APPENDIX I
AUXILIARY RESULTS

Here, we provide necessary inequalities for analysis purpose

Proposition 1. ***Azuma-Hoeffding Inequality***

Let N be a positive integer and C be a positive real number. Suppose the random variables $X_1, ..., X_N$ constitute a martingale difference sequence with respect to the filtration $\{\mathcal{F}_n\}_{n=0}^N$, i.e. $\mathbb{E}[X_n|\mathcal{F}_{n-1}] = 0$ almost surely for every $n \in [N]$. In addition, suppose $X_n \in [-C, C]$ almost surely for every $n \in [N]$. For any fixed confidence level $\delta \in (0, 1)$, it holds that[1]

$$\begin{aligned} \Pr\left[\frac{1}{N}\sum_{n=1}^{N} X_n \geq C\sqrt{\frac{\log(1/\delta)}{2N}}\right] \leq \delta \\ \Pr\left[\frac{1}{N}\sum_{n=1}^{N} X_n \leq -C\sqrt{\frac{\log(1/\delta)}{2N}}\right] \leq \delta \end{aligned} \tag{1}$$

Proposition 2. ***Multiplicative Chernoff inequality***

We give a simplified extension of multiplicative Chernoff inequality[2]*. Suppose random variables $\{X_n\}_{n=1}^N$ are independent, and that $\Pr\left(X_n \in [0, C] \text{ for all } n \in [N]\right) = 1$ for some $C \in \mathbb{R}_+$. Denote $\mu = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} X_n\right]$. The following concentration inequalities hold for any fixed confidence level $\delta \in (0, 1)$*

$$\begin{aligned} \Pr\left(\frac{1}{N}\sum_{n=1}^{N} X_n - \mu \geq \sqrt{\frac{3C\mu}{N}\log\frac{1}{\delta}}\right) \leq \delta \\ \Pr\left(\frac{1}{N}\sum_{n=1}^{N} X_n - \mu \leq -\sqrt{\frac{2C\mu}{N}\log\frac{1}{\delta}}\right) \leq \delta. \end{aligned} \tag{2}$$

Proposition 3. ***Multiplicative Weight Update for Gains***

Let $\{\boldsymbol{l}(s)\}_{s=1}^T$ be an arbitrary sequence of vectors and $\eta \leq 1$, where $\boldsymbol{l}(s) = (l_i(s))_{i\in[I]} \in [-1, 1]^I$ for each $s \in [T]$. Consider the sequence of weight vectors $\boldsymbol{\vartheta}(1), \ldots, \boldsymbol{\vartheta}(T)$, where $\boldsymbol{\vartheta}(s) = (\vartheta_i(s))_{i\in[I]} \in \Delta^I$ is defined as

$$\vartheta_i(s) = \frac{\exp\left[\eta\sum_{t=1}^{s-1} l_i(t)\right]}{\sum_{j=1}^{I}\exp\left[\eta\sum_{t=1}^{s-1} l_j(t)\right]}$$

for each $s \in [T]$ and $i \in [I]$. Then, for any $i \in [I]$, it holds that[3]

$$\frac{1}{T}\sum_{s=1}^{T} l_i(s) \leq \frac{1}{T}\sum_{s=1}^{T}\sum_{j=1}^{I}\vartheta_j(s) l_j(s) + \frac{\log I}{T\eta} + \eta \tag{3}$$

[1] Mohri M., Rostamizadeh A., and Talwalkar A., "Concentration Inequalities," in *Foundations of Machine Learning*, second ed. USA: MIT press, 2018, pp. 441–442.

[2] Mohri M., Rostamizadeh A., and Talwalkar A., "Concentration Inequalities," in *Foundations of Machine Learning*, second ed. USA: MIT press, 2018, pp. 439–440.

[3] Arora, Hazan, Kale. (2012). The Multiplicative Weights Update Method: a Meta Algorithm and Applications. Theory of Computing [electronic only]. 8. 10.4086/toc.2012.v008a006.

APPENDIX II
PROOFS

A. Proof of Lemma 1

Proof. Suppose there is a non-anticipatory feasible policy that achieves the expected optimum $\mathbb{E}\left[\lambda_*^B\right]$ in (BP-C), i.e. $\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{I} W_i^k(t)A_i^k(t)D_i(t)X^k(t)\right] = \mathbb{E}[\text{opt(BP-C)}]$. Define $\boldsymbol{x} = \left\{x^k(t)\right\}_{t\in[T]}$ as $x^k(t) = \Pr(X^k(t) = 1)$. We claim that $\boldsymbol{x}$ is feasible to (LP-E), and verifying the claims about the feasibility and the objective value proves the claim.

