Online Resource Allocation with Samples

Negin Golrezaei
Sloan School of Management, Massachusetts Institute of Technology, golrezaei@mit.edu

Patrick Jaillet
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, jaillet@mit.edu

Zijie Zhou
Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, 02139, zhou98@mit.edu

The problem of online allocation of scarce resources is one of the most important problems faced by governments (e.g., during a pandemic crisis), hospitals, e-commerce, etc. We study an online resource allocation problem where we have uncertainty about demand and the reward of each type of demand (agents) for resources. While dealing with demand uncertainty in resource allocation problems has been the topic of many papers in the literature, the challenge of not knowing rewards has been barely explored. The lack of knowledge about agents’ rewards is inspired by the problem of allocating new resources (e.g., newly developed vaccines or drugs) with unknown effectiveness/value. For such settings, we assume that we have the ability to test the market before the allocation period starts. During the test period, we collect some limited information, called sample information, about agents’ expected rewards, as well as, the size of the market for each type of agent. We study how to optimally exploit the sample information in our online resource allocation problem under adversarial arrival processes. We present an asymptotically optimal protection level algorithm that achieves $1 - \Theta(1/\sqrt{m})$ competitive ratio, where $m$ is the number of resources. Via presenting an upper bound on the competitive ratio of any randomized and deterministic algorithms, we show that our competitive ratio of $1 - \Theta(1/\sqrt{m})$ is tight. We further demonstrate the efficacy of our proposed algorithm using a dataset that contains the number of COVID-19 related hospitalized patients across different age groups.

1. Introduction

In online resource allocation problems, the goal is to allocate a limited number of resources to heterogeneous demand/agents that arrive over time. These problems are notorious to be challenging mainly because of the demand uncertainty and having a limited number of resources. These problems, however, get even more challenging for newly developed resources (e.g., new drugs and products). For such resources, the effectiveness/value of resources for different types of agents may not be fully known. To overcome this additional challenge, businesses, for example, offer free products in an exchange for honest feedback [productreviewmom.com 2022], and pharmaceutical companies test potential treatments/drugs in human volunteers [Pfizer 2022]. These practices raise the following key question: can and to what extent such feedback improve the efficiency of resource allocation?
To answer this question, we consider a decision-maker who aims to allocate her resources to two types of unit-demand agents with unknown (expected) rewards, where type 1 has a higher expected reward than type 2. The total number of agents (i.e., the market size), as well as, the number of agents of type 1 and 2, denoted by $h$ and $\ell$ respectively, are chosen adversarially, and hence are unknown to the decision-maker. Before the allocation period starts, the decision-maker tests the market by, for example, making a public announcement and offering resources for free. We assume that with probability $p \in (0, 1)$, each of the $h + \ell$ agents sees and reacts to the announcement, and gets one unit of the resource. These agents then provide feedback on their realized reward for the resource in return. This test procedure supplies the decision-maker with some information about agents’ expected rewards, as well as, the size of the market for each type of agent. We refer to this information as sample information.

After the test period ends, the remaining agents arrive over time according to an adversarially-chosen order. For each arriving agent, the decision-maker has to make an irrevocable decision about accepting them and allocating them one unit of the resource or rejecting them. The decision-maker who has $m$ units of the resource when the allocation period starts makes acceptance/rejection decisions while being uncertain about the number, type of future agents, and their expected rewards. The decision-maker is also uncertain about which types of agents earn higher rewards upon receiving the resource. For such a demanding setting, our goal is to design efficient resource allocation algorithms that can optimally utilize the sample information under any possible arrival sequence. In other words, we measure the performance of algorithms in terms of their competitive ratio, which is the expected ratio of the reward of the algorithm to the reward of the optimal clairvoyant algorithm that knows the arrival sequence in advance, where the expectation is with respect to the sample information; see Section 2 for the formal definition of competitive ratio.

Before presenting our contributions, we make two remarks about our model. First, our model bears resemblance to the proposed models in Correa et al. (2021), Kaplan et al. (2022) for secretary and online bipartite matching problems, respectively. In Correa et al. (2021), each of the secretaries is placed in a sample set with probability $p$, where the value of the sampled secretaries will be disclosed to the decision-maker before the decision period starts. In Kaplan et al. (2022) that studies an online bipartite matching problem, each agent will be placed in a sample with probability $p$ independently. While at a high level, these works seek to design algorithms that can optimally take advantage of samples, the nature of their considered problems is different from ours, preventing us to use their designed algorithms for our setting. Second, our model is a special case of the single-leg revenue management problem, which has been widely studied in the literature; see, e.g., Littlewood (1972), Amaruchkul et al. (2007), Ball and Queyranne (2009), Gallego et al. (2009), Ferreira et al. (2018), Jasin (2015), Hwang et al. (2021), Golrezaei and Yao (2021). In all of these aforementioned
works, while the decision-maker may be uncertain about the demand (i.e., the number and the order of the arrivals), the obtained reward of different types of demand upon receiving the resource is fully known to the decision-maker. This is in sharp contrast with our model in which these rewards are not known to the decision-maker, adding extra complexity to our problem.

1.1. Our Contributions
In addition to our modeling contribution, our work makes the following contributions.

**Importance of sample information.** We demonstrate the importance of using sample information by presenting a naive protection level algorithm that does not use the sample information at all; see Algorithm [1]. The algorithm does not know which type of agents has a higher expected reward. With probability $1/2$, it assumes that type 1 agent has a higher reward and aims to protect this type of agent by assigning them a protection level. That is, the algorithm keeps accepting any arriving agents until the remaining number of resources hits the protection level $x \in [0, m]$. After that, only type 1 agents get accepted. On the other hand, with probability $1/2$, the algorithm assumes that type 2 has a higher reward, and hence it aims to protect this type of agent by assigning them a protection level of $x$. We show that the competitive ratio of this naive algorithm cannot exceed $1/(2 - r_2/r_1)$, regardless of how large the initial number of resources $m$ is. (See Theorem [1].) Here, $r_i$ is the expected reward of type $i \in \{1, 2\}$ and $r_1 > r_2$. Put differently, our result shows that as $m$ goes to infinity, the competitive ratio of the naive algorithm is bounded away from 1.

**Asymptotically optimal protection level algorithm.** The failure of the naive algorithm highlights the importance of using the sample information. Driven by that, we design a simple, yet effective protection level algorithm that utilizes the sample information; see Algorithm [2]. The algorithm uses the sample information to obtain an estimate for the expected reward of each type of agents. If the estimated reward of type 1 is greater than that of type 2, the algorithm protects type 1 agents, otherwise type 2 agents will be protected. In each of these cases, the algorithm uses the sample information to estimate the protection level for the type that has a higher estimated average reward.

In Theorem [2] we present the competitive ratio of our proposed algorithm for any finite value of $m$. In addition, in Theorem [3] we show that our algorithm is asymptotically optimal as $m$ goes to infinity. More precisely, we show that the asymptotic competitive ratio of our algorithm is in the order of $1 - \Theta(1/\sqrt{m})$. This result shows that the sample information can be extremely useful to improve the performance of resource allocation algorithms. This is because in the absence of the sample information, and even when the expected rewards of agents are fully known, as shown in [Ball and Queyranne (2009)], the competitive ratio of any algorithm cannot exceed $1/(2 - r_2/r_1)$ no matter how large $m$ is. Here, we show that we can break the barrier of $1/(2 - r_2/r_1)$ by taking
advantage of the sample information in a very challenging setting, where the expected reward of agents is not known to the decision-maker. This is mainly because the sample information can be used to infer some knowledge about the number of agents of different types in the online arrival sequence, allowing the decision-maker to learn some information about the adversarially-chosen arrival sequence.

We now comment on our technical contributions in characterizing the competitive ratio of our algorithm. The proof of Theorem 2 is quite involved and is divided into three main cases, where each case bounds the competitive ratio of the algorithm when the total number of agents of type 1 and 2 falls into a certain region. The most challenging regime is the one in which the total number of type 1 agents (i.e., $h$) is large. In this case, the algorithm may lose reward in three aspects: not protecting type 1 agents, over-protecting type 1 agents, and under-protecting type 1 agents. Recall that in our setting, the decision-maker does not even know which type has the higher reward, and hence our algorithm may wrongfully protect type 2 agents. In addition, because $h$ is large, even when the right type is protected, by over- and under-protecting protected agents, our algorithm can suffer from reward loss. At a high level, we overcome these challenges, by showing that either the algorithm protects the right (high-reward) type with high probability, or the loss of the algorithm is small if it protects the wrong (low-reward) type. We then show that when protecting the right type, thanks to our protection levels, the loss due to over- and under-protection is small.

**Upper bound on the competitive ratio of any deterministic and randomized algorithms.** Our proposed algorithm obtains the worst-case asymptotic competitive ratio of $1 - \Theta(1/\sqrt{m})$. Such a superb performance makes us wonder if it is possible to design an algorithm with even a better asymptotic competitive ratio. In Theorems 4 and 5, we show that even when the decision-maker is fully aware of the expected reward of agents, no deterministic and randomized algorithms can obtain an asymptotic competitive ratio better than $1 - \Theta(1/\sqrt{m})$. To show the upper bound of $1 - \Theta(1/\sqrt{m})$ on the competitive ratio for any deterministic algorithm, in the proof of Theorem 4, we consider a family of arrivals, wherein this family, a very large number of type 2 agents (i.e., the type with a lower reward) arrive first, followed by some number of type 1 agents. Under this family of arrivals, any deterministic algorithms have to decide how many type 2 agents to accept based on the number of type 1 agents in the sample. We show that due to the lack of precise knowledge about the number of type 1 agents, no deterministic algorithm can do better than $1 - \Theta(1/\sqrt{m})$ on the constructed family of arrivals. To show the same result for any randomized algorithms, we first derive a variation of Yao's Lemma that could be of independent interest; see Lemma 14. We then devise a distribution over the family of the arrival sequence that we considered in Theorem 4 and show the upper bound using Lemma 14.
Case study. We perform a case study in Section 5 where using the “Laboratory-Confirmed COVID-19-Associated Hospitalizations” dataset which contains the number of bi-weekly cases of COVID-19-associated hospitalizations in the US from March 7th, 2020 to February 5th, 2022. The dataset is obtained from the following website: [gis.cdc.gov/grasp/covidnet/covid19_5.html](https://gis.cdc.gov/grasp/covidnet/covid19_5.html).

We study how to use our algorithm to allocate limited hospital resources (e.g., a certain medicine) to different types of COVID-19 patients while having access to some sample information. We show that the average competitive ratio of our algorithm in various realistic settings exceed 0.88 and our algorithm substantially outperforms the naive algorithm.

1.2. Other Related Works

In the previous section, we review some of the works on single-leg revenue management and works that study the online algorithm design with sample information. Here, we discuss other streams of works that are related to our paper.

Online algorithm design with machine learned advice. The sample information in our setting provides some information to the decision-maker regarding the adversarially-chosen arrival sequence, allowing us to significantly improve the worst-case guarantee of our algorithm. Improving the worst-case guarantee of online algorithms with the help of extra information has been the topic of the recent literature on algorithm design with machine-learned advice. See, for example, Antoniadis et al. (2020) for using advice on the maximum value of secretaries in the online secretary problems, Lattanzi et al. (2020) for using advice on the weights of jobs in online scheduling problems, and Lykouris and Vassilvitskii (2018) for using the advice in the online caching problem. Our work contributes to this line of work by presenting the first algorithm that optimally exploits rather unstructured advice obtained through the sample information in an online resource allocation problem.

Multi-armed bandits. Our setting is also related to the vast literature on multi-armed bandits; see, for example, Thompson (1933), Auer et al. (2002a,b), Balseiro et al. (2019), Van Parys and Golrezaei (2020), Niaziadeh et al. (2021). In this literature, it is assumed that there are some arms/options with unknown expected reward, and the goal of the decision-maker is to identify the best arm by suffering from minimal regret, where the regret is computed against an algorithm that knows the reward of arms in advance. In our setting, similar to the bandit settings, the rewards of agents are unknown. However, unlike the bandit setting, the algorithm aims to partially learn these rewards via the sample information, rather than the feedback it receives throughout the allocation period. In fact, our upper bound results, which are obtained under the setting when the rewards are fully known, show that using the feedback during the allocation period does not improve the asymptotic competitive ratio of our algorithm. This is because in our setting, the order of arrivals...
and the number of agents of different types are chosen adversarially. Hence, in the worst case, the feedback throughout the allocation periods does not add any value, as the decision-maker is not able to acquire the right feedback at the right time. What feedback the decision-maker can receive is mainly determined by the adversary.

**Online resource allocation.** Our work is also related to the literature on online resource allocation. Devanur and Hayes (2009), Feldman et al. (2010), and Agrawal et al. (2014) study this problem in a stochastic setting with i.i.d. demand arrivals. In these works, the authors use the primal-dual technique to design algorithms with sub-linear regret, where the algorithms aim to learn the optimal dual variables associated with resource constraints. The primal-dual technique is also effective for adversarial demand arrivals even though attaining sub-linear regret in adversarial settings is generally impossible (Mehta et al. 2007, Buchbinder et al. 2007, Golrezaei et al. 2014). See also Balseiro et al. (2020) for a recent work that shows how the primal-dual technique results in well-performing algorithms for various demand processes. While in our work, we do not rely on the primal-dual technique, our work contributes to this literature by presenting a model that—with the help of the sample information—bridges the gap between stochastic and adversarial arrivals, allowing us to bypass the negative results in the adversarial settings.

**Outline.** In Section 2, we present our model and the definition of competitive ratio. In Section 3, we introduce the naive algorithm, as well as, our main algorithm that uses the sample information. In Section 4, we state the asymptotic upper bound among all deterministic and randomized algorithms. Section 5 provides a case study to show how our algorithm performs in realistic settings. Section 6 provides the proof of our main theorem. (The proof of other statements is presented in the appendix.)

2. Model

We consider a decision-maker who would like to allocate (identical) units of a resource to two types of unit-demand agents. The expected reward of agents of type $i \in \{1, 2\}$ upon receiving one unit of the resource is $r_i \in (0, 1)$, where without loss of generality, we assume that $r_1 > r_2$ and define $\alpha = \frac{r_2}{r_1} \in (0, 1)$. The decision-maker, however, is not aware of the expected reward of agents; he does not even know which type of agents attains a higher expected reward. As stated in the introduction, this setting captures scenarios where we would like to allocate new drugs, services, and products to customers. Given the lack of knowledge about the expected rewards, before the allocation period starts, the decision-maker aims to collect some information about the unknown expected rewards of agents during a *test period*.

During the test period, the decision-maker aims to outreach the market by, for example, making a public announcement. Let $h, \ell \geq 0$ be respectively the market size of agents of types 1 and 2, i.e.,
the total number of agents of types 1 and 2 that are interested in the resource. The market size $h$ and $\ell$, which are unknown to the decision-maker, can take any arbitrary values; that is, they are chosen adversarially. We assume that during the test period, with probability $p \in (0, 1)$, each of $h + \ell$ agents sees and responds to the outreach program runs by the decision-maker, where $p$ is known to the decision-maker. The assumption that the sampling probability $p$ is known to the decision-maker is motivated by the fact that the outreach program is designed by the decision-maker himself. Nonetheless, in Section 5, via our case study, we investigate the robustness of our proposed algorithm to the lack of exact knowledge about $p$.

Any agent who responds to the outreach program gets one unit of the resource and reveals her realized reward to the decision-maker. Here, without loss of generality, we assume that the realized reward of type $i \in \{0, 1\}$ agents is either 0 or 1. Let $s_i, i \in \{1, 2\}$, be the (random) number of agents of type $i$ that are reached/sampled during the test period. Note that $s_1$ is drawn from a binomial distribution with parameters $h$ and $p$ (i.e., $s_1 \sim \text{Bin}(h, p)$) and $s_2 \sim \text{Bin}(\ell, p)$. The set of realized rewards of $s_i$ sampled agents is denoted by $\rho_i \in \{0, 1\}^{s_i}$, where $\rho_i = \{\rho_{i,1}, \rho_{i,2}, \ldots, \rho_{i,s_i}\}$ and $\rho_{i,j}, j \in [s_i]$, is drawn from a Bernoulli distribution with the success probability of $r_i$. Throughout the paper, we denote $(s_1, s_2)$ and $(\rho_1, \rho_2)$ with $s$ and $\rho$, respectively. In addition, we refer to $\psi := (s, \rho)$ as the sample information that decision-maker obtains during the test period.

Having described the test period, we are now ready to explain the allocation period. Let $m \geq 2$ be the number of available resources at beginning of the allocation period. During this period, the rest of market, i.e., $h - s_1$ type 1 agents and $\ell - s_2$ type 2 agents arrive one by one over time in an arbitrary order. We denote the number of agents of type $i$ during the allocation period with $n_i$ and we further denote $(n_1, n_2)$ with $n$. Let $z_t \in \{1, 2\}, t = 1, 2, \ldots$, be type of the agent in time period $t$ within the allocation period. The agent type $z_t$ is observable to the decision-maker at the time of the decision (i.e., time period $t$), but given the online nature of the problem during the allocation period, $z_{t'}$ for any $t' > t$ is not observable at time period $t$. We define $I = (z_t)_{t \geq 1}$ as the online arrival sequence and recall that $n_i = |\{t \leq |I| : z_t = i\}|$.

Upon the arrival of the agent of type $z_t$ in time period $t$, the decision-maker has to make an irrevocable acceptance/rejection decision regarding that agent. If the decision-maker accepts the agent, he allocates the agent one unit of the resource. Otherwise, no resource will not be allocated to the agent, and that agent will not come back. If the agent gets accepted, she reveals her realized reward to the decision-maker. That being said, as it becomes more clear in Section 3, our algorithm does not use this additional information, and despite that, manages to have an asymptotically optimal competitive ratio. (See the definition of our competitive ratio below.) We refer readers to our earlier discussion about this in Section 1.2, where we review works on multi-armed bandits.
The goal of the decision-maker during the allocation period is to yield high total rewards while being uncertain about the number and the order of agents, as well as, their expected rewards. For an algorithm $A$, online arrival sequence $I$, and the sample information, $\psi = (s, \rho)$, let $\text{REW}_A(I, \psi)$ be the cumulative expected reward of algorithm $A$ across all the time periods. Further, let $\text{OPT}(I)$ be the optimal clairvoyant cumulative expected reward that can be obtained from $I$ using $m$ units of the resource. We measure the performance of an algorithm $A$ using the following competitive ratio (CR) definition, which compares our algorithm to the optimal clairvoyant benchmark:

$$CR_A = \inf_{(h, \ell)} \mathbb{E}_\psi \left[ \inf_I \frac{\mathbb{E}[\text{REW}_A(I, \psi)]}{\text{OPT}(I)} \right].$$

Here, the inner expectation is with respect to (w.r.t.) any randomness in algorithm $A$, and the outer expectation is w.r.t. the arrival sequence $I$ and the sample information $\psi$. Note that given $(h, \ell)$, the arrival sequence $I$ and more precisely, the number of agents of type $i \in \{1, 2\}$ in $I$ (denoted by $n_i$), is random. Further, observe that in our definition of competitive ratio, we take infimum over (i) the size of the market, i.e., $h, \ell$, and (ii) the order of arrivals in $I$.

3. Online Resource Allocation Algorithm with Samples

In this section, we present an online resource allocation algorithm that is asymptotically optimal as the number of resources $m$ goes to infinity; that is, the competitive ratio of our algorithm converges to one as $m$ goes to one. Before presenting our algorithm, to shed light on the importance of using the sample information $\psi = (s, \rho)$, we present a naive algorithm that ignores the sample information. We show that, unlike our algorithm, the competitive ratio of naive algorithm does not go to one as $m$ goes to infinity.

3.1. Naive Protection Level Algorithm

In our setting, the decision-maker is not aware the expected reward of type $i \in \{1, 2\}$ agents, i.e., $r_i$. The decision-maker can only learn about the expected reward of type $i \in \{1, 2\}$ agents through the sample information. In the naive algorithm (Algorithm 1), however, the sample information is completely overlooked. With a probability of $1/2$, the algorithm assumes that type 1 is the type with a higher expected reward and assigns a protection level of $x$ to type 1, where $x \in \{0, 1, 2, \ldots, m\}$ is the input to the algorithm. That is, the algorithm keeps accepting any type of agents until the remaining number of resources is $x$. After that, the algorithm only accepts type 1 agents. Similarly, with a probability of $1/2$, the algorithm assumes that type 2 agents has a higher expected reward than type 1 agents and then gives a protection level of $x$ to type 2 agents. See Algorithm 1.

In the following theorem, we show that for any protection level $x \in \{0, 1, \ldots, m\}$, the competitive ratio of the naive algorithm is at most $\frac{1}{2-\alpha}$ and can be as small as 0, where $\alpha = r_2/r_1$. We note that the upper bound on the competitive ratio of the naive algorithm does not depend on $m$ and more...
Algorithm 1 Naive Algorithm

Input: The number of resources $m$ and protection level $x \in \{0, 1, \ldots, m\}$.

1. With probability $1/2$, we set the protection level $x$ for type 1. In this case, we keep accepting any agents until we have accepted $m - x$ agents. After that, we only accept type 1 agents.

2. With probability $1/2$, set the protection level $x$ for type 2. In this case, we keep accepting any agents until we have accepted $m - x$ agents. After that, we reject all arriving type 1 agents.

Importantly does not go to one as the number of resources goes to one. We further highlight that as shown in [Ball and Queyranne (2009)], $\frac{1}{2} - \alpha$ is the best competitive ratio of a resource allocation algorithm in a setting where the expected rewards $r_i, i \in \{1, 2\}$ are known to the decision-maker.

In the setting studied in [Ball and Queyranne (2009)], however, the decision-maker does not obtain any information regarding the number of type $i$ agents in the arrival sequence, i.e., $n_i$.

Theorem 1 (Competitive Ratio of Naive Protection Level Algorithm). Consider the model presented in Section 2, where the expected rewards of $r_i, i \in \{1, 2\}$ is unknown to the decision-maker. Then, for any $m > 0$, the competitive ratio of the naive algorithm (Algorithm 1) is at most $\frac{1}{2} - \alpha$, where $\alpha = \frac{r_2}{r_1}$.

