Nearly Optimal Competitive Ratio for Online Allocation Problems with Two-sided Resource Constraints and Finite Requests

Qixin Zhang [*1,2] **Wenbing Ye** [*2] **Zaiyi Chen** [*2] **Haoyuan Hu** [2] **Enhong Chen** [3] **Yu Yang** [1]

Abstract

In this paper, we investigate the online allocation problem of maximizing the overall revenue subject to both lower and upper bound constraints. Compared to the extensively studied online problems with only resource upper bounds, the two-sided constraints affect the prospects of resource consumption more severely. As a result, only limited violations of constraints or pessimistic competitive bounds could be guaranteed. To tackle the challenge, we define a measure of feasibility ξ^* to evaluate the hardness of this problem, and estimate this measurement by an optimization routine with theoretical guarantees. We propose an online algorithm adopting a constructive framework, where we initialize a threshold price vector using the estimation, then dynamically update the price vector and use it for decision-making at each step. It can be shown that the proposed algorithm is $\big(1 - O(\frac{\varepsilon}{\xi^*-\varepsilon})\big)$ or $\big(1 - O(\frac{\varepsilon}{\xi^*-\sqrt{\varepsilon}})\big)$ competitive with high probability for ξ^* known or unknown respectively. To the best of our knowledge, this is the first result establishing a nearly optimal competitive algorithm for solving two-sided constrained online allocation problems with a high probability of feasibility.

1. Introduction

Online resource allocation is a prominent paradigm for sequential decision making during a finite horizon subject to the resource constraints, increasingly attracting the wide attention of researchers and practitioners in theoretical computer science (Mehta et al., 2007; Devanur & Jain, 2012; Devanur et al., 2019), operations research (Agrawal et al., 2014; Li & Ye, 2021) and machine learning communities (Balseiro et al., 2020; Li et al., 2020). In these settings, the requests arrive online and we need to serve each request via one of the available channels, which consumes a certain amount of resources and generates a corresponding service charge. The objective of the decision maker is to maximize the cumulative revenue subject to the resource capacity constraints. Such problem frequently appears in many applications including online advertising (Mehta et al., 2007; Buchbinder et al., 2007), online combinatorial auctions (Chawla et al., 2010), online linear programming (Agrawal et al., 2014; Buchbinder & Naor, 2009), online routing (Buchbinder & Naor, 2006), online multi-leg flight seats and hotel rooms allocation (Talluri et al., 2004), etc.

The aforementioned online resource allocation framework only considers the capacity (upper bound) constraints for resources. As a measure of fairness for resources expenditure, the requirements for guaranteeing a certain amount of resource allocation play important roles in real-world applications (Haitao Cui et al., 2007; Zhang et al., 2020) ranging from contractual obligations and group-level fairness to load balance. We give several examples in Appendix A for completeness. Recently, some attempts come out to alleviate the difficulties introduced by lower bound constraints. For instance, Lobos et al. (2021) propose an online mirror descent method to address this new online allocation problem with $O(\sqrt{T})$ asymptotic regret as well as $O(\sqrt{T})$ violation of lower bounds in expectation, where T is the number of requests. Meanwhile, Balseiro et al. (2021) consider a more general regularized setting to satisfy fairness requirements by non-separable penalties, which also leads to $O(\sqrt{T})$ asymptotic regret for the *regularized reward*. That is, Balseiro et al. (2021) only guarantee a gap of the order $O(\sqrt{T})$ for the sum of the regret of revenues and the violation of lower-bound constraints. All these studies consider the lower bound requirements as *soft* threshold, i.e., there is no guarantee on the satisfaction of lower bound constraints, which remains an open problem for online resource allocation.

In this paper, we investigate the online resource allocation problem under stochastic setting, where requests are drawn independently from some *unknown* distribution, and arrive

[*]Equal contribution [1]School of Data Science City University of Hong Kong, Kowloon Hong Kong, China [2] Cainiao Network, Hang Zhou, China [3]University of Science and Technology of China. Correspondence to: Yu Yang <yuyang@cityu.edu.hk>.

Proceedings of the 40^{th} International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023.

sequentially. The goal is to maximize the total revenue subject to the two-sided resource constraints. In the beginning, the decision maker is endowed with a limited and unreplenishable amount of resources, and agrees a minimum consumption on each resource. The decision maker can access the information of the current request and requests that have been processed before, but it is unable to get the information of future requests until their arrivals. Once observing an online request, the decision is made irrevocably to assign one available channel to serve the request, where we assume there is an oracle that determines the revenue and the amount of consumed resources depending on (request, channel) pair for convenience. A formal definition can be found in Section 3.

To address the challenges brought by two-sided resource constraints, we first define a problem dependent quantity γ to represent i). the maximum fraction of revenue or resource consumption per request, ii). the pessimism about the satisfaction of lower bound constraints. Generally, large γ will rule out good competitive ratio across different settings (Mehta et al., 2007; Buchbinder et al., 2007; Agrawal et al., 2014; Devanur et al., 2019). Meanwhile, we find that the margins between lower and upper bounds directly affect the hardness of the online allocation problem in stochastic setting. Inspired by the Slater's condition and its applications (Slater, 2014; Boyd et al., 2004), we define a measure of feasibility ξ^* to evaluate the hardness of constraints and present an estimator with sufficient accuracy. In the end, the proposed algorithm integrates several estimators to compensate undiscovered information. Our contribution can be summarized as follows:

1. Under the gradually improved assumptions, we propose three algorithms that return $1 - O(\frac{\varepsilon}{\xi^* - \varepsilon})$ competitive solutions *satisfying* the two-sided constraints w.h.p., if γ is at most $O\left(\frac{\varepsilon^2}{\ln(K/\varepsilon)}\right)$, where $\xi^* \gg \varepsilon$ is the measure of feasibility and K is the number of resources. To the best of our knowledge, this is the first result establishing a competitive algorithm for solving two-sided constrained online allocation problems *feasibly*, which is nearly optimal according to (Devanur et al., 2019).
2. To tackle the unknown parameter ξ^*, we propose an optimization routine in Algorithm 5. Through merging the estimate method into the previous framework, a new algorithm is proposed in Algorithm 6 with a solution achieving $1 - O(\frac{\varepsilon}{\xi^* - \sqrt{\varepsilon}})$ competitive ratio when ξ^* is an unknown and problem dependent constant.
3. Our analytical tools can be used to strengthen the existing models (Mehta et al., 2007; Devanur et al., 2019; Lobos et al., 2021) for online resource allocation problems.

2. Related Works

Online allocation problems have been extensively studied in theoretical computer science and operations research communities. In this section, we overview the related literature.

When the incoming requests are adversarially chosen, there is a stream of literature investigating online allocation problems. Mehta et al. (2007) and Buchbinder et al. (2007) first study the AdWords problem, a special case of online allocation, and provide an algorithm that obtains a $(1 - 1/e)$ approximation to the offline optimal allocation, which is optimal under the adversarial input model. However, the adversarial assumption may be too pessimistic about the requests. To consider another application scenarios, Devanur & Hayes (2009) propose the random permutation model, where an adversary first selects a sequence of requests which are then presented to the decision maker in random order. This model is more general than the stochastic i.i.d. setting in which requests are drawn independently and at random from an unknown distribution. In this new stochastic model, Devanur & Hayes (2009) revisit the AdWords problem and present a dual training algorithm with two phases: a training phase where data is used to estimate the dual variables by solving a linear program and an exploitation phase where actions are taken using the estimated dual variables. Their algorithm is guaranteed to obtain a $1 - o(1)$ competitive ratio, which is problem dependent. Feldman et al. (2010) show that this training-based algorithm could resolve more general linear online allocation problems. Pushing these ideas one step further, Agrawal et al. (2014) consider primal and dual algorithm that dynamically updates dual variables by periodically solving a linear program using the data collected so far. Meanwhile, Kesselheim et al. (2014) take the same policy and only consider renewing the primal variables. Recently, Devanur et al. (2019) take other innovative techniques to geometrically update the price vector via some decreasing potential function derived from probability inequalities. These algorithms also obtain $1 - o(1)$ approximation guarantees under some mild assumptions. While the algorithms described above usually require solving large linear problems periodically, there is a recent line of work seeking simple algorithms that does not need to solve a large linear programming. Balseiro et al. (2020) study a simple dual mirror descent algorithm for online allocation problems with concave reward functions and stochastic inputs, which attains $O(\sqrt{KT})$ regret, where K and T are the number of resources and requests respectively, i.e., updating dual variables via mirror descent algorithm and avoids solving large auxiliary linear programming. Simultaneously, Li et al. (2020) present a similar fast algorithm that updates the dual variable via projected gradient descent in every round for linear rewards. It is worth noting that all of these literature only consider the online allocation with capacity constraints.

3. Preliminaries and Assumptions

In this section, we introduce some essential concepts and assumptions that are adopted in our framework.

3.1. Two-sided Resource Allocation Framework

We consider the following framework for offline resource allocation problems. Let $\mathcal{K}$ be the set of K kinds of resources. There are J different types of requests. Each request $j \in \mathcal{J}$ ($|\mathcal{J}| = J$) could be served via some channel $i \in \mathcal{I}$, which will consume a_{ijk} amount of resource $k \in \mathcal{K}$ and generate w_{ij} amount of revenue. For each resource $k \in \mathcal{K}$, L_k and U_k denote the lower bound requirement and capacity, respectively. The objective of the two-sided resource allocation is to maximize the revenue subject to the two-sided resource constraints. The following is the offline integer linear programming for resource allocation where the entire sequence of T requests is given in advance:

$$\begin{aligned} W_R = \max_x & \sum_{i\in\mathcal{I},t\in[T]} w_{iR(t)}x_{it} \\ s.t.\ L_k \le & \sum_{i\in\mathcal{I},t\in[T]} a_{iR(t)k}x_{it} \le U_k, \forall k\in\mathcal{K} \\ & \sum_{i\in\mathcal{I}} x_{it} \le 1, \forall t \in [T] \\ & x_{it} \in \{0,1\}, \forall i\in\mathcal{I}, t\in[T] \end{aligned} \tag{1}$$

where $[T] = \{1, 2, \dots, T\}$ and the function $R(t) : [T] \to \mathcal{J}$ to record the types of requests, i.e., the type of t-th request is $R(t)$. As a common relaxation in online resource allocation literature (Agrawal et al., 2014; Devanur et al., 2019; Li et al., 2020; Balseiro et al., 2020), not picking any channel is permitted in ILP (1). We denote the no picking channel option by $\perp \in \mathcal{I}$, where $a_{\perp jk} = 0$ and $w_{\perp j} = 0$ for all $j \in \mathcal{J}$ and $k \in \mathcal{K}$.

3.2. Assumptions and Concepts

To facilitate the analysis in following sections, we hereby explain the detailed assumptions for two-sided online resource allocation problem.

In practice, it's impossible to know all incoming requests ahead of time. We therefore assume:

Assumption 1. The requests arrive sequentially and are independently drawn from some unknown distribution $\mathcal{P} : \mathcal{J} \to [0, 1]$, where $\mathcal{P}(j)$ denotes the arrival probability of request $j \in \mathcal{J}$ and we denote $p_j = \mathcal{P}(j), \forall j \in \mathcal{J}$.

To demonstrate the performance of our proposed algorithms, we also need to make some extra requirements about the parameters, which are widely adopted in literature.

Assumption 2. In stochastic settings, we make the following reasonable assumptions:

i. When the algorithm is initialized, we know the lower and upper bound requirements L_k and U_k regarding every resource $k \in \mathcal{K}$ and the number of requests T.
ii. Without loss of generality, the revenues w_{ij} and consumption of resources a_{ijk} are finite, non-negative and revealed when each request arrives, $\forall i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}$. Moreover, we know $\bar{w} = \sup_{i\in\mathcal{I},j\in\mathcal{J}} w_{ij}$ and $\bar{a}_k = \sup_{i\in\mathcal{I},j\in\mathcal{J}} a_{ijk}, \forall k \in \mathcal{K}$.

Under Assumption 1- 2, we denote the expected Linear Programming problem with lower bound $L_k + \beta T\bar{a}_k$ for every resource $k \in \mathcal{K}$ by $E(\beta)$, i.e.,

$$\begin{aligned} W_\beta = \max_x & \sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j w_{ij}x_{ij} \\ s.t.\ L_k + \beta T\bar{a}_k \le & \sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk}x_{ij} \le U_k, \forall k\in\mathcal{K} \\ & \sum_{i\in\mathcal{I}} x_{ij} \le 1, \forall j \in \mathcal{J} \\ & x_{ij} \ge 0, \forall i\in\mathcal{I}, j\in\mathcal{J} \end{aligned} \tag{2}$$

where β is a deviation parameter of lower bound constraints and W_β is the optimal value for problem $E(\beta)$. Meanwhile, we denote the optimal solution for $E(\beta)$ as $\{x(\beta)^*_{ij}, \forall i \in \mathcal{I}, j \in \mathcal{J}\}$. Under Assumption 1, if we take the same policy for every request $j \in \mathcal{J}$ in ILP (1), i.e., $x_{it} = x(\beta)^*_{ij}$ when $R(t) = j$, it can be verified that $\mathbb{E}\left[\sum_{i\in\mathcal{I},j\in[T]} a_{iR(t)k}x_{it}\right] = \sum_{i\in\mathcal{I},t\in\mathcal{J}} Tp_j a_{ijk} x(\beta)^*_{ij}$ and $\mathbb{E}\left[\sum_{i\in\mathcal{I},t\in[T]} w_{iR(t)}x_{it}\right] = \sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j w_{ij} x(\beta)^*_{ij}$. Therefore, when $\beta = 0$, we could view the LP (2) as a relaxed version of the expectation of ILP (1). Moreover, W_0 is an upper bound of the expectation of the optimal reward W_R in ILP (1).

Lemma 3.1. $W_0 \ge \mathbb{E}[W_R]$.

The proof of the above lemma is deferred to Appendix C. As a result, if an algorithm attains $(1 - o(1))W_0$, we can infer that this algorithm also achieves at least $(1 - o(1))\mathbb{E}[W_R]$.

We also make an assumption on the margin of resource constraints.

Assumption 3. (Strong feasible condition)

There exists a $\xi > 0$ making the linear constraints of the following problem feasible,

$$\begin{aligned} \xi^* = \max_{\xi\ge 0,x} & \ \xi \\ s.t.\ L_k + \xi T\bar{a}_k \le & \sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk}x_{ij} \le U_k, \forall k\in\mathcal{K} \\ & \sum_{i\in\mathcal{I}} x_{ij} \le 1, \forall j \in \mathcal{J} \\ & x_{ij} \ge 0, \forall i\in\mathcal{I}, j\in\mathcal{J}. \end{aligned} \tag{3}$$

The motivation for adding $\xi T\bar{a}_k$ to the lower bound is to *measure* the size of the original feasible space with lower bound L_k and upper bound U_k. Intuitively, it is relatively easy to satisfy the lower-bound requirements in the *wide* feasible space at the end of T-round decisions. Therefore, we call the optimal ξ^* the **measure of feasibility** and assume $\xi^* \gg \varepsilon$ for simplicity, where $\varepsilon > 0$ is a predefined error parameter.

Due to limited space, we present some frequently used concentration inequalities to predigest the theoretical analysis in Appendix B.

4. Competitive Algorithms for Online Resource Allocation with Two-sided Constraints

In this section, we propose a series of online algorithms for resource allocation with two-sided constraints by progressively weakening the following assumptions.

4.1 and 4.2. With known distribution;

4.3. With known optimal objective;

4.4. Completely unknown distribution.

4.1. With Known Distribution

In this section, we assume that we have the complete knowledge of the distribution $\mathcal{P}$. We first propose a high-level overview of our algorithm and outcomes as follows.

High-level Overview:

1. With the knowledge of $\mathcal{P}$, we could directly solve the expected problem $E(\tau)$ and obtain the optimal solution $\{x(\tau)^*_{ij}, \forall i \in \mathcal{I}, j \in \mathcal{J}\}$, where the deviation parameter $\tau = \frac{\varepsilon}{1-\varepsilon}$. For each fixed request $j \in \mathcal{J}$, if $x(\tau)^*_{i_1 j} \geq x(\tau)^*_{i_2 j}$, we tend to assign this type of request to channel i_1 rather than i_2. Motivated by this intuition, we design the Algorithm $\widetilde{P}$ in Algorithm 1, which assigns request $j \in \mathcal{J}$ to channel $i \in \mathcal{I}$ with probability $(1-\varepsilon)x(\tau)^*_{ij}$.
2. In the Algorithm $\widetilde{P}$, if we define the r.v. $X_k^{\widetilde{P}}$ for resource k consumed by one request sampled from distribution $\mathcal{P}$ and r.v. $Y^{\widetilde{P}}$ for the revenue, it is easy to obtain that $(1-\varepsilon)\frac{L_k}{T} + \varepsilon\bar{a}_k \leq \mathbb{E}(X_k^{\widetilde{P}}) \leq (1-\varepsilon)\frac{U_k}{T}$ and $\mathbb{E}[Y^{\widetilde{P}}] = (1-\varepsilon)\frac{W_\tau}{T}$. Because the expectation of resource consumption about one sample is restricted to the interval $[\frac{L_k}{T}, \frac{U_k}{T}]$, in Theorem 4.3 we could prove that the Algorithm $\widetilde{P}$ generates a feasible solution with high probability for online resource problem with two-sided constraints ILP (1) via the Bernstein inequalities in Lemma B.1. This is the reason for enlarging the lower bound L_k by $\tau T\bar{a}_k$ and scaling the solution with a factor $1-\varepsilon$. Meanwhile, we also verify that the accumulative revenue will be no less than $(1-2\varepsilon)W_\tau$ w.p. $1-\varepsilon$ in Lemma 4.1.
3. Since we have lifted the lower resource constraints from L_k to $L_k + \tau T\bar{a}_k$ in Algorithm $\widetilde{P}$, it would cause the change of baseline when analyzing the accumulative revenues. We have to derive the relationship between W_τ and W_0. By taking a sensitive analysis in Section 4.2, we obtain $W_\tau \geq \left(1 - \frac{\tau}{\xi^*}\right)W_0$ in Theorem 4.2 under the Assumption 3, where ξ^* is the measure of feasibility. Finally, we could prove the Algorithm $\widetilde{P}$ achieves an objective value at least $(1 - O(\epsilon))W_0$ in Theorem 4.3.

Algorithm 1 Algorithm $\widetilde{P}$

1: **Input:** $\tau = \frac{\varepsilon}{1-\varepsilon}, \mathcal{P}$
2: **Output:** $\{x_{ij}\}_{i\in\mathcal{I}, j\in[T]}$
3: $\{x(\tau)^*_{ij}\}_{i\in\mathcal{I}, j\in\mathcal{J}} = \arg\max_{\{\mathbf{x}\}} E(\tau)$.
4: When a type $j \in \mathcal{J}$ request comes, we assign this request to channel i with probability $(1-\varepsilon)x(\tau)^*_{ij}$. That is, If assigning the type j request to channel i, we set $x_{ij} = 1$, otherwise $x_{ij} = 0$.

Similar to (Devanur et al., 2019), we first consider the competitive ratio that Algorithm 1 could achieve for a surrogate LP problem $E(\tau)$ with optimal objective W_τ according to Definition (2).

Lemma 4.1. *Under Assumption 1-3, if* $\forall\varepsilon > 0$ *and* $\gamma = \max\left(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{T\bar{a}_k - L_k}, \frac{\bar{w}}{W_\tau}\right) = O(\frac{\varepsilon^2}{\ln(K/\varepsilon)})$, *Algorithm 1 achieves an objective value at least* $(1-2\varepsilon)W_\tau$ *and satisfies the constraints w.p.* $1-\varepsilon$.

The proof is deferred to Appendix E. From Lemma 4.1, we have verified the cumulative revenue is at least $(1-2\varepsilon)W_\tau$, w.p. $1-\varepsilon$. Thus, in order to compare the revenue with W_0, we should derive the relationship between W_τ and W_0. However, due to the deviation $\tau T\bar{a}_k$, it is hard to directly obtain this relationship. We will tackle this sensitive problem in the next subsection.

4.2. Sensitive Analysis

In this subsection, we demonstrate the relationship between W_τ and W_0. Before introducing the details, we first notice the difference between the problem $E(\tau)$ and the problem $E(0)$. Specifically, we enlarge the lower bound for every resource $k \in \mathcal{K}$ by an extra amount of $\tau T\bar{a}_k$.

Through a sensitive analysis, we finally found that the measure of feasibility ξ^* controls the decline ratio of the $E(\tau)$ objective value. The result can be summarized by follows.

Theorem 4.2. *Under the strong feasible condition in As-*

sumption 3, the optimal objective of $E(\tau)$ satisfies that

$$W_\tau \geq \left(1 - \frac{\tau}{\xi^*}\right) W_0,$$

where $\tau = \frac{\varepsilon}{1-\varepsilon}$.

We prove Theorem 4.2 from a geometrical perspective.

Proof. Let $x_{ij}^{'} = (1 - \frac{\tau}{\xi^*})x(0)_{ij}^* + \frac{\tau}{\xi^*}x(\xi^*)_{ij}^*$, $\forall i \in \mathcal{I}$ and $j \in \mathcal{J}$, then $x_{ij}^{'}$ is a convex combination between the optimal solutions of $E(0)$ and $E(\xi^*)$. According to the constraint set of problem $E(0)$ and $E(\xi)$, we can verify that $x_{ij}^{'}$ is non-negative and satisfies $\sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk} x_{ij} \leq U_k, \forall k \in \mathcal{K}, \sum_{i\in\mathcal{I}} x_{ij} \leq 1, \forall j \in \mathcal{J}$.

