Submitted to Math Programming manuscript No.
(will be inserted by the editor)

Piecewise Static Policies for Two-stage Adjustable Robust Linear Optimization

Omar El Housni · Vineet Goyal

Received: date / Accepted: date

Abstract In this paper, we consider two-stage adjustable robust linear optimization problems under uncertain constraints and study the performance of piecewise static policies. These are a generalization of static policies where we divide the uncertainty set into several pieces and specify a static solution for each piece. We 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. This is quite surprising as piecewise static policies are significantly more general than static policies. The proof is based on a combinatorial argument and the structure of piecewise static policies.

Keywords Robust optimization · Adaptive optimization

1 Introduction

In many real world problems, parameters are uncertain at the optimization phase. Stochastic optimization has been studied extensively in the literature to address the parameter uncertainty (see Dantzing [16], Beale [2], Prekopa [28], Shapiro [29], Shapiro et al. [30]). In stochastic optimization, uncertainty is modeled by probability distributions and the goal is to optimize over an expected objective. By and large, computing the optimal solution is intractable (see Shapiro and Nemirovski [31], Dyer and Stougie [18], Hanasusanto et al. [25]). Moreover, in many cases, we do not have sufficient historical data to estimate a joint multivariate probability distribution.

Robust optimization is another approach that has been studied extensively in the literature. In a robust optimization approach, the uncertain parameters are assumed to belong to a given uncertainty set $\mathcal{U}$ and an adversary selects the uncertain parameters from $\mathcal{U}$. The goal is to optimize the worst case objective

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

(see Ben-Tal and Nemirovski [5], El Ghaoui and Lebret [19], Bertsimas and Sim [13,14], Goldfarb and Iyengar [22]). Computing a static robust solution that is feasible for all scenarios is tractable for large class of robust optimization problems. However, an optimal adjustable (dynamic) solution is significantly more challenging to compute.

In this paper, we consider the following two-stage adjustable robust linear optimization problem, $\Pi_{\mathsf{AR}}(\mathcal{U})$ under uncertain constraint coefficients:

$$\begin{aligned} z_{\mathsf{AR}}(\mathcal{U}) = \max \ & \boldsymbol{c}^T\boldsymbol{x} + \min_{\boldsymbol{B}\in\mathcal{U}} \max_{\boldsymbol{y}(\boldsymbol{B})} \boldsymbol{d}^T\boldsymbol{y}(\boldsymbol{B}) \\ & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{B}) \leq \boldsymbol{h} \\ & \boldsymbol{x} \in \mathbb{R}_+^{n_1},\ \boldsymbol{y}(\boldsymbol{B}) \in \mathbb{R}_+^{n_2}, \end{aligned} \tag{1.1}$$

where $\boldsymbol{A} \in \mathbb{R}^{m\times n_1}$, $\boldsymbol{c} \in \mathbb{R}_+^{n_1}$, $\boldsymbol{d} \in \mathbb{R}_+^{n_2}$, $\boldsymbol{h} \in \mathbb{R}_+^{m}$. Also, $\mathcal{U} \subseteq R_+^{m\times n_2}$ is a full dimensional compact convex *down-monotone* uncertainty set in the non-negative orthant. Following Bertsimas et al. [10], we can assume without loss of generality that $\mathcal{U}$ is *down-monotone* and $n_1 = n_2 = n$ (A set $\mathcal{S} \subseteq \mathbb{R}_+^n$ is *down-monotone* if $\boldsymbol{s} \in \mathcal{S}$, $\boldsymbol{t} \in \mathbb{R}_+^n$ and $\boldsymbol{t} \leq \boldsymbol{s}$ implies $\boldsymbol{t} \in \mathcal{S}$). Note that $\boldsymbol{x}$ represents the first-stage decisions and $\boldsymbol{y}(\boldsymbol{B})$ represents the second-stage decisions after observing the uncertain constraint matrix $\boldsymbol{B} \in \mathcal{U}$.

The above formulation models many interesting applications including revenue management and resource allocation problems with uncertain demand. For instance, in a resource allocation application, the right hand side $\boldsymbol{h}$ can model the fixed resource capacities and the uncertain coefficients in $\boldsymbol{B}$ model the uncertain requirements of resources for demand. The goal is to find an optimal allocation of resources that maximizes the worst case profit (see Wiesemann [33]). When $m = 1$, the above problem reduces to a fractional knapsack problem with uncertain item sizes. The stochastic version of the knapsack problem has been widely studied in the literature (see Dean et al. [17], Goel and Indyk [21], Goyal and Ravi [23]).

In general, it is intractable to compute an adjustable robust solution for (1.1). In fact, Awasthi et al. [1] show that the two-stage adjustable robust problem (1.1) is $\Omega(\log n)$-hard to approximate if the uncertainty set of constraint coefficients belongs to the non-negative orthant. In other words, there is no polynomial time algorithm that approximates the optimal adjustable solution within a factor better than $\log n$. Therefore, the goal is to construct approximate policies with good performance. A static solution approach, where we give a single solution feasible for all scenarios, has been widely studied in the literature. We can formulate the static robust optimization problem $\Pi_{\mathsf{Rob}}(\mathcal{U})$ to approximate (1.1) as follows.

$$\begin{aligned} z_{\mathsf{Rob}}(\mathcal{U}) = \max \ & \boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{d}^T\boldsymbol{y} \\ & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \leq \boldsymbol{h} \ \ \forall \boldsymbol{B} \in \mathcal{U} \\ & \boldsymbol{x} \in \mathbb{R}_+^n,\ \boldsymbol{y} \in \mathbb{R}_+^n. \end{aligned} \tag{1.2}$$

As we mention earlier, an optimal static solution can be computed efficiently for large class of problems (see Bertsimas et al. [7], Ben-Tal et al. [3]). Ben-Tal and Nemirovski [6] show that a static solution is optimal for (1.1) if the uncertainty set is constraint-wise where each constraint is selected independently from a compact convex set $\mathcal{U}_i$ (i.e. $\mathcal{U}$ is a Cartesian product of $\mathcal{U}_i$, $i = 1, \ldots, m$). Bertsimas et al.[10] generalize the result of [6] and show that a static solution is near optimal for several interesting families of $\mathcal{U}$. In particular, they give a tight characterization on the performance of the static solution related to the measure of non-convexity of a transformation of the uncertainty set $\mathcal{U}$. While a static solution provides a good approximation in many cases, it can be as bad as a factor m away from the optimal adjustable solution in general.

