Electronic Companion to Thomä, Schiffer, and Wiesemann: A Note on Piecewise Affine Decision Rules for Robust, Stochastic, and Data-Driven Optimization

EC.1. Proof of Lemma 1

Lemma 1 generalizes Lemma 4.3 of Georghiou et al. (2015), which covers the special case where $G = 1$ and thus $[\underline{\boldsymbol{\theta}}_1, \overline{\boldsymbol{\theta}}_1] = [\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]$. The key difference in our setting is that $[\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g] \neq [\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]$ in general. In this case, the outer approximation of Georghiou et al. (2015), which utilizes $\mathrm{conv}(F([\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]))$ instead of $\mathrm{conv}(F([\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g]))$, contains realizations $\boldsymbol{\xi}'$ that are not contained in the convex hull $\mathrm{conv}(F([\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g]))$. In particular, components ξ'_{ij} corresponding to intervals $[z_{i,j-1}, z_{ij}]$ that are (partially) outside of the box $[\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g]$, that is, $z_{i,j-1} < \underline{\theta}_{gi}$ or $z_{ij} > \overline{\theta}_{gi}$, may be smaller or larger than the respective components of any realization in the convex hull, respectively, that is, $\xi'_{ij} < \xi''_{ij}$ (when $z_{i,j-1} < \underline{\theta}_{gi}$) or $\xi'_{ij} > \xi''_{ij}$ (when $z_{ij} > \overline{\theta}_{gi}$) for all $\boldsymbol{\xi}'' \in \mathrm{conv}(F([\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g]))$. Lemma 1 addresses this issue by tightening the formulation for these components.

Proof of Lemma 1. We prove the lemma by first showing that $\mathrm{conv}(F([\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g]))$ decomposes by components $i \in [I]$ and then dealing with each component separately. By rectangularity of $[\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g]$, we find

$$\mathrm{conv}(F([\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g])) = \mathrm{conv}(F(\bigtimes_{i\in[I]} [\underline{\theta}_{gi}, \overline{\theta}_{gi}])) \overset{\text{(a)}}{=} \mathrm{conv}(\bigtimes_{i\in[I]} F_i([\underline{\theta}_{gi}, \overline{\theta}_{gi}])) \overset{\text{(b)}}{=} \bigtimes_{i\in[I]} \mathrm{conv}(F_i([\underline{\theta}_{gi}, \overline{\theta}_{gi}])),$$

where (a) follows from F decomposing by components θ_i, and (b) follows from standard convexity properties of Cartesian products.

For the remainder of the proof, we fix $i \in [I]$ and consider $\mathrm{conv}(F_i([\underline{\theta}_{gi}, \overline{\theta}_{gi}]))$. Our arguments are based on the intuition that F_i maps to edges of the hyperrectangle spanned by the breakpoints visualized in Figure 1 (b). Let $\hat{Z} = \{z_{ij} : \underline{\theta}_{gi} \leq z_{ij} \leq \overline{\theta}_{gi},\ j \in [J_i]\} \cup \{\underline{\theta}_{gi}, \overline{\theta}_{gi}\}$ be the union of all breakpoints contained in $[\underline{\theta}_{gi}, \overline{\theta}_{gi}]$ and the bounds of the interval. Geometrically, $F_i(\hat{Z})$ coincides with the extreme points of $F_i([\underline{\theta}_{gi}, \overline{\theta}_{gi}])$, and thus $\mathrm{conv}(F_i([\underline{\theta}_{gi}, \overline{\theta}_{gi}])) = \mathrm{conv}(F_i(\hat{Z}))$. Let $H\colon [0,1]^{J_i} \to \mathbb{R}^{J_i}$ be the affine transformation defined component-wise via $H_j(\zeta_j) = F_{ij}(\underline{\theta}_i) + \zeta_j(F_{ij}(\overline{\theta}_i) - F_{ij}(\underline{\theta}_i))$, and let $\boldsymbol{e}^j = \sum_{j'\in[j]} \boldsymbol{e}_{j'}$ be the vector with ones in the first j components and zero otherwise. One readily verifies that

$$H(\boldsymbol{e}^j) = \begin{cases} F_i(\overline{\theta}_{gi}) & \text{if } z_{ij} \geq \overline{\theta}_{gi} \\ F_i(z_{ij}) & \text{if } \underline{\theta}_{gi} \leq z_{ij} \leq \overline{\theta}_{gi} \\ F_i(\underline{\theta}_{gi}) & \text{if } z_{ij} \leq \underline{\theta}_{gi}. \end{cases}$$

Thus, $F_i(\hat{Z}) = H(\{\boldsymbol{e}^j : j \in [J_i]\})$ and consequently

$$\mathrm{conv}(F_i(\hat{Z})) = \mathrm{conv}(H(\{\boldsymbol{e}^j : j \in [J_i]\})) = H(\mathrm{conv}(\{\boldsymbol{e}^j : j \in [J_i]\})),$$

where the second identity follows from H being affine. Note that

$$\mathrm{conv}(\{\boldsymbol{e}^j : j \in [J_i]\}) = \{\boldsymbol{\zeta} \in [0,1]^{J_i} : \zeta_j \geq \zeta_{j+1}\ \forall j \in [J_i - 1]\},$$

where the "⊆" direction is immediate and the "⊇" direction follows from $\boldsymbol{\zeta} = \sum_{j \in [J_i]_0} (\zeta_j - \zeta_{j+1}) \boldsymbol{e}^j$, where we use $\zeta_0 = 1$ and $\zeta_{J_i+1} = 0$ for notational convenience. Substituting $\boldsymbol{\xi}' = H(\boldsymbol{\zeta})$ and noting that $H_j^{-1}(\xi'_j) = \frac{\xi'_j - F_{ij}(\underline{\theta}_{gi})}{F_{ij}(\overline{\theta}_{gi}) - F_{ij}(\underline{\theta}_{gi})}$ yields

$$
H(\operatorname{conv}(\{\boldsymbol{e}^j \,:\, j \in [J_i]\})) = \left\{ \boldsymbol{\xi}' \in \mathbb{R}^{J_i} : \begin{array}{ll} F_{ij}(\underline{\theta}_{gi}) \le \xi'_j \le F_{ij}(\overline{\theta}_{gi}) & \forall j \in [J_i] \\ \dfrac{\xi'_j - F_{ij}(\underline{\theta}_{gi})}{F_{ij}(\overline{\theta}_{gi}) - F_{ij}(\underline{\theta}_{gi})} \ge \dfrac{\xi'_{j+1} - F_{i,j+1}(\underline{\theta}_{gi})}{F_{i,j+1}(\overline{\theta}_{gi}) - F_{i,j+1}(\underline{\theta}_{gi})} & \forall j \in [J_i - 1] \end{array} \right\},
$$

which is the i^{th} component of the right-hand-side in the statement of the lemma. □

EC.2. Proof of Theorem 1

We prove a slightly more general version of Theorem 1, where the constraint matrices $\boldsymbol{A}_g$ may depend on the uncertainty realizations. In particular, we study the problem

$$
\begin{array}{ll} \underset{\boldsymbol{x} \in \mathcal{F}}{\text{minimize}} & \mathbb{E}_{\mathbb{P}} \left[\boldsymbol{c}(\tilde{\boldsymbol{\xi}})^\top \boldsymbol{x}(\tilde{\boldsymbol{\xi}}) \right] \\ \text{subject to} & \boldsymbol{A}_g(\boldsymbol{\xi}) \boldsymbol{x}(\boldsymbol{\xi}) \le \boldsymbol{b}_g(\boldsymbol{\xi}) \quad \forall g \in [G], \boldsymbol{\xi} \in \Xi_g \end{array} \tag{$\mathcal{P}_{\text{A}}(\mathbb{P}, \Xi)$}
$$

and its lifted version

$$
\begin{array}{ll} \underset{\boldsymbol{x}' \in \mathcal{F}'}{\text{minimize}} & \mathbb{E}_{\mathbb{P}'} \left[\boldsymbol{c}(R(\tilde{\boldsymbol{\xi}}'))^\top \boldsymbol{x}'(\tilde{\boldsymbol{\xi}}') \right] \\ \text{subject to} & \boldsymbol{A}_g(R(\boldsymbol{\xi}')) \boldsymbol{x}'(\boldsymbol{\xi}') \le \boldsymbol{b}_g(R(\boldsymbol{\xi}')) \quad \forall g \in [G], \boldsymbol{\xi}' \in \Xi'_g. \end{array} \tag{$\mathcal{P}'_{\text{A}}(\mathbb{P}', \Xi')$}
$$

These problems are significantly more general than the problems studied in the main part of the paper and, among others, also include robust optimization with uncertain objectives $\boldsymbol{c}(\boldsymbol{\xi})$ via the epigraph reformulation from Example 2. In contrast to $\mathcal{P}(\mathbb{P}, \Xi)$ studied in the main paper, affine relaxations of $\mathcal{P}_{\text{A}}(\mathbb{P}, \Xi)$ no longer yield tractable counterparts in general (Ben-Tal et al. 2004, Example 3.1). Thus, finding optimal affine policies for these problems is generally intractable. From a theoretical perspective, it is nonetheless interesting to see that the limitations of lifting policies using outer approximations $\overline{\Xi}'$ extend to this more general setting.

THEOREM 1′. *For $G = 1$, there exists a subset $\overline{\Psi}' \subseteq \overline{\Xi}'$ with $R(\overline{\Psi}') = \Xi$ such that every feasible solution $\boldsymbol{x}'$ in $\operatorname{Aff} \mathcal{P}'_{\text{A}}(\mathbb{P}', \overline{\Xi}')$ has a corresponding feasible solution $\boldsymbol{x}$ in $\operatorname{Aff} \mathcal{P}_{\text{A}}(\mathbb{P}, \Xi)$ with $\boldsymbol{x}'(\boldsymbol{\xi}') = \boldsymbol{x}(R(\boldsymbol{\xi}'))$ for all $\boldsymbol{\xi}' \in \overline{\Psi}'$.*

REMARK EC.1. Theorem 1′ still holds when we replace the requirement $G = 1$ with the weaker requirement $[\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g] = [\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]$ for all $g \in [G]$. As we are not aware of any interesting problems with $G > 1$ that meet this weaker requirement, we focus on the relevant cases where $G = 1$.

Proof of Theorem 1′. Consider the affine approximation

$$
\overline{F}_{ij}(\theta_i) \;=\; (z_{ij} - z_{i,j-1}) \cdot \frac{\theta_i - \underline{\theta}_i}{\overline{\theta}_i - \underline{\theta}_i} \qquad \forall i \in [I], j \in [J_i]
$$

of the folding operator F and the associated affine approximation $\overline{L} = \overline{F} \circ \boldsymbol{E}$ of the lifting operator L. We prove the statement of the theorem in three steps: (i) we show that $\overline{\Psi}' = \overline{L}(\Xi)$ satisfies $\overline{\Psi}' \subseteq \overline{\Xi}'$; (ii) we verify that $R(\overline{\Psi}') = \Xi$; and (iii) we show that for every feasible solution $\boldsymbol{x}'$ in $\mathrm{Aff}\,\mathcal{P}'_{\mathrm{A}}(\mathbb{P}', \overline{\Xi}')$, the corresponding solution $\boldsymbol{x} = \boldsymbol{x}' \circ \overline{L}$ is feasible in $\mathrm{Aff}\,\mathcal{P}_{\mathrm{A}}(\mathbb{P}, \Xi)$ and satisfies $\boldsymbol{x}'(\boldsymbol{\xi}') = \boldsymbol{x}(R(\boldsymbol{\xi}'))$ for all $\boldsymbol{\xi}' \in \overline{\Psi}'$. Since $G = 1$, we drop the subscript g throughout the proof.

In view of Step (i), we show that $\overline{\Psi}' \subseteq \overline{\Xi}'$ by verifying that $\overline{\Psi}'$ is contained in each of the two sets on the right-hand side of (3). We have $\overline{\Psi}' \subseteq \{\boldsymbol{\xi}' : F^+(\boldsymbol{\xi}') \in \Theta\}$ whenever $F^+(\overline{\Psi}') \subseteq \Theta$, that is, whenever $F^+ \circ \overline{F} \circ \boldsymbol{E}\Xi \subseteq \Theta$. To this end, we note that F^+ is a left-inverse of $\overline{F}$ since

$$F_i^+(\overline{F}(\boldsymbol{\theta})) \stackrel{\text{(a)}}{=} \underline{\theta}_i + \sum_{j \in [J_i]} (z_{ij} - z_{i,j-1}) \cdot \frac{\theta_i - \underline{\theta}_i}{\overline{\theta}_i - \underline{\theta}_i} \stackrel{\text{(b)}}{=} \underline{\theta}_i + (\overline{\theta}_i - \underline{\theta}_i) \cdot \frac{\theta_i - \underline{\theta}_i}{\overline{\theta}_i - \underline{\theta}_i} = \theta_i \qquad \forall i \in [I], \boldsymbol{\theta} \in \Theta,$$

where (a) uses the definitions of F^+ and $\overline{F}$ and (b) contracts the sum. The set $\overline{\Psi}'$ is therefore contained in the first set on the right-hand side of (3) whenever $\boldsymbol{E}\Xi \subseteq \Theta$, which holds by construction. To see that $\overline{\Psi}' = \overline{F}(\Theta)$ is also contained in the second set on the right-hand side of (3), we show that any $\overline{F}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$, satisfies the conditions in the representation of $\mathrm{conv}(F([\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]))$ offered by Lemma 1. To see that $F_{ij}(\underline{\theta}_i) \leq \overline{F}_{ij}(\theta_i) \leq F_{ij}(\overline{\theta}_i)$, we observe that

$$F_{ij}(\underline{\theta}_i) \stackrel{\text{(a)}}{=} 0 \stackrel{\text{(b)}}{\leq} (z_{ij} - z_{i,j-1}) \cdot \frac{\theta_i - \underline{\theta}_i}{\overline{\theta}_i - \underline{\theta}_i} \stackrel{\text{(c)}}{\leq} (z_{ij} - z_{i,j-1}) \stackrel{\text{(d)}}{=} F_{ij}(\overline{\theta}_i) \qquad \forall \boldsymbol{\theta} \in \Theta, i \in [I], j \in [J_i],$$

where (a) and (d) follow from the definition of F_{ij} on the bounds $\underline{\theta}_i$ and $\overline{\theta}_i$ respectively; (b) and (c) follow from $z_{ij} \geq z_{i,j-1}$ and $\underline{\theta}_i \leq \theta_i \leq \overline{\theta}_i$. To see that the second condition in the representation of $\mathrm{conv}(F([\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]))$ is satisfied as well, we note that

$$\frac{\overline{F}_{ij}(\theta_i) - F_{ij}(\underline{\theta}_i)}{F_{ij}(\overline{\theta}_i) - F_{ij}(\underline{\theta}_i)} \stackrel{\text{(a)}}{=} \frac{\overline{F}_{ij}(\theta_i)}{z_{ij} - z_{i,j-1}} \stackrel{\text{(b)}}{=} \frac{z_{ij} - z_{i,j-1}}{z_{ij} - z_{i,j-1}} \cdot \frac{\theta_i - \underline{\theta}_i}{\overline{\theta}_i - \underline{\theta}_i} \stackrel{\text{(c)}}{=} \frac{\theta_i - \underline{\theta}_i}{\overline{\theta}_i - \underline{\theta}_i} \stackrel{\text{(d)}}{=} \frac{\overline{F}_{i,j+1}(\theta_i) - F_{i,j+1}(\underline{\theta}_i)}{F_{i,j+1}(\overline{\theta}_i) - F_{i,j+1}(\underline{\theta}_i)},$$

where (a) is due to the definition of F_{ij} and the bounds $\underline{\theta}_i$ and $\overline{\theta}_i$, (b) follows from the definition of $\overline{F}$, (c) simplifies the term, and (d) follows by replacing j with $j+1$ in our chain of equations (a)–(c). This concludes Step (i).

As for Step (ii), we note that $R = \boldsymbol{E}^+ \circ F^+$ and that Step (i) has shown that F^+ is a left-inverse of $\overline{F}$. Thus, R is a left-inverse of $\overline{L}$, which implies that $R(\overline{\Psi}') = R(\overline{L}(\Xi)) = \Xi$.

In view of Step (iii), finally, fix any $\boldsymbol{x}'$ feasible in $\mathrm{Aff}\,\mathcal{P}'_{\mathrm{A}}(\mathbb{P}', \overline{\Xi}')$ and set $\boldsymbol{x} = \boldsymbol{x}' \circ \overline{L}$. Let $\boldsymbol{\xi}' \in \overline{\Psi}'$. By definition of $\overline{\Psi}' = \overline{L}(\Xi)$, there exists some $\boldsymbol{\xi} \in \Xi$ such that $\boldsymbol{\xi}' = \overline{L}(\boldsymbol{\xi})$. With this we find

$$\boldsymbol{x}(R(\boldsymbol{\xi}')) \stackrel{\text{(a)}}{=} \boldsymbol{x}'(\overline{L}(R(\boldsymbol{\xi}'))) \stackrel{\text{(b)}}{=} \boldsymbol{x}'(\overline{L}(R(\overline{L}(\boldsymbol{\xi})))) \stackrel{\text{(c)}}{=} \boldsymbol{x}'(\overline{L}(\boldsymbol{\xi})) \stackrel{\text{(d)}}{=} \boldsymbol{x}'(\boldsymbol{\xi}'),$$

where (a) is due to the definition of $\boldsymbol{x}$, (b) and (d) follow from $\boldsymbol{\xi}' = \overline{L}(\boldsymbol{\xi})$, and (c) holds since R is a left-inverse of $\overline{L}$. As both $\boldsymbol{x}'$ and $\overline{L}$ are affine, so is $\boldsymbol{x}$. Further, $\boldsymbol{x}$ is feasible in $\mathrm{Aff}\,\mathcal{P}_{\mathrm{A}}(\mathbb{P}, \Xi)$ since

$$\boldsymbol{A}(\boldsymbol{\xi})\boldsymbol{x}(\boldsymbol{\xi}) \stackrel{\text{(a)}}{=} \boldsymbol{A}(R(\overline{L}(\boldsymbol{\xi}))\boldsymbol{x}'(\overline{L}(\boldsymbol{\xi})) \stackrel{\text{(b)}}{\leq} \boldsymbol{b}(R(\overline{L}(\boldsymbol{\xi}))) \stackrel{\text{(c)}}{=} \boldsymbol{b}(\boldsymbol{\xi}) \qquad \forall \boldsymbol{\xi} \in \Xi,$$

where (a) is due to the definition of $\boldsymbol{x}$ and R being a left-inverse of $\overline{L}$, (b) holds since $\boldsymbol{x}'$ is feasible in $\mathrm{Aff}\,\mathcal{P}'_{\mathrm{A}}(\mathbb{P}', \overline{\Xi}')$, and (c) follows from R being a left-inverse of $\overline{L}$. □

EC.3. Proof of Observation 1

Proof of Observation 1. We prove the inequality (4) by showing that for all $\boldsymbol{\xi}' \in \overline{\Xi}'$, we have

$$\begin{aligned} d(F^+(\boldsymbol{\xi}'), [\boldsymbol{z}^-, \boldsymbol{z}^+]) &= \min_{\boldsymbol{\theta}' \in [\boldsymbol{z}^-, \boldsymbol{z}^+]} \left\| F^+(\boldsymbol{\xi}') - \boldsymbol{\theta}' \right\|_1 = \sum_{i \in [I]} \min_{\theta'_i \in [z_i^-, z_i^+]} |F_i^+(\boldsymbol{\xi}') - \theta'_i| \\ &= \sum_{i \in [I]} (F_i^+(\boldsymbol{\xi}') - z_i^+)_+ + (z_i^- - F_i^+(\boldsymbol{\xi}'))_+ \\ &\overset{\text{(i)}}{=} \sum_{i \in [I]} \left(\sum_{j \in J_i^+} F_{ij}(F_i^+(\boldsymbol{\xi}')) + z_i^- - \underline{\theta}_i - \sum_{j \in J_i^-} F_{ij}(F_i^+(\boldsymbol{\xi}')) \right) \\ &\overset{\text{(ii)}}{\leq} \sum_{i \in [I]} \left(\sum_{j \in J_i^+} \xi'_{ij} + z_i^- - \underline{\theta}_i - \sum_{j \in J_i^-} \xi'_{ij} \right). \end{aligned}$$

Here, the first equality holds by construction of d, and the subsequent two equalities are immediate. In the remainder of the proof, we show that (i) and (ii) hold. Since $F_{ij}(F_i^+(\boldsymbol{\xi}')) = \xi'_{ij}$ for all $\boldsymbol{\xi}' \in \Xi'$, our above argument also implies that (4) is tight for all $\boldsymbol{\xi}' \in \Xi'$.

