Supplementary Material

EC.1. Proof of Lemma 1

Proof of Lemma 1. Recall that $V^{\mathrm{on}}(\mathcal{I}) = \sum_{t=1}^{T} r_{s_t^{\mathrm{on}}}$. We decompose $\mathrm{REG}_0 = \mathbb{E}_\Theta\left[V^{\mathrm{off}}(\mathcal{I},\Theta,B)\right] - V^{\mathrm{on}}(\mathcal{I})$ in (3) into the regret from time period T to period $K+2$ and the regret during the last $K+1$ time periods:

$$
\begin{aligned}
&\mathbb{E}_\Theta\left[V^{\mathrm{off}}(\mathcal{I},\Theta,B_T) - V^{\mathrm{off}}(\mathcal{I},(\Theta_{K+1},\ldots,\Theta_1),B_{K+1}) + V^{\mathrm{off}}(\mathcal{I},(\Theta_{K+1},\ldots,\Theta_1),B_{K+1}) - \sum_{t=1}^{T} r_{s_t^{\mathrm{on}}}\right]\\
&= \sum_{t=K+2}^{T} \mathbb{E}_{(\Theta_t,\ldots,\Theta_1),B_t}\left[V^{\mathrm{off}}(\mathcal{I},(\Theta_t,\ldots,\Theta_1),B_t) - V^{\mathrm{off}}(\mathcal{I},(\Theta_{t-1},\ldots,\Theta_1),B_{t-1}) - r_{s_t^{\mathrm{on}}}\right]\\
&\quad + \mathbb{E}_{(\Theta_{K+1},\ldots,\Theta_1),B_{K+1}}\left[V^{\mathrm{off}}(\mathcal{I},(\Theta_{K+1},\ldots,\Theta_1),B_{K+1}) - \sum_{t=1}^{K+1} r_{s_t^{\mathrm{on}}}\right]\\
&\le a_{\max}\cdot M\cdot r_{\max}\cdot \sum_{t=K+2}^{T} \mathbb{P}\left(V^{\mathrm{off}}(\mathcal{I},(\Theta_t,\ldots,\Theta_1),B_t) - V^{\mathrm{off}}(\mathcal{I},(\Theta_{t-1},\ldots,\Theta_1),B_{t-1}) > r_{s_t^{\mathrm{on}}}\right) + \sum_{s\in S} r_s.
\end{aligned}
$$

In particular, the first part of the last inequality comes from the explanation right below the statement of Lemma 1, and the second part follows from the fact that

$$V^{\mathrm{off}}(\mathcal{I},(\Theta_{K+1},\ldots,\Theta_1),B_{K+1}) - \sum_{t=1}^{K+1} r_{s_t^{\mathrm{on}}} \le \sum_{s\in S} r_s. \tag{EC.1}$$

To see (EC.1), note that with a K-period delay, the online policy has full knowledge of $(\Theta_{K+1},\ldots,\Theta_1)$ so that the gap between $V^{\mathrm{off}}(\mathcal{I},(\Theta_{K+1},\ldots,\Theta_1),B_{K+1})$ and $\sum_{t=1}^{K+1} r_{s_t^{\mathrm{on}}}$ is due to integrality. This gap has an upper bound $\sum_{s\in S} r_s$ by rounding each variable of the optimal solution to the LP, which is independent of T and K. The proof of Lemma 1 is completed. □

EC.2. Proofs of Statements in the Multisecretary Problem

Derivation of (8). We first characterize the optimal solutions to the offline and online LP Problems (6) and (7). Because both B_t and $D_j[t,1]$ are integers, Problem (6) has an optimal integer solution $\left\{X^*_{t,j},\ X^*_{t,\emptyset j}\right\}_{j\in[N]}$ for each t:

$$X^*_{t,j} = \min\left\{\left(B_t - \sum_{\ell=j+1}^{N} D_\ell[t,1]\right)^+,\ D_j[t,1]\right\} \quad\text{and}\quad X^*_{t,\emptyset j} = D_j[t,1] - X^*_{t,j}. \tag{EC.2}$$

Similarly, Problem (7) has an optimal solution $\{X_{t,j},\ X_{t,\emptyset j}\}_{j\in[N]}$:

$$X_{t,j} = \min\left\{\left(B_t - \sum_{\ell=j+1}^{N}(D_\ell[t,t-K]+\lambda_\ell[t-K-1,1])\right)^+,\ D_j[t,t-K]+\lambda_j[t-K-1,1]\right\}, \tag{EC.3}$$

$$X_{t,\emptyset j} = D_j[t,t-K]+\lambda_j[t-K-1,1]-X_{t,j}.$$

Recall that for each $t=T,T-1,\dots,K+2$, the algorithm accepts the candidate with type Θ_t at time t iff $X_{t,\Theta_t}\ge X_{t,\emptyset\Theta_t}$, which is equivalent to that at least half of type Θ_t candidates are accepted in the LP solution:

$$X_{t,\Theta_t} \ge \frac{1}{2}\left(D_{\Theta_t}[t,t-K]+\lambda_{\Theta_t}[t-K-1,1]\right). \tag{EC.4}$$

Combining with (EC.3), (EC.4) is equivalent to

$$\begin{aligned}B_t &\ge \sum_{\ell=\Theta_t+1}^{N}(D_\ell[t,t-K]+\lambda_\ell[t-K-1,1]) + \frac{1}{2}\left(D_{\Theta_t}[t,t-K]+\lambda_{\Theta_t}[t-K-1,1]\right)\\ &= \left(1-F_{\Theta_t}+\frac{\lambda_{\Theta_t}}{2}\right)(t-K-1)+\left(\sum_{\ell=\Theta_t+1}^{N}D_\ell[t,t-K]+\frac{1}{2}\cdot D_{\Theta_t}[t,t-K]\right).\end{aligned}$$

Similarly, the algorithm rejects the candidate with type Θ_t at time t iff

$$X_{t,\Theta_t} < \frac{1}{2}\left(D_{\Theta_t}[t,t-K]+\lambda_{\Theta_t}[t-K-1,1]\right). \tag{EC.5}$$

Combining with (EC.3), (EC.5) is equivalent to

$$B_t < \left(1-F_{\Theta_t}+\frac{\lambda_{\Theta_t}}{2}\right)(t-K-1)+\left(\sum_{\ell=\Theta_t+1}^{N}D_\ell[t,t-K]+\frac{1}{2}\cdot D_{\Theta_t}[t,t-K]\right).$$

The derivation of (8) is completed. ☐

Proof of Theorem 1. Combining (5) with the fact that $a_{\max}=1$, $M=1$, $r_{\max}=r_N$, and when the decision for Θ_t incurs regret is exactly when incorrect acceptance or rejection occurs, the regret REG of Algorithm 1 can be upper bounded by

$$r_N\cdot\sum_{t=K+2}^{T}\mathbb{P}\left(V^{\text{off}}(\mathcal{I},(\Theta_t,\dots,\Theta_1),B_t)-V^{\text{off}}(\mathcal{I},(\Theta_{t-1},\dots,\Theta_1),B_{t-1})>r_{s_t^{\text{on}}}\right). \tag{EC.6}$$

