

## A Proofs of preliminaries

### Proof of Lemma 2.2

**Proof.** We have

$$\begin{aligned}
z_{\text{AR}}(\mathbf{B}) &= \min_{\mathbf{x} \geq \mathbf{0}} \mathbf{c}^T \mathbf{x} + \max_{\mathbf{h} \in \mathcal{U}} \min_{\substack{\mathbf{B}\mathbf{y} \geq \mathbf{h} - \mathbf{A}\mathbf{x} \\ \mathbf{y} \geq \mathbf{0}}} \mathbf{d}^T \mathbf{y} \\
&= \min_{\mathbf{x} \geq \mathbf{0}} \mathbf{c}^T \mathbf{x} + \max_{\mathbf{h} \in \mathcal{U}} \max_{\substack{\mathbf{B}^T \mathbf{w} \leq \mathbf{d} \\ \mathbf{w} \geq \mathbf{0}}} (\mathbf{h} - \mathbf{A}\mathbf{x})^T \mathbf{w} \\
&= \min_{\mathbf{x} \geq \mathbf{0}} \mathbf{c}^T \mathbf{x} + \max_{\mathbf{w} \in \mathcal{W}} -(\mathbf{A}\mathbf{x})^T \mathbf{w} + \max_{\substack{\mathbf{R}\mathbf{h} \leq \mathbf{r} \\ \mathbf{h} \geq \mathbf{0}}} \mathbf{h}^T \mathbf{w} \\
&= \min_{\mathbf{x} \geq \mathbf{0}} \mathbf{c}^T \mathbf{x} + \max_{\mathbf{w} \in \mathcal{W}} -(\mathbf{A}\mathbf{x})^T \mathbf{w} + \min_{\substack{\mathbf{R}^T \boldsymbol{\lambda} \geq \mathbf{w} \\ \boldsymbol{\lambda} \geq \mathbf{0}}} \mathbf{r}^T \boldsymbol{\lambda} \\
&= z_{\text{d-AR}}(\mathbf{B}).
\end{aligned} \tag{11}$$

where the second equality holds by taking the dual of the inner minimization problem, the third equality follows from switching the two max, and the fourth one by taking the dual of the second maximization problem.  $\square$

### Proof of Lemma 2.4

**Proof.** We restate the same proof in [8] in our setting. First, since the adjustable problem is a relaxation of the affine problem then  $z_{\text{d-Aff}}(\mathbf{B}) \leq z_{\text{d-AR}}(\mathbf{B})$ .

Now let's prove the other inequality. Consider  $\mathcal{W} = \{\mathbf{w} \in \mathbb{R}_+^m \mid \mathbf{B}^T \mathbf{w} \leq \mathbf{d}\}$  which is a simplex. Note that  $\mathbf{0}$  is always an extreme point of the simplex  $\mathcal{W}$  and denote  $\mathbf{w}^1, \mathbf{w}^2, \dots, \mathbf{w}^m$  the remaining  $m$  points. In particular, we have for any  $\mathbf{w} \in \mathcal{W}$

$$\mathbf{w} = \sum_{j=1}^m \alpha_j \mathbf{w}^j = \mathbf{Q}\boldsymbol{\alpha}$$

where  $\sum_{j=1}^m \alpha_j \leq 1$  and  $\mathbf{Q} = [\mathbf{w}^1 | \mathbf{w}^2 | \dots | \mathbf{w}^m]$ . Note that  $\mathbf{Q}$  is invertible since  $\mathbf{w}^1, \mathbf{w}^2, \dots, \mathbf{w}^m$  are linearly independent. Hence,  $\boldsymbol{\alpha} = \mathbf{Q}^{-1} \mathbf{w}$ . Denote  $\mathbf{x}^*, \boldsymbol{\lambda}^*(\mathbf{w})$  for  $\mathbf{w} \in \mathcal{W}$  the optimal adjustable solution of the adjustable problem (4). We define the following affine solution  $\mathbf{x} = \mathbf{x}^*$  and for  $\mathbf{w} \in \mathcal{W}$ ,  $\boldsymbol{\lambda}(\mathbf{w}) = \mathbf{P}\mathbf{Q}^{-1} \mathbf{w}$  where  $\mathbf{P} = [\boldsymbol{\lambda}^*(\mathbf{w}^1) | \boldsymbol{\lambda}^*(\mathbf{w}^2) | \dots | \boldsymbol{\lambda}^*(\mathbf{w}^m)]$ . In particular, we have

$$\boldsymbol{\lambda}(\mathbf{w}) = \sum_{j=1}^m \alpha_j \boldsymbol{\lambda}^*(\mathbf{w}^j).$$

