Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Achievable Performance of Blind Policies in Heavy Traffic

Nikhil Bansal, Bart Kamphorst, Bert Zwart

informs
http://pubsonline.informs.org/journal/moor/

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 43, No. 3, August 2018, pp. 949–964
ISSN 0364-765X (print), ISSN 1526-5471 (online)

Achievable Performance of Blind Policies in Heavy Traffic

Nikhil Bansal,[a, b] **Bart Kamphorst,**[a] **Bert Zwart**[a, b]

[a] Centrum Wiskunde and Informatica, 1090 GB Amsterdam, Netherlands; [b] Technische Universiteit Eindhoven, 5600 MB Eindhoven, Netherlands

Contact: n.bansal@tue.nl (NB); b.kamphorst@cwi.nl (BK); b.zwart@cwi.nl, http://orcid.org/0000-0001-9336-0096 (BZ)

Received: December 23, 2015
Revised: December 23, 2016
Accepted: March 26, 2017
Published Online in Articles in Advance: March 8, 2018

MSC2010 Subject Classification: Primary: 60K25; secondary: 68M20, 90B22, 90B36
OR/MS Subject Classification: Primary: probability; secondary: applications

https://doi.org/10.1287/moor.2017.0890

Abstract. For a GI/GI/1 queue, we show that the average sojourn time under the (blind) Randomized Multilevel Feedback algorithm is no worse than that under the Shortest Remaining Processing Time algorithm times a logarithmic function of the system load. Moreover, it is verified that this bound is tight in heavy traffic, up to a constant multiplicative factor. We obtain this result by combining techniques from two disparate areas: competitive analysis and applied probability.

Funding: The research of Nikhil Bansal is partly supported by the NWO VIDI [Grant 639.022.211] and the ERC consolidator [Grant 617951]. The research of Bart Kamphorst is supported by the NWO free competition research programme 613.001.219. The research of Bert Zwart is partly supported by the NWO VICI [Grant 639.033.413].

Keywords: GI/GI/1 queue • response time • heavy traffic • competitive ratio • blind policies • shortest remaining processing time • randomized multilevel feedback

1. Introduction

One of the most relevant and widely studied measures of quality of service in a GI/GI/1 queue is the average sojourn time, also known as response time or flow time, defined as the average time spent by a job from its arrival in the system until its completion (Bansal and Harchol-Balter [5], Bansal [3], Bansal and Gamarnik [4], Becchetti and Leonardi [6], Boxma and Zwart [8], Lin et al. [15], Nair et al. [19], Nuyens et al. [21], Wierman and Nuyens [28], Wierman and Zwart [29], Wierman et al. [30]). We consider the most basic setting of a single machine with preemption, i.e., jobs can be interrupted arbitrarily and resumed later without any penalty. It is well known that the Shortest Remaining Processing Time (SRPT) policy, that at any time works on the job with the least remaining processing time, is the optimal policy for every problem instance (or equivalently, for every sample path) for minimizing the sojourn time (Schrage [26]). However, to run SRPT, one needs the exact knowledge of all job sizes. This information may not be available in many settings; specifically, jobs sizes may only be known approximately, or may not be known at all (Lu et al. [17]). For this reason, one may have to be content with more generally applicable policies.

In this paper, we are interested in policies that do not require the knowledge of job sizes in their scheduling decisions. We refer to such policies as *blind* policies. More formally, in a blind policy, the scheduler is only aware of the existence of a job and how much processing it has received thus far. The size of the job becomes known to the scheduler only when it terminates and leaves the system. Observe that the class of blind policies contains several well-studied policies, such as Processor Sharing (also known as Round Robin) (Kleinrock [14]), Foreground-Background (Nuyens and Wierman [20]), and First In, First Out (Asmussen [2]).

It is natural to ask how much this inability to use the knowledge of sizes can hurt performance. In particular, how much can the average sojourn time between SRPT and an optimal blind policy differ for a given GI/GI/1 queue? As an illustration, let us consider the M/M/1 queue. In this setting, all blind policies are identical due to the memoryless nature of the job size distribution. More precisely, Conway et al. [9] state that any blind policy has an average sojourn time equal to $\mathbb{E}[B]/(1-\rho)$, where $\mathbb{E}[B]$ is the average job size and ρ is the load of the system. On the other hand, if job sizes are known upon arrival, then Bansal [3] derives that the average sojourn time $\bar{T}_{\mathrm{SRPT}}$ under M/M/1/SRPT is

$$\bar{T}_{\mathrm{SRPT}} = (1+o(1))\frac{1}{\log(e/(1-\rho))}\frac{\mathbb{E}[B]}{1-\rho}, \tag{1}$$

where $o(1)$ vanishes as ρ approaches one. That is, the SRPT policy performs better by a factor $\log(e/(1-\rho))$ in heavy traffic. So while SRPT can be arbitrarily better than a blind policy for a M/M/1 queue as the load approaches one, this improvement factor grows quite mildly.

