E-Companion: Proofs

Proof of Observation 1. If there is no $\boldsymbol{x} \in \mathcal{X}$ and $\boldsymbol{y} \in \mathcal{Y}$ such that $\boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y} \leq \boldsymbol{h}$, then problem (3) and problem $\mathcal{PO}_K$ are both infeasible and the statement holds true. In the remainder of the proof, we thus assume that there is $\boldsymbol{x} \in \mathcal{X}$ and $\boldsymbol{y} \in \mathcal{Y}$ such that $\boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y} \leq \boldsymbol{h}$.

Any feasible solution $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_{k \in \mathcal{K}})$ to problem $\mathcal{PO}_K$ is also feasible in (3), and it has the same objective value in both problems since all second-stage policies $\boldsymbol{y}^k$ satisfy the second-stage constraints in (3). Conversely, take any feasible solution $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_{k \in \mathcal{K}})$ to problem (3). Since Ξ is nonempty and the problem is feasible, at least one policy $\boldsymbol{y}^{k^\star}$ must satisfy the second-stage constraints. By replacing any policy $\boldsymbol{y}^k$ that violates the second-stage constraints with $\boldsymbol{y}^{k^\star}$ if necessary, we obtain a feasible solution to problem $\mathcal{PO}_K$ that attains the same objective value. □

Proof of Theorem 1. Assume that $\dim \mathcal{Y} \leq \operatorname{rk} \boldsymbol{Q}$ and consider problem $\mathcal{PO}_K$ with $K = |\mathcal{Y}|$ policies. Note that $|\mathcal{Y}| < \infty$ because $\mathcal{Y} \subseteq \{0,1\}^M$. For this choice of K, problem $\mathcal{PO}_K$ has the same optimal objective value as $\mathcal{PO}$. Indeed, one readily verifies that any feasible solution $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_{k \in \mathcal{K}})$ to problem $\mathcal{PO}_K$ can be transformed into a feasible solution to problem $\mathcal{PO}$ that attains the same objective value and vice versa. We can reformulate the objective function of $\mathcal{PO}_K$ as

$$\max_{\boldsymbol{\xi} \in \Xi} \left[\boldsymbol{\xi}^\top \boldsymbol{C}\, \boldsymbol{x} + \min_{\boldsymbol{\lambda} \in \Delta_K} \sum_{k \in \mathcal{K}} \boldsymbol{\lambda}_k \cdot \boldsymbol{\xi}^\top \boldsymbol{Q}\, \boldsymbol{y}^k \right],$$

where $\Delta_K = \{\boldsymbol{\lambda} \in \mathbb{R}^K_+ : \mathbf{e}^\top \boldsymbol{\lambda} = 1\}$ denotes the unit simplex in $\mathbb{R}^K$. Since the terms in the objective function are linear in $\boldsymbol{\xi}$ and linear in $\boldsymbol{\lambda}$, we can employ the classical min-max theorem to exchange the order of the maximum and minimum operators. Problem $\mathcal{PO}_K$ is thus equivalent to

$$\begin{aligned} &\text{minimize} \quad \max_{\boldsymbol{\xi} \in \Xi} \left[\boldsymbol{\xi}^\top \boldsymbol{C}\, \boldsymbol{x} + \sum_{k \in \mathcal{K}} \boldsymbol{\lambda}_k \cdot \boldsymbol{\xi}^\top \boldsymbol{Q}\, \boldsymbol{y}^k \right] \\ &\text{subject to} \quad \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y}^k \in \mathcal{Y},\ k \in \mathcal{K},\ \boldsymbol{\lambda} \in \Delta_K \\ &\qquad\qquad\quad \boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k \leq \boldsymbol{h} \qquad \forall k \in \mathcal{K}. \end{aligned} \tag{EC.1}$$

We claim that problem (EC.1) has the same optimal value as

$$\begin{aligned}
&\text{minimize} && \max_{\boldsymbol{\xi}\in\Xi}\left[\boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x} + \sum_{k\in\mathcal{D}} \boldsymbol{\lambda}_k \cdot \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k\right] \\
&\text{subject to} && \boldsymbol{x}\in\mathcal{X},\ \ \boldsymbol{y}^k\in\mathcal{Y},\, k\in\mathcal{D},\ \ \boldsymbol{\lambda}\in\Delta_{D+1} \\
& && \boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k \leq \boldsymbol{h} \qquad\qquad \forall k\in\mathcal{D},
\end{aligned} \tag{EC.2}$$

