Supplementary Material

1. Proof of Lemma 1 Our analysis requires the following claim on bounding the dual prices.

CLAIM 1. *For any phase* $m \in \{1,2,\ldots,M\}$ *and any* $s \in [\tau^{(m-2)}+1 : \tau^{(m-1)}-1]$,

$$\phi_i(s+1) \le \gamma + \frac{\gamma}{\sqrt{2W^{(m-2)}}} \qquad \forall i \in \mathcal{I}$$

almost surely.

Proof of Claim 1 We prove by induction for a fixed $i \in \mathcal{I}$. Fix some $m \in \{1,2,\ldots,M\}$, it is evident that $\phi_i(\tau^{(m-2)}+1) = 0 \le \frac{r_{\max}}{v_{\min}} + \frac{v_{\max}}{\sqrt{2W^{(m-2)}}}$. Suppose for some $s \in [\tau^{(m-2)}+1 : \tau^{(m-1)}-1]$, we have $\phi_i(s) \le \frac{r_{\max}}{v_{\min}} + \frac{v_{\max}}{\sqrt{2W^{(m-2)}}}$. Then it suffices to show that $\phi_i(s+1) \le \frac{r_{\max}}{v_{\min}} + \frac{v_{\max}}{\sqrt{2W^{(m-2)}}}$.

If $\phi_i(s) \le \frac{r_{\max}}{v_{\min}}$, then by (11), we evidently have

$$\phi_i(s+1) \le \frac{r_{\max}}{v_{\min}} + \frac{v_{\max}}{\sqrt{2W^{(m-2)}}} \le \gamma + \frac{\gamma}{\sqrt{2W^{(m-2)}}}.$$

In the complementary case when $\frac{r_{\max}}{v_{\min}} < \phi_i(s) \le \frac{r_{\max}}{v_{\min}} + \frac{v_{\max}}{\sqrt{2W^{(m-2)}}}$, we first show $v_{i,j(s),\tilde{k}(s)} = 0$ by contradiction. Recall that $\tilde{k}(s) = \text{argmax}_{k\in\mathcal{K}} \left\{ r_{j(s),k} - \sum_{i\in\mathcal{I}} \phi_i(s) v_{i,j(s),k} \right\}$, and there exists a null action $k_{\text{null}} \in \mathcal{K}$ such that $r_{j,k_{\text{null}}} = v_{i,j,k_{\text{null}}} = 0$. Hence, the virtual action $\tilde{k}(s)$ satisfies

$$r_{j(s),\tilde{k}(s)} - \sum_{i\in\mathcal{I}} \phi_i(s) v_{i,j(s),\tilde{k}(s)} \ge r_{j(s),k_{\text{null}}} - \sum_{i\in\mathcal{I}} \phi_i(s) v_{i,j(s),k_{\text{null}}} = 0.$$

In this case, if $v_{i,j(s),\tilde{k}(s)} \ne 0$, we have

$$r_{j(s),\tilde{k}(s)} \ge \sum_{i\in\mathcal{I}} \phi_i(s) v_{i,j(s),\tilde{k}(s)} > \frac{r_{\max}}{v_{\min}} v_{\min} > r_{\max},$$

which contradicts the fact that $r_{\max} = \max_{i,j,k}\{r_{ijk}\}$. Hence, by (11), we have

$$\phi_i(s+1) \le \phi_i(s) \le \frac{r_{\max}}{v_{\min}} + \frac{v_{\max}}{\sqrt{2W^{(m-2)}}} \le \gamma + \frac{\gamma}{\sqrt{2W^{(m-2)}}}.$$

$\square$

Proof of Lemma 1 **Step 1 (validating inequality (18)):** It suffices to show the following two inequalities:

$$\begin{aligned} &\sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \left(\lambda_*^{S(s)} - \sum_{j\in\mathcal{J}} p_{jt} r_{j,\kappa(\boldsymbol{\phi}(s),j)} \right) \\ \le &\sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{i\in\mathcal{I}} \phi_i(s) \left(c_i - \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)} \right) + \gamma W^{(m-1)} V^{(m-1)}, \end{aligned} \qquad \text{(SM-1)}$$

and with probability at least $1-\delta$,

$$\sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{i\in\mathcal{I}} \phi_i(s)\left(c_i - \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)}\right) \le W^{(m-1)} \epsilon_F^{(m-1)}. \tag{SM-2}$$

To show (SM-1), we define the following Lagrangian function regarding (LP-S(t)):

$$\begin{aligned}
L(\boldsymbol{\psi}(t),\boldsymbol{y}(t),t) &= \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} p_{jt} r_{jk} y_{jk}(t) - \sum_{i\in\mathcal{I}} \psi_i(t)\left(\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} p_{jt} v_{ijk} y_{jk}(t) - c_i\right)\\
&= \sum_{i\in\mathcal{I}} \psi_i(t) c_i + \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} p_{jt}\left(r_{jk} - \sum_{i\in\mathcal{I}} \psi_i(t) v_{ijk}\right) y_{jk}(t)
\end{aligned}$$

where $\boldsymbol{\psi}(t)$ are the dual prices corresponding to the constraints at time step t. By strong duality, we have

$$\begin{aligned}
\lambda_*^{S(t)} &= \min_{\boldsymbol{\psi}(t)\in[0,2\gamma]^{|\mathcal{I}|}} \max_{\boldsymbol{y}(t):\sum_{k\in\mathcal{K}} y_{jk}(t)\le 1} L(\boldsymbol{\psi}(t),\boldsymbol{y}(t),t)\\
&= \min_{\boldsymbol{\psi}(t)\in[0,2\gamma]^{|\mathcal{I}|}} \max_{\boldsymbol{y}(t):\sum_{k\in\mathcal{K}} y_{jk}(t)\le 1} \left\{\sum_{i\in\mathcal{I}} \psi_i(t) c_i + \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} p_{jt}\left(r_{jk} - \sum_{i\in\mathcal{I}} \psi_i(t) v_{ijk}\right) y_{jk}(t)\right\}\\
&= \min_{\boldsymbol{\psi}(t)\in[0,2\gamma]^{|\mathcal{I}|}} \left\{\sum_{i\in\mathcal{I}} \psi_i(t) c_i + \sum_{j\in\mathcal{J}} p_{jt} \max_{\boldsymbol{y}(t):\sum_{k\in\mathcal{K}} y_{jk}(t)\le 1} \left\{\sum_{k\in\mathcal{K}}\left(r_{jk} - \sum_{i\in\mathcal{I}} \psi_i(t) v_{ijk}\right) y_{jk}(t)\right\}\right\}.
\end{aligned} \tag{SM-3}$$

It is worth noticing that we set the dual variables

$$\psi_i(t) \le \frac{r_{\max}}{v_{\min}} \le \gamma \qquad \forall i\in\mathcal{I},$$

since otherwise $r_{jk} - \sum_{i\in\mathcal{I}} \psi_i(t) v_{ijk} \le 0$ for all j,k. In this case, $y_{jk}(t)=0$ by the maximization argument, and $\psi_i(t)$ does not get larger by the minimization argument. Under Algorithm 2, we have:

$$\begin{aligned}
&\sum_{t=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \left(\lambda_*^{S(t)} - \sum_{j\in\mathcal{J}} p_{jt} r_{j,\kappa(\boldsymbol{\phi}(t),j)}\right)\\
=& \sum_{t=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \min_{\boldsymbol{\psi}(t)\in[0,2\gamma]^{|\mathcal{I}|}} \max_{\boldsymbol{y}(t):\sum_{k\in\mathcal{K}} y_{jk}(t)\le 1} L(\boldsymbol{\psi}(t),\boldsymbol{y}(t),t) - \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{j\in\mathcal{J}} p_{js} r_{j,\kappa(\boldsymbol{\phi}(s),j)}\\
=& \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \left(\frac{1}{W^{(m-1)}} \sum_{t=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \min_{\boldsymbol{\psi}(t)\in[0,2\gamma]^{|\mathcal{I}|}} \max_{\boldsymbol{y}(t):\sum_{k\in\mathcal{K}} y_{jk}(t)\le 1} L(\boldsymbol{\psi}(t),\boldsymbol{y}(t),t) - \sum_{j\in\mathcal{J}} p_{js} r_{j,\kappa(\boldsymbol{\phi}(s),j)}\right)
\end{aligned}$$