The performance of static policies has also been studied for other models. Bertsimas et al. [11] study two-stage and multi-stage adjustable robust linear optimization problems with covering constraints and the uncertainty in the right hand side and relate the performance of static solutions to a measure of symmetry of the uncertainty set. More general policies such as affine policies or piecewise static policies for dynamic optimization problems under uncertainty have been considered in the literature. For instance, Charnes et al. [15] and Garstka and Wets [20] were the first to study affine decision rules in the context of stochastic optimization where recourse decisions are restricted to be affine decisions of the uncertain parameters. Ben-Tal et al. [4] consider affine policies for adjustable robust problems that have been subsequently studied extensively in the literature (see Kuhn et al. [27], Bertsimas et al. [12], Iancu et al. [26], Bertsimas and Goyal [9], Skaf and Boyd [32]).

Piecewise static policies is another solution approach that has been studied in the literature. A piecewise static policy is a generalization of the static policy where the uncertainty set is divided into several pieces and we specify a static policy for each piece. Bertsimas and Caramanis [8] consider a piecewise static solution approach (also referred to as *finite K-adaptability*) where they propose a hierarchy of increasing adaptability that bridges the gap between the static robust formulation, and the fully adaptable formulation. Wiesemann et al. [24] consider a K-adaptable solution approach for two-stage robust integers optimization problems.

1.1 Our contributions

In this paper, we consider the piecewise static solution approach for (1.1). In particular, we consider a piecewise policy with p pieces (or subsets): $\mathcal{U}_1, \ldots, \mathcal{U}_p$ of $\mathcal{U}$ such that

$$\mathcal{U} = \bigcup_{1 \leq i \leq p} \mathcal{U}_i,$$

where each $\mathcal{U}_i$ is convex, compact and down-monotone uncertainty subset. Note that $\mathcal{U}_i$ are not necessarily disjoint. We can formulate the two-stage piecewise robust linear optimization problem $\Pi_{\mathsf{PR}}(\mathcal{U}_1, \ldots, \mathcal{U}_p)$ as follows:

$$\begin{aligned} z_{\mathsf{PR}}(\mathcal{U}_1, \ldots, \mathcal{U}_p) = \max \ & \boldsymbol{c}^T \boldsymbol{x} + \ \min(\boldsymbol{d}^T \boldsymbol{y}_1, \boldsymbol{d}^T \boldsymbol{y}_2, \ldots, \boldsymbol{d}^T \boldsymbol{y}_p) \\ & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}_i \boldsymbol{y}_i \leq \boldsymbol{h} \qquad \forall i \in [p] \ , \ \forall \boldsymbol{B}_i \in \mathcal{U}_i \\ & \boldsymbol{x} \in \mathbb{R}^n_+, \ \boldsymbol{y}_i \in \mathbb{R}^n_+ \ \ \forall i \in [p]. \end{aligned} \tag{1.3}$$

We show that the performance of the optimal piecewise static policy for given pieces is related to the maximum of the measures of non-convexity of transformations of the pieces $\mathcal{U}_i$; thereby extending the bound in [10] for piecewise static policies. Note that if the pieces $\mathcal{U}_i$ are given explicitly, we can efficiently compute an optimal piecewise static policy provide we can solve linear optimization over each $\mathcal{U}_i$ efficiently. However, one of the main challenges in designing a good piecewise static policy, is to construct good pieces of the uncertainty set. 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 general.

Our main contribution in this paper is to show that even if we ignore the computational complexity of computing optimal pieces, surprisingly the performance of piecewise static policies with a polynomial number of pieces is not significantly better than a static policy in general. In particular, we show that there is no piecewise static policy with polynomial number of pieces that gives an approximation bound better than $O(m^{1-\epsilon})$ for any $\epsilon > 0$ for general uncertainty sets $\mathcal{U} \subseteq R_+^{m\times n}$ where the approximation bound for the static policy is m. We prove this by constructing a family of instances of $\mathcal{U}$ for any $\epsilon > 0$, such that the performance of the static policy is m and the performance of any piecewise policy with polynomial number of pieces is $\Omega(m^{1-\epsilon})$. Our proof is based on a combinatorial argument and structural results about piecewise static policies.

Outline. The rest of the paper is organized as follows. We present the preliminaries in Section 2. In Section 3, we present the structural results for piecewise static policies. Finally, we present the lower bound on the performance of piecewise static policies in Section 4.

2 Preliminaries: Static Policies

In this section, we present some preliminaries and definitions for our results. As we mention earlier, Bertsimas et al. [10] give a tight characterization on the performance of a static solution as compared to the optimal adjustable solution for problem (1.1). They relate this performance to the measure of non-convexity of a transformation of the uncertainty set. We first introduce the following definitions.

Definition 1 (Transformation $T(U, \cdot)$). For any $\boldsymbol{h} > \boldsymbol{0}$ and convex compact full-dimensional down-monotone set $\mathcal{U} \subseteq R_+^{m\times n}$, we define the following transformation:

$$T(\mathcal{U}, \boldsymbol{h}) = \{\boldsymbol{B}^T\boldsymbol{\mu} \mid \boldsymbol{h}^T\boldsymbol{\mu} = 1, \boldsymbol{B} \in \mathcal{U}, \boldsymbol{\mu} \geq \boldsymbol{0}\}.$$

Definition 2 (Measure of non-convexity). For any down-monotone compact set $\mathcal{S} \subseteq R_+^n$, the measure of non-convexity $\kappa(\mathcal{S})$ is defined as follows:

$$\kappa(\mathcal{S}) = \min\{\alpha \mid \mathsf{conv}(\mathcal{S}) \subseteq \alpha \cdot \mathcal{S}\}.$$

Definition 3 For any convex compact full-dimensional down-monotone set $\mathcal{U}$, let,

$$\rho(\mathcal{U}) = \max_{\boldsymbol{h}>\boldsymbol{0}} \kappa(T(\mathcal{U},\boldsymbol{h})).$$

For any $\mathcal{U} \subset \mathbb{R}_+^{m\times n}$, Bertsimas et al.[10] give the following characterization of $\mathsf{conv}(T(U,\cdot))$.