Note that the probability in (EC.6) is w.r.t. $(\Theta_{t-1},\ldots,\Theta_1)$ and B_t. It is straightforward to verify that incorrect acceptance for Θ_t occurs iff $X^*_{t,\Theta_t}=0$ and (8) holds, and incorrect rejection for Θ_t occurs iff $X^*_{t,\emptyset\Theta_t}=0$ and the reverse of (8) holds. From (EC.2),

$$\begin{aligned} X^*_{t,\Theta_t}=0 &\iff \sum_{\ell=\Theta_t+1}^{N} D_\ell[t,1]\ge B_t,\\ X^*_{t,\emptyset\Theta_t}=0 &\iff \sum_{\ell=\Theta_t}^{N} D_\ell[t,1]\le B_t. \end{aligned} \tag{EC.7}$$

Combining (EC.7) and (8), it follows that incorrect acceptance implies

$$\sum_{\ell=\Theta_t+1}^{N} D_\ell[t-K-1,1]-(1-F_{\Theta_t})(t-K-1)\ge \frac{\lambda_{\Theta_t}}{2}\cdot(t-K-1)+\frac{1}{2}\cdot D_{\Theta_t}[t,t-K],$$

and incorrect rejection implies

$$\left(1-F_{\Theta_t}+\lambda_{\Theta_t}\right)(t-K-1)-\sum_{\ell=\Theta_t}^{N} D_\ell[t-K-1,1]\ge \frac{\lambda_{\Theta_t}}{2}\cdot(t-K-1)+\frac{1}{2}\cdot D_{\Theta_t}[t,t-K].$$

Hence, conditional on $(\Theta_t,\ldots,\Theta_{t-K})$, the probability of incorrect acceptance for Θ_t is at most

$$\mathbb{P}\left(\sum_{\ell=\Theta_t+1}^{N} D_\ell[t-K-1,1]-(1-F_{\Theta_t})(t-K-1)\ge \frac{\lambda_{\Theta_t}}{2}\cdot(t-K-1)+\frac{1}{2}\cdot D_{\Theta_t}[t,t-K]\right), \tag{EC.8}$$

and the probability of incorrect rejection for Θ_t is at most

$$\mathbb{P}\left(\left(1-F_{\Theta_t}+\lambda_{\Theta_t}\right)(t-K-1)-\sum_{\ell=\Theta_t}^{N} D_\ell[t-K-1,1]\ge \frac{\lambda_{\Theta_t}}{2}\cdot(t-K-1)+\frac{1}{2}\cdot D_{\Theta_t}[t,t-K]\right). \tag{EC.9}$$

Furthermore, from Hoeffding's inequality, both (EC.8) and (EC.9) can be upper bounded by

$$\exp\left(-\frac{t-K-1}{2}\cdot\left(\lambda_{\Theta_t}+\frac{D_{\Theta_t}[t,t-K]}{t-K-1}\right)^2\right). \tag{EC.10}$$

Note that B_t is the only state variable in the multisecretary problem and the upper bound (EC.10) holds for any value of B_t. Thus, conditional on Θ_t, it follows that

$$\begin{aligned} &\mathbb{P}\left(V^{\text{off}}(\mathcal{I},(\Theta_t,\ldots,\Theta_1),B_t)-V^{\text{off}}(\mathcal{I},(\Theta_{t-1},\ldots,\Theta_1),B_{t-1})>r_{s_t^{\text{on}}}\right)\\ &\le \mathbb{E}\left[\exp\left(-\frac{t-K-1}{2}\cdot\left(\lambda_{\Theta_t}+\frac{D_{\Theta_t}[t,t-K]}{t-K-1}\right)^2\right)\middle|\Theta_t\right], \end{aligned} \tag{EC.11}$$

where the probability and expectation above are w.r.t. $(\Theta_{t-1},\ldots,\Theta_{t-K})$. Furthermore,

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(-\frac{t-K-1}{2}\cdot\left(\lambda_{\Theta_t}+\frac{D_{\Theta_t}[t,t-K]}{t-K-1}\right)^2\right)\middle|\Theta_t\right]\\
=\;&\exp\left(-\frac{\lambda_{\Theta_t}^2}{2}\cdot(t-K-1)\right)\cdot\mathbb{E}\left[\exp\left(-\lambda_{\Theta_t}\cdot D_{\Theta_t}[t,t-K]\right)\cdot\exp\left(-\frac{\left(D_{\Theta_t}[t,t-K]\right)^2}{2\left(t-K-1\right)}\right)\middle|\Theta_t\right]\\
\le\;&\exp\left(-\frac{\lambda_{\Theta_t}^2}{2}\cdot(t-K-1)\right)\cdot\mathbb{E}\left[\exp\left(-\lambda_{\Theta_t}\cdot D_{\Theta_t}[t,t-K]\right)\middle|\Theta_t\right]\\
=\;&\exp\left(-\frac{\lambda_{\Theta_t}^2}{2}\cdot(t-K-1)\right)\cdot e^{-\lambda_{\Theta_t}}\cdot\mathbb{E}\left[\exp\left(-\lambda_{\Theta_t}\cdot D_{\Theta_t}[t-1,t-K]\right)\middle|\Theta_t\right] &\text{(EC.12)}\\
=\;&\exp\left(-\frac{\lambda_{\Theta_t}^2}{2}\cdot(t-K-1)\right)\cdot e^{-\lambda_{\Theta_t}}\cdot\left(\lambda_{\Theta_t}e^{-\lambda_{\Theta_t}}+(1-\lambda_{\Theta_t})\right)^K, &\text{(EC.13)}
\end{aligned}$$

where the inequality comes from the fact that the term $\exp\left(-\frac{\left(D_{\Theta_t}[t,t-K]\right)^2}{2(t-K-1)}\right)$ is no greater than one and (EC.12) is from the fact that $D_{\Theta_t}[t,t]=1$ conditional on Θ_t.