where $\mathcal{D}=\{1,\ldots,D+1\}$, $D=\dim\mathcal{Y}$ and Δ_{D+1} is the unit simplex in $\mathbb{R}^{D+1}$. Note that the optimal value of (EC.2) provides an upper bound on the optimal value of (EC.1) since $D+1=\dim\mathcal{Y}+1\leq |\mathcal{Y}|=K$. We now show that both problems indeed attain the same optimal value. To this end, fix an optimal solution $(\boldsymbol{x}^\star,\{\boldsymbol{y}^{\star,k}\}_{k\in\mathcal{K}},\boldsymbol{\lambda}^\star)$ to problem (EC.1) and define $\boldsymbol{y}'\in\mathbb{R}^M$ as $\boldsymbol{y}'=\sum_{k\in\mathcal{K}}\boldsymbol{\lambda}^\star_k\cdot \boldsymbol{y}^{\star,k}$. By construction, $\boldsymbol{y}'$ is contained in the convex hull of the K policies $\boldsymbol{y}^{\star,k}\in\mathbb{R}^M$, $k\in\mathcal{K}$. From Carathéodory's theorem we then conclude that $\boldsymbol{y}'$ is indeed contained in the convex hull of $D+1$ policies $\boldsymbol{y}^{\star,k_1},\ldots,\boldsymbol{y}^{\star,k_{D+1}}$, $k_1,\ldots,k_{D+1}\in\mathcal{K}$ (not necessarily all different). Choose $\boldsymbol{\lambda}'\in\Delta_{D+1}$ such that $\boldsymbol{y}'=\sum_{d\in\mathcal{D}}\boldsymbol{\lambda}'_d\cdot\boldsymbol{y}^{\star,k_d}$. We readily verify that $(\boldsymbol{x}^\star,\{\boldsymbol{y}^{\star,k_d}\}_{d\in\mathcal{D}},\boldsymbol{\lambda}')$ is a feasible solution to problem (EC.2) that attains the same objective value as $(\boldsymbol{x}^\star,\{\boldsymbol{y}^{\star,k}\}_{k\in\mathcal{K}},\boldsymbol{\lambda}^\star)$ in problem (EC.1). We thus conclude that both problems attain the same optimal value.

Employing the classical min-max theorem, we can reformulate problem (EC.2) as

$$\begin{aligned}
&\text{minimize} && \max_{\boldsymbol{\xi}\in\Xi}\left[\boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x} + \min_{k\in\mathcal{D}}\ \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k\right] \\
&\text{subject to} && \boldsymbol{x}\in\mathcal{X},\ \ \boldsymbol{y}^k\in\mathcal{Y},\, k\in\mathcal{D} \\
& && \boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k \leq \boldsymbol{h} \qquad\qquad \forall k\in\mathcal{D},
\end{aligned}$$

which we readily identify as the $(D+1)$-adaptability problem. We have thus shown the statement of the theorem for $\dim\mathcal{Y}\leq\operatorname{rk}\boldsymbol{Q}$. We can proceed analogously for the case $\operatorname{rk}\boldsymbol{Q}<\dim\mathcal{Y}$ if we define $\boldsymbol{y}'\in\mathbb{R}^M$ as $\boldsymbol{y}'=\sum_{k\in\mathcal{K}}\boldsymbol{\lambda}^\star_k\cdot\boldsymbol{Q}\boldsymbol{y}^{\star,k}$. This concludes the proof. $\square$

Proof of Theorem 2. For fixed $\boldsymbol{x}$ and $\boldsymbol{y}^k$, $k \in \mathcal{K}$, we can employ an epigraph formulation to express the evaluation of the objective function in $\mathcal{PO}_K$ as

$$\begin{array}{ll} \text{maximize} & \boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x} + \tau \\ \text{subject to} & \boldsymbol{\xi} \in \mathbb{R}^Q,\ \tau \in \mathbb{R} \\ & \boldsymbol{A}\boldsymbol{\xi} \le \boldsymbol{b} \\ & \tau \le \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \qquad \forall k \in \mathcal{K}. \end{array}$$

Strong LP duality implies that this LP has the same optimal value as its dual problem,

$$\begin{array}{ll} \text{minimize} & \boldsymbol{b}^\top \boldsymbol{\alpha} \\ \text{subject to} & \boldsymbol{\alpha} \in \mathbb{R}^R_+,\ \boldsymbol{\beta} \in \mathbb{R}^K_+ \\ & \boldsymbol{A}^\top \boldsymbol{\alpha} = \boldsymbol{C}\boldsymbol{x} + \sum_{k \in \mathcal{K}} \boldsymbol{\beta}_k \boldsymbol{Q}\boldsymbol{y}^k \\ & \mathbf{e}^\top \boldsymbol{\beta} = 1. \end{array}$$

The result now follows if we replace the bilinear terms $\boldsymbol{\beta}_k \boldsymbol{y}^k$ with auxiliary variables $\boldsymbol{z}^k \in \mathbb{R}^M_+$, $k \in \mathcal{K}$, subject to the constraints that

$$\boldsymbol{z}^k = \boldsymbol{\beta}_k \boldsymbol{y}^k \quad \Longleftrightarrow \quad \boldsymbol{z}^k \le \boldsymbol{y}^k,\ \boldsymbol{z}^k \le \boldsymbol{\beta}_k \mathbf{e},\ \boldsymbol{z}^k \ge (\boldsymbol{\beta}_k - 1)\mathbf{e} + \boldsymbol{y}^k.$$

This reformulation exploits the fact that $\mathbf{0} \le \boldsymbol{\beta}$, $\boldsymbol{y}^k \le \mathbf{e}$ and that $\boldsymbol{y}^k$ is binary. $\square$

Proof of Theorem 3. For ease of exposition, we deferred the proof of the first statement to Corollary 1 in the main text. In the following, we focus on the proof of the second statement.

We recall that the strongly $\mathcal{NP}$-hard 0/1 Integer Programming (IP) feasibility problem is defined as follows (Garey and Johnson 1979).

0/1 INTEGER PROGRAMMING FEASIBILITY.

Instance. Given are $\boldsymbol{A} \in \mathbb{R}^{R \times Q}$ and $\boldsymbol{b} \in \mathbb{R}^R$.