To show Theorem 1, we consider two choices for the total number of type 1 and type 2 agents: $(h, \ell)$ and $(\hat{h}, \hat{\ell})$. Here, $h = 0$, and $\ell \gg m$ while $\hat{h} \gg m$, $\hat{\ell} \gg \ell$. We define two random arrival sequences $I_1$ and $I_2$, where in both sequences, type 2 agents arrive first, followed by type 1 agents. The number of type 1 and type 2 agents in sequence $I_1$ is $n_1 \sim \text{Bin}(h, 1 - p)$ and $n_2 \sim \text{Bin}(\ell, 1 - p)$ respectively. The number of type 1 and type 2 agents in sequence $I_2$ is $\tilde{n}_1 \sim \text{Bin}(\hat{h}, 1 - p)$ and $\tilde{n}_2 \sim \text{Bin}(\hat{\ell}, 1 - p)$ respectively. By our construction, we have $n_1 = 0$, $n_2 > m$ with probability 1. Furthermore, $\tilde{n}_1 > m$ and $\tilde{n}_2 > n_2 > m$ with probability 1. As the naive algorithm does not use the sample information, it cannot differentiate $I_1$ and $I_2$ before the arrival of the type 2 agents. Therefore, for both arrival sequences, the naive algorithm should decide about the number of type 2 agents it accepts among the first $m$ arrivals. If the naive algorithm selects a small protection level $x$, it loses many type 1 agents in $I_2$. If it selects a large protection level $x$, then it loses many type 2 agents in $I_1$. This trade-off then leads to the upper bound of $\frac{1}{2} - \alpha$ on the competitive ratio of the naive algorithm.

3.2. Asymptotically Optimal Protection Level Algorithm

The naive algorithm (Algorithm 1) and in particular, our upper bound on its competitive ratio, presented in Theorem 1, highlights the necessity of using the sample information $\psi = (s, \rho)$. Recall that in our setting, the decision-maker receives some information about the number of agents of each type in the online arrival sequence. The decision-maker further obtains some partial information.

Electronic copy available at: https://ssrn.com/abstract=4054796
about the expected reward of each type. Recall that the decision-maker observes \( s_i \) samples, \( i \in \{1, 2\} \), for each type \( i \) agents where \( s_1 \sim \text{Bin}(h, p) \) and \( s_2 \sim \text{Bin}(\ell, p) \). In addition, the decision-maker observes the realized reward of \( s_i \) agents of type \( i \in \{1, 2\} \), denoted by \( \rho_i = (\rho_{i,1}, \rho_{i,2}, \ldots, \rho_{i,s_i}) \).

Our proposed algorithm (Algorithm 2) takes advantage of both \( s = (s_1, s_2) \) and \( \rho = (\rho_1, \rho_2) \). The algorithm uses these pieces of information to obtain an estimate of the expected reward of each type \( i \), denoted by \( b_{ri} \). In particular, if \( s_i > 0 \), the algorithm simply uses the sample average of the realized reward observed in \( \rho_i \) as an estimate for \( r_i \); that is,

\[
\hat{r}_i = \frac{\sum_{j=1}^{s_i} \rho_{i,j}}{s_i}, \quad i \in \{1, 2\}.
\]

When \( s_i = 0 \), \( \hat{r}_i \) is randomly drawn from a uniform distribution in \([0, 1]\).

Having access to these estimates, the algorithm follows a protection level policy to protect the type that has the highest estimated expected reward. More specifically, if \( \hat{r}_1 > \hat{r}_2 \), then the algorithm assigns a protection level of \( x_1 = \min\{m, s_1 \frac{1-p}{p}\} \) for type 1 agents, where we note \( \mathbb{E}[s_1 \frac{1-p}{p}] \) is equal to the expected number of type 1 agents in the online arrival sequence, i.e., \( n_1 \). Recall that \( n_1 = h - s_1 \) and \( s_1 \sim \text{Bin}(h, p) \). Under this policy, the algorithm keeps accepting any agents until the remaining number of resources is \( x_1 \). After that, the algorithm only accepts type 1 agents. On the other hand, when \( \hat{r}_2 \geq \hat{r}_1 \), the algorithm assigns a protection level of \( x_2 = \min\{m, s_2 \frac{1-p}{p}\} \) for type 2 agents. The description of the algorithm can be found in Algorithm 2.

**Algorithm 2** Online Resource Allocation Algorithm with Samples

**Input:** The number of resources \( m \) and sample information \( \psi = (s, \rho) \), where \( s = (s_1, s_2) \), \( \rho = (\rho_1, \rho_2) \), and for any \( i \in \{1, 2\} \), \( \rho_i = \{\rho_{i,1}, \rho_{i,2}, \ldots, \rho_{i,s_i}\} \).

1. If the number of samples for type \( i \in \{1, 2\} \), i.e., \( s_i \), is positive, define

\[
\hat{r}_i = \frac{\sum_{j=1}^{s_i} \rho_{i,j}}{s_i}, \quad i \in \{1, 2\}.
\]

as an estimate of the expected reward of type \( i \in \{1, 2\} \). Otherwise, \( \hat{r}_i \sim \text{uniform}(0, 1) \), where \( \text{uniform}(0, 1) \) is the uniform distribution in \([0, 1]\).

2. If \( \hat{r}_1 > \hat{r}_2 \), set the protection level \( x_1 = \min\{m, s_1 \frac{1-p}{p}\} \) for type 1. In this case, we keep accepting arriving agents until the remaining number of resources is \( \lfloor x_1 \rfloor \). Then, we reject all arriving type 2 agents.

3. If \( \hat{r}_1 \leq \hat{r}_2 \), set the protection level \( x_2 = \min\{m, s_2 \frac{1-p}{p}\} \) for type 2. In this case, we keep accepting arriving agents until the remaining number of resources is \( \lfloor x_2 \rfloor \). Then, we reject all arriving type 1 agents.
We highlight that in the described algorithm, the adversary’s choice of \( h \) and \( \ell \) influences the algorithm’s estimate for the expected rewards, as well as, the protection levels \( x_1 \) and \( x_2 \). Furthermore, the adversary’s choices impact the algorithm’s decision about protecting a certain type. Note that when the estimated expected rewards \( \hat{r}_i \), \( i \in \{1, 2\} \), are noisy, the algorithm may end up protecting a wrong type (i.e., the type with a lower reward). Nonetheless, as we show in Theorems 2 and 3, Algorithm 2 manages to perform very well despite these challenges. Theorem 2 presents a lower bound on the competitive ratio of Algorithm 2 for any value of \( m \), where \( m \) is the number of resources while Theorem 3 shows how this lower bound scales with \( m \). In particular, Theorem 3 shows that the asymptotic competitive ratio of Algorithm 2 is \( 1 - \Theta(\frac{1}{\sqrt{m}}) \). Comparing this bound with the result in Theorem 1 for the naive algorithm (Algorithm 1) highlights the substantial benefits of using sample information \( \psi \) as we do in Algorithm 2.

**Theorem 2 (Competitive ratio of Algorithm 2).** Consider the model presented in Section 2, where the expected rewards of \( r_i \), \( i \in \{1, 2\} \) is unknown to the decision-maker. Let \( h_0 = \min \{ y : \frac{1}{\sqrt{m}} \geq \frac{y p + \sqrt{y}}{1 - p} \geq h_0 \} = \Theta(m) \), \( h_1 = h_0(1 - p) - \sqrt{p(1 - p)h_0} - \frac{\beta}{\sqrt{h_0}} = \Theta(m) \), \( V = 1 - 2(1 - p)^{\sqrt{m}} = 1 - \Theta((1 - p)^{\sqrt{m}}) \), \( W = \min \left\{ 1 - \frac{1}{\sqrt{m}}, \frac{h_1}{m} \right\} = 1 - \Theta\left(\frac{1}{\sqrt{m}}\right) \), \( \ell_0 = \min \{ y : (1 - p)y - \sqrt{y} \geq m \} = \Theta(m) \), \( \ell_1 = \sqrt{p(1 - p)\ell_0} + \frac{\beta}{\sqrt{m} + \ell_0p} = \Theta(m) \), \( \alpha = \frac{\ell_2}{1 - \ell_0} \), and \( \beta = 0.4215 \cdot \frac{p^2 + (1 - p)^2}{p(1 - p)} \). Then, for any \( m \geq 2 \) and \( p \in (0, 1) \), the competitive ratio of Algorithm 2 denoted by \( CR_A \) is lower bounded by:

\[
CR_A \geq \min \left\{ \min \left\{ CR_1, CR_2, CR_3 \right\}, m \geq m_1, \min \left\{ CR_1, CR_2, CR_3 \right\}, m < m_1 \right\},
\]

where \( m_1 = \min_{y \geq 1/p^4} \left\{ y : r_1 - r_2 > \frac{2}{\sqrt{y}/4^{1/4}p - 1} \right\} \), and

\[
CR_1 = \min \left\{ \left( 1 - \frac{1 - p}{\sqrt{m}} \right)^{+}, 1 - \frac{r_1 - r_2}{(r_1 - r_2) + r_2\sqrt{m}} \right\},
\]

\[
CR_2 = \min \left\{ \left( 1 - \frac{1}{\sqrt{m}} \right) \left( 1 - \frac{1}{m} \right), \frac{h_1}{m} \right\},
\]

\[
CR_3 = (1 - \frac{1}{m})^2 \left( 1 - \frac{1}{\sqrt{mp - m^{1/4}}^2} \right)^2 W^2,
\]

\[
CR_3 = V \cdot \min \left\{ (1 - \frac{1}{m^2})W, \frac{1}{2} \left( 1 - \frac{1}{m^2} \right)W + \frac{1}{2} \min \left\{ (1 - \frac{1}{\ell_0} \alpha, 1 - \frac{1 - p}{\ell_1} \right\} \right\}.
\]

Here, \( y^+ = \max\{y, 0\} \).

**Remark 1.** We provide the exact formula for \( h_0 \), \( \ell_0 \), and \( m_1 \) here. It can be shown that \( h_0 = \frac{m - \sqrt{m}}{1 - p} + \frac{1}{4p^2} + \frac{1}{\sqrt{4p^2 + \frac{m - \sqrt{m}}{p(1 - p)}}} \), \( \ell_0 = \min \{ y : (1 - p)y - \sqrt{y} \geq m \} \) can be written as \( \frac{m}{1 - p} + \frac{1 + \sqrt{4m(1 - p) + 1}}{2(1 - p^2)} \), and \( m_1 = \min_{y \geq 1/p^4} \left\{ y : r_1 - r_2 > \frac{2}{\sqrt{y}/4^{1/4}p - 1} \right\} \), which a constant that depends on \( r_1 \), \( r_2 \), and \( p \) can be written as \( m_1 = \left( \frac{b + \sqrt{b^2 + 4ac}}{2a} \right)^4 \), where \( a = (4 - 3p)p, b = (1 - p)(4 - 3p)^{1/2} \), and \( c = \frac{4(1 - p)^2}{(r_1 - r_2)^2} \).
The proof of Theorem 2 is quite involved, which is presented in Section 6. To obtain a lower bound on the competitive ratio of Algorithm 2, we first show that it suffices to only consider ordered arrival sequences in which any type 2 agents arrive before any type 1 agents; see Lemma 1. We then focus on three cases based on the total number of type 1 and type 2 agents (i.e., $h$ and $\ell$). In the first case, $h$ is less than $\sqrt{m}$ and $\ell$ can take any values, and in the last two cases, $h$ is greater than or equal to $\sqrt{m}$. What distinguishes case 2 and case 3 is the number type 2 agents chosen by the adversary, i.e., $\ell$. While in case 2, $\ell$ is less than $\sqrt{m}$, in case 3, $\ell$ is at least $\sqrt{m}$.

The quantity $\overline{CR}_i$, $i \in [2]$, is the lower bound on the competitive ratio of Algorithm 2 under case $i$. Further, $\overline{CR}_3$ is the lower bound on the competitive ratio of Algorithm 2 under case 3 when the number of resources $m$ is at least a constant $m_1$ while $\overline{CR}_3$ is the lower bound on the competitive ratio of Algorithm 2 under case 3 when $m$ is less than $m_1$. In each case, we consider three factors which can make some loss, namely protecting the wrong (low-reward) type agents, over-protecting the right (high-reward) type agents, and under-protecting the right type agents. To bound the loss caused by wrongfully protecting type 2 agents, we show that either this event happens with low probability, or its loss is small. To bound the loss caused by over- or under-protecting, we use concentration inequalities to show that either situation does not have a significant amount of loss.

To provide insights into the lower bound on Theorem 2 before we derive the asymptotic competitive ratio of Algorithm 2, we present the following example.

Example 1. Consider a setting in which $r_1 = 0.9$, and $r_2$ is uniformly drawn from the interval $(0.5, 0.9)$; that is, $r_2 \sim \text{Uniform}(0.5, 0.9)$. We then consider two regimes for the sampling probability:

1. Low sampling probability: $p \sim \text{Uniform}(0.2, 0.35)$.
2. High sampling probability: $p \sim \text{Uniform}(0.35, 0.5)$.

We let the number of resources $m = \{10, 35, \ldots, 335, 360\}$. For each value of $m$, and each of the two regimes, we generate 300 instances where an instance is determined by $r_2$ and $p$. We then compute three quantities for each of these instances: (i) the true competitive ratio of Algorithm 2 (ii) the lower bound on competitive ratio of Algorithm 2 which is presented in Theorem 2 and (ii) $1/(2 - r_2/r_1) = 1/(2 - \alpha)$. The true competitive ratio of the algorithm is obtained by considering all the possible values for $h$ and $\ell$, and only focusing on ordered arrival sequences in which type 2 agents arrive first. Note that focusing on ordered arrival sequences is without loss of generality shown in Lemma 1. The last quantity $1/(2 - \alpha)$ is the optimal competitive ratio when $r_i, i \in [2]$ is known, but sample information is not available. We use this quantity as the benchmark.

Figures 1a and 1b respectively, show the average of these three quantities versus $m$ in low and high sampling probability regimes. These figures show that our lower bound gets tighter as $m$ increases. Nonetheless, our lower bound is quite loose when $m$ is small. This is mainly because of our loose lower bound for $\overline{CR}_1$. Recall that $\overline{CR}_1$ is the lower bound on the competitive ratio
of the algorithm when \( h \leq \sqrt{m} \). To characterize CR\(_1\), we present a universal lower bound for any realization \( s \), which leads to a loose bound; see Lemma 2 and its proof. We further observe that by increasing \( m \), the competitive ratio of the algorithm improves. What is quite interesting is that the true competitive ratio of the algorithm exceeds the benchmark when \( m \geq 35 \). Recall that the benchmark is the optimal competitive ratio when \( r_i, i \in [2] \) is known, but sample information is not available. This highlights that for large enough \( m \), the value of the sample information outweighs the drawbacks of not knowing the true rewards. Q.E.D.

Example 1 shows that the true competitive ratio of Algorithm 2, as well as, its lower bound improves as the number of resources \( m \) increases. This raises the following question: how does the competitive ratio of Algorithm 2 scale with \( m \) as \( m \) goes to infinity? Theorem 3 answers this question.

**Theorem 3 (Asymptotic Competitive Ratio of Algorithm 2).** Consider the model presented in Section 3, where the expected rewards of \( r_i, i \in \{1, 2\} \) is unknown to the decision-maker. As \( m \) goes to infinity, the worst-case competitive ratio of Algorithm 2 scales with \( 1 - \Theta(1/\sqrt{m}) \).

### 4. Upper Bound of Competitive Ratio Among All Algorithms
In our problem setting, the average rewards of agents are unknown, but the decision-maker can learn about them with the help of the sample information. In this section, we relax the problem a little bit by considering setting in which the decision-maker knows \( r_1 \) and \( r_2 \) in advance. We present an upper bound on the CR of any deterministic and randomized algorithms in the aforementioned relaxed setting when the number of resources \( m \) goes to infinity. Clearly, any upper bound for this relaxed setting is also a valid upper bound for our original setting in Section 2.
First, in Theorem 4 we present an upper bound on the CR of any deterministic algorithm under the relaxed setting. Second, in Theorem 5 we show that the same upper bound holds for any randomized algorithm. These theorems show that the worst-case CR of any randomized and deterministic algorithm is upper bounded by \(1 - \Theta(1/\sqrt{m})\), implying the asymptotic optimality of our Algorithm 2 (Recall that by Theorem 3 the asymptotic CR of Algorithm 2 is \(1 - \Theta(1/\sqrt{m})\).)

In other words, the asymptotic CR of our algorithm—which does not use any feedback throughout the allocation period to tune its estimates for \(r_i\)'s—matches the upper bound on the CR of any deterministic and randomized algorithm in the relaxed setting where \(r_i\)'s are known in advance. This implies that asymptotically, feedback during the allocation period does not add any value in improving the CR of algorithms. This is mainly because of the adversarial nature of arrival sequences. Under adversarial arrivals, the decision-maker cannot control for what type of agents he receives feedback; this is governed by the adversarially-chosen order.

**Theorem 4 (Upper Bound on Competitive Ratio of Any Deterministic Algorithm).**

Consider a special case of our model, presented in Section 3, where the expected rewards of \(r_i\), \(i \in \{1, 2\}\) is known to the decision-maker. As the number of resources \(m\) goes to infinity, any deterministic algorithm has competitive ratio of at most \(1 - \Theta(1/\sqrt{m})\).

To show Theorem 4, we construct the following input family \(\mathcal{F}\): Let \(h = \frac{1}{2}pm\) and \(\bar{h} = pm\). The input family \(\mathcal{F}\) contains all \((h, \ell)\) such that \(h \in [h, \bar{h}]\) and \(\ell = \frac{10000 \cdot m}{p}\). For any \(h \in [h, \bar{h}]\), we then denote \(I_h\) as a random arrival sequence under which \(n_2\) type 2 agents arrive followed by \(n_1\) type 1 agents, where we recall that \(n_1 \sim \text{Bin}(h, 1 - p), n_2 \sim \text{Bin}(\ell, 1 - p)\). We characterize an upper bound on the CR of any deterministic algorithm under the family \(\mathcal{F}\). In this family, because \(\ell >> m\), the decision-maker knows that there will be more than \(m\) type 2 agents showing up. Therefore, the number of type 2 agents in the sample does not impact the acceptance/rejection decisions. Then, given that the online arrival sequences \(I_h\) are all ordered, any deterministic algorithm has to decide about how many type 2 agents they accept provided that they observe \(s_1\) samples from type 1 agents. Put differently, any deterministic algorithm can be represented by a mapping that maps \(s_1\) to the number of type 2 agents it accepts. The proof of Theorem 4 then shows that the best CR under any such mapping is \(1 - \Theta(1/\sqrt{m})\). The main challenge in showing this result is characterizing the optimal mapping. To overcome this, instead of characterizing the optimal mapping, we first construct a specific mapping and show that under this mapping, the CR is at most \(1 - \Theta(1/\sqrt{m})\). We then compare any other mapping with this specific mapping to obtain the desired result.

Next, we derive an upper bound of CR among all randomized algorithm.
Theorem 5 (Upper Bound on Competitive Ratio of Any Randomized Algorithms). Consider a special case of our model, presented in Section 2, where the expected rewards of $r_i$, $i \in \{1, 2\}$ is known to the decision-maker. Then, any randomized algorithm has the CR of at most $1 - \Theta(\frac{1}{\sqrt{m}})$.

To show Theorem 5, unfortunately, we cannot use Von Neuman/Yao principle Seiden (2000). This is because in our setting, even when the input $(h, \ell)$ is realized, due to our sampling procedure, the online arrival sequence is still random. This is different from Von Neuman/Yao principle, because Von Neuman/Yao principle can only be applied to the model without any randomness. Nonetheless, in Lemma 5 we derive a result similar to the Von Neuman/Yao principle that can be applied to our setting. We then apply Lemma 5 by constructing a distribution over the input family $F$ introduced above. This leads to the desired upper bound.

5. A Case Study on COVID-19 Dataset

Dataset. In this section, we numerically evaluate our algorithm (Algorithm 2) using the “Laboratory-Confirmed COVID-19-Associated Hospitalizations” dataset. This dataset contains the number of bi-weekly cases of COVID-19-associated hospitalizations in the US from March 7th, 2020 to February 5th, 2021 across five age groups of patients: 0-4 years, 5-17 years, 18-49 years, 50-64 years, and 65+ years. Because the number of 0-17 years old patients is small, we discard them in our analysis. We then consider two groups/types of patients: middle-age patients (18-64 years) and senior patients (65+ years). In our studies, the middle-age patients are considered to be the low-reward agents (i.e., type 2 is our setting), and senior patients are considered to be high-reward agents (i.e., type 1 agent in our setting). This assumption is partly motivated by high death-rate of senior patients when contracting COVID-19 The-New-York-Times (2022).

Simulation setup. In our studies, given the description of our dataset, we consider the allocation problem of a resource (e.g., certain medicine) over the course of two weeks, where each resource can be assigned to at most one patient and each patient needs one unit of the resource. For each such periods, we determine the number of hospitalized patients of the two aforementioned types (i.e., $h$ and $\ell$) using our dataset. (Note that we have 83 periods in our dataset.) The value of $h$ and $\ell$ over 83 periods can be found in Figure 2. At the beginning of each of the periods, we observe the (realized) effectiveness of the resource for a sample of hospitalized patients, where each of the $(h + \ell)$ hospitalized patients falls into the sample with probability $p$. We assume that the leftover units of the resource cannot carry over to the next period. In addition, to be consistent with our setup, we assume that no learning happens across periods. These sampled patients are, for example, the patients that arrive at the beginning of the two-week period during which rationing the resource has not started yet. Let $n_1$ and $n_2$ be the number of hospitalized patients of types 1
and 2 who did not fall into the sample. We assume that the order over these $n_1 + n_2$ patients are completely random. That is, we consider uniform permutations over these patients, modeling more realistic scenarios where the order over patients are not chosen adversarially.

We consider 9 problem classes where each problem class is determined by two parameters $(p, \gamma)$. Here, $p \in \{0.1, 0.15, 0.2\}$ is the sampling probability, and $\gamma \in \{0.3, 0.5, 0.7\}$ determines the scarcity of the resource. In particular, we set $m = \gamma \cdot (\hat{h} + \hat{\ell})$, where $\hat{h}$ and $\hat{\ell}$ are respectively the number of type 1 and type 2 patients in the previous period (i.e., the last two weeks). For each problem class, we generate 1000 instances, each instance is determined by $(n_1, n_2)$ and an order over $n_1 + n_2$ patients. For each of the 9 problem classes, we evaluate the performance of three algorithms: naive Algorithm $\text{1}$, Algorithm $\text{2}$ that has access to the true value of $p$, and Algorithm $\text{2}$ that has only access to the noisy estimate of $p$, denoted by $\hat{p}$, where $\hat{p} \sim \text{Uniform}(p, \bar{p})$. Here, $\mathbb{E}[\hat{p}] = p$ and standard error of $(\hat{p})$ is set to $0.3p$. For example, if $p = 0.1$, we have $\bar{p} \approx 0.15$, $p \approx 0.05$. By evaluating the third algorithm, we investigate how sensitive Algorithm $\text{2}$ is to estimation errors in $p$, as the decision-maker may not have access to the exact knowledge of $p$ in practice. In all of our evaluations, we assume that $r_1 = 0.6$ and $r_2 = 0.2$.