We first verify the feasibility to (LP-E). Since the proposed policy satisfies the reusable resource constraints, the inequality $\sum_{\tau=1}^{t}\sum_{k\in\mathcal{K}} \mathbf{1}\left(D_i(\tau) \geq t-\tau+1\right) A_i^k(\tau)X^k(\tau) \leq c_i$ holds for all $i \in [I], t \in [T]$. Taking expectation over $X^k(\tau), D_i(\tau)$, and $A_i^k(\tau)$ for $\tau \in [t]$ on the left-hand side gives

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{\tau=1}^{t}\sum_{k\in\mathcal{K}} \mathbf{1}\left(D_i(\tau) \geq t-\tau+1\right) A_i^k(\tau)X^k(\tau)\right]\\
=&\sum_{\tau=1}^{t}\sum_{k\in\mathcal{K}} \mathbb{E}\left[\mathbf{1}\left(D_i(\tau) \geq t-\tau+1\right) A_i^k(\tau)X^k(\tau)\right]\\
=&\sum_{\tau=1}^{t}\sum_{k\in\mathcal{K}} \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}\left(D_i(\tau) \geq t-\tau+1\right) A_i^k(\tau) \mid X^k(\tau)\right] X^k(\tau)\right]\\
=&\sum_{\tau=1}^{t}\sum_{k\in\mathcal{K}} \mathbb{E}\left[\mathbf{1}\left(D_i \geq t-\tau+1\right) A_i^k\right] x(\tau) \leq c_i
\end{aligned}
$$

Hence, the claim about the objective value is shown, and the Lemma is proved. □

B. Proof of Lemma 2

Proof. We consider a fixed $i \in [I]$, and apply the multiplicative Chernoff inequality with $X_t = \sum_{i=1}^{I}\sum_{k\in\mathcal{K}} w_i^k v_i^k y_*^k$ for $t \in [\tau^{(q-1)}]$, and $C = \sum_{i=1}^{I} \bar{J}_i$. Suppose $\lambda_* = \sum_{i=1}^{I}\sum_{k\in\mathcal{K}} w_i^k v_i^k y_*^k$, and $\hat{\lambda}_* = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{k\in\mathcal{K}} J_i^k(t)x_*^k(t)$, $\hat{\gamma}_* = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{k\in\mathcal{K}} M_i^k(t)x_*^k(t)$

$$
\hat{\lambda}_* - \lambda_* \geq -\sqrt{\frac{2C\mu}{\tau^{(q-1)}}\log\frac{1}{\delta}} \qquad w.p. 1-\delta
$$

$$
\hat{\lambda}_* + \sum_{i=1}^{I} \bar{J}_i \cdot \sqrt{\frac{2\log(1/\delta)}{\tau^{(q-1)}}} \geq \lambda_* \qquad w.p. 1-\delta
$$

Both , then

$$
\bar{J}\hat{\gamma}_* + C\sqrt{\frac{2\log(1/\delta)}{\tau^{(q-1)}}} \geq \lambda_* \qquad w.p. 1-\delta
$$

□

C. Proof of Lemma 3

Proof. The proof relies on a crucial application of Proposition 3, with a judicious choice of $l(1), \ldots, l(\tau)$ (where we set $\tau = \tau^{(q-1)}$) that underpins the construction of Algorithm 2. Now, for each $s \in [\tau^{(q-1)}]$, we define

$$
l_i(s) = \begin{cases} M_i^{k^{(q)}}(s), & \forall i \in [I] \\ N_i^{k^{(q)}}(s), & \forall i \in [I] \end{cases} \tag{4}
$$

It is evident that $|l_i(s)| \leq 1, \forall i \in [I], s \in [\tau^{(q-1)}]$. In addition, under the specification of $\{l(s)\}_{s=1}^{\tau}$ in (4), it can be directly verified that the MWU weigh vector $\vartheta(s)$ in (**??**) is equal to $\left\{\phi^{(q)}(s)\right\}$ for each $s \in [\tau^{(q-1)}]$. Applying Proposition 3 gives us the following inequalities, which holds with certainty

$$
\frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} M_i^{k^{(q)}}(s) \leq \Xi^{(q)} + \frac{\log 2I}{\tau^{(q-1)}\eta} + \eta \quad \forall i \in [I] \tag{5}
$$

$$
\frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} N_i^{k^{(q)}}(s) \leq \Xi^{(q)} + \frac{\log 2I}{\tau^{(q-1)}\eta} + \eta \quad \forall i \in [I] \tag{6}
$$

$$
\tag{7}
$$

where

$$\Xi^{(q)} = \frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{l\in[I]} \left[\phi_l^{(q)}(s) M_i^{k(s)} + \psi_l^{(q)}(s) N_i^{k(s)}\right] \tag{8}$$

Moreover, if we take $\eta = \sqrt{\frac{\log 2I}{\tau^{(q-1)}}}$ in (5), we yield a conservative upperbound for $\frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} M_i^{k^{(q)}}(s)$, that is

$$\frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} M_i^{k^{(q)}}(s) \le \Xi^{(q)} + 2\sqrt{\frac{\log 2I}{\tau^{(q-1)}}} \quad \forall i \in [I]$$

but $\eta = \sqrt{\frac{\log 2I}{\tau^{(q-1)}}}$ has to be less than 1 if we want to yield this upperbound, which means $I \le \frac{1}{2}\exp\{\tau^{(q-1)}\}$. Then, to prove this lemma, the following 2 inequalities have to be satisfied:

