \dots, m$ describe the geometric structure of \mathcal{U} . In particular, we have the following lemma.

Lemma 3 *Let \mathcal{U} be a permutation invariant set and $\gamma(\cdot)$ be as defined in (3.1). Then,*

$$\gamma(k) \cdot \sum_{i=1}^k \mathbf{e}_i \in \mathcal{U}, \quad \forall k = 1, \dots, m$$

We present the proof of Lemma 3 in Appendix B. For the sake of simplicity, we denote $\gamma(m)$ by γ in the rest of the paper. From the above lemma, we know that $\gamma \cdot \mathbf{e} \in \mathcal{U}$.

3.1 Our Piecewise affine policy for Permutations Invariant Sets

For any permutation invariant set \mathcal{U} , we consider the following dominating uncertainty set, $\hat{\mathcal{U}}$ of the form (2.2) with $\mathbf{v} = \gamma \mathbf{e}$, i.e.,

$$\hat{\mathcal{U}} = \beta \cdot \text{conv}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \gamma \mathbf{e}) \quad (3.2)$$

where β is the scaling factor such that $\hat{\mathcal{U}}$ dominates \mathcal{U} . The above dominating set is motivated by the symmetry of the permutation invariant set, \mathcal{U} . In this section, we show that we can efficiently compute the minimum β such that $\hat{\mathcal{U}}$ in (3.2) dominates \mathcal{U} . In particular, we give an efficiently computable closed form expression to compute β for any permutation invariant set \mathcal{U} .

From Corollary 1 we know that to construct a piecewise affine policy with an approximation bound of 2β , it is sufficient to find β such that

$$\frac{1}{\beta} \max_{\mathbf{h} \in \mathcal{U}} \sum_{i=1}^m (h_i - \beta\gamma)^+ \leq 1. \quad (3.3)$$

Furthermore, any β that satisfy (3.3) implies that $2\beta \cdot \text{conv}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \gamma \mathbf{e})$ dominates \mathcal{U} . So, we will concentrate on finding the minimum β that satisfies (3.3), i.e.,

$$\min \left\{ \beta \geq 1 \mid \frac{1}{\beta} \max_{\mathbf{h} \in \mathcal{U}} \sum_{i=1}^m (h_i - \beta\gamma)^+ \leq 1 \right\}. \quad (3.4)$$

The following lemma characterizes the structure of the optimal solution for the maximization problem in (3.3) for a fixed β .

Lemma 4 *Consider the maximization problem in (3.3) for a fixed β . There exists an optimal solution \mathbf{h}^* such that*

$$\mathbf{h}^* = \gamma(k) \cdot \sum_{i=1}^k \mathbf{e}_i,$$

for some $k = 1, \dots, m$.

We present the proof of Lemma 4 in Appendix C. The following lemma characterizes the optimal β for (3.4).

Lemma 5 *Let \mathcal{U} a permutation invariant uncertainty set satisfying Assumption 1. Then the optimal solution for (3.4) is given by*

$$\beta = \max_{k=1, \dots, m} \frac{\gamma(k)}{\gamma + \frac{1}{k}}. \quad (3.5)$$

Proof Using Lemma 4, we can reformulate (3.4) as follows.

$$\min \left\{ \beta \geq 1 \mid \frac{1}{\beta} \max_{k=1, \dots, m} \sum_{i=1}^k (\gamma(k) - \beta\gamma) \leq 1 \right\},$$

i.e.,

$$\min \left\{ \beta \geq 1 \mid \beta \geq \frac{\gamma(k)}{\gamma + \frac{1}{k}}, \forall k = 1, \dots, m \right\}.$$

Therefore,

$$\beta = \max_{k=1, \dots, m} \frac{\gamma(k)}{\gamma + \frac{1}{k}}.$$

□

The above lemma computes the minimum β that satisfies (3.3). Therefore, from Corollary 1, we have the following theorem.

Theorem 3 *Let \mathcal{U} be a permutation invariant set satisfying Assumption 1. Let $\gamma = \gamma(m)$ be as defined in (3.1) and β be as defined in (3.5),*

$$\hat{\mathcal{U}} = \beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \gamma \mathbf{e}),$$

and let $\hat{\mathbf{x}}, \hat{\mathbf{y}}(\hat{\mathbf{h}})$ for $\hat{\mathbf{h}} \in \hat{\mathcal{U}}$ be an optimal solution for $\Pi_{\text{AR}}(\hat{\mathcal{U}})$ (1.1). Then the following piecewise affine solution

$$\begin{aligned} \mathbf{x} &= 2\hat{\mathbf{x}} \\ \mathbf{y}(\mathbf{h}) &= \frac{1}{\beta} \sum_{i=1}^m (h_i - \beta\gamma)^+ \hat{\mathbf{y}}(\beta \mathbf{e}_i) + \hat{\mathbf{y}}(\beta\gamma \mathbf{e}) \quad \forall \mathbf{h} \in \mathcal{U}, \end{aligned} \quad (3.6)$$

gives a 2β -approximation for $\Pi_{\text{AR}}(\mathcal{U})$ (1.1). Moreover, the set $2 \cdot \hat{\mathcal{U}}$ dominates \mathcal{U} .

Therefore, for any permutation invariant uncertainty set, \mathcal{U} , we can compute the piecewise-affine policy for $\Pi_{\text{AR}}(\mathcal{U})$ (1.1) efficiently; in fact, even more efficiently than an affine policy over \mathcal{U} in many cases.

3.2 Examples

We present the approximation bounds for several permutation invariant uncertainty sets that are commonly used in the literature and practice including norm balls, intersection of norm balls and budget of uncertainty sets. In particular, we show that for these sets, the performance bounds of our piecewise affine policy are significantly better than the best known performance bounds for affine policy.

Proposition 1 (Hypersphere) *Consider the uncertainty set $\mathcal{U} = \{\mathbf{h} \in \mathbb{R}_+^m \mid \|\mathbf{h}\|_2 \leq 1\}$ which is the intersection of the unit hypersphere and the nonnegative orthant. Then,*

$$\hat{\mathcal{U}} = m^{\frac{1}{4}} \cdot \text{conv} \left(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \frac{1}{\sqrt{m}} \mathbf{e} \right),$$

dominates \mathcal{U} and our piecewise affine solution (3.6) gives $O(m^{\frac{1}{4}})$ approximation to (1.1).

Proof We have for $k = 1, \dots, m$,

$$\gamma(k) = \frac{1}{k} \cdot \max \left\{ \sum_{i=1}^k h_i \mid \mathbf{h} \in \mathcal{U} \right\} = \frac{1}{\sqrt{k}}.$$

In particular, $\gamma = \frac{1}{\sqrt{m}}$. From Lemma 5 we get,

$$\begin{aligned} \beta &= \max_{k=1, \dots, m} \frac{\gamma(k)}{\gamma(m) + \frac{1}{k}} \\ &= \max_{k=1, \dots, m} \frac{\frac{1}{\sqrt{k}}}{\frac{1}{\sqrt{m}} + \frac{1}{k}}. \end{aligned}$$

The maximum of this problem occurs for $k = \sqrt{m}$. Then, $\beta = \frac{m^{\frac{1}{4}}}{2}$. We conclude from Theorem 3 that $\hat{\mathcal{U}}$ dominates \mathcal{U} and our piecewise affine policy gives $O(m^{\frac{1}{4}})$ approximation to the adjustable problem (1.1). \square

Proposition 2 (p-norm ball) Consider the p -norm ball uncertainty set $\mathcal{U} = \{\mathbf{h} \in \mathbb{R}_+^m \mid \|\mathbf{h}\|_p \leq 1\}$ where $p \geq 1$. Then

$$\hat{\mathcal{U}} = 2\beta \cdot \text{conv} \left(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, m^{-\frac{1}{p}} \mathbf{e} \right)$$

dominates \mathcal{U} with

$$\beta = \frac{1}{p} (p-1)^{\frac{p-1}{p}} \cdot m^{\frac{p-1}{p^2}} = O(m^{\frac{p-1}{p^2}}).$$

Our piecewise affine solution (3.6) gives $O(m^{\frac{p-1}{p^2}})$ approximation to (1.1).