Question. Is there a vector $\boldsymbol{\xi} \in \{0,1\}^Q$ such that $\boldsymbol{A}\boldsymbol{\xi} \le \boldsymbol{b}$?

We aim to reduce the IP feasibility problem to evaluating the objective function of an instance of problem $\mathcal{P}_K$. To this end, fix an instance $(\boldsymbol{A}, \boldsymbol{b})$ of the IP feasibility problem, and let $\Xi = \{\boldsymbol{\xi} \in \mathbb{R}^Q : \boldsymbol{A}\boldsymbol{\xi} \leq \boldsymbol{b},\ \mathbf{0} \leq \boldsymbol{\xi} \leq \mathbf{e}\}$ denote the LP relaxation of the IP problem's feasible region. Set $\boldsymbol{y}^k = \mathbf{e}_k$, $k = 1, \ldots, 2Q$, where $\mathbf{e}_k$ denotes the k-th unit basis vector in $\mathbb{R}^{2Q}$. We claim that the answer to the IP feasibility problem is affirmative if and only if

$$\max_{\boldsymbol{\xi} \in \Xi} \min_{k \in \mathcal{K}} \left\{ (\boldsymbol{\xi}^\top,\ \mathbf{e}^\top - \boldsymbol{\xi}^\top)\, \boldsymbol{y}^k \ :\ \boldsymbol{y}^k_q \leq 2\boldsymbol{\xi}_q,\ \boldsymbol{y}^k_{Q+q} \leq 2(1 - \boldsymbol{\xi}_q),\ q = 1, \ldots, Q \right\} \tag{EC.3}$$

is greater than or equal to 1. Note that (EC.3) can be interpreted as the objective function of an instance of problem $\mathcal{P}_K$ if we set $N = 0$, $M = K = 2Q$ and $\mathcal{Y} = \{\boldsymbol{y} \in \{0,1\}^{2Q} : \mathbf{e}^\top \boldsymbol{y} = 1\}$.

To prove our claim, assume first that the answer to the IP feasibility problem is affirmative, that is, there is $\boldsymbol{\xi}^\star \in \Xi$ such that $\boldsymbol{\xi}^\star \in \{0,1\}^Q$. The policy $\boldsymbol{y}^q$, $q = 1, \ldots, Q$, satisfies the constraints in (EC.3) for $\boldsymbol{\xi} = \boldsymbol{\xi}^\star$ if and only if $\boldsymbol{\xi}^\star_q = 1$, in which case the objective function within the minimization evaluates to $(\boldsymbol{\xi}^{\star\top},\ \mathbf{e}^\top - \boldsymbol{\xi}^{\star\top})\, \mathbf{e}_q = \boldsymbol{\xi}^\star_q = 1$. Likewise, the policy $\boldsymbol{y}^{Q+q}$, $q = 1, \ldots, Q$, satisfies the constraints in (EC.3) for $\boldsymbol{\xi} = \boldsymbol{\xi}^\star$ if and only if $\boldsymbol{\xi}^\star_q = 0$, in which case the objective function within the minimization evaluates to $(\boldsymbol{\xi}^{\star\top},\ \mathbf{e}^\top - \boldsymbol{\xi}^{\star\top})\, \mathbf{e}_{Q+q} = 1 - \boldsymbol{\xi}^\star_q = 1$. We thus conclude that the minimum in (EC.3) evaluates to 1 for $\boldsymbol{\xi}^\star$, which implies that the value of the overall expression (EC.3) is greater than or equal to 1.

Assume now that the value of the expression in (EC.3) is greater than or equal to 1. Fix any maximizer $\boldsymbol{\xi}^\star$ for this expression and consider any component $\boldsymbol{\xi}^\star_q$ of $\boldsymbol{\xi}^\star$, $q = 1, \ldots, Q$. If $\boldsymbol{\xi}^\star_q \geq 1/2$, then $\boldsymbol{y}^q = \mathbf{e}_q$ is a feasible second-stage policy, and we conclude that $(\boldsymbol{\xi}^{\star\top},\ \mathbf{e}^\top - \boldsymbol{\xi}^{\star\top})\, \mathbf{e}_q = \boldsymbol{\xi}^\star_q \geq 1$. If $\boldsymbol{\xi}^\star_q \leq 1/2$, on the other hand, then $\boldsymbol{y}^{Q+q} = \mathbf{e}_{Q+q}$ is a feasible second-stage policy, and we conclude that $(\boldsymbol{\xi}^{\star\top},\ \mathbf{e}^\top - \boldsymbol{\xi}^{\star\top})\, \mathbf{e}_{Q+q} = 1 - \boldsymbol{\xi}^\star_q \geq 1$, that is, $\boldsymbol{\xi}^\star_q \leq 0$. Since $\boldsymbol{\xi}^\star_q \in [0,1]$ by construction of Ξ, we thus conclude that $\boldsymbol{\xi}^\star_q \in \{0,1\}$. Since q was chosen arbitrarily, the answer to the IP feasibility problem is affirmative. This concludes the proof. $\square$

Proof of Theorem 4. Consider the following instance of problem $\mathcal{P}$ where no first-stage decision is taken.