The performance of SRPT as a function of load can be dramatically different for heavy-tailed distributions. Bansal and Gamarnik [4] and Lin et al. [15] show that the growth factor of the average sojourn time in heavy traffic can be much smaller than $1/(1-\rho)$, even in M/GI/1 queues. For example, if the job sizes follow a Pareto(β) distribution with $\beta \in (1,2)$, then the growth factor of the average sojourn time $\bar{T}_{\mathrm{SRPT}}$ is $\mathbb{E}[B]\log(1/(1-\rho))$, up to constant factors depending on β. On the other hand, Kleinrock [14] states that Processor Sharing has an average sojourn time of $\mathbb{E}[B]/(1-\rho)$ in any M/GI/1 queue. As this example illustrates, it is conceivable that for a general distribution, the gap between blind policies and SRPT can be much larger than in the M/M/1 case.

Another subfield of computer science where the performance improvement of SRPT over blind policies has been studied is *competitive analysis* (Borodin and El-Yaniv [7], Fiat and Woeginger [11], Pruhs et al. [22]), which generally regards worst-case analyses of algorithms. The study of competitive analysis of blind scheduling policies was initiated by Motwani et al. [18], who showed that no blind deterministic algorithm[1] can have a better competitive ratio than $\Omega(m^{1/3})$ for the problem of minimizing the average sojourn time, where m is the number of jobs in an instance. Motwani et al. [18] also showed that no blind randomized algorithm can have a competitive ratio better than $\Omega(\log(m))$. In a significant breakthrough, Kalyanasundaram and Pruhs [13] gave an elegant and nontrivial randomized algorithm that they called Randomized Multilevel Feedback (RMLF) and proved that it has competitive ratio of at most $O(\log(m)\log(\log(m)))$. Later, Becchetti and Leonardi [6] showed that RMLF is, in fact, an $O(\log(m))$-competitive randomized algorithm, and hence the best possible (up to constant factors). Additional background on multilevel algorithms can be found in Kleinrock [14], and an analysis of the average sojourn time under such algorithms is performed in Aalto and Ayesta [1].

The result in Becchetti and Leonardi [6] is derived under the assumption that job sizes are bounded from below by a strictly positive constant, an assumption that is removed in this paper. This "extended" version of RMLF is denoted by eRMLF.

The insights from applied probability and competitive analysis concerning the relation between blind policies and SRPT can be combined when m is taken as the number of jobs in a regeneration cycle, which has an expected value of the order $1/(1-\rho)$. We make this precise in our main theorem and its proof. The main theorem shows that, for a GI/GI/1 queue, the gap between SRPT and the best blind policy $\mathscr{A}$ for that system is at most $\log(1/(1-\rho))$, up to constant factors. More specifically, we show that this growth factor is a guaranteed upper bound on the gap between SRPT and the eRMLF algorithm. That is, we show that

$$\mathbb{E}[\bar{T}_{\mathscr{A}}] \le \mathbb{E}[\bar{T}_{\mathrm{eRMLF}}] = O\left(\log\left(\frac{1}{1-\rho}\right)\right) \cdot \bar{T}_{\mathrm{SRPT}} \tag{2}$$

as ρ grows to one. Note that the eRMLF algorithm makes random decisions, and as such, the outcome of $\bar{T}_{\mathrm{eRMLF}}$ is a stochastic random variable for any given instance. The same may hold for the optimal blind policy $\mathscr{A}$. Also, we emphasise that the implementation of the RMLF algorithm does not depend on the distributions of interarrival times and job sizes, and is therefore applicable to every GI/GI/1 queue; this property may not hold for $\mathscr{A}$ in general.

The second main contribution of this paper is the proof of (2) itself. It involves a novel combination of techniques from competitive analysis and applied probability. Using a renewal argument, we consider the average sojourn time $\mathbb{E}[\bar{T}_{\mathrm{RMLF}}]$ of jobs in a general busy period, and subsequently, distinguish two types of busy periods (small and large) by the number of jobs. For small busy periods, we apply a worst-case performance bound of RMLF from the study of competitive analysis. For large busy periods, we derive the heavy-traffic behaviour of moments of two functionals: the busy period duration and the number of jobs in a busy period. In particular, we show that the κ-th moment of both of these functionals behaves like $O((1-\rho)^{1-2\kappa})$ for $\kappa \ge 1$; a new result, which may facilitate future instances where competitive analysis and regenerative process theory can be combined to obtain information about algorithms under uncertainty. To prove these bounds, we rely on properties of ladder-height distributions derived in Asmussen [2] and Lotov [16].