Meanwhile,

$$\begin{aligned}
&\sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk} x_{ij}^{'} \\
&\geq \sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk} \left((1 - \frac{\tau}{\xi^*})x(0)_{ij}^* + \frac{\tau}{\xi^*}x(\xi^*)_{ij}^*\right) \quad (4)\\
&\geq (1 - \frac{\tau}{\xi^*})L_k + \frac{\tau}{\xi^*}(L_k + \xi^* T\bar{a}_k) \\
&\geq L_k + \tau T \bar{a}_k, \forall k \in \mathcal{K}.
\end{aligned}$$

Thus $x_{ij}^{'}$ is feasible to $E(\tau)$, and

$$W_\tau \geq \sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j w_{ij} x_{ij}^{'} \geq (1 - \frac{\tau}{\xi^*})W_0.$$

This completes the proof of Theorem 4.2. □

In practical problems, the influence of ξ^* on competitive ratio may be far better than the worst case bound in Theorem 4.2, since constraints usually represent different resource requirements and only affect a part of requests. It is worth mentioning that our initial proof of Theorem 4.2 was based on analyzing a factor-revealing fractional linear programming problem motivated by (Jain et al., 2003). Although the above geometric proof of Theorem 4.2 is simpler, the factor-revealing prospective tells more relationship about the dual value $\sum_{k\in\mathcal{K}} \bar{a}_k \beta_k$, W_0 and ξ^*, which plays a key role in deriving Theorem 4.5 (See the end of Lemma G.1). We put this analysis in Appendix D due to the space limit and wish to provide more insights.

Combining Lemma 4.1 with Theorem 4.2, we can show that the cumulative revenue obtained by Algorithm $\widetilde{P}$ is larger than $(1 - 2\varepsilon)W_\tau \geq (1 - 2\varepsilon)(1 - \frac{\tau}{\xi^*})W_0 \geq \big(1 - (2 + \frac{1}{\xi^*})\varepsilon\big)W_0$. Therefore, we have the following theorem.

Theorem 4.3. *Under Assumption 1-3, if $\varepsilon > 0$, $\tau = \frac{\varepsilon}{1-\varepsilon}$ and $\gamma = O\big(\frac{\varepsilon^2}{\ln(K/\varepsilon)}\big)$, Algorithm 1 achieves an objective value of at least $\big(1 - (2 + \frac{1}{\xi^*})\varepsilon\big)W_0$ and satisfies the constraints w.p. $1 - \varepsilon$.*

Although $\widetilde{P}$ is an impractical algorithm owing to the complete knowledge of distribution $\mathcal{P}$, it builds a bridge to the desired competitive ratio. In the following subsections, we will release the unavailable knowledge by replacing $\widetilde{P}$ in a constructive way.

4.3. With Known W_τ

In this section, we only assume the knowledge of the optimal value W_τ and abandon the assumption of knowing the distribution $\mathcal{P}$. Based on this assumption, we design an algorithm A only using W_τ in Algorithm 2, which achieves at least $\big(1-(2+\frac{1}{\xi^*})\varepsilon\big)W_0$ objective and satisfies the two-sided constraints w.p. $1 - \varepsilon$.

Intuitively, we hope to construct a new algorithm A through Algorithm $\widetilde{P}$, which performs no worse than $\widetilde{P}$ and only need the knowledge of W_τ. We still use the r.v. X_{jk}^A for resource k consumed by the j-th request, which is determined by algorithm A, and r.v. Y_j^A for the revenue. In general, the key to our theoretical analysis is bounding the probabilities of three **bad** events: i) the violation of lower-bound constraints, such as, $\sum_{j=1}^T X_{jk}^A \leq L_k$; ii) the violation of capacity constraints, e.g., $\sum_{j=1}^T X_{jk}^A \geq U_k$; iii) the cumulative revenue is less than $(1 - O(\epsilon))W_0$, namely, $\sum_{j=1}^T Y_j^A \leq (1 - O(\epsilon))W_0$.

Note that, under the strong knowledge of the distribution $\mathcal{P}$, we have presented a simple yet powerful algorithm $\widetilde{P}$ (Algorithm 1) and also verified its effectiveness in the Theorem 4.3, i.e.,

$$\begin{aligned}
&P(\sum_{j=1}^T Y_j^{\widetilde{P}} \leq (1 - O(\epsilon))W_0) + \sum_{k=1}^K P(\sum_{j=1}^T X_{jk}^{\widetilde{P}} \notin [L_k, U_k]) \\
&\leq (2K + 1)\exp(-\frac{\epsilon^2}{2\gamma}) \leq \epsilon.
\end{aligned}$$

Therefore, to devise efficient algorithms with weaker assumptions, a natural idea is to construct a new algorithm A throughout Algorithm $\widetilde{P}$, which only needs the knowledge of W_τ and performs no worse than $\widetilde{P}$. Next, we demonstrate how to put this thought into effect. First, we consider a one-step hybrid algorithm $M\widetilde{P}^{T-1}$, which runs M for the first request and $\widetilde{P}$ for the rest $T - 1$ requests. Due to the theoretical result of $\widetilde{P}$ on T-round allocation problems, if selecting an appropriate M, we may expect an excellent performance of $M\widetilde{P}^{T-1}$. Then, after complicated computations (see Appendix F), the probability of all bad events of

$M\widetilde{P}^{T-1}$ is bounded by $\mathcal{F}(M\widetilde{P}^{T-1})$:

$$\mathcal{F}(M\widetilde{P}^{T-1}) = \Big(\mathbb{E}\Big[\sum_{k\in\mathcal{K}} \exp(\frac{-\ln(1-\varepsilon)}{\bar{a}_k}(X_{1k}^M - \frac{U_k}{T})) + \sum_{k\in\mathcal{K}} \exp(-\frac{\ln(1-\varepsilon)}{\bar{a}_k}(\frac{L_k}{T} - X_{1k}^M)) + \exp(\frac{-\ln(1-\varepsilon)}{\bar{w}}(\frac{(1-2\epsilon)W_\tau}{T} - Y_1^M))\Big]\Big)\exp(-\frac{T-1}{T}\frac{\varepsilon^2}{2\gamma}).$$

Notably, when $M=\widetilde{P}$, $\mathcal{F}(\widetilde{P}^T) \le (2K+1)\exp(-\frac{\epsilon^2}{2\gamma}) \le \epsilon$ (see Appendix E). As a result, we can conclude that the best choice $\widetilde{M}$ satisfies $\mathcal{F}(\widetilde{M}\widetilde{P}^{T-1}) = \min_M \mathcal{F}(M\widetilde{P}^{T-1}) \le \mathcal{F}(\widetilde{P}^T) \le \epsilon$. Since we have finite J different types of requests, we also can show

$$\begin{aligned}&\min_M \mathcal{F}(M\widetilde{P}^{T-1})\\ &= \Big(\mathbb{E}\Big[\min_{i\in\mathcal{I}}\Big(\sum_{k\in\mathcal{K}} \exp(\frac{-\ln(1-\varepsilon)}{\bar{a}_k}(a_{i1k} - \frac{U_k}{T})) + \sum_{k\in\mathcal{K}} \exp(-\frac{\ln(1-\varepsilon)}{\bar{a}_k}(\frac{L_k}{T} - a_{i1k})) + \exp(\frac{-\ln(1-\varepsilon)}{\bar{w}}(\frac{(1-2\epsilon)W_\tau}{T} - w_{i1}))\Big)\Big]\Big)\exp(-\frac{T-1}{T}\frac{\varepsilon^2}{2\gamma}).\end{aligned} \tag{5}$$

From line 6 in Algorithm A (Algorithm 2) and Equation (5), we could check that $\widetilde{M}$ is equal to the decision of A for the first request, so $\mathcal{F}(A\widetilde{P}^{T-1}) = \mathcal{F}(\widetilde{M}\widetilde{P}^{T-1}) \le \mathcal{F}(\widetilde{P}^T) \le \epsilon$. Therefore, we ensure that Algorithm $A\widetilde{P}^{T-1}$ achieves an objective value at least $(1-O(\epsilon))W_0$ and satisfies the constraints w.p. $1-\epsilon$.

Like the previous policy, based on $A\widetilde{P}^{T-1}$, we also construct another two-step hybrid algorithm $AM\widetilde{P}^{T-2}$ which runs A for the first request, M for the second one and $\widetilde{P}$ for the rest. Similarly, the probability of its bad events is less than some $\mathcal{F}(AM\widetilde{P}^{T-2})$. Moreover, we know the best $\widetilde{M_1}$ satisfying $\mathcal{F}(A\widetilde{M_1}\widetilde{P}^{T-2}) \le \mathcal{F}(A\widetilde{P}^{T-1}) \le \epsilon$ and $\widetilde{M_1}$ is the same with the second iteration of A. Continuing these iterations, we finally get a T-step hybrid method A^T satisfying the constraints and attaining an objective value at least $(1-O(\epsilon))W_0$ w.p. $1-\epsilon$. This A^T is in accord with all iterations of our Algorithm 2.

Theorem 4.4. *Under Assumption 1-3, if $\varepsilon > 0$, τ and γ are defined as Theorem 4.3, the Algorithm 2 achieves an objective value at least $\big(1-(2+\frac{1}{\xi^*})\varepsilon\big)W_0$ and satisfies the constraints w.p. $1-\varepsilon$.*

The proof is deferred to Appendix F.

4.4. With Known ξ^* but Unknown Distribution

In this section, we consider the completely unknown distribution setting. Under this setting, we first divide the incoming T requests into multiple stages and run an inner

Algorithm 2 Algorithm A

1: **Input:** ε, W_τ
2: **Output:** $\{x_{ij}\}_{i\in\mathcal{K},j\in[T]}$
3: Set $c_{1k} = \frac{-\ln(1-\varepsilon)}{\bar{a}_k}, \forall k\in\mathcal{K}$ and $c_2 = \frac{-\ln(1-\varepsilon)}{\bar{w}}$
4: Initiate $\phi_k^0 = 1, \varphi_k^0 = 1, \forall k\in\mathcal{K}$, and $\psi^0 = 1$
5: **for** $j = 1,\ldots,T$ **do**
6: compute the optimal i^* by
$$i^* = \arg\min_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} \phi_k^{j-1}\exp\Big(c_{1k}\big(a_{ijk} - \frac{U_k}{T}\big)\Big) + \sum_{k\in\mathcal{K}} \varphi_k^{j-1}\exp\Big(c_{1k}\big(\frac{L_k}{T} - a_{ijk}\big)\Big) + \psi^{j-1}\exp\Big(c_2\big(\frac{(1-2\varepsilon)W_\tau}{T} - w_{ij}\big)\Big)$$
7: Set $X_{jk}^A = a_{i^*jk}$, $Y_j^A = w_{i^*j}$
8: Update $\phi_k^j = \phi_k^{j-1}\exp\big(c_{1k}(X_{jk}^A - \frac{U_k}{T})\big), \forall k\in\mathcal{K}$
9: Update $\varphi_k^j = \varphi_k^{j-1}\exp\big(c_{1k}(\frac{L_k}{T} - X_{jk}^A)\big), \forall k\in\mathcal{K}$
10: Update $\psi^j = \psi^{j-1}\exp\big(c_2(\frac{(1-2\varepsilon)W_\tau}{T} - Y_j^A)\big)$
11: **end for**

loop similar to Algorithm 2 with the estimate of the optimal objective value in each stage. More specifically, the differences between the inner loop and Algorithm 2 are three-fold: i). choose the deviation parameter $\beta = \varepsilon$ instead of τ, where ε is an error parameter defined in Theorem 4.5; ii). pessimistically reduce the estimated objective value from W^{r-1} to Z^r in each stage for boosting the chance of success, where W^{r-1} is estimated by Algorithm 3; iii). consider the relative error of objective and the maximum relative error of constraints as $\varepsilon_{y,r}$ and $\varepsilon_{x,r}$ separately in each stage, instead of using the global error parameter ε. We consider that $\epsilon_{y,r}$ describes the error we care about the most (objective) and $\epsilon_{x,r}$ describes the rest errors (constraints and objective estimates). Distinguishing these error terms allows for more refined results, but will lead to the competitive ratio of the same order in our settings. In Algorithm 4, we name the proposed algorithm by A_1 to facilitate the analysis.

The high-level ideas can be summarized as follows. There are three types of errors that have to be restrained, 1). the error $|Z_r - W_\epsilon|$ from estimating deviant objective value, 2). the error $\epsilon_{x,r}$ from guaranteeing the satisfaction of constraints, and 3). the error $\epsilon_{y,r}$ from Algorithm A_1 affected by the randomness in the objective. We use an iterative learn-and-predict strategy to periodically reduce relative errors caused by insufficient samples. Let $\widehat{W}^r$ be the objective obtained by Algorithm A_1 in stage r. According to the previous sections, if objective Z_r is reachable, we could prove that the loss $\ell_r := |\widehat{W}^r - Z^r|$ is upper bounded by $O(\frac{t_r}{T}\epsilon_{y,r}Z_r)$ in each stage. We also need Z_r to be a good estimate for the objective. $Z^r \ge (1-O(\epsilon_{x,r-1}))W_\epsilon$ en-

Algorithm 3 Objective_Estimator($\{\text{request}_j\}_r, t_r, \beta$)

1: **Input:** requests from r-th stage $\{\text{request}_j\}_r$, request number t_r, deviation parameter β.
2: **Output:** W^r
3: Solve $E(\beta)$ as

$$\begin{aligned} W^r = & \max_x \sum_{i\in\mathcal{I}}\sum_{j=1}^{t_r} w_{ij}x_{ij} \\ s.t.\ & \frac{t_r}{T}(L_k+\beta T\bar{a}_k) \le \sum_{i\in\mathcal{I}}\sum_{j=1}^{t_r} a_{ijk}x_{ij} \\ & \le \frac{t_r}{T}U_k, \forall k\in\mathcal{K} \\ & \sum_{i\in\mathcal{I}} x_{ij} \le 1, \forall j\in[t_r] \\ & x_{ij}\ge 0, \forall i\in\mathcal{I}, j\in[t_r] \end{aligned} \tag{6}$$

sures that the objective is not underestimated, and $Z_r \le W_\epsilon$ ensures that it is a reachable target and facilitates the proof for $r = 0, \ldots, l-1$. Besides the normal stages, a warm-up stage $r = -1$ is added to provide an estimate Z_{-1}. Then the cumulative loss $\ell_{-1} + \sum_{r=0}^{l-1} \ell_r$ will be no greater than $O\big(\frac{t_{-1}}{T}W_0 + \sum_r \frac{t_r(\epsilon_{x,r-1}+\epsilon_{y,r})}{T}W_\epsilon\big)$. The algorithm is designed to balance the relative error $\epsilon_{x,r}$, $\epsilon_{y,r}$ and their impact on the requests in each stage.

More precisely, Algorithm A_1 geometrically divides T requests into $l+1$ stages, where the number of requests are $t_r = \varepsilon 2^r T$ for $r = 0, \ldots, l-1$ and $t_{-1} = \varepsilon T$. In the initial stage $r = -1$, we use the first $t_r = \varepsilon T$ requests to estimate W_ε and obtain W^{-1}, assuming that none of the requests are served in worst case. In stage $r \in \{0, 1, \ldots, l-1\}$, the requests from stage $r-1$ are used to provide more and more accurate estimate W^{r-1} of W_ε. By Union Bound, we set the failure probability as $\delta = \frac{\varepsilon}{3l}$ and reduce the estimate W^{r-1} to $Z^r = \frac{TW^{r-1}}{t_r(1+O(\varepsilon_{x,r-1}))}$, which is promised between $[(1-O(\varepsilon_{x,r-1}))W_\varepsilon, W_\varepsilon]$ w.p. $1-2\delta$.

Then in each stage r, define the error parameters for objective and constraints as $\varepsilon_{y,r} = O\left(\sqrt{\frac{T\ln(K/\delta)}{t_r Z_r}}\right)$ and $\varepsilon_{x,r} = O\left(\sqrt{\frac{\gamma_1 T\ln(K/\delta)}{t_r}}\right)$ respectively. We construct a surrogate Algorithm $\widetilde{P}_2$ that achieves $\frac{t_r}{T}(1-\varepsilon_{y,r})Z^r$ cumulative revenue with the consumption of every resource k between $\left[\frac{t_r(1+\varepsilon_{x,r})}{T}\big(L_k + (\varepsilon - \frac{\varepsilon_{x,r}}{1+\varepsilon_{x,r}})T\bar{a}_k\big), \frac{t_r(1+\varepsilon_{x,r})}{T}U_k\right]$ with probability at least $1-\delta$. Similar to previous sections, the connection between the inner loop of Algorithm A_1 and Algorithm $\widetilde{P}_2$ is built to minimize the upper bound of failure probabilities. Finally, with probability at least $1-\delta$, the cumulative revenue is at least $\sum_{r=0}^{l-1}\frac{t_r Z_r}{T}(1-\varepsilon_{y,r})$ and the cumulative consumed resource of $k\in\mathcal{K}$ is between $\sum_{r=0}^{l-1}\frac{t_r(1+\varepsilon_{x,r})}{T}(L_k + (\varepsilon - \frac{\varepsilon_{x,r}}{1+\varepsilon_{x,r}})T\bar{a}_k)$ and $\sum_{r=0}^{l-1}\frac{t_r(1+\varepsilon_{x,r})}{T}U_k$. Letting $\gamma_1 = O(\frac{\varepsilon^2}{\ln(K/\varepsilon)})$, we could keep $\sum_{r=0}^{l-1}\frac{t_r}{T}(1+\varepsilon_{x,r})U_k \le U_k$,$\sum_{r=0}^{l-1}\frac{t_r(1+\varepsilon_{x,r})}{T}(L_k + (\varepsilon - \frac{\varepsilon_{x,r}}{1+\varepsilon_{x,r}})T\bar{a}_k) \ge L_k$ and $\sum_{r=0}^{l-1}\frac{t_r Z_r}{T}(1-\varepsilon_{y,r}) \ge \big(1-O(\frac{\varepsilon}{\xi^*-\varepsilon})\big)W_0$.

The most tricky part of Algorithm A_1 is estimating W^r in Algorithm 3. Since the lower bound L_k cannot upper bound the mean of X_{jk} as U_k does, it brings a great challenge to the theoretical analysis. To address this issue, we solve a biased LP problem, uplifting L_k by $\epsilon T\bar{a}_k$, at the end of each stage. It helps the satisfaction of the lower bound in the next stage, and the influence on the objective can be determined by Theorem 4.2. This completes the high-level overview of the analysis of Algorithm A_1. The theoretical result of Algorithm A_1 is presented in Theorem 4.5 with explicit requirements of parameters.

Theorem 4.5. *Under Assumption 1-3, if* $\varepsilon > 0$ $\tau_1 = \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}$ *such that* $\tau_1 + \varepsilon \le \xi^*$ *and* $\gamma_1 = \max\big(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{(1-\varepsilon)T\bar{a}_k - L_k}, \frac{\bar{w}}{W_{\varepsilon+\tau_1}}\big) = O\big(\frac{\varepsilon^2}{\ln(K/\varepsilon)}\big)$, *Algorithm* A_1 *defined in Algorithm 4 achieves an objective value of at least* $\big(1-O(\frac{\varepsilon}{\xi^*-\varepsilon})\big)W_0$ *and satisfies the constraints w.p.* $1-\varepsilon$.

The proof of the theorem is deferred to Appendix G.

Remark:

1. It can be shown that the problem dependent parameter ξ^* restricts the capacity of Algorithm A_1 by affecting ε, τ_1, and γ_1. This is consistent with our intuition that the problem with two-sided constraints becomes harder if ξ^* decreases.
2. It is worth noting that the competitive ratio obtained in this paper reflects the high-probability performance of the proposed algorithms under a finite T, so it is difficult to compare with the regret bound for dual-mirror-descent methods rigorously due to the different settings. But if T is given and the linear growth of lower and upper bound is assumed, i.e. L_k and U_k are both $O(T)$, ξ^* turns to be a problem-dependent constant, which leads the proposed algorithm towards an $\widetilde{O}(\sqrt{T\ln(KT)})$ regret w.h.p., where we hide the potential $\log\log$-term in $\widetilde{O}(\cdot)$. In real-world applications with massive constraints, such as guaranteed advertising delivery (Zhang et al., 2020) with hundreds of thousands of ad providers, the proposed algorithm could exceed existing $O(\sqrt{KT})$ averaged regrets (Balseiro et al., 2020; 2021; Lobos et al., 2021) with respect to the number of constraints K. More comparisons can also be found in (Devanur et al., 2019).
3. Although an LP is solved at the beginning of each

Algorithm 4 Algorithm A_1

Input: $\varepsilon, \gamma_1, \xi^*$
Output: $\{x_{ij}\}_{i\in\mathcal{K},j\in[T]}$
Set $l = \log_2(\frac{1}{\varepsilon})$, $t_r = \varepsilon 2^r T$, $t_{-1} = \varepsilon T$ and $\delta = \frac{\varepsilon}{3l}$
for $r = 0$ to l-1 **do**
  Set W^{r-1} =Objective_Estimator($\{\text{request}_j\}_{r-1}, t_{r-1}, \varepsilon$)
  Set $Z_r = \frac{TW^{r-1}}{(1+(2+\frac{1}{\xi^*-\varepsilon})\varepsilon_{x,r-1})t_{r-1}}$
  Set $\varepsilon_{x,r} = \sqrt{\frac{4T\gamma_1\ln(\frac{2K+1}{\delta})}{t_r}}$, $\varepsilon_{y,r} = \sqrt{\frac{4T\ln(\frac{2K+1}{\delta})\bar{w}}{Z_r t_r}}$
  Set $c_{1k,r} = \frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}$ and $c_{2,r} = \frac{\ln(1+\varepsilon_{y,r})}{\bar{w}}$
  Initialize $\phi_k^0 = \exp\left(\frac{-(t_r-1)\varepsilon_{x,r}^2}{4\gamma_1 T}\right)$,
  $\varphi_k^0 = \exp\left(\frac{-(t_r-1)\varepsilon_{x,r}^2}{4\gamma_1 T}\right)$, $\psi^0 = \exp\left(\frac{-(t_r-1)\varepsilon_{y,r}^2 Z^r}{4\bar{w}T}\right)$
  for $j = 1, \ldots, t_r$ **do**
    compute the optimal i^* by

$$i^* = \arg\min_{i\in\mathcal{I}} \left\{ \sum_{k\in\mathcal{K}} \phi_k^{j-1} \exp\left(c_{1k,r}\left(-\frac{(1+\varepsilon_{x,r})U_k}{T} + a_{ijk}\right)\right) + \sum_{k\in\mathcal{K}} \varphi_k^{j-1} \exp\left(c_{1k,r}\left(\bar{a}_k - a_{ijk} - \frac{(1+\varepsilon_{x,r})((1-\varepsilon)T\bar{a}_k - L_k)}{T}\right)\right) + \psi^{j-1} \exp\left(c_{2,r}\left(\frac{(1-\varepsilon_{y,r})Z^r}{T} - w_{ij}\right)\right) \right\}$$

    Set $X_{jk}^{A_1} = a_{i^*jk}$, $Y_j^{A_1} = w_{i^*j}$, $Z_{jk}^{A_1} = \bar{a}_k - a_{i^*jk}$
    Update
    $\phi_k^j = \phi_k^{j-1} \exp\left(c_{1k,r}(X_{tk}^{A_1} - \frac{(1+\varepsilon_{x,r})U_k}{T}) + \frac{\varepsilon_{x,r}^2}{4T\gamma_1}\right)$,
    $\varphi_k^j = \varphi_k^{j-1} \exp\left(-c_{1k,r}(\frac{(1+\varepsilon_{x,r})((1-\varepsilon)T\bar{a}_k - L_k)}{T} - Z_{jk}^{A_1}) + \frac{\varepsilon_{x,r}^2}{4T\gamma_1}\right)$,
    $\psi^j = \psi^{j-1} \exp\left(c_{2,r}(\frac{(1-\varepsilon_{y,r})Z_r}{T} - Y_j^A) + \frac{\varepsilon_{y,r}^2 Z_r}{4T\bar{w}}\right)$
  end for
end for

stage, according to the proof of Theorem G.2, we can use the revenue obtained from previous stage as a good approximation of W^{r-1} before Algorithm 3 returning it in practice for $r = 1, \ldots, l-1$. Thus, the online property of Algorithm 4 will not be hurt severely.