Lemma 1 (Bertsimas et al. [10]) *For any* $\boldsymbol{h} > \boldsymbol{0}$,

$$\mathsf{conv}(T(\mathcal{U},\boldsymbol{h})) = \mathsf{conv}\left(\bigcup_{1\le i\le m}\left\{\frac{1}{h_i}\boldsymbol{B}^T\boldsymbol{e}_i \;\middle|\; \boldsymbol{B}\in\mathcal{U}\right\}\right).$$

Consider the following one-stage adjustable robust problem, $\Pi^I_{\mathsf{AR}}(\mathcal{U},\boldsymbol{h})$, corresponding to (1.1).

$$z^I_{\mathsf{AR}}(\mathcal{U},\boldsymbol{h}) = \min_{\boldsymbol{B}\in\mathcal{U}}\max_{\boldsymbol{y}\ge\boldsymbol{0}} \{\boldsymbol{d}^T\boldsymbol{y} \mid \boldsymbol{B}\boldsymbol{y}\le\boldsymbol{h}\}. \tag{2.1}$$

The one-stage problem is related to the separation problem for the two-stage adjustable robust optimization problem (1.1). Similarly, we can consider the following one-stage robust problem, $\Pi^I_{\mathsf{Rob}}(\mathcal{U},\boldsymbol{h})$, corresponding to (1.2).

$$z^I_{\mathsf{Rob}}(\mathcal{U},\boldsymbol{h}) = \max_{\boldsymbol{y}\ge\boldsymbol{0}} \{\boldsymbol{d}^T\boldsymbol{y} \mid \boldsymbol{B}\boldsymbol{y}\le\boldsymbol{h}\ \forall\boldsymbol{B}\in\mathcal{U}\}. \tag{2.2}$$

Bertsimas et al. [10] give the following reformulations of (2.1) and (2.2).

Lemma 2 (Bertsimas et al. [10]) $\Pi^I_{\mathsf{AR}}(\mathcal{U},\boldsymbol{h})$ (2.1) *can be reformulated as*

$$z^I_{\mathsf{AR}}(\mathcal{U},\boldsymbol{h}) = \min\{\lambda \mid \lambda\boldsymbol{b}\ge\boldsymbol{d}, \boldsymbol{b}\in T(\mathcal{U},\boldsymbol{h})\}.$$

Lemma 3 (Bertsimas et al. [10]) $\Pi^I_{\mathsf{Rob}}(\mathcal{U},\boldsymbol{h})$ *can be formulated as*

$$z^I_{\mathsf{Rob}}(\mathcal{U},\boldsymbol{h}) = \min\{\lambda \mid \lambda\boldsymbol{b}\ge\boldsymbol{d},\ \boldsymbol{b}\in\mathsf{conv}(T(\mathcal{U},\boldsymbol{h})\}.$$

Furthermore, they show that

$$z_{\mathsf{Rob}}(\mathcal{U}) \le z_{\mathsf{AR}}(\mathcal{U}) \le \rho(\mathcal{U})\cdot z_{\mathsf{Rob}}(\mathcal{U}),$$

where $\rho(\mathcal{U})$ is the tight bound that characterizes the performance of the static policy. Note that $\rho(\mathcal{U})$ can be as bad as m in general. The worst case instance for $\rho(\mathcal{U})$ is the diagonal uncertainty set

$$\mathcal{U} = \left\{\mathsf{diag}(\boldsymbol{x}) \;\middle|\; \sum_{i=1}^m x_i\le 1,\ \boldsymbol{x}\ge\boldsymbol{0}\right\}. \tag{2.3}$$

For this example of uncertainty set we have, $z_{\mathsf{AR}}(\mathcal{U}) = m\cdot z_{\mathsf{Rob}}(\mathcal{U})$. We refer the reader to Bertsimas et al. [10] for more details.

3 Structural results on piecewise static policy

In this section, we introduce the piecewise static policies for the two-stage adjustable robust optimization problem (1.1) and study the structural properties and performance of these policies. We first introduce the following definition.

Definition 4 (Convex cover) Let $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_p$ subsets of $\mathcal{U}$ such that $\mathcal{U}_i$ is convex, compact and down-monotone set. We say that $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_p$ is a convex cover of $\mathcal{U}$ if $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2 \ldots \cup \mathcal{U}_p$.

Note that different pieces are not necessarily disjoint. We only require that the union of pieces covers $\mathcal{U}$.

3.1 Performance of piecewise static policy

Let $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2 \ldots \cup \mathcal{U}_p$ be a convex cover of $\mathcal{U}$. We relate the performance of the optimal piecewise static solution to the maximum of the measures of non-convexity of the transformations $T(\mathcal{U}_i, \cdot)$. Consider the following reformulation of the two-stage piecewise static robust linear optimization problem (1.3).

$$
\begin{aligned}
z_{\mathsf{PR}}(\mathcal{U}_1, \ldots, \mathcal{U}_p) = \max \;\; & \boldsymbol{c}^T \boldsymbol{x} + \; z \\
& \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}_i \boldsymbol{y}_i \leq \boldsymbol{h} \qquad \forall i \in [p] \;,\; \forall \boldsymbol{B}_i \in \mathcal{U}_i \\
& z \leq \boldsymbol{d}^T \boldsymbol{y}_i \qquad \forall i \in [p] \\
& \boldsymbol{x} \in \mathbb{R}^n_+, \; \boldsymbol{y}_i \in \mathbb{R}^n_+ \; \forall i \in [p], \; z \in \mathbb{R}.
\end{aligned}
\tag{3.1}
$$

We can compute the solution of this problem efficiently if the number of pieces is small and linear optimization is efficient over each piece.

Let $(\boldsymbol{x}^*, (\boldsymbol{y}_1^*, \boldsymbol{y}_2^*, \ldots, \boldsymbol{y}_p^*))$ be an optimal solution of (3.1). Then $(\boldsymbol{x}^*, \boldsymbol{y(B)})$, where $\boldsymbol{y(B)} = \boldsymbol{y}_i^*$ if $\boldsymbol{B} \in \mathcal{U}_i$, is a feasible solution for the adjustable problem (1.1). Therefore,

$$
z_{\mathsf{PR}}(\mathcal{U}_1, \ldots, \mathcal{U}_p) \leq z_{\mathsf{AR}}(\mathcal{U}). \tag{3.2}
$$