**Performance Evaluation.** Table $\text{1}$ displays the worst-case and average CR of the three algorithms across the 82 allocation periods for each of our 9 problem classes, where the CR of an algorithm is defined as the ratio of its realized reward to the optimal in-hindsight cumulative reward. Interestingly, for all problem classes, the average CR of our algorithm and its noisy version is at least 0.88 while the average CR of the naive algorithm can be as low as 0.66. Overall, we observe that Algorithm $\text{2}$ with true value of $p$ and Algorithm $\text{2}$ with noisy $\hat{p}$ significantly outperform the naive algorithm that does not take advantage of the sample information. In addition, from the performance of Algorithm $\text{2}$ with $\hat{p}$, we can find that our algorithm is robust to the lack of exact

![Figure 2](image.png)  

**Figure 2** Number of middle-age patients and senior patients over weeks
knowledge of $p$. Finally, as a general observation, the performance of Algorithm 2 and its noisy variation enhance as $p$ increases.

Table 1  Competitive Ratio of Three Algorithms. The standard error of average CRs of the three algorithms is less than $0.001$.

<table>
<thead>
<tr>
<th>Value of $p$</th>
<th>0.1</th>
<th>0.15</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value of $\gamma$</td>
<td>$\gamma = 0.3$</td>
<td>$\gamma = 0.5$</td>
<td>$\gamma = 0.7$</td>
</tr>
<tr>
<td>Naive</td>
<td>Avg. CR</td>
<td>0.656</td>
<td>0.693</td>
</tr>
<tr>
<td>Alg. 1</td>
<td>Worst CR</td>
<td>0.567</td>
<td>0.573</td>
</tr>
<tr>
<td>Alg. 2 with true $p$</td>
<td>Avg. CR</td>
<td>0.977</td>
<td>0.953</td>
</tr>
<tr>
<td>Alg. 2 with noisy $p$</td>
<td>Worst CR</td>
<td>0.791</td>
<td>0.763</td>
</tr>
</tbody>
</table>

6. Proof of Theorem 2

To provide the worst-case CR of Algorithm 2, we begin by presenting the following lemma. This lemma states that the worst-case CR of Algorithm 2 is obtained when in the arrival sequence, type 2 agents (i.e., type with the lowest average reward) arrive first, followed by type 1 agents. The proof of all the lemmas in this section is presented in Appendix [B].

**Lemma 1 (Worst order).** For any realization $n_1, n_2$, where $n_i$ is the number of type $i$ agents in the arrival sequence, let $I_{order}$ be an ordered arrival sequence under which $n_2$ type 2 agents arrived first, followed by $n_1$ type 1 agents. Let $I$ be any arrival sequence that contains $n_2$ type 2 agents and $n_1$ type 1 agents. Then, we have

$$\frac{E[REW_A(I_{order}, \psi)]}{OPT(I_{order})} \leq \frac{E[REW_A(I, \psi)]}{OPT(I)}.$$  \hspace{1cm} (2)

where $A$ represents Algorithm 2.

By Lemma 1 to characterize the worst-case CR of Algorithm 2, it suffices to consider ordered sequences under which type 2 agents arrive before any type 1 agents. This then allows us to rewrite the CR of Algorithm 2 denoted by $A$, as follows

$$CR_A = \inf_{(h, \ell)} E(\psi) \left[ \inf_I \frac{E[REW_A(I, \psi)]}{OPT(I)} \right] = \inf_{(h, \ell)} E(\psi) \left[ \frac{REW_A(n, \psi)}{OPT(n)} \right],$$

where $h$ and $\ell$ are the number of types 1 and 2 agents the adversary picks and $n = (n_1, n_2)$ is the number of types 1 and 2 agents in an online ordered arrival sequence $I$. Further, with a slight abuse of notation, $OPT(n)$ is the optimal clairvoyant cumulative expected reward under an online ordered arrival sequence $I$ with $n_i$ type $i$ agents. Similarly, $REW_A(n, \psi)$ is the (expected) reward of Algorithm 2 under the sample information $\psi$ and an online ordered arrival sequence $I$ with $n_i$ type $i$ agents.
To bound \( CR_A \), we define the following \textit{good event}, denoted by, \( \mathcal{G}(\psi;(h,\ell)) \), as the event under which Algorithm [2] correctly assumes type 1 agents have a higher expected reward than type 2 agents:

\[
\mathcal{G}(\psi;(h,\ell)) = \{ \hat{r}_1 > \hat{r}_2 \} = \left\{ \frac{\sum_{j=1}^{s_1^1} \rho_{1,j}}{s_1} > \frac{\sum_{j=1}^{s_2^2} \rho_{2,j}}{s_2} \right\}.
\]

Then, for a given realization of \( \mathbf{s} \) or equivalently \( \mathbf{n} \), we have

\[
\text{REW}_A(\mathbf{n}, \psi) = E[\text{REW}_A(\mathbf{n}, \psi) \cdot 1(\mathcal{G}(\psi;(h,\ell)))] + E[\text{REW}_A(\mathbf{n}, \psi) \cdot 1(\mathcal{G}^c(\psi;(h,\ell)))]
\]

\[
= E[\text{REW}_A(\mathbf{n}, \psi) | \mathcal{G}(\psi;(h,\ell))] \cdot Pr(\mathcal{G}(\psi;(h,\ell)))
\]

\[
+ E[\text{REW}_A(\mathbf{n}, \psi) | \mathcal{G}^c(\psi;(h,\ell))] \cdot Pr(\mathcal{G}^c(\psi;(h,\ell)))
\],

where \( 1(\mathcal{A}) \) is 1 if an event \( \mathcal{A} \) occurs and zero otherwise. Further, \( \mathcal{G}(\psi;(h,\ell)) \) is the good event defined in Equation [(3)]. The outer expectation is with respect to \( \mathbf{\rho} \) and any randomness in the algorithm. The expression \( E[\text{REW}_A(\mathbf{n}, \psi) | \mathcal{G}(\psi;(h,\ell))] \) inside the outer expectation presents the reward of Algorithm [2] under a specific realization of \( \mathbf{s} \) when the good event \( \mathcal{G}(\psi;(h,\ell)) \) happens. Note that for a given realization of \( \mathbf{s} \) and under the good event, the algorithm’s action and hence its reward is deterministic and does not depend on \( \mathbf{\rho} \). This is because when the good event happens, the algorithm assigns a protection level \( x_1 \) to type 1, where \( x_1 \) only depends on \( s_1 \), not \( \mathbf{\rho} \). Finally, in \( Pr(\mathcal{G}(\psi;(h,\ell))) \), we take expectation with respect to \( \mathbf{\rho} \) and any randomness in the algorithm in defining \( \hat{r}_i, i \in [2] \), given a realization of \( \mathbf{s} \).

We consider the following three cases based on the number of type 1 and 2 agents chosen by the adversary (i.e., \( h \) and \( \ell \)):

\textit{Case 1:} \( h < \sqrt{m}, \ell \geq 0 \). In this case, the number of type 1 agents is small (less than \( \sqrt{m} \)). That is, \( (h,\ell) \in \mathcal{R}_1 \), where \( \mathcal{R}_1 = \{(h,\ell) : h < \sqrt{m}, \ell \geq 0 \} \).

\textit{Case 2:} \( h \geq \sqrt{m}, \ell < \sqrt{m} \). In this case, while the number of type 1 agents is large, the number of type 2 agents is small. That is, \( (h,\ell) \in \mathcal{R}_2 \), where \( \mathcal{R}_2 = \{(h,\ell) : h \geq \sqrt{m}, \ell < \sqrt{m} \} \).

\textit{Case 3:} \( h \geq \sqrt{m}, \ell \geq \sqrt{m} \). In this case, the numbers of both type 1 and 2 agents are large. That is, \( (h,\ell) \in \mathcal{R}_3 \), where \( \mathcal{R}_3 = \{(h,\ell) : h \geq \sqrt{m}, \ell \geq \sqrt{m} \} \).

Observe that the worst case \( CR \) of Algorithm [2] can be written as:

\[
\min_{i \in [3]} \left\{ \inf_{(h,\ell) \in \mathcal{R}_i} E[\psi \left[ \frac{\text{REW}_A(\mathbf{n}, \psi)}{\text{OPT}(\mathbf{n})} \right] \right\},
\]

where by Equation [(4)], we have

\[
\inf_{(h,\ell) \in \mathcal{R}_i} E[\psi \left[ \frac{\text{REW}_A(\mathbf{n}, \psi)}{\text{OPT}(\mathbf{n})} \right] = \inf_{(h,\ell) \in \mathcal{R}_i} \left\{ E_{\mathbf{s}} \left[ \frac{E[\text{REW}_A(\mathbf{n}, \psi) | \mathcal{G}(\psi;(h,\ell))] \cdot Pr(\mathcal{G}(\psi;(h,\ell)))}{\text{OPT}(\mathbf{n})} \right] \right\}
\]

\[
+ E_{\mathbf{s}} \left[ \frac{E[\text{REW}_A(\mathbf{n}, \psi) | \mathcal{G}^c(\psi;(h,\ell))] \cdot Pr(\mathcal{G}^c(\psi;(h,\ell)))}{\text{OPT}(\mathbf{n})} \right] \right\}
\]
In the rest of the proof, we use the following shorthand notation to simplify the exposition:

\[
\mathbb{E}[\text{REW}_A(n, \psi) | \mathcal{G}(\psi; (h, \ell))] = \text{REW}_A(n, \psi; \mathcal{G}), \quad \mathbb{E}[\text{REW}_A(n, \psi) | \mathcal{G}^c(\psi; (h, \ell))] = \text{REW}_A(n, \psi; \mathcal{G}^c).
\]

The proof will be completed by presenting the following three main lemmas: Lemmas \[2\] \[3\] and \[6\]. Each lemma provides the lower bound of \( \inf_{(h, \ell) \in \mathcal{R}_i} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \) for one of the regions \( \mathcal{R}_i \), \( i \in [3] \).

**Lemma 2 (Case 1: Region \( \mathcal{R}_1 \)).** Let \( \mathcal{R}_1 = \{(h, \ell) : h < \sqrt{m}, \ell \geq 0\} \). Then, we have

\[
\inf_{(h, \ell) \in \mathcal{R}_1} \mathbb{E}_s \left[ \frac{\mathbb{E}[\text{REW}_A(n, \psi)]}{\text{OPT}(n)} \right] \geq CR_1 = \min \left\{ \left(1 - \frac{1-p}{p\sqrt{m}}\right), 1 - \frac{r_1-r_2}{(r_1-r_2)+r_2\sqrt{m}} \right\}.
\]

To show Lemma \[2\] by Equation \[6\], we bound \( \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \) and \( \frac{\text{REW}_A(n, \psi; G^c)}{\text{OPT}(n)} \) for any realization of \( s \) by considering different ranges for the number of agents of type \( i \in [2] \) in the online arrival sequence.

**Lemma 3 (Case 2: region \( \mathcal{R}_2 \)).** Let \( \mathcal{R}_2 = \{(h, \ell) : h \geq \sqrt{m}, \ell < \sqrt{m}\} \). Then, we have

\[
\inf_{(h, \ell) \in \mathcal{R}_2} \mathbb{E}_s \left[ \frac{\mathbb{E}[\text{REW}_A(n, \psi)]}{\text{OPT}(n)} \right] \geq CR_2 = \min \left\{ \left(1 - \frac{1}{\sqrt{m}}\right), h_1 \right\},
\]

where \( h_0 = \min \{y \geq 0 : \frac{p(m-\sqrt{m})}{1-p} \geq yp + \sqrt{g} \} \), \( h_1 = h_0(1-p) - \frac{p(1-p)}{\sqrt{h_0}} - \frac{\beta}{\sqrt{h_0}} \), and \( \beta = 0.4215 \cdot \frac{p^2+(1-p)^2}{p(1-p)} \).

**Proof of Lemma \[3\]** We follow similar steps in the proof of Lemma \[2\]. That is, we bound

\[
\inf_{(h, \ell) \in \mathcal{R}_2} \mathbb{E}_s \left[ \frac{\mathbb{E}[\text{REW}_A(n, \psi)]}{\text{OPT}(n)} \right] \geq \inf_{(h, \ell) \in \mathcal{R}_2} \min \left\{ \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right], \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G^c)}{\text{OPT}(n)} \right] \right\}.
\]

In light of Equation \[7\], we divide the rest of the proof into two parts where in the first part, for any realization of \( s \), we bound \( \frac{\text{REW}_A(n, \psi; G^c)}{\text{OPT}(n)} \) and in the second part, we bound \( \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \). While bounding the former ratio is straightforward, bounding the latter one is quite involved. This is because when the good event \( \mathcal{G}(\psi; (h, \ell)) \) happens, Algorithm \[2\] may still over-protect or under-protect type 1 agents given the randomness in the sample information. As \( h \) is large in this case, the number of type 1 agents that the algorithm over-protects or under-protects can be large too, leading to a low reward for the algorithm under some realization of \( s \). Nonetheless, we show that the expected reward of the algorithm, in this case, is still large once we take expectation with respect to \( s \).
Part 1: bounding $\frac{\text{REW}_A(n, \psi; G^C)}{\text{OPT}(n)}$. Recall that for fixed realization of $s$, $\text{REW}_A(n, \psi; G^C)$ is the reward of our algorithm when there are $n_i$ type $i \in \{1, 2\}$ agents in the online arrival sequence and $\hat r_1 \leq \hat r_2$. In this case, the algorithm assigns a protection level of $x_2 = \min\{m, \frac{1-p}{p}s_2\}$, and hence we have

$$\text{REW}_A(n, \psi; G^C) \geq \min\{m, n_2\} r_2 + \min\{n_1, (m-n_2)^+\} r_1.$$ 

As $\ell < \sqrt{m}$, we have $n_2 \leq \ell < \sqrt{m} < m$. Hence,

$$\frac{\text{REW}_A(n, \psi; G^C)}{\text{OPT}(n)} \geq \frac{\text{REW}_A(n, \psi; G^C)}{n_1 r_1 + \min\{m-n_1, n_2\} r_2} = \frac{n_2 r_2 + \min\{m-n_2, n_1\} r_1}{n_1 r_1 + \min\{m-n_1, n_2\} r_2} \geq \min\left\{\frac{n_1 r_1 + n_2 r_2}{n_1 r_1 + (m-n_1) r_2}, \frac{(m-n_2) r_1 + n_2 r_2}{mr_1}\right\} = 1 - \frac{1}{\sqrt{m}}.$$ 

The second inequality is because when $\min\{m-n_2, n_1\} = n_1$, we have $\min\{m-n_2, n_1\} = n_2$, and when $\min\{m-n_2, n_1\} = m-n_2$, we have $\min\{m-n_2, n_1\} = m-n_1$.

Part 2: bounding $\inf_{(h, \ell) \in R_2} \mathbb{E}_s\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}\right]$. We begin by partitioning region $R_2$ based on the number of type 1 agents, (i.e. $h$):

$$R_2 = \{(h, \ell) : (h, \ell) \in R_2, h \geq h_0\} \quad R_2 = \{(h, \ell) : (h, \ell) \in R_2, h < h_0\},$$

where $h_0 = \min\{y \geq 0 : \frac{y(m-\sqrt{m})}{1-p} \geq yp + \sqrt{y}\}$. Then, $\inf_{(h, \ell) \in R_2} \mathbb{E}_s\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}\right]$ is equal to the minimum of $\inf_{(h, \ell) \in R_2} \mathbb{E}_s\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}\right]$ and $\inf_{(h, \ell) \in R_2} \mathbb{E}_s\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}\right]$. We bound the first ratio (i.e., the one concerns $R_2$) in Lemma 4 and the second ratio (i.e., the one concerns $R_2$) will be bounded in Lemma 5.

**Lemma 4.** Let $R_2 = \{(h, \ell) : (h, \ell) \in R_2, h < h_0\}$, where $R_2 = \{(h, \ell) : h \geq \sqrt{m}, \ell < \sqrt{m}\}$, $h_0 = \min\{y \geq 0 : \frac{y(m-\sqrt{m})}{1-p} \geq yp + \sqrt{y}\}$.

$$\inf_{(h, \ell) \in R_2} \mathbb{E}_s\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}\right] \geq (1 - \frac{1}{\sqrt{m}})(1 - \frac{1}{m}).$$

Lemma 4 gives the lower bound of the ratio conditional on the good event happens. In this case, the algorithm gives certain protection level to type 1 agents. Then, we show that when $(h, \ell) \in R_2$, with high probability, the number of type 2 agents is less than the total number of resources minus the protection level. Thus, in this regime, we do not reject any type 2 agents with high probability. Conditional on this event, we can obtain a lower bound of the ratio for any realization $s$.

Next, we lower bound $\inf_{(h, \ell) \in R_2} \mathbb{E}_s\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}\right]$ in Lemma 5.
Lemma 5. Let $\mathcal{R}_2 = \{(h, \ell) : (h, \ell) \in \mathcal{R}_2, h \geq h_0\}$, where $\mathcal{R}_2 = \{(h, \ell) : h \geq \sqrt{m}, \ell < \sqrt{m}\}$, where $h_1 = h_0(1 - p) - \sqrt{p(1 - p)h_0} - \frac{\beta}{\sqrt{h_0}}$, $\beta = 0.4215 \cdot \frac{p^2 + (1-p)^2}{p(1-p)}$, $h_0 = \min\{y \geq 0 : \frac{p(m - \sqrt{m})}{1 - p} \geq yp + \sqrt{y}\}$. Then,

$$\inf_{(h, \ell) \in \mathcal{R}_2} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \right] \geq \min \left\{1 - \frac{1}{\sqrt{m}}, \frac{h_1}{m} \right\}.$$ 

To show Lemma 5, notice that given the good event happens, the algorithm may over- or under-protect type 1 agents. If it under-protects type 1 agents, we further split the analysis into two cases: the number of type 2 agents is less than, or is larger than the total number of resources minus the protection level. In both cases, we can easily find a lower bound for any realization $s$. If the algorithm over-protects type 1 agents, then it accepts all arriving type 1 agents, but some resources are wasted due to over-protecting. In this case, we lower bound the ratio by a linear function with respect to the number of arriving type 1 agents, $n_1$. Then, we used the results in Nagaev and Chebotarev (2011) and Berend and Kontorovich (2013) to estimate the conditional expectation of $n_1$ given that the algorithm over-protects type 1 agents.

Finally, by Lemmas 4 and 5, we have the following inequality, which is the desired result.

$$\inf_{(h, \ell) \in \mathcal{R}_2} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \right] \geq CR_2 = \min \left\{1 - \frac{1}{\sqrt{m}}, \frac{h_1}{m} \right\};$$

Lemma 6 (Case 3: region $\mathcal{R}_3$). Let $\mathcal{R}_3 = \{(h, \ell) : h \geq \sqrt{m}, \ell \geq \sqrt{m}\}$. Then, $\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \right]$ is at least

$$\min \left\{(1 - \frac{1}{m})^2(1 - \frac{1}{\sqrt{m} - \frac{1}{m}}) W, \quad m \geq m_1, \right. \right.$$

$$\left. \left. V \cdot \min \left\{1 - \frac{1}{m^2} \right\} W, \frac{1}{2}(1 - \frac{1}{m^2}) W + \frac{1}{2} \min \left\{1 - \frac{1}{m^2} \right\} \right\} m < m_1, \right.$$ 

where $V = 1 - 2(1-p)^{\sqrt{m}}$, $W = \min \left\{1 - \frac{1}{\sqrt{m}} \cdot \frac{h_1}{m} \right\}$, $h_1 = h_0(1 - p) - \sqrt{p(1 - p)h_0^*} - \frac{\beta}{\sqrt{h_0}^*}$, $\beta = 0.4215 \cdot \frac{p^2 + (1-p)^2}{p(1-p)}$, $\ell_0 = \min\{y \geq 0 : \frac{p(m - \sqrt{m})}{1 - p} \geq yp + \sqrt{y}\}$, $m_1 = \min_{y \geq 1/p^4} \left\{y : r_1 - r_2 > \frac{2}{y^{1/8} \sqrt{y^{1/4} - 1}} \right\}$, $\ell_1 = \sqrt{p(1-p)} \ell_0 + \frac{\beta}{\sqrt{m}} + \ell_0 p$, and $\ell_0 = \min\{y : (1 - p) y - \sqrt{y} \geq m\}$.

Proof of Lemma 6. We split the analysis into two cases based on the initial number of resources $m$. In the first case, we assume that $m \geq m_1$, and in the second case, we assume that $m < m_1$.

Here, $m_1 = \min_{y \geq 1/p^4} \left\{y : r_1 - r_2 > \frac{2}{y^{1/8} \sqrt{y^{1/4} - 1}} \right\}$ is a constant that only depends on $p, r_1$, and $r_2$.

Case 1 is considered in Lemma 7, while case 2 is studied in Lemma 8.

Lemma 7. Let $\mathcal{R}_3 = \{(h, \ell) : h \geq \sqrt{m}, \ell \geq \sqrt{m}\}$. Let $m_1 = \min_{y \geq 1/p^4} \left\{y : r_1 - r_2 > \frac{2}{y^{1/8} \sqrt{y^{1/4} - 1}} \right\}$. When the initial number of resources $m \geq m_1$, we have

$$\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \right] \geq \left(1 - \frac{1}{m} \right)^2(1 - \frac{1}{(\sqrt{mp} - m^{1/4})^2}) W.$$
To show Lemma 7 we observe that when the initial number of resources \( m \geq m_1 \), the good event happens with high probability. Then, conditional on the algorithm protects type 1 agents, we give the lower bound of the ratio by the similar methodology introduced in Lemma 4 and Lemma 5: we discuss whether the algorithm over- or under-protects type 1 agents, and whether the number of type 2 agents in the arriving sequence is less than, or is larger than the total number of resources minus the protection level.

