On the Optimality of Affine Policies for Budgeted Uncertainty Sets

Omar El Housni

Dept of Industrial Engineering and Operations Research, Columbia University, New York NY 10027, oe2148@columbia.edu,
http://www.columbia.edu/ oe2148/

Vineet Goyal

Dept of Industrial Engineering and Operations Research, Columbia University, New York NY 10027, vg2277@columbia.edu,
http://www.columbia.edu/ vg2277/

In this paper, we study the performance of affine policies for two-stage adjustable robust optimization problem with uncertain right hand side belonging to a budgeted uncertainty set. This is an important class of uncertainty sets widely used in practice where we can specify a budget on the adversarial deviations of the uncertain parameters from the nominal values to adjust the level of conservatism. The two-stage adjustable robust optimization problem is hard to approximate within a factor better than $\Omega(\frac{\log n}{\log \log n})$ even for budget of uncertainty sets where n is the number of decision variables. Affine policies, where the second-stage decisions are constrained to be an affine function of the uncertain parameters, provide a tractable approximation for the problem and exhibit good empirical performance. We show that affine policies give an $O(\frac{\log n}{\log \log n})$-approximation for the two-stage adjustable robust problem for this class of uncertainty sets. This matches the hardness of approximation and therefore, surprisingly affine policies provide an optimal approximation for the problem (up to some constant factor). We also show strong theoretical performance bounds for affine policy for significantly more general class of intersection of budgeted sets including disjoint constrained budgeted sets, permutation invariant sets and general intersection of budgeted sets. Our analysis relies on showing the existence of a near-optimal feasible affine policy that satisfies certain nice structural properties. Based on these structural properties, we also present an alternate algorithm to compute near-optimal affine solution that is significantly faster than computing the optimal affine policy by solving a large linear program.

Key words: robust optimization; affine policies; budget of uncertainty
MSC2000 subject classification: Primary: 90C15 , 90C47

1. Introduction. In this paper, we consider the following two-stage adjustable robust linear optimization problem with uncertain right hand side:

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

where $\boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{B} \in \mathbb{R}^{m \times n}_+, \boldsymbol{c} \in \mathbb{R}^n_+$ and $\boldsymbol{d} \in \mathbb{R}^n_+$. The right-hand-side $\boldsymbol{h}$ is uncertain and belongs to a compact convex uncertainty set $\mathcal{U} \subseteq \mathbb{R}^m_+$. The goal in this problem is to select the first-stage decision $\boldsymbol{x}$, and the second-stage recourse decision, $\boldsymbol{y}(\boldsymbol{h})$, as a function of the uncertain right hand side realization, $\boldsymbol{h}$ such that the worst-case cost over all realizations of $\boldsymbol{h} \in \mathcal{U}$ is minimized. We assume that $\mathcal{U} \subseteq [0,1]^m$ and $\forall i \in [m], \boldsymbol{e}_i \in \mathcal{U}$. This assumption is without loss of generality since we can scale the constraint matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ to satisfy the assumption without changing the optimal.

This model has been widely considered in the literature (see for example [7, 9, 17, 19, 22, 28, 29]). It captures many important applications including set cover, facility location, capacity planning

and network design problems under uncertain demand. Here the right hand side, $\boldsymbol{h}$ models the uncertain demand and the covering constraints capture the requirement of satisfying the uncertain demand. However, the adjustable robust optimization problem (1.1) is intractable in general. In fact, Feige et al. [20] show that the two-stage adjustable problem (1.1) can not be approximated within a ratio better than $\Omega(\frac{\log n}{\log \log n})$ under a reasonable complexity assumption, namely, 3SAT can not be solved in time $2^{O(\sqrt{n})}$ on instances of size n.

In view of the intractability, several approximation policies (or *decision rules*) have been considered in the literature for (1.1) including static, piecewise static, affine and piecewise affine policies. In a static policy, we compute a single optimal solution $(\boldsymbol{x}, \boldsymbol{y})$ that is feasible for all realizations of the uncertain right hand side. Bertsimas et al. [10] relate the performance of static solution to the symmetry of the uncertainty set and show that it provides a good approximation to the adjustable problem if the uncertainty set verifies some symmetry properties. However, static policy is too conservative in general and the performance of static solutions can be arbitrarily large for a general convex uncertainty set with the worst-case performance being $\Omega(m)$.

Ben-Tal et al. [4] introduce affine policy approximation for (1.1), where they restrict the second-stage decision, $\boldsymbol{y}(\boldsymbol{h})$ to being an affine function of the uncertain right-hand-side $\boldsymbol{h}$, i.e., $\boldsymbol{y}(\boldsymbol{h}) = \boldsymbol{Ph} + \boldsymbol{q}$ for some decision variables $\boldsymbol{P} \in \mathbb{R}^{n \times m}$ and $\boldsymbol{q} \in \mathbb{R}^n$. Affine policy can be computed efficiently for a large class of uncertainty sets and therefore, provide a tractable approximation for the two-stage problem. Furthermore, the empirical performance of affine policies has been observed to be near-optimal for a large class of instances even though theoretically, optimality of affine policies is known in only a few settings. Bertsimas et al. [11] and Iancu et al. [25] show that affine policy is optimal for multi-stage adjustable problems with a single uncertain parameter at each stage. Bertsimas and Goyal [9] show that affine policy is optimal for the two-stage adjustable problem (1.1) if the uncertainty set $\mathcal{U}$ is a simplex. However, the worst-case performance of an optimal affine solution is $\Theta(\sqrt{m})$ times the cost of an optimal dynamic solution for (1.1) [9]. Therefore, there is a significant gap between the worst-case performance of affine policies and the observed empirical performance.

More general decision rules have been considered in the literature for two-stage problems; piecewise affine policies (Ben-Tal et al. [3]), binary decision rules (Bertsimas and Georghiou [8]), adjustable solutions via iterative splitting of uncertainty sets (Postek and Den Hertog [26]), k-adaptibility (Hanasusanto et al. [24]), extended affine decision rules (Chen et al. [15]), etc. While these decision rules can improve in some instances over affine policies, they become computationally very challenging especially for large size instances. For a more extensive review of the literature, we refer the reader to Bertsimas et al. [6] and Ben-Tal et al. [2].

In this paper, we study the performance of affine policies for the two-stage adjustable problem (1.1). One of our important goals in this work is to address the stark contrast between the observed near-optimal empirical performance and the worst-case approximation bound of $\Theta(\sqrt{m})$ [9]. Towards this, we consider the class of budget of uncertainty sets and intersection of budget of uncertainty sets that was introduced in Bertsimas and Sim [12]. This is widely used class of uncertainty sets in practice where the decision maker can specify a budget on the sum of adversarial deviations of the uncertain parameter from the nominal values. In particular, a budget of uncertainty set can be formulated as follows:

$$\mathcal{U} = \left\{ \boldsymbol{h} \in [0,1]^m \,\middle|\, \sum_{i=1}^{m} w_i h_i \le 1 \right\}, \tag{1.2}$$

where $w_i \in [0,1]$ for all $i \in [m]$. It is known that the two-stage adjustable problem (1.1) is hard to approximate under this class of uncertainty set. In particular, Feige et al. [20] show that the two-stage adjustable problem (1.1) where $\mathcal{U}$ is the budget of uncertainty set (1.2) is hard to approximate

within a factor $\Omega(\frac{\log n}{\log \log n})$ even when all w_i are equal and $\boldsymbol{A}, \boldsymbol{B}$ are 0-1 matrices. The analysis in Bertsimas and Goyal [9] shows that affine policy gives $O(\sqrt{m})$-approximation to the adjustable problem (1.1) under this class of uncertainty sets. Bertsimas and Bidkhori [5] provide a geometric bound $O(\frac{k^2+mk}{k^2+m})$ in the special case of (1.2) where all $w_i = 1/k$ which is also $O(\sqrt{m})$ in the worst-case for $k = \sqrt{m}$.

The above two-stage adjustable robust problem (1.1) has also been considered in the context of combinatorial optimization problems such as facility location and network design where the constraint matrices, $\boldsymbol{A}, \boldsymbol{B} \in \{0,1\}^{m \times n}$ and the first-stage and second-stage decisions are constrained to be binary (see for instance, Dhamdhere et al. [18], Feige et al. [20], Gupta et al. [22] and [23]). Feige et al. [20] and Gupta et al. [22] give an $O(\log n)$-approximation for (1.1) for the special case when $\boldsymbol{A} = \boldsymbol{B} \in \{0,1\}^{m \times n}$, first and second-stage costs are proportional, i.e., $\boldsymbol{d} = \lambda \cdot \boldsymbol{c}$ for some constant $\lambda \geq 1$ and a budget of uncertainty set with $w_i = 1/k$. Gupta et al. [23] also consider more general uncertainty set, namely, intersection of *p-system* and *q-knapsack*[1] and give $O(pq \log n)$-approximation for the two-stage problem. However, we would like to note that the focus of this stream of work is to design approximation algorithms for combinatorial optimization problems where the decisions are constrained to be binary. Moreover, the algorithms are not restricted to and do not necessarily give decision rules or functional policy approximations for the two-stage problem.

In contrast, the focus of our work is to analyze the performance of affine policies for the two-stage adjustable robust problem (1.1) that are widely used in practice and exhibit strong empirical performance.

1.1. Our contributions. Our main contributions are the following.

(a) **Optimal Approximation for Budget of Uncertainty Sets.** We show that affine policy gives an optimal approximation for the two-stage adjustable robust problem for budgeted uncertainty sets. In particular, affine policy gives an $O(\frac{\log n}{\log \log n})$-approximation to the two-stage adjustable problem (1.1) where $\mathcal{U}$ is a budget of uncertainty set (1.2). This performance bound matches the hardness of approximation [20]; thereby, showing that surprisingly affine policies give an optimal approximation (up to some constant factor) for (1.1) for budget of uncertainty sets. In other words, there is no polynomial time algorithm with worst-case approximation guarantee better than affine policies by more than some constant factor. This bound significantly improves over the previous known bound of $O(\sqrt{m})$ [9, 5] for budget of uncertainty sets.

Our analysis relies on constructing a feasible affine solution whose worst-case cost is within a factor $O(\frac{\log n}{\log \log n})$ of the optimal cost. In particular, we partition the components of $\mathcal{U}$ into *inexpensive* and *expensive* components based on a threshold and construct an affine solution that covers the inexpensive components using a linear solution and the expensive components using a static solution. We show that for an appropriately chosen threshold that depends on the optimal cost, such an affine solution gives an $O(\frac{\log n}{\log \log n})$-approximation for the two-stage problem for the budget of uncertainty set.

Therefore, in addition to establishing the performance bound that matches the hardness of approximation, our analysis shows there is a near-optimal affine solution whose structure is closely related to *threshold policies* that are widely used in many applications. This structural property might be of independent interest and also gives an alternate faster algorithm for computing near-optimal affine solutions for budget of uncertainty sets as we discuss later.

---

[1] p-system and q-knapsack are generalizations of matroid constraints. For more details, we refer the reader to [23]

(b) **Intersection of Budgeted Sets.** We also consider a more general family of uncertainty sets, namely the following intersection of budget of uncertainty sets:

$$\mathcal{U} = \left\{ \boldsymbol{h} \in [0,1]^m \ \Big| \ \sum_{i \in S_\ell} w_{\ell i} h_i \leq 1 \ \forall \ell \in [L] \right\}, \tag{1.3}$$

where $\boldsymbol{w}_l \in [0,1]^m$, and $S_\ell \subseteq [m]$ for all $\ell \in [L]$. The set (1.3) is defined by the intersection of L budget constraints. These are an important generalization of the budget of uncertainty set (1.2) that are widely used in practice. They capture for instance CLT sets [1] and *inclusion-constrained budgeted* sets [21].

  (i) We first consider the case when the family of subsets S_ℓ for $\ell \in [L]$ are disjoint. We refer to this class of sets as *disjoint constrained budgeted sets*. These are essentially Cartesian product of L budget of uncertainty sets. We show that affine policy is near-optimal and gives an $O\left(\log^2 n / \log\log n\right)$-approximation to (1.1) for this class of sets. We would like to note that the bound is independent of L. Similar to the case of budget of uncertainty sets, our analysis is based on constructing a near-optimal affine solution by partitioning components of $\mathcal{U}$ into *inexpensive* and *expensive* components using appropriate thresholds for each of the L budgeted sets in the Cartesian product. However, in this case, we are able to relate the performance of our affine solution to only a lower bound of $z_{\mathsf{AR}}(\mathcal{U})$. In particular, we use an *online algorithm* for the *fractional covering problem* to both construct thresholds (and therefore, a feasible affine solution) as well as the lower bound of the optimal.

  (ii) For general intersection of L budgets. We show that affine policy gives $O\left(\log L \log n / \log\log n\right)$ to (1.1) for the case where $\mathcal{U}$ is *permutation invariant*. We say that $\mathcal{U}$ is permutation invariant if for any $\boldsymbol{h} \in \mathcal{U}$ and any τ permutation of $[m]$, then $\boldsymbol{h}^\tau \in \mathcal{U}$ where $h_i^\tau = h_{\tau(i)}$. This class captures many important sets such as CLT sets. The performance of affine policy depends on L in this case but degrades gracefully. For general intersection of budgeted sets, we show a worst-case bound of $O\left(L \log n / \log\log n\right)$ on the performance of affine policy for (1.1). We summarize our results in Table 1.

(c) **Faster Algorithm to compute Near-Optimal Affine Solutions.** Based on the structural properties of the near-optimal affine policies constructed for analysis of performance, we present an alternate algorithm to compute an approximate affine policy for (1.1) for budget of uncertainty sets that is significantly faster than computing optimal affine policy by solving a large LP. In particular, our construction partitions the components into *inexpensive* and *expensive* based on a threshold depending on the optimal cost and shows the existence of a near-optimal affine solution that covers *inexpensive* component using a linear solution and the *expensive* components using a static solution.

From an algorithmic perspective, while we do not know the optimal cost and therefore, the threshold, we can still use this structure of a near-optimal affine solution to construct a good affine solution. In particular, we consider all partitions of $[m]$ into expensive and inexpensive components. Since our analysis partition is based on a threshold, we only have $(m+1)$ possibilities. For each possible partition, the affine solution can be computed very efficiently by solving a single LP covering problem with $O(n)$ variables as opposed to $O(nm)$ variables in the optimal affine policy. Our algorithm scales very well and is significantly faster than computing affine policies. For instance, our algorithm takes less than a second to compute our approximate affine policy whereas the optimal affine policy takes few minutes on average for $m = n = 100$. Moreover, our solution is within 8% to 15% of the optimal affine solution and the sub-optimality gap does not increase with dimension in our numerical experiments. We would like to note that since our approximate affine is based on the construction of affine policy in our analysis, the worst-case approximation bound for the faster algorithm is also $O(\frac{\log n}{\log\log n})$.

