

On the power and limitations of affine policies in two-stage adaptive optimization

Dimitris Bertsimas · Vineet Goyal

Received: 31 October 2009 / Accepted: 17 January 2011
© Springer and Mathematical Optimization Society 2011

Abstract We consider a two-stage adaptive linear optimization problem under right hand side uncertainty with a min–max objective and give a sharp characterization of the power and limitations of affine policies (where the second stage solution is an affine function of the right hand side uncertainty). In particular, we show that the worst-case cost of an optimal affine policy can be $\Omega(m^{1/2-\delta})$ times the worst-case cost of an optimal fully-adaptable solution for any $\delta > 0$, where m is the number of linear constraints. We also show that the worst-case cost of the best affine policy is $O(\sqrt{m})$ times the optimal cost when the first-stage constraint matrix has non-negative coefficients. Moreover, if there are only $k \leq m$ uncertain parameters, we generalize the performance bound for affine policies to $O(\sqrt{k})$, which is particularly useful if only a few parameters are uncertain. We also provide an $O(\sqrt{k})$ -approximation algorithm for the general case without any restriction on the constraint matrix but the solution is not an affine function of the uncertain parameters. We also give a tight characterization of the conditions under which an affine policy is optimal for the above model. In particular, we show that if the uncertainty set, $\mathcal{U} \subseteq \mathbb{R}_+^m$ is a simplex, then an affine policy is optimal. However, an affine policy is suboptimal even if \mathcal{U} is a convex combination of only $(m + 3)$ extreme points (only two more extreme points than a simplex) and

Research is partially supported by NSF Grants DMI-0556106, and EFRI-0735905.

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

V. Goyal
Department of Industrial Engineering and Operations Research,
Columbia University, New York, NY 10027, USA
e-mail: vgoyal@ieor.columbia.edu

the worst-case cost of an optimal affine policy can be a factor $(2 - \delta)$ worse than the worst-case cost of an optimal fully-adaptable solution for any $\delta > 0$.

Keywords Robust optimization · Affine control policies

Mathematics Subject Classification (2000) 90C99 · 49K35

1 Introduction

In this paper, we study the power and limitations of affine policies in solving adaptive optimization problems. In an affine policy (or *linear decision rule*), we restrict to solutions that are an affine function of the uncertain parameters. The problem of finding the best such policy is tractable. Therefore, these are widely used to solve multi-stage dynamic optimization problems and perform extremely well in practice as observed empirically. However, to the best of our knowledge, there are no theoretical performance bounds for affine policies in general.

We consider the following two-stage problem, $\Pi_{Adapt}(\mathcal{U})$ with an uncertain right hand side.

$$\begin{aligned} z_{Adapt}(\mathcal{U}) &= \min c^T x + \max_{b \in \mathcal{U}} d^T y(b) \\ Ax + By(b) &\geq b, \quad \forall b \in \mathcal{U} \\ x, y(b) &\geq 0, \end{aligned} \tag{1}$$

where $A \in \mathbb{R}^{m \times n_1}$, $B \in \mathbb{R}^{m \times n_2}$, $c \in \mathbb{R}_+^{n_1}$, $d \in \mathbb{R}_+^{n_2}$, $\mathcal{U} \subseteq \mathbb{R}_+^m$ is a compact, convex and full-dimensional uncertainty set of possible values of the right hand side of the constraints. For any $b \in \mathcal{U}$, $y(b)$ denotes the value of the second-stage variables in the scenario when the right hand side is b . We assume that (1) is feasible unless stated otherwise.

The above model includes set covering formulations considered in the recent work on approximation algorithms for two-stage robust combinatorial problems such as set cover, facility location and Steiner trees (Dhamdhere et al. [10]). For instance, by setting A and B to be the element set incidence matrix, we obtain a two-stage set covering problem. Our model is more general than the set covering model, as there are no restrictions on the coefficients of the constraint matrices. For instance, we can also model constraints of the form $y \leq \theta \cdot x$ that arise in network design problems. This can be done by setting $A = \theta \cdot I$, $B = -I$, and $b = 0$, where I is an identity matrix of an appropriate dimension. Note that our model does not admit a complete recourse, unlike the set covering models, which makes the problem significantly more difficult to solve. However, we can verify whether there exists a feasible first-stage solution that can be extended to a feasible second-stage solution for all $b \in \mathcal{U}$ (for some $b^0 \in \mathcal{U}$ such that $b_j^0 > 0$ for all $j = 1, \dots, m$, there exists a feasible solution (x, y) satisfying $Ax + By \geq b^0$. Then $(\tau x, \tau y)$ for some large enough τ is a feasible solution for all $b \in \mathcal{U}$ as \mathcal{U} is compact). Although the above models are quite general, the assumed nonnegativity of the right-hand side of the constraints prevents

us from modeling upper bounds on variables as well as packing constraints such as $\sum_{j=1}^{n_1} x_j \leq 1$. Moreover, the assumed nonnegativity of the objective coefficients does not allow us to model maximization problems.

The problem $\Pi_{Adapt}(\mathcal{U})$ is intractable in general. In fact, it is hard to approximate within a factor better than $O(\log m)$ under certain complexity assumptions via a reduction from the following max-min fractional set cover problem,

$$\max_{b \in \mathcal{U}} \min_{y(b)} \{d^T y(b) \mid By(b) \geq b\},$$

where $d \in \mathbb{R}_+^{n_2}$, $B \in \{0, 1\}^{m \times n_2}$ is an element-set incidence matrix and $\mathcal{U} \subseteq \mathbb{R}_+^m$ is a convex uncertainty set. Feige et al. [11] prove that the above max-min fractional set cover problem can not be approximated within a factor better than $O(\log m)$ in polynomial time unless $\text{NP} \subseteq \text{TIME} \left(2^{O(\sqrt{n})}\right)$, where n refers to the problem input size.

The hardness of $\Pi_{Adapt}(\mathcal{U})$ follows as setting $A = 0$ and B as an element-set incidence matrix reduces $\Pi_{Adapt}(\mathcal{U})$ to a max-min fractional set cover problem. Therefore, the problem $\Pi_{Adapt}(\mathcal{U})$ can not be approximated within a factor better than $O(\log m)$ in polynomial time unless $\text{NP} \subseteq \text{TIME} \left(2^{O(\sqrt{n})}\right)$. However, an exact or an approximate solution to $\Pi_{Adapt}(\mathcal{U})$ can be computed efficiently in several cases when the uncertainty set \mathcal{U} satisfies certain properties.

If the uncertainty set \mathcal{U} is a polytope with a small (polynomial) number of extreme points, $\Pi_{Adapt}(\mathcal{U})$ can be solved efficiently. Instead of considering the constraint $Ax + By(b) \geq b$ for all $b \in \mathcal{U}$, it suffices to consider the constraint only for all the extreme points of \mathcal{U} . Thus, the resulting expanded formulation of $\Pi_{Adapt}(\mathcal{U})$ has only a small (polynomial) number of variables and constraints which can be solved efficiently. Dhamdhare et al. [10] consider the problem $\Pi_{Adapt}(\mathcal{U})$ where the constraint matrix A and B are equal and define a combinatorial problem such as set cover, Steiner tree and facility location, and \mathcal{U} has a small number of extreme points. They give approximation algorithms with similar performance bounds as the deterministic versions. Feige et al. [11] extend to a special case of the uncertainty set with an exponential number of extreme points and give a polylogarithmic approximation for the set cover problem in this setting. Khandekar et al. [14] consider a similar uncertainty set with an exponential number of extreme points as [11] and give constant factor approximations for several network design problems such as Steiner tree and uncapacitated facility location. In most of these papers, the algorithm and the analysis is very specific to the problem and does not provide insights for a tractable solution for the general two-stage problem $\Pi_{Adapt}(\mathcal{U})$.

Bertsimas and Goyal [7] show that a static robust solution (that can be computed efficiently) is a 2-approximation to $\Pi_{Adapt}(\mathcal{U})$ if the uncertainty set is *symmetric* or *positive* (see Definition 1). To the best of our knowledge, this is the first such worst-case bound on the performance of a tractable solution approach in such generality. The assumptions of symmetry are mild and often satisfied in practice. However, non-symmetric uncertainty sets occur in modeling in several important applications (such as modeling at most k resource failures) and the cost of an optimal static solution can be $\Omega(m)$ worse [7] as compared to the fully-adaptable two-stage solution.

- Definition 1** (i) A set $P \subset \mathbb{R}^n$ is *symmetric*, if there exists some $u_0 \in P$, such that, for any $z \in \mathbb{R}^n$, $(u_0 + z) \in P \Leftrightarrow (u_0 - z) \in P$.
- (ii) A convex set $P \subset \mathbb{R}_+^n$ is *positive*, if there exists a convex symmetric set $S \subset \mathbb{R}_+^n$ such that $P \subset S$ and the point of symmetry of S is contained in P .

Affine policies have been studied extensively in the literature for two-stage and multi-stage adaptive optimization problems. They were introduced in the context of stochastic optimization in Rockafellar and Wets [17] and then later in robust optimization in Ben-Tal et al. [4] and also extended to linear systems theory in Ben-Tal et al. [2] and Ben-Tal et al. [3]. Affine policies have also been considered extensively in control theory of dynamic systems (see Bemporad [1], Kerrigan and Maciejowski [13], Löfberg [16], Bertsimas and Brown [6], Skaf and Boyd [18], Ben-Tal and Nemirovski [5] and the references therein). In all these papers, the authors show how to reformulate the multi-stage linear optimization problem such that an optimal affine policy can be computed efficiently by solving convex optimization problems such as linear, quadratic, conic and semi-definite. Kerrigan and Maciejowski citeKM04 first consider the problem of theoretically analyzing the properties of such policies and show that, under suitable conditions, the affine policy has certain desirable properties such as stability and robust invariance. Goulart and Kerrigan [12] show that under certain conditions, the class of policies that are affine in the uncertain parameters in the current state are equivalent to the class of affine policies with memory of prior states. Bertsimas et al. [9] show that an affine policy is optimal for a multi-stage problem where there is a single decision variable with upper and lower bounds and a single uncertain parameter in each stage. Bertsimas et al. [8] consider linear dynamical systems under uncertainty and show using numerical experiments on classical multi-stage inventory management problems that affine policies give near-optimal solutions. Kuhn et al. [15] consider one-stage and multi-stage stochastic optimization problems, give efficient algorithms to compute the primal and dual approximations using affine policies, and determine a posteriori performance bounds for affine policies on an instance by instance basis in contrast to the a priori performance bounds we derive for a class of instances. To the best of our knowledge, there are no a priori performance bounds known for affine policies in a general setting such as the one considered in $\Pi_{Adapt}(\mathcal{U})$.

Our main contributions are as follows.

1. **(Optimality for Simplex sets)** We show that an affine policy is optimal if the uncertainty set $\mathcal{U} \subset \mathbb{R}_+^m$ is a simplex, i.e., it is a convex combination of $(m + 1)$ affinely independent points in \mathbb{R}_+^m .
2. **(Suboptimality)** We show that the above optimality result for affine policies is almost tight, i.e., an affine policy is suboptimal even when \mathcal{U} is a convex combination of 0 and $(m + 2)$ non-zero extreme points in \mathbb{R}_+^m . In particular, given any $\delta > 0$, we present an example where \mathcal{U} is a convex combination of 0 and $(m + 2)$ other points in \mathbb{R}_+^m and the worst-case cost of an optimal affine policy is a factor $(2 - \delta)$ greater than the worst-case cost of an optimal fully-adaptable two-stage solution.

3. **(Lower Bound)** We show that surprisingly, the best affine policy can be $\Omega(m^{1/2-\delta})$ times the optimal cost of a fully-adaptable two-stage solution for $\Pi_{Adapt}(\mathcal{U})$ for any $\delta > 0$. In particular, we construct a family of examples where the uncertainty set \mathcal{U} has an exponential number of extreme points and the worst-case cost of an optimal affine policy is $\Omega(m^{1/2-\delta})$ times the optimal cost of $\Pi_{Adapt}(\mathcal{U})$. This lower bound is in contrast to strong empirical evidence of the performance of affine policies.

4. **(Upper Bound)** For a general convex, compact and full-dimensional uncertainty set $\mathcal{U} \subseteq \mathbb{R}_+^m$, we give a tight upper bound on the performance of affine policies for the case when the constraint matrix corresponding to the first-stage decision variables, $A \in \mathbb{R}_+^{n_1 \times m}$, i.e., $A_{ij} \geq 0$ for all $i = 1, \dots, m$ and $j = 1, \dots, n_1$. In particular, we show that the worst-case cost of an optimal affine policy for $\Pi_{Adapt}(\mathcal{U})$ is $O(\sqrt{m})$ times the worst-case cost of an optimal fully-adaptable two-stage solution if $A \in \mathbb{R}_+^{n_1 \times m}$. Furthermore, if only $k \leq m$ parameters are uncertain, then we show that the worst-case cost of an optimal affine policy is $O(\sqrt{k})$ times the optimal which is a good approximation when there are only a small number of uncertain parameters. We also give an $O(\sqrt{k})$ -approximation algorithm for the general case when there is no condition on A . However, the solution in the general case is not an affine policy.

The rest of the paper is organized as follows. In Sect. 2, we discuss the case when the uncertainty set is a simplex and present the proof of optimality of affine policies. We present an example for suboptimality of affine policies when the uncertainty has only $(m + 3)$ extreme points in Sect. 3. In Sect. 4, we present a family of examples that show a lower bound of $\Omega(m^{1/2-\delta})$ for any $\delta > 0$ on the performance of affine policies for general convex uncertainty sets. In Sect. 5, we present an upper bound of $O(\sqrt{m})$ on the performance of affine policies for the case when the constraint matrix $A \geq 0$ and also the performance bound of $O(\sqrt{k})$ when there are only $k \leq m$ uncertain parameters. Finally, in Sect. 6, we give an $O(\sqrt{k})$ -approximation algorithm (that is not an affine policy) for the two-stage adaptive problem when there is no non-negativity restriction on the constraint matrix A where k again is the number of uncertain parameters.