To compute an upper bound for $z_{\mathsf{AR}}(\mathcal{U})$ in terms of $z_{\mathsf{PR}}(\mathcal{U}_1, \ldots, \mathcal{U}_p)$, consider the following one stage piecewise static problem $\Pi^I_{\mathsf{PR}}((\mathcal{U}_1, \ldots, \mathcal{U}_p), \boldsymbol{h})$:

$$
z^I_{\mathsf{PR}}((\mathcal{U}_1, .., \mathcal{U}_p), \boldsymbol{h}) = \max_{\boldsymbol{y}_i \geq \boldsymbol{0}} \left\{ z \;\middle|\; \boldsymbol{B}_i \boldsymbol{y}_i \leq \boldsymbol{h} \; \forall \boldsymbol{B}_i \in \mathcal{U}_i, \; z \leq \boldsymbol{d}^T \boldsymbol{y}_i, \; \forall i \in [p] \right\}
\tag{3.3}
$$

Lemma 4 *For the one stage piecewise static problem $\Pi^I_{\mathsf{PR}}((\mathcal{U}_1, \ldots, \mathcal{U}_p), \boldsymbol{h})$,*

$$
z^I_{\mathsf{PR}}((\mathcal{U}_1, \ldots, \mathcal{U}_p), \boldsymbol{h}) = \min_{1 \leq i \leq p} z^I_{\mathsf{Rob}}(U_i, \boldsymbol{h}).
$$

Lemma 4 follows directly from (3.3). The following theorem relates the performance of a piecewise static solution to the measures of non-convexity of $T(\mathcal{U}_i, \boldsymbol{h})$.

Theorem 1 *For any convex cover of* $\mathcal{U}$ *such that* $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2 \ldots \cup \mathcal{U}_p$, *we have,*

$$z_{\mathsf{AR}}(\mathcal{U}) \leq \max(\rho(\mathcal{U}_1), \ldots, \rho(\mathcal{U}_p)) \cdot z_{\mathsf{PR}}(\mathcal{U}_1, \ldots, \mathcal{U}_p).$$

Furthermore, the bound is tight.

Proof Denote $\hat{\lambda}_\ell$, $\hat{\boldsymbol{b}}_\ell \in \mathsf{conv}(T(\mathcal{U}_\ell, \boldsymbol{h}))$ the solutions of the one stage piecewise static problem $\Pi^I_{\mathsf{PR}}((\mathcal{U}_1, \ldots, \mathcal{U}_p), \boldsymbol{h})$ under the formulations of Lemma 4 and Lemma 3, where $\ell \in [p]$. We have $z^I_{\mathsf{PR}}((\mathcal{U}_1, \ldots, \mathcal{U}_p), \boldsymbol{h}) = \hat{\lambda}_\ell$ and $\hat{\lambda}_\ell \hat{\boldsymbol{b}}_\ell \geq \boldsymbol{d}$, i.e.

$$\kappa_\ell \hat{\lambda}_\ell \cdot \frac{\hat{\boldsymbol{b}}_\ell}{\kappa_\ell} \geq \boldsymbol{d}$$

where $\kappa_\ell = \kappa(T(\mathcal{U}_\ell, \boldsymbol{h}))$. Since,

$$\frac{\hat{\boldsymbol{b}}_\ell}{\kappa_\ell} \in T(\mathcal{U}_\ell, \boldsymbol{h}) \subseteq T(\mathcal{U}, \boldsymbol{h}),$$

then $(\kappa_\ell \hat{\lambda}_\ell)$ is a feasible solution for $\Pi^I_{\mathsf{AR}}(\mathcal{U}, \boldsymbol{h})$ under the formulation of Lemma 2, i.e. $\kappa_\ell \hat{\lambda}_\ell \geq z^I_{\mathsf{AR}}(\mathcal{U}, \boldsymbol{h})$.
Moreover, we know that $\max(\rho(\mathcal{U}_1), \ldots, \rho(\mathcal{U}_p)) \geq \kappa_\ell$. Then,

$$\max(\rho(\mathcal{U}_1), \ldots, \rho(\mathcal{U}_p)) \cdot z^I_{\mathsf{PR}}((\mathcal{U}_1, \ldots, \mathcal{U}_p), \boldsymbol{h}) \geq z^I_{\mathsf{AR}}(\mathcal{U}, \boldsymbol{h}).$$

Therefore,

$$\begin{aligned} z_{\mathsf{AR}}(\mathcal{U}, \boldsymbol{h}) &= \boldsymbol{c}^T \boldsymbol{x}^* + z^I_{\mathsf{AR}}(\mathcal{U}, \boldsymbol{h} - \boldsymbol{A}\boldsymbol{x}^*) \\ &\leq \boldsymbol{c}^T \boldsymbol{x}^* + \max(\rho(\mathcal{U}_1), \ldots, \rho(\mathcal{U}_p)) \cdot z^I_{\mathsf{PR}}((\mathcal{U}_1, \ldots, \mathcal{U}_p), \boldsymbol{h} - \boldsymbol{A}\boldsymbol{x}^*) \\ &\leq \max(\rho(\mathcal{U}_1), \ldots, \rho(\mathcal{U}_p)) \cdot (\boldsymbol{c}^T \boldsymbol{x}^* + z^I_{\mathsf{PR}}((\mathcal{U}_1, \ldots, \mathcal{U}_p), \boldsymbol{h} - \boldsymbol{A}\boldsymbol{x}^*)) \\ &\leq \max(\rho(\mathcal{U}_1), \ldots, \rho(\mathcal{U}_p)) \cdot z_{\mathsf{PR}}(\mathcal{U}_1, \ldots, \mathcal{U}_p). \end{aligned}$$

The last inequality follows from the definition (3.3) of the one stage piecewise static problem. The tightness of the bound follows from the tightness of the bound for static policies [10]. □

3.2 Examples of Piecewise Static Policies

We present several examples to illustrate the performance bound for piecewise static policies. In particular, we consider the diagonal uncertainty set defined in (2.3) for which the performance of static policies is the worst possible as compared to the optimal fully adjustable solution. We first show that without loss of generality, we can consider pieces of the following form for any convex cover of $\mathcal{U}$ (2.3).

$$\mathcal{V}(\tau_1, \tau_2, \ldots, \tau_m) = \left\{ \mathsf{diag}(\boldsymbol{x}) \;\middle|\; \sum_{j=1}^{m} x_j \leq 1,\ 0 \leq x_j \leq \tau_j \ \forall j \in [m] \right\}. \quad (3.4)$$

In particular, we have the following structural lemma.

