Information Processing Letters 180 (2023) 106334

Contents lists available at ScienceDirect

Information Processing Letters

journal homepage: www.elsevier.com/locate/ipl

The loss of serving in the dark ☆

Yossi Azar [a,1], Ilan Reuven Cohen [b,*,2], Iftah Gamzu [c,3]

[a] *Blavatnik School of Computer Science, Tel-Aviv University, Israel*
[b] *Faculty of Engineering, Bar-Ilan University, Israel*
[c] *Amazon, Israel*

ARTICLE INFO

Article history:
Received 25 February 2022
Received in revised form 21 July 2022
Accepted 3 October 2022
Available online 14 October 2022
Communicated by Ranko Lazic

Keywords:
Oblivious algorithms
Randomized algorithms
Online algorithms
Scheduling
Prompt mechanisms

ABSTRACT

We study the following balls and bins stochastic process: There is a buffer with B bins, and there is a stream of balls $X = \langle X_1, X_2, \dots, X_T \rangle$ such that X_i is the number of balls that arrive before time i but after time $i-1$. Once a ball arrives, it is stored in one of the unoccupied bins. If all the bins are occupied then the ball is thrown away. In each time step, we select a bin uniformly at random, clear it, and gain its content. Once the stream of balls ends, all the remaining balls in the buffer are cleared and added to our gain. We are interested in analyzing the expected gain of this randomized process with respect to that of an optimal gain-maximizing strategy, which gets the same online stream of balls, and clears a ball from a bin, if exists, at any step. We name this gain ratio the loss of serving in the dark.

In this paper, we determine the exact loss of serving in the dark. We prove that the expected gain of the randomized process is worse by a factor of $\rho + \epsilon$ from that of the optimal gain-maximizing strategy where $\epsilon = O(1/B^{1/3})$ and $\rho = \max_{\alpha>1} \alpha e^{\alpha}/((\alpha-1)e^{\alpha} + e - 1) \approx 1.69996$. We also demonstrate that this bound is essentially tight as there are specific ball streams for which the above-mentioned gain ratio tends to ρ. Our stochastic process occurs naturally in packets scheduling and mechanisms design applications.

1. Introduction

Consider the fundamental packets scheduling scenario in which there is an online stream of packets with arbitrary values arriving to a network device that can accommodate B packets. The device can transmit one packet in each time-step, and several packets may arrive between the transmission time-steps. The goal is to maximize the overall value of transmitted packets. A trivial greedy algorithm for this scenario keeps the B packets with the highest values, i.e., in the arrival of a new packet when the buffer is full, it discards the lowest valued packet among the new packet and the buffer's packets, and in a transmission time-step, it transmits the packet with the highest value in the buffer. It is easy to verify that the greedy algorithm is optimal. However, it inspects the values of the packets prior to their transmission. We are interested in algorithms whose transmission decisions are *value-oblivious*. Such algorithms have the property that if one focuses on any single packet then for all possible values it is either transmitted in the same time-step or rejected. Value-oblivious transmission algorithms are ben-

☆ An extended abstract of this work appears in the proceedings of STOC2013 [6].

* Corresponding author.
E-mail addresses: azar@tau.ac.il (Y. Azar), ilan-reuven.cohen@biu.ac.il (I.R. Cohen), iftah.gamzu@yahoo.com (I. Gamzu).

1 Supported in part by the Israel Science Foundation (grant No. 2304/20).
2 This research was supported by the Israel Science Foundation (grant No. 1737/21).
3 This work was done while the author was affiliated with Tel Aviv university and prior to joining Amazon.

https://doi.org/10.1016/j.ipl.2022.106334

eficial in game-theoretic settings that require fairness in the following informal sense: one would not like a high value packet to have a higher priority in transmission compared to a low value packet if both packets are to be transmitted. This fairness property is necessary for the design of *prompt mechanisms*, where a packet should learn its payment immediately upon its transmission. We note that value-oblivious transmission algorithms may inspect the values of packets on their arrival. Therefore, one can assume without loss of generality that any value-oblivious transmission algorithm keeps the B packets with the highest values at any point in time. One example of a value-oblivious transmission algorithm is the FIFO algorithm, which keeps the highest value packets and transmits the earliest packet in the buffer. The algorithm is known to be 2 competitive and this bound is tight for this algorithm [22]. A natural question is whether one can design a value-oblivious transmission algorithm with a better competitive ratio, and we answer this question affirmatively.

We consider a simple randomized algorithm that transmits a packet from the buffer uniformly at random. The core of analyzing this algorithm can be reduced using the zero-one principle [7], to the following natural *balls and bins stochastic process*: There is a buffer with B bins, and there is a stream of balls $X = \langle X_1, X_2, \ldots, X_T \rangle$ such that X_i is the number of balls that arrive before time i but after time $i-1$. Once a ball arrives, it is stored in one of the unoccupied bins, i.e., a bin that does not hold a ball. If all the bins are occupied then the ball is thrown away. In each time step, we select a bin uniformly at random, clear it, and gain its content. In particular, if that bin is occupied with a ball then our gain is one; otherwise, our gain is zero. Once the stream of balls ends, all the remaining balls in the buffer are cleared and added to our gain. We are interested in analyzing the expected gain of this randomized process with respect to that of an *optimal gain-maximizing strategy*, which gets the same online stream of balls, and clears a ball from a bin, if exists, at any step. We name this gain ratio the *loss of serving in the dark* since the bins are selected without knowledge about their content.

1.1. Our results

Determining the exact loss of serving in the dark. We prove that the expected gain of the randomized process is worse by a factor of $\rho + \epsilon$ from that of the optimal gain-maximizing strategy where $\epsilon = O(1/B^{1/3})$, and ρ is defined by the following algebraic expression:

$$\rho = \max_{\alpha > 1} \frac{\alpha e^{\alpha}}{(\alpha - 1)e^{\alpha} + e - 1} \approx 1.69996$$

We also demonstrate that this bound is essentially tight as there are specific ball streams for which the above-mentioned gain ratio tends to ρ. As a corollary, we attain that the asymptotic loss of serving in the dark is *exactly* ρ. These findings are presented in Section 2.

Application 1: Packets Scheduling. The stochastic process occurs naturally in many applications. As described before, one such example is value-oblivious transmission algorithm. The above result implies that the random transmission algorithm has a competitive ratio of $\rho + \epsilon$. Note that in the randomized algorithm, a packet might remain in the buffer for a long time. Nevertheless, one can easily validate that with high probability, the delay of a packet is at most logarithmic factor more than its delay in the FIFO algorithm. Thus, one can reject a packet after it stayed in the buffer for $O(B \log B)$ steps without degrading the competitive ratio.