2 Optimality of affine policies for simplex uncertainty sets

In general, the optimal second-stage solution, $y^*(b)$, of the problem: $\Pi_{Adapt}(\mathcal{U})$, is a piecewise linear function of the uncertain right hand side b for all $b \in \mathcal{U}$ where the number of pieces can be exponentially many (see [1]). However, if $\mathcal{U} \subseteq \mathbb{R}_+^m$ is a simplex, we show that there is an optimal two-stage solution where the second-stage solution $y(b)$ is an affine function of $b \in \mathcal{U}$.

Theorem 1 Consider the problem $\Pi_{Adapt}(\mathcal{U})$ such that \mathcal{U} is a simplex, i.e.,

$$\mathcal{U} = \text{conv}(b^1, \dots, b^{m+1}),$$

where $b^j \in \mathbb{R}_+^m$ for all $j = 1, \dots, m$ such that b^1, \dots, b^{m+1} are affinely independent. Then, there is an optimal two-stage solution $\hat{x}, \hat{y}(b)$ for all $b \in \mathcal{U}$ such that $\hat{y}(b)$ is an affine function of b , i.e., for all $b \in \mathcal{U}$,

$$\hat{y}(b) = Pb + q,$$

where $P \in \mathbb{R}^{n_2 \times m}, q \in \mathbb{R}^{n_2}$.

Proof Consider an optimal solution $x^*, y^*(b)$ for all $b \in \mathcal{U}$ of $\Pi_{Adapt}(\mathcal{U})$. We will construct an alternate solution $\tilde{x}, \tilde{y}(b)$ for all $b \in \mathcal{U}$ such that $\tilde{y}(b)$ is an affine function of b and the worst case cost of the new solution is equal to $z_{Adapt}(\mathcal{U})$.

Let

$$Q = \left[(b^1 - b^{m+1}) \dots (b^m - b^{m+1}) \right].$$

Since b^1, \dots, b^{m+1} are affinely independent, $(b^1 - b^{m+1}), \dots, (b^m - b^{m+1})$ are linearly independent and Q is an invertible full-rank matrix. For any $b \in \mathcal{U}$,

$$b = \sum_{j=1}^{m+1} \alpha_j b^j,$$

where $\alpha_j \geq 0$ for all $j = 1, \dots, m + 1$ and $\alpha_1 + \dots + \alpha_{m+1} = 1$. Therefore, we can write the above equation as:

$$\begin{aligned} b &= \sum_{j=1}^m \alpha_j b^j + \left(1 - \sum_{j=1}^m \alpha_j \right) b^{m+1} \\ &= \sum_{j=1}^m \alpha_j (b^j - b^{m+1}) + b^{m+1} \\ &= Q \cdot \alpha + b^{m+1}, \quad \alpha = (\alpha_1, \dots, \alpha_m)^T. \end{aligned}$$

Since Q is invertible, we have,

$$Q^{-1} (b - b^{m+1}) = \alpha,$$

which implies that the convex combination multipliers for any $b \in \mathcal{U}$ can be expressed as an affine function of b . Let

$$Y = \left[(y^*(b^1) - y^*(b^{m+1})) \dots (y^*(b^m) - y^*(b^{m+1})) \right]. \tag{2}$$

Then, we can define an affine solution as follows.

$$\tilde{x} = x^*.$$

For all $b \in \mathcal{U}$, where $b = \alpha_1 b^1 + \dots + \alpha_{m+1} b^{m+1}$ for some convex multipliers $0 \leq \alpha_j \leq 1, j = 1, \dots, m + 1$,

$$\begin{aligned} \tilde{y}(b) &= Y Q^{-1} (b - b^{m+1}) + y^*(b^{m+1}) \\ &= Y \alpha + y^*(b^{m+1}), \text{ (where } \alpha = (\alpha_1, \dots, \alpha_m)) \\ &= \sum_{j=1}^{m+1} \alpha_j y^*(b^j), \end{aligned}$$

where the last equality follows from substituting the value of Y from (2). The solution $\tilde{x}, \tilde{y}(b)$ for all $b \in \mathcal{U}$ is clearly feasible for $\Pi_{Adapt}(\mathcal{U})$ and the worst case cost can be bounded as follows.

$$c^T \tilde{x} + \max_{b \in \mathcal{U}} d^T \tilde{y}(b) = c^T \tilde{x} + \max_{b \in \mathcal{U}} d^T \left(\sum_{j=1}^{m+1} \alpha_j y^*(b^{m+1}) \right), \left(\text{where } b = \sum_{j=1}^{m+1} \alpha_j b^j \right) \tag{3}$$

$$\leq c^T x^* + \max_{j=1, \dots, m+1} d^T y^*(b^j) \tag{4}$$

$$\leq z_{Adapt}(\mathcal{U}), \tag{5}$$

where (3) follows as $\tilde{y}(b) = \sum_{j=1}^{m+1} \alpha_j y^*(b^j)$. Inequality (4) follows as

$$d^T y^*(b^j) \leq \max_{j=1, \dots, m+1} d^T y^*(b^j),$$

and $\alpha_1 + \dots + \alpha_{m+1} = 1$. The last inequality follows as $b^j \in \mathcal{U}$ for all $j = 1, \dots, m + 1$. Therefore, the worst case cost of the solution $\tilde{x}, \tilde{y}(b), \forall b \in \mathcal{U}$ is equal to the optimal cost of $\Pi_{Adapt}(\mathcal{U})$, which implies that the best affine second-stage solution is optimal. \square

Note that when \mathcal{U} is a simplex, the second-stage solution $y(b)$ for all $b \in \mathcal{U}$ is described completely by only describing the solution at $(m + 1)$ extreme points, b^1, \dots, b^{m+1} . Therefore, to construct an affine solution we need to find $P \in \mathbb{R}^{n_2 \times m}, q \in \mathbb{R}^{n_2}$ such that,

$$P b^j + q = y^*(b^j), \quad \forall j = 1, \dots, m + 1,$$

where $y^*(b^j)$ is an optimal fully-adaptable second-stage solution. Such P, q always exist as the number of degrees of freedom is at least the number of constraints and in the proof of Theorem 1, we construct one such solution.

3 Suboptimality of affine policies for uncertainty sets with $(m + 2)$ non-zero extreme points

In this section, we show that an affine policy is suboptimal even if \mathcal{U} is full-dimensional and a convex combination of 0 and $(m + 2)$ other points in \mathbb{R}_+^m . This implies that the optimality of affine policies is almost tight, i.e., an affine policy is optimal only if there are only $(m + 1)$ non-zero extreme points in the uncertainty set. For simplex uncertainty sets, the degrees of freedom in an affine policy is exactly equal to the degrees of freedom required to specify an optimal solution at $(m + 1)$ extreme points of the simplex. If the uncertainty set has even one more non-zero extreme point, affine policies can be suboptimal due to a mismatch between the required and available degrees of freedom. In fact, for any $\delta > 0$, we show that if \mathcal{U} is a convex combination of 0 and $(m + 2)$ other points, the cost of an optimal affine policy can be a factor $(2 - \delta)$ worse than the cost of an optimal fully-adaptable solution. For any $\delta > 0$, we consider the following instance \mathcal{I} of $\Pi_{Adapt}(\mathcal{U})$ where m is an even integer such that $m > \frac{200}{\delta^2}$, $n_1 = n_2 = m$ and $c \in \mathbb{R}_+^{n_1}$, $d \in \mathbb{R}_+^{n_2}$, $A \in \mathbb{R}^{m \times n_1}$, $B \in \mathbb{R}^{m \times n_2}$ and the uncertainty set \mathcal{U} are defined as follows.

$$\begin{aligned}
 c &= 0, \\
 d &= (1, \dots, 1)^T. \\
 &\text{For all } i, j = 1, \dots, m \\
 A_{ij} &= 0, \\
 B_{ij} &= \begin{cases} 1, & \text{if } i = j, \\ \frac{1}{\sqrt{m}}, & \text{otherwise,} \end{cases} \\
 \mathcal{U} &= \text{conv} \left(\{b^0, b^1, \dots, b^{m+2}\} \right),
 \end{aligned} \tag{6}$$

where,

$$\begin{aligned}
 b^0 &= 0, \\
 b^j &= e_j, \quad \forall j = 1, \dots, m, \\
 b^{m+1} &= \left(\underbrace{\frac{1}{\sqrt{m}}, \dots, \frac{1}{\sqrt{m}}}_{m/2}, \underbrace{0, \dots, 0}_{m/2} \right), \\
 b^{m+2} &= \left(\underbrace{0, \dots, 0}_{m/2}, \underbrace{\frac{1}{\sqrt{m}}, \dots, \frac{1}{\sqrt{m}}}_{m/2} \right).
 \end{aligned}$$

Here, e_j denotes the j th unit vector in \mathbb{R}^m whose j th coordinate is one and all others are zero.

Theorem 2 Given any $\delta > 0$, consider the instance \mathcal{I} of $\Pi_{Adapt}(\mathcal{U})$ defined in (6) where the uncertainty set is \mathcal{U} . Then,

$$z_{Aff}(\mathcal{U}) > (2 - \delta) \cdot z_{Adapt}(\mathcal{U}).$$

We first show that the optimal worst-case cost of the instance \mathcal{I} defined in (6) is at most one.

Lemma 1 For the instance \mathcal{I} defined in (6),

$$z_{Adapt}(\mathcal{U}) \leq 1.$$

Proof We construct a fully-adaptable solution whose worst-case cost is 1 which implies that $z_{Adapt}(\mathcal{U}) \leq 1$. Since $A = 0$, we can assume without loss of generality that the first-stage decision variables $x = 0$ in all feasible solutions. Let us first specify the second-stage solution at the extreme points of \mathcal{U} :

$$y(b^j) = \begin{cases} b^j, & \text{if } j \in \{0, 1, \dots, m\}, \\ \frac{1}{m} \cdot e, & \text{if } j \in \{m + 1, m + 2\}. \end{cases}$$

For any $b \in \mathcal{U}$, b can be expressed as a convex combination of the extreme points, i.e.,

$$b = \sum_{j=0}^{m+2} \alpha_j b^j,$$

for some $\alpha_j \geq 0$ for all $j = 0, \dots, m + 2$ and $\alpha_0 + \dots + \alpha_{m+2} = 1$. Then, the second-stage solution, $y(b)$ is defined as

$$y(b) = \sum_{j=0}^{m+2} \alpha_j y(b^j).$$

We next verify that the solution is feasible. For the extreme point b^0 , $By(b^0) = 0 \geq b^0$. For any b^j , $j = 1, \dots, m$,

$$By(b^j) = Bb^j = Be_j = \frac{1}{\sqrt{m}} \cdot e + \left(1 - \frac{1}{\sqrt{m}}\right) \cdot e_j \geq e_j.$$

Therefore, the solution is feasible for b^j , $j = 0, \dots, m$. Also, for $j = m + 1, m + 2$,

$$\begin{aligned} By(b^j) &= B \cdot \left(\frac{1}{m}, \dots, \frac{1}{m}\right) \\ &= \left(\frac{1}{m} + \frac{m-1}{m} \cdot \frac{1}{\sqrt{m}}\right) \cdot e \end{aligned}$$

$$\begin{aligned}
 &= \left(\frac{\sqrt{m} + m - 1}{m^{3/2}} \right) \cdot e \\
 &\geq \left(\frac{m}{m^{3/2}} \right) \cdot e \\
 &= \frac{1}{\sqrt{m}} \cdot e \\
 &\geq b^j,
 \end{aligned}$$

which implies that the solution is feasible for b^{m+1} and b^{m+2} . Therefore, the solution is feasible for all extreme points. For any $b \in \mathcal{U}$ such that,

$$b = \sum_{j=0}^{m+2} \alpha_j b^j,$$

for some $\alpha_j \geq 0$ for all $j = 0, \dots, m + 2$ and $\alpha_0 + \dots + \alpha_{m+2} = 1$,

$$\begin{aligned}
 By(b) &= B \left(\sum_{j=0}^{m+2} \alpha_j y(b^j) \right) \\
 &= \sum_{j=0}^{m+2} \alpha_j \cdot By(b^j) \\
 &\geq \sum_{j=0}^{m+2} \alpha_j b^j \\
 &= b.
 \end{aligned}$$

Hence, the solution is feasible for all $b \in \mathcal{U}$. Therefore,

$$\begin{aligned}
 z_{Adapt}(\mathcal{U}) &\leq \max\{d^T y(b) \mid b \in \mathcal{U}\} \\
 &= \max\{d^T y(b^j) \mid j = 0, \dots, m + 2\} \tag{7}
 \end{aligned}$$

$$= \max \left(\max\{d^T b^j \mid j = 0, \dots, m\}, d^T \left(\frac{1}{m} \cdot e \right) \right) \tag{8}$$

$$= 1, \tag{9}$$

where (7) follows as the worst-case cost of the solution is achieved at an extreme point of \mathcal{U} . Equation (8) follows as $y(b^j) = b^j$, $j = 0, \dots, m$ and $y(b^j) = \frac{1}{m} \cdot e$ for $j = m + 1, m + 2$. Equation (9) follows as $d^T b^0 = 0$, $d^T b^j = d^T e_j = 1$ for all $j = 1, \dots, m$ and $d^T \left(\frac{1}{m} \cdot e \right) = 1$. \square

Next we show that the worst-case cost of an optimal affine second-stage solution is at least $(2 - \delta)$. We first show the existence of a *symmetric* optimal affine solution for the instance \mathcal{I} . Consider any permutation $\tau \in S^m$, where S^m is the set of permutations of $\{1, \dots, m\}$. We define the following notations.

- (i) For any $x \in \mathbb{R}^m$, $x^\tau = (x_{\tau(1)}, \dots, x_{\tau(m)})$.
- (ii) For any $Y \in \mathbb{R}^{m \times m}$, $Y_{ij}^\tau = Y_{\tau(i)\tau(j)}$, for all $i, j \in \{1, \dots, m\}$.

Definition 2 A set $U \subseteq \mathbb{R}^m$ is said to be *permutation-invariant* with respect to a permutation $\tau \in S^m$ if, $x \in P \iff x^\tau \in P$.

Let

$$\Gamma = \left\{ \tau \in S^m \mid i \leq \frac{m}{2} \iff \tau(i) \leq \frac{m}{2} \right\}. \tag{10}$$

We first show that the uncertainty set \mathcal{U} defined in (6) is *permutation-invariant* with respect to any permutation in Γ .