Recall that $\lambda_{\min}=\min_{j\in[N]}\lambda_j$ and $\rho_0=\lambda_{\min}e^{-\lambda_{\min}}+(1-\lambda_{\min})<1$. Note that the function $g(x)\triangleq xe^{-x}+(1-x)$ is non-increasing in x on $[0,1]$. Hence, combining (EC.6), (EC.11), and (EC.13), the expected regret is no greater than

$$\begin{aligned}
&r_N\cdot\sum_{t=K+2}^{T}\sum_{j\in[N]}\lambda_j\left(e^{-\lambda_j}\cdot\exp\left(-\frac{\lambda_j^2}{2}\cdot(t-K-1)\right)\cdot\left(\lambda_je^{-\lambda_j}+(1-\lambda_j)\right)^K\right)\\
\le\;& r_N\cdot\sum_{t=K+2}^{T}\sum_{j\in[N]}\lambda_j\left(e^{-\lambda_{\min}}\cdot\exp\left(-\frac{\lambda_{\min}^2}{2}\cdot(t-K-1)\right)\cdot\left(\lambda_{\min}e^{-\lambda_{\min}}+(1-\lambda_{\min})\right)^K\right)\\
=\;& r_N\cdot e^{-\lambda_{\min}}\cdot\rho_0^K\cdot\sum_{t=K+2}^{T}\exp\left(-\frac{\lambda_{\min}^2}{2}\cdot(t-K-1)\right)\\
\le\;& r_N\cdot e^{-\lambda_{\min}}\cdot\rho_0^K\cdot\sum_{t=1}^{\infty}\exp\left(-\frac{\lambda_{\min}^2t}{2}\right)\\
\le\;& r_N\cdot e^{-\lambda_{\min}}\cdot\rho_0^K\cdot\frac{\exp(-\lambda_{\min}^2/2)}{1-\exp(-\lambda_{\min}^2/2)}\\
=\;& r_N\cdot e^{-\lambda_{\min}}\cdot\rho_0^K\cdot\frac{1}{\exp(\lambda_{\min}^2/2)-1}\\
\le\;& \rho_0^K\cdot\frac{2\cdot r_N\cdot e^{-\lambda_{\min}}}{\lambda_{\min}^2},
\end{aligned}$$

where the last inequality comes from the fact that $e^x\ge 1+x$ for $x\ge 0$. The proof of Theorem 1 is completed. □

EC.3. Proofs of Statements in the General Online Resource Allocation Problem

Proof of Theorem 2. We first prove that $X^*_{t,s_t^{\text{on}}} < a_{\max} + 1$ is a necessary condition for

$$V^{\text{off}}(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t) - V^{\text{off}}(\mathcal{I}, (\Theta_{t-1}, \ldots, \Theta_1), B_{t-1}) > r_{s_t^{\text{on}}}.$$

Indeed, if $X^*_{t,s_t^{\text{on}}} \geq a_{\max} + 1$, then $B_{t,i}$ for each resource $i \in [M]$ satisfying $a_{is_t^{\text{on}}} > 0$ is at least $(a_{\max} + 1)a_{is_t^{\text{on}}}$. This implies that the remaining resource after action s_t^{on} is at least $a_{\max} \cdot a_{is_t^{\text{on}}} \geq a_{\max}$, because $a_{\min}$ is normalized to be 1. Thus, action s_t^{on} does not result in an update of feasible action sets. That is, $S^t = S^{t-1}$, and $V^{\text{off}}(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t)$ and $V^{\text{off}}(\mathcal{I}, (\Theta_{t-1}, \ldots, \Theta_1), B_{t-1})$ have the same coefficient matrix. From this, we prove that $V^{\text{off}}(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t) = V^{\text{off}}(\mathcal{I}, (\Theta_{t-1}, \ldots, \Theta_1), B_{t-1}) + r_{s_t^{\text{on}}}$. More specifically, for any given sample path Θ, we define $\omega^{\text{on}}_{t,s}$ as the indicator of whether action $s \in S_j$ is chosen for arrival Θ_t with customer type j, namely, $\omega^{on}_{t,s} = \mathbb{I}\{\Theta_t = j, s_t^{\text{on}} = s\}$. Let us consider the following variant of the offline Problem (19):

$$\begin{aligned} \max \quad & \sum_{s \in S_t} r_s x_s && && \text{(EC.14)} \\ \text{s.t.} \quad & \sum_{s \in S_t} a_{is} x_s \leq B_{t,i}, && \forall i \in [M], \\ & \sum_{s \in S_{t,j}} x_s = D_j[t,1], && \forall j \in [N], \\ & x_s \geq \omega_s, && \forall s \in S_t. \end{aligned}$$

Let $\hat{V}^{\text{off}}(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t, \{\omega_s\})$ denote the optimal value of Problem (EC.14) with right-hand side vector B_t, $\{D_j[t,1]\}$, and $\{\omega_s\}$. It is easy to see that the optimal value of the offline LP Problem (19) is $\hat{V}^{\text{off}}(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t, 0) = V^{\text{off}}(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t)$. After action s_t^{on} is chosen for Θ_t, the optimal value that can be obtained afterwards reduces to $\hat{V}^{\text{off}}\left(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t, \{\omega^{\text{on}}_{t,s}\}\right)$, Note that $\hat{V}^{\text{off}}\left(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t, \{\omega^{\text{on}}_{t,s}\}\right) = V^{\text{off}}(\mathcal{I}, (\Theta_{t-1}, \ldots, \Theta_1), B_{t-1}) + r_{s_t^{\text{on}}}$, because $S^t = S^{t-1}$. Thus, for any given t, the regret of the decision s_t^{on} for Θ_t is

$$\begin{aligned} \text{Diff} \triangleq\ & \hat{V}^{\text{off}}(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t, 0) - \hat{V}^{\text{off}}\left(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t, \{\omega^{\text{on}}_{t,s}\}\right) \\ =\ & V^{\text{off}}(\mathcal{I}, (\Theta_t, \ldots, \Theta_1), B_t) - \left(V^{\text{off}}(\mathcal{I}, (\Theta_{t-1}, \ldots, \Theta_1), B_{t-1}) + r_{s_t^{\text{on}}}\right). \end{aligned}$$

Furthermore, for an optimal solution $\{X^*_{t,s}\}_{s\in S_t}$ of Problem (19), because $X^*_{t,s^{\text{on}}_t} \geq a_{\max}+1 \geq 1$, we have that $\{X^*_{t,s}\}_{s\in S_t}$ is a feasible solution of Problem (EC.14) with $\omega_s = \omega^{\text{on}}_{t,s}$, implying $\text{Diff} = 0$. Therefore, if $\text{Diff} > 0$, then we must have $X^*_{t,s^{\text{on}}_t} < a_{\max}+1$.