	Uncertainty sets	Our Bounds
1.	Budget of uncertainty set (1.2)	$O\left(\dfrac{\log n}{\log\log n}\right)^{*}$
2.	Disjoint Intersection of Budgeted sets (4.1)	$O\left(\dfrac{\log^{2} n}{\log\log n}\right)$
3.	Permutation Invariant Intersection of Budgeted sets (1.3)	$O\left(\dfrac{\log L\cdot\log n}{\log\log n}\right)$
4.	General Intersection of Budgeted sets (1.3)	$O\left(\dfrac{L\cdot\log n}{\log\log n}\right)$

TABLE 1. Our performance bounds for affine policy under different uncertainty sets including budget of uncertainty set and intersection of budgeted sets. * denotes that this bound mathches the approximation hardness of (1.1).

2. Affine policies: Preliminaries. Affine policies (also known as *linear decision rules*) are widely used in the literature of robust optimization. They were introduced by Ben-Tal et al. [4] for the two-stage adjustable problem (1.1). In an affine solution, we restrict the second stage decision $\boldsymbol{y}(\boldsymbol{h})$ to be an affine function of the uncertain parameter $\boldsymbol{h}$, i.e.,

$$\boldsymbol{y}(\boldsymbol{h}) = \boldsymbol{Ph} + \boldsymbol{q},$$

and we optimize over the variables $\boldsymbol{P}$ and $\boldsymbol{q}$. The affine problem is formulated as:

$$
\begin{aligned}
z_{\mathsf{Aff}}(\mathcal{U}) = \min_{\boldsymbol{x},\boldsymbol{P},\boldsymbol{q}} \; & \boldsymbol{c}^{T}\boldsymbol{x} + \max_{\boldsymbol{h}\in\mathcal{U}} \boldsymbol{d}^{T}\left(\boldsymbol{Ph}+\boldsymbol{q}\right) \\
& \boldsymbol{Ax} + \boldsymbol{B}\left(\boldsymbol{Ph}+\boldsymbol{q}\right) \;\geq\; \boldsymbol{h}, \;\; \forall \boldsymbol{h}\in\mathcal{U} \\
& \boldsymbol{Ph} + \boldsymbol{q} \;\geq\; \boldsymbol{0}, \;\; \forall \boldsymbol{h}\in\mathcal{U} \\
& \boldsymbol{x}\in\mathbb{R}^{n}_{+}.
\end{aligned}
\tag{2.1}
$$

Affine policy has been widely used as an approximation to (1.1) due to its tractability. In fact, Ben-Tal et al. [4] show that affine problems have an equivalent standard LP formulation when the uncertainty set is described by a polyhedral set. The size of the LP is polynomial in the size of the input parameters (i.e., number of variables and constraints in (1.1) and number of constraints in $\mathcal{U}$). For completeness, we briefly discuss the tractability and compact LP formulation of affine policies. Consider the following polyhedral uncertainty set

$$\mathcal{U} = \{\boldsymbol{h}\in\mathbb{R}^{m}_{+} \mid \boldsymbol{Rh}\leq\boldsymbol{r}\}, \tag{2.2}$$

where $\boldsymbol{R}\in\mathbb{R}^{L\times m}$ and $\boldsymbol{r}\in\mathbb{R}^{L}$. The affine problem (2.1) can be reformulated as the following epigraph formulation:

$$
\begin{aligned}
z_{\mathsf{Aff}}(\mathcal{U}) = \min \; & \boldsymbol{c}^{T}\boldsymbol{x} + z \\
& z \geq \boldsymbol{d}^{T}\left(\boldsymbol{Ph}+\boldsymbol{q}\right), \;\; \forall \boldsymbol{h}\in\mathcal{U} \\
& \boldsymbol{Ax} + \boldsymbol{B}\left(\boldsymbol{Ph}+\boldsymbol{q}\right) \geq \boldsymbol{h}, \;\; \forall \boldsymbol{h}\in\mathcal{U} \\
& \boldsymbol{Ph}+\boldsymbol{q} \geq \boldsymbol{0}, \;\; \forall \boldsymbol{h}\in\mathcal{U} \\
& \boldsymbol{x}\in\mathbb{R}^{n}_{+}, \; \boldsymbol{P}\in\mathbb{R}^{n\times m}, \; \boldsymbol{q}\in\mathbb{R}^{n}, \; z\in\mathbb{R}.
\end{aligned}
$$

Note that this formulation can have infinitely many constraints but the separation problem is tractable. For example, the separation problem for the first set of constraints is: Given $z, \boldsymbol{x}, \boldsymbol{P}, \boldsymbol{q}$ decide if

$$z - \boldsymbol{d}^{T}\boldsymbol{q} \geq \max_{\boldsymbol{h}\geq\boldsymbol{0}}\{\boldsymbol{d}^{T}\boldsymbol{Ph} \mid \boldsymbol{Rh}\leq\boldsymbol{r}\}. \tag{2.3}$$

This can be done efficiently by solving the above maximization LP. Moreover, Ben-Tal et al. [4] show that we can formulate (2.1) as a compact LP using standard techniques from duality. For

instance, consider the first set of constraints (2.3), by taking the dual of the maximization problem, the constraint becomes

$$z - \boldsymbol{d}^T \boldsymbol{q} \geq \min_{\boldsymbol{v} \geq \boldsymbol{0}} \{\boldsymbol{r}^T \boldsymbol{v} \mid \boldsymbol{R}^T \boldsymbol{v} \geq \boldsymbol{P}^T \boldsymbol{d}\}.$$

We can then drop the min and introduce $\boldsymbol{v}$ as a variable. Hence, we obtain the following linear constraints:

$$z - \boldsymbol{d}^T \boldsymbol{q} \geq \boldsymbol{r}^T \boldsymbol{v}, \qquad \boldsymbol{R}^T \boldsymbol{v} \geq \boldsymbol{P}^T \boldsymbol{d}, \qquad \boldsymbol{v} \geq \boldsymbol{0}.$$

We can apply the same techniques for the other constraints. For completeness, we restate the compact LP formulation of Ben-Tal et al. [4] adapted to our problem in the following lemma. The proof of the lemma is deferred to Appendix C.

LEMMA 2.1. *The affine problem* (2.1) *can be formulated as the following LP*

$$
\begin{aligned}
z_{\mathsf{Aff}}(\mathcal{U}) = \min \; & \boldsymbol{c}^T \boldsymbol{x} + z \\
& z - \boldsymbol{d}^T \boldsymbol{q} \geq \boldsymbol{r}^T \boldsymbol{v} \\
& \boldsymbol{R}^T \boldsymbol{v} \geq \boldsymbol{P}^T \boldsymbol{d} \\
& \boldsymbol{A} \boldsymbol{x} + \boldsymbol{B} \boldsymbol{q} \geq \boldsymbol{V}^T \boldsymbol{r} \\
& \boldsymbol{R}^T \boldsymbol{V} \geq \boldsymbol{I}_m - \boldsymbol{B} \boldsymbol{P} \\
& \boldsymbol{q} \geq \boldsymbol{U}^T \boldsymbol{r} \\
& \boldsymbol{U}^T \boldsymbol{R} + \boldsymbol{P} \geq \boldsymbol{0} \\
& \boldsymbol{x} \in \mathbb{R}^n_+, \; \boldsymbol{v} \in \mathbb{R}^L_+, \; \boldsymbol{U} \in \mathbb{R}^{L \times n}_+, \; \boldsymbol{V} \in \mathbb{R}^{L \times m}_+ \\
& \boldsymbol{P} \in \mathbb{R}^{n \times m}, \; \boldsymbol{q} \in \mathbb{R}^n, \; z \in \mathbb{R}.
\end{aligned}
\tag{2.4}
$$

3. Performance analysis for budget of uncertainty sets. In this section, we consider the class of budget of uncertainty sets (1.2) given by

$$\mathcal{U} = \left\{ \boldsymbol{h} \in [0,1]^m \; \middle| \; \sum_{i=1}^m w_i h_i \leq 1 \right\}.$$

As we mention earlier, this class is widely used in the literature of robust optimization both in theory and practice. It provides the flexibility to adjust the level of conservatism in terms of probabilistic bounds on constraint violations. A special case of this class is when w_i are all equal to $\frac{1}{k}$ for some parameter $k \in \mathbb{N}$. In particular, in this case we have

$$\mathcal{U} = \left\{ \boldsymbol{h} \in [0,1]^m \; \middle| \; \sum_{i=1}^m h_i \leq k \right\}.
\tag{3.1}$$

The parameter k is the *budget of uncertainty* that controls the conservatism of the uncertainty model. This special class (3.1) of budgeted sets is also known as the *cardinality constrained set* or *k-ones polytope.*

The two-stage adjustable problem (1.1) is known to be $\Omega(\frac{\log n}{\log \log n})$-hard to approximate under the class of budget of uncertainty sets even in the special case of (3.1) (Feige et al. [20]). We show that surprisingly the performance bound for affine policy matches this hardness of approximation In particular, we show that affine policy gives $O(\frac{\log n}{\log \log n})$-approximation for (1.1) under budget of uncertainty sets (1.2).

THEOREM 3.1. *Consider the two-stage adjustable problem* (1.1) *where* $\mathcal{U}$ *is the budget of uncertainty set* (1.2) . *Then,*

$$z_{\mathsf{Aff}}(\mathcal{U}) = O\left(\frac{\log n}{\log \log n}\right) \cdot z_{\mathsf{AR}}(\mathcal{U}).$$

Our analysis significantly improves over the previous best known bound of $O(\sqrt{m})$ for the performance of affine policies for budget of uncertainty sets. Furthermore, since our performance bound matches the hardness of approximation, affine policy provides an optimal approximation for (1.1) for budget of uncertainty sets. In particular, there is no polynomial time algorithm whose worst-case approximation is better than affine policies by more than a constant factor. Note that the above statements relate to the worst-case performance. For particular instances, it may be possible to get better solutions than affine policies.

3.1. Construction of our affine solution. Our analysis is based on constructing a *good* approximate affine solution that has a worst-case cost being $O\left(\frac{\log n}{\log\log n}\right)$ times the optimal cost $z_{\text{AR}}(\mathcal{U})$. In this section, we present the construction of our affine solution and consequently prove Theorem 3.1. Let us first introduce the following notations.

Notations. We consider an optimal solution $\boldsymbol{x}^*, \boldsymbol{y}^*(\boldsymbol{h})$ where $\boldsymbol{h} \in \mathcal{U}$, for the adjustable problem (1.1). Let OPT be the optimal cost for (1.1) and $\text{OPT}_1, \text{OPT}_2$ respectively the first stage cost and the second stage cost associated with $\boldsymbol{x}^*, \boldsymbol{y}^*(\boldsymbol{h})$, i.e.

$$\text{OPT}_1 = \boldsymbol{c}^T \boldsymbol{x}^*$$
$$\text{OPT}_2 = \max_{\boldsymbol{h} \in \mathcal{U}} \boldsymbol{d}^T \boldsymbol{y}^*(\boldsymbol{h})$$
$$\text{OPT} = \text{OPT}_1 + \text{OPT}_2 = z_{\text{AR}}(\mathcal{U}).$$

We would like to remark that this split is not unique since there might be other optimal solutions for (1.1). We denote $z(\boldsymbol{h})$, the cost of covering the requirement $\boldsymbol{h}$ in the second-stage, i.e.

$$z(\boldsymbol{h}) = \min_{\boldsymbol{y} \geq \boldsymbol{0}} \left\{ \boldsymbol{d}^T \boldsymbol{y} \,\middle|\, \boldsymbol{B} \boldsymbol{y} \geq \boldsymbol{h} \right\}. \tag{3.2}$$

We refer to problem (3.2) as the *fractional covering problem*. For any $\mathcal{W}$ a subset of $[m]$, we denote $\mathbb{1}(\mathcal{W}) \in \mathbb{R}^m$ the indicator of $\mathcal{W}$, i.e.

$$\mathbb{1}(\mathcal{W}) = \sum_{i \in \mathcal{W}} \boldsymbol{e}_i.$$

For simplicity, we use the following notation

$$z(\mathbb{1}(\mathcal{W})) = z(\mathcal{W}).$$

Our construction. For all $i \in [m]$, recall $z(\boldsymbol{e}_i)$ the optimal cost to cover component $\boldsymbol{e}_i$ in the second stage as defined in (3.2). Let $\boldsymbol{v}_i$ be the optimal corresponding solution, i.e.

$$\boldsymbol{v}_i \in \arg\min_{\boldsymbol{y} \geq \boldsymbol{0}} \left\{ \boldsymbol{d}^T \boldsymbol{y} \,\middle|\, \boldsymbol{B} \boldsymbol{y} \geq \boldsymbol{e}_i \right\}.$$

We split the components $\{1, 2, \ldots, m\}$ based on a threshold into two sets $\mathcal{I}$ and its complement $\mathcal{I}^c$:

$$\mathcal{I} = \left\{ i \in [m] \,\middle|\, \frac{z(\boldsymbol{e}_i)}{w_i} \leq \beta \cdot \text{OPT} \right\}$$
$$\mathcal{I}^c = [m] \setminus \mathcal{I},$$

where

$$\beta = \frac{8 \log n}{\log \log n}.$$

We cover $\mathcal{I}$ (*inexpensive components*) using a linear solution and $\mathcal{I}^c$ (*expensive components*) using a static solution.

Linear part. We cover the components of $\mathcal{I}$ using the following linear solution for any $\boldsymbol{h} \in \mathcal{U}$,

$$\boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h}) = \sum_{i \in \mathcal{I}} h_i \boldsymbol{v}_i. \tag{3.3}$$

Static part. We use a static solution to cover the remaining components $\boldsymbol{e}_i$ where $i \in \mathcal{I}^c$. In particular, we consider the following static problem

$$(\boldsymbol{x}_{\mathsf{Sta}}, \boldsymbol{y}_{\mathsf{Sta}}) \in \operatorname*{arg\,min}_{\boldsymbol{x}, \boldsymbol{y} \geq \boldsymbol{0}} \left\{ \boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{d}^T \boldsymbol{y} \;\middle|\; \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \geq \sum_{i \in \mathcal{I}^c} \boldsymbol{e}_i \right\}, \tag{3.4}$$

and denote

$$z_{\mathsf{Sta}} = \boldsymbol{c}^T \boldsymbol{x}_{\mathsf{Sta}} + \boldsymbol{d}^T \boldsymbol{y}_{\mathsf{Sta}}.$$

Therefore our candidate affine solution is

$$\begin{aligned} \boldsymbol{x} &= \boldsymbol{x}_{\mathsf{Sta}} \\ \boldsymbol{y}(\boldsymbol{h}) &= \boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h}) + \boldsymbol{y}_{\mathsf{Sta}}, \; \forall \boldsymbol{h} \in \mathcal{U}. \end{aligned} \tag{3.5}$$

Feasibility. We first show that our candidate solution (3.5) is feasible for the adjustable problem (1.1). The proof is a direct consequence of our construction. In particular, we have the following lemma.