Let us first check the feasibility of the solution

$$\mathbf{R}^T \boldsymbol{\lambda}(\mathbf{w}) = \sum_{j=1}^m \alpha_j \mathbf{R}^T \boldsymbol{\lambda}^*(\mathbf{w}^j) \geq \sum_{j=1}^m \alpha_j \mathbf{w}^j = \mathbf{w}$$

where the inequality follows from the feasibility of the adjustable solution. Therefore,

$$\begin{aligned}
z_{\text{d-Aff}}(\mathbf{B}) &\leq \mathbf{c}^T \mathbf{x} + \max_{\mathbf{w} \in \mathcal{W}} (-\mathbf{A}\mathbf{x})^T \mathbf{w} + \mathbf{r}^T \boldsymbol{\lambda}(\mathbf{w}) \\
&= \mathbf{c}^T \mathbf{x}^* + \max_{\boldsymbol{\alpha}} (-\mathbf{A}\mathbf{x}^*)^T \mathbf{w} + \sum_{j=1}^m \alpha_j \mathbf{r}^T \boldsymbol{\lambda}^*(\mathbf{w}^j) \\
&= \mathbf{c}^T \mathbf{x}^* + \max_{\boldsymbol{\alpha}} \sum_{j=1}^m \alpha_j \left( (-\mathbf{A}\mathbf{x}^*)^T \mathbf{w}^j + \mathbf{r}^T \boldsymbol{\lambda}^*(\mathbf{w}^j) \right) \\
&\leq \mathbf{c}^T \mathbf{x}^* + \max_{\mathbf{w} \in \mathcal{W}} \left( (-\mathbf{A}\mathbf{x}^*)^T \mathbf{w} + \mathbf{r}^T \boldsymbol{\lambda}^*(\mathbf{w}) \right) \max_{\boldsymbol{\alpha}} \sum_{j=1}^m \alpha_j = z_{\text{d-AR}}(\mathbf{B})
\end{aligned}$$

where the last inequality holds because  $\sum_{j=1}^m \alpha_j \leq 1$ . We conclude that  $z_{\text{d-Aff}}(\mathbf{B}) = z_{\text{d-AR}}(\mathbf{B})$ .  $\square$

## Proof of Lemma 2.5

**Proof.** First the inequality  $z_{d-\text{AR}}(\mathbf{B}) \leq z_{d-\text{Aff}}(\mathbf{B})$  is straightforward since the adjustable problem(1) is a relaxation of the affine problem (2). On the other hand, since  $\mathcal{W} \subseteq \kappa \cdot \mathcal{S}$  then,

$$z_{d-\text{Aff}}(\mathbf{B}) \leq \kappa \cdot z_{d-\text{Aff}}(\mathbf{B}, \mathcal{S})$$

where we denote  $z_{d-\text{Aff}}(\mathbf{B}, \mathcal{S})$  the dualized affine problem over  $\mathcal{S}$  (it's the same problem as  $z_{d-\text{Aff}}(\mathbf{B})$  where we only replace  $\mathcal{W}$  by  $\mathcal{S}$ ). Since  $\mathcal{S}$  is a simplex, from Lemma 2.4, we have  $z_{d-\text{Aff}}(\mathbf{B}, \mathcal{S}) = z_{d-\text{AR}}(\mathbf{B}, \mathcal{S})$ . Moreover,  $z_{d-\text{AR}}(\mathbf{B}, \mathcal{S}) \leq z_{d-\text{AR}}(\mathbf{B})$  because  $\mathcal{S} \subseteq \mathcal{W}$ . We conclude that

$$z_{d-\text{AR}}(\mathbf{B}) \leq z_{d-\text{Aff}}(\mathbf{B}) \leq \kappa \cdot z_{d-\text{AR}}(\mathbf{B}).$$

Furthermore, since  $z_{d-\text{AR}}(\mathbf{B}) = z_{\text{AR}}(\mathbf{B})$  from Lemma 2.2 and  $z_{d-\text{Aff}}(\mathbf{B}) = z_{\text{Aff}}(\mathbf{B})$  from Lemma 2.3, then

$$z_{\text{AR}}(\mathbf{B}) \leq z_{\text{Aff}}(\mathbf{B}) \leq \kappa \cdot z_{\text{AR}}(\mathbf{B}).$$