$$\begin{aligned}
&\le \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \left(\frac{1}{W^{(m-1)}} \sum_{t=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \max_{\boldsymbol{y}(t):\sum_{k\in\mathcal{K}} y_{jk}(t)\le 1} L(\boldsymbol{\phi}(s),\boldsymbol{y}(t),t) - \sum_{j\in\mathcal{J}} p_{js} r_{j,\kappa(\boldsymbol{\phi}(s),j)} \right) \\
&\le \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \left(\sum_{i\in\mathcal{I}} \phi_i(s) c_i + \frac{1}{W^{(m-1)}} \sum_{t=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{j\in\mathcal{J}} p_{jt} \max_{\boldsymbol{y}(t):\sum_{k\in\mathcal{K}} y_{jk}(t)\le 1} \left\{ \sum_{k\in\mathcal{K}} \left(r_{jk} - \sum_{i\in\mathcal{I}} \phi_i(s) v_{ijk} \right) y_{jk}(t) \right\} \right) \\
&\quad - \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{j\in\mathcal{J}} p_{js} r_{j,\kappa(\boldsymbol{\phi}(s),j)} \\
&\le \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \left(\sum_{i\in\mathcal{I}} \phi_i(s) c_i + \sum_{j\in\mathcal{J}} \frac{p_{js}}{W^{(m-1)}} \sum_{t=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \max_{\boldsymbol{y}(t):\sum_{k\in\mathcal{K}} y_{jk}(t)\le 1} \left\{ \sum_{k\in\mathcal{K}} \left(r_{jk} - \sum_{i\in\mathcal{I}} \phi_i(s) v_{ijk} \right) y_{jk}(t) \right\} \right. \\
&\qquad\qquad \left. + \frac{r_{\max}}{W^{(m-1)}} \sum_{t=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{j\in\mathcal{J}} |p_{jt} - p_{js}| \right) - \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{j\in\mathcal{J}} p_{js} r_{j,\kappa(\boldsymbol{\phi}(s),j)} \\
&\le \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \left(\sum_{i\in\mathcal{I}} \phi_i(s) c_i + \sum_{j\in\mathcal{J}} p_{js} \left(r_{j,\kappa(\boldsymbol{\phi}(s),j)} - \sum_{i\in\mathcal{I}} \phi_i(s) v_{i,j,\kappa(\boldsymbol{\phi}(s),j)} \right) \right) - \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{j\in\mathcal{J}} p_{js} r_{j,\kappa(\boldsymbol{\phi}(s),j)} \\
&\quad + \frac{r_{\max}}{W^{(m-1)}} \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{t=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{q=\tau^{(m-2)}+1}^{\tau^{(m-1)}-1} |p_{j,q+1} - p_{jq}| \qquad \text{(SM-4)} \\
&\le \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{i\in\mathcal{I}} \phi_i(s) \left(c_i - \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)} \right) + \gamma W^{(m-1)} V^{(m-1)}. \qquad \text{(SM-5)}
\end{aligned}$$

Step (SM-4) holds by Line 4 in Algorithm 2 that defines $\kappa(\boldsymbol{\phi}(s),j) \in \operatorname{argmax}_{k\in\mathcal{K}} \left\{ r_{jk} - \sum_{i\in\mathcal{I}} \phi_i(s) v_{ijk} \right\}$. Step (SM-5) follows from the definition of $V^{(m-1)}$.

To show (SM-2) holds with probability at least $1-\delta$, we refer to the update rule of the dual prices in (11) and obtain:

$$\begin{aligned}
||\boldsymbol{\phi}(s+1)||_2^2 \le& ||\boldsymbol{\phi}(s) + \frac{1}{\sqrt{2W^{(m-2)}}} (\boldsymbol{v}_{j(s),\tilde{k}(s)} - \boldsymbol{c})||_2^2 \\
\le& ||\boldsymbol{\phi}(s)||_2^2 + \frac{1}{2W^{(m-2)}} ||\boldsymbol{v}_{j(s),\tilde{k}(s)} - \boldsymbol{c}||_2^2 - \frac{2}{\sqrt{2W^{(m-2)}}} \sum_{i\in\mathcal{I}} \phi_i(s) (c_i - v_{i,j(s),\tilde{k}(s)}).
\end{aligned}$$

Rearranging terms, we have

$$\begin{aligned}
\sum_{i\in\mathcal{I}} \phi_i(s) \left(c_i - v_{i,j(s),\tilde{k}(s)} \right) \le& \frac{\sqrt{2W^{(m-2)}}}{2} \left(||\boldsymbol{\phi}(s)||_2^2 - ||\boldsymbol{\phi}(s+1)||_2^2 \right) + \frac{1}{2\sqrt{2W^{(m-2)}}} ||\boldsymbol{v}_{j(s),\tilde{k}(s)} - \boldsymbol{c}||_2^2 \\
\le& \frac{\sqrt{2W^{(m-2)}}}{2} \left(||\boldsymbol{\phi}(s)||_2^2 - ||\boldsymbol{\phi}(s+1)||_2^2 \right) + \frac{|\mathcal{I}|\gamma^2}{2\sqrt{2W^{(m-2)}}}.
\end{aligned}$$

Summing over $s \in [\tau^{(m-2)}+1 : \tau^{(m-1)}]$, we have

$$\begin{aligned}\sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{i\in\mathcal{I}} \phi_i(s)\left(c_i - v_{i,j(s),\tilde{k}(s)}\right) \le& \frac{\sqrt{2W^{(m-2)}}}{2} \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \left(||\boldsymbol{\phi}(s)||_2^2 - ||\boldsymbol{\phi}(s+1)||_2^2\right) + \frac{|\mathcal{I}|\gamma^2 W^{(m-1)}}{2\sqrt{2W^{(m-2)}}}\\ \le& \frac{|\mathcal{I}|\gamma^2 W^{(m-1)}}{2\sqrt{2W^{(m-2)}}}. \end{aligned} \tag{SM-6}$$

The last inequality holds by the telescoping expression. Fix any $\tau \ge \tau^{(m-2)}+1$, consider $\{v_{i,j(s),\tilde{k}(s)} - \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)}\}_{s=\tau^{(m-2)}+1}^{\tau}$, which is a martingale difference sequence with respect to the filtration $\{\mathcal{F}(s)\}_{s=\tau^{(m-2)}+1}^{\tau}$ defined as $\mathcal{F}(s) = \sigma(\{j(t)\}_{t=\tau^{(m-2)}+1}^{s})$. Crucially, the conditional expectation $\mathbb{E}[v_{i,j(s),\tilde{k}(s)}|\mathcal{F}(s-1)] = \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)}$ since $\phi_i(s)$ is $\mathcal{F}(s-1)$-measurable. Therefore, by applying Azuma-Hoeffding Inequality, we have

$$\begin{aligned}&\Pr\left(\sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} v_{i,j(s),\tilde{k}(s)} \ge \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)} + \gamma\sqrt{2W^{(m-1)} \log\frac{2W^{(m-2)}}{\delta}}\right)\\ &\le \Pr\left(\exists\tau \in [\tau^{(m-2)}+1 : \tau^{(m-2)}+2W^{(m-2)}] : \sum_{s=\tau^{(m-2)}+1}^{\tau} v_{i,j(s),\tilde{k}(s)} \ge \sum_{s=\tau^{(m-2)}+1}^{\tau} \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)} \right.\\ &\qquad\left. + \gamma\sqrt{2(\tau-\tau^{(m-2)}))\log\frac{2W^{(m-2)}}{\delta}}\right)\\ &\le \sum_{\tau=1}^{2W^{(m-2)}} \frac{\delta}{2W^{(m-2)}} \le \delta. \end{aligned} \tag{SM-7}$$