Lemma 5 (Structure of Piecewise static policies) *Let $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2 \ldots \cup \mathcal{U}_p$ a convex cover of the diagonal uncertainty set* (2.3). *For all $i \in [p]$ we define, $\mathcal{V}_i = \mathcal{V}(\tau_{i1}, \tau_{i2}, \ldots, \tau_{im})$, where for all $i \in [p]$ and $j \in [m]$,*

$$\tau_{ij} = \max_{\mathsf{diag}(\boldsymbol{x}) \in \mathcal{U}_i} \boldsymbol{e}_j^T \boldsymbol{x}.$$

Then, $\forall i \in [p]$, $\mathcal{U}_i \subseteq \mathcal{V}_i \subseteq \mathcal{U}$ and $\kappa(T(\mathcal{V}_i, \boldsymbol{h})) \leq \kappa(T(\mathcal{U}_i, \boldsymbol{h}))$.

Proof Let $i \in [p]$. We have $\forall$ $\mathsf{diag}(\boldsymbol{x}) \in \mathcal{U}_i$, $x_j \leq \tau_{ij}$ for $j = 1, \ldots, m$. Then, $\mathcal{U}_i \subseteq \mathcal{V}_i \subseteq \mathcal{U}$. Now, we will show that for all $i \in [p]$, $\mathsf{conv}(T(\mathcal{U}_i, \boldsymbol{h})) = \mathsf{conv}(T(\mathcal{V}_i, \boldsymbol{h}))$. First, since $\mathcal{U}_i \subseteq \mathcal{V}_i$, clearly, $\mathsf{conv}(T(\mathcal{U}_i, \boldsymbol{h})) \subseteq \mathsf{conv}(T(\mathcal{V}_i, \boldsymbol{h}))$. Consider any $\boldsymbol{b} \in T(\mathcal{V}_i, \boldsymbol{h})$. Then,

$$\boldsymbol{b} = \mathsf{diag}(\boldsymbol{x})^T \boldsymbol{\mu},$$

where $\sum_{k=1}^m \mu_k h_k = 1$ and $\mathsf{diag}(\boldsymbol{x}) \in \mathcal{V}_i$. Therefore,

$$\boldsymbol{b} = \sum_{k=1}^{m} \mu_k h_k \cdot \frac{x_k}{h_k} \boldsymbol{e}_k$$

For all $k \in [m]$, we have $x_k \leq \tau_{ik}$ and we know that $\mathcal{U}_i$ is *down-monotone*. Therefore, $x_k \boldsymbol{e}_k \in \mathcal{U}_i$ and $\frac{x_k}{h_k} \boldsymbol{e}_k \in T(\mathcal{U}_i, \boldsymbol{h})$. Hence $\boldsymbol{b} \in \mathsf{conv}(T(\mathcal{U}_i, \boldsymbol{h}))$ and $\mathsf{conv}(T(\mathcal{V}_i, \boldsymbol{h})) \subseteq \mathsf{conv}(T(\mathcal{U}_i, \boldsymbol{h}))$. Therefore,

$$\begin{aligned}\mathsf{conv}(T(\mathcal{V}_i, \boldsymbol{h})) = \mathsf{conv}(T(\mathcal{U}_i, \boldsymbol{h})) &\subseteq \kappa(T(\mathcal{U}_i, \boldsymbol{h})) \cdot T(U_i, h) \\ &\subseteq \kappa(T(\mathcal{U}_i, \boldsymbol{h})) \cdot T(\mathcal{V}_i, \boldsymbol{h}),\end{aligned}$$

which implies $\kappa(T(\mathcal{V}_i, \boldsymbol{h})) \leq \kappa(T(\mathcal{U}_i, \boldsymbol{h}))$. □

In the following lemma, we show that we can compute the measure of non-convexity of $T(\mathcal{V}(\tau_1, \tau_2, \ldots, \tau_m), \boldsymbol{h})$ where $\mathcal{V}(\tau_1, \tau_2, \ldots, \tau_m)$ is defined in (3.4).

Lemma 6 *Let,*

$$\mathcal{U} = \mathcal{V}(\tau_1, \tau_2, \ldots, \tau_m)$$

where $\mathcal{V}(\tau_1, \tau_2, \ldots, \tau_m)$ is defined in (3.4) *such that $\forall i \in [m]$, $0 \leq \tau_i \leq 1$ and $\sum_{i=1}^m \tau_i \geq 1$. Then for all $\boldsymbol{h} > \boldsymbol{0}$,*

$$\kappa(T(\mathcal{U}, \boldsymbol{h})) = \sum_{i=1}^{m} \tau_i.$$

The proof of Lemma 6 is presented in Appendix A. We now present two examples of convex covers of the diagonal uncertainty set $\mathcal{U}$ (2.3) and give the performance of the corresponding piecewise static policy for each example.

Example 1. For all $j = 1, \ldots, m$ let,

$$\mathcal{U}_j = \left\{ \mathsf{diag}(\boldsymbol{x}) \;\middle|\; \sum_{i=1}^{m} x_i \leq 1,\ 0 \leq x_j \leq \frac{1}{m} \right\}.$$

Note that $\bigcup_{1 \leq j \leq m} \mathcal{U}_j$ is a convex cover of $\mathcal{U}$ with m number of pieces. From Lemma 6, we have the following.

Proposition 1 *For the cover defined in Example 1, the performance of piecewise static policy is*

$$\rho = m - 1 + \frac{1}{m}.$$

Example 2. Let $\mathcal{S}_m$ be the set of permutations in $\{1, 2, \ldots, m\}$ and let $\boldsymbol{\tau} = (1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{m})$.
For all $\sigma \in \mathcal{S}_m$ let,

$$\mathcal{U}_\sigma = \left\{ \mathsf{diag}(\boldsymbol{x}) \;\middle|\; 0 \leq x_i \leq \tau_{\sigma(i)} \; \forall i \in [m] \, , \; \sum_{i=1}^{m} x_i \leq 1 \right\},$$

Note that $\underset{\sigma \in \mathcal{S}_m}{\cup} U_\sigma$ is a convex cover of $\mathcal{U}$ with $m!$ number of pieces. From Lemma 6, we have the following.

Proposition 2 *For the cover defined in Example 2, the performance of piecewise static policy is*

$$\rho = \sum_{i=1}^{m} \frac{1}{i} = O(\log(m)).$$

We would like to note that for the cover in Example 1, the number of pieces is polynomial and the performance bound for the piecewise static policy is $\Omega(m)$ which is the same order as the approximation bound for static policies. For Example 2, the performance bound for the piecewise static policy is $O(\log m)$ which is significantly better. However the number of pieces is exponential. Since it is difficult to compute a piecewise static policy with exponentially many pieces, it motivates us to consider the problem of finding piecewise static policies with a polynomial number of pieces that have a significantly better performance than the static policy.