□

## B Hoeffding's inequality

**Hoeffding's inequality**[18]. Let  $Z_1, \dots, Z_n$  be independent bounded random variables with  $Z_i \in [a, b]$  for all  $i \in [n]$  and denote  $Z = \frac{1}{n} \sum_{i=1}^n Z_i$ . Therefore,

$$\mathbb{P}(Z - \mathbb{E}(Z) \leq -\tau) \leq \exp\left(\frac{-2n\tau^2}{(b-a)^2}\right).$$

## C Proof of Theorem 2.6

**Proof.** Denote  $\tilde{\mathcal{W}} = \{\mathbf{w} \in \mathbb{R}_+^m \mid \tilde{\mathbf{B}}^T \mathbf{w} \leq \bar{d} \cdot \mathbf{e}\}$  and  $\mathcal{S} = \{\mathbf{w} \in \mathbb{R}_+^m \mid \sum_{i=1}^m w_i \leq \bar{d}\}$ . Our goal is to sandwich  $\tilde{\mathcal{W}}$  between two simplices and use Lemma 2.5. Using the following tail inequality for Gaussian random variables  $\tilde{G} \sim \mathcal{N}(\mu, \sigma^2)$ ,  $\mathbb{P}(|\tilde{G} - \mu| \geq t) \leq 2e^{-\frac{t^2}{2\sigma^2}}$ , we have

$$\begin{aligned} \mathbb{P}(\tilde{B}_{ij} \leq \sqrt{6 \log(mn)}) &= 1 - \mathbb{P}\left(|\tilde{G}_{ij}| \geq \sqrt{6 \log(mn)}\right) \\ &\geq 1 - 2 \exp\left(\frac{-6 \log(mn)}{2}\right) = 1 - \frac{2}{(mn)^3} \geq 1 - \frac{1}{(mn)^2} \end{aligned}$$

Therefore by taking a union bound,

$$\mathbb{P}\left(\tilde{B}_{ij} \leq \sqrt{6 \log(mn)} \quad \forall i \in [n], \forall j \in [m]\right) \geq \left(1 - \frac{1}{(mn)^2}\right)^{mn} \geq 1 - \frac{1}{mn}$$

where the last inequality follows from Bernoulli's inequality. Therefore for any  $w \in \mathcal{S}$ , we have with probability at least  $1 - \frac{1}{mn}$ ,

$$\sum_{j=1}^m \tilde{B}_{ji} w_j \leq \sqrt{6 \log(mn)} \sum_{j=1}^m w_j \leq \sqrt{6 \log(mn)} \cdot \bar{d} \quad \forall i \in [n]$$

Hence, with probability at least  $1 - \frac{1}{mn}$  we have,  $\mathcal{S} \subseteq \sqrt{6 \log(mn)} \cdot \tilde{\mathcal{W}}$ .

Now, we want to find a simplex that includes  $\tilde{\mathcal{W}}$ . We follow a similar approach to the proof of Theorem 2.1. Consider any  $\mathbf{w} \in \tilde{\mathcal{W}}$ . We have similarly to equation (7)

$$\sum_{j=1}^m \left(\frac{\sum_{i=1}^n \tilde{B}_{ji}}{n}\right) \cdot w_j \leq \bar{d}. \quad (12)$$

We have the following concentration inequality for non-negative random variables (see Theorem 7 in [12]),

$$\mathbb{P}\left(\frac{\sum_{i=1}^n \tilde{B}_{ji}}{n} \geq \mu - \tau\right) \geq 1 - \exp\left(\frac{-n\tau^2}{2\mathbb{E}(\tilde{B}_{11}^2)}\right) = 1 - \exp\left(\frac{-n\tau^2}{2}\right) = 1 - \frac{1}{m^2}$$

where  $\tau = 2\sqrt{\frac{\log m}{n}}$  and  $\mu = \mathbb{E}[\tilde{B}_{ji}] = \sqrt{\frac{2}{\pi}}$  is the expectation of a folded standard normal distribution. Then, union bound over  $j = 1, \dots, m$  gives us

$$\mathbb{P}\left(\frac{\sum_{i=1}^n \tilde{B}_{ji}}{n} \geq \mu - \tau \quad \forall j = 1, \dots, m\right) \geq \left(1 - \frac{1}{m^2}\right)^m \geq 1 - \frac{1}{m}.$$

where the last inequality follows from Bernoulli's inequality. Therefore, combining this result with inequality (12), we have with probability at least  $1 - \frac{1}{m}$ ,  $\tilde{\mathcal{W}} \subseteq \frac{1}{\mu - \tau} \mathcal{S}$ . Denote,  $\mathcal{S}' = \frac{1}{\sqrt{6 \log(mn)}} \mathcal{S}$ . Then, we have with probability at least  $1 - \frac{1}{m}$ ,  $\mathcal{S}' \subseteq \tilde{\mathcal{W}} \subseteq \kappa \cdot \mathcal{S}'$  where

$$\kappa = \frac{\sqrt{6 \log(mn)}}{\sqrt{\frac{2}{\pi}} - 2\sqrt{\frac{\log m}{n}}} = O\left(\sqrt{\log m + \log n}\right),$$

for sufficiently large values of  $m$  and  $n$ . We finally use Lemma 2.5 to conclude.  $\square$

## D Proofs of Theorem 3.1

To prove Theorem 3.1, we introduce the following Lemma which shows a deterministic bad instance where the optimal affine solution is  $\Theta(\sqrt{m})$  away from the optimal adjustable solution.