Plugging (SM-7) into (SM-6), we have with probability at least $1-\delta$,

$$\begin{aligned}&\sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{i\in\mathcal{I}} \phi_i(s)\left(c_i - \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)}\right)\\ \le& \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{i\in\mathcal{I}} \phi_i(s)\left(c_i - v_{i,j(s),\tilde{k}(s)} + \gamma\sqrt{\frac{2\log\frac{2W^{(m-2)}}{\delta}}{W^{(m-1)}}}\right)\\ \le& \frac{|\mathcal{I}|\gamma^2 W^{(m-1)}}{2\sqrt{2W^{(m-2)}}} + \gamma\sqrt{\frac{2\log\frac{2W^{(m-2)}}{\delta}}{W^{(m-1)}}} \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{i\in\mathcal{I}} \phi_i(s)\\ \le& \frac{|\mathcal{I}|\gamma^2 W^{(m-1)}}{2\sqrt{2W^{(m-2)}}} + \gamma\sqrt{\frac{2\log\frac{2W^{(m-2)}}{\delta}}{W^{(m-1)}}} \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{i\in\mathcal{I}} \left(\gamma + \frac{\gamma}{\sqrt{2W^{(m-2)}}}\right)\\ \le& \frac{|\mathcal{I}|\gamma^2 W^{(m-1)}}{2\sqrt{2W^{(m-2)}}} + |\mathcal{I}|\gamma^2\sqrt{2W^{(m-1)}\log\frac{2W^{(m-2)}}{\delta}} + \frac{|\mathcal{I}|\gamma^2\sqrt{2W^{(m-1)}\log\frac{2W^{(m-2)}}{\delta}}}{\sqrt{2W^{(m-2)}}} \end{aligned}$$

$$\begin{aligned}
&\leq \frac{|\mathcal{I}|\gamma^2 W^{(m-1)}}{2\sqrt{2W^{(m-2)}}} + |\mathcal{I}|\gamma^2\sqrt{2W^{(m-1)}\log\frac{2W^{(m-2)}}{\delta}} + |\mathcal{I}|\gamma^2\sqrt{2\log\frac{2W^{(m-2)}}{\delta}} \qquad \text{(SM-8)}\\
&\leq W^{(m-1)}\epsilon_F^{(m-1)}.
\end{aligned}$$

Inequality (SM-8) holds since $W^{(m-1)} \leq 2W^{(m-2)}$.

Step 2 (validating inequalities (19)): By the update of dual prices (11), we have

$$v_{i,j(s),\tilde{k}(s)} - c_i \leq \sqrt{2W^{(m-2)}}\,(\phi_i(s+1) - \phi_i(s))$$

for any i, s. Summing over $s \in [\tau^{(m-2)}+1 : \tau^{(m-1)}]$, we have

$$\begin{aligned}
\sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} (v_{i,j(s),\tilde{k}(s)} - c_i) &\leq \sqrt{2W^{(m-2)}} \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} (\phi_i(s+1) - \phi_i(s))\\
&\leq \sqrt{2W^{(m-2)}}\phi_i(\tau^{(m-1)}) \leq \gamma\sqrt{2W^{(m-2)}} + \gamma
\end{aligned}$$

with certainty for all $i \in \mathcal{I}$, by Claim 1. By applying Azuma-Hoeffding Inequality in a similar manner as (SM-7), for all $i \in \mathcal{I}$,

$$\begin{aligned}
&\Pr\left(\sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} v_{i,j(s),\tilde{k}(s)} \leq \sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)} - \gamma\sqrt{2W^{(m-1)}\log\frac{2|\mathcal{I}|W^{(m-2)}}{\delta}}\right)\\
\leq &\Pr\Bigg(\exists \tau \in [\tau^{(m-2)}+1 : \tau^{(m-2)} + 2W^{(m-2)}] : \sum_{s=\tau^{(m-2)}+1}^{\tau} v_{i,j(s),\tilde{k}(s)} \leq \sum_{s=\tau^{(m-2)}+1}^{\tau} \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)}\\
&\qquad\qquad - \gamma\sqrt{2W^{(m-1)}\log\frac{2|\mathcal{I}|W^{(m-2)}}{\delta}}\Bigg)\\
\leq &\sum_{\tau=1}^{2W^{(m-2)}} \frac{\delta}{2|\mathcal{I}|W^{(m-2)}} \leq \frac{\delta}{|\mathcal{I}|}.
\end{aligned}$$

Hence, with probability of at least $1 - \frac{\delta}{|\mathcal{I}|}$,

$$\sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \left(\sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)} - c_i\right) \leq \gamma\sqrt{2W^{(m-2)}} + \gamma + \gamma\sqrt{2W^{(m-1)}\log\frac{2|\mathcal{I}|W^{(m-2)}}{\delta}}.$$

Taking a union bound over all $i \in \mathcal{I}$, with probability of at least $1 - \delta$,

$$\begin{aligned}
\sum_{s=\tau^{(m-2)}+1}^{\tau^{(m-1)}} \sum_{j\in\mathcal{J}} p_{js} v_{i,j,\kappa(\boldsymbol{\phi}(s),j)} &\leq W^{(m-1)} c_i + \gamma\sqrt{2W^{(m-2)}} + \gamma + \gamma\sqrt{2W^{(m-1)}\log\frac{2|\mathcal{I}|W^{(m-2)}}{\delta}} \quad \forall i \in \mathcal{I}\\
&\leq W^{(m-1)} c_i (1 + \epsilon_E^{(m-1)}) \quad \forall i \in \mathcal{I}.
\end{aligned}$$

$\square$

2. Proof of Lemma 2 Recall that τ is an arbitrary but fixed time index lying in $[s+1,t]$. The proof of the Lemma consists of lower bounding $\hat{\lambda}_*^{S(s+1:t)}$ in terms of $\lambda_*^{S(\tau)}$, and lower bounding $\lambda_*^{S(\tau)}$ in terms of $\hat{\lambda}_*^{S(s+1:t)}$.

Lower bounding $\hat{\lambda}_*^{S(s+1:t)}$ in terms of $\lambda_t^{S(\tau)}$. For notational ease, we define

$$\tilde{\epsilon}_A^{(s+1:t)} = \frac{\epsilon_A^{(s+1:t)}}{\gamma} = \frac{\gamma}{c_{\min}} V^{(s+1:t)} + \frac{2\gamma}{c_{\min}} \sqrt{\frac{2}{t-s} \log \frac{|\mathcal{I}|}{\delta}} + \frac{2\gamma}{c_{\min}(t-s)} \log \frac{|\mathcal{I}|}{\delta},$$

and recall that

$$\hat{p}_j^{(s+1:t)} = \sum_{q=s+1}^{t} \frac{\mathbf{1}(j(q)=j)}{t-s}.$$

Consider $\hat{y}_{jk}^{(s+1:t)} = \frac{y_{jk}^*(\tau)}{1+\tilde{\epsilon}_A^{(s+1:t)}}$, where $\{y_{jk}^*(\tau)\}_{j,k}$ is an optimal solution to the fluid relaxation LP-S(t). We claim that $\hat{y}_{jk}^{(s+1:t)}$ is feasible to LP-RS$^{(s+1:t)}$ with probability at least $1-\delta$, and the inequality

$$\lambda_*^{S(\tau)} \leq (1+\tilde{\epsilon}_A^{(s+1:t)}) \sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}} \hat{p}_j^{(s+1:t)} r_{ijk} \hat{y}_{jk}^{(s+1:t)} + \epsilon_B^{(s+1:t)} \tag{SM-9}$$

holds with probability at least $1-\delta/|\mathcal{I}|$. Combining the two claims tells us that the inequalities

$$\lambda_*^{S(\tau)} \leq \hat{\lambda}_*^{S(s+1:t)} (1+\tilde{\epsilon}_A^{(s+1:t)}) + \epsilon_B^{(s+1:t)} \leq \hat{\lambda}_*^{S(s+1:t)} + \epsilon_A^{(s+1:t)} + \epsilon_B^{(s+1:t)}$$

hold with probability $\geq 1-2\delta/|\mathcal{I}|$. To complete the proof for the lower bounding, it suffices to establish (a) inequality (SM-9) and (b) the feasibility of $\hat{y}_{jk}^{(s+1:t)}$.

