Nurse Staffing under Absenteeism: A Distributionally Robust Optimization Approach

Minseok Ryu and Ruiwei Jiang

Appendix A: Proof of Lemma 1

Proof: We rewrite formulation (1a)–(1c) as

$$V(\tilde{w},\tilde{y},\tilde{d}) \;=\; \min_{z\in\mathbb{Z}_+^{I\times|P_i|},\, x\in\mathbb{Z}_+^J} \;\; \sum_{j=1}^{J} c_j^{\mathrm{x}} x_j$$

$$\text{s.t.}\;\; x_j + \sum_{i\in[I]:j\in P_i} z_{ij} \geq \tilde{d}_j - \tilde{w}_j, \;\; \forall j\in[J], \qquad -\sum_{j\in P_i} z_{ij} \geq -\tilde{y}_i, \;\; \forall i\in[I], \tag{18a}$$

We note that the constraint matrix of the above formulation is totally unimodular (TU), and so the conclusion follows. To see the TU property, we consider the following constraint matrix:

$$\begin{bmatrix} \left(x_j + \sum_{i\in[I]:j\in P_i} z_{ij}\right) \\ \left(-\sum_{j\in P_i} z_{ij}\right) \end{bmatrix}.$$

It follows that (a) each entry of this matrix is -1, 0, or 1, (b) this matrix has at most two nonzero entries in each column, and (c) the entries sum up to be zero for any column containing two nonzero entries. Hence, the constraint matrix is TU based on Proposition 2.6 in Nemhauser and Wolsey (1999). The conclusion follows because $\tilde{d}_j - \tilde{w}_j$ and $-\tilde{y}_i$ are integers for all $j\in[J]$ and for all $i\in[I]$, respectively. □

Appendix B: Verifying Assumption 1

We present necessary and sufficient conditions for Assumption 1 in the following proposition.

Proposition 7. *For any given w and y, $\mathcal{D}(w,y)$ is non-empty if and only if the following three conditions are satisfied:*

1. $f_j(w_j)\in[0,w_j]$ for all $j\in[J]$;

2. $g_i(y_i)\in[0,y_i]$ for all $i\in[I]$;

3. The following polyhedron defined for every $j\in[J]$ is nonempty:

$$\left\{p_j\in\mathbb{R}_+^{d_j^U-d_j^L+1} \;:\; \sum_{k=d_j^L}^{d_j^U} k^q p_{jk} = \mu_{jq}, \;\forall q\in[Q], \qquad \sum_{k=d_j^L}^{d_j^U} p_{jk} = 1\right\}. \tag{19}$$

Proof: (**Necessity**) Suppose that $\mathcal{D}(w,y)\neq\emptyset$. Then, there exists a $\mathbb{P}\in\mathcal{P}(\Xi)$ such that $\mathbb{E}_{\mathbb{P}}[\tilde{d}_j^q]=\mu_{jq}$ for all $j\in[J]$ and $q\in[Q]$, $\mathbb{E}_{\mathbb{P}}[\tilde{w}_j]=f_j(w_j)$ for all $j\in[J]$, and $\mathbb{E}_{\mathbb{P}}[\tilde{y}_i]=g_i(y_i)$ for all $i\in[I]$. It follows that, for all $j\in[J]$, we have $f_j(w_j)\leq \operatorname{esssup}_{\Xi}\{\tilde{w}_j\}\leq w_j$ and $f_j(w_j)\geq \operatorname{essinf}_{\Xi}\{\tilde{w}_j\}\geq 0$, leading to $f_j(w_j)\in[0,w_j]$. Likewise, it holds that $g_i(y_i)\in[0,y_i]$ for all $i\in[I]$. In addition, for all $j\in[J]$ and $k\in[d_j^{\mathrm{L}},d_j^{\mathrm{U}}]_{\mathbb{Z}}$, we let $\bar{p}_{jk}=\mathbb{P}\{\tilde{d}_j=k\}$. It follows that, for all $q\in[Q]$, $\sum_{k=d_j^{\mathrm{L}}}^{d_j^{\mathrm{U}}}\bar{p}_{jk}=\sum_{k=d_j^{\mathrm{L}}}^{d_j^{\mathrm{U}}}\mathbb{P}\{\tilde{d}_j=k\}=1$. Moreover,

$$\sum_{k=d_j^{\mathrm{L}}}^{d_j^{\mathrm{U}}} k^q\bar{p}_{jk} \;=\; \sum_{k=d_j^{\mathrm{L}}}^{d_j^{\mathrm{U}}} k^q\mathbb{P}\{\tilde{d}_j=k\} \;=\; \mathbb{E}_{\mathbb{P}}[\tilde{d}_j^q] \;=\; \mu_{jq}.$$

Hence, $\bar{p}_{jk}$ constitutes a feasible solution to the polyhedron (19) This holds for all $j \in [J]$ and proves the necessity of the three conditions.

(**Sufficiency**) Suppose that the three conditions are satisfied. For all $j \in [J]$, as $f_j(w_j) \in [0, w_j] \equiv \text{conv}([0, w_j]_{\mathbb{Z}})$ by condition 1, there exists a $\mathbb{P}_{\tilde{w}_j} \in \mathcal{P}([0, w_j]_{\mathbb{Z}})$ such that $f_j(w_j) = \mathbb{E}_{\mathbb{P}_{\tilde{w}_j}}[\tilde{w}_j]$. Likewise, for all $i \in [I]$, there exists a $\mathbb{P}_{\tilde{y}_i} \in \mathcal{P}([0, y_i]_{\mathbb{Z}})$ such that $g_i(y_i) = \mathbb{E}_{\mathbb{P}_{\tilde{y}_i}}[\tilde{y}_i]$. In addition, for all $j \in [J]$, there exist p_{jk} such that $\sum_{k=d_j^{\mathrm{L}}}^{d_j^{\mathrm{U}}} k^q p_{jk} = \mu_{jq}$ for all $q \in [Q]$ and $\sum_{k=d_j^{\mathrm{L}}}^{d_j^{\mathrm{U}}} p_{jk} = 1$. Defining $\mathbb{P}_{\tilde{d}_j} \in \mathcal{P}([d_j^{\mathrm{L}}, d_j^{\mathrm{U}}]_{\mathbb{Z}})$ such that $\mathbb{P}_{\tilde{d}_j}\{\tilde{d}_j = k\} = p_{jk}$ for all $k \in [d_j^{\mathrm{L}}, d_j^{\mathrm{U}}]_{\mathbb{Z}}$, we have $\mathbb{E}_{\mathbb{P}_{\tilde{d}_j}}[\tilde{d}_j^q] = \mu_{jq}$.

Therefore, the probability distribution

$$\mathbb{P} := \Pi_{j=1}^J \mathbb{P}_{\tilde{w}_j} \times \Pi_{i=1}^I \mathbb{P}_{\tilde{y}_i} \times \Pi_{j=1}^J \mathbb{P}_{\tilde{d}_j}$$

satisfies (2a)–(2b) and hence $\mathbb{P} \in \mathcal{D}(w, y)$. It follows that $\mathcal{D}(w, y) \neq \emptyset$ and the proof is completed. □

Appendix C: Proof of Proposition 1

Proof: First, denoting $\tilde{\xi} := (\tilde{w}, \tilde{y}, \tilde{d})$, we present $\max_{\mathbb{P} \in \mathcal{D}(w,y)} \mathbb{E}_{\mathbb{P}}[V(\tilde{w}, \tilde{y}, \tilde{d})]$ as the following optimization problem:

$$\begin{aligned}
\max_{p \geq 0} \quad & \sum_{\tilde{\xi} \in \Xi} p_{\tilde{\xi}} V(\tilde{\xi}) \\
\text{s.t.} \quad & \sum_{\tilde{\xi} \in \Xi} p_{\tilde{\xi}} \tilde{w}_j = f_j(w_j), \quad \forall j \in [J], && \text{(20a)} \\
& \sum_{\tilde{\xi} \in \Xi} p_{\tilde{\xi}} \tilde{y}_i = g_i(y_i), \quad \forall i \in [I], && \text{(20b)} \\
& \sum_{\tilde{\xi} \in \Xi} p_{\tilde{\xi}} \tilde{d}_j^q = \mu_{jq}, \quad \forall j \in [J], \ \forall q \in [Q], && \text{(20c)} \\
& \sum_{\tilde{\xi} \in \Xi} p_{\tilde{\xi}} = 1, && \text{(20d)}
\end{aligned}$$

where decision variables $p_{\tilde{\xi}}$ represent the probability of the random variables being realized as $\tilde{\xi}$, and constraints (20a)–(20d) describe the ambiguity set $\mathcal{D}(w, y)$ defined in (2a)–(2b). The dual of this formulation is

$$\begin{aligned}
\min_{\gamma, \lambda, \rho, \theta} \quad & \sum_{j=1}^J \sum_{q=1}^Q \mu_{jq} \rho_{jq} + \sum_{j=1}^J f_j(w_j) \gamma_j + \sum_{i=1}^I g_i(y_i) \lambda_i + \theta && \text{(21a)} \\
\text{s.t.} \quad & \theta + \sum_{j=1}^J \sum_{q=1}^Q \rho_{jq} \tilde{d}_j^q + \sum_{j=1}^J \gamma_j \tilde{w}_j + \sum_{i=1}^I \lambda_i \tilde{y}_i \geq V(\tilde{\xi}), \quad \forall \tilde{\xi} \in \Xi, && \text{(21b)}
\end{aligned}$$

where dual variables γ_j, λ_i, ρ_{jq}, and θ are associated with primal constraints (20a)–(20d), respectively, and dual constraints (21b) are associated with primal variables $p_{\tilde{\xi}}$. By Assumption 1, strong duality holds between the primal and dual formulations because they are both linear programs. As the objective function aims to minimize the value of θ, we observe by constraints (21b) that $\theta = \max_{\tilde{\xi} \in \Xi}\{V(\tilde{\xi}) - \sum_{j=1}^J \sum_{q=1}^Q \rho_{jq} \tilde{d}_j^q - \sum_{j=1}^J \gamma_j \tilde{w}_j - \sum_{i=1}^I \lambda_i \tilde{y}_i\}$. Hence, $\max_{\mathbb{P} \in \mathcal{D}(w,y)} \mathbb{E}_{\mathbb{P}}[V(\tilde{\xi})]$ equals the optimal value of the following min-max optimization problem:

$$\min_{\gamma, \lambda, \rho} \max_{\tilde{\xi} \in \Xi} \left\{ V(\tilde{\xi}) - \sum_{j=1}^J \left[\sum_{q=1}^Q \rho_{jq} \tilde{d}_j^q + \gamma_j \tilde{w}_j \right] - \sum_{i=1}^I \lambda_i \tilde{y}_i \right\} + \sum_{j=1}^J \left[\sum_{q=1}^Q \mu_{jq} \rho_{jq} + f_j(w_j) \gamma_j \right] + \sum_{i=1}^I g_i(y_i) \lambda_i. \quad \text{(22a)}$$

Second, in view of the dual formulation (4a)–(4c) of $V(\tilde{\xi})$, we rewrite the maximum term in (22a) as

$$\begin{aligned}
&\max_{(\tilde{w},\tilde{y},\tilde{d})\in\Xi}\ \max_{(\alpha,\beta)\in\Lambda}\left\{\sum_{j=1}^{J}(\tilde{d}_j-\tilde{w}_j)\alpha_j+\sum_{i=1}^{I}\tilde{y}_i\beta_i-\sum_{j=1}^{J}\left[\sum_{q=1}^{Q}\rho_{jq}\tilde{d}_j^q+\gamma_j\tilde{w}_j\right]-\sum_{i=1}^{I}\lambda_i\tilde{y}_i\right\}\\
=\ &\max_{(\alpha,\beta)\in\Lambda}\ \max_{(\tilde{w},\tilde{y},\tilde{d})\in\Xi}\left\{\sum_{j=1}^{J}(\tilde{d}_j-\tilde{w}_j)\alpha_j+\sum_{i=1}^{I}\tilde{y}_i\beta_i-\sum_{j=1}^{J}\left[\sum_{q=1}^{Q}\rho_{jq}\tilde{d}_j^q+\gamma_j\tilde{w}_j\right]-\sum_{i=1}^{I}\lambda_i\tilde{y}_i\right\}\\
=\ &\max_{(\alpha,\beta)\in\Lambda}\left\{\sum_{j=1}^{J}\max_{\tilde{w}_j\in[0,w_j]_{\mathbb{Z}}}\Big\{(-\alpha_j-\gamma_j)\tilde{w}_j\Big\}+\sum_{i=1}^{I}\max_{\tilde{y}_i\in[0,y_i]_{\mathbb{Z}}}\Big\{(\beta_i-\lambda_i)\tilde{y}_i\Big\}+\sum_{j=1}^{J}\max_{\tilde{d}_j\in[d_j^{\mathrm{L}},d_j^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{\alpha_j\tilde{d}_j-\sum_{q=1}^{Q}\rho_{jq}\tilde{d}_j^q\Big\}\right\}.
\end{aligned}$$

Finally, as $(-\alpha_j-\gamma_j)\tilde{w}_j$ is linear in $\tilde{w}_j$, we have

$$\max_{\tilde{w}_j\in[0,w_j]_{\mathbb{Z}}}\Big\{(-\alpha_j-\gamma_j)\tilde{w}_j\Big\}=\max\Big\{0,(-\alpha_j-\gamma_j)w_j\Big\}=\Big[(-\alpha_j-\gamma_j)w_j\Big]_+.$$

Similarly, we have $\max_{\tilde{y}_i\in[0,y_i]_{\mathbb{Z}}}\{(\beta_i-\lambda_i)\tilde{y}_i\}=[(\beta_i-\lambda_i)y_i]_+$. This completes the proof. □

Appendix D: Proof of Lemma 2

Proof: As $\max_{(\alpha,\beta)\in\Lambda}F(\alpha,\beta)$ is to maximize a convex function over a polyhedron, we only need to analyze the extreme directions and extreme points of Λ.