**Lemma D.1.** *Consider the two-stage adjustable problem (1) where:  $n = m$ ,  $\mathbf{c} = \mathbf{0}$ ,  $\mathbf{d} = \mathbf{e}$ ,  $\mathbf{A} = \mathbf{0}$ ,*

$$B_{ij} = \begin{cases} 1 & \text{if } i = j \\ \frac{1}{\sqrt{m}} & \text{if } i \neq j \end{cases} \quad (13)$$

and the uncertainty set is defined as

$$\mathcal{U} = \text{conv}(\mathbf{0}, \mathbf{e}_1, \dots, \mathbf{e}_m, \boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_m) \quad (14)$$

where  $\boldsymbol{\nu}_i = \frac{1}{\sqrt{m}}(\mathbf{e} - \mathbf{e}_i)$  for  $i = 1, \dots, m$ . Then,  $z_{\text{Aff}}(\mathbf{B}) = \Omega(\sqrt{m}) \cdot z_{\text{AR}}(\mathbf{B})$ .

**Proof.** First, let us prove that  $z_{\text{AR}}(\mathbf{B}) \leq 1$ . It is sufficient to define an adjustable solution only for the extreme points of  $\mathcal{U}$  because the constraints are linear. We define the following solution for all  $i = 1, \dots, m$ .

$$\mathbf{x} = \mathbf{0}, \quad \mathbf{y}(\mathbf{0}) = \mathbf{0}, \quad \mathbf{y}(\mathbf{e}_i) = \mathbf{e}_i, \quad \mathbf{y}(\boldsymbol{\nu}_i) = \frac{1}{m} \mathbf{e}.$$

We have  $\mathbf{B}\mathbf{y}(\mathbf{0}) = \mathbf{0}$  and for  $i \in [m]$

$$\mathbf{B}\mathbf{y}(\mathbf{e}_i) = \mathbf{e}_i + \frac{1}{\sqrt{m}}(\mathbf{e} - \mathbf{e}_i) \geq \mathbf{e}_i$$

and

$$\mathbf{B}\mathbf{y}(\boldsymbol{\nu}_i) = \frac{1}{m} \mathbf{B}\mathbf{e} = \left(\frac{1}{m} + \frac{m-1}{m\sqrt{m}}\right) \mathbf{e} \geq \frac{1}{\sqrt{m}} \mathbf{e} \geq \boldsymbol{\nu}_i$$

Therefore, the solution defined above is feasible. Moreover, the cost of our feasible solution is 1 because for all  $i \in [m]$ , we have

$$\mathbf{d}^T \mathbf{y}(\mathbf{e}_i) = \mathbf{d}^T \mathbf{y}(\boldsymbol{\nu}_i) = 1.$$

Hence,  $z_{\text{AR}}(\mathbf{B}) \leq 1$ . Now, it is sufficient to prove that  $z_{\text{Aff}}(\mathbf{B}) = \Omega(\sqrt{m})$ . From Lemma 8 in Bertsimas and Goyal [8], since our instance is symmetric, i.e.  $\mathcal{U}$  and  $\mathcal{W}$  are permutation invariant, where  $\mathcal{W}$  is the dualized uncertainty set, there exists an optimal solution for the affine problem (2) of the following form  $\mathbf{y}(\mathbf{h}) = \mathbf{P}\mathbf{h} + \mathbf{q}$  for  $\mathbf{h} \in \mathcal{U}$  where

$$\mathbf{P} = \begin{pmatrix} \theta & \mu & \dots & \mu \\ \mu & \theta & \dots & \mu \\ \vdots & \vdots & \ddots & \vdots \\ \mu & \mu & \dots & \theta \end{pmatrix} \quad (15)$$

and  $\mathbf{q} = \lambda \mathbf{e}$ .