However, sometimes ξ^* is inaccessible in practice when Algorithm 4 is initialized. We will address the unknown measurement ξ^* in the next section.

5. Exploring the Measure of Feasibility

As shown in previous sections, the constant ξ^* measures the feasibility of the original problem, and plays a central role in Algorithm A_1. ξ^* can hardly be obtained in practice. Motivated by Algorithm 3, we propose Algorithm 5 for estimating ξ^* to provide a more complete analysis. In this section, we investigate the strong feasible condition, i.e., Assumption 3. Recalling the definition of $E(\beta)$ in Section 3, we can conclude the following results.

Algorithm 5 Feas_Estimator($\{\text{request}_j\}_r,\ t_r,\ \gamma_2,\ \delta$)

Input: requests from r-th stage $\{\text{request}_j\}_r$, request number t_r, relative error $\varepsilon_{x,r}$, problem dependent quantity γ_2, failure probability δ.
Output: $\widehat{\xi} = \xi_{\max} - 2\varepsilon_{x,r}$
Compute $\varepsilon_{x,r} = \sqrt{\frac{4\gamma_2 T\ln(\frac{K}{\delta})}{t_r}}$
Set

$$\begin{aligned} \xi_{\max} = &\max_{\xi,x}\ \xi \\ s.t.\ &\frac{t_r}{T}(L_k + \xi T\bar{a}_k) \le \sum_{i\in\mathcal{I}} \sum_{j=t_r+1}^{t_{r+1}} a_{ijk}x_{ij} \le \frac{t_r}{T}U_k,\ \forall k\in\mathcal{K} \\ &\sum_{i\in\mathcal{I}} x_{ij} \le 1, \forall j = t_r+1, \ldots, t_{r+1} \\ &x_{ij} \ge 0,\ \forall i\in\mathcal{I},\ j = t_r+1, \ldots, t_{r+1} \end{aligned} \tag{7}$$

Algorithm 6 Algorithm A_2

Input: $\varepsilon, L_k, U_k, \gamma_1, \gamma_2 = O\big(\frac{\epsilon^2}{\ln(K/\varepsilon)}\big), \bar{a}_k$
Output: $\{x_{ij}\}_{i\in\mathcal{K},j\in[T]}$
Set $l = \log_2(\frac{1}{\varepsilon})$, $t_r = \varepsilon 2^r T$, $t_{-1} = \varepsilon T$ and $\delta = \frac{\varepsilon}{3l+2}$
Set $\widehat{\xi_0}$=Feas_Estimator($\{\text{request}_j\}_{-1},\ t_{-1},\ \gamma_2,\ \delta$)
Execute the Algorithm $A_1(\epsilon,\ \gamma_1,\ \widehat{\xi_0})$

Proposition 5.1. *According to the definition of ξ^* in LP* (3)*, we have*

1. *When $\xi^* \ge 0$, the problem $E(0)$ is feasible. Otherwise, it is infeasible.*
2. *When $\xi^* > 0$ and $0 < \xi \le \xi^*$, the problem $E(\xi)$ is feasible.*

We omit the proof of the proposition for simplicity. From Proposition 5.1, with the knowledge of ξ^*, we could easily check the feasibility of $E(0)$. Next, we show how to estimate the ξ^* from served requests in Algorithm 5.

Theorem 5.2. *Under Assumption 1-3, if $\gamma_2 = \max(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{T\bar{a}_k - L_k}) = O(\frac{\epsilon^2}{\ln(K/\varepsilon)})$, Algorithm 5 with t_r i.i.d. requests outputs $\widehat{\xi}$ such that*

$$\widehat{\xi} \in [\xi^* - 4\epsilon_{x,r}, \xi^*]$$

w.p. $1 - 2\delta$, where $\epsilon_{x,r} = \sqrt{\frac{4\gamma_2 T\ln(K/\delta)}{t_r}}$.

The proof is deferred to Appendix H.

Now $\widehat{\xi_r}$ can be viewed as a good estimate for ξ^* from stage 0. Based on the estimation of ξ^*, we propose an algorithm A_2 in Algorithm 6. Like the previous methods, we geometrically divide T requests into $l+1$ stages for $r = -1, 0, \ldots, l-1$, where the initial stage $r = -1$ and the first stage $r = 0$ have εT requests. Besides the estimate W^{-1}, the estimate $\widehat{\xi_0}$ for ξ^* is obtained by Algorithm 5 at the end of the initial stage $r = -1$. Then we consider the new expected problem $E\left(\widehat{\xi}_0\right)$ and run Algorithm A_1 defined in Algorithm 4 for the rest requests.

Theorem 5.3. *Under Assumption 1-3, if* $\varepsilon > 0$, $\tau_1 = \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}$ *such that* $\tau_1 + 4\sqrt{\varepsilon} + \varepsilon \leq \xi^*$ *and* $\gamma_1 = \max\left(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{(1-\epsilon)T\bar{a}_k - L_k}, \frac{\bar{w}}{W_{\epsilon+\tau_1}}\right) = O\left(\frac{\varepsilon^2}{\ln(K/\varepsilon)}\right)$, *Algorithm 6 achieves an objective value of at least* $\left(1 - O\left(\frac{\varepsilon}{\xi^* - 4\sqrt{\varepsilon} - \epsilon}\right)\right)W_0$ *and satisfies the constraints, w.p.* $1 - \varepsilon$.

The proof can be found in Appendix I.

Until now, we have dropped the knowledge of the entire distribution $\mathcal{P}$, the objective value W_τ for problem $E(\tau)$ and the strong feasible constant ξ^* step by step. In practice, the only hyperparameter ϵ can be well approximated by solving a polynomial approximation of the transcendental equation between γ_1 and ϵ in Theorem 5.3. We regard the proposed Algorithm 6 as a practical algorithm for two-sided constrained online resource allocation problems.

6. Conclusion

In this paper, we have developed a method for online allocation problems with two-sided resource constraints, which has a wide range of real-world applications. By designing a factor-revealing linear fractional programming, a measure of feasibility ξ^* is defined to facilitate our theoretical analysis. We prove that Algorithm 4 holds a nearly optimal competitive ratio if the measurement is known and large enough compared with the error parameter, i.e. $\xi^* \gg \varepsilon$. An estimator is also presented in the paper for the unknown ξ^* scenario. We will investigate more efficient extensions of this work in the near future.

Acknowledgements

Zhang and Yang's research is supported in part by the Hong Kong Research Grants Council (ECS 21214720), City University of Hong Kong (Project 9610465), Hong Kong Institute for Data Science, and Alibaba Group through Alibaba Innovative Research (AIR) Program. Finally, we are also grateful to helpful comments from Qiang Ha, Tao Li, Zengde Deng, Xiaojun Tan and Jiayuan Han.

References

Agrawal, S., Wang, Z., and Ye, Y. A dynamic near-optimal algorithm for online linear programming. *Operations Research*, 62(4):876–890, 2014.

Balseiro, S., Lu, H., and Mirrokni, V. Dual mirror descent for online allocation problems. In *International Conference on Machine Learning*, pp. 613–628. PMLR, 2020.

Balseiro, S., Lu, H., and Mirrokni, V. Regularized online allocation problems: Fairness and beyond. In *International Conference on Machine Learning*, pp. 630–639. PMLR, 2021.

Bernstein, S. *The theory of probabilities*. Gastehizdat Publishing House, Moscow, 1946.

Boyd, S., Boyd, S. P., and Vandenberghe, L. *Convex optimization*. Cambridge university press, 2004.

Buchbinder, N. and Naor, J. Improved bounds for online routing and packing via a primal-dual approach. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pp. 293–304. IEEE, 2006.

Buchbinder, N. and Naor, J. Online primal-dual algorithms for covering and packing. *Mathematics of Operations Research*, 34(2):270–286, 2009.

Buchbinder, N., Jain, K., and Naor, J. S. Online primal-dual algorithms for maximizing ad-auctions revenue. In *European Symposium on Algorithms*, pp. 253–264. Springer, 2007.

Chawla, S., Hartline, J. D., Malec, D. L., and Sivan, B. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pp. 311–320, 2010.

Devanur, N. R. and Hayes, T. P. The adwords problem: online keyword matching with budgeted bidders under random permutations. In *Proceedings of the 10th ACM conference on Electronic commerce*, pp. 71–78, 2009.

Devanur, N. R. and Jain, K. Online matching with concave returns. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pp. 137–144, 2012.

Devanur, N. R., Jain, K., Sivan, B., and Wilkens, C. A. Near optimal online algorithms and fast approximation algorithms for resource allocation problems. *Journal of the ACM (JACM)*, 66(1):1–41, 2019.

Feldman, J., Henzinger, M., Korula, N., Mirrokni, V. S., and Stein, C. Online stochastic packing applied to display ad allocation. In *ESA (1)*, pp. 182–194, 2010.

Haitao Cui, T., Raju, J. S., and Zhang, Z. J. Fairness and channel coordination. *Management science*, 53(8):1303–1314, 2007.

Jain, K., Mahdian, M., Markakis, E., Saberi, A., and Vazirani, V. V. Greedy facility location algorithms analyzed using dual fitting with factor-revealing lp. *Journal of the ACM (JACM)*, 50(6):795–824, 2003.

Kesselheim, T., Tönnis, A., Radke, K., and Vöcking, B. Primal beats dual on online packing lps in the random-order model. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pp. 303–312, 2014.

Li, X. and Ye, Y. Online linear programming: Dual convergence, new algorithms, and regret bounds. *Operations Research*, 2021.

Li, X., Sun, C., and Ye, Y. Simple and fast algorithm for binary integer and online linear programming. In *NeurIPS*, 2020.

Lobos, A., Grigas, P., and Wen, Z. Joint online learning and decision-making via dual mirror descent. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139, pp. 7080–7089. PMLR, 2021.

Mehta, A., Saberi, A., Vazirani, U., and Vazirani, V. Adwords and generalized online matching. *Journal of the ACM (JACM)*, 54(5):22–es, 2007.

Slater, M. Lagrange multipliers revisited. In *Traces and emergence of nonlinear programming*, pp. 293–306. Springer, 2014.

Talluri, K. T., Van Ryzin, G., and Van Ryzin, G. *The theory and practice of revenue management*, volume 1. Springer, 2004.

Zhang, H., Zhang, L., Xu, L., Ma, X., Wu, Z., Tang, C., Xu, W., and Yang, Y. A request-level guaranteed delivery advertising planning: Forecasting and allocation. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2980–2988, 2020.

A. The Examples of Resource Lower Bounds

In this section, we give three practical instances showing the necessity of lower bound constraints in real-world applications.

1. **Guaranteed Advertising Delivery**: In the online advertising scenario, the advertising publishers will sell the ad impressions in advance with the promise to provide each advertiser an agree-on number of target impressions over a fixed future time period, which is usually written in the contracts. Furthermore, the advertising platform considers other constraints, such as advertisers' budgets and impressions inventories, and simultaneously maximizes multiple accumulative objectives regarding different interested parties, e.g. Gross Merchandise Volume (GMV) for ad providers and publisher's revenue. This widely used guaranteed delivery advertising model is generally formulated as an online resource allocation problem with two-sided constraints (Zhang et al., 2020).

2.1. **Fair Channel Constraints**: We consider the online orders assignment in an e-commerce platform, where the platform allocates the orders (or called packages in logistics) to different warehouses providers. Due to the governmental regulations and contracts between platforms and providers, the platform usually tends to set lower bounds to the daily accepted orders for special channels, such as new/small-scale providers or those in developing areas. Besides, it can be shown that when providers are concerned about fairness, the platform can use a simple wholesale price above its marginal cost to coordinate this channel in terms of both achieving the maximum channel profit and attaining the maximum channel utility (Haitao Cui et al., 2007).

2.2. **Timeliness Achievement Constraints**: The delivery time for orders is highly related with the customers' shopping experience. Thus, the online shopping platforms also take the time-effectiveness of parcel shipment into account. For every parcel, platforms usually use the timeline achievement rate $r_{uov} \in [0, 1]$ to denote the probability of arriving at the destination in required u days if we assign the order of type o to one channel v, which could be estimated by the historical data and features of order and channel/logistics providers. As we know, long delivery time will impair the consumer's shopping experience, but reducing delivery time means increasing cost. In order to balance the customers' shopping experience and transportation costs, platforms always set a predefined lower threshold to the average timeline achievement rate of daily orders, which can be modeled as lower bound constraints in order assignment task.

B. Useful Concentrations

In this section, we present the classical concentration inequalities for completeness.

Lemma B.1. *(Bernstein, 1946)*

i. Suppose that $|X| \leq c$ and $\mathbb{E}[X] = 0$. For any $t > 0$,

$$\mathbb{E}\big[\exp(tX)\big] \leq \exp\left(\frac{\sigma^2}{c^2}\big(e^{tc} - 1 - ct\big)\right)$$

where $\sigma^2 = Var(X)$.

ii. If $X_1, X_2, \ldots, X_n$ are independent r.v., $\mathbb{E}[X_i] = \mu$ and $\mathrm{P}(|X_i - \mu| \leq c) = 1$,$\forall i = 1, \ldots, n$, then $\forall \varepsilon > 0$ following inequality holds

$$\mathrm{P}\left(\left|\frac{\sum_{i=1}^n X_i}{n} - \mu\right| \geq \varepsilon\right) \leq 2\exp\left(-\frac{n\varepsilon^2}{2\sigma^2 + \frac{2c\varepsilon}{3}}\right)$$

where $\sigma^2 = \frac{\sum_{i=1}^n Var(X_i)}{n}$.

The first result, Lemma B.1.i., is a well-known intermediate result of Bennett's inequality. We go a few steps further to construct the algorithm.

C. Proof for Lemma 3.1

Lemma 3.1. $W_0 \geq \mathbb{E}\left[W_R\right]$.

Proof. : We consider two linear programming problem, the linear relaxation of sampled integer linear programming (1), i.e.

$$\begin{aligned} \max_{x} \quad & \sum_{i\in\mathcal{I},j\in[T]} w_{ij}x_{ij} \\ s.t.\ L_k \leq & \sum_{i\in\mathcal{I},j\in[T]} a_{ijk}x_{ij} \leq U_k, \forall k\in\mathcal{K} \\ & \sum_{i\in\mathcal{I}} x_{ij} \leq 1, \forall j\in[T] \\ & x_{ij}\geq 0, \forall i\in\mathcal{I}, j\in[T] \end{aligned} \tag{S.I I}$$

and the linear programming that considers the samples from the type of requests perspective, i.e.

$$\begin{aligned} \max_{x} \quad & \sum_{i\in\mathcal{I},j\in\mathcal{J}} \big|\{j'|j'\in[T], j'=j\}\big| w_{ij}x_{ij} \\ s.t.\ L_k \leq & \sum_{i\in\mathcal{I},j\in\mathcal{J}} \big|\{j'|j'\in[T], j'=j\}\big| a_{ijk}x_{ij} \leq U_k, \forall k\in\mathcal{K} \\ & \sum_{i\in\mathcal{I}} x_{ij} \leq 1, \forall j\in\mathcal{J} \\ & x_{ij}\geq 0, \forall i\in\mathcal{I}, j\in\mathcal{J} \end{aligned} \tag{S.I II}$$

where $|\cdot|$ denote the cardinality of a given set.We use the W_{R_1} and W_{R_2} to denote the optimal value of Sample Instance (S.I I) and (S.I II)) respectively. Because the LP (S.I I) is a relaxation version of ILP (1), we know that $W_{R_1}\geq W_R$. For any optimal solution $\{x^*_{ij}\ \forall i\in\mathcal{I}, j\in[T]\}$ of LP (S.I I), it's easy to verify that the solution $\Big\{x_{ij}|x_{ij}=\frac{\sum_{j'=j,j'\in[T]} x^*_{ij'}}{\big|\{j'|j'\in[T]:j'=j\}\big|}, \forall i\in\mathcal{I}, j\in\mathcal{J}\Big\}$ is feasible for LP (S.I II), so that $W_{R_2}\geq W_{R_1}$. Moreover, the average of optimal solution for all possible LP (S.I II) is a feasible solution for LP (2), $\beta=0$, whose optimal solution is W_0. Thus $W_0\geq \mathbb{E}[W_{R_2}]\geq \mathbb{E}[W_{R_1}]\geq \mathbb{E}[W_R]$. ☐

D. Proof of Theorem 4.2

Theorem 4.2. *Under the strong feasible condition in Assumption 3, the optimal objective of $E(\tau)$ satisfies that*

$$W_\tau \geq \left(1-\frac{\tau}{\xi^*}\right)W_0,$$

where $\tau=\frac{\varepsilon}{1-\varepsilon}$.

Proof. The dual problem of the expected instance, i.e., problem $E(0)$, is

$$\begin{aligned} \min_{\alpha,\beta,\rho} \quad & \sum_{k\in\mathcal{K}}\alpha_k U_k - \sum_{k\in\mathcal{K}}\beta_k L_k + \sum_{j\in\mathcal{J}}\rho_j \\ s.t.\quad & \sum_{k\in\mathcal{K}}(\alpha_k-\beta_k)Tp_j a_{ijk} - Tp_j w_{ij} + \rho_j \geq 0, \\ & \forall i\in\mathcal{I}, j\in\mathcal{J}, \\ & \alpha_k,\beta_k,\rho_j\geq 0, k\in\mathcal{K}, j\in\mathcal{J}, \end{aligned} \tag{8}$$

and the dual problem of $E(\tau)$ is

$$\begin{aligned} \min_{\alpha,\beta,\rho} \quad & \sum_{k\in\mathcal{K}}\alpha_k U_k - \sum_{k\in\mathcal{K}}\beta_k (L_k+\tau T\bar{a}_k) + \sum_{j\in\mathcal{J}}\rho_j \\ s.t.\quad & \sum_{k\in\mathcal{K}}(\alpha_k-\beta_k)Tp_j a_{ijk} - Tp_j w_{ij} + \rho_j \geq 0, \\ & \forall i\in\mathcal{I}, j\in\mathcal{J}, \\ & \alpha_k,\beta_k,\rho_j\geq 0, k\in\mathcal{K}, j\in\mathcal{J}. \end{aligned} \tag{9}$$

It can be observed that LP (8) and LP (9) share the same feasible set, while LP (9) has an extra term $-\sum_{k\in\mathcal{K}}\tau T\bar{a}_k\beta_k$ in the objective. It is necessary to study the relationship between $\sum_{k\in\mathcal{K}} \tau T\bar{a}_k\beta_k$ and $\sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} \rho_j$ under the dual constraints if we want to obtain the ratio of W_τ to W_0. Motivated by (Jain et al., 2003), we propose a Factor-Revealing Linear Programming method for this analysis.