First, the extreme directions of Λ are $(\alpha,\beta)=(0,-e_i)$ for all $i\in[I]$, where e_i represents the i^{th} standard basis vector. As $\tilde{y}_i\geq 0$, moving along any of these extreme directions (i.e., decreasing the value of any β_i) does not increase the value of $F(\alpha,\beta)$. Hence, we can omit these extreme directions in the attempt of maximizing $F(\alpha,\beta)$ and accordingly $\bar{\beta}_i=\min\{-\bar{\alpha}_j: j\in P_i\}=-\max\{\bar{\alpha}_j: j\in P_i\}$ without loss of optimality. This proves property (b) in the claim. In addition, there exists an extreme point of Λ that is optimal to $\max_{(\alpha,\beta)\in\Lambda}F(\alpha,\beta)$.

Second, we prove, by contradiction, that each extreme point of Λ satisfies property (a) in the claim. Suppose that there exists an extreme point $(\bar{\alpha},\bar{\beta})$ such that property (a) fails, i.e., $\bar{\alpha}_{j^*}\notin\{0,c_1^{\mathrm{x}},\ldots,c_{j^*}^{\mathrm{x}}\}$ for some $j^*\in[J]$. Consider the set $\mathcal{I}(j^*):=\{i\in[I]:-\bar{\beta}_i=\bar{\alpha}_{j^*}\}$. We discuss the following two cases. In each case, we shall construct two points in Λ such that their midpoint is $(\bar{\alpha},\bar{\beta})$, which provides a desired contradiction.

1. If $\mathcal{I}(j^*)=\emptyset$, then $-\bar{\beta}_i>\bar{\alpha}_{j^*}$ for all i such that $j^*\in P_i$. Defining $\epsilon:=(1/2)\min\Big\{-\bar{\beta}_i-\bar{\alpha}_{j^*},\forall i\in[I]: j^*\in P_i,\ \bar{\alpha}_{j^*},\ \min\big\{|\bar{\alpha}_{j^*}-c_\ell^{\mathrm{x}}|:\ell\in[j^*]\big\}\Big\}>0$, we construct two points $(\bar{\alpha}^+,\bar{\beta})$ and $(\bar{\alpha}^-,\bar{\beta})$ such that $\bar{\alpha}_{j^*}^+=\bar{\alpha}_{j^*}+\epsilon$, $\bar{\alpha}_{j^*}^-=\bar{\alpha}_{j^*}-\epsilon$, and $\bar{\alpha}_j^+=\bar{\alpha}_j^-=\bar{\alpha}_j$ for all $j\neq j^*$. Then, it is clear that $(\bar{\alpha}^+,\bar{\beta}),(\bar{\alpha}^-,\bar{\beta})\in\Lambda$. But $(\bar{\alpha},\bar{\beta})=(1/2)(\bar{\alpha}^+,\bar{\beta})+(1/2)(\bar{\alpha}^-,\bar{\beta})$, which contradicts the fact that $(\bar{\alpha},\bar{\beta})$ is an extreme point of Λ.
2. If $\mathcal{I}(j^*)\neq\emptyset$, then we define $\mathcal{J}(j^*):=\bigcup_{i\in\mathcal{I}(j^*)}\{j\in P_i:\bar{\alpha}_j=-\bar{\beta}_i\}$. It follows that $\bar{\alpha}_j=\bar{\alpha}_{j^*}$ for all $j\in\mathcal{J}(j^*)$. Hence, for each $i\in\mathcal{I}(j^*)$, $\bar{\alpha}_j=\bar{\alpha}_{j^*}$ for all $j\in P_i\cap\mathcal{J}(j^*)$ and $\bar{\alpha}_j<\bar{\alpha}_{j^*}$ for all $j\in P_i\setminus\mathcal{J}(j^*)$. We define $\epsilon:=(1/2)\min\Big\{\min\big\{\bar{\alpha}_{j^*}-\bar{\alpha}_j: i\in\mathcal{I}(j^*), j\in P_i\setminus\mathcal{J}(j^*)\big\},\ \min\big\{-\bar{\beta}_i-\bar{\alpha}_{j^*}: i\notin\mathcal{I}(j^*),\ -\bar{\beta}_i>\bar{\alpha}_{j^*}\big\},\ \bar{\alpha}_{j^*},\ \min\big\{|\bar{\alpha}_{j^*}-c_\ell^{\mathrm{x}}|:\ell\in[j^*]\big\}\Big\}$. Then $\epsilon>0$ because it is the minimum of a finite number of positive reals. Here we adopt the convention that $\min\{a:a\in A\}=\infty$ if $A=\emptyset$. For example, if there does not exist an $i\notin\mathcal{I}(j^*)$ such that $-\bar{\beta}_i>\bar{\alpha}_{j^*}$, then $\min\{-\bar{\beta}_i-\bar{\alpha}_{j^*}: i\notin\mathcal{I}(j^*),\ -\bar{\beta}_i>\bar{\alpha}_{j^*}\}=\infty$. We construct two points $(\bar{\alpha}^+,\bar{\beta}^+)$ and $(\bar{\alpha}^-,\bar{\beta}^-)$ such that

$$\bar{\alpha}_j^+=\begin{cases}\bar{\alpha}_{j^*}+\epsilon & \forall j\in\mathcal{J}(j^*)\\ \bar{\alpha}_j & \text{otherwise}\end{cases},\quad \bar{\alpha}_j^-=\begin{cases}\bar{\alpha}_{j^*}-\epsilon & \forall j\in\mathcal{J}(j^*)\\ \bar{\alpha}_j & \text{otherwise}\end{cases},$$

$$\bar{\beta}_i^+ = \begin{cases} -(\bar{\alpha}_{j^*} + \epsilon) & \forall i \in \mathcal{I}(j^*) \\ \bar{\beta}_i & \text{otherwise} \end{cases}, \quad \bar{\beta}_i^- = \begin{cases} -(\bar{\alpha}_{j^*} - \epsilon) & \forall i \in \mathcal{I}(j^*) \\ \bar{\beta}_i & \text{otherwise} \end{cases}.$$

It is clear that $(\bar{\alpha}, \bar{\beta}) = (1/2)(\bar{\alpha}^+, \bar{\beta}^+) + (1/2)(\bar{\alpha}^-, \bar{\beta}^-)$. To finish the proof, it remains to show that $(\bar{\alpha}^+, \bar{\beta}^+), (\bar{\alpha}^-, \bar{\beta}^-) \in \Lambda$. To see this, we check constraints (4b) and (4c). For constraints (4c), we have $\bar{\alpha}_j^+, \bar{\alpha}_j^- \in (0, c_j^{\mathrm{x}})$ for all $j \in \mathcal{J}(j^*)$ by the definition of ϵ. Additionally, for all $j \notin \mathcal{J}(j^*)$, we have $\bar{\alpha}_j^+ = \bar{\alpha}_j^- = \bar{\alpha}_j \in [0, c_j^{\mathrm{x}}]$. Hence, constraints (4c) are indeed satisfied and it remains to check constraints (4b). For each $i \in \mathcal{I}(j^*)$, $-\bar{\beta}_i^+ = \bar{\alpha}_{j^*} + \epsilon = \bar{\alpha}_j^+$ for all $j \in P_i \cap \mathcal{J}(j^*)$ and $-\bar{\beta}_i^+ = \bar{\alpha}_{j^*} + \epsilon \geq \bar{\alpha}_{j^*} \geq \bar{\alpha}_j = \bar{\alpha}_j^+$ for all $j \in P_i \setminus \mathcal{J}(j^*)$, where the first inequality is because $\epsilon > 0$, and the second inequality follows from the definition of $\mathcal{J}(j^*)$. Meanwhile, $-\bar{\beta}_i^- = \bar{\alpha}_{j^*} - \epsilon = \bar{\alpha}_j^-$ for all $j \in P_i \cap \mathcal{J}(j^*)$, and $-\bar{\beta}_i^- = \bar{\alpha}_{j^*} - \epsilon \geq \bar{\alpha}_j = \bar{\alpha}_j^-$ for all $j \in P_i \setminus \mathcal{J}(j^*)$, where the inequality follows from the definition of ϵ and the last equality is because $j \notin \mathcal{J}(j^*)$. It follows that constraints (4b) are indeed satisfied for all $i \in \mathcal{I}(j^*)$. For each $i \notin \mathcal{I}(j^*)$, $\bar{\beta}_i^+ = \bar{\beta}_i^- = \bar{\beta}_i$ and $-\bar{\beta}_i \neq \bar{\alpha}_{j^*}$. We discuss the following two sub-cases to complete the proof.

(a) If $-\bar{\beta}_i > \bar{\alpha}_{j^*}$, then $-\bar{\beta}_i^+ = -\bar{\beta}_i \geq \bar{\alpha}_{j^*} + \epsilon \geq \bar{\alpha}_j^+$, where the first inequality follows from the definition of ϵ. In addition, by construction $-\bar{\beta}_i^- = -\bar{\beta}_i > \bar{\alpha}_{j^*} \geq \bar{\alpha}_j^-$ for all $j \in P_i$.

(b) If $-\bar{\beta}_i < \bar{\alpha}_{j^*}$, then $j \notin \mathcal{J}(j^*)$ for all $j \in P_i$ because otherwise $-\bar{\beta}_i \geq \bar{\alpha}_j = \bar{\alpha}_{j^*}$. It follows that $\bar{\alpha}_j^+ = \bar{\alpha}_j^- = \bar{\alpha}_j$ and so $-\bar{\beta}_i^+ = -\bar{\beta}_i \geq \bar{\alpha}_j = \bar{\alpha}_j^+$ and $-\bar{\beta}_i^- = -\bar{\beta}_i \geq \bar{\alpha}_j = \bar{\alpha}_j^-$. □

Appendix E: Proof of Theorem 1

Proof: First, pick any $(\alpha, \beta) \in \Lambda$ that satisfies the optimality conditions (a)–(b) stated in Lemma 2. We shall show that there exists a feasible solution (t, s, r, p) to formulation (7) that attains the same objective function value as $F(\alpha, \beta)$.