Lemma 2 *The uncertainty set \mathcal{U} in the instance \mathcal{I} defined in (6) is permutation-invariant with respect to any permutation in Γ .*

Proof Consider any $b \in \mathcal{U}$ and any permutation $\tau \in \Gamma$. We prove that $b^\tau \in \mathcal{U}$ by a case analysis.

Case 1 (*b is an extreme point of \mathcal{U}*). In this case, $b = b^j$ for some $j \in \{0, 1, \dots, m+2\}$. If $b = b^0$, $b^\tau = (0, \dots, 0) = b^0$. For any $j = 1, \dots, m$, if $b = e^j$, then $b^\tau = (e^j)^\tau = e^{\tau(j)} \in \mathcal{U}$. If $b = b^{m+1}$, then

$$b_{\tau(j)} = \begin{cases} 1, & \text{if } j \leq m/2, \\ 0, & \text{otherwise,} \end{cases}$$

as $j \leq m/2 \iff \tau(j) \leq m/2$. Therefore, $b^\tau = (b^{m+1})^\tau = b^{m+1}$. Similarly, $(b^{m+2})^\tau = b^{m+2}$. Therefore, if b is an extreme point $b^\tau \in \mathcal{U}$.

Case 2 (*b is not an extreme point of \mathcal{U}*). In this case $b = \alpha_0 b^0 + \dots + \alpha_{m+2} b^{m+2}$ for some $\alpha_j \geq 0$ for all $j = 0, \dots, m+2$ such that $\alpha_0 + \dots + \alpha_{m+2} = 1$. It is easy to observe that,

$$b^\tau = \sum_{j=0}^{m+2} \alpha_j (b^j)^\tau.$$

Since $(b^j)^\tau \in \mathcal{U}$ for all $j = 0, \dots, m+2$ (from Case 1), $b^\tau \in \mathcal{U}$.

Using a similar argument as above, we can show that if $b^\tau \in \mathcal{U}$, then $(b^\tau)^{\tau^{-1}} \in \mathcal{U}$ as $\tau \in \Gamma$ implies $\tau^{-1} \in \Gamma$. Therefore, $b \in \mathcal{U}$ as $(b^\tau)^{\tau^{-1}} = b$. □

Lemma 3 *There exists an optimal affine solution,*

$$\hat{y}(b) = \hat{P}b + \hat{q},$$

where $\hat{P} \in \mathbb{R}^{m \times m}$ and $\hat{q} \in \mathbb{R}^m$ such that $\hat{q}_i = \hat{q}_j$ for all $i, j \in \{1, \dots, m\}$.

Proof Consider an optimal affine solution:

$$\tilde{y}(b) = \tilde{P}b + \tilde{q},$$

for all $b \in \mathcal{U}$. For any $\tau \in \Gamma$ (10), consider the following permuted instance $\mathcal{I}' = \mathcal{I}^\tau$ where $c' = c^\tau = 0$, $A' = A^\tau = 0$, $B' = B^\tau = B$ and $\mathcal{U}' = \mathcal{U}^\tau = \mathcal{U}$. Note that the permuted instance is constructed by permuting the index of constraints and variables. Therefore, it is easy to observe that the following permuted solution,

$$\tilde{y}^\tau(b) = \tilde{P}^\tau b + \tilde{q}^\tau,$$

for all $b \in \mathcal{U}^\tau = \mathcal{U}$ is an optimal solution for the permuted instance \mathcal{I}^τ .

Furthermore, \mathcal{I}^τ is same as the original instance \mathcal{I} . Therefore, $\tilde{y}^\tau(b)$ is an optimal solution of the original instance. Therefore, the following solution:

$$\begin{aligned} \hat{y}(b) &= \frac{1}{|\Gamma|} \sum_{\tau \in \Gamma} \tilde{y}^\tau(b) \\ &= \frac{1}{|\Gamma|} \sum_{\tau \in \Gamma} (\tilde{P}^\tau b + \tilde{q}^\tau) \\ &= \frac{1}{|\Gamma|} \left(\sum_{\tau \in \Gamma} \tilde{P}^\tau \right) b + \frac{1}{|\Gamma|} \sum_{\tau \in \Gamma} \tilde{q}^\tau, \end{aligned}$$

is an optimal affine solution for the instance \mathcal{I} as it is a convex combination of optimal affine solutions. Let

$$\bar{P} = \frac{1}{|\Gamma|} \left(\sum_{\tau \in \Gamma} \tilde{P}^\tau \right), \quad \bar{q} = \frac{1}{|\Gamma|} \sum_{\tau \in \Gamma} \tilde{q}^\tau.$$

Therefore, for any $j = 1, \dots, m/2$,

$$\begin{aligned} \bar{q}_j &= \frac{1}{|\Gamma|} \sum_{i=1}^{m/2} \sum_{\tau \in \Gamma: \tau(i)=j} \tilde{q}_i \\ &= \frac{1}{|\Gamma|} \sum_{i=1}^{m/2} \binom{m}{2} \tilde{q}_i, \end{aligned} \tag{11}$$

where (11) follows as there are exactly $(m/2 - 1)!$ permutations in Γ that map i to j for any $i, j \in \{1, \dots, m/2\}$. Similarly, for any $j = m/2 + 1, \dots, m$,

$$\bar{q}_j = \frac{1}{|\Gamma|} \sum_{i=m/2+1}^n \binom{m}{2} \tilde{q}_i.$$

Therefore,

$$\bar{q} = \left(\underbrace{\beta_1, \dots, \beta_1}_{m/2}, \underbrace{\beta_2, \dots, \beta_2}_{m/2} \right),$$

for some $\beta_1, \beta_2 \in \mathbb{R}$. Now consider the following permutation σ where

$$\sigma(j) = \left(\frac{m}{2} + j\right) \bmod m, \quad \forall j = 1, \dots, m.$$

It is easy to observe that \mathcal{U} is permutation-invariant with respect to the permutation σ . Therefore, using a similar argument as before \overline{P}^σ and \overline{q}^σ define an optimal affine solution for the original instance \mathcal{I} . Let

$$\hat{P} = \frac{1}{2} (\overline{P} + \overline{P}^\sigma); \hat{q} = \frac{1}{2} (\overline{q} + \overline{q}^\sigma).$$

Therefore, the solution

$$\hat{y}(b) = \hat{P}b + \hat{q},$$

is an optimal affine solution as it is a convex combination of two optimal affine solutions of \mathcal{I} . Now for any $j = 1, \dots, m$,

$$\begin{aligned} \hat{q}_j &= \frac{\overline{q}_j + \overline{q}_{(j+m/2) \bmod m}}{2} \\ &= \frac{\beta_1 + \beta_2}{2}. \end{aligned}$$

□

Proof of Theorem 2 Let $\hat{y}(b) = \hat{P}b + \hat{q}$ for all $b \in \mathcal{U}$ be an optimal affine solution that satisfies the symmetric property in Lemma 3 and let $\hat{q}_j = \beta$ for all $j = 1, \dots, m$ for some $\beta \in \mathbb{R}$. Now,

$$z_{Aff}(\mathcal{U}) = \max\{d^T(\hat{P}b + \hat{q}) \mid b \in \mathcal{U}\}.$$

For the sake of contradiction, suppose that

$$z_{Aff}(\mathcal{U}) \leq (2 - \delta). \tag{12}$$

Claim $\beta \leq (2 - \delta)/m$.

Suppose not. Consider $b = b^0 = 0$, $\hat{y}(b) = \hat{P}b + \hat{q} = \hat{q}$. Therefore,

$$z_{Aff}(\mathcal{U}) \geq d^T \hat{y}(b^0) = d^T \hat{q} = m \cdot \beta > 2 - \delta,$$

a contradiction to (12).

Claim For all $j = 1, \dots, m$, $\hat{P}_{jj} \geq 1 - \frac{2}{\sqrt{m}} - \frac{2}{m}$.

Consider $b = e_j$ for $j = 1, \dots, m$, $\hat{y}(b) = \hat{P}e_j + \hat{q}$. Now,

$$B(\hat{P}b^j + \hat{q}) = B(\hat{P}e_j + \hat{q}) \geq e_j.$$

Therefore, $(B\hat{P}e_j + B\hat{q})_j \geq 1$ which implies that,

$$(\hat{P}_{jj} + \beta) + \sum_{i \neq j} \frac{1}{\sqrt{m}} (\hat{P}_{ij} + \beta) \geq 1.$$

Therefore, either

$$(\hat{P}_{jj} + \beta) \geq 1 - \frac{2}{\sqrt{m}} \text{ or } \left(\sum_{i \neq j} (\hat{P}_{ij} + \beta) \right) \geq 2.$$

Suppose for some $j' \in \{1, \dots, m\}$,

$$\left(\sum_{i \neq j'} (\hat{P}_{ij'} + \beta) \right) \geq 2. \quad (13)$$

Then,

$$\begin{aligned} d^T \hat{y}(b^{j'}) &= \sum_{i=1}^m \hat{y}_i(b^{j'}) \\ &= \sum_{i=1}^m (\hat{P}_{ij'} + \beta) \\ &\geq \sum_{i \neq j'} (\hat{P}_{ij'} + \beta) \end{aligned} \quad (14)$$

$$\geq 2, \quad (15)$$

where (14) follows as $P_{j'j'} + \beta \geq 0$ and (15) follows from (13). This implies that $z_{Aff}(\mathcal{U}) \geq 2$; a contradiction to (12). Thus, we can assume without loss of generality that for all $j = 1, \dots, m$,

$$(\hat{P}_{jj} + \beta) \geq 1 - \frac{2}{\sqrt{m}},$$

which implies

$$\begin{aligned} \hat{P}_{jj} &\geq 1 - \frac{2}{\sqrt{m}} - \beta \\ &> 1 - \frac{2}{\sqrt{m}} - \frac{2}{m}, \end{aligned}$$

where the last inequality follows from Claim 3.

Claim $\hat{P}_{ij} \geq -(2 - \delta)/m$ for all $i, j \in \{1, \dots, m\}$.

Consider $b = e_j$ for $j = 1, \dots, m$. Now, $\hat{y}(b) = \hat{P}e_j + \hat{q}$. For any $i = 1, \dots, m$, $\hat{y}_i(b) = \hat{P}_{ij} + \beta$. Since $\hat{y}(b) \geq 0$,

$$\hat{P}_{ij} \geq -\beta \geq -\frac{2 - \delta}{m}.$$

Now, consider $b = b^{m+1}$. For $i = 1, \dots, m/2$,

$$\begin{aligned} y_i(b^{m+1}) &= (\hat{P}b^{m+1} + \hat{q})_i \\ &= \left(\sum_{j=1}^{m/2} \hat{P}_{ij} \cdot \frac{1}{\sqrt{m}} \right) + \beta \\ &= \frac{\hat{P}_{ii}}{\sqrt{m}} + \sum_{j=1, j \neq i}^{m/2} \frac{\hat{P}_{ij}}{\sqrt{m}} + \beta \\ &\geq \frac{1}{\sqrt{m}} - \frac{2}{m} - \frac{2}{m^{3/2}} - \left(\frac{m}{2} - 1 \right) \cdot \frac{2 - \delta}{m^{3/2}} \tag{16} \\ &\geq \frac{1}{\sqrt{m}} - \frac{2}{m} - \frac{2}{m^{3/2}} - \frac{2 - \delta}{2\sqrt{m}} \\ &= \left(\frac{1}{\sqrt{m}} - \frac{2 - \delta}{2\sqrt{m}} \right) - \frac{2}{m} - \frac{2}{m^{3/2}} \\ &= \frac{\delta}{2\sqrt{m}} - \frac{2}{m} - \frac{2}{m^{3/2}} \\ &= \frac{1}{\sqrt{m}} \left(\frac{\delta}{2} - \frac{2}{\sqrt{m}} - \frac{2}{m} \right) \\ &> \frac{1}{\sqrt{m}} \left(\frac{\delta}{2} - \frac{\delta}{\sqrt{50}} - \frac{\delta^2}{100} \right) \tag{17} \\ &> \frac{\delta}{3\sqrt{m}}, \end{aligned}$$

where (16) follows from Claims 3 and 3 and Inequality (17) follows as $m > 200/\delta^2$. Therefore,

$$d^T y(b^{m+1}) > \frac{m}{2} \cdot \frac{\delta}{3\sqrt{m}} = \frac{\delta \cdot \sqrt{m}}{6} > 2,$$

a contradiction to (12). Hence, $z_{Aff}(\mathcal{U}) > (2 - \delta) \cdot z_{Adapt}(\mathcal{U})$ even when there are only $m + 3$ extreme points in the right hand side uncertainty set. \square

4 A large gap example for affine policies

In this section, we show that the worst-case cost of an affine policy in $\Pi_{Adapt}(\mathcal{U})$ can be arbitrarily large as compared to the optimal cost. In particular, we show that for any $\delta > 0$ the cost of an optimal affine policy can be $\Omega(m^{1/2-\delta})$ times the cost of an optimal fully-adaptable solution for the following instance \mathcal{I} . Let $n_1 = n_2 = m$, where,

$$m^\delta > 200, \tag{18}$$

and $c \in \mathbb{R}_+^{n_1}$, $d \in \mathbb{R}_+^{n_2}$, $A \in \mathbb{R}^{m \times n_1}$, $B \in \mathbb{R}^{m \times n_2}$ and the uncertainty set \mathcal{U} are defined as follows.

$$\begin{aligned} c &= 0, \\ d &= (1, \dots, 1)^T. \\ &\text{For all } i, j = 1, \dots, m \\ A_{ij} &= 0, \\ B_{ij} &= \begin{cases} 1, & \text{if } i = j, \\ \theta_0, & \text{otherwise,} \end{cases} \\ \mathcal{U} &= \text{conv} \left(\{b^0, b^1, \dots, b^N\} \right), \end{aligned} \tag{19}$$

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

$$\begin{aligned} b^0 &= 0, \\ b^j &= e_j, \quad \forall j = 1, \dots, m, \\ b^{m+1} &= \frac{1}{\sqrt{m}} \cdot e, \\ b^{m+2} &= \theta_0 \cdot \left(\underbrace{1, \dots, 1}_r, 0, \dots, 0 \right), \end{aligned}$$

where exactly r coordinates are non-zero, each equal to θ_0 . Extreme points b^j , $j \geq m + 3$ are permutations of the non-zero coordinates of b^{m+2} . Therefore, \mathcal{U} has exactly $\binom{m}{r}$ extreme points of the form of b^{m+2} . Note that all the non-zero extreme points of \mathcal{U} are roughly on the boundary of the unit hypersphere centered at zero. Therefore, intuitively a unit hypersphere intersected with the non-negative orthant is the worst-case example for affine policies in view of a tight upper bound of $O(\sqrt{m})$ on their performance presented in the next section.