This paper is organized as follows. A detailed model description and notation is introduced in Section 2. Section 3 clarifies the concept of a competitive ratio and describes the RMLF algorithm. Additionally, Section 3.2 relaxes the constraints on RMLF while preserving the competitive ratio. The main result, Theorem 3, is presented in Section 4, whereas its proof is given in Section 5. Propositions required for the main theorem are proven in Section 6. Finally, Section 7 concludes the paper.

2. Preliminaries

This section introduces a general framework for sequences of GI/GI/1 queues, so that we may analyse their limiting behaviour in further sections. In particular, the model allows for a heavy-traffic analysis of the average sojourn time and various other functionals.

Sequence of Queues

Consider a sequence of GI/GI/1 queues, indexed by $n \geq 1$, where jobs arrive sequentially with independent and identically distributed (i.i.d.) sizes $B_i^{(n)}$, $i \in \{1, 2, \ldots\}$, chosen from a distribution $F_B^{(n)}$. The jobs are then processed by a single server with unit speed. The time between two consecutive job arrivals is given by the i.i.d. interarrival times $A_i^{(n)}$, $i \in \{1, 2, \ldots\}$, chosen from a distribution $F_A^{(n)}$. All jobs and interarrival times are assumed to be positive, i.e., the support of $F_A^{(n)}$ and $F_B^{(n)}$ is contained in $(0, \infty)$. For notational convenience, we define $A^{(n)} := A_1^{(n)}$ and $B^{(n)} := B_1^{(n)}$.

In order for every queue to be stable, we require $\mathbb{E}[A^{(n)}] > \mathbb{E}[B^{(n)}]$ for all $n \geq 1$. The traffic intensity of the n-th system is denoted by $\rho^{(n)} := \mathbb{E}[B^{(n)}]/\mathbb{E}[A^{(n)}] \in [\rho_0, 1)$, and is interpreted as the fraction of time that the server is busy. As is customary in the literature on heavy-traffic analysis (Gromoll [12], Puha et al. [23]), we assume $\lim_{n\to\infty} \rho^{(n)} = 1$. The mean change in backlog between two consecutive arrivals is represented by $\mu^{(n)} := \mathbb{E}[A^{(n)}] - \mathbb{E}[B^{(n)}] = \mathbb{E}[A^{(n)}](1 - \rho^{(n)})$.

Furthermore, we require that the interarrival times have finite variance for all n and additionally that $\limsup_{n\to\infty} \mathbb{E}[(A^{(n)})^2] < \infty$. Since a queue can only form when a job arrives to a nonempty system, we pose the final requirement that for some $\delta > 0$ and $\gamma > 0$ independent of n, the system satisfies $\mathbb{P}(B^{(n)} - A^{(n)} \geq \delta) \geq \gamma$.

Example 1. To interpret some of our obtained results, one may compare them to a M/GI/1 queue that is sent into heavy traffic in a natural manner. Specifically, assume that the A_i's and B_i's have unit mean and consider interarrival times $A_i^{(r)} = A_i/r$, $i \in \{1, 2, \ldots\}$, $r \in (0, 1)$. This model experiences a load of $\rho = \mathbb{E}[B]/\mathbb{E}[A^{(r)}] = r$ and is exposed to heavy traffic as r tends to one due to decreasing interarrival times. The model fits in the framework described above by letting $A_i^{(n)} = A_i/(1 - 1/n)$, $B_i^{(n)} = B_i$, and $\rho^{(n)} = 1 - 1/n$, and is referred to as the *Example Model*. All further remarks on the Example Model are emphasised by superscripts r for all related variables and functionals; e.g., the notation $A_i^{(r)}$ indicates the i-th interarrival time in this Example Model.

Queueing Functionals

The sojourn time of a job is the amount of time it spends in the system; namely, its service completion time minus its arrival time. Given a scheduling policy π, we denote the average sojourn time of a generic job by $\mathbb{E}[\bar{T}_\pi^{(n)}]$ or just $\bar{T}_\pi^{(n)}$ if the policy is deterministic. The steady-state cumulative amount of work in the system is represented by $V^{(n)}$, whose distribution has an atom at zero that corresponds to the times when the server is idle. The steady-state duration of such an idle period is denoted by $I^{(n)}$. Idle periods are ended by the arrival of a new job, which initiates a busy period. A busy period finishes at the earliest subsequent time for which the system is empty again. The steady-state duration of a busy period is represented by $P^{(n)}$, whereas the total number of arrivals during a busy period is denoted by $N^{(n)}$. Finally, the steady-state cumulative amount of work in the system *at an arrival instance* is represented by $W^{(n)}$.