$$\Pr\left(\Xi^{(q)} \ge -\sqrt{\frac{\log(1/\delta)}{2\tau^{(q-1)}}}\right) \ge 1-\delta \quad \text{or} \quad \Pr\left(\Xi^{(q)} \le \sqrt{\frac{\log(1/\delta)}{2\tau^{(q-1)}}}\right) \ge 1-\delta \tag{9}$$

$$\Pr\left(\frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{i=1}^{I} w_i v_i^{k(\phi^{(q)},\psi^{(q)})}(s) \ge \lambda_* - \epsilon_B^{(q)}\right) \ge 1-2\delta \tag{10}$$

$$\Pr\left(\frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} v_i^{k(\phi^{(q)},\psi^{(q)})}(s) \le \epsilon_A^{(q)}\right) \ge 1-2\delta \tag{11}$$

where

$$\epsilon_A^{(q)} = \bar{V}\left(1 + \sqrt{\frac{\log(1/\delta)}{2\tau^{(q-1)}}} - 2\sqrt{\frac{\log 2I}{\tau^{(q-1)}}} + \sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}}\right) \tag{12}$$

$$\epsilon_B^{(q)} = C\sqrt{\frac{2\log(1/\delta)}{\tau^{(q-1)}}} + C\sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}} \tag{13}$$

First of all, we prove (9). Based on (8), notice that weight vectors $\{\phi^{(q)}(s), \psi^{(q)}(s)\}^I$ are all within $[0,1]$. We apply Proposition 1, and give (9).

Secondly, we prove (10). Given (**??**), (9) and Proposition 1 for the martingale difference sequence with respect to the filtration $\{\mathcal{F}(s)\}_{s=1}^{\tau^{(q-1)}}$ defined as $\{\mathcal{F}(s)\} = \sigma(\{\hat{\lambda}_*^{(q)}\} \cup \{J_i(s)\})$. The expectation $\mathbb{E}[J_i(s)|\mathcal{F}(s-1)]$ is only taken over the randomness in $V_i(s)$, and note that $\{\phi(s), \psi(s)\}^I$ are $\mathcal{F}(s-1)$-measurable. Now we proceed with $\{J_i^{k^{(q)}}(s) - w_i^{k(\phi^{(q)},\psi^{(q)})}(s) \cdot v_i^{k(\phi^{(q)},\psi^{(q)})}(s)\}_{s=1}^{\tau^{(q-1)}}$ for all $i \in [I]$. By applying Azuma-Hoeffding inequality, we have

$$\Pr\left(\frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} J_i^{k^{(q)}}(s) - \frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} w_i^{k(\phi^{(q)},\psi^{(q)})}(s) \cdot v_i^{k(\phi^{(q)},\psi^{(q)})}(s) \ge \bar{J}_i\sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}}\right) \le \frac{\delta}{I}$$

for a fixed i. Reformulate and sum up all events for all $i \in [I]$, we have

$$\Pr\left(\frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{i=1}^{I} J_i^{k^{(q)}}(s) - \frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{i=1}^{I} w_i v_i^{k(\phi^{(q)},\psi^{(q)})}(s) \ge \sum_{i\in[I]} \bar{J}_i\sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}}\right) \le \delta$$

$$\Pr\left(\frac{\bar{J}}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{i=1}^{I} M_i^{k^{(q)}}(s) - \frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{i=1}^{I} w_i v_i^{k(\phi^{(q)},\psi^{(q)})}(s) \ge C\sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}}\right) \le \delta$$

$$\Pr\left(\lambda_* - C\sqrt{\frac{2\log(1/\delta)}{\tau^{(q-1)}}} - \frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{i=1}^{I} w_i v_i^{k(\phi^{(q)},\psi^{(q)})}(s) \ge C\sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}}\right) \le 2\delta$$

$$\Pr\left(\lambda_* - C\sqrt{\frac{2\log(1/\delta)}{\tau^{(q-1)}}} - C\sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}} \le \frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{i=1}^{I} w_i v_i^{k(\phi^{(q)},\psi^{(q)})}(s)\right) \ge 1-2\delta$$

$$\Pr\left(\lambda_* - \epsilon_B^{(q)} \le \frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{i=1}^{I} w_i v_i^{k(\phi^{(q)},\psi^{(q)})}(s)\right) \ge 1-2\delta$$

Therefore, inequality (10) is proved.

Inequality (11) can be proved in a different way. We proceed with $\{N_i^{k^{(q)}}(s) - N_i^{k(\phi^{(q)},\psi^{(q)})}(s)\}_{s=1}^{\tau^{(q-1)}}$ for all $i \in [I]$. By applying Azuma-Hoeffding inequality, we have

$$\Pr\left(\frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} V_i^{k^{(q)}}(s) - \frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} v_i^{k(\phi^{(q)},\psi^{(q)})}(s) \leq -\bar{V}_i\sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}}\right) \leq \frac{\delta}{I}$$

for an arbitrary i. Reformulate and take a union bound for all $i \in [I]$, we have

$$\Pr\left(\frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} V_i^{k^{(q)}}(s) \leq \frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} v_i^{k(\phi^{(q)},\psi^{(q)})}(s) - \bar{V}_i\sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}}\right) \leq \delta \tag{14}$$