(a) Inequality (SM-9). We first lower bound the following expected reward, where the expectation is over the randomness in the empirical estimate $\hat{p}_j^{(s+1:t)}$:

$$\begin{aligned}
&\mathbb{E}\left[\sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}} \hat{p}_j^{(s+1:t)} r_{ijk} \hat{y}_{jk}^{(s+1:t)}\right] \\
=&\mathbb{E}\left[\sum_{q=s+1}^{t} \sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}} \frac{\mathbf{1}(j(q)=j)}{(t-s)\left(1+\tilde{\epsilon}_A^{(s+1:t)}\right)} r_{ijk} y_{jk}^*(\tau)\right] \\
\geq&\sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}} p_{j\tau} r_{ijk} \frac{y_{jk}^*(\tau)}{1+\tilde{\epsilon}_A^{(s+1:t)}} - \sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}} \left(p_{j\tau} - \sum_{q=s+1}^{t} \frac{p_{jq}}{t-s}\right) \frac{r_{ijk} y_{jk}^*(\tau)}{1+\tilde{\epsilon}_A^{(s+1:t)}} \\
\geq&\frac{\lambda_*^{S(\tau)} - r_{\max} V^{(s+1:t)}}{1+\tilde{\epsilon}_A^{(s+1:t)}} \geq \frac{\lambda_*^{S(\tau)} - \gamma V^{(s+1:t)}}{1+\tilde{\epsilon}_A^{(s+1:t)}}.
\end{aligned}$$

In passing, we note that the lemma assumption of $\tau \in [s+1, t]$ is crucial for the bound $\sum_{j\in\mathcal{J}}\left(p_{j\tau} - \sum_{q=s+1}^{t}\frac{p_{jq}}{t-s}\right) \le V^{(s+1,t)}$ to hold. Now, since $\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\frac{\mathbf{1}(j(q)=j)}{1+\tilde{\epsilon}_A^{(s+1:t)}} r_{ijk} y^*_{jk}(\tau) \le \frac{r_{\max}}{1+\tilde{\epsilon}_A^{(s+1:t)}}$, applying the multiplicative Chernoff bounds (from Lemma 8 in Appendix B), we have

$$\begin{aligned}
\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\hat{p}_j^{(s+1:t)} r_{ijk}\hat{y}_{jk}^{(s+1:t)} \ge& \frac{\lambda_*^{S(\tau)} - \gamma V^{(s+1:t)}}{1+\tilde{\epsilon}_A^{(s+1:t)}} - \frac{1}{1+\tilde{\epsilon}_A^{(s+1:t)}}\sqrt{\frac{2\lambda_*^{S(\tau)} r_{\max}}{t-s}\log\frac{T}{\delta}} \quad \text{w.p. } 1-\frac{\delta}{T} \\
\ge& \frac{\lambda_*^{S(\tau)} - \gamma V^{(s+1:t)}}{1+\tilde{\epsilon}_A^{(s+1:t)}} - \frac{\gamma}{1+\tilde{\epsilon}_A^{(s+1:t)}}\sqrt{\frac{2}{t-s}\log\frac{T}{\delta}} \\
=& \frac{\lambda_*^{S(\tau)} - \epsilon_B^{(s+1:t)}}{1+\tilde{\epsilon}_A^{(s+1:t)}}. \qquad \text{(SM-11)}
\end{aligned}$$

Rearranging (SM-11) gives us (SM-9).

(b) Feasibility of $\hat{y}_{jk}^{(s+1:t)}$. Similar to part (a), we first bound the expectation:

$$\begin{aligned}
&\mathbb{E}\left[\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\hat{p}_j^{(s+1:t)} v_{ijk}\hat{y}_{jk}^{(s+1:t)}\right] \\
=&\mathbb{E}\left[\sum_{q=s+1}^{t}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\frac{\mathbf{1}(j(q)=j)}{(t-s)\left(1+\tilde{\epsilon}_A^{(s+1:t)}\right)} v_{ijk} y^*_{jk}(\tau)\right] \\
\le&\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} p_{j\tau} v_{ijk}\frac{y^*_{jk}(\tau)}{1+\tilde{\epsilon}_A^{(s+1:t)}} + \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\left(\sum_{q=s+1}^{t}\frac{p_{jq}}{t-s} - p_{j\tau}\right)\frac{v_{ijk} y^*_{jk}(\tau)}{1+\tilde{\epsilon}_A^{(s+1:t)}} \\
\le&\frac{c_i + v_{\max} V^{(s+1:t)}}{1+\tilde{\epsilon}_A^{(s+1:t)}} \le \frac{c_i + \gamma V^{(s+1:t)}}{1+\tilde{\epsilon}_A^{(s+1:t)}}.
\end{aligned}$$

Since $\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\frac{\mathbf{1}(j(t)=j)}{1+\tilde{\epsilon}_A^{(s+1:t)}} v_{ijk} y^*_{jk}(t) \le \frac{\gamma}{1+\tilde{\epsilon}_A^{(s+1:t)}}$, applying the multiplicative Chernoff bounds in Lemma 8 in Appendix B, we have

$$\begin{aligned}
&\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\hat{p}_j^{(s+1:t)} v_{ijk}\hat{y}_{jk}^{(s+1:t)} \\
\le&\frac{c_i + \gamma V^{(s+1:t)}}{1+\tilde{\epsilon}_A^{(s+1:t)}} + \frac{2\gamma}{1+\tilde{\epsilon}_A^{(s+1:t)}}\sqrt{\frac{2}{t-s}\log\frac{|\mathcal{I}|}{\delta}} + \frac{2\gamma}{(t-s)\left(1+\tilde{\epsilon}_A^{(s+1:t)}\right)}\log\frac{|\mathcal{I}|}{\delta} \quad \text{w.p. } 1-\frac{\delta}{|\mathcal{I}|} \\
\le& c_i\frac{1+\tilde{\epsilon}_A^{(s+1:t)}}{1+\tilde{\epsilon}_A^{(s+1:t)}} = c_i \quad \text{almost surely.} \qquad \text{(SM-12)}
\end{aligned}$$

Taking the union bound of (SM-12) yields the desired feasibility.

Lower bounding $\hat{\lambda}_*^{S(s+1:t)}$ in terms of $\lambda_\tau^{S(\tau)}$. The proof involves considering the following dual formulation of (LP-S(τ)):

$$
\begin{aligned}
\text{(LP-S-D}(\tau)) : \quad & \min_{\boldsymbol{\alpha}_\tau^{(2)},\boldsymbol{\beta}_\tau^{(2)}} \sum_{j\in\mathcal{J}} p_{j\tau}\beta_{j\tau}^{(2)} + \sum_{i\in\mathcal{I}} c_i\alpha_{i\tau}^{(2)} & \\
& \text{s.t. } \beta_{j\tau}^{(2)} + \sum_{i\in\mathcal{I}} v_{ijk}\alpha_{i\tau}^{(2)} \geq r_{jk} & \forall j\in\mathcal{J},\ k\in\mathcal{K} \\
& \alpha_{i\tau}^{(2)} \geq 0 & \forall i\in\mathcal{I} \\
& \beta_{j\tau}^{(2)} \geq 0 & \forall j\in\mathcal{J}
\end{aligned}
$$

Similarly, we consider the following dual formulation to the sample average approximation model (LP-RS)$^{(s+1:t)}$:

$$
\begin{aligned}
\text{(LP-RS-D)}^{(s+1:t)} : \quad & \min_{\boldsymbol{\alpha}^{(2)},\boldsymbol{\beta}^{(2)}} \sum_{j\in\mathcal{J}} \hat{p}_j^{(s+1:t)}\beta_j^{(s+1:t)} + \sum_{i\in\mathcal{I}} c_i\alpha_i^{(s+1:t)} & \\
& \text{s.t. } \beta_j^{(s+1:t)} + \sum_{i\in\mathcal{I}} v_{ijk}\alpha_i^{(s+1:t)} \geq r_{jk} & \forall j\in\mathcal{J},\ k\in\mathcal{K} \\
& \alpha_i^{(s+1:t)} \geq 0 & \forall i\in\mathcal{I} \\
& \beta_j^{(s+1:t)} \geq 0 & \forall j\in\mathcal{J}
\end{aligned}
$$