As for (i), recall the following property of F,

$$\underline{\theta}_i + \sum_{j' \in [j]} F_{ij'}(\theta_i) = \min\{z_{ij}, \theta_i\} \qquad \forall i \in [I], j \in [J_i], \boldsymbol{\theta} \in \Theta, \tag{EC.1}$$

and note that $F^+(\boldsymbol{\xi}') \in \Theta$ for any $\boldsymbol{\xi}' \in \overline{\Xi}'$. With this we find

$$\begin{aligned} (F_i^+(\boldsymbol{\xi}') - z_i^+)_+ &= F_i^+(\boldsymbol{\xi}') - \min\{F_i^+(\boldsymbol{\xi}'), z_i^+\} \overset{\text{(a)}}{=} \min\{F_i^+(\boldsymbol{\xi}'), \overline{\theta}_i\} - \min\{F_i^+(\boldsymbol{\xi}'), z_i^+\} \\ &\overset{\text{(b)}}{=} \sum_{j \in [J_i]} F_{ij}(F_i^+(\boldsymbol{\xi}')) + \underline{\theta}_i - \sum_{j \in [J_i], z_{ij} \leq z_i^+} F_{ij}(F_i^+(\boldsymbol{\xi}')) - \underline{\theta}_i \overset{\text{(c)}}{=} \sum_{j \in J_i^+} F_{ij}(F_i^+(\boldsymbol{\xi}')), \end{aligned}$$

where (a) holds since $F^+(\boldsymbol{\xi}') \in \Theta \subseteq [\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]$, (b) follows from $\overline{\theta}_i = z_{iJ_i}$ and applying (EC.1) to each of the minima, and (c) is due to $\{j \in [J_i] : \; z_{ij} \leq z_i^+\} = [J_i] \setminus J_i^+$. We further find

$$(z_i^- - F_i^+(\boldsymbol{\xi}'))_+ = z_i^- - \min\{z_i^-, F_i^+(\boldsymbol{\xi}')\} \overset{\text{(a)}}{=} z_i^- - \sum_{j \in J_i^-} F_{ij}(F_i^+(\boldsymbol{\xi}')) - \underline{\theta}_i,$$

where (a) holds due to (EC.1).

To prove (ii), we will show that

$$\sum_{j \in J_i^-} \xi'_{ij} \leq \sum_{j \in J_i^-} F_{ij}(F_i^+(\boldsymbol{\xi}')) \quad \text{and} \quad \sum_{j \in J_i^+} \xi'_{ij} \geq \sum_{j \in J_i^+} F_{ij}(F_i^+(\boldsymbol{\xi}')) \qquad \forall i \in [I], \boldsymbol{\xi}' \in \overline{\Xi}'. \tag{EC.2}$$

In view of the first inequality in (EC.2), we show the more general property

$$\sum_{j' \in [j]} \xi'_{ij'} \overset{\text{(a)}}{\leq} \min\{F_i^+(\boldsymbol{\xi}'), z_{ij}\} - \underline{\theta}_i \overset{\text{(b)}}{=} \sum_{j' \in [j]} F_{ij'}(F_i^+(\boldsymbol{\xi}')) \qquad \forall i \in [I], j \in [J_i], \boldsymbol{\xi}' \in \overline{\Xi}', \tag{EC.3}$$

where the identity (b) follows from (EC.1). For (a), we observe that

$$F_i^+(\boldsymbol{\xi}') - \underline{\theta}_i \overset{\text{(c)}}{=} \sum_{j \in [J_i]} \xi'_{ij} \overset{\text{(d)}}{\geq} \sum_{j' \in [j]} \xi'_{ij'} \quad \text{and} \quad z_{ij} - \underline{\theta}_i \overset{\text{(e)}}{=} \sum_{j' \in [j]} z_{ij'} - z_{i,j'-1} \overset{\text{(f)}}{\geq} \sum_{j' \in [j]} \xi'_{ij'},$$

where (c) uses the definition of F^+, (d) holds since $\xi'_{ij} \geq F_{ij}(\underline{\theta}_i) = 0$, (e) follows from contracting the sum, and (f) holds since $\xi'_{ij'} \leq F_{ij'}(\overline{\theta}_i) = z_{ij'} - z_{i,j'-1}$. The second inequality in (EC.2), finally, holds since

$$\begin{aligned}\sum_{j \in J_i^+} \xi'_{ij} = \sum_{j \in [J_i]} \xi'_{ij} - \sum_{j \in [J_i] \setminus J_i^+} \xi'_{ij} &\overset{\text{(a)}}{=} \sum_{j \in [J_i]} F_{ij}(F_i^+(\boldsymbol{\xi}')) - \sum_{j \in [J_i] \setminus J_i^+} \xi'_{ij} \\ &\overset{\text{(b)}}{\geq} \sum_{j \in [J_i]} F_{ij}(F_i^+(\boldsymbol{\xi}')) - \sum_{j \in [J_i] \setminus J_i^+} F_{ij}(F_i^+(\boldsymbol{\xi}')) = \sum_{j \in J_i^+} F_{ij}(F_i^+(\boldsymbol{\xi}')),\end{aligned}$$

where (a) follows from the definition of F and F^+ and (b) follows from (EC.3). □

EC.4. Proof of Observation 2

Proof of Observation 2. The proofs of (i) and (ii) largely follow the same arguments. In the following, we explicitly show (ii) and point out which parts need to be adjusted for the proof of (i). Our proof employs the primal-worst-dual-best strong duality argument of Zhen et al. (2025), which extends the duality argument of Beck and Ben-Tal (2009) to problem classes that encompass $\text{Aff}\,\mathcal{P}'(\mathbb{P}', \overline{\Xi}^\star)$. Under this scheme, the dual of $\text{Aff}\,\mathcal{P}'(\mathbb{P}', \overline{\Xi}^\star)$ is

$$\begin{aligned}\underset{\boldsymbol{Y},\boldsymbol{y},\boldsymbol{Z},\boldsymbol{z},\boldsymbol{\xi}',\boldsymbol{s}}{\text{maximize}} \quad & -\sup_{\boldsymbol{X},\boldsymbol{x}} \left(\langle \boldsymbol{X}, \boldsymbol{Y} \rangle + \langle \boldsymbol{x}, \boldsymbol{y} \rangle - \mathbb{E}_{\mathbb{P}'}\left[\boldsymbol{c}(R(\tilde{\boldsymbol{\xi}}'))^\top (\boldsymbol{X}\tilde{\boldsymbol{\xi}}' + \boldsymbol{x}) \right] \right) \\ & - \sum_{g \in [G], k \in [K_g]} \sup_{\boldsymbol{X},\boldsymbol{x}} \Big(\langle \boldsymbol{X}, \boldsymbol{Z}_{gk} \rangle + \langle \boldsymbol{x}, \boldsymbol{z}_{gk} \rangle \\ & \qquad - s_{gk} \left(\boldsymbol{A}_{gk}(\boldsymbol{X}\boldsymbol{\xi}'_{gk} + \boldsymbol{x}) - b_{gk}(R(\boldsymbol{\xi}'_{gk})) \right) \Big) \\ \text{subject to} \quad & \boldsymbol{Y} + \sum_{g \in [G], k \in [K_g]} \boldsymbol{Z}_{gk} = \boldsymbol{0} \\ & \boldsymbol{y} + \sum_{g \in [G], k \in [K_g]} \boldsymbol{z}_{gk} = \boldsymbol{0} \\ & \boldsymbol{\xi}'_{gk} \in \overline{\Xi}^\star_g, \quad s_{gk} \geq 0 \qquad \forall g \in [G], k \in [K_g],\end{aligned} \tag{EC.4}$$

where we use the affine representation $\boldsymbol{x}'(\boldsymbol{\xi}') = \boldsymbol{X}\boldsymbol{\xi}' + \boldsymbol{x}$.

By assumption, each set $\overline{\Xi}^\star_g$ is bounded, and $\text{Aff}\,\mathcal{P}'(\mathbb{P}', \overline{\Xi}^\star)$ affords a Slater point. Thus, strong duality holds by Theorem 4 (ii) of Zhen et al. (2025). Together with the assumed feasibility of $\text{Aff}\,\mathcal{P}'(\mathbb{P}', \overline{\Xi}^\star)$, this, in particular, implies that the dual has a finite optimal value. The suprema in the objective function only yield finite values, when $\boldsymbol{Y}$, $\boldsymbol{y}$, $\boldsymbol{Z}$, and $\boldsymbol{z}$ omit certain conditions. In particular, for the first term of the objective function in (EC.4), we find

$$\sup_{\boldsymbol{X},\boldsymbol{x}} \left(\langle \boldsymbol{X}, \boldsymbol{Y} \rangle + \langle \boldsymbol{x}, \boldsymbol{y} \rangle - \mathbb{E}_{\mathbb{P}'}\left[\boldsymbol{c}(R(\tilde{\boldsymbol{\xi}}'))^\top (\boldsymbol{X}\tilde{\boldsymbol{\xi}}' + \boldsymbol{x}) \right] \right)$$

$$\begin{aligned}
&\overset{(a)}{=} \sup_{\boldsymbol{X},\boldsymbol{x}} \left(\langle \boldsymbol{X}, \boldsymbol{Y} - \mathbb{E}_{\mathbb{P}'}[c(R(\tilde{\boldsymbol{\xi}}'))\tilde{\boldsymbol{\xi}}'^{\top}] \rangle + \langle \boldsymbol{x}, \boldsymbol{y} - \mathbb{E}_{\mathbb{P}'}[\boldsymbol{c}(R(\tilde{\boldsymbol{\xi}}'))] \rangle \right) \\
&\overset{(b)}{=} \begin{cases} 0 & \text{if } \boldsymbol{Y} = \mathbb{E}_{\mathbb{P}'}[c(R(\tilde{\boldsymbol{\xi}}'))\tilde{\boldsymbol{\xi}}'^{\top}] \text{ and } \boldsymbol{y} = \mathbb{E}_{\mathbb{P}'}[\boldsymbol{c}(R(\tilde{\boldsymbol{\xi}}'))] \\ \infty & \text{otherwise,} \end{cases}
\end{aligned}$$

where (a) follows from the linearity of the expectation operator, and (b) holds since $\boldsymbol{X}$ and $\boldsymbol{x}$ are free variables. For the second term of the objective function in (EC.4), we find

$$\begin{aligned}
&\sup_{\boldsymbol{X},\boldsymbol{x}} \left(\langle \boldsymbol{X}, \boldsymbol{Z}_{gk} \rangle + \langle \boldsymbol{x}, \boldsymbol{z}_{gk} \rangle - s_{gk} \left(\boldsymbol{A}_{gk}(\boldsymbol{X}\boldsymbol{\xi}'_{gk} + \boldsymbol{x}) - b_{gk}(R(\boldsymbol{\xi}'_{gk})) \right) \right) \\
&\overset{(a)}{=} \sup_{\boldsymbol{X},\boldsymbol{x}} \left(\langle \boldsymbol{X}, \boldsymbol{Z}_{gk} - s_{gk}\boldsymbol{A}_{gk}^{\top}\boldsymbol{\xi}'^{\top}_{gk} \rangle + \langle \boldsymbol{x}, \boldsymbol{z}_{gk} - s_{gk}\boldsymbol{A}_{gk}^{\top} \rangle + s_{gk} b_{gk}(R(\boldsymbol{\xi}'_{gk})) \right) \\
&\overset{(b)}{=} \begin{cases} s_{gk} b_{gk}(R(\boldsymbol{\xi}'_{gk})) & \text{if } \boldsymbol{Z}_{gk} = s_{gk}\boldsymbol{A}_{gk}^{\top}\boldsymbol{\xi}'^{\top}_{gk} \text{ and } \boldsymbol{z}_{gk} = s_{gk}\boldsymbol{A}_{gk}^{\top} \\ \infty & \text{otherwise,} \end{cases}
\end{aligned}$$

where (a) reorders terms and (b) holds since $\boldsymbol{X}$ and $\boldsymbol{x}$ are free variables. Using this, we substitute $\boldsymbol{Y}$, $\boldsymbol{y}$, $\boldsymbol{Z}$, and $\boldsymbol{z}$ in (EC.4) to arrive at the equivalent problem

$$\begin{aligned}
\underset{\boldsymbol{\xi}',\boldsymbol{s}}{\text{maximize}} \quad & -\sum_{g\in[G],k\in[K_g]} s_{gk} b_{gk}(R(\boldsymbol{\xi}'_{gk})) \\
\text{subject to} \quad & \sum_{g\in[G],k\in[K_g]} s_{gk}\boldsymbol{A}_{gk}^{\top}\boldsymbol{\xi}'^{\top}_{gk} = -\mathbb{E}_{\mathbb{P}'}[\boldsymbol{c}(R(\tilde{\boldsymbol{\xi}}'))\tilde{\boldsymbol{\xi}}'^{\top}] \\
& \sum_{g\in[G],k\in[K_g]} s_{gk}\boldsymbol{A}_{gk}^{\top} = -\mathbb{E}_{\mathbb{P}'}[\boldsymbol{c}(R(\tilde{\boldsymbol{\xi}}'))] \\
& \boldsymbol{\xi}'_{gk} \in \overline{\Xi}^{\star}_g, \quad s_{gk} \geq 0 \qquad \forall g \in [G], k \in [K_g].
\end{aligned} \tag{EC.5}$$

Problem (EC.5) contains bi-linear terms in $\boldsymbol{\xi}'$ and $\boldsymbol{s}$ and is thus non-convex in general. Next, we eliminate these bi-linear terms via another variable substitution. In particular, we reformulate the problem to have only bi-linearities of the form $s_{gk}\boldsymbol{\xi}'_{gk}$. For the term $s_{gk}b_{gk}(R(\boldsymbol{\xi}'_{gk}))$ in the objective of (EC.5), note that $b_{gk} \circ R$ is an affine function as both b_{gk} and R are affine. Thus, $b_{gk}(R(\cdot)) - b_{gk}(R(\boldsymbol{0}))$ is linear, which implies that $s_{gk}b_{gk}(R(\boldsymbol{\xi}'_{gk})) = b_{gk}(R(s_{gk}\boldsymbol{\xi}'_{gk})) + (s_{gk} - 1)b_{gk}(R(\boldsymbol{0}))$ for all $\boldsymbol{\xi}'_{gk} \in \overline{\Xi}^{\star}_g$ and $s_{gk} \geq 0$. For the term $s_{gk}\boldsymbol{A}_{gk}^{\top}\boldsymbol{\xi}'^{\top}_{gk}$ in the first constraint of (EC.5), we find that $s_{gk}\boldsymbol{A}_{gk}^{\top}\boldsymbol{\xi}'^{\top}_{gk} = \boldsymbol{A}_{gk}^{\top}(s_{gk}\boldsymbol{\xi}'_{gk})^{\top}$ by the linearity of matrix multiplications and transpositions. Substituting $\boldsymbol{\xi}^{\star}_{gk} = s_{gk}\boldsymbol{\xi}'_{gk}$, where $\boldsymbol{\xi}^{\star}_{gk} \in \{s_{gk}\boldsymbol{\xi}' : \boldsymbol{\xi}' \in \overline{\Xi}^{\star}_g\} = s_{gk}\overline{\Xi}^{\star}_g$, we find that (EC.5) is equivalent to

$$\begin{aligned}
\underset{\boldsymbol{\xi}^{\star},\boldsymbol{s}}{\text{maximize}} \quad & -\sum_{g\in[G],k\in[K_g]} b_{gk}(R(\boldsymbol{\xi}^{\star}_{gk})) + (s_{gk} - 1)b_{gk}(R(\boldsymbol{0})) \\
\text{subject to} \quad & \sum_{g\in[G],k\in[K_g]} \boldsymbol{A}_{gk}^{\top}\boldsymbol{\xi}^{\star\top}_{gk} = -\mathbb{E}_{\mathbb{P}'}[\boldsymbol{c}(R(\tilde{\boldsymbol{\xi}}'))\tilde{\boldsymbol{\xi}}'^{\top}] \\
& \sum_{g\in[G],k\in[K_g]} s_{gk}\boldsymbol{A}_{gk}^{\top} = -\mathbb{E}_{\mathbb{P}'}[\boldsymbol{c}(R(\tilde{\boldsymbol{\xi}}'))] \\
& \boldsymbol{\xi}^{\star}_{gk} \in s_{gk}\overline{\Xi}^{\star}_g, \quad s_{gk} \geq 0 \qquad \forall g \in [G], k \in [K_g].
\end{aligned}$$