LEMMA 3.1. *The affine solution in* (3.5) *is feasible for the adjustable problem* (1.1).

Proof. We have,

$$\boldsymbol{B}\boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h}) = \sum_{i \in \mathcal{I}} h_i \boldsymbol{B} \boldsymbol{v}_i \geq \sum_{i \in \mathcal{I}} h_i \boldsymbol{e}_i,$$

and

$$\boldsymbol{A}\boldsymbol{x}_{\mathsf{Sta}} + \boldsymbol{B}\boldsymbol{y}_{\mathsf{Sta}} \geq \sum_{i \in \mathcal{I}^c} \boldsymbol{e}_i.$$

Therefore, the solution in (3.5) verifies

$$\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{h}) \geq \sum_{i \in \mathcal{I}} h_i \boldsymbol{e}_i + \sum_{i \in \mathcal{I}^c} \boldsymbol{e}_i \geq \boldsymbol{h},$$

where the last inequality holds because $h_i \in [0, 1]$ for all $i \in [m]$. □

We would like to remark that the construction of $\mathcal{I}$ (linear part) and $\mathcal{I}^c$ (static part) depends on all the parameters of the instance, i.e., $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{c}, \boldsymbol{d}$ and $\mathcal{U}$ and not only the uncertainty set $\mathcal{U}$. This is in contrast to the analysis in [9] where the construction of affine policies depends only on $\mathcal{U}$.

3.2. Performance analysis. We analyze separately the cost of the static and linear parts. For the linear part, the cost analysis is a direct consequence of our construction. In fact, we leave only inexpensive scenarios to the linear part, i.e., scenarios $\boldsymbol{e}_i$ such that $z(\boldsymbol{e}_i)/w_i$ is less than the threshold $\beta \cdot \mathrm{OPT}$. We know that for all $\boldsymbol{h} \in \mathcal{U}$, we have $\sum_{i=1}^m w_i h_i \leq 1$. Hence, the cost of linear part is at most $\beta \cdot \mathrm{OPT}$. In particular, we have the following lemma.

LEMMA 3.2 **(Cost of Linear part).** *The cost of the linear part* $\boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h})$ *defined in* (3.3) *is at most* $\beta \cdot \mathrm{OPT}$ *for any* $\boldsymbol{h} \in \mathcal{U}$.

Proof. We have for all $\boldsymbol{h} \in \mathcal{U}$,

$$\boldsymbol{d}^T \boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h}) = \sum_{i \in \mathcal{I}} h_i \boldsymbol{d}^T \boldsymbol{v}_i = \sum_{i \in \mathcal{I}} h_i z(\boldsymbol{e}_i) \leq \beta \cdot \mathrm{OPT} \cdot \sum_{i \in \mathcal{I}} w_i h_i \leq \beta \cdot \mathrm{OPT},$$

where the first inequality holds because $z(\boldsymbol{e}_i) \leq w_i \beta \cdot \mathrm{OPT}$ for all $i \in \mathcal{I}$ and the second inequality follows as $\sum_{i \in \mathcal{I}} w_i h_i \leq 1$ for all $\boldsymbol{h} \in \mathcal{U}$. □

The key part is to analyze the cost of the static part. In fact, we show that the cost of the static part is also $O(\beta) \cdot \text{OPT}$. This relies on a structural result on fractional covering problems. Intuitively, let us explain the structural result in the special case (3.1) of the budget of uncertainty set. We show that if the cost of covering every single component $\boldsymbol{e}_i$, for $i \in \mathcal{J}$ is *expensive*, i.e. $z(\boldsymbol{e}_i) > \beta \cdot \text{OPT}/k$ and the cost of covering any k components is *inexpensive*, i.e. less than 2OPT. Then, the cost of covering all components of $\mathcal{J}$ is not too costly and can not exceed $\beta \cdot \text{OPT}$. The formal general statement is given in the following lemma.

LEMMA 3.3 (**Structural Result**). *Consider $\boldsymbol{B} \in \mathbb{R}_+^{m \times n}$, $\boldsymbol{d} \in \mathbb{R}_+^n$ and $\mathcal{J} \subseteq [m]$. Let $z(\boldsymbol{h})$ be the cost of covering $\boldsymbol{h}$ as defined in (3.2). Suppose there exists $\gamma > 0$ and $w_i \in (0,1]$, $\forall i \in \mathcal{J}$ such that the following two conditions are satisfied:*
1. *for all $i \in \mathcal{J}$,*

$$\frac{z(\boldsymbol{e}_i)}{w_i} > 4\gamma \cdot \frac{\log n}{\log \log n},$$

2. *for all $\mathcal{W} \subseteq \mathcal{J}$,*

$$\sum_{i \in \mathcal{W}} w_i \leq 1 \ \ implies \ \ z(\mathcal{W}) \leq \gamma.$$

Then,

$$z(\mathcal{J}) \leq 4\gamma \cdot \frac{\log n}{\log \log n}.$$

We present the proof of Lemma 3.3 later in Section 3.3. But let us first use the structural result to show that the cost of the static part is $O(\beta) \cdot \text{OPT}$ and consequently prove Theorem 3.1. In particular, we have the following lemma.

LEMMA 3.4 (**Cost of Static part**). *The cost z_{Sta} of the static part $(\boldsymbol{x}_{\text{Sta}}, \boldsymbol{y}_{\text{Sta}})$ defined in (3.4) is $O(\beta) \cdot \text{OPT}$.*

Proof. Consider the following set

$$\mathcal{J}_1 = \left\{ i \in [m] \ \middle| \ (\boldsymbol{A}\boldsymbol{x}^*)_i \geq \frac{1}{2} \right\}.$$

We have for all $i \in \mathcal{J}_1$, $2\boldsymbol{A}\boldsymbol{x}^* \geq \boldsymbol{e}_i$ and we know that $2\boldsymbol{c}^T\boldsymbol{x}^* = 2\text{OPT}_1$. In other words, we cover the components of $\mathcal{J}_1$ using the static solution $2\boldsymbol{x}^*$ with a cost 2OPT_1. Now, define the set of remaining components

$$\mathcal{J}_2 = \mathcal{I}^c \setminus \mathcal{J}_1.$$

We apply the structural Lemma 3.3 with $\mathcal{J}_2$ and $\gamma = 2\text{OPT}$. In fact, because $\mathcal{J}_2 \subseteq \mathcal{I}^c$, we have for all $i \in \mathcal{J}_2$,

$$\frac{z(\boldsymbol{e}_i)}{w_i} > \beta \cdot \text{OPT} = 4\gamma \cdot \frac{\log n}{\log \log n},$$

i.e., the first condition of Lemma 3.3 is satisfied. Moreover, for any $\mathcal{W} \subseteq \mathcal{J}_2$ such that $\sum_{i \in \mathcal{W}} w_i \leq 1$, we have $\boldsymbol{h} = \mathbb{1}(\mathcal{W}) \in \mathcal{U}$. By feasibility of the optimal solution, we know that

$$\boldsymbol{A}\boldsymbol{x}^* + \boldsymbol{B}\boldsymbol{y}^*(\boldsymbol{h}) \geq \boldsymbol{h}.$$

Furthermore, for all $i \in \mathcal{W}$, we have $i \notin \mathcal{J}_1$ and therefore $2(\boldsymbol{A}\boldsymbol{x}^*)_i < 1$ which implies

$$2\boldsymbol{A}\boldsymbol{x}^* < \mathbb{1}(\mathcal{W}) = \boldsymbol{h}.$$

Hence,

$$2\boldsymbol{B}\boldsymbol{y}^*(\boldsymbol{h}) \geq \boldsymbol{h}.$$

This means that $2\boldsymbol{y}^*(\boldsymbol{h})$ is a feasible solution for the covering problem (3.2) with requirement $\boldsymbol{h}$, therefore,

$$z(\mathcal{W}) = z(\boldsymbol{h}) \leq 2\boldsymbol{d}^T \boldsymbol{y}^*(\boldsymbol{h}) \leq 2\mathrm{OPT}_2 \leq 2\mathrm{OPT} = \gamma.$$

i.e. condition 2 of Lemma 3.3 is also satisfied. Therefore from the structural result, we have

$$z(\mathcal{J}_2) \leq 4\gamma \frac{\log n}{\log \log n} = \beta \cdot \mathrm{OPT}.$$

Denote $\boldsymbol{y}$ an optimal solution corresponding to $z(\mathcal{J}_2)$, i.e. $\boldsymbol{d}^T \boldsymbol{y} = z(\mathcal{J}_2)$. Hence, we have $(2\boldsymbol{x}^*, \boldsymbol{y})$ is a feasible solution for the static problem in (3.4), therefore

$$z_{\mathsf{Sta}} \leq 2\boldsymbol{c}^T \boldsymbol{x}^* + \boldsymbol{d}^T \boldsymbol{y} = 2\mathrm{OPT}_1 + z(\mathcal{J}_2) \leq 2\mathrm{OPT}_1 + \beta \cdot \mathrm{OPT} = O(\beta) \cdot \mathrm{OPT}.$$

$\square$

Proof of Theorem 3.1. Lemma 3.1 show that our affine solution (3.5) is feasible for the adjustable problem (1.1). Lemma 3.2 and 3.4 show that the cost of the affine solution (3.5) is less than $\beta \cdot \mathrm{OPT} + O(\beta) \cdot \mathrm{OPT} = O(\beta) \cdot \mathrm{OPT}$ which implies that

$$z_{\mathsf{Aff}}(\mathcal{U}) = O\left(\frac{\log n}{\log \log n}\right) \cdot z_{\mathsf{AR}}(\mathcal{U}).$$

$\square$

3.3. Proof of the Structural Result. Suppose the two conditions of Lemma 3.3 are satisfied. Denote $\eta = \frac{4\log n}{\log \log n}$. The goal is to show that $z(\mathcal{J}) \leq \eta\gamma$. For all $i \in \mathcal{J}$, denote $\boldsymbol{v}_i$ the optimal solution corresponding to $z(\boldsymbol{e}_i)$. We have $\|\boldsymbol{v}_i\|_0 = 1$, i.e. $z(\boldsymbol{e}_i) = d_\ell v_{i\ell}$ where

$$\ell = \underset{\substack{j=1,\ldots,n \\ B_{ij} \neq 0}}{\arg\min} \frac{d_j}{B_{ij}}.$$

In particular, we have for all $j \in [n]$ such that $B_{ij} \neq 0$

$$\frac{d_j}{B_{ij}} \geq \frac{d_\ell}{B_{i\ell}} = d_\ell v_{i\ell} = z(\boldsymbol{e}_i) > \eta\gamma w_i,$$

i.e. for all $j \in [n]$,

$$d_j \geq \eta\gamma \cdot \max_{i \in \mathcal{J}}(w_i B_{ij}).$$

Denote for $j \in [n]$,

$$\hat{d}_j = \frac{d_j}{\eta\gamma \cdot \max_{i \in \mathcal{J}}(w_i B_{ij})},$$

and for all $i \in \mathcal{J}, j \in [n]$,

$$\hat{B}_{ij} = \frac{w_i B_{ij}}{\max_{i \in \mathcal{J}}(w_i B_{ij})}.$$

In particular, we have for all $i \in \mathcal{J}, j \in [n]$, $\hat{B}_{ij} \in [0,1]$ and for all $j \in [n]$, $\hat{d}_j > 1$. For any $\mathcal{W} \subseteq \mathcal{J}$, consider the following problem

$$\hat{z}(\mathcal{W}) = \min_{\boldsymbol{y} \geq \boldsymbol{0}}\left\{ \hat{\boldsymbol{d}}^T \boldsymbol{y} \,\middle|\, \hat{\boldsymbol{B}}\boldsymbol{y} \geq \sum_{i \in \mathcal{W}} w_i \boldsymbol{e}_i \right\}. \tag{3.6}$$

We show that for any $\mathcal{W} \subseteq \mathcal{J}$, $\hat{z}(\mathcal{W})$ is just a scaling of $z(\mathcal{W})$. In particular, we have

CLAIM 3.1. $z(\mathcal{W}) = \eta\gamma \cdot \hat{z}(\mathcal{W})$.

We present the proof of Claim 3.1 in Appendix A. To show that $z(\mathcal{J}) \leq \eta\gamma$. We show equivalently that $\hat{z}(\mathcal{J}) \leq 1$. For the sake of contradiction, suppose $\hat{z}(\mathcal{J}) > 1$. Our goal is to construct a scenario that contradicts condition 2 of the lemma. We use ideas on dual rounding and randomized solutions from [20] and [22]. In particular, let the dual problem of $\hat{z}(\mathcal{J})$ be

$$\hat{\Delta}_{\mathcal{J}} = \max_{\mathbf{z} \geq \mathbf{0}} \left\{ \sum_{i \in \mathcal{J}} w_i z_i \;\middle|\; \sum_{i \in \mathcal{J}} \hat{B}_{ij} z_i \leq \hat{d}_j \;\; \forall j \in [n] \right\}. \tag{3.7}$$

Denote $\mathbf{z}^*$ the optimal solution for the dual problem (3.7). By strong LP duality, we have

$$\hat{\Delta}_{\mathcal{J}} = \hat{z}(\mathcal{J}),$$

and therefore,

$$\hat{\Delta}_{\mathcal{J}} = \sum_{i \in \mathcal{J}} w_i z_i^* > 1.$$

We define the following randomized solution for all $i \in \mathcal{J}$

$$Z_i = \lfloor z_i^* \rfloor + \xi_i,$$

where ξ_i, for $i \in \mathcal{J}$ are independent Bernoulli variables with parameter $z_i^* - \lfloor z_i^* \rfloor$, i.e.

$$\xi_i = \mathsf{Ber}(z_i^* - \lfloor z_i^* \rfloor).$$

CLAIM 3.2.   *With probability at least* $1 - O(1/n)$,

$$\left(\frac{2Z_i}{\eta}, i \in \mathcal{J} \right),$$

is a feasible solution to the dual problem (3.7).

We show that $\left(\frac{2Z_i}{\eta}, i \in \mathcal{J} \right)$ satisfies the constraints of (3.7) with high probability by using Chernoff bound concentration inequalities. The proof of Claim 3.2 is presented in Appendix A. Furthermore, we show that the cost of our randomized solution $\left(\frac{2Z_i}{\eta}, i \in \mathcal{J} \right)$ is greater than $\frac{1}{\eta}$ with a constant probability. In particular, we have the following claim.

CLAIM 3.3. $\mathbb{P}\left(\sum_{i \in \mathcal{J}} w_i Z_i > \frac{1}{2} \right) \geq 1 - e^{-\frac{1}{8}}$.