We have  $\mathbf{y}(\mathbf{0}) = \lambda \mathbf{e} \geq \mathbf{0}$  hence

$$\lambda \geq 0. \quad (16)$$

We know that

$$z_{\text{Aff}}(\mathbf{B}) \geq \mathbf{d}^T \mathbf{y}(\mathbf{0}) = \lambda m. \quad (17)$$

**Case 1:** If  $\lambda \geq \frac{1}{6\sqrt{m}}$ , then from (17) we have  $z_{\text{Aff}}(\mathbf{B}) \geq \frac{\sqrt{m}}{6}$ .

**Case 2:** If  $\lambda \leq \frac{1}{6\sqrt{m}}$ . We have

$$\mathbf{y}(\mathbf{e}_1) = (\theta + \lambda)\mathbf{e}_1 + (\mu + \lambda)(\mathbf{e} - \mathbf{e}_1).$$

By feasibility of the solution, we have  $\mathbf{B}\mathbf{y}(\mathbf{e}_1) \geq \mathbf{e}_1$ , hence

$$(\theta + \lambda) + \frac{1}{\sqrt{m}}(m-1)(\mu + \lambda) \geq 1$$

Therefore  $\theta + \lambda \geq \frac{1}{2}$  or  $\frac{1}{\sqrt{m}}(m-1)(\mu + \lambda) \geq \frac{1}{2}$ .

**Case 2.1:** Suppose  $\frac{1}{\sqrt{m}}(m-1)(\mu + \lambda) \geq \frac{1}{2}$ . Therefore,

$$z_{\text{Aff}}(\mathbf{B}) \geq \mathbf{d}^T \mathbf{y}(\mathbf{e}_1) = \theta + \lambda + (m-1)(\mu + \lambda) \geq \frac{\sqrt{m}}{2}.$$

where the last inequality holds because  $\theta + \lambda \geq 0$  as  $\mathbf{y}(\mathbf{e}_1) \geq \mathbf{0}$ .

**Case 2.2:** Now suppose we have the other inequality i.e.  $\theta + \lambda \geq \frac{1}{2}$ . Recall that we have  $\lambda \leq \frac{1}{6\sqrt{m}}$  as well. Therefore,

$$\theta \geq \frac{1}{2} - \frac{1}{6\sqrt{m}} \geq \frac{1}{3}.$$

We have,

$$\mathbf{y}(\boldsymbol{\nu}_1) = \frac{1}{\sqrt{m}} ((\theta + (m-2)\mu)(\mathbf{e} - \mathbf{e}_1) + (m-1)\mu\mathbf{e}_1) + \lambda\mathbf{e}.$$

In particular we have ,

$$\begin{aligned} z_{\text{Aff}}(\mathbf{B}) &\geq \mathbf{d}^T \mathbf{y}(\boldsymbol{\nu}_1) = \frac{1}{\sqrt{m}} ((m-1)\theta + (m-1)^2\mu) + \lambda m \\ &\geq \frac{m-1}{\sqrt{m}} \left( \frac{1}{3} + (m-1)\mu \right). \end{aligned} \quad (18)$$

where the last inequality follows from  $\lambda \geq 0$  and  $\theta \geq \frac{1}{3}$ .

**Case 2.2.1:** If  $\mu \geq 0$  then from (18)

$$z_{\text{Aff}}(\mathbf{B}) \geq \frac{m-1}{3\sqrt{m}} = \Omega(\sqrt{m}).$$

**Case 2.2.2:** Now suppose that  $\mu < 0$ , by non-negativity of  $\mathbf{y}(\boldsymbol{\nu}_1)$  we have

$$\frac{m-1}{\sqrt{m}}\mu + \lambda \geq 0$$

i.e.

$$\mu \geq \frac{-\lambda\sqrt{m}}{m-1}$$

and from (18)

$$\begin{aligned} z_{\text{Aff}}(\mathbf{B}) &\geq \frac{m-1}{\sqrt{m}} \left( \frac{1}{3} + (m-1)\mu \right) \\ &\geq \frac{m-1}{\sqrt{m}} \left( \frac{1}{3} - \lambda\sqrt{m} \right) \\ &\geq \frac{m-1}{\sqrt{m}} \left( \frac{1}{3} - \frac{1}{6} \right) = \frac{m-1}{6\sqrt{m}} = \Omega(\sqrt{m}). \end{aligned}$$

We conclude that in all cases  $z_{\text{Aff}}(\mathbf{B}) = \Omega(\sqrt{m})$  and consequently  $z_{\text{Aff}}(\mathbf{B}) = \Omega(\sqrt{m}) \cdot z_{\text{AR}}(\mathbf{B})$ .  $\square$