Proof We have for $k = 1, \dots, m$,

$$\gamma(k) = \frac{1}{k} \cdot \max \left\{ \sum_{i=1}^k h_i \mid \mathbf{h} \in \mathcal{U} \right\} = k^{\frac{-1}{p}}.$$

In particular, $\gamma = m^{\frac{-1}{p}}$. From Lemma 5 we get,

$$\begin{aligned} \beta &= \max_{k=1, \dots, m} \frac{\gamma(k)}{\gamma(m) + \frac{1}{k}} \\ &= \max_{k=1, \dots, m} \frac{k^{\frac{-1}{p}}}{m^{\frac{-1}{p}} + \frac{1}{k}} \\ &= \frac{1}{p} (p-1)^{\frac{p-1}{p}} \cdot m^{\frac{p-1}{p^2}} = O\left(m^{\frac{p-1}{p^2}}\right). \end{aligned}$$

We conclude from Theorem 3 that $\hat{\mathcal{U}}$ dominates \mathcal{U} and our piecewise affine policy gives $O(m^{\frac{p-1}{p^2}})$ approximation to the adjustable problem (1.1). \square

Proposition 3 (Intersection of two norm balls) Consider \mathcal{U} the intersection of the norm balls $\mathcal{U}_1 = \{\mathbf{h} \in \mathbb{R}_+^m \mid \|\mathbf{h}\|_p \leq 1\}$ and $\mathcal{U}_2 = \{\mathbf{h} \in \mathbb{R}_+^m \mid \|\mathbf{h}\|_q \leq r\}$ where $p > q \geq 1$ and $m^{\frac{1}{q} - \frac{1}{p}} \geq r \geq 1$. Then

$$\hat{\mathcal{U}} = \beta \cdot \text{conv} \left(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \left(rm^{-\frac{1}{q}} \right) \mathbf{e} \right),$$

where

$$\beta = \min(\beta_1, \beta_2), \quad \beta_1 = r^{\frac{1-p}{p}} m^{\frac{p-1}{pq}}, \quad \text{and} \quad \beta_2 = r^{\frac{1}{q}} m^{\frac{q-1}{q^2}}.$$

Our piecewise affine solution (3.6) gives 2β approximation to (1.1).

Proof To prove that $\hat{\mathcal{U}}$ dominates $\mathcal{U}_1 \cap \mathcal{U}_2$, it is sufficient to consider \mathbf{h} in the boundary of \mathcal{U}_1 or \mathcal{U}_2 and find $\alpha_1, \alpha_2, \dots, \alpha_{m+1} \geq 0$ with $\alpha_1 + \dots + \alpha_{m+1} = 1$ such that for all $i \in [m]$,

$$h_i \leq \beta \left(\alpha_i + rm^{-\frac{1}{q}} \alpha_{m+1} \right).$$

Case 1: $\beta = \beta_1$.

Let $\mathbf{h} \in \mathcal{U}_1$ such that $\|\mathbf{h}\|_p = 1$, we take $\alpha_i = \frac{h_i^p}{p}$ for $i \in [m]$ and $\alpha_{m+1} = \frac{p-1}{p}$. First, we have $\sum_{i=1}^{m+1} \alpha_i = 1$ and for all $i \in [m]$,

$$\begin{aligned} \beta \left(\alpha_i + rm^{-\frac{1}{q}} \alpha_{m+1} \right) &= \beta_1 \left(\frac{h_i^p}{p} + \frac{p-1}{p} rm^{-\frac{1}{q}} \right) \\ &\geq \beta_1 (h_i^p)^{\frac{1}{p}} \left(rm^{-\frac{1}{q}} \right)^{\frac{p-1}{p}} = h_i, \end{aligned}$$

where the inequality follows from the weighted AM-GM inequality. Therefore $\hat{\mathcal{U}}$ dominates $\mathcal{U}_1 \cap \mathcal{U}_2$.

Case 2: $\beta = \beta_2$.

Let $\mathbf{h} \in \mathcal{U}_2$ such that $\|\mathbf{h}\|_q = r$, we take $\alpha_i = \frac{h_i^q}{r^q q}$ for $i \in [m]$ and $\alpha_{m+1} = \frac{q-1}{q}$. First, we have $\sum_{i=1}^{m+1} \alpha_i = 1$ and for all $i \in [m]$,

$$\begin{aligned} \beta \left(\alpha_i + rm^{-\frac{1}{q}} \alpha_{m+1} \right) &= \beta_2 \left(\frac{h_i^q}{r^q q} + \frac{q-1}{q} rm^{-\frac{1}{q}} \right) \\ &\geq \beta_2 \left(\frac{h_i^q}{r^q} \right)^{\frac{1}{q}} \left(rm^{-\frac{1}{q}} \right)^{\frac{q-1}{q}} = h_i, \end{aligned}$$

where the inequality followed from the weighted AM-GM inequality. Therefore, $\hat{\mathcal{U}}$ dominates $\mathcal{U}_1 \cap \mathcal{U}_2$. \square

We also consider a permutation invariant uncertainty set that is the intersection of an ellipsoid and the non-negative orthant, i.e.,

$$\mathcal{U} = \left\{ \mathbf{h} \in \mathbb{R}_+^m \mid \mathbf{h}^T \Sigma \mathbf{h} \leq 1 \right\} \quad (3.7)$$

where $\Sigma \succeq \mathbf{0}$. For \mathcal{U} to be a permutation invariant set satisfying Assumption 1, Σ must be of the following form

$$\Sigma = \begin{pmatrix} 1 & a & \dots & a \\ a & 1 & \dots & a \\ \vdots & \vdots & \ddots & \vdots \\ a & a & \dots & 1 \end{pmatrix} \quad (3.8)$$

where $0 \leq a \leq 1$.

Proposition 4 (Permutation invariant ellipsoid) *Consider the uncertainty set \mathcal{U} defined in (3.7) where Σ is defined in (3.8). Then*

$$\hat{\mathcal{U}} = \beta \cdot \text{conv}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \gamma \mathbf{e}),$$

dominates \mathcal{U} with

$$\beta = \left(\frac{a}{2} + \frac{(1-a)^{\frac{1}{2}}}{(am^2 + (1-a)m)^{\frac{1}{4}}} \right)^{-1} = O\left(m^{\frac{2}{5}}\right)$$

and

$$\gamma = \frac{1}{\sqrt{(am^2 + (1-a)m)}}.$$

Our piecewise affine policy (3.6) gives $O\left(m^{\frac{2}{5}}\right)$ approximation to the adjustable robust problem (1.1).

The proof of Proposition 4 is presented in Appendix D.

Proposition 5 (Budget of uncertainty set) *Consider the budget of uncertainty set*

$$\mathcal{U} = \left\{ \mathbf{h} \in \mathbb{R}_+^m \mid \sum_{i=1}^m h_i \leq k, \forall i \in [m], h_i \leq 1 \right\}.$$

Then,

$$\hat{\mathcal{U}} = \beta \cdot \text{conv}\left(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \frac{k}{m} \mathbf{e}\right)$$

where $\beta = \min\left(k, \frac{m}{k}\right)$. In particular, our piecewise affine policy (3.6) gives 2β approximation to the adjustable problem (1.1).

The proof of Proposition 5 is presented in Appendix E.

Table 1 summarizes the performance bounds for our piecewise affine policy and the bounds for affine policy in [6] for these sets. Our piecewise affine policy performs significantly better than the known bounds for affine policy for many interesting sets. For instance, our policy gives $O(m^{\frac{1}{4}})$ -approximation for the hypersphere uncertainty set and $O(m^{\frac{p-1}{p^2}})$ -approximation for the p -norm ball while affine policy gives $O(m^{\frac{1}{2}})$ -approximation for hypersphere and

$O(m^{\frac{1}{p}})$ -approximation for the p -norm ball [6]. We would like to mention that Table 1 presents the best known performance bounds in the literature for affine policies [6]. It is possible that for some uncertainty sets, the actual performance bound for affine policy is better significantly lower.