Application 2: Prompt mechanisms for bounded capacity auctions. We use the stochastic process to establish a natural randomized selection mechanism for bounded capacity auctions. A bounded capacity auction is a single-item periodic auction for bidders that arrive online, and the number of participating bidders is bounded, e.g., when the auction room has a limited size. We show that the random selection mechanism is truthful, supports prompt payments (see the definition of prompt payments in [13]), and achieves an expected competitive ratio of $\rho + \epsilon$. This finding surpasses the 2-competitive FIFO algorithm.

1.2. Our approach and techniques

An essential component in our approach is to utilize a deterministic fractional process, designed in a natural way to correspond to the randomized process, as a proxy for the analysis of the loss of serving in the dark. As we do not know how to analyze the loss of serving in the dark directly, we make the following two steps which are combined to yield the desired result:

(1) *Analyzing the fractional process against the optimal one* – We characterize the ball stream with the worst gain ratio between the fractional process and the optimal one. This characterization defines the stream uniquely, i.e., depending only on its length, and reduces the problem of finding the worst gain ratio between the two previously-mentioned processes to that of analyzing a specific algebraic expression, which was previously identified with ρ.

(2) *Analyzing the randomized process against the fractional one* – Ideally, we would have liked to show that the expected gain of the randomized process and the gain of the fractional one are essentially equal. Kurtz's theorem [25] informally says that the solutions of a stochastic process behave similar to the solutions to the differential equation of its fractional counterpart (see, e.g., [21]). Unfortunately, we cannot apply this theorem in our setting due to the hard constraint on the number of bins that induces overflows, i.e., some balls need to be rejected if there are more balls than bins. Specifically, one can demonstrate that there is a drift between the randomized and fractional processes. Nevertheless, we are able to bound their gain difference by bounding the drift in their loads. This is achieved by applying Azuma's inequality to a martingale process defined with respect to the two previously-mentioned processes.

1.3. Related work

A classical and well-known balls and bins scenario is when B balls are placed into B bins, where the optimization criteria is the fraction of full bins, namely, bins that got at least one ball. A simple result demonstrates that if the balls are placed independently and uniformly at random then the expected fraction of full bins is $1 - 1/e$. This result has a similar flavor to our result in the sense that if this process could have been done in the light, i.e., one could deterministically place each ball in any bin, then the fraction of full bins would have been 1; however, since this process is done in the dark, i.e., the balls are placed in a random way, then there is a loss of gain. There are other randomized ball and bins stochastic processes that have been analyzed using various techniques such as martingales and Azuma's inequality. Due to the ever-growing line of work in this context, it is beyond the scope of this writing to do justice and present an exhaustive survey of previous work. We refer the reader to directly related papers [20,24,29,5,2,30,31,14] and to the references therein for a more comprehensive review of the literature.

Various problems are related to prompt mechanisms for bounded capacity auctions. A closely related one is the dynamic auction with expiring items problem. Hajiaghayi et al. [19] introduced the problem and proposed a 2-competitive truthful mechanism for it, and Cole et al. [13] presented a 2-competitive truthful mechanism which is also prompt, see also [4,9,11,16,18,28,26] for other variants and results. Numerous papers deal with different aspects of packets scheduling, see [1,3,8,10,12,15,17,22,23,27,33] for a more comprehensive review.

2. The stochastic process and its analysis

In this section, we prove the next theorem that determines the loss of serving in the dark.

Theorem 2.1. *The expected gain of the randomized process is worse by a factor of $\rho + \epsilon$ from that of the optimal gain-maximizing strategy for $\epsilon = O(1/B^{1/3})$.*

We also show that the above gain ratio is essentially tight, resulting in the following corollary.

Corollary 2.2. *The loss of serving in the dark is asymptomatically $\rho \approx 1.69996$.*

2.1. Notation and terminology

Given a buffer with B bins, and a stream of balls $X = \langle X_1, X_2, \dots X_T \rangle$, we use the following notation with respect to some strategy ALG for clearing the balls: Let $G_i^{\text{ALG}}(X)$ be the *gain* of ALG at time i, and let $L_i^{\text{ALG}}(X)$ be the *load* of the buffer at time i, namely, the number of balls in the buffer just before ALG clears some bin at time i. Let $O_i^{\text{ALG}}(X)$ be the *overflow* at time i, that is, the number of balls thrown away at time i. Note that $L_i^{\text{ALG}}(X) = \min\{L_{i-1}^{\text{ALG}}(X) - G_{i-1}^{\text{ALG}}(X) + X_i, B\}$ and $O_i^{\text{ALG}}(X) = \max\{0, L_{i-1}^{\text{ALG}}(X) - G_{i-1}^{\text{ALG}}(X) + X_i - B\}$. Let $G^{\text{ALG}}(X) = \sum_{i=1}^{T-1} G_i^{\text{ALG}}(X) + L_T^{\text{ALG}}(X)$ be the *overall gain* of ALG. By this definition, once the stream ends, all the remaining balls in the buffer are cleared and added to the gain. Also note that we can alternatively define $G^{\text{ALG}}(X) = \sum_{i=1}^{T} X_i - \sum_{i=1}^{T} O_i^{\text{ALG}}(X)$.

It is easy to see that for the optimal gain-maximizing strategy OPT, $G_i^{\text{OPT}}(X) = 1$ iff $L_i^{\text{OPT}}(X) \geq 1$, and 0 otherwise. Turning to our randomized process RND, the gain of RND at step i is a random variable, which depends on the current load (which is also a random variable), since the RND chooses uniformly at random, we have for any specific realization of the random variables before step i, $G_i^{\text{RND}}(X) = 1$ w.p. $L_i^{\text{RND}}(X)/B$ and otherwise 0. Let $G^{\text{RND}}(X) = \mathbb{E}[\sum_i G_i^{\text{RND}}(X) + L_T^{\text{RND}}(X)]$ the expected gain of RND on X. With these definitions in mind, our goal is to determine the exact *loss of serving in the dark* defined as $\hat{\rho} = \sup_X \hat{\rho}(X) = \sup_X G^{\text{OPT}}(X)/G^{\text{RND}}(X)$.

The fractional process. As we do not know how to bound $\hat{\rho}$ directly, we define a *deterministic* fractional process that will be used as a proxy for the analysis. We analyze the gain of this fractional process and prove that it is far by a factor of ρ from the gain of the optimal gain-maximizing strategy. This fractional process is designed in a natural way to correspond to the randomized process. Unfortunately, we observe that the gain of this process is not the expected gain of the randomized process, but rather dominates it. Nevertheless, we still establish that it is within a $1 + \epsilon$ factor away from the expected gain of randomized process. Combining these two results together enables us to prove the claimed $\hat{\rho}$.

The *fractional process* is defined so that if the buffer is loaded with L (fractional) balls, then an L/B fraction of a ball is cleared from the buffer. Formally, the gain of the fractional process FRC at time i is $G_i^{\text{FRC}}(X) = L_i^{\text{FRC}}(X)/B$, while its load is $L_i^{\text{FRC}}(X) = \min\{\beta \cdot L_{i-1}^{\text{FRC}}(X) + X_i, B\}$, where $\beta = 1 - 1/B$.