### Proof of Theorem 3.1

**Proof.** Denote  $\mathcal{W} = \{\mathbf{w} \in \mathbb{R}_+^m \mid \mathbf{B}^T \mathbf{w} \leq \bar{d}\mathbf{e}\}$  and  $\tilde{\mathcal{W}} = \{\mathbf{w} \in \mathbb{R}_+^m \mid \tilde{\mathbf{B}}^T \mathbf{w} \leq \bar{d}\mathbf{e}\}$  where  $\mathbf{B}$  is defined in (13) and  $\tilde{\mathbf{B}}$  is defined in (8). Since for all  $i, j$  in  $\{1, \dots, m\}$  we have  $\tilde{B}_{ij} \leq B_{ij}$ . Hence, for any  $\mathbf{w} \in \mathcal{W}$ , we have  $\tilde{\mathbf{B}}^T \mathbf{w} \leq \mathbf{B}^T \mathbf{w} \leq \bar{d}\mathbf{e}$ . Therefore  $\mathbf{w} \in \tilde{\mathcal{W}}$  and consequently  $\mathcal{W} \subseteq \tilde{\mathcal{W}}$ .

Now, suppose  $\mathbf{w} \in \tilde{\mathcal{W}}$ , we have for all  $i = 1, \dots, m$

$$w_i + \frac{1}{\sqrt{m}} \sum_{\substack{j=1 \\ j \neq i}}^m \tilde{u}_{ji} w_j \leq \bar{d}. \quad (19)$$

By taking the sum over  $i$ , dividing by  $m$  and rearranging, we get

$$\sum_{i=1}^m w_i \left( \frac{1}{m} + \frac{1}{m\sqrt{m}} \sum_{\substack{j=1 \\ j \neq i}}^m \tilde{u}_{ij} \right) \leq \bar{d}. \quad (20)$$

Here, similarly to the proof of Lemma 2.1 we apply Hoeffding's inequality [18](see appendix B), with  $\tau = \sqrt{\frac{\log m}{m-1}}$

$$\mathbb{P}\left(\frac{\sum_{j \neq i}^m \tilde{u}_{ij}}{m-1} \geq \frac{1}{2} - \tau\right) \geq 1 - \exp(-2(m-1)\tau^2) = 1 - \frac{1}{m^2}$$

and we take a union bound over  $j = 1, \dots, m$

$$\mathbb{P}\left(\frac{\sum_{i=1}^n \tilde{u}_{ij}}{m-1} \geq \frac{1}{2} - \epsilon \quad \forall j = 1, \dots, m\right) \geq \left(1 - \frac{1}{m^2}\right)^m \geq 1 - \frac{1}{m}. \quad (21)$$

where the last inequality follows from Bernoulli's inequality. Therefore, we conclude from (20) and (21), that with probability at least  $1 - \frac{1}{m}$  we have  $\beta \sum_{i=1}^m w_i \leq \bar{d}$  where  $\beta = \frac{1}{m} + \frac{m-1}{m\sqrt{m}}(\frac{1}{2} - \tau) \geq \frac{1}{4\sqrt{m}}$  for  $m$  sufficiently large. Note from (19) that for all  $i$  we have  $w_i \leq \bar{d}$ . Hence with probability at least  $1 - \frac{1}{m}$ , we have for all  $i = 1, \dots, m$

$$\mathbf{B}_i^T \mathbf{w} = w_i + \frac{1}{\sqrt{m}} \sum_{\substack{j=1 \\ j \neq i}}^m w_j \leq \bar{d} + \frac{\bar{d}}{\beta\sqrt{m}} \leq 5 \cdot \bar{d}$$

Therefore,  $\mathbf{w} \in 5 \cdot \mathcal{W}$  for any  $\mathbf{w}$  in  $\mathcal{W}$  and consequently we have with probability at least  $1 - \frac{1}{m}$ ,  $\tilde{\mathcal{W}} \subseteq 5 \cdot \mathcal{W}$ . All together we have proved with probability at least  $1 - \frac{1}{m}$   $\mathcal{W} \subseteq \tilde{\mathcal{W}} \subseteq 5 \cdot \mathcal{W}$ . This implies with probability at least  $1 - \frac{1}{m}$ , that  $z_{d-\text{Aff}}(\tilde{\mathbf{B}}) \geq z_{d-\text{Aff}}(\mathbf{B})$  and  $z_{d-\text{AR}}(\mathbf{B}) \geq \frac{z_{d-\text{AR}}(\tilde{\mathbf{B}})}{5}$ . We know from Lemma 2.3 and Lemma 2.2 that the dualized and primal are the same both for the adjustable problem and affine problem. Hence, with probability at least  $1 - \frac{1}{m}$ , we have  $z_{\text{Aff}}(\tilde{\mathbf{B}}) \geq z_{\text{Aff}}(\mathbf{B})$  and  $z_{\text{AR}}(\mathbf{B}) \geq \frac{z_{\text{AR}}(\tilde{\mathbf{B}})}{5}$ .