Scheduling Policies

A scheduling policy π is an algorithm or a rule that specifies which job receives service at any time in the system. For the GI/GI/1 queue under consideration, such a policy prescribes the behaviour of a single server under the relaxation that jobs can be *preempted*; that is, jobs can be interrupted at any point during their execution and can be resumed later from this point without any penalty. Of the large class of scheduling policies that apply to this system, we consider only those policies π that satisfy the following two criteria (quoted from Wierman and Zwart [29], after Stolyar and Ramanan [27]):

1. π is *nonanticipative*: a scheduling decision at time t does not depend on information about jobs that arrive beyond time t.

2. π is *nonlearning*: the scheduling decisions cannot depend on information about previous busy periods. That is, a scheduling decision on a sample path cannot change when the history before the current busy period is changed.

Of special interest are those scheduling policies π that additionally obey the following characteristic:

3. π is *blind*: the scheduling decisions do not depend on the sizes of the jobs. That is, the scheduling decisions on a sample path up to time t cannot change when the sizes of jobs that have not finished at that time are altered (in such a way that the jobs remain unfinished).

Policies that satisfy all above criteria are very common: First In, First Out (FIFO), Processor Sharing and Foreground-Background (FB) are all blind policies within the specified subclass of scheduling policies. On the other hand, policies like Shortest Job First or Shortest Remaining Processing Time (SRPT) are *nonblind* elements of the specified subclass as they require knowledge of the job sizes when making a scheduling decision.

We let $\mathscr{A}^{(n)}$ denote a blind policy that minimizes the average sojourn time over the space of all blind policies for the n-th GI/GI/1 queue. In general, $\mathscr{A}^{(n)}$ could depend on the distributions $F_A^{(n)}$ and $F_B^{(n)}$ that specify the GI/GI/1 queue. We denote RMLF and eRMLF to be the randomized blind algorithms that are formalized in Sections 3.1 and 3.2, respectively. The implementation of the RMLF and eRMLF algorithms does not depend on $F_A^{(n)}$ and $F_B^{(n)}$ and is therefore independent of the system index n.

Finally, we call a scheduling policy π *work-conserving* if it always has the server working at unit speed whenever work is present in the system. One can easily verify that all above policies, including $\mathscr{A}^{(n)}$ are work-conserving.

Asymptotic Relations

We use the standard notation that for two functions $f(n)$ and $g(n)$, $f(n) = O(g(n))$ and $f(n) = o(g(n))$ if $\limsup_{n\to\infty} f(n)/g(n) < \infty$ and $\limsup_{n\to\infty} f(n)/g(n) = 0$, respectively. Similarly, $f(n) = \Omega(g(n))$ means $\liminf_{n\to\infty} f(n)/g(n) > 0$ and $f(n) = \Theta(g(n))$ is equivalent to $0 < \liminf_{n\to\infty} f(n)/g(n) \le \limsup_{n\to\infty} f(n)/g(n) < \infty$.

Final notational conventions in this paper are the floor-function $\lfloor x \rfloor := \sup\{m \in \mathbb{N}: m \le x\}$ and the indicator function $\mathbb{1}([\textit{logical expression}])$ that takes value 1 if the logical expression is true, and value 0 otherwise.

3. Competitive Analysis of Scheduling Policies

In this section, we describe some relevant definitions and results from the area of competitive analysis, which deals with the worst-case analysis of algorithms. We restrict our presentation here only to the competitive analysis of scheduling algorithms with respect to average sojourn time. Subsequently, we introduce the original RMLF algorithm and its extension eRMLF.

A scheduling problem instance $\mathcal{J}$ consists of a collection of jobs specified by their sizes and their arrival times. We say that an instance has size m, i.e., $|\mathcal{J}| = m$, if it consists of m jobs. For an instance $\mathcal{J}$, we denote the optimal average sojourn time possible for this instance by $\mathbb{E}[\bar{T}_{\mathrm{OPT}}(\mathcal{J})]$, which for our purposes, is same as $\bar{T}_{\mathrm{SRPT}}(\mathcal{J})$.