We use a concentration bound to prove Claim 3.3. The proof is presented in Appendix A. Putting Claim 3.2 and Claim 3.3 together, we have that

$$\left(\frac{2Z_i}{\eta}, i \in \mathcal{J} \right),$$

is feasible with high probability and has a cost $\sum_{i \in \mathcal{J}} w_i \frac{2Z_i}{\eta}$ strictly greater than $\frac{1}{\eta}$ with a non-zero constant probability. Therefore, there exists a solution for problem (3.7) with a cost at least $\frac{1}{\eta}$. Let us assume that $\left(\frac{2Z_i}{\eta}, i \in \mathcal{J} \right)$ is such a solution. Let us order $w_i Z_i$ in an increasing order, i.e.,

$$w_{(1)} Z_{(1)} \geq w_{(2)} Z_{(2)} \geq \ldots \geq w_{(|\mathcal{J}|)} Z_{(|\mathcal{J}|)}.$$

We know that $\sum_{i \in \mathcal{J}} w_i Z_i > \frac{1}{2}$. Denote L the index such that

$$\sum_{i=1}^{L-1} w_{(i)} Z_{(i)} \leq \frac{1}{2} \qquad \text{and} \qquad \sum_{i=1}^{L} w_{(i)} Z_{(i)} > \frac{1}{2}.$$

Note that $Z_{(i)}$ are integral and $Z_{(L)} \neq 0$. Hence for all $i = 1, \ldots, L$, $Z_{(i)} \geq 1$. Therefore,

$$\sum_{i=1}^{L-1} w_{(i)} \leq \sum_{i=1}^{L-1} w_{(i)} Z_{(i)} < \frac{1}{2}.$$

Note that if $L = 1$, $\sum_{i=1}^{L} w_{(i)} = w_{(1)} \leq 1$ because all w_i are in $[0, 1]$. On the other hand, if $L \geq 2$, then

$$w_{(L)} \leq w_{(L)} Z_{(L)} \leq w_{(1)} Z_{(1)} \leq \sum_{i=1}^{L-1} w_{(i)} Z_{(i)} < \frac{1}{2},$$

and therefore,

$$\sum_{i=1}^{L} w_{(i)} = \sum_{i=1}^{L-1} w_{(i)} + w_{(L)} \leq \frac{1}{2} + \frac{1}{2} = 1.$$

Therefore, in both cases we have $\sum_{i=1}^{L} w_{(i)} \leq 1$. Denote $\mathcal{W} \subseteq \mathcal{J}$ the set of indices corresponding to the top L of $w_{(i)} Z_{(i)}$. In particular, we have

$$\sum_{i \in \mathcal{W}} w_i Z_i > \frac{1}{2},$$

and consequently

$$\sum_{i \in \mathcal{W}} \frac{2 Z_i}{\eta} > \frac{1}{\eta}.$$

Consider the following problem

$$\hat{\Delta}_{\mathcal{W}} = \max_{\mathbf{z} \geq \mathbf{0}} \left\{ \sum_{i \in \mathcal{W}} w_i z_i \ \middle| \ \sum_{i \in \mathcal{W}} \hat{B}_{ij} z_i \leq \hat{d}_j \ \forall j \in [n] \right\} \tag{3.8}$$

and its dual

$$\hat{z}(\mathcal{W}) = \min_{\mathbf{y} \geq \mathbf{0}} \left\{ \hat{\mathbf{d}}^T \mathbf{y} \ \middle| \ \hat{\mathbf{B}} \mathbf{y} \geq \sum_{i \in \mathcal{W}} w_i \mathbf{e}_i \right\}. \tag{3.9}$$

We have shown the existence of a solution $(\frac{2 Z_i}{\eta})_{i \in \mathcal{W}}$ to problem (3.8) with a cost strictly greater than $\frac{1}{\eta}$. Hence, by LP duality

$$\hat{z}(\mathcal{W}) > \frac{1}{\eta}.$$

Note that $z(\mathcal{W}) = \eta \gamma \cdot \hat{z}(\mathcal{W})$. Hence, $z(\mathcal{W}) > \gamma$ which contradicts condition 2 of our lemma.

4. Intersection of Disjoint budget constraints. In this section, we consider more general uncertainty sets that are defined by intersection of budget constraints that model many practical settings. We first consider the case where the budget constraints are *disjoint*. In particular, consider $S_1, S_2, \ldots, S_L$ a partition of $\{1, 2, \ldots, m\}$, i.e.

$$\bigcup_{\ell=1}^{L} S_\ell = [m] \qquad \text{and} \qquad S_i \cap S_j = \emptyset, \ \forall i \neq j.$$

We define *disjoint constrained budgeted sets* as follows

$$
\mathcal{U} = \left\{ \boldsymbol{h} \in [0,1]^m \ \middle| \ \sum_{i \in S_\ell} w_{\ell i} h_i \le 1 \ \forall \ell \in [L] \right\}. \tag{4.1}
$$

where $S_\ell, \ell = 1, \ldots, L$ is a partition of $[m]$. This is an important class of uncertainty sets that generalizes the budget of uncertainty set (1.2). These are essentially Cartesian product of L budget of uncertainty sets. A special case of this class of uncertainty sets where all $w_{\ell i}$ are equal has been considered for example in Gupta et al. [23] and Feige et al. [20]. Recall for $L = 1$, namely the budget of uncertainty set (1.2), affine policy gives the optimal approximation to (1.1) (see Theorem 3.1). Our result in this section show that the performance of affine policy remains near-optimal for the more general class (4.1). In particular, we have the following theorem.

THEOREM 4.1. *Consider the two-stage adjustable problem* (1.1) *where* $\mathcal{U}$ *is the disjoint constrained budgeted set* (4.1). *Then,*

$$
z_{\mathsf{Aff}}(\mathcal{U}) = O\left(\frac{\log^2 n}{\log \log n} \right) \cdot z_{\mathsf{AR}}(\mathcal{U}).
$$

Our analysis relies on constructing a feasible affine solution for (1.1) and relating the worst-case performance to a lower bound of (1.1). In particular, we consider the *online fractional covering problem* and use an online algorithm with $O(\log n)$-*competitive ratio* to both construct a feasible affine and also a lower bound for (1.1). The performance bound of our feasible affine is related the competitive ratio of the online algorithm. We first introduce some preliminaries before discussing our construction and analysis.

4.1. Online fractional covering. For any $\boldsymbol{B} \in \mathbb{R}_+^{m \times n}$ and $\boldsymbol{d} \in \mathbb{R}_+^n$, consider the (offline) fractional covering problem defined in (3.2):

$$
\min_{\boldsymbol{y} \ge \boldsymbol{0}} \left\{ \boldsymbol{d}^T \boldsymbol{y} \ \middle| \ \boldsymbol{B}\boldsymbol{y} \ge \boldsymbol{h} \right\},
$$

for any requirement $\boldsymbol{h} \in \{0,1\}^m$. The *online fractional covering problem* is an online version of the covering problem where the requirements are revealed online in a sequential manner. In particular, at each step we get a new constraint $\sum_{j=1}^n B_{ij} y_j \ge 1$ for some i and the algorithm needs to augment the current solution to satisfy the new requirement in each step.

This problem has been studied in the literature. We refer the reader to Buchbinder and Naor [14] for an extensive discussion of the problem. Buchbinder and Naor [13] give an online algorithm $\mathcal{A}$ for the online fractional covering problem that is $O(\log n)$-*competitive* (see Theorem 4.1 in [13]). In other words, the cost of the solution given by $\mathcal{A}$ for any set and sequence of requirements is at most $O(\log n)$ times the cost of the optimal solution of the corresponding offline covering problem where all the requirements are known upfront. In particular, for any sequence of requirements τ, we have

$$
\max_\tau \frac{\mathcal{A}(\tau)}{z(\tau)} = O(\log n),
$$

where $\mathcal{A}(\tau)$ is the online covering cost and $z(\tau)$ is the offline covering cost. Note that the competitive ratio guarantee also holds for the case where in each step, we get a subset of constraints instead of just a single constraint. We consider the algorithm $\mathcal{A}$ of Buchbinder and Naor [13] for our analysis. Let us introduce some notations that we will use for our construction and analysis.

Notations. Consider a sequence of subsets of constraints given by $(S_1, \ldots, S_R)$ where $S_r \subseteq [m]$ and $S_r \cap S_{r'} = \emptyset$ for all $r \neq r'$. In particular, in step r we get subset S_r of constraints

$$\sum_{j=1}^n B_{ij} y_j \geq 1 \qquad \forall i \in S_r.$$

For brevity of notations, let

$$\boldsymbol{h}_r = \mathbb{1}(S_r).$$

In particular, the sequence $(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_R)$ verifies $\boldsymbol{h}_r \in \{0, 1\}^m$ for all $r \in [R]$ and $\sum_{r=1}^R \boldsymbol{h}_r \leq \boldsymbol{e}$. We introduce the following definitions.

1. **Online cost.** We denote $\mathcal{A}(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r)$ the (online) cost of covering the sequence $(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r)$ using the online algorithm $\mathcal{A}$.

2. **Online augmenting cost.** We denote $\mathcal{A}(\boldsymbol{h}_{r+1} | \boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r)$ the extra cost to cover $\boldsymbol{h}_{r+1}$ using the online algorithm $\mathcal{A}$ when the algorithm have already covered the sequence $(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r)$. By definition, the online augmenting cost is given by

$$\mathcal{A}(\boldsymbol{h}_{r+1} | \boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r) = \mathcal{A}(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r, \boldsymbol{h}_{r+1}) - \mathcal{A}(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r). \tag{4.2}$$

3. **Greedy augmenting cost.** We denote $\mathsf{Aug}(\boldsymbol{h}_{r+1} | \boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r)$ the optimal cost to cover $\boldsymbol{h}_{r+1}$ given that the sequence $(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r)$ was already covered by the online algorithm $\mathcal{A}$. In particular, the greedy augmenting cost is given by

$$\mathsf{Aug}(\boldsymbol{h}_{r+1} | \boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r) = \min_{\boldsymbol{y} \geq \boldsymbol{0}} \left\{ \boldsymbol{d}^T \boldsymbol{y} \mid \boldsymbol{B} \left(\boldsymbol{y} + \boldsymbol{y}_r^{\mathcal{A}} \right) \geq \boldsymbol{h}_{r+1} \right\}, \tag{4.3}$$

where $\boldsymbol{y}_r^{\mathcal{A}}$ is the online solution corresponding to the cost $\mathcal{A}(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r)$.

4. **Offline cost.** Denote $z(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r)$ the optimal (offline) cost to cover $(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r)$ as defined in (3.2), i.e.,

$$z(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r) = z\left(\sum_{i=1}^r \boldsymbol{h}_i \right) = \min_{\boldsymbol{y} \geq \boldsymbol{0}} \left\{ \boldsymbol{d}^T \boldsymbol{y} \mid \boldsymbol{B} \boldsymbol{y} \geq \sum_{i=1}^r \boldsymbol{h}_i \right\}.$$

Since $\mathcal{A}$ is $O(\log n)$-competitive. Then for any sequence $(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r)$,

$$\mathcal{A}(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_r) = O(\log n) \cdot z\left(\sum_{i=1}^r \boldsymbol{h}_i \right). \tag{4.4}$$

4.2. Construction of our affine solution. Similar to the proof of Theorem 3.1, we construct a feasible affine solution where we split the components of $[m]$ into two subsets and cover one using a linear solution and the other one using a static solution. Consider $1, 2, \ldots, L$ the blocks of components of the the disjoint constrained budgeted set (4.1). For each block of components, we construct a threshold using the online fractional covering algorithm $\mathcal{A}$. This threshold defines the *expensive* components that we cover using a static solution and *inexpensive* components that we cover using a linear solution.

Construction of thresholds. Define the following set

$$\mathcal{T} = \left\{ i \in [m] \,\middle|\, (\boldsymbol{A} \boldsymbol{x}^*)_i \geq \frac{1}{2} \right\}$$

and its complement

$$\mathcal{T}^c = [m] \setminus \mathcal{T}.$$

For $\ell = 1, \dots, L$, denote

$$\hat{S}_\ell = S_\ell \cap \mathcal{T}^c.$$

Let us define the following sets for all $\ell \in [L]$,

$$\mathcal{U}_\ell = \left\{ \boldsymbol{h} \in \{0,1\}^m \,\bigg|\, \sum_{i \in \hat{S}_\ell} w_{\ell i} h_i \leq 1, \ h_i = 0 \ \forall i \notin \hat{S}_\ell \right\}.$$

Note that $\oplus_{\ell=1}^L \mathcal{U}_\ell \subseteq \mathcal{U}$. We construct a greedy sequence $(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_L)$ where each $\boldsymbol{a}_\ell$ is chosen from some set $\mathcal{U}_\ell$ such that it maximizes the online augmenting cost (4.2) of the sequence. Algorithm 1 describes the procedure in details.

Algorithm 1 Computing a greedy scenario $\boldsymbol{a}$

1: Initialize $\mathcal{L} = \{1, 2, \dots, L\}$.
2: **for** $\ell = 1, \dots, L$ **do**
3:     $(s, \boldsymbol{b}) = \underset{s \in \mathcal{L}, \, \boldsymbol{b} \in \mathcal{U}_s}{\arg\max} \ \mathcal{A}(\boldsymbol{b} \,|\, \boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_{\ell-1})$
4:     Set $\boldsymbol{a}_\ell = \boldsymbol{b}$, update $\mathcal{L} = \mathcal{L} \setminus s$
5: **end for**

We suppose wlog that the sequence $(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_L)$ belongs respectively to $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_L$ and let

$$\boldsymbol{a} = \sum_{\ell=1}^L \boldsymbol{a}_\ell \in \mathcal{U}.$$

For all $\ell \in [L]$, denote ν_ℓ the cost of covering the sequence $(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_\ell)$ using the online algorithm $\mathcal{A}$, i.e.

$$\nu_\ell = \mathcal{A}(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_\ell).$$

Note that from (4.4), we have

$$\nu_L \leq O(\log n) \cdot z(\boldsymbol{a}).$$

Moreover, since $2(\boldsymbol{A}\boldsymbol{x}^*)_i < 1$ for all $i \in \mathcal{T}^c$. Then $2\boldsymbol{B}\boldsymbol{y}^*(\boldsymbol{a}) \geq \boldsymbol{a}$. Hence,

$$z(\boldsymbol{a}) \leq 2\boldsymbol{d}^T y^*(\boldsymbol{a}) \leq 2\mathrm{OPT}_2 \leq 2\mathrm{OPT}$$

and consequently

$$\nu_L \leq O(\log n) \cdot \mathrm{OPT}. \tag{4.5}$$

Now, we are ready to construct our feasible affine solution that has a cost $O\left(\frac{\log^2 n}{\log \log n}\right)$ times $z_{\mathsf{AR}}(\mathcal{U})$ using a linear and a static part. Recall for all $i \in [m]$, $z(\boldsymbol{e}_i)$ the cost of covering component $\boldsymbol{e}_i$ in the second stage and $\boldsymbol{v}_i$ an optimal corresponding solution. For all $\ell = 1, \dots, L$, we consider the following sets of components

$$\mathcal{I}_\ell = \left\{ i \in S_\ell \,\bigg|\, \frac{z(\boldsymbol{e}_i)}{w_{\ell i}} \leq \beta \cdot (\nu_\ell - \nu_{\ell-1}) \right\},$$

where

$$\beta = \frac{8 \log n}{\log \log n},$$

and

$$\nu_\ell - \nu_{\ell-1} = \mathcal{A}(\boldsymbol{a}_\ell | \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{\ell-1}),$$

as defined in (4.2). Denote

$$\mathcal{I} = \bigcup_{\ell=1}^{L} \mathcal{I}_\ell.$$

and $\mathcal{I}^c$ its complement, i.e. $\mathcal{I}^c = [m] \setminus \mathcal{I}$.