Our analysis involves considering $(\boldsymbol{\alpha}_\tau^{(2)*},\boldsymbol{\beta}_\tau^{(2)*})$, an optimal solution to (LP-S-D(τ)). To proceed, we first note that for all j,τ, it holds that $\beta_{j\tau}^{(2)*}\leq r_{\max}\leq\gamma$, by optimality. In addition, $(\boldsymbol{\alpha}_\tau^{(2)*},\boldsymbol{\beta}_\tau^{(2)*})$ is feasible to (LP-RS-D)$^{(s+1:t)}$, since (LP-S(τ)), (LP-RS-D)$^{(s+1:t)}$ have the same feasible region. Consequently, we derive the following bounds, where all inequalities and equalites except (SM-13) holds with certainty:

$$
\begin{aligned}
& \hat{\lambda}_*^{S(s+1:t)} \\
\leq & \sum_{j\in\mathcal{J}} \hat{p}_j^{(s+1:t)}\beta_{j\tau}^{(2)*} + \sum_{i\in\mathcal{I}} c_i\alpha_{i\tau}^{(2)*} \\
\leq & \sum_{j\in\mathcal{J}} \left[\frac{1}{t-s}\sum_{q=s+1}^{t} p_{js}\right]\beta_{j\tau}^{(2)*} + \sum_{i\in\mathcal{I}} c_i\alpha_{i\tau}^{(2)*} + 2r_{\max}\sqrt{\frac{2}{t-s}\log\frac{1}{\delta}} + \frac{2r_{\max}}{t-s}\log\frac{1}{\delta} \quad \text{w.p. } 1-\delta \qquad \text{(SM-13)} \\
\leq & \sum_{j\in\mathcal{J}} p_{j\tau}\beta_{j\tau}^{(2)*} + r_{\max}V^{(s+1:t)} + \sum_{i\in\mathcal{I}} c_i\alpha_{i\tau}^{(2)*} + 2r_{\max}\sqrt{\frac{2}{t-s}\log\frac{1}{\delta}} + \frac{2r_{\max}}{t-s}\log\frac{1}{\delta} \\
\leq & \lambda_\tau^{S(\tau)} + \epsilon_C^{(s+1:t)}.
\end{aligned}
$$

Step (SM-13) is by the multiplicative Chernoff bounds in Lemma 8 in Appendix B. Altogether, the desired lower bounds are established and the Lemma is proved. □

3. Proof of Theorem 2

Our proof consists of 5 steps. In step 1, we construct a linear program (LP-E) with its optimum denoted as λ_*^E, which is a fluid approximation of (IP-C). We show

$$\lambda_*^E \geq \text{opt(IP-C)},$$

and therefore, λ_*^E is an upper bound of the expected reward achieved by any non-anticipatory policy. In the following steps, to upper bound opt(IP-C) with $\lambda_*^{S(t)}$, it suffices to upper bound λ_*^E using $\lambda_*^{S(t)}$. In step 2, we partition the planning horizon into episodes of length W, and demonstrate that the sum of optimums of "episode-wise (LP-E)s" is no smaller than $T\lambda_*^E$. In step 3, we construct duals for each "episode-wise (LP-E)" and (LP-S)$_t$. In step 4, we define an intermediate quantity based on the duals constructed in the previous step. By comparing this intermediate quantity with the objective values of both duals, we bridge the gap between the optimal objective of each "episode-wise (LP-E)" and the sum of $\lambda_*^{S(t)}$ over each episode. In the last step, we sum up the results of all episodes and derive our Theorem.

Step 1: The intractability of opt(IP-C) motivates us to consider a fluid approximation, dubbed (LP-E). We provide (LP-E) in the following, where the realization of the customer arrivals, their usage duration and outcomes exactly follow the expectation. The linear program (LP-E) is

$$\begin{aligned}
&\max \frac{1}{T}\sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} p_{jt} r_{jk} x_{jk}(t) \\
&\text{s.t.} \sum_{\tau=\max\{1,t-d_{\max}\}}^{t}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} p_{j\tau}\mathbb{E}[A_{ijk}\mathbf{1}(D_{ijk}\geq t-\tau+1)]x_{jk}(\tau)\leq c_i && \forall i\in\mathcal{I},\ t\in[T] \\
&\sum_{k\in\mathcal{K}} x_{jk}(t)\leq 1 && \forall j\in\mathcal{J},\ t\in[T] \\
&x_{jk}(t)\geq 0 && \forall j\in\mathcal{J},\ k\in\mathcal{K},\ t\in[T].
\end{aligned}$$

We denote the optimal objective value of (LP-E) as λ_*^E. We next show $\lambda_*^E \geq$ opt(IP-C).

Let π be a non-anticipatory feasible policy that achieves the expected optimum in (IP-C), i.e. $\mathbb{E}[\sum_{t=1}^{T}\sum_{k\in\mathcal{K}} R_{j(t),k}(t)X_k^\pi(t)] = \mathbb{E}[T\lambda_*^C]$. Define $x = \{x_{jk}(t)\}_{j\in\mathcal{J},k\in\mathcal{K},t\in\{1,\ldots,T\}}$ where $x_{jk}(t) = \Pr(X_k^\pi(t) = 1 | j(t) = j)$. We claim that x is feasible to (LP-E), with objective value equal to $\mathbb{E}[\sum_{t=1}^{T}\sum_{k\in\mathcal{K}} R_{j(t),k}(t)X_k^\pi(t)] = \mathbb{E}[T\lambda_*^C]$. Thus, verifying the claims about the feasibility and the objective value proves the claim. We first verify the feasibility to (LP-E). Since the policy π satisfies the reusable resource constraints, the inequality $\sum_{\tau=1}^{t}\sum_{k\in\mathcal{K}}\mathbf{1}\left(D_{ij(\tau)k}(\tau)\geq t-\tau+1\right)A_{ij(\tau)k}(\tau)X_k^\pi(\tau)\leq c_i$ holds for all $i\in\mathcal{I}, t\in[T]$. Taking expectation (which is over $X_k^\pi(\tau)$, $D_{ijk}(\tau)$, $A_{ij(\tau)k}(\tau)$ and $j(\tau)$) for $\tau = 1, \dots, t$ on the left hand side gives

$$\begin{aligned}
&\mathbb{E}\left[\sum_{\tau=1}^{t}\sum_{k\in\mathcal{K}}\mathbf{1}\left(D_{ij(\tau)k}(\tau)\geq t-\tau+1\right)A_{ij(\tau)k}(\tau)\,X_k^\pi(\tau)\right]\\
=&\sum_{\tau=1}^{t}\sum_{k\in\mathcal{K}}\mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[\mathbf{1}\left(D_{ij(\tau)k}(\tau)\geq t-\tau+1\right)A_{ij(\tau)k}(\tau)\mid j(\tau),X_k^\pi(\tau)\right]X_k^\pi(\tau)\mid j(\tau)\right]\right]\\
=&\sum_{\tau=1}^{t}\sum_{k\in\mathcal{K}}\sum_{j\in\mathcal{J}}p_{j\tau}\mathbb{E}\left[\mathbb{E}\left[\mathbf{1}\left(D_{ij(\tau)k}(\tau)\geq t-\tau+1\right)A_{ij(\tau)k}(\tau)\mid j(\tau)=j,X_k^\pi(\tau)\right]X_k^\pi(\tau)\mid j(t)=j\right]\\
=&\sum_{\tau=1}^{t}\sum_{k\in\mathcal{K}}\sum_{j\in\mathcal{J}}p_{j\tau}\mathbb{E}\left[\mathbf{1}\left(D_{ijk}\geq t-\tau+1\right)A_{ijk}\right]x_{jk}(\tau)\leq c_i
\end{aligned}$$

Similarly, by taking expectation over each of the reward constraints, we have

$$\mathbb{E}\left[\sum_{t=1}^{T}\sum_{k\in\mathcal{K}}R_{j(t)k}(t)X_k^\pi(t)\right]=\sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}p_{jt}r_{jk}x_{jk}(t).$$

Hence, the claim about the objective value is shown, and we have proved $\lambda_*^E \geq$ opt(IP-C).