For a deterministic algorithm π, we let $\bar{T}_\pi(\mathcal{J})$ denote the average sojourn time when the instance $\mathcal{J}$ is executed according to the algorithm π. We say that the algorithm π has competitive ratio $c(m)$ if

$$\sup_{\mathcal{J}:|\mathcal{J}|\le m} \frac{\bar{T}_\pi(\mathcal{J})}{\bar{T}_{\mathrm{OPT}}(\mathcal{J})} \le c(m).$$

Thus the competitive ratio of an algorithm (possibly a function of m) is the worst-case ratio over all input instances of length at most m of the sojourn time achieved by π and the optimal sojourn time on that instance. Observe that the definition of the competitive ratio is rather strict, in that even if an algorithm is close to optimal on all but one input instance, its competitive ratio will be lower bounded by its performance on the bad input instance.

For this purpose, it is useful to consider randomized algorithms. A randomized algorithm $\tilde{\pi}$ can toss coins internally and base its decisions on this outcome of these internal random variables. Such an algorithm can thus be viewed as a probability distribution over blind deterministic algorithms π_i (Borodin and El-Yaniv [7]). It then follows that the average sojourn time of instance $\mathcal{J}$ under a randomized algorithm $\tilde{\pi}$ equals $\mathbb{E}[\bar{T}_{\tilde{\pi}}(\mathcal{J})] = \mathbb{E}_i[\bar{T}_{\pi_i}(\mathcal{J})]$, where the expectation is over the internal random choices of the algorithm. We say that $\tilde{\pi}$ has competitive ratio $c(m)$ if

$$\sup_{\mathcal{J}:|\mathcal{J}|\le m} \frac{\mathbb{E}[\bar{T}_{\tilde{\pi}}(\mathcal{J})]}{\bar{T}_{\mathrm{OPT}}(\mathcal{J})} \le c(m). \tag{3}$$

Observe that the expectation is only over the random choices made by the algorithm, and the competitive ratio is still determined by the worst-possible instance. However, the competitive ratio of a blind randomized algorithm can be substantially lower, e.g., in situations where no single blind deterministic algorithm is good for all instances, but a suitable combination of algorithms is close to optimal for all instances.

3.1. Randomized Multilevel Feedback Algorithm

This section introduces Kalyanasundaram and Pruhs [13]'s Randomized Multilevel Feedback (RMLF) algorithm. As the name suggests, it is a randomized version of the Multilevel Feedback (MLF) algorithm proposed by Corbató et al. [10]. Both algorithms are blind and can therefore only learn the size of a job upon completion.

The general idea of MLF and RMLF is to prioritize potential short jobs (e.g., jobs that have not received much service) and reduce the priority of a job as it receives more service. This prioritisation is embodied by assigning every job J_j to a virtual high priority queue Q_i, and move it to a lower priority queue Q_{i+1} once it has received

Figure 1. Formal statement of RMLF algorithm.

Algorithm RMLF: At all times, the collection of released but uncompleted jobs is partitioned into queues, $Q_0, Q_1, \dots$ We say that Q_i is lower than Q_j for $i < j$. For each job, $J_j \in Q_i, U_{i,j} \in [2^i, 2^{i+1}]$ when it entered Q_i. RMLF maintains the invariant that it is always running the earliest released job in the lowest nonempty queue.

When a job J_h is released at time r_h, RMLF takes the following actions:

- Job J_h is enqueued on Q_0.
- The target $U_{0,h}$ is set to $\max\{1, 2-\beta_h\}$.
- If, just prior to r_h, it was the case that Q_0 was empty, and that RMLF was running a job J_j, RMLF then takes the following actions:
    —Job J_j is preempted. Note that J_j remains at the front of its queue.
    —RMLF begins running J_h.

If at some time t, a job $J_j \in Q_{i-1}$ is being run when $w_j(t)$ becomes equal to $U_{i-1,j}$, then RMLF takes the following actions:

- Job J_j is dequeued from Q_{i-1}.
- Job J_j is enqueued on Q_i.
- The target $U_{i,j}$ is set to $2U_{i-1,j} = 2^i \max\{1, 2-\beta_j\}$.

Whenever a job is completed, it is removed from its queue.

$U_{i,j}$ units of service. The performance of the algorithm may suffer from a poor choice of the so-called targets $U_{i,j}$; in particular, if the job sizes are slightly above their targets, then jobs are moved to lower priority queues just prior to completion. The improvement of RMLF over MLF is due to randomization of the targets, thereby reducing the possibility of such events over general instances.