Linear part. We cover the component of $\mathcal{I}$ using the following linear solution for any $\boldsymbol{h} \in \mathcal{U}$,

$$\boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h}) = \sum_{i \in \mathcal{I}} h_i \boldsymbol{v}_i. \tag{4.6}$$

Static part. We use a static solution to cover the remaining components $\boldsymbol{e}_i$ where $i \in \mathcal{I}^c$. In particular, similar to (3.4) we consider the following static problem

$$(\boldsymbol{x}_{\mathsf{Sta}}, \boldsymbol{y}_{\mathsf{Sta}}) \in \underset{\boldsymbol{x}, \boldsymbol{y} \geq \boldsymbol{0}}{\arg\min} \left\{ \boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{d}^T \boldsymbol{y} \,\middle|\, \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \geq \sum_{i \in \mathcal{I}^c} \boldsymbol{e}_i \right\} \tag{4.7}$$

and denote $z_{\mathsf{Sta}} = \boldsymbol{c}^T \boldsymbol{x}_{\mathsf{Sta}} + \boldsymbol{d}^T \boldsymbol{y}_{\mathsf{Sta}}$. Our affine solution is given by

$$\begin{aligned} \boldsymbol{x} &= \boldsymbol{x}_{\mathsf{Sta}} \\ \boldsymbol{y}(\boldsymbol{h}) &= \boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h}) + \boldsymbol{y}_{\mathsf{Sta}}, \qquad \forall \boldsymbol{h} \in \mathcal{U}. \end{aligned} \tag{4.8}$$

We can show that the affine solution (4.8) is feasible for (1.1). In particular, we have

LEMMA 4.1 **(Feasibility).**   *The affine solution in (4.8) is feasible for the adjustable problem (1.1).*

The proof is similar to the proof of Lemma 3.1.

4.3. Cost analysis.   In the following two lemmas, we analyze the cost of the linear and static parts in our affine solution (4.8).

LEMMA 4.2 **(Cost of Linear part).**   *The cost of the linear part $\boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h})$ defined in (4.8) is $O(\beta \log n) \cdot \mathrm{OPT}$ for any $\boldsymbol{h} \in \mathcal{U}$.*

Proof. We have,

$$\begin{aligned} \boldsymbol{d}^T \boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h}) = \sum_{\ell=1}^{L} \sum_{i \in \mathcal{I}_\ell} h_i \boldsymbol{d}^T \boldsymbol{v}_i = \sum_{\ell=1}^{L} \sum_{i \in \mathcal{I}_\ell} h_i z(\boldsymbol{e}_i) &\leq \sum_{\ell=1}^{L} \beta \cdot (\nu_\ell - \nu_{\ell-1}) \sum_{i \in \mathcal{I}_\ell} w_{\ell i} h_i \\ &\leq \beta \sum_{\ell=1}^{L} (\nu_\ell - \nu_{\ell-1}) \\ &= \beta \cdot \nu_L \\ &\leq O(\beta \log n) \cdot \mathrm{OPT}. \end{aligned}$$

where the first inequality holds because $z(\boldsymbol{e}_i) \leq \beta w_{\ell i} \cdot (\nu_\ell - \nu_{\ell-1})$ for all $i \in \mathcal{I}_\ell$ and $\ell \in [L]$, the second one because $\sum_{i \in \mathcal{I}_\ell} w_{\ell i} h_i \leq 1$ for any $\boldsymbol{h} \in \mathcal{U}$ and the third inequality follows from (4.5). $\square$

LEMMA 4.3 **(Cost of Static part).**   *The cost of the static part $(\boldsymbol{x}_{\mathsf{Sta}}, \boldsymbol{y}_{\mathsf{Sta}})$ defined in (4.8) is $O(\beta \log n) \cdot \mathrm{OPT}$.*

Proof. First, we have for all $i \in \mathcal{T}$, $2\boldsymbol{A}\boldsymbol{x}^* \geq \boldsymbol{e}_i$ and $2\boldsymbol{c}^T\boldsymbol{x}^* = 2 \cdot \mathrm{OPT}_1$. So, we cover the components of $\mathcal{T}$ using the static solution $2\boldsymbol{x}^*$ for a cost $2\mathrm{OPT}_1$. Denote $\boldsymbol{y}_{\mathcal{A}}$ the solution provided by the online algorithm $\mathcal{A}$ that covers the sequence $(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_L)$. Note that while covering the sequence $(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_L)$, the online algorithm $\mathcal{A}$ might also have some extra coverage. Let consider

$$\mathcal{J}_1 = \left\{ i \in [m] \,\middle|\, (\boldsymbol{B}\boldsymbol{y}_{\mathcal{A}})_i \geq \frac{1}{2} \right\}.$$

We have for all $i \in \mathcal{J}_1$, $2\boldsymbol{B}\boldsymbol{y}_{\mathcal{A}} \geq \boldsymbol{e}_i$ and

$$2\boldsymbol{d}^T\boldsymbol{y}_{\mathcal{A}} = 2\nu_L = O(\log n) \cdot \mathrm{OPT}.$$

So, we cover the components of $\mathcal{J}_1$ using the static solution $2\boldsymbol{y}_{\mathcal{A}}$ for a cost $O(\log n) \cdot \mathrm{OPT}$. Now, we focus on the set of the remaining components. Denote

$$\mathcal{J}_2 = \mathcal{I}^c \setminus \{\mathcal{T} \cup \mathcal{J}_1\}$$

and for $\ell = 1, \ldots, L$ denote

$$\mathcal{V}_\ell = \mathcal{J}_2 \cap S_\ell.$$

For each $\ell \in [L]$, we apply the structural result in Lemma 3.3 for the subset $\mathcal{V}_\ell$ with $\gamma = 2(\nu_\ell - \nu_{\ell-1})$ and $w_{\ell i}$. The first condition of Lemma 3.3 is satisfied because

$$\frac{z(\boldsymbol{e}_i)}{w_{\ell i}} > \beta(\nu_\ell - \nu_{\ell-1}) = 4\gamma \frac{\log n}{\log \log n}.$$

Consider any $\mathcal{W} \subseteq \mathcal{V}_\ell$ such that $\sum_{i \in \mathcal{W}} w_{\ell i} \leq 1$. We have $\mathbb{1}(\mathcal{W}) \in \mathcal{U}_\ell$. Moreover, $\boldsymbol{y}_{\mathcal{A}}$ covers less than $\frac{1}{2}\mathbb{1}(\mathcal{W})$. Therefore,

$$\frac{1}{2}z(\mathcal{W}) \leq \mathsf{Aug}\left(\mathbb{1}(\mathcal{W}) \,\middle|\, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_L\right).$$

Furthermore,

$$\begin{aligned}
\mathsf{Aug}\left(\mathbb{1}(\mathcal{W}) \,\middle|\, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_L\right) &\leq \mathsf{Aug}\left(\mathbb{1}(\mathcal{W}) \,\middle|\, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{\ell-1}\right) \\
&\leq \mathcal{A}\left(\mathbb{1}(\mathcal{W}) \,\middle|\, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{\ell-1}\right) \\
&\leq \mathcal{A}\left(\boldsymbol{a}_\ell \,\middle|\, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{\ell-1}\right) \\
&= \nu_\ell - \nu_{\ell-1},
\end{aligned}$$

where the first inequality holds because the cost of covering $\mathbb{1}(\mathcal{W})$ given the online solution for $(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{\ell-1})$ is smaller than the cost of covering $\mathbb{1}(\mathcal{W})$ given the online solution for $(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_L)$. The second inequality holds because the cost of the online algorithm is less than the offline cost and the third one follows from Step 3 in the construction of the greedy scenario $\boldsymbol{a}$ in Algorithm 1. Hence,

$$z(\mathcal{W}) \leq 2(\nu_\ell - \nu_{\ell-1}) = \gamma.$$

Therefore, the second condition of Lemma 3.3 is also satisfied and consequently from Lemma 3.3 we get,

$$z(\mathcal{V}_\ell) \leq 4\gamma \frac{\log n}{\log \log n} = \beta(\nu_\ell - \nu_{\ell-1}).$$

By taking the sum over all $\ell = 1, \ldots, L$, we get

$$\sum_{\ell=1}^{L} z(\mathcal{V}_\ell) \leq \beta \nu_L = O(\beta \log n) \cdot \mathrm{OPT}.$$

Putting all together, we have

$$
\begin{aligned}
z_{\mathsf{Sta}} &\leq 2\boldsymbol{c}^T\boldsymbol{x}^* + 2\boldsymbol{d}^T\boldsymbol{y}_{\mathcal{A}} + \sum_{\ell=1}^{L} z(\mathcal{V}_\ell) \\
&\leq 2\mathrm{OPT}_1 + O(\log n)\cdot\mathrm{OPT} + O(\beta\log n)\cdot\mathrm{OPT} \\
&= O(\beta\log n)\cdot\mathrm{OPT}.
\end{aligned}
$$

□

Proof of Theorem 4.1. Lemma 4.1 show that our affine solution (4.8) is feasible for the adjustable problem (1.1). Lemma 4.2 and 4.3 show that the cost of the feasible affine solution is less than

$$
O(\beta\log n)\cdot\mathrm{OPT} + O(\beta\log n)\cdot\mathrm{OPT} = O\left(\frac{\log^2 n}{\log\log n}\right)\cdot\mathrm{OPT}
$$

which implies that

$$
z_{\mathsf{Aff}}(\mathcal{U}) = O\left(\frac{\log^2 n}{\log\log n}\right)\cdot z_{\mathsf{AR}}(\mathcal{U}).
$$

□

We would like to note that Gupta et al. [23] give $O(\log n)$-approximation to (1.1) in the special case $\boldsymbol{A}, \boldsymbol{B} \in \{0,1\}^{m\times n}$, $\boldsymbol{d} = \lambda\boldsymbol{c}$ and $w_{\ell i} = w$ are all ℓ, i for some constant w. Therefore, for this special case the bound of [23] is stronger than our bound in Theorem 4.1. However, their algorithm does not give a functional policy approximation. Here, our focus is different, namely, to analyze the performance of affine policies that are widely used in practice and exhibit strong empirical performance. Our analysis shows that the performance of affine policies for disjoint constrained budgeted sets is near-optimal and nearly matches the hardness of the problem. Note that our bound in Theorem 4.1 is not necessarily tight. It is an interesting open question to study if affine policies also give an optimal approximation for this more general class of budgeted uncertainty sets.

5. General Intersection of budgeted uncertainty sets. In this section, we consider the general intersection of budget of uncertainty sets given by (1.3)

$$
\mathcal{U} = \left\{ \boldsymbol{h} \in [0,1]^m \;\Big|\; \sum_{i\in S_\ell} w_{\ell i} h_i \leq 1 \;\; \forall \ell \in [L] \right\},
$$

where $\boldsymbol{w}_\ell \in [0,1]^m$ and S_ℓ for $\ell \in [L]$ is a general family of subsets of $[m]$. This class is a generalization of the single budget of uncertainty set (1.2). It captures many important sets including CLT sets considered in Bertsimas and Bandi [1] and *inclusion-constrained budgeted* sets considered in Gounaris et al. [21]. In this section , we study the performance of affine policies for intersection of budget of uncertainty sets (1.3) and show strong theoretical guarantees. We start by the case when the set (1.3) verifies some symmetric properties (*permutation invariant sets*) and then we give our results for the general form (1.3).

5.1. Permutation Invariant Sets. We consider intersection of budgeted sets that are *permutation invariant*.

DEFINITION 5.1 (**Permutation Invariant Sets**). We say that $\mathcal{U}$ is a *permutation invariant set* if $\boldsymbol{x} \in \mathcal{U}$ implies that for any permutation τ of $\{1, 2, \ldots, m\}$, $\boldsymbol{x}^\tau \in \mathcal{U}$ where $x_i^\tau = x_{\tau(i)}$.

This class of sets captures many important sets including CLT sets that have been considred in Bertsimas and Bandi [1]. Note that a CLT set is given by

$$\mathcal{U} = \left\{ \boldsymbol{h} \in [0,1]^m \,\Big|\, \sum_{i \in \mathcal{S}} h_i \leq \Gamma \; \forall \mathcal{S} \subseteq [m], |\mathcal{S}| = k \right\}, \tag{5.1}$$

for some $k \in \mathbb{N}$. The following theorem gives our performance bound for affine policies under the class of intersection of budgeted sets that are permutation invariant.

THEOREM 5.1. *Consider the two-stage adjustable problem* (1.1) *where* $\mathcal{U}$ *is the intersection of* L *budget constraints* (1.3). *Suppose that* $\mathcal{U}$ *is permutation invariant set. Then,*

$$z_{\mathsf{Aff}}(\mathcal{U}) = O\left(\log L \cdot \frac{\log n}{\log \log n} \right) \cdot z_{\mathsf{AR}}(\mathcal{U}).$$

Proof. Our proof relies on a geometric property that we show for budgeted uncertainty sets that are permutation invariant. In particular, we show that for any $\mathcal{U}$ permutation invariant, there exists a (single) budget of uncertainty set $\mathcal{V}$ of the form (1.2) such that

$$\frac{1}{4 \log L} \cdot \mathcal{V} \subseteq \mathcal{U} \subseteq 2\mathcal{V}. \tag{5.2}$$

Since $\mathcal{U}$ is permutation invariant,

$$\gamma \boldsymbol{e} \in \arg\min \left\{ \boldsymbol{e}^T \boldsymbol{h} \,\Big|\, \boldsymbol{h} \in \mathcal{U} \right\},$$

for some $\gamma \in [0,1]$. Consider

$$\xi_i \overset{i.i.d.}{\sim} \mathsf{Ber}(\gamma) \qquad i = 1, \ldots, m$$

i.e. $\xi_1, \xi_2, \ldots, \xi_m$ are i.i.d. Bernoulli random variables of parameter γ. Let

$$\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_m).$$

Consider the following budget of uncertainty set

$$\tilde{\mathcal{V}} = \left\{ \boldsymbol{h} \in [0,1]^m \,\bigg|\, \sum_{i=1}^m h_i \leq \sum_{i=1}^m \xi_i \right\}.$$

Note that $\tilde{\mathcal{V}}$ is random depending on the realization of $\xi_1, \xi_2, \ldots, \xi_m$. We show that

$$\mathbb{P}\left(\frac{1}{4 \log L} \cdot \tilde{\mathcal{V}} \subseteq \mathcal{U} \subseteq 2\tilde{\mathcal{V}} \right) > \epsilon,$$

for some constant $\epsilon > 0$ which implies the existence of $\tilde{\mathcal{V}}$ such that (5.2) is verified. For that purpose, we show first that the right inclusion holds with a constant probability and then the left one holds with high probability.