2.2. Analysis of the fractional process

First, we observe that FRC and RND are monotone with respect to their current load and accumulated gain.

Observation 2.3. *If* ALG *is either* FRC *or* RND *and suppose at some step of* ALG *we remove some positive amount of balls from its load, and continue the process,*

- *If we discard the removed balls, then the total gain of* ALG *can not increase.*
- *If we add the removed balls into its gain, then the total gain of* ALG *can not decrease.*

Next, we show that it is sufficient to consider *valid* ball streams for which the optimal strategy has no overflow nor subflow. An *overflow* is a situation in which OPT cannot store all arriving balls in the buffer and therefore has to throw some of them away, while a *subflow* is a situation in which there are no balls in OPT's buffer to be cleared.

Lemma 2.4. *Given any ball stream X there is a valid ball stream X', with an equal or smaller number of steps, for which the optimal strategy does not have an overflow nor a subflow and $\hat{\rho}(X') \geq \hat{\rho}(X)$.*

Proof. By removing overflow balls from X for which OPT overflows, and by removing steps in which OPT does not clear a ball would not modify OPT total gain. While by Observation 2.3, RND total gain can not increase in this modified sequence. □

The following observation relating to valid ball streams will be utilized later.

Observation 2.5. *A ball stream $X = \langle X_1, \ldots, X_T \rangle$ is valid if and only if $1 \le \left(\sum_{i=1}^{k} X_i\right) - (k-1) \le B$, for any $k \in \{1, \ldots, T\}$. In addition, given a valid ball stream $X = \langle X_1, \ldots, X_T \rangle$, the gain of the optimal strategy is $G^{\mathrm{OPT}}(X) = \sum_{i=1}^{T} X_i$, and its load in each step k is $L_k^{\mathrm{OPT}}(X) = \left(\sum_{i=1}^{k} X_i\right) - (k-1)$.*

In what follows, in Theorem 2.6, we establish an upper bound $\max_X \hat{\rho}(X) = \max_X G^{\mathrm{OPT}}(X)/G^{\mathrm{RND}}(X)$ under the assumption that X is valid. By Lemma 2.4, we know that this is sufficient to upper bound $\hat{\rho}$ for such streams.

Theorem 2.6. *$\rho(X) \le \rho$ for any valid ball stream X.*

Consider a bounded length ball stream. It is clear that there is a ball stream with the worst gain ratio for this length as the number of relevant ball streams is finite. We next characterize the bounded length ball stream with the worst gain ratio between OPT and FRC. Our characterization defines the stream uniquely (i.e., depending only on its length). Subsequently, we compute the exact gain ratio for this stream. The next lemma identifies an important property of the stream under consideration.

Lemma 2.7. *Given a valid ball stream X with a maximal gain ratio $\rho(X)$, we may assume that if at step i, $L_i^{\mathrm{FRC}}(X) = B$ then also $L_i^{\mathrm{OPT}}(X) = B$.*

Proof. Given a sequence X and step i where FRC is fully loaded but OPT is not. As a consequence, we can modify the stream by shifting balls from the remainder of the stream to this step, or by adding balls to this step if the remainder of the stream is empty. Since OPT was not full, the ball stream stays valid. Furthermore, the gain of OPT may only increase while the gain of FRC cannot increase by Observation 2.3, and hence, the gain ratio may only increase. □

Corollary 2.8. *Given a valid ball stream X with a maximal gain ratio $\rho(X)$, we may assume that if at step i, $L_i^{\mathrm{OPT}}(X) < B$ then $L_i^{\mathrm{FRC}}(X) < B$, and thus, $O_i^{\mathrm{FRC}}(X) = 0$.*

For the purpose of characterizing the ball stream with the worst gain ratio between OPT and FRC, we subsequently focus on analyzing valid ball streams that maximize the number of balls thrown away by FRC, namely, the total sum of the overflows. By Corollary 2.8, we know that when an overflow of FRC occurs then OPT must be full. We define a *block* as a substream between two consecutive occasions where FRC is full. Let block(X, k, d) be a block that starts on step k with a length of d. Note that replacing one block with another block does not influence the load states of FRC before the beginning or after the end of the block. Similarly to before, we say that a block is *valid* if OPT does not have an overflow nor a subflow in that block. The following observation summarizes the properties of a block(X, k, d).

Observation 2.9. *For a block(X, k, d):*

1. $L_k^{\mathrm{FRC}}(X) = B$, $L_k^{\mathrm{OPT}}(X) = B$, $L_{k+d}^{\mathrm{FRC}}(X) = B$, *and* $L_{k+d}^{\mathrm{OPT}}(X) = B$.
2. $\sum_{i=1}^{k} X_i = B + (k-1)$, *and* $\sum_{i=k+1}^{k+d} X_i = d$.
3. $L_{k+j}^{\mathrm{OPT}}(X) = B - j + \sum_{i=k+1}^{k+j} X_i$, *for* $0 \le j \le d$.

Note that the only overflow of FRC in a block may occur in the last step of the block. Accordingly, and in conjunction with the definition of the fractional process, we observe the following.

Observation 2.10. *For a block(X, k, d):*

1. $L_{k+j}^{\mathrm{FRC}}(X) = B \cdot \beta^j + \sum_{i=1}^{j} X_{k+i} \cdot \beta^{j-i}$, *for* $0 \le j < d$.
2. $O_{k+d}^{\mathrm{FRC}}(X) = \left(B \cdot \beta^d + \sum_{i=1}^{d} X_{k+i} \cdot \beta^{d-i}\right) - B$.

In order to characterize a valid block with a maximum overflow of FRC, we first define a *forward shift* procedure. This procedure simply moves a ball inside the block from one step to the consecutive step. We next prove that the overflow of FRC strictly increases after a forward shift. This implies that in a valid block with a maximum overflow, one may not apply any forward shift while keeping the block valid. This characterizes the block with the maximum overflow uniquely (given its length). Formally, a forward shift $fs(X, k, j)$ is defined within a block(X, k, d) such that $0 < j < d$ and $X_{k+j} \ge 1$, and results in a ball stream X' in which $X'_i = X_i$ for all $i \notin \{k+j, k+j+1\}$, $X'_{k+j} = X_{k+j} - 1$, and $X'_{k+j+1} = X_{k+j+1} + 1$. We say that a forward shift is *admissible* if it keeps the validity of the block. The following observation identifies a condition for the validity of a block after a forward shift.