Theorem 3 *For the instance \mathcal{I} defined in (19) defined the uncertainty set is \mathcal{U} ,*

$$z_{Aff}(\mathcal{U}) = \Omega \left(m^{1/2-\delta} \right) \cdot z_{Adapt}(\mathcal{U}),$$

for any given $\delta > 0$.

We first show that the worst case cost of an optimal fully-adaptable two-stage solution is at most one.

Lemma 4 For the instance \mathcal{I} defined in (19), $z_{Adapt}(\mathcal{U}) \leq 1$.

Proof We show this by constructing a fully-adaptable solution whose worst-case cost is 1 which implies that $z_{Adapt}(\mathcal{U}) \leq 1$. Since $A = 0$, we can assume without loss of generality that the first-stage decision variables $x = 0$ in all feasible solutions. Let us first specify the second-stage solution at the extreme points of \mathcal{U} :

$$y(b^j) = \begin{cases} b^j, & \text{if } j \in \{0, 1, \dots, m\}, \\ \frac{1}{m} \cdot e, & \text{if } j \geq m + 1. \end{cases}$$

For any $b \in \mathcal{U}$, b can be expressed as a convex combination of the extreme points, i.e.,

$$b = \sum_{j=0}^N \alpha_j b^j,$$

for some $\alpha_j \geq 0$ for all $j = 0, \dots, N$ and $\alpha_0 + \dots + \alpha_N = 1$. Then, the second-stage solution, $y(b)$ is defined as:

$$y(b) = \sum_{j=0}^N \alpha_j y(b^j).$$

It is easy to observe that the solution is feasible. Let us first verify the feasibility for all the extreme points. For extreme point b^0 , $y(b^0) = b^0 = 0$ and $By(b^0) = 0 = b^0$. For b^j , $j = 1, \dots, m$,

$$By(b^j) = Bb^j = Be_j \geq e_j.$$

Therefore, the solution is feasible for b^j , $j = 0, \dots, m$. For any, $j \geq m + 1$,

$$\begin{aligned} By(b^j) &= B \cdot \left(\frac{1}{m} \cdot e \right) \\ &= \left(\frac{1}{m} + \frac{m-1}{m} \cdot \frac{1}{m^{(1-\delta)/2}} \right) \cdot e \\ &= \left(\frac{m^{(1-\delta)/2} + m-1}{m^{(3-\delta)/2}} \right) \cdot e \\ &\geq \left(\frac{m}{m^{(3-\delta)/2}} \right) \cdot e \\ &= \left(\frac{1}{m^{(1-\delta)/2}} \right) \cdot e \\ &\geq b^j, \end{aligned}$$

which implies that the solution is feasible for b^j . Therefore, the solution is feasible for all extreme points. Using an argument similar to the proof of Lemma 1, we can show that the solution is feasible for all $b \in \mathcal{U}$.

Therefore,

$$\begin{aligned} z_{Adapt}(\mathcal{U}) &\leq \max\{d^T y(b) \mid b \in \mathcal{U}\} \\ &= \max\{d^T y(b^j) \mid j = 0, \dots, N\} \\ &= \max\left(\max\{d^T b^j \mid j = 0, \dots, m\}, d^T \left(\frac{1}{m} \cdot e\right)\right) \tag{20} \\ &= 1, \tag{21} \end{aligned}$$

where (20) follows as $y(b^j) = b^j, j = 0, \dots, m$ and $y(b^j) = 1/m \cdot e$ for all $j \geq m + 1$ and (21) follows as $d^T b^0 = 0, d^T b^j = d^T e_j = 1$ for all $j = 1, \dots, m$ and $d^T (1/m \cdot e) = 1$. \square

In the following lemmas, we show that there exists an optimal affine solution satisfying certain symmetric properties for the instance of $\Pi_{Adapt}(\mathcal{U})$ in Theorem 3.

Lemma 5 *The uncertainty set \mathcal{U} in (19) is permutation-invariant with respect to any permutation in S^m .*

Proof Consider any permutation $\sigma \in S^m$. We need to show that $b \in \mathcal{U} \iff b^\sigma \in \mathcal{U}$. Suppose $b \in \mathcal{U}$. We show that $b^\sigma \in \mathcal{U}$ by a case analysis.

1. If $b = b^0$, then clearly $b^\sigma = b \in \mathcal{U}$.
2. If $b = b^j = e_j$ for $j = 1, \dots, m$, then $b^\sigma = e_{\sigma(j)} \in \mathcal{U}$.
3. If $b = b^{m+1} = (1/\sqrt{m}, \dots, 1/\sqrt{m})$, then $b^\sigma = b \in \mathcal{U}$.
4. If $b = b^j$ for $j = m + 2, \dots, N$, then b has exactly $\lceil m^{1-\delta} \rceil$ non-zero coordinates each equal to θ_0 . Therefore, b^σ also satisfies the same structure and since all possible combinations of the non-zero coordinates are extreme points of \mathcal{U} , $b^\sigma \in \mathcal{U}$.
5. If b is not an extreme point of \mathcal{U} , then $b = \alpha_0 b^0 + \dots + \alpha_N b^N$ for some $\alpha_j \geq 0, j = 0, \dots, N$ and $\alpha_0 + \dots + \alpha_N = 1$. Therefore,

$$b^\sigma = \sum_{j=0}^N \alpha_j (b^j)^\sigma.$$

Since $(b^j)^\sigma \in \mathcal{U}$ for all $j = 0, \dots, N$, b^σ is a convex combination of the extreme points of \mathcal{U} which implies $b^\sigma \in \mathcal{U}$.

Therefore, $b \in \mathcal{U}$ implies $b^\sigma \in \mathcal{U}$.

Conversely, suppose $v = b^\sigma \in \mathcal{U}$. By the same argument as above, we can show that $v^{(\sigma^{-1})} \in \mathcal{U}$. Now,

$$v^{(\sigma^{-1})} = (b^\sigma)^{(\sigma^{-1})} = b,$$

which implies $b \in \mathcal{U}$. \square

We next show that if we consider a *permuted instance* of the instance in Theorem 3, we obtain exactly the same instance. For any $\sigma \in S^m$, consider the following permuted instance $\mathcal{I}' = \mathcal{I}(\sigma)$ of the instance defined in (19) where

$$\begin{aligned} c' &= 0, \\ d' &= d^\sigma = (1, \dots, 1)^T, \\ A' &= A^\sigma = 0, \\ B' &= B^\sigma, \\ \mathcal{U}' &= \mathcal{U}^\sigma. \end{aligned} \tag{22}$$

Note that for any $i, j = 1, \dots, m$, $B_{ij}^\sigma = B_{\sigma(i),\sigma(j)}$.

Lemma 6 *The permuted instance $\mathcal{I}(\sigma)$ defined in (22) is the same as the original instance.*

Proof Note that B is symmetric where $B_{ii} = 1$ for all $i = 1, \dots, m$ and $B_{ij} = \theta_0$ for all $i \neq j$. Therefore, $B_{ii}^\sigma = B_{\sigma(i),\sigma(i)} = 1$ for all $i = 1, \dots, m$. Also for any $i \neq j$, $B_{ij}^\sigma = B_{\sigma(i),\sigma(j)} = \theta_0$ since $i \neq j$ implies $\sigma(i) \neq \sigma(j)$. Therefore, $B^\sigma = B$. In Lemma 5, we show that \mathcal{U} is permutation-invariant and thus, $\mathcal{U}^\sigma = \mathcal{U}$. Therefore, the permuted instance is exactly similar to the original instance. \square

Lemma 7 *Let $\sigma \in S^n$, and let*

$$y(b) = Pb + q,$$

be an optimal affine solution. Then the following solution,

$$y^\sigma(b) = P^\sigma b + q^\sigma,$$

where $P_{ij}^\sigma = P_{\sigma(i),\sigma(j)}$ for all $i, j \in \{1, 2, \dots, m\}$, and $q_j^\sigma = q_{\sigma(j)}$ for all $j = 1, \dots, m$, is an optimal affine solution for the original instance. Note that $y^\sigma(b)$ is not the permutation of coordinates of the solution vector $y(b)$ by σ .

Proof Let us first prove that the new solution is feasible. Suppose not. Therefore, there exists $b \in \mathcal{U}$ and $i' \in \{1, \dots, m\}$ such that,

$$(By^\sigma(b))_{i'} < b_{i'}.$$

Let us permute the constraints by σ^{-1} . Therefore,

$$\left(B(\sigma^{-1})y^\sigma(b) \right)_{\sigma^{-1}(i')} < b_{\sigma^{-1}(i')},$$

where $B(\sigma^{-1})_{ij} = B_{\sigma^{-1}(i),j}$ for all $i, j \in \{1, \dots, m\}$. Let us also permute the columns of $B(\sigma^{-1})$ by σ^{-1} so that we get back the original constraint matrix as $B_{\sigma^{-1}(i),\sigma^{-1}(j)} = B_{ij}$ for all $i, j \in \{1, \dots, m\}$. Since we permute the columns of

$B(\sigma^{-1})$, we must also permute the solution vector $y^\sigma(b)$ by the same permutation. Therefore,

$$\left(B \left(y^\sigma(b) \right)^{(\sigma^{-1})} \right)_{\sigma^{-1}(i')} < b_{\sigma^{-1}(i')}, \quad (23)$$

where,

$$\left(y^\sigma(b) \right)^{(\sigma^{-1})} = P'b + (q^\sigma)^{\sigma^{-1}} = P'b + q,$$

and $P' = (P^\sigma)(\sigma^{-1})$, i.e., $P'_{ij} = P_{i,\sigma(j)}$ for all $i, j \in \{1, \dots, m\}$. If we also permute the columns of P' by σ^{-1} , we can express the above equation as,

$$\left(y^\sigma(b) \right)^{(\sigma^{-1})} = Pb^{\sigma^{-1}} + q = y(b^{\sigma^{-1}}). \quad (24)$$

Simplifying (23) using (24), we have that

$$\left(By \left(b^{\sigma^{-1}} \right) \right)_{\sigma^{-1}(i')} < b_{\sigma^{-1}(i')} = b_{i'}^{\sigma^{-1}},$$

which is a contradiction as $y(b^{\sigma^{-1}})$ is a feasible solution for $b^{\sigma^{-1}}$. Furthermore,

$$d^T y^\sigma(b) = \left(d^{\sigma^{-1}} \right)^T \left(y^\sigma(b) \right)^{\sigma^{-1}} = d^T y(b^{\sigma^{-1}}),$$

as $\left(y^\sigma(b) \right)^{\sigma^{-1}} = y(b^{\sigma^{-1}})$ from (24). Hence, the worst-case of the solution $y^\sigma(b)$ is same as the worst-case cost of the optimal affine solution. \square

We next show that there is an optimal affine solution for the instance in Theorem 3 that satisfies certain symmetry properties.

Lemma 8 *There exists an optimal affine solution,*

$$\hat{y}(b) = \hat{P}b + \hat{q},$$

for all $b \in \mathcal{U}$ such that,

1. $\hat{P}_{ij} = \mu$ for all $i \neq j$ for some $\mu \in \mathbb{R}$ and $\hat{P}_{jj} = \theta$ for all $j = 1, \dots, m, \theta \in \mathbb{R}$.
2. $\hat{q}_j = \lambda$, for all $j = 1, \dots, m, \lambda \in \mathbb{R}$.

Proof Consider an optimal affine solution:

$$y(b) = Pb + q,$$

where $P \in \mathbb{R}^{m \times m}$ and $q \in \mathbb{R}^m$. Let S^m denote the set of all permutations of $\{1, \dots, m\}$. From Lemma 7, we know that for any $\sigma \in S^m$, the solution,

$$y^\sigma(b) = P^\sigma b + q^\sigma,$$

for all $b \in \mathcal{U}$ is also an optimal solution for the problem. We know that any convex combination of a set of optimal solutions is also optimal. Consider the following solution,

$$\hat{y}(b) = \hat{P}b + \hat{q},$$

where

$$\begin{aligned} \hat{P} &= \frac{1}{|S^m|} \cdot \sum_{\sigma \in S^m} P^\sigma, \\ \hat{q} &= \frac{1}{|S^m|} \cdot \sum_{\sigma \in S^m} q^\sigma. \end{aligned}$$