Plugging this fact into (5), we have

$$\begin{aligned}\text{REG} \leq\ & a_{\max}\cdot M\cdot r_{\max}\cdot \sum_{t=K+2}^{T} \mathbb{P}\left(X^*_{t,s^{\text{on}}_t} < a_{\max}+1\right)\\ =\ & a_{\max}\cdot M\cdot r_{\max}\cdot \sum_{t=K+2}^{T}\sum_{j\in[N]} \lambda_j\cdot \mathbb{P}\left(X^*_{t,s^{\text{on}}_t} < a_{\max}+1|\Theta_t=j\right).\end{aligned} \tag{EC.15}$$

Recall that $C_1 = (a_{\max}+1)C_{\max}+a_{\max}$. We next bound $\mathbb{P}\left(X^*_{t,s^{\text{on}}_t} < a_{\max}+1|\Theta_t=j\right)$:

$$\begin{aligned}&\mathbb{P}\left(X^*_{t,s^{\text{on}}_t} < a_{\max}+1|\Theta_t=j\right)\\ \leq\ & \mathbb{P}\left(X^*_{t,s^{\text{on}}_t} < a_{\max}+1+X_{t,s^{\text{on}}_t}-\frac{D_j[t,t-K]+\lambda_j(t-K-1)}{C_{\max}+1}\middle|\Theta_t=j\right)\\ \leq\ & \mathbb{P}\left(\max_{s\in S_t}\left|X^*_{t,s}-X_{t,s}\right| > \frac{D_j[t,t-K]+\lambda_j(t-K-1)}{C_{\max}+1}-(a_{\max}+1)\middle|\Theta_t=j\right)\\ \leq\ & \mathbb{P}\left(\max_{j\in[N]}\left|D_j[t-K-1,1]-\lambda_j(t-K-1)\right| > \frac{D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1}{\kappa_t(C_{\max}+1)}\middle|\Theta_t=j\right)\\ \leq\ & \mathbb{P}\left(\left|D_{j^*_t}[t-K-1,1]-\lambda_{j^*_t}(t-K-1)\right| > \frac{D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1}{\kappa_{\max}(C_{\max}+1)}\middle|\Theta_t=j\right)\\ \triangleq\ & \mathbb{P}\left(E^j_t\right).\end{aligned} \tag{EC.16}$$

Here, $\kappa_t = \kappa(J_t)$, $j^*_t \in \arg\max_{j\in[N]}\{|D_j[t-K-1,1]-\lambda_j(t-K-1)|\}$, and E^j_t represents the following event conditional on $\Theta_t = j$:

$$\left|D_{j^*_t}[t-K-1,1]-\lambda_{j^*_t}(t-K-1)\right| > \frac{D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1}{\kappa_{\max}(C_{\max}+1)}.$$

In addition, the first inequality comes from (21), the third inequality comes from (22), and the fourth inequality is due to the monotonicity of κ_t in t.

Note that (EC.16) is independent of B_t. Conditional on $(\Theta_{t-1},\ldots,\Theta_{t-K})$, we further decompose $\mathbb{P}\left(E^j_t|\Theta_{t-1},\ldots,\Theta_{t-K}\right)$ into two parts:

$$\mathbb{P}\left(E^j_t|\Theta_{t-1},\ldots,\Theta_{t-K}\right) = A^{j,1}_t + A^{j,2}_t, \tag{EC.17}$$

where

$$\begin{aligned} A_t^{j,1} \triangleq\ & \mathbb{I}\left(D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1<0\right) \\ & \cdot \mathbb{P}\left(E_t^j \middle| D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1<0\right), \\ A_t^{j,2} \triangleq\ & \mathbb{I}\left(D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1\geq 0\right) \\ & \cdot \mathbb{P}\left(E_t^j \middle| D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1\geq 0\right). \end{aligned}$$

It is easy to see that $\mathbb{P}\left(E_t^j \middle| D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1<0\right)=1$. It implies that

$$\begin{aligned} & \sum_{t=K+2}^{T} \mathbb{E}\left[A_t^{j,1}\right] \\ = & \sum_{t=K+2}^{T} \mathbb{E}\Bigg[\mathbb{I}\left(D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1<0\right) \\ & \qquad \cdot \mathbb{P}\left(E_t^j \middle| D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1<0\right)\Bigg] \\ = & \sum_{t=K+2}^{T} \mathbb{P}\left(D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1<0\right)\cdot 1, \\ = & \sum_{m=1}^{T-K-1} \mathbb{P}\left(\mathrm{Bin}(K,\lambda_j)<C_1-\lambda_j m\right), \end{aligned}$$

where the expectation above is w.r.t. $(\Theta_{t-1},\ldots,\Theta_{t-K})$, and the last equality holds because $D_j[t-1,t-K]$ follows a binomial distribution with parameters K and λ_j (denoted by $\mathrm{Bin}(K,\lambda_j)$). It is straightforward to see that

$$\begin{aligned} \sum_{m=1}^{T-K-1} \mathbb{P}\left(\mathrm{Bin}(K,\lambda_j)<C_1-\lambda_j m\right) \leq\ & \left\lceil \frac{C_1}{\lambda_j} \right\rceil \cdot \mathbb{P}\left(\mathrm{Bin}(K,\lambda_j)\leq C_1\right) \\ \leq\ & \left\lceil \frac{C_1}{\lambda_{\min}} \right\rceil \cdot \mathbb{P}\left(\mathrm{Bin}(K,\lambda_j)\leq C_1\right). \end{aligned} \tag{EC.18}$$

For $K > \frac{C_1}{\lambda_{\min}}$, Hoeffding's inequality can yield the following upper bound on $\mathbb{P}\left(\mathrm{Bin}(K,\lambda_j)\leq C_1\right)$:

$$\begin{aligned} \mathbb{P}\left(\mathrm{Bin}(K,\lambda_j)\leq C_1\right) \leq\ & \exp\left(-2K\left(\lambda_j-\frac{C_1}{K}\right)^2\right) \\ \leq\ & \exp\left(-2K\left(\lambda_{\min}-\frac{C_1}{K}\right)^2\right) \\ =\ & \exp\left(-2K\lambda_{\min}^2+4\lambda_{\min}C_1-2\frac{C_1^2}{K}\right) \end{aligned}$$

$$\leq \exp\left(-2\lambda_{\min}^2\left(K-2\frac{C_1}{\lambda_{\min}}\right)\right).$$

In addition, $\mathbb{P}\left(\mathrm{Bin}(K,\lambda_j)\leq C_1\right)$ has a trivial upper bound 1 for $0\leq K\leq \frac{C_1}{\lambda_{\min}}$. Thus, for all $K\geq 0$, we have

$$\mathbb{P}\left(\mathrm{Bin}(K,\lambda_j)\leq C_1\right)\leq \left(\exp\left(-2\lambda_{\min}^2\right)\right)^{K-2\frac{C_1}{\lambda_{\min}}}=\rho_1^{K-2\frac{C_1}{\lambda_{\min}}}, \tag{EC.19}$$

where $\rho_1=\exp(-2\lambda_{\min}^2)$. It implies that

$$\sum_{t=K+2}^{T}\mathbb{E}\left[A_t^{j,1}\right]\leq \left\lceil\frac{C_1}{\lambda_{\min}}\right\rceil\cdot\rho_1^{K-2\frac{C_1}{\lambda_{\min}}}. \tag{EC.20}$$