Observation 2.11. *For $0 < j < d$, if $L_{k+j}^{\mathrm{OPT}}(X) \ge 2$ and $X_{k+j} \ge 1$ then $fs(X, k, j)$ is admissible.*

Proof. For $X' = fs(X, k, j)$, we have $L_{k+r}^{\mathrm{OPT}}(X') = L_{k+r}^{\mathrm{OPT}}(X)$ for $r \ne j$ and $L_{k+j}^{\mathrm{OPT}}(X') = L_{k+j}^{\mathrm{OPT}}(X) - 1 \ge 1$. □

Lemma 2.12. *Applying a $fs(X, k, j)$ increases the overflow of the block(X, k, d) in FRC.*

Proof. For $0 < j < d$, let $X' = fs(X, k, j)$, then by Observation 2.10(2), we have

$$O_{k+d}^{\mathrm{FRC}}(X') = O_{k+d}^{\mathrm{FRC}}(X) + \beta^{d-j-1} - \beta^{d-j} > O_{k+d}^{\mathrm{FRC}}(X) \quad \square$$

Corollary 2.13. *Given a valid ball stream X, if we apply an admissible forward shift within some block of X then we get a valid ball stream X' for which $\rho(X') > \rho(X)$.*

As a result of the last lemma, we can now characterize the worst valid block in terms of overflow for every block length.

Lemma 2.14. *Given a valid ball stream X with a maximal gain ratio $\rho(X)$ that contains some block(X, k, d) then $X_{k+j} = 0$ for $j < \min\{B, d\}$, $X_{k+j} = 1$ for $B \le j < d$, and $X_{k+d} = \min\{B, d\}$.*

Proof. Assume for contradiction that a block with a different structure participates in the ball stream with the worst gain ratio. Let j be the first index it differs from the proposed structure, we will show that in this case $L^{\text{OPT}}_{k+j}(X) \geq 2$ and $X_{k+j} \geq 1$, so by Observation 2.11 $fs(X,k,j)$ is admissible and by Corollary 2.13, we get a contradiction. If $j < \min\{d, B\}$, we have $X_{k+j} \geq 1$, so by Observation 2.9(3), $L^{\text{OPT}}_{k+j}(X,k) = B - j + X_{k+j} \geq 2$. For $B \leq j < d$, we have $L^{\text{OPT}}_{k+j}(X,k) = X_{k+j}$ since we assume that j is the first index that differs from the proposed structure. Therefore $X_{k+j} \geq 1$ since the block is valid, and if $X_{k+j} \neq 1$ then $L^{\text{OPT}}_{k+j}(X) = X_{k+j} \geq 2$. Finally, $X_{k+d} = \min\{B, d\}$ since by Observation 2.9(1), $L^{\text{OPT}}_{k+d}(X,k) = B$. $\square$

Proof of Theorem 2.6. We first analyze the maximal gain ratio for a valid block(X,k,d). Using Lemma 2.14 and Observation 2.10(2), one can derive that if $d \leq B$ then $O^{\text{FRC}}_{k+d}(X) = B \cdot \beta^d + d - B$ and if $d > B$ then

$$\begin{aligned} O^{\text{FRC}}_{k+d}(X) &= B \cdot \beta^d + \left(\sum_{i=B}^{d-1} \beta^{d-i}\right) + B \cdot \beta^{d-d} - B \\ &= B \cdot \beta^d + \left(\sum_{i=1}^{d-B} \beta^{i}\right) \\ &= B \cdot \beta^d + \beta \cdot B \cdot \left(1 - \beta^{d-B}\right), \end{aligned}$$

where the last equality follows from $\beta = 1 - 1/B$. Recall that the gain of OPT in a valid block of length d is d. So the worst gain ratio for a valid block(X,k,d) is $d/(d - O^{\text{FRC}}_{k+d}(X))$. If we substitute the formulas above in the last expression then the worst gain ratio happens when $d > B$. Specifically, we get by substituting $\alpha = d/B$ that $d/(d - O^{\text{FRC}}_{k+d}(X)) =$

$$\frac{d}{d - B \cdot \beta^d - \beta \cdot B \cdot \left(1 - \beta^{d-B}\right)} = \frac{d/B}{d/B - \beta^d - \beta \cdot \left(1 - \beta^{d-B}\right)} \xrightarrow[B \to \infty]{} \frac{\alpha}{\alpha - e^{-\alpha} - (1 - e^{1-\alpha})}$$

since $\beta^d = (1 - 1/B)^{B \cdot \alpha} \to e^{-\alpha}$ and $\beta \cdot (1 - \beta^{d-B}) = (1 - 1/B) \cdot (1 - (1 - 1/B)^{B \cdot (\alpha - 1)}) \to 1 - e^{1-\alpha}$. One can analytically verify that the worst gain ratio is attained for $\alpha \approx 1.429$, and its value is $\rho \approx 1.69996$.

We turn to bound $\rho(X)$ for any valid ball stream X. Given a ball stream X, we divide it into three parts: the first part consists of all ball arrivals until the first overflow of FRC, the second part consists of all the complete (valid) blocks defined by FRC, and the last part consists of the remainder of the ball stream. Similarly to Lemma 2.14, one can demonstrate that to maximize the overflow, if the length of the first part of X (until the first overflow) is d', in each of the first $d' - 1$ steps, a single ball arrives, and then B balls arrive in step d'. In particular, the number of balls that are thrown away in this overflow can be easily shown to be $O^{\text{FRC}}_{d'}(X) = \sum_{i=1}^{d'-1} \beta^i = \beta \cdot B \cdot (1 - \beta^{d'-1})$. Regarding the last part of the stream, we may assume that it is empty since we know that FRC has no overflows within it, and thus, it does not lose any gain with respect to OPT. This implies that after the end of the last block, the gain of FRC and OPT is exactly the content of their buffer, namely, B balls. To sum up, the gain of FRC in the first and last parts is $d' - \beta \cdot B \cdot (1 - \beta^{d'-1}) + B$, while the gain of OPT is $d' + B$. One can analytically verify that the worst gain ratio happens when $d' \approx 1.146B$, and its value is roughly 1.466. Now, if there are K blocks in the ball stream with parameters k_i, d_i to indicate the starting index and length of each block i, respectively, then we get that $G^{\text{OPT}}(X) = d' + B + \sum_{i=1}^{K} d_i$ and $G^{\text{FRC}}(X) = d' - O^{\text{FRC}}_{d'}(X) + B + \sum_{i=1}^{K} \left(d_i - O^{\text{FRC}}_{k_i+d_i}(X)\right)$ we have

$$\begin{aligned} \frac{G^{\text{OPT}}(X)}{G^{\text{FRC}}(X)} &\leq \max_i \left\{ \frac{d' + B}{d' - O^{\text{FRC}}_{d'}(X) + B}, \frac{d_i}{d_i - O^{\text{FRC}}_{k_i+d_i}(X)} \right\} \\ &\leq \rho \quad \square \end{aligned}$$

2.3. Analyzing the randomized process against the fractional one

Our main tool to compare the randomized process against the fractional one is to bound their loads' difference. Specifically, we show that if the difference at step k is 0, the difference at step $k + B$ is at most $\epsilon \cdot B$, with high probability. We first generalize the definition of a ball stream X in which $X_i \in \mathbb{R}_+$. For our random process, X_i indicates that it gets $\lfloor X_i \rfloor$ balls deterministically and an additional ball with probability $X_i - \lfloor X_i \rfloor = \bar{X}_i$ at step i. On the other hand, the fractional process gets exactly X_i balls. Using this definition, in order to upper bound the ratio of FRC(X) and RND(X), it is sufficient to consider streams where FRC does not have an overflow. Let X be such a stream.