$$\begin{aligned} &\text{maximize} && \min_{\boldsymbol{y}\in\{0,1\}^Q} \{0 : \boldsymbol{y}_q - \boldsymbol{\xi}_q \le 1/2,\ \boldsymbol{\xi}_q - \boldsymbol{y}_q \le 1/2,\ q=1,\ldots,Q\} \\ &\text{subject to} && \boldsymbol{\xi}\in\mathbb{R}^Q \\ & && 0\le \boldsymbol{\xi}_q \le 1,\ q=1,\ldots,Q \end{aligned} \tag{EC.4}$$

The minimization in the objective function of (EC.4) has a nonempty feasible region. In fact, for any $\boldsymbol{\xi}\in[0,1]^Q$, the second-stage policy $\boldsymbol{y}(\boldsymbol{\xi})$ defined through $[\boldsymbol{y}(\boldsymbol{\xi})]_q = \mathbb{I}(\boldsymbol{\xi}_q \ge 1/2),\ q=1,\ldots,Q$, satisfies the second-stage constraints in the objective function of (EC.4). We thus conclude that the optimal value of problem (EC.4) is 0.

We claim that the optimal value of the K-adaptability problem associated with (EC.4),

$$\begin{aligned} &\text{minimize} && \max_{\boldsymbol{\xi}\in[0,1]^Q} \min_{k\in\mathcal{K}} \left\{0 : \boldsymbol{y}^k_q - \boldsymbol{\xi}_q \le 1/2,\ \boldsymbol{\xi}_q - \boldsymbol{y}^k_q \le 1/2, q=1,\ldots,Q\right\} \\ &\text{subject to} && \boldsymbol{y}^k\in\{0,1\}^Q,\ k\in\mathcal{K}, \end{aligned} \tag{EC.5}$$

is unbounded whenever $K<|\mathcal{Y}|=2^Q$. In fact, assume that a solution $\{\boldsymbol{y}^k\}_{k\in\mathcal{K}}$ to (EC.5) does not contain the policy $\boldsymbol{y}^\star\in\{0,1\}^Q$. Then the objective function in (EC.5) is unbounded since there is no policy $\boldsymbol{y}^k$, $k\in\mathcal{K}$, that satisfies the second-stage constraints in the objective function of (EC.5) for the parameter realization $\boldsymbol{\xi}=\boldsymbol{y}^\star$. We thus conclude that the problems (EC.4) and (EC.5) only attain the same optimal values if $K=|\mathcal{Y}|$. □

Proof of Proposition 1. We first show that $\{\Xi(\boldsymbol{\ell})\}_{\boldsymbol{\ell}\in\mathcal{L}}$ is a cover of Ξ, that is, $\bigcup_{\boldsymbol{\ell}\in\mathcal{L}}\Xi(\boldsymbol{\ell})=\Xi$. To this end, fix any $\boldsymbol{\xi}\in\Xi$ and choose $\boldsymbol{\ell}\in\mathcal{L}$ such that

$$\boldsymbol{\ell}_k = \begin{cases} 0 & \text{if } \boldsymbol{T}\boldsymbol{x}+\boldsymbol{W}\boldsymbol{y}^k\le\boldsymbol{H}\boldsymbol{\xi}, \\ \min\left\{l\in\{1,\ldots,L\} : [\boldsymbol{T}\boldsymbol{x}+\boldsymbol{W}\boldsymbol{y}^k]_l > [\boldsymbol{H}\boldsymbol{\xi}]_l\right\} & \text{otherwise} \end{cases} \qquad \forall k\in\mathcal{K}.$$

By construction, we have $\boldsymbol{\xi}\in\Xi(\boldsymbol{\ell})$. Conversely, the definition of $\Xi(\boldsymbol{\ell})$ implies that $\Xi(\boldsymbol{\ell})\subseteq\Xi$ for all $\boldsymbol{\ell}\in\mathcal{L}$. We thus conclude that $\{\Xi(\boldsymbol{\ell})\}_{\boldsymbol{\ell}\in\mathcal{L}}$ is indeed a cover of Ξ. Note that $\{\Xi(\boldsymbol{\ell})\}_{\boldsymbol{\ell}\in\mathcal{L}}$ is in general not a partition of Ξ since the sets $\Xi(\boldsymbol{\ell})$ overlap whenever a policy $\boldsymbol{y}^k$ violates multiple constraints.

Our findings imply that the objective function of problem $\mathcal{P}_K$ satisfies

$$\begin{aligned}
&\max_{\boldsymbol{\xi}\in\Xi}\left[\boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x}+\min_{k\in\mathcal{K}}\left\{\boldsymbol{\xi}^\top\boldsymbol{Q}\,\boldsymbol{y}^k : \boldsymbol{T}\boldsymbol{x}+\boldsymbol{W}\boldsymbol{y}^k\le \boldsymbol{H}\boldsymbol{\xi}\right\}\right]\\
&=\max_{\boldsymbol{\xi}\in\bigcup_{\boldsymbol{\ell}\in\mathcal{L}}\Xi(\boldsymbol{\ell})}\left[\boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x}+\min_{k\in\mathcal{K}}\left\{\boldsymbol{\xi}^\top\boldsymbol{Q}\,\boldsymbol{y}^k : \boldsymbol{T}\boldsymbol{x}+\boldsymbol{W}\boldsymbol{y}^k\le \boldsymbol{H}\boldsymbol{\xi}\right\}\right]\\
&=\max_{\boldsymbol{\ell}\in\mathcal{L}}\max_{\boldsymbol{\xi}\in\Xi(\boldsymbol{\ell})}\left[\boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x}+\min_{k\in\mathcal{K}}\left\{\boldsymbol{\xi}^\top\boldsymbol{Q}\,\boldsymbol{y}^k : \boldsymbol{T}\boldsymbol{x}+\boldsymbol{W}\boldsymbol{y}^k\le \boldsymbol{H}\boldsymbol{\xi}\right\}\right].
\end{aligned}$$