In order to derive the competitive ratio of the cumulative revenue obtained by the Algorithm $\widetilde{P}$ to W_0, we need to find a number $c \in (0,1)$ which makes $W_\tau \geq (1-c)W_0$ always hold. Considering the dual LP (8) and LP (9), if we can show that $\sum_{k\in\mathcal{K}} \frac{\varepsilon}{1-\varepsilon} T\bar{a}_k\beta_k \leq c(\sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} \rho_j)$ for any dual feasible solution, it will give us that $W_\tau \geq (1-c)W_0$. Hence this question can be translated to solving the following linear fractional programming

$$
\begin{aligned}
\max_{\alpha,\beta,\rho}\ & \frac{\sum_{k\in\mathcal{K}} \frac{\varepsilon}{1-\varepsilon} T\bar{a}_k\beta_k}{\sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} \rho_j} \\
s.t.\ & \sum_{k\in\mathcal{K}} (\alpha_k - \beta_k) T p_j a_{ijk} - T p_j w_{ij} + \rho_j \geq 0, \\
& \forall i \in \mathcal{I}, j \in \mathcal{J} \\
& \alpha_k, \beta_k, \rho_j \geq 0, k \in \mathcal{K}, j \in \mathcal{J}.
\end{aligned} \tag{10}
$$

Since $\sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} \rho_j \geq W_0 > 0$ for any dual feasible solution, we can do the following transformation

$$
\begin{aligned}
\widetilde{\alpha}_k &= \frac{\alpha_k}{\sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} \rho_j} \\
\widetilde{\beta}_k &= \frac{\beta_k}{\sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} \rho_j} \\
\widetilde{\rho}_j &= \frac{\rho_j}{\sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} \rho_j} \\
z &= \frac{1}{\sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} \rho_j}.
\end{aligned}
$$

In this way, we transfer the linear fractional programming (10) into

$$
\begin{aligned}
\max_{\widetilde{\alpha},\widetilde{\beta},\widetilde{\rho}}\ & \sum_{k\in\mathcal{K}} \frac{\varepsilon}{1-\varepsilon} T\bar{a}_k\widetilde{\beta}_k \\
s.t.\ & \sum_{k\in\mathcal{K}} (\widetilde{\alpha}_k - \widetilde{\beta}_k) T p_j a_{ijk} - T p_j w_{ij} z + \widetilde{\rho}_j \geq 0, \\
& \forall i \in \mathcal{I}, j \in \mathcal{J} \\
& \sum_{k\in\mathcal{K}} \widetilde{\alpha}_k U_k - \sum_{k\in\mathcal{K}} \widetilde{\beta}_k L_k + \sum_{j\in\mathcal{J}} \widetilde{\rho}_j = 1 \\
& \widetilde{\alpha}_k, \widetilde{\beta}_k, \widetilde{\rho}_j, z \geq 0, k \in \mathcal{K}, j \in \mathcal{J}.
\end{aligned} \tag{11}
$$

We investigate the dual problem of LP (11) as follows

$$
\begin{aligned}
\min_{t,d}\ & t \\
s.t.\ & \sum_{i\in\mathcal{I}} d_{ij} \leq t \\
& \sum_{i\in\mathcal{I}, j\in\mathcal{J}} d_{ij} T p_j a_{ijk} \leq t U_k \\
& \sum_{i\in\mathcal{I}, j\in\mathcal{J}} d_{ij} T p_j a_{ijk} \geq t L_k + \frac{\varepsilon}{1-\varepsilon} T\bar{a}_k \\
& \forall d_{ij} \geq 0, t \in \mathbb{R}, \forall i \in \mathcal{I}, j \in \mathcal{J}.
\end{aligned} \tag{12}
$$

By the first constraint of LP (12), it can be observed that $t \geq 0$ holds. With the auxiliary variable z_{ij} which makes

$d_{ij} = tz_{ij}$,we reformulate LP (12) into

$$
\begin{aligned}
t^* = & \min_{t,z} t \\
s.t.\ & t\bigg(\sum_{i\in\mathcal{I}} z_{ij} - 1\bigg) \le 0 \\
& t\bigg(\sum_{i\in\mathcal{I}, j\in\mathcal{J}} z_{ij} T p_j a_{ijk} - U_k\bigg) \le 0 \\
& t\bigg(\sum_{i\in\mathcal{I}, j\in\mathcal{J}} z_{ij} T p_j a_{ijk} - L_k\bigg) \ge \frac{\varepsilon}{1-\varepsilon} T\bar{a}_k \\
& \forall z_{ij} \ge 0, t \ge 0, \forall i \in \mathcal{I}, j \in \mathcal{J}.
\end{aligned}
\tag{13}
$$

According to the strong feasible condition Assumption 3, there exists a feasible solution $\{x'_{ij} | x'_{ij} \ge 0\}$ to LP (3) such that $\sum_{i\in\mathcal{I},j\in\mathcal{J}} x'_{ij} T p_j a_{ijk} \le U_k$, $\sum_{i\in\mathcal{I},j\in\mathcal{J}} T p_j a_{ijk} x'_{ij} \ge L_k + \xi^* T a_k$ and $\sum_{i\in\mathcal{I}} x'_{ij} \le 1, \forall k \in \mathcal{K}$. Therefore, $t = \frac{\tau}{\xi^*}$ and $z_{ij} = x'_{ij}\ \forall i \in \mathcal{I},\ j \in \mathcal{J}$ is a feasible solution to LP (13). As a result, we have $t^* \le \frac{\tau}{\xi^*}$. Since LP (10) has the same optimum as LP (13), we have that

$$
\begin{aligned}
& \sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k \bigg(L_k + \frac{\varepsilon}{1-\varepsilon} T\bar{a}_k\bigg) + \sum_{j\in\mathcal{J}} \rho_j \\
&= \sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} \rho_j - \sum_{k\in\mathcal{K}} \frac{\varepsilon}{1-\varepsilon} T\bar{a}_k \beta_k \\
&\ge \bigg(1 - \frac{\tau}{\xi^*}\bigg)\bigg(\sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} \rho_j\bigg)
\end{aligned}
$$

under the constraint of LP (8). Therefore, $W_\tau \ge (1 - \frac{\tau}{\xi^*}) W_0$. □

The factor-revealing linear fractional programming analysis shows the way we develop the definition of ξ^* and enlightens the design of feasibility estimator in Algorithm 5. Besides, the proof framework is reused in proving Lemma G.1.

E. Proof of Lemma 4.1 and Theorem 4.3

We restate the Lemma 4.1 as follows.

Lemma 4.1. *Under Assumption 1-3, if* $\forall \varepsilon > 0$ *and* $\gamma = \max\left(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{T\bar{a}_k - L_k}, \frac{\bar{w}}{W_\tau}\right) = O(\frac{\varepsilon^2}{\ln(K/\varepsilon)})$*, Algorithm 1 achieves an objective value at least* $(1-2\varepsilon)W_\tau$ *and satisfies the constraints w.p.* $1-\varepsilon$*.*

Proof. We first prove that, for every resource k, the consumed resource is below the capacity U_k w.h.p..

$$
\begin{aligned}
& P\left(\sum_{j=1}^{T} X_{jk}^{\widetilde{P}} \ge U_k\right) \\
&= P\left(\sum_{j=1}^{T} \Big(X_{jk}^{\widetilde{P}} - \mathbb{E}[X_{jk}^{\widetilde{P}}]\Big) \ge U_k - T\mathbb{E}[X_{jk}^{\widetilde{P}}]\right) \\
&\le \exp\left\{-\frac{\Big(U_k - T\mathbb{E}[X_{jk}^{\widetilde{P}}]\Big)^2}{2T\sigma^2 + \frac{2}{3}\bar{a}_k\Big(U_k - T\mathbb{E}[X_{jk}^{\widetilde{P}}]\Big)}\right\} \\
&= \exp\left\{-\frac{U_k - T\mathbb{E}[X_{jk}^{\widetilde{P}}]}{2\frac{T\sigma^2}{U_k - T\mathbb{E}[X_{jk}^{\widetilde{P}}]} + \frac{2}{3}\bar{a}_k}\right\}
\end{aligned}
\tag{14}
$$

$$\leq \exp\left\{-\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\frac{\bar{a}_k}{U_k}}\right\}$$
$$\leq \exp\left\{-\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma}\right\}$$
$$\leq \frac{\varepsilon}{2K+1}$$

where the first equality follows from $\mathbb{E}[X_{1j}^{\widetilde{P}}] = \mathbb{E}[X_{2j}^{\widetilde{P}}] = \cdots = \mathbb{E}[X_{Tj}^{\widetilde{P}}]$; the first inequality from Lemma B.1 and setting $\sigma^2 = Var(X_{jk}^{\widetilde{P}})$; in the second inequality, it can verified that $\sigma^2 \leq \mathbb{E}[(X_{jk}^{\widetilde{P}})^2] \leq \bar{a}_k \mathbb{E}[X_{jk}^{\widetilde{P}}] \leq \frac{(1-\varepsilon)\bar{a}_k U_k}{T}$, $U_k - T\mathbb{E}[X_{jk}^{\widetilde{P}}] \geq \varepsilon U_k$ and $\frac{T\sigma^2}{U_k - T\mathbb{E}[X_{ij}^{\widetilde{P}}]} \leq \bar{a}_k \frac{1-\varepsilon}{\varepsilon}$ so that $-\frac{U_k - T\mathbb{E}[X_{jk}^{\widetilde{P}}]}{2\frac{T\sigma^2}{U_k - T\mathbb{E}[X_{jk}^{\widetilde{P}}]} + \frac{2}{3}\bar{a}_k} \leq -\frac{\varepsilon U_k}{2\bar{a}_k \frac{1-\varepsilon}{\varepsilon} + \frac{2}{3}\bar{a}_k} = -\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\frac{\bar{a}_k}{U_k}}$; the third inequality from $\gamma \geq \frac{\bar{a}_k}{U_k}$; the final inequality from $\gamma = O(\frac{\varepsilon^2}{\ln(K/\varepsilon)})$.

Next, we verify that the Algorithm $\widetilde{P}$ satisfies the lower resource bound with high probability.

$$\begin{aligned}
&P\left(\sum_{j=1}^{T} X_{jk}^{\widetilde{P}} \leq L_k\right) \\
&= P\left(\sum_{j=1}^{T}\left(\mathbb{E}[X_{jk}^{\widetilde{P}}] - X_{jk}\right) \geq T\mathbb{E}[X_{jk}^{\widetilde{P}}] - L_k\right) \\
&\leq \exp\left\{-\frac{\left(T\mathbb{E}[X_{jk}^{\widetilde{P}}] - L_k\right)^2}{2T\sigma^2 + \frac{2}{3}\bar{a}_k\left(T\mathbb{E}[X_{jk}^{\widetilde{P}}] - L_k\right)}\right\} \\
&= \exp\left\{-\frac{(T\mathbb{E}[X_{jk}^{\widetilde{P}}] - L_k)}{2\frac{T\sigma^2}{T\mathbb{E}(X_{jk}) - L_k} + \frac{2}{3}\bar{a}_k}\right\} \\
&\leq \exp\left\{-\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\frac{\bar{a}_k}{T\bar{a}_k - L_k}}\right\} \\
&\leq \frac{\varepsilon}{2K+1}
\end{aligned} \tag{15}$$

where the first inequality from Lemma B.1 and setting $\sigma^2 = Var(X_{jk}^{\widetilde{P}})$; in the second inequality, we could verify that $\sigma^2 = Var(X_{jk}^{\widetilde{P}}) = Var(\bar{a}_k - X_{jk}^{\widetilde{P}}) \leq \mathbb{E}[(\bar{a}_k - X_{jk}^{\widetilde{P}})^2] \leq \bar{a}_k \mathbb{E}[\bar{a}_k - X_{jk}^{\widetilde{P}}] \leq \frac{(1-\varepsilon)\bar{a}_k(T\bar{a}_k - L_k)}{T}$, $T\mathbb{E}[X_{jk}^{\widetilde{P}}] - L_k \geq \varepsilon(T\bar{a}_k - L_k)$, and $\frac{T\sigma^2}{T\mathbb{E}[X_{jk}^{\widetilde{P}}] - L_k} \leq \bar{a}_k \frac{1-\varepsilon}{\varepsilon}$ so that $-\frac{(T\mathbb{E}[X_{jk}^{\widetilde{P}}] - L_k)}{2\frac{T\sigma^2}{T\mathbb{E}[X_{jk}^{\widetilde{P}}] - L_k} + \frac{2}{3}\bar{a}_k} \leq -\frac{\varepsilon(T\bar{a}_k - L_k)}{2\bar{a}_k\frac{1-\varepsilon}{\varepsilon} + \frac{2}{3}\bar{a}_k} = -\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\frac{\bar{a}_k}{T\bar{a}_k - L_k}}$; the final inequality from $\gamma = O(\frac{\varepsilon^2}{\ln(\frac{K}{\varepsilon})})$.

Therefore, from the previous outcomes, the consumed resource k satisfies our lower and upper bound requirements, w.h.p. Next, we investigate the revenue the Algorithm $\widetilde{P}$ brings.

$$\begin{aligned}
&P\left(\sum_{j=1}^{T} Y_j^{\widetilde{P}} \leq (1-2\varepsilon)W_\tau\right) \\
&= P\left(\sum_{j=1}^{T}\left(\mathbb{E}[Y_j^{\widetilde{P}}] - Y_j^{\widetilde{P}}\right) \geq T\mathbb{E}[Y_j^{\widetilde{P}}] - (1-2\varepsilon)W_\tau\right) \\
&\leq \exp\left\{-\frac{\left(T\mathbb{E}[Y_j^{\widetilde{P}}] - (1-2\varepsilon)W_\tau\right)^2}{2T\sigma_1^2 + \frac{2}{3}\bar{w}\left(T\mathbb{E}[Y_j^{\widetilde{P}}] - (1-2\varepsilon)W_\tau\right)}\right\}
\end{aligned} \tag{16}$$

$$
\begin{aligned}
&= \exp\left\{-\frac{T\mathbb{E}[Y_j^{\widetilde{P}}] - (1-2\varepsilon)W_\tau}{2\frac{T\sigma_1^2}{T\mathbb{E}[Y_j^{\widetilde{P}}]-(1-2\varepsilon)W_\tau} + \frac{2}{3}\bar{w}}\right\} \\
&\leq \exp\left\{-\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\frac{\bar{w}}{W_\tau}}\right\} \\
&\leq \frac{\varepsilon}{2K+1}
\end{aligned}
$$

where the first equality follows from $\mathbb{E}[Y_1^{\widetilde{P}}] = \mathbb{E}[Y_2^{\widetilde{P}}] = \cdots = \mathbb{E}[Y_T^{\widetilde{P}}]$; the first inequality from Lemma B.1 and setting $\sigma_1^2 = Var(Y_j^{\widetilde{P}})$; in the second inequality, we could easily verify that $\sigma_1^2 \leq \mathbb{E}[(Y_i^{\widetilde{P}})^2] \leq \bar{w}\mathbb{E}[Y_j^{\widetilde{P}}] \leq \frac{(1-\varepsilon)\bar{w}W_\tau}{T}$ and $\mathbb{E}[Y_j^{\widetilde{P}}] = (1-\varepsilon)\frac{W_\tau}{T}$ so that $-\frac{T\mathbb{E}[Y_j^{\widetilde{P}}]-(1-2\varepsilon)W_\tau}{2\frac{T\sigma_1^2}{T\mathbb{E}[Y_j^{\widetilde{P}}]-(1-2\varepsilon)W_\tau}+\frac{2}{3}\bar{w}} \leq -\frac{\varepsilon W_\tau}{2\bar{w}\frac{1-\varepsilon}{\varepsilon}+\frac{2}{3}\bar{w}} = -\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\frac{\bar{w}}{W_\tau}}$; the final inequality follows from $\gamma = O(\frac{\varepsilon^2}{\ln(\frac{K}{\varepsilon})})$.

From equation (14)-(16),

$$
\begin{aligned}
P(\sum_{j=1}^{T} Y_j^{\widetilde{P}} \leq (1-2\varepsilon)W_\tau) + \sum_{k\in\mathcal{K}} P(\sum_{j=1}^{T} X_{jk}^{\widetilde{P}} \notin [L_k, U_k]) \\
\leq (2K+1)\frac{\varepsilon}{2K+1} \leq \varepsilon
\end{aligned}
\tag{17}
$$

when $\gamma = O(\frac{\varepsilon^2}{\ln(\frac{K}{\varepsilon})})$, where $\gamma = \max(\frac{\bar{w}}{W_\tau}, \frac{\bar{a}_k}{T\bar{a}_k - L_k}, \frac{\bar{a}_k}{U_k})$. □

Theorem 4.3. *Under Assumption 1-3, if $\varepsilon > 0$, $\tau = \frac{\varepsilon}{1-\varepsilon}$ and $\gamma = O\big(\frac{\varepsilon^2}{\ln(K/\varepsilon)}\big)$, Algorithm 1 achieves an objective value of at least $\big(1-(2+\frac{1}{\xi^*})\varepsilon\big)W_0$ and satisfies the constraints w.p. $1-\varepsilon$.*

With Assumption 3, Theorem 4.2 and Lemma 4.1, the cumulative revenue is larger than $(1-2\varepsilon)W_\tau \geq (1-2\varepsilon)(1-\frac{\tau}{\xi^*})W_0 \geq (1-(2+\frac{1}{\xi^*})\varepsilon)W_0$. We finish the proof of Theorem 4.3.

F. Proof of Theorem 4.4

Theorem 4.4. *Under Assumption 1-3, if $\varepsilon > 0$, τ and γ are defined as Theorem 4.3, the Algorithm 2 achieves an objective value at least $\big(1-(2+\frac{1}{\xi^*})\varepsilon\big)W_0$ and satisfies the constraints w.p. $1-\varepsilon$.*

Proof. We consider the good event defined by

$$
\begin{aligned}
G :=& \left\{\sum_{j=1}^{s} X_{jk}^{A} + \sum_{j=s+1}^{T} X_{jk}^{\widetilde{P}} \leq U_k, \forall k \in \mathcal{K}\right\} \cap \left\{\sum_{j=1}^{s} X_{jk}^{A} + \sum_{j=s+1}^{T} X_{jk}^{\widetilde{P}} \geq L_k, \forall k \in \mathcal{K}\right\} \\
&\cap \left\{\sum_{j=1}^{s} Y_j^{A} + \sum_{j=s+1}^{T} Y_j^{\widetilde{P}} \geq (1-2\varepsilon)W_\tau\right\} \\
=& G_1 \cap G_2 \cap G_3,
\end{aligned}
$$

which means the hybrid Algorithm $A^s\widetilde{P}^{T-s}$ can achieve at least $(1-2\varepsilon)W_\tau$ revenue while satisfying the two-side constraints. We will show that the probability of the complement event G^c can be bounded by some moment generating functions.

For the first bad event G_1^c, we have

$$\begin{aligned}
&P(\sum_{j=1}^{s} X_{jk}^{A} + \sum_{j=s+1}^{T} X_{jk}^{\widetilde{P}} \geq U_k) \\
&\leq \min_{t>0} \mathbb{E}\left[\exp(t(\sum_{j=1}^{s} X_{jk}^{A} + \sum_{j=s+1}^{T} X_{jk}^{\widetilde{P}} - U_k))\right] \\
&= \min_{t>0} \mathbb{E}\left[\exp(t(\sum_{j=1}^{s} X_{jk}^{A} - \frac{s}{T}U_k) + t(\sum_{j=s+1}^{T} X_{jk}^{\widetilde{P}} - \frac{T-s}{T}U_k))\right] \\
&= \min_{t>0} \mathbb{E}\left[\phi_k^s(t)\exp(t(\sum_{j=s+1}^{T}(X_{jk}^{\widetilde{P}} - \mathbb{E}[X_{jk}^{\widetilde{P}}])) + \frac{T-s}{T}t(T\mathbb{E}[X_{jk}^{\widetilde{P}}] - U_k))\right] \\
&\leq \min_{t>0} \mathbb{E}\left[\phi_k^s(t)\exp((T-s)\frac{\sigma^2}{\bar{a}_k^2}(e^{t\bar{a}_k} - 1 - t\bar{a}_k) + \frac{-(T-s)t\varepsilon U_k}{T})\right] \\
&\leq \min_{t>0} \mathbb{E}\left[\phi_k^s(t)\exp(\frac{T-s}{T}\frac{(1-\varepsilon)U_k}{\bar{a}_k}(e^{t\bar{a}_k} - 1 - t\bar{a}_k - t\frac{\varepsilon}{1-\varepsilon}\bar{a}_k))\right] \\
&\leq \mathbb{E}\left[\phi_k^s(\frac{-\ln(1-\varepsilon)}{\bar{a}_k})\exp(-\frac{(1-\varepsilon)(T-s)U_k}{T\bar{a}_k}((1+\eta)\ln(1+\eta) - \eta))\right] \\
&\leq \mathbb{E}\left[\phi_k^s(\frac{-\ln(1-\varepsilon)}{\bar{a}_k})\exp(-\frac{T-s}{T}\frac{\varepsilon^2}{2\gamma(1-\frac{2}{3}\varepsilon)})\right]
\end{aligned} \tag{18}$$

where the first inequality follows from $\exp(t(\sum_{j=1}^{s} X_{jk}^{A} + \sum_{j=s+1}^{T} X_{jk}^{\widetilde{P}} - U_k)) \geq 1$ when $\sum_{j=1}^{s} X_{jk}^{A_j} + \sum_{j=s+1}^{T} X_{jk}^{\widetilde{P}} \geq U_k$; in the second equality, we set $\phi_k^s(t) = \exp(t(\sum_{j=1}^{s} X_{jk}^{A_j} - \frac{s}{T}U_k))$; the second inequality from Lemma B.1 and $T\mathbb{E}[X_{jk}^{\widetilde{P}}] \leq (1-\varepsilon)U_k$; the third inequality from $\sigma^2 = Var(X_{jk}^{\widetilde{P}}) \leq \frac{(1-\varepsilon)\bar{a}_k U_k}{T}$; in the fourth inequality, we set $t = \frac{-\ln(1-\varepsilon)}{\bar{a}_k}, \eta = \frac{\varepsilon}{1-\varepsilon}$; then the last inequality from $(1+\eta)\ln(1+\eta) - \eta \geq \frac{\eta^2}{2+\frac{2}{3}\eta}$ and the definition of γ.