Therefore the solution defined by \hat{P} and \hat{q} is a convex combination of the solutions defined by all the permutations which are each optimal implying that the new solution is also optimal. Furthermore, for any $j = 1, \dots, m$,

$$\begin{aligned} \hat{P}_{jj} &= \frac{1}{|S^m|} \cdot \sum_{\sigma \in S^m} P_{jj}^\sigma \\ &= \frac{1}{|S^m|} \cdot \sum_{\sigma \in S^m} P_{\sigma(j),\sigma(j)} \\ &= \frac{1}{|S^m|} \cdot \sum_{k=1}^m \sum_{\sigma:\sigma(j)=k} P_{kk} \\ &= \frac{1}{|S^m|} \cdot \sum_{k=1}^m (m-1)! P_{kk} \tag{25} \\ &= \frac{(m-1)!}{m!} \cdot \sum_{k=1}^m P_{kk} \\ &= \frac{1}{m} \cdot \sum_{k=1}^m P_{kk}, \end{aligned}$$

where (25) follows as $|\{\sigma \in S^m \mid \sigma(j) = k\}| = (m-1)!$. The final expression of \hat{P}_{jj} does not depend on j and thus, $\hat{P}_{jj} = \theta$ for all $j = 1, \dots, m$ where

$$\theta = \frac{1}{m} \cdot \sum_{k=1}^m P_{kk}.$$

For any $i, j \in \{1, \dots, m\}, i \neq j$,

$$\hat{P}_{ij} = \frac{1}{|S^m|} \cdot \sum_{\sigma \in S^m} P_{ij}^\sigma$$

$$\begin{aligned}
 &= \frac{1}{|S^m|} \cdot \sum_{\sigma \in S^m} P_{\sigma(i),\sigma(j)} \\
 &= \frac{1}{|S^m|} \cdot \sum_{k=1}^m \sum_{l=1, \dots, m, l \neq k} \left(\sum_{\sigma: \sigma(i)=k, \sigma(j)=l} P_{kl} \right) \\
 &= \frac{1}{|S^m|} \cdot \sum_{k=1}^m \sum_{l=1, \dots, m, l \neq k} (m-2)! P_{kl} \tag{26} \\
 &= \frac{(m-2)!}{m!} \cdot \sum_{k=1}^m \sum_{l=1, \dots, m, l \neq k} P_{kl} \\
 &= \frac{1}{m(m-1)} \cdot \sum_{k=1}^m \sum_{l=1, \dots, m, l \neq k} P_{kl},
 \end{aligned}$$

where (26) follows as $|\{\sigma \in S^m \mid \sigma(i) = k, \sigma(j) = l\}| = (m-2)!$. Again, the expression of \hat{P}_{ij} does not depend on i, j and thus, $\hat{P}_{ij} = \mu$ for all $i, j = 1, \dots, m, i \neq j$ where,

$$\mu = \frac{1}{m(m-1)} \cdot \sum_{k=1}^m \sum_{l=1, \dots, m, l \neq k} P_{kl}.$$

Using a similar argument, we can show that $\hat{q}_j = \lambda$ for all $j = 1, \dots, m$ where,

$$\lambda = \frac{1}{m} \cdot \sum_{k=1}^m q_k.$$

□

Proof of Theorem 3 Let $\hat{y}(b) = \hat{P}b + \hat{q}$ for all $b \in \mathcal{U}$ be an optimal affine solution satisfying that for all $i, j \in \{1, \dots, m\}, i \neq j, \hat{P}_{ij} = \mu, P_{jj} = \theta$ and $\hat{q}_j = \lambda$. Such a solution exists from Lemma 8. Let

$$z_{Aff}(\mathcal{U}) = \max\{d^T(\hat{P}b + \hat{q}) \mid b \in \mathcal{U}\}.$$

For the sake of contradiction, suppose that

$$z_{Aff}(\mathcal{U}) \leq \frac{m^{1/2-\delta}}{4}. \tag{27}$$

Claim $0 \leq \lambda \leq \frac{1}{m^{1/2+\delta}}$.

Note that $\hat{y}(b^0) = \hat{y}(0) = \lambda \cdot e$. Since $\hat{y}(b^0) \geq 0, \lambda \geq 0$. Now, suppose $\lambda > \frac{1}{m^{1/2+\delta}}$. Then,

$$z_{Aff}(\mathcal{U}) \geq d^T \hat{y}(b^0) = d^T (\lambda \cdot e) = \lambda \cdot e^T e = \lambda \cdot m > m^{1/2-\delta},$$

a contradiction to (27).

Claim $\theta \geq \frac{1}{3}$.

Note that

$$y(e_1) = \left(\theta + \lambda, \underbrace{\mu + \lambda, \dots, \mu + \lambda}_{(m-1)} \right).$$

Since $By(e_1) \geq e_1$,

$$(\theta + \lambda) + (m - 1) \cdot (\mu + \lambda) \cdot \theta_0 \geq 1.$$

Therefore, either

$$(\theta + \lambda) \geq \frac{1}{2} \text{ or } (m - 1) \cdot (\mu + \lambda) \cdot \theta_0 \geq \frac{1}{2}.$$

Suppose

$$(m - 1) \cdot (\mu + \lambda) \cdot \theta_0 \geq \frac{1}{2}. \tag{28}$$

Then,

$$\begin{aligned} z_{Aff}(\mathcal{U}) \geq d^T y(e_1) &= e^T \left(\theta + \lambda, \underbrace{\mu + \lambda, \dots, \mu + \lambda}_{m-1} \right) \\ &= \theta + \lambda + (m - 1) \cdot (\mu + \lambda) \\ &\geq (m - 1) \cdot (\mu + \lambda) \end{aligned} \tag{29}$$

$$\begin{aligned} &\geq \frac{m^{(1-\delta)/2}}{2} \\ &> \frac{m^{1/2-\delta}}{2}, \end{aligned} \tag{30}$$

where (29) follows as $\theta + \lambda \geq 0$ since $y(e_1) \geq 0$ and (30) follows from (28). This implies a contradiction to (27). Thus, we can assume that

$$(\theta + \lambda) \geq \frac{1}{2},$$

which implies that

$$\begin{aligned}
 \theta &\geq \frac{1}{2} - \lambda \\
 &\geq \frac{1}{2} - \frac{1}{m^{1/2+\delta}} \\
 &> \frac{1}{3},
 \end{aligned}
 \tag{31}$$

where (31) follows from Claim 4.

Claim $-\frac{1}{m^{1+\delta/2}} \leq \mu < 0$.

Suppose $\mu \geq 0$. Consider $b = b^{m+1}$. For any $i = 1, \dots, m$,

$$\begin{aligned}
 y_i(b^{m+1}) &= \left(\hat{P}b^{m+1} + \hat{q} \right)_i \\
 &= \theta \cdot \frac{1}{\sqrt{m}} + (m-1) \cdot \mu \frac{1}{\sqrt{m}} + \lambda \\
 &\geq \theta \cdot \frac{1}{\sqrt{m}} + \lambda
 \end{aligned}
 \tag{32}$$

$$\geq \theta \cdot \frac{1}{\sqrt{m}}
 \tag{33}$$

$$\geq \frac{1}{3\sqrt{m}},
 \tag{34}$$

where (32) follows since we assumed that $\mu \geq 0$. Inequality (33) follows as $\lambda \geq 0$ from Claim 4 and (34) follows from Claim 4. Now,

$$z_{Aff}(\mathcal{U}) \geq d^T y(b^{m+1}) = \sum_{i=1}^m y_i(b^{m+1}) > m \cdot \frac{1}{3\sqrt{m}} = \frac{\sqrt{m}}{3},$$

a contradiction to (27). Therefore, $\mu < 0$.

Consider the extreme point b^{m+2} where the first $r = \lceil m^{1-\delta} \rceil$ coordinates are non-zero and each equal to θ_0 . Now,

$$y(b^{m+2}) = \begin{bmatrix} \theta_0 \cdot (\theta + (r-1)\mu) + \lambda \\ \vdots \\ \theta_0 \cdot (\theta + (r-1)\mu) + \lambda \\ \theta_0 \cdot r\mu + \lambda \\ \vdots \\ \theta_0 \cdot r\mu + \lambda \end{bmatrix}.$$

Since $y(b^{m+2}) \geq 0$, $\theta_0 \cdot r\mu + \lambda \geq 0$ and thus,

$$\begin{aligned} \mu &\geq -\frac{\lambda}{\theta_0 \cdot r} \\ &\geq -\frac{1}{m^{1/2+\delta}} \cdot m^{(1-\delta)/2} \cdot \frac{1}{\lceil m^{1-\delta} \rceil} \\ &\geq -\frac{1}{m^{3\delta/2}} \cdot \frac{1}{m^{1-\delta}} \\ &= -\frac{1}{m^{1+\delta/2}}, \end{aligned} \tag{35}$$

where (35) follows from the bound on λ in Claim 4 and substituting values of θ_0 and r . Now, consider $b = b^{m+1}$. For any $i = 1, \dots, m$,

$$\begin{aligned} y_i(b^{m+1}) &= \left(\hat{p}b^{m+1} + \hat{q} \right)_i \\ &= \theta \cdot \frac{1}{\sqrt{m}} + (m-1) \cdot \mu \cdot \frac{1}{\sqrt{m}} + \lambda \\ &\geq \frac{1}{3\sqrt{m}} - (m-1) \cdot \frac{1}{m^{(3+\delta)/2}} \end{aligned} \tag{36}$$

$$\begin{aligned} &> \frac{1}{3\sqrt{m}} - m \cdot \frac{1}{m^{(3+\delta)/2}} \\ &= \frac{1}{3\sqrt{m}} - \frac{1}{m^{(1+\delta)/2}} \\ &= \frac{1}{\sqrt{m}} \cdot \left(\frac{1}{3} - \frac{1}{m^{\delta/2}} \right) \\ &> \frac{1}{\sqrt{m}} \cdot \frac{1}{4}, \end{aligned} \tag{37}$$

where (36) follows as $\theta \geq 1/3$ and $\mu \geq -1/m^{1+\delta/2}$ and (37) follows as $m^\delta > 200$ (18) which implies $m^{\delta/2} > 12$. Therefore,

$$z_{Aff}(\mathcal{U}) \geq d^T y(b^{m+1}) = \sum_{i=1}^m y_i(b^{m+1}) > m \cdot \frac{1}{4\sqrt{m}} = \frac{\sqrt{m}}{4},$$

which is a contradiction to (27). □

We would like to note that the large gap instance defined in (19) satisfies the restrictions of $c \geq 0, d \geq 0, A \geq 0, B \geq 0$ and $\mathcal{U} \subset \mathbb{R}_+^m$. Therefore, the lower bound on the performance of affine policies we prove above clearly holds for the more general problem without some of these restrictions.

5 Performance guarantee for affine policies when $A \geq 0$

In this section, we prove that an optimal affine policy is within $O(\sqrt{m})$ from the optimal two-stage adaptive objective when the right hand side is uncertain and the constraint matrix corresponding to the first-stage decision variables, $A \geq 0$. In view of the example in Sect. 4 where we show a gap of $\Omega(m^{1/2-\delta})$ for any constant $\delta > 0$ for affine policies, it follows that the approximation factor $O(\sqrt{m})$ is tight. We later consider the case where there are only $k \leq m$ uncertain parameters and show that the performance of affine policies is only $O(\sqrt{k})$ times worse than the cost of an optimal fully-adaptable solution. This bound is particularly useful in applications where there are only a small number of uncertain parameters. Furthermore, the assumption that $A \geq 0$ is satisfied in many applications including network design, facility location, capacity planning and other applications of the set-covering problem.

Theorem 4 Consider the problem $\Pi_{Adapt}(\mathcal{U})$ where the right hand side uncertainty set $\mathcal{U} \subseteq \mathbb{R}_+^m$ is convex, compact and full-dimensional and $A \geq 0$. Then

$$z_{Aff}(\mathcal{U}) \leq 3\sqrt{m} \cdot z_{Adapt}(\mathcal{U}),$$

i.e., the worst case cost of an optimal affine policy is at most $3\sqrt{m}$ times the worst case cost of an optimal fully adaptable solution.

We construct a feasible affine solution for $\Pi_{Adapt}(\mathcal{U})$ starting from an optimal fully-adaptable solution such that the worst case cost of the affine solution is at most $3\sqrt{m} \cdot z_{Adapt}(\mathcal{U})$. For all $j = 1, \dots, m$, let

$$\begin{aligned} \mu_j &= \max\{b_j \mid b \in \mathcal{U}\}, \\ \beta^j &\in \operatorname{argmax}\{b_j \mid b \in \mathcal{U}\}. \end{aligned} \quad (38)$$

It is easy to observe that $\mu_j > 0$ for all $j = 1, \dots, m$, since $\mathcal{U} \subseteq \mathbb{R}_+^m$ is full-dimensional. To motivate the construction we consider the following case where,

$$\sum_{j=1}^m \frac{b_j}{\mu_j} \leq \sqrt{m}, \quad \forall b \in \mathcal{U}. \quad (39)$$

In this case we construct a feasible affine (in fact a linear) solution as follows. Let $x^*, y^*(b)$ for all $b \in \mathcal{U}$ denote an optimal fully-adaptable solution for $\Pi_{Adapt}(\mathcal{U})$. Let

$$P = \left[\frac{1}{\mu_1} \cdot y^*(\beta^1), \dots, \frac{1}{\mu_m} \cdot y^*(\beta^m) \right].$$

Consider the solution:

$$\tilde{x} = \sqrt{m} \cdot x^*, \quad \tilde{y}(b) = Pb, \quad \forall b \in \mathcal{U}. \quad (40)$$

The solution is feasible as for any $b \in \mathcal{U}$,

$$\begin{aligned}
 A\tilde{x} + B\tilde{y}(b) &= A(\sqrt{m} \cdot x^*) + B(Pb) \\
 &= \sqrt{m} \cdot Ax^* + B \left(\sum_{j=1}^m \frac{b_j}{\mu_j} \cdot y^*(\beta^j) \right) \\
 &= \left(\sqrt{m} - \sum_{j=1}^m \frac{b_j}{\mu_j} \right) \cdot Ax^* + \sum_{j=1}^m \frac{b_j}{\mu_j} \cdot (Ax^* + By^*(\beta^j)) \\
 &\geq \sum_{j=1}^m \frac{b_j}{\mu_j} \cdot (Ax^* + By^*(\beta^j)) \tag{41} \\
 &\geq \sum_{j=1}^m \frac{b_j}{\mu_j} \cdot \beta^j \tag{42} \\
 &= \sum_{j=1}^m b_j \cdot \left(\frac{1}{\mu_j} \cdot \beta^j \right) \\
 &\geq \sum_{j=1}^m b_j \cdot e_j \tag{43} \\
 &= b,
 \end{aligned}$$

where (41) follows, since $Ax^* \geq 0$ and thus, $Ax^* \geq 0$. Inequality (42) follows from the feasibility of $x^*, y^*(\beta^j)$ for $\beta^j, j = 1, \dots, m$ and (43) follows as $\beta_i^j = \mu_j$ and $\beta_i^j \geq 0$ for all $i = 1, \dots, m$. The cost of the solution for any $b \in \mathcal{U}$ is

$$\begin{aligned}
 c^T \tilde{x} + d^T (Pb) &= \sqrt{m} \cdot c^T x^* + d^T \left(\sum_{j=1}^m \frac{b_j}{\mu_j} y^*(\beta^j) \right) \\
 &= \sqrt{m} \cdot c^T x^* + \sum_{j=1}^m \frac{b_j}{\mu_j} \cdot d^T y^*(\beta^j) \\
 &\leq \sqrt{m} \cdot c^T x^* + \left(\sum_{j=1}^m \frac{b_j}{\mu_j} \right) \cdot \max_{j=1, \dots, m} d^T y^*(\beta^j) \\
 &\leq \sqrt{m} \cdot c^T x^* + \sqrt{m} \cdot \max_{j=1, \dots, m} d^T y^*(\beta^j) \tag{44} \\
 &= \sqrt{m} \cdot \left(c^T x^* + \max_{j=1, \dots, m} d^T y^*(\beta^j) \right) \\
 &\leq \sqrt{m} \cdot z_{Adapt}(\mathcal{U}), \tag{45}
 \end{aligned}$$

where (44) follows from (39) and (45) follows as $c^T x^* + d^T y^*(\beta^j) \leq z_{Adapt}(\mathcal{U})$ since $\beta^j \in \mathcal{U}$ for all $j = 1, \dots, m$. Therefore, the worst case cost of the affine solution (40)

is at most \sqrt{m} times the worst case cost of an optimal fully-adaptable solution if \mathcal{U} satisfies (39).