introduce 3 on the left side of (14) gives

$$1 - \Xi^{(q)} - 2\sqrt{\frac{\log 2I}{\tau^{(q-1)}}} \leq \frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}}\left(\frac{V_i^{k^{(q)}}(s)}{\bar{V}_i}\right) \quad \forall i \in [I]$$

$$\Pr\left(1 - \Xi^{(q)} - 2\sqrt{\frac{\log 2I}{\tau^{(q-1)}}} + \sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}} \leq \frac{1}{\bar{V}_i\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} v_i^{k(\phi^{(q)},\psi^{(q)})}(s)\right) \leq \delta$$

introduce (9) to eliminate $\Xi^{(q)}$, and take $\eta = \sqrt{\frac{\log 2I}{\tau^{(q-1)}}}$, we have

$$\Pr\left(1 + \sqrt{\frac{\log(1/\delta)}{2\tau^{(q-1)}}} - 2\sqrt{\frac{\log 2I}{\tau^{(q-1)}}} + \sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}} \leq \frac{1}{\bar{V}_i\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} v_i^{k(\phi^{(q)},\psi^{(q)})}(s)\right) \leq 2\delta$$

suppose $\bar{V}_i\left(1 + \sqrt{\frac{\log(1/\delta)}{2\tau^{(q-1)}}} - 2\sqrt{\frac{\log 2I}{\tau^{(q-1)}}} + \sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}}\right) = \epsilon_A^{(q)}$, for all $i \in [I]$, we have

$$\Pr\left(\frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} v_i^{k(\phi^{(q)},\psi^{(q)})}(s) \leq \epsilon_A^{(q)}\right) \geq 1 - 2\delta$$

Therefore, inequality (10) is proved. □

D. Proof of Lemma 5

Proof. The Lemma is proved by established three steps. We start by observing that

$$\mathbb{E}[L(t)|\mathcal{H}(\tau^{(q-1)})] \geq 1 - \sum_{i\in[I]}\mathbb{E}\left[\mathbf{1}\left(\sum_{\tau=\max\{t-\bar{d},1\}}^{t-1}\hat{A}_i(\tau)\mathbf{1}(\hat{D}_i(\tau) \geq t-\tau+1) > c_i - \bar{a}\right)\middle|\mathcal{H}(\tau^{(q-1)})\right]$$

we denote the expectation term $\mathbb{E}\left[\mathbf{1}\left(\sum_{\tau=\max\{t-\bar{d},1\}}^{t-1}\hat{A}_i(\tau)\mathbf{1}(\hat{D}_i(\tau) \geq t-\tau+1) > c_i - \bar{a}_i\right) \mid \mathcal{H}(\tau^{(q-1)})\right]$ as $G_i^{(q)}(t)$. Firstly, we demonstrate that for any $t \in \left\{\tau^{(q-1)} + 1 + \bar{d}, \ldots, \tau^{(q)}\right\}$, any $i \in [I]$ and any fixed $\varepsilon > 0$, the inequality

$$G_i^{(q)}(t) \leq \frac{1}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}-1}} \cdot \exp\left[\frac{\varepsilon}{\bar{a}_i} \cdot \frac{c_{\min}}{\epsilon_A^{(q)} + \bar{V}\beta} \cdot \frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}} v_i^{k(\phi^{(q)},\psi^{(q)})}(s)\right] \tag{15}$$

holds with at least $1 - 2\delta$ probability. Secondly, by setting $\varepsilon = \frac{\beta}{\epsilon_A^{(q)}+1}$, we demonstrate the inequality

$$\frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}} \cdot \exp\left[\frac{\varepsilon}{\bar{a}_i} \cdot \frac{c_{\min}}{\epsilon_A^{(q)} + \bar{V}\beta} \cdot \epsilon_A^{(q)}\right] \leq \frac{2\sqrt{\xi}}{I} \tag{16}$$

which holds with certainty since inequality (16) only involves deterministic parameters. To show the Lemma, we just have to prove (15) and (16). Inequality (15) is shown by the following string of calculations

$$\begin{aligned}
&G_i^{(q)}(t)\\
&=\mathbb{E}\left[\mathbf{1}\left(\sum_{\tau=\max\{t-\bar{d},1\}}^{t-1}\hat{A}_i(\tau)\mathbf{1}(\hat{D}_i(\tau)\geq t-\tau+1)>c_i-\bar{a}_i\right)\middle|\mathcal{H}(\tau^{(q-1)})\right]\\
&=\mathbb{E}\left\{\mathbf{1}\left[(1+\varepsilon)^{\sum_{\tau=\max\{t-\bar{d},1\}}^{t-1}\frac{\hat{A}_i(\tau)}{\bar{a}_i}\mathbf{1}\left(\hat{D}_i(\tau)\geq t-\tau+1\right)}>(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}-1}\right]\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right\}
\end{aligned}$$