To this end, for all $j \in [J]$ and $k \in [j]$, we let $t_{jk} = 1$ if $\alpha_j = c_k^{\mathrm{x}}$, and $t_{jk} = 0$ otherwise. In addition, for all $i \in [I]$, if $\alpha_j = 0$ for all $j \in P_i$, then we let $s_{ik} = 0$ for all $k \in [J(i)]$; and otherwise, i.e. there exists a $k^* \in [J(i)]$ such that $\max\{\alpha_j : j \in P_i\} = c_{k^*}^{\mathrm{x}}$, we let $s_{ik^*} = 1$ and all other $s_{ik} = 0$. Also, we define r and p as in (7b). By construction, (t, s, r, p) satisfies (7b). It follows that the objective function value of (t, s, r, p) equals

$$\begin{aligned}
&\sum_{j=1}^{J} \left(c_j^{\mathrm{r}} r_j + \sum_{k=1}^{j} c_{jk}^{\mathrm{t}} t_{jk} \right) + \sum_{i=1}^{I} \left(c_i^{\mathrm{p}} p_i + \sum_{k=1}^{J(i)} c_{ik}^{\mathrm{s}} s_{ik} \right) \\
&= \sum_{j=1}^{J} \Bigg[\mathbb{1}_{\{0\}}(\alpha_j) \left\{ [-\gamma_j w_j]_+ + \max_{\tilde{d}_j \in [d_j^{\mathrm{L}}, d_j^{\mathrm{U}}]_{\mathbb{Z}}} \left\{ -\sum_{q=1}^{Q} \rho_{jq} \tilde{d}_j^q \right\} \right\} + \\
&\quad \sum_{k=1}^{j} \mathbb{1}_{\{c_k^{\mathrm{x}}\}}(\alpha_j) \left\{ [(-c_k^{\mathrm{x}} - \gamma_j) w_j]_+ + \max_{\tilde{d}_j \in [d_j^{\mathrm{L}}, d_j^{\mathrm{U}}]_{\mathbb{Z}}} \left\{ c_k^{\mathrm{x}} \tilde{d}_j - \sum_{q=1}^{Q} \rho_{jq} \tilde{d}_j^q \right\} \right\} \Bigg] + \\
&\quad \sum_{i=1}^{I} \left[\left\{ \mathbb{1}_{\{0\}} \left(\max_{j \in P_i} \{\alpha_j\} \right) [-\lambda_i y_i]_+ \right\} + \sum_{k=1}^{J(i)} \left\{ \mathbb{1}_{\{c_k^{\mathrm{x}}\}} \left(\max_{j \in P_i} \{\alpha_j\} \right) [(-c_k^{\mathrm{x}} - \lambda_i) y_i]_+ \right\} \right] \\
&= \sum_{j=1}^{J} \left\{ \Big[(-\alpha_j - \gamma_j) w_j \Big]_+ + \max_{\tilde{d}_j \in [d_j^{\mathrm{L}}, d_j^{\mathrm{U}}]_{\mathbb{Z}}} \left\{ \alpha_j \tilde{d}_j - \sum_{q=1}^{Q} \rho_{jq} \tilde{d}_j^q \right\} \right\} + \sum_{i=1}^{I} \Big[(\beta_i - \lambda_i) y_i \Big]_+ = F(\alpha, \beta),
\end{aligned}$$

where the first equality follows from the definition of (t, s, r, p) and the second equality follows from the optimality conditions stated in Lemma 2.

Second, pick any feasible solution (t,s,r,p) to formulation (7). We construct an $(\alpha,\beta)\in\Lambda$ that satisfies the optimality conditions (a)–(b) in Lemma 2 and $F(\alpha,\beta)$ equals the objective function value (7a) of (t,s,r,p). Specifically, for all $j\in[J]$, we let $\alpha_j=\sum_{k=1}^{j}c_k^{\mathrm{x}}t_{jk}$ and, for all $i\in[I]$, $\beta_i=-\sum_{k=1}^{J(i)}c_k^{\mathrm{x}}s_{ik}$. Then, for all $i\in[I]$ and $j\in P_i$,

$$\begin{aligned}\beta_i+\alpha_j&=-\sum_{k=1}^{J(i)}c_k^{\mathrm{x}}s_{ik}+\sum_{k=1}^{j}c_k^{\mathrm{x}}t_{jk}\le-\sum_{k=1}^{J(i)}c_k^{\mathrm{x}}\Big(\sum_{\ell\in P_i:\ell\ge k}t_{\ell k}\Big)+\sum_{k=1}^{j}c_k^{\mathrm{x}}t_{jk}\\&=-\sum_{\ell\in P_i}\sum_{k=1}^{\ell}c_k^{\mathrm{x}}t_{\ell k}+\sum_{k=1}^{j}c_k^{\mathrm{x}}t_{jk}=-\sum_{\ell\in P_i\setminus\{j\}}\sum_{k=1}^{\ell}c_k^{\mathrm{x}}t_{\ell k}\le 0,\end{aligned}$$

where the first inequality is due to constraints (6c), and the last inequality is due to $c_k^{\mathrm{x}},t_{\ell k}\ge 0$. Next, we have $\alpha_j\in\{0,c_1^{\mathrm{x}},\ldots,c_j^{\mathrm{x}}\}$ for all $j\in[J]$ due to constraints (7b), and $\beta_i\in\{0,-c_1^{\mathrm{x}},\ldots,-c_{J(i)}^{\mathrm{x}}\}$ for all $i\in[I]$. Hence, $(\alpha,\beta)\in\Lambda$ and satisfies optimality condition (a). It remains to show that the constructed (α,β) satisfies optimality condition (b).

For all $i\in[I]$, if $\sum_{k=1}^{J(i)}s_{ik}=0$, i.e., $\beta_i=0$, then $t_{jk}=0$ for all $j\in P_i$ and $k\in[j]$ due to constraints (6e). It follows that $\alpha_j=0$ for all $j\in P_i$ and so $\beta_i=-\max\{\alpha_j:j\in P_i\}=0$. On the other hand, if $\sum_{k=1}^{J(i)}s_{ik}=1$, there exists a $k^*\in[J(i)]$ with $s_{ik^*}=1$, i.e., $\beta_i=-c_{k^*}^{\mathrm{x}}$. By constraints (6c) and (6d), $\sum_{j\in P_i:j\ge k^*}t_{jk^*}\ge 1$ and $t_{j\ell}=0$ for all $j\in P_i:j\ge k^*+1$ and $\ell\in[k^*+1,j]_{\mathbb{Z}}$. It follows that $\max\{\alpha_j:j\in P_i\}=c_{k^*}^{\mathrm{x}}$ and so $\beta_i=-\max\{\alpha_j:j\in P_i\}=-c_{k^*}^{\mathrm{x}}$. Hence, (α,β) satisfies the optimality conditions (b). Finally,

$$\begin{aligned}F(\alpha,\beta)=&\sum_{j=1}^{J}\left\{\Big[\Big(-\sum_{k=1}^{j}c_k^{\mathrm{x}}t_{jk}-\gamma_j\Big)w_j\Big]_+ +\max_{\tilde d_j\in[d_j^{\mathrm{L}},d_j^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{\Big(\sum_{k=1}^{j}c_k^{\mathrm{x}}t_{jk}\Big)\tilde d_j-\sum_{q=1}^{Q}\rho_{jq}\tilde d_j^{q}\Big\}\right\}\\&+\sum_{i=1}^{I}\Big[\Big(-\sum_{k=1}^{J(i)}c_k^{\mathrm{x}}s_{ik}-\lambda_i\Big)y_i\Big]_+\\=&\sum_{j=1}^{J}\Big(c_j^{\mathrm{r}}r_j+\sum_{k=1}^{j}c_{jk}^{\mathrm{t}}t_{jk}\Big)+\sum_{i=1}^{I}\Big(c_i^{\mathrm{p}}p_i+\sum_{k=1}^{J(i)}c_{ik}^{\mathrm{s}}s_{ik}\Big)\end{aligned}$$

by the definition of (α,β) and constraints (7b). This completes the proof. □

Appendix F: Proof of Proposition 2

Proof: Let $G(\gamma,\lambda,\rho)$ be the objective function of problem (5a)–(5b), $(\hat\gamma,\hat\lambda,\hat\rho)$ be any feasible solution, and S^* be the set of optimal solution to problem $\max_{(\alpha,\beta)\in\Lambda}F(\alpha,\beta)$ for the given $(\hat\gamma,\hat\lambda,\hat\rho)$. Suppose that there exists a $j\in[J]$ such that $\hat\gamma_j<-c_j^{\mathrm{x}}$. Then, $-\hat\gamma_j-\alpha_j^*>0$ and $[(-\hat\gamma_j-\alpha_j^*)w_j]_+=(-\hat\gamma_j-\alpha_j^*)w_j$ for all $(\alpha^*,\beta^*)\in S^*$ because $\alpha_j^*\le c_j^{\mathrm{x}}$ by (4c). Additionally, due to Lemma 2, we can replace polyhedron Λ by the (compact) set of its extreme points $\mathrm{ex}(\Lambda)$ without loss of optimality, i.e., $\max_{(\alpha,\beta)\in\Lambda}F(\alpha,\beta)=\max_{(\alpha,\beta)\in\mathrm{ex}(\Lambda)}F(\alpha,\beta)$. It then follows from Theorem 2.87 in Ruszczyński (2006) that, for all subgradient $\varpi\in\partial G(\hat\gamma,\hat\lambda,\hat\rho)$, the entry in ϖ with regard to variable γ_j at $(\hat\gamma,\hat\lambda,\hat\rho)$ equals $f_j(w_j)-w_j$, i.e., $\varpi(\gamma_j)|_{(\hat\gamma,\hat\lambda,\hat\rho)}=f_j(w_j)-w_j\le 0$. Noting that $\varpi(\gamma_j)|_{(\hat\gamma,\hat\lambda,\hat\rho)}\le 0$ holds valid whenever $\hat\gamma_j<-c_j^{\mathrm{x}}$, we can increase $\hat\gamma_j$ to $-c_j^{\mathrm{x}}$ without any loss of optimality.

Now suppose that $\hat\gamma_j>0$. Then, we have $-\hat\gamma_j-\alpha_j^*<0$ and $[(-\hat\gamma_j-\alpha_j^*)w_j]_+=0$ for all $(\alpha^*,\beta^*)\in S^*$ because $\alpha_j^*\ge 0$ by (4c). It follows from a similar implication as in the previous paragraph that, for all subgradient $\varpi\in\partial G(\hat\gamma,\hat\lambda,\hat\rho)$, we have $\varpi(\gamma_j)|_{(\hat\gamma,\hat\lambda,\hat\rho)}=f_j(w_j)\ge 0$. Noting that this holds valid whenever $\hat\gamma_j>0$, we can decrease $\hat{\gamma}_j$ to 0 without any loss of optimality. Therefore, there exists an optimal solution $(\gamma^*, \lambda^*, \rho^*)$ to problem (5a)–(5b) such that $\gamma_j^* \in [-c_j^x, 0]$ for all $j \in [J]$.

Following a similar proof, we can show that there exists an optimal solution $(\gamma^*, \lambda^*, \rho^*)$ to problem (5a)–(5b) such that $\lambda_i^* \in [-c_{J(i)}^x, 0]$ for all $i \in [I]$. We omit the details for the sake of saving space. □

Appendix G: Proof of Theorem 3

Proof: In each iteration of Algorithm 1, we solve a relaxation of the (DRNS) reformulation (9a)–(9d). It follows that, if the algorithm stops in an iteration and returns a solution (u^*, v^*) then (u^*, v^*) satisfies all the constraints (9c) because of Step 5. Then, (u^*, v^*) is feasible to formulation (9a)–(9d) and meanwhile optimal to its relaxation. Hence, (u^*, v^*) is optimal to formulation (9a)–(9d), i.e., optimal to (DRNS).

It remains to show that Algorithm 1 stops in a finite number of iterations. To see this, we notice that the set $\mathcal{H}$ contains a finite number of elements. Indeed, binary variables t and s only have a finite number of possible values. Although r and p are continuous variables, they also only have a finite number of possible values due to constraints (7b). □

Appendix H: Proof of Lemma 3

Proof: Pick any $i \in [I]$. We note that $\sum_{\ell=1}^{J(i)} s_{i\ell} \in \{0, 1\}$ due to constraints (6b) and discuss the following three cases. First, if $\sum_{\ell=1}^{J(i)} s_{i\ell} = 0$, i.e., $s_{i\ell} = 0$ for all $\ell \in [J(i)]$, then $t_{jk} = 0$ for all $j \in P_i$ and $k \in [j]$ by constraints (6e). In this case, inequalities (10) are valid because both left-hand and right-hand sides of (10) are 0. Second, if $\sum_{\ell=1}^{J(i)} s_{i\ell} = 1$ and $\sum_{\ell=k}^{J(i)} s_{i\ell} = 1$, then inequalities (10) reduce to $\sum_{\ell=k}^{j} t_{j\ell} \leq 1$ for all $j \in P_i$ with $j \geq k$, which are valid due to constraints (6a). Third, if $\sum_{\ell=1}^{J(i)} s_{i\ell} = 1$ and $\sum_{\ell=k}^{J(i)} s_{i\ell} = 0$, then there exists a $k^* < k$ such that $s_{ik^*} = 1$. By constraints (6d), $t_{j\ell} = 0$ for all $j \in P_i$ with $j \geq k^* + 1$ and for all $\ell \in [k^* + 1, j]_{\mathbb{Z}}$. It follows that both left-hand and right-hand sides of (10) are 0, and so inequalities (10) are valid.

Finally, inequalities (10) imply constraints (6e) because

$$t_{jk} \leq \sum_{\ell=k}^{j} t_{j\ell} \leq \sum_{\ell=k}^{J(i)} s_{i\ell} \leq \sum_{\ell=1}^{J(i)} s_{i\ell}, \ \forall i \in [I], \forall j \in P_i, \forall k \in [j].$$

□

Appendix I: Proof of Proposition 3

We first show that $\overline{\mathcal{H}}_1 \subseteq \text{conv}(\mathcal{H}_1)$. To prove this, we pick any fractional solution $(\hat{s}, \hat{t})$ that satisfies constraints (11b)–(11d), (12a), and (12b) in $\overline{\mathcal{H}}_1$, and find a finite number of points $\{(s^1, t^1), \ldots, (s^n, t^n)\}$ such that: (i) each (s^i, t^i) is binary-valued and satisfies constraints (11b)–(11g) in $\mathcal{H}_1$ and (ii) their convex combination produces $(\hat{s}, \hat{t})$, i.e., we find nonnegative weights $\{\pi^i : i \in [n]\}$ with $\sum_{i=1}^n \pi^i = 1$ such that $\hat{s}_k = \sum_{i=1}^n \pi^i s_k^i$ for all $k \in [J]$ and $\hat{t}_{jk} = \sum_{i=1}^n \pi^i t_{jk}^i$ for all $j \in [J]$ and $k \in [j]$.