CLAIM 5.1. $\mathbb{P}\left(\mathcal{U} \subseteq 2\tilde{\mathcal{V}} \right) \geq 1 - e^{-\frac{1}{8}}.$

Let us prove the above claim. Note that

$$\gamma m = \max_{\boldsymbol{h} \in \mathcal{U}} \boldsymbol{e}^T \boldsymbol{h}.$$

Suppose that

$$\gamma m \le 2 \boldsymbol{e}^T \boldsymbol{\xi}.$$

Then for all $\boldsymbol{h} \in \mathcal{U}$,

$$\boldsymbol{e}^T \boldsymbol{h} \le 2 \boldsymbol{e}^T \boldsymbol{\xi}$$

i.e., for all $\boldsymbol{h} \in \mathcal{U}$

$$\boldsymbol{h} \in 2 \tilde{\mathcal{V}}.$$

Hence, $\gamma m \le 2 \boldsymbol{e}^T \boldsymbol{\xi}$ implies that $\mathcal{U} \subseteq 2 \tilde{\mathcal{V}}$. Therefore,

$$\begin{aligned}
\mathbb{P}\left(\mathcal{U} \subseteq 2\tilde{\mathcal{V}}\right) &\ge \mathbb{P}\left(2\boldsymbol{e}^T\boldsymbol{\xi} \ge \gamma m\right) \\
&= \mathbb{P}\left(\sum_{i=1}^{m} \xi_i \ge \frac{1}{2}\gamma m\right).
\end{aligned}$$

We know that $\mathbb{E}\left(\sum_{i=1}^{m} \xi_i\right) = \gamma m$. Therefore, from the Chernoff inequality in Lemma B.2, we have

$$\mathbb{P}\left(\sum_{i=1}^{m} \xi_i \ge \frac{1}{2}\gamma m\right) \ge 1 - \exp\left(-\frac{\gamma m}{8}\right) \ge 1 - e^{-\frac{1}{8}}.$$

where the last inequality holds because $\gamma m \ge 1$ since $\boldsymbol{e}_i \in \mathcal{U}$ for all i and $\mathcal{U}$ is convex.

CLAIM 5.2. $\mathbb{P}\left(\tilde{\mathcal{V}} \subseteq 4\log L \cdot \mathcal{U}\right) \ge 1 - \frac{1}{L}$.

Note that $\boldsymbol{\xi}$ is an extreme point of $\tilde{\mathcal{V}}$ and that all pareto extreme points of $\tilde{\mathcal{V}}$ are just permutation of $\boldsymbol{\xi}$. Moreover, we know that $\mathcal{U}$ is permutation invariant set, hence if $\mathcal{U}$ contains $\boldsymbol{\xi}$ then $\mathcal{U}$ contains all pareto extreme points of $\mathcal{U}$ and consequently contains $\mathcal{U}$ by down-monotonicity. Therefore,

$$\begin{aligned}
\mathbb{P}\left(\tilde{\mathcal{V}} \subseteq 4\log L \cdot \mathcal{U}\right) &\ge \mathbb{P}\left(\boldsymbol{\xi} \in 4\log L \cdot \mathcal{U}\right) \\
&= \mathbb{P}\left(\boldsymbol{\nu}_\ell^T \boldsymbol{\xi} \le 4\log L, \ \forall \ell \in [L]\right) \\
&= 1 - \mathbb{P}\left(\exists \ell \in [L], \ \boldsymbol{\nu}_\ell^T \boldsymbol{\xi} > 4\log L\right) \\
&\ge 1 - \sum_{\ell=1}^{L} \mathbb{P}\left(\boldsymbol{\nu}_\ell^T \boldsymbol{\xi} > 4\log L\right),
\end{aligned}$$

where the last inequality follows from a union bound. We have $\mathbb{E}(\boldsymbol{\nu}_\ell^T \boldsymbol{\xi}) = \boldsymbol{\nu}_\ell^T \gamma \boldsymbol{e} \le 1 \le \log L$ because $\gamma \boldsymbol{e}$ is a feasible point in $\mathcal{U}$. Therefore, from Lemma B.1 with $\delta = 3$.

$$\mathbb{P}\left(\boldsymbol{\nu}_\ell^T \boldsymbol{\xi} > 4\log L\right) \le \left(\frac{e^3}{4^4}\right)^{\log L} \le \left(e^{-2}\right)^{\log L} = \frac{1}{L^2}.$$

We conclude that

$$\mathbb{P}\left(\tilde{\mathcal{V}} \subseteq 4\log L \cdot \mathcal{U}\right) \ge 1 - \sum_{\ell=1}^{L} \frac{1}{L^2} = 1 - \frac{1}{L}.$$

Hence, from Claim 5.1 and Claim 5.2, there exists a budget of uncertainty set $\tilde{\mathcal{V}}$ with a non zero probability that verifies the inclusion in (5.2). Therefore,

$$
\begin{aligned}
z_{\mathsf{Aff}}(\mathcal{U}) &\leq 2 \cdot z_{\mathsf{Aff}}(\tilde{\mathcal{V}}) \\
&= 2 \cdot O\left(\frac{\log n}{\log \log n}\right) \cdot z_{\mathsf{AR}}(\tilde{\mathcal{V}}) \\
&\leq 2 \cdot O\left(\frac{\log n}{\log \log n}\right) \cdot 4 \log L \cdot z_{\mathsf{AR}}(\mathcal{U}) = O\left(\frac{\log n}{\log \log n} \cdot \log L\right) \cdot z_{\mathsf{AR}}(\mathcal{U}),
\end{aligned}
$$

where the first inequality holds because $\mathcal{U} \subseteq 2\tilde{\mathcal{V}}$, the first equality follows from Theorem 3.1 because $\tilde{\mathcal{V}}$ is a budget of uncertainty set, and finally the last inequality holds because $\tilde{\mathcal{V}} \subseteq 4 \log L \cdot \mathcal{U}$. □

We would like to note that the result of Theorem 5.1 extends as well to the class of intersection of budgeted sets that are *scaled permutation invariant*. We say that $\mathcal{U}$ is a scaled permutation invariant set if there exists $\boldsymbol{\lambda} \in \mathbb{R}_+^m$ and $\mathcal{V}$ a permutation invariant set such that

$$
\mathcal{U} = \mathsf{diag}(\boldsymbol{\lambda}) \cdot \mathcal{V}.
$$

In fact, for a given scaled permutation invariant set $\mathcal{U}$, it is possible to scale the two-stage adjustable problem (1.1) and get a new problem where the uncertainty set is permutation invariant. Indeed, suppose $\mathcal{U} = \mathsf{diag}(\boldsymbol{\lambda}) \cdot \mathcal{V}$ where $\mathcal{V}$ is a permutation invariant set; by multiplying the constraint matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ by $\mathsf{diag}(\boldsymbol{\lambda})^{-1}$, we get a new problem where the uncertainty set now is permutation invariant. The performance of affine policy is not affected by this scaling and the bound given by Theorem 5.1 still hold.

5.2. General Intersection of Budgets. Consider the general intersection of budgeted sets (1.3) which is given by

$$
\mathcal{U} = \left\{ \boldsymbol{h} \in [0,1]^m \ \Big| \ \sum_{i \in S_\ell} w_{\ell i} h_i \leq 1 \ \forall \ell \in [L] \right\}.
$$

We show that affine policy gives a worst-case bound of $O\left(L \cdot \frac{\log n}{\log \log n}\right)$ where L is the number of constraints in $\mathcal{U}$. In particular, we have the following theorem.

THEOREM 5.2. *Consider the two-stage adjustable problem* (1.1) *where $\mathcal{U}$ is the intersection of L budgeted sets given by* (1.3). *Then,*

$$
z_{\mathsf{Aff}}(\mathcal{U}) = O\left(L \cdot \frac{\log n}{\log \log n}\right) \cdot z_{\mathsf{AR}}(\mathcal{U}).
$$

Proof. Denote for all $\ell \in [L]$,

$$
\boldsymbol{w}_\ell = \sum_{i \in S_\ell} w_{\ell i} \boldsymbol{e}_i
$$

and let

$$
\bar{\boldsymbol{w}} = \frac{1}{L} \sum_{\ell=1}^{L} \boldsymbol{w}_\ell.
$$

Consider the following budget of uncertainty set

$$
\mathcal{V} = \left\{ \boldsymbol{h} \in [0,1]^m \ \Big| \ \bar{\boldsymbol{w}}^T \boldsymbol{h} \leq 1 \right\}.
$$

We show that $\mathcal{U} \subseteq \mathcal{V} \subseteq L \cdot \mathcal{U}$. Suppose $\boldsymbol{h} \in \mathcal{U}$. Then, for any $\ell \in [L]$, we have $\boldsymbol{w}_\ell^T \boldsymbol{h} \leq 1$. Therefore, by summing up all these inequalities and dividing by L, we get $\bar{\boldsymbol{w}}^T \boldsymbol{h} \leq 1$, i.e. $\boldsymbol{h} \in \mathcal{V}$. Hence $\mathcal{U} \subseteq \mathcal{V}$.

Conversely, suppose $\boldsymbol{h} \in \mathcal{V}$. For any $\ell \in [L]$, we have $\boldsymbol{w}_\ell^T \boldsymbol{h} \le \sum_{\ell=1}^L \boldsymbol{w}_\ell^T \boldsymbol{h} \le L$, hence $\boldsymbol{h} \in L \cdot \mathcal{U}$ and consequently $\mathcal{V} \subseteq L \cdot \mathcal{U}$. Therefore,

$$
\begin{aligned}
z_{\mathsf{Aff}}(\mathcal{U}) &\le z_{\mathsf{Aff}}(\mathcal{V}) \\
&= O\left(\frac{\log n}{\log\log n}\right) \cdot z_{\mathsf{AR}}(\mathcal{V}) \\
&\le O\left(\frac{\log n}{\log\log n}\right) \cdot L \cdot z_{\mathsf{AR}}(\mathcal{U}),
\end{aligned}
$$

where the first inequality holds because $\mathcal{U} \subseteq \mathcal{V}$, the second one is a consequence of Theorem 3.1 since $\mathcal{V}$ is a budget of uncertainty set of the form (1.2), and finally the last inequality holds because $\mathcal{V} \subseteq L \cdot \mathcal{U}$.

$\square$

6. Faster algorithm for near-optimal affine solutions. In this section, we present an algorithm to compute an approximate affine policy for (1.1) under budget of uncertainty sets, that is significantly faster than solving the optimization program (2.4) that computes the optimal affine policy. Our algorithm is based on the analysis of the performance of affine policies that shows the existence of a good affine solution that satisfies certain nice structural properties. In particular, our construction of approximate affine solution in Section 3.1 partitions the components into *expensive* and *inexpensive* components based on a threshold. We cover the *inexpensive* components using a linear solution and the *expensive* components using a static solution. In particular, we show that there exists an affine solution with such a structure and cost at most $O(\log n / \log\log n)$ times the optimal optimal cost of (1.1) for some partition of components into *expensive* and *inexpensive*. Based on this structure, we give a faster algorithm to compute an approximate affine solution for budget of uncertainty sets.

6.1. Our Algorithm. Let $\mathcal{U}$ be the budget of uncertainty set (1.2) given by

$$
\mathcal{U} = \left\{ \boldsymbol{h} \in [0,1]^m \;\middle|\; \sum_{i=1}^m w_i h_i \le 1 \right\}.
$$

Recall from the proof of Theorem 3.1, we construct our candidate affine solution by partitioning the components of $[m]$ into two subsets $\mathcal{I}$ and its complement $\mathcal{I}^c$ where

$$
\mathcal{I} = \left\{ i \in [m] \;\middle|\; \frac{z(\boldsymbol{e}_i)}{w_i} \le \beta \cdot \mathrm{OPT} \right\}.
$$

Since $\mathcal{I}$ is a threshold set (determined by threshold on $z(\boldsymbol{e}_i)/w_i$), there are only $(m+1)$ possibilities for $\mathcal{I}$. For each $\mathcal{I}$, the affine solution covers components of $\mathcal{I}$ using a linear solution and components of $\mathcal{I}^c$ using a static solution. Based on this, we consider the following algorithm. For all $i \in [m]$ denote

$$
\theta_i = \frac{z(\boldsymbol{e}_i)}{w_i}.
$$

Assume wlog that

$$
\theta_1 \le \theta_2 \le \ldots \le \theta_m.
$$

We consider an affine solution corresponding to each choice of $\mathcal{I}$. In particular, for all $j = 0, \ldots, m$, consider $\mathcal{I} = \{1, \ldots, j\}$. Cover components of $\mathcal{I}$ using the linear solution (3.3) given by

$$
\boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h}) = \sum_{i \in \mathcal{I}} h_i \boldsymbol{v}_i
$$

and cover components of $\mathcal{I}^c$ using the static solution (3.4) given by

$$(\boldsymbol{x}_{\mathsf{Sta}}, \boldsymbol{y}_{\mathsf{Sta}}) \in \arg\min_{\boldsymbol{x}, \boldsymbol{y} \geq \boldsymbol{0}} \left\{ \boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{d}^T \boldsymbol{y} \;\middle|\; \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \geq \sum_{i \in \mathcal{I}^c} \boldsymbol{e}_i \right\}.$$

The cost of this affine solution is given by:

$$\boldsymbol{c}^T \boldsymbol{x}_{\mathsf{Sta}} + \boldsymbol{d}^T \boldsymbol{y}_{\mathsf{Sta}} + \max_{\boldsymbol{h} \in \mathcal{U}} \boldsymbol{d}^T \boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h})$$

Our algorithm considers all $(m+1)$ possibilities for $\mathcal{I}$ and the corresponding affine solutions and returned the one with the minimum worst-case cost. Algorithm 2 describes the procedure in detail.

Algorithm 2 Computing Approximate Affine Policy

1: **for** $i = 0, \ldots, m$ **do**
2:     $\mathcal{I} = \{1, \ldots, i\}, \mathcal{I}^c = [m] \setminus \mathcal{I}$
3:     $y_{\mathsf{Lin}}(\boldsymbol{h}) = \sum_{i \in \mathcal{I}} h_i \boldsymbol{v}_i$
4:     $(\boldsymbol{x}_{\mathsf{Sta}}, \boldsymbol{y}_{\mathsf{Sta}}) \in \arg\max \left\{ \boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{d}^T \boldsymbol{y} \;\middle|\; \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \geq \sum_{i \in \mathcal{I}^c} \boldsymbol{e}_i, \; \boldsymbol{x}, \boldsymbol{y} \geq \boldsymbol{0} \right\}$
5:     $z_i = \boldsymbol{c}^T \boldsymbol{x}_{\mathsf{Sta}} + \boldsymbol{d}^T \boldsymbol{y}_{\mathsf{Sta}} + \max_{\boldsymbol{h} \in \mathcal{U}} \boldsymbol{d}^T \boldsymbol{y}_{\mathsf{Lin}}(\boldsymbol{h})$
6: **end for**
7: **return** $\min_{i=0,\ldots,m} z_i$

At each each iteration our algorithm solves a static problem to compute $(\boldsymbol{x}_{\mathsf{Sta}}, \boldsymbol{y}_{\mathsf{Sta}})$ and a fractional knapsack problem to compute the cost of the linear solution $\boldsymbol{y}_{\mathsf{Lin}}$. This can be done much faster than solving the affine problem (2.1) as we observe in our numerical experiments. In fact, the affine problem requires solving the LP (2.4) whose size grows quadratically with the size of the problem. On the other hand, the size of our static problems is linear. Moreover, solving the $(m+1)$ static problems in Algorithm 2 can be done in parallel.