4 Lower bound for polynomial pieces

In this section, we show that, surprisingly there is no piecewise static policy with polynomial number of pieces that gives an approximation bound significantly better than the static policies in general. In particular, we consider the diagonal uncertainty set (2.3). Bertsimas et al. [10] present family of instances where $z_{\mathsf{AR}}(\mathcal{U}) = m \cdot z_{\mathsf{Rob}}(\mathcal{U})$ for the uncertainty set (2.3). We show that for any fixed $\epsilon > 0$, there is no piecewise static policy with polynomial number of pieces with approximation bound as $O(m^{1-\epsilon})$. Our proof is based on a combinatorial argument that exploits the structural result for piecewise policies for (2.3) derived in the previous section. We have the following theorem.

Theorem 2 (Main result) *For any given $0 < \epsilon < 1$ and $k \in \mathbb{N}$, there are instances of uncertainty set $\mathcal{U} \subset \mathbb{R}_+^{m \times n}$ with sufficiently large m such that for any convex cover $(\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_p)$ of $\mathcal{U}$ with $p \leq (\max(m, n))^k$ pieces,*

$$\max(\rho(\mathcal{U}_1), \ldots, \rho(\mathcal{U}_p)) > m^{1-\epsilon}.$$

Proof Consider the diagonal uncertainty set $\mathcal{U} \subset \mathbb{R}_+^{m\times m}$ defined in (2.3) for m sufficiently large. Consider $(\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_p)$ a convex cover of $\mathcal{U}$ such that $p \le m^k$. We can assume without loss of generality $p = m^k$. Suppose for the sake of contradiction,

$$\max(\rho(\mathcal{U}_1), \ldots, \rho(\mathcal{U}_p)) \le m^{1-\epsilon}. \tag{4.1}$$

From Lemma 5, it is sufficient to consider $\mathcal{U}_i$ of the following form for all $i \in [p]$:

$$\mathcal{U}_i = \left\{ \mathsf{diag}(\boldsymbol{x}) \;\middle|\; \sum_{i=1}^m x_i \le 1,\ 0 \le x_j \le \tau_{ij} \ \ \forall j \in [m] \right\}$$

From Lemma 6, for all $i \in [p]$, $\forall \boldsymbol{h} > \mathbf{0}$,

$$\kappa(T(\mathcal{U}_i, \boldsymbol{h})) = \sum_{j=1}^m \tau_{ij} \le m^{1-\epsilon}$$

where the last inequality follows from the assumption (4.1). Let

$$\beta = \left\lfloor \frac{1}{\epsilon} \right\rfloor.$$

We define the following discrete set

$$\mathcal{W} = \left\{ \mathsf{diag}\left(\frac{a_1}{\gamma}, \ldots, \frac{a_m}{\gamma} \right) \;\middle|\; \sum_{i=1}^m a_i = \gamma,\ a_i \in \{0,1\}, \forall i \in [m] \right\},$$

where $\gamma = \beta k + k + 1$. Note that $\mathcal{W}$ is a discrete subset of $\mathcal{U}$ with cardinality

$$|\mathcal{W}| = \binom{m}{\gamma} = \binom{m}{\beta k + k + 1} = \Theta\left(m^{\beta k + k + 1} \right).$$

We have

$$\mathcal{W} \subseteq \mathcal{U} = \bigcup_{1 \le i \le p} \mathcal{U}_i.$$

Hence there exists $1 \le \ell \le m^k$ such that $\mathcal{U}_\ell$ contains at least $\frac{|\mathcal{W}|}{m^k}$ elements of $\mathcal{W}$. In particular, there exists $\hat{\mathcal{W}} \subseteq \mathcal{W}$ such that $\hat{\mathcal{W}} \subseteq \mathcal{U}_\ell$ and

$$|\hat{\mathcal{W}}| \ge \frac{|\mathcal{W}|}{m^k} = \Theta(m^{\beta k + 1}). \tag{4.2}$$

Then $\forall j \in [m]$ and $\forall \boldsymbol{a} \in \hat{\mathcal{W}}$,

$$\frac{a_j}{\beta k + k + 1} \le \tau_{\ell j}.$$

Therefore, $\forall j \in [m]$,

$$\frac{\max_{\boldsymbol{a} \in \hat{\mathcal{W}}} \boldsymbol{e}_j^T \boldsymbol{a}}{\beta k + k + 1} \le \tau_{\ell j},$$

which implies

$$\begin{aligned}\sum_{j=1}^{m} \max_{\boldsymbol{a}\in\hat{\mathcal{W}}} \boldsymbol{e}_j^T \boldsymbol{a} &\leq (\beta k + k + 1) \sum_{j=1}^{m} \tau_{\ell j} \\ &\leq (\beta k + k + 1) m^{1-\epsilon} \\ &< (\beta k + k + 1) m^{\frac{\beta}{\beta+1}},\end{aligned}$$

where the last inequality follows from $\frac{1}{\beta+1} < \epsilon$. Denote $t = (\beta k + k + 1) m^{\frac{\beta}{\beta+1}}$ and

$$\mathcal{S} = \{ j \in [m] \mid \exists \boldsymbol{a} \in \hat{\mathcal{W}},\ a_j = 1 \}.$$

Then, $|\mathcal{S}| \leq \lfloor t \rfloor$. We have,

$$\hat{\mathcal{W}} \subseteq \left\{ \mathsf{diag}\left(\frac{a_1}{\gamma}, \ldots, \frac{a_m}{\gamma} \right) \;\middle|\; \sum_{i\in\mathcal{S}} a_i = \gamma,\ a_i \in \{0,1\}, \forall i \in [m] \right\}.$$

Therefore,

$$|\hat{\mathcal{W}}| \leq \binom{|\mathcal{S}|}{\gamma} = \Theta\left(|\mathcal{S}|^\gamma \right).$$

We have,

$$\begin{aligned}|\mathcal{S}|^\gamma = |\mathcal{S}|^{\beta k + k + 1} &\leq \lfloor t \rfloor^{\beta k + k + 1} \\ &\leq t^{\beta k + k + 1} \\ &= \left((\beta k + k + 1) m^{\frac{\beta}{\beta+1}} \right)^{(\beta k + k + 1)} \\ &= (\beta k + k + 1)^{(\beta k + k + 1)} \cdot m^{\beta k + \frac{\beta}{\beta+1}}\end{aligned}$$

Then,

$$|\hat{\mathcal{W}}| \leq \Theta\left(m^{\beta k + \frac{\beta}{\beta+1}} \right).$$

On the other hand, $|\hat{\mathcal{W}}| \geq \Theta\left(m^{\beta k + 1} \right)$ (4.2) which is a contradiction for m sufficiently large. □

The above theorem implies that if we restrict to piecewise policies with a polynomial number of pieces, we can not get significantly better policies than static in general. This is quite surprising since piecewise static policies are more general than a single static solution.