We now note that by definition of the set $\Xi(\boldsymbol{\ell})$, the indices $k\in\mathcal{K}$ for which the associated second-stage policies $\boldsymbol{y}^k$ satisfy the constraints $\boldsymbol{T}\boldsymbol{x}+\boldsymbol{W}\boldsymbol{y}^k\le\boldsymbol{H}\boldsymbol{\xi}$ are precisely those indices $k\in\mathcal{K}$ for which $\boldsymbol{\ell}_k=0$. This concludes the proof. $\square$

The proof of Proposition 2 requires the following auxiliary result, which we prove first.

Lemma 1 *Consider the uncertainty sets $\Xi(\boldsymbol{\ell})$ and $\Xi_\epsilon(\boldsymbol{\ell})$ defined in (6) and (6_ϵ), respectively. Fix $\boldsymbol{\ell}\in\mathcal{L}$ and assume that $\Xi(\boldsymbol{\ell})\neq\emptyset$. Then the sets $\Xi_\epsilon(\boldsymbol{\ell})$ converge to $\Xi(\boldsymbol{\ell})$ in Hausdorff distance as ϵ approaches 0, that is, we have*

$$\lim_{\epsilon\downarrow 0} d^{\mathrm{H}}(\Xi_\epsilon(\boldsymbol{\ell}),\Xi(\boldsymbol{\ell}))=0,$$

where the Hausdorff distance between two sets X and Y is defined as

$$d^{\mathrm{H}}(X,Y)=\max\left\{\sup_{\boldsymbol{x}\in X}\inf_{\boldsymbol{y}\in Y}\|\boldsymbol{x}-\boldsymbol{y}\|,\ \sup_{\boldsymbol{y}\in Y}\inf_{\boldsymbol{x}\in X}\|\boldsymbol{x}-\boldsymbol{y}\|\right\}.$$

Proof. Note that $\Xi_\epsilon(\boldsymbol{\ell})=\Xi(\boldsymbol{\ell})$ for all $\epsilon>0$ if $\boldsymbol{\ell}=\boldsymbol{0}$. In the remainder, we thus assume that $\boldsymbol{\ell}\neq\boldsymbol{0}$. We first show that $\Xi_\epsilon(\boldsymbol{\ell})\neq\emptyset$ for sufficiently small $\epsilon>0$. To this end, select any $\overline{\boldsymbol{\xi}}\in\Xi(\boldsymbol{\ell})$ and define

$$\epsilon'=\min\left\{[\boldsymbol{T}\boldsymbol{x}+\boldsymbol{W}\boldsymbol{y}^k]_{\boldsymbol{\ell}_k}-[\boldsymbol{H}\overline{\boldsymbol{\xi}}]_{\boldsymbol{\ell}_k} : k\in\mathcal{K},\ \boldsymbol{\ell}_k\neq 0\right\}.$$

By construction, $\epsilon'>0$ and $\overline{\boldsymbol{\xi}}\in\Xi_\epsilon(\boldsymbol{\ell})$ for all $\epsilon\in(0,\epsilon']$, that is, $\Xi_\epsilon(\boldsymbol{\ell})\neq\emptyset$ for sufficiently small ϵ. From now on, we always assume that ϵ is chosen so that $\Xi_\epsilon(\boldsymbol{\ell})$ is nonempty.

By definition of $\Xi_\epsilon(\boldsymbol{\ell})$, we have that $\Xi_\epsilon(\boldsymbol{\ell})\subseteq\Xi(\boldsymbol{\ell})$. Hence, we only have to prove that

$$\forall\kappa>0\,\exists\overline{\epsilon}>0:\ \sup_{\boldsymbol{\xi}\in\Xi(\boldsymbol{\ell})}\inf_{\boldsymbol{\xi}'\in\Xi_\epsilon(\boldsymbol{\ell})}\|\boldsymbol{\xi}-\boldsymbol{\xi}'\|\le\kappa\ \forall\epsilon\in(0,\overline{\epsilon}].$$

Assume to the contrary that there is $\kappa > 0$ such that

$$\forall i \in \mathbb{N}\, \exists \epsilon_i \in (0, 1/i],\, \boldsymbol{\xi}^i \in \Xi(\boldsymbol{\ell}) \,:\, \left\| \boldsymbol{\xi}^i - \boldsymbol{\xi}' \right\| > \kappa \;\; \forall \boldsymbol{\xi}' \in \Xi_\epsilon(\boldsymbol{\ell}).$$

Since $\Xi(\boldsymbol{\ell}) \subseteq \Xi$ and Ξ is compact, we can apply the Bolzano-Weierstrass Theorem to conclude that the sequence $\boldsymbol{\xi}^i$ has an accumulation point $\boldsymbol{\xi}^\star \in \mathrm{cl}\, \Xi(\boldsymbol{\ell})$. However, by construction $\boldsymbol{\xi}^\star$ satisfies $\|\boldsymbol{\xi}^\star - \boldsymbol{\xi}'\| \geq \kappa$ for all $\boldsymbol{\xi}' \in \bigcup_{i \in \mathbb{N}} \Xi_{\epsilon_i}(\boldsymbol{\ell}) = \Xi(\ell)$, which contradicts the statement that $\boldsymbol{\xi}^\star \in \mathrm{cl}\, \Xi(\boldsymbol{\ell})$. We thus conclude that our assumption was wrong, that is, the assertion in the statement of the lemma indeed holds. $\square$