**Lemma 8.** Let \( \mathcal{R}_3 = \{(h, \ell) : h \geq \sqrt{m}, \ell \geq \sqrt{m}\} \), \( m_1 = \min_{y \geq 1/p^4} \left\{ y : r_1 - r_2 > \frac{2}{y^{1/8} \sqrt{y^{1/4} p - 1}} \right\} \), and \( \alpha = \frac{r_2}{r_1} \). When the initial number of resources \( m < m_1 \), we have

\[
\inf_{(h, \ell) \in \mathcal{R}_3} E\left[ \frac{\mathbb{E}[\text{REW}_A(n, \psi)]}{\text{OPT}(n)} \right] \geq V \cdot \min \left\{ \left(1 - \frac{1}{m^2}\right) W, \frac{1}{2} \left(1 - \frac{1}{m^2}\right) W + \frac{1}{2} \min \left\{ (1 - \frac{1}{\ell_0^2}) \alpha, 1 - \frac{1 - p}{pm} \ell_1 \right\} \right\}.
\]

To show Lemma 8 we should use a different technique rather than the one we use in Lemma 7. This is because if the initial number of resources \( m < m_1 \), the bad event is no longer a rare event. Then, if the good event happens, the ratio is derived in Lemma 7. If the bad event happens, the algorithm gives protection level to type 2 agents. Then, we first consider that if the algorithm under-protects type 2 agents, the algorithm accepts all type 2 agents, and give the remaining resources to type 1 agents. In this case, the ratio is bounded by \( \alpha \). Otherwise, if the algorithm over-protects type 2 agents, the algorithm rejects some type 1 agents and wastes some resources. Then, we use concentration inequality to bound the loss.

### 7. Concluding Remarks and Future Directions

In this paper, we consider an online resource allocation problem in which the decision-maker has uncertainty about the arrival process, as well as, the obtained rewards upon the allocation. The decision-maker has access to the sample information that is often acquired through an initial test period. We study to optimally exploit the sample information that provides partial knowledge about the arrival process and rewards. We propose a protection-level algorithm that achieves the competitive ratio of \( 1 - \Theta(1/\sqrt{m}) \) and show that the obtained competitive ratio is asymptotically tight, where \( m \) is the initial number of resources.

Our result shows that the sample information can be significantly beneficial, and hence opens up several new directions for future research. One natural direction is to study how to optimize the number of resources \( m \) upon receiving the sample information. While in many settings, due to a long lead production time, it is not possible to react to the sample information to optimize the inventory decisions, for some other settings when products/resources are produced domestically, the decision-maker can better plan his inventory decisions using the sample information. Another research question of interest is to study how to generalize our setting to the network revenue...
management problem where there are multiple types of agents and multiple types of resources. The network revenue management problem under adversarial arrival setting does not admit a constant competitive ratio. Thus, it is interesting to explore if one can even obtain a constant competitive ratio under adversarial arrival setting in the presence of sample information.

References


In the appendix, we provide the proof of Theorem 4 and most of the Lemmas in Section 6. The rest of the proofs can be found in https://www.dropbox.com/sh/kar1hxmg9tlctmw/AAA1RcYK1f34HeKM_N326DRXa?
dl=0

Appendix A: Proof of Theorem 4

We construct the following input family $F$: let $h = \frac{1}{2}pm$ and $\tilde{h} = pm$. The input family $F$ contains all $(h, \ell)$ such that $h \in [h, \tilde{h}]$ and $\ell = \frac{10000 - m}{p}$. For any $h \in [h, \tilde{h}]$, we then denote $I_h$ as a random arrival sequence under which $n_2$ type 2 agents arrive followed by $n_1$ type 1 agents, where we recall that $n_1 \sim \text{Bin}(h, 1 - p)$, $n_2 \sim \text{Bin}(\ell, 1 - p)$. We denote $I_h$ as the random arrival sequences, where in $I_h$, $n_2$ type 2 agents arrive followed by $n_1$ type 1 agents.

Our goal is to characterize an upper bound on the competitive ratio of any deterministic policy under the family $F$. Consider a specific input $h \in [h, \tilde{h}]$. The decision maker gets the sample $(s_1, s_2)$ at the beginning, where $s_1 \sim \text{Bin}(h, p)$, and $s_2 \sim \text{Bin}(\ell, p)$. As we select $l$ to be a super large constant, we have with probability $1$, $s_2 >> m$. Therefore, under the family $F$, the decision-maker knows that there will be more than $m$ type 2 agents showing up. Based on the structure that all type 2 agents arrive before all type 1 agents, any deterministic algorithm has to decide how many type 2 agents they accept if they observe $(s_1, s_2)$. As $s_2$ is always larger than $m$, it does not impact the decisions. Put differently, any deterministic algorithm can be represented by a mapping $z(\cdot): \{0, 1, \ldots, \tilde{h}\} \rightarrow \{0, 1, \ldots, m\}$, such that once they observe $(s_1, s_2)$, they will accept $m - z(s_1)$ type 2 agents. (As the number of type 1 agents, $h$, is less than $m$, the number of type 1 agents in the online arrival sequence $n_1 < m$.) Let $A$ be an algorithm associated with mapping $z$. We define the following loss function for algorithm $A$ under arrival sequence $I$ and sample information $s$:

$$\text{LOSS}_A(I, s) = \text{OPT}(I) - \text{REW}_A(I, s).$$

(Note that since $r_1$ and $r_2$ are known, we can replace $\psi = (s, \rho)$ with $s$.) In the rest of this proof, we use the short-hand of $\text{LOSS}(I, s, z)$ and $\text{REW}(I, s, z)$ in place of $\text{LOSS}_A(I, s)$ and $\text{REW}_A(I, s)$, respectively. Then, the competitive ratio of a deterministic algorithm $A$ with mapping $z(\cdot)$ on family $F$ is at least

$$\text{CR}(z) = 1 - \sup_{(h, \ell) \in F} \mathbb{E}_s \left[ \frac{\text{LOSS}(I, s, z)}{\text{OPT}(I)} \right].$$

In what follows, we provide an upper bound on $\text{CR}(z)$ in Equation (8). To do so, we consider the following two cases:

- **Case 1**: $n_1 \geq z(s_1)$. In this case, the number of class 1 agent arriving online (which is a random variable) is larger than the remaining number of resources after we accept $m - z(s_1)$ type 2 agents. Therefore, the decision-maker accepts $m - z(s_1)$ type 2 agents and $z(s_1)$ type 1 agents. In this case, the loss (i.e., the difference between the optimal clairvoyant cumulative expected reward and what the decision-maker gets which is defined in Equation (8)) is $(n_1 - z(s_1))(r_1 - r_2)$.

- **Case 2**: $n_1 < z(s_1)$. In this case, the number of class 1 agent arriving online is less than the remaining number of resources after we accept $m - z(s_1)$ type 2 agents. Therefore, the decision-maker accepts $m - z(s_1)$ type 2 agents and $n_1$ type 1 agents. Therefore, we waste $z(s_1) - n_1$ units of resources. In this case, the loss, defined in Equation (8), is $(z(s_1) - n_1)r_2$. 

Electronic copy available at: https://ssrn.com/abstract=4054796
Putting these two cases together, we have

\[
\text{LOSS}(I, s, z) = (n_1 - z(s_1))(r_1 - r_2)1(n_1 \geq z(s_1)) + (z(s_1) - n_1)r_21(n_1 < z(s_1)) \\
= (n_1 - z(s_1))r_1(1(n_1 \geq z(s_1)) + (z(s_1) - n_1)r_2).
\]

Hence, an upper bound on the comparative ratio of any deterministic algorithm on family \( F \), denoted by \( \overline{CR} \) is at least

\[
\overline{CR} = \max_z CR(z) \\
= 1 - \min_z \max_{h \in [h, \bar{h}]} E_s \left[ \frac{\text{LOSS}(I, s, z)}{\text{OPT}(I)} \right] \\
= 1 - \min_z \max_{h \in [h, \bar{h}]} E_s \left[ \frac{(n_1 - z(s_1))r_1(1(n_1 \geq z(s_1)) + (z(s_1) - n_1)r_2)}{n_1r_1 + (m - n_1)r_2} \right],
\]

where \( CR(z) \) is defined in Equation \([9]\).

Although we can solve this minmax optimization problem by some solver because the objective function is convex in every dimension, we cannot calculate the value of the optimal solution in a closed form. Therefore, we develop the analysis below to show that \( \overline{CR} \) is \( 1 - \Theta(\frac{1}{\sqrt{m}}) \).

In our analysis, we first focus on a specific mapping: \( z(s_1) = \frac{1 - p}{p} s_1 \) for any \( s_1 \in \{0, 1, \ldots, \bar{h}\} \). (Since \( s_1 \leq \bar{h} = mp \), we have \( z(s_1) \leq m \).) For this mapping, we calculate the competitive ratio of its associated deterministic algorithm. Second, we compare any other feasible mappings with this mapping to show our result. We want to highlight that this mapping (i.e., \( z(s_1) = \frac{1 - p}{p} s_1 \)) is not the optimal mapping: The optimal mapping should depend on \( r_1 \) and \( r_2 \), but we cannot find its closed form. This mapping is considered as a benchmark with which we compare all other mappings.

**Lemma 9.** Consider the following mapping \( z : \{0, 1, \ldots, \bar{h}\} \to \{0, 1, \ldots, m\} \) such that \( z(s_1) = \frac{1 - p}{p} s_1 \) for any \( s_1 \in \{0, 1, \ldots, \bar{h}\} \). Then, for any value of \( m \)

\[
\max_{h \in [h, \bar{h}]} E_s \left[ \frac{\text{LOSS}(I, s, z)}{\text{OPT}(I)} \right] \leq \frac{r_1\sqrt{1 - p}}{r_2} \frac{1}{\sqrt{m}} + \frac{\sqrt{2}\beta r_1}{p\sqrt{pr_2^2}} \frac{1}{m\sqrt{m}},
\]

and

\[
\max_{h \in [h, \bar{h}]} E_s \left[ \frac{\text{LOSS}(I, s, z)}{\text{OPT}(I)} \right] \geq \frac{p}{2\sqrt{2}} \sqrt{1 - p} \frac{1}{\sqrt{m}} - \frac{\beta}{p\sqrt{p}} \frac{1}{m\sqrt{m}},
\]

where \( \beta = 0.4215 \cdot \frac{p^2 + (1 - p)^2}{p(1 - p)} \). This implies that \( CR(z) = 1 - \Theta(\frac{1}{\sqrt{m}}) \).

By Lemma \([9]\) we have under that particular mapping \( z \), we have \( CR(z) = 1 - \Theta(\frac{1}{\sqrt{m}}) \) over input family \( F \). Now, we want to argue that there not exists any other mapping \( \hat{z}(\cdot) \) such that the competitive ratio over \( F \) is \( 1 - o(\frac{1}{\sqrt{m}}) \). To do so, we compare all other mappings \( \hat{z} : \{0, \ldots, \bar{h}\} \to \{0, 1, \ldots, m\} \), with \( z(s_1) = \frac{1 - p}{p} s_1 \), and show the best policy can only achieve \( 1 - \Theta(\frac{1}{\sqrt{m}}) \) competitive ratio.

Consider any arbitrary mapping \( \hat{z}(\cdot) \). By Equation \([10]\), we have

\[
\text{CR}(\hat{z}) = 1 - \max_{h \in [h, \bar{h}]} E_s \left[ \frac{\text{LOSS}(I, s, \hat{z})}{\text{OPT}(I)} \right] \leq 1 - \max_{h \in [h, \bar{h}]} E_s \left[ \frac{(n_1 - \hat{z}(s_1))r_11(n_1 \geq \hat{z}(s_1)) + (\hat{z}(s_1) - n_1)r_2}{mr_2} \right] \\
= 1 - \frac{1}{m} \max_{h \in [h, \bar{h}]} E_s \left[ \frac{(n_1 - \hat{z}(s_1))r_11(n_1 \geq \hat{z}(s_1)) + (\hat{z}(s_1) - n_1)}{r_2} \right].
\]
Let \( \hat{y}(s_1) = s_1 + \hat{z}(s_1) \) for any \( s_1 \in \{0, 1, \ldots, \hat{h} \} \). Then, since \( n_1 + s_1 = h \), we have

\[
\text{CR}(\hat{z}) \leq 1 - \frac{1}{m} \max_{h \in [\hat{h}, \hat{h}]} \mathbb{E}_s \left[ (h - \hat{y}(s_1)) \frac{r_1}{r_2} 1(h \geq \hat{y}(s_1)) + (\hat{y}(s_1) - h) \right].
\]

Our goal to show that \( \text{CR}(\hat{z}) \leq 1 - O(\frac{1}{\sqrt{m}}) \). We show this by contradiction. Contrary to our claim, suppose that there exists a feasible mapping \( y_0(\cdot) \) with \( y_0(s_1) = s_1 + z_0(s_1) \) under which

\[
\lim_{m \to \infty} \max_{h \in [\hat{h}, \hat{h}]} \frac{\mathbb{E}_s [y_0(s_1) - h]}{\sqrt{m}} = 0.
\]

Note that if Equation (11) holds, we have \( \text{CR}(z_0) = 1 - O(\frac{1}{\sqrt{m}}) \). Now observe that \( \left((h - y_0(s_1)) \frac{r_1}{r_2} 1(h \geq y_0(s_1))\right) \geq 0 \), and hence if Equation (11) holds, we must have

\[
\lim_{m \to \infty} \max_{h \in [\hat{h}, \hat{h}]} \frac{\mathbb{E}_s [y_0(s_1) - h]}{\sqrt{m}} \leq 0.
\]

Next we consider two cases. In the first case, we assume that \( \lim_{m \to \infty} \max_{h \in [\hat{h}, \hat{h}]} \frac{\mathbb{E}_s [y_0(s_1) - h]}{\sqrt{m}} < 0 \) and we reach a contradiction under this assumption. In the second case, assume that \( \lim_{m \to \infty} \max_{h \in [\hat{h}, \hat{h}]} \frac{\mathbb{E}_s [y_0(s_1) - h]}{\sqrt{m}} = 0 \) and again reach a contradiction.

**Case 1:** \( \lim_{m \to \infty} \max_{h \in [\hat{h}, \hat{h}]} \frac{\mathbb{E}_s [y_0(s_1) - h]}{\sqrt{m}} < 0 \). Suppose that \( \lim_{m \to \infty} \max_{h \in [\hat{h}, \hat{h}]} \frac{\mathbb{E}_s [y_0(s_1) - h]}{\sqrt{m}} = \eta < 0 \). Then, we show that \( y_0(\cdot) \) does not satisfy Equation (11), which is a contradiction. The left hand side of Equation (11) can be written as

\[
\lim_{m \to \infty} \max_{h \in [\hat{h}, \hat{h}]} \frac{\mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ (h - y_0(s_1)) \frac{r_1}{r_2} 1(h \geq y_0(s_1)) + (y_0(s_1) - h) \right]}{\sqrt{m}}
\]

\[
\geq \lim_{m \to \infty} \max_{h \in [\hat{h}, \hat{h}]} \frac{\mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ (h - y_0(s_1)) \frac{r_1}{r_2} + (y_0(s_1) - h) \right]}{\sqrt{m}} = - \frac{r_1}{r_2} \eta + \eta = - \frac{r_1 - r_2}{r_2} \eta > 0,
\]

which implies that Equation (11) does not hold and hence it is a contradiction.

**Case 2:** \( \lim_{m \to \infty} \max_{h \in [\hat{h}, \hat{h}]} \frac{\mathbb{E}_s [y_0(s_1) - h]}{\sqrt{m}} = 0 \). Here, we compare \( y_0(\cdot) \) and \( y(\cdot) \) where we recall \( y(s_1) = s_1 + z(s_1) = \frac{1}{p} s_1 \) for any \( s_1 \in \{0, 1, \ldots, \hat{h}\} \). We then show a contradiction. In Lemma 9, we have shown that if \( y(s_1) = \frac{1}{p} s_1 \) (or equivalently \( z(s_1) = \frac{1}{p} - \frac{1}{p} s_1 \)), then for any \( h \in [\hat{h}, \hat{h}] \), we have

\[
\lim_{m \to \infty} \frac{\mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ (h - y(s_1)) \frac{r_1}{r_2} 1(h \geq y(s_1)) + (y(s_1) - h) \right]}{\sqrt{m}} = \lim_{m \to \infty} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ (h - \frac{s_1}{p}) \frac{r_1}{r_2} 1(h \geq \frac{s_1}{p}) + (\frac{s_1}{p} - h) \right] > 0.
\]

Let Equation (11) takes its maximum at \( h = h_0 \in [\hat{h}, \hat{h}] \). Also, note that Equation (11) holds for any \( h \), including \( h_0 \). Hence, by subtracting Equation (13) from Equation (11) evaluated at \( h_0 \), we have

\[
\lim_{m \to \infty} \frac{1}{\sqrt{m}} \left( \mathbb{E}_{s_1 \sim \text{Bin}(h_0, p)} \left[ (h_0 - \frac{s_1}{p}) \frac{r_1}{r_2} 1(h_0 \geq \frac{s_1}{p}) + (\frac{s_1}{p} - h_0) \right] - \mathbb{E}_{s_1 \sim \text{Bin}(h_0, p)} \left[ (h_0 - y_0(s_1)) \frac{r_1}{r_2} 1(h_0 \geq y_0(s_1)) + (y_0(s_1) - h_0) \right] \right) > 0.
\]

Now recall that under this case, we have \( \lim_{m \to \infty} \frac{\mathbb{E}_{s_1 \sim \text{Bin}(h_0, p)} [y_0(s_1) - h_0]}{\sqrt{m}} = 0 \) where \( s_1 \sim \text{Bin}(h_0, p) \). This implies that \( \lim_{m \to \infty} \frac{\mathbb{E}_{s_1 \sim \text{Bin}(h_0, p)} [y_0(s_1) - h_0]}{\sqrt{m}} = 0 \), as \( \mathbb{E}[S_1] = h_0 p \). Therefore, Equation (14) can be written as

\[
\lim_{m \to \infty} \frac{1}{\sqrt{m}} \left( \mathbb{E}_{s_1 \sim \text{Bin}(h_0, p)} \left[ (h_0 - \frac{s_1}{p}) \frac{r_1}{r_2} 1(h_0 \geq \frac{s_1}{p}) - (h_0 - y_0(s_1)) \frac{r_1}{r_2} 1(h_0 \geq y_0(s_1)) \right] \right) > 0.
\]
Next, let sequence \( q_d(m) = \frac{|y_0(d) - \frac{d}{n}|}{\sqrt{m}} \) for \( d \in \{0, 1, \ldots\} \). We define set \( \mathcal{D} \) as follows

\[
\mathcal{D} = \{ d \in \{0, 1, \ldots\} : \lim_{m \to \infty} q_d(m) > 0 \}.
\]

By definition \( q_d(m) \) is non-negative for all \( d \), and hence we define the complement of set \( \mathcal{D} \) as follows

\[
\mathcal{D}^c = \{ d \in \{0, 1, \ldots\} : \lim_{m \to \infty} q_d = 0 \}.
\] (16)

Then, we can write the inner term of Equation (15) as

\[
E_{s_1 \sim \text{Bin}(h_0, p)} \left[ \left( h_0 - \frac{s_1}{p} \right) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq \frac{s_1}{p}) - (h_0 - y_0(s_1)) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq y_0(s_1)) \right]
\]

\[
= \sum_{d \in \mathcal{D}} \Pr(s_1 = d) \left[ \left( h_0 - \frac{d}{p} \right) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d)) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq y_0(d)) \right]
\]

\[
+ \sum_{d \in \mathcal{D}^c} \Pr(s_1 = d) \left[ \left( h_0 - \frac{d}{p} \right) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d)) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq y_0(d)) \right].
\]

We first show that

\[
\lim_{m \to \infty} \frac{1}{\sqrt{m}} \sum_{d \in \mathcal{D}^c} \Pr(s_1 = d) \left[ \left( h_0 - \frac{d}{p} \right) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d)) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq y_0(d)) \right] = 0. \tag{17}
\]

Observe that \( \sum_{d \in \mathcal{D}^c} \Pr(s_1 = d) \leq 1 \), and hence we have

\[
\lim_{m \to \infty} \frac{1}{\sqrt{m}} \sum_{d \in \mathcal{D}^c} \Pr(s_1 = d) \left[ \left( h_0 - \frac{d}{p} \right) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d)) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq y_0(d)) \right]
\]

\[
\leq \lim_{m \to \infty} \frac{1}{\sqrt{m}} \max_{d \in \mathcal{D}^c} \left[ \left( h_0 - \frac{d}{p} \right) \frac{r_1}{r_2} \mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d)) \mathbf{1}(h_0 \geq y_0(d)) \right]. \tag{18}
\]

The inner term of the above equation can then be written as

\[
\left( h_0 - \frac{d}{p} \right) \mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d)) \mathbf{1}(h_0 \geq y_0(d))
\]

\[
= \left( \left( h_0 - \frac{d}{p} \right) \mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d)) \mathbf{1}(h_0 \geq y_0(d)) \right) \mathbf{1} \left( \frac{d}{p} < y_0(d) \right)
\]

\[
+ \left( \left( h_0 - \frac{d}{p} \right) \mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d)) \mathbf{1}(h_0 \geq y_0(d)) \right) \mathbf{1} \left( \frac{d}{p} \geq y_0(d) \right).
\]

Notice that Notice that

\[
\left( \left( h_0 - \frac{d}{p} \right) \mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d)) \mathbf{1}(h_0 \geq y_0(d)) \right) \mathbf{1} \left( \frac{d}{p} < y_0(d) \right)
\]

\[
= \left( h_0 - \frac{d}{p} + h_0 - y_0(d) \right) \mathbf{1}(h_0 \geq y_0(d)) + \left( h_0 - \frac{d}{p} \right) \mathbf{1} \left( \frac{d}{p} \leq h_0 \leq y_0(d) \right) \mathbf{1} \left( \frac{d}{p} < y_0(d) \right)
\]

\[
= \left( y_0(d) - \frac{d}{p} \right) \mathbf{1}(h_0 \geq y_0(d)) + \left( h_0 - \frac{d}{p} \right) \mathbf{1} \left( \frac{d}{p} \leq h_0 \leq y_0(d) \right) \mathbf{1} \left( \frac{d}{p} < y_0(d) \right)
\]

\[
\leq \left( y_0(d) - \frac{d}{p} \right) + \left( y_0(d) - \frac{d}{p} \right) \mathbf{1} \left( \frac{d}{p} < y_0(d) \right)
\]

\[
\leq 2 \left( y_0(d) - \frac{d}{p} \right) \mathbf{1} \left( \frac{d}{p} < y_0(d) \right).
\]

Similarly, we have

\[
\left( \left( h_0 - \frac{d}{p} \right) \mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d)) \mathbf{1}(h_0 \geq y_0(d)) \right) \mathbf{1} \left( \frac{d}{p} \geq y_0(d) \right) \leq 2 \left( y_0(d) - \frac{d}{p} \right) \mathbf{1} \left( \frac{d}{p} \geq y_0(d) \right).
\]
Therefore, the inner term of Equation (18) is upper bounded by
\[(h_0 - \frac{d}{p})\mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d))\mathbf{1}(h_0 \geq y_0(d)) \leq 2|y_0(d) - \frac{d}{p}|.\] (19)

Using this in Equation (19) gives us
\[
\lim_{m \to \infty} \frac{1}{\sqrt{m}} \sum_{d \in \mathcal{D}} \frac{r_1}{\sqrt{m}} \max_{d \in \mathcal{D}^c} \left[ (h_0 - \frac{d}{p})\mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d))\mathbf{1}(h_0 \geq y_0(d)) \right]
\leq \lim_{m \to \infty} \frac{1}{\sqrt{m}} \sum_{d \in \mathcal{D}} \frac{r_1}{\sqrt{m}} \max_{d \in \mathcal{D}^c} \left[ 2|y_0(d) - \frac{d}{p}| \right]
= \frac{2r_1}{r_2} \max_{d \in \mathcal{D}^c} \frac{|y_0(d) - \frac{d}{p}|}{\sqrt{m}},
\]

which is what we wanted to show. Note that the last equation holds because of definition of \(\mathcal{D}^c\) in Equation (16).