5 Conclusions

In this paper, we consider piecewise static policies to approximate two-stage adjustable linear optimization problems under uncertainty. We relate the performance of the piecewise static policy to the measure of non-convexity of a transformation of the uncertainty pieces. We show that there is no piecewise policy with polynomially many pieces that performs significantly better than a static solution in general. This is quite surprising as piecewise static policy is a significant generalization of the static policy but still does not give a better approximation for the adjustable robust problem in general when restricted to a polynomial number of pieces. We would like to note that although we show that the performance of a piecewise static policy with polynomially many pieces is similar to the static policy in the worst-case, the piecewise static policy can be better than static in many cases. It is an interesting open question to study design of good piecewise policies with a given number of pieces.

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

References

1. P. Awasthi, V. Goyal, and B. Y. Lu. On the adaptivity gap in two-stage robust linear optimization under uncertain constraints. 2015.
2. E. M. Beale. On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 173–184, 1955.
3. A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University Press, 2009.
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. A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.
6. A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations research letters*, 25(1):1–13, 1999.
7. D. Bertsimas, D. B. 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 V. Goyal. On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical programming*, 134(2):491–531, 2012.
10. D. Bertsimas, V. Goyal, and B. Y. Lu. A tight characterization of the performance of static solutions in two-stage adjustable robust linear optimization. *Mathematical Programming*, pages 1–39, 2014.
11. D. Bertsimas, V. Goyal, and X. A. 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. A. Iancu, and P. A. Parrilo. Optimality of affine policies in multistage robust optimization. *Mathematics of Operations Research*, 35(2):363–394, 2010.
13. D. Bertsimas and M. Sim. Robust discrete optimization and network flows. *Mathematical programming*, 98(1):49–71, 2003.
14. D. Bertsimas and M. Sim. The price of robustness. *Operations research*, 52(1):35–53, 2004.

15. A. Charnes, W. W. Cooper, and G. H. Symonds. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4(3):235–263, 1958.
16. G. B. Dantzig. Linear programming under uncertainty. *Management science*, 1(3-4):197–206, 1955.
17. B. C. Dean, M. X. Goemans, and J. Vondrák. Adaptivity and approximation for stochastic packing problems. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 395–404. Society for Industrial and Applied Mathematics, 2005.
18. M. Dyer and L. Stougie. Computational complexity of stochastic programming problems. *Mathematical Programming*, 106(3):423–432, 2006.
19. L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18(4):1035–1064, 1997.
20. S. J. Garstka and R. J.-B. Wets. On decision rules in stochastic programming. *Mathematical Programming*, 7(1):117–143, 1974.
21. A. Goel and P. Indyk. Stochastic load balancing and related problems. In *Foundations of Computer Science, 1999. 40th Annual Symposium on*, pages 579–586. IEEE, 1999.
22. D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.
23. V. Goyal and R. Ravi. A ptas for the chance-constrained knapsack problem with random item sizes. *Operations Research Letters*, 38(3):161–164, 2010.
24. G. A. Hanasusanto, D. Kuhn, and W. Wiesemann. K-adaptability in two-stage robust binary programming, 2014.
25. G. A. Hanasusanto, D. Kuhn, and W. Wiesemann. A comment on" computational complexity of stochastic programming problems". 2015.
26. D. A. Iancu, M. Sharma, and M. Sviridenko. Supermodularity and affine policies in dynamic robust optimization. *Operations Research*, 61(4):941–956, 2013.
27. D. Kuhn, W. Wiesemann, and A. Georghiou. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*, 130(1):177–209, 2011.
28. A. Prékopa. Stochastic programming, volume 324 of mathematics and its applications, 1995.
29. A. Shapiro. Stochastic programming approach to optimization under uncertainty. *Mathematical Programming*, 112(1):183–220, 2008.
30. A. Shapiro, D. Dentcheva, et al. *Lectures on stochastic programming: modeling and theory*, volume 16. SIAM, 2014.
31. A. Shapiro and A. Nemirovski. On complexity of stochastic programming problems. In *Continuous optimization*, pages 111–146. Springer, 2005.
32. J. Skaf and S. P. Boyd. Design of affine controllers via convex optimization. *Automatic Control, IEEE Transactions on*, 55(11):2476–2487, 2010.
33. W. Wiesemann, D. Kuhn, and B. Rustem. Robust resource allocations in temporal networks. *Mathematical programming*, 135(1-2):437–471, 2012.

A Proof of Lemma 6

First, note that for $h > 0$,

$$T(\mathcal{U}, \boldsymbol{h}) = \left\{ \left(\frac{y_1}{h_1}, \frac{y_2}{h_2}, \ldots, \frac{y_m}{h_m} \right) \;\middle|\; (y_1, y_2, \ldots, y_m) \in T(\mathcal{U}, \boldsymbol{e}) \right\}.$$

Then we can easily prove that $\kappa(T(\mathcal{U}, \boldsymbol{h})) = \kappa(T(\mathcal{U}, \boldsymbol{e}))$. In fact, let $\boldsymbol{x} \in \mathsf{conv}(T(\mathcal{U}, \boldsymbol{h}))$. Then,
$\sum_{i=1}^m x_i h_i \boldsymbol{e}_i \in \mathsf{conv}(T(\mathcal{U}, \boldsymbol{e}))$. Therefore,

$$\frac{1}{\kappa(T(\mathcal{U}, \boldsymbol{e}))} \cdot \left(\sum_{i=1}^m x_i h_i \boldsymbol{e}_i \right) \in T(\mathcal{U}, \boldsymbol{e}).$$