Proof of Proposition 2. In view of the first statement, Lemma 1 implies that for sufficiently small $\epsilon > 0$, the approximate uncertainty sets $\Xi_\epsilon(\boldsymbol{\ell})$, $\boldsymbol{\ell} \in \mathcal{L}$, are nonempty if and only if the exact uncertainty sets $\Xi(\boldsymbol{\ell})$ are. Thus, we have $\mathrm{dom}\,(6_\epsilon) = \mathrm{dom}\,(6)$ for sufficiently small $\epsilon > 0$.

As for the second statement, fix any $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_{k \in \mathcal{K}}) \in \mathrm{dom}\,(6)$ and assume that $\epsilon > 0$ is small enough so that $\mathrm{dom}\,(6_\epsilon) = \mathrm{dom}\,(6)$. Denote by φ_ϵ and φ the objective values of $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_{k \in \mathcal{K}})$ in problems (6_ϵ) and (6), respectively. We then have that

$$\begin{aligned}
|\varphi_\epsilon - \varphi| &\overset{\text{(a)}}{=} \max_{\boldsymbol{\ell} \in \mathcal{L}} \max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{\ell})} \left[\boldsymbol{\xi}^\top \boldsymbol{C}\, \boldsymbol{x} + \min_{\substack{k \in \mathcal{K}: \\ \boldsymbol{\ell}_k = 0}} \boldsymbol{\xi}^\top \boldsymbol{Q}\, \boldsymbol{y}^k \right] - \max_{\boldsymbol{\ell}' \in \mathcal{L}} \max_{\boldsymbol{\xi}' \in \Xi_\epsilon(\boldsymbol{\ell})} \left[\boldsymbol{\xi}'^\top \boldsymbol{C}\, \boldsymbol{x} + \min_{\substack{k' \in \mathcal{K}: \\ \boldsymbol{\ell}'_{k'} = 0}} \boldsymbol{\xi}'^\top \boldsymbol{Q}\, \boldsymbol{y}^{k'} \right] \\
&= \max_{\boldsymbol{\ell} \in \mathcal{L}} \max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{\ell})} \min_{\boldsymbol{\ell}' \in \mathcal{L}} \min_{\boldsymbol{\xi}' \in \Xi_\epsilon(\boldsymbol{\ell}')} \left[\boldsymbol{\xi}^\top \boldsymbol{C}\, \boldsymbol{x} + \min_{\substack{k \in \mathcal{K}: \\ \boldsymbol{\ell}_k = 0}} \boldsymbol{\xi}^\top \boldsymbol{Q}\, \boldsymbol{y}^k \right] - \left[\boldsymbol{\xi}'^\top \boldsymbol{C}\, \boldsymbol{x} + \min_{\substack{k' \in \mathcal{K}: \\ \boldsymbol{\ell}'_{k'} = 0}} \boldsymbol{\xi}'^\top \boldsymbol{Q}\, \boldsymbol{y}^{k'} \right] \\
&\overset{\text{(b)}}{\leq} \max_{\boldsymbol{\ell} \in \mathcal{L}} \max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{\ell})} \min_{\boldsymbol{\xi}' \in \Xi_\epsilon(\boldsymbol{\ell})} \left[\boldsymbol{\xi}^\top \boldsymbol{C}\, \boldsymbol{x} + \min_{\substack{k \in \mathcal{K}: \\ \boldsymbol{\ell}_k = 0}} \boldsymbol{\xi}^\top \boldsymbol{Q}\, \boldsymbol{y}^k \right] - \left[\boldsymbol{\xi}'^\top \boldsymbol{C}\, \boldsymbol{x} + \min_{\substack{k' \in \mathcal{K}: \\ \boldsymbol{\ell}_{k'} = 0}} \boldsymbol{\xi}'^\top \boldsymbol{Q}\, \boldsymbol{y}^{k'} \right] \\
&\overset{\text{(c)}}{\leq} \max_{\boldsymbol{\ell} \in \mathcal{L}} \max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{\ell})} \min_{\boldsymbol{\xi}' \in \Xi_\epsilon(\boldsymbol{\ell})} \max_{\substack{k \in \mathcal{K}: \\ \boldsymbol{\ell}_k = 0}} \left[\boldsymbol{\xi}^\top \boldsymbol{C}\, \boldsymbol{x} + \boldsymbol{\xi}^\top \boldsymbol{Q}\, \boldsymbol{y}^k \right] - \left[\boldsymbol{\xi}'^\top \boldsymbol{C}\, \boldsymbol{x} + \boldsymbol{\xi}'^\top \boldsymbol{Q}\, \boldsymbol{y}^k \right] \\
&\overset{\text{(d)}}{\leq} \left(\max_{\boldsymbol{\ell} \in \mathcal{L}} \max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{\ell})} \min_{\boldsymbol{\xi}' \in \Xi_\epsilon(\boldsymbol{\ell})} \|\boldsymbol{\xi} - \boldsymbol{\xi}'\| \right) \cdot \left(\max_{k \in \mathcal{K}} \left\| \boldsymbol{C}\boldsymbol{x} + \boldsymbol{Q}\boldsymbol{y}^k \right\| \right),
\end{aligned}$$