Step 2: For an arbitrary $W \in \{d_{\max}, \dots, T\}$, we partition the planning horizon into $\frac{T}{W}$ episodes each consisting of W time steps. We define the episode-wise DM's expected LPs for each $m \in [1, \dots, \frac{T}{W}]$ as follows:

$$\begin{aligned}
&\text{(LP-EW)}^{(m)}\ \max\ \sum_{t=(m-1)W+1}^{mW}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}p_{jt}r_{jk}z_{jk}(t)\\
&\text{s.t.}\ \sum_{\tau=\max\{(m-1)W+1,t-d_{\max}\}}^{t}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}p_{j\tau}\mathbb{E}[A_{ijk}\mathbf{1}(D_{ijk}\geq t-\tau+1)]\quad z_{jk}(\tau)\leq c_i\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\forall i\in\mathcal{I},\ t\in[(m-1)W+1:mW]\\
&\sum_{k\in\mathcal{K}}z_{jk}(t)\leq 1\qquad\qquad\qquad\qquad\qquad\forall j\in\mathcal{J},\ t\in[(m-1)W+1:mW]\\
&z_{jk}(t)\geq 0\qquad\qquad\qquad\qquad\qquad\qquad\forall j\in\mathcal{J},\ k\in\mathcal{K},\ t\in[(m-1)W+1:mW].
\end{aligned}$$

Let the optimum of (LP-EW)$^{(m)}$ be $W\lambda_*^{EW(m)}$ and let the corresponding optimal solution be $z_{jk}^*(t)$. To compare $\sum_{m=1}^{T/W} W\lambda^{EW(m)}$ (i.e., sum of optimal objective values of (LP-EW)$^{(m)}$ over all episodes) and $T\lambda_*^E$ (i.e., optimal objective value of (LP-E($[1:T]$))), we observe that $\bar{z}_{jk}(t) = x_{jk}^*(t)\ \forall t \in [(m-1)W + 1 : mW]$ is a feasible solution for each (LP-EW)$^{(m)}$, and

$$\sum_{m=1}^{T/W}W\lambda^{EW(m)}=\sum_{m=1}^{T/W}\sum_{t=(m-1)W+1}^{mW}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}p_{jt}r_{jk}\bar{z}_{jk}(t)$$

$$
\begin{aligned}
&=\sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} p_{jt}r_{jk}x^*_{jk}(t)\\
&=T\lambda^E_*.
\end{aligned}
$$

Hence, we have $T\lambda^E_* \le \sum_{m=1}^{T/W} W\lambda^{EW(m)}_*$.

Step 3: In this step, we first construct the duals of (LP-EW)$^{(m)}$ and (LP-S(t)). We firstly present the dual of (LP-EW)$^{(m)}$:

$$
\text{(LP-EW-D0)}^{(m)}: \min_{\boldsymbol{\alpha}^{(0)},\boldsymbol{\beta}^{(0)}} \sum_{t=(m-1)W+1}^{mW}\left(\sum_{j\in\mathcal{J}}\beta^{(0)}_{jt}+\sum_{i\in\mathcal{I}}c_i\alpha^{(0)}_{it}\right)
$$

$$
\begin{aligned}
\text{s.t. }\ &\beta^{(0)}_{jt}+\sum_{i\in\mathcal{I}}p_{jt}\sum_{\tau=t}^{\min\{t+d_{\max},mW\}}\mathbb{E}[A_{ijk}\mathbf{1}(D_{ijk}\ge\tau-t+1)]\alpha^{(0)}_{i\tau}\ge p_{jt}r_{jk} && \forall j\in\mathcal{J},\ k\in\mathcal{K},\\
& && t\in[(m-1)W+1:mW]\\
&\alpha^{(0)}_{it}\ge 0 && \forall i\in\mathcal{I},\ t\in[(m-1)W+1:mW]\\
&\beta^{(0)}_{jt}\ge 0 && \forall j\in\mathcal{J},\ t\in[(m-1)W+1:mW].
\end{aligned}
$$

The optimal objective value of (LP-EW-D0)$^{(m)}$ is $W\lambda^{EW(m)}_*$, which is the same as its primal's. Let $(\boldsymbol{\alpha}^{(0)*},\boldsymbol{\beta}^{(0)*})$ denote the optimal solution of (LP-EW-D0)$^{(m)}$. We assert that the following LP achieves the same optimal objective value as (LP-EW-D0)$^{(m)}$:

$$
\text{(LP-EW-D}(m)): \min_{\boldsymbol{\alpha}^{(1)},\boldsymbol{\beta}^{(1)}} \sum_{t=(m-1)W+1}^{mW}\left(\sum_{j\in\mathcal{J}}p_{jt}\beta^{(1)}_{jt}+\sum_{i\in\mathcal{I}}c_i\alpha^{(1)}_{it}\right)
$$

$$
\begin{aligned}
\text{s.t. }\ &\beta^{(1)}_{jt}+\sum_{i\in\mathcal{I}}\sum_{\tau=t}^{\min\{t+d_{\max},mW\}}\mathbb{E}[A_{ijk}\mathbf{1}(D_{ijk}\ge\tau-t+1)]\alpha^{(1)}_{i\tau}\ge r_{jk} && \forall j\in\mathcal{J},\ k\in\mathcal{K},\\
& && t\in[(m-1)W+1:mW]\\
&\alpha^{(1)}_{it}\ge 0 && \forall i\in\mathcal{I},\ t\in[(m-1)W+1:mW]\\
&\beta^{(1)}_{jt}\ge 0 && \forall j\in\mathcal{J},\ t\in[(m-1)W+1:mW].
\end{aligned}
$$

To verify our assertion, we first claim that $W\lambda^{EW(m)}_* \le$ opt((LP-EW-D)$^{(m)}$). Letting $(\boldsymbol{\alpha}^{(1)*},\boldsymbol{\beta}^{(1)*})$ be an optimal solution to (LP-EW-D)$^{(m)}$, we observe that the solution $(\bar{\boldsymbol{\alpha}}^{(0)},\bar{\boldsymbol{\beta}}^{(0)})$, defined as

$$
\bar{\alpha}^{(0)}_{it}=\alpha^{(1)*}_{it}\quad \forall i\in\mathcal{I},\ t\in[(m-1)W+1:mW]
$$

$$
\bar{\beta}^{(0)}_{jt}=p_{jt}\beta^{(1)*}_{jt}\quad \forall j\in\mathcal{J}\ t\in[(m-1)W+1:mW]
$$

is a feasible solution to (LP-EW-D0)$^{(m)}$. In addition, we have

$$\sum_{t=(m-1)W+1}^{mW}\left(\sum_{j\in\mathcal{J}} p_{jt}\beta^{(1)*}_{jt}+\sum_{i\in\mathcal{I}} c_i\alpha^{(1)*}_{it}\right)=\sum_{t=(m-1)W+1}^{mW}\left(\sum_{j\in\mathcal{J}}\bar{\beta}^{(0)}_{jt}+\sum_{i\in\mathcal{I}} c_i\bar{\alpha}^{(0)}_{it}\right).$$

Hence the claimed inequality is shown.

We also claim that $W\lambda^{EW(m)}_{*}\geq$ opt((LP-EW-D)$^{(m)}$). Observe that the solution $(\bar{\boldsymbol{\alpha}}^{(1)},\bar{\boldsymbol{\beta}}^{(1)})$, defined as

$$\bar{\alpha}^{(1)}_{it}=\alpha^{(0)*}_{it}\quad \forall i\in\mathcal{I},\ t\in[(m-1)W+1:mW]$$

and

$$\bar{\beta}^{(1)}_{jt}=\begin{cases}\beta^{(0)*}_{jt}/p_{jt} & \text{if } p_{jt}>0\\ r_{\max} & \text{if } p_{jt}=0\end{cases}\qquad \forall j\in\mathcal{J},\ t\in[(m-1)W+1:mW]$$

is a feasible solution to (LP-EW-D0)$^{(m)}$. We further observe that if $p_{jt}=0$ for some j,t, then $\beta^{(0)*}_{jt}=0$ since the first set of constraints on $\beta^{(0)}_{jt}$ reduces to $\beta^{(0)}_{jt}\geq 0$. Therefore, we have

$$\begin{aligned}&\sum_{t=(m-1)W+1}^{mW}\left(\sum_{j\in\mathcal{J}} p_{jt}\bar{\beta}^{(1)*}_{jt}+\sum_{i\in\mathcal{I}} c_i\bar{\alpha}^{(1)*}_{it}\right)\\
=&\sum_{t=(m-1)W+1}^{mW}\left(\sum_{j\in\mathcal{J}}\left(\mathbf{1}(p_{jt}>0)\cdot p_{jt}\frac{\beta^{(0)*}_{jt}}{p_{jt}}+\mathbf{1}(p_{jt}=0)\cdot p_{jt}r_{\max}\right)+\sum_{i\in\mathcal{I}} c_i\alpha^{(0)*}_{it}\right)\\
=&\sum_{t=(m-1)W+1}^{mW}\left(\sum_{j\in\mathcal{J}}\left(\mathbf{1}(p_{jt}>0)\cdot\beta^{(0)*}_{jt}+\mathbf{1}(p_{jt}=0)\cdot 0\right)+\sum_{i\in\mathcal{I}} c_i\alpha^{(0)*}_{it}\right)\\
=&\sum_{t=(m-1)W+1}^{mW}\left(\sum_{j\in\mathcal{J}}\left(\mathbf{1}(p_{jt}>0)\cdot\beta^{(0)*}_{jt}+\mathbf{1}(p_{jt}=0)\cdot\beta^{(0)*}_{jt}\right)+\sum_{i\in\mathcal{I}} c_i\alpha^{(0)*}_{it}\right)\\
=&\sum_{t=(m-1)W+1}^{mW}\left(\sum_{j\in\mathcal{J}}\beta^{(0)*}_{jt}+\sum_{i\in\mathcal{I}} c_i\alpha^{(0)*}_{it}\right),\end{aligned}$$

and the claimed inequality is verified. Since (LP-EW-D)$^{(m)}$ has a simpler form and we only make use of the objective values for comparison between LPs, we will refer to (LP-EW-D)$^{(m)}$ as the dual of (LP-EW(m)) in forthcoming analysis.