We conclude the proof by showing that the condition $\boldsymbol{\xi}^\star_{gk} \in s_{gk}\overline{\Xi}^\star_g$ is conic representable for conic representable sets $\Xi_g = \{\boldsymbol{\xi} \in \mathbb{R}^N : \boldsymbol{V}_g\boldsymbol{\xi} \succeq_{\mathcal{K}_g} \boldsymbol{d}_g\}$. Indeed, using the definitions of $\overline{\Xi}^\star_g$ and Ξ_g, we find that

$$
\begin{aligned}
s\overline{\Xi}^\star_g = \Big\{ s\boldsymbol{\xi}' : \ & \boldsymbol{\xi}' \in \mathbb{R}^{N'}_+,\ \boldsymbol{\xi} \in \mathbb{R}^N, \\
& \boldsymbol{V}_g\boldsymbol{\xi} \succeq_{\mathcal{K}_g} \boldsymbol{d}_g,\ F^+(\boldsymbol{\xi}') = \boldsymbol{E}\boldsymbol{\xi} \\
& F_{ij}(\underline{\theta}_{gi}) \le \xi'_{ij} \le F_{ij}(\overline{\theta}_{gi}) && \forall i \in [I], j \in [J_i] \\
& \frac{\xi'_{ij} - F_{ij}(\underline{\theta}_{gi})}{F_{ij}(\overline{\theta}_{gi}) - F_{ij}(\underline{\theta}_{gi})} \ge \frac{\xi'_{i,j+1} - F_{i,j+1}(\underline{\theta}_{gi})}{F_{i,j+1}(\overline{\theta}_{gi}) - F_{i,j+1}(\underline{\theta}_{gi})} && \forall i \in [I], j \in [J_i] \ \text{with}\ \underline{\theta}_{gi} < z_{ij} < \overline{\theta}_{gi} \\
& d'(\boldsymbol{\xi}', [z^-, z^+]) \le \overline{d}_g([z^-, z^+]) && \forall [z^-, z^+] \in \mathcal{Z} \Big\}
\end{aligned}
\tag{EC.6}
$$

for any $s \ge 0$. Scaling all conditions in (EC.6) by s and using the linearity of matrix multiplications, (EC.6) is equivalent to

$$
\begin{aligned}
\Big\{ s\boldsymbol{\xi}' : \ & \boldsymbol{\xi}' \in \mathbb{R}^{N'}_+,\ \boldsymbol{\xi} \in \mathbb{R}^N, \\
& \boldsymbol{V}_g(s\boldsymbol{\xi}) \succeq_{\mathcal{K}_g} s\boldsymbol{d}_g,\ sF^+(\boldsymbol{\xi}') = \boldsymbol{E}(s\boldsymbol{\xi}) \\
& sF_{ij}(\underline{\theta}_{gi}) \le s\xi'_{ij} \le sF_{ij}(\overline{\theta}_{gi}) && \forall i \in [I], j \in [J_i] \\
& \frac{s\xi'_{ij} - sF_{ij}(\underline{\theta}_{gi})}{F_{ij}(\overline{\theta}_{gi}) - F_{ij}(\underline{\theta}_{gi})} \ge \frac{s\xi'_{i,j+1} - sF_{i,j+1}(\underline{\theta}_{gi})}{F_{i,j+1}(\overline{\theta}_{gi}) - F_{i,j+1}(\underline{\theta}_{gi})} && \forall i \in [I], j \in [J_i] \ \text{with}\ \underline{\theta}_{gi} < z_{ij} < \overline{\theta}_{gi} \\
& sd'(\boldsymbol{\xi}', [z^-, z^+]) \le s\overline{d}_g([z^-, z^+]) && \forall [z^-, z^+] \in \mathcal{Z} \Big\}.
\end{aligned}
\tag{EC.7}
$$

Using that F^+ is affine, we find that $sF^+(\boldsymbol{\xi}') = F^+(s\boldsymbol{\xi}') + (s-1)F^+(\boldsymbol{0})$. Similarly, $d'(\cdot, [z^-, z^+])$ being affine implies that $sd'(\boldsymbol{\xi}', [z^-, z^+]) = d'(s\boldsymbol{\xi}', [z^-, z^+]) + (s-1)d'(\boldsymbol{0}, [z^-, z^+])$. Thus, substituting $\hat{\boldsymbol{\xi}} = s\boldsymbol{\xi}$ and $\boldsymbol{\xi}^\star = s\boldsymbol{\xi}'$ in (EC.7) yields the equivalent formulation

$$
\begin{aligned}
\Big\{ \boldsymbol{\xi}^\star : \ & \boldsymbol{\xi}^\star \in \mathbb{R}^{N'}_+,\ \hat{\boldsymbol{\xi}} \in \mathbb{R}^N, \\
& \boldsymbol{V}_g\hat{\boldsymbol{\xi}} \succeq_{\mathcal{K}_g} s\boldsymbol{d}_g,\ F^+(\boldsymbol{\xi}^\star) + (s-1)F^+(\boldsymbol{0}) = \boldsymbol{E}\hat{\boldsymbol{\xi}} \\
& sF_{ij}(\underline{\theta}_{gi}) \le \xi^\star_{ij} \le sF_{ij}(\overline{\theta}_{gi}) && \forall i \in [I], j \in [J_i] \\
& \frac{\xi^\star_{ij} - sF_{ij}(\underline{\theta}_{gi})}{F_{ij}(\overline{\theta}_{gi}) - F_{ij}(\underline{\theta}_{gi})} \ge \frac{\xi^\star_{i,j+1} - sF_{i,j+1}(\underline{\theta}_{gi})}{F_{i,j+1}(\overline{\theta}_{gi}) - F_{i,j+1}(\underline{\theta}_{gi})} && \forall i \in [I], j \in [J_i] \ \text{with}\ \underline{\theta}_{gi} < z_{ij} < \overline{\theta}_{gi} \\
& d'(\boldsymbol{\xi}^\star, [z^-, z^+]) + (s-1)d'(\boldsymbol{0}, [z^-, z^+]) \le s\overline{d}_g([z^-, z^+]) && \forall [z^-, z^+] \in \mathcal{Z} \Big\},
\end{aligned}
\tag{EC.8}
$$

which is manifestly conic representable. Finally, note that the last condition of (EC.8) resembles the affine cuts (7) since

$$
\begin{aligned}
& d'(\boldsymbol{\xi}^\star, [z^-, z^+]) + (s-1)d'(\boldsymbol{0}, [z^-, z^+]) \\
\overset{\text{(a)}}{=} & \sum_{i \in [I]} \left(\sum_{j \in J_i^+} \xi^\star_{ij} + z_i^- - \underline{\theta}_i - \sum_{j \in J_i^-} \xi^\star_{ij} \right) + (s-1)(z_i^- - \underline{\theta}_i)
\end{aligned}
$$

$$\overset{\text{(b)}}{=} s \sum_{i \in [I]} \left(\sum_{j \in J_i^+} \frac{1}{s} \xi_{ij}^\star + z_i^- - \underline{\theta}_i - \sum_{j \in J_i^-} \frac{1}{s} \xi_{ij}^\star \right)$$
$$\overset{\text{(c)}}{=} s d'(\boldsymbol{\xi}^\star / s, [\boldsymbol{z}^-, \boldsymbol{z}^+]),$$

where (a) and (c) follow from the definition of d' and (b) follows from simplifying the term.

The proof for (i) follows the same line of arguments if we replace $\overline{\Xi}^\star$ with $\overline{\Xi}'$ and drop the constraints in our representation of $s\overline{\Xi}_g^\star$ that pertain to the cuts $d'(\boldsymbol{\xi}', [\boldsymbol{z}^-, \boldsymbol{z}^+]) \le \overline{d}_g([\boldsymbol{z}^-, \boldsymbol{z}^+])$, $[\boldsymbol{z}^-, \boldsymbol{z}^+] \in \mathcal{Z}$. □

EC.5. Proof of Proposition 1

Proof of Proposition 1. We show strong NP-hardness of the separation problem (8) via a reduction from Minimum Vertex Cover, which is known to be strongly NP-hard (Karp 1972). To this end, consider an undirected connected graph $(\mathcal{V}, \mathcal{E})$ with n vertices and m edges, and let $\boldsymbol{W} \in \mathbb{R}^{n \times m}$ be the incidence matrix that satisfies $W_{ij} = 1$ if the i^{th} vertex is incident to the j^{th} edge; $W_{ij} = 0$ otherwise. The optimal value of Minimum Vertex Cover is then given by the optimal value of the binary linear program

$$\min\{\boldsymbol{e}^\top \boldsymbol{y} : \boldsymbol{W}^\top \boldsymbol{y} \ge \boldsymbol{e},\, \boldsymbol{y} \in \{0,1\}^n\}. \tag{EC.9}$$

In the remainder of the proof, we define a valid instance to the separation problem (8) whose optimal value coincides with that of problem (EC.9), modulo some affine transformation.

Consider the instance of problem $\mathcal{P}(\mathbb{P}, \Xi)$ with $G = 1$ support set of the form $\Xi = \{\boldsymbol{\xi} \in \mathbb{R}_+^N : \overline{\boldsymbol{W}}\boldsymbol{\xi} \le 2\boldsymbol{e}\}$, $N = m + n$ and $\overline{\boldsymbol{W}} = (\boldsymbol{W}, \mathbf{1})$, as well as the embedding $\boldsymbol{E} = \mathbf{1}$ satisfying $I = N$ and $\Theta = \Xi$. We omit the other problem data as it is irrelevant for the separation problem (8). We choose the bounding box $[\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}] = [-2\boldsymbol{e}, 2\boldsymbol{e}]$, which is not tight but allows for a more intuitive proof; the extension to a tight support is tedious but straightforward. Our folding operator F employs a single breakpoint at 0 in every dimension, that is, we have $J_i = 2$ and $z_{i0} = -2 < z_{i1} = 0 < z_{i2} = 2$ for all $i \in [I]$. We wish to separate the lifted vector $\boldsymbol{\xi}' \in \mathbb{R}^{2I}$ with components $\xi_{ij}' = 1$ for all $i \in [I]$, $j \in [J_i]$ and the scaling factor $s = 1$. Throughout the proof, we omit the indices g and k for better readability.

To confirm that our choice of Ξ, Θ, $\mathcal{Z}$, $\boldsymbol{\xi}'$, and s induces a valid instance of the separation problem (8), we show that $\boldsymbol{\xi}' \in \overline{\Xi}'$. By the definition of $\overline{\Xi}'$ in (3) and the representation of $\operatorname{conv}(F([\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g]))$ in Lemma 1, this holds whenever the following three conditions are satisfied:

1. $F^+(\boldsymbol{\xi}') \in \Theta$: This holds since $F^+(\boldsymbol{\xi}') = \mathbf{0} \in \Theta$ by definition of F^+, $\boldsymbol{\xi}'$ and Θ.
2. $F_{ij}(\underline{\theta}_i) \le \xi_{ij}' \le F_{ij}(\overline{\theta}_i)$ for all $i \in [I]$ and $j \in [J_i]$: This holds since $F_{ij}(\underline{\theta}_i) = 0$ and $F_{ij}(\overline{\theta}_i) = 2$ by construction of Z, as well as $\xi_{ij}' = 1$, for all $i \in [I]$ and $j \in [J_i]$.
3. $\frac{\xi_{ij}' - F_{ij}(\underline{\theta}_i)}{F_{ij}(\overline{\theta}_i) - F_{ij}(\underline{\theta}_i)} \ge \frac{\xi_{i,j+1}' - F_{i,j+1}(\underline{\theta}_i)}{F_{i,j+1}(\overline{\theta}_i) - F_{i,j+1}(\underline{\theta}_i)}$ for all $i \in [I]$ and $j \in [J_i]$ with $\underline{\theta}_i < z_{ij} < \overline{\theta}_i$: This holds since $F_{ij}(\underline{\theta}_i) = 0$ and $F_{ij}(\overline{\theta}_i) = 2$ by construction of Z, as well as $\xi_{ij}' = 1$, for all $i \in [I]$ and $j \in [J_i]$.

We now prove the main claim by transforming our instance of the separation problem to the MINIMUM VERTEX COVER instance through a series of equivalent problems. To this end, we first reformulate the distances d' and $\overline{d}$ in the separation problem. For d', we find that

$$d'(\boldsymbol{\xi}'/s, [\boldsymbol{z}^-, \boldsymbol{z}^+]) \overset{\text{(a)}}{=} \sum_{i \in [I]} \left(\sum_{j \in J_i^+} \xi'_{ij} + z_i^- - \underline{\theta}_i - \sum_{j \in J_i^-} \xi'_{ij} \right)$$
$$= \sum_{i \in [I]} \left(\sum_{j \in J_i^+} 1 + z_i^- + 2 - \sum_{j \in J_i^-} 1 \right) \overset{\text{(b)}}{=} \sum_{i \in [I]} \left(\frac{z_i^- - z_i^+}{2} + 2 \right),$$

where (a) immediately follows from the definition of d' and $s = 1$, while (b) is due to $|J_i^+| = 1 - \frac{z_i^+}{2}$ and $|J_i^-| = \frac{z_i^-}{2} + 1$. For $\overline{d}$, we find that

$$\overline{d}([\boldsymbol{z}^-, \boldsymbol{z}^+]) = \max_{\boldsymbol{\theta} \in \Theta} \min_{\boldsymbol{\theta}' \in [\boldsymbol{z}^-, \boldsymbol{z}^+]} \|\boldsymbol{\theta} - \boldsymbol{\theta}'\|_1$$
$$\overset{\text{(a)}}{=} \max_{\boldsymbol{\theta} \in \Theta} \sum_{i \in [I]} \min_{\theta' \in [z_i^-, z_i^+]} |\theta_i - \theta'| \overset{\text{(b)}}{=} \max_{\boldsymbol{\theta} \in \Theta} \sum_{i \in [I]} \begin{cases} z_i^- - \theta_i & \text{if } z_i^- = z_i^+ = 2 \\ \theta_i - z_i^+ & \text{if } z_i^+ \leq 0 \\ 0 & \text{otherwise,} \end{cases}$$

where (a) follows from the properties of the 1-norm and (b) is due to a case distinction and $\Theta \subseteq [0, 2]^I$. Combining both reformulations, we equivalently express the separation problem as

$$\max_{[\boldsymbol{z}^-, \boldsymbol{z}^+] \in \mathcal{Z}} d'(\boldsymbol{\xi}'/s, [\boldsymbol{z}^-, \boldsymbol{z}^+]) - \overline{d}([\boldsymbol{z}^-, \boldsymbol{z}^+])$$
$$\overset{\text{(a)}}{=} \max_{[\boldsymbol{z}^-, \boldsymbol{z}^+] \in \mathcal{Z}} \sum_{i \in [I]} \left(\frac{z_i^- - z_i^+}{2} + 2 \right) - \max_{\boldsymbol{\theta} \in \Theta} \sum_{i \in [I]} \begin{cases} 2 - \theta_i & \text{if } z_i^- = z_i^+ = 2 \\ \theta_i - z_i^+ & \text{if } z_i^+ \leq 0 \\ 0 & \text{otherwise} \end{cases}$$
$$\overset{\text{(b)}}{=} \max_{[\boldsymbol{z}^-, \boldsymbol{z}^+] \in \mathcal{Z}} \min_{\boldsymbol{\theta} \in \Theta} \sum_{i \in [I]} \begin{cases} \frac{z_i^- - z_i^+}{2} + \theta_i & \text{if } z_i^- = z_i^+ = 2 \\ \frac{z_i^- - z_i^+}{2} + 2 + z_i^+ - \theta_i & \text{if } z_i^+ \leq 0 \\ \frac{z_i^- - z_i^+}{2} + 2 & \text{otherwise} \end{cases}$$
$$\overset{\text{(c)}}{=} \max_{\boldsymbol{z}^+ \in \{0,2\}^I} \min_{\boldsymbol{\theta} \in \Theta} \sum_{i \in [I]} \begin{cases} -\frac{z_i^+}{2} + 2 + z_i^+ - \theta_i & \text{if } z_i^+ = 0 \\ -\frac{z_i^+}{2} + 2 & \text{if } z_i^+ = 2 \end{cases}$$
$$\overset{\text{(d)}}{=} \max_{\boldsymbol{z}^+ \in \{0,2\}^I} \min_{\boldsymbol{\theta} \in \Theta} \sum_{i \in [I]} \left(-\frac{z_i^+}{2} + 2 - \theta_i \left(1 - \frac{z_i^+}{2} \right) \right)$$
$$\overset{\text{(e)}}{=} \max_{\overline{\boldsymbol{\delta}} \in \{0,1\}^I} \min_{\boldsymbol{\theta} \in \Theta} \sum_{i \in [I]} \left(1 + \overline{\delta}_i - \theta_i \overline{\delta}_i \right)$$
$$\overset{\text{(f)}}{=} \max\{\boldsymbol{e}^\top \overline{\boldsymbol{\delta}} + \min\{-\overline{\boldsymbol{\delta}}^\top \boldsymbol{\theta} : \overline{\boldsymbol{W}} \boldsymbol{\theta} \leq 2\boldsymbol{e}, \boldsymbol{\theta} \geq \boldsymbol{0}\} : \overline{\boldsymbol{\delta}} \in \{0,1\}^I\} + I$$
$$\overset{\text{(g)}}{=} \max\{\boldsymbol{e}^\top \overline{\boldsymbol{\delta}} - \max\{\overline{\boldsymbol{\delta}}^\top \boldsymbol{\theta} : \overline{\boldsymbol{W}} \boldsymbol{\theta} \leq 2\boldsymbol{e}, \boldsymbol{\theta} \geq \boldsymbol{0}\} : \overline{\boldsymbol{\delta}} \in \{0,1\}^I\} + I. \tag{EC.10}$$

Here, (a) follows from the earlier reformulations of d' and $\overline{d}$, (b) merges the two terms, and (c) holds since there always exists an optimal solution $[\boldsymbol{z}^-, \boldsymbol{z}^+]$ satisfying $\boldsymbol{z}^- = \boldsymbol{0}$. Indeed, if $z_i^- = 2$ for some $i \in [I]$, implying that $z_i^+ = 2$ as well (case 1), then setting $z_i^- = 0$ (and thus switching to case 3) weakly increases the objective value. Likewise, if $z_i^- = -2$ for some $i \in [I]$, then the expressions in cases 2 and 3 imply that increasing z_i^-—and, possibly in case 2, z_i^+—to 0 increases the objective value. Equation (d) holds by construction, and (e) substitutes $\overline{\delta}_i = 1 - z_i^+/2$. Equation (f) reformulates the problem and replaces Θ with its definition, and (g) negates the inner optimization problem.