Next,for the bad event G_2^c, we have

$$\begin{aligned}
&P(\sum_{j=1}^{s} X_{jk}^{A} + \sum_{j=s+1}^{T} X_{jk}^{\widetilde{P}} \leq L_k) \\
&\leq \min_{t>0} \mathbb{E}\left[\exp(t(L_k - \sum_{j=1}^{s} X_{jk}^{A} - \sum_{j=s+1}^{T} X_{jk}^{\widetilde{P}}))\right] \\
&= \min_{t>0} \mathbb{E}\left[\exp(t(\frac{s}{T}L_k - \sum_{j=1}^{s} X_{jk}^{A}) + t(\frac{T-s}{T}L_k - \sum_{j=s+1}^{n} X_{jk}^{\widetilde{P}}))\right] \\
&= \min_{t>0} \mathbb{E}\left[\varphi_k^s(t)\exp(t\sum_{j=s+1}^{T}(\mathbb{E}[X_{jk}^{\widetilde{P}}] - X_{jk}^{\widetilde{P}}) + \frac{T-s}{T}t(L_k - T\mathbb{E}[X_{jk}^{\widetilde{P}}]))\right] \\
&\leq \min_{t>0} \mathbb{E}\left[\varphi_k^s(t)\exp((T-s)\frac{\sigma^2}{\bar{a}_k^2}(e^{t\bar{a}_k} - 1 - t\bar{a}_k) - \frac{(T-s)\varepsilon(T\bar{a}_k - L_k))}{T}t\right] \\
&\leq \min_{t>0} \mathbb{E}\left[\varphi_k^s(t)\exp(\frac{(1-\varepsilon)(T-s)(T\bar{a}_k - L_k)}{\bar{a}_k}(e^{t\bar{a}_k} - 1 - t\bar{a}_k - t\frac{\varepsilon}{1-\varepsilon}\bar{a}_k))\right] \\
&\leq \mathbb{E}\left[\varphi_k^s(\frac{-\ln(1-\varepsilon)}{\bar{a}_k})\exp(-\frac{(1-\varepsilon)(T-s)(T\bar{a}_k - L_k)}{\bar{a}_k}((1+\eta)\ln(1+\eta) - \eta))\right] \\
&\leq \mathbb{E}\left[\varphi_k^s(\frac{-\ln(1-\varepsilon)}{\bar{a}_k})\exp(-\frac{T-s}{T}\frac{\varepsilon^2}{2\gamma(1-\frac{2}{3}\varepsilon)})\right]
\end{aligned} \tag{19}$$

where in the second equality, we set $\varphi_k^s(t) = \exp(t(\frac{s}{T}L_k - \sum_{j=1}^s X_{jk}^A))$; the second inequality from Lemma B.1 and $T\mathbb{E}[X_{jk}^{\widetilde{P}}] \geq (1-\varepsilon)L_k + \varepsilon T\bar{a}_k$; the third inequality from $\sigma^2 = Var(X_{jk}^{\widetilde{P}}) = Var(\bar{a}_k - X_{jk}^{\widetilde{P}}) \leq \bar{a}_k\mathbb{E}[\bar{a}_k - X_{jk}^{\widetilde{P}}] \leq \bar{a}_k\frac{(1-\varepsilon)(T\bar{a}_k - L_k)}{T}$; in the fourth inequality, we set $t = \frac{-\ln(1-\varepsilon)}{\bar{a}_k}, \eta = \frac{\varepsilon}{1-\varepsilon}$; The last inequality from $(1+\eta)\ln(1+\eta) - \eta \geq \frac{\eta^2}{2+\frac{2}{3}\eta}$ and the definition of γ.

Finally, we bound the probability of event G_3^c by

$$
\begin{aligned}
&P\left(\sum_{j=1}^{s} Y_j^A + \sum_{j=s+1}^{T} Y_j^{\widetilde{P}} \leq (1-2\varepsilon)W_\tau\right) \\
&\leq \min_{t>0} \mathbb{E}\left[\exp(t((1-2\varepsilon)W_\tau - \sum_{j=1}^{s} Y_j^A - \sum_{j=s+1}^{T} Y_j^{\widetilde{P}}))\right] \\
&= \min_{t>0} \mathbb{E}\left[\exp(t(\frac{s}{T}(1-2\varepsilon)W_\tau - \sum_{j=1}^{s} Y_j^A) + t(\frac{T-s}{T}(1-2\varepsilon)W_\tau - \sum_{j=s+1}^{T} Y_j^{\widetilde{P}}))\right] \\
&= \min_{t>0} \mathbb{E}\left[\psi^s(t)\exp(t\sum_{j=s+1}^{T}(\mathbb{E}[Y_j^{\widetilde{P}}] - Y_j^{\widetilde{P}}) + \frac{T-s}{T}t((1-2\varepsilon)W_\tau - T\mathbb{E}[Y_j^{\widetilde{P}}]))\right] \\
&\leq \min_{t>0} \mathbb{E}\left[\psi^s(t)\exp((T-s)\frac{\sigma_1^2}{\bar{w}^2}(e^{t\bar{w}} - 1 - t\bar{w}) + \frac{-(T-s)t\varepsilon W_\tau}{T})\right] \\
&\leq \min_{t>0} \mathbb{E}\left[\psi^s(t)\exp(\frac{(1-\varepsilon)(T-s)W_\tau}{T\bar{w}}(e^{t\bar{w}} - 1 - t\bar{w} - \frac{\varepsilon}{1-\varepsilon}t\bar{w}))\right] \\
&\leq \mathbb{E}\left[\psi^s(\frac{-\ln(1-\varepsilon)}{\bar{w}})\exp(-\frac{(1-\varepsilon)(T-s)W_\tau}{T\bar{w}}((1+\eta)\ln(1+\eta) - \eta))\right] \\
&\leq \mathbb{E}\left[\psi^s(\frac{-\ln(1-\varepsilon)}{\bar{w}})\exp(-\frac{T-s}{T}\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma})\right]
\end{aligned}
\tag{20}
$$

where in the second equality, we set $\psi^s(t) = \exp(t(\frac{s}{T}(1-2\varepsilon)W_\tau - \sum_{j=1}^s Y_j^A))$; the second inequality from Lemma B.1 and $T\mathbb{E}[Y_j^{\widetilde{P}}] = (1-\varepsilon)W_\tau$; the third inequality from $\sigma_1^2 = Var(Y_j^{\widetilde{P}}) \leq \bar{w}\mathbb{E}[Y_j^{\widetilde{P}}] \leq \frac{(1-\varepsilon)\bar{w}W_\tau}{T}$; in the fourth inequality, we set $t = \frac{-\ln(1-\varepsilon)}{\bar{w}}, \eta = \frac{\varepsilon}{1-\varepsilon}$; The last inequality from $(1+\eta)\ln(1+\eta) - \eta \geq \frac{\eta^2}{2+\frac{2}{3}\eta}$ and the definition of γ.

With the inequalities (18)-(20) and union bound in probability theory, we can show that $P(G^c) \leq \mathcal{F}(A^s\widetilde{P}^{T-s})$ where $\mathcal{F}(A^s\widetilde{P}^{T-s})$ is defined by

$$
\begin{aligned}
\mathcal{F}(A^s\widetilde{P}^{T-s}) =& \mathbb{E}\Bigg[\sum_{k\in\mathcal{K}}\phi_k^s(\frac{-\ln(1-\varepsilon)}{\bar{a}_k})\exp(-\frac{T-s}{T}\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma}) + \sum_{k\in\mathcal{K}}\varphi_k^s(-\frac{\ln(1-\varepsilon)}{\bar{a}_k})\exp(-\frac{T-s}{T}\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma}) \\
&+ \psi^s(\frac{-\ln(1-\varepsilon)}{\bar{w}})\exp(-\frac{T-s}{T}\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma})\Bigg]
\end{aligned}
$$

In Lemma 4.1, we have proven that $\mathcal{F}(\widetilde{P}^T) = (2K+1)\exp\left(-\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma}\right) \leq \varepsilon$, and we will show that $\mathcal{F}(A^s\widetilde{P}^{T-s}) \leq \mathcal{F}(A^{s-1}\widetilde{P}^{T-s+1})$ in the Lemma F.1. Thus, we have that $\mathcal{F}(A^T) \leq \mathcal{F}(\widetilde{P}^T) \leq \varepsilon$ by induction. Substituting $\tau = \frac{\varepsilon}{1-\varepsilon}$ and $W_\tau \geq (1-\frac{\tau}{\xi^*})W_0$ in Theorem 4.3, we complete the proof of Theorem 4.4 . □

Lemma F.1. $\mathcal{F}(A^s\widetilde{P}^{T-s}) \leq \mathcal{F}(A^{s-1}\widetilde{P}^{T-s+1})$

Proof. By the definition of $\mathcal{F}(A^s\widetilde{P}^{T-s})$, we have that

$$\begin{aligned}\mathcal{F}(A^s\widetilde{P}^{T-s}) = &\Bigg(\mathbb{E}\Big[\sum_{k\in\mathcal{K}}\phi_k^{s-1}(\frac{-\ln(1-\varepsilon)}{\bar{a}_k})\exp((-\frac{\ln(1-\varepsilon)}{\bar{a}_k})(X_{jk}^A-\frac{U_k}{T}))\\&+\sum_{k\in\mathcal{K}}\varphi_k^{s-1}(-\frac{\ln(1-\varepsilon)}{\bar{a}_k})\exp((-\frac{\ln(1-\varepsilon)}{\bar{a}_k})(\frac{L_k}{T}-X_{jk}^A))\\&+\psi^{s-1}(\frac{-\ln(1-\varepsilon)}{\bar{w}}))\exp((\frac{-\ln(1-\varepsilon)}{\bar{w}})(\frac{(1-2\varepsilon)W_\tau}{T}-Y_s^A))\Big]\Bigg)\exp(-\frac{T-s}{T}\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma}).\end{aligned} \tag{21}$$

According to algorithm A in Algorithm 2, we allocate the s-th request to the channel i^* where

$$\begin{aligned}i^* = \arg\min_{i\in\mathcal{I}}&\sum_{k\in\mathcal{K}}\phi_k^{s-1}(-\frac{\ln(1-\varepsilon)}{\bar{a}_k})\exp(-\frac{\ln(1-\varepsilon)}{\bar{a}_k}(a_{isk}-\frac{U_k}{T}))\\&+\sum_{k\in\mathcal{K}}\varphi_k^{s-1}(-\frac{\ln(1-\varepsilon)}{\bar{a}_k})\exp(-\frac{\ln(1-\varepsilon)}{\bar{a}_k}(\frac{L_k}{T}-a_{isk}))\\&+\psi^{s-1}(-\frac{\ln(1-\varepsilon)}{\bar{w}})\exp(-\frac{\ln(1-\varepsilon)}{\bar{w}}(\frac{(1-2\varepsilon)W_\tau}{T}-w_{is}))\end{aligned} \tag{22}$$

which means that

$$\begin{aligned}\mathcal{F}(A^s\widetilde{P}^{T-s}) \le &\Big(\mathbb{E}\Big[\sum_{k\in\mathcal{K}}\phi_k^{s-1}(\frac{-\ln(1-\varepsilon)}{\bar{a}_k})\underbrace{\exp((-\frac{\ln(1-\varepsilon)}{\bar{a}_k})(X_{sk}^{\widetilde{P}}-\frac{U_k}{T}))}_{①}\\&+\sum_{k\in\mathcal{K}}\varphi_k^{s-1}(-\frac{\ln(1-\varepsilon)}{\bar{a}_k})\underbrace{\exp((-\frac{\ln(1-\varepsilon)}{\bar{a}_k})(\frac{L_k}{T}-X_{sk}^{\widetilde{P}}))}_{②}\\&+\psi^{s-1}(\frac{-\ln(1-\varepsilon)}{\bar{w}})\underbrace{\exp((\frac{-\ln(1-\varepsilon)}{\bar{w}})(\frac{(1-2\varepsilon)W_\tau}{T}-Y_s^{\widetilde{P}}))}_{③}\Big]\Big)\exp\big(-\frac{T-s}{T}\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma}\big).\end{aligned} \tag{23}$$

For the term ①, we have

$$\begin{aligned}① &= \mathbb{E}\left[\exp\left(-\frac{\ln(1-\varepsilon)}{\bar{a}_k}\Big((X_{sk}^{\widetilde{P}}-\mathbb{E}[X_{sk}^{\widetilde{P}}])+(\mathbb{E}[X_{sk}^{\widetilde{P}}]-\frac{U_k}{T})\Big)\right)\right]\\&\le \mathbb{E}\left[\exp\left(\frac{\sigma^2}{\bar{a}_k^2}\Big(e^{-\ln(1-\varepsilon)}-1+\ln(1-\varepsilon)\Big)+\frac{\varepsilon U_k}{T\bar{a}_k}\ln(1-\varepsilon)\right)\right]\\&\le \mathbb{E}\left[\exp\left(\frac{(1-\varepsilon)U_k}{T\bar{a}_k}\Big(\frac{1}{1-\varepsilon}-1+\ln(1-\varepsilon)+\frac{\varepsilon}{1-\varepsilon}\ln(1-\varepsilon)\Big)\right)\right]\\&= \mathbb{E}\left[\exp\left(-\frac{(1-\varepsilon)U_k}{T\bar{a}_k}\Big((1+\eta)\ln(1+\eta)-\eta\Big)\right)\right]\\&\le \exp\Big(-\frac{1}{T}\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma}\Big)\end{aligned} \tag{24}$$

where the first inequality follows from Lemma B.1 and $\mathbb{E}[X_{sk}^{\widetilde{P}}] \le (1-\varepsilon)U_k/T$, the second from $\sigma^2 = Var(X_{sk}^{\widetilde{P}}) \le \frac{(1-\varepsilon)\bar{a}_k U_k}{T}$. Next setting $\eta = \frac{\varepsilon}{1-\varepsilon}$, the last inequality follows from $(1+\eta)\ln(1+\eta)-\eta \ge \frac{\eta^2}{2+\frac{2}{3}\eta}$ and the definition of γ.

For the term ②, we have

$$\begin{aligned}
② =& \mathbb{E}\left[\exp\Big(-\frac{\ln(1-\varepsilon)}{\bar{a}_k}\Big((\mathbb{E}[X_{sk}^{\widetilde{P}}] - X_{sk}^{\widetilde{P}}) + (\frac{L_k}{T} - \mathbb{E}[X_{sk}^{\widetilde{P}}])\Big)\Big)\right] \\
\le& \mathbb{E}\left[\exp\Big(\frac{\sigma^2}{\bar{a}_k^2}\Big(e^{-\ln(1-\varepsilon)} - 1 + \ln(1-\varepsilon)\Big) + \frac{\varepsilon(T\bar{a}_k - L_k)}{T\bar{a}_k}\ln(1-\varepsilon)\Big)\right] \\
\le& \mathbb{E}\left[\exp\Big(\frac{(1-\varepsilon)(T\bar{a}_k - L_k)}{T\bar{a}_k}\Big(\frac{1}{1-\varepsilon} - 1 + \ln(1-\varepsilon) + \frac{\varepsilon}{1-\varepsilon}\ln(1-\varepsilon)\Big)\Big)\right] \\
=& \mathbb{E}\left[\exp\Big(-\frac{(1-\varepsilon)(T\bar{a}_k - L_k)}{T\bar{a}_k}\Big((1+\eta)\ln(1+\eta) - \eta\Big)\Big)\right] \\
\le& \exp\Big(-\frac{1}{T}\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma}\Big)
\end{aligned} \tag{25}$$

where the first inequality follows from Lemma B.1 and $\mathbb{E}[X_{sk}^{\widetilde{P}}] \le \frac{(1-\varepsilon)L_k + \varepsilon T\bar{a}_k}{T}$, the second from $\sigma^2 = Var(X_{sk}^{\widetilde{P}}) \le \bar{a}_k\frac{(1-\varepsilon)(T\bar{a}_k - L_k)}{T}$. Next setting $\eta = \frac{\varepsilon}{1-\varepsilon}$, the last inequality follows from $(1+\eta)\ln(1+\eta) - \eta \ge \frac{\eta^2}{2+\frac{2}{3}\eta}$ and the definition of γ.

For the term ③, we have

$$\begin{aligned}
③ =& \mathbb{E}\left[\exp\Big(-\frac{\ln(1-\varepsilon)}{\bar{w}}\Big((\mathbb{E}[Y_s^{\widetilde{P}}] - Y_s^{\widetilde{P}}) + (\frac{(1-2\varepsilon)W_\tau}{T} - \mathbb{E}[Y_s^{\widetilde{P}}])\Big)\Big)\right] \\
\le& \mathbb{E}\left[\exp\Big(\frac{\sigma_1^2}{\bar{w}^2}\Big(e^{-\ln(1-\varepsilon)} - 1 + \ln(1-\varepsilon)\Big) + \frac{\varepsilon W_\tau}{T\bar{w}}\ln(1-\varepsilon)\Big)\right] \\
\le& \mathbb{E}\left[\exp\Big(\frac{(1-\varepsilon)W_\tau}{T\bar{w}}\Big(\frac{1}{1-\varepsilon} - 1 + \ln(1-\varepsilon) + \frac{\varepsilon}{1-\varepsilon}\ln(1-\varepsilon)\Big)\Big)\right] \\
=& \mathbb{E}\left[\exp\Big(-\frac{(1-\varepsilon)W_\tau}{T\bar{w}}\Big((1+\eta)\ln(1+\eta) - \eta\Big)\Big)\right] \\
\le& \exp\Big(-\frac{1}{T}\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma}\Big)
\end{aligned} \tag{26}$$

where the first inequality follows from Lemma B.1 and $\mathbb{E}[Y_s^{\widetilde{P}}] = \frac{(1-\varepsilon)W_\tau}{T}$, the second from $\sigma_1^2 = Var(Y_s^{\widetilde{P}}) \le \frac{(1-\varepsilon)\bar{w}W_\tau}{T}$. Next setting $\eta = \frac{\varepsilon}{1-\varepsilon}$, the last inequality follows from $(1+\eta)\ln(1+\eta) - \eta \ge \frac{\eta^2}{2+\frac{2}{3}\eta}$ and the definition of γ. According to the inequality (24)-(26), we can show that

$$\begin{aligned}
\mathcal{F}(A^s\widetilde{P}^{T-s}) \le& \Big(\mathbb{E}\Big[\sum_{k\in\mathcal{K}}\phi_k^{s-1}(\frac{-\ln(1-\varepsilon)}{\bar{a}_k}) + \sum_{k\in\mathcal{K}}\varphi_k^{s-1}(-\frac{\ln(1-\varepsilon)}{\bar{a}_k}) \\
&+ \psi^{s-1}(\frac{-\ln(1-\varepsilon)}{\bar{w}})\Big]\Big)\exp(-\frac{T-s+1}{T}\frac{\varepsilon^2}{2(1-\frac{2}{3}\varepsilon)\gamma}) \\
=& \mathcal{F}(A^{s-1}\widetilde{P}^{T-s+1})
\end{aligned} \tag{27}$$

which completes the proof.

□

G. Proof of Theorem 4.5

G.1. Concentration of Z^r

In the first step, we study the relationship between Z^r and W_ε.

Lemma G.1. *Under Assumption 1-3, if $\tau_1 + \varepsilon \le \xi^*$ and $\gamma_1 = \max\big(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{(1-\varepsilon)T\bar{a}_k - L_k}, \frac{\bar{w}}{W_{\varepsilon+\tau_1}}\big) = O\big(\frac{\varepsilon^2}{\ln(K/\varepsilon)}\big)$, for given measure of feasibility ξ^* , we have*

$$W^r \le \frac{t_r W_\varepsilon}{T}\Big(1 + \big(2 + \frac{1}{\xi^* - \varepsilon}\big)\varepsilon_{x,r}\Big)$$

with probability at least $1-\delta$, where the predefined parameter $\varepsilon > 0$, $\tau_1 = \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}$, $\delta = \frac{\varepsilon}{3l}$, $l = \log_2(\frac{1}{\varepsilon})$ and $\varepsilon_{x,r} =$

$\sqrt{\frac{4T\gamma_1 \ln(\frac{2K+1}{\delta})}{t_r}}$.

Proof. We consider the definition of W^r:

$$\begin{aligned} W^r = &\max_x \sum_{i\in\mathcal{I}, j\in\mathcal{S}_r} w_{ij}x_{ij} \\ & s.t.\ \frac{t_r}{T}(L_k + \varepsilon T\bar{a}_k) \le \sum_{i\in\mathcal{I}, j\in\mathcal{S}_r} a_{ijk}x_{ij} \le \frac{t_r}{T}U_k, \forall k \in \mathcal{K} \\ & \sum_{i\in\mathcal{I}} x_{ij} \le 1, \forall j \in \mathcal{S}_r \\ & x_{ij} \ge 0, \forall i \in \mathcal{I}, j \in \mathcal{S}_r \end{aligned} \tag{28}$$

where we use $\mathcal{S}_r$ to denote the request set in stage r. The dual of LP (28) is

$$\begin{aligned} W^r = &\min_{\alpha,\beta,\rho} \sum_{k\in\mathcal{K}} \alpha_k \frac{t_r}{T}U_k - \sum_{k\in\mathcal{K}} \beta_k \frac{t_r}{T}(L_k + \varepsilon T\bar{a}_k) + \sum_{j\in\mathcal{S}_r} \rho_j \\ & s.t.\ \sum_{k\in\mathcal{K}} (\alpha_k - \beta_k)a_{ijk} - w_{ij} + \rho_j \ge 0\ \forall i \in \mathcal{I}, j \in \mathcal{S}_r \\ & \alpha_k, \beta_k, \rho_j \ge 0, k \in \mathcal{K}, j \in \mathcal{S}_r \end{aligned} \tag{29}$$

Comparing to the dual of LP (28) with the dual of problem $E(\varepsilon)$, which is

$$\begin{aligned} &\min_{\alpha,\beta,\rho} \sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k (L_k + \varepsilon T\bar{a}_k) + \sum_{j\in\mathcal{J}} Tp_j\rho_j \\ & s.t.\ \sum_{k\in\mathcal{K}} (\alpha_k - \beta_k)a_{ijk} - w_{ij} + \rho_j \ge 0\ \forall i \in \mathcal{I}, j \in \mathcal{J} \\ & \alpha_k, \beta_k, \rho_j \ge 0, k \in \mathcal{K}, j \in \mathcal{J} \end{aligned} \tag{30}$$

we can observe that the constraints of LP (29) is a subset of those of LP (30). We denote the primal and dual optimal solution of $E(\varepsilon)$ as $\{x^*_{ij}\}$ and $\{\alpha^*_k, \beta^*_k, \rho^*_k\}$. So $\{\alpha^*_k, \beta^*_k, \rho^*_k\}$ is feasible for LP (29).