For the general case when (39) is not satisfied, the basic idea is that the set of coordinates $\{1, \dots, m\}$ can be partitioned into two sets J_1, J_2 with $J_1 \cup J_2 = \{1, \dots, m\}$, $J_1 \cap J_2 = \emptyset$ such that

1. For all $b \in \mathcal{U}$,

$$\sum_{j \in J_1} \frac{b_j}{\mu_j} \leq \sqrt{m}.$$

2. There exists $\beta \in \mathbb{R}_+^m$ that is a sum of at most $2\sqrt{m}$ points in \mathcal{U} such that $\beta_j \geq b_j$ for all $j \in J_2, b \in \mathcal{U}$.

We construct the partition J_1, J_2 and $\beta \in \mathbb{R}_+^m$ satisfying the above properties iteratively where we initialize β to zero and J_1 to $\{1, \dots, m\}$. In each iteration, we select a $b \in \mathcal{U}$ such that $\sum_{j \in J_1} b_j / \mu_j$ is strictly greater than \sqrt{m} (if such a point exists) and add it to β . If for any coordinate $j \in J_1, \beta_j$ is at least μ_j , this implies that the current β dominates the j th coordinate of all $b \in \mathcal{U}$. Therefore, we remove j from J_1 and add it to J_2 . Since in each iteration we are adding a point from \mathcal{U} whose scaled sum over coordinates in J_1 is large, we argue that the algorithm will find β and the partition J_1, J_2 in at most $2\sqrt{m}$ iterations which we use to prove the required properties. The algorithm to construct β and the partition J_1, J_2 is described in Fig. 1. Note that the construction of β and the partition $[J_1, J_2]$ is only to show the existence of a good affine solution for $\Pi_{Adapt}(\mathcal{U})$ and is not required to compute the best affine solution which can be found directly in polynomial time.

Lemma 9 *Let β, J_1, J_2 be the vector and the partition that Algorithm \mathcal{A} produces. For all $b \in \mathcal{U}$,*

$$\sum_{j \in J_1} \frac{b_j}{\mu_j} \leq \sqrt{m},$$

and $b_j \leq \beta_j$ for all $j \in J_2$.

Proof The first property follows from the fact that Algorithm \mathcal{A} terminates when the condition in Step 2 is not satisfied which implies that,

$$\sum_{j \in J_1} \frac{b_j}{\mu_j} \leq \sqrt{m},$$

for all $b \in \mathcal{U}$. Furthermore, for any $j \in J_2, \beta_j \geq \mu_j$ from Step 2d(i). □

Next we show that the number of iterations required to compute β in Algorithm \mathcal{A} is at most $2\sqrt{m}$.

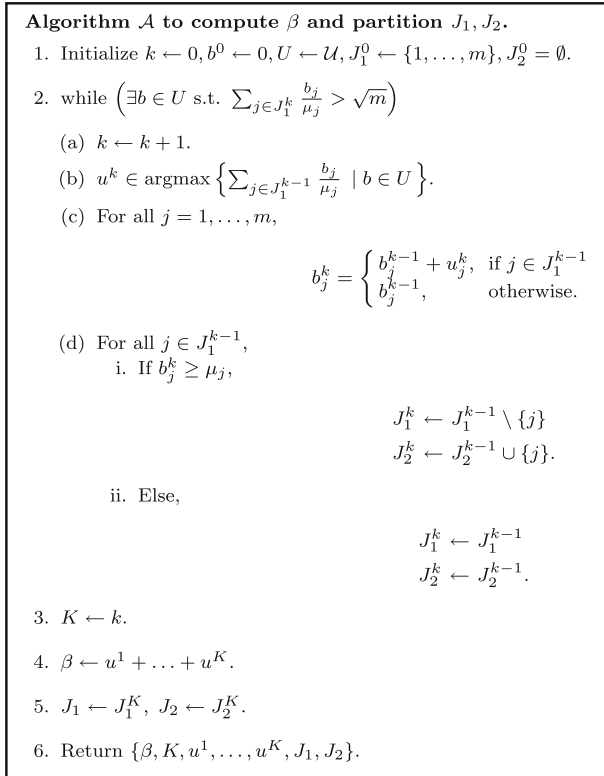


Fig. 1 Algorithm \mathcal{A} for computing β and partition J_1, J_2

Lemma 10 *The number of iterations, K in Algorithm \mathcal{A} is at most $2\sqrt{m}$.*

Proof Let b^1, \dots, b^K be the sequence of vectors constructed in Algorithm \mathcal{A} . We first argue that $b_j^K \leq 2\mu_j$ for all $j = 1, \dots, m$. For any $j = 1, \dots, m$, suppose $k(j)$ is the last iteration when $j \in J_1$ before Step 2d in Algorithm \mathcal{A} . Therefore, $b_j^{k(j)-1} \leq \mu_j$. Also, $u_j^{k(j)} \leq \mu_j$. Therefore,

$$b^{k(j)} = b^{k(j)-1} + u_j^{k(j)} \leq 2\mu_j.$$

Now, $b_j^K = b_j^k = b_j^{k(j)}$ for all $k \geq k(j)$ which implies $b_j^K \leq 2\mu_j$ for all $j = 1, \dots, m$. Therefore,

$$\sum_{j=1}^m \frac{b_j^K}{\mu_j} \leq 2m.$$

Alternatively,

$$\begin{aligned} \sum_{j=1}^m \frac{b_j^K}{\mu_j} &= \sum_{j=1}^m \sum_{k=1}^K \frac{b_j^k - b_j^{k-1}}{\mu_j} \\ &= \sum_{k=1}^K \sum_{j=1}^m \frac{b_j^k - b_j^{k-1}}{\mu_j} \\ &= \sum_{k=1}^K \sum_{j \in J_1^{k-1}} \frac{u_j^k}{\mu_j} \end{aligned} \tag{46}$$

$$\begin{aligned} &\geq \sum_{k=1}^K \sqrt{m} \\ &= K \cdot \sqrt{m}, \end{aligned} \tag{47}$$

where (46) follows as $b_j^k = b_j^{k-1}$ for all $j \notin J_1^{k-1}$ and (47) follows from the choice of u^k in each iteration $k = 1, \dots, K$ in \mathcal{A} . Therefore, we have that

$$K \cdot \sqrt{m} \leq \sum_{j=1}^m \frac{b_j^K}{\mu_j} \leq 2m,$$

which implies that $K \leq 2\sqrt{m}$. □

Proof of Theorem 4 Suppose $x^*, y^*(b)$ for all $b \in \mathcal{U}$ is an optimal solution for $\Pi_{Adapt}(\mathcal{U})$. Also, let $\beta, K, u^1, \dots, u^K, J_1, J_2$ be returned by Algorithm \mathcal{A} and let

$$\hat{y} = \frac{2\sqrt{m}}{K} \sum_{k=1}^K y^*(u^k). \tag{48}$$

Consider the following affine solution: $\tilde{x} = 3\sqrt{m} \cdot x^*$ and for all $b \in \mathcal{U}$,

$$\tilde{y}(b) = \sum_{j \in J_1} \frac{1}{\mu_j} \cdot y^*(\beta^j) \cdot b_j + \hat{y}.$$

Let us first verify the feasibility of the affine solution. For any $b \in \mathcal{U}$,

$$\begin{aligned} A\tilde{x} + B\tilde{y}(b) &= A(3\sqrt{m} \cdot x^*) + B \left(\sum_{j \in J_1} y^*(\beta^j) \cdot \frac{b_j}{\mu_j} + \hat{y} \right) \\ &\geq \left(2\sqrt{m} + \sum_{j \in J_1} \frac{b_j}{\mu_j} \right) \cdot Ax^* + B \left(\sum_{j \in J_1} y^*(\beta^j) \cdot \frac{b_j}{\mu_j} + \hat{y} \right) \end{aligned} \tag{49}$$

$$\begin{aligned}
 &= \sum_{j \in J_1} \frac{b_j}{\mu_j} \cdot (Ax^* + By^*(\beta^j)) + 2\sqrt{m} \cdot Ax^* + B\hat{y} \\
 &\geq \sum_{j \in J_1} \frac{b_j}{\mu_j} \cdot \beta^j + 2\sqrt{m} \cdot Ax^* + B\hat{y} \tag{50}
 \end{aligned}$$

$$\geq \sum_{j \in J_1} b_j \cdot e_j + 2\sqrt{m} \cdot Ax^* + B\hat{y} \tag{51}$$

$$= \sum_{j \in J_1} b_j \cdot e_j + 2\sqrt{m} \cdot Ax^* + \sum_{k=1}^K \frac{2\sqrt{m}}{K} \cdot By^*(u^k) \tag{52}$$

$$\begin{aligned}
 &= \sum_{j \in J_1} b_j \cdot e_j + \sum_{k=1}^K \frac{2\sqrt{m}}{K} \cdot (Ax^* + By^*(u^k)) \\
 &\geq \sum_{j \in J_1} b_j \cdot e_j + \sum_{k=1}^K \frac{2\sqrt{m}}{K} u^k \tag{53}
 \end{aligned}$$

$$\geq \sum_{j \in J_1} b_j \cdot e_j + \sum_{k=1}^K u^k \tag{54}$$

$$\begin{aligned}
 &= \sum_{j \in J_1} b_j \cdot e_j + \beta \\
 &\geq b, \tag{55}
 \end{aligned}$$

where (49) follows as $\sum_{j \in J_1} \frac{b_j}{\mu_j} \leq \sqrt{m}$ from Lemma 9 and $Ax^* \geq 0$. Inequality (50) follows as the solution $x^*, y^*(\beta^j)$ is feasible for β^j for all $j = 1, \dots, m$ since $\beta^j \in \mathcal{U}$. Inequality (51) follows as $\beta^j \geq 0$ and $\beta_j^j = \mu_j$. Equation (52) follows from substituting the value of \hat{y} from (48). Inequality (53) follows from the feasibility of $x^*, y^*(u^k)$ as $u^k \in \mathcal{U}$ for all $k = 1, \dots, K$ and (54) follows as $K \leq 2\sqrt{m}$ from Lemma 10. Finally (55) follows as $\beta \geq 0$ and for all $j \notin J_1, \beta_j \geq b_j$. Therefore, the solution is feasible for all $b \in \mathcal{U}$.

We next argue that the cost of the affine solution $\tilde{x}, \tilde{y}(b)$ for all $b \in \mathcal{U}$ is $O(\sqrt{m})$ times $z_{Adapt}(\mathcal{U})$. For any $b \in \mathcal{U}$,

$$\begin{aligned}
 c^T \tilde{x} + d^T \tilde{y}(b) &= c^T (3\sqrt{m} \cdot x^*) + d^T \left(\sum_{j \in J_1} y^*(\beta^j) \cdot \frac{b_j}{\mu_j} + \hat{y} \right) \\
 &= 3\sqrt{m} \cdot c^T x^* + \sum_{j \in J_1} \frac{b_j}{\mu_j} \cdot d^T y^*(\beta^j) + d^T \hat{y} \\
 &\leq 3\sqrt{m} \cdot c^T x^* + \sum_{j \in J_1} \frac{b_j}{\mu_j} \cdot \max_{j \in J_1} d^T y^*(\beta^j) + d^T \hat{y} \\
 &\leq \sqrt{m} \cdot \left(c^T x^* + \max_{j \in J_1} d^T y^*(\beta^j) \right) + 2\sqrt{m} \cdot c^T x^* + d^T \hat{y} \tag{56}
 \end{aligned}$$

$$\leq \sqrt{m} \cdot z_{Adapt}(\mathcal{U}) + 2\sqrt{m} \cdot c^T x^* + d^T \hat{y} \tag{57}$$

$$= \sqrt{m} \cdot z_{Adapt}(\mathcal{U}) + \sum_{k=1}^K \frac{2\sqrt{m}}{K} \cdot (c^T x^* + d^T y^*(u^k)) \tag{58}$$

$$\leq \sqrt{m} \cdot z_{Adapt}(\mathcal{U}) + \sum_{k=1}^K \frac{2\sqrt{m}}{K} \cdot z_{Adapt}(\mathcal{U}) \tag{59}$$

$$= 3\sqrt{m} \cdot z_{Adapt}(\mathcal{U}),$$

where (56) follows from Lemma 9, (57) follows from the fact that $\beta^j \in \mathcal{U}$ for all $j = 1, \dots, m$ and thus, $z_{Adapt}(\mathcal{U}) \geq c^T x^* + d^T y^*(\beta^j)$. Equation (58) follows from substituting the value of \hat{y} from (48) and (59) follows the fact that $u^k \in \mathcal{U}$ and thus, $z_{Adapt}(\mathcal{U}) \geq c^T x^* + d^T y^*(u^k)$ for all $k = 1, \dots, K$. Therefore, the cost of the affine solution is at most $3\sqrt{m} \cdot z_{Adapt}(\mathcal{U})$ for all $b \in \mathcal{U}$ which implies

$$z_{Aff}(\mathcal{U}) \leq 3\sqrt{m} \cdot z_{Adapt}(\mathcal{U}).$$

□

Several comments are in order. The upper bound of $O(\sqrt{m})$ on the performance of affine policies is tight (up to a constant factor) since from Theorem 3, we have that for any $\delta > 0$, there exists an instance of $\Pi_{Adapt}(\mathcal{U})$ such that the worst case cost of an optimal affine policy is $\Omega(m^{1/2-\delta})$ times the worst case cost of an optimal fully adaptable solution. Furthermore, while the proof of Theorem 4 is existential where we only show the existence of a good affine policy, an optimal affine policy can be computed efficiently in polynomial time. Also note that the performance of affine policies is strictly better as compared to the performance of a single robust solution which can be $\Omega(m)$ factor worse than the optimal.