Denote $R_k = L^{\text{RND}}_k(X)$ and $F_k = L^{\text{FRC}}_k(X)$, $O_k = O^{\text{RND}}_k(X)$, $Z_k = R_k - F_k$. Our goal is to bound Z_k w.h.p. By our assumption, $F_{k+1} = \beta \cdot F_k + X_k$, by linearity of expectation $\mathbb{E}[R_{k+1} + O_k | R_k] = \beta \cdot R_k + X_k$ and,

$$\begin{aligned} &\mathbb{E}[Z_{k+1} + O_k - Z_k | Z_k] \\ &= \beta \cdot R_k + X_k - (\beta \cdot F_k + X_k) - (R_k - F_k) \\ &= -\frac{Z_k}{B}. \end{aligned} \tag{1}$$

Next, we characterize the sequence Z_k. Let $\delta^Z_k = Z_{k+1} - Z_k$, we observe that $\delta^Z_k + O_k = Z_{k+1} + O_k - Z_k \in \{a - 1, a, a+1\}$ with probabilities p_1, p_2, p_3, respectively, where $a = F_k/B - \bar{X}_k$ and $p_1 = (1 - \bar{X}_k) \cdot R_k/B$, $p_3 = \bar{X}_k \cdot (1 - R_k/B)$ and $p_2 = 1 - p_1 - p_3$. And accordingly, $\delta^Z_k \in \{a - 1 - o_1, a - o_2, a + 1 - o_3\}$ with the same probabilities, where $o_i = \max\{R_k + \lfloor X_k \rfloor - B + i - 2, 0\}$. Note that $-1 \leq a \leq 1$ and $-1 - o_1 \leq -o_2 \leq 1 - o_3$. Next, we observe,

Observation 2.15. *For a sequence X where* FRC *does not overflow: If $Z_k \geq 2$ then $Z_{k+1} \geq 0$. If $Z_k \leq -2$ then $Z_{k+1} \leq 0$ and $O_k = 0$. If $-2 \leq Z_k \leq 2$ then $-4 \leq Z_{k+1} \leq 4$.*

Proof. We have, $Z_k + a + 1 \geq Z_{k+1} \geq Z_k + a - 1 - o_1$. If $o_1 = 0$ the observations hold since $-1 \leq a \leq 1$, and if $o_1 \geq 1$, $R_{k+1} = B$ and therefore $Z_{k+1} \geq 0$. Finally, for $Z_k \leq -2$ we have, $-2 \geq Z_k \geq Z_k + \beta \cdot F_k + X_k - B \geq Z_k + F_k - 1 + X_k - B \geq R_k + \lfloor X_k \rfloor - 1 - B$ which yields $R_k + \lfloor X_k \rfloor \leq B - 1$, and therefore $O_k = 0$. $\square$

By Equation (1), the absolute value of Z_k behaves "almost" as a super-martingale (if the sign of Z_{k+1} is the same as Z_k and $O_k = 0$). Furthermore, if there is overflow in RND, the difference only decreases, and the sign stays the same if Z_k is larger than a constant (2).

To bound $|Z_k|$ formally, we define a martingale Y_k which stochastically dominates Z_k. Specifically, we define a sequence of pairs (Y_k, Z_k), where Y_{k+1} depends only on (Y_k, Z_k), $\mathbb{E}[Y_{k+1}|(Y_k, Z_k)] = Y_k$, and $|Y_k| \geq |Z_k| - 4$ for each pair. Bounding $|Y_k|$ using Azuma's inequality would bound $|Z_k|$ as well. We define:

- **For** $Z_k \geq 2$**,** we define $\delta_k^Y \in \{a-1, a, a+1\}$ with probabilities $p_1 - x_1, p_2 - x_2, p_3 + x_3$, respectively, where $0 \leq x_1 \leq p_1, 0 \leq x_2 \leq p_2, x_3 = x_1 + x_2$ and $E[\delta_k^Y] = 0$.[4] We couple (δ_k^Y, δ_k^Z) such that $\delta_k^Y \geq \delta_k^Z$ (it is possible since $o_i \geq 0$). Set $(Y_{k+1}, Z_{k+1}) = (Y_k, Z_k) + (s \cdot \delta_k^Y, \delta_k^Z)$, where $s = 1$ if $Y_k \geq 0$ and -1 otherwise. We have $|Y_{k+1}| = |Y_k + s \cdot \delta_k^Y| \geq |Y_k| + \delta_k^Y \geq |Z_k| + \delta_k^Z - 4 = |Z_{k+1}| - 4$, where the first inequality is by s definition, the second inequality is by our inductive assumption and by our coupling, and the last equality is since $|Z_{k+1}| = Z_{k+1}$ by Observation 2.15.
- **For** $Z_k \leq -2$**,** we define $\delta_k^Y \in \{a-1, a, a+1\}$ with probabilities $p_1 + x_1, p_2 - x_2, p_3 - x_3$ respectively, where $0 \leq x_3 \leq p_3, 0 \leq x_2 \leq p_2, x_1 = x_3 + x_2$ and $E[\delta_k^Y] = 0$.[5] We couple (δ_k^Y, δ_k^Z) such that $\delta_k^Y \leq \delta_k^Z$ (note that, $o_i = 0$ in this case). Set $(Y_{k+1}, Z_{k+1}) = (Y_k, Z_k) + (s \cdot \delta_k^Y, \delta_k^Z)$, where $s = 1$ if $Y_k \leq 0$ and -1 otherwise. And, similarly to the first case: $|Y_{k+1}| = |Y_k + s \cdot \delta_k^Y| \geq |Y_k| - \delta_k^Y \geq |Z_k| - \delta_k^Z - 4 = |Z_{k+1}| - 4$.
- **For** $|Z_k| \leq 2$**,** we set $(Y_{k+1}, Z_{k+1}) = (Y_k, Z_k + \delta_k^Z)$ by Observation 2.15 $|Y_{k+1}| + 4 \geq 4 \geq |Z_{k+1}|$.