Next, we bound $\sum_{t=K+2}^{T}\mathbb{E}\left[A_t^{j,2}\right]$ from above. From Hoeffding's inequality, for all $t\geq K+2$, we have

$$\begin{aligned}&\mathbb{P}\left(E_t^j\middle| D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1\geq 0\right)\\ &\leq 2\exp\left(-2\frac{\left(D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1\right)^2}{(t-K-1)\cdot\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right),\end{aligned}$$

which follows that

$$\begin{aligned}&\sum_{t=K+2}^{T}\mathbb{E}\left[A_t^{j,2}\right]\\ &\leq \sum_{t=K+2}^{T}\mathbb{E}\Bigg[\mathbb{I}\left(D_j[t-1,t-K]+\lambda_j(t-K-1)\geq C_1\right)\\ &\qquad\cdot 2\exp\left(-2\frac{\left(D_j[t-1,t-K]+\lambda_j(t-K-1)-C_1\right)^2}{(t-K-1)\cdot\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\Bigg]\\ &\leq 2\sum_{t=K+2}^{T}\mathbb{E}\left[\exp\left(-2\frac{\left(\lambda_j(t-K-1)+(D_j[t-1,t-K]-C_1)\right)^2}{(t-K-1)\cdot\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\right].\end{aligned}$$

In addition,

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-2\frac{\left(\lambda_j(t-K-1)+(D_j[t-1,t-K]-C_1)\right)^2}{(t-K-1)\cdot\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\right]\\ &=\exp\left(-\frac{2\lambda_j^2(t-K-1)}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\mathbb{E}\Bigg[\exp\left(-\frac{4\lambda_j\left(D_j[t-1,t-K]-C_1\right)}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\\ &\qquad\cdot\exp\left(-\frac{2\left(D_j[t-1,t-K]-C_1\right)^2}{(t-K-1)\cdot\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\Bigg]\end{aligned}$$

$$
\begin{aligned}
&\le \exp\left(-\frac{2\lambda_j^2(t-K-1)}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\mathbb{E}\left[\exp\left(-\frac{4\lambda_j\left(D_j[t-1,t-K]-C_1\right)}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\right]\\
&= \exp\left(-\frac{2\lambda_j^2(t-K-1)}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\exp\left(\frac{4\lambda_j\cdot C_1}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\left(\lambda_j\exp\left(-\frac{4\lambda_j}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)+(1-\lambda_j)\right)^K\\
&\le \exp\left(-\frac{2\lambda_{\min}^2(t-K-1)}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\exp\left(\frac{4\lambda_{\max}\cdot C_1}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\rho_2^K, \qquad \text{(EC.21)}
\end{aligned}
$$

where $\rho_2=\lambda_{\min}\exp\left(-\frac{4\lambda_{\min}}{\kappa_{\max}^2(C_{\max}+1)^2}\right)+(1-\lambda_{\min})$ and the last inequality comes from the fact that the function $f(x)\triangleq x\exp\left(-\frac{4x}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)+(1-x)$ is non-increasing in $x\in[0,1]$. Thus,

$$
\begin{aligned}
\sum_{t=K+2}^{T}\mathbb{E}\left[A_t^{j,2}\right] &\le 2\sum_{t=K+2}^{T}\left(\exp\left(-\frac{2\lambda_{\min}^2(t-K-1)}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\exp\left(\frac{4\lambda_{\max}\cdot C_1}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\rho_2^K\right)\\
&\le 2\cdot\frac{\exp\left(\frac{-2\lambda_{\min}^2}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)}{1-\exp\left(\frac{-2\lambda_{\min}^2}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)}\cdot\exp\left(\frac{4\lambda_{\max}\cdot C_1}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\cdot\rho_2^K\\
&= \frac{2}{\exp\left(\frac{2\lambda_{\min}^2}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)-1}\cdot\exp\left(\frac{4\lambda_{\max}\cdot C_1}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\cdot\rho_2^K\\
&\le \frac{2}{1+\frac{2\lambda_{\min}^2}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}-1}\cdot\exp\left(\frac{4\lambda_{\max}\cdot C_1}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\cdot\rho_2^K\\
&= \frac{\kappa_{\max}^2\cdot(C_{\max}+1)^2}{\lambda_{\min}^2}\cdot\exp\left(\frac{4\lambda_{\max}\cdot C_1}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\cdot\rho_2^K, \qquad \text{(EC.22)}
\end{aligned}
$$

where the last inequality comes from the fact that $e^x\ge 1+x$ for $x\ge 0$.

Therefore, by (EC.20) and (EC.22), (EC.15) is upper bounded by

$$
\begin{aligned}
&a_{\max}\cdot M\cdot r_{\max}\cdot\sum_{t=K+2}^{T}\sum_{j\in[N]}\lambda_j\cdot\mathbb{P}\left(X^*_{t,s_t^{\text{on}}}<a_{\max}+1|\Theta_t=j\right)\\
&\le a_{\max}\cdot M\cdot r_{\max}\cdot\sum_{j\in[N]}\lambda_j\left(\left\lceil\frac{C_1}{\lambda_{\min}}\right\rceil\cdot\rho_1^{K-2\frac{C_1}{\lambda_{\min}}}+\frac{\kappa_{\max}^2\cdot(C_{\max}+1)^2}{\lambda_{\min}^2}\cdot\exp\left(\frac{4\lambda_{\max}\cdot C_1}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\cdot\rho_2^K\right)\\
&\le a_{\max}\cdot M\cdot r_{\max}\cdot\left(\left\lceil\frac{C_1}{\lambda_{\min}}\right\rceil\cdot\rho_1^{K-2\frac{C_1}{\lambda_{\min}}}+\frac{\kappa_{\max}^2\cdot(C_{\max}+1)^2}{\lambda_{\min}^2}\cdot\exp\left(\frac{4\lambda_{\max}\cdot C_1}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\cdot\rho_2^K\right)\\
&\le a_{\max}\cdot M\cdot r_{\max}\cdot\left(\left\lceil\frac{C_1}{\lambda_{\min}}\right\rceil\cdot\rho_1^{-2\frac{C_1}{\lambda_{\min}}}+\frac{\kappa_{\max}^2\cdot(C_{\max}+1)^2}{\lambda_{\min}^2}\cdot\exp\left(\frac{4\lambda_{\max}\cdot C_1}{\kappa_{\max}^2\cdot(C_{\max}+1)^2}\right)\right)\cdot\rho^K,
\end{aligned}
$$

where $\rho=\max\{\rho_1,\rho_2\}$. The proof of Theorem 2 is completed. □

EC.4. Proofs of Theorems 3 and 4 in Section 6

Proof of Theorem 3. Recall that the estimated arrival probability of type $j \in [N]$ in period t is

$$\hat{\lambda}_{t,j} = \frac{D_j[T, t-K]}{T-t+1+K}.$$

Additionally define the following event:

$$F_{t,j} \triangleq \left\{ \left| \hat{\lambda}_{t,j} - \lambda_j \right| \le \eta \right\},$$

where $\eta = \frac{\lambda_{\min}}{3} \min\left\{ \frac{1}{\kappa_{\max}(C_{\max}+1)},\ 1 \right\}$ as defined in (26). Let $F^c_{t,j}$ be the complement set of $F_{t,j}$. By Hoeffding's inequality, for all $j \in [N]$,

$$\mathbb{P}\left(F^c_{t,j}\right) = \mathbb{P}\Big(\left| D_j[T, t-K] - \lambda_j (T-t+1+K) \right| > \eta (T-t+1+K) \Big) \le 2e^{-2\eta^2 (T-t+1+K)}. \tag{EC.23}$$