To this end, we develop Algorithm 2, which iteratively constructs binary points $\{(s^i, t^i) : i \in [n]\}$ and weights $\{\pi^i : i \in [n]\}$. This algorithm initializes by setting an incumbent $(\hat{s}^1, \hat{t}^1) := (\hat{s}, \hat{t})$ and, in each iteration $i \in [n]$, update the incumbent $(\hat{s}^{i+1}, \hat{t}^{i+1}) := (\hat{s}^i - \pi^i s^i, \ \hat{t}^i - \pi^i t^i)$ as it constructs a new binary point (s^i, t^i)

with weight π^i, till the incumbent becomes $(0,0)$. For ease of exposition, we assume in Algorithm 2 that optimizing over an empty set yields zero, i.e., $\max\{f(x): x \in \emptyset\} = 0$. In addition, we provide an example of implementing Algorithm 2 at the end of this section. We show the correctness of Algorithm 2 by proving its properties in Theorem 6. Specifically, Theorem 6 shows that each point (s^i, t^i) found by Algorithm 2 lies in $\mathcal{H}_1$ and has a weight $\pi^i \in [0,1]$. In addition, after finding all n points and their weights, the incumbent reduces to zero, that is, $(\hat{s}, \hat{t}) - \sum_{i=1}^n \pi^i \cdot (s^i, t^i) = (0,0)$.

THEOREM 6. *For each binary point (s^i, t^i) produced in Algorithm 2, the following properties hold:*

1. *(s^i, t^i) satisfies constraints (11b)–(11g) in $\mathcal{H}_1$.*
2. *In line 6, the weight $\pi^i \in (0,1]$.*
3. *In line 7, the updated fractional solution $(\hat{s}^{i+1}, \hat{t}^{i+1})$ satisfies constraints (11b)–(11d), (12a), and (12b) in $\overline{\mathcal{H}}_1$.*

In a finite number of iterations, Algorithm 2 terminates with the following properties:

4. *After we execute line 7 for the last time, $\hat{s}^n_k = 0$ for all $k \in [J]$ and $\hat{t}^n_{jk} = 0$ for all $j \in [J]$ and $k \in [j]$.*
5. *$\pi^n \in [0,1)$.*

Therefore, $(\hat{s}, \hat{t})$ is a convex combination of $\{(s^1, t^1), \dots, (s^n, t^n)\}$ with weights $\pi^1, \dots, \pi^n$, respectively.

Proof: See Appendix J. □

It follows from Theorem 6 that $\overline{\mathcal{H}}_1 \subseteq \text{conv}(\mathcal{H}_1)$. On the other hand, $\text{conv}(\mathcal{H}_1) \subseteq \overline{\mathcal{H}}_1$ because $\mathcal{H}_1 \subseteq \overline{\mathcal{H}}_1$ by construction. Therefore, $\overline{\mathcal{H}}_1 = \text{conv}(\mathcal{H}_1)$ and this completes the proof. □

Appendix J: Proof of Theorem 6

Proof: We first prove Property 3 and then use it to prove all other properties.

(Property 3) We prove by mathematical induction. It is clear that $(\hat{s}^1, \hat{t}^1) \equiv (\hat{s}, \hat{t})$ satisfies Property 3. It remains to show that $(\hat{s}^{i+1}, \hat{t}^{i+1})$ satisfies constraints (11b)–(11d), (12a), and (12b) under the assumption that they are satisfied by $(\hat{s}^i, \hat{t}^i)$. Before getting into details, we rewrite the construction of $(\hat{s}^{i+1}, \hat{t}^{i+1})$ based on line 5 as follows:

$$\hat{s}^{i+1}_k = \begin{cases} \hat{s}^i_k & \text{if } k \in [m-1] \\ \hat{s}^i_m - \pi^i & \text{if } k = m \\ \hat{s}^i_k & \text{if } k \in [m+1, J]_{\mathbb{Z}} \end{cases}, \quad \hat{t}^{i+1}_{jk} = \begin{cases} \hat{t}^i_{jk} & \text{if } k \in [m-1], j \in [k, J]_{\mathbb{Z}} : k < \text{Index(j)} \\ \hat{t}^i_{jk} - \pi^i & \text{if } k \in [m-1], j \in [k, J]_{\mathbb{Z}} : k = \text{Index(j)} \\ 0 & \text{if } k \in [m-1], j \in [k, J]_{\mathbb{Z}} : k > \text{Index(j)} \\ \hat{t}^i_{jm} - \pi^i & \text{if } k = m, j \in A_m \\ \hat{t}^i_{jm} & \text{if } k = m, j \in B_m \\ \hat{t}^i_{jk} & \text{if } k \in [m+1, J]_{\mathbb{Z}}, j \in [k, J]_{\mathbb{Z}} \end{cases}. \tag{24}$$

First, we consider constraints (12b). We notice that $\hat{s}^{i+1}_k \geq 0$ for all $k \in [J]$ because $\pi^i \leq \hat{s}^i_m$ by construction. Likewise, $\hat{t}^{i+1}_{jk} \geq 0$ for all $j \in [J]$ and $k \in [j]$.

Second, we consider constraint (11c). By (24), we have $\sum_{k=1}^J \hat{s}^{i+1}_k = \sum_{k=1}^J \hat{s}^i_k - \pi^i \leq \sum_{k=1}^J \hat{s}^i_k \leq 1$ because $\pi^i \geq 0$ by construction.

Third, we consider constraints (11b). By (24), for all $j \in [J]$, we have $\sum_{k=1}^j \hat{t}^{i+1}_{jk} = \sum_{k=1}^j \hat{t}^i_{jk} - \mathbb{1}(\text{Index}(j) \neq 0)\pi^i \leq \sum_{k=1}^j \hat{t}^i_{jk} \leq 1$ because $\pi^i \geq 0$ by construction.

Algorithm 2 Finding binary points $\{(s^1,t^1),\dots,(s^n,t^n)\}$ and their weights $\{\pi^1,\dots,\pi^n\}$

Initialization: $i=1$, and $\hat{s}^1_k=\hat{s}_k,\ \forall k\in[J],\ \hat{t}^1_{jk}=\hat{t}_{jk},\ \forall j\in[J],\forall k\in[j]$.

for $m=1:J$ **do**

  while $\hat{s}^i_m>0$ **do**

    Find:

$$j^*=\min\left\{j\in[m,J]_{\mathbb{Z}}:\ \hat{t}^i_{jm}>0\right\},$$

$$A_m=\{j^*\}\cup\left\{j\in[m,J]_{\mathbb{Z}}:\ \sum_{\ell=m}^{j}\hat{t}^i_{j\ell}=\sum_{\ell=m}^{J}\hat{s}^i_\ell\right\},\quad B_m=[m,J]_{\mathbb{Z}}\setminus A_m,$$

$$\text{Index}(j)=\begin{cases} m & \forall j\in A_m,\\ \max\left\{k\in[\min\{j,m-1\}]:\hat{t}^i_{jk}>0\right\} & \forall j\in B_m\cup[m-1].\end{cases}$$

    Construct a binary point (s^i,t^i)**:**

$$s^i_k=\begin{cases}1 & \text{if } k=m\\ 0 & \text{otherwise}\end{cases},\ \forall k\in[J],\quad \text{and}\quad t^i_{jk}=\begin{cases}1 & \text{if } k=\text{Index}(j)\neq 0\\ 0 & \text{otherwise}\end{cases},\ \forall j\in[J],\forall k\in[j].$$

    Construct a weight π^i**:**

$$\pi^i=\min\left\{\sum_{\ell=m}^{J}\hat{s}^i_\ell-M_m,\quad \hat{s}^i_m,\quad \hat{t}^i_{j,\text{Index}(j)},\ \forall j\in[J]:\text{Index}(j)\neq 0\right\},$$

$$\text{where } M_m=\max\left\{\sum_{\ell=m}^{j}\hat{t}^i_{j\ell}:j\in B_m\right\}.$$

    Update the fractional solution $(\hat{s}^{i+1},\hat{t}^{i+1})$**:**

$$\hat{s}^{i+1}_k=\hat{s}^i_k-\pi^i s^i_k,\ \forall k\in[J]\quad\text{and}\quad \hat{t}^{i+1}_{jk}=\hat{t}^i_{jk}-\pi^i t^i_{jk},\ \forall j\in[J],\forall k\in[j].$$

    Update: $i=i+1$.

  end while

end for

Set: $n=i$, $\pi^n=1-(\pi^1+\cdots+\pi^{n-1})$, $s^n_k=0,\ \forall k\in[J]$, and $t^n_{jk}=0,\ \forall j\in[J],\forall k\in[j]$.

Fourth, we consider constraints (12a). For fixed $k\in[J]$ and $j\in[k,J]_{\mathbb{Z}}$, we rewrite each term in (12a) as follows:

$$\sum_{\ell=k}^{J}\hat{s}^{i+1}_\ell=\begin{cases}\sum_{\ell=k}^{J}\hat{s}^i_\ell-\pi^i & \text{if } k\in[m]\\ \sum_{\ell=k}^{J}\hat{s}^i_\ell & \text{if } k\in[m+1,J]_{\mathbb{Z}}\end{cases},\tag{25a}$$

$$\sum_{\ell=k}^{j}\hat{t}^{i+1}_{j\ell}=\begin{cases}\sum_{\ell=k}^{j}\hat{t}^i_{j\ell}-\pi^i & \text{if } k\in[m-1],j\in[k,m-1]_{\mathbb{Z}}:k\le\text{Index}(j)\\ 0 & \text{if } k\in[m-1],j\in[k,m-1]_{\mathbb{Z}}:k>\text{Index}(j)\\ \sum_{\ell=k}^{j}\hat{t}^i_{j\ell}-\pi^i & \text{if } j\in A_m,k\in[m]\\ \sum_{\ell=k}^{j}\hat{t}^i_{j\ell} & \text{if } j\in A_m,k\in[m+1,j]_{\mathbb{Z}}\\ \sum_{\ell=k}^{j}\hat{t}^i_{j\ell}-\pi^i & \text{if } j\in B_m,k\in[j]:k\le\text{Index}(j)\\ \sum_{\ell=k}^{j}\hat{t}^i_{j\ell} & \text{if } j\in B_m,k\in[j]:k>\text{Index}(j)\end{cases}.\tag{25b}$$

We prove by enumerating all 6 cases of $\sum_{\ell=k}^{j} \hat{t}_{j\ell}^{i+1}$ in (25b). In cases 1 and 3, (12a) is satisfied because $\sum_{\ell=k}^{j} \hat{t}_{j\ell}^{i} - \pi^i \le \sum_{\ell=k}^{J} \hat{s}_\ell^i - \pi^i$, which follows from the induction assumption that $(\hat{s}^i, \hat{t}^i)$ satisfies (12a). Likewise, (12a) is satisfied in cases 4 and 5. In case 2, (12a) is satisfied because $\pi^i \le \hat{s}_m^i \le \sum_{\ell=k}^{J} \hat{s}_\ell^i$. In case 6, we discuss the following two sub-cases: (i) If $k \ge m+1$, then (12a) is satisfied because $\sum_{\ell=k}^{j} \hat{t}_{j\ell}^i \le \sum_{\ell=k}^{J} \hat{s}_\ell^i$ by the induction assumption. (ii) If $\text{Index}(j) < k \le m$, then $\sum_{\ell=k}^{j} \hat{t}_{j\ell}^i = \sum_{\ell=m}^{j} \hat{t}_{j\ell}^i$ by definition of $\text{Index}(j)$. It follows that $\pi^i \le \sum_{\ell=m}^{J} \hat{s}_\ell^i - M_m \le \sum_{\ell=k}^{J} \hat{s}_\ell^i - \sum_{\ell=m}^{j} \hat{t}_{j\ell}^i$, where the first inequality is due to the definition of π^i and the second inequality is due to that of M_m. This proves that $(\hat{s}^{i+1}, \hat{t}^{i+1})$ satisfies (12a).