5.1 Performance bound when the number of uncertain parameters is $k \leq m$

We show that if the number of uncertain parameters in $\Pi_{Adapt}(\mathcal{U})$ is $k \leq m$, then we can prove an improved performance bound of $O(\sqrt{k})$ on the performance of affine policies. In many applications, only a small number of parameters are uncertain where our result shows that affine policies provide a good approximation. In particular, we consider the following problem $\Pi_{Adapt}^k(\mathcal{U})$ where only k right hand side coefficients are uncertain.

$$z_{Adapt}^k(\mathcal{U}) = \min c^T x + \max_{b \in \mathcal{U}} d^T y(b, b^0) \tag{60}$$

$$A_1 x + B_1 y(b, b^0) \geq b, \quad \forall b \in \mathcal{U} \tag{61}$$

$$A_2 x + B_2 y(b, b^0) \geq b^0 \tag{62}$$

$$x, y(b, b^0) \geq 0, \tag{63}$$

where $A_1 \in \mathbb{R}_+^{k \times n_1}$, $A_2 \in \mathbb{R}_+^{(m-k) \times n_2}$, $B_1 \in \mathbb{R}_+^{k \times n_2}$, $B_2 \in \mathbb{R}_+^{(m-k) \times n_2}$, $b^0 \in \mathbb{R}_+^{m-k}$, $c \in \mathbb{R}_+^{n_1}$, $d \in \mathbb{R}_+^{n_2}$, $\mathcal{U} \subseteq \mathbb{R}_+^k$ is a convex uncertainty set of possible values of b . As before,

for any $b \in \mathcal{U}$, $y(b, b^0)$ denotes the value of the second-stage variables in the scenario when the right hand side is (b, b^0) . We show that an affine policy is an $O(\sqrt{k})$ approximation for the above problem. In particular, we prove the following theorem.

Theorem 5 *Let $z_{Aff}^k(\mathcal{U})$ denote the cost of an optimal affine policy and let $z_{Adapt}^k(\mathcal{U})$ denote the cost of an optimal fully-adaptable solution for $\Pi_{Adapt}^k(\mathcal{U})$ where the number of uncertain right hand side coefficients is $k \leq m$. Then,*

$$z_{Aff}^k(\mathcal{U}) = O(\sqrt{k}) \cdot z_{Adapt}^k(\mathcal{U}).$$

We show the existence of a good affine solution for $\Pi_{Adapt}^k(\mathcal{U})$ by constructing one such affine solution from a fully-adaptable solution. In particular, we consider the following two problems constructed from $\Pi_{Adapt}^k(\mathcal{U})$.

$$\begin{aligned} z(\Pi_1) &= \min c^T x + \max_{b \in \mathcal{U}} d^T y(b) \\ A_1 x + B_1 y(b) &\geq b, \quad \forall b \in \mathcal{U} \\ x, y(b) &\geq 0. \end{aligned} \tag{\Pi_1}$$

$$\begin{aligned} z(\Pi_2) &= \min c^T x + d^T y \\ A_2 x + B_2 y &\geq b^0 \\ x, y &\geq 0. \end{aligned} \tag{\Pi_2}$$

Lemma 11 *The optimal cost of both Π_1 and Π_2 are at most the optimal cost of $\Pi_{Adapt}^k(\mathcal{U})$, i.e.,*

$$z(\Pi_1) \leq z_{Adapt}^k(\mathcal{U}), \quad z(\Pi_2) \leq z_{Adapt}^k(\mathcal{U}).$$

Proof Consider $x^*, y^*(b, b^0)$ for all $b \in \mathcal{U}$ be an optimal full-adaptable solution for $\Pi_{Adapt}^k(\mathcal{U})$. Clearly, it is a feasible solution for Π_1 as:

$$A_1 x^* + B_1 y^*(b, b^0) \geq b, \quad \forall b \in \mathcal{U}.$$

Therefore,

$$z(\Pi_1) \leq c^T x^* + \max_{b \in \mathcal{U}} d^T y^*(b, b^0) = z_{Adapt}^k(\mathcal{U}).$$

Also, consider any $b^1 \in \mathcal{U}$. Then the solution $x^*, y(b^1, b^0)$ is a feasible solution for Π_2 as:

$$A_2 x^* + B_2 y^*(b^1, b^0) \geq b^0.$$

Therefore,

$$z(\Pi_2) \leq c^T x^* + d^T y^*(b^1, b^0) \leq c^T x^* + \max_{b \in \mathcal{U}} d^T y^*(b, b^0) = z_{Adapt}^k(\mathcal{U}).$$

□

Proof of Theorem 5 Let $x^1, y^1(b) = Pb + q$ for all $b \in \mathcal{U}$ be an optimal affine solution for Π_1 and let x^2, y^2 be an optimal solution for Π_2 . We construct an affine solution for $\Pi_{Adapt}^k(\mathcal{U})$ as follows.

$$\tilde{x} = x^1 + x^2, \tilde{y}(b, b^0) = y^1(b) + y^2 = Pb + q + y^2, \quad \forall b \in \mathcal{U}. \tag{61}$$

We first show that the above solution is feasible.

$$\begin{aligned} A_1 \tilde{x} + B_1 \tilde{y}(b, b^0) &= A_1(x^1 + x^2) + B_1(y^1(b) + y^2) \\ &\geq A_1 x^1 + B_1 y^1(b) \\ &\geq b, \end{aligned} \tag{62}$$

where (62) follows as $A_1 x^2, B_1 y^2 \geq 0$ and the last inequality follows from feasibility of $x^1, y^1(b)$ for Π_1 . Also,

$$A_2(x^1 + x^2) + B_2(y^1(b) + y^2) \geq A_2 x^2 + B_2 y^2 \geq b^0,$$

where the first inequality follows as $A_2 x^1, B_2 y^1(b) \geq 0$ and the second inequality follows from the feasibility of x^2, y^2 for Π_2 . Therefore, the solution (61) is feasible for $\Pi_{Adapt}^k(\mathcal{U})$. Now,

$$\begin{aligned} c^T \tilde{x} + \max_{b \in \mathcal{U}} d^T \tilde{y}(b, b^0) &= c^T(x^1 + x^2) + \max_{b \in \mathcal{U}} d^T(y^1(b) + y^2) \\ &= \left(c^T x^1 + \max_{b \in \mathcal{U}} d^T(Pb + q) \right) + \left(c^T x^2 + d^T y^2 \right) \\ &= z_{Aff}(\Pi_1) + z(\Pi_2) \end{aligned} \tag{63}$$

$$\leq 3\sqrt{k} \cdot z(\Pi_1) + z(\Pi_2) \tag{64}$$

$$\leq 3\sqrt{k} \cdot z_{Adapt}^k(\mathcal{U}) + z_{Adapt}^k(\mathcal{U}) \tag{65}$$

$$= O(\sqrt{k}) \cdot z_{Adapt}^k(\mathcal{U}),$$

where (63) follows from as $x^1, y^1(b)$ is an optimal affine solution for Π_1 and x^2, y^2 is an optimal solution for Π_2 . Also, $z_{Aff}(\Pi_1)$ denotes the cost of an optimal affine solution. Inequality (64) follows from Theorem 4 which shows that the cost of an optimal affine solution is at most $3\sqrt{m}$ times the cost of an optimal fully adaptable solution, where m is the number of constraints. Note that the number of constraints in Π_1 is exactly k and $z(\Pi_1)$ denotes the cost of an optimal fully-adaptable solution. Finally, inequality (65) follows from Lemma 11. □

6 An $O(\sqrt{m})$ -approximation for the general case

In this section, we present $O(\sqrt{m})$ -approximation algorithm for $\Pi_{Adapt}(\mathcal{U})$ even when the assumption on $A \geq 0$ is relaxed using an argument similar in spirit to the proof of Theorem 4. The goal is to compute a first stage solution x such that for any $b \in \mathcal{U}$, the optimal second-stage solution does not have a high cost. Note that the algorithm does not propose a functional form such as an affine policy for the second stage solution instead just computes a good first stage solution in an efficient manner.

The basic idea is to construct an uncertainty set \mathcal{U}^0 from the given uncertainty set \mathcal{U} in $\Pi_{Adapt}(\mathcal{U})$, such that \mathcal{U}^0 has a small number of extreme points and it dominates \mathcal{U} . In other words, for all $b \in \mathcal{U}$, there exists a $b' \in \mathcal{U}^0$ such that $b \leq b'$. In particular, we consider the following uncertainty set.

$$\mathcal{U}^0 = \text{conv} \left\{ 2\sqrt{m} \cdot \beta^1, \dots, 2\sqrt{m} \cdot \beta^m, 2\beta \right\}, \tag{66}$$

where $\beta^j, j = 1, \dots, m$ are defined in (38) and $\beta = u^1 + \dots + u^K$ is the point returned by Algorithm \mathcal{A} . We show that \mathcal{U}^0 satisfies the above mentioned properties and the worst case cost of an optimal fully-adaptable solution on \mathcal{U}^0 is $O(\sqrt{m})$ times $z_{Adapt}(\mathcal{U})$. Furthermore, since \mathcal{U}^0 is a convex combination of only $(m + 1)$ extreme points which are affinely independent, $\Pi_{Adapt}(\mathcal{U}^0)$ can be solved optimally using an affine policy (see Theorem 1) and an optimal first stage solution for $\Pi_{Adapt}(\mathcal{U}^0)$ is a good approximation as a first stage solution of $\Pi_{Adapt}(\mathcal{U})$. In particular, we prove the following theorem.

Theorem 6 *Let \mathcal{U}^0 be as defined in (66) and let $\tilde{x} \in \mathbb{R}_+^n$ be an optimal first stage solution for $\Pi_{Adapt}(\mathcal{U}^0)$. Then \tilde{x} is an $O(\sqrt{m})$ -approximation for the first stage solution of $\Pi_{Adapt}(\mathcal{U})$, i.e., the worst case cost of the following solution for $\Pi_{Adapt}(\mathcal{U})$: \tilde{x} is the first stage solution and for all $b \in \mathcal{U}$,*

$$y(b) = \text{argmin}\{d^T y \mid By \geq b - A\tilde{x}, y \geq 0\},$$

is at most $O(\sqrt{m}) \cdot z_{Adapt}(\mathcal{U})$.

We first show that \mathcal{U}^0 dominates \mathcal{U} .

Lemma 12 *For any $b \in \mathcal{U}$, there exists $b' \in \mathcal{U}^0$ such that $b \leq b'$.*

Proof Consider any $b \in \mathcal{U}$ and consider the partition $[J_1, J_2]$ of $\{1, \dots, m\}$ computed by Algorithm \mathcal{A} . We know that

$$\sum_{j \in J_1} \frac{b_j}{\mu_j} \leq \sqrt{m},$$

from Lemma 9. Therefore,

$$\lambda = \sum_{j \in J_1} \frac{b_j}{2\sqrt{m}\mu_j} \leq \frac{1}{2}. \tag{67}$$

Also for all $j \in J_2$,

$$b_j \leq \beta_j \leq 2\beta_j,$$

since $\beta \geq 0$. Let $v^j = 2\sqrt{m} \cdot \beta^j$ for all $j = 1, \dots, m$ and $v^{m+1} = 2\beta$. Therefore, $\mathcal{U}^0 = \text{conv}(v^1, \dots, v^{m+1})$. We can express b as follows.

$$\begin{aligned} b &= \sum_{j=1}^m b_j \cdot e_j \\ &= \sum_{j \in J_1} \frac{b_j}{2\sqrt{m}\mu_j} \cdot 2\sqrt{m}\mu_j \cdot e_j + \sum_{j \in J_2} b_j \cdot e_j \\ &\leq \sum_{j \in J_1} \frac{b_j}{2\sqrt{m}\mu_j} \cdot 2\sqrt{m} \cdot \beta^j + \sum_{j \in J_2} b_j \cdot e_j \end{aligned} \tag{68}$$

$$\begin{aligned} &= \sum_{j \in J_1} \frac{b_j}{2\sqrt{m}\mu_j} \cdot v^j + \sum_{j \in J_2} b_j \cdot e_j \\ &\leq \sum_{j \in J_1} \frac{b_j}{2\sqrt{m}\mu_j} \cdot v^j + \sum_{j \in J_2} \beta_j \cdot e_j \end{aligned} \tag{69}$$

$$\leq \sum_{j \in J_1} \frac{b_j}{2\sqrt{m}\mu_j} \cdot v^j + \beta \tag{70}$$

$$= \sum_{j \in J_1} \frac{b_j}{2\sqrt{m}\mu_j} \cdot v^j + \frac{1}{2} \cdot v^{m+1}, \tag{71}$$

where (68) follows as $\beta^j \geq \mu_j \cdot e_j$ for all $j = 1, \dots, m$. Inequalities (69) and (70) follow as $\beta_j \geq b_j$ for all $j \in J_2$ and $\beta \geq 0$. Equation (71) follows as $v^{m+1} = 2\beta$. If $\lambda = 0$, then $b_j = 0$ for all $j \in J_1$ which implies that $b \leq 1/2 \cdot v^{m+1} \leq v^{m+1}$. Therefore, we can assume that $\lambda > 0$ and (71) can be further modified as follows.

$$\begin{aligned} b &\leq \sum_{j \in J_1} \frac{b_j}{2\sqrt{m}\mu_j} \cdot v^j + \frac{1}{2} \cdot v^{m+1} \\ &\leq \frac{1}{2\lambda} \cdot \left(\sum_{j \in J_1} \frac{b_j}{2\sqrt{m}\mu_j} \cdot v^j \right) + \frac{1}{2} \cdot v^{m+1} \end{aligned} \tag{72}$$

$$\begin{aligned} &= \sum_{j \in J_1} \frac{b_j}{4\lambda\sqrt{m}\mu_j} \cdot v_j + \frac{1}{2} \cdot v^{m+1} \\ &= \sum_{j=1}^{m+1} \alpha_j \cdot v^j, \end{aligned} \tag{73}$$

where (72) follows as $2\lambda \leq 1$ from (67) and for all $j = 1, \dots, m + 1$,

$$\alpha_j = \begin{cases} \frac{b_j}{4\lambda\sqrt{m}\mu_j}, & j \in J_1, \\ \frac{1}{2}, & j = m + 1, \\ 0, & \text{otherwise.} \end{cases}$$

Now,

$$\begin{aligned} \sum_{j=1}^{m+1} \alpha_j &= \sum_{j \in J_1} \alpha_j + \alpha_{m+1} \\ &= \sum_{j \in J_1} \frac{b_j}{4\lambda\sqrt{m}\mu_j} + \frac{1}{2} \\ &= \frac{1}{2\lambda} \cdot \left(\sum_{j \in J_1} \frac{b_j}{2\sqrt{m}\mu_j} \right) + \frac{1}{2} \\ &= \frac{1}{2\lambda} \cdot \lambda + \frac{1}{2} \\ &= 1, \end{aligned} \tag{74}$$

where (74) follows from (67). Therefore, $b' = \alpha_1 \cdot v^1 + \dots + \alpha_{m+1} \cdot v^{m+1}$ is a convex combination of extreme points of \mathcal{U}^0 which implies $b' \in \mathcal{U}^0$ and from (73) we have that $b \leq b'$. \square

We next show that $z_{Adapt}(\mathcal{U}^0)$ is at most a factor of $O(\sqrt{m})$ worse than $z_{Adapt}(\mathcal{U})$.