By Markov inequality, we have

$$\begin{aligned}
&\leq\frac{1}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}-1}}\mathbb{E}\left[(1+\varepsilon)^{\sum_{\tau=\max\{t-\bar{d},1\}}^{t-1}\frac{\hat{A}_i(\tau)}{\bar{a}_i}\mathbf{1}\left(\hat{D}_i(\tau)\geq t-\tau+1\right)}\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]\\
&=\frac{1}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}-1}}\prod_{\tau=\max\{t-\bar{d},1\}}^{t-1}\mathbb{E}\left[(1+\varepsilon)^{\frac{\hat{A}_i(\tau)}{\bar{a}_i}\mathbf{1}\left(\hat{D}_i(\tau)\geq t-\tau+1\right)}\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]
\end{aligned}$$

By the fact that $(1+\varepsilon)^a\leq 1+\varepsilon\cdot a$ for all $a\in[0,1],\varepsilon>0$

$$\begin{aligned}
&\leq\frac{1}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}-1}}\prod_{\tau=\max\{t-\bar{d},1\}}^{t-1}\left(1+\varepsilon\cdot\mathbb{E}\left[\frac{\hat{A}_i(\tau)}{\bar{a}_i}\mathbf{1}\left(\hat{D}_i(\tau)\geq t-\tau+1\right)\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]\right)\\
&\leq\frac{1}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}-1}}\prod_{\tau=\max\{t-\bar{d},1\}}^{t-1}\left(1+\varepsilon\cdot\mathbb{E}[B(\tau)]\mathbb{E}\left[\frac{\tilde{A}_i(\tau)}{\bar{a}_i}\mathbf{1}\left(\tilde{D}_i(\tau)\geq t-\tau+1\right)\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]\right)
\end{aligned}$$

By the inequality $1+\varepsilon\leq e^{\varepsilon}$ which holds for all $\varepsilon>0$

$$\begin{aligned}
&\leq\frac{1}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}-1}}\prod_{\tau=\max\{t-\bar{d},1\}}^{t-1}\exp\left(\varepsilon\cdot\mathbb{E}[B(\tau)]\mathbb{E}\left[\frac{\tilde{A}_i(\tau)}{\bar{a}_i}\mathbf{1}\left(\tilde{D}_i(\tau)\geq t-\tau+1\right)\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]\right)\\
&=\frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}}\exp\left(\frac{\varepsilon}{\bar{a}_i}\mathbb{E}[B(\tau)]\sum_{\tau=\max\{t-\bar{d},1\}}^{t-1}\mathbb{E}\left[\tilde{A}_i(\tau)\mathbf{1}\left(\tilde{D}_i(\tau)\geq t-\tau+1\right)\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]\right)\\
&\leq\frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}}\exp\left(\frac{\varepsilon}{\bar{a}_i}\mathbb{E}[B(\tau)]\sum_{\tau=t-\bar{d}+1}^{t}\sum_{s=t-\tau+1}^{\bar{d}}\mathbb{E}\left[\tilde{A}_i(\tau)\mathbf{1}\left(\tilde{D}_i(\tau)=s\right)\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]\right)\\
&=\frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}}\exp\left(\frac{\varepsilon}{\bar{a}_i}\mathbb{E}[B(\tau)]\sum_{s=1}^{\bar{d}}\sum_{\tau=t-s+1}^{t}\mathbb{E}\left[\tilde{A}_i(\tau)\mathbf{1}\left(\tilde{D}_i(\tau)=s\right)\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]\right)
\end{aligned}$$

Note that $\{(\tilde{A}_i(\tau),\tilde{D}_i(\tau))\}_{\tau=\max\left\{t-\bar{d},1\right\}}^{t}$ are i.i.d. conditioned on $\mathcal{H}\left(\tau^{(q-1)}\right)$, leads to

$$\begin{aligned}
&=\frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}}\exp\left(\frac{\varepsilon}{\bar{a}_i}\frac{c_{\min}}{\epsilon_A^{(q)}+\bar{V}\beta}\sum_{s=1}^{\bar{d}}\sum_{\tau=t-s+1}^{t}\mathbb{E}\left[\tilde{A}_i(t)\mathbf{1}\left(\tilde{D}_i(t)=s\right)\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]\right)\\
&=\frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}}\exp\left(\frac{\varepsilon}{\bar{a}_i}\frac{c_{\min}}{\epsilon_A^{(q)}+\bar{V}\beta}\sum_{s=1}^{\bar{d}}\mathbb{E}\left[\tilde{A}_i(t)\cdot s\cdot\mathbf{1}\left(\tilde{D}_i(t)=s\right)\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]\right)\\
&=\frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}}\exp\left(\frac{\varepsilon}{\bar{a}_i}\frac{c_{\min}}{\epsilon_A^{(q)}+\bar{V}\beta}\mathbb{E}\left[\tilde{A}_i(t)\tilde{D}_i(t)\middle|\mathcal{H}\left(\tau^{(q-1)}\right)\right]\right)\\
&=\frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}}\exp\left(\frac{\varepsilon}{\bar{a}_i}\frac{c_{\min}}{\epsilon_A^{(q)}+\bar{V}\beta}\frac{1}{\tau^{(q-1)}}\sum_{s=1}^{\tau^{(q-1)}}v_i^{k(\phi^{(q)},\psi^{(q)})}(s)\right)
\end{aligned}$$