Finally, we consider constraints (11d). By (24), we rewrite the right-hand side of constraints (11d) as follows:

$$\sum_{j=k}^{J} \hat{t}_{jk}^{i+1} = \begin{cases} \sum_{j=k}^{J} (\hat{t}_{jk}^i - \mathbb{1}_{\text{Index}(j)}(k)\pi^i) & \text{if } k \in [m-1] \\ \sum_{j=m}^{J} \hat{t}_{jm}^i - \pi^i & \text{if } k = m \text{ and } A_m = \{j^*\} \\ \sum_{j=m}^{J} \hat{t}_{jm}^i - \sum_{j \in A_m} \pi^i & \text{if } k = m \text{ and } A_m \supset \{j^*\} \\ \sum_{j=k}^{J} \hat{t}_{jk}^i & \text{if } k \in [m+1, J]_{\mathbb{Z}} \end{cases}, \tag{26}$$

where $\mathbb{1}_{\text{Index}(j)}(k) = 1$ if $k = \text{Index}(j)$ and $\mathbb{1}_{\text{Index}(j)}(k) = 0$ otherwise. We prove by enumerating all 4 cases of $\sum_{j=k}^{J} \hat{t}_{jk}^{i+1}$ in (26). In case 1, we notice that (i) $\hat{s}_k^i = 0$, $\forall k \in [m-1]$ due to line 3, (ii) $\hat{t}_{jk}^i \ge 0$ for all $j \in [k, J]_{\mathbb{Z}}$, and (iii) $\pi^i \le \hat{t}_{j,\text{Index}(j)}^i$, $\forall j \in [J] : \text{Index}(j) \ne 0$ by the definition of π^i. Therefore, (11d) is satisfied:

$$\hat{s}_k^{i+1} = \hat{s}_k^i = 0 \le \sum_{j=k}^{J} (\hat{t}_{jk}^i - \mathbb{1}_{\text{Index}(j)}(k)\pi^i) = \sum_{j=k}^{J} \hat{t}_{jk}^{i+1}.$$

In case 2, (11d) is satisfied because $\hat{s}_m^i - \pi^i \le \sum_{j=m}^{J} \hat{t}_{jm}^i - \pi^i$, which follows from the induction assumption that $(\hat{s}^i, \hat{t}^i)$ satisfies (11d). In case 3, we notice that $\sum_{\ell=m}^{j} \hat{t}_{j\ell}^i = \sum_{\ell=m}^{J} \hat{s}_\ell^i$ for all $j \in A_m \setminus \{j^*\} \ne \emptyset$. As $\sum_{\ell=m+1}^{j} \hat{t}_{j\ell}^i \le \sum_{\ell=m+1}^{J} \hat{s}_\ell^i$ because $(\hat{s}^i, \hat{t}^i)$ satisfies (12a) by the induction assumption, we have $\hat{t}_{jm}^i \ge \hat{s}_m^i$, $\forall j \in A_m \setminus \{j^*\}$. Then, $\hat{s}_m^i - \pi^i \le \sum_{j \in A_m \setminus \{j^*\}} (\hat{t}_{jm}^i - \pi^i) + (\hat{t}_{j^*m}^i - \pi^i)$, where $\hat{t}_{jm}^i - \pi^i \ge 0$ for all $j \in A_m \setminus \{j^*\}$ and $\hat{t}_{j^*m}^i - \pi^i \ge 0$ by the definition of π^i. It follows that (11d) is satisfied:

$$\hat{s}_k^{i+1} = \hat{s}_m^i - \pi^i \le \sum_{j \in A_m} (\hat{t}_{jm}^i - \pi^i) \le \sum_{j=m}^{J} \hat{t}_{jm}^i - \sum_{j \in A_m} \pi^i = \sum_{j=k}^{J} \hat{t}_{jk}^{i+1}.$$

In case 4, (11d) is satisfied because $\hat{s}_k^i \le \sum_{j=k}^{J} \hat{t}_{jk}^i$, which follows from the induction assumption that $(\hat{s}^i, \hat{t}^i)$ satisfies (11d).

(Property 1) By construction, (s^i, t^i) produced in line 5 satisfies constraints (11b), (11c), and (11g). It remains to show constraints (11d)–(11f), which reduce to the following inequalities as we set $s_m^i = 1$:

$$1 \le \sum_{j=m}^{J} t_{jm}^i, \tag{27a}$$

$$t_{j\ell}^i \le 0, \ \forall \ell \in [m+1, J]_{\mathbb{Z}}, \forall j \in [\ell, J]_{\mathbb{Z}}. \tag{27b}$$

To that end, we note from line 4 that for all $j \in [J]$, $\text{Index}(j) \in [0, m]_{\mathbb{Z}}$. This implies that, in line 5, $t_{j\ell}^i = 0$ for all $\ell \in [m+1, J]_{\mathbb{Z}}$ and $j \in [\ell, J]_{\mathbb{Z}}$. This proves (27b). To see (27a), we notice that the construction of (s^i, t^i) implies that $\hat{s}_m^i > 0$ in line 3. Then, Property 3 implies that $\sum_{j=m}^{J} \hat{t}_{jm}^i \ge \hat{s}_m^i > 0$. It follows that, in line 5, we are able to find a non-empty A_m and so there exists a $j \in [m, J]_{\mathbb{Z}}$ such that $t_{jm}^i = 1$. This proves (27a).

(Property 2) To show that weight $\pi^i > 0$, it suffices to check that each term in the definition of π^i in line 6 is positive. First, by Property 3 and the definition of B_m in line 4, we have $\sum_{\ell=m}^{j} \hat{t}_{j\ell}^i < \sum_{\ell=m}^{J} \hat{s}_\ell^i$, $\forall j \in B_m$, which implies that $\sum_{\ell=m}^{J}\hat{s}^i_\ell - M_m = \sum_{\ell=m}^{J}\hat{s}^i_\ell - \max_{j\in B_m}\{\sum_{\ell=m}^{j}\hat{t}^i_{j\ell}\} > 0$. Second, $\hat{s}^i_m > 0$ due to line 3. It remains to show that $\hat{t}^i_{j,\text{Index}(j)} > 0,\ \forall j\in[J]:\text{Index}(j)\neq 0$. By noting that $[J] = A_m \cup B_m \cup [m-1]$, we consider the following three cases: (i) $j\in B_m\cup[m-1]$, (ii) $j = j^*$, and (iii) $j\in A_m\setminus\{j^*\}\neq\emptyset$. In case (i), $\text{Index}(j)\neq 0$ implies that the set $\{k\in[\min\{j,m-1\}]:\hat{t}^i_{jk}>0\}\neq\emptyset$, and thus $\hat{t}^i_{j,\text{Index}(j)} > 0$. In case (ii), we have $\hat{t}^i_{j^*,\text{Index}(j^*)} = \hat{t}^i_{j^*m} > 0$ by the definition of j^*. In case (iii), $\sum_{\ell=m}^{j}\hat{t}^i_{j\ell} = \sum_{\ell=m}^{J}\hat{s}^i_\ell$ as $j\in A_m\setminus\{j^*\}$. Since $\sum_{\ell=m+1}^{j}\hat{t}^i_{j\ell} \le \sum_{\ell=m+1}^{J}\hat{s}^i_\ell$ by Property 3, $\hat{t}^i_{jm}\ge\hat{s}^i_m>0$. This indicates that $\hat{t}^i_{j,\text{Index}(j)} = \hat{t}^i_{jm} > 0$. Finally, $\pi^i\le 1$ because each term in the definition of π^i is less than or equal to 1 by construction.

(Property 4) We claim that, for all $m\in[J]$, the while loop in line 3 terminates in a finite number of iterations, i.e., $\hat{s}^i_m$ becomes zero after a finite number of updates in line 7. It follows from this claim that Algorithm 2 terminates in a finite number of iterations. Accordingly, as $\hat{s}^n_k = 0, \forall k\in[J]$ at termination, we have $\hat{t}^n_{jk} = 0, \forall j\in[J], \forall k\in[j]$ because $(\hat{s}^n, \hat{t}^n)$ satisfies constraints (12a) by Property 3.

To see the claim, we consider the value of π^i returned by the minimum comparison in line 6. In particular, we distinguish between two cases: (A) $\pi^i = \hat{s}^i_m$ or $\pi^i = \hat{t}^i_{j,\text{Index}(j)}$ for a $j\in[J]$ with $\text{Index}(j)\neq 0$, and (B) $\pi^i = \sum_{\ell=m}^{J}\hat{s}^i_\ell - M_m$, and in addition, $\pi^i < \hat{s}^i_m$ and $\pi^i < \hat{t}^i_{j,\text{Index}(j)}$ for all $j\in[J]$ with $\text{Index}(j)\neq 0$. We notice that $\hat{s}^i_m$ will be updated to zero in line 7 after case (A) takes place for a finite number of times. This is because we only have a finite number of positive values among $\{\hat{s}^i_k : k\in[m]\}$ and $\{\hat{t}^i_{jk} : k\in[m], j\in[k,J]_{\mathbb{Z}}\}$. Now suppose that case (B) takes place. In this case, as $\pi^i = \sum_{\ell=m}^{J}\hat{s}^i_\ell - M_m < \hat{s}^i_m$, we know that $M_m > 0$ and so there exists a $j\in B_m$ such that $M_m = \sum_{\ell=m}^{j}\hat{t}^i_{j\ell}$. Then, after the update in line 7, we have $\sum_{\ell=m}^{j}\hat{t}^{i+1}_{j\ell} = \sum_{\ell=m}^{j}\hat{t}^i_{j\ell} = M_m$ and $\sum_{\ell=m}^{J}\hat{s}^{i+1}_\ell = \sum_{\ell=m}^{J}\hat{s}^i_\ell - \pi^i = M_m$. It follows that $\sum_{\ell=m}^{j}\hat{t}^{i+1}_{j\ell} = \sum_{\ell=m}^{J}\hat{s}^{i+1}_\ell$, increasing the cardinality of A_m by 1. Since $|A_m|\le J-m+1$, this indicates that case (B) can take place for at most $(J-m+1)$ times before case (A) takes place. Therefore, the claim holds valid.

(Property 5) By construction of the binary point in line 5, vectors $\hat{s}^{i+1}$ and $\hat{s}^i$ are only different in the m^{th} entry, i.e., $\hat{s}^{i+1}_m = \hat{s}^i_m - \pi^i$ and $\hat{s}^{i+1}_k = \hat{s}^i_k$ for all $k\neq m$. It follows that in line 10, i.e., when the for loop terminates, $0 < \sum_{i=1}^{n-1}\pi^i = \sum_{m=1}^{J}\hat{s}_m \le 1$, where the first inequality is due to Property 2 and the last inequality is due to Property 3. It follows that $\pi^n \equiv 1 - \sum_{i=1}^{n-1}\pi^i \in [0,1)$. □

Appendix K: Proof of Theorem 4

Proof: Under Structure 1, $\max_{(\alpha,\beta)\in\Lambda} F(\alpha,\beta)$ can be reformulated as the following linear program:

$$\begin{aligned}
\max_{(\alpha,\beta)\in\Lambda} F(\alpha,\beta) &= \max_{(t,s,r,p)\in\overline{\mathcal{H}}_1} \sum_{j=1}^{J}\left(c^{\text{r}}_j r_j + \sum_{k=1}^{j} c^{\text{t}}_{jk} t_{jk}\right) + c^{\text{p}}_1 p + \sum_{k=1}^{J} c^{\text{s}}_{1k} s_k \\
&= c^{\text{p}}_1 + \sum_{j=1}^{J} c^{\text{r}}_j + \max_{s\ge 0:(11c)}\left[\sum_{k=1}^{J}(c^{\text{s}}_{1k} - c^{\text{p}}_1)s_k + \left\{\max_{t\ge 0}\sum_{j=1}^{J}\sum_{k=1}^{j}(c^{\text{t}}_{jk} - c^{\text{r}}_j)t_{jk}\right\}\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{s.t. } (11b), (11d), (12a).
\end{aligned} \tag{28}$$

An optimal solution (s^*, t^*) to (28) is integral as $\overline{\mathcal{H}}_1 = \text{conv}(\mathcal{H}_1)$ is an integral polyhedron. Specifically, by constraints (11c), we have $s^*\in\{0\}\cup\{e_k : k\in[J]\}$, where e_k is the k^{th} standard basis vector, and for given s^*, the inner maximization problem in (28) finds an optimal t^*, which is also integral. We rewrite (28) by discussing two cases of s^*: (i) $s^* = 0$ and (ii) $s^*\in\{e_k : k\in[J]\}$.

In case (i), we have $t^*_{jk} = 0$ for all $j \in [J]$ and $k \in [j]$ by constraints (12a). Then, (28) reduces to $c^{\mathrm{p}}_1 + \sum_{j=1}^J c^{\mathrm{r}}_j$.