Hence,

$$\begin{aligned} W^r &\le \sum_{k\in\mathcal{K}} \alpha^*_k \frac{t_r}{T}U_k - \sum_{k\in\mathcal{K}} \beta^*_k \frac{t_r}{T}(L_k + \varepsilon T\bar{a}_k) + \sum_{j\in\mathcal{S}_r} \rho^*_j \\ &= \underbrace{\sum_{k\in\mathcal{K}} \alpha^*_k (\frac{t_r}{n}U_k - \sum_{j\in\mathcal{S}_r, i\in\mathcal{I}} a_{ijk}x^*_{ij})}_{①} + \underbrace{\sum_{k\in\mathcal{K}} \beta^*_k (\sum_{j\in\mathcal{S}_r, i\in\mathcal{I}} a_{ijk}x^*_{ij} - \frac{t_r}{T}(L_k + \varepsilon T\bar{a}_k))}_{②} \\ &+ \underbrace{\sum_{j\in\mathcal{S}_r} (\rho^*_j + \sum_{i\in\mathcal{I}, k\in\mathcal{K}} (\alpha^*_k - \beta^*_k)a_{ijk}x^*_{ij})}_{③} \end{aligned} \tag{31}$$

We have divided the equation (31) into three parts. Next, we will derive the relationship between W^r and W_ε by controlling these three parts. To facilitate the analysis, we first present the KKT conditions(Boyd et al., 2004) for the problem $E(\varepsilon)$ as follows

$$\begin{aligned} &\sum_{k\in\mathcal{K}} (\alpha^*_k - \beta^*_k)a_{ijk}x^*_{ij} - w_{ij}x^*_{ij} + \rho^*_t x^*_{ij} = 0, \forall i \in \mathcal{I}, j \in \mathcal{S}_r \\ &\rho^*_j (\sum_{i\in\mathcal{I}} x^*_{ij} - 1) = 0, \forall j \in \mathcal{S}_r \\ &\alpha^*_k (\sum_{ij} Tp_j a_{ijk}x^*_{ij} - U_k) = 0, \forall k \in \mathcal{K} \\ &\beta^*_k (L_k + \varepsilon T\bar{a}_k - \sum_{ij} Tp_j a_{ijk}x^*_{ij}) = 0, \forall k \in \mathcal{K}. \end{aligned} \tag{32}$$

For part ①, according to the KKT conditions, if $\sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk} x^*_{ij} < U_k$, then $\alpha^*_k = 0$. Thus we only consider the resource k making $\sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk} x^*_{ij} = U_k$. According to the Assumption 1, we have $\mathbb{E}(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}} a_{ijk}x^*_{ij}) = \frac{t_r}{T}U_k$ $\forall j \in \mathcal{S}_r$. Thus, we can show that

$$\begin{aligned}
&P\left(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}} a_{ijk}x^*_{ij} \le (1-\varepsilon_{x,r})\frac{t_r}{T}U_k\right)\\
&= P\left(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}} a_{ijk}x^*_{ij} - \mathbb{E}[\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}} a_{ijk}x^*_{ij}] \le -\varepsilon_{x,r}\frac{t_r}{T}U_k\right)\\
&\le \exp\left(-\frac{t_r U_k^2 \varepsilon_{x,r}^2/T^2}{2Var(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}} a_{ijk}x^*_{ij})/t_r + \frac{2}{3}\bar{a}_k U_k \varepsilon_{x,r}/T}\right)\\
&\le \exp\left(-\frac{\frac{t_r}{T}\varepsilon_{x,r}^2}{2(1+\frac{\varepsilon_{x,r}}{3})\frac{\bar{a}_k}{U_k}}\right)\\
&\le \exp\left(-\frac{\frac{t_r}{T}\varepsilon_{x,r}^2}{2(1+\frac{\varepsilon_{x,r}}{3})\gamma_1}\right)\\
&\le \frac{\delta}{2K+1}
\end{aligned} \tag{33}$$

where the first inequality follows from the Bernstein inequality in Lemma B.1, the second inequality from $Var(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}} a_{ijk}x^*_{ij})/t_r \le \bar{a}_k\mathbb{E}[\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}} a_{ijk}x_{ij^*}]/t_r = \bar{a}_k\frac{U_k}{T}$, the third inequality from the definition of γ_1 and the last from the definition of $\varepsilon_{x,r}$.

For part ②, we only consider the k making $\sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk}x^*_{ij} = L_k + \varepsilon T\bar{a}_k$, which means $\mathbb{E}\Big[\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}(\bar{a}_k - a_{ijk}x^*_{ij})\Big] = \frac{t_r}{T}\Big((1-\varepsilon)T\bar{a}_k - L_k\Big)$. Using Bernstein inequality, we have

$$\begin{aligned}
&P\left(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}(\bar{a}_k - a_{ijk}x^*_{ij}) \le (1-\varepsilon_{x,r})\frac{t_r}{T}((1-\varepsilon)T\bar{a}_k - L_k)\right)\\
&= P\left(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}(\bar{a}_k - a_{ijk}x^*_{ij}) - \mathbb{E}\Big[\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}(\bar{a}_k - a_{ijk}x^*_{ij})\Big] \le -\varepsilon_{x,r}\frac{t_r}{T}((1-\varepsilon)T\bar{a}_k - L_k)\right)\\
&\le \exp\left(-\frac{t_r\Big((1-\varepsilon)T\bar{a}_k - L_k\Big)^2\varepsilon_{x,r}^2/T^2}{2Var\Big(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}(\bar{a}_k - a_{ijk}x^*_{ij})\Big)/t_r + \frac{2}{3}\bar{a}_k\Big((1-\varepsilon)T\bar{a}_k - L_k\Big)\varepsilon_{x,r}/T}\right)\\
&\le \exp\left(-\frac{\frac{t_r}{T}\varepsilon_{x,r}^2}{2(1+\frac{\varepsilon_{x,r}}{3})\frac{\bar{a}_k}{(1-\varepsilon)T\bar{a}_k - L_k}}\right)\\
&\le \exp\left(-\frac{\frac{t_r}{T}\varepsilon_{x,r}^2}{2(1+\frac{\varepsilon_{x,r}}{3})\gamma_1}\right)\\
&\le \frac{\delta}{2K+1}
\end{aligned} \tag{34}$$

where the second inequality from $Var\Big(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}(\bar{a}_k - a_{ijk}x^*_{ij})\Big)/t_r \le \bar{a}_k\mathbb{E}[\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}(\bar{a}_k - a_{ijk}x^*_{ij})]/t_r = \bar{a}_k\frac{(1-\varepsilon)T\bar{a}_k - L_k}{T}$, the third inequality from the definition of γ_1 and the last from the definition of $\varepsilon_{x,r}$.

For the last part ③, from KKT conditions, it's easy to verify that $\sum_{i\in\mathcal{I},k\in\mathcal{K}}(\alpha^*_k - \beta^*_k)a_{ijk}x^*_{ij} - \sum_{i\in\mathcal{I}} w_{ij}x^*_{ij} + \rho^*_j\sum_{i\in\mathcal{I}} x^*_{ij} = \sum_{i\in\mathcal{I},k\in\mathcal{K}}(\alpha^*_k - \beta^*_k)a_{ijk}x^*_{ij} - \sum_{i\in\mathcal{I}} w_{ij}x^*_{ij} + \rho^*_j = 0$, so $\sum_{i\in\mathcal{I},k\in\mathcal{K}}(\alpha^*_k - \beta^*_k)a_{ijk}x^*_{ij} + \rho^*_j = \sum_{i\in\mathcal{I}} w_{ij}x^*_{ij} \in [0,\bar{w}]$. Moreover, $\mathbb{E}[\rho^*_j + \sum_{i\in\mathcal{I},k\in\mathcal{K}}(\alpha^*_k - \beta^*_k)a_{ijk}x^*_{ij}] = \frac{\sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j w_{ij}x^*_{ij}}{T} = \frac{W_\varepsilon}{T}$ $\forall j \in \mathcal{S}_r$. Therefore, following the similar analysis as equation (33), we have

$$\begin{aligned}
&P\left(\sum_{j\in\mathcal{S}_r}\left(\rho_j^*+\sum_{i\in\mathcal{I},k\in\mathcal{K}}(\alpha_k^*-\beta_k^*)a_{ijk}x_{ij}^*\right)\geq\frac{t_r}{T}W_\varepsilon(1+\varepsilon_{x,r})\right)\\
&=P(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}w_{ij}x_{ij}^*-\mathbb{E}[\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}w_{ij}x_{ij}^*]\geq\varepsilon_{x,r}\frac{t_r}{T}W_\varepsilon)\\
&\leq\exp\left(-\frac{t_rW_\varepsilon^2\varepsilon_{x,r}^2/T^2}{2Var(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}w_{ij}x_{ij}^*)/t_r+\frac{2}{3}\bar{a}_kW_\varepsilon\varepsilon_{x,r}/T}\right)\\
&\leq\exp(-\frac{\frac{t_r}{T}\varepsilon_{x,r}^2}{2(1+\frac{\varepsilon_{x,r}}{3})\frac{\bar{w}}{W_\varepsilon}})\\
&\leq\exp\left(-\frac{\frac{t_r}{T}\varepsilon_{x,r}^2}{2(1+\frac{\varepsilon_{x,r}}{3})\gamma_1}\right)\\
&\leq\frac{\delta}{2K+1}
\end{aligned}\tag{35}$$

where the first inequality follows from the Bernstein inequality in Lemma B.1, the second inequality from $Var(\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}w_{ij}x_{ij}^*)/t_r\leq\bar{w}\mathbb{E}[\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}w_{ij}x_{ij}^*]/t_r=\bar{w}\frac{W_\varepsilon}{T}$, the third inequality from the definition of γ_1 and the last from the definition of $\varepsilon_{x,r}$.

Based on inequality (32)-(34), we have shown that the following inequalities holds with probability at least $1-\delta$

$$\begin{aligned}
&\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}a_{ijk}x_{ij}^*\geq(1-\varepsilon_{x,r})\frac{t_r}{T}U_k,\forall k\in\mathcal{K}\\
&\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}(\bar{a}_k-a_{ijk}x_{ij}^*)\geq(1-\varepsilon_{x,r})\frac{t_r}{T}((1-\varepsilon)T\bar{a}_k-L_k),\forall k\in\mathcal{K}\\
&\sum_{j\in\mathcal{S}_r}(\rho_j^*+\sum_{i\in\mathcal{I},k\in\mathcal{K}}(\alpha_k^*-\beta_k^*)a_{ijk}x_{ij}^*)\leq\frac{t_r}{T}W_\varepsilon(1+\varepsilon_{x,r}),\forall k\in\mathcal{K}.
\end{aligned}\tag{36}$$

Therefore, with probability at least $1-\delta$, we have

$$\begin{aligned}
&①+②+③\\
&=\sum_{k\in\mathcal{K}}\alpha_k^*(\frac{t_r}{T}U_k-\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}a_{ijk}x_{ij}^*)+\sum_{k\in\mathcal{K}}\beta_k^*\big(\frac{t_r}{T}\big((1-\varepsilon)T\bar{a}_k-L_k\big)-\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}(\bar{a}_k-a_{ijk}x_{ij}^*)\big)+③\\
&\leq\varepsilon_{x,r}\sum_{k\in\mathcal{K}}\left(\alpha_k^*\frac{t_r}{T}U_k+\beta_k^*\Big(\frac{t_r}{T}\big((1-\varepsilon)T\bar{a}_k-L_k\big)\Big)\right)+③\\
&\leq\varepsilon_{x,r}\sum_{k\in\mathcal{K}}\left(\alpha_k^*\frac{t_r}{T}U_k+\beta_k^*\Big(\frac{t_r}{T}\big((1-\varepsilon)T\bar{a}_k-L_k\big)\Big)\right)+\frac{t_r}{T}W_\varepsilon(1+\varepsilon_{x,r})\\
&\leq\varepsilon_{x,r}\big(\frac{t_r}{T}W_\varepsilon+\sum_{k\in\mathcal{K}}\beta_k^*t_r\bar{a}_k\big)+\frac{t_r}{T}W_\varepsilon(1+\varepsilon_{x,r})\\
&\leq\big(1+(2+\frac{1}{\xi^*-\varepsilon})\varepsilon_{x,r}\big)\frac{t_r}{T}W_\varepsilon
\end{aligned}\tag{37}$$

where the first inequality from $\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}a_{ijk}x_{ij}^*\geq(1-\varepsilon_{x,r})\frac{t_r}{T}U_k$ and $\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}}(\bar{a}_k-a_{ijk}x_{ij}^*)\geq(1-\varepsilon_{x,r})\frac{t_r}{T}((1-\varepsilon)T\bar{a}_k-L_k)$; the second inequality from $\sum_{j\in\mathcal{S}_r}(\rho_j^*+\sum_{i\in\mathcal{I},k\in\mathcal{K}}(\alpha_k^*-\beta_k^*)a_{ijk}x_{ij}^*)\leq\frac{t_r}{T}W_\varepsilon(1+\varepsilon_{x,r})$; the third inequality from $W_\varepsilon=\sum_{k\in\mathcal{K}}\alpha_k^*U_k-\sum_{k\in\mathcal{K}}\beta_k^*(L_k+\varepsilon T\bar{a}_k)+\sum_{j\in\mathcal{J}}Tp_j\rho_j^*\geq\sum_{k\in\mathcal{K}}\alpha_k^*U_k-\sum_{k\in\mathcal{K}}\beta_k^*(\varepsilon T\bar{a}_k+L_k)$; the final inequality from $\sum_{k\in\mathcal{K}}\beta_k^*T\bar{a}_k\leq\frac{1}{\xi^*-\varepsilon}W_\varepsilon$, which can be shown if we follow the proof of Theorem 4.2 in Appendix D and regard the problem $E(\varepsilon)$ as an LP with $\xi^*-\varepsilon$ measure of feasibility. □

Theorem G.2. *Under Assumption 1-3, if $\tau_1+\varepsilon\leq\xi^*$ and $\gamma_1=\max\big(\frac{\bar{a}_k}{U_k},\frac{\bar{a}_k}{(1-\varepsilon)T\bar{a}_k-L_k},\frac{\bar{w}}{W_{\varepsilon+\tau_1}}\big)=O\big(\frac{\varepsilon^2}{\ln(K/\varepsilon)}\big)$, with* *probability* $1-2\delta$, *we have*

$$\frac{t_r W_\varepsilon}{T}\left(1-\big(2+\frac{1}{\xi^*-\varepsilon}\big)\varepsilon_{x,r}\right) \le W^r \le \frac{t_r W_\varepsilon}{T}\left(1+\big(2+\frac{1}{\xi^*-\varepsilon}\big)\varepsilon_{x,r}\right)$$

where the predefined parameter $\varepsilon > 0$, $\tau_1 = \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}$, $\delta = \frac{\varepsilon}{3l}$, $l = \log_2(\frac{1}{\varepsilon})$ and $\varepsilon_{x,r} = \sqrt{\frac{4T\gamma_1 \ln(\frac{2K+1}{\delta})}{t_r}}$.

Proof. **RHS:** we have proven the right hand side in Lemma G.1.

LHS: For every request in stage r, we consider to imitate Algorithm 1 to design an algorithm $\widetilde{P}_1$. In Algorithm $\widetilde{P}_1$, we first solve the LP problem $E(\varepsilon + \frac{\varepsilon_{x,r}}{1-\varepsilon_{x,r}})$ to get the LP solution x_{ij}^{1*} and then assign request $j \in \mathcal{J}$ to channel $i \in \mathcal{I}$ with probability $(1-\varepsilon_{x,r})x_{ij}^{1*}$. Following the similar analysis in the proof of Lemma 4.1, we can prove that , with probability $1-\delta$,

$$\begin{aligned}
&P\left(\sum_{j=1}^{t_r} X_{jk}^{\widetilde{P}_1} \ge \frac{t_r}{T}U_k\right) \le \exp\left(-\frac{t_r\varepsilon_{x,r}^2}{2(1-\frac{2}{3}\varepsilon_{x,r})T\frac{\bar{a}_k}{U_k}}\right)\\
&P\left(\sum_{j=1}^{t_r} Y_j^{\widetilde{P}_1} \le (1-(2+\frac{1}{\xi^*-\varepsilon})\varepsilon_{x,r})\frac{t_r}{T}W_\varepsilon\right) \le \exp\left(-\frac{t_r\varepsilon_{x,r}^2}{2(1-\frac{2}{3}\varepsilon_{x,r})T\frac{\bar{w}}{W_{\tau_1+\varepsilon}}}\right)\\
&P\left(\sum_{j=1}^{t_r} X_{jk}^{\widetilde{P}_1} \le \frac{t_r}{T}L_k\right) \le \exp\left(-\frac{t_r\varepsilon_{x,r}^2}{2(1-\frac{2}{3}\varepsilon_{x,r})T\frac{\bar{a}_k}{(1-\varepsilon-\tau_1)T\bar{a}_k-L_k}}\right)
\end{aligned} \tag{38}$$

where the second inequality from the truth that the problem $E(\varepsilon)$ is satisfied with the strong feasible condition with the measure parameter $\xi^*-\varepsilon$ and $W_{\frac{\varepsilon_{x,r}}{1-\varepsilon_{x,r}}+\varepsilon} \ge W_{\tau_1+\varepsilon}$. Therefore, we have that

$$\begin{aligned}
&P\left(\sum_{j=1}^{t_r} Y_j^{\widetilde{P}_1} \le \left(1-(2+\frac{1}{\xi^*-\varepsilon})\varepsilon_{x,r}\right)\frac{t_r}{T}W_\varepsilon\right) + \sum_{k\in\mathcal{K}} P\left(\sum_{j=1}^{t_r} X_{jk}^{\widetilde{P}_1} \notin [\frac{t_r}{T}L_k, \frac{t_r}{T}U_k]\right)\\
&\le (2K+1)\exp(-\frac{t_r\varepsilon_{x,r}^2}{4T\gamma_1})\\
&\le \delta,
\end{aligned} \tag{39}$$

which means that $W^r \ge (1-(2+\frac{1}{\xi^*-\varepsilon})\varepsilon_{x,r})\frac{t_r}{T}W_\tau$, w.p. $1-\delta$. This completes the proof. ☐

G.2. Proof with ξ^* but without the Knowledge of Distribution

Theorem 4.5. *Under Assumption 1-3, if* $\varepsilon > 0$ $\tau_1 = \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}$ *such that* $\tau_1 + \varepsilon \le \xi^*$ *and* $\gamma_1 = \max\left(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{(1-\varepsilon)T\bar{a}_k - L_k}, \frac{\bar{w}}{W_{\varepsilon+\tau_1}}\right) = O\left(\frac{\varepsilon^2}{\ln(K/\varepsilon)}\right)$, *Algorithm* A_1 *defined in Algorithm 4 achieves an objective value of at least* $\left(1-O(\frac{\varepsilon}{\xi^*-\varepsilon})\right)W_0$ *and satisfies the constraints w.p.* $1-\varepsilon$.

Now, we prove that, at each stage r, Algorithm A_1 returns a solution whose cumulative revenue is at least $\frac{t_r Z^r}{T}(1-\varepsilon_{y,r})$. Meanwhile, the consumed amount of every resource k is between $\frac{t_r}{T}(L_k + (\varepsilon - \frac{\varepsilon_{x,r}}{1+\varepsilon_{x,r}})T\bar{a}_k)(1+\varepsilon_{x,r})$ and $\frac{t_r U_k}{T}(1+\varepsilon_{x,r})$ with probability at least $1-\delta$.

First step: We design a surrogate Algorithm $\widetilde{P}_2$ that allocates request j to channel i with probability $x(\varepsilon)_{ij}^*$.

Lemma G.3. *In the* r*-th stage, if* $\gamma_1 = O(\frac{\varepsilon^2}{\ln(\frac{K}{\varepsilon})})$ *and* $Z^r \le W_\varepsilon$, *the Algorithm* $\widetilde{P}_2$ *returns a solution satisfying the* $\sum_{j=t_r+1}^{t_{r+1}} Y_j^{\widetilde{P}_2} \ge (1-\varepsilon_{y,r})\frac{t_r}{n}Z^r$ *and* $\sum_{j=t_r+1}^{t_{r+1}} X_j^{\widetilde{P}_2} \in \left[\frac{t_r}{T}\Big((1+\varepsilon_{x,r})L_k + (\varepsilon(1+\varepsilon_{x,r})-\varepsilon_{x,r})T\bar{a}_k\Big), \frac{(1+\varepsilon_{x,r})t_r}{T}U_k\right]$ *w.p.* $1-\delta$, *where* $\delta = \frac{\varepsilon}{3l}$, $\varepsilon_{x,r} = \sqrt{\frac{4T\gamma_1\ln(\frac{2K+1}{\delta})}{t_r}}$ *and* $\varepsilon_{y,r} = \sqrt{\frac{4T\ln(\frac{2K+1}{\delta})\bar{w}}{Z_r t_r}}$.