Moreover, we know from Lemma D.1 that  $z_{\text{Aff}}(\mathbf{B}) \geq \Omega(\sqrt{m}) \cdot z_{\text{AR}}(\mathbf{B})$ . Therefore,  $z_{\text{Aff}}(\tilde{\mathbf{B}}) \geq \Omega(\sqrt{m}) z_{\text{AR}}(\tilde{\mathbf{B}})$  with probability at least  $1 - \frac{1}{m}$ . □

## E LP and MIP formulations for the empirical section

**LP formulation for the affine problem.** The affine problem (2) can be formulated as the following LP

$$\begin{aligned} z_{\text{Aff}}(\mathbf{B}) &= \min \mathbf{c}^T \mathbf{x} + z \\ z - \mathbf{d}^T \mathbf{q} &\geq \mathbf{r}^T \mathbf{v} \\ \mathbf{R}^T \mathbf{v} &\geq \mathbf{P}^T \mathbf{d} \\ \mathbf{A} \mathbf{x} + \mathbf{B} \mathbf{q} &\geq \mathbf{V}^T \mathbf{r} \\ \mathbf{R}^T \mathbf{V} &\geq \mathbf{I}_m - \mathbf{B} \mathbf{P} \\ \mathbf{q} &\geq \mathbf{U}^T \mathbf{r} \\ \mathbf{U}^T \mathbf{R} + \mathbf{P} &\geq \mathbf{0} \\ \mathbf{x} \in \mathbb{R}_+^n, \mathbf{v} \in \mathbb{R}_+^L, \mathbf{U} \in \mathbb{R}_+^{L \times n}, \mathbf{V} \in \mathbb{R}_+^{L \times m} \end{aligned} \quad (22)$$

**Proof.** The affine problem (2) can be reformulated as follows

$$\begin{aligned} z_{\text{Aff}}(\mathbf{B}) &= \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} + z \\ z &\geq \mathbf{d}^T (\mathbf{P} \mathbf{h} + \mathbf{q}) \quad \forall \mathbf{h} \in \mathcal{U} \\ \mathbf{A} \mathbf{x} + \mathbf{B} (\mathbf{P} \mathbf{h} + \mathbf{q}) &\geq \mathbf{h} \quad \forall \mathbf{h} \in \mathcal{U} \\ \mathbf{P} \mathbf{h} + \mathbf{q} &\geq \mathbf{0} \quad \forall \mathbf{h} \in \mathcal{U} \\ \mathbf{x} &\in \mathbb{R}_+^n \end{aligned}$$

We use standard duality techniques to derive formulation (22). The first constraint is equivalent to

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

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

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

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

$$\begin{aligned} z - \mathbf{d}^T \mathbf{q} &\geq \mathbf{r}^T \mathbf{v} \\ \mathbf{R}^T \mathbf{v} &\geq \mathbf{P}^T \mathbf{d} \\ \mathbf{v} &\in \mathbb{R}_+^L \end{aligned}$$

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

$$\mathbf{Ax} + \mathbf{Bq} \geq \max_{\substack{\mathbf{R}\mathbf{h} \leq \mathbf{r} \\ \mathbf{h} \geq \mathbf{0}}} \mathbf{h}(\mathbf{I}_m - \mathbf{BP})$$

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

$$\begin{aligned} \mathbf{Ax} + \mathbf{Bq} &\geq \mathbf{V}^T \mathbf{r} \\ \mathbf{R}^T \mathbf{V} &\geq \mathbf{I}_m - \mathbf{BP} \\ \mathbf{V} &\in \mathbb{R}_+^{L \times m} \end{aligned}$$