In case (ii), $s_{k^*} = 1$ for some $k^* \in [J]$ and $s_k = 0$ for all $k \in [J] \setminus \{k^*\}$. By constraints (12a), we have $t_{jk} = 0$ if $j, k \geq k^* + 1$. Then, we can rewrite (28) as follows:

$$c^{\mathrm{p}}_1 + \sum_{j=1}^J c^{\mathrm{r}}_j + (c^{\mathrm{s}}_{1k^*} - c^{\mathrm{p}}_1) + \max_{t \geq 0} \sum_{j=1}^J \sum_{k=1}^{\min\{j,k^*\}} (c^{\mathrm{t}}_{jk} - c^{\mathrm{r}}_j) t_{jk}$$
$$\text{s.t.} \sum_{k=1}^{\min\{j,k^*\}} t_{jk} \leq 1, \ \forall j \in [J], \quad 1 \leq \sum_{j=k^*}^J t_{jk^*}.$$

By constraint (11b), $t_{j^*k^*} = 1$ for some $j^* \in [k^*, J]_{\mathbb{Z}}$ and $t_{j^*k} = 0$ for all $k \in [j^*] \setminus \{k^*\}$. Then, we recast (28) as

$$c^{\mathrm{p}}_1 + \sum_{j=1}^J c^{\mathrm{r}}_j + (c^{\mathrm{s}}_{1k^*} - c^{\mathrm{p}}_1) + (c^{\mathrm{t}}_{j^*k^*} - c^{\mathrm{r}}_{j^*}) + \sum_{j \in [J] \setminus \{j^*\}} \max_{t \geq 0} \left\{ \sum_{k=1}^{\min\{j,k^*\}} (c^{\mathrm{t}}_{jk} - c^{\mathrm{r}}_j) t_{jk} \,\middle|\, \sum_{k=1}^{\min\{j,k^*\}} t_{jk} \leq 1 \right\}$$
$$= \sum_{j=1}^J c^{\mathrm{r}}_j + c^{\mathrm{s}}_{1k^*} + (c^{\mathrm{t}}_{j^*k^*} - c^{\mathrm{r}}_{j^*}) + \sum_{j \in [J] \setminus \{j^*\}} \max_{k \in [\min\{j,k^*\}]} \left\{ c^{\mathrm{t}}_{jk} - c^{\mathrm{r}}_j \right\}.$$

Therefore, we have

$$\max_{(\alpha,\beta) \in \Lambda} F(\alpha, \beta) =$$
$$\max \left\{ c^{\mathrm{p}}_1 + \sum_{j=1}^J c^{\mathrm{r}}_j, \ \max_{\substack{j^* \in [J], \\ k^* \in [j^*]}} \left\{ \sum_{j=1}^J c^{\mathrm{r}}_j + c^{\mathrm{s}}_{1k^*} + (c^{\mathrm{t}}_{j^*k^*} - c^{\mathrm{r}}_{j^*}) + \sum_{j \in [J] \setminus \{j^*\}} \max_{k \in [\min\{j,k^*\}]} \left\{ c^{\mathrm{t}}_{jk} - c^{\mathrm{r}}_j \right\} \right\} \right\}.$$

This completes the proof. □

Appendix L: Proof of Proposition 4

Proof: By Theorem 4, constraints (9c) are equivalent to

$$\theta \geq \max \left\{ c^{\mathrm{p}}_1 + \sum_{j=1}^J c^{\mathrm{r}}_j, \ \max_{\substack{j^* \in [J], \\ k^* \in [j^*]}} \left\{ \sum_{j=1}^J c^{\mathrm{r}}_j + c^{\mathrm{s}}_{1k^*} + (c^{\mathrm{t}}_{j^*k^*} - c^{\mathrm{r}}_{j^*}) + \sum_{j \in [J] \setminus \{j^*\}} \left(\max_{k \in [\min\{j,k^*\}]} \left\{ (c^{\mathrm{t}}_{jk} - c^{\mathrm{r}}_j) \right\} \right) \right\} \right\}.$$

where c^{r}_j, c^{t}_{jk}, c^{p}_1, and c^{s}_{1k} are computed by (9e)–(9h). Let $\chi_{jk^*} := \max_{k \in [\min\{j,k^*\}]} \left\{ (c^{\mathrm{t}}_{jk} - c^{\mathrm{r}}_j) \right\}$ for all $j \in [J]$ and $k^* \in [J]$. Then, constraints (9c) can be rewritten as follows:

$$\theta \geq c^{\mathrm{p}}_1 + \sum_{j=1}^J c^{\mathrm{r}}_j, \tag{29a}$$

$$\theta \geq \sum_{j=1}^J c^{\mathrm{r}}_j + c^{\mathrm{s}}_{1k^*} + (c^{\mathrm{t}}_{j^*k^*} - c^{\mathrm{r}}_{j^*}) + \sum_{j \in [J] \setminus \{j^*\}} \chi_{jk^*}, \ \forall j^* \in [J], \forall k^* \in [j^*], \tag{29b}$$

$$\chi_{jk^*} \geq (c^{\mathrm{t}}_{jk} - c^{\mathrm{r}}_j), \ \forall j \in [J], \forall k^* \in [J], \forall k \in [\min\{j,k^*\}]. \tag{29c}$$

Moreover, we define the following auxiliary variables:

$$\left.\begin{aligned} \zeta^{\mathrm{e}}_j &:= \left[-\gamma_j w^{\mathrm{L}}_j - \textstyle\sum_{\ell=1}^{w^{\mathrm{U}}_j - w^{\mathrm{L}}_j} \varphi_{j\ell} \right]_+ \\ \eta^{\mathrm{e}}_j &:= \max_{\tilde{d}_j \in [d^{\mathrm{L}}_j, d^{\mathrm{U}}_j]_{\mathbb{Z}}} \left\{ - \textstyle\sum_{q=1}^Q \tilde{d}^q_j \rho_{jq} \right\} \end{aligned}\right\} \ \forall j \in [J], \tag{30a}$$

$$\left.\begin{aligned} \zeta^{\mathrm{x}}_{jk} &:= \left[(-c^{\mathrm{x}}_k - \gamma_j) w^{\mathrm{L}}_j - \textstyle\sum_{\ell=1}^{w^{\mathrm{U}}_j - w^{\mathrm{L}}_j} (c^{\mathrm{x}}_k u_{j\ell} + \varphi_{j\ell}) \right]_+ \\ \eta^{\mathrm{x}}_{jk} &:= \max_{\tilde{d}_j \in [d^{\mathrm{L}}_j, d^{\mathrm{U}}_j]_{\mathbb{Z}}} \left\{ c^{\mathrm{x}}_k \tilde{d}_j - \textstyle\sum_{q=1}^Q \tilde{d}^q_j \rho_{jq} \right\} \end{aligned}\right\} \ \forall j \in [J], \forall k \in [j], \tag{30b}$$

$$\phi_1^{\mathrm{e}} := \Big[-\lambda_1 y_1^{\mathrm{L}} - \sum_{\ell=1}^{y_1^{\mathrm{U}}-y_1^{\mathrm{L}}} \nu_{1\ell}\Big]_+, \tag{30c}$$

$$\phi_{1k}^{\mathrm{x}} := \Big[(-c_k^{\mathrm{x}} - \lambda_1) y_1^{\mathrm{L}} - \sum_{\ell=1}^{y_1^{\mathrm{U}}-y_1^{\mathrm{L}}} (c_k^{\mathrm{x}} v_{1\ell} + \nu_{1\ell})\Big]_+, \ \forall k \in [J], \tag{30d}$$

Since $c_j^{\mathrm{r}} = \zeta_j^{\mathrm{e}} + \eta_j^{\mathrm{e}}$, $c_{jk}^{\mathrm{t}} = \zeta_{jk}^{\mathrm{x}} + \eta_{jk}^{\mathrm{x}}$, $c_1^{\mathrm{p}} = \phi_1^{\mathrm{e}}$, and $c_{1k}^{\mathrm{s}} = \phi_{1k}^{\mathrm{x}}$, constraints (29) can be rewritten as follows:

$$\theta \geq \phi_1^{\mathrm{e}} + \sum_{j=1}^{J} (\zeta_j^{\mathrm{e}} + \eta_j^{\mathrm{e}}), \tag{31a}$$

$$\theta \geq \sum_{j=1}^{J} (\zeta_j^{\mathrm{e}} + \eta_j^{\mathrm{e}}) + \phi_{1k^*}^{\mathrm{x}} + (\zeta_{j^*k^*}^{\mathrm{x}} + \eta_{j^*k^*}^{\mathrm{x}}) - (\zeta_{j^*}^{\mathrm{e}} + \eta_{j^*}^{\mathrm{e}}) + \sum_{j \in [J]\setminus\{j^*\}} \chi_{jk^*}, \ \forall j^* \in [J], \forall k^* \in [j^*], \tag{31b}$$

$$\chi_{jk^*} \geq (\zeta_{jk}^{\mathrm{x}} + \eta_{jk}^{\mathrm{x}}) - (\zeta_j^{\mathrm{e}} + \eta_j^{\mathrm{e}}), \ \forall j \in [J], \forall k^* \in [j], \forall k \in [\min\{j, k^*\}]. \tag{31c}$$

Replacing constraints (9c) with (30)–(31) in the formulation (9a)–(9d) leads to the claimed reformulation of (DRNS). This completes the proof. □

Appendix M: Proof of Proposition 5

We start by proving the following technical lemma.

LEMMA 4. *Consider sets $A_i \subseteq \mathbb{R}^{k_i}$ for all $i \in [I]$ and let $A := \Pi_{i=1}^I A_i$. Then, $conv(A) = \Pi_{i=1}^I conv(A_i)$.*

Proof: First, as $A = \Pi_{i=1}^I A_i$ and $A_i \subseteq \mathrm{conv}(A_i)$, we have $A \subseteq \Pi_{i=1}^I \mathrm{conv}(A_i)$ and hence $\mathrm{conv}(A) \subseteq \prod_{i=1}^I \mathrm{conv}(A_i)$ because $\prod_{i=1}^I \mathrm{conv}(A_i)$ is convex and $\mathrm{conv}(A)$ by definition is the smallest convex set that contains A. Second, to show that $\Pi_{i=1}^I \mathrm{conv}(A_i) \subseteq \mathrm{conv}(A)$, we pick any $a \in \Pi_{i=1}^I \mathrm{conv}(A_i)$ and prove that $a \in \mathrm{conv}(A)$. To this end, we denote $a := [a_1, \ldots, a_I]^\top$, where $a_i \in \mathrm{conv}(A_i)$ for all $i \in [I]$. Then, for all $i \in [I]$, there exist $\{\lambda_i^{n_i}\}_{n_i=1}^{N_i}$ and $\{a_i^{n_i}\}_{n_i=1}^{N_i}$ such that each $\lambda_i^{n_i} \geq 0$, each $a_i^{n_i} \in A_i$, $\sum_{n_i=1}^{N_i} \lambda_i^{n_i} = 1$, and $\sum_{n_i=1}^{N_i} \lambda_i^{n_i} a_i^{n_i} = a_i$ for all $i \in [I]$.

Denote set $\mathcal{N} := \{(n_1, \ldots, n_I) : n_i \in [N_i], \ \forall i \in [I]\}$, vector $a^n := [a_1^{n_1}, \ldots, a_I^{n_I}]^\top$ for all $n := [n_1, \ldots, n_I]^\top \in \mathcal{N}$, and scalar $\lambda^n := \Pi_{i=1}^I \lambda_i^{n_i}$ for all $n \in \mathcal{N}$. Then, $\lambda^n \geq 0$ and $a^n \in A$ for all $n \in \mathcal{N}$. In addition, $\sum_{n \in \mathcal{N}} \lambda^n = \sum_{n \in \mathcal{N}} \Pi_{i=1}^I \lambda_i^{n_i} = (\lambda_1^1 + \cdots + \lambda_1^{N_1})(\lambda_2^1 + \cdots + \lambda_2^{N_2}) \cdots (\lambda_I^1 + \cdots + \lambda_I^{N_I}) = 1$. Furthermore, for all $i \in [I]$, we have

$$\begin{aligned}
\sum_{n \in \mathcal{N}} \lambda^n a_i^{n_i} &= \sum_{m_i=1}^{N_i} \sum_{n \in \mathcal{N}: n_i = m_i} \lambda^n a_i^{n_i} \\
&= \sum_{m_i=1}^{N_i} a_i^{m_i} \sum_{n \in \mathcal{N}: n_i = m_i} \lambda^n \\
&= \sum_{m_i=1}^{N_i} a_i^{m_i} \lambda_i^{m_i} \\
&= a_i,
\end{aligned}$$

where third equality is because, for fixed $i \in [I]$ and $m_i \in [N_i]$, $\sum_{n \in \mathcal{N}: n_i = m_i} \lambda^n = (\lambda_1^1 + \cdots + \lambda_1^{N_1}) \cdots (\lambda_{m_i-1}^1 + \cdots + \lambda_{m_i-1}^{N_{m_i-1}}) \lambda_i^{m_i} (\lambda_{m_i+1}^1 + \cdots + \lambda_{m_i+1}^{N_{m_i+1}}) \cdots (\lambda_I^1 + \cdots + \lambda_I^{N_I}) = \lambda_i^{m_i}$, and the last inequality follows from the definitions of $\{\lambda_i^{m_i}\}_{m_i=1}^{N_i}$ and $\{a_i^{m_i}\}_{m_i=1}^{N_i}$. It follows that $a \equiv [a_1, \ldots, a_I]^\top = \sum_{n \in \mathcal{N}} \lambda^n [a_1^{n_1}, \ldots, a_I^{n_I}]^\top \equiv \sum_{n \in \mathcal{N}} \lambda^n a^n$

and hence $a \in \text{conv}(A)$. This completes the proof. □

We are now ready to present the main proof of this section.