We now provide a mathematical representation of the RMLF algorithm. Assume first that there is a universal lower bound on the job sizes in every instance $\mathcal{I}$; say, with value 2. For every instance of size m, the j-th job J_j is released at time r_j and has size B_j. The process $w_j(t)$ denotes the amount of time that RMLF has run J_j before time t. For some symbolic constant θ, fixed at $\theta := 4/3$, we define the independent exponentially distributed variables β_j with $\mathbb{P}(\beta_j \le x) = 1 - \exp[-\theta x \ln j]$. Finally, the targets are defined as $U_{i,j} = 2^i \max\{1, 2-\beta_j\}$ for all $i \in \{1, 2, \dots\}$, $j \in \{1, \dots, m\}$. RMLF is then formalized in Figure 1, similar to Kalyanasundaram and Pruhs [13], Becchetti and Leonardi [6].

Kalyanasundaram and Pruhs [13] proved that the RMLF algorithm is $O(\log(m)\log(\log(m)))$-competitive. This result was later strengthened by Becchetti and Leonardi [6] to a competitive ratio of $O(\log(m))$.

Theorem 1 (Becchetti and Leonardi [6]). *The* RMLF *algorithm is* $\log(m)$*-competitive. That is,*

$$\mathbb{E}[\bar{T}_{\mathrm{RMLF}}(\mathcal{I})] \le C_1 \log(m) \bar{T}_{\mathrm{SRPT}}(\mathcal{I}) \tag{4}$$

for all instances $\mathcal{I}$ of size at most m and a universal constant C_1.

The competitive ratio lower bound of $\Omega(\log(m))$ as shown by Motwani et al. [18] implies that, up to multiplicative factors, this is the best bound possible for randomized algorithms in the current model. Note that this competitive ratio is significantly lower than the best-possible ratio for blind deterministic algorithms: $\Omega(m^{1/3})$.

In the next section, we propose a variant on RMLF that makes the assumption of a universal lower bound on job sizes obsolete.

3.2. Extending the RMLF Algorithm

In a general GI/GI/1 queue, there may not be a strictly positive lower bound on the job sizes. The RMLF algorithm is not directly applicable in that case. This problem is solved in an extension of the RMLF algorithm, which we will refer to as the eRMLF algorithm. The eRMLF algorithm defines queues $\tilde{Q}_1, \tilde{Q}_2, \dots$ that are identical to the queues $Q_1, Q_2, \dots$ of the RMLF algorithm, but splits the first queue Q_0 into many queues $\tilde{Q}_0, \tilde{Q}_{-1}, \dots$ Additionally, it considers a "new job" queue $\tilde{Q}^*$. The concept of the eRMLF algorithm is described below; the formal statement is presented in the appendix.

Let a problem instance $\tilde{\mathcal{I}}$ for eRMLF be given. A target $\tilde{U}_{*,j} = 2^{z_j^*} \max\{1, 2-\tilde{\beta}_j\}$ is assigned to every job $\tilde{J}_j$ upon arrival, where $\tilde{\beta}_j$ is an exponentially distributed random variable and $z_j^* \in \mathbb{Z}$ depends on the current state of the system. When the target has been assigned to the new job, it receives service in $\tilde{Q}^*$ until either the job is completed, the obtained service equals the target, or a new job arrives. Once either of the latter two events happens, the job in $\tilde{Q}^*$ is assigned to a queue $\tilde{Q}_z$, $z \in \mathbb{Z}$.

If there are no jobs in queue $\tilde{Q}^*$, the eRMLF algorithm serves the queues $\tilde{Q}_z$ in a similar fashion as the RMLF algorithm. Moreover, at any time, the problem instance $\tilde{\mathcal{I}}$ can be converted to a problem instance $\mathcal{I}$ for RMLF by a scaling argument, and under this scaling, the sojourn times of all jobs are identical for both algorithms. From this perspective, it is only natural that eRMLF inherits the competitive ratio of RMLF.

Theorem 2. *The* eRMLF *algorithm is* $\log(m)$*competitive. That is,*

$$\mathbb{E}[\bar{T}_{\mathrm{eRMLF}}(\mathcal{I})] \le C_1 \log(m) \bar{T}_{\mathrm{SRPT}}(\mathcal{I}) \tag{5}$$

for all instances $\mathcal{I}$ *of size at most* m *for a universal constant* C_1*. This constant is identical to the constant* C_1 *in Theorem* 1.

The proof of Theorem 2 is given in the appendix.

4. Main Result

We are now ready to present the main result, Theorem 3. The main result states that the average sojourn time under SRPT is at most a factor $\log(1/(1-\rho^{(n)}))$ better than that under eRMLF in heavy traffic.