Similarly, the last constraint

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

is equivalent to

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

□

#### MIP formulation for the separation adjustable problem.

The separation problem (10) can be formulated as the following MIP

$$\begin{aligned} \max \quad & \sum_{i=1}^m \sum_{j=-\Delta_{\mathcal{W}}}^s \sum_{k=-\Delta_{\mathcal{U}}}^s \frac{1}{2^{j+k}} \cdot \gamma_{ijk} - (\mathbf{A}\hat{\mathbf{x}})^T \mathbf{w} \\ \mathbf{w} = \quad & \sum_{i=1}^m \sum_{j=-\Delta_{\mathcal{W}}}^s \frac{\beta_{ij}}{2^j} \cdot \mathbf{e}_i \\ \mathbf{h} = \quad & \sum_{i=1}^m \sum_{k=-\Delta_{\mathcal{U}}}^s \frac{\alpha_{ik}}{2^k} \cdot \mathbf{e}_i \\ \gamma_{ijk} &\leq \beta_{ij} & \forall i \in [m], j \in [-\Delta_{\mathcal{U}}, s], k \in [-\Delta_{\mathcal{W}}, s] \\ \gamma_{ijk} &\leq \alpha_{ik} & \forall i \in [m], j \in [-\Delta_{\mathcal{U}}, s], k \in [-\Delta_{\mathcal{W}}, s] \\ \gamma_{ijk} + 1 &\geq \alpha_{ik} + \beta_{ij} & \forall i \in [m], j \in [-\Delta_{\mathcal{U}}, s], k \in [-\Delta_{\mathcal{W}}, s] \\ \alpha_{ik}, \beta_{ij}, \gamma_{ijk} &\in \{0, 1\} & \forall i \in [m], j \in [-\Delta_{\mathcal{U}}, s], k \in [-\Delta_{\mathcal{W}}, s] \\ \mathbf{R}\mathbf{h} &\leq \mathbf{r} \\ \mathbf{B}^T \mathbf{w} &\leq \mathbf{d} \end{aligned} \tag{23}$$

where  $s = \lceil \log_2 \left( \frac{m}{\epsilon} \right) \rceil$ ,  $\Delta_{\mathcal{W}}$  is an upper bound on any component of  $w \in \mathcal{W}$ ,  $\Delta_{\mathcal{U}}$  is an upper bound on any component of  $h \in \mathcal{U}$  and  $\epsilon$  is the accuracy of the problem.

**Proof.** The separation problem (10) is equivalent to solving the following problem for given  $\hat{\mathbf{x}}$

$$\max_{\substack{\mathbf{h} \in \mathcal{U} \\ \mathbf{w} \in \mathcal{W}}} \mathbf{h}^T \mathbf{w} - (\mathbf{A}\hat{\mathbf{x}})^T \mathbf{w}$$

The constraints of the above problem are linear and the second term in the objective function is linear as well. So we will focus only on the first term  $\mathbf{h}^T \mathbf{w}$  which is a bilinear function and write it in terms of linear constraints and binary variables. Let us write  $\mathbf{h} = \sum_{i=1}^m h_i \mathbf{e}_i$ . For all  $i \in [m]$  we digitize the component  $h_i$  as follows

$$h_i = \sum_{k=-\Delta_{\mathcal{U}}}^s \frac{\alpha_{ik}}{2^k}$$

where  $s = \lceil \log_2 \left( \frac{m}{\epsilon} \right) \rceil$ ,  $\Delta_{\mathcal{U}}$  is an upper bound on any  $h_i$  and  $\alpha_{ik}$  are binary variables. This digitization gives an approximation to  $h_i$  within  $\frac{\epsilon}{m}$  which translates to an accuracy of  $\epsilon$  in the objective function. We have

$$\mathbf{h} = \sum_{i=1}^m \sum_{k=-\Delta_{\mathcal{U}}}^s \frac{\alpha_{ik}}{2^k} \cdot \mathbf{e}_i$$

Similarly, we have

$$\mathbf{w} = \sum_{i=1}^m \sum_{j=-\Delta_{\mathcal{W}}}^s \frac{\beta_{ij}}{2^j} \cdot \mathbf{e}_i$$

where  $\Delta_{\mathcal{W}}$  is an upper bound on any component of  $w \in \mathcal{W}$ . Therefore, the first term in the objective function becomes

$$\sum_{i=1}^m \sum_{j=-\Delta_{\mathcal{W}}}^s \sum_{k=-\Delta_{\mathcal{U}}}^s \frac{1}{2^{j+k}} \cdot \alpha_{ik} \beta_{ij}$$

The final step is to linearize the term  $\alpha_{ik} \beta_{ij}$ . We set,  $\alpha_{ik} \beta_{ij} = \gamma_{ijk}$  where again  $\gamma_{ijk}$  is a binary variable. Since all the variables here are binary we can express  $\gamma_{ijk}$  using only linear constraints as follows

$$\begin{aligned} \gamma_{ijk} &\leq \beta_{ij} \\ \gamma_{ijk} &\leq \alpha_{ik} \\ \gamma_{ijk} + 1 &\geq \alpha_{ik} + \beta_{ij} \end{aligned}$$

which leads to formulation (23).

□