We would like to note that since our approximate affine solution is based on the construction of affine policy in our analysis, the worst-case approximation bound for the Algorithm 2 is also $O(\frac{\log n}{\log \log n})$.

6.2. Numerical experiments. We study the empirical performance of our algorithm in computing affine policies for budget of uncertainty sets both from the perspective of computation time and the quality of the solution.

Experimental setup. We use the same test instances as in Ben-Tal et al. [3]. In particular, we choose $n = m$, $\boldsymbol{c} = \boldsymbol{d} = \boldsymbol{e}$ and $\boldsymbol{A} = \boldsymbol{B}$ where $\boldsymbol{B}$ is randomly generated as $\boldsymbol{B} = \boldsymbol{I}_m + \boldsymbol{G}$, where $\boldsymbol{I}_m$ is the identity matrix and $\boldsymbol{G}$ is a random normalized Gaussian, i.e. $G_{ij} = |Y_{ij}|/\sqrt{m}$ where Y_{ij} are i.i.d. standard Gaussians. Let consider the following budget of uncertainty sets:

$$\mathcal{U}_1 = \left\{ \boldsymbol{h} \in [0,1]^m \;\middle|\; \sum_{i=1}^m h_i \leq k \right\} \tag{6.1}$$

$$\mathcal{U}_2 = \left\{ \boldsymbol{h} \in [0,1]^m \;\middle|\; \sum_{i=1}^m w_i h_i \leq 1 \right\}. \tag{6.2}$$

For our numerical experiments, we choose $k = c\sqrt{m}$ with c a random uniform constant between 1 and 2 for the first uncertainty set $\mathcal{U}_1$. For the second uncertainty set $\mathcal{U}_2$, we choose $\boldsymbol{w}$ a normalized Gaussian vector, i.e., $w_i = G_i/\|\boldsymbol{G}\|_2$ where G_i are i.i.d. standard Gaussians. We consider values of m from $m = 10$ to $m = 100$ in increments of 10 and consider 20 instances for each value of m.

We compute the optimal affine solution by solving the LP formulation (2.4). We compute our approximate affine solution returned by Algorithm 2. We denote $z_{\mathsf{Alg}}(\mathcal{U})$ and $z_{\mathsf{Aff}}(\mathcal{U})$ respectively the cost of our affine solution returned by Algorithm 2 and the cost of the optimal affine solution. For each m from $m = 10$ to $m = 100$, we report the average ratio $z_{\mathsf{Alg}}(\mathcal{U})/z_{\mathsf{Aff}}(\mathcal{U})$, the running time of Algorithm 2 in seconds ($T_{\mathsf{Alg}}(s)$) and the running time of the optimal affine policy in seconds ($T_{\mathsf{aff}}(s)$). We present the results of our computational experiments in Table 2. The numerical results are obtained using Gurobi 7.0.2 on a 16-core server with 2.93GHz processor and 56GB RAM.

Results. We observe from Table 2 that our algorithm is significantly faster than the optimal affine policy. In fact, Algorithm 2 scales very well and the average running time is less than a second even for large values of m. On the other hand, computing the optimal affine solution becomes computationally challenging for large values of m. For example, for $m = 100$, the average running time is more than 4 minutes for $\mathcal{U}_1$ and more than 11 minutes for $\mathcal{U}_2$. Furthermore, we observe that the gap between our affine solution and the optimal one is only between 8% and 15%. Moreover, this gap does not increase with the dimension of m, thereby confirming that our affine solution performs well even for large values of m.

m	$T_{\mathsf{aff}}(s)$	$T_{\mathsf{Alg}}(s)$	$z_{\mathsf{Alg}}/z_{\mathsf{Aff}}$
10	0.009	0.004	1.146
20	0.176	0.011	1.106
30	0.587	0.024	1.143
40	2.395	0.039	1.145
50	9.718	0.063	1.097
60	17.40	0.087	1.155
70	52.36	0.118	1.101
80	108.8	0.155	1.128
90	188.7	0.205	1.133
100	270.7	0.247	1.146

(a) Budget of uncertainty (6.1)

m	$T_{\mathsf{aff}}(s)$	$T_{\mathsf{Alg}}(s)$	$z_{\mathsf{Alg}}/z_{\mathsf{Aff}}$
10	0.011	0.009	1.093
20	0.171	0.015	1.093
30	0.987	0.032	1.096
40	4.006	0.049	1.095
50	13.01	0.075	1.085
60	34.58	0.109	1.094
70	85.95	0.137	1.092
80	183.01	0.178	1.090
90	396.8	0.221	1.083
100	719.5	0.281	1.080

(b) Budget of uncertainty (6.2)

TABLE 2. Comparison on the performance and computation time of the optimal affine policy and our algorithm. For 20 instances, we compute $z_{\mathsf{Alg}}(\mathcal{U})/z_{\mathsf{Aff}}(\mathcal{U})$ for $\mathcal{U}$ the budget of uncertainty sets (6.1) and (6.2). Here, $T_{\mathsf{Alg}}(s)$ denotes the running time for our piecewise affine policy and $T_{\mathsf{aff}}(s)$ denotes the running time for affine policy in seconds. These results are obtained using Gurobi 7.0.2 on a 16-core server with 2.93GHz processor and 56GB RAM.

7. Conclusion. In this paper, we consider a two-stage adjustable robust problem (1.1) under budget of uncertainty set and show that surprisingly affine policies give an optimal approximation for problem matching the hardness of approximation. We also present strong theoretical performance for more general intersection of budget of uncertainty sets that significantly improve over the worst case bound of $\Theta(\sqrt{m})$. Budgeted uncertainty sets are an important class of uncertainty sets that are widely used in practice. Therefore, our improved bounds for affine policies for this class of uncertainty sets also address the stark contrast between the near-optimal empirical performance and the worst case performance of $\Theta(\sqrt{m})$.

Furthermore, our analysis shows that there exists a near-optimal affine solution satisfying a nice structural property where the scenarios are partitioned into *inexpensive* and *expensive* based on

a threshold and the affine solution covers the inexpensive components using a linear solution and the expensive components using a static solution. This structure is closely related to threshold policies that are widely used in many applications, and allows us to design an alternate algorithm for computing near-optimal affine solutions for budget of uncertainty sets that is significantly faster than solving a large LP. This structural property might be of independent interest in itself and could also provide insights into designing improved policies for more general uncertainty sets.

References

[1] Bandi C, Bertsimas D (2012) Tractable stochastic analysis in high dimensions via robust optimization. *Mathematical programming* 134(1):23–70.

[2] Ben-Tal A, El Ghaoui L, Nemirovski A (2009) *Robust optimization* (Princeton University press).

[3] Ben-Tal A, El Housni O, Goyal V (2018) A tractable approach for designing piecewise affine policies in two-stage adjustable robust optimization. *Optimization Online* .

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

[5] Bertsimas D, Bidkhori H (2014) On the performance of affine policies for two-stage adaptive optimization: a geometric perspective. *Mathematical Programming* 1–18.

[6] Bertsimas D, Brown D, Caramanis C (2011) Theory and applications of robust optimization. *SIAM review* 53(3):464–501.

[7] Bertsimas D, de Ruiter FJ (2016) Duality in two-stage adaptive linear optimization: Faster computation and stronger bounds. *INFORMS Journal on Computing* 28(3):500–511.

[8] Bertsimas D, Georghiou A (2015) Design of near optimal decision rules in multistage adaptive mixed-integer optimization. *Operations Research* 63(3):610–627.

[9] Bertsimas D, Goyal V (2012) On the Power and Limitations of Affine Policies in Two-Stage Adaptive Optimization. *Mathematical Programming* 134(2):491–531.

[10] Bertsimas D, Goyal V, Sun X (2011) A geometric characterization of the power of finite adaptability in multistage stochastic and adaptive optimization. *Mathematics of Operations Research* 36(1):24–54.

[11] Bertsimas D, Iancu D, Parrilo P (2010) Optimality of Affine Policies in Multi-stage Robust Optimization. *Mathematics of Operations Research* 35:363–394.

[12] Bertsimas D, Sim M (2004) The Price of Robustness. *Operations Research* 52(2):35–53.

[13] Buchbinder N, Naor J (2009) Online primal-dual algorithms for covering and packing. *Mathematics of Operations Research* 34(2):270–286.

[14] Buchbinder N, Naor JS, et al. (2009) The design of competitive online algorithms via a primal–dual approach. *Foundations and Trends® in Theoretical Computer Science* 3(2–3):93–263.

[15] Chen X, Sim M, Sun P, Zhang J (2008) A linear decision-based approximation approach to stochastic programming. *Operations Research* 56(2):344–357.

[16] Chernoff H (1952) A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics* 493–507.

[17] Dhamdhere K, Goyal V, Ravi R, Singh M (2005) How to pay, come what may: Approximation algorithms for demand-robust covering problems. *Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on*, 367–376 (IEEE).

[18] Dhamdhere K, Goyal V, Ravi R, Singh M (2005) How to pay, come what may: Approximation algorithms for demand-robust covering problems. *Annual Symposium on Foundations of Computer Science*, volume 46, 367.

[19] El Housni O, Goyal V (2017) Beyond worst-case: A probabilistic analysis of affine policies in dynamic optimization. *Advances in Neural Information Processing Systems 30*, 4756–4764 (Curran Associates, Inc.).

[20] Feige U, Jain K, Mahdian M, Mirrokni V (2007) Robust combinatorial optimization with exponential scenarios. *Lecture Notes in Computer Science* 4513:439–453.

[21] Gounaris CE, Repoussis PP, Tarantilis CD, Wiesemann W, Floudas CA (2014) An adaptive memory programming framework for the robust capacitated vehicle routing problem. *Transportation Science* 50(4):1239–1260.

[22] Gupta A, Nagarajan V, Ravi R (2014) Thresholded covering algorithms for robust and max–min optimization. *Mathematical Programming* 146(1-2):583–615.

[23] Gupta A, Nagarajan V, Ravi R (2016) Robust and maxmin optimization under matroid and knapsack uncertainty sets. *ACM Transactions on Algorithms (TALG)* 12(1):10.

[24] Hanasusanto GA, Kuhn D, Wiesemann W (2015) K-adaptability in two-stage robust binary programming. *Operations Research* 63(4):877–891.

[25] Iancu D, Sharma M, Sviridenko M (2013) Supermodularity and affine policies in dynamic robust optimization. *Operations Research* 61(4):941–956.

[26] Postek K, Hertog Dd (2016) Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. *INFORMS Journal on Computing* 28(3):553–574.

[27] Raghavan P (1986) Probabilistic construction of deterministic algorithms: approximating packing integer programs. *Foundations of Computer Science, 1986., 27th Annual Symposium on*, 10–18 (IEEE).

[28] Xu G, Burer S (2018) A copositive approach for two-stage adjustable robust optimization with uncertain right-hand sides. *Computational Optimization and Applications* 70(1):33–59.

[29] Zhen J, Den Hertog D, Sim M (2016) Adjustable robust optimization via fourier-motzkin elimination. *Optimization Online* .

Appendix A: Proof of Claims in Section 3.3

Proof of Claim 3.1.

$$
\begin{aligned}
z(\mathcal{W}) &= \min_{\boldsymbol{y} \geq \boldsymbol{0}} \left\{ \boldsymbol{d}^T \boldsymbol{y} \ \middle|\ \boldsymbol{B}\boldsymbol{y} \geq \sum_{i \in \mathcal{W}} \boldsymbol{e}_i \right\} \\
&= \min_{\boldsymbol{y} \geq \boldsymbol{0}} \left\{ \boldsymbol{d}^T \boldsymbol{y} \ \middle|\ \sum_{j=1}^{n} B_{ij} y_j \geq 1, \ \forall i \in \mathcal{W} \right\} \\
&= \min_{\boldsymbol{y} \geq \boldsymbol{0}} \left\{ \boldsymbol{d}^T \boldsymbol{y} \ \middle|\ \sum_{j=1}^{n} \frac{w_i B_{ij}}{\max_{i \in \mathcal{W}}(w_i B_{ij})} \cdot \max_{i \in \mathcal{W}}(w_i B_{ij}) \cdot y_j \geq w_i, \ \forall i \in \mathcal{W} \right\} \\
&= \min_{\boldsymbol{y} \geq \boldsymbol{0}} \left\{ \boldsymbol{d}^T \boldsymbol{y} \ \middle|\ \sum_{j=1}^{n} \hat{B}_{ij} \cdot \max_{i \in \mathcal{W}}(w_i B_{ij}) \cdot y_j \geq w_i, \ \forall i \in \mathcal{W} \right\} \\
&= \min_{\boldsymbol{y} \geq \boldsymbol{0}} \left\{ \sum_{j=1}^{n} \frac{d_j}{\max_{i \in \mathcal{W}}(w_i B_{ij})} y_j \ \middle|\ \sum_{j=1}^{n} \hat{B}_{ij} y_j \geq w_i, \ \forall i \in \mathcal{W} \right\} \\
&= \min_{\boldsymbol{y} \geq \boldsymbol{0}} \left\{ \eta\gamma \sum_{j=1}^{n} \hat{d}_j y_j \ \middle|\ \sum_{j=1}^{n} \hat{B}_{ij} y_j \geq w_i, \ \forall i \in \mathcal{W} \right\} \\
&= \eta\gamma \cdot \hat{z}(\mathcal{W}).
\end{aligned}
$$

□

Proof of Claim 3.2. Consider $j \in [n]$, by the feasibility of the solution $\boldsymbol{z}^*$ we have,

$$
\sum_{i \in \mathcal{J}} \hat{B}_{ij} \cdot z_i^* \leq \hat{d}_j,
$$

and consequently

$$\sum_{i \in \mathcal{J}} \hat{B}_{ij} \frac{2\lfloor z_i^* \rfloor}{\eta} \leq \sum_{i \in \mathcal{J}} \hat{B}_{ij} \frac{2z_i^*}{\eta} \leq \frac{2\hat{d}_j}{\eta}.$$