Dualizing the inner maximization in (EC.10) yields

$$\begin{aligned} &\max\{\boldsymbol{e}^\top \overline{\boldsymbol{\delta}} - \min\{2\boldsymbol{e}^\top \boldsymbol{y} : \overline{\boldsymbol{W}}^\top \boldsymbol{y} \geq \overline{\boldsymbol{\delta}},\ \boldsymbol{y} \geq \boldsymbol{0}\},\ \overline{\boldsymbol{\delta}} \in \{0,1\}^I\} + I \\ &\overset{\text{(a)}}{=} -\min\{2\boldsymbol{e}^\top \boldsymbol{y} - \boldsymbol{e}^\top \boldsymbol{\delta} - \boldsymbol{e}^\top \boldsymbol{\gamma} : \boldsymbol{W}^\top \boldsymbol{y} \geq \boldsymbol{\gamma},\ \boldsymbol{y} \geq \boldsymbol{\delta},\ \boldsymbol{y} \geq \boldsymbol{0},\ (\boldsymbol{\gamma}, \boldsymbol{\delta}) \in \{0,1\}^{(m+n)}\} + I. \end{aligned} \tag{EC.11}$$

Strong LP duality and finiteness of the optimal objective values hold since $\boldsymbol{\theta} = \boldsymbol{0}$ is a primal feasible solution and $\boldsymbol{y} = \boldsymbol{e}$ is a dual feasible solution for any $\overline{\boldsymbol{\delta}}$. Equation (a) negates the maximum and replaces $\overline{\boldsymbol{\delta}}$ with $(\boldsymbol{\gamma}, \boldsymbol{\delta})$. Thus, the values of (EC.10) and (EC.11) coincide.

We next claim that there is always an optimal solution $(\boldsymbol{y}, \boldsymbol{\gamma}, \boldsymbol{\delta})$ to problem (EC.11) with $\boldsymbol{\gamma} = \boldsymbol{e}$. Indeed, assume that $\gamma_e = 0$ for some edge $e = \{v, w\} \in \mathcal{E}$, and observe that $2y_v - \delta_v - \gamma_e \geq 0$ since $y_v \geq \delta_v$. By setting $y_v = \delta_v = \gamma_e = 1$, we obtain another feasible solution whose contribution $2y_v - \delta_v - \gamma_e = 0$ to the objective function of (EC.11) is weakly less than before. Fixing $\boldsymbol{\gamma} = \boldsymbol{e}$ simplifies the optimization problem (EC.11) to

$$-\min\{2\boldsymbol{e}^\top \boldsymbol{y} - \boldsymbol{e}^\top \boldsymbol{\delta} : \boldsymbol{W}^\top \boldsymbol{y} \geq \boldsymbol{e},\ \boldsymbol{y} \geq \boldsymbol{\delta},\ \boldsymbol{y} \geq \boldsymbol{0},\ \boldsymbol{\delta} \in \{0,1\}^n\} + I + m. \tag{EC.12}$$

We conclude the proof by arguing that (EC.9) attains the same optimal value as the minimization problem inside (EC.12), that is, that (EC.9) $= -$(EC.12) $+ I + m$. To this end, observe first that any solution $\boldsymbol{y}$ feasible in (EC.9) can be augmented to a feasible solution $(\boldsymbol{y}, \boldsymbol{\delta})$ to (EC.12) that attains the same objective value in the minimization problem inside (EC.12) by setting $\boldsymbol{\delta} = \boldsymbol{y}$. Conversely, we claim that there is always an optimal solution $(\boldsymbol{y}, \boldsymbol{\delta})$ to (EC.12) with $\boldsymbol{y} \in \{0,1\}^n$; the partial solution $\boldsymbol{y}$ then attains the same objective value in (EC.9) as $(\boldsymbol{y}, \boldsymbol{\delta})$ does in the minimization problem inside (EC.12). Assume to the contrary that all optimal solutions $(\boldsymbol{y}, \boldsymbol{\delta})$ to (EC.12) are non-binary, and fix an optimal solution $(\boldsymbol{y}, \boldsymbol{\delta})$ to (EC.12) with the fewest number of non-binary components in $\boldsymbol{y}$. Fix any $v \in \mathcal{V}$ with $y_v \in (0,1)$. Then there must be $\{v, w\} \in \mathcal{E}$ such that $\boldsymbol{W}^\top \boldsymbol{y} \geq \boldsymbol{e}$ imposes the covering condition $y_v + y_w \geq 1$, which implies that either $y_v \geq 1/2$ or $y_w \geq 1/2$ or both. Assume without loss of generality that $y_v \geq 1/2$. In that case, setting $y_v = \delta_v = 1$ results in another feasible solution with fewer non-binary components that is at least as good as the previous one. This leads to a contradiction, and we thus conclude that there is indeed always an optimal solution $(\boldsymbol{y}, \boldsymbol{\delta})$ to (EC.12) with $\boldsymbol{y} \in \{0,1\}^n$. □

EC.6. Proof of Theorem 2

We structure the proof of Theorem 2 into multiple lemmas. We first use the permutation invariance of Assumptions 1 and 2 to identify analytical representations for the distance function $\overline{d}$. We split this simplification into three lemmas. Lemma EC.1 generalizes Lemma 3 of Ben-Tal et al. (2020) to show that for each $i \in [I]$, Θ contains realizations with one-norm $\eta(i)$ that are constant on their first i components and zero on their remaining components. Lemma EC.2 shows that there always exists an optimal solution to the separation problem that is symmetric around the origin. Finally, Lemma EC.3 combines these two results to derive an analytical representation for $\overline{d}$. Next, we use this representation to identify a small subset of solutions guaranteed to contain an optimal solution for the separation problem in Lemma EC.4. We conclude the proof of part (i) of Theorem 2 by describing how to efficiently enumerate this subset in Algorithm 1 and prove its validity in Lemma EC.5. For part (ii) of Theorem 2, Lemma EC.6 shows that the additional property in the statement of the theorem further reduces the set of possible solutions to the set of squares. We omit the index $k \in [K]$ of the separation problem considered and use $\boldsymbol{\xi}^\star = \boldsymbol{\xi}'/s$ for increased readability throughout the proof. Further, we assume that $[\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]$ is a tight bounding box for Θ, that is $\underline{\theta}_i \boldsymbol{e}_i, \overline{\theta}_i \boldsymbol{e}_i \in \Theta$ for all $i \in [I]$. The extension to general bounding boxes $[\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]$ is straightforward, but it requires additional tedious case distinctions leading to a less intuitive proof.

LEMMA EC.1. *Let Θ satisfy property (i) of Assumption 2. Then for each $i \in [I]$ there exists a $\lambda(i)$ such that $\sum_{i' \in [i]} \lambda(i) \boldsymbol{e}_{i'} \in \Theta$, and $i\lambda(i) = \eta(i)$.*

Proof of Lemma EC.1. Fix any $\boldsymbol{\theta} \in \arg\max_{\boldsymbol{\theta} \in \Theta} \sum_{i' \in [i]} |\theta_{i'}|$. Let $\hat{\boldsymbol{\theta}}$ be the vector with absolute values of $\boldsymbol{\theta}$ in each component, that is, $\hat{\theta}_{i'} = |\theta_{i'}|$. By axis symmetry of Θ, we have $\hat{\boldsymbol{\theta}} \in \Theta$. Thus, we assume w.l.o.g. that $\boldsymbol{\theta} \geq 0$. The remainder of the proof follows from Lemma 3 of Ben-Tal et al. (2020), which considers the special case of Lemma EC.1 where Θ is non-negative. □

LEMMA EC.2. *Let Z satisfy Assumption 1 and let Θ satisfy property (i) of Assumption 2. Then there exists an optimal solution $[\boldsymbol{z}^-, \boldsymbol{z}^+]$ for the separation problem that satisfies $\boldsymbol{z}^- = -\boldsymbol{z}^+$.*

Proof of Lemma EC.2. We split the proof into two parts. First, we show that there always exists an optimal solution $[\boldsymbol{z}^-, \boldsymbol{z}^+]$ to the separation problem that satisfies $\boldsymbol{z}^- \leq \boldsymbol{0} \leq \boldsymbol{z}^+$. We then use this property to prove the actual claim. We prove both parts using contradiction arguments.

Let $[\boldsymbol{z}^-, \boldsymbol{z}^+]$ be an optimal solution to the separation problem with a minimal number of indices $i \in [I]$ such that $z_i^+ < 0$. Assume there was such an index $i^\star$. Define $\hat{\boldsymbol{z}}^+$ via

$$\hat{z}_i^+ = \begin{cases} z_i^+ & \text{if } i \neq i^\star \\ 0 & \text{if } i = i^\star. \end{cases}$$

Then $[\boldsymbol{z}^-, \hat{\boldsymbol{z}}^+] \in \mathcal{Z}$, as $z_i^- \leq z_i^+ \leq \hat{z}_i^+$ for all $i \in [I]$. Fix any $\boldsymbol{\theta} \in \arg\max_{\boldsymbol{\theta} \in \Theta} d(\boldsymbol{\theta}, [\boldsymbol{z}^-, \hat{\boldsymbol{z}}^+])$ maximizing the distance from the new rectangle. By $z_{i^\star}^- \leq z_{i^\star}^+ < \hat{z}_{i^\star}^+ = 0$ we have $|\theta_{i^\star}| - \hat{z}_{i^\star}^+ > |\theta_{i^\star}| + z_{i^\star}^-$, which by maximality of $\boldsymbol{\theta}$ and symmetry of Θ implies $\theta_{i^\star} \geq 0$. With this we find

$$\overline{d}([\boldsymbol{z}^-, \hat{\boldsymbol{z}}^+]) \overset{\text{(a)}}{=} \sum_{i \in [I]} \left(\theta_i - \hat{z}_i^+\right)_+ + \left(z_i^- - \theta_i\right)_+$$
$$\overset{\text{(b)}}{=} \sum_{i \in [I]} \left(\theta_i - z_i^+\right)_+ + \left(z_i^- - \theta_i\right)_+ + z_{i^\star}^+ \overset{\text{(c)}}{\leq} \overline{d}([\boldsymbol{z}^-, \boldsymbol{z}^+]) + z_{i^\star}^+,$$

Here, (a) follows from the choice of $\boldsymbol{\theta}$, (b) is due to $\theta_{i^\star} \geq 0$, and (c) holds since $\boldsymbol{\theta}$ is a feasible, though not necessarily optimal, solution for the maximization problem represented by $\overline{d}([\boldsymbol{z}^-, \boldsymbol{z}^+])$. Similarly we find

$$d'(\boldsymbol{\xi}^\star, [\boldsymbol{z}^-, \hat{\boldsymbol{z}}^+]) \overset{\text{(a)}}{=} \sum_{i \in [I]} \sum_{j \in \hat{J}_i^+} \xi_{ij}^\star + z_i^- - \underline{\theta}_i - \sum_{j \in J_i^-} \xi_{ij}^\star$$
$$\overset{\text{(b)}}{=} \sum_{i \in [I]} \sum_{j \in J_i^+} \xi_{ij}^\star + z_i^- - \underline{\theta}_i - \sum_{j \in J_i^-} \xi_{ij}^\star - \sum_{j \in J_{i^\star}^+ \setminus \hat{J}_{i^\star}^+} \xi_{i^\star j}^\star$$
$$\overset{\text{(c)}}{=} d'(\boldsymbol{\xi}^\star, [\boldsymbol{z}^-, \boldsymbol{z}^+]) - \sum_{j \in J_{i^\star}^+ \setminus \hat{J}_{i^\star}^+} \xi_{i^\star j}^\star$$
$$\overset{\text{(d)}}{\geq} d'(\boldsymbol{\xi}^\star, [\boldsymbol{z}^-, \boldsymbol{z}^+]) + z_{i^\star}^+,$$

where we use $\hat{J}_i^+ = \{j \in [J_i] \colon \hat{z}_i^+ < z_{ij}\}$ analogously to J_i^+. Here, (a) and (c) are due to the definition of d', (b) follows from $z_{i^\star}^+ < \hat{z}_{i^\star}^+$ and splitting the sum, and (d) follows from $\xi_{ij}^\star \leq F_{ij}(\overline{\theta}_i) \leq z_j - z_{j-1}$ by Lemma 1, contracting the sum, and $\hat{z}_{i^\star}^+ = 0$. Combining the newly found inequalities for $\overline{d}$ and d' yields $d'(\boldsymbol{\xi}^\star, [\boldsymbol{z}^-, \hat{\boldsymbol{z}}^+]) - \overline{d}([\boldsymbol{z}^-, \hat{\boldsymbol{z}}^+]) \geq d'(\boldsymbol{\xi}^\star, [\boldsymbol{z}^-, \boldsymbol{z}^+]) - \overline{d}([\boldsymbol{z}^-, \boldsymbol{z}^+])$. This contradicts the assumption that $[\boldsymbol{z}^-, \boldsymbol{z}^+]$ had a minimal number of indices $i \in [I]$ such that $z_i^+ < 0$. The case for indices $i \in [I]$ with $z_i^- > 0$ follows analogously, and we thus conclude that the separation problem always has an optimal solution $[\boldsymbol{z}^-, \boldsymbol{z}^+]$ with $\boldsymbol{z}^- \leq \mathbf{0} \leq \boldsymbol{z}^+$.

We now use this result to prove the claim of the lemma. Let $[\boldsymbol{z}^-, \boldsymbol{z}^+]$ be an optimal solution to the separation problem that satisfies $\boldsymbol{z}^- \leq \mathbf{0} \leq \boldsymbol{z}^+$ with a minimal number of indices $i \in [I]$ such that $z_i^+ > -z_i^-$. Assume there was such an index $i^\star$. Define $\hat{\boldsymbol{z}}^+$ via

$$\hat{z}_i^+ = \begin{cases} z_i^+ & \text{if } i \neq i^\star \\ -z_i^- & \text{if } i = i^\star. \end{cases}$$

By $z_{i^\star}^- \leq 0 \leq -z_{i^\star}^- = \hat{z}_{i^\star}^+$, we have $[\boldsymbol{z}^-, \hat{\boldsymbol{z}}^+] \in \mathcal{Z}$. Fix any $\boldsymbol{\theta} \in \arg\max_{\boldsymbol{\theta} \in \Theta} d(\boldsymbol{\theta}, [\boldsymbol{z}^-, \hat{\boldsymbol{z}}^+])$ maximizing the distance from the new rectangle. Then by $\hat{z}_{i^\star}^+ = -z_{i^\star}^-$ we have $|\theta_{i^\star}| - \hat{z}_{i^\star}^+ = |\theta_{i^\star}| + z_{i^\star}^-$. By the symmetry of Θ we can assume w.l.o.g. that $\theta_{i^\star} \leq 0$. We thus have

$$\left(\theta_{i^\star} - \hat{z}_{i^\star}^+\right)_+ + \left(z_{i^\star}^- - \theta_{i^\star}\right)_+ = \left(z_{i^\star}^- - \theta_{i^\star}\right)_+ = \left(\theta_{i^\star} - z_{i^\star}^+\right)_+ + \left(z_{i^\star}^- - \theta_{i^\star}\right)_+.$$

With this we find that

$$\begin{aligned}\overline{d}([z^-,\hat{z}^+]) &= \sum_{i\in[I]} \left(\theta_i - \hat{z}_i^+\right)_+ + \left(z_i^- - \theta_i\right)_+ \\ &= \sum_{i\in[I]} \left(\theta_i - z_i^+\right)_+ + \left(z_i^- - \theta_i\right)_+ \le \overline{d}([z^-,z^+]).\end{aligned}$$

Using a similar argument as above for d', we additionally find that

$$\begin{aligned}d'(\boldsymbol{\xi}^\star,[z^-,\hat{z}^+]) &\overset{\text{(a)}}{=} \sum_{i\in[I]} \sum_{j\in\hat{J}_i^+} \xi_{ij}^\star + z_i^- - \underline{\theta}_i - \sum_{j\in J_i^-} \xi_{ij}^\star \\ &\overset{\text{(b)}}{=} \sum_{i\in[I]} \sum_{j\in J_i^+} \xi_{ij}^\star + z_i^- - \underline{\theta}_i - \sum_{j\in J_i^-} \xi_{ij}^\star + \sum_{j\in \hat{J}_{i^\star}^+\setminus J_{i^\star}^+} \xi_{i^\star j}^\star \\ &\overset{\text{(c)}}{=} d'(\boldsymbol{\xi}^\star,[z^-,z^+]) + \sum_{j\in \hat{J}_{i^\star}^+\setminus J_{i^\star}^+} \xi_{i^\star j}^\star \\ &\overset{\text{(d)}}{\ge} d'(\boldsymbol{\xi}^\star,[z^-,z^+]).\end{aligned}$$

Here, (a) and (c) are due to the definition of d', (b) follows from $\hat{z}_{i^\star}^+ \le z_{i^\star}^+$ and splitting the sum, and (d) holds since $\boldsymbol{\xi}^\star \ge \mathbf{0}$. Combining the inequalities for $\overline{d}$ and d' yields $d'(\boldsymbol{\xi}^\star,[z^-,\hat{z}^+]) - \overline{d}([z^-,\hat{z}^+]) \ge d'(\boldsymbol{\xi}^\star,[z^-,z^+]) - \overline{d}([z^-,z^+])$. This contradicts the assumption that $[z^-,z^+]$ had a minimal number of violated indices. The case for indices i with $z_i^+ < -z_i^-$ follows analogously. We thus conclude that the separation problem always has an optimal solution $[z^-,z^+]$ with $z^- = -z^+$, which completes the proof. □

LEMMA EC.3. *Let Z satisfy Assumption 1 and let Θ satisfy property (i) of Assumption 2. Let $[z^-,z^+] \in \mathcal{Z}$ with $z^- = -z^+$ and z^+ being non-descending, that is, $\forall i \in [I-1]$: $z_i^+ \le z_{i+1}^+$. Then $\overline{d}([z^-,z^+]) = \max_{i\in[I]}\{\eta(i) - \sum_{i'\in[i]} z_{i'}^+\}$.*

Proof of Lemma EC.3. We prove the claim by showing each of the inequalities "≤" and "≥" separately. For the "≥" direction, recall that by Lemma EC.1, there is some $\lambda(i)$ such that $\sum_{i'\in[i]} \lambda(i)\boldsymbol{e}_{i'} \in \Theta$, and $i\lambda(i) = \eta(i)$ for each $i \in [I]$. Then

$$\overline{d}([z^-,z^+]) \overset{\text{(a)}}{\ge} \max_{i\in[I]} \sum_{i'\in[i]} (\lambda(i) - z_{i'}^+)_+ \overset{\text{(b)}}{\ge} \max_{i\in[I]} \sum_{i'\in[i]} (\lambda(i) - z_{i'}^+) \overset{\text{(c)}}{=} \max_{i\in[I]}(\eta(i) - \sum_{i'\in[i]} z_{i'}^+),$$

where (a) follows from $\sum_{i'\in[i]} \lambda(i)\boldsymbol{e}_{i'} \in \Theta$ and the definition of $\overline{d}$, (b) follows from $z^+ \ge \mathbf{0}$ and relaxing the $(\cdot)_+$ operator, and (c) follows from $i\lambda(i) = \eta(i)$. This concludes the "≥" direction.