However, we present a family examples where an optimal affine policy gives an $\Omega(\sqrt{m})$ -approximation, while our policy gives a significantly better $O(m^{\frac{1}{4}})$ -approximation for the adjustable robust problem (1.1). In particular, we consider the worst-case examples for affine policy in [10]. For any $\delta > 0$, Bertsimas and Goyal [10] give a family of examples such that

$$z_{\text{Aff}}(\mathcal{U}) = \Omega(m^{\frac{1}{2}-\delta}) \cdot z_{\text{AR}}(\mathcal{U}).$$

The uncertainty set in the family of examples is a polytope with exponential number of extreme points which is close to the hypersphere. The instance in [10] is given by

$$\begin{aligned} \mathbf{c} &= \mathbf{0} & \mathbf{d} &= \mathbf{e} \\ \mathbf{A} &= \mathbf{0} \\ B_{ij} &= \begin{cases} 1 & \text{if } i = j \\ \theta_0 & \text{otherwise.} \end{cases} & \text{for all } i, j = 1, \dots, m & \quad (3.9) \\ \mathcal{U} &= \text{conv}(\mathbf{h}^0, \mathbf{h}^1, \dots, \mathbf{h}^N) \end{aligned}$$

where $\theta_0 = \frac{1}{m^{\frac{1-\delta}{2}}}$, $r = \lceil m^{1-\delta} \rceil$, $N = \binom{m}{r} + m + 2$ and

$$\begin{aligned} \mathbf{h}^0 &= \mathbf{0} \\ \mathbf{h}^j &= \mathbf{e}_j \quad \forall j = 1, \dots, m \\ \mathbf{h}^{m+1} &= \frac{1}{\sqrt{m}} \cdot \mathbf{e} \\ \mathbf{h}^{m+2} &= \theta_0 \cdot \underbrace{[1, \dots, 1, 0 \dots, 0]}_r \end{aligned}$$

where exactly r coordinates are non-zero, each equal to θ_0 . Extreme points \mathbf{h}^j , $j \geq m+3$ are permutations of the non-zero coordinates of \mathbf{h}^{m+2} . Therefore, \mathcal{U} has exactly $\binom{m}{r}$ extreme points of the form of \mathbf{h}^{m+2} . Note that all the non-zero extreme points of \mathcal{U} are roughly on the boundaries of the unit hypersphere or exactly on the boundary of the hypersphere when $m^{1-\delta}$ is integer.

Lemma 6 *Our piecewise affine policy (2.6) gives an $O(m^{\frac{1}{4}})$ -approximation for the adjustable robust problem (1.1) for instance (3.9).*

Proof Let \mathcal{S} be the unit hypersphere intersected with the non-negative orthant. For any extreme point $\mathbf{h}^j \in \mathcal{U}$, $j = 0, \dots, N$, $\|\mathbf{h}^j\|_2 \leq \sqrt{2}$. Therefore, $\mathcal{U} \subseteq \sqrt{2} \cdot \mathcal{S}$. From Proposition 1, we know that $m^{\frac{1}{4}} \cdot \text{conv}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \frac{1}{\sqrt{m}}\mathbf{e})$ dominates \mathcal{S} . Let

$$\hat{\mathcal{U}} = \sqrt{2} \cdot m^{\frac{1}{4}} \cdot \text{conv}\left(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \frac{1}{\sqrt{m}}\mathbf{e}\right).$$

Therefore, $\hat{\mathcal{U}}$ dominates \mathcal{U} . Moreover,

$$\text{conv}\left(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m, \frac{1}{\sqrt{m}}\mathbf{e}\right) \subseteq \mathcal{U},$$

which implies that the scaling factor

$$\beta(\mathcal{U}, \hat{\mathcal{U}}) = O\left(m^{\frac{1}{4}}\right).$$

Therefore, our piecewise affine policy gives $O(m^{\frac{1}{4}})$ -approximation for instance (3.9) using $\hat{\mathcal{U}}$ as a dominating set. \square

Since the performance of the affine policy is $\Omega(m^{\frac{1}{2}-\delta})$ for the instance (3.9), our policy provides a significant improvement. We would like to note that since $\hat{\mathcal{U}}$ is a simplex, an affine policy is optimal for $\Pi_{\text{AR}}(\hat{\mathcal{U}})$. In particular, we have the following

$$z_{\text{AR}}(\mathcal{U}) \leq z_{\text{AR}}(\hat{\mathcal{U}}) = z_{\text{Aff}}(\hat{\mathcal{U}}) \leq O(m^{1/4}) \cdot z_{\text{AR}}(\mathcal{U}),$$

where the first inequality follows as $\hat{\mathcal{U}}$ dominates \mathcal{U} and the last inequality follows from Lemma 6. Moreover, from [10], we know that for instance (3.9),

$$z_{\text{Aff}}(\mathcal{U}) = \Omega(m^{1/2-\delta}) \cdot z_{\text{AR}}(\mathcal{U}).$$

Therefore,

$$z_{\text{Aff}}(\mathcal{U}) = \Omega(m^{1/4-\delta}) \cdot z_{\text{Aff}}(\hat{\mathcal{U}}),$$

which is quite surprising since $\hat{\mathcal{U}}$ dominates \mathcal{U} . We would like to emphasize that $\hat{\mathcal{U}}$ does not contain \mathcal{U} and this is crucial to get a significant improvement for our piecewise affine policy constructed through the dominating set.

4 General uncertainty set

In this section, we consider the case of general uncertainty sets. The main challenge in our framework of constructing the piecewise affine policy is the choice of the dominating simplex, $\hat{\mathcal{U}}$. More specifically, the choice of β and $\mathbf{v} \in \mathcal{U}$ such that $\beta \cdot \text{conv}(\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v})$ dominates \mathcal{U} . For a permutation invariant set, \mathcal{U} , we choose $\mathbf{v} = \gamma\mathbf{e}$ and we can efficiently find β using Lemma 5 to construct the dominating set. However, this does not extend to general sets and we need a new procedure to find those parameters.

Corollary 1 shows that to construct a good piecewise affine policy over \mathcal{U} , it is sufficient to find β and $\mathbf{v} \in \mathcal{U}$ such that for all $\mathbf{h} \in \mathcal{U}$

$$\frac{1}{\beta} \sum_{i=1}^m (h_i - \beta v_i)^+ \leq 1. \quad (4.1)$$

In this section, we present an iterative algorithm to find such β and $\mathbf{v} \in \mathcal{U}$ satisfying (4.1). In each iteration t , the algorithm maintains a candidate

Algorithm 1 Computing β and \mathbf{v} for general uncertainty sets

```

1: Initialize  $t = 0$ ,  $\mathbf{u}^0 = 0$ 
2: while  $\left\{ \max_{\mathbf{h} \in \mathcal{U}} \sum_{i=1}^m (h_i - u_i^t)^+ > t \right\}$  do
3:    $\mathbf{h}^t \in \operatorname{argmax}_{\mathbf{h} \in \mathcal{U}} \sum_{i=1}^m (h_i - u_i^t)^+$ 
4:   for  $i = 1, \dots, m$  do
5:     if  $u_i^t = 1$  then  $h_i^t = 0$ 
6:     end if
7:      $u_i^{t+1} = \min(1, u_i^t + h_i^t)$ 
8:   end for
9:    $t = t + 1$ 
10: end while
11: return  $\beta = t$ ,  $\mathbf{v} = \frac{\mathbf{u}^t}{\beta}$ .

```

solution, β^t and $\mathbf{v}^t \in \mathcal{U}$. Let $\mathbf{u}^t = \beta^t \cdot \mathbf{v}^t$. The algorithm solves the following maximization problem:

$$\max_{\mathbf{h} \in \mathcal{U}} \sum_{i=1}^m (h_i - u_i^t)^+ \quad (4.2)$$

The algorithm stops if the optimal value is at most β^t . Otherwise, let \mathbf{h}^t be an optimal solution for (4.2). The candidate solutions are updated as follows

$$\begin{aligned} \beta^{t+1} &= \beta^t + 1 \\ u_i^{t+1} &= \min(1, u_i^t + h_i^t). \end{aligned}$$

This corresponds to updating $\mathbf{v}^{t+1} = \frac{1}{\beta^{t+1}} \cdot \mathbf{u}^{t+1}$. Algorithm 1 presents the steps in detail.