Hence,

$$\mathbb{P}\left(\sum_{i \in \mathcal{J}} \hat{B}_{ij} \cdot \frac{2Z_i}{\eta} > \hat{d}_j\right) = \mathbb{P}\left(\sum_{i \in \mathcal{J}} \hat{B}_{ij} \cdot \frac{2(\lfloor z_i^* \rfloor + \xi_i)}{\eta} > \hat{d}_j\right)$$

$$\leq \mathbb{P}\left(\sum_{i \in \mathcal{J}} \hat{B}_{ij} \frac{2\xi_i}{\eta} > (1 - \frac{2}{\eta})\hat{d}_j\right)$$

$$= \mathbb{P}\left(\sum_{i \in \mathcal{J}} \hat{B}_{ij} \xi_i > (\frac{\eta}{2} - 1)\hat{d}_j\right).$$

Now, we apply the Chernoff inequality in Lemma B.1 with $\delta = \frac{\eta}{2} - 2$ and $\Xi = \sum_{i \in \mathcal{J}} \hat{B}_{ij} \xi_i$. Note that $\delta = 2 \cdot \frac{\log n}{\log \log n} - 2 > 0$ for sufficiently large n. Moreover, we have for all $i \in \mathcal{J}, j \in \mathcal{J}, \hat{B}_{ij} \in [0, 1]$ and

$$\mathbb{E}\left(\sum_{i \in \mathcal{J}} \hat{B}_{ij} \xi_i\right) = \sum_{i \in \mathcal{J}} \hat{B}_{ij}(z_i^* - \lfloor z_i^* \rfloor) \leq \sum_{i \in \mathcal{J}} \hat{B}_{ij} z_i^* \leq \hat{d}_j.$$

Therefore the Chernoff bound gives

$$\mathbb{P}\left(\sum_{i \in \mathcal{J}} \hat{B}_{ij} \xi_i > (\frac{\eta}{2} - 1)\hat{d}_j\right) \leq \left(\frac{e^{\frac{\eta}{2} - 1}}{(\frac{\eta}{2} - 2)^{\frac{\eta}{2} - 2}}\right)^{\hat{d}_j}.$$

Recall $\eta = 4 \frac{\log n}{\log \log n}$. Hence the RHS is equivalent to

$$\left(\frac{e^{\frac{\eta}{2} - 1}}{(\frac{\eta}{2} - 2)^{\frac{\eta}{2} - 2}}\right)^{\hat{d}_j} = O\left(\frac{e^{\frac{\eta \hat{d}_j}{2}}}{(\frac{\eta}{2})^{\frac{\hat{d}_j \eta}{2}}}\right) = O\left(\exp\left(\hat{d}_j \frac{\eta}{2}\left(1 - \log(\frac{\eta}{2})\right)\right)\right)$$

$$= O\left(\exp\left(-\hat{d}_j \frac{\eta}{2} \log \eta\right)\right)$$

$$= O\left(\exp\left(-2\hat{d}_j \frac{\log n}{\log \log n} \log\left(4\frac{\log n}{\log \log n}\right)\right)\right)$$

$$= O\left(\exp\left(-2\hat{d}_j \log n\right)\right) = O\left(\frac{1}{n^{2\hat{d}_j}}\right) \leq \frac{c}{n^2},$$

for some constant c. The last inequality holds because $\hat{d}_j \geq 1$. Hence,

$$\mathbb{P}\left(\sum_{i \in \mathcal{J}} \hat{B}_{ij} \cdot \frac{2Z_i}{\eta} > \hat{d}_j\right) \leq \frac{c}{n^2}.$$

Therefore by a union bound we have

$$\mathbb{P}\left(\sum_{i \in \mathcal{J}} \hat{B}_{ij} \frac{2Z_i}{\eta} > \hat{d}_j, \exists j \in [n]\right) \leq \sum_{i=1}^{n} \mathbb{P}\left(\sum_{i \in \mathcal{J}} \hat{B}_{ij} \frac{2Z_i}{\eta} > \hat{d}_j\right) \leq \frac{c}{n}.$$

Therefore,

$$\mathbb{P}\left(\sum_{i \in \mathcal{J}} \hat{B}_{ij} \frac{2Z_i}{\eta} \leq \hat{d}_j, \forall j \in [n]\right) \geq 1 - \frac{c}{n} = 1 - O(\frac{1}{n}).$$

$\square$

Proof of Claim 3.3. We have,

$$\sum_{i \in \mathcal{J}} w_i Z_i = \lambda + \sum_{i \in \mathcal{J}} w_i \xi_i,$$

where

$$\lambda = \sum_{i \in \mathcal{J}} w_i \lfloor z_i^* \rfloor.$$

We apply the Chernoff inequality in Lemma B.2 with $\delta = \frac{1}{2}$ and $\Xi = \lambda + \sum_{i \in \mathcal{J}} w_i \xi_i$. Note that

$$\mathbb{E}(\Xi) = \sum_{i \in \mathcal{J}} w_i z_i^* > 1.$$

Hence

$$
\begin{aligned}
\mathbb{P}\left(\sum_{i \in \mathcal{J}} w_i Z_i > \frac{1}{2}\right) &= \mathbb{P}\left(\Xi > \frac{1}{2}\right) \\
&\geq \mathbb{P}\left(\Xi > \frac{\mathbb{E}(\Xi)}{2}\right) \\
&\geq 1 - e^{-\frac{\mathbb{E}(\Xi)}{8}} \geq 1 - e^{-\frac{1}{8}}.
\end{aligned}
$$

$\square$

Appendix B: Chernoff bounds

LEMMA B.1 **(Chernoff Bound 1)**. *Let $\xi_1, \xi_2, \ldots, \xi_r$ be independent Bernoulli trials. Denote $\Xi = \sum_{i=1}^{r} \alpha_i \xi_i$ where $\alpha_1, \ldots, \alpha_r$ are reals in $[0,1]$. Let $s > 0$ such that $\mathbb{E}(\xi) \leq s$. Then for any $\delta > 0$,*

$$\mathbb{P}(\Xi \geq (1+\delta)s) \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^s.$$

The Chernoff bound in Lemma B.1 is a slight variant of the Raghavan-Spencer inequality (Theorem 1 in [27]). The proof is along the same lines as in [27]. For completeness, we are providing it below.

Proof. From Markov's inequality we have for all $t > 0$,

$$\mathbb{P}(\Xi \geq (1+\delta)s) = \mathbb{P}(e^{t\Xi} \geq e^{t(1+\delta)s}) \leq \frac{\mathbb{E}(e^{t\Xi})}{e^{t(1+\delta)s}}.$$

Denote p_i the parameter of the Bernoulli ξ_i. By independence, we have

$$\mathbb{E}(e^{t\Xi}) = \prod_{i=1}^{r} \mathbb{E}(e^{t\alpha_i \xi_i}) = \prod_{i=1}^{r} (p_i e^{t\alpha_i} + 1 - p_i) \leq \prod_{i=1}^{r} \exp\left(p_i(e^{t\alpha_i} - 1)\right)$$

where the inequality holds because $1 + x \leq e^x$ for all $x \in \mathbb{R}$. By taking $t = \ln(1+\delta) > 0$, the right hand side becomes

$$
\begin{aligned}
\prod_{i=1}^{r} \exp\left(p_i((1+\delta)^{\alpha_i} - 1)\right) &\leq \prod_{i=1}^{r} \exp\left(p_i \delta \alpha_i\right) \\
&= \exp\left(\delta \cdot \mathbb{E}(\Xi)\right) \leq e^{\delta s},
\end{aligned}
$$

where the first inequality holds because $(1+x)^\alpha \leq 1 + \alpha x$ for any $x \geq 0$ and $\alpha \in [0,1]$ and the second one because $s \geq \mathbb{E}(\Xi) = \sum_{i=1}^{r} \alpha_i p_i$. Hence, we have

$$\mathbb{E}(e^{t\Xi}) \leq e^{\delta s}.$$

On the other hand,

$$e^{t(1+\delta)s} = (1+\delta)^{(1+\delta)s}.$$

Therefore,

$$\mathbb{P}(\Xi \ge (1+\delta)s) \le \left(\frac{e^{\delta s}}{(1+\delta)^{(1+\delta)s}}\right) = \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{s}.$$

$\square$

LEMMA B.2 **(Chernoff Bound 2)**. *Let $\xi_1, \xi_2, \ldots, \xi_r$ be independent Bernoulli trials. Denote*

$$\Xi = \lambda + \sum_{i=1}^{r} \alpha_i \xi_i$$

where $\lambda \in \mathbb{R}_+$ and $\alpha_1, \ldots, \alpha_r$ are reals in $(0,1]$. Denote $\mu = \mathbb{E}(\Xi)$. Then for any $0 < \delta < 1$,

$$\mathbb{P}(\Xi > (1-\delta)\mu) > 1 - e^{\frac{-\delta^2 \mu}{2}}.$$

This is a slight variant of the lower tail Chernoff bound [16]. The proof is along the same lines as in [16]. For completeness, we are providing it below.

Proof. We show equivalently that

$$\mathbb{P}(\Xi \le (1-\delta)\mu) \le e^{\frac{-\delta^2 \mu}{2}}.$$

From Markov's inequality we have for all $t < 0$,

$$\mathbb{P}(\Xi \le (1-\delta)\mu) = \mathbb{P}(e^{t\Xi} \ge e^{t(1-\delta)\mu}) \le \frac{\mathbb{E}(e^{t\Xi})}{e^{t(1-\delta)\mu}}.$$

Denote p_i the parameter of the Bernoulli ξ_i. By independence, we have

$$\mathbb{E}(e^{t\Xi}) = e^{t\lambda} \prod_{i=1}^{r} \mathbb{E}(e^{t\alpha_i \xi_i}) = e^{t\lambda} \prod_{i=1}^{r} (p_i e^{t\alpha_i} + 1 - p_i) \le e^{t\lambda} \prod_{i=1}^{r} \exp\left(p_i(e^{t\alpha_i} - 1)\right),$$

where the inequality holds because $1 + x \le e^x$ for all $x \in \mathbb{R}$. We take $t = \ln(1-\delta) < 0$. We have $t \le -\delta$, hence

$$e^{t\lambda} \le e^{-\delta\lambda}.$$

Moreover,

$$\prod_{i=1}^{r} \exp\left(p_i(e^{t\alpha_i} - 1)\right) = \prod_{i=1}^{r} \exp\left(p_i((1-\delta)^{\alpha_i} - 1)\right) \le \prod_{i=1}^{r} \exp\left(-p_i \delta \alpha_i\right),$$

where the inequality holds because $(1-x)^\alpha \le 1 - \alpha x$ for any $0 < x < 1$ and $\alpha \in [0,1]$. Therefore,

$$\mathbb{E}(e^{t\Xi}) \le e^{-\delta\lambda} \prod_{i=1}^{r} \exp\left(-p_i \delta \alpha_i\right) = e^{-\delta\mu}.$$

On the other hand,

$$e^{t(1-\delta)\mu} = (1-\delta)^{(1-\delta)\mu}.$$

Therefore,

$$\mathbb{P}(\Xi \le (1-\delta)\mu) \le \left(\frac{e^{-\delta\mu}}{(1-\delta)^{(1-\delta)\mu}}\right) = \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}.$$

Finally, we have for any $0 < \delta < 1$,

$$\ln(1-\delta) \geq -\delta + \frac{\delta^2}{2}$$

which implies

$$(1-\delta) \cdot \ln(1-\delta) \geq -\delta + \frac{\delta^2}{2}$$

and consequently

$$\left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}} \right)^{\mu} \leq e^{\frac{-\delta^2 \mu}{2}}.$$

$\square$

Appendix C: Proof of Lemma 2.1

The affine problem (2.1) has the following epigraph formulation

$$
\begin{aligned}
z_{\mathsf{Aff}}(\mathcal{U}) = \min \ & \boldsymbol{c}^T \boldsymbol{x} + z \\
& z \geq \boldsymbol{d}^T \left(\boldsymbol{P}\boldsymbol{h} + \boldsymbol{q} \right), \quad \forall \boldsymbol{h} \in \mathcal{U} \\
& \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B} \left(\boldsymbol{P}\boldsymbol{h} + \boldsymbol{q} \right) \geq \boldsymbol{h}, \quad \forall \boldsymbol{h} \in \mathcal{U} \\
& \boldsymbol{P}\boldsymbol{h} + \boldsymbol{q} \geq \boldsymbol{0}, \quad \forall \boldsymbol{h} \in \mathcal{U} \\
& \boldsymbol{x} \in \mathbb{R}_+^n, \ \boldsymbol{P} \in \mathbb{R}^{n \times m}, \ \boldsymbol{q} \in \mathbb{R}^n, \ z \in \mathbb{R}.
\end{aligned}
$$

We use standard duality techniques to derive formulation (2.4). The first constraint is equivalent to

$$z - \boldsymbol{d}^T \boldsymbol{q} \geq \max_{\substack{\boldsymbol{R}\boldsymbol{h} \leq \boldsymbol{r} \\ \boldsymbol{h} \geq \boldsymbol{0}}} \boldsymbol{d}^T \boldsymbol{P}\boldsymbol{h}$$

By taking the dual of the maximization problem, the constraint is equivalent

$$z - \boldsymbol{d}^T \boldsymbol{q} \geq \min_{\substack{\boldsymbol{R}^T \boldsymbol{v} \geq \boldsymbol{P}^T \boldsymbol{d} \\ \boldsymbol{v} \geq \boldsymbol{0}}} \boldsymbol{r}^T \boldsymbol{v}$$

We can then drop the min and introduce $\boldsymbol{v}$ as a variable, hence we obtain the following linear constraints

$$z - \boldsymbol{d}^T \boldsymbol{q} \geq \boldsymbol{r}^T \boldsymbol{v}$$

$$\boldsymbol{R}^T \boldsymbol{v} \geq \boldsymbol{P}^T \boldsymbol{d}$$

$$\boldsymbol{v} \in \mathbb{R}_+^L$$

We use the same technique for the second sets of constraints, i.e.

$$\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{q} \geq \max_{\substack{\boldsymbol{R}\boldsymbol{h} \leq \boldsymbol{r} \\ \boldsymbol{h} \geq \boldsymbol{0}}} \boldsymbol{h}(\boldsymbol{I}_m - \boldsymbol{B}\boldsymbol{P})$$

By taking the dual of the maximization problem for each row and dropping the min we get the following compact formulation of these constraints

$$\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{q} \geq \boldsymbol{V}^T \boldsymbol{r}$$

$$\boldsymbol{R}^T \boldsymbol{V} \geq \boldsymbol{I}_m - \boldsymbol{B}\boldsymbol{P}$$

$$\boldsymbol{V} \in \mathbb{R}_+^{L \times m}$$

Similarly, the last constraint

$$q \geq \max_{\substack{Rh \leq r \\ h \geq 0}} -Ph$$

is equivalent to

$$q \geq U^T r$$
$$U^T R + P \geq 0$$
$$U \in \mathbb{R}_+^{L \times n}.$$

Putting all together, we get the formulation (2.4).