Theorem 3. *For a* GI/GI/1 *queue, the* eRMLF *algorithm satisfies the relation*

$$\mathbb{E}[\bar{T}^{(n)}_{\mathrm{eRMLF}}] = O\left(\log\left(\frac{1}{1-\rho^{(n)}}\right)\right) \cdot \bar{T}^{(n)}_{\mathrm{SRPT}} \tag{6}$$

as $n \to \infty$*, provided that* $\sup_{n\in\{1,2,\ldots\}} \mathbb{E}[(B^{(n)})^{\alpha}] < \infty$ *for some* $\alpha > 2$.

The proof of the theorem is postponed until the next section. It relies on techniques from competitive analysis and applied probability.

As a consequence of Theorem 3, the blind policy $\mathscr{A}^{(n)}$ that minimizes the average sojourn time for the n-th system also satisfies the above performance bound. We emphasise the fact that the implementation of eRMLF does not depend on the interarrival and job-size distributions, whereas this may not be true for the optimal blind policy $\mathscr{A}^{(n)}$. This property may pose a considerable advantage over a system-dependent optimal blind policy with similar mean performance, for example, when the input distributions are only approximately known. Also, we note that Theorem 3 remains true if eRMLF is replaced by RMLF, provided that the support of the job-size distribution $F_B^{(n)}$ is uniformly bounded away from zero (i.e., $B_i^{(n)} \ge B_{\min}$ for some $B_{\min} > 0$ independent of i and n).

We conclude this section with some remarks.

Remark 1. Recall that the average sojourn time under any blind policy in an M/M/1 queue is $\mathbb{E}[B^{(n)}]/(1-\rho^{(n)})$, whereas the average sojourn time under SRPT (Bansal and Harchol-Balter [5]) is

$$\bar{T}^{(n)}_{\mathrm{SRPT}} = (1+o(1)) \frac{1}{\log(e/(1-\rho^{(n)}))} \frac{\mathbb{E}[B^{(n)}]}{1-\rho^{(n)}}. \tag{7}$$

In this case, our result is tight up to a multiplicative factor.

Remark 2. There may be sequences of GI/GI/1 queues for which $\mathbb{E}[\bar{T}^{(n)}_{\mathrm{eRMLF}}]$ has a worse heavy-traffic scaling than $\mathbb{E}[\bar{T}^{(n)}_{\mathscr{A}^{(n)}}]$. For example, it is known that the FB policy minimizes the average sojourn time over all blind policies in a M/GI/1 queue if $F_B^{(n)}$ has a decreasing failure rate (Righter et al. [25]). Moreover, if $F_B^{(n)}(x) = 1 - x^{-\beta}, x \ge 1, \beta \in (1,\infty)/\{2\}$, then $\bar{T}^{(n)}_{\mathrm{FB}} = \Theta(\bar{T}^{(n)}_{\mathrm{SRPT}})$ displays the best-possible scaling in heavy traffic (Nuyens and Wierman [20], Lin et al. [15]). The heavy-traffic behaviour of $\mathbb{E}[\bar{T}^{(n)}_{\mathrm{eRMLF}}]$ is unknown for any GI/GI/1 queue and could scale worse than $\bar{T}^{(n)}_{\mathrm{FB}}$ (although no worse than $\log(1/(1-\rho)) \cdot \bar{T}^{(n)}_{\mathrm{FB}}$ by Theorem 3).

On the other hand, the optimal blind policy $\mathscr{A}^{(n)}$ may not be robust under different input distributions $F_A^{(n)}$ and $F_B^{(n)}$. Continuing the FB example, we see that it is optimal if $F_B^{(n)}$ is the Pareto distribution, yet $\bar{T}^{(n)}_{\mathrm{FB}} = \Theta((1-\rho)^{-2}) = \Theta((1-\rho)^{-1}) \cdot \bar{T}^{(n)}_{\mathrm{SRPT}}$ if $F_B^{(n)} = \mathbb{1}(x \le 1)$ is deterministic (Nuyens and Wierman [20], Lin et al. [15]).

5. Proof of the Main Theorem

The current section presents the proof of Theorem 3.

5.1. Proof Strategy

The competitive ratio of the eRMLF algorithm provides an upper bound on the suboptimality of eRMLF. Specifically, it guarantees an upper bound of $O(\log(m))$ on the ratio of the average sojourn time under eRMLF over the average sojourn time under SRPT, for instances of length at most m. Unfortunately, a general GI/GI/1 queue corresponds to an infinite-length problem instance, and hence the competitive ratio result cannot be applied directly.