The correctness of the algorithm is straightforward. In fact, the algorithm will stop when the inequality (4.1) is verified for all $\mathbf{h} \in \mathcal{U}$. The number of iterations β is finite since \mathcal{U} is compact. The following theorem shows that \mathbf{v} returned by the algorithm belongs to \mathcal{U} and the corresponding piecewise affine policy is a $O(\sqrt{m})$ -approximation for the adjustable problem (1.1).

Theorem 4 *Suppose Algorithm 1 returns β , \mathbf{v} . Then $\mathbf{v} \in \mathcal{U}$. Furthermore, the piecewise affine policy (2.6) with parameters β and \mathbf{v} gives a $O(\sqrt{m})$ -approximation for the adjustable problem (1.1).*

Proof Suppose Algorithm 1 returns β, \mathbf{v} . Note that β is the number of iterations in Algorithm 1. First, we have

$$\mathbf{u}^\beta \leq \sum_{t=0}^{\beta-1} \mathbf{h}^t.$$

Moreover $\frac{1}{\beta} \cdot \sum_{t=0}^{\beta-1} \mathbf{h}^t \in \mathcal{U}$ since \mathcal{U} is convex. Therefore $\mathbf{v} = \frac{\mathbf{u}^\beta}{\beta} \in \mathcal{U}$ by down-monotonicity of \mathcal{U} .

Let us prove that $\beta = O(\sqrt{m})$. First, note that, when we set $h_i^t = 0$ for $u_i^t = 1$. The objective of the maximization problem in the algorithm does not change and \mathbf{h}^t still belongs to \mathcal{U} by down-monotonicity. Then, for any $t = 0, \dots, \beta - 1$

$$\sum_{i=1}^m (h_i^t - u_i^t)^+ > t.$$

Moreover, $h_i^t \geq 0$ and $u_i^t \geq 0$, hence $h_i^t \geq (h_i^t - u_i^t)^+$ and therefore for all $t = 0, \dots, \beta - 1$

$$\sum_{i=1}^m h_i^t > t.$$

Then,

$$\sum_{t=0}^{\beta-1} \sum_{i=1}^m h_i^t > \sum_{t=0}^{\beta-1} t = \frac{1}{2}\beta(\beta - 1). \quad (4.3)$$

Note that, if $u_i^t = 1$ at some iteration t , then $h_i^{t'} = 0$ for any $t' \geq t$. Hence, for any $i \in [m]$,

$$\sum_{t=0}^{\beta-1} h_i^t \leq u_i^\beta + 1 \leq 2. \quad (4.4)$$

Hence, from (4.3) and from (4.4) we get, $2m > \frac{1}{2}\beta(\beta - 1)$, i.e., $\beta \cdot (\beta - 1) \leq 4m$, which implies, $\beta = O(\sqrt{m})$. \square

We would like to note that the maximization problem (4.2) that Algorithm 1 solves in each iteration t is not a convex optimization problem. However, (4.2) can be formulated as the following MIP:

$$\begin{aligned} \max \quad & \sum_{i=1}^m z_i \\ & z_i \leq (h_i - u_i^t) + (1 - x_i) \quad \forall i \in [m], \\ & z_i \leq x_i \quad \forall i \in [m] \\ & z_i \geq 0, \quad \forall i \in [m] \\ & x_i \in \{0, 1\} \quad \forall i \in [m] \\ & \mathbf{h} \in \mathcal{U}. \end{aligned} \quad (4.5)$$

Therefore, the procedure to find β and $\mathbf{v} \in \mathcal{U}$ to construct the dominating set for \mathcal{U} is computationally much more challenging than the case of permutation invariant sets. However, we would like to note that the computation of β and \mathbf{v} only depends on the uncertainty set and not on the problem instance. Therefore, we can compute this offline and use it to efficiently construct a good piecewise affine policy.

Connection to Bertsimas and Goyal [10]. We would like to note that Algorithm 1 is quite analogous to the explicit construction of good affine policies

in [10]. The analysis of the $O(\sqrt{m})$ -approximation bound for affine policies is based on the following projection result (which is a restatement of Lemma 8 and Lemma 9 in [10]).

Theorem 5 [Bertsimas and Goyal 2011] *Consider any uncertainty set \mathcal{U} satisfying Assumption 1. There exists $\beta \leq \sqrt{m}$, $\mathbf{v} \in \mathcal{U}$ such that*

$$\sum_{j:\beta v_j < 1} h_j \leq \beta, \forall \mathbf{h} \in \mathcal{U}.$$

Suppose $J = \{j \mid \beta v_j < 1\}$. The affine solution in [10] covers $\beta \mathbf{v}$ using the static component and the components J using a linear solution. The linear solution does not exploit the coverage of βv_i for $i \in J$ from the static solution. The approximation factor is $O(\beta)$ since for all $\mathbf{h} \in \mathcal{U}$, $\sum_{j \in J} h_j \leq \beta$.

Our piecewise affine solution given by Algorithm 1 finds analogous β , $\mathbf{v} \in \mathcal{U}$ such that

$$\sum_{i=1}^m (h_i - \beta v_i)_+ \leq \beta, \forall \mathbf{h} \in \mathcal{U}.$$

In the piecewise affine solution, the static component covers $\beta \mathbf{v}$ and the remaining demand $(\mathbf{h} - \beta \mathbf{v})_+$ is covered by a piecewise-linear function that exploits the coverage of $\beta \mathbf{v}$. This allows us to improve significantly as compared to the affine policy for a large family of uncertainty sets. We would like to note again that our policy is not necessarily an optimal one and there can be examples where affine policy is better than our policy.

5 Worst case example for the domination policy

From Theorem 4, we know that our piecewise affine policy gives an $O(\sqrt{m})$ -approximation for the adjustable robust problem (1.1). In this section, we show that this bound is tight for a budget of uncertainty set. In particular, we consider the following budget of uncertainty set with a budget equal to \sqrt{m} :

$$\mathcal{U} = \left\{ \mathbf{h} \in \mathbb{R}_+^m \mid \sum_{i=1}^m h_i = \sqrt{m}, 0 \leq h_i \leq 1 \quad \forall i \in [m] \right\}. \quad (5.1)$$

We show that our dominating simplex based piecewise affine policy gives an $\Omega(\sqrt{m})$ -approximation to the adjustable robust problem (1.1). The lower bound of $\Omega(\sqrt{m})$ holds even when we consider more general dominating sets than simplex. We show that for any $\epsilon > 0$, there is no polynomial number of points in \mathcal{U} such that the convex hull of those points scaled by $m^{\frac{1}{2}-\epsilon}$ dominates \mathcal{U} . In particular, we have the following theorem.

Theorem 6 *Given any $0 < \epsilon < 1/2$, and $k \in \mathbb{N}$, consider the budget of uncertainty set, \mathcal{U} (5.1) with m sufficiently large. Let $P(m) \leq m^k$. For any $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{P(m)} \in \mathcal{U}$, the set*

$$\hat{\mathcal{U}} = m^{\frac{1}{2}-\epsilon} \cdot \text{conv}(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{P(m)}),$$

does not dominate \mathcal{U} .

Proof Suppose for a sake of contradiction that there exists $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{P(m)} \in \mathcal{U}$ such that $\hat{\mathcal{U}} = m^{\frac{1}{2}-\epsilon} \cdot \text{conv}(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{P(m)})$ dominates \mathcal{U} .