We next present the dual of (LP-S)$_t$. For notational ease, we recycle the variables $(\boldsymbol{\alpha}^{(0)*},\boldsymbol{\beta}^{(0)*})$ from (LP-EW-D0)$^{(m)}$ as (LP-EW-D0)$^{(m)}$ is no longer useful:

$$\text{(LP-S-D0}(t)\text{)}:\ \min_{\boldsymbol{\alpha}^{(0)}_t,\boldsymbol{\beta}^{(0)}_t}\ \sum_{j\in\mathcal{J}}\beta^{(0)}_{jt}+\sum_{i\in\mathcal{I}} c_i\alpha^{(0)}_{it}$$

$$
\begin{aligned}
\text{s.t. } & \beta_{jt}^{(0)} + \sum_{i \in \mathcal{I}} p_{jt} v_{ijk} \alpha_{it}^{(0)} \geq p_{jt} r_{jk} && \forall j \in \mathcal{J},\ k \in \mathcal{K} \\
& \alpha_{it}^{(0)} \geq 0 && \forall i \in \mathcal{I} \\
& \beta_{jt}^{(0)} \geq 0 && \forall j \in \mathcal{J}
\end{aligned}
$$

The optimal objective value of (LP-S-D0(t)) is $\lambda_*^{S(t)}$, which can be shown to be the same with the following LP:

$$
\begin{aligned}
\text{(LP-S-D}(t)\text{)}: \quad \min_{\boldsymbol{\alpha}_t^{(2)}, \boldsymbol{\beta}_t^{(2)}} \quad & \sum_{j \in \mathcal{J}} p_{jt} \beta_{jt}^{(2)} + \sum_{i \in \mathcal{I}} c_i \alpha_{it}^{(2)} \\
\text{s.t. } & \beta_{jt}^{(2)} + \sum_{i \in \mathcal{I}} v_{ijk} \alpha_{it}^{(2)} \geq r_{jk} && \forall j \in \mathcal{J},\ k \in \mathcal{K} \\
& \alpha_{it}^{(2)} \geq 0 && \forall i \in \mathcal{I} \\
& \beta_{jt}^{(2)} \geq 0 && \forall j \in \mathcal{J}.
\end{aligned}
$$

We prove the above LP has the same optimal objective value as (LP-S-D0(t)) for all t in a similar manner as showing (LP-EW-D)$^{(m)}$ has the same optimal objective value as (LP-EW-D0)$^{(m)}$. We first claim that $\lambda_*^{S(t)} \leq$ opt((LP-S-D(t))). For any $t \in [T]$, letting $(\boldsymbol{\alpha}^{(2)*}, \boldsymbol{\beta}^{(2)*})$ be an optimal solution to (LP-S-D(t)), we observe that the solution $(\bar{\boldsymbol{\alpha}}^{(0)}, \bar{\boldsymbol{\beta}}^{(0)})$, defined as

$$\bar{\alpha}_{it}^{(0)} = \alpha_{it}^{(2)*} \quad \forall i \in \mathcal{I}$$

$$\bar{\beta}_{jt}^{(0)} = p_{jt} \beta_{jt}^{(2)*} \quad \forall j \in \mathcal{J}$$

is a feasible solution to (LP-S-D0(t)). For the objective, we have

$$\sum_{j \in \mathcal{J}} p_{jt} \beta_{jt}^{(2)*} + \sum_{i \in \mathcal{I}} c_i \alpha_{it}^{(2)*} = \sum_{j \in \mathcal{J}} \bar{\beta}_{jt}^{(0)} + \sum_{i \in \mathcal{I}} c_i \bar{\alpha}_{it}^{(0)}$$

for each t. Hence the claimed inequality is shown.

We also claim that $\lambda_*^{S(t)} \geq$ opt((LP-S-D(t))). Let $(\boldsymbol{\alpha}^{(0)*}, \boldsymbol{\beta}^{(0)*})$ denote the optimal solution of (LP-S-D0(t)). Observe that the solution $(\bar{\boldsymbol{\alpha}}^{(1)}, \bar{\boldsymbol{\beta}}^{(1)})$ for each $t \in [T]$, defined as

$$\bar{\alpha}_{it}^{(1)} = \alpha_{it}^{(0)*} \quad \forall i \in \mathcal{I}$$

and

$$\bar{\beta}_{jt}^{(1)} = \begin{cases} \beta_{jt}^{(0)*} / p_{jt} & \text{if } p_{jt} > 0 \\ r_{\max} & \text{if } p_{jt} = 0 \end{cases} \quad \forall j \in \mathcal{J}$$

is a feasible solution to (LP-S-D0(t)). We further observe that if $p_{jt}=0$ for some j,t, then $\beta^{(0)*}_{jt}=0$ since the first set of constraints on $\beta^{(0)}_{jt}$ reduces to $\beta^{(0)}_{jt}\geq 0$. Therefore, we have

$$\begin{aligned}\sum_{j\in\mathcal{J}} p_{jt}\bar{\beta}^{(2)*}_{jt}+\sum_{i\in\mathcal{I}} c_i\bar{\alpha}^{(2)*}_{it} &= \sum_{j\in\mathcal{J}}\left(\mathbf{1}(p_{jt}>0)\cdot p_{jt}\frac{\beta^{(0)*}_{jt}}{p_{jt}}+\mathbf{1}(p_{jt}=0)\cdot p_{jt}r_{\max}\right)+\sum_{i\in\mathcal{I}} c_i\alpha^{(0)*}_{it}\\ &= \sum_{j\in\mathcal{J}}\left(\mathbf{1}(p_{jt}>0)\cdot \beta^{(0)*}_{jt}+\mathbf{1}(p_{jt}=0)\cdot 0\right)+\sum_{i\in\mathcal{I}} c_i\alpha^{(0)*}_{it}\\ &= \sum_{j\in\mathcal{J}}\beta^{(0)*}_{jt}+\sum_{i\in\mathcal{I}} c_i\alpha^{(0)*}_{it},\end{aligned}$$

and the claimed inequality is verified. In further analysis, we will refer to (LP-S-D(t)) as the dual of (LP-S(t)).

Step 4: To proceed, for any m and $s\in[(m-1)W+1:mW]$ we define

$$W\lambda^{EWS(m,s)} := \sum_{t=(m-1)W+1}^{mW}\left(\sum_{j\in\mathcal{J}} p_{jt}\beta^{(2)*}_{js}+\sum_{i\in\mathcal{I}} c_i\alpha^{(2)*}_{is}\right),$$

and assert that it holds for all $m\in\left[\frac{T}{W}\right]$,

$$W\lambda^{EW(m)}_{*}-r_{max}d_{max}\overset{(\dagger)}{\leq}\sum_{s=(m-1)W+1}^{MW}\lambda^{EWS(m,s)}\overset{(\ddagger)}{\leq}\sum_{s=(m-1)W+1}^{MW}\lambda^{S(s)}_{*}+r_{\max}W\sum_{s=(m-1)W+1}^{MW}||p_{s+1}-p_s||_1.$$