Proof of Proposition 5: Since $\mathcal{H}_{\text{D}}$ is separable in index i, we have $\mathcal{H}_{\text{D}} = \prod_{i=1}^{I} \mathcal{H}_i^{\text{D}}$, where each $\mathcal{H}_i^{\text{D}}$ is defined as

$$\begin{aligned}
\mathcal{H}_i^{\text{D}} := \Bigg\{ (t, s_i, r, p_i) : \; & r_j + \sum_{k=1}^{j} t_{jk} = 1, \; \forall j \in P_i, \quad p_i + \sum_{k=1}^{J(i)} s_{ik} = 1, && \text{(32a)} \\
& \sum_{k=1}^{j} t_{jk} \le 1, \; \forall j \in P_i, && \text{(32b)} \\
& \sum_{k=1}^{J(i)} s_{ik} \le 1, && \text{(32c)} \\
& s_{ik} \le \sum_{j \in P_i : j \ge k} t_{jk}, \; \forall k \in [J(i)], && \text{(32d)} \\
& s_{ik} + t_{j\ell} \le 1, \; \forall k \in [J(i)], \forall j \in P_i : j \ge k+1, \forall \ell \in [k+1, J(i)]_{\mathbb{Z}}, \\
& t_{jk} \le \sum_{\ell=1}^{J(i)} s_{i\ell}, \; \forall j \in P_i, \forall k \in [j], \\
& s_{ik} \in \mathbb{B}, \; \forall k \in [J(i)], \quad t_{jk} \in \mathbb{B}, \; \forall j \in P_i, \forall k \in [j] \Bigg\}.
\end{aligned}$$

Following a similar proof in Section 4.1, we can show that incorporating inequalities (10) produces the convex hull of $\mathcal{H}_i^{\text{D}}$, i.e.,

$$\begin{aligned}
\overline{\mathcal{H}}_i^{\text{D}} := \Bigg\{ (t, s_i, r, p_i) : \; & (32a) - (32d), \\
& \sum_{\ell=k}^{j} t_{j\ell} \le \sum_{\ell=k}^{J(i)} s_{i\ell}, \; \forall j \in P_i, \forall k \in [j], \qquad (33) \\
& s_{ik} \in \mathbb{R}_+, \forall k \in [J(i)], \quad t_{jk} \in \mathbb{R}_+, \; \forall j \in P_i, \forall k \in [j] \Bigg\}.
\end{aligned}$$

Let $\overline{\mathcal{H}}_{\text{D}} = \prod_{i=1}^{I} \overline{\mathcal{H}}_i^{\text{D}}$. Then, it follows from Lemma 4 that $\overline{\mathcal{H}}_{\text{D}} = \text{conv}(\mathcal{H}_{\text{D}})$. Following a similar proof as that of Theorem 4, we have

$$\begin{aligned}
\max_{(\alpha,\beta)\in\Lambda} F(\alpha,\beta) &= \max_{(t,s,r,p)\in\overline{\mathcal{H}}_{\text{D}}} \sum_{j=1}^{J} \left(c_j^{\text{r}} r_j + \sum_{k=1}^{j} c_{jk}^{\text{t}} t_{jk} \right) + \sum_{i=1}^{I} \left(c_i^{\text{p}} p_i + \sum_{k=1}^{J(i)} c_{ik}^{\text{s}} s_{ik} \right) \\
&= \sum_{i=1}^{I} \left[\max_{(t,s_i,r,p_i)\in\overline{\mathcal{H}}_i^{\text{D}}} \sum_{j\in P_i} \left(c_j^{\text{r}} r_j + \sum_{k=1}^{j} c_{jk}^{\text{t}} t_{jk} \right) + \left(c_i^{\text{p}} p_i + \sum_{k=1}^{J(i)} c_{ik}^{\text{s}} s_{ik} \right) \right] \\
&= \sum_{i=1}^{I} \left[c_i^{\text{p}} + \sum_{j\in P_i} c_j^{\text{r}} + \max_{s_i \ge 0 : (32c)} \sum_{k=1}^{J(i)} (c_{ik}^{\text{s}} - c_i^{\text{p}}) s_{ik} + \left(\max_{t\ge 0} \sum_{j\in P_i} \sum_{k=1}^{j} (c_{jk}^{\text{t}} - c_j^{\text{r}}) t_{jk} \right) \right] \qquad (34) \\
& \qquad \text{s.t. } (32b), (32d), (33).
\end{aligned}$$

For every $i \in [I]$, an optimal solution (s_i^*, t^*) to (34) is integral as $\overline{\mathcal{H}}_i^{\text{D}} = \text{conv}(\mathcal{H}_i^{\text{D}})$ is an integral polyhedron. Following a similar proof in Theorem 4, we have

$$\max_{(\alpha,\beta)\in\Lambda} F(\alpha,\beta) = \sum_{i=1}^{I} \left[\max \left\{ c_i^{\text{p}} + \sum_{j\in P_i} c_j^{\text{r}}, \right. \right.$$

$$\max_{j^*\in P_i, k^*\in[j^*]}\left\{\sum_{j\in P_i} c^{\mathrm{r}}_j + c^{\mathrm{s}}_{ik^*} + (c^{\mathrm{t}}_{j^*k^*} - c^{\mathrm{r}}_{j^*}) + \sum_{j\in P_i\setminus\{j^*\}}\left(\max_{k\in[\min\{j,k^*\}]}\left\{c^{\mathrm{t}}_{jk} - c^{\mathrm{r}}_j\right\}\right)\right\}\Bigg].$$

We define $\chi_{jk^*} := \max_{k\in[\min\{j,k^*\}]}\left\{(c^{\mathrm{t}}_{jk} - c^{\mathrm{r}}_j)\right\}$ for all $j\in[J]$ and $k^*\in[J]$. Then, we rewrite (34) as follows:

$$\begin{aligned}
\min\ & \sum_{i=1}^{I}\theta_i\\
\text{s.t. } & \theta_i \ge c^{\mathrm{p}}_i + \sum_{j\in P_i} c^{\mathrm{r}}_j,\ \forall i\in[I],\\
& \theta_i \ge \sum_{j\in P_i} c^{\mathrm{r}}_j + c^{\mathrm{s}}_{ik^*} + (c^{\mathrm{t}}_{j^*k^*} - c^{\mathrm{r}}_{j^*}) + \sum_{j\in P_i\setminus\{j^*\}}\chi_{jk^*},\ \forall i\in[I], \forall j^*\in P_i, \forall k^*\in[j^*],\\
& \chi_{jk^*} \ge (c^{\mathrm{t}}_{jk} - c^{\mathrm{r}}_j),\ \forall j\in[J], \forall k^*\in[J], \forall k\in[\min\{j,k^*\}].
\end{aligned}$$

Moreover, we define the following auxiliary variables:

$$\left.\begin{aligned}
\zeta^{\mathrm{e}}_j &:= \left[-\gamma_j w^{\mathrm{L}}_j - \textstyle\sum_{\ell=1}^{w^{\mathrm{U}}_j - w^{\mathrm{L}}_j}\varphi_{j\ell}\right]_+\\
\eta^{\mathrm{e}}_j &:= \max_{\tilde d_j\in[d^{\mathrm{L}}_j, d^{\mathrm{U}}_j]_{\mathbb{Z}}}\left\{-\textstyle\sum_{q=1}^{Q}\tilde d^{q}_j\rho_{jq}\right\}
\end{aligned}\right\}\ \forall j\in[J], \tag{35a}$$

$$\left.\begin{aligned}
\zeta^{\mathrm{x}}_{jk} &:= \left[(-c^{\mathrm{x}}_k - \gamma_j) w^{\mathrm{L}}_j - \textstyle\sum_{\ell=1}^{w^{\mathrm{U}}_j - w^{\mathrm{L}}_j}(c^{\mathrm{x}}_k u_{j\ell} + \varphi_{j\ell})\right]_+\\
\eta^{\mathrm{x}}_{jk} &:= \max_{\tilde d_j\in[d^{\mathrm{L}}_j, d^{\mathrm{U}}_j]_{\mathbb{Z}}}\left\{c^{\mathrm{x}}_k\tilde d_j - \textstyle\sum_{q=1}^{Q}\tilde d^{q}_j\rho_{jq}\right\}
\end{aligned}\right\}\ \forall j\in[J], \forall k\in[j], \tag{35b}$$

$$\phi^{\mathrm{e}}_i := \left[-\lambda_i y^{\mathrm{L}}_i - \sum_{\ell=1}^{y^{\mathrm{U}}_i - y^{\mathrm{L}}_i}\nu_{i\ell}\right]_+,\ \forall i\in[I], \tag{35c}$$

$$\phi^{\mathrm{x}}_{ik} := \left[(-c^{\mathrm{x}}_k - \lambda_i) y^{\mathrm{L}}_i - \sum_{\ell=1}^{y^{\mathrm{U}}_i - y^{\mathrm{L}}_i}(c^{\mathrm{x}}_k v_{i\ell} + \nu_{i\ell})\right]_+,\ \forall i\in[I], \forall k\in[J(i)], \tag{35d}$$

Since $c^{\mathrm{r}}_j = \zeta^{\mathrm{e}}_j + \eta^{\mathrm{e}}_j$, $c^{\mathrm{t}}_{jk} = \zeta^{\mathrm{x}}_{jk} + \eta^{\mathrm{x}}_{jk}$, $c^{\mathrm{p}}_i = \phi^{\mathrm{e}}_i$, and $c^{\mathrm{s}}_{ik} = \phi^{\mathrm{x}}_{ik}$, we can rewrite (34) as follows:

$$\begin{aligned}
\min\ & \sum_{i=1}^{I}\theta_i\\
\text{s.t. } & (35),\\
& \theta_i \ge \phi^{\mathrm{e}}_i + \sum_{j\in P_i}(\zeta^{\mathrm{e}}_j + \eta^{\mathrm{e}}_j),\ \forall i\in[I],\\
& \theta_i \ge \sum_{j\in P_i}(\zeta^{\mathrm{e}}_j + \eta^{\mathrm{e}}_j) + \phi^{\mathrm{x}}_{ik^*} + (\zeta^{\mathrm{x}}_{j^*k^*} + \eta^{\mathrm{x}}_{j^*k^*}) - (\zeta^{\mathrm{e}}_{j^*} + \eta^{\mathrm{e}}_{j^*}) + \sum_{j\in P_i\setminus\{j^*\}}\chi_{jk^*},\ \forall i\in[I], \forall j^*\in P_i, \forall k^*\in[j^*],\\
& \chi_{jk^*} \ge (\zeta^{\mathrm{x}}_{jk} + \eta^{\mathrm{x}}_{jk}) - (\zeta^{\mathrm{e}}_j + \eta^{\mathrm{e}}_j),\ \forall j\in[J], \forall k^*\in[J], \forall k\in[\min\{j,k^*\}].
\end{aligned}$$

This completes the proof. □

Appendix N: Proof of Theorem 5

Proof: Each trajectory of states $t_{[1]}, (t_{[1]}, t_{[2]}), \ldots, (t_{[1]}, t_{[I]})$ in the DP corresponds to an S-T path in the network $(\mathcal{N}, \mathcal{A})$, where the objective function value of the trajectory $V_I(t_{[1]}, t_{[I]}) + \sum_{k\preceq 1\curlyvee I}(c^{\mathrm{s}}_{Ik} - c^{\mathrm{p}}_I)h_k(t_{[I]}, t_{[1]})$ equals the length of the S-T path by definition of the arc lengths $c_{[m,n]}$. Likewise, each S-T path in $(\mathcal{N}, \mathcal{A})$ corresponds to a trajectory of states in the DP and the length of the path equals the objective function value of the trajectory. This proves that the length of the longest path in $(\mathcal{N}, \mathcal{A})$ equals $\max_{(\alpha,\beta)\in\Lambda} F(\alpha,\beta)$ and completes the proof. □