Proof. Following the same technique in the proof of Lemma 4.1, we use Bernstein inequality to bound the probability of upper bound violation as follows

$$
\begin{aligned}
&P\left(\sum_{j=t_r+1}^{t_{r+1}} X_{jk}^{\widetilde{P}_2} \geq (1+\varepsilon_{x,r})\frac{t_r}{T}U_k\right)\\
&\leq \exp\left(-\frac{\frac{t_r}{T}\varepsilon_{x,r}^2}{2(1+\frac{\varepsilon_{x,r}}{3})\frac{\bar{a}_k}{U_k}}\right)\\
&\leq \exp\left(-\frac{\frac{t_r}{T}\varepsilon_{x,r}^2}{2(1+\frac{\varepsilon_{x,r}}{3})\gamma_1}\right)\\
&\leq \frac{\delta}{2K+1}.
\end{aligned}
$$

For the lower bound,

$$
\begin{aligned}
&P\left(\sum_{j=1+t_r}^{t_{r+1}} X_{jk}^{\widetilde{P}_2} \leq \frac{t_r}{T}(L_k + (\varepsilon - \frac{\varepsilon_{x,r}}{1+\varepsilon_{x,r}})T\bar{a}_k)(1+\varepsilon_{x,r})\right)\\
&= P\left(\sum_{j=1+t_r}^{t_{r+1}} (\bar{a}_k - X_{jk}^{\widetilde{P}_2}) \geq (1+\varepsilon_{x,r})\frac{t_r}{T}((1-\varepsilon)T\bar{a}_k - L_k)\right)\\
&\leq \exp\left(-\frac{\frac{t_r}{T}\varepsilon_{x,r}^2}{2(1+\frac{1}{3}\varepsilon_{x,r})\frac{\bar{a}_k}{(1-\varepsilon)T\bar{a}_k - L_k}}\right)\\
&\leq \exp\left(-\frac{\frac{t_r}{T}\varepsilon_{x,r}^2}{2(1+\frac{\varepsilon_{x,r}}{3})\gamma_1}\right)\\
&\leq \frac{\delta}{2K+1}.
\end{aligned}
$$

For the accumulative revenue in r-th stage,

$$
\begin{aligned}
&P\left(\sum_{j=t_r+1}^{t_{r+1}} Y_j^{\widetilde{P}_2} \leq (1-\varepsilon_{y,r})\frac{t_r}{T}Z^r\right)\\
&= P\left(\sum_{j=t_r+1}^{t_{r+1}} \left(\mathbb{E}[Y_j^{\widetilde{P}_2}] - Y_j^{\widetilde{P}_2}\right) \geq \frac{t_r}{T}\left(T\mathbb{E}[Y_j^{\widetilde{P}_2}] - (1-\varepsilon_{y,r})Z^r\right)\right)\\
&\leq \exp\left(-\frac{\left(T\mathbb{E}[Y_j^{\widetilde{P}_2}] - (1-\varepsilon_{y,r})Z^r\right)^2 t_r}{T\left(2T\sigma_1^2 + \frac{2}{3}\bar{w}\left(T\mathbb{E}[Y_j^{\widetilde{P}_2}] - (1-\varepsilon_{y,r})Z^r\right)\right)}\right)\\
&= \exp\left(-\frac{\frac{t_r}{T}\left(T\mathbb{E}[Y_j^{\widetilde{P}_2}] - (1-\varepsilon_{y,r})Z^r\right)}{2\frac{T\sigma_1^2}{T\mathbb{E}[Y_j^{\widetilde{P}_2}] - (1-2\varepsilon_{y,r})Z^r} + \frac{2}{3}\bar{w}}\right)\\
&\leq \exp\left(-\frac{t_r\left(\varepsilon_{y,r}T\mathbb{E}[Y_j^{\widetilde{P}_2}]\right)^2}{T\left(2T\mathbb{E}[Y_j^{\widetilde{P}_2}]\bar{w} + \frac{2}{3}\varepsilon_{y,r}\bar{w}T\mathbb{E}[Y_j^{\widetilde{P}_2}]\right)}\right)\\
&\leq \exp\left(-\frac{t_r\varepsilon_{y,r}^2}{2(1+\frac{\varepsilon_{y,r}}{3})T\frac{\bar{w}}{Z^r}}\right)\\
&\leq \frac{\delta}{2K+1}
\end{aligned}
$$

where the second inequality follows $\sigma_1^2 = Var(Y_j^{\widetilde{P}_2}) \leq \bar{w}\mathbb{E}[Y_j^{\widetilde{P}_2}]$ and $T\mathbb{E}[Y_j^{\widetilde{P}_2}] - (1-\varepsilon_{y,r})Z^r \geq \varepsilon_{y,r}T\mathbb{E}[Y_j^{\widetilde{P}_2}]$, the third inequality follows $Z^r \leq W_\varepsilon = T\mathbb{E}[Y_j^{\widetilde{P}_2}]$ according to the condition of the lemma, and the last inequality from $\varepsilon_{y,r} = \sqrt{\frac{4T\ln(\frac{2K+1}{\delta})\bar{w}}{Z^r t_r}}$.

Therefore,

$$\sum_{k\in\mathcal{K}} P\left(\sum_{j=t_r+1}^{t_{r+1}} X_{jk} \notin \left[\frac{t_r}{T}\Big((1+\varepsilon_{x,r})L_k - (\varepsilon(1+\varepsilon_{x,r})-\varepsilon_{x,r})T\bar{a}_k\Big), \frac{(1+\varepsilon_{x,r})t_r}{T}U_r\right]\right)$$
$$+ P\left(\sum_{j=t_r+1}^{t_{r+1}} Y_j^{\widetilde{P}_2} \leq (1-\varepsilon_{y,r})\frac{t_r}{T}Z^r\right) \leq \delta.$$

□

Second Step: Applying the same technique in the proof of Theorem 4.4 , we derive a potential function to bound the failure probability of hybrid Algorithm $A_1^s\widetilde{P}_2^{t_r-s}$ for request in stage r.

We begin with the moment generating function for the event that the consumed resource $k \in \mathcal{K}$ is larger than $(1+\varepsilon_{x,r})\frac{t_r}{T}U_k$. It can be shown that

$$\begin{aligned}
&P\left(\sum_{j=1+t_r}^{s+t_r} X_{jk}^{A_1} + \sum_{j=s+t_r}^{t_{r+1}} X_{jk}^{\widetilde{P}_2} \geq \frac{(1+\varepsilon_{x,r})t_rU_k}{T}\right) \\
&\leq \min_{t>0}\mathbb{E}\left[\exp(t(\sum_{j=1+t_r}^{s+t_r} X_{jk}^{A_1} + \sum_{j=s+t_r}^{t_{r+1}} X_{jk}^{\widetilde{P}_2} - \frac{(1+\varepsilon_{x,r})t_rU_k}{T}))\right] \\
&\leq \min_{t>0}\mathbb{E}\left[\exp(t(\sum_{j=1+t_r}^{s+t_r} X_{jk}^{A_1} - \frac{(1+\varepsilon_{x,r})sU_k}{T}) + t(\sum_{j=s+t_r}^{t_{r+1}} X_{jk}^{\widetilde{P}_2} - \frac{(1+\varepsilon_{x,r})(t_r-s)U_k}{T}))\right] \\
&\leq \min_{t>0}\mathbb{E}\left[\phi_k^s(t)\exp(t(\sum_{j=s+t_r}^{t_{r+1}}(X_{jk}^{A_1} - \mathbb{E}[X_{jk}^{\widetilde{P}_2}])) + \frac{t_r-s}{T}t(T\mathbb{E}[X_{jk}^{\widetilde{P}_2}] - (1+\varepsilon_{x,r})U_k))\right] \\
&\leq \min_{t>0}\mathbb{E}\left[\phi_k^s(t)\exp((t_r-s)\frac{Var(X_{jk}^{\widetilde{P}_2})}{\bar{a}_k^2}(e^{t\bar{a}_k}-1-t\bar{a}_k) + \frac{-(t_r-s)t\varepsilon_{x,r}U_k}{T})\right] \\
&\leq \mathbb{E}\left[\phi_k^s\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp(-\frac{(t_r-s)U_k}{T\bar{a}_k}((1+\eta)\ln(1+\eta)-\eta))\right] \\
&\leq \mathbb{E}\left[\phi_k^s\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp(-\frac{t_r-s}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1})\right]
\end{aligned} \tag{40}$$

where $\phi_k^s(t) = \exp(t(\sum_{j=1+t_r}^{s+t_r} X_{jk}^{A_1} - \frac{(1+\varepsilon_{x,r})sU_k}{T}))$. It should be noted that most of the above analysis is similar as the derivation of inequality (18) except that $Var(X_{jk}^{\widetilde{P}_2}) \leq \bar{a}_k\mathbb{E}[X_{ik}^{\widetilde{P}_2}] \leq \bar{a}_k\frac{U_k}{T}$.

Next for the lower bound, we set $Z_{jk}^{\widetilde{P}_2} = \bar{a}_k - X_{jk}^{\widetilde{P}_2}$ and $Z_{jk}^{A_1} = \bar{a}_k - X_{jk}^{A_1}$, then we have

$$\begin{aligned}
&P\left(\sum_{j=1+t_r}^{s+t_r} Z_{jk}^{A_1} + \sum_{j=s+t_r}^{t_{r+1}} Z_{jk}^{\widetilde{P}_2} \geq \frac{(1+\varepsilon_{x,r})t_r\Big((1-\varepsilon)T\bar{a}_k - L_k\Big)}{T}\right) \\
&\leq \min_{t>0}\mathbb{E}\left[\exp\left(t\Big(\sum_{j=1+t_r}^{s+t_r} Z_{jk}^{A_1} + \sum_{j=s+t_r}^{t_{r+1}} Z_{jk}^{\widetilde{P}_2} - \frac{(1+\varepsilon_{x,r})t_r\Big((1-\varepsilon)T\bar{a}_k - L_k\Big)}{T}\Big)\right)\right]
\end{aligned} \tag{41}$$

$$
\begin{aligned}
&\le \min_{t>0} \mathbb{E}\left[\exp\left(t\Big(\sum_{j=1+t_r}^{s+t_r} Z_{jk}^{A_1} - \frac{(1+\varepsilon_{x,r})s\Big((1-\varepsilon)T\bar{a}_k - L_k\Big)}{T}\Big)\right.\right.\\
&\qquad\qquad \left.\left. +t\Big(\sum_{j=s+t_r}^{t_{r+1}} Z_{jk}^{\widetilde{P}_2} - \frac{(1+\varepsilon_{x,r})(t_r-s)\Big((1-\varepsilon)T\bar{a}_k - L_k\Big)}{T}\Big)\right)\right]\\
&\le \min_{t>0} \mathbb{E}\left[\varphi_k^s(t)\exp\left(t\Big(\sum_{j=s+t_r}^{t_{r+1}} (Z_{jk}^{\widetilde{P}_2} - \mathbb{E}[Z_{jk}^{\widetilde{P}_2}])\Big) + \frac{t_r-s}{T} t\Big(T\mathbb{E}[Z_{jk}^{\widetilde{P}_2}] - (1+\varepsilon_{x,r})((1-\varepsilon)T\bar{a}_k - L_k)\Big)\right)\right]\\
&\le \min_{t>0} \mathbb{E}\left[\varphi_k^s(t)\exp\left((t_r-s)\frac{\sigma^2}{\bar{a}_k^2}(e^{t\bar{a}_k} - 1 - t\bar{a}_k) + \frac{-(t_r-s)t\varepsilon_{x,r}\Big((1-\varepsilon)T\bar{a}_k - L_k\Big)}{T}\right)\right]\\
&\le \mathbb{E}\left[\varphi_k^s\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(-\frac{(t_r-s)(1-\varepsilon)(T\bar{a}_k - L_k)}{T\bar{a}_k}\Big((1+\eta)\ln(1+\eta) - \eta\Big)\right)\right]\\
&\le \mathbb{E}\left[\varphi_k^s\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)\right]
\end{aligned}
$$

where the third inequality follows from setting $\varphi_k^s(t) = \exp\Big(t(\sum_{j=1+t_r}^{s+t_r} Z_{jk}^{A_1} - \frac{(1+\varepsilon_{x,r})s((1-\varepsilon)T\bar{a}_k - L_k)}{T})\Big)$; the fifth inequality from $T\mathbb{E}[Z_{jk}^{\widetilde{P}_2}] \le (1-\varepsilon)T\bar{a}_k - L_k$ and $\sigma^2 = Var(Z_{jk}^{\widetilde{P}_2}) \le \bar{a}_k\mathbb{E}[Z_{jk}^{\widetilde{P}_2}]$.

Then we consider the revenue and have that

$$
\begin{aligned}
&P(\sum_{j=1+t_r}^{s+t_r} Y_j^{A_1} + \sum_{j=s+1+t_r}^{t_{r+1}} Y_j^{\widetilde{P}_2} \le (1-\varepsilon_{y,r})\frac{t_r}{T}Z^r)\\
&\le \min_{t>0} \mathbb{E}\left[\exp\left(t\Big((1-\varepsilon_{y,r})\frac{t_r}{T}Z^r - \sum_{j=1+t_r}^{s+t_r} Y_j^{A_1} - \sum_{j=s+1+t_r}^{t_{r+1}} Y_j^{\widetilde{P}_2})\Big)\right)\right]\\
&\le \min_{t>0} \mathbb{E}\left[\exp\left(t\Big(\frac{s}{T}(1-\varepsilon_{y,r})Z^r - \sum_{j=1+t_r}^{s+t_r} Y_j^{A_1}\Big) + t\Big(\frac{t_r-s}{T}(1-\varepsilon_{y,r})Z^r - \sum_{j=s+1+t_r}^{t_{r+1}} Y_j^{\widetilde{P}_2}\Big)\right)\right]\\
&\le \min_{t>0} \mathbb{E}\left[\psi^s(t)\exp\left(t\sum_{j=s+1+t_r}^{t_{r+1}}(\mathbb{E}[Y_j^{\widetilde{P}_2}] - Y_j^{\widetilde{P}_2}) + \frac{t_r-s}{T}t\Big((1-\varepsilon_{y,r})Z^r - \mathbb{E}[Y_j^{\widetilde{P}_2}]\Big)\right)\right] \qquad (42)\\
&\le \min_{t>0} \mathbb{E}\left[\psi^s(t)\exp\left((t_r-s)\frac{\sigma_1^2}{\bar{w}^2}(e^{t\bar{w}} - 1 - t\bar{w}) - (t_r-s)t\varepsilon_{y,r}\mathbb{E}[Y_j^{\widetilde{P}_2}]\right)\right]\\
&\le \mathbb{E}\left[\psi^s\left(\frac{\ln(1+\varepsilon_{y,r})}{\bar{w}}\right)\exp\left(-\frac{(t_r-s)\mathbb{E}[Y_j^{\widetilde{P}_2}]}{\bar{w}}\Big((1+\eta)\ln(1+\eta) - \eta\Big)\right)\right]\\
&\le \mathbb{E}\left[\psi^s\left(\frac{\ln(1+\varepsilon_{y,r})}{\bar{w}}\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{y,r}^2 Z^r}{4\bar{w}}\right)\right]
\end{aligned}
$$

where the third inequality follows from $\psi^s(t) = \exp(t(\frac{s}{T}(1-\varepsilon_{y,r})Z^r - \sum_{j=1+t_r}^{s+t_r} Y_j^{A_1}))$; the fourth inequality from $Z^r \le T\mathbb{E}[Y_j^{\widetilde{P}_2}]$; the fifth inequality from $\sigma_1^2 = Var(Y_j^{\widetilde{P}_2}) \le \bar{w}\mathbb{E}[Y_j^{\widetilde{P}_2}]$.

With the inequalities (40)-(42), we can bound the failure probability of hybrid Algorithm $A_1^s\widetilde{P}_2^{t_r-s}$ in stage r by

$\mathcal{F}_r(A_1^s\widetilde{P}_2^{t_r-s})$ which is defined as

$$\begin{aligned}\mathcal{F}_r(A_1^s\widetilde{P}_2^{t_r-s}) =&\mathbb{E}\Big[\phi_k^s\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)+\varphi_k^s\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)\\&+\psi^s\left(\frac{\ln(1+\varepsilon_{y,r})}{\bar{w}}\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{y,r}^2Z^r}{4\bar{w}}\right)\Big]\end{aligned} \tag{43}$$

Lemma G.4. $\mathcal{F}_r(A_1^s\widetilde{P}_2^{t_r-s})\le\mathcal{F}_r(A_1^{s-1}\widetilde{P}_2^{t_r-s+1})$

Proof. By the definition of $\mathcal{F}_r(A^s\widetilde{P}_2^{T-s})$, we have that

$$\begin{aligned}&\mathcal{F}_r(A_1^s\widetilde{P}_2^{t_r-s})\\&=\mathbb{E}\Big[\phi_k^s\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)+\varphi_k^s\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)\\&\quad+\psi^s\left(\frac{\ln(1+\varepsilon_{y,r})}{\bar{w}}\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{y,r}^2Z^r}{4\bar{w}}\right)\Big]\\&=\mathbb{E}\Big[\phi_k^{s-1}\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\Big(X_{sk}^{A_1}-\frac{(1+\varepsilon_{x,r})U_k}{T}\Big)\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)\\&\quad+\varphi_k^{s-1}\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\Big(Z_{sk}^{A_1}-\frac{(1+\varepsilon_{x,r})((1-\varepsilon)T\bar{a}_k-L_k)}{T}\Big)\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)\\&\quad+\psi^{s-1}\left(\frac{\ln(1+\varepsilon_{y,r})}{\bar{w}}\right)\exp\left(\frac{\ln(1+\varepsilon_{y,r})}{\bar{w}}\Big(\frac{(1-\varepsilon_{y,r})Z^r}{T}-Y_s^{A_1}\Big)\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{y,r}^2Z^r}{4\bar{w}}\right)\Big]\end{aligned}$$

According to algorithm A in Algorithm 2, we allocate the s-th request to the channel i^* which minimize the $\mathcal{F}_r(A_1^s\widetilde{P}_2^{t_r-s})$. Thus we have

$$\begin{aligned}&\mathcal{F}_r(A_1^s\widetilde{P}_2^{t_r-s})\\&\le\mathbb{E}\Big[\phi_k^{s-1}\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\Big(X_{sk}^{\widetilde{P}_2}-\frac{(1+\varepsilon_{x,r})U_k}{T}\Big)\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)\\&\quad+\varphi_k^{s-1}\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\Big(Z_{sk}^{\widetilde{P}_2}-\frac{(1+\varepsilon_{x,r})((1-\varepsilon)T\bar{a}_k-L_k)}{T}\Big)\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)\\&\quad+\psi^{s-1}\left(\frac{\ln(1+\varepsilon_{y,r})}{\bar{w}}\right)\exp\left(\frac{\ln(1+\varepsilon_{y,r})}{\bar{w}}\Big(\frac{(1-\varepsilon_{y,r})Z^r}{T}-Y_s^{\widetilde{P}_2}\Big)\right)\exp\left(-\frac{t_r-s}{T}\frac{\varepsilon_{y,r}^2Z^r}{4\bar{w}}\right)\Big]\end{aligned}$$

Following the similar analysis in the inequality (40)-(42), we can show that

$$\begin{aligned}\mathcal{F}_r(A_1^s\widetilde{P}_2^{t_r-s})\le&\mathbb{E}\Big[\phi_k^{s-1}\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(-\frac{t_r-s+1}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)\\&+\varphi_k^{s-1}\left(\frac{\ln(1+\varepsilon_{x,r})}{\bar{a}_k}\right)\exp\left(-\frac{t_r-s+1}{T}\frac{\varepsilon_{x,r}^2}{4\gamma_1}\right)\\&+\psi^{s-1}\left(\frac{\ln(1+\varepsilon_{y,r})}{\bar{w}}\right)\exp\left(-\frac{t_r-s+1}{T}\frac{\varepsilon_{y,r}^2Z^r}{4\bar{w}}\right)\Big]\\\le&\mathcal{F}_r(A_1^{s-1}\widetilde{P}_2^{t_r-s+1}),\end{aligned}$$

which completes the proof. □

In Lemma G.3, we have proven that $\mathcal{F}_r(\widetilde{P}_2^{t_r})\le\delta$ and we will show that $\mathcal{F}_r(A_1^s\widetilde{P}_2^{t_r-s})\le\mathcal{F}_r(A_1^{s-1}\widetilde{P}_2^{t_r-s+1})$ in Lemma G.4.