Similar to (21) and (EC.15), we have

$$X_{t, s^{\text{on}}_t} \ge \frac{D_{\Theta_t}[t, t-K] + \hat{\lambda}_{t,\Theta_t}[t-K-1, 1]}{C_{\max}+1},$$

and

$$\text{REG} \le a_{\max} \cdot M \cdot r_{\max} \cdot \sum_{t=K+2}^{T} \sum_{j \in [N]} \lambda_j \cdot \mathbb{P}\left(X^*_{t, s^{\text{on}}_t} < a_{\max} + 1 \middle| \Theta_t = j \right). \tag{EC.24}$$

We next bound $\mathbb{P}\left(X^*_{t, s^{\text{on}}_t} < a_{\max} + 1 \middle| \Theta_t = j \right)$. Similar to (EC.16),

$$\begin{aligned}
&\mathbb{P}\left(X^*_{t, s^{\text{on}}_t} < a_{\max} + 1 \middle| \Theta_t = j \right) \\
&\le \mathbb{P}\left(\left| D_{j^*_t}[t-K-1, 1] - \hat{\lambda}_{t, j^*_t}(t-K-1) \right| > \frac{D_j[t-1, t-K] + \hat{\lambda}_{t,j}(t-K-1) - C_1}{\kappa_{\max}(C_{\max}+1)} \middle| \Theta_t = j \right) \\
&\triangleq \mathbb{P}(A) \\
&= \mathbb{P}\left(A,\ F^c_{t,j^*_t}\right) + \mathbb{P}\left(A,\ F_{t,j^*_t},\ F^c_{t,j}\right) + \mathbb{P}\left(A,\ F_{t,j^*_t},\ F_{t,j}\right) \\
&= \mathbb{P}\left(A,\ F^c_{t,j^*_t}\right) + \mathbb{P}\left(A,\ F_{t,j^*_t},\ F^c_{t,j}\right) + \mathbb{P}\left(F_{t,j^*_t},\ F_{t,j}\right) \cdot \mathbb{P}\left(A \middle| F_{t,j^*_t},\ F_{t,j}\right) \\
&\le \mathbb{P}\left(F^c_{t,j^*_t}\right) + \mathbb{P}\left(F^c_{t,j}\right) + \mathbb{P}\left(A \middle| F_{t,j^*_t},\ F_{t,j}\right),
\end{aligned}$$

where A represents the conditional event

$$\left\{ \left| D_{j^*_t}[t-K-1, 1] - \hat{\lambda}_{t, j^*_t}(t-K-1) \right| > \frac{D_j[t-1, t-K] + \hat{\lambda}_{t,j}(t-K-1) - C_1}{\kappa_{\max}(C_{\max}+1)} \middle| \Theta_t = j \right\}.$$

From (EC.23),

$$\mathbb{P}\left(F^c_{t,j^*_t}\right) \le\ 2e^{-2\eta^2(T-t+1+K)}, \quad \mathbb{P}\left(F^c_{t,j}\right) \le\ 2e^{-2\eta^2(T-t+1+K)}.$$

In addition, it is straightforward to verify that conditional on F_{t,j^*_t},

$$\begin{aligned}
&\left|D_{j^*_t}[t-K-1,1] - \hat{\lambda}_{t,j^*_t}(t-K-1)\right| \\
\le &\left|D_{j^*_t}[t-K-1,1] - \lambda_{j^*_t}(t-K-1)\right| + \left|\left(\hat{\lambda}_{t,j^*_t} - \lambda_{j^*_t}\right)(t-K-1)\right| \\
\le &\left|D_{j^*_t}[t-K-1,1] - \lambda_{j^*_t}(t-K-1)\right| + \eta(t-K-1),
\end{aligned} \tag{EC.25}$$

and conditional on $F_{t,j}$,

$$\hat{\lambda}_{t,j} \ge\ \lambda_j - \eta \ge\ \eta\cdot\kappa_{\max}(C_{\max}+1) + \frac{\lambda_j}{3}. \tag{EC.26}$$

(EC.25) and (EC.26) together imply that

$$\begin{aligned}
&\mathbb{P}\left(A\middle|F_{t,j^*_t},\ F_{t,j}\right) \\
&\le \mathbb{P}\left(\left|D_{j^*_t}[t-K-1,1] - \lambda_{j^*_t}(t-K-1)\right| > \frac{D_j[t-1,t-K] + \lambda_j(t-K-1)/3 - C_1}{\kappa_{\max}(C_{\max}+1)}\right).
\end{aligned}$$

Moreover, similar to (EC.20) and (EC.22), we can derive

$$\begin{aligned}
&\sum_{t=K+2}^{T} \mathbb{P}\left(\left|D_{j^*_t}[t-K-1,1] - \lambda_{j^*_t}(t-K-1)\right| > \frac{D_j[t-1,t-K] + \lambda_j(t-K-1)/3 - C_1}{\kappa_{\max}(C_{\max}+1)}\right) \\
&\le\ \left\lceil\frac{3C_1}{\lambda_{\min}}\right\rceil\cdot\hat{\rho}_1^{K-2\frac{C_1}{\lambda_{\min}/3}} + \frac{\kappa^2_{\max}\cdot(C_{\max}+1)^2}{(\lambda_{\min}/3)^2}\cdot\exp\left(\frac{4\cdot(\lambda_{\max}/3)\cdot C_1}{\kappa^2_{\max}\cdot(C_{\max}+1)^2}\right)\cdot\hat{\rho}_2^K,
\end{aligned}$$

where $\hat{\rho}_1$ and $\hat{\rho}_2$ are defined in (26). Combining all the above, (EC.24) is upper bounded by