By Caratheodory's theorem, we know that any point in \mathcal{U} can be expressed as a convex combination of at most $m+1$ extreme points of \mathcal{U} . Therefore

$$\hat{\mathcal{U}} \subseteq m^{\frac{1}{2}-\epsilon} \cdot \text{conv}(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{Q(m)}),$$

where $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{Q(m)}$ are extreme points of \mathcal{U} and

$$Q(m) \leq (m+1) \cdot P(m) = O(m^{k+1}).$$

Consider any $I \subseteq \{1, 2, \dots, m\}$ such that $|I| = \sqrt{m}$. Let \mathbf{h} be an extreme point of \mathcal{U} corresponding to I , i.e., $h_i = 1$ if $i \in I$ and $h_i = 0$ otherwise. There exists $\hat{\mathbf{h}} \in \hat{\mathcal{U}}$ such that $\mathbf{h} \leq \hat{\mathbf{h}}$. Let

$$\hat{\mathbf{h}} = m^{\frac{1}{2}-\epsilon} \sum_{j=1}^{Q(m)} \alpha_j \mathbf{y}_j,$$

where $\sum_{j=1}^{Q(m)} \alpha_j = 1$ and $\alpha_j \geq 0$ for all $j = 1, 2, \dots, Q(m)$. Therefore,

$$1 \leq m^{\frac{1}{2}-\epsilon} \sum_{j=1}^{Q(m)} \alpha_j y_{ji}, \quad \forall i \in I.$$

Summing over $i \in I$, we have,

$$\sqrt{m} = |I| \leq m^{\frac{1}{2}-\epsilon} \sum_{i \in I} \sum_{j=1}^{Q(m)} \alpha_j y_{ji}.$$

Therefore,

$$\begin{aligned} m^\epsilon &\leq \sum_{j=1}^{Q(m)} \alpha_j \sum_{i \in I} y_{ji}, \\ &\leq \sum_{j=1}^{Q(m)} \alpha_j \max_{j=1,2,\dots,Q(m)} \sum_{i \in I} y_{ji} \\ &= \max_{j=1,2,\dots,Q(m)} \sum_{i \in I} y_{ji}. \end{aligned}$$

Let $\mathcal{F} = \{I \subseteq \{1, 2, \dots, m\} \mid |I| = \sqrt{m}\}$. Note that the cardinality of \mathcal{F} is

$$|\mathcal{F}| = \binom{m}{\sqrt{m}}.$$

For any $I \in \mathcal{F}$ there exists $\mathbf{y} \in \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{Q(m)}\}$ such that

$$\sum_{i \in I} y_i \geq m^\epsilon.$$

Therefore, there exists $\mathbf{y} \in \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{Q(m)}\}$ and $\tilde{\mathcal{F}} \subseteq \mathcal{F}$ such that

$$|\tilde{\mathcal{F}}| \geq \frac{1}{Q(m)} \binom{m}{\sqrt{m}}, \text{ and} \quad (5.2)$$

$$\sum_{i \in I} y_i \geq m^\epsilon, \forall I \in \tilde{\mathcal{F}}.$$

Note that \mathbf{y} is an extreme point of \mathcal{U} . Hence, \mathbf{y} has exactly \sqrt{m} ones and the remaining components are zeros. The maximum cardinality of the subsets $I \subseteq [m]$ that satisfy (5.2) is

$$\sum_{k=m^\epsilon}^{k=\sqrt{m}} \binom{\sqrt{m}}{k} \cdot \binom{m-\sqrt{m}}{\sqrt{m}-k}.$$

By over counting, the above sum can be upper-bounded by

$$\binom{\sqrt{m}}{m^\epsilon} \cdot \binom{m-m^\epsilon}{\sqrt{m}-m^\epsilon}.$$

Therefore,

$$\binom{\sqrt{m}}{m^\epsilon} \cdot \binom{m-m^\epsilon}{\sqrt{m}-m^\epsilon} \geq |\tilde{\mathcal{F}}| \geq \frac{1}{Q(m)} \binom{m}{\sqrt{m}}$$

Then,

$$\frac{\binom{\sqrt{m}}{m^\epsilon} \cdot \binom{m-m^\epsilon}{\sqrt{m}-m^\epsilon}}{\binom{m}{\sqrt{m}}} \geq \frac{1}{Q(m)}. \quad (5.3)$$

which is a contradiction (see Appendix F). \square

6 Computational study

In this section, we present a computational study to compare the performance of our policy with affine policies both in terms of objective value and computation times.

Experimental setup. We consider three different uncertainty sets for our computational experiments, namely, *i*) hypersphere uncertainty set (1.2), *ii*) p -norm ball with $p = 3$, and *iii*) p -norm ball with $p = 3/2$ defined in Proposition 2. The test instances of the adjustable robust problem (1.1) are constructed as follows. We choose $n = m$, $\mathbf{c} = \mathbf{d} = \mathbf{e}$ and $\mathbf{A} = \mathbf{B}$ where \mathbf{B} is randomly generated as

$$\mathbf{B} = \mathbf{I}_m + \mathbf{G},$$

where \mathbf{I}_m is the identity matrix and \mathbf{G} is a random normalized gaussian. In particular, for the hypersphere uncertainty set, $G_{ij} = |Y_{ij}|/\sqrt{m}$, for the 3-norm ball, $G_{ij} = |Y_{ij}|/m^{\frac{1}{3}}$ and for the $\frac{3}{2}$ -norm ball uncertainty set, $G_{ij} = |Y_{ij}|/m^{\frac{2}{3}}$,

m	Avg	Max	Min	5%	10%	25%	50%	$T_{p\text{-aff}}(s)$	$T_{\text{aff}}(s)$
10	0.955	1.021	0.835	1.004	0.997	0.978	0.965	0.108	0.517
20	1.115	1.156	1.054	1.145	1.142	1.134	1.118	0.139	1.162
30	1.218	1.259	1.174	1.248	1.240	1.233	1.220	0.183	2.709
40	1.289	1.345	1.242	1.317	1.314	1.303	1.290	0.212	9.300
50	1.347	1.378	1.312	1.375	1.372	1.358	1.348	0.247	21.178
60	1.400	1.443	1.368	1.423	1.419	1.410	1.400	0.323	47.485
70	1.446	1.476	1.416	1.466	1.461	1.455	1.449	0.330	112.863
80	1.486	1.517	1.463	1.502	1.498	1.492	1.485	0.404	211.143
90	1.524	1.553	1.503	1.541	1.537	1.531	1.523	0.485	434.059
100	1.555	1.579	1.538	1.572	1.569	1.564	1.555	0.591	965.120

Table 2 Comparison on the performance and computation time of affine policy and our piecewise affine policy for the hypersphere uncertainty set. For 100 instances, we compute $\frac{z_{\text{Aff}}(U)}{z_{p\text{-aff}}(U)}$ and present the average, min, max ratios and the percentiles 5%, 10%, 25%, 50%. Here, $T_{p\text{-aff}}(s)$ denotes the running time for our piecewise affine policy and $T_{\text{aff}}(s)$ denotes the running time for affine policy in seconds.

m	Avg	Max	Min	5%	10%	25%	50%	$T_{p\text{-aff}}(s)$	$T_{\text{aff}}(s)$
10	0.975	1.051	0.895	1.040	1.026	0.993	0.972	0.157	1.729
20	1.085	1.162	1.018	1.123	1.114	1.101	1.086	0.184	4.063
30	1.161	1.203	1.111	1.197	1.187	1.177	1.160	0.192	9.597
40	1.220	1.253	1.168	1.247	1.241	1.230	1.221	0.218	21.053
50	1.267	1.298	1.224	1.291	1.285	1.277	1.267	0.266	53.371
60	1.310	1.336	1.284	1.330	1.327	1.320	1.312	0.304	130.348
70	1.346	1.371	1.322	1.367	1.362	1.355	1.346	0.356	287.014
80	1.378	1.400	1.349	1.395	1.393	1.386	1.379	0.444	616.678
90	1.408	1.428	1.384	1.423	1.420	1.414	1.408	0.535	1144.954
100	1.436	1.451	1.419	1.445	1.444	1.440	1.436	0.637	2214.153

Table 3 Comparison on the performance and computation time of affine policy and our piecewise affine policy for the 3-norm ball uncertainty set.

m	Avg	Max	Min	5%	10%	25%	50%	$T_{p\text{-aff}}(s)$	$T_{\text{aff}}(s)$
10	0.910	0.963	0.857	0.950	0.943	0.927	0.910	0.146	1.516
20	1.031	1.071	0.990	1.057	1.050	1.041	1.033	0.179	3.665
30	1.111	1.136	1.088	1.128	1.126	1.118	1.110	0.200	8.593
40	1.175	1.192	1.156	1.189	1.188	1.181	1.175	0.222	20.355
50	1.225	1.245	1.211	1.236	1.233	1.230	1.226	0.315	57.618
60	1.267	1.280	1.255	1.277	1.274	1.271	1.268	0.294	109.963
70	1.304	1.318	1.293	1.310	1.309	1.306	1.304	0.352	224.721
80	1.335	1.346	1.323	1.341	1.339	1.337	1.335	0.448	419.427
90	1.363	1.373	1.355	1.369	1.369	1.365	1.363	0.568	824.759
100	1.387	1.394	1.378	1.393	1.392	1.390	1.387	0.680	1595.490

Table 4 Comparison on the performance and computation time of affine policy and our piecewise affine policy for the 3/2-norm ball uncertainty set.