Appendix O: Proof of Proposition 6

Proof: Taking the dual of the longest-path formulation yields

$$\begin{aligned}
&\min_{\pi} \ \pi_{\mathtt{S}} - \pi_{\mathtt{T}} + \sum_{i=1}^{I}(c^{\mathrm{r}}_i + c^{\mathrm{p}}_i) \\
&\text{s.t. } \pi_{\mathtt{S}} - \pi_{t_{[1]}} \geq (c^{\mathrm{t}}_{11} - c^{\mathrm{r}}_1)t_{11}, \ \forall t_{[1]} \in \mathcal{B}_1, \\
&\quad \pi_{t_{[1]}} - \pi_{(t_{[1]}, e_{2\ell})} \geq \begin{cases} (c^{\mathrm{s}}_{11} - c^{\mathrm{p}}_1)t_{11} & \text{if } \ell = 0 \\ (c^{\mathrm{t}}_{2\ell} - c^{\mathrm{r}}_2) + (c^{\mathrm{s}}_{1\ell} - c^{\mathrm{p}}_1) & \text{if } \ell \neq 0 \end{cases} \quad \forall t_{[1]} \in \mathcal{B}_1, \forall \ell \preceq 2, \\
&\quad \pi_{(t_{[1]}, e_{(i-1)k})} - \pi_{(t_{[1]}, e_{i\ell})} \geq \begin{cases} 0, & \text{if } k = 0, \ell = 0 \\ c^{\mathrm{s}}_{(i-1)k} - c^{\mathrm{p}}_{i-1}, & \text{if } k \neq 0, \ell = 0 \\ (c^{\mathrm{t}}_{i\ell} - c^{\mathrm{r}}_i) + (c^{\mathrm{s}}_{(i-1)(k \curlyvee \ell)} - c^{\mathrm{p}}_{i-1}), & \text{if } \ell \neq 0, \end{cases} \\
&\qquad\qquad \forall t_{[1]} \in \mathcal{B}_1, \forall i \in [3, I]_{\mathbb{Z}}, \forall k \preceq i-1, \forall \ell \preceq i, \\
&\quad \pi_{(t_{[1]}, e_{Ik})} - \pi_{\mathtt{T}} \geq \begin{cases} (c^{\mathrm{s}}_{I1} - c^{\mathrm{p}}_I)t_{11} & \text{if } k = 0 \\ c^{\mathrm{s}}_{Ik} - c^{\mathrm{p}}_I & \text{if } k \neq 0 \end{cases} \quad \forall t_{[1]} \in \mathcal{B}_1, \forall k \preceq I,
\end{aligned}$$

where dual variables π are associated with the (primal) flow balance constraints and all dual constraints are associated with the primal variables x. The strong duality holds valid because the longest-path formulation is finitely optimal. The claimed reformulation of (DRNS) then follows from a similar proof as that of Proposition 4. □

Appendix P: Formulation for DRO$^{\mathrm{exo}}$

We formulate DRO$^{\mathrm{exo}}$, which models constant nurse show-up rates (exogenous of the staffing levels) as follows:

$$\begin{aligned}
(\mathrm{DRO}^{\mathrm{exo}}): \quad & \min_{w,y} \ \sum_{j=1}^{J} c^{\mathrm{w}}_j w_j + \sum_{i=1}^{I} c^{\mathrm{y}}_i y_i + \max_{\mathbb{Q} \in \mathcal{D}^{\mathrm{exo}}} \mathbb{E}_{\mathbb{Q}}\left[V\left(\tilde{u} \circ w, \tilde{p} \circ y, \tilde{d}\right)\right] \\
& \text{s.t. } (3b)\text{–}(3c),
\end{aligned}$$

where $\tilde{u} \in [0,1]^J$ and $\tilde{p} \in [0,1]^I$ are the show-up rates of unit and pool nurses, respectively, $\circ$ represents the element-wise product, and $\mathbb{Q}$ is the probability distribution of $(\tilde{u}, \tilde{p}, \tilde{d})$. We formulate the ambiguity set $\mathcal{D}^{\mathrm{exo}}$ for $\mathbb{Q}$ as

$$\begin{aligned}
\mathcal{D}^{\mathrm{exo}} \ = \ \Big\{ \mathbb{Q} \in \mathcal{P}(\Gamma) : \ & \mathbb{E}_{\mathbb{Q}}[\tilde{d}^q_j] = \mu_{jq}, \ \forall j \in [J], \ \forall q \in [Q], \\
& \mathbb{E}_{\mathbb{Q}}[\tilde{u}_j] = \tau^{\mathrm{w}}_j, \ \forall j \in [J], \qquad \mathbb{E}_{\mathbb{Q}}[\tilde{p}_i] = \tau^{\mathrm{y}}_i, \ \forall i \in [I] \Big\},
\end{aligned}$$

where $\Gamma := [0,1]^J \times [0,1]^I \times \Xi_{\tilde{d}}$ represents the support of $(\tilde{u}, \tilde{p}, \tilde{d})$ and $(\tau^{\mathrm{w}}, \tau^{\mathrm{y}})$ are the mean show-up rates. We note that, unlike $\mathcal{D}(w, y)$ in DRNS, $\mathcal{D}^{\mathrm{exo}}$ is independent of the staffing levels.

DRO$^{\mathrm{exo}}$ can be viewed as a special case of DRNS, wherein we set $f_j(w_j) = \tau^{\mathrm{w}}_j w_j$ for all $j \in [J]$ and $g_i(y_i) = \tau^{\mathrm{y}}_i y_i$ for all $i \in [I]$. As a result, we can solve DRO$^{\mathrm{exo}}$ using the same solution approaches (e.g., Algorithm 1) as DRNS.

Appendix Q: Derivation for (17)

Using the standard duality techniques as in the proof of Proposition 1, we recast the DRNS$^{(Q)}$ model as the following min-max reformulation:

$$\min_{(w,y)\in\mathcal{W},\gamma,\lambda,\rho}\ \max_{(\alpha,\beta)\in\Lambda} \underline{F}(\alpha,\beta)+\sum_{i=1}^{I}\left(c_i^{\mathrm{y}}y_i+\sum_{q=1}^{Q}g_{iq}(y_i)\lambda_{iq}\right)+\sum_{j=1}^{J}\left[c_j^{\mathrm{w}}w_j+\sum_{q=1}^{Q}\mu_{jq}\rho_{jq}+\sum_{q=1}^{Q}f_{jq}(w_j)\gamma_{jq}\right],$$

where $\mathcal{W}:=\{(w,y):\text{(3b)–(3c)}\}$ and

$$\begin{aligned}\underline{F}(\alpha,\beta)\ :=\ &\sum_{j=1}^{J}\Big[\max_{\tilde{w}_j\in[0,w_j]_{\mathbb{Z}}}\Big\{-\alpha_j\tilde{w}_j-\sum_{q=1}^{Q}\gamma_{jq}\tilde{w}_j^q\Big\}+\max_{\tilde{d}_j\in[d_j^{\mathrm{L}},d_j^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{\alpha_j\tilde{d}_j-\sum_{q=1}^{Q}\rho_{jq}\tilde{d}_j^q\Big\}\Big]+\\ &\sum_{i=1}^{I}\max_{\tilde{y}_i\in[0,y_i]_{\mathbb{Z}}}\Big\{\beta_i\tilde{y}_i-\sum_{q=1}^{Q}\lambda_{iq}\tilde{y}_i^q\Big\}.\end{aligned}$$

The computation of $\underline{F}(\alpha,\beta)$ involves a set of integer programs with polynomial objective functions and changing feasible regions pertaining to the staffing levels (w,y), yielding an intractable formulation.

Fortunately, we can relax the decision-dependent support sets $[0,w_j]_{\mathbb{Z}}$ and $[0,y_i]_{\mathbb{Z}}$ to be decision-*in*dependent $[0,w_j^{\mathrm{U}}]_{\mathbb{Z}}$ and $[0,y_i^{\mathrm{U}}]_{\mathbb{Z}}$, respectively, by replacing (w,y) with their upper bounds $(w^{\mathrm{U}},y^{\mathrm{U}})$. This produces an upper bound for $\underline{F}(\alpha,\beta)$. In addition, the integer programs in $\underline{F}(\alpha,\beta)$ can now be solved by enumerating a linear number of staffing levels, i.e., $\tilde{w}_j\in[0,w_j^{\mathrm{U}}]_{\mathbb{Z}}$ and $\tilde{y}_i\in[0,y_i^{\mathrm{U}}]_{\mathbb{Z}}$. This bounds $\underline{F}(\alpha,\beta)$ as

$$\begin{aligned}&\underline{F}(\alpha,\beta)\le F(\alpha,\beta)\\ &:=\sum_{j=1}^{J}\Big[\max_{\tilde{w}_j\in[0,w_j^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{-\alpha_j\tilde{w}_j-\sum_{q=1}^{Q}\gamma_{jq}\tilde{w}_j^q\Big\}+\max_{\tilde{d}_j\in[d_j^{\mathrm{L}},d_j^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{\alpha_j\tilde{d}_j-\sum_{q=1}^{Q}\rho_{jq}\tilde{d}_j^q\Big\}\Big]+\\ &\quad\ \sum_{i=1}^{I}\max_{\tilde{y}_i\in[0,y_i^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{\beta_i\tilde{y}_i-\sum_{q=1}^{Q}\lambda_{iq}\tilde{y}_i^q\Big\}\\ &=\sum_{j=1}^{J}\Bigg[\mathbb{1}_{\{0\}}(\alpha_j)\Bigg\{\max_{\tilde{w}_j\in[0,w_j^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{-\sum_{q=1}^{Q}\gamma_{jq}\tilde{w}_j^q\Big\}+\max_{\tilde{d}_j\in[d_j^{\mathrm{L}},d_j^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{-\sum_{q=1}^{Q}\rho_{jq}\tilde{d}_j^q\Big\}\Bigg\}+\\ &\quad\ \sum_{k=1}^{j}\mathbb{1}_{\{c_k^{\mathrm{x}}\}}(\alpha_j)\Bigg\{\max_{\tilde{w}_j\in[0,w_j^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{-c_k^{\mathrm{x}}\tilde{w}_j-\sum_{q=1}^{Q}\gamma_{jq}\tilde{w}_j^q\Big\}+\max_{\tilde{d}_j\in[d_j^{\mathrm{L}},d_j^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{c_k^{\mathrm{x}}\tilde{d}_j-\sum_{q=1}^{Q}\rho_{jq}\tilde{d}_j^q\Big\}\Bigg\}\Bigg]+\\ &\quad\ \sum_{i=1}^{I}\Bigg[\mathbb{1}_{\{0\}}\Big(\max_{j\in P_i}\{\alpha_j\}\Big)\Bigg\{\max_{\tilde{y}_i\in[0,y_i^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{-\sum_{q=1}^{Q}\lambda_{iq}\tilde{y}_i^q\Big\}\Bigg\}+\\ &\quad\ \sum_{k=1}^{J(i)}\mathbb{1}_{\{c_k^{\mathrm{x}}\}}\Big(\max_{j\in P_i}\{\alpha_j\}\Big)\Bigg\{\max_{\tilde{y}_i\in[0,y_i^{\mathrm{U}}]_{\mathbb{Z}}}\Big\{-c_k^{\mathrm{x}}\tilde{y}_i-\sum_{q=1}^{Q}\lambda_{iq}\tilde{y}_i^q\Big\}\Bigg\}\Bigg]\\ &=\sum_{j=1}^{J}\Bigg[\mathbb{1}_{\{0\}}(\alpha_j)c_j^{\mathrm{r}}+\sum_{k=1}^{j}\mathbb{1}_{\{c_k^{\mathrm{x}}\}}(\alpha_j)c_{jk}^{\mathrm{t}}\Bigg]+\sum_{i=1}^{I}\Bigg[\mathbb{1}_{\{0\}}\Big(\max_{j\in P_i}\{\alpha_j\}\Big)c_i^{\mathrm{p}}+\sum_{k=1}^{J(i)}\mathbb{1}_{\{c_k^{\mathrm{x}}\}}\Big(\max_{j\in P_i}\{\alpha_j\}\Big)c_{ik}^{\mathrm{s}}\Bigg],\end{aligned}$$

where the last equality follows from the definition of $(c_j^{\mathrm{r}},c_{jk}^{\mathrm{t}},c_i^{\mathrm{p}},c_{ik}^{\mathrm{s}})$ in (17). Finally, we observe that $F(\alpha,\beta)$ is jointly convex in (α,β) and hence the problem $\max_{(\alpha,\beta)\in\Lambda}F(\alpha,\beta)$ resorts to the extreme points of Λ. Following Lemma 2, as well as the subsequent construction of binary variables (t,s,r,p) and set $\mathcal{H}$, we have

$$\max_{(\alpha,\beta)\in\Lambda}F(\alpha,\beta)\ =\ \max_{(t,s,r,p)\in\mathcal{H}}\ \sum_{j=1}^{J}\left(c_j^{\mathrm{r}}r_j+\sum_{k=1}^{j}c_{jk}^{\mathrm{t}}t_{jk}\right)+\sum_{i=1}^{I}\left(c_i^{\mathrm{p}}p_i+\sum_{k=1}^{J(i)}c_{ik}^{\mathrm{s}}s_{ik}\right).$$