For the "≤" direction, fix any $\boldsymbol{\theta} \in \arg\max_{\boldsymbol{\theta}\in\Theta} d(\boldsymbol{\theta},[z^-,z^+])$. By the symmetry of Θ and $z^+ = -z^-$, we assume w.l.o.g. that $\boldsymbol{\theta} \ge \mathbf{0}$. First, we show that there always exists a maximizing $\boldsymbol{\theta}$ with non-ascending ordering, that is, $\theta_i \ge \theta_{i+1}$ for all $i \in [I-1]$. To see this, assume there were two indices $i_1 \le i_2$ with $\theta_{i_1} \le \theta_{i_2}$. We show $\left(\theta_{i_1} - z_{i_1}^+\right)_+ + \left(\theta_{i_2} - z_{i_2}^+\right)_+ \le \left(\theta_{i_2} - z_{i_1}^+\right)_+ + \left(\theta_{i_1} - z_{i_2}^+\right)_+$, which implies that swapping the values of θ_{i_1} and θ_{i_2} does not decrease $d(\boldsymbol{\theta},[z^-,z^+])$.

For $\theta_{i_1} \geq z^+_{i_1}$ the claim follows from

$$\left(\theta_{i_1} - z^+_{i_1}\right)_+ + \left(\theta_{i_2} - z^+_{i_2}\right)_+ \overset{\text{(a)}}{=} \theta_{i_1} - z^+_{i_1} + \left(\theta_{i_2} - \theta_{i_1} + \theta_{i_1} - z^+_{i_2}\right)_+$$
$$\overset{\text{(b)}}{\leq} \theta_{i_1} + \theta_{i_2} - \theta_{i_1} - z^+_{i_1} + \left(\theta_{i_1} - z^+_{i_2}\right)_+ \overset{\text{(c)}}{=} \left(\theta_{i_2} - z^+_{i_1}\right)_+ + \left(\theta_{i_1} - z^+_{i_2}\right)_+,$$

where (a) follows from $\theta_{i_1} \geq z^+_{i_1}$, (b) is due to $\theta_{i_1} \leq \theta_{i_2}$ and the piecewise linearity of $(\cdot)_+$, and (c) holds since $z^+_{i_1} \leq \theta_{i_1} \leq \theta_{i_2}$.

For $\theta_{i_1} \leq z^+_{i_1}$ the claim follows from

$$\left(\theta_{i_1} - z^+_{i_1}\right)_+ + \left(\theta_{i_2} - z^+_{i_2}\right)_+ \overset{\text{(a)}}{=} \left(\theta_{i_2} - z^+_{i_2}\right)_+ \overset{\text{(b)}}{\leq} \left(\theta_{i_2} - z^+_{i_1}\right)_+ \overset{\text{(c)}}{=} \left(\theta_{i_2} - z^+_{i_1}\right)_+ + \left(\theta_{i_1} - z^+_{i_2}\right)_+,$$

where (a) follows from $\theta_{i_1} \leq z^+_{i_1}$, (b) holds since $z^+_{i_1} \leq z^+_{i_2}$, and (c) is due to $\theta_{i_1} \leq z^+_{i_1} \leq z^+_{i_2}$. Thus, by the permutation invariance of Θ, we can w.l.o.g assume that $\boldsymbol{\theta}$ is in non-ascending order.

Let $i^\star \in [I]$ be the last index with $\theta_{i^\star} \geq z^+_{i^\star}$. Such an index always exists by $\overline{\theta}_1 \geq z^+_1$ and $\overline{\theta}_1 \boldsymbol{e}_1 \in \Theta$. By the ordering of $\boldsymbol{\theta}$ and $\boldsymbol{z}^+$, we have $\theta_i \geq z^+_i$ for all $i \leq i^\star$. We conclude the proof by showing that $\overline{d}([\boldsymbol{z}^-, \boldsymbol{z}^+]) \leq \eta(i^\star) - \sum_{i \in [i^\star]} z^+_i$. In particular, we find

$$\overline{d}([\boldsymbol{z}^-, \boldsymbol{z}^+]) \overset{\text{(a)}}{=} \sum_{i \in I} \left(\theta_i - z^+_i\right)_+ + \left(z^-_i - \theta_i\right)_+$$
$$\overset{\text{(b)}}{=} \sum_{i \in I} \left(\theta_i - z^+_i\right)_+ \overset{\text{(c)}}{=} \sum_{i \in [i^\star]} \theta_i - z^+_i \overset{\text{(d)}}{\leq} \eta(i^\star) - \sum_{i \in [i^\star]} z^+_i,$$

where (a) follows from the optimality of $\boldsymbol{\theta}$, (b) holds since $\boldsymbol{\theta} \geq \boldsymbol{0}$, (c) is due to the choice of $i^\star$, and (d) follows from the definition of η and $\boldsymbol{\theta} \in \Theta$. □

LEMMA EC.4. *Let Z satisfy Assumption 1 and let Θ satisfy Assumption 2. Then there always exists an index pair $(i, j) \in [I] \times [J]$, $j \geq \frac{J}{2}$, such that there is an optimal solution $[\boldsymbol{z}^-, \boldsymbol{z}^+]$ for the separation problem that satisfies $z^+_{\pi(i')} = \begin{cases} z_j & \text{if } i' \leq i \\ z_{j+1} & \text{if } i' > i \end{cases}$ and $\boldsymbol{z}^- = -\boldsymbol{z}^+$, where π is an non-ascending ordering of $\{\delta_{i'j}\}_{i' \in [I]}$ with $\delta_{i'j} = z_{j+1} - z_j + \xi^\star_{i',j+1} - \xi^\star_{i',J-j}$.*

Proof of Lemma EC.4. The proof proceeds in two parts. We first prove the characteristics of Lemma EC.4 for a potentially arbitrary ordering of $\{\delta_{ij}\}_{i \in [I]}$. Then, we use the resulting solution to construct another solution with the same objective value, following a non-ascending ordering of $\{\delta_{ij}\}_{i \in [I]}$.

For the first part of the proof, with a potentially arbitrary ordering of $\{\delta_{ij}\}_{i \in [I]}$, let $[\boldsymbol{z}^-, \boldsymbol{z}^+]$ with $\boldsymbol{z}^+ = -\boldsymbol{z}^-$ be an optimal solution to the separation problem. By Lemma EC.2, such a solution always exists. Let σ be an ordering of the indices such that $z^+_{\sigma(i)} \leq z^+_{\sigma(i+1)}$ for all $i \in [I-1]$. Then,

$$\overline{d}([\boldsymbol{z}^-, \boldsymbol{z}^+]) \overset{\text{(a)}}{=} \overline{d}([\sigma(\boldsymbol{z}^-), \sigma(\boldsymbol{z}^+)]) \overset{\text{(b)}}{=} \max_{i \in [I]} \eta(i) - \sum_{i' \in [i]} z^+_{\sigma(i')}.$$

Here (a) follows from permutational invariance of Θ and (b) is due to Lemma EC.3. Let $i^\star \in [I]$ be the maximal index such that $\overline{d}([z^-, z^+]) = \eta(i^\star) - \sum_{i' \in [i^\star]} z^+_{\sigma(i')}$ and let $j^\star \in [J]$ be the maximal index such that

$$z_{j^\star} \leq \eta(i^\star) - \eta(i^\star - 1) \tag{EC.13}$$

$$\text{and} \quad z_{j^\star} \leq z^+_{\sigma(i^\star+1)}. \tag{EC.14}$$

With this we construct $\hat{z}^+$ via

$$\hat{z}^+_{\sigma(i)} = \begin{cases} z_{j^\star} & \text{if } i \leq i^\star \text{ or } z_{j^\star} > \eta(i^\star+1) - \eta(i^\star) \\ z_{j^\star+1} & \text{if } i > i^\star \text{ and } z_{j^\star} \leq \eta(i^\star+1) - \eta(i^\star), \end{cases}$$

we define $\hat{z}^- = -\hat{z}^+$, and we claim that

$$d'(\boldsymbol{\xi}^\star, [\hat{z}^-, \hat{z}^+]) - \overline{d}([\hat{z}^-, \hat{z}^+]) \geq d'(\boldsymbol{\xi}^\star, [z^-, z^+]) - \overline{d}([z^-, z^+]). \tag{EC.15}$$

To prove (EC.15), we first simplify $\overline{d}([\hat{z}^-, \hat{z}^+])$ and show that

$$\overline{d}([\hat{z}^-, \hat{z}^+]) = \eta(i^\star) - \sum_{i \in [i^\star]} \hat{z}^+_{\sigma(i)} = \eta(i^\star) - i^\star z_{j^\star}. \tag{EC.16}$$

By Lemma EC.3, there exists an index i' such that $\overline{d}([\hat{z}^-, \hat{z}^+]) = \eta(i') - \sum_{i \in [i']} \hat{z}^+_{\sigma(i)}$. We claim that $i' = i^\star$, which immediately implies (EC.16).

First, consider the case where $i' > i^\star$. Then

$$\begin{aligned} \eta(i') - \sum_{i \in [i']} \hat{z}^+_{\sigma(i)} &\overset{\text{(a)}}{=} \eta(i^\star) - \sum_{i \in [i^\star]} \hat{z}^+_{\sigma(i)} + \eta(i') - \eta(i^\star) - \sum_{i \in [i'] \setminus [i^\star]} \hat{z}^+_{\sigma(i)} \\ &\overset{\text{(b)}}{\leq} \eta(i^\star) - \sum_{i \in [i^\star]} \hat{z}^+_{\sigma(i)} + (i' - i^\star)(\eta(i^\star+1) - \eta(i^\star) - \hat{z}^+_{\sigma(i^\star+1)}) \overset{\text{(c)}}{<} \eta(i^\star) - \sum_{i \in [i^\star]} \hat{z}^+_{\sigma(i)}, \end{aligned}$$

where (a) follows from adding a smart zero and splitting the sum, (b) is due to the non-increasing differences of η and the definition of $\hat{z}^+$, and (c) holds since $\hat{z}^+_{\sigma(i^\star+1)} > \eta(i^\star+1) - \eta(i^\star)$, which follows from a case distinction in the definition of $\hat{z}^+$. By this definition, the claim is immediate for the case $\hat{z}^+_{\sigma(i^\star+1)} = z_{j^\star}$. For the case $\hat{z}^+_{\sigma(i^\star+1)} = z_{j^\star+1}$, we consider another case distinction over (EC.13) and (EC.14). If (EC.13) is binding, then the claim follows from the non-increasing differences of η. If (EC.14) is binding, then the claim follows from $z_{j^\star+1} > z^+_{\sigma(i^\star+1)}$ and $z^+_{\sigma(i^\star+1)} > \eta(i^\star+1) - \eta(i^\star)$, which holds since $\eta(i^\star+1) - \sum_{i \in [i^\star+1]} z^+_{\sigma(i)} < \eta(i^\star) - \sum_{i \in [i^\star]} z^+_{\sigma(i)}$ by the optimality and maximality of the index $i^\star$.

Next, consider the case where $i' < i^\star$. Then

$$\begin{aligned} \eta(i') - \sum_{i \in [i']} \hat{z}^+_{\sigma(i)} &\overset{\text{(a)}}{=} \eta(i^\star) - \sum_{i \in [i^\star]} \hat{z}^+_{\sigma(i)} + \eta(i') - \eta(i^\star) + \sum_{i \in [i^\star] \setminus [i']} \hat{z}^+_{\sigma(i)} \\ &\overset{\text{(b)}}{\leq} \eta(i^\star) - \sum_{i \in [i^\star]} \hat{z}^+_{\sigma(i)} + (i^\star - i')(z_{j^\star} - (\eta(i^\star) - \eta(i^\star - 1))) \overset{\text{(c)}}{\leq} \eta(i^\star) - \sum_{i \in [i^\star]} \hat{z}^+_{\sigma(i)}, \end{aligned}$$

where (a) follows from adding a smart zero and splitting the sum, (b) is due to the non-increasing differences of η and the definition of $\hat{z}^+$, and (c) follows from (EC.13).

To conclude our proof of (EC.15), we show the following relational property between $\hat{z}^+$ and z^+:

$$z^+_{\sigma(i)} \le \hat{z}^+_{\sigma(i)} \qquad \text{if } i \le i^\star \tag{EC.17}$$

$$\text{and} \quad z^+_{\sigma(i)} \ge \hat{z}^+_{\sigma(i)} \qquad \text{if } i \ge i^\star + 1. \tag{EC.18}$$

For (EC.17) it is sufficient to show that $z^+_{\sigma(i^\star)} \le z_{j^\star}$, as $z^+_{\sigma(i)} \le z^+_{\sigma(i^\star)}$ for all $i \le i^\star$ by definition of σ and $\hat{z}^+_{\sigma(i)} = z_{j^\star}$ for all $i \le i^\star$ by definition of $\hat{z}^+$, respectively. Consider the case where (EC.13) is the binding constraint for $j^\star$. Then by the maximality of $i^\star$, we find $\eta(i^\star) - \sum_{i\in[i^\star]} z^+_{\sigma(i)} \ge \eta(i^\star - 1) - \sum_{i\in[i^\star-1]} z^+_{\sigma(i)}$, which implies that $z^+_{\sigma(i^\star)} \le \eta(i^\star) - \eta(i^\star - 1)$. This implies (EC.17) as $j^\star$ is the maximal index such that $z_{j^\star} \le \eta(i^\star) - \eta(i^\star - 1)$. Now consider the case where (EC.14) is the binding constraint for $j^\star$. Then $z^+_{\sigma(i^\star)} \le z^+_{\sigma(i^\star+1)} = z_{j^\star}$ by the definition of σ.

For (EC.18) it is sufficient to show $z^+_{\sigma(i^\star+1)} \ge \hat{z}^+_{\sigma(i^\star+1)}$, as $z^+_{\sigma(i)} \ge z^+_{\sigma(i^\star+1)}$ for all $i \ge i^\star + 1$ by definition of σ and $\hat{z}^+_{\sigma(i)} = \hat{z}^+_{\sigma(i^\star+1)}$ for all $i \ge i^\star + 1$ by definition of $\hat{z}^+$, respectively. Consider the case $z_{j^\star} > \eta(i^\star + 1) - \eta(i^\star)$. Then $\hat{z}^+_{\sigma(i^\star+1)} = z_{j^\star} \le z^+_{\sigma(i^\star+1)}$ by (EC.14). Now consider the case $z_{j^\star} \le \eta(i^\star + 1) - \eta(i^\star)$. Then $\hat{z}^+_{\sigma(i^\star+1)} = z_{j^\star+1}$. Further, we have $z^+_{\sigma(i^\star+1)} > \eta(i^\star + 1) - \eta(i^\star) \ge z_{j^\star}$ by optimality and maximality of $i^\star$. Thus $z^+_{\sigma(i^\star+1)} \ge z_{j^\star+1} = \hat{z}^+_{\sigma(i^\star+1)}$.

Next, we find that

$$\begin{aligned}
& d'(\boldsymbol{\xi}^\star, [\hat{z}^-, \hat{z}^+]) - \overline{d}([\hat{z}^-, \hat{z}^+]) \\
\overset{\text{(a)}}{=} & \sum_{i\in[I]} \left(\sum_{j\in\hat{J}_i^+} \xi^\star_{ij} + \hat{z}^-_i - \underline{\theta}_i - \sum_{j\in\hat{J}_i^-} \xi^\star_{ij} \right) - \eta(i^\star) + \sum_{i\in[i^\star]} \hat{z}^+_{\sigma(i)} \\
\overset{\text{(b)}}{=} & \sum_{i\in[I]} \left(\sum_{j\in J_i^+} \xi^\star_{ij} + z^-_i - \underline{\theta}_i - \sum_{j\in J_i^-} \xi^\star_{ij} \right) - \eta(i^\star) + \sum_{i\in[i^\star]} \hat{z}^+_{\sigma(i)} \\
& + \sum_{i\in[I]\setminus[i^\star]} \left(\sum_{j\in\hat{J}^+_{\sigma(i)}\setminus J^+_{\sigma(i)}} \xi^\star_{\sigma(i)j} + \hat{z}^-_{\sigma(i)} - z^-_{\sigma(i)} - \sum_{j\in\hat{J}^-_{\sigma(i)}\setminus J^-_{\sigma(i)}} \xi^\star_{\sigma(i)j} \right) \\
& - \sum_{i\in[i^\star]} \left(\sum_{j\in J^+_{\sigma(i)}\setminus\hat{J}^+_{\sigma(i)}} \xi^\star_{\sigma(i)j} - \hat{z}^-_{\sigma(i)} + z^-_{\sigma(i)} - \sum_{j\in J^-_{\sigma(i)}\setminus\hat{J}^-_{\sigma(i)}} \xi^\star_{\sigma(i)j} \right) \\
\overset{\text{(c)}}{=} & \sum_{i\in[I]} \left(\sum_{j\in J_i^+} \xi^\star_{ij} + z^-_i - \underline{\theta}_i - \sum_{j\in J_i^-} \xi^\star_{ij} \right) - \eta(i^\star) + \sum_{i\in[i^\star]} z^+_{\sigma(i)} \\
& + \sum_{i\in[I]\setminus[i^\star]} \left(\sum_{j\in\hat{J}^+_{\sigma(i)}\setminus J^+_{\sigma(i)}} \xi^\star_{\sigma(i)j} + \hat{z}^-_{\sigma(i)} - z^-_{\sigma(i)} - \sum_{j\in\hat{J}^-_{\sigma(i)}\setminus J^-_{\sigma(i)}} \xi^\star_{\sigma(i)j} \right)
\end{aligned}$$

$$
\begin{aligned}
&- \sum_{i \in [i^\star]} \left(\sum_{j \in J^+_{\sigma(i)} \setminus \hat{J}^+_{\sigma(i)}} \xi^\star_{\sigma(i)j} - \sum_{j \in J^-_{\sigma(i)} \setminus \hat{J}^-_{\sigma(i)}} \xi^\star_{\sigma(i)j} \right) \\
&\overset{(d)}{\geq} \sum_{i \in [I]} \left(\sum_{j \in J^+_i} \xi^\star_{ij} + z^-_i - \underline{\theta}_i - \sum_{j \in J^-_i} \xi^\star_{ij} \right) - \eta(i^\star) + \sum_{i \in [i^\star]} z^+_{\sigma(i)} \\
&= d'(\boldsymbol{\xi}^\star, [z^-, z^+]) - \overline{d}([z^-, z^+]),
\end{aligned}
$$

where $\hat{J}^+_i = \{j \in [J_i] : \hat{z}^+_i < z_{ij}\}$ and $\hat{J}^-_i = \{j \in [J_i] : \hat{z}^-_i \geq z_{ij}\}$ analogous to J^+_i and J^-_i. Here, (a) is due to (EC.16) and the definition of d', (b) follows from (EC.17) and (EC.18) as well as separating the sum, (c) follows from $z^+ = -z^-$ and $\hat{z}^+ = -\hat{z}^-$ and combining the sums over $[i^\star]$ in the first and last term, and (d) holds since the second row is non-negative by $\hat{z}^-_{\sigma(i)} - z^-_{\sigma(i)} = \sum_{j \in \hat{J}^-_{\sigma(i)} \setminus J^-_{\sigma(i)}} z_{j+1} - z_j$ and $0 \leq \xi^\star_{ij} \leq z_j - z_{j-1}$ for all $i \in [I], j \in [J_i]$, and the third row is non-negative by $\xi^\star_{ij} \geq \xi^\star_{i,J-j}$ $i \in [I], j \in [J/2]$, and $J^+_i = \{J - j : j \in J^-_i\}$ and $\hat{J}^+_i = \{J - j : j \in \hat{J}^-_i\}$ by $z^+ = -z^-$ and $\hat{z}^+ = -\hat{z}^-$, respectively. Thus $[\hat{z}^-, \hat{z}^+]$ is an optimal solution to (8).