By Equation (15) and the inequality above we have
\[
\lim_{m \to \infty} \frac{1}{\sqrt{m}} \sum_{d \in \mathcal{D}} \text{Pr}(s_1 = d) \left[ (h_0 - \frac{d}{p})\mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d))\mathbf{1}(h_0 \geq y_0(d)) \right] > 0. \quad (20)
\]

Notice that for \(d \in \mathcal{D}\), we have \(\lim_{m \to \infty} \frac{|y_0(d) - \frac{d}{p}|}{\sqrt{m}} > 0\). We define \(c_1, c_2 > 0\) be the constant such that
\[
\sup_{d \in \mathcal{D}} \lim_{m \to \infty} \frac{|y_0(d) - \frac{d}{p}|}{\sqrt{m}} = c_1 \quad \text{and} \quad \inf_{d \in \mathcal{D}} \lim_{m \to \infty} \frac{|y_0(d) - \frac{d}{p}|}{\sqrt{m}} = c_2.
\]

Recall that \(\mathcal{D} = \{d \in \{0, 1, \ldots\} : \lim_{m \to \infty} q(d)(m) > 0\}\), where \(q(d)(m) = \frac{|y_0(d) - \frac{d}{p}|}{\sqrt{m}}\) for \(d \in \{0, 1, \ldots\}\).

By Equation (15), we have
\[
(h_0 - \frac{d}{p})\mathbf{1}(h_0 \geq \frac{d}{p}) - (h_0 - y_0(d))\mathbf{1}(h_0 \geq y_0(d)) \leq 2\frac{r_1}{r_2} |y_0(d) - \frac{d}{p}|.
\]

Therefore, Equation (20) can be upper bounded as
\[
0 < \lim_{m \to \infty} \frac{1}{\sqrt{m}} \sum_{d \in \mathcal{D}} \text{Pr}(s_1 = d)2\frac{r_1}{r_2} |y_0(d) - \frac{d}{p}| \leq 2c_1 \frac{r_1}{r_2} \lim_{m \to \infty} \sum_{d \in \mathcal{D}} \text{Pr}(s_1 = d),
\]

which implies that
\[
\lim_{m \to \infty} \sum_{d \in \mathcal{D}} \text{Pr}(s_1 = d) > 0. \quad (22)
\]

We are now ready to show the contraction. Recall that in the current case, we assumed that \(\lim_{m \to \infty} \frac{\mathbb{E}_s[y_0(s_1) - h_0]}{\sqrt{m}} = 0\). However, we have
\[
\lim_{m \to \infty} \frac{\mathbb{E}_s[y_0(s_1) - h_0]}{\sqrt{m}} = \lim_{m \to \infty} \frac{1}{\sqrt{m}} \left( \sum_{d \in \mathcal{D}} \text{Pr}(s_1 = d)(y_0(d) - h_0) + \sum_{d \in \mathcal{D}^c} \text{Pr}(s_1 = d)(y_0(d) - h_0) \right)
= \lim_{m \to \infty} \frac{1}{\sqrt{m}} \left( \sum_{d \in \mathcal{D}} \text{Pr}(s_1 = d)(y_0(d) - \frac{d}{p} - \frac{d}{p} - h_0) + \sum_{d \in \mathcal{D}^c} \text{Pr}(s_1 = d)(\frac{d}{p} - h_0) \right)
= \lim_{m \to \infty} \frac{1}{\sqrt{m}} \left( \sum_{d \in \mathcal{D}} \text{Pr}(s_1 = d)(y_0(d) - \frac{d}{p} + \frac{d}{p} - h_0) + \sum_{d \in \mathcal{D}^c} \text{Pr}(s_1 = d)(\frac{d}{p} - h_0) \right)
\]
\[
\beta \quad \text{A.1. Proof of Lemma 9}
\]

\[
\frac{1}{\sqrt{m}} \left( \sum_{d \in [h_0]} \left( \frac{d}{p} - h_0 \right) + \sum_{d \in D} \Pr(s_1 = d)(y_0(d) - \frac{d}{p}) \right)
\]

\[
\geq \lim_{m \to \infty} \sum_{d \in D} \Pr(s_1 = d)c_2
\]

\[
> 0.
\]

The first inequality holds because \(D^c = \{d \in \{0, 1, \ldots \} : \lim_{m \to \infty} q_d(m) = 0 \}\), where \(q_d(m) = \frac{\lceil y_0(s_1) \rceil - m}{\sqrt{m}}\) for \(d \in \{0, 1, \ldots \}\). The third equation is due to definition of \(D^c\). The last inequality is from Equation (22). This is a contradiction to \(\lim_{m \to \infty} \max_{s \in [h, h]} \frac{E_s[|y_0(s_1) - \beta|]}{\sqrt{m}} = 0\).

Put the two cases together, we conclude that there does not exist such mapping \(y_0(\cdot)\) with which the competitive ratio on family of \(F\) is \(1 - o(\frac{1}{\sqrt{m}})\).

**A.1. Proof of Lemma 9**

The proof has two parts. In the first part, we present an upper bound on \(\max_{s \in [h, h]} E_s \left[ \frac{\text{loss}(I, s, z)}{\text{OPT}(I)} \right] \) and in the second part, we present a lower bound on the same quantity.

**First part: upper bound.** By definition, \(z(s_1) = \frac{1}{p} s_1\) for any \(s_1 \in \{0, 1, \ldots, \tilde{h}\}\), and hence we have \(n_1 - z(s_1) = h - s_1 - \frac{1}{p} s_1 = h - \frac{1}{p} s_1\). Similarly, \(z(s_1) - n_1 = \frac{1}{p} s_1 - h\). Then, by Equation (10), we have

\[
\max_{s \in [h, h]} E_s \left[ \frac{\text{loss}(I, s, z)}{\text{OPT}(I)} \right] \leq \max_{s \in [h, h]} E_s \left[ \frac{(n_1 - z(s_1))r_1 1(n_1 \geq z(s_1)) + (z(s_1) - n_1)r_2}{mp^2} \right]
\]

\[
= \frac{1}{m} \max_{s \in [h, h]} E_s \left[ (h - \frac{1}{p} s_1) \frac{r_1}{r_2} 1(h - \frac{1}{p} s_1 \geq 0) + (\frac{1}{p} s_1 - h) \right]
\]

\[
= \frac{1}{m} \max_{s \in [h, h]} E_s \left[ (h - \frac{1}{p} s_1) \frac{r_1}{r_2} 1(h - \frac{1}{p} s_1 \geq 0) \right]
\]

\[
= \frac{r_1}{mp^2} \max_{s \in [h, h]} E_s \left[ (hp - s_1)1(hp - s_1 \geq 0) \right].
\]

The second to last equation follows because \(E_s[\frac{1}{p} s_1 - h] = 0\). Recall that \(s_1 \sim \text{Bin}(h, p)\). Next, we bound \(E_s[(hp - s_1)1(hp - s_1 \geq 0)]\). We will show that

\[
\max_{s \in [h, h]} E_s [(hp - s_1)1(hp - s_1 \geq 0)] \leq \frac{p^2}{2} \sqrt{1 - p\sqrt{h}} + \frac{\sqrt{2\beta}}{\sqrt{mp^2}}.
\]

(23)

where \(\beta = 0.4215 \cdot \frac{p^2 (1-p)^2}{p(1-p)}\). This confirms that

\[
\max_{s \in [h, h]} E_s \left[ \frac{\text{loss}(I, s, z)}{\text{OPT}(I)} \right] \leq \frac{r_1}{mp^2} \left( \frac{p^2}{2} \sqrt{1 - p\sqrt{h}} + \frac{\sqrt{2\beta}}{\sqrt{mp^2}} \right)
\]

\[
= \frac{r_1}{2p\sqrt{m}} + \frac{\sqrt{2\beta} r_1}{2p\sqrt{m}},
\]

which is the desired result.

It remains to show Equation (23). By Nagaev and Chebotarev (2011), we have

\[
\left| E_s [(hp - s_1)1(hp - s_1 \geq 0)] - \frac{1}{2} E_s [hp - s_1] \right| \leq \frac{\beta}{\sqrt{h}},
\]

(24)
In addition, Berend and Kontorovich (2013) provides a sharp estimate of $\mathbb{E}_s[|hp - s_1|]$: 
\[
\frac{1}{\sqrt{2}} \sqrt{p(1-p)h} \leq \mathbb{E}_s[|hp - s_1|] \leq \sqrt{p(1-p)h}.
\] (25)
This leads to 
\[
\mathbb{E}_s[|hp - s_1|1(hp - s_1 \geq 0)] \leq \frac{1}{2} \mathbb{E}_s[|hp - s_1|] + \frac{\beta}{\sqrt{h}} \leq \frac{1}{2} \sqrt{p(1-p)h} + \frac{\beta}{\sqrt{h}}.
\] (26)
Therefore, for large enough $m$ (recall that $h = mp$ and $\bar{h} = mp/2$), we have 
\[
\max_{h \in [\bar{h}, h]} \mathbb{E}_s[|hp - s_1|1(hp - s_1 \geq 0)] \leq \sqrt{p(1-p)h} + 0.4215 \frac{\beta}{\sqrt{h}} = \frac{p}{2} \sqrt{1-p} \sqrt{m} + \frac{\sqrt{2} \beta}{\sqrt{pm}}
\] (27)
which is the desired result.

Second part: lower bound. Here, we present a lower bound on $\max_{h \in [\bar{h}, h]} \mathbb{E}_s[L(I, s, z)_{OPT(I)}]$, following similar steps in the first part. By definition,
\[
\max_{h \in [\bar{h}, h]} \mathbb{E}_s \left[ \frac{\text{LOSS}(I, s, z)}{\text{OPT}(I)} \right] \geq \max_{h \in [\bar{h}, h]} \mathbb{E}_s \left[ \frac{(n_1 - z(s_1))r_1 1(n_1 \geq z(s_1)) + (z(s_1) - n_1)r_2}{\sqrt{m} r_1} \right]
\]
\[
\geq \frac{1}{m} \max_{h \in [\bar{h}, h]} \mathbb{E}_s \left[ (h - \frac{1}{p} s_1)1(h - \frac{1}{p} s_1 \geq 0) + (\frac{1}{p} s_1 - h) \right]
\]
\[
= \frac{1}{mp} \max_{h \in [\bar{h}, h]} \mathbb{E}_s \left[ (h - \frac{1}{p} s_1)1(h - \frac{1}{p} s_1 \geq 0) \right]
\]
\[
\geq \frac{1}{mp} \max_{h \in [\bar{h}, h]} \left( \frac{1}{2} \mathbb{E}_s[|hp - s_1|] - \frac{\beta}{\sqrt{h}} \right)
\]
\[
\geq \frac{1}{mp} \max_{h \in [\bar{h}, h]} \left( \frac{1}{2} \sqrt{p(1-p)h} - \frac{\beta}{\sqrt{h}} \right)
\]
\[
= \frac{1}{mp} \left( \frac{1}{2 \sqrt{2}} \sqrt{p(1-p)pm} - \frac{\beta}{\sqrt{pm}} \right)
\]
\[
= \frac{p}{2 \sqrt{2}} \sqrt{(1-p)} \frac{1}{\sqrt{m}} - \frac{\beta}{p \sqrt{p} \sqrt{m}}.
\]
The second last inequality is due to Equation [24]. The last inequality is due to Equation (25).

Appendix B: Proof of Statements in Section 6

B.1. Proof of Lemma 1

We will show that for any realization of $(n_1, n_2)$ and sample information $\psi$, we have $\text{REW}_A(I_{order}, \psi) \leq \text{REW}_A(I, \psi)$. This implies that $\frac{\text{REW}_A(I_{order}, \psi)}{\text{OPT}(I_{order})} \leq \frac{\text{REW}_A(I, \psi)}{\text{OPT}(I)}$, as $\text{OPT}(I) = \text{OPT}(I_{order})$. The last inequality is the desired result.

It remains to show that $\text{REW}_A(I_{order}, \psi) \leq \text{REW}_A(I, \psi)$. First consider a case where under $\psi$, $\hat{r}_1 > \hat{r}_2$. In this case, Algorithm 2 assigns a protection level of $x_1 = s_1(1-p)/p$ to type 1 agents. Then, it is clear that (i) the number of accepted type 2 agents under $I_{order}$ is larger than or equal to that under $I$, and (ii) the number of accepted type 1 agents under $I_{order}$ is smaller than or equal to that under $I$. Hence in this case, Algorithm 2 under $I_{order}$ obtains higher reward than $I$. On the other hand, when $\hat{r}_1 \leq \hat{r}_2$, Algorithm 2 assigns a protection level of $x_2 = s_2(1-p)/p$ to type 2 agents. Again, in this case, we have (i) the number of accepted type 2 agents under $I_{order}$ is larger than or equal to that under $I$, and (ii) the number of accepted type 1 agents under $I_{order}$ is smaller than or equal to that under $I$. Hence, we have $\text{REW}_A(I_{order}, \psi) \leq \text{REW}_A(I, \psi)$.
B.2. Proof of Lemma 4

Define event $\mathcal{E}_0$ as $\{n_2 \leq m - \frac{1-p}{p} s_1\}$. We provide a lower bound on $\mathbb{E}_s \left[ \frac{\text{rew}_A(n, \psi; \mathcal{G})}{\text{opt}(n)} \right]$ as follows:

$$
\inf_{(h, \ell) \in \mathbb{R}_2} \mathbb{E}_s \left[ \frac{\text{rew}_A(n, \psi; \mathcal{G})}{\text{opt}(n)} \right] \\
= \inf_{(h, \ell) \in \mathbb{R}_2} \mathbb{E}_s \left[ \frac{\text{rew}_A(n, \psi; \mathcal{G})}{\text{opt}(n)} \right] \mathbb{E}_s \left[ \mathcal{E}_0^c \right] + \mathbb{E}_s \left[ \frac{\text{rew}_A(n, \psi; \mathcal{G})}{\text{opt}(n)} \right] \mathbb{P}(\mathcal{E}_0) \\
\geq \inf_{(h, \ell) \in \mathbb{R}_2} \mathbb{E}_s \left[ \frac{\text{rew}_A(n, \psi; \mathcal{G})}{\text{opt}(n)} \right] \mathbb{P}(\mathcal{E}_0) \\
\left(28\right)
$$

In light of Equation (28), we will provide a lower bound for $\mathbb{E}_s \left[ \frac{\text{rew}_A(n, \psi; \mathcal{G})}{\text{opt}(n)} \right]$ and $\mathbb{P}(\mathcal{E}_0)$.

**Part 1: bounding $\mathbb{E}_s \left[ \frac{\text{rew}_A(n, \psi; \mathcal{G})}{\text{opt}(n)} \right]$.** By Equation (29) in the proof of Lemma 2 for a fixed realization of $s$ (or equivalently $n$), we have

$$
\text{rew}_A(n, \psi; \mathcal{G}) = \min \left\{ n_2, \left( m - \frac{1-p}{p} s_1 \right)^+ \right\} \cdot r_2 \\
+ \min \left\{ m - \min \left\{ n_2, \left( m - \frac{1-p}{p} s_1 \right)^+ \right\}, n_1 \right\} \cdot r_1 \\
\left(29\right)
$$

Conditional on $\mathcal{E}_0 = \{n_2 \leq m - \frac{1-p}{p} s_1\}$, we have $\min\{n_2, (m - \frac{1-p}{p} s_1)^+\} = n_2$, and $\min\{m - \min\{n_2, (m - \frac{1-p}{p} s_1)^+\}, n_1\} \geq \min\{m - n_2, n_1\}$. Then, we have

$$
\frac{\text{rew}_A(n, \psi; \mathcal{G})}{\text{opt}(n)} \geq \frac{\text{rew}_A(n, \psi; \mathcal{G})}{\text{opt}(n)} \\
= \frac{n_1 r_1 + \min\{m - n_2, n_1\} r_2}{n_2 r_1 + \min\{m - n_2, n_1\} r_2} \\
= \frac{n_2 r_2 + \min\{m - n_2, n_1\} r_2}{n_1 r_1 + \min\{m - n_2, n_1\} r_2} \\
\geq \min \left\{ \frac{n_1 r_1 + n_2 r_2}{n_1 r_1 + \min\{m - n_2, n_1\} r_2} \right\} \\
= \min \left\{ \frac{(m - n_2) r_1 + n_2 r_2}{(m - n_1) r_2} \right\} \\
\left(30\right)
$$

The second inequality is because when $\min\{m - n_2, n_1\} = n_1$, we have $\min\{m - n_1, n_2\} = n_2$, and when $\min\{m - n_2, n_1\} = m - n_2$, we have $\min\{m - n_1, n_2\} = m - n_1$.

**Part 2: bounding $\mathbb{P}(\mathcal{E}_0)$.** We want to show that $\mathbb{P}(\mathcal{E}_0)$ is a high probability event for $(h, \ell) \in \mathbb{R}_2$. As $\ell \in \sqrt{m}$ for any $(h, \ell) \in \mathbb{R}_2$, we have $n_2 \leq \sqrt{m}$. Then, because $\mathcal{E}_0 = \{n_2 \leq m - \frac{1-p}{p} s_1\}$, we have

$$
\mathbb{P}(\mathcal{E}_0) = \mathbb{P} \left( \frac{1-p}{p} s_1 \leq m - n_2 \right) \geq \mathbb{P} \left( \frac{1-p}{p} s_1 \leq m - \sqrt{m} \right) = \mathbb{P} \left( s_1 \leq \frac{p(m - \sqrt{m})}{1-p} \right) \\
\left(31\right)
$$

Next, we show that with high probability, the upper bound of $s_1$ is no more than $\frac{p(m - \sqrt{m})}{1-p}$. Because $s_1 \sim \text{Bin}(h, p)$, by Hoeffding’s inequality, we have

$$
\mathbb{P}(\left| s_1 - hp \right| \leq \sqrt{h}) \geq 1 - \frac{1}{h^2} \geq 1 - \frac{1}{m}. \\
$$

This implies that $s_1 \leq hp + \sqrt{h}$ happens with high probability. We then have

$$
1 - \frac{1}{m} \leq \mathbb{P}(s_1 \leq hp + \sqrt{h}) \leq \mathbb{P}(s_1 \leq ho + \sqrt{h_0}) \leq \mathbb{P} \left( s_1 \leq \frac{p(m - \sqrt{m})}{1-p} \right), \\
\left(31\right)
$$

Electronic copy available at: https://ssrn.com/abstract=4054796
where the second inequality holds because $hp + \sqrt{h}$ is an increasing function, and $h < h_0$ in region $R_2$. These imply that $hp + \sqrt{p} \leq h_0p + \sqrt{h_0}$. The last inequality holds because $\frac{p(m+\sqrt{m})}{1-p} \geq h_0p + \sqrt{h_0}$.

Recall that $h_0 = \min\{y \geq 0 : \frac{p(m+\sqrt{m})}{1-p} \geq yp + \sqrt{y}\}$.

By Equation (30) and (31), we have

$$\inf_{(h, \ell) \in R_2} \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \mathbb{E}_0 \Pr(E_0) \geq (1 - \frac{1}{\sqrt{m}})(1 - \frac{1}{m})$$

which is the desired result.

**B.3. Proof of Lemma 5**

We define event $E_1 = \{n_1 > \frac{1-p}{p}s_1\}$. We then have

$$\inf_{(h, \ell) \in R_2} \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \geq \inf_{(h, \ell) \in R_2} \left\{ \min \left\{ \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1 \right\}, \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1^C \right\} \right\}$$

(33)

We will provide the lower bound of $\inf_{(h, \ell) \in R_2} \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1^C$ in the first part of the proof, and the lower bound of $\inf_{(h, \ell) \in R_2} \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1^C$ in the second part of the proof.

*Part 1: bounding* $\inf_{(h, \ell) \in R_2} \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1^C$. We then have

$$\inf_{(h, \ell) \in R_2} \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1^C \geq \inf_{(h, \ell) \in R_2} \min \left\{ \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1, E_0^C \right\}, \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1, E_0 \right\}$$

where we recall $E_0^C = \{n_2 > m - \frac{1-p}{p}s_1\}$.