Then, we assume $\varepsilon = \frac{\beta}{\bar{\epsilon}_A^{(q)}+1}$ (recall $\beta \leq 1$, and $\bar{\epsilon}_A^{(q)} = \epsilon_A^{(q)}/\bar{V}$).

$$
\begin{aligned}
&\frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}} \exp\left[\frac{\varepsilon}{\bar{a}_i} \cdot \frac{c_{\min}}{\epsilon_A^{(q)} + \bar{V}\beta} \cdot \frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} v_i^{k(\phi^{(q)},\psi^{(q)})}(s)\right] \\
\leq& \frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}} \exp\left[\frac{\varepsilon}{\bar{a}_i} \cdot \frac{c_{\min}}{\epsilon_A^{(q)} + \bar{V}\beta} \cdot \epsilon_A^{(q)}\right] \\
\leq& \frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}} \exp\left[\frac{\varepsilon}{\bar{a}_i} \cdot \frac{c_i}{1+\beta/\bar{\epsilon}_A^{(q)}}\right] \\
\leq& \frac{1+\varepsilon}{(1+\varepsilon)^{\frac{c_i}{\bar{a}_i}}} \exp\left[\frac{\varepsilon}{\bar{a}_i} \cdot \frac{c_i}{1+\varepsilon}\right] \\
=& (1+\varepsilon)\left[\frac{e^{\varepsilon}}{(1+\varepsilon)^{1+\varepsilon}}\right]^{\frac{c_i}{\bar{a}_i(1+\varepsilon)}} \\
=& (1+\varepsilon)\exp\left[\frac{1}{\xi} \cdot \frac{\varepsilon}{1+\varepsilon} - \frac{1}{\xi} \cdot \log(1+\varepsilon)\right] \\
\approx& (1+\varepsilon)\exp\left[\frac{1}{\xi} \cdot \frac{\varepsilon}{1+\varepsilon} - \frac{1}{\xi} \cdot \frac{\varepsilon(2+\varepsilon)}{(1+\varepsilon)^2}\right] \\
\leq& (1+\varepsilon)\exp\left[-\frac{\varepsilon}{(1+\varepsilon)^2} \cdot \frac{1}{\xi}\right] \\
\leq& 2\exp\left[-\frac{\sqrt{\xi}\log(I/\sqrt{\xi})}{(1+\varepsilon)^2(\bar{\epsilon}_A^{(q)}+1)} \cdot \frac{1}{\xi}\right] \\
\leq& 2\exp\left[-\frac{\log(I/\sqrt{\xi})}{8\sqrt{\xi}}\right]
\end{aligned} \tag{17}
$$

Suppose $8\sqrt{\xi} \leq 1$ and the fact that $\xi \geq 0$, (17) can be upper bounded as

$$
G_i^{(q)}(t) \leq 2\exp\left(-\log\frac{I}{\sqrt{\xi}}\right) = 2\frac{\sqrt{\xi}}{I} \qquad w.p. \quad 1-2\delta
$$

Therefore, we have

$$
\mathbb{E}[L(t)|\mathcal{H}(\tau^{(q-1)})] \geq 1 - \sum_{i\in[I]} G_i^{(q)}(t) = 1 - 2\sqrt{\xi} \qquad w.p. \quad 1-2\delta
$$

□

E. Proof of Lemma 6

Proof. By the coupling argument that constructs $\hat{J}_i(t)$, for every $t \in \left\{\bar{\tau}^{(q-1)}+1,\ldots,\tau^{(q)}\right\}$ and every $i \in [I]$, with certainty we have

$$
\begin{aligned}
\sum_{i=1}^{I}\mathbb{E}\left[\hat{J}_i(t) \mid \mathcal{H}\left(\tau^{(q-1)}\right)\right] &= \sum_{i=1}^{I}\mathbb{E}\left[B(t)\tilde{W}_i(t)\tilde{A}_i(t)\tilde{D}_i \mid \mathcal{H}\left(\tau^{(q-1)}\right)\right] \\
&= \mathbb{E}[B(t)] \cdot \sum_{i=1}^{I}\mathbb{E}\left[\tilde{W}_i(t)\tilde{A}_i(t)\tilde{D}_i \mid \mathcal{H}\left(\tau^{(q-1)}\right)\right] \\
&= \frac{c_{\min}}{\epsilon_A^{(q)} + \bar{V}\beta} \cdot \frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{i=1}^{I} w_i v_i^{k(\phi^{(q)},\psi^{(q)})}(s)
\end{aligned}
$$

Applying Lemma 3, we know that the inequality

$$
\begin{aligned}
&\frac{c_{\min}}{\epsilon_A^{(q)} + \bar{V}\beta} \cdot \frac{1}{\tau^{(q-1)}} \sum_{s=1}^{\tau^{(q-1)}} \sum_{i=1}^{I} w_i v_i^{k(\phi^{(q)},\psi^{(q)})}(s) \\
\geq& \frac{c_{\min}}{\epsilon_A^{(q)} + \bar{V}\beta} \cdot (\lambda_* - \epsilon_B^{(q)}) \\
\geq& \frac{c_{\min}}{2\bar{V}}(\lambda_* - \epsilon_B^{(q)})
\end{aligned}
$$

holds for all $i \in [I]$ with probability at least $1-2\delta$. Thus, we obtain the desired lower bound. □