Thus we have $\mathcal{F}_r(A_1^{t_r}) \leq \mathcal{F}_r(\widetilde{P}_2^{t_r}) \leq \delta$ by induction. Meanwhile, according to Theorem G.2 we have that

$$(1-(4+\frac{2}{\xi^*-\varepsilon})\varepsilon_{x,r-1})W_\varepsilon \leq Z_r \leq W_\varepsilon$$

with probability $1-2\delta$. During the stage r, the Algorithm A_1 return a solution satisfying

$$\sum_{j=t_r+1}^{t_{r+1}} X_{jk}^{A_1} \leq \frac{(1+\varepsilon_{x,r})t_r}{T}U_r, \forall k \in \mathcal{K}$$

$$\sum_{j=t_r+1}^{t_{r+1}} (\bar{a}_k - X_{jk}^{A_1}) \leq \frac{(1+\varepsilon_{x,r})t_r}{T}((1-\varepsilon)T\bar{a}_k - L_k), \forall k \in \mathcal{K}$$

$$\sum_{j=t_r+1}^{t_{r+1}} Y_j^{A_1} \geq (1-\varepsilon_{y,r})\frac{t_r}{T}Z^r \geq (1-\varepsilon_{y,r})\frac{t_r}{T}(1-(4+\frac{2}{\xi^*-\varepsilon})\varepsilon_{x,r-1})W_\varepsilon$$

with probability at least $1-3\delta$, since

$$P\Bigg(\Big\{\sum_{j=t_r+1}^{t_{r+1}} X_{jk}^{A_1} \in \Big[\frac{t_r}{T}\Big((1+\varepsilon_{x,r})L_k - (\varepsilon(1+\varepsilon_{x,r})-\varepsilon_{x,r})T\bar{a}_k\Big), \frac{(1+\varepsilon_{x,r})t_r}{T}U_r\Big], \forall k \in \mathcal{K}\Big\}$$
$$\bigcap\Big\{\sum_{j=t_r+1}^{t_{r+1}} Y_j^{A_1} \geq (1-\varepsilon_{y,r})\frac{t_r}{T}Z^r\Big\}\bigcap\Big\{Z^r \in \Big[\frac{t_rW_\varepsilon}{T}\Big(1-(4+\frac{1}{\xi^*-\varepsilon})\varepsilon_{x,r}\Big), \frac{t_rW_\varepsilon}{T}\Big(1+(\frac{1}{\xi^*-\varepsilon})\varepsilon_{x,r}\Big)\Big]\Big\}\Bigg)$$
$$\geq (1-\delta)(1-2\delta) \geq 1-3\delta.$$

Now considering all the stages, for the upper bound, we have

$$\sum_{r=0}^{l-1}\sum_{j=t_r+1}^{t_{r+1}} X_{jk}^{A_1} \leq \sum_{r=0}^{l-1}\frac{(1+\varepsilon_{x,r})t_r}{T}U_k \leq U_k. \tag{44}$$

For the lower bound, we have

$$\begin{aligned}(1-\varepsilon)T\bar{a}_k - \sum_{r=0}^{l-1}\sum_{j=t_r+1}^{t_{r+1}} X_{jk}^{A_1} &= \sum_{r=0}^{l-1}\sum_{j=t_r+1}^{t_{r+1}} (\bar{a}_k - X_{jk}^{A_1}) \\ &\leq \sum_{r=0}^{l-1}\frac{(1+\varepsilon_{x,r})t_r}{T}((1-\varepsilon)T\bar{a}_k - L_k) \\ &\leq (1-\varepsilon)T\bar{a}_k - L_k.\end{aligned} \tag{45}$$

which is equivalent to

$$\sum_{r=0}^{l-1}\sum_{j=t_r+1}^{t_{r+1}} X_{jk}^{A_1} \geq L_k.$$

And for the revenue, we have

$$\begin{aligned}\sum_{r=0}^{l-1}\sum_{j=t_r+1}^{t_{r+1}} Y_j^{A_1} &\geq \sum_{r=0}^{l-1}(1-\varepsilon_{y,r})\frac{t_r(1-(4+\frac{2}{\xi^*-\varepsilon})\varepsilon_{x,r-1}}{T})W_\varepsilon \\ &\geq \sum_{r=0}^{l-1}(1-\varepsilon_{y,r})\frac{t_r(1-(4+\frac{2}{\xi^*-\varepsilon})\varepsilon_{x,r-1})}{T}(1-\frac{\varepsilon}{\xi^*})W_0 \\ &\geq (1-O(\frac{\varepsilon}{\xi^*-\varepsilon}))W_0.\end{aligned} \tag{46}$$

since $\delta = \frac{\varepsilon}{3l}$, the inequalities (44)-(46) hold with probability at least $1-\varepsilon$, which completes the proof of Theorem 4.5.

H. Proof of Theorem 5.2

Theorem 5.2. *Under Assumption 1-3, if* $\gamma_2 = \max(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{T\bar{a}_k - L_k}) = O(\frac{\epsilon^2}{\ln(K/\varepsilon)})$*, Algorithm 5 with* t_r *i.i.d. requests outputs* $\widehat{\xi}$ *such that*

$$\widehat{\xi} \in [\xi^* - 4\epsilon_{x,r}, \xi^*]$$

w.p. $1 - 2\delta$*, where* $\epsilon_{x,r} = \sqrt{\frac{4\gamma_2 T \ln(K/\delta)}{t_r}}$.

Proof. **RHS:** This side takes the same techniques as in Lemma G.1. First, the dual of LP (3) in Section 3.1 is

$$\begin{aligned}
\min_{\alpha,\beta,\rho} \ & \sum_{k\in\mathcal{K}} \alpha_k U_k - \sum_{k\in\mathcal{K}} \beta_k L_k + \sum_{j\in\mathcal{J}} T p_j \rho_j \\
s.t. \ & \sum_{k\in\mathcal{K}} (\alpha_k - \beta_k) a_{ijk} + \rho_j \geq 0 \ \forall i \in \mathcal{I}, j \in \mathcal{J} \\
& \sum_{k\in\mathcal{K}} T\bar{a}_k \beta_k = 1 \\
& \alpha_k, \beta_k, \rho_j \geq 0, k \in \mathcal{K}, j \in \mathcal{J}.
\end{aligned} \tag{47}$$

We denote the optimal solution of LP (3) in Section 3.1 and LP (47) as (x^*_{ij}, ξ^*) and $(\alpha^*_k, \beta^*_k, \rho^*_k)$ respectively.

According to the KKT conditions(Boyd et al., 2004), we have that

$$\begin{aligned}
& \sum_{k\in\mathcal{K}} (\alpha^*_k - \beta^*_k) a_{ijk} x^*_{ij} + \rho^*_j x^*_{ij} = 0 \\
& \sum_{k\in\mathcal{K}} T\bar{a}_k \beta^*_k = 1 \\
& \rho^*_j (\sum_{i\in\mathcal{I}} x^*_{ij} - 1) = 0 \\
& \alpha^*_k (\sum_{ij} T p_j a_{ijk} x^*_{ij} - U_k) = 0 \\
& \beta^*_k (L_k + \xi^* T\bar{a}_k - \sum_{ij} T p_j a_{ijk} x^*_{ij}) = 0
\end{aligned} \tag{48}$$

Similarly, the dual of sampled LP (7) in Algorithm 5 is

$$\begin{aligned}
\min_{\alpha,\beta,\rho} \ & \sum_{k\in\mathcal{K}} \alpha_k \frac{t_r}{T} U_k - \sum_{k\in\mathcal{K}} \beta_k \frac{t_r}{T} L_k + \sum_{j\in\mathcal{S}_r} \rho_j \\
s.t. \ & \sum_{k\in\mathcal{K}} (\alpha_k - \beta_k) a_{ijk} + \rho_j \geq 0 \ \forall i \in \mathcal{I}, j \in \mathcal{S}_r \\
& \sum_{k\in\mathcal{K}} t_r \bar{a}_k \beta_k = 1 \\
& \alpha_k, \beta_k, \rho_j \geq 0, k \in \mathcal{K}, j \in \mathcal{S}_r
\end{aligned} \tag{49}$$

where $\mathcal{S}_r$ denotes the request set in stage r. Since $(\alpha^*_k, \beta^*_k, \rho^*_k)$ is a feasible solution to the LP (47), the solution

$(\frac{T\alpha_k^*}{t_r}, \frac{T\beta_k^*}{t_r}, \frac{T\rho_k^*}{t_r})$ is feasible for the dual of sample LP (49), we have that

$$\begin{aligned}\widehat{\xi} + 2\epsilon_{x,r} =& \frac{T}{t_r}\left(\sum_{k\in\mathcal{K}} \alpha_k^* \frac{t_r}{T} U_k - \sum_{k\in\mathcal{K}} \beta_k^* \frac{t_r}{T} L_k + \sum_{j\in\mathcal{S}_r} \rho_j^*\right) \\ =& \underbrace{\frac{T}{t_r}\sum_{k\in\mathcal{K}} \alpha_k^* \left(\frac{t_r}{T}U_k - \sum_{j\in\mathcal{S}_r, i\in\mathcal{I}} a_{ijk}x_{ij}^*\right)}_{①} + \underbrace{\frac{T}{t_r}\sum_{k\in\mathcal{K}} \beta_k^* \left(\sum_{j\in\mathcal{S}_r, i\in\mathcal{I}} a_{ijk}x_{ij}^* - \frac{t_r}{T}L_k\right)}_{②} \\ &+ \underbrace{\frac{T}{t_r}\sum_{j\in\mathcal{S}_r}\left(\rho_j^* + \sum_{i\in\mathcal{I},k\in\mathcal{K}} (\alpha_k^* - \beta_k^*) a_{ijk}x_{ij}^*\right)}_{③} \\ =& \underbrace{\frac{T}{t_r}\sum_{k\in\mathcal{K}} \alpha_k^* \left(\frac{t_r}{T}U_k - \sum_{j\in\mathcal{S}_r, i\in\mathcal{I}} a_{ijk}x_{ij}^*\right)}_{①} + \underbrace{\frac{T}{t_r}\sum_{k\in\mathcal{K}} \beta_k^* \left(\sum_{j\in\mathcal{S}_r, i\in\mathcal{I}} a_{ijk}x_{ij}^* - \frac{t_r}{T}L_k\right)}_{②}\end{aligned} \tag{50}$$

where the final equality follows from the KKT conditions (48), i.e. $\sum_{k\in\mathcal{K}}(\alpha_k^* - \beta_k^*)a_{ijk}x_{ij}^* + \rho_j^* x_{ij}^* = 0$ and $\rho_j^*(\sum_{i\in\mathcal{I}} x_{ij}^* - 1) = 0$, so that $\sum_{i\in\mathcal{I},k\in\mathcal{K}}(\alpha_k^* - \beta_k^*)a_{ijk}x_{ij}^* + \sum_{i\in\mathcal{I}} \rho_j^* x_{ij}^* = \sum_{i\in\mathcal{I},k\in\mathcal{K}}(\alpha_k^* - \beta_k^*)a_{ijk}x_{ij}^* + \rho_j^* = 0$.

For those k such that $L_k + \xi^* T\bar{a}_k < \sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk}x_{ij}^* < U_k$, we know that they have no effect to $\widehat{\xi}$ following the complementary slackness in (48). For part ①, we only consider the resource k making $\sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk}x_{ij}^* = U_k$. By Lemma B.1, it is easy to get that

$$P(\sum_{j\in[t_r],i\in\mathcal{I}} a_{ijk}x_{ij}^* \le (1-\epsilon_{x,r})\frac{t_r}{T}U_k) \le \exp\left(-\frac{\frac{t_r}{T}\epsilon_{x,r}^2}{2(1+\frac{\epsilon_{x,r}}{3})\frac{\bar{a}_k}{U_k}}\right) \tag{51}$$

where $\mathbb{E}(\sum_{i\in\mathcal{I}} a_{ijk}x_{ij}^*) = \frac{U_k}{T}, \forall j \in \mathcal{J}$.

Similarly, for part ②, we only consider the constraints k making $\sum_{i\in\mathcal{I},j\in\mathcal{J}} Tp_j a_{ijk}x_{ij}^* = L_k + \xi^* T\bar{a}_k$. Before that, we redefine the r.v. $Y_{jk} = (1+\xi^*)\bar{a}_k - \sum_{i\in\mathcal{I}} a_{ijk}x_{ij}^*$. Since $\xi^* \in [0,1]$ from Assumption 3, we know that $|Y_{jk}| \le (1+\xi^*)\bar{a}_k$ and $E(Y_{jk}) = \frac{T\bar{a}_k - L_k}{T}$. Therefore, by Lemma B.1,

$$P(\sum_{j\in\mathcal{S}_r} Y_{jk} \le (1-\epsilon_{x,r})\frac{t_r}{T}(T\bar{a}_k - L_k)) \le \exp\left(-\frac{\frac{t_r}{T}\epsilon_{x,r}^2}{2(1+\frac{\epsilon_{x,r}}{3})\frac{(1+\xi^*)\bar{a}_k}{T\bar{a}_k - L_k}}\right) \tag{52}$$

Since $\gamma_2 = \max(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{T\bar{a}_k - L_k}) = O(\frac{\epsilon^2}{\ln(\frac{K}{\varepsilon})})$, and both lower and upper bound are achieved only if $U_k = T\bar{a}_k - L_k$, we have that

$$\begin{aligned}&\sum_{j\in\mathcal{S}_r, i\in\mathcal{I}} a_{ijk}x_{ij}^* \ge (1-\epsilon_{x,r})\frac{t_r}{T}U_k \\ &\sum_{j\in\mathcal{S}_r} ((1+\xi^*)\bar{a}_k - \sum_{i\in\mathcal{I}} a_{ijk}x_{ij}^*) \ge (1-\epsilon_{x,r})\frac{t_r}{T}(T\bar{a}_k - L_k)\end{aligned} \tag{53}$$

w.p. at least $1-\delta$.

Therefore, with probability at least $1-\delta$ we have that

$$
\begin{aligned}
&① + ②\\
&= \frac{T}{t_r}\left(\sum_{k\in\mathcal{K}}\alpha_k^*\Big(\frac{t_r}{T}U_k - \sum_{j\in\mathcal{S}_r, i\in\mathcal{I}} a_{ijk}x_{ij}^*\Big) + \sum_{k\in\mathcal{K}}\beta_k^*\Big(\sum_{j\in\mathcal{S}_r, i\in\mathcal{I}} a_{ijk}x_{ij}^* - \frac{t_r}{T}L_k\Big)\right)\\
&= \frac{T}{t_r}\left(\sum_{k\in\mathcal{K}}\alpha_k^*\Big(\frac{t_r}{T}U_k - \sum_{j\in\mathcal{S}_r, i\in\mathcal{I}} a_{ijk}x_{ij}^*\Big) + \sum_{k\in\mathcal{K}}\beta_k^*\Big(\frac{t_r}{T}(T\bar{a}_k - L_k) - \sum_{j\in\mathcal{S}_r}((1+\xi^*)\bar{a}_k - \sum_{i\in\mathcal{I}} a_{ijk}x_{ij}^*)\Big)\right.\\
&\left.+\sum_{k\in\mathcal{K}}\beta_k^*\xi^* t_r\bar{a}_k\right)\\
&\le \frac{T}{t_r}\left(\epsilon_{x,r}\sum_{k\in\mathcal{K}}\alpha_k^*\frac{t_r}{T}U_k + \epsilon_{x,r}\sum_{k\in\mathcal{K}}\beta_k^*\frac{t_r}{T}(T\bar{a}_k - L_k) + \sum_{k\in\mathcal{K}}\beta_k^*\xi^* t_r\bar{a}_k\right)\\
&= \frac{T}{t_r}\left(\epsilon_{x,r}(\sum_{k\in\mathcal{K}}\alpha_k^*\frac{t_r}{T}U_k - \sum_{k\in\mathcal{K}}\beta_k^*\frac{t_r}{T}L_k) + (\epsilon_{x,r}+\xi^*)\sum_{k\in\mathcal{K}}\beta_k^* t_r\bar{a}_k\right)\\
&\le \frac{T}{t_r}\left(\epsilon_{x,r}\xi^*\frac{t_r}{T} + (\epsilon_{x,r}+\xi^*)\frac{t_r}{T}\right)\\
&= \xi^* + (\xi^*+1)\epsilon_{x,r}\\
&\le \xi^* + 2\epsilon_{x,r}
\end{aligned}
\tag{54}
$$

where the first inequality from $\sum_{j\in\mathcal{S}_r,i\in\mathcal{I}} a_{ijk}x_{ij}^* \ge (1-\epsilon_{x,r})\frac{t_r}{T}U_k$ and $\sum_{j\in\mathcal{S}_r}((1+\xi^*)\bar{a}_k - \sum_{i\in\mathcal{I}} a_{ijk}x_{ij}^*) \ge (1-\epsilon_{x,r})\frac{t_r}{T}(T\bar{a}_k - L_k)$; the second inequality from $\sum_{k\in\mathcal{K}}\beta_k^* T\bar{a}_k = 1$ and $\xi^* = \sum_{k\in\mathcal{K}}\alpha_k^*U_k - \sum_{k\in\mathcal{K}}\beta_k^*L_k + \sum_{j\in\mathcal{J}} Tp_j\rho_j^* \ge \alpha_k^*U_k - \sum_{k\in\mathcal{K}}\beta_k^*L_k$; the last inequality follows from the fact $\xi^*\le 1$.

LHS: We design an algorithm $\widetilde{P}_3$ by allocating request j to channel i with probability $(1-\epsilon_{x,r})x_{ij}^*$, where (x_{ij}^*, ξ^*) is the optimal solution for LP (7). Following the very similar proofs in Section 4.1 and letting $\gamma_2 = \max(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{T\bar{a}_k - L_k})$, $\varepsilon_{x,r} = \sqrt{\frac{4\gamma_2 T\ln(K/\delta)}{t_r}}$, we have that

$$P\Big(\sum_{j=1}^{t_r} X_{jk}^{\widetilde{P}_3} \ge \frac{t_r}{T}U_k\Big) \le \exp\left(-\frac{t_r\epsilon_{x,r}^2/T}{2(1-\frac{2}{3}\epsilon_{x,r})\frac{\bar{a}_k}{U_k}}\right) \le \frac{\delta}{2K}$$

where the second inequality follows from the definition of t_r and $\gamma_2 = O\big(\frac{\epsilon^2}{\ln(K/\varepsilon)}\big)$, which result in $\varepsilon_{x,r} < 1$. Defining $Y_{jk}^{\widetilde{P}_3} = (1-\epsilon_{x,r})(1+\xi^*)\bar{a}_k - X_{jk}^{\widetilde{P}_3}$, we have that $E(Y_{jk}^{\widetilde{P}_3}) \le \frac{(1-\epsilon_{x,r})(T\bar{a}_k - L_k)}{T}$ and $|Y_{jk}| \le (1-\epsilon_{x,r})(1+\xi^*)\bar{a}_k$, since $|X_{jk}^{\widetilde{P}_3}| \le (1-\epsilon_{x,r})\bar{a}_k$. Therefore, we have

$$P\Big(\sum_{j=1}^{t_r} Y_{jk}^{\widetilde{P}_3} \ge \frac{t_r}{T}(T\bar{a}_k - L_k)\Big) \le \exp\left(-\frac{t_r\epsilon_{x,r}^2/T}{2(1-\frac{2}{3}\epsilon_{x,r})\frac{(1-\epsilon_{x,r})(1+\xi^*)\bar{a}_k}{T\bar{a}_k - L_k}}\right) \le \frac{\delta}{2K}$$

Thus, with probability at least $1-\delta$, we could find a solution whose consumed resource for each k is in $\Big[\frac{t_r}{T}(L_k + (\xi^* - 2\epsilon_{x,r})T\bar{a}_k), \frac{t_r}{T}U_k\Big]$ by Algorithm $\widetilde{P}_3$. According to the definition of $\widehat{\xi}$ in Algorithm 5, we have that $\widehat{\xi} \ge \xi^* - 4\epsilon_{x,r}$.

In conclusion, we have that $\xi^* - 4\epsilon_{x,r} \le \widehat{\xi} \le \xi^*$, w.p. $1-2\delta$.

□

Now $\widehat{\xi_r}$ can be viewed as an good estimate for ξ^* from 0, if we have enough data.

I. Proof of Theorem 5.3

Proof. We mainly consider two events, namely,

$$G_1 = \left\{ \sum_{j=1}^{T} Y_j^{A_2} \geq (1 - O(\frac{\varepsilon}{\xi^* - 4\sqrt{\varepsilon} - \varepsilon}))W_0, \sum_{j=1}^{T} X_{jk}^{A_2} \in [L_k, U_k] \forall k \in \mathcal{K} \right\},$$
$$G_2 = \left\{ \xi^* - 4\sqrt{\varepsilon} \leq \widehat{\xi}_0 \leq \xi^* \right\}.$$

Step 1.

When initializing Algorithm 6, we use the first εT incoming requests to estimate the optimal measure of feasibility ξ^*. From the Theorem 5.2, we have that $\mathrm{P}\big(\{\xi^* - 4\sqrt{\varepsilon} \leq \widehat{\xi}_0 \leq \xi^*\}\big) \geq 1 - 2\delta$, choosing $\delta = \frac{\varepsilon}{3l+2}$.

Step 2.

We investigate the conditional event $G_1|G_2$. Under the assumption $\frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}} + 4\sqrt{\varepsilon} + \varepsilon \leq \xi^*$, if $\xi^* - 4\sqrt{\varepsilon} \leq \widehat{\xi}_0$, we have $\widehat{\xi}_0 \geq \xi^* - 4\sqrt{\varepsilon} \geq \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}} + \varepsilon$. Besides, we know that the domain

$$\begin{aligned}
&L_k + \widehat{\xi}_0 T \bar{a}_k \leq \sum_{i \in \mathcal{I}, j \in \mathcal{J}} T p_j a_{ijk} x_{ij} \leq U_k, \forall k \in \mathcal{K} \\
&\sum_{i \in \mathcal{I}} x_{ij} \leq 1, \forall j \in \mathcal{J} \\
&x_{ij} \geq 0, \forall i \in \mathcal{I}, j \in \mathcal{J},
\end{aligned}$$

is feasible under the assumption.

When $\gamma_3 = O\big(\frac{\varepsilon^2}{\ln(K/\varepsilon)}\big)$, according to the Theorem 4.5, we have

$$\mathrm{P}\left(\left\{ \sum_{j=1}^{T} Y_j^{A_2} \geq (1 - O(\frac{\varepsilon}{\widehat{\xi}_0 - \varepsilon}))W_0 \right\} \bigcap \left\{ \sum_{j=1}^{T} X_{jk}^{A_2} \in [L_k, U_k] \right\}\right) \geq 1 - 3l\delta,$$

where $l = \log_2\left(\frac{1}{\varepsilon}\right)$. Due to $\widehat{\xi}_0 \geq \xi^* - 4\sqrt{\varepsilon}$, we also could derive that $\mathrm{P}(G_1|G_2) \geq 1 - 3l\delta$.

Step 3.

Now we can verify that

$$\begin{aligned}
\mathrm{P}(G_1^c) &= \mathrm{P}(G_1^c|G_2)\mathrm{P}(G_2) + \mathrm{P}(G_1^c|G_2^c)\mathrm{P}(G_2^c) \\
&\leq \mathrm{P}(G_1^c|G_2) + \mathrm{P}(G_2^c) \\
&\leq 3l\delta + 2\delta = \epsilon.
\end{aligned}$$

Therefore, $\mathrm{P}(G_1) \geq 1 - \varepsilon$, if $\varepsilon \geq 0$ and $\tau_1 = \frac{\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}$ such that $\tau_1 + 4\sqrt{\varepsilon} + \varepsilon \leq \xi^*$ and $\gamma_3 = O(\frac{\varepsilon^2}{\ln(\frac{K}{\varepsilon})}) \leq \max(\frac{\bar{a}_k}{U_k}, \frac{\bar{a}_k}{(1-\epsilon)T\bar{a}_k - L_k}, \frac{\bar{w}}{W_{\epsilon+\tau_1}})$.

□