To provide a lower bound for $\mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1, E_0^C$, we note that under event $E_0^C$, we have $n_2 > m - \frac{1-p}{p}s_1$, and hence

$$\mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1, E_0^C \geq \frac{(m - \frac{1-p}{p}s_1)^+ r_2 + \min\{m, \frac{1-p}{p}s_1\} r_1}{mr_1} \geq \frac{(m - n_2) r_1}{mr_1} \geq 1 - \frac{1}{\sqrt{m}}$$

(34)

Then, we give a lower bound to $\mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1, E_0$. Conditional on $E_1$ and $E_0$, we have $\frac{1-p}{p}s_1 \geq m - n_2$ and $\frac{1-p}{p}s_1 \leq n_1$. That is, the number of type 1 agents in the online arrival sequence is greater than or equal to the protection level of Algorithm 2 for type 1 agents (i.e., $\frac{1-p}{p}s_1 \leq n_1$), and in addition, the algorithm ends up rejecting some of type 2 agents as $\frac{1-p}{p}s_1 \geq m - n_2$. Therefore, we have

$$\mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1, E_0 \geq \frac{(m - \frac{1-p}{p}s_1)^+ r_2 + \min\{m, \frac{1-p}{p}s_1\} r_1}{mr_1} \geq \frac{(m - n_2) r_1}{mr_1} \geq 1 - \frac{1}{\sqrt{m}}$$

where the last inequality holds because $n_2 \leq \ell \leq \sqrt{m}$ for any $(h, \ell) \in R_2$.

*Part 2: bounding* $\inf_{(h, \ell) \in R_2} \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] | E_1^C$. Recall that for a fixed realization of $s$ (or equivalently $n$),

$$\text{REW}_A(n, \psi; G) = \min \left\{ n_2, \left( m - \frac{1-p}{p}s_1 \right)^+ \right\} \cdot r_2 + \min \left\{ m - \min \left\{ n_2, \left( m - \frac{1-p}{p}s_1 \right)^+ \right\}, n_1 \right\} \cdot r_1$$

Electronic copy available at: https://ssrn.com/abstract=4054796
Conditional on $\mathcal{E}_1^c = \{n_1 \leq \frac{1-p}{p}s_1\}$, we do not receive enough number of type 1 agents and thus, we have

\[
\min \left\{ m - \min \left\{ n_2, \left( m - \frac{1-p}{p}s_1 \right)^+ \right\}, n_1 \right\} = \min \left\{ \max \left\{ m - n_2, \min\{m, \frac{1-p}{p}s_1\} \right\}, n_1 \right\} = n_1.
\]

That is, Algorithm 2 accepts all type 1 agents in the online arrival sequence. Then, we have

\[
E_s \left[ \frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \Big| \mathcal{E}_1^c \right] \geq E_s \left[ \frac{\min \left\{ n_2, \left( m - \frac{1-p}{p}s_1 \right)^+ \right\} \cdot r_2 + n_1 r_1}{mr_1} \Big| \mathcal{E}_1^c \right] \geq \frac{1}{m} E_s \left[ n_1 \Big| \mathcal{E}_1^c \right].
\]

We will show that $E_s \left[ n_1 \Big| \mathcal{E}_1^c \right] \geq h_1$, where $h_1 = h_0(1-p) - \sqrt{p(1-p)h_0} - \frac{\beta}{\sqrt{n_0}}$ and $\beta = 0.4215 \cdot \frac{\sqrt{s + (1-p)^2}}{p(1-p)}$.

Since $\mathcal{E}_1^c = \{n_1 \leq \frac{1-p}{p}s_1\} = \{h(1-p) - n_1 \geq 0\}$, we have

\[
E_s \left[ n_1 \Big| \mathcal{E}_1^c \right] = h(1-p) - E_s \left[ h(1-p) - n_1 \Big| h(1-p) - n_1 \geq 0 \right] \geq h(1-p) - E_s \left[ \left| h(1-p) - n_1 \right| \right] - \frac{\beta}{\sqrt{h}} \geq h(1-p) - \sqrt{p(1-p)h} - \frac{\beta}{\sqrt{h}} \geq h_0(1-p) - \sqrt{p(1-p)h} - \frac{\beta}{\sqrt{h}} = h_1,
\]

where the first inequality holds because by Nagaev and Chebotarev [2011], we have $\left| E_s \left[ (h(1-p) - n_1) \Big| \{h(1-p) - n_1 \geq 0\} - E_s \left[ \left| h(1-p) - n_1 \right| \right]\right] \leq \frac{\beta}{\sqrt{n}}$. The second inequality follows from Berend and Kontorovich [2013] that shows $E_s \left[ \left| h(1-p) - n_1 \right| \right] \leq \sqrt{p(1-p)h}$. The third inequality holds because when $h > h_0$, $h(1-p) - \sqrt{p(1-p)h} - \frac{\beta}{\sqrt{h}}$ is a non-decreasing function for any $p$. To see why note that

\[
\frac{\partial(h(1-p) - \sqrt{p(1-p)h} - \frac{\beta}{\sqrt{h}})}{\partial h} \geq (1-p) - \frac{1}{2} \sqrt{p(1-p)} \frac{1}{\sqrt{h}} \geq 0. \tag{35}
\]

where the last inequality holds for any $h \geq \frac{1}{4} \frac{p}{1-p}$, and we have $h_0 > \frac{m - \sqrt{m}}{1-p} > \frac{1}{4} \frac{p}{1-p}$ for all $m \geq 2$.

**B.4. Proof of Lemma [7]**

We start by defining the event

\[
\mathcal{E}_2(h, \ell) = \left\{ s_1 \in [hp - \sqrt{h}, hp + \sqrt{h}] \cap \{s_2 \in [\ell p - \sqrt{\ell}, \ell p + \sqrt{\ell}] \right\}.
\]

Under event $\mathcal{E}_2(h, \ell)$, both $s_1$ and $s_2$ are concentrated around their average. Recall that $s_1 \sim \text{Bin}(h, p)$ and $s_2 \sim \text{Bin}(\ell, p)$. By Equation (6), we have

\[
\inf \limits_{(h, \ell) \in \mathcal{R}_3} E_{\psi} \left[ \frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \right] \geq \inf \limits_{(h, \ell) \in \mathcal{R}_3} E_s \left[ \frac{\text{REW}_A(n, \psi; \mathcal{G}) \cdot \text{Pr}(\mathcal{G}(\psi; (h, \ell)))}{\text{OPT}(n)} \Big| \mathcal{E}_2(h, \ell) \right] \cdot \text{Pr}(\mathcal{E}_2(h, \ell)). \tag{37}
\]

We lower bound $\text{Pr}(\mathcal{E}_2(h, \ell))$ in Part 1 of the proof. Next, we explain how to handle the conditional expectation in Equation (37). We start by defining the following event

\[
\mathcal{E}_3 = \left\{ r_1 \in [r_1 - \frac{1}{\sqrt{s_1}}, r_1 + \frac{1}{\sqrt{s_1}}] \right\} \cap \left\{ r_2 \in [r_2 - \frac{1}{\sqrt{s_2}}, r_2 + \frac{1}{\sqrt{s_2}}] \right\}.
\]
Under event $E_3$, our estimates for $r_1$ and $r_2$ are concentrated around their true values. The conditional expectation in Equation (37) is lower bounded as follows:

$$\inf_{(h, \ell) \in R_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G) \cdot \Pr(G(\psi; (h, \ell)))}{\text{OPT}(n)} \right]_{E_2(h, \ell)} \geq \inf_{(h, \ell) \in R_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G) \cdot \Pr(G(\psi; (h, \ell)))}{\text{OPT}(n)} \right]_{E_2(h, \ell), E_3} \Pr(E_3|E_2(h, \ell))$$

(38)

Given (38), we lower bound $\Pr(E_3|E_2(h, \ell))$ in part 2 of the theorem. In part 3, conditioned on $E_2(h, \ell), E_3$, we show that $\Pr(G(\psi; (h, \ell)))$, for any realization $s$, equals to 1 in Part 3. Finally, in part 4, we lower bound $\inf_{(h, \ell) \in R_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G) \cdot \Pr(G(\psi; (h, \ell)))}{\text{OPT}(n)} \right]_{E_2(h, \ell), E_3}$.

Part 1: bounding $\Pr(E_2(h, \ell))$. By Hoeffding’s inequality, we have

$$\Pr(s_1 \in [hp - \sqrt{h}, hp + \sqrt{h}]) \geq 1 - \frac{1}{h^2} \geq 1 - \frac{1}{m},$$

$$\Pr(s_2 \in [\ell p - \sqrt{\ell}, \ell p + \sqrt{\ell}]) \geq 1 - \frac{1}{\ell^2} \geq 1 - \frac{1}{m}.$$ 

The last steps of both equations are because in region $R_3$, we have $h \geq \sqrt{m}$, $\ell \geq \sqrt{m}$. Then, because $s_1$ and $s_2$ are independent of each other, we have

$$\Pr(E_2(h, \ell)) = \Pr(s_1 \in [hp - \sqrt{h}, hp + \sqrt{h}]) \cdot \Pr(s_2 \in [\ell p - \sqrt{\ell}, \ell p + \sqrt{\ell}]) \geq (1 - \frac{1}{m})^2.$$ 

Part 2: bounding $\Pr(E_3|E_2(h, \ell))$. Conditioned on $s_1$, we know $\hat{r}_1 = \frac{\text{Bin}(s_1, r_1)}{s_1}$, and similarly conditioned on $s_2$, $\hat{r}_2 = \frac{\text{Bin}(s_2, r_2)}{s_2}$. Then, Hoeffding’s inequality implies that

$$\Pr\left(\hat{r}_1 \in [r_1 - \frac{1}{\sqrt{s_1}}, r_1 + \frac{1}{\sqrt{s_1}}] \mid s_1 = \hat{s}_1\right) \geq 1 - \frac{1}{\hat{s}_1^2}.$$ 

(39)

Now consider event $E_2(h, \ell)$ that implies that $s_1 \geq hp - \sqrt{h}$. As shown in the proof of Lemma 4, $hp - \sqrt{h}$ is an increasing function for any $h \geq \sqrt{m}$, and hence under event $E_2(h, \ell)$, $s_1 \geq hp - \sqrt{h} \geq \sqrt{mp} - m^{1/4}$. Applying this lower bound on $s_1$ to Equation (39) leads to

$$\Pr\left(\hat{r}_1 \in [r_1 - \frac{1}{\sqrt{s_1}}, r_1 + \frac{1}{\sqrt{s_1}}] \mid E_2(h, \ell)\right) \geq 1 - \frac{1}{(\sqrt{mp} - m^{1/4})^2}.$$ 

Note that the same inequality holds $r_2$ and its estimate. As a result, we have

$$\Pr(E_3|E_2(h, \ell)) \geq (1 - \frac{1}{(\sqrt{mp} - m^{1/4})^2}).$$

Part 3: bounding $\Pr(G(\psi; (h, \ell)))$ conditioned on events $E_2(h, \ell)$ and $E_3$. We show that conditional on $E_2(h, \ell)$ and $E_3$, for any realization $s$, $\Pr(G(\psi; (h, \ell))) = 1$. We show this by proving $\hat{r}_1 \geq \hat{r}_2$ always happens under events $E_2(h, \ell)$ and $E_3$. Under events $E_2(h, \ell)$ and $E_3$, we know that (i) $s_1, s_2 \geq \sqrt{mp} - m^{1/4}$ and (ii) $\hat{r}_1 \in [r_1 - \frac{1}{\sqrt{s_1}}, r_1 + \frac{1}{\sqrt{s_1}}]$ and $\hat{r}_2 \in [r_2 - \frac{1}{\sqrt{s_2}}, r_2 + \frac{1}{\sqrt{s_2}}]$. To show the result, we confirm that the lower bound on $\hat{r}_1$ is smaller than the upper bound on $\hat{r}_2$. That is,

$$r_2 + \frac{1}{\sqrt{s_2}} \leq r_1 - \frac{1}{\sqrt{s_1}} \quad \text{for any } s_1, s_2 \geq \sqrt{mp} - m^{1/4}.$$ 

Given the lower bound on $s_1$ and $s_2$, it suffices to show that

$$r_2 + \frac{1}{\sqrt{mp} - m^{1/4}} \leq r_1 - \frac{1}{\sqrt{mp} - m^{1/4}} \iff r_1 - r_2 \geq \frac{2}{m^{1/8} \sqrt{m^{1/4}p} - 1}.$$
Be definition of $m_1$, the above inequality holds for any $m \geq m_1$. Recall that $m_1 = \min_{y \geq 1/p^t} \{y : r_1 - r_2 > \frac{2}{y^{1/b} \sqrt{y^{1/p^t} - 1}}\}$.

Part 4: bounding $\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} | \mathcal{E}_2(h, \ell), \mathcal{E}_3 \right]$. First, observe that the ratio $\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}$ is independent with $\mathcal{E}_3$ because both the reward under the good event $\mathbb{E}[\text{REW}_A(n, i.e., \psi) | G(\psi; (h, \ell))]$, and the optimal reward, i.e., $\text{OPT}(n)$, do not depend on $\tilde{r}_1$ and $\tilde{r}_2$. We can simply delete $\mathcal{E}_3$, and bound $\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} | \mathcal{E}_2(h, \ell) \right]$. We begin by partition region $\mathcal{R}_3$ based the number of type 1 agents, (i.e., $h$):

$$\mathcal{R}_3^1 = \{(h, \ell) : (h, \ell) \in \mathcal{R}_3, h < h_0\} \quad \text{and} \quad \mathcal{R}_3^2 = \{(h, \ell) : (h, \ell) \in \mathcal{R}_3, h \geq h_0\},$$

where $h_0 = \min\{y \geq 0 : \frac{p(n^{-\sqrt{m}})}{1-p} \geq yp + \sqrt{y}\}$. Then, $\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} | \mathcal{E}_2(h, \ell) \right]$ is equal to the minimum of $\inf_{(h, \ell) \in \mathcal{R}_3^1} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} | \mathcal{E}_2(h, \ell) \right]$ and $\inf_{(h, \ell) \in \mathcal{R}_3^2} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} | \mathcal{E}_2(h, \ell) \right]$.

**Lemma 10.** Let $\mathcal{R}_3^1 = \{(h, \ell) : (h, \ell) \in \mathcal{R}_3, h < h_0\}$. We then have

$$\inf_{(h, \ell) \in \mathcal{R}_3^1} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} | \mathcal{E}_2(h, \ell) \right] \geq 1 - \frac{1}{m},$$

where $h_0 = \min\{y \geq 0 : \frac{p(n^{-\sqrt{m}})}{1-p} \geq yp + \sqrt{y}\}$ and event $\mathcal{E}_2(h, \ell)$ is defined in Equation (36).

The proof of Lemma 10 is similar to the proof of Lemma 4.

**Lemma 11.** Let $\mathcal{R}_3^2 = \{(h, \ell) : (h, \ell) \in \mathcal{R}_3, h \geq h_0\}$. We then have

$$\inf_{(h, \ell) \in \mathcal{R}_3^2} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} | \mathcal{E}_2(h, \ell) \right] \geq \frac{h_1}{m},$$

where $h_1 = h_0(1-p) - \sqrt{p(1-p)h_0 - \frac{\beta}{\sqrt{h_0}}}, \beta = 0.4215 \frac{p^2 + (1-p)^2}{p(1-p)}$, and $h_0 = \min\{y \geq 0 : \frac{p(n^{-\sqrt{m}})}{1-p} \geq yp + \sqrt{y}\}$. Further, event $\mathcal{E}_2(h, \ell)$ is defined in Equation (36).

The proof of Lemma 11 is similar to the proof of Lemma 5. Lemmas 10 and 11 imply that

$$\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \geq \min \left\{ 1 - \frac{1}{\sqrt{m}}, \frac{h_1}{m} \right\} = W.$$

Finally, by the results in Part 1, Part 2, Part 3, and Part 4, we have

$$\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \geq (1 - \frac{1}{m})^2 (1 - \frac{1}{(\sqrt{mp} - m^{1/4})^2})^2 W,$$

which is the desired result.

**B.5. Proof of Lemma 8**

In Equation (6), we have

$$\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi)}{\text{OPT}(n)} \right] = \inf_{(h, \ell) \in \mathcal{R}_3} \left\{ \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G) \cdot \Pr(G(\psi; (h, \ell)))}{\text{OPT}(n)} \right] + \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G^2) \cdot (1 - \Pr(G(\psi; (h, \ell))))}{\text{OPT}(n)} \right] \right\}.$$

(40)
Define event $E_5 = \{s_1 \neq 0\} \cap \{s_2 \neq 0\}$. We can lower bound Equation (40) by
\[
\inf_{(h, \ell) \in \mathbb{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi)}{\text{OPT}(n)} \right] \geq \inf_{(h, \ell) \in \mathbb{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi)}{\text{OPT}(n)} \right] \mathbb{P}(E_5)
\]
\[
\geq \inf_{(h, \ell) \in \mathbb{R}_3} \left\{ \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G) \cdot \mathbb{P}(G(h, \ell))}{\text{OPT}(n)} \right] \mathbb{P}(E_5) \right\}
\]
\[
+ \mathbb{E}_s \left[ \text{REW}_A(n, \psi; G^c) \cdot (1 - \mathbb{P}(G(h, \ell))) \right] \mathbb{P}(E_5) \mathbb{P}(E_5). \tag{41}
\]
Note that
\[
\mathbb{P}(E_5) = 1 - \mathbb{P}(E_5^c) \geq 1 - \mathbb{P}(s_1 = 0) - \mathbb{P}(s_2 = 0)
\]
\[
= 1 - \binom{h}{0}(1-p)^h - \binom{\ell}{0}(1-p)^\ell \geq 1 - 2 \left( \frac{\sqrt{m}}{0} \right)(1-p)^\sqrt{m} = 1 - 2(1-p)^\sqrt{m}. \tag{42}
\]
We denote $V$ as $1 - 2(1-p)^\sqrt{m}$. Furthermore, conditioned on event $E_5$, it is obvious that the probability of the good event $G(h, \ell) = \{\hat{r}_1 \geq \hat{r}_2\}$ is greater than or equal to $1/2$. This is because when event $E_5$ holds, (i) our estimate for $r_i$, $i \in \{1, 2\}$ is simply the sample average of the $s_i$ realized rewards, and (ii) we have $r_1 > r_2$. In light of Equation (41), if $\mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] < \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G^c)}{\text{OPT}(n)} \right]$, then Equation (41) is lower bounded by setting the probability of good event to be $\frac{1}{2}$. In light of Equation (43),
\[
\inf_{(h, \ell) \in \mathbb{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi)}{\text{OPT}(n)} \right] \mathbb{P}(E_5) \geq \inf_{(h, \ell) \in \mathbb{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \mathbb{P}(E_5). \tag{43}
\]
Otherwise, Equation (41) is lower bounded by setting the probability of good event to be $\frac{1}{2}$.
\[
\inf_{(h, \ell) \in \mathbb{R}_3} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi)}{\text{OPT}(n)} \right] \mathbb{P}(E_5) \geq \inf_{(h, \ell) \in \mathbb{R}_3} \frac{1}{2} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \mathbb{P}(E_5) + \frac{1}{2} \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G^c)}{\text{OPT}(n)} \right] \mathbb{P}(E_5). \tag{44}
\]
Given Equations (43) and (44), we need to bound $\mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \mathbb{P}(E_5)$ and $\mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G^c)}{\text{OPT}(n)} \right] \mathbb{P}(E_5)$. Notice that the former ratio can be lower bounded by
\[
\mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \mathbb{P}(E_5) \geq \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \mathbb{P}(E_5, E_2(h, \ell))
\]
\[
= \mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \mathbb{P}(E_2(h, \ell)) \mathbb{P}(E_5)
\]
where event $E_2(h, \ell) = \{s_1 \in [hp - \sqrt{\ell}, hp + \sqrt{\ell}] \cap \{s_2 \in [lp - \sqrt{\ell}, lp + \sqrt{\ell}] \}$. The equality holds because $E_2(h, \ell) \subset E_5$. Then, by Part 1 and Part 4 of the proof of Lemma 7, we have the former ratio is bounded by
\[
\mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)} \right] \mathbb{P}(E_5) \geq (1 - \frac{1}{m})^2 W, \tag{45}
\]
where $W = \min \left\{ 1 - \frac{1}{\sqrt{m}}, \frac{h_1}{m} \right\}$. Bounding the latter ratio, i.e., $\mathbb{E}_s \left[ \frac{\text{REW}_A(n, \psi; G^c)}{\text{OPT}(n)} \right] \mathbb{P}(E_5)$, is quite involved because when the bad event $G^c(h, \ell)$ happens, Algorithm 2 may over-protect or under-protect type 2 agents given the randomness in the sample information. Over-protecting type 2 agents leads to reject a certain number of type 1 agents and waste some of the resources. As $\ell$ is large in this case, the number of wasted resources can be large too. Under-protecting type 2 agents leads to accept too many type 1 agents. Nonetheless, we provide the lower
bound of the ratio below and show that it is still not too small in this case once we take expectation with
respect to \( s \).

We begin by partitioning region \( \mathcal{R}_3 \) based on the number of type 2 agents chosen by the adversary, (i.e. \( \ell \)):

\[
\mathcal{R}_3 = \{(h, \ell) : (h, \ell) \in \mathcal{R}_3, \ell > \ell_0 \}, \quad \mathcal{R}_3 = \{(h, \ell) : (h, \ell) \in \mathcal{R}_3, \ell \leq \ell_0 \},
\]

where \( \ell_0 = \min \{ y : (1 - p)y - \sqrt{y} \geq m \} \). Then, \( \inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_n \left[ \frac{REW_A(n, \psi; \mathcal{G}^c)}{\text{OPT}(n)} | \mathcal{E}_5 \right] \) is equal to the minimum of \( \inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_n \left[ \frac{REW_A(n, \psi; \mathcal{G}^c)}{\text{OPT}(n)} | \mathcal{E}_5 \right] \) and \( \inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_n \left[ \frac{REW_A(n, \psi; \mathcal{G}^c)}{\text{OPT}(n)} | \mathcal{E}_5 \right] \). We bound the first ratio (i.e., the one concerns \( \mathcal{R}_3 \)) in Lemma 12 and the second ratio (i.e., the one concerns \( \mathcal{R}_3 \)) will be bounded in Lemma 13.

**Lemma 12.** Let \( \mathcal{R}_3 = \{(h, \ell) : (h, \ell) \in \mathcal{R}_3, \ell > \ell_0 \} \), where \( \mathcal{R}_3 = \{(h, \ell) : h \geq \sqrt{m}, \ell \geq \sqrt{m} \} \), and \( \ell_0 = \min \{ y : (1 - p)y - \sqrt{y} > m \} \). Then,

\[
\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_n \left[ \frac{REW_A(n, \psi; \mathcal{G}^c)}{\text{OPT}(n)} | \mathcal{E}_5 \right] \geq (1 - \frac{1}{\ell_0^2}) \alpha,
\]

where \( \alpha = \frac{c_4}{\ell_1} \).