Lemma 2.16. *For $\epsilon = O(1/B^{1/3})$, if $Z_k = 0$ then the probability that load difference between* FRC *and* RND *at step $k+i$, for $i \leq B$ is at most ϵB with probability of at least $1 - \epsilon$.*

Proof. By Azuma's Inequality, since Y_k is a martingale and $a - 1 \leq Y_{k+1} - Y_k \leq a + 1$, for $(Y_k, Z_k) = (0, 0)$ we have for $t = \epsilon \cdot B$:

$$\begin{aligned}
\mathbb{P}\left(|Z_{k+i}(X)| \geq t\right) &\leq \mathbb{P}\left(|Y_{k+i}(X)| \geq t - 4\right) \\
&\leq 2 \cdot \exp\left(-\frac{2 \cdot (t-4)^2}{2^2 \cdot k}\right) \\
&\leq 2 \cdot \exp\left(-\frac{(t-4)^2}{2B}\right) \leq \epsilon,
\end{aligned}$$

where the first inequality is by the dominance of $|Y_k|$ on $|X_k| - 4$, the second inequality is by union bounding the positive and negative bound in Azuma's inequality when the difference of each step is bounded by $a + 1 - (a - 1) = 2$, the last inequality is since $B = \Omega(1/\epsilon^3)$. $\square$

Lemma 2.17. *For $\epsilon = O(1/B^{1/3})$, if $Z_k = 0$ then the total expected gain difference between* FRC *and* RND *in the next B steps is at most $2\epsilon \cdot B$, and the expected absolute load difference between* FRC *and* RND *after the next B steps is at most $2\epsilon \cdot B$.*

Proof. If $Z_k = 0$ the expected gain of RND, then for $i \leq B$

$$\begin{aligned}
\mathbb{E}[G_{k+i}^{\mathrm{RND}}(X)] &\geq \mathbb{E}\left[G_{k+i}^{\mathrm{RND}}(X) \,\middle|\, |Z_{k+i}| \leq \epsilon B\right] \cdot \mathbb{P}[|Z_{k+i}| \leq \epsilon B] \\
&\geq \mathbb{E}\left[L_{k+i}^{\mathrm{RND}}(X)/B \,\middle|\, |Z_{k+i}| \leq \epsilon B\right] \cdot (1 - \epsilon) \\
&\geq (L_{k+i}^{\mathrm{FRC}}(X)/B - \epsilon) \cdot (1 - \epsilon) \geq G_{k+i}^{\mathrm{FRC}}(X) - 2\epsilon,
\end{aligned}$$

where the first inequality is by the law of total probability, the second inequality is by Lemma 2.17, the last inequality is since the gain in each step is at most 1. Similarly, we bound the expected load difference

$$\begin{aligned}
&\mathbb{E}[|L_{k+B}^{\mathrm{RND}}(X) - L_{k+B}^{\mathrm{FRC}}(X)|] \\
&\quad \leq \mathbb{E}\left[|L_{k+B}^{\mathrm{RND}}(X) - L_{k+B}^{\mathrm{FRC}}(X)| \,\middle|\, |Z_i| \leq \epsilon B\right] \\
&\qquad \cdot \mathbb{P}[|Z_i| \leq \epsilon B] + B \cdot \mathbb{P}[|Z_i| > \epsilon B] \\
&\quad \leq \epsilon B \cdot 1 + B \cdot \epsilon \leq 2\epsilon \cdot B,
\end{aligned}$$

where the first inequality is by the law of total probability and since the maximal difference is B, the second inequality is by Lemma 2.16. $\square$

Lemma 2.18. *For $\epsilon = O(1/B^{1/3})$, and for any sequence X of length T: $G^{\mathrm{RND}}(X) \geq G^{\mathrm{FRC}}(X) - 4\epsilon \cdot T$.*

Proof. Given a ball stream of length T, partition it to T/B equal parts of length B. Intuitively, we would like to apply the above lemma to each of the partition parts and attain that the gains of FRC and RND are close to one another. Unfortunately, this is not the case since the loads at the beginning of each part may not be the same. We use the modifications exhibited in Observation 2.3 to manipulate the loads of FRC and RND to be equal after each part. We inductively assume that loads of RND and FRC are the same at the beginning of a part under consideration. Then, we apply Lemma 2.17 that implies that load expected loads difference between FRC and RND after B steps is at most $2\epsilon \cdot B$. At the end of each part, we manipulate the processes as described in Observation 2.3 by either decreasing the load of FRC (to that of RND) or decreasing the load of RND (to that of FRC). In any case, this manipulation may decrease the gain of FRC by at most $2\epsilon \cdot B$ in expectation

[4] Such values exist since for $x_1 = x_2 = x_3 = 0$, $E[\delta_k^Y] = E[Z_{k+1} + O_k - Z_k | Z_k] = -\beta \cdot Z_k \leq 0$, and for $x_1 = p_1$, $x_2 = p_2$, $x_3 = p_1 + p_2$, $E[\delta_k^Y] = a + 1 \geq 0$.

[5] Such values exist since for $x_1 = x_2 = x_3 = 0$, $E[\delta_k^Y] = E[Z_{k+1} + O_k - Z_k | Z_k] = -\beta \cdot Z_k \geq 0$, and for $x_3 = p_3$, $x_2 = p_2$, $x_1 = p_3 + p_2$, $E[\delta_k^Y] = a - 1 \leq 0$.

(and would not increase the gain of RND). This enables us to apply the inductive step while having an additional gain difference loss of at most $2\epsilon \cdot B$ in each part, in addition, by Lemma 2.17 the expected gain loss of RND vs FRC (after the manipulation) is at most $2\epsilon \cdot B$ in each part. In conclusion, by summing over all the parts, we attain that $G^{\mathrm{RND}}(X) \geq G^{\mathrm{FRC}}(X) - T/B \cdot 4\epsilon \cdot B = G^{\mathrm{FRC}}(X) - 4\epsilon \cdot T$. □

We are now ready to complete the proof of main theorem of the paper.

Proof of Theorem 2.1. We know from Lemma 2.4 that it is sufficient to consider valid ball streams for the purpose of bounding $\hat{\rho} = \max_X G^{\mathrm{OPT}}(X)/G^{\mathrm{RND}}(X)$. Now, let $\epsilon' = \epsilon/20$, we have

$$G^{\mathrm{RND}}(X) \geq G^{\mathrm{FRC}}(X) - 4\epsilon' \cdot T \geq \frac{G^{\mathrm{OPT}}(X)}{\rho} - 4\epsilon' \cdot G^{\mathrm{OPT}}(X) = G^{\mathrm{OPT}}(X)\left(\frac{1}{\rho} - \frac{\epsilon}{5}\right)$$

where the first inequality is due to Lemma 2.18 with ϵ', the second inequality is since by Observation 2.5, we have $G^{\mathrm{OPT}}(X) \geq T$, and the third inequality holds by Theorem 2.6. Hence, $\hat{\rho} \leq 1\Big/\left(\frac{1}{\rho} - \frac{\epsilon}{5}\right) \leq \rho + \epsilon$. □

Finally, we prove that the bound ρ is essentially tight.