To conclude the proof, we show that there is an optimal solution $[\overline{z}^-, \overline{z}^+]$ following the ordering π of $\{\delta_{ij^\star}\}_{i \in [I]}$. Define $\overline{z}^+$ via $\overline{z}^+_{\pi(i)} = \hat{z}^+_{\sigma(i)}$, $i \in [I]$, and $\overline{z}^- = -\overline{z}^+$. If $z_{j^\star} > \eta(i^\star + 1) - \eta(i^\star)$, then we have $\overline{z}^+_{\pi(i)} = \hat{z}^+_{\sigma(i)} = z_{j^\star}$ for all $i \in [I]$, which implies that $\overline{z}^+ = \hat{z}^+$, making $[\overline{z}^-, \overline{z}^+]$ an optimal solution to (8). Now consider the case where $z_{j^\star} \leq \eta(i^\star + 1) - \eta(i^\star)$. As $\overline{d}$ is permutation invariant, $\overline{d}([\overline{z}^-, \overline{z}^+]) = \overline{d}([\hat{z}^-, \hat{z}^+])$. For d', we find that

$$
\begin{aligned}
d'(\boldsymbol{\xi}^\star, [\overline{z}^-, \overline{z}^+]) &= \sum_{i \in [I]} \left(\sum_{j \in \overline{J}^+_i} \xi^\star_{ij} + \overline{z}^-_i - \underline{\theta}_i - \sum_{j \in \overline{J}^-_i} \xi^\star_{ij} \right) \\
&\overset{(a)}{=} \sum_{i \in [I]} \left(\sum_{j \in [J] \setminus [j^\star+1]} \xi^\star_{ij} - z_{j^\star+1} - \underline{\theta}_i - \sum_{j \in [J-(j^\star+1)]} \xi^\star_{ij} \right) \\
&\quad + \sum_{i \in [i^\star]} \left(\xi^\star_{\pi(i), j^\star+1} - z_{j^\star} + z_{j^\star+1} - \xi^\star_{\pi(i), J-j^\star} \right) \\
&\overset{(b)}{=} \sum_{i \in [I]} \left(\sum_{j \in [J] \setminus [j^\star+1]} \xi^\star_{ij} - z_{j^\star+1} - \underline{\theta}_i - \sum_{j \in [J-(j^\star+1)]} \xi^\star_{ij} \right) + \sum_{i \in [i^\star]} \delta_{\pi(i)j^\star} \\
&\overset{(c)}{\geq} \sum_{i \in [I]} \left(\sum_{j \in [J] \setminus [j^\star+1]} \xi^\star_{ij} - z_{j^\star+1} - \underline{\theta}_i - \sum_{j \in [J-(j^\star+1)]} \xi^\star_{ij} \right) + \sum_{i \in [i^\star]} \delta_{\sigma(i)j^\star} \\
&\overset{(d)}{=} d'(\boldsymbol{\xi}^\star, [\hat{z}^-, \hat{z}^+]).
\end{aligned}
$$

Here, (a) follows from splitting the sums and the definition of $[\overline{z}^-, \overline{z}^+]$, (b) is due to the definition of $\delta_{ij^\star}$, (c) holds since π is a non-ascending ordering of $\{\delta_{ij^\star}\}_{i \in [I]}$, and (d) follows from combining the arguments from (a) and (b). Thus, $[\overline{z}^-, \overline{z}^+]$ is an optimal solution to (8) that satisfies all properties claimed by the lemma. □

Lemma EC.4 provides us with the key structural property to efficiently find an optimal solution for the separation problem, as it guarantees that one of the $\frac{IJ}{2}$ solution candidates with the properties specified by the lemma is optimal. Using Algorithm 1, we can efficiently enumerate all of these candidates.

Data: A scaled incumbent solution $\boldsymbol{\xi}^\star = \boldsymbol{\xi}'/s$ to problem Aff $\mathcal{P}'(\mathbb{P}', \overline{\Xi}^\star)$

Result: An optimal solution $[z^-, z^+]$ for the separation problem

$o^\star = o = 0$

for $j = J-1$ **to** $\frac{J}{2}$ **do**

  Set $\delta_{ij} = z_{j+1} - z_j + \xi^\star_{i,j+1} - \xi^\star_{i,J-j}$ for all $i \in [I]$

  Let π be a non-ascending order of the indices of $\{\delta_{ij}\}_{i\in[I]}$

  for $i = 1$ **to** I **do**

$$o = o + \delta_{\pi(i)j} - \begin{cases} z_{j+1} - z_j & \text{if } \eta(i) - \eta(i-1) \geq z_{j+1} \\ \eta(i) - \eta(i-1) - z_j & \text{if } z_{j+1} > \eta(i) - \eta(i-1) \geq z_j \\ 0 & \text{otherwise} \end{cases}$$

    if $o > o^\star$ **then**

      $j^\star = j, i^\star = i, \pi^\star = \pi, o^\star = o$

    end

  end

end

return $[z^-, z^+]$ with $z^+_{\pi^\star(i)} = \begin{cases} z_{j^\star} & \text{if } i \leq i^\star \\ z_{j^\star+1} & \text{if } i > i^\star \end{cases} \quad \forall i \in [I]$ and $z^- = -z^+$

Algorithm 1: Permutation invariant cut separation

LEMMA EC.5. *Let Z satisfy Assumption 1 and let Θ satisfy Assumption 2. Then, Algorithm 1 determines an optimal solution for the separation problem in time $O(IJ\log(I))$.*

Proof of Lemma EC.5. For the runtime complexity note that each operation in the inner loop of Algorithm 1 can be executed in $O(1)$, and thus the entire inner loop runs in time $O(I)$. Sorting the values $\{\delta_{ij}\}_{i\in[I]}$ for a single index j takes time $O(I\log(I))$ using standard sorting algorithms such as merge sort. Thus, the entire runtime of Algorithm 1 is bounded by $O(IJ\log(I))$.

To prove the optimality of the algorithm, we show that it iterates through all solutions of the form introduced by Lemma EC.4. In particular, we show that o maintains the objective value of the solution $[z^-, z^+]$ induced by i, j and π via the construction in Lemma EC.4, that is, $o = d'(\boldsymbol{\xi}^\star, [z^-, z^+]) - \overline{d}(\boldsymbol{\xi}^\star, [z^-, z^+])$. Throughout the remainder of the proof, we use the abbreviations $d'_{ij} = d'(\boldsymbol{\xi}^\star, [z^-, z^+])$ and $\overline{d}_{ij} = \overline{d}(\boldsymbol{\xi}^\star, [z^-, z^+])$. Note that we omit the explicit dependence of $[z^-, z^+]$ on i and j as well as the dependence of π on j. We prove the result of the lemma by showing that

$$d'_{ij} = d'_{i-1,j} + \delta_{\pi(i)j} \qquad \forall i \in [I], j \in [J] \qquad \text{(EC.19)}$$

$$\text{and } \overline{d}_{ij} = \overline{d}_{i-1,j} + \begin{cases} z_{j+1} - z_j & \text{if } \eta(i) - \eta(i-1) \geq z_{j+1} \\ \eta(i) - \eta(i-1) - z_j & \text{if } z_{j+1} > \eta(i) - \eta(i-1) \geq z_j \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in [I], j \in [J]. \tag{EC.20}$$

By construction, we have $d'_{Ij} = d'_{0,j+1}$ and $\overline{d}_{Ij} = \overline{d}_{0,j+1}$. Thus, combining (EC.19) and (EC.20) immediately implies the claim.

Equation (EC.19) follows from the construction of $[z^-, z^+]$ in Lemma EC.4. Specifically, we find

$$\begin{aligned} d'_{ij} &\overset{\text{(a)}}{=} \sum_{i' \in [i]} \left(\sum_{j' \in [J] \setminus [j]} \xi^\star_{\pi(i')j'} - z_j - \underline{\theta}_{i'} - \sum_{j' \in [J-j]} \xi^\star_{\pi(i')j'} \right) \\ &\quad + \sum_{i' \in [I] \setminus [i]} \left(\sum_{j' \in [J] \setminus [j+1]} \xi^\star_{\pi(i')j'} - z_{j+1} - \underline{\theta}_{i'} - \sum_{j' \in [J-(j+1)]} \xi^\star_{\pi(i')j'} \right) \\ &\overset{\text{(b)}}{=} \sum_{i' \in [i-1]} \left(\sum_{j' \in [J] \setminus [j]} \xi^\star_{\pi(i')j'} - z_j - \underline{\theta}_{i'} - \sum_{j' \in [J-j]} \xi^\star_{\pi(i')j'} \right) \\ &\quad + \sum_{i' \in [I] \setminus [i-1]} \left(\sum_{j' \in [J] \setminus [j+1]} \xi^\star_{\pi(i')j'} - z_{j+1} - \underline{\theta}_{i'} - \sum_{j' \in [J-(j+1)]} \xi^\star_{\pi(i')j'} \right) \\ &\quad + (\xi^\star_{\pi(i),j+1} + z_{j+1} - z_j - \xi^\star_{\pi(i),J-j}) \overset{\text{(c)}}{=} d'_{i-1,j} + \delta_{\pi(i)j}, \end{aligned}$$

where (a) and (c) follow from the construction of $[z^-, z^+]$ and the definition of δ_{ij} and (b) follows from shifting the sums.

To show equation (EC.20), we first simplify the expression for $\overline{d}_{ij}$ and prove the claim via a case distinction. Recall that Lemma EC.3 implies the existence of an index $i^\star$ for each $i \in [I]$ and $j \in [J], j \geq \frac{J}{2}$, such that

$$\overline{d}_{ij} = \eta(i^\star) - \sum_{i' \in [i^\star]} z^+_{i'} \overset{\text{(a)}}{=} \eta(i^\star) - i^\star z_{j+1} + \min\{i, i^\star\}(z_{j+1} - z_j), \tag{EC.21}$$

where (a) follows from the construction of z^+. Let $i^\star \in [I]$ be the maximal index satisfying (EC.21) and let $i^- \in [I]$ be the maximal index such that $\overline{d}_{i-1,j} = \eta(i^-) - i^- z_{j+1} + \min\{i-1, i^-\}(z_{j+1} - z_j)$. By construction of $i^\star$, we have

$$\begin{aligned} &\eta(i) - \eta(i-1) \overset{\text{(a)}}{\leq} \eta(i^\star + 1) - \eta(i^\star) \overset{\text{(b)}}{<} z_j && \text{if } i^\star < i \\ &\eta(i+1) - \eta(i) \overset{\text{(a)}}{\leq} \eta(i^\star + 1) - \eta(i^\star) \overset{\text{(b)}}{<} z_{j+1} && \text{if } i^\star \leq i \\ &\eta(i) - \eta(i-1) \overset{\text{(a)}}{\geq} \eta(i^\star) - \eta(i^\star - 1) \overset{\text{(c)}}{\geq} z_j && \text{if } i^\star \geq i \\ &\eta(i+1) - \eta(i) \overset{\text{(a)}}{\geq} \eta(i^\star) - \eta(i^\star - 1) \overset{\text{(c)}}{\geq} z_{j+1} && \text{if } i^\star > i, \end{aligned}$$

where the inequalities (a) are due to the non-increasing differences of η, the inequalities (b) follow from $i^\star$ being a maximal index satisfying (EC.21), and the inequalities (c) are due to the optimality of $i^\star$, that is, $\eta(i^\star) - \sum_{i' \in [i^\star]} z^+_{i'} = \max_{\hat{i} \in [I]} (\eta(\hat{i}) - \sum_{i' \in [\hat{i}]} z^+_{i'})$ by Lemma EC.3. Negating the above implications yields

$$\eta(i) - \eta(i-1) \geq z_j \qquad \Rightarrow \qquad i^\star \geq i \tag{EC.22a}$$

$$\eta(i+1) - \eta(i) \geq z_{j+1} \quad \Rightarrow \quad i^\star > i \tag{EC.22b}$$

$$\eta(i) - \eta(i-1) < z_j \quad \Rightarrow \quad i^\star < i \tag{EC.22c}$$

$$\eta(i+1) - \eta(i) < z_{j+1} \quad \Rightarrow \quad i^\star \leq i. \tag{EC.22d}$$

Using the same argument, we can show that the implications (EC.22a)–(EC.22d) hold for i^- when substituting i by $i-1$. We next consider each of the three cases in the case distinction of (EC.20) separately.

Case 1. Let $\eta(i) - \eta(i-1) \geq z_{j+1}$. We find that

$$\begin{aligned}
\overline{d}_{ij} &= \eta(i^\star) - i^\star z_{j+1} + \min\{i, i^\star\}(z_{j+1} - z_j) \\
&\overset{(a)}{=} \eta(i^\star) - i^\star z_{j+1} + i(z_{j+1} - z_j) \\
&\overset{(b)}{=} \eta(i^-) - i^- z_{j+1} + (i-1)(z_{j+1} - z_j) + (z_{j+1} - z_j) \\
&\overset{(c)}{=} \eta(i^-) - i^- z_{j+1} + \min\{i-1, i^-\}(z_{j+1} - z_j) + (z_{j+1} - z_j) = \overline{d}_{j,i-1} + (z_{j+1} - z_j),
\end{aligned}$$

where (a) and (c) follow from $i^\star \geq i$ by (EC.22a) and $i^- \geq i$ by (EC.22b), respectively. For (b), we use

$$\eta(i^-) - i^- z_{j+1} + (i-1)(z_{j+1} - z_j) \geq \eta(i^\star) - i^\star z_{j+1} + (i-1)(z_{j+1} - z_j)$$

and

$$\eta(i^\star) - i^\star z_{j+1} + i(z_{j+1} - z_j) \geq \eta(i^-) - i^- z_{j+1} + i(z_{j+1} - z_j)$$

that follow from $i^\star, i^- \geq i$ and the optimality of i^- and $i^\star$, respectively.

Case 2. Let $z_{j+1} > \eta(i) - \eta(i-1) \geq z_j$. Then $i^\star \leq i$ by (EC.22d) and the non-increasing differences of η, $i^- \leq i-1$ by (EC.22d), $i^\star \geq i$ by (EC.22a), and $i^- \geq i-1$ by (EC.22a) and the non-increasing differences of η. We thus have $i^\star = i$ and $i^- = i-1$. With this, we find that

$$\begin{aligned}
\overline{d}_{ij} &= \eta(i^\star) - i^\star z_{j+1} + \min\{i, i^\star\}(z_{j+1} - z_j) \\
&\overset{(a)}{=} \eta(i) - i z_{j+1} + i(z_{j+1} - z_j) \\
&\overset{(b)}{=} \eta(i-1) - (i-1) z_{j+1} + (i-1)(z_{j+1} - z_j) + \eta(i) - \eta(i-1) - z_j \\
&\overset{(c)}{=} \eta(i^-) - i^- z_{j+1} + \min\{i-1, i^-\}(z_{j+1} - z_j) + \eta(i) - \eta(i-1) - z_j \\
&= \overline{d}_{j,i-1} + \eta(i) - \eta(i-1) - z_j,
\end{aligned}$$

where (a) and (c) follow from $i^\star = i$ and $i^- = i-1$, respectively, and (b) follows from adding $\eta(i-1) - \eta(i-1)$.

Case 3. Let $\eta(i) - \eta(i-1) < z_j$. We find that

$$\begin{aligned}
\overline{d}_{ij} &= \eta(i^\star) - i^\star z_{j+1} + \min\{i, i^\star\}(z_{j+1} - z_j) \\
&\overset{(a)}{=} \eta(i^\star) - i^\star z_{j+1} + i^\star(z_{j+1} - z_j) \\
&\overset{(b)}{=} \eta(i^-) - i^- z_{j+1} + i^-(z_{j+1} - z_j) \\
&\overset{(c)}{=} \eta(i^-) - i^- z_{j+1} + \min\{i-1, i^-\}(z_{j+1} - z_j) = \overline{d}_{j,i-1},
\end{aligned}$$

where (a) and (c) follow from $i^\star \leq i-1$ by (EC.22c) and $i^- \leq i-1$ by (EC.22d), respectively. For (b), we use

$$\eta(i^-) - i^- z_{j+1} + i^-(z_{j+1} - z_j) \geq \eta(i^\star) - i^\star z_{j+1} + i^\star(z_{j+1} - z_j)$$

and

$$\eta(i^\star) - i^\star z_{j+1} + i^\star(z_{j+1} - z_j) \geq \eta(i^-) - i^- z_{j+1} + i^-(z_{j+1} - z_j)$$

that follow from $i^\star, i^- \leq i-1$ and the optimality of i^- and $i^\star$, respectively. □

LEMMA EC.6. *Let the Assumptions 1 and 2 both hold. If for every $j \in [J]$ there is at most one $i \in [I]$ such that $\eta(i) - \eta(i-1) \in [z_{j-1}, z_j)$, then the separation problem* (8) *has a square solution $[z_j \boldsymbol{e}, -z_j \boldsymbol{e}]$ for some $j \in [J/2]$.*

Proof of Lemma EC.6. The proof follows from Step 1 of Algorithm 1. For any iteration i, j of the algorithm, let o_{ij} be the value of o after Step 1 and recall that $o_{0j} = o_{I,j+1}$. Let $i^\star$, $j^\star$ and π be an optimal solution returned by the algorithm. We claim that

$$o_{i^\star j^\star} \leq \begin{cases} o_{0j^\star} & \text{if } \eta(i^\star) - \eta(i^\star - 1) \geq z_{j^\star+1} \\ o_{Ij^\star} & \text{if } \eta(i^\star) - \eta(i^\star - 1) < z_{j^\star+1} \end{cases}$$

For the first case, note that by the non-increasing differences of η, each $i \leq i^\star$ satisfies $\eta(i) - \eta(i-1) \geq z_{j^\star+1}$. We thus have

$$o_{i^\star j^\star} \overset{\text{(a)}}{=} o_{0j^\star} + \sum_{i \in [i^\star]} \xi^\star_{\pi(i), j^\star+1} - \xi^\star_{\pi(i), J-j^\star} \overset{\text{(b)}}{\leq} o_{0j^\star},$$

where (a) follows from Step 1. For (b) we use

$$\xi^\star_{\pi(i), j^\star+1} = (z_{j^\star+1} - z_{j^\star}) \frac{\xi^\star_{\pi(i), j^\star+1}}{z_{j^\star+1} - z_{j^\star}} \overset{\text{(c)}}{\leq} (z_{j^\star+1} - z_{j^\star}) \frac{\xi^\star_{\pi(i), J-j^\star}}{z_{J-j^\star} - z_{J-(j^\star+1)}} \overset{\text{(d)}}{=} \xi^\star_{\pi(i), J-j^\star},$$

where (c) follows from $\boldsymbol{\xi}^\star \in \overline{\Xi}'$, $j^\star \geq J/2$, and Lemma 1 and (d) holds since $z_j = -z_{J-j}$, $j \in [J]$, by Assumption 1.