In this case, we show that with high probability, the number of type 2 agents in the online arrival sequence \( n_2 \geq m \). Then, the algorithm accepts \( m \) type 2 agents, and the ratio is bounded by \( \alpha \).

**Lemma 13.** Let \( \mathcal{R}_3 = \{(h, \ell) : (h, \ell) \in \mathcal{R}_3, \ell \leq \ell_0 \} \), where \( \mathcal{R}_3 = \{(h, \ell) : h \geq \sqrt{m}, \ell \geq \sqrt{m} \} \), and \( \ell_0 = \min \{ y : (1 - p)y - \sqrt{y} \geq m \} \). Let \( \beta = 0.4215 \cdot \frac{p^2 + (1 - p)^2}{m(1 - p)}, \ell_1 = \sqrt{p(1 - p)\ell_0} + \frac{\beta}{m} + \ell_0p \), then,

\[
\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_n \left[ \frac{REW_A(n, \psi; \mathcal{G}^c)}{\text{OPT}(n)} | \mathcal{E}_5 \right] \geq \min \left\{ \alpha, 1 - \frac{1 - p}{pm} \ell_1 \right\}.
\]

To show Lemma 13, we define event \( \mathcal{E}_4 = \{ n_2 \geq \frac{1 - p}{\ell_0} n_2 \} \). Under this event the number of type 2 agents is greater than or equal to the protection level assigned to this type of agents. Then, we provide the lower bound of \( \inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_n \left[ \frac{REW_A(n, \psi; \mathcal{G}^c)}{\text{OPT}(n)} | \mathcal{E}_5 \right] \) by further conditioning on either \( \mathcal{E}_4 \) or \( \mathcal{E}_4^c \). For the part with \( \mathcal{E}_4 \), we do not reject type 1 agents unless there is no resource remaining. Therefore, the lower bound can be bounded by \( \alpha \). For the part with \( \mathcal{E}_4^c \), we directly calculate how many resources we waste, and then provide a lower bound.

Finally, combining the results from Lemmas 12 and 13, we have

\[
\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_n \left[ \frac{REW_A(n, \psi; \mathcal{G}^c)}{\text{OPT}(n)} | \mathcal{E}_5 \right] \geq \min \left\{ (1 - \frac{1}{\ell_0^2}) \alpha, 1 - \frac{1 - p}{pm} \ell_1 \right\}.
\]

Combining the results from Equations (43), and (45), we have if

\[
(1 - \frac{1}{m})^2 W \leq \min \left\{ (1 - \frac{1}{\ell_0^2}) \alpha, 1 - \frac{1 - p}{pm} \ell_1 \right\},
\]

\[
\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_n \left[ \frac{REW_A(n, \psi; \mathcal{G}^c)}{\text{OPT}(n)} | \mathcal{E}_5 \right] \geq (1 - \frac{1}{m})^2 W.
\]

Otherwise, by Equations (44), and (46), we have

\[
\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_n \left[ \frac{REW_A(n, \psi; \mathcal{G}^c)}{\text{OPT}(n)} | \mathcal{E}_5 \right] \geq \frac{1}{2} (1 - \frac{1}{m})^2 W + \frac{1}{2} \min \left\{ (1 - \frac{1}{\ell_0^2}) \alpha, 1 - \frac{1 - p}{pm} \ell_1 \right\}.
\]

Therefore, combining the results in Equations (41), (42), (47), (48), we have

\[
\inf_{(h, \ell) \in \mathcal{R}_3} \mathbb{E}_n \left[ \frac{REW_A(n, \psi; \mathcal{G}^c)}{\text{OPT}(n)} | \mathcal{E}_5 \right] \geq V \cdot \min \left\{ (1 - \frac{1}{m^2}) W, \frac{1}{2} (1 - \frac{1}{m^2}) W + \frac{1}{2} \min \left\{ (1 - \frac{1}{\ell_0^2}) \alpha, 1 - \frac{1 - p}{pm} \ell_1 \right\} \right\},
\]

where recall that \( V = 1 - 2(1 - p)^m \).
B.6. Proof of Lemma 13

We have

\[
\inf_{(h,t) \in \mathcal{R}_k} \mathbb{E}_s \left[ \frac{\text{REW}_A (n, \psi; G^C)}{\text{OPT}(n)} \right] = \inf_{(h,t) \in \mathcal{R}_k} \min \left\{ \mathbb{E}_s \left[ \frac{\text{REW}_A (n, \psi; G^C)}{\text{OPT}(n)} \right] \mid \{n_2 \geq m\} \right\},
\]

\[
\mathbb{E}_s \left[ \frac{\text{REW}_A (n, \psi; G^C)}{\text{OPT}(n)} \right] \{n_2 < m\} \right\}
\]

(49)

We have discussed in Lemma 12 that \( \mathbb{E}_s \left[ \frac{\text{REW}_A (n, \psi; G^C)}{\text{OPT}(n)} \right] \{n_2 \geq m\} \geq \alpha \). Then, given Equation (49), we provide a lower bound to \( \mathbb{E}_s \left[ \frac{\text{REW}_A (n, \psi; G^C)}{\text{OPT}(n)} \right] \{n_2 < m\} \).

Conditional on event \( \{n_2 < m\} \), the algorithm accepts all \( n_2 \) type 2 agents. If \( n_2 \) is larger than the protection level \( \min \{m, \frac{1}{p} s_2\} \), the algorithm accepts as many type 1 agents as possible with the remaining \( m - n_2 \) units of the resource. If \( n_2 \) is smaller than the protection level \( \min \{m, \frac{1}{p} s_2\} \), the algorithm rejects some type 1 agents because we can accept at most \( (m - \frac{1}{p} s_2)^+ \) type 1 agents. Then, we have

\[
\mathbb{E}_s \left[ \frac{\text{REW}_A (n, \psi; G^C)}{\text{OPT}(n)} \right] \{n_2 < m\} \geq \mathbb{E}_s \left[ \frac{n_2 r_2 + \min \{n_1, m - \max \{n_2, \min \{m, \frac{1}{p} s_2\}\}\} r_1}{\min \{m, n_1\} r_1 + \min \{n_2, (m - n_1)^+\} r_2} \right].
\]

We define the event \( \mathcal{E}_4 = \{n_2 \geq \frac{1}{p} s_2\} \). Then, we have

\[
\mathbb{E}_s \left[ \frac{n_2 r_2 + \min \{n_1, m - \max \{n_2, \min \{m, \frac{1}{p} s_2\}\}\} r_1}{\min \{m, n_1\} r_1 + \min \{n_2, (m - n_1)^+\} r_2} \right] \geq \min \left\{ \mathbb{E}_s \left[ \frac{n_2 r_2 + \min \{n_1, m - \max \{n_2, \min \{m, \frac{1}{p} s_2\}\}\} r_1}{\min \{m, n_1\} r_1 + \min \{n_2, (m - n_1)^+\} r_2} \right] \mid \mathcal{E}_4 \right\},
\]

\[
\mathbb{E}_s \left[ \frac{n_2 r_2 + \min \{n_1, m - \max \{n_2, \min \{m, \frac{1}{p} s_2\}\}\} r_1}{\min \{m, n_1\} r_1 + \min \{n_2, (m - n_1)^+\} r_2} \right] \mid \mathcal{E}_4 \right\}.
\]

(50)

Conditioned on \( \mathcal{E}_4 \), for any realization \( s \), we have

\[
\frac{n_2 r_2 + \min \{n_1, m - \max \{n_2, \min \{m, \frac{1}{p} s_2\}\}\} r_1}{\min \{m, n_1\} r_1 + \min \{n_2, (m - n_1)^+\} r_2} \geq \frac{n_2 r_2 + \min \{n_1, (m - n_2)\} r_1}{\min \{m, n_1\} r_1 + \min \{n_2, (m - n_1)^+\} r_2}.
\]

\[
\geq \frac{mr_2}{mr_1} = \alpha.
\]

When \( \mathcal{E}_4 \) happens, the number of arriving type 2 agents is larger than the protection level, and the algorithm accepts all the arriving agents until there is no resource left. If OPT does not use all resources, then the algorithm accepts everyone, and hence the above ratio will be one. If OPT uses all resources, then the algorithm also uses all resources, and thus, this bound clearly holds.

Conditioned on \( \mathcal{E}_4 \), we have \( n_2 < \frac{1}{p} s_2 \). Then,

\[
\frac{n_2 r_2 + \min \{n_1, m - \max \{n_2, \min \{m, \frac{1}{p} s_2\}\}\} r_1}{\min \{m, n_1\} r_1 + \min \{n_2, (m - n_1)^+\} r_2} \geq \frac{n_2 r_2 + \min \{n_1, (m - \frac{1}{p} s_2)^+\} r_1}{\min \{m, n_1\} r_1 + \min \{n_2, (m - n_1)^+\} r_2}.
\]

\[
\geq \frac{n_2 r_2 + (m - \frac{1}{p} s_2)^+ r_1}{mr_1}.
\]
The inequality is because if $n_1 < m - \frac{1-p}{p} s_2$, then the algorithm accepts $n_2$ type 2 agents and $n_1$ type 1 agents, and the ratio is 1. Then, we have

\[
\mathbb{E}_a \left[ \frac{n_2 r_2 + \min\{n_1, m - \max\{n_2, \min\{m, \frac{1-p}{p} s_2\}\}\} r_1}{\min\{m, n_1\} r_1 + \min\{n_2, (m-n_1)^+\} r_2} \right] \geq \mathbb{E}_a \left[ \frac{n_2 r_2 + (m - \frac{1-p}{p} s_2)^+ r_1}{mr_1} \right] \geq \mathbb{E}_a \left[ \frac{(m - \frac{1-p}{p} s_2)^+ r_1}{mr_1} \right] \geq \mathbb{E}_a \left[ \frac{(m - \frac{1-p}{p} s_2) r_1}{mr_1} \right] = 1 - \frac{1-p}{pm} \mathbb{E}_a \left[ s_2 \mathcal{C}_4 \right].
\] (51)

Since $\mathcal{C}_4 = \{s_2 \geq \ell p\}$, we have

\[
\mathbb{E}_a \left[ s_2 \mathbb{1}_{s_2 \geq \ell p} \right] = \mathbb{E}_a \left[ s_2 \mathbb{1}_{s_2 \geq \ell p} \right] + \ell p \leq \mathbb{E}_a \left[ |s_2 - \ell p| + \frac{\beta}{\sqrt{\ell}} + \ell p \right] \leq \sqrt{p(1-p)} \ell + \frac{\beta}{\sqrt{\ell}} + \ell p \leq \sqrt{p(1-p)} \ell_0 + \frac{\beta}{\sqrt{m}} + \ell_0 p = \ell_1,
\]

where the first inequality holds because by Nagaev and Chebotarev (2011), we have $\mathbb{E}_a \left[ s_2 \mathbb{1}_{s_2 \geq \ell p} \right] - \mathbb{E}_a \left[ |s_2 - \ell p| \right] \leq \frac{\beta}{\sqrt{\ell}}$. The second inequality follows from Berend and Kontorovich (2013) that shows $\mathbb{E}_a \left[ |s_2 - \ell p| \right] \leq \sqrt{p(1-p)\ell}$.

Finally, we have

\[
\inf_{(h,t) \in \mathcal{R}_3} \mathbb{E}_a \left[ \frac{\text{REW}_A(n, \psi; \mathcal{C})}{\text{OPT}(n)} \right] \geq \min \left\{ \alpha, 1 - \frac{1-p}{pm} \ell_1 \right\}.
\]
Appendix C: Proof of Theorem 5

We construct the following random input distribution by taking advantage of input family $F$ that we used in the proof of Theorem 4. Recall that under a given input $h, \ell$ in family $F$, $I_h$ is the random arrival sequence such that $n_2$ type 2 agents arrive first followed by $n_1$ type 1 agents, where $n_1 \sim \text{Bin}(h, 1-p), n_2 \sim \text{Bin}(\ell, 1-p)$, and $\ell = \frac{n_000000}{p}$, and $h \in [\hat{h}, \tilde{h}]$ with $\hat{h} = \frac{1}{2}pm$ and $\tilde{h} = pm$. Then, in our random input distribution, we choose one of the feasible $h \in [\hat{h}, \tilde{h}]$ uniformly at random; that is $\Pr(\text{Unif}[\hat{h}, \tilde{h}] = h) = \frac{1}{\tilde{h} - \hat{h} + 1}$ for any $h \in [\hat{h}, \tilde{h}]$.

Observe that due to our sampling procedure, our model is random. Put differently, even when the input $h \in [\hat{h}, \tilde{h}]$ is realized, the online arrival sequence $I = I_h$ is still random. That prevents us from applying the Von Neuman/Yao principle [Seiden (2000)] to our setting. Nonetheless, we derive a result similar to the Von Neuman/Yao principle that can be applied to our setting.

**Lemma 14 (From Deterministic to Randomized Inputs).** Consider a setting where the adversary chooses a meta input $X \in \mathcal{X}$, and then based on the meta input $X$, a random input $\tilde{X}$ from a distribution $D(X)$ is realized. For any random or deterministic algorithm $A$ and meta input $X \in \mathcal{X}$, let $C_A(\tilde{X}) \in [0,1]$ be the (realized) reward of algorithm $A$ under input $\tilde{X}$ over the reward of the optimal in-hindsight algorithm that knows $\tilde{X}$ in advance, where $\tilde{X} \sim D(X)$. Then, $\mathbb{E}_{\tilde{X} \sim D(X)}[C_A(\tilde{X})]$ is the competitive ratio of algorithm $A$ under meta input $X$, where the expectation is with respect to $\tilde{X}$ and any randomness in algorithm $A$.

Now, let $\mathcal{A}_d$ be the set of all deterministic algorithms. Let $P$ be a probability distribution over any deterministic algorithms. Define $A \sim P$ as an algorithm chosen according to probability distribution $P$. By [Yao (1977), Ball and Queyranne (2009)], any randomized algorithm may be viewed as a random choice $A \sim P$ among deterministic algorithms. Further, let $Q$ be a probability distribution over the meta inputs in $\mathcal{X}$, and let $X \sim Q$ denote a random meta input chosen according to $Q$. Then, for any distribution $P$ over algorithms, distribution $Q$ over meta inputs, we have

$$\min_{\mathcal{A} \in \mathcal{A}_d} \mathbb{E}_{A \sim P} [\mathbb{E}_{\tilde{X} \sim D(X)}[C_A(\tilde{X})]] \leq \max_{\mathcal{A} \in \mathcal{A}_d} \mathbb{E}_{X \sim Q} [\mathbb{E}_{\tilde{X} \sim D(X)}[C_A(\tilde{X})]]].$$

(52)

We now ready to show the result using Lemma 14. Here, our meta inputs are $(h, \ell) \in F$ and our inputs are online arrival sequence $I_h$. Furthermore, the distribution over our meta inputs, i.e., $Q$ in Lemma 14, is uniform distribution. Recall that $\Pr(\text{Unif}[\hat{h}, \tilde{h}] = h) = \frac{1}{\tilde{h} - \hat{h} + 1}$ for any $h \in [\hat{h}, \tilde{h}]$.

Our goal here is to upper bound the competitive ratio of any randomized algorithm on family $F$, defined above. That is, for any distribution $P$ over any feasible deterministic algorithms $A$, we would like to bound the following quantity

$$\text{CR}_{\text{rand}} := \min_{I \in F} \mathbb{E}_{A \sim P} [\mathbb{E}_I[\text{CR}_A(I)]] .$$

By Lemma 14 we have

$$\text{CR}_{\text{rand}} \leq \max_{A \in \mathcal{A}_d} \mathbb{E}_{\tilde{X} \sim Q} [\mathbb{E}_I[\text{CR}_A(I)]] = \max_{A \in \mathcal{A}_d} \mathbb{E}_{\tilde{X} \sim Q} \left[ \mathbb{E}_I \left[ \frac{\text{REW}_A(I)}{\text{OPT}(I)} \right] \right],$$

where $\mathcal{A}_d$ is the set of all deterministic algorithms. Since $Q$ is a uniform distribution over any $h \in [\hat{h}, \tilde{h}]$, we have

$$\text{CR}_{\text{rand}} \leq \frac{1}{h - \hat{h} + 1} \cdot \max_{A \in \mathcal{A}_d} \sum_{h=\hat{h}}^{\tilde{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ \frac{\text{REW}_A(I_h, s_1)}{\text{OPT}(I_h)} \right].$$

(53)
Recall that under $\hat{I}_h$, REW$_A$ does not depend on $s_2 \gg m$, and hence we denote REW$_A(I, s)$ with REW$_A(I, s_1)$. We would like to show that for any deterministic algorithm $A \in \mathcal{A}$, we have

$$\lim_{m \to \infty} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ \text{REW}_A \left( I_{h_2}, s_i \right) \right] = 1 - \Theta(\frac{1}{\sqrt{m}}).$$

By Equation (11) in the proof of Theorem 4, it suffices to show that there does not exist a mapping $y_0(\cdot) : \{0, \ldots, \hat{h}\} \to \{0, 1, \ldots, m\}$, such that

$$\lim_{m \to \infty} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ (h - y_0(s_1)) \frac{r_1}{r_2} 1(h \geq y_0(s_1)) + (y_0(s_1) - h) \right] \leq 0. \quad (54)$$

Recall that under family $F$, any deterministic algorithm can be presented with such a mapping.

We show this by contradiction. Contrary to our claim, suppose there exists a mapping $y_0(\cdot)$ such that Equation (54) holds. This and the fact that $(h - y_0(s_1)) \frac{r_1}{r_2} 1(h \geq y_0(s_1)) \geq 0$, imply that

$$\lim_{m \to \infty} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ y_0(s_1) - h \right] \leq 0. \quad (55)$$

Next we consider two cases. In the first case, we assume that

$$\lim_{m \to \infty} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ y_0(s_1) - h \right] < 0$$

and we reach a contradiction under this assumption. In the second case, assume that

$$\lim_{m \to \infty} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ y_0(s_1) - h \right] = 0$$

and again reach a contradiction.

**Case 1:** $\lim_{m \to \infty} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ y_0(s_1) - h \right] < 0$. Suppose that

$$\lim_{m \to \infty} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ y_0(s_1) - h \right] = \eta < 0.$$ Then, we show that $y_0$ does not satisfy Equation (54), which is a contradiction. The left hand side of Equation (54) can be written as

$$\lim_{m \to \infty} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ (h - y_0(s_1)) \frac{r_1}{r_2} 1(h \geq y_0(s_1)) + (y_0(s_1) - h) \right]$$

$$\geq \lim_{m \to \infty} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ (h - y_0(s_1)) \frac{r_1}{r_2} + (y_0(s_1) - h) \right]$$

$$= \frac{r_1}{r_2} \eta + \eta$$

$$= -\frac{r_1}{r_2} \eta$$

$$> 0,$$

which is a contradiction of Equation (54).

**Case 2:** $\lim_{m \to \infty} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ y_0(s_1) - h \right] = 0$. Similar to the proof of Theorem 4, we compare $y_0(\cdot)$ and $y(\cdot)$, where $y(s_1) = \frac{s_1}{p}$ for any $s_1 \in \{0, 1, \ldots, \hat{h}\}$, and then show a contradiction. Equation (13) in the proof of Theorem 4 has shown that if $y(s_1) = \frac{s_1}{p}$, then for any $h \in [\hat{h}, \bar{h}]$,

$$\lim_{m \to \infty} \mathbb{E}_{s_1 \sim \text{Bin}(h, p)} \left[ (h - y(s_1)) \frac{r_1}{r_2} 1(h \geq \frac{s_1}{p}) + (y(s_1) - h) \right] > 0. \quad (56)$$

By subtracting Equation (56) from Equation (54), we have

$$\lim_{m \to \infty} \frac{1}{\sqrt{m}} \frac{1}{h-h+1} \sum_{h=\hat{h}}^{\hat{h}} \left( \mathbb{E}_{s_1} \left[ (h - y_0(s_1)) \frac{r_1}{r_2} 1(h \geq y_0(s_1)) + (y_0(s_1) - h) \right] \right. - \left. \mathbb{E}_{s_1} \left[ (h - y_0(s_1)) \frac{r_1}{r_2} 1(h \geq \frac{s_1}{p}) + (y_0(s_1) - h) \right] \right) > 0. \quad (57)$$
As in this case, we have \( \lim_{m \to \infty} \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} E_a \left[ \frac{E_a[\xi_1 - y_0(\xi_1)]}{\sqrt{m}} \right] = 0 \), Equation \((57)\) can be written as
\[
\lim_{m \to \infty} \frac{1}{\sqrt{m}} \left( \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} E_a \left[ (h - \frac{s_1}{p}) \frac{r_1}{r_2} \right] \right) > 0. \tag{58}
\]
Using the same set \( D \) defined in the proof of Theorem \( 3 \) and following similar steps, we can show that
\[
\lim_{m \to \infty} \frac{1}{\sqrt{m}} \left( \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} \sum_{r \in D \cap \{h\}} \Pr(s_1 = d) \left[ (h - \frac{d}{p}) \frac{r_1}{r_2} \right] \right) = 0. \tag{59}
\]
Therefore, we have
\[
\lim_{m \to \infty} \frac{1}{\sqrt{m}} \left( \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} \sum_{r \in D \cap \{h\}} \Pr(s_1 = d) \left[ (h - \frac{d}{p}) \frac{r_1}{r_2} 1(h \geq \frac{d}{p}) \right] \right) > 0. \tag{60}
\]
By the above inequality and following similar steps in the proof of Theorem \( 4 \), we can show that
\[
\lim_{m \to \infty} \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} \sum_{r \in D \cap \{h\}} \Pr(s_1 = d) > 0. \tag{61}
\]
We are now ready to show the contradiction using the assumption in this case, i.e.,
\[
\lim_{m \to \infty} \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} \frac{E_a [y_0(s_1) - h]}{\sqrt{m}} = 0.
\]
We have that
\[
\begin{align*}
\lim_{m \to \infty} & \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} \frac{E_a [y_0(s_1) - h]}{\sqrt{m}} \\
= & \lim_{m \to \infty} \frac{1}{\sqrt{m}} \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} \left( \sum_{r \in D \cap \{h\}} \Pr(s_1 = d)(y_0(d) - h) + \sum_{r \in D \cap \{h\}} \Pr(s_1 = d)\left(\frac{d}{p} - h\right) \right) \\
= & \lim_{m \to \infty} \frac{1}{\sqrt{m}} \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} \left( \sum_{r \in D \cap \{h\}} \Pr(s_1 = d)(y_0(d) - \frac{d}{p} - h) + \sum_{r \in D \cap \{h\}} \Pr(s_1 = d)\left(\frac{d}{p} - h\right) \right) \\
= & \lim_{m \to \infty} \frac{1}{\sqrt{m}} \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} \left( \sum_{r \in \{h\}} \Pr(s_1 = d)(y_0(d) - \frac{d}{p}) + \sum_{r \in D \cap \{h\}} \Pr(s_1 = d)\left(\frac{d}{p} - h\right) \right) \\
= & \lim_{m \to \infty} \frac{1}{\sqrt{m}} \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} \Pr(s_1 = d)(y_0(d) - \frac{d}{p}) \\
> & \lim_{m \to \infty} \frac{1}{h-\frac{1}{h}+1} \sum_{h=\frac{k}{h}} \Pr(s_1 = d)c_2 \\
> & 0,
\end{align*}
\]
The last inequality is from Equation \((61)\). This is a contradiction to the assumption in case 2. Therefore, there does not exist a mapping \( y_0(\cdot) \) with which we can achieve a competitive ratio of \( 1 - o\left(\frac{1}{\sqrt{m}}\right)\)
C.1. Proof of Lemma 14

Let $P(A)$ be the probability of choosing an algorithm $A$. Further, let $Q(X)$ be the probability of choosing meta input $X \in \mathcal{X}$. We start from the right hand side of Equation [52]:

$$\max_{A \in A_d} E_{X \sim Q}[E_{X \sim D(X)}[C_A(X)]] = \sum_{A \in A} P(A) \cdot \max_{A \in A_d} E_{X \sim Q}[E_{X \sim D(X)}[C_A(X)]]$$

$$\geq \sum_{A \in A} P(A) \cdot E_{X \sim Q}[E_{X \sim D(X)}[C_A(X)]]$$

$$= \sum_{A \in A} P(A) \cdot \sum_{X \in \mathcal{X}} Q(X) \cdot E_{X \sim D(X)}[C_A(X)]$$

$$= \sum_{X \in \mathcal{X}} Q(X) \cdot \sum_{A \in A} P(A) \cdot E_{X \sim D(X)}[C_A(X)]$$

$$\geq \min_{X \in \mathcal{X}} E_{A \sim P}[E_{X \sim D(X)}[C_A(X)]]$$

where the last inequality is the desired result.