To demonstrate (†), we compare the optimum values of (LP-S-D(s)) and (LP-EW-D)$^{(m)}$. For any m, we fix an optimal solution $(\boldsymbol{\alpha}^{(2)*}_s,\boldsymbol{\beta}^{(2)*}_s)$ of (LP-S-D(s)) for some $s\in[(m-1)W+1:mW]$. We claim that the solution $(\bar{\alpha}^{(1)},\bar{\beta}^{(1)})$, which is defined as

$$\bar{\alpha}^{(1)}_{it}=\alpha^{(2)*}_{is}\quad \forall i\in\mathcal{I},\ t\in[(m-1)W+1:mW]$$

and

$$\bar{\beta}^{(1)}_{j,t}=\begin{cases}\beta^{(2)*}_{js} & \forall j\in\mathcal{J},\ t\in[(m-1)W+1:mW-d_{\max}]\\ r_{\max} & \forall j\in\mathcal{J}, t\in[mW-d_{\max}+1:mW]\end{cases},$$

is feasible for (LP-EW-D)$^{(m)}$. To verify the feasibility, it suffices to verify for the first set of constraints in (LP-EW-D)$^{(m)}$, since it is evident that $\bar{\alpha}^{(1)},\bar{\beta}^{(1)}\geq 0$. For any $j\in\mathcal{J},k\in\mathcal{K}$ and $t\in[(m-1)W+1:mW-d_{\max}]$, we have

$$\beta^{(1)}_{jt}+\sum_{i\in\mathcal{I}}\sum_{\tau=t}^{\min\{t+d_{\max},mW\}}\mathbb{E}[A_{ijk}\mathbf{1}(D_{ijk}\geq\tau-t+1)]\alpha^{(1)}_{i\tau}$$

$$
\begin{aligned}
&=\beta_{js}^{(2)*}+\sum_{i\in\mathcal{I}}\sum_{\tau=t}^{t+d_{\max}}\mathbb{E}[A_{ijk}\mathbf{1}(D_{ijk}\geq\tau-t+1)]\alpha_{js}^{(2)*}\\
&=\beta_{js}^{(2)*}+\sum_{i\in\mathcal{I}}\sum_{\tau=1}^{d_{\max}}\mathbb{E}[A_{ijk}\mathbf{1}(D_{ijk}\geq\tau)]\alpha_{js}^{(2)*}\\
&=\beta_{js}^{(2)*}+\sum_{i\in\mathcal{I}}\mathbb{E}[A_{ijk}\sum_{\tau=1}^{\infty}\tau\mathbf{1}(D_{ijk}=\tau)]\alpha_{js}^{(2)*}\\
&=\beta_{js}^{(2)*}+\sum_{i\in\mathcal{I}}\mathbb{E}[A_{ijk}D_{ijk}]\alpha_{js}^{(2)*}\\
&=\beta_{js}^{(2)*}+\sum_{i\in\mathcal{I}}v_{ijk}\alpha_{js}^{(2)*}\\
&\geq r_{jk}\quad\forall j\in\mathcal{J},\ k\in\mathcal{K}.
\end{aligned}
$$

The constraints for any $j\in\mathcal{J},k\in\mathcal{K}$ and $t\in[mW-d_{\max}+1,mW]$ clearly hold since they reduce to $r_{\max}+\sum_{i\in\mathcal{I}}\sum_{\tau=t}^{\min\{t+d_{\max},mW\}}\mathbb{E}[A_{ijk}\mathbf{1}(D_{ijk}\geq\tau-t+1)]\alpha_{js}^{(2)*}\geq r_{jk}$, and $\alpha_{js}^{(2)*}\geq 0$. The feasibility establishes that

$$
W\lambda_*^{EW(m)}-W\lambda^{EWS(ms)}\leq r_{\max}d_{\max}\quad\forall m\in[\frac{T}{W}],\ s\in[(m-1)W+1:mW],
$$

since $W\lambda^{EWS(m,s)}+r_{\max}d_{\max}$ is the objective value of the constructed solution $(\bar{\alpha}^{(1)},\bar{\beta}^{(1)})$ for (LP-EW-D)$^{(m)}$. Furthermore, we have the claimed inequality $(\dagger)$,

$$
W\lambda_*^{EW(m)}-\sum_{s=(m-1)W+1}^{MW}\lambda^{EWS(ms)}\leq r_{\max}d_{\max}\quad\forall m\in\left[\frac{T}{W}\right].
$$

We next prove the inequality $(\ddagger)$ as follows:

$$
\begin{aligned}
\sum_{s=(m-1)W+1}^{MW}(\lambda^{EWS(ms)}-\lambda_*^{S(s)})=&\frac{1}{W}\left[\sum_{s=(m-1)W+1}^{MW}\sum_{t=(m-1)W+1}^{MW}(\sum_{j\in\mathcal{J}}p_{jt}\beta_{js}^{(2)*}+\sum_{i\in\mathcal{I}}c_i\alpha_{is}^{(2)*})\right.\\
&\left.-\sum_{s=(m-1)W+1}^{MW}(\sum_{j\in\mathcal{J}}p_{js}\beta_{js}^{(2)*}+\sum_{i\in\mathcal{I}}c_i\alpha_{is}^{(2)*})\right]\\
=&\frac{1}{W}\sum_{s=(m-1)W+1}^{MW}\sum_{t=(m-1)W+1}^{MW}(\sum_{j\in\mathcal{J}}p_{jt}\beta_{js}^{(2)*}-\sum_{j\in\mathcal{J}}p_{js}\beta_{js}^{(2)*})\\
=&\frac{1}{W}\sum_{s=(m-1)W+1}^{MW}\sum_{t=(m-1)W+1}^{MW}\sum_{j\in\mathcal{J}}\beta_{js}^{(2)*}(p_{jt}-p_{js})\\
\leq&\frac{r_{\max}}{W}\sum_{s=(m-1)W+1}^{MW}\sum_{t=(m-1)W+1}^{MW}||\boldsymbol{p}_t-\boldsymbol{p}_s||_1
\end{aligned}
$$

$$\begin{aligned}
&\leq \frac{r_{\max}}{W} \sum_{s=(m-1)W+1}^{MW} \sum_{t=(m-1)W+1}^{MW} \sum_{q=\min\{s,t\}}^{\max\{s,t\}-1} \left\|\boldsymbol{p}_{q+1}-\boldsymbol{p}_q\right\|_1 \\
&\leq \frac{r_{\max}}{W} \sum_{s=(m-1)W+1}^{MW} \sum_{t=(m-1)W+1}^{MW} \sum_{q=(m-1)W+1}^{MW} ||\boldsymbol{p}_{q+1}-\boldsymbol{p}_q||_1 \\
&= r_{\max} W \sum_{q=(m-1)W+1}^{MW} \left\|\boldsymbol{p}_{q+1}-\boldsymbol{p}_q\right\|_1 .
\end{aligned}$$

The first 3 equations holds by the definition of $\lambda^{EWS(ms)}$ and $\lambda_*^{S(s)}$. The first inequality stands since $\beta_{js}^{(2)*} \leq r_{\max}$ for any j, s and $\sum_{j\in\mathcal{J}} p_{jt} - p_{js} \leq \sum_{j\in\mathcal{J}} |p_{jt} - p_{js}| = ||\boldsymbol{p}_t - \boldsymbol{p}_s||_1$. The second inequality follows from a telescoping sum.

Step 5: To derive the final result, we sum up the previous results over the $\frac{T}{W}$ episodes.

$$\begin{aligned}
T\lambda_*^E &\leq \sum_{m=1}^{T/W} W\lambda_*^{EW(m)} \\
&\leq \sum_{m=1}^{T/W} \sum_{s=(m-1)W+1}^{MW} (\lambda^{EWS(ms)} + r_{\max} d_{\max}) \\
&\leq \sum_{m=1}^{T/W} \left\{ \sum_{s=(m-1)W+1}^{MW} \lambda_*^{S(s)} + r_{\max} W \sum_{t=(m-1)W+1}^{MW} ||\boldsymbol{p}_{t+1}-\boldsymbol{p}_t||_1) \right\} + r_{\max} d_{\max} \frac{T}{W} \\
&= \left[\sum_{t=1}^{T} \lambda_*^{S(t)} \right] + r_{\max} W V + r_{\max} d_{\max} \frac{T}{W}.
\end{aligned}$$

The first inequality follows from step 1. The second and third inequalities stands by step 3. The last inequality holds by the definition of $V = \sum_{t=1}^{T-1} ||\boldsymbol{p}_{t+1} - \boldsymbol{p}_t||_1$. Since $V = o(T)$, by letting $W = \sqrt{d_{\max} T/V}$, we have

$$T\text{opt(IP-C)} \leq \sum_{t=1}^{T} \lambda_*^{S(t)} + r_{\max} \sqrt{d_{\max} VT} \leq \sum_{t=1}^{T} \lambda_*^{S(t)} + O(V^{\frac{1}{3}} T^{\frac{2}{3}}),$$

where $O(\cdot)$ hides a multiplicative factor in terms of $r_{\max}, d_{\max}$. □