F. Proof of Lemma 7

Proof. Indeed, the random variables $\{L(t)\hat{J}.(t)\}_{t=\bar{\tau}^{(q-1)}+1}^{\tau^{(q)}}$ are correlated even when we condition on $\mathcal{H}\left(\tau^{(q-1)}\right)$. Instead, we apply Lemma 7 and partial proof of Lemma 6 in Zhang and Chi (2022) [4] on suitable subsets of $\{L(t)\hat{J}.(t)\}_{t=\bar{\tau}^{(q-1)}+1}^{\tau^{(q)}}$ that partition $\{L(t)\hat{J}.(t)\}_{t=\bar{\tau}^{(q-1)}+1}^{\tau^{(q)}}$. To this end, we define $N = \left\lceil \frac{\tau^{(q)}-\tau^{(q-1)}}{\bar{d}+1} \right\rceil$. For $l \in [\bar{d}+1]$ and $n \in [N]$, we define a new time index

$$k(n,l) = \bar{\tau}^{(q-1)} + l + (n-1)\cdot\left(\bar{d}+1\right)$$

Clearly, we have $\{k(n,l)\}_{n\in[N],l\in[\bar{d}+1]} \supseteq \left\{\bar{\tau}^{(q-1)}+1,\ldots,\tau^{(q)}\right\}$. Crucially, we observe that for any $l \in [\bar{d}]$, the random variables in the collection

$$\Gamma(l) = \left\{L(k(n,l))\cdot\hat{J}.(k(n,l))\right\}_{n=1}^{N}$$

are independent and identically distributed conditioned on $\mathcal{H}\left(\tau^{(q-1)}\right)$. We first show the conditional independence. For every t, the random variable $L(t)\hat{J}.(t)$ is $\sigma\left(\left\{(\hat{W}_i(\tau),\hat{A}_i(\tau),\hat{D}_i(\tau))\right\}_{\tau=t-\bar{d}}^{t}\right)$-measurable. More precisely, by the definition of $L(t),\hat{J}.(t)$, we know that there is a deterministic function g_i such that $L(t)\hat{J}.(t) = g_i\left(\{(\hat{W}_i(\tau),\hat{A}_i(\tau),\hat{D}_i(\tau))\}_{\tau=t-\bar{d}}^{t}\right)$, where g_i does not vary with t and only depends on i. Since the time indexes in $\Gamma(l)$ are at least $\bar{d}+1$ time steps apart, we know that for any two distinct $n,n' \in [N]$, the time indexes sets $\left\{k(n,l)-\bar{d},\ldots,k(n,l)\right\}$ and $\left\{k(n',l)-\bar{d},\ldots,k(n',l)\right\}$ are disjoint. By observing that $\{(\hat{W}_i(\tau),\hat{A}_i(\tau),\hat{D}_i(\tau))\}_{\tau=\tau^{(q-1)}-\bar{d}}^{\tau^{(q)}}$ are independent conditioned on $\mathcal{H}\left(\tau^{(q-1)}\right)$, we know that the random variables in $\Gamma(l)$ are independent conditioned on $\mathcal{H}\left(\tau^{(q-1)}\right)$.

The identically distributed part follows from the fact that, for any $t \in \left\{\bar{\tau}^{(q-1)}+1,\ldots,\tau^{(q)}\right\}$, we know that $\left\{t-\bar{d},\ldots,t\right\} \subset \left\{\tau^{(q-1)}+1,\ldots,\tau^{(q)}\right\}$. In addition, by the coupling argument on the construction of $\hat{W},\hat{A},\hat{D}$, we know that $\{(\hat{W}(t),\hat{A}(t),\hat{D}(t))\}_{t=\tau^{(q-1)}+1}^{(q)}$ are identically distributed conditioned on $\mathcal{H}\left(\tau^{(q-1)}\right)$, since all the time indexes in $\left\{\tau^{(q-1)}+1,\ldots,\tau^{(q)}\right\}$ belong to phase q.

We know that $\{L(t)\hat{J}.(t)\}_{t=\tau^{(q-1)}+1}^{\tau^{(q)}}$ are identically distributed conditioned on $\mathcal{H}\left(\tau^{(q-1)}\right)$, which in particular implies that the random variables in $\Gamma(l)$ are identically distributed conditioned on $\mathcal{H}\left(\tau^{(q-1)}\right)$.