For the second case, note that the non-increasing differences of η and the assumption that there is at most one i with $\eta(i) - \eta(i-1) \in [z_{j^\star}, z_{j^\star+1})$ imply that each $i > i^\star$ satisfies $\eta(i) - \eta(i-1) < z_{j^\star}$. We thus have

$$o_{Ij^\star} \overset{\text{(a)}}{=} o_{i^\star j^\star} + \sum_{i \in [i^\star]} \xi^\star_{\pi(i), j^\star+1} + z_{j^\star+1} - z_{j^\star} - \xi^\star_{\pi(i), J-j^\star} \overset{\text{(b)}}{\geq} o_{i^\star j^\star},$$

where (a) is due to Step 1 and (b) follows from $0 \leq \xi^\star_{\pi(i), j^\star+1} - \xi^\star_{\pi(i), J-j^\star} \leq z_{j^\star+1} - z_{j^\star}$ by $\boldsymbol{\xi}^\star \in \overline{\Xi}'$, $j^\star \geq J/2$, Lemma 1, and Assumption 1. □

EC.7. Proof of Proposition 2

Proof of Proposition 2. Property (i) of Assumption 2 immediately follows from the permutation invariance of the norm $\|\cdot\|_p$. Property (ii) of Assumption 2, on the other hand, follows from the closed-form expression of $\eta(i)$. Indeed, Lemma EC.1 implies that $\eta(i) = i\lambda(i)$, where $\lambda(i) = i^{-\frac{1}{p}}\delta$ is the maximal value satisfying $\left\|\sum_{i' \in [i]} \lambda(i)\boldsymbol{e}_{i'}\right\|_p \leq \delta$. We thus have $\eta(i) = i^{1-\frac{1}{p}}\delta$. Since $p \geq 1$, the exponent of i in this expression is always between zero and one. This makes η a concave function of i, which implies non-increasing differences. □

EC.8. Proof of Proposition 3

Proof of Proposition 3. We first show the permutation invariance of Θ. Let $\boldsymbol{\xi} \in \Theta = \bigcap_{h \in [n]} \Theta_h$ and $\boldsymbol{S} \in \mathcal{S}$. Then $\boldsymbol{S\xi} \in \Theta_h$ for all $h \in [n]$ by the symmetry and permutation invariance of each Θ_h. Thus $\boldsymbol{S\xi} \in \bigcap_{h \in [n]} \Theta_h = \Theta$.

We next show that $\eta(i) = \min_{h \in [n]} \eta_h(i)$, which implies property (ii) of Assumption 2, as the minimum preserves non-increasing differences. Indeed, Lemma EC.1 implies $\eta_h(i) = i\lambda_h(i)$ for all $h \in [n], i \in [I]$, where $\lambda_h(i)$ is the maximal value such that $\lambda_h(i) \sum_{i' \in [i]} \boldsymbol{e}_{i'} \in \Theta_h$. Analogously, $\eta(i) = i\lambda(i)$, where $\lambda(i)$ is the maximal value such that $\lambda(i) \sum_{i' \in [i]} \boldsymbol{e}_{i'} \in \Theta = \bigcap_{h \in [n]} \Theta_h$. The intersection immediately implies that $\lambda(i) \leq \min_{h \in [n]} \lambda_h(i)$. By convexity and the fact that $\boldsymbol{0} \in \Xi_h$, we further observe that $\min_{h' \in [n]} \lambda_{h'}(i) \sum_{i' \in [i]} \boldsymbol{e}_{i'} \in \Xi_h$ for all $h \in [n]$. We thus have $\min_{h \in [n]} \lambda_h(i) \sum_{i' \in [i]} \boldsymbol{e}_{i'} \in \bigcap_{h \in [n]} \Theta_h = \Theta$, which implies that $\lambda(i) = \min_{h \in [n]} \lambda_h(i)$ and $\eta(i) = \min_{h \in [n]} \eta_h(i)$. □

EC.9. Proof of Proposition 4

Proof of Proposition 4. We use the reformulations offered by Observation 2 to show the inequalities from the statement of the proposition. In particular, we show that (i) any optimal solution to the reformulation of $\operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^\star)$ is a feasible solution to the reformulation of $\operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{A}})$ and (ii) any optimal solution to the reformulation of $\operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{A}})$ is a feasible solution to the reformulation of $\operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{P}})$.

(i) $\operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^\star) \leq \operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{A}})$: Note that the reformulations only differ in the affine cuts (7). More specifically, the reformulation of $\operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{A}})$ only adds the subsets $\mathcal{Z}_g^{\mathrm{A}} \subseteq \mathcal{Z}$ of cuts for each $g \in [G]$ found by the separation algorithm. Thus, the solution space of $\operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^\star)$ is a subset of the solution space of $\operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{A}})$, which immediately implies the inequality.

(ii) $\operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{A}}) \leq \operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{P}})$: Let $(\boldsymbol{\xi}', \boldsymbol{s})$ be a feasible solution of $\operatorname{Aff} \mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{A}})$. Note that the reformulations only differ in the last set of constraints in (6) and the affine cuts (7), which separate along indices $g \in [G], k \in [K_g]$. For the remainder of the proof, we fix one index pair (g, k). For $s_{gk} = 0$, the constraints of interest trivially hold. Thus, we w.l.o.g. assume that $s_{gk} > 0$ and define $\boldsymbol{\xi}^\star = \frac{1}{s_{gk}} \boldsymbol{\xi}'_{gk}$. Then, the last constraint of (6) is equivalent to $\boldsymbol{\xi}^\star \in \overline{\Xi}'_g$, which implies that $F^+(\boldsymbol{\xi}^\star) \in \Theta_g$ and $\boldsymbol{\xi}^\star \in \operatorname{conv}(F([\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g]))$.

Using $\Theta_g \subseteq \Theta_g^{\mathrm{P}}$ and consequently $[\underline{\boldsymbol{\theta}}_g, \overline{\boldsymbol{\theta}}_g] \subseteq [\underline{\boldsymbol{\theta}}_g^{\mathrm{P}}, \overline{\boldsymbol{\theta}}_g^{\mathrm{P}}]$, we find that $F^+(\boldsymbol{\xi}^\star) \in \Theta_g^{\mathrm{P}}$ and $\boldsymbol{\xi}^\star \in \mathrm{conv}(F([\underline{\boldsymbol{\theta}}_g^{\mathrm{P}}, \overline{\boldsymbol{\theta}}_g^{\mathrm{P}}]))$, which implies that $(\boldsymbol{\xi}', \boldsymbol{s})$ satisfies all constraints in the reformulation (6) of Aff $\mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{P}})$.

For the cuts (7) we recall that $\overline{d}_g$ depends on the embedded support set Θ_g. We denote by $\overline{d}_g^{\mathrm{P}}([z^-, z^+]) = \max_{\boldsymbol{\theta} \in \Theta_g^{\mathrm{P}}} d(\boldsymbol{\theta}, [z^-, z^+])$ the maximal distance to the permutation invariant Θ_g^{P}. Note that $\overline{d}_g([z^-, z^+]) \leq \overline{d}_g^{\mathrm{P}}([z^-, z^+])$ by $\Theta_g \subseteq \Theta_g^{\mathrm{P}}$. Thus, by

$$d'(\boldsymbol{\xi}^\star, [z^-, z^+]) \leq \overline{d}_g([z^-, z^+]) \leq \overline{d}_g^{\mathrm{P}}([z^-, z^+])$$

all cuts induced via $[z^-, z^+] \in \mathcal{Z}_g^{\mathrm{A}}$ are also satisfied by the reformulation of Aff $\mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{P}})$. By Lemma EC.4 we can w.l.o.g. restrict the reformulation of Aff $\mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{P}})$ to a subset $\mathcal{Z}^{\mathrm{P}} \subseteq \mathcal{Z}$ satisfying the specific structure described in the lemma. Further, any violated cut in the reformulation of Aff $\mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{P}})$ induced by some $[z^-, z^+] \in \mathcal{Z}^{\mathrm{P}}$ is guaranteed to be found by Algorithm 1. Note that the order in which cuts are enumerated by Algorithm 1 solely depends on the current solution $\boldsymbol{\xi}^\star$, and recall that after each iteration of Step 1 of the algorithm we have $o = d'(\boldsymbol{\xi}^\star, [z^-, z^+]) - \max_{i^\star \in [I]} \{\eta_g(i^\star) - \sum_{i' \in [i^\star]} z_{i'}^+\}$, where $\eta_g(i) = \max_{\boldsymbol{\theta} \in \Theta_g} \sum_{i' \in [i]} |\boldsymbol{\theta}_{i'}|$, which follows from the proof of Lemma EC.5. Denote by o^{P} the value for the permutation invariant case, and let $\eta_g^{\mathrm{P}}(i)$ be the maximal l_1-norm of the first i components of Θ_g^{P}. We then have

$$\eta_g(i) = \max_{\boldsymbol{\theta} \in \Theta} \sum_{i' \in [i]} |\boldsymbol{\theta}_{i'}| \leq \max_{\boldsymbol{\theta} \in \Theta^{\mathrm{P}}} \sum_{i' \in [i]} |\boldsymbol{\theta}_{i'}| = \eta_g^{\mathrm{P}}(i),$$

which implies that $o \geq o^{\mathrm{P}}$. Thus, any violated cut for the reformulation of Aff $\mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{P}})$ found by Algorithm 1 is also found as a violated cut for the reformulation of Aff $\mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{A}})$. Consequently, at termination any feasible solution to the reformulation of Aff $\mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{A}})$ is also a feasible solution to the reformulation of Aff $\mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\mathrm{P}})$. □

EC.10. Proof of Proposition 5

The proof of Proposition 5 builds on the piecewise affine policies designed by Ben-Tal et al. (2020) and Thomä et al. (2024). To keep our work self-contained, we briefly summarize their key results in the context of our problem setting.

Piecewise Affine Policies Via Domination. The key observation for piecewise affine policies via domination is that for any two uncertainty realizations $\boldsymbol{\xi}$ and $\hat{\boldsymbol{\xi}}$, where $\hat{\boldsymbol{\xi}}$ dominates $\boldsymbol{\xi}$ in the sense that $\boldsymbol{\xi} \leq \hat{\boldsymbol{\xi}}$, Assumption 3 (ii) implies that $\boldsymbol{b}_g(\hat{\boldsymbol{\xi}}) \leq \boldsymbol{b}_g(\boldsymbol{\xi})$ for all $g \in [G]$. Thus, any fixed solution $\boldsymbol{x}$ that is feasible for $\hat{\boldsymbol{\xi}}$ is also feasible for $\boldsymbol{\xi}$. We use this property to construct approximate policies by replacing the uncertainty support Ξ via a dominating support $\hat{\Xi}$, such that for each $\boldsymbol{\xi} \in \Xi$ there is some $\hat{\boldsymbol{\xi}} \in \hat{\Xi}$ with $\boldsymbol{\xi} \leq \hat{\boldsymbol{\xi}}$. Assumption 4 extends the domination properties described above to the embedded support Θ. Analogous to the dominating polytopes in the original works, we define a *dominating polytope* $\hat{\Theta} = \mathrm{conv}(\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_I)$ for Θ such that $\boldsymbol{v}_0 \in \mathbb{R}^I$, $\beta(\boldsymbol{v}_0) \in \mathbb{R}_+$, $\boldsymbol{v}_i = \boldsymbol{v}_0 + \rho \boldsymbol{e}_i$ with $\rho \geq 0$ and $\sum_{i \in [I]} (\theta_i - v_{0i})_+ \leq \rho$ for all $\boldsymbol{\theta} \in \Theta$. Note that the last property implies that $\hat{\Theta}$ dominates Θ (*cf.* Thomä et al. 2024, Condition (5)). When replacing the embedded support with a dominating polytope $\hat{\Theta}$, all worst-case realizations are contained within the vertices $\boldsymbol{v}_0, \ldots, \boldsymbol{v}_I$. Thus, the dominating problem is equivalent to

$$\begin{array}{lll} \underset{\boldsymbol{x}_0,\ldots,\boldsymbol{x}_I}{\text{minimize}} & \boldsymbol{c}^\top \boldsymbol{x}_{00} & \\ \text{subject to} & \boldsymbol{A}_g \boldsymbol{x}_i \leq \boldsymbol{b}_g(\boldsymbol{E}^+ \boldsymbol{v}_i) & \forall g \in [G], i \in [I]_0 \\ & \boldsymbol{x}_{it} = \boldsymbol{x}_{i't} & \forall i, i' \in [I]_0, t \in [T]_0 \text{ with } \boldsymbol{v}_i^t = \boldsymbol{v}_{i'}^t, \end{array} \qquad (\hat{\mathcal{P}}(\hat{\Theta}))$$

which follows from straightforward minor modifications to Lemma 2 of Thomä et al. (2024). Intuitively, $\hat{\mathcal{P}}(\hat{\Theta})$ determines one solution $\boldsymbol{x}_i$ to the embedded problem for each vertex $\boldsymbol{v}_i$. The last set of constraints couples the solutions $\boldsymbol{x}_i$ to ensure nonanticipativity. Note that the here-and-now decisions $\boldsymbol{x}_{i0}$ are identical for all $i \in [I]_0$ as $\boldsymbol{v}_i^0$ is the empty vector, and thus $\boldsymbol{v}_i^0 = \boldsymbol{v}_{i'}^0$ holds for any $i, i' \in [I]_0$. The solution $(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_I)$ to $\hat{\mathcal{P}}(\hat{\Theta})$ induces a feasible piecewise affine solution $\boldsymbol{x}(\boldsymbol{\xi})$ to $\mathcal{P}(\mathbb{P}, \Xi)$ with the same objective value via $\boldsymbol{x}(\boldsymbol{\xi}) = \boldsymbol{x}_0 + \frac{1}{\rho} \sum_{i \in [I]} ([\boldsymbol{E}(\boldsymbol{\xi})]_i - v_{0i})_+ (\boldsymbol{x}_i - \boldsymbol{x}_0)$. This policy is further guaranteed to be within a factor $\beta(\boldsymbol{v}_0) + \rho$ of optimal solutions to $\mathcal{P}(\mathbb{P}, \Xi)$. Since these guarantees are central to our proof of Proposition 5, we first formalize them by adapting Theorem 1 of Thomä et al. (2024) to the dominating problem $\hat{\mathcal{P}}(\hat{\Theta})$.

LEMMA EC.7. *Let $\mathcal{P}(\mathbb{P}, \Xi)$ be a problem instance that satisfies Assumption 3, let $\boldsymbol{E}$ be such that Assumption 4 holds, and let $\hat{\Theta} = \operatorname{conv}(\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_I)$ be a dominating polytope for Θ. Then,*

$$\mathcal{P}(\mathbb{P}, \Xi) \leq \hat{\mathcal{P}}(\hat{\Theta}) \leq (\beta(\boldsymbol{v}_0) + \rho) \mathcal{P}(\mathbb{P}, \Xi).$$

Proof of Lemma EC.7. First, note that either both problems are bounded or both problems are unbounded. This follows, as unboundedness of either of the problems implies the existence of a feasible ray of decisions $\boldsymbol{x}$ with negative objective value, implying that both problems are unbounded (*cf.* Part 1 of the original proof). For the remainder of the proof, we assume both problems are bounded and show each of the inequalities separately.

For the first inequality, let $(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_I)$ be an optimal solution to $\hat{\mathcal{P}}(\hat{\Theta})$ and consider the piecewise affine policy $\boldsymbol{x}(\boldsymbol{\xi}) = \boldsymbol{x}_0 + \frac{1}{\rho} \sum_{i \in [I]} ([\boldsymbol{E}(\boldsymbol{\xi})]_i - v_{0i})_+ (\boldsymbol{x}_i - \boldsymbol{x}_0)$ on $\mathcal{P}(\mathbb{P}, \Xi)$. The nonanticipativity constraints of $\hat{\mathcal{P}}(\hat{\Theta})$ imply that the here-and-now decisions of $\boldsymbol{x}(\boldsymbol{\xi})$ are given by $\boldsymbol{x}(\boldsymbol{\xi})_0 = \boldsymbol{x}_{00}$. Thus, $\mathbb{E}_{\mathbb{P}}[\boldsymbol{c}(\tilde{\boldsymbol{\xi}})^\top \boldsymbol{x}(\tilde{\boldsymbol{\xi}})] = \boldsymbol{c}^\top \boldsymbol{x}_{00} = \hat{\mathcal{P}}(\hat{\Xi})$ by part (i) of Assumption 3. Further, $\boldsymbol{x}(\boldsymbol{\xi})$ is feasible for $\mathcal{P}(\mathbb{P}, \Xi)$ as for any $g \in [G]$ and $\boldsymbol{\xi} \in \Xi_g$, we have

$$\begin{aligned} \boldsymbol{A}_g \boldsymbol{x}(\boldsymbol{\xi}) &\overset{\text{(a)}}{=} \boldsymbol{A}_g \left(\left(1 - \frac{1}{\rho} \sum_{i \in [I]} ([\boldsymbol{E}(\boldsymbol{\xi})]_i - v_{0i})_+ \right) \boldsymbol{x}_0 + \frac{1}{\rho} \sum_{i \in [I]} ([\boldsymbol{E}(\boldsymbol{\xi})]_i - v_{0i})_+ \boldsymbol{x}_i \right) \\ &\overset{\text{(b)}}{\leq} \left(\left(1 - \frac{1}{\rho} \sum_{i \in [I]} ([\boldsymbol{E}(\boldsymbol{\xi})]_i - v_{0i})_+ \right) \boldsymbol{b}_g(\boldsymbol{E}^+ \boldsymbol{v}_0) + \frac{1}{\rho} \sum_{i \in [I]} ([\boldsymbol{E}(\boldsymbol{\xi})]_i - v_{0i})_+ \boldsymbol{b}_g(\boldsymbol{E}^+ \boldsymbol{v}_i) \right) \end{aligned}$$

$$\overset{\text{(c)}}{=} \boldsymbol{b}_g\left(\boldsymbol{E}^+\left(\boldsymbol{v}_0 + \frac{1}{\rho}\sum_{i\in[I]}([\boldsymbol{E}(\boldsymbol{\xi})]_i - v_{0i})_+(\boldsymbol{v}_i - \boldsymbol{v}_0)\right)\right)$$
$$\overset{\text{(d)}}{=} \boldsymbol{b}_g(\boldsymbol{E}^+(\max\{\boldsymbol{E}(\boldsymbol{\xi}), \boldsymbol{v}_0\})) \overset{\text{(e)}}{\le} \boldsymbol{b}_g(\boldsymbol{E}^+(\boldsymbol{E}(\boldsymbol{\xi}))) \overset{\text{(f)}}{=} \boldsymbol{b}_g(\boldsymbol{\xi}).$$

Here, (a) follows from the definition of $\boldsymbol{x}$, and (b) follows from $\boldsymbol{A}_g\boldsymbol{x}_i \le \boldsymbol{b}_g(\boldsymbol{E}^+\boldsymbol{v}_i)$ for all $i \in [I]_0$ by feasibility of $(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_I)$ in $\hat{\mathcal{P}}(\hat{\Theta})$, linearity, and the factors pre-multiplying $\boldsymbol{b}_g$ being non-negative. In particular, the choice of ρ implies that $1 - \frac{1}{\rho}\sum_{i\in[I]}([\boldsymbol{E}(\boldsymbol{\xi})]_i - v_{0i})_+ \ge 0$. Finally, (c) follows from the linearity of $\boldsymbol{b}_g$ and $\boldsymbol{E}^+$, (d) follows from rearranging the terms, using the element-wise maximum, (e) follows from the dominating property implied by Assumptions 3 (ii) and Assumption 4, and (f) follows from $\boldsymbol{E}^+$ being a left inverse of $\boldsymbol{E}$.