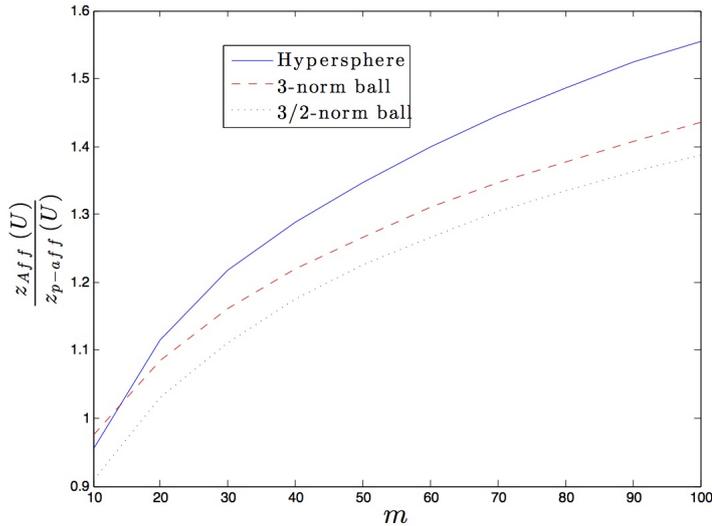


Fig. 3 The ratio $r = \frac{z_{\text{Aff}}(\mathcal{U})}{z_{\text{p-aff}}(\mathcal{U})}$ for different values of m and different uncertainty sets: hypersphere, 3-norm ball and 3/2-norm ball.

where Y_{ij} is i.i.d. standard gaussian. We consider values of m from $m = 10$ to $m = 100$ in increments of 10 and consider 100 instances for each value of m .

We construct the piecewise affine policy based on the dominating simplex $\hat{\mathcal{U}}$ as described in Proposition 1 and Proposition 2. Let $z_{\text{p-aff}}(\mathcal{U})$ denote the worst-case objective value of our piecewise affine policy. Note that the piecewise affine policy over \mathcal{U} is computed by solving the adjustable robust problem over $\hat{\mathcal{U}}$ and $z_{\text{p-aff}}(\mathcal{U}) = z_{\text{AR}}(\hat{\mathcal{U}})$. We report the ratio $r = \frac{z_{\text{Aff}}(\mathcal{U})}{z_{\text{p-aff}}(\mathcal{U})}$ in Tables 2, 3 and 4 for the hypersphere, 3-norm ball and 3/2-norm ball uncertainty sets respectively. In particular, for each value of m , we report the average ratio Avg, the maximum ratio Max, the minimum ratio Min, the quantiles 5%, 10%, 25%, 50% for the ratio r , and the running time of our policy $T_{\text{p-aff}}(s)$ and the running time of affine policy $T_{\text{aff}}(s)$.

Results. We observe in the computational experiments that the piecewise affine policy performs significantly better than affine policy for our family of test instances. The gap between our piecewise affine policy and affine policy increases as m increases. In Figure 3, we observe that the ratio $r = \frac{z_{\text{Aff}}(\mathcal{U})}{z_{\text{p-aff}}(\mathcal{U})}$ increases significantly with m which implies that our policy has a significant improvement over affine policy for large values of m . We also observe that the gap for the hypersphere uncertainty set is larger than the one for the norm-balls. This matches the theoretical bounds presented in Table 1. We would like to note that for small values of m (in particular, $m = 10$), the performance of affine policy is better than our policy. This is because our policy is not an optimal piecewise affine policy and there are instances where an affine policy

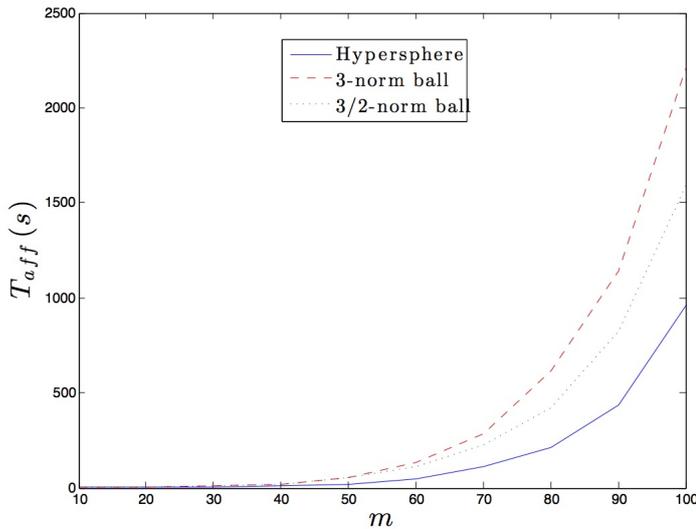


Fig. 4 Running time of affine policy in seconds for different values of m and different uncertainty sets: hypersphere, 3-norm ball and 3/2-norm ball.

performs significantly better. However, for large values of m , the performance of our policy is significantly better for all three classes of uncertainty sets.

Furthermore, our policy scales very well and the average running time is less than one second even for large values of m . On the other hand, computing the optimal affine policy over \mathcal{U} becomes computationally challenging as m increases. For instance, the average running time for computing an optimal affine policy for $m = 100$ is around 16 minutes for the hypersphere uncertainty set, around 36 minutes for the 3-norm ball and around 26 minutes for the 3/2-norm ball. Whereas, our policy can be computed in less than 1 second for all instances on average. Figure 4 gives the running time of affine policy for different values of m and as we can observe this running time becomes very high for large values of m .

7 Conclusions

In this paper, we present a new tractable framework for designing good piecewise affine policies for two-stage adjustable robust optimization problem. Our framework is based on approximating the uncertainty set by a simplex and constructing a piecewise affine policy based on the map from the uncertainty set to the simplex. We show that our piecewise affine policy performs significantly better than affine policy for many important uncertainty sets both theoretically and numerically. To the best of our knowledge, this is the first tractable framework for designing piecewise affine policies with significantly better theoretical guarantees than affine policies in many cases. While our policy improves over affine policy in many cases, we show that the worst case

performance bound for our policy is $\Theta(\sqrt{m})$ for the case of budget of uncertainty set with budget equal to \sqrt{m} . Therefore, it is an interesting open question to design piecewise affine policies that significantly improve over affine in the worst-case.

Acknowledgements O. El Housni and V. Goyal are supported by NSF grants CMMI 1201116 and CMMI 1351838.

References

1. A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University press, 2009.
2. A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.
3. A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.
4. A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25(1):1–14, 1999.
5. A. Ben-Tal and A. Nemirovski. Robust optimization—methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002.
6. D. Bertsimas and H. Bidkhori. On the performance of affine policies for two-stage adaptive optimization: a geometric perspective. *Mathematical Programming*, 153(2):577–594, 2015.
7. D. Bertsimas, D. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.
8. D. Bertsimas and C. Caramanis. Finite adaptability in multistage linear optimization. *Automatic Control, IEEE Transactions on*, 55(12):2751–2766, 2010.
9. D. Bertsimas and I. Dunning. Multistage robust mixed integer optimization with adaptive partitions. *Under Review*, 2015.
10. 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.
11. D. Bertsimas, V. Goyal, and X. Sun. A geometric characterization of the power of finite adaptability in multistage stochastic and adaptive optimization. *Mathematics of Operations Research*, 36(1):24–54, 2011.
12. D. Bertsimas, D. Iancu, and P. Parrilo. Optimality of Affine Policies in Multi-stage Robust Optimization. *Mathematics of Operations Research*, 35:363–394, 2010.
13. D. Bertsimas and M. Sim. Robust Discrete Optimization and Network Flows. *Mathematical Programming Series B*, 98:49–71, 2003.
14. D. Bertsimas and M. Sim. The Price of Robustness. *Operations Research*, 52(2):35–53, 2004.
15. G. Dantzig. Linear programming under uncertainty. *Management Science*, 1:197–206, 1955.
16. L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18:1035–1064, 1997.
17. O. El Housni and V. Goyal. Piecewise static policies for two-stage adjustable robust linear optimization problems under uncertainty. 2015.
18. U. Feige, K. Jain, M. Mahdian, and V. Mirrokni. Robust combinatorial optimization with exponential scenarios. *Lecture Notes in Computer Science*, 4513:439–453, 2007.
19. D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.
20. D. Iancu, M. Sharma, and M. Sviridenko. Supermodularity and affine policies in dynamic robust optimization. *Operations Research*, 61(4):941–956, 2013.
21. P. Kall and S. Wallace. *Stochastic programming*. Wiley New York, 1994.
22. A. Prékopa. *Stochastic programming*. Kluwer Academic Publishers, Dordrecht, Boston, 1995.