After establishing the claim that the random variables in $\Gamma(l)$ are conditionally i.i.d. for any l, we apply the conditional Chernoff inequality (Lemma 7 in Zhang and Chi (2022)) on the random variables in $\Gamma(l)$, along with $\mathcal{F} = \mathcal{H}\left(\tau^{(q-1)}\right)$ and $C = \sum_i \bar{J}_i$. Summing (11) over $t \in \{k(n,l)\}_{n=1}^{N}$ gives us that, with probability at least $1-4\delta$, it holds that

$$\sum_{n=1}^{N}\sum_{i=1}^{I}\mathbb{E}\left[\hat{J}_i(k(n,l))L(k(n,l)) \mid \mathcal{H}\left(\tau^{(q-1)}\right)\right] = \mu_-(l) \geq \frac{c_{\min}\left(1-2\sqrt{\xi}\right)\left(\lambda_* - \epsilon_B^{(q)}\right)}{\epsilon_A^{(q)}+\bar{V}\beta}$$

for all $l \in [\bar{d}+1]$. Observe that $\mu_-(l) \leq N\sum_i \bar{J}_i$ for all $l \in [\bar{d}+1]$ a.s. The conditional Chernoff inequality shows us that with probability $\geq 1-5\delta$, we have

$$\sum_{n=1}^{N}\sum_{i=1}^{I}\hat{J}_i(k(n,l))L(t(n;l)) \geq \mu_-(l) - \sqrt{2C\mu_-\log\left(\frac{\bar{d}+1}{\delta}\right)} \geq \mu_-(l) - C\sqrt{2N\log\left(\frac{\bar{d}+1}{\delta}\right)} \tag{18}$$

for all $l \in [\bar{d}+1]$. Summing (18) over l gives

$$\begin{aligned}
&\sum_{t=\tau^{(q-1)}+1}^{\tau^{(q)}}\sum_i \hat{J}_i(t)L(t)\\
\geq&\frac{c_{\min}\left(1-2\sqrt{\xi}\right)\left(\lambda_*-\epsilon_B^{(q)}\right)}{2\bar{V}}\left(\tau^{(q)}-\bar{\tau}^{(q-1)}\right) - \left(\bar{d}+1\right)C\sqrt{2N\log\left(\frac{\bar{d}+1}{\delta}\right)}\\
\geq&\frac{c_{\min}\left(1-2\sqrt{\xi}\right)\left(\lambda_*-\epsilon_B^{(q)}\right)}{2\bar{V}}\left(\tau^{(q)}-\tau^{(q-1)}\right) - \left(\bar{d}+1\right)C\sqrt{2\frac{\tau^{(q)}-\tau^{(q-1)}}{\bar{d}+1}\log\left(\frac{\bar{d}+1}{\delta}\right)}\\
\geq&\frac{c_{\min}\left(1-2\sqrt{\xi}\right)}{2\bar{V}}\lambda_*\left(\tau^{(q)}-\tau^{(q-1)}\right) - \frac{c_{\min}\left(1-2\sqrt{\xi}\right)}{2\bar{V}}\epsilon_B^{(q)}\left(\tau^{(q)}-\tau^{(q-1)}\right)\\
&\qquad - C\sqrt{2(\tau^{(q)}-\tau^{(q-1)})(\bar{d}+1)\log\left(\frac{\bar{d}+1}{\delta}\right)}
\end{aligned}$$

[4] Xilin, Zhang and Cheungwang Chi. Online Resource Allocation for Reusable Resources. 2022, [Online]. Available: `https://arxiv.org/abs/2212.02855`

With the fact that $\epsilon_B^{(q)} = C\sqrt{\frac{2\log(1/\delta)}{\tau^{(q-1)}}} + C\sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}}$, we have

$$
\begin{aligned}
\geq & \frac{c_{\min}\left(1-2\sqrt{\xi}\right)}{2\bar{V}} \lambda_* \left(\tau^{(q)} - \tau^{(q-1)}\right) - \frac{c_{\min}\left(1-2\sqrt{\xi}\right) C}{2\bar{V}} \left(\sqrt{\frac{2\log(1/\delta)}{\tau^{(q-1)}}} + \sqrt{\frac{\log(I/\delta)}{2\tau^{(q-1)}}}\right) \left(\tau^{(q)} - \tau^{(q-1)}\right) \\
& - C\sqrt{2(\tau^{(q)} - \tau^{(q-1)})(\bar{d}+1)\log\left(\frac{\bar{d}+1}{\delta}\right)} \\
= & \frac{c_{\min}\left(1-2\sqrt{\xi}\right)}{2\bar{V}} \lambda_* \left(\tau^{(q)} - \tau^{(q-1)}\right) - \\
& C\sqrt{\tau^{(q)} - \tau^{(q-1)}} \left[\frac{c_{\min}\left(1-2\sqrt{\xi}\right)}{2\bar{V}} \left(\sqrt{2\log(1/\delta)} + \sqrt{\frac{\log(I/\delta)}{2}}\right) + \sqrt{2(\bar{d}+1)\log\left(\frac{\bar{d}+1}{\delta}\right)}\right]
\end{aligned}
$$

If phase q goes to infinity, $\tau^{(q)} - \tau^{(q-1)}$ and $\epsilon_A^{(q)}$ expand as well, we shall have

$$
\begin{aligned}
& \sum_{t=\tau^{(q-1)}+1}^{\tau^{(q)}} \sum_i \hat{J}_i(t) L(t) \\
\geq & \frac{c_{\min}\left(1-2\sqrt{\xi}\right)}{2\bar{V}} \left(\tau^{(q)} - \tau^{(q-1)}\right) \lambda_* - O\left(\sqrt{\tau^{(q)} - \tau^{(q-1)}}\right)
\end{aligned}
$$

□