where (a) holds because $\varphi \geq \varphi_\epsilon$ since $\Xi_\epsilon(\boldsymbol{\ell}) \subseteq \Xi(\boldsymbol{\ell})$ for all $\boldsymbol{\ell} \in \mathcal{L}$, (b) and (c) follow since we restrict the minimizations to $\boldsymbol{\ell}' = \boldsymbol{\ell}$ and to the minimizer in the second pair of parentheses, respectively, and (d) is due to the Cauchy-Schwarz inequality. Lemma 1 implies that the first product term in the last expression can be made arbitrarily small by choosing $\epsilon > 0$ appropriately. The second product term, on the other hand, is bounded over $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_{k\in\mathcal{K}}) \in \mathrm{dom}\,(6)$ since $\mathcal{X}$ and $\mathcal{Y}$ are bounded. For sufficiently small $\epsilon > 0$, we can thus bound the difference $|\varphi_\epsilon - \varphi|$ uniformly over $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_{k\in\mathcal{K}}) \in \mathrm{dom}\,(6)$ by an arbitrarily small constant, which concludes the proof. $\square$

Proof of Theorem 5. The objective function of the approximate problem (6_ϵ) is identical to

$$\max_{\boldsymbol{\ell}\in\mathcal{L}} \max_{\boldsymbol{\xi}\in\Xi_\epsilon(\boldsymbol{\ell})} \left[\boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x} + \min_{\boldsymbol{\lambda}\in\Delta_K(\boldsymbol{\ell})} \sum_{k\in\mathcal{K}} \boldsymbol{\lambda}_k \cdot \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k\right],$$

where $\Delta_K(\boldsymbol{\ell}) = \{\boldsymbol{\lambda} \in \mathbb{R}^K_+ : \mathbf{e}^\top\boldsymbol{\lambda} = 1,\ \boldsymbol{\lambda}_k = 0\ \forall k \in \mathcal{K} : \boldsymbol{\ell}_k \neq 0\}$. Note that $\Delta_K(\boldsymbol{\ell}) = \emptyset$ if and only if $\boldsymbol{\ell} > \mathbf{0}$. If $\Xi_\epsilon(\boldsymbol{\ell}) = \emptyset$ for all $\boldsymbol{\ell} \in \mathcal{L}_+$, then the problem is equivalent to

$$\max_{\boldsymbol{\ell}\in\partial\mathcal{L}} \max_{\boldsymbol{\xi}\in\Xi_\epsilon(\boldsymbol{\ell})} \left[\boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x} + \min_{\boldsymbol{\lambda}\in\Delta_K(\boldsymbol{\ell})} \sum_{k\in\mathcal{K}} \boldsymbol{\lambda}_k \cdot \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k\right],$$

and we can apply the classical min-max theorem (since $\Delta_K(\boldsymbol{\ell})$ is nonempty for all $\boldsymbol{\ell} \in \partial\mathcal{L}$) to obtain the equivalent reformulation

$$\max_{\boldsymbol{\ell}\in\partial\mathcal{L}} \min_{\boldsymbol{\lambda}\in\Delta_K(\boldsymbol{\ell})} \max_{\boldsymbol{\xi}\in\Xi_\epsilon(\boldsymbol{\ell})} \left[\boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x} + \sum_{k\in\mathcal{K}} \boldsymbol{\lambda}_k \cdot \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k\right],$$

which in turn is equivalent to

$$\min_{\substack{\boldsymbol{\lambda}(\boldsymbol{\ell})\in\Delta_K(\boldsymbol{\ell}),\\ \boldsymbol{\ell}\in\partial\mathcal{L}}} \max_{\boldsymbol{\ell}\in\partial\mathcal{L}} \max_{\boldsymbol{\xi}\in\Xi_\epsilon(\boldsymbol{\ell})} \left[\boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x} + \sum_{k\in\mathcal{K}} \boldsymbol{\lambda}_k(\boldsymbol{\ell}) \cdot \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k\right].$$

If, on the other hand, $\Xi_\epsilon(\boldsymbol{\ell}) \neq \emptyset$ for some $\boldsymbol{\ell} \in \mathcal{L}_+$, then the objective function in (6_ϵ) evaluates to $+\infty$. Using an epigraph reformulation, we thus conclude that (6_ϵ) is equivalent to the problem

$$\begin{array}{lll}
\text{minimize} & \tau & \\
\text{subject to} & \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y}^k \in \mathcal{Y},\ k \in \mathcal{K}, & \\
& \tau \in \mathbb{R},\ \boldsymbol{\lambda}(\boldsymbol{\ell}) \in \Delta_K(\boldsymbol{\ell}),\ \boldsymbol{\ell} \in \partial\mathcal{L} & \\
& \tau \geq \boldsymbol{\xi}^\top \boldsymbol{C}\,\boldsymbol{x} + \sum_{k\in\mathcal{K}} \boldsymbol{\lambda}_k(\boldsymbol{\ell}) \cdot \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k & \forall \boldsymbol{\ell} \in \partial\mathcal{L},\ \forall \boldsymbol{\xi} \in \Xi_\epsilon(\boldsymbol{\ell}) \\
& \Xi_\epsilon(\boldsymbol{\ell}) = \emptyset & \forall \boldsymbol{\ell} \in \mathcal{L}_+.
\end{array} \tag{EC.6}$$