23. A. Shapiro. Stochastic programming approach to optimization under uncertainty. *Mathematical Programming, Series B*, 112(1):183–220, 2008.
24. A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming: modeling and theory*. Society for Industrial and Applied Mathematics, 2009.
25. A. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations research*, 21(5):1154–1157, 1973.

A Proof of Theorem 1

Proof Let $(\hat{\mathbf{x}}, \hat{\mathbf{y}}(\hat{\mathbf{h}}), \hat{\mathbf{h}} \in \hat{\mathcal{U}})$ be an optimal solution for $z_{\text{AR}}(\hat{\mathcal{U}})$. For each $\mathbf{h} \in \mathcal{U}$, let $\tilde{\mathbf{y}}(\mathbf{h}) = \hat{\mathbf{y}}(\hat{\mathbf{h}})$ where $\hat{\mathbf{h}} \in \hat{\mathcal{U}}$ dominates \mathbf{h} . Therefore, for any $\mathbf{h} \in \mathcal{U}$,

$$\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\tilde{\mathbf{y}}(\mathbf{h}) = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}(\hat{\mathbf{h}}) \geq \hat{\mathbf{h}} \geq \mathbf{h},$$

i.e., $(\hat{\mathbf{x}}, \tilde{\mathbf{y}}(\mathbf{h}), \mathbf{h} \in \mathcal{U})$ is a feasible solution for $z_{\text{AR}}(\mathcal{U})$. Therefore,

$$z_{\text{AR}}(\mathcal{U}) \leq \mathbf{c}^T \hat{\mathbf{x}} + \max_{\mathbf{h} \in \mathcal{U}} \mathbf{d}^T \tilde{\mathbf{y}}(\mathbf{h}) \leq \mathbf{c}^T \hat{\mathbf{x}} + \max_{\hat{\mathbf{h}} \in \hat{\mathcal{U}}} \mathbf{d}^T \hat{\mathbf{y}}(\hat{\mathbf{h}}) = z_{\text{AR}}(\hat{\mathcal{U}}).$$

Conversely, let $(\mathbf{x}^*, \mathbf{y}^*(\mathbf{h}), \mathbf{h} \in \mathcal{U})$ be an optimal solution of $z_{\text{AR}}(\mathcal{U})$. Then, for any $\hat{\mathbf{h}} \in \hat{\mathcal{U}}$, since $\frac{\hat{\mathbf{h}}}{\beta} \in \mathcal{U}$, we have,

$$\mathbf{A}\mathbf{x}^* + \mathbf{B}\mathbf{y}^*\left(\frac{\hat{\mathbf{h}}}{\beta}\right) \geq \frac{\hat{\mathbf{h}}}{\beta},$$

Therefore, $(\beta\mathbf{x}^*, \beta\mathbf{y}^*\left(\frac{\hat{\mathbf{h}}}{\beta}\right), \hat{\mathbf{h}} \in \mathcal{U})$ is feasible for $\Pi_{\text{AR}}(\hat{\mathcal{U}})$. Therefore,

$$z_{\text{AR}}(\hat{\mathcal{U}}) \leq \mathbf{c}^T \beta\mathbf{x}^* + \max_{\hat{\mathbf{h}} \in \hat{\mathcal{U}}} \mathbf{d}^T \beta\mathbf{y}^*\left(\frac{\hat{\mathbf{h}}}{\beta}\right) \leq \beta \cdot \left(\mathbf{c}^T \mathbf{x}^* + \max_{\mathbf{h} \in \mathcal{U}} \mathbf{d}^T \mathbf{y}^*(\mathbf{h}) \right) = \beta \cdot z_{\text{AR}}(\mathcal{U}).$$

□

B Proof of Lemma 3

Proof Suppose $k \in [m]$. Let us consider

$$\mathbf{h} \in \operatorname{argmax}_{\mathbf{h} \in \mathcal{U}} \sum_{i=1}^k h_i.$$

Without loss of generality, we can suppose that $h_i = 0$ for $i = k+1, \dots, m$. Denote, \mathcal{S}_k the set of permutations of $\{1, 2, \dots, k\}$. We define $\mathbf{h}^\sigma \in \mathbb{R}_+^m$ such that $h_i^\sigma = h_{\sigma(i)}$ for $i = 1, \dots, k$ and $h_i^\sigma = 0$ otherwise. Since \mathcal{U} is a permutation invariant set, we have $\mathbf{h}^\sigma \in \mathcal{U}$ for any $\sigma \in \mathcal{S}_k$. The convexity of \mathcal{U} implies that

$$\frac{1}{k!} \sum_{\sigma \in \mathcal{S}_k} \mathbf{h}^\sigma \in \mathcal{U}.$$

We have,

$$\sum_{\sigma \in \mathcal{S}_k} h_i^\sigma = \begin{cases} (k-1)! \cdot \sum_{j=1}^k h_j & \text{if } i = 1, \dots, k \\ 0 & \text{otherwise,} \end{cases}$$

and $\sum_{j=1}^k h_j = k \cdot \gamma(k)$ by definition. Therefore,

$$\frac{1}{k!} \sum_{\sigma \in \mathcal{S}_k} \mathbf{h}^\sigma = \gamma(k) \cdot \sum_{i=1}^k \mathbf{e}_i \in \mathcal{U}.$$

□

C Proof of Lemma 4

Proof Consider, $\tilde{\mathbf{h}} \in \mathcal{U}$ an optimal solution for the maximization problem in (3.3) for fixed β . We will construct $\mathbf{h}^* \in \mathcal{U}$ another optimal solution of (3.3) that verifies the properties in the lemma. First, denote $I = \{i \mid \tilde{h}_i > \beta\gamma\}$ and $|I| = k$. Since, \mathcal{U} is permutation invariant, we can suppose without loss of generality that $I = \{1, 2, \dots, k\}$. We define,

$$h_i^* = \begin{cases} \gamma(k) & \text{if } i = 1, \dots, k \\ 0 & \text{otherwise.} \end{cases}$$

From Lemma 3, we have $\mathbf{h}^* \in \mathcal{U}$. Moreover,

$$\begin{aligned} \sum_{i=1}^m (\tilde{h}_i - \beta\gamma)^+ &= \sum_{i=1}^k \tilde{h}_i - \beta\gamma k \leq k \cdot \gamma(k) - \beta\gamma k \\ &= \sum_{i=1}^k (\gamma(k) - \beta\gamma) = \sum_{i=1}^k (h_i^* - \beta\gamma) \\ &\leq \sum_{i=1}^k (h_i^* - \beta\gamma)^+ = \sum_{i=1}^m (h_i^* - \beta\gamma)^+ \end{aligned}$$

where the first inequality follows from the definition of the coefficients $\gamma(\cdot)$. Therefore, \mathbf{h}^* and $\tilde{\mathbf{h}}$ have the same objective value in (3.3) and consequently \mathbf{h}^* is also optimal for the maximization problem (3.3). Moreover, from the first inequality, we have $\gamma(k) - \beta\gamma > 0$, i.e., $|\{i \mid h_i^* > \beta\gamma\}| = k$. Therefore, \mathbf{h}^* verifies the properties of the lemma. \square