Then,

$$\frac{1}{\kappa T((\mathcal{U},\boldsymbol{e}))}\cdot \boldsymbol{x}\in T(\mathcal{U},\boldsymbol{h}),$$

which implies,

$$\mathsf{conv}(T(\mathcal{U},\boldsymbol{h}))\subseteq \kappa(T(\mathcal{U},\boldsymbol{e}))\cdot T(\mathcal{U},\boldsymbol{h}),$$

and finally $\kappa(T(\mathcal{U},\boldsymbol{h}))\leq \kappa(T(\mathcal{U},\boldsymbol{e}))$. Similarly, we also have $\kappa(T(\mathcal{U},\boldsymbol{h}))\geq \kappa(T(\mathcal{U},\boldsymbol{e}))$. Now, it's sufficient to show that $\kappa(T(\mathcal{U},\boldsymbol{e}))=\sum_{i=1}^m \tau_i$. Let first show that

$$\mathsf{conv}(T(\mathcal{U},\boldsymbol{e}))=\left\{(x_1,x_2,\ldots,x_m)\in[0,1]^m \;\middle|\; \sum_{i=1}^m \frac{x_i}{\tau_i}\leq 1\right\}. \tag{A.1}$$

Let $\boldsymbol{x}\in \mathsf{conv}(T(\mathcal{U},\boldsymbol{e}))$. From Lemma 1, we have $\boldsymbol{x}=\sum_{i=1}^m \lambda_i a_i \boldsymbol{e}_i$, where $\sum_{i=1}^m \lambda_i=1$, $\lambda_i\in[0,1]$ and $0\leq a_i\leq \tau_i,\ \forall i\in[m]$. We have,

$$\sum_{i=1}^m \frac{x_i}{\tau_i}=\sum_{i=1}^m \lambda_i\cdot\frac{a_i}{\tau_i}\leq \sum_{i=1}^m \lambda_i=1.$$

Conversely, let $\boldsymbol{x}\in\mathbb{R}_+^m$ such that,

$$\sum_{i=1}^m \frac{x_i}{\tau_i}\leq 1.$$

We have

$$\boldsymbol{x}=\sum_{j=1}^m \lambda_j a_j \boldsymbol{e}_j,$$

where for all $j\in[m]$,

$$\lambda_j=\frac{\frac{x_j}{\tau_j}}{\sum_{i=1}^m \frac{x_i}{\tau_i}} \qquad \text{and} \qquad a_j=\tau_j\sum_{i=1}^m \frac{x_i}{\tau_i}.$$

We have $\sum_{j=1}^m \lambda_j=1$ and $a_j\leq \tau_j\ \forall j\in[m]$. Then, $\boldsymbol{x}\in\mathsf{conv}(T(\mathcal{U},\boldsymbol{e}))$.

Now, we would like to find a lower bound for $\kappa(T(\mathcal{U},\boldsymbol{e}))$. Let $\alpha\geq 1$ such that $\mathsf{conv}(T(\mathcal{U},\boldsymbol{e}))\subseteq\alpha\cdot T(\mathcal{U},\boldsymbol{e})$. From (A.1), we have

$$\left(\frac{\tau_1^2}{\sum_{i=1}^m \tau_i},\frac{\tau_2^2}{\sum_{i=1}^m \tau_i},\ldots,\frac{\tau_m^2}{\sum_{i=1}^m \tau_i}\right)\in\mathsf{conv}(T(\mathcal{U},\boldsymbol{e}))$$

Then, there exists $\mathsf{diag}(\boldsymbol{x})\in\mathcal{U}$ and $\boldsymbol{\mu}\in\mathbb{R}_+^m$, $\sum_{i=1}^m \mu_i=1$, such that

$$\left(\frac{\tau_1^2}{\sum_{i=1}^m \tau_i},\frac{\tau_2^2}{\sum_{i=1}^m \tau_i},\ldots,\frac{\tau_m^2}{\sum_{i=1}^m \tau_i}\right)=\alpha\cdot\mathsf{diag}(\boldsymbol{x})^T\boldsymbol{\mu},$$

i.e. $\forall 1\leq i\leq m$,

$$\frac{\tau_i^2}{\sum_{j=1}^m \tau_j}=\alpha\mu_i x_i$$

From Cauchy-Shwartz inequality we have,

$$\sum_{i=1}^m \frac{\tau_i^2}{\mu_i}\geq\left(\sum_{i=1}^m \tau_i\right)^2,$$

Then,

$$\alpha\left(\sum_{i=1}^m \tau_i\right)\left(\sum_{i=1}^m x_i\right)\geq\left(\sum_{i=1}^m \tau_i\right)^2,$$

i.e.

$$\alpha\left(\sum_{i=1}^{m} x_i\right) \geq \left(\sum_{i=1}^{m} \tau_i\right),$$

therefore,

$$\alpha \geq \sum_{i=1}^{m} \tau_i,$$

where the last inequality follows from $\sum_{i=1}^{m} x_i \leq 1$. To finish our proof we show that,

$$\mathsf{conv}(T(\mathcal{U}, \boldsymbol{e})) \subseteq \left(\sum_{i=1}^{m} \tau_i\right) \cdot T(\mathcal{U}, \boldsymbol{e}).$$

Let $\boldsymbol{x} \in \mathsf{conv}(T(\mathcal{U}, \boldsymbol{e}))$, we have from (A.1),

$$\sum_{i=1}^{m} \frac{x_i}{\tau_i} \leq 1.$$

For all $1 \leq j \leq m$, let define,

$$\mu_j = \frac{\frac{x_j}{\tau_j}}{\sum_{i=1}^{m} \frac{x_i}{\tau_i}} \qquad \text{and} \qquad b_j = \tau_j \frac{\sum_{i=1}^{m} \frac{x_i}{\tau_i}}{\sum_{i=1}^{m} \tau_i}.$$

Then

$$\boldsymbol{x} = \left(\sum_{i=1}^{m} \tau_i\right) \cdot \mathsf{diag}(\boldsymbol{b})^T \boldsymbol{\mu}$$

We have $\forall j \in [m]$,

$$b_j \leq \frac{\tau_j}{\sum_{i=1}^{m} \tau_i} \leq \tau_j$$

where the second inequality holds because $\sum_{i=1}^{m} \tau_i \geq 1$. Furthermore,

$$\sum_{j=1}^{m} b_j = \sum_{i=1}^{m} \frac{x_i}{\tau_i} \leq 1.$$

Therefore, $\mathsf{diag}(\boldsymbol{b}) \in \mathcal{U}$. Since $\sum_{j=1}^{n} \mu_j = 1$, $\mathsf{diag}(\boldsymbol{b})^T \boldsymbol{\mu} \in T(\mathcal{U}, \boldsymbol{e})$. We conclude that

$$\boldsymbol{x} \in \left(\sum_{i=1}^{m} \tau_i\right) \cdot T(\mathcal{U}, \boldsymbol{e}).$$

□