The semi-infinite constraint associated with $\boldsymbol{\ell} \in \partial\mathcal{L}$ is satisfied if and only if the optimal value of

$$
\begin{array}{lll}
\text{maximize} & \left[\boldsymbol{C}\boldsymbol{x} + \sum_{k\in\mathcal{K}} \boldsymbol{\lambda}_k(\boldsymbol{\ell})\boldsymbol{Q}\boldsymbol{y}^k\right]^\top \boldsymbol{\xi} & \\
\text{subject to} & \boldsymbol{\xi} \in \mathbb{R}^Q & \\
& \boldsymbol{A}\boldsymbol{\xi} \leq \boldsymbol{b} & \\
& \boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k \leq \boldsymbol{H}\boldsymbol{\xi} & \forall k \in \mathcal{K} : \boldsymbol{\ell}_k = 0 \\
& [\boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k]_{\boldsymbol{\ell}_k} \geq [\boldsymbol{H}\boldsymbol{\xi}]_{\boldsymbol{\ell}_k} + \epsilon & \forall k \in \mathcal{K} : \boldsymbol{\ell}_k \neq 0
\end{array}
$$

does not exceed τ. Strong linear programming duality implies that this problem attains the same optimal value as its dual problem, which is given by

$$
\begin{array}{ll}
\text{minimize} & \boldsymbol{b}^\top\boldsymbol{\alpha} - \sum_{\substack{k\in\mathcal{K}:\\ \boldsymbol{\ell}_k=0}} (\boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k)^\top \boldsymbol{\beta}^k + \sum_{\substack{k\in\mathcal{K}:\\ \boldsymbol{\ell}_k\neq 0}} \left([\boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k]_{\boldsymbol{\ell}_k} - \epsilon\right)\gamma_k \\
\text{subject to} & \boldsymbol{\alpha} \in \mathbb{R}^R_+,\ \boldsymbol{\beta}^k \in \mathbb{R}^L_+,\ k \in \mathcal{K},\ \boldsymbol{\gamma} \in \mathbb{R}^K_+ \\
& \boldsymbol{A}^\top\boldsymbol{\alpha} - \sum_{\substack{k\in\mathcal{K}:\\ \boldsymbol{\ell}_k=0}} \boldsymbol{H}^\top\boldsymbol{\beta}^k + \sum_{\substack{k\in\mathcal{K}:\\ \boldsymbol{\ell}_k\neq 0}} \boldsymbol{H}_{\boldsymbol{\ell}_k}\gamma_k = \boldsymbol{C}\boldsymbol{x} + \sum_{k\in\mathcal{K}} \boldsymbol{\lambda}_k(\boldsymbol{\ell})\boldsymbol{Q}\boldsymbol{y}^k.
\end{array}
$$

Strong duality holds because the dual problem is always feasible. Indeed, one can show that the compactness of Ξ implies that $\{\boldsymbol{A}^\top\boldsymbol{\alpha} : \boldsymbol{\alpha} \geq \boldsymbol{0}\} = \mathbb{R}^Q$. Note that the first constraint set in problem (7) ensures that the optimal value of this dual problem does not exceed τ for all $\boldsymbol{\ell} \in \partial\mathcal{L}$.

The last constraint in (EC.6) is satisfied for $\boldsymbol{\ell} \in \mathcal{L}_+$ whenever the linear program

$$
\begin{array}{lll}
\text{maximize} & 0 & \\
\text{subject to} & \boldsymbol{\xi} \in \mathbb{R}^Q & \\
& \boldsymbol{A}\boldsymbol{\xi} \leq \boldsymbol{b} & \\
& [\boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k]_{\boldsymbol{\ell}_k} \geq [\boldsymbol{H}\boldsymbol{\xi}]_{\boldsymbol{\ell}_k} + \epsilon & \forall k \in \mathcal{K}
\end{array}
$$

is infeasible. Strong LP duality implies that this is the case whenever the dual problem,

$$
\begin{array}{ll}
\text{minimize} & \boldsymbol{b}^\top\boldsymbol{\alpha} + \sum_{\substack{k\in\mathcal{K}:\\ \boldsymbol{\ell}_k\neq 0}} \left([\boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k]_{\boldsymbol{\ell}_k} - \epsilon\right)\gamma_k \\
\text{subject to} & \boldsymbol{\alpha} \in \mathbb{R}^R_+,\ \boldsymbol{\gamma} \in \mathbb{R}^K_+ \\
& \boldsymbol{A}^\top\boldsymbol{\alpha} + \sum_{k\in\mathcal{K}} \boldsymbol{H}_{\boldsymbol{\ell}_k}\gamma_k = \boldsymbol{0},
\end{array}
$$

is unbounded. Note that strong duality holds because $\boldsymbol{\alpha} = \mathbf{0}$ and $\boldsymbol{\gamma} = \mathbf{0}$ are feasible in the dual problem. Since the feasible region of the dual problem constitutes a cone, the dual problem is unbounded if and only if it admits a feasible solution that achieves an objective value of -1 or less. This is precisely what the last constraint set in problem (7) requires. $\square$