Proof of Corollary 2.2. For a large enough B, by considering a ball stream where the $block(X, k, d)$ with the worst gain ratio (as presented in Theorem 2.6, a block with $d \approx 1.429B$), appears over and over again. It is easy to validate that the gain ratio on this sequence between OPT and RND converges to ρ. □

3. Applications

3.1. Application 1: value-oblivious packets scheduling

We show that the competitiveness of value-oblivious transmission algorithm, that described in the introduction, can be analyzed by the stochastic process described. The value-oblivious transmission algorithm selects uniformly at random a packet to transmit among the packets in the buffer. We compare it to an unconstrained optimal algorithm which transmits the highest valued packet in the buffer. We claim that it is sufficient to analyze our random transmission algorithm with respect to packets whose values are restricted to the set $\{0, 1\}$. This follows from the zero-one principle below, while observing that our algorithm is indeed a comparison-based algorithm. We note that an algorithm is *comparison-based* if all its decisions are based on the relative order between the values of the bids with no regard to their actual values.

Theorem 3.1. (The zero-one principle [7]) *Let $\mathcal{A}$ be a (deterministic or randomized) comparison-based algorithm. $\mathcal{A}$ is a c-approximation algorithm if and only if A achieves c-approximation for all input whose values are restricted to $\{0, 1\}$ for every possible way of breaking ties between equal values.*

Observe that the expected gain of our algorithm in each transmission is the number of 1-valued packets divided by the overall number of packets. Hence, for the sake of analysis, we may assume that there are exactly B packets in every transmission. Moreover, since our algorithm keeps the packets with the highest values at any step, we may assume that the input stream is $X = \langle X_1, X_2, \ldots X_T \rangle$, where X_i represents the number of 1-valued packets that arrived before transmission i but after transmission $i - 1$. One can easily validate that there is a correspondence between the competitive ratio of our randomized transmission algorithm and the loss of serving in the dark ratio. As a result, we obtain the following theorem.

Theorem 3.2. *The competitive ratio of the randomized transmission algorithm is $\rho + \epsilon$ for $\epsilon = O(1/B^{1/3})$.*

3.2. Application 2: prompt mechanisms for bounded capacity auctions

We consider the problem of developing *prompt truthful mechanisms* for *periodic bounded capacity auctions*. In the underlying scenario, there is a single item with unlimited supply, and a stream of buyers, arriving dynamically over time, each of which is interested in purchasing one instance of the item. An instance of the item is offered for sale in a bounded capacity auction *periodically*, that is, over and over again. A *bounded capacity auction* is a single-item auction in which the number of participating bidders is bounded by a fixed $B \in \mathbb{N}_+$, e.g., when the auction room has a limited size. Since the auction has bounded capacity, it is common that the auction cannot accommodate all the buyers. In such a case, some of the buyers must be indefinitely rejected. These buyers cannot participate in any auction after their rejection. We remark that the auction events continue even if the stream of buyers ceases. This implies that one can sell items to all pending buyers once the stream ends. Our goal is to design prompt truthful mechanisms that maximizes the social welfare, i.e., the sum of the values of the buyers that purchase an item. In a *prompt mechanism* [13], a buyer that purchases an item learns her payment immediately after she wins the auction. In the above scenario, the private information of each buyer is her positive value for purchasing an item. Each buyer declares her bid for purchasing the item once she arrives. This model falls within the scope of online single-parameter setting (see, e.g., [32]). It is well-known that developing a truthful mechanism in this setting is roughly equivalent to designing a monotone algorithm. An algorithm is monotone if a winning buyer, namely, a buyer that purchases an item, remains a winner if she raises her bid. An algorithm for our problem can be logically split into two parts: (1) Once a buyer arrives, the algorithm needs to decide whether to reject that buyer or to keep her active. Note that there must be at most B active buyers at any time, and therefore, it may happen that in order to keep a new buyer active, the algorithm has to reject another active buyer. (2) When an auction occurs, the algorithm needs to decide which of the active bidders wins the item.

One practical motivation for studying the above problem relates to buffer management issues arising in context of network devices such as switches and routers. In this application domain, there is an incoming stream of (strategic) packets with intrinsic values, and there is a network device that can accommodate a bounded number of packets at any time. The device can transmit one packet in each time-slot. The goal is to design a truthful mechanism maximizing the overall value of transmitted packets, while charging any packet for the given service once it is transmitted.

We consider the following algorithms:

The Greedy algorithm. The greedy algorithm keeps the buyers with highest bids at any step. Specifically, once a new buyer arrives, if there are less than B active buyers then the new buyer is kept as active; if the bid of the new buyer is higher than the minimal bid of the current B active buyers then she is kept as active and the active buyer with minimal bid is rejected; otherwise, she is rejected. On the other hand, when an auction occurs, the item is sold to an active buyer with a maximum bid.

One can easily verify that this simple algorithm achieves optimal outcome and that it is monotone. Therefore, the algorithm can underlie an optimal truthful mechanism. However, this algorithm does not support prompt payments. In particular, there are simple input instances for which a buyer may have to wait indefinitely to learn her payment.

The FIFO algorithm. This algorithm activates and rejects buyers in an identical way to the greedy algorithm, that is, it keeps the buyers with highest bids at any step. However, when an auction occurs, the item is sold to the active buyer that arrived earliest. One can easily verify that this algorithm is monotone and supports prompt payments. Moreover, it is 2-competitive [22].

The *randomized selection* algorithm. The proposed algorithm activates and rejects buyers in an identical way to the greedy algorithm, that is, it keeps the buyers with highest bids at any step. When an auction occurs, one buyer is selected uniformly at random from all active buyers, and the item is sold to her. We prove that this algorithm is universally truthful (i.e. it is truthful for any possible coins flip), supports prompt payments, and achieves an expected competitive ratio of $\rho+\epsilon$ which is strictly better than 2. Recall that $\rho \approx 1.69996$.

We begin by proving that the randomized selection algorithm is truthful and supports prompt payments.

Lemma 3.3. *The randomized selection mechanism is monotone and supports prompt payments.*

Proof. Consider two input instances that are identical with the exception that in the first instance the bid of buyer i is v and in the second instance her bid is $\tilde{v}$, where $\tilde{v} \geq v$. For the sake of monotonicity, we need to prove that if our algorithm selects i as a winner when her bid is v then it also selects her as a winner in the latter instance. This clearly happens. Specifically, notice that buyer i is kept active when she arrives since $\tilde{v} \geq v$. Moreover, she cannot be rejected by buyers arriving later since she was not rejected when her bid was v. This implies that she must win the same auction in the latter instance.

As for the promptness of the corresponding mechanism, observe that the selection of a buyer as a winner only depends on the bids provided by buyers arriving before she wins. In particular, a buyer is kept active as long as her bid is sufficiently large with respect to the other buyers, and the selection of the winner in each auction is determined by the random selection process which is independent of the bids of the buyers. Thus, we can calculate the payment of a winning buyer immediately after she wins the item. □

By the same arguments described in Section 3.1, the competitive ratio of the mechanism is equivalent to the loss of serving in the dark ratio presented in Section 2. Therefore, we conclude:

Theorem 3.4. *The randomized selection mechanism is truthful, supports prompt payments and its competitive ratio is $\rho+\epsilon$ for $\epsilon = O(1/B^{1/3})$.*

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ilan Reuven Cohen reports financial support was provided by Israel Science Foundation. Yossi Azar reports financial support was provided by Israel Science Foundation.