D Proof of Proposition 4

Proof To prove that $\hat{\mathcal{U}}$ dominates \mathcal{U} , it is sufficient to take \mathbf{h} in the boundaries of \mathcal{U} , i.e.,

$$a \sum_{i=1}^m h_i \sum_{j=1}^m h_j + (1-a) \sum_{i=1}^m h_i^2 = 1, \quad (\text{D.1})$$

and find $\alpha_1, \alpha_2, \dots, \alpha_{m+1}$ nonnegative reals with $\sum_{i=1}^{m+1} \alpha_i = 1$ such that for all $i \in [m]$,

$$h_i \leq \beta(\alpha_i + \gamma\alpha_{m+1}).$$

By taking all h_i equal in (D.1), we get

$$\gamma = \frac{1}{\sqrt{(am^2 + (1-a)m)}}.$$

We choose for $i \in [m]$,

$$\alpha_i = \frac{1}{2} \left((1-a)h_i^2 + ah_i \sum_{j=1}^m h_j \right)$$

and $\alpha_{m+1} = \frac{1}{2}$. First, we have $\sum_{i=1}^{m+1} \alpha_i = 1$ and for all $i \in [m]$,

$$\begin{aligned} \beta(\alpha_i + \gamma\alpha_{m+1}) &= \frac{\beta}{2} \left((1-a)h_i^2 + ah_i \sum_{j=1}^m h_j + \frac{1}{\sqrt{am^2 + (1-a)m}} \right) \\ &\geq \frac{\beta}{2} \left((1-a)h_i^2 + \frac{1}{\sqrt{am^2 + (1-a)m}} + ah_i \right) \\ &\geq \frac{\beta}{2} \left(2 \left(\frac{(1-a)}{\sqrt{am^2 + (1-a)m}} \right)^{\frac{1}{2}} h_i + ah_i \right) = h_i \end{aligned}$$

where the first inequality holds because $\sum_{j=1}^m h_j \geq 1$ which is a direct consequence of $\mathbf{h}^T \Sigma \mathbf{h} = 1$ and $a \leq 1$. The second one follows from AM-GM inequality. Finally, we can verify by case analysis on the values of a that

$$\left(\frac{a}{2} + \frac{(1-a)^{\frac{1}{2}}}{(am^2 + (1-a)m)^{\frac{1}{4}}} \right)^{-1} = O\left(m^{\frac{2}{5}}\right).$$

□

E Proof of Proposition 5

Proof To prove that $\hat{\mathcal{U}}$ dominates \mathcal{U} , it is sufficient to take \mathbf{h} in the boundaries of \mathcal{U} , i.e., $\sum_{i=1}^m h_i = k$ and find $\alpha_1, \alpha_2, \dots, \alpha_{m+1}$ non-negative reals with $\sum_{i=1}^{m+1} \alpha_i = 1$ such that for all $i \in [m]$,

$$h_i \leq \beta \left(\alpha_i + \frac{k}{m} \alpha_{m+1} \right).$$

First case: If $\beta = k$, we choose $\alpha_i = \frac{h_i}{k}$ for $i \in [m]$ and $\alpha_{m+1} = 0$. We have $\sum_{i=1}^{m+1} \alpha_i = 1$ and for all $i \in [m]$,

$$\beta \left(\alpha_i + \frac{k}{m} \alpha_{m+1} \right) = k \frac{h_i}{k} \geq h_i.$$

Second case: If $\beta = \frac{m}{k}$, we choose $\alpha_i = 0$ for $i \in [m]$ and $\alpha_{m+1} = 1$. We have $\sum_{i=1}^{m+1} \alpha_i = 1$ and for all $i \in [m]$,

$$\beta \left(\alpha_i + \frac{k}{m} \alpha_{m+1} \right) = 1 \geq h_i.$$

□

F Proof of Theorem 6

Proof Let us find the order of the left hand side ratio in inequality (5.3). We have,

$$\begin{aligned} \frac{\binom{\sqrt{m}}{m^\epsilon} \cdot \binom{m-m^\epsilon}{\sqrt{m}-m^\epsilon}}{\binom{m}{\sqrt{m}}} &= \frac{(\sqrt{m})! \times (m-m^\epsilon)! \times (m-\sqrt{m})! \times (\sqrt{m})!}{(\sqrt{m}-m^\epsilon)! \times (m^\epsilon)! \times m! \times (\sqrt{m}-m^\epsilon)! \times (m-\sqrt{m})!} \\ &= \left(\frac{(\sqrt{m})!}{(\sqrt{m}-m^\epsilon)!} \right)^2 \cdot \frac{(m-m^\epsilon)!}{(m^\epsilon)! \times m!}. \end{aligned}$$

By Stirling's approximation, we have

$$\begin{aligned}(\sqrt{m})! &= \Theta \left(m^{\frac{1}{4}} \left(\frac{\sqrt{m}}{e} \right)^{\sqrt{m}} \right). \\(\sqrt{m} - m^\epsilon)! &= \Theta \left((\sqrt{m} - m^\epsilon)^{\frac{1}{2}} \left(\frac{\sqrt{m} - m^\epsilon}{e} \right)^{\sqrt{m} - m^\epsilon} \right). \\(m - m^\epsilon)! &= \Theta \left((m - m^\epsilon)^{\frac{1}{2}} \left(\frac{m - m^\epsilon}{e} \right)^{m - m^\epsilon} \right). \\(m)! &= \Theta \left(m^{\frac{1}{2}} \left(\frac{m}{e} \right)^m \right). \\(m^\epsilon)! &= \Theta \left(m^{\frac{1}{2}\epsilon} \left(\frac{m^\epsilon}{e} \right)^{m^\epsilon} \right).\end{aligned}$$

All together,

$$\frac{\binom{\sqrt{m}}{m^\epsilon} \cdot \binom{m - m^\epsilon}{\sqrt{m} - m^\epsilon}}{\binom{m}{\sqrt{m}}} = \Theta \left(\frac{(\sqrt{m})^{2\sqrt{m}} \cdot (m - m^\epsilon)^{(m - m^\epsilon)}}{m^{\frac{1}{2}\epsilon} \cdot (\sqrt{m} - m^\epsilon)^{2(\sqrt{m} - m^\epsilon)} \cdot m^m \cdot m^{\epsilon m^\epsilon}} \right).$$

We have

$$(m - m^\epsilon)^{(m - m^\epsilon)} = \Theta \left(m^{(m - m^\epsilon)} \cdot e^{-m^\epsilon + \frac{m^{2\epsilon}}{m}} \right),$$

and

$$(\sqrt{m} - m^\epsilon)^{2(\sqrt{m} - m^\epsilon)} = \Theta \left((\sqrt{m})^{2(\sqrt{m} - m^\epsilon)} \cdot e^{-2m^\epsilon + 2\frac{m^{2\epsilon}}{\sqrt{m}}} \right),$$

WLOG, we can suppose that $\epsilon < \frac{1}{4}$, therefore

$$\begin{aligned}\frac{\binom{\sqrt{m}}{m^\epsilon} \cdot \binom{m - m^\epsilon}{\sqrt{m} - m^\epsilon}}{\binom{m}{\sqrt{m}}} &= \Theta \left(\frac{e^{m^\epsilon - 2\frac{m^{2\epsilon}}{\sqrt{m}} + \frac{m^{2\epsilon}}{m}}}{m^{\epsilon m^\epsilon + \frac{1}{2}\epsilon}} \right) \\ &= \Theta \left(\frac{e^{m^\epsilon}}{m^{\epsilon m^\epsilon + \frac{1}{2}\epsilon}} \right).\end{aligned}$$

We have,

$$\Theta \left(\frac{Q(m)e^{m^\epsilon}}{m^{\epsilon m^\epsilon + \frac{1}{2}\epsilon}} \right) \geq 1,$$

but the later inequality contradicts

$$\lim_{m \rightarrow \infty} \frac{Q(m)e^{m^\epsilon}}{m^{\epsilon m^\epsilon + \frac{1}{2}\epsilon}} = 0.$$

□