$$\begin{aligned}
&a_{\max}\cdot M\cdot r_{\max}\cdot\sum_{t=K+2}^{T}\sum_{j\in[N]}\lambda_j\cdot\mathbb{P}\left(X^*_{t,s^{\mathrm{on}}_t} < a_{\max}+1|\Theta_t = j\right) \\
&\le a_{\max}\cdot M\cdot r_{\max}\cdot\sum_{t=K+2}^{T} 4e^{-2\eta^2(T-t+1+K)} \\
&\quad + a_{\max}\cdot M\cdot r_{\max}\cdot\left(\left\lceil\frac{3C_1}{\lambda_{\min}}\right\rceil\cdot\hat{\rho}_1^{K-2\frac{C_1}{\lambda_{\min}/3}} + \frac{\kappa^2_{\max}\cdot(C_{\max}+1)^2}{(\lambda_{\min}/3)^2}\cdot\exp\left(\frac{4\cdot(\lambda_{\max}/3)\cdot C_1}{\kappa^2_{\max}\cdot(C_{\max}+1)^2}\right)\cdot\hat{\rho}_2^K\right) \\
&\le a_{\max}\cdot M\cdot r_{\max}\cdot\left(\frac{2}{\eta^2} + \left\lceil\frac{3C_1}{\lambda_{\min}}\right\rceil\cdot\hat{\rho}_1^{-2\frac{C_1}{\lambda_{\min}/3}} + \frac{\kappa^2_{\max}\cdot(C_{\max}+1)^2}{(\lambda_{\min}/3)^2}\cdot\exp\left(\frac{4\cdot(\lambda_{\max}/3)\cdot C_1}{\kappa^2_{\max}\cdot(C_{\max}+1)^2}\right)\right)\cdot\hat{\rho}^K,
\end{aligned}$$

where the last inequality comes from the fact that

$$\sum_{t=K+2}^{T} 4e^{-2\eta^2(T-t+1+K)} \le\ 4e^{-2\eta^2 K}\frac{e^{-2\eta^2}}{1-e^{-2\eta^2}} \le\ \frac{2}{\eta^2}\cdot\hat{\rho}_3^K.$$

The proof of Theorem 3 is completed. □

Proof of Theorem 4. As discussed in Section 6.3, $T = (K+1)(\lceil T/(K+1)\rceil - 1) + R$ and there are $\lceil T/(K+1)\rceil$ batches. It suffices to consider the regret incurred in the first $\lceil T/(K+1)\rceil - 1$ batches. Particularly, the n-th batch consists of $\Theta_{T-(n-1)(K+1)}, \Theta_{T-(n-1)(K+1)-1}, \dots, \Theta_{T-n(K+1)+1}$ for $n = 1, \dots, \lceil T/(K+1)\rceil - 1$, and the decision for $\Theta_{T-(n-1)(K+1)-(K-m)}$ in the n-th batch is equivalent to m-period delay based on the expected demand $\lambda[T-n(K+1), 1]$ for $m = K, K-1, \dots, 0$. Each arrival in the first $\lceil T/(K+1)\rceil - 1$ batches can be specified by some $n \in [\lceil T/(K+1)\rceil - 1]$ and $m \in \{0\} \cup [K]$:

$$\begin{aligned} t = \ & t(n,m) \triangleq T - (n-1)(K+1) - (K-m) \\ & \quad = (K+1)\left(\lceil T/(K+1)\rceil - 1 - n\right) + (m+1) + R. \end{aligned} \tag{EC.27}$$

Namely, the arrival at time $t(n,m)$ is in the n-th batch with m-period delay.

By the characterization of (n,m) and (EC.15), we have

$$\mathrm{REG} \le \ a_{\max} \cdot M \cdot r_{\max} \cdot \sum_{n=1}^{\lceil T/(K+1)\rceil - 1} \sum_{m=0}^{K} \sum_{j\in[N]} \lambda_j \cdot \mathbb{P}\left(X^*_{t(n,m), s^{\mathrm{on}}_{t(n,m)}} < a_{\max} + 1 | \Theta_{t(n,m)} = j\right).$$

Similar to (EC.16), $\mathbb{P}\left(X^*_{t(n,m), s^{\mathrm{on}}_{t(n,m)}} < a_{\max} + 1 | \Theta_{t(n,m)} = j\right)$ can be bounded by

$$\mathbb{P}\left(X^*_{t(n,m), s^{\mathrm{on}}_{t(n,m)}} < a_{\max} + 1 | \Theta_{t(n,m)} = j\right) \le \ \mathbb{P}\left(E^j_{t(n,m)}\right),$$

where $E^j_{t(n,m)}$ represents the following event conditional on $\Theta_{t(n,m)} = j$:

$$\begin{aligned} & \left| D_{j^*_{t(n,m)}}[T-n(K+1), 1] - \lambda_{j^*_{t(n,m)}}(T - n(K+1)) \right| \\ > \ & \frac{D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1}{\kappa_{\max}(C_{\max}+1)}. \end{aligned}$$

Here, $j^*_{t(n,m)} \in \arg\max_{j\in[N]} \{|D_j[T-n(K+1), 1] - \lambda_j(T-n(K+1))|\}$ and $t(n,m) - m - 1 = T - n(K+1)$ from (EC.27). Similar to (EC.17), define

$$\mathbb{P}\left(E^j_{t(n,m)} | \Theta_{t(n,m)-1}, \dots, \Theta_{t(n,m)-m}\right) = \ A^{j,1}_{t(n,m)} + A^{j,2}_{t(n,m)},$$

where

$$\begin{aligned} A^{j,1}_{t(n,m)} & \triangleq \mathbb{I}\left(D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1 < 0\right) \\ & \quad \cdot \mathbb{P}\left(E^j_{t(n,m)} \middle| D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1 < 0\right), \\ A^{j,2}_{t(n,m)} & \triangleq \mathbb{I}\left(D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1 \ge 0\right) \\ & \quad \cdot \mathbb{P}\left(E^j_{t(n,m)} \middle| D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1 \ge 0\right). \end{aligned}$$

It is easy to see $\mathbb{P}\left(E^j_{t(n,m)} \middle| D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1 < 0\right) = 1$. It implies that

$$\begin{aligned}
&\sum_{n=1}^{\lceil T/(K+1)\rceil -1} \sum_{m=0}^{K} \mathbb{E}\left[A^{j,1}_{t(n,m)}\right] \\
= &\sum_{n=1}^{\lceil T/(K+1)\rceil -1} \sum_{m=0}^{K} \mathbb{P}\left(D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1 < 0\right)\cdot 1 \\
= &\sum_{m=0}^{K} \sum_{n=1}^{\lceil T/(K+1)\rceil -1} \mathbb{P}\left(\mathrm{Bin}(m,\lambda_j) < C_1 - \lambda_j(T-n(K+1))\right) \\
= &\sum_{m=0}^{K} \sum_{n=0}^{\lceil T/(K+1)\rceil -2} \mathbb{P}\left(\mathrm{Bin}(m,\lambda_j) < C_1 - \lambda_j(n(K+1)+R)\right) \\
\le &\sum_{m=0}^{K} \min\left\{\left\lceil \frac{T}{K+1}\right\rceil - 1,\ \left\lceil \frac{C_1}{\lambda_{\min}(K+1)}\right\rceil\right\} \mathbb{P}\left(\mathrm{Bin}(m,\lambda_j) < C_1 - \lambda_j R\right),
\end{aligned}$$

where the last equality comes from the expression of $T = (K+1)(\lceil T/(K+1)\rceil - 1) + R$ and the last inequality is similar to (EC.18). Thus, by applying Hoeffding's inequality and using a similar argument to (EC.19),

$$\sum_{n=1}^{\lceil T/(K+1)\rceil -1} \sum_{m=0}^{K} \mathbb{E}\left[A^{j,1}_{t(n,m)}\right] \le \min\left\{\left\lceil \frac{T}{K+1}\right\rceil - 1,\ \left\lceil \frac{C_1}{\lambda_{\min}(K+1)}\right\rceil\right\} \sum_{m=0}^{K} \rho_1^{m+2R-2\frac{C_1}{\lambda_{\min}}}, \tag{EC.28}$$

where $\rho_1 = \exp(-2\lambda_{\min}^2)$.