Data availability

No data was used for the research described in the article.

References

[1] W. Aiello, Y. Mansour, S. Rajagopolan, A. Rosén, Competitive queue policies for differentiated services, J. Algorithms 55 (2) (2005) 113–141.
[2] N. Alon, J.H. Spencer, The Probabilistic Method, second edition, Wiley, New York, 2000.
[3] N. Andelman, Y. Mansour, Competitive management of non-preemptive queues with multiple values, in: Proceedings 17th International Conference on Distributed Computing, 2003, pp. 166–180.
[4] N. Andelman, Y. Mansour, A. Zhu, Competitive queueing policies for qos switches, in: Proceedings 14th Annual ACM-SIAM Symposium on Discrete Algorithms, 2003, pp. 761–770.
[5] Y. Azar, A.Z. Broder, A.R. Karlin, E. Upfal, Balanced allocations, SIAM J. Comput. 29 (1) (1999) 180–200.
[6] Y. Azar, I.R. Cohen, I. Gamzu, The loss of serving in the dark, in: Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing, 2013, pp. 951–960.
[7] Y. Azar, Y. Richter, The zero-one principle for switching networks, in: Proceedings 36th Annual ACM Symposium on Theory of Computing, 2004, pp. 64–71.
[8] N. Bansal, L. Fleischer, T. Kimbrel, M. Mahdian, B. Schieber, M. Sviridenko, Further improvements in competitive guarantees for qos buffering, in: Proceedings 31st International Colloquium on Automata, Languages and Programming, 2004, pp. 196–207.
[9] Y. Bartal, F.Y.L. Chin, M. Chrobak, S.P.Y. Fung, W. Jawor, R. Lavi, J. Sgall, T. Tichý, Online competitive algorithms for maximizing weighted throughput of unit jobs, in: Proceedings 21st Annual Symposium on Theoretical Aspects of Computer Science, 2004, pp. 187–198.
[10] M. Bienkowski, M. Chrobak, C. Dürr, M. Hurand, A. Jez, L. Jez, G. Stachowiak, Collecting weighted items from a dynamic queue, in: Proceedings 20th Annual ACM-SIAM Symposium on Discrete Algorithms, 2009, pp. 1126–1135.

[11] F.Y.L. Chin, S.P.Y. Fung, Online scheduling with partial job values: does timesharing or randomization help?, Algorithmica 37 (3) (2003) 149–164.
[12] M. Chrobak, W. Jawor, J. Sgall, T. Tichý, Improved online algorithms for buffer management in qos switches, ACM Trans. Algorithms 3 (4) (2007).
[13] R. Cole, S. Dobzinski, L. Fleischer, Prompt mechanisms for online auctions, in: Proceedings 1st International Symposium on Algorithmic Game Theory, 2008, pp. 170–181.
[14] D.P. Dubhashi, A. Panconesi, Concentration of Measure for the Analysis of Randomized Algorithms, Cambridge University Press, 2009.
[15] M. Englert, M. Westermann, Lower and upper bounds on fifo buffer management in qos switches, in: Proceedings 14th Annual European Symposium on Algorithms, 2006, pp. 352–363.
[16] M. Englert, M. Westermann, Considering suppressed packets improves buffer management in qos switches, in: Proceedings 18th Annual ACM-SIAM Symposium on Discrete Algorithms, 2007, pp. 209–218.
[17] A. Fiat, Y. Mansour, U. Nadav, Competitive queue management for latency sensitive packets, in: Proceedings 19th Annual ACM-SIAM Symposium on Discrete Algorithms, 2008, pp. 228–237.
[18] B. Hajek, On the competitiveness of online scheduling of unit-length packets with hard deadlines in slotted time, in: Proceedings Conference on Information Sciences and Systems, 2001, pp. 434–438.
[19] M.T. Hajiaghayi, R.D. Kleinberg, M. Mahdian, D.C. Parkes, Online auctions with re-usable goods, in: Proceedings 6th ACM Conference on Electronic Commerce, 2005, pp. 165–174.
[20] N.L. Johnson, S. Kotz, Urn Models and Their Applications, John Wiley & Sons, 1977.
[21] R.M. Karp, U.V. Vazirani, V.V. Vazirani, An optimal algorithm for online bipartite matching, in: Proceedings 22nd Annual ACM Symposium on Theory of Computing, 1990, pp. 352–358.
[22] A. Kesselman, Z. Lotker, Y. Mansour, B. Patt-Shamir, B. Schieber, M. Sviridenko, Buffer overflow management in qos switches, SIAM J. Comput. 33 (3) (2004) 563–583.
[23] A. Kesselman, Y. Mansour, R. van Stee, Improved competitive guarantees for qos buffering, Algorithmica 43 (1–2) (2005) 63–80.
[24] V.F. Kolchin, B.A. Sevastyanov, V.P. Chistyakov, Random Allocations, John Wiley & Sons, 1978.
[25] T.G. Kurtz, Solutions of ordinary differential equations as limits of pure jump Markov processes, J. Appl. Probab. 7 (1970) 49–58.
[26] R. Lavi, N. Nisan, Online ascending auctions for gradually expiring items, in: Proceedings 16th Annual ACM-SIAM Symposium on Discrete Algorithms, 2005, pp. 1146–1155.
[27] F. Li, J. Sethuraman, C. Stein, An optimal online algorithm for packet scheduling with agreeable deadlines, in: Proceedings 16th Annual ACM-SIAM Symposium on Discrete Algorithms, 2005, pp. 801–802.
[28] F. Li, J. Sethuraman, C. Stein, Better online buffer management, in: Proceedings 18th Annual ACM-SIAM Symposium on Discrete Algorithms, 2007, pp. 199–208.
[29] C. McDiarmid Concentration, in: Probabilistic Methods for Algorithmic Discrete Mathematics, Springer, 1998.
[30] M. Mitzenmacher, A.W. Richa, R. Sitaraman, The Power of Two Random Choices: A Survey of Techniques and Results, Handbook of Randomized Computing, vol. 1, Springer, 2001, pp. 255–312.
[31] M. Mitzenmacher, E. Upfal, Probability and Computing - Randomized Algorithms and Probabilistic Analysis, Cambridge University Press, 2005.
[32] D.C. Parkes, Online mechanisms, in: N. Nisan, T. Roughgarden, E. Tardos, V. Vazirani (Eds.), Algorithmic Game Theory, Cambridge University Press, 2007, chapter 16.
[33] M. Schmidt, Packet buffering: randomization beats deterministic algorithms, in: Proceedings 22nd Annual Symposium on Theoretical Aspects of Computer Science, 2005, pp. 293–304.