Lemma 13 $z_{Aff}(\mathcal{U}^0) = z_{Adapt}(\mathcal{U}^0) \leq 4\sqrt{m} \cdot z_{Adapt}(\mathcal{U})$.

Proof Note that since \mathcal{U}^0 is a simplex, an affine policy is optimal and can be computed in polynomial time as shown in Theorem 1. Therefore, $z_{Aff}(\mathcal{U}^0) = z_{Adapt}(\mathcal{U}^0)$. To prove the second inequality, consider an optimal solution $x^*, y^*(b)$ for all $b \in \mathcal{U}$ for $\Pi_{Adapt}(\mathcal{U})$. We construct the following approximate solution for $\Pi_{Adapt}(\mathcal{U}^0)$: $\hat{x} = 4\sqrt{m} \cdot x^*$ and for all $j = 1, \dots, m$, $\hat{y}(v^j) = 4\sqrt{m} \cdot y^*(\beta^j)$ and

$$\hat{y}(v^{m+1}) = \frac{4\sqrt{m}}{K} \cdot \left(y^*(u^1) + \dots + y^*(u^K) \right).$$

For any $b \in \mathcal{U}^0$, $b = \alpha_1 \cdot v^1 + \dots + \alpha_{m+1} \cdot v^{m+1}$, where $\alpha_j \geq 0$ for all $j = 1, \dots, m + 1$ and $\alpha_1 + \dots + \alpha_{m+1} = 1$. Therefore,

$$\hat{y}(b) = \sum_{j=1}^{m+1} \alpha_j \cdot \hat{y}(v^j).$$

We show the feasibility of the solution \hat{x} , $\hat{y}(b)$ for all extreme points of \mathcal{U}^0 and the feasibility of the solution at other points follows from a standard argument as in Lemma 1. For any $j = 1, \dots, m$,

$$\begin{aligned} A\hat{x} + B\hat{y}(v^j) &= 4\sqrt{m} \cdot (Ax^* + By^*(\beta^j)) \\ &\geq 4\sqrt{m} \cdot \beta^j \\ &\geq 2\sqrt{m} \cdot \beta^j \\ &= v^j. \end{aligned}$$

For $j = m + 1$,

$$\begin{aligned} A\hat{x} + B\hat{y}(v^{m+1}) &= A(4\sqrt{m} \cdot x^*) + B\left(\frac{4\sqrt{m}}{K} \cdot (y^*(u^1) + \dots + y^*(u^K))\right) \\ &= \frac{4\sqrt{m}}{K} \cdot \left(\sum_{k=1}^K (Ax^* + By^*(u^k))\right) \\ &\geq \frac{2\sqrt{m}}{K} \cdot 2 \left(\sum_{k=1}^K u^k\right) \\ &= \frac{2\sqrt{m}}{K} \cdot v^{m+1} \\ &\geq v^{m+1}, \end{aligned} \tag{75}$$

where (75) follows as $K \leq 2\sqrt{m}$ from Lemma 10. Therefore, the solution \hat{x} , $\hat{y}(b)$ is feasible for all $b \in \mathcal{U}^0$. For any $j = 1, \dots, m$,

$$\begin{aligned} c^T \hat{x} + d^T \hat{y}(v^j) &= 4\sqrt{m} \cdot (c^T x^* + d^T y^*(\beta^j)) \\ &\leq 4\sqrt{m} \cdot z_{Adapt}(\mathcal{U}), \end{aligned} \tag{76}$$

where (76) follows from the optimality of x^* , $y^*(b)$ for all $b \in \mathcal{U}$ for $\Pi_{Adapt}(\mathcal{U})$. Also,

$$\begin{aligned} c^T \hat{x} + d^T \hat{y}(v^{m+1}) &= 4\sqrt{m} \cdot c^T x^* + \frac{4\sqrt{m}}{K} \cdot \sum_{k=1}^K d^T y^*(u^k) \\ &= \frac{4\sqrt{m}}{K} \cdot \left(\sum_{k=1}^K (c^T x^* + d^T y^*(u^k))\right) \\ &\leq \frac{4\sqrt{m}}{K} \cdot \left(\sum_{k=1}^K z_{Adapt}(\mathcal{U})\right) \end{aligned} \tag{77}$$

$$= 4\sqrt{m} \cdot z_{Adapt}(\mathcal{U}), \tag{78}$$

where (77) follows from the optimality of $x^*, y^*(b)$ for all $b \in \mathcal{U}$ for $\Pi_{Adapt}(\mathcal{U})$ and the fact that $u^k \in \mathcal{U}$ for all $k = 1, \dots, K$. Therefore, we have the following.

$$\begin{aligned} z_{Adapt}(\mathcal{U}^0) &\leq \max_{b \in \mathcal{U}^0} (c^T \hat{x} + d^T \hat{y}(b)) \\ &= \max_{j=1, \dots, m+1} (c^T \hat{x} + d^T \hat{y}(v^j)) \end{aligned} \tag{79}$$

$$\leq 4\sqrt{m} \cdot z_{Adapt}(\mathcal{U}), \tag{80}$$

where (79) follows as the worst case cost is achieved at an extreme point of \mathcal{U}^0 and (80) follows from (76) and (78). □

Proof of Theorem 6 We need to show that an optimal first stage solution of $\Pi_{Adapt}(\mathcal{U}^0)$ is a good approximation for $\Pi_{Adapt}(\mathcal{U})$. Let $\tilde{x}, \tilde{y}(b)$ for all $b \in \mathcal{U}^0$ be an optimal solution for $\Pi_{Adapt}(\mathcal{U}^0)$ that can be computed in polynomial time. Let the first stage solution for $\Pi_{Adapt}(\mathcal{U})$ be $\bar{x} = \tilde{x}$ and for any $b \in \mathcal{U}$,

$$\bar{y}(b) = \operatorname{argmin}\{d^T y(b) \mid By \geq b - A\bar{x}\}.$$

We claim that the worst case cost of this solution is at most $z_{Adapt}(\mathcal{U}^0)$. From Lemma 12, we know that for any $b \in \mathcal{U}$, there exists $b' \in \mathcal{U}^0$ such that $b \leq b'$. Therefore, $\bar{y}(b) = \tilde{y}(b')$ is a feasible solution for $b \in \mathcal{U}$ as:

$$\begin{aligned} A\bar{x} + B\bar{y}(b) &= A\tilde{x} + B\tilde{y}(b') \\ &\geq b' \\ &\geq b. \end{aligned}$$

Therefore,

$$\min\{d^T y \mid Ax + By \geq b - A\bar{x}\} \leq d^T \tilde{y}(b'),$$

which implies that

$$\begin{aligned} c^T \bar{x} + \min\{d^T y \mid Ax + By \geq b - A\bar{x}\} &\leq c^T \tilde{x} + d^T \tilde{y}(b') \\ &\leq z_{Adapt}(\mathcal{U}^0) \\ &\leq 4\sqrt{m} \cdot z_{Adapt}(\mathcal{U}), \end{aligned}$$

where the last inequality follows from Lemma 13.

7 Conclusions

In this paper, we give a tight characterization of the performance of affine policies in the context of the two-stage adaptive optimization problem defined in (1). In particular, we show that the performance of an optimal affine policy can be a factor $\Omega(m^{1/2-\delta})$

worse as compared to an optimal full-adaptable solution for any $\delta > 0$. The uncertainty sets in the family of examples that achieve this lower bound have an exponential number of extreme points that lie on or near the surface of a unit hypersphere in the non-negative orthant. Therefore, the intersection of a hypersphere with the non-negative orthant seems to be essentially the worst-case example for affine policies. It would be interesting to find an explicit connection between the performance of affine policies and some geometric property of the worst-case set, i.e., the intersection of a unit hypersphere with the non-negative orthant. We would like to note that since we prove a lower bound on the performance ratio of affine policies for the model in (1), the lower bound clearly holds for more general models without some of the restrictions in (1).

We also present a matching upper bound of $O(\sqrt{m})$ on the performance of affine policies for (1) when the constraint coefficients for the first-stage decision variables are non-negative. Furthermore, if the number of uncertain right hand side coefficients is $k \leq m$, where m is the number of linear constraints in the model, we show that the affine policies are a $O(\sqrt{k})$ -approximation for the adaptive problem if all the constraint coefficients are non-negative. We also present an $O(\sqrt{m})$ -approximation algorithm (that is not an affine policy) if the non-negativity condition on the constraint coefficients is relaxed. Analyzing the performance of affine policies for more general models than (1) is an interesting open question.

To the best of our knowledge, our results provide the first bound on the performance of affine policies in such generality. Moreover, we are not aware of any other efficient algorithms for the general two-stage adaptive problem with a performance guarantee better than $O(\sqrt{m})$. Since the hardness of approximation is known to be $O(\log m)$ [11], reducing the gap is an interesting open problem. It would be interesting to study other tractable policies for adaptive optimization such as polynomial policies and piecewise affine policies both from a theoretical perspective as well as designing better policies that work well in practice.

Acknowledgments We thank the associate editor and the reviewers for their suggestions that improved the paper.

References

1. Bemporad, A., Borrelli, F., Morari, M.: Min–max control of constrained uncertain discrete-time linear systems. *IEEE Trans. Autom. Control* **48**(9), 1600–1606 (2003)
2. Ben-Tal, A., Boyd, S., Nemirovski, A.: Control of uncertainty-affected discrete time linear systems via convex programming. Technical report, Minerva Optimization Center, Technion, Haifa, Israel. *SIAM J. Control Optim.* (submitted, 2005)
3. Ben-Tal, A., Boyd, S., Nemirovski, A.: Extending the scope of robust optimization: comprehensive robust counterparts of uncertain problems. *Math. Program.* **107**(1), 63–89 (2006)
4. Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A.: Adjustable robust solutions of uncertain linear programs. *Math. Program.* **99**(2), 351–376 (2004)
5. Ben-Tal, A., Nemirovski, A.: Selected topics in robust convex optimization. *Math. Program.* **112**(1), 125–158 (2008)
6. Bertsimas, D., Brown, D.B.: Constrained stochastic LQC: a tractable approach. *IEEE Trans. Autom. Control* **52**(10), 1826–1841 (2007)
7. Bertsimas D., Goyal, V.: On the Power of Robust Solutions in Two-Stage Stochastic and Adaptive Optimization Problems. *Math. Oper. Res.* **35**(2), 284–305 (under review)

8. Bertsimas, D., Iancu, D.A., Parrilo, P.A.: A hierarchy of near-optimal policies for multi-stage adaptive optimization. Submitted to IEEE Trans. Autom. Control (under review)
9. Bertsimas, D., Iancu, D.A., Parrilo, P.A.: Optimality of affine policies in multistage robust optimization. *Math. Oper. Res.* **35**(2), 363–394 (2010)
10. Dhamdhere, K., Goyal, V., Ravi, R., Singh, M.: How to pay, come what may: Approximation algorithms for demand-robust covering problems. In: Annual Symposium on Foundations of Computer Science, vol. 46, pp. 367–376 (2005)
11. Feige, U., Jain, K., Mahdian, M., Mirrokni, V.: Robust combinatorial optimization with exponential scenarios. *Lect. Notes Comput. Sci.* **4513**, 439–453 (2007)
12. Goulart, P.J., Kerrigan, E.C.: Relationships between affine feedback policies for robust control with constraints. In: Proceedings of the 16th IFAC World Congress on Automatic Control (2005)
13. Kerrigan, E.C., Maciejowski, J.M.: On robust optimization and the optimal control of constrained linear systems with bounded state disturbances. In: Proceedings of the European Control Conference (2003)
14. Khandekar, R., Kortsarz, G., Mirrokni, V., Salavatipour, M.: Two-stage robust network design with exponential scenarios. In: Proceedings of the 16th Annual European Symposium on Algorithms, pp. 589–600. Springer, New York (2008)
15. Kuhn, D., Wiesemann, W., Georghiou, A.: Primal and dual linear decision rules in stochastic and robust optimization. *Math. Program.* 1–33 (2009)
16. Lofberg, J.: Approximations of closed-loop minimax MPC. In: Proceedings of the 42nd IEEE Conference on Decision and Control, vol. 2 (2003)
17. Rockafellar, R.T., Wets, R.J.B.: The optimal recourse problem in discrete time: L^1 -multipliers for inequality constraints. *SIAM J. Control Optim.* **16**, 16–36 (1978)
18. Skaf, J., Boyd S.: Design of affine controllers via convex optimization. *IEEE Trans. Autom. Control* **55**, 2476–2487 (2010)