For the second inequality, let $\boldsymbol{x}^\star$ be an optimal solution to $\mathcal{P}(\mathbb{P}, \Xi)$ and consider $\boldsymbol{x}_i = (\beta(\boldsymbol{v}_0) + \rho)\boldsymbol{x}^\star\left(\boldsymbol{E}^+\left(\frac{1}{\beta(\boldsymbol{v}_0)+\rho}\boldsymbol{v}_i\right)\right)$ for each $i \in [I]_0$. First, we find

$$\frac{1}{\beta(\boldsymbol{v}_0)+\rho}\boldsymbol{v}_i \overset{\text{(a)}}{=} \frac{1}{\beta(\boldsymbol{v}_0)+\rho}(\boldsymbol{v}_0 + \rho\boldsymbol{e}_i) \overset{\text{(b)}}{=} \frac{\beta(\boldsymbol{v}_0)}{\beta(\boldsymbol{v}_0)+\rho}\frac{1}{\beta(\boldsymbol{v}_0)}\boldsymbol{v}_0 + \frac{\rho}{\beta(\boldsymbol{v}_0)+\rho}\boldsymbol{e}_i \overset{\text{(c)}}{\in} \Theta_g$$

for all $i \in [I]$ and $g \in [G]$. Here, (a) and (b) are immediate, and (c) follows from $\boldsymbol{e}_i \in \Theta_g$ by Assumption 4, $\frac{1}{\beta(\boldsymbol{v}_0)}\boldsymbol{v}_0 \in \Theta_g$ by definition of β, and convexity of Θ_g. Similarly, $\frac{1}{\beta(\boldsymbol{v}_0)+\rho}\boldsymbol{v}_0 \le \frac{1}{\beta(\boldsymbol{v}_0)}\boldsymbol{v}_0 \in \Theta_g$ implies $\frac{1}{\beta(\boldsymbol{v}_0)+\rho}\boldsymbol{v}_0 \in \Theta_g$ by down-monotonicity from Assumption 4. We now show that $(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_I)$ is a feasible solution for $\hat{\mathcal{P}}(\hat{\Theta})$. For the first set of constraints, we find

$$\boldsymbol{A}_g\boldsymbol{x}_i \overset{\text{(a)}}{=} (\beta(\boldsymbol{v}_0)+\rho)\boldsymbol{A}_g\boldsymbol{x}^\star\left(\boldsymbol{E}^+\left(\frac{1}{\beta(\boldsymbol{v}_0)+\rho}\boldsymbol{v}_i\right)\right)$$
$$\overset{\text{(b)}}{\le} (\beta(\boldsymbol{v}_0)+\rho)\boldsymbol{b}_g\left(\boldsymbol{E}^+\left(\frac{1}{\beta(\boldsymbol{v}_0)+\rho}\boldsymbol{v}_i\right)\right) \overset{\text{(c)}}{\le} \boldsymbol{b}_g(\boldsymbol{E}^+\boldsymbol{v}_i)$$

for all $i \in [I]$ and $g \in [G]$. Here, (a) follows by definition of $\boldsymbol{x}_i$, (b) follows by $\boldsymbol{E}^+\left(\frac{1}{\beta(\boldsymbol{v}_0)+\rho}\boldsymbol{v}_i\right) \in \Xi_g$ and feasibility of $\boldsymbol{x}^\star$, and (c) follows from $\boldsymbol{b}_g$ having non-positive coefficients, $\beta(\boldsymbol{v}_0) + \rho \ge 1$, and linearity of $\boldsymbol{E}^+$. The non-anticipativity condition of $\hat{\mathcal{P}}(\hat{\Theta})$ holds by non-anticipativity of $\boldsymbol{x}^\star$ and $\boldsymbol{E}$ being information preserving. Finally, the objective value satisfies $\boldsymbol{c}^\top\boldsymbol{x}_{00} = (\beta(\boldsymbol{v}_0)+\rho)\boldsymbol{c}^\top\boldsymbol{x}_0^\star = (\beta(\boldsymbol{v}_0)+\rho)\mathcal{P}(\mathbb{P}, \Xi)$ by Assumption 3 (i). □

We now use these policies and their approximation guarantees to prove Proposition 5.

Proof of Proposition 5. In the following, we explicitly prove the claim for $\text{Aff}\,\mathcal{P}'(\mathbb{P}, \overline{\Xi}^{\text{A}})$. Specifically, we show that $\text{Aff}\,\mathcal{P}'(\mathbb{P}', \overline{\Xi}^{\text{A}}) \le (\beta(\boldsymbol{z}) + \max_{g\in[G]} \overline{d}_g([\boldsymbol{0}, \boldsymbol{z}])) \cdot \mathcal{P}(\mathbb{P}, \Xi)$ for each $\boldsymbol{z} \in Z^{\text{A}}$. For the remainder of the proof, we w.l.o.g. fix one $\boldsymbol{z} \in Z^{\text{A}}$. Consider the polytope $\hat{\Theta} = \text{conv}(\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_I)$ as described above with $\boldsymbol{v}_0 = \boldsymbol{z}$ and $\rho = \max_{g\in[G]} \overline{d}_g([\boldsymbol{0}, \boldsymbol{z}])$. Then, $\hat{\Theta}$ is a valid dominating polytope for Θ since

$$\max_{\boldsymbol{\theta}\in\Theta}\sum_{i\in[I]}(\theta_i - z_i)_+ \overset{\text{(a)}}{=} \max_{\boldsymbol{\theta}\in\Theta}\sum_{i\in[I]}\min_{\theta_i'\in[0,z_i]}|\theta_i - \theta_i'| = \max_{\boldsymbol{\theta}\in\Theta}\min_{\boldsymbol{\theta}'\in[\boldsymbol{0},\boldsymbol{z}]}\|\boldsymbol{\theta} - \boldsymbol{\theta}'\|_1$$
$$\overset{\text{(b)}}{=} \max_{\boldsymbol{\theta}\in\Theta} d(\boldsymbol{\theta}, [\boldsymbol{0}, \boldsymbol{z}]) \overset{\text{(c)}}{=} \max_{g\in[G]}\max_{\boldsymbol{\theta}\in\Theta_g} d(\boldsymbol{\theta}, [\boldsymbol{0}, \boldsymbol{z}]) \overset{\text{(d)}}{=} \max_{g\in[G]} \overline{d}_g([\boldsymbol{0}, \boldsymbol{z}]) \overset{\text{(e)}}{=} \rho,$$

where (a) follows from $\Theta \geq \mathbf{0}$ by Assumption 4, (b) follows by definition of d, (c) follows by definition of $\Theta = \bigcup_{g \in [G]} \Theta_g$, (d) follows by definition of $\overline{d}_g$, and (e) follows by definition of ρ. Thus, $\hat{\mathcal{P}}(\hat{\Theta})$ is a $\beta(\boldsymbol{z}) + \max_{g \in [G]} \overline{d}_g([\mathbf{0}, \boldsymbol{z}])$ approximation of $\mathcal{P}(\mathbb{P}, \Xi)$ by Lemma EC.7. Let $(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_I)$ be an optimal solution to $\hat{\mathcal{P}}(\hat{\Theta})$ and consider the affine policy $\boldsymbol{x}'(\boldsymbol{\xi}') = \boldsymbol{x}_0 + \frac{1}{\rho} \sum_{i \in [I], j \in J_i^+} \xi'_{ij} (\boldsymbol{x}_i - \boldsymbol{x}_0)$ on $\overline{\Xi}^{\mathrm{A}}$. The nonanticipativity constraints of $\hat{\mathcal{P}}(\hat{\Theta})$ imply that the here-and-now decisions of $\boldsymbol{x}'(\boldsymbol{\xi}')$ are given by $\boldsymbol{x}'(\boldsymbol{\xi}')_0 = \boldsymbol{x}_{00}$. Thus, $\mathbb{E}_{\mathbb{P}'}[\boldsymbol{c}(R(\tilde{\boldsymbol{\xi}}'))^\top \boldsymbol{x}'(\tilde{\boldsymbol{\xi}}')] = \boldsymbol{c}^\top \boldsymbol{x}_{00} = \hat{\mathcal{P}}(\hat{\Theta})$ by part (i) of Assumption 3. Finally, $\boldsymbol{x}'(\boldsymbol{\xi}')$ is feasible for $\mathrm{Aff}\, \mathcal{P}'(\mathbb{P}', \overline{\Xi}^{\mathrm{A}})$ as for any $g \in [G]$ and $\boldsymbol{\xi}' \in \Xi'_g$, we have

$$
\begin{aligned}
\boldsymbol{A}_g \boldsymbol{x}'(\boldsymbol{\xi}') &\overset{\text{(a)}}{=} \boldsymbol{A}_g \left(\left(1 - \frac{1}{\rho} \sum_{i \in [I], j \in J_i^+} \xi'_{ij} \right) \boldsymbol{x}_0 + \frac{1}{\rho} \sum_{i \in [I], j \in J_i^+} \xi'_{ij} \boldsymbol{x}_i \right) \\
&\overset{\text{(b)}}{\leq} \left(\left(1 - \frac{1}{\rho} \sum_{i \in [I], j \in J_i^+} \xi'_{ij} \right) \boldsymbol{b}_g(\boldsymbol{E}^+ \boldsymbol{v}_0) + \frac{1}{\rho} \sum_{i \in [I], j \in J_i^+} \xi'_{ij} \boldsymbol{b}_g(\boldsymbol{E}^+ \boldsymbol{v}_i) \right) \\
&\overset{\text{(c)}}{=} \boldsymbol{b}_g \left(\boldsymbol{E}^+ \left(\boldsymbol{v}_0 + \frac{1}{\rho} \sum_{i \in [I], j \in J_i^+} \xi'_{ij} (\boldsymbol{v}_i - \boldsymbol{v}_0) \right) \right) \\
&\overset{\text{(d)}}{\leq} \boldsymbol{b}_g(\boldsymbol{E}^+(F^+(\boldsymbol{\xi}'))) \overset{\text{(e)}}{=} \boldsymbol{b}_g(R(\boldsymbol{\xi}')).
\end{aligned}
$$

Here, (a) follows from the definition of $\boldsymbol{x}'$, and (b) follows from $\boldsymbol{A}_g \boldsymbol{x}_i \leq \boldsymbol{b}_g(\boldsymbol{E}^+ \boldsymbol{v}_i)$ for all $i \in [I]_0$ by feasibility of $(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_I)$ in $\hat{\mathcal{P}}(\hat{\Theta})$, linearity, and the factors pre-multiplying $\boldsymbol{b}_g$ being non-negative. In particular, the cuts (7) for $\boldsymbol{z}$ imply that $1 - \frac{1}{\rho} \sum_{i \in [I], j \in J_i^+} \xi'_{ij} \geq 0$ by $\sum_{i \in [I], j \in J_i^+} \xi'_{ij} = d'(\boldsymbol{\xi}', [\mathbf{0}, \boldsymbol{z}]) \leq \overline{d}_g([\mathbf{0}, \boldsymbol{z}]) \leq \rho$. Finally, (c) follows from the linearity of $\boldsymbol{b}_g$ and $\boldsymbol{E}^+$, (d) follows from $\boldsymbol{v}_0 + \frac{1}{\rho} \sum_{i \in [I], j \in J_i^+} \xi'_{ij} (\boldsymbol{v}_i - \boldsymbol{v}_0) = \boldsymbol{z} + \sum_{i \in [I], j \in J_i^+} \xi'_{ij} \boldsymbol{e}_i \geq F^+(\boldsymbol{\xi}')$ and the dominating property implied by Assumption 3 (ii) and Assumption 4, and (e) follows from the definition of R. We thus observe that

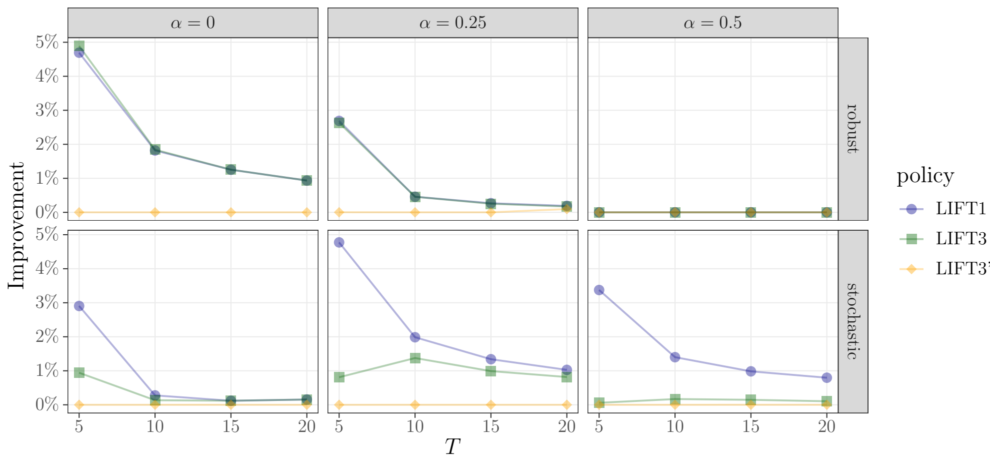
$$
\mathrm{Aff}\, \mathcal{P}'(\mathbb{P}', \overline{\Xi}^{\mathrm{A}}) \leq \hat{\mathcal{P}}(\hat{\Theta}) \leq (\beta(\boldsymbol{z}) + \max_{g \in [G]} \overline{d}_g([\mathbf{0}, \boldsymbol{z}])) \cdot \mathcal{P}(\mathbb{P}, \Xi)
$$

as desired. The case for $\mathrm{Aff}\, \mathcal{P}'(\mathbb{P}, \overline{\Xi}^\star)$ follows analogously by setting $\overline{\Xi}^{\mathrm{A}} = \overline{\Xi}^\star$ and $Z^{\mathrm{A}} = Z$. □

EC.11. Dynamic Cut Generation

The experiments in Section 6 do not use the full dynamic cut generation procedure described in Section 4; instead, they limit our policies to square cuts $[z_j \boldsymbol{e}, -z_j \boldsymbol{e}]$ for all $j \in [J/2]$. Figure EC.1 shows that dynamically adding more cuts when computing these policies may improve objective values by up to 5% for small instances and by up to 2% for medium-sized instances. However, Table EC.1 indicates that these improvements come at the cost of significant increases in solution times.

Figure EC.1 Dynamic cut objective improvements.



Note. Relative objective improvements of different piecewise affine policies with dynamic cut generation over their respective counterparts as used in Section 6 for different serial demand correlations α and time horizons T.

Table EC.1 Dynamic cut solution times.

T	LIFT1	LIFT1 CG	LIFT3	LIFT3 CG	LIFT3'	LIFT3' CG
5	0.01	0.11	0.01	0.14	0.02	0.03
10	0.04	1.20	0.09	4.18	0.08	3.43
15	0.12	6.82	0.52	23.39	0.61	28.32
20	0.31	18.61	1.18	58.08	1.59	62.44

Average solution times (in seconds) for the different policies from Figure EC.1 without and with dynamic cut generation (CG).

References

Beck A, Ben-Tal A (2009) Duality in robust optimization: Primal worst equals dual best. *Operations Research Letters* 37(1):1–6, ISSN 0167-6377, URL `http://dx.doi.org/10.1016/j.orl.2008.09.010`.

Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* 99(2):351–376, URL `http://dx.doi.org/10.1007/s10107-003-0454-y`.

Ben-Tal A, Housni OE, Goyal V (2020) A tractable approach for designing piecewise affine policies in two-stage adjustable robust optimization. *Mathematical Programming* 182(1-2):57–102, URL `http://dx.doi.org/10.1007/s10107-019-01385-0`.

Georghiou A, Wiesemann W, Kuhn D (2015) Generalized decision rule approximations for stochastic programming via liftings. *Mathematical Programming* 152(1-2):301–338, URL `http://dx.doi.org/10.1007/s10107-014-0789-6`.

Karp RM (1972) Reducibility among combinatorial problems. Miller RE, Thatcher JW, Bohlinger JD, eds., *Complexity of Computer Computations*, 85–103 (Springer US), ISBN 978-1-4684-2001-2, URL `http://dx.doi.org/10.1007/978-1-4684-2001-2_9`.

Thomä S, Walther G, Schiffer M (2024) Designing tractable piecewise affine policies for multi-stage adjustable robust optimization. *Mathematical Programming* 208(1–2):661–716, ISSN 1436-4646, URL http://dx.doi.org/10.1007/s10107-023-02053-0.

Zhen J, Kuhn D, Wiesemann W (2025) A unified theory of robust and distributionally robust optimization via the primal-worst-equals-dual-best principle. *Operations Research* 73(2):862–878, ISSN 1526-5463, URL http://dx.doi.org/10.1287/opre.2021.0268.