Appendix D: Proof of Theorem 1

We let $(h, \ell)$, $(\tilde{h}, \tilde{\ell})$ be that $h = 0$, $\ell >> m$, $\tilde{h} >> m$, $\tilde{\ell} >> \ell$. We define two random arrival sequences $I_1$ and $I_2$, where in both sequences, type 2 agents arrive first, followed by type 1 agents. The number of type 1 and type 2 agents in sequence $I_1$ is $n_1 \sim \text{Bin}(h, 1 - p)$ and $n_2 \sim \text{Bin}(\ell, 1 - p)$ respectively. The number of type 1 and type 2 agents in sequence $I_2$ is $\tilde{n}_1 \sim \text{Bin}(\tilde{h}, 1 - p)$ and $\tilde{n}_2 \sim \text{Bin}(\tilde{\ell}, 1 - p)$ respectively. By our construction, we have $n_1 = 0$, $n_2 > m$ with probability 1. Furthermore, $\tilde{n}_1 > m$ and $\tilde{n}_2 > n_2 > m$ with probability 1.

Now consider the online arrival sequence $I_1$. Under $I_1$, with a probability of 1/2, the naive algorithm $A$ treats type 1 as the type with a higher expected reward, and gives protection level $x$ to type 1 agents. This means that, under $I_1$, we accept $m - x$ type 2 agents with probability 1/2. Further, with a probability of 1/2, it treats type 2 as the type with a higher expected reward and accept $m$ type 2 agents. Therefore, we have

$$E[\text{REW}_A(I_1)] = \frac{1}{2}(m - x)r_2 + \frac{1}{2}mr_2 = \frac{1}{2} \left( \frac{m - x}{m} + 1 \right).$$

Now consider the online arrival sequence $I_2$. Under $I_2$, the naive algorithm $A$ with a probability of 1/2, it treats type 1 as high-fare type, the naive algorithm accepts $m - x$ type 2 agents and $x$ type 1 agents. With a probability of 1/2, it accept $m$ type 2 agents and 0 type 1 agents. Therefore, we have

$$E[\text{REW}_A(I_2)] = \frac{1}{2}(m - x)r_2 + \frac{1}{2}mr_2 + \frac{1}{2}mr_1 = \frac{1}{2} \left( \frac{m - x}{m} + \frac{x}{m} + \alpha \right),$$

where $\alpha = r_2/r_1$.

Then, the CR of the naive algorithm $A$ is upper bounded by

$$\min_{x \in \{0, 1, 2, \ldots, m\}} \left\{ \frac{1}{2} \left( \frac{m - x}{m} + 1 \right), \frac{1}{2} \left( \frac{m - x}{m} + \frac{x}{m} + \alpha \right) \right\}. $$

By letting two terms equal to each other, we get $\frac{x}{m} = \frac{\alpha - 1}{2 - \alpha}$, which implies that the CR of the naive algorithm $A$ is at most $\frac{1}{2 - \alpha}$, as desired.
Appendix E: Proof of Theorem 3

Because $m_1$ is a constant number that depends on $r_1$, $r_2$, and $p$, to show the asymptotic result, we only need to consider the case that $m \geq m_1$. We can write the CR as

$$CR_A \geq \min \{CR_1, CR_2, CR_3\}.$$ 

To show the result, we will show how $CR_1, CR_2$, and $CR_3$ scale $m$. First it is easy to see that $CR_1 = \min \left\{ \left(1 + \frac{1-p}{2m}\right) + 1 - \frac{r_1 + r_2}{\sqrt{m}} \right\}$ scales with $1 - \Theta(1/\sqrt{m})$. Similarly, it is easy to see that $CR_2 = \min \left\{ (1 - \frac{1}{2m})(1 - \frac{1}{m}), \frac{m}{h} \right\}$ scales with $1 - \Theta(1/\sqrt{m})$. This is because $h_0 = \frac{m - \sqrt{m}}{1-p} \frac{1}{\sqrt{m}} + \frac{1}{2p}$, and $h_1 = h_0(1 - p) - \sqrt{p(1 - p)} h_0 - \frac{q}{\sqrt{m}} = \Theta(m)$. Then, we have $\frac{h}{m} = 1 - \Theta(1/\sqrt{m})$.

To complete the proof, we show how

$$CR_3 = (1 - \frac{1}{m})^2 (1 - \frac{1}{\sqrt{m}p})^2 W$$

scales with $m$. To do so, firstly, it is clear that $(1 - \frac{1}{m})^2 (1 - \frac{1}{\sqrt{mp}})^2 = 1 - \Theta(1/m)$. Secondly, by the analysis for $CR_3$, we know that $W = \min \left\{ (1 - \frac{1}{\sqrt{m}}), \frac{m}{h} \right\}$ also scales with $1 - \Theta(1/\sqrt{m})$. Putting these together, we conclude that $CR_A = 1 - \Theta(1/\sqrt{m})$, as desired.

Appendix F: Proof of Lemmas 10 and 11

F.0.1. Proof of Lemma 10 Recall that $E_0 = \{n_2 > m - \frac{1-p}{p} s_1\}$. We provide a lower bound of $\mathbb{E}_{\alpha} \left[ \text{REW}_{\alpha}(n, \psi; \mathcal{G}) \right] | E_2(h, \ell)$ by considering the conditional expectation on $E_0$:

$$\inf_{(h, \ell) \in \mathbb{R}_+^2} \mathbb{E}_{\alpha} \left[ \text{REW}_{\alpha}(n, \psi; \mathcal{G}) \right] | E_2(h, \ell) \geq \inf_{(h, \ell) \in \mathbb{R}_+^2} \mathbb{E}_{\alpha} \left[ \text{REW}_{\alpha}(n, \psi; \mathcal{G}) \right] | E_2(h, \ell), E_0 \right) \Pr(E_0 | E_2(h, \ell)).$$

In light of Equation (62), we will provide a lower bound for $\mathbb{E}_{\alpha} \left[ \text{REW}_{\alpha}(n, \psi; \mathcal{G}) \right] | E_2(h, \ell), E_0 \right)$ and $\Pr(E_0 | E_2(h, \ell))$.

**Part 1: bounding $\mathbb{E}_{\alpha} \left[ \text{REW}_{\alpha}(n, \psi; \mathcal{G}) \right] | E_2(h, \ell), E_0$**. By Equation (??) in the proof of Lemma 2 for a fixed realization of $s$ (or equivalently $n$), we have

$$\text{REW}_{\alpha}(n, \psi; \mathcal{G}) = \min \left\{ n_2, \left( m - \frac{1-p}{p} s_1 \right)^+ \right\} \cdot r_2 + \min \left\{ m - \min \left\{ n_2, \left( m - \frac{1-p}{p} s_1 \right)^+ \right\}, n_1 \right\} \cdot r_1$$

Conditional on $E_0 = \{n_2 \leq m - \frac{1-p}{p} s_1\}$, we have $\min \{n_2, (m - \frac{1-p}{p} s_1)^+\} = n_2$, and $\min \{m - \min \{n_2, (m - \frac{1-p}{p} s_1)^+\}, n_1\} = \min \{m - n_2, n_1\}$. Then, we have

$$\text{REW}_{\alpha}(n, \psi; \mathcal{G}) \geq \frac{\text{REW}_{\alpha}(n, \psi; \mathcal{G})}{\text{OPT}(n)} \geq \frac{\text{REW}_{\alpha}(n, \psi; \mathcal{G})}{n_1 r_1 + \min \{m - n_1, n_2\} r_2} = \frac{n_2 r_2 + \min \{m - n_2, n_1\} r_1}{n_1 r_1 + \min \{m - n_1, n_2\} r_2} \geq \frac{n_1 r_1 + n_2 r_2}{n_1 r_1 + (m - n_1) r_2} \geq \frac{(m - n_2) r_1 + n_2 r_2}{n_1 r_1 + (m - n_1) r_2} \geq \frac{(m - n_2) r_1 + n_2 r_2}{n_1 r_1 + (m - n_1) r_2} \geq \frac{(m - n_2) r_1 + n_2 r_2}{m r_1} = 1 - \frac{1}{\sqrt{m}}.$$


The second inequality is because when \(\min\{m - n_2, n_1\} = n_1\), we have \(\min\{m - n_1, n_2\} = n_2\), and when \(\min\{m - n_2, n_1\} = m - n_2\), we have \(\min\{m - n_1, n_2\} = m - n_1\).

**Part 2: bounding \(\Pr(\mathcal{E}_0|\mathcal{E}_2(h, \ell))\).** We want to show that \(\Pr(\mathcal{E}_0|\mathcal{E}_2(h, \ell))\) equals 1 for \((h, \ell) \in \mathcal{R}_3^1\). As \(\ell \in \sqrt{m}\) for any \((h, \ell) \in \mathcal{R}_3^1\), we have \(n_2 \leq \sqrt{m}\). Then, because \(\mathcal{E}_0 = \{n_2 \leq m - \frac{1-p}{p}s_1\}\), we have

\[
\Pr(\mathcal{E}_0|\mathcal{E}_2(h, \ell)) = \Pr\left(1 - \frac{p}{p}s_1 \leq m - n_2|\mathcal{E}_2(h, \ell)\right) \geq \Pr\left(1 - \frac{p}{p}s_1 \leq m - \sqrt{m}|\mathcal{E}_2(h, \ell)\right)
\]

\[
= \Pr\left(s_1 \leq \frac{p(m - \sqrt{m})}{1 - p}|\mathcal{E}_2(h, \ell)\right).
\]

Next, we show that conditional on \(\mathcal{E}_2(h, \ell)\), the upper bound of \(s_1\) is no more than \(\frac{p(m - \sqrt{m})}{1 - p}\). If \(\mathcal{E}_2(h, \ell)\) happens, we have \(s_1 \leq hp + \sqrt{h}\). We then have

\[
s_1 \leq hp + \sqrt{h} \Rightarrow s_1 \leq h_0 p + \sqrt{h_0} \leq s_1 \Rightarrow \frac{p(m - \sqrt{m})}{1 - p},
\]

where the second inequality holds because \(hp + \sqrt{h}\) is an increasing function, and \(h < h_0\) in region \(\mathcal{R}_2\). These imply that \(hp + \sqrt{p} \leq h_0 p + \sqrt{h_0}\). The last inequality holds because \(\frac{p(m - \sqrt{m})}{1 - p} \geq h_0 p + \sqrt{h_0}\). Recall that

\[
h_0 = \min\{y \geq 0 : \frac{p(m - \sqrt{m})}{1 - p} \geq yp + \sqrt{y}\}\]

This means that

\[
\Pr(\mathcal{E}_0^c|\mathcal{E}_2(h, \ell)) = 1.
\]

By Equations (64) and (65), we have

\[
\inf_{(h, \ell) \in \mathcal{R}_3^1} \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \mid \mathcal{E}_0\right] \Pr(\mathcal{E}_0) \geq (1 - \frac{1}{\sqrt{m}}),
\]

**F.0.2. Proof of Lemma 11** Recall that \(\mathcal{E}_1 = \{n_1 > \frac{1-p}{p}s_1\}\). We bound \(\mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \mid \mathcal{E}_2(h, \ell)\right]\) by considering the conditional expectation on \(\mathcal{E}_1\) and \(\mathcal{E}_1^c\) respectively.

\[
\inf_{(h, \ell) \in \mathcal{R}_3^1} \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \mid \mathcal{E}_2(h, \ell)\right] \geq \inf_{(h, \ell) \in \mathcal{R}_3^1} \left\{ \min \left\{ \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \mid \mathcal{E}_2(h, \ell), \mathcal{E}_1\right] \right\} \right\}
\]

\[
\mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \mid \mathcal{E}_2(h, \ell), \mathcal{E}_1\right]
\]

We will provide the lower bound of \(\inf_{(h, \ell) \in \mathcal{R}_3^1} \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \mid \mathcal{E}_2(h, \ell), \mathcal{E}_1\right]\) in the first part of the proof, and the lower bound of \(\inf_{(h, \ell) \in \mathcal{R}_3^1} \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \mid \mathcal{E}_2(h, \ell), \mathcal{E}_1^c\right]\) in the second part of the proof.

**Part 1: bounding \(\inf_{(h, \ell) \in \mathcal{R}_3^1} \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \mid \mathcal{E}_2(h, \ell), \mathcal{E}_1\right]\).** We then have

\[
\inf_{(h, \ell) \in \mathcal{R}_3^1} \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \mid \mathcal{E}_2(h, \ell), \mathcal{E}_1\right] \geq \inf_{(h, \ell) \in \mathcal{R}_3^1} \min \left\{ \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; \mathcal{G})}{\text{OPT}(n)} \mid \mathcal{E}_2(h, \ell), \mathcal{E}_1^c, \mathcal{E}_0^c\right] \right\}
\]

where we recall \(\mathcal{E}_0^c = \{n_2 > m - \frac{1-p}{p}s_1\}\).
To provide a lower bound for $\mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}|\mathcal{E}_2(h, \ell), \mathcal{E}_1, \mathcal{E}_0^C]\right]$, we note that under event $\mathcal{E}_0^C$, we have $n_2 > m - \frac{1-p}{p}s_1$, and hence

$$\mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}|\mathcal{E}_2(h, \ell), \mathcal{E}_1, \mathcal{E}_0^C\right] \geq \frac{(m - \frac{1-p}{p}s_1)^+ r_2 + \min\{m, \frac{1-p}{p}s_1\} r_1}{mr_1} \geq \frac{(m-n_2)r_1}{mr_1} \geq 1 - \frac{1}{\sqrt{m}}. \quad (68)$$

Then, we give a lower bound to $\mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}|\mathcal{E}_2(h, \ell), \mathcal{E}_1, \mathcal{E}_0\right]$. Conditional on $\mathcal{E}_1$ and $\mathcal{E}_0$, we have $\frac{1-p}{p}s_1 \geq m-n_2$ and $\frac{1-p}{p}s_1 \leq n_1$. That is, the number of type 1 agents in the online arrival sequence is greater than or equal to the protection level of Algorithm 2 for type 1 agents (i.e., $\frac{1-p}{p}s_1 \leq n_1$), and in addition, the algorithm ends up rejecting some of type 2 agents as $\frac{1-p}{p}s_1 \geq m-n_2$. Therefore, we have

$$\mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}|\mathcal{E}_2(h, \ell), \mathcal{E}_1, \mathcal{E}_0\right] \geq \frac{(m - \frac{1-p}{p}s_1)^+ r_2 + \min\{m, \frac{1-p}{p}s_1\} r_1}{mr_1} \geq \frac{(m-n_2)r_1}{mr_1} \geq 1 - \frac{1}{\sqrt{m}},$$

where the last inequality holds because $n_2 \leq \ell \leq \sqrt{m}$ for any $(h, \ell) \in \mathbb{R}_0^2$.

Part 2: bounding $\inf_{(h, \ell) \in \mathbb{R}_0^2} \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}|\mathcal{E}_2(h, \ell), \mathcal{E}_1^C\right]$. Recall that for a fixed realization of $s$ (or equivalently $n$),

$$\text{REW}_A(n, \psi; G) = \min\left\{n_2, \left(m - \frac{1-p}{p}s_1\right)^+, \min\{m-n_2, \left(m - \frac{1-p}{p}s_1\right)^+\}, n_1\right\} \cdot r_1. \quad (69)$$

Conditional on $\mathcal{E}_1^C = \{n_1 \leq \frac{1-p}{p}s_1\}$, we do not receive enough number of type 1 agents and thus, we have

$$\min\left\{m - \min\left\{n_2, \left(m - \frac{1-p}{p}s_1\right)^+\right\}, n_1\right\} \geq \min\left\{\max\left\{m-n_2, \frac{1-p}{p}s_1\right\}, n_1\right\} = n_1.$$

That is, Algorithm 2 accepts all type 1 agents in the online arrival sequence. Then, we have

$$\mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}|\mathcal{E}_2(h, \ell), \mathcal{E}_1^C\right] \geq \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}|\mathcal{E}_2(h, \ell), \mathcal{E}_0^C\right] \geq \mathbb{E}_a\left[\frac{\text{REW}_A(n, \psi; G)}{\text{OPT}(n)}|\mathcal{E}_2(h, \ell), \mathcal{E}_1\right].$$

We will show that $\mathbb{E}_a\left[n_1|\mathcal{E}_2(h, \ell), \mathcal{E}_1^C\right] \geq h_1$, where $h_1 = h_0(1-p) - \sqrt{p(1-p)}h_0 - \frac{\beta}{\sqrt{h_0}}$ and $\beta = 0.4215 \cdot \frac{\mu^2 + (1-p)^2}{p(1-p)}$.

Since $\mathcal{E}_1^C = \{n_1 \leq \frac{1-p}{p}s_1\} = \{h(1-p) - n_1 \geq 0\}$, we have

$$\mathbb{E}_a[n_1|\mathcal{E}_1^C, \mathcal{E}_2(h, \ell)]$$

$$= h(1-p) - \mathbb{E}_a[h(1-p) - n_1|\{h(1-p) - n_1 \geq 0\}, \mathcal{E}_2(h, \ell)]$$

$$= h(1-p) - \mathbb{E}_a[h(1-p) - n_1|\{h(1-p) - n_1 \geq 0\} \cap \{h(1-p) - \sqrt{h} \leq n_1 \leq h(1-p) + \sqrt{h}\}]$$

$$= h(1-p) - \mathbb{E}_a[h(1-p) - n_1|\{0 \leq h(1-p) - n_1 \leq \sqrt{h}\}]$$

$$\geq h(1-p) - \mathbb{E}_a[h(1-p) - n_1|\{h(1-p) - n_1 \geq 0\}]$$

$$= \mathbb{E}_a[n_1|\mathcal{E}_1^C].$$
The inequality is because
\[
\mathbb{E}_s[h(1-p) - n_1 \{0 \leq h(1-p) - n_1 \leq \sqrt{h}\}] \\
\leq \mathbb{E}_s[h(1-p) - n_1 \{0 \leq h(1-p) - n_1 \leq \sqrt{h}\}] + \mathbb{E}_s[h(1-p) - n_1 \{h(1-p) - n_1 \geq \sqrt{h}\}] \\
= \mathbb{E}_s[h(1-p) - n_1 \{h(1-p) - n_1 \geq 0\}].
\]
Next, we have
\[
\mathbb{E}_s[n_1 | \mathcal{E}_1^c] = h(1-p) - \mathbb{E}_s[h(1-p) - n_1 | h(1-p) - n_1 \geq 0] \\
\geq h(1-p) - \mathbb{E}_s[|h(1-p) - n_1|] - \frac{\beta}{\sqrt{h}} \\
\geq h(1-p) - \sqrt{p(1-p)}h - \frac{\beta}{\sqrt{h}} \\
\geq h_0(1-p) - \sqrt{p(1-p)}h_0 - \frac{\beta}{\sqrt{h_0}} \\
= h_1,
\]
where the first inequality holds because by Nagaev and Chebotarev (2011), we have \(\mathbb{E}_s[(h(1-p) - n_1) | (h(1-p) - n_1 \geq 0)] - \mathbb{E}_s[|h(1-p) - n_1|] \leq \frac{\beta}{\sqrt{h}}\). The second inequality follows from Berend and Kontorovich (2013) that shows \(\mathbb{E}_s[|h(1-p) - n_1|] \leq \sqrt{p(1-p)}h\). The third inequality holds because when \(h > h_0\), \(h(1-p) - \sqrt{p(1-p)}h - \frac{\beta}{\sqrt{h}}\) is a non-decreasing function for any \(p\). To see why note that
\[
\frac{\partial (h(1-p) - \sqrt{p(1-p)}h - \frac{\beta}{\sqrt{h}})}{\partial h} \geq (1-p) - \frac{1}{2} \sqrt{p(1-p)} \frac{1}{\sqrt{h}} \geq 0.
\]
where the last inequality holds for any \(h \geq \frac{1}{4} \frac{p}{1-p}\), and we have \(h_0 > \frac{m - \sqrt{m}}{1-p} > \frac{1}{4} \frac{p}{1-p}\) for all \(m \geq 2\).