Next, we bound $\sum_{n=1}^{\lceil T/(K+1)\rceil -1} \sum_{m=0}^{K} \mathbb{E}\left[A^{j,2}_{t(n,m)}\right]$ from above: by Hoeffding's inequality,

$$\begin{aligned}
&\mathbb{P}\left(E^j_{t(n,m)} \middle| D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1 \ge 0\right) \\
\le &\ 2\exp\left(-2\frac{\left[D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1\right]^2}{(T-n(K+1))\cdot \kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right),
\end{aligned}$$

which implies

$$\mathbb{E}\left[A^{j,2}_{t(n,m)}\right] \le\ 2\mathbb{E}\left[\exp\left(-2\frac{\left[D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1\right]^2}{(T-n(K+1))\cdot \kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right)\right].$$

In addition, similar to (EC.21),

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(-2\frac{\left[D_j[t(n,m)-1, t(n,m)-m] + \lambda_j(T-n(K+1)) - C_1\right]^2}{(T-n(K+1))\cdot \kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right)\right] \\
\le &\ \exp\left(-\frac{2\lambda_{\min}^2(T-n(K+1))}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right) \exp\left(\frac{4\lambda_{\max}\cdot C_1}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right)\rho_2^m,
\end{aligned}$$

where $\rho_2 = \lambda_{\min} \exp\left(-\frac{4\lambda_{\min}}{\kappa_{\max}^2 (C_{\max}+1)^2}\right) + (1-\lambda_{\min})$. Thus, we have

$$\begin{aligned}
&\sum_{n=1}^{\lceil T/(K+1)\rceil - 1} \mathbb{E}\left[A_{t(n,m)}^{j,2}\right] \\
\le\ & 2 \sum_{n=1}^{\lceil T/(K+1)\rceil - 1} \left(\exp\left(-\frac{2\lambda_{\min}^2 (T - n(K+1))}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right) \exp\left(\frac{4\lambda_{\max} \cdot C_1}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right) \rho_2^m \right) \\
=\ & 2 \left(\sum_{n=0}^{\lceil T/(K+1)\rceil - 2} \exp\left(-\frac{2\lambda_{\min}^2 (n(K+1)+R)}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right) \right) \exp\left(\frac{4\lambda_{\max} \cdot C_1}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right) \rho_2^m \\
=\ & \frac{2\left(1 - \exp\left(-\frac{2\lambda_{\min}^2 (K+1)}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\left(\left\lceil \frac{T}{K+1}\right\rceil - 1\right)\right)\right) \exp\left(-\frac{2\lambda_{\min}^2}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2} R\right)}{1 - \exp\left(-\frac{2\lambda_{\min}^2}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}(K+1)\right)} \cdot \exp\left(\frac{4\lambda_{\max} \cdot C_1}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right) \cdot \rho_2^m \\
\le\ & \frac{2\left(1 - \exp\left(-\frac{2\lambda_{\min}^2 (K+1)}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\left(\left\lceil \frac{T}{K+1}\right\rceil - 1\right)\right)\right)}{1 - \exp\left(-\frac{2\lambda_{\min}^2}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right)} \cdot \exp\left(\frac{4\lambda_{\max} \cdot C_1}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right) \cdot \rho_3^R \cdot \rho_2^m \\
=\ & \frac{2\left(1 - \rho_3^{(K+1)\left(\left\lceil \frac{T}{K+1}\right\rceil - 1\right)}\right)}{1 - \rho_3} \cdot \exp\left(\frac{4\lambda_{\max} \cdot C_1}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right) \cdot \rho_3^R \cdot \rho_2^m,
\end{aligned} \tag{EC.29}$$

where $\rho_3 = \exp\left(-\frac{2\lambda_{\min}^2}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right)$.

Therefore, by combining (EC.28) and (EC.29), we conclude that

$$\begin{aligned}
&\sum_{n=1}^{\lceil T/(K+1)\rceil - 1} \sum_{m=0}^{K} \sum_{j \in [N]} \lambda_j \cdot \mathbb{P}\left(X_{t(n,m), s_{t(n,m)}}^* < a_{\max} + 1 \middle| \Theta_{t(n,m)} = j\right) \\
\le\ & \min\left\{ \left\lceil \frac{T}{K+1} \right\rceil - 1,\ \left\lceil \frac{C_1}{\lambda_{\min}(K+1)} \right\rceil \right\} \cdot \sum_{m=0}^{K} \rho_1^{m - 2\frac{C_1}{\lambda_{\min}}} \cdot \rho_1^{2R} \\
&+ \frac{2\left(1 - \rho_3^{(K+1)\left(\left\lceil \frac{T}{K+1}\right\rceil - 1\right)}\right)}{1 - \rho_3} \cdot \exp\left(\frac{4\lambda_{\max} \cdot C_1}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right) \cdot \sum_{m=0}^{K} \rho_2^m \cdot \rho_3^R \\
\le\ & \left(c_1 \cdot \sum_{m=0}^{K} \rho_1^{m - 2\frac{C_1}{\lambda_{\min}}} + \frac{2c_2 \cdot \exp\left(\frac{4\lambda_{\max} \cdot C_1}{\kappa_{\max}^2 \cdot (C_{\max}+1)^2}\right)}{1 - \rho_3} \cdot \sum_{m=0}^{K} \rho_2^m \right) \cdot \rho^R,
\end{aligned}$$

where $c_1 = \min\left\{ \left\lceil \frac{T}{K+1} \right\rceil - 1,\ \left\lceil \frac{C_1}{\lambda_{\min}(K+1)} \right\rceil \right\}$, $c_2 = 1 - \rho_3^{(K+1)\left(\left\lceil \frac{T}{K+1}\right\rceil - 1\right)}$ and $\rho = \max\{\rho_1^2, \rho_3\}$. The proof of Theorem 4 is completed. □