The key idea of the proof is that a GI/GI/1 queue is a regenerative process, and as such, one would like to analyse individual busy periods rather than the infinite problem instance. This approach is justified by the fact that for a single server, any two work-conserving scheduling policies π_1 and π_2 generate the same busy periods, i.e., $V_{\pi_1}^{(n)}(t) \equiv V_{\pi_2}^{(n)}(t)$. This means that the server is simultaneously active under both policies, and hence, in particular, that every busy period under π_1 can be compared to the same busy period under π_2.

Still, regarding every busy period as an individual problem instance does not bound the problem instance length. One way to circumvent the unbounded problem instances is by discriminating between "small" busy periods with at most $N_0^{(n)}$ jobs, and "large" busy periods. Busy periods with at most $N_0^{(n)}$ jobs can be analysed with the competitive ratio, yielding a bound of $O(\log(N_0^{(n)}))$. This leaves us with the analysis of large busy periods.

Since the GI/GI/1 queue induces a distribution over problem instances, the probability of experiencing busy periods with at least $N_0^{(n)}$ jobs can be made arbitrarily small by choosing the threshold $N_0^{(n)}$ properly. The combined sojourn time of all the jobs in such a large busy period is dominated by the product of the number of jobs $N^{(n)}$ in the busy period and the duration $P^{(n)}$ of the busy period. Therefore, the contribution of large busy periods to the overall average sojourn time is at most $\mathbb{E}[N^{(n)}P^{(n)}\mathbb{1}(N^{(n)} > N_0^{(n)})]/\mathbb{E}[N^{(n)}]$. We will show that, for an appropriate choice of $N_0^{(n)}$, the contribution of the large busy periods is $o(\log(N_0^{(n)}))$.

The second part of this section formalizes the above strategy. In the analysis of the expectation $\mathbb{E}[N^{(n)}P^{(n)}\mathbb{1}(N^{(n)} > N_0^{(n)})]$, we greatly rely on Hölder's inequality for decoupling the given expectation into individual moments of $P^{(n)}$ and $N^{(n)}$. The behaviour of these moments is then the subject of Propositions 1 and 2, which are proven in Section 6.

5.2. Small and Large Busy Periods

We begin by specifying the threshold that distinguishes small and large busy periods based on the number of jobs. Fix $s \in (\alpha/(\alpha-1), 2)$, and $\zeta > (4+2s)/(2-s)$. The threshold $N_0^{(n)}$ is now defined as $N_0^{(n)} := (1-\rho^{(n)})^{-\zeta}$.

Let $T_{\mathrm{eRMLF},i}^{(n)}, T_{\mathrm{SRPT},i}^{(n)}, i \in \{1, \ldots, N^{(n)}\}$ be the sojourn time of job i under algorithm eRMLF and SRPT, respectively. Using the fact that a GI/GI/1 queue is a regenerative process, we only need to consider a general busy period when analysing the average sojourn time (Asmussen [2, theorem VI.1.2, proposition X.1.3]):

$$\mathbb{E}[\bar{T}_{\mathrm{eRMLF}}^{(n)}] = \frac{1}{\mathbb{E}[N^{(n)}]}\mathbb{E}\left[\sum_{i=1}^{N^{(n)}} T_{\mathrm{eRMLF},i}^{(n)}\right]. \tag{8}$$

Discriminating between small and large busy periods then yields

$$\mathbb{E}[\bar{T}_{\mathrm{eRMLF}}^{(n)}] = \frac{1}{\mathbb{E}[N^{(n)}]}\mathbb{E}\left[\sum_{i=1}^{N^{(n)}} T_{\mathrm{eRMLF},i}^{(n)}\mathbb{1}(N^{(n)} \le N_0^{(n)})\right] + \frac{1}{\mathbb{E}[N^{(n)}]}\mathbb{E}\left[\sum_{i=1}^{N^{(n)}} T_{\mathrm{eRMLF},i}^{(n)}\mathbb{1}(N^{(n)} > N_0^{(n)})\right]. \tag{9}$$

As described in the strategy, we will bound the first term by means of the competitive ratio of eRMLF and show that the second term vanishes asymptotically as $n \to \infty$. These analyses are the subjects of the following two subsections.

5.3. Small Busy Periods: Competitive Ratio

The first term in (9) considers busy periods with at most $N_0^{(n)}$ jobs. Theorem 2 ensures that, for any problem instance $\mathcal{I}$ with $N^{(n)} \le N_0^{(n)}$ jobs, the average sojourn time $\mathbb{E}[\bar{T}_{\mathrm{eRMLF}}(\mathcal{I})]$ is bounded by $C_1\log(N_0^{(n)})\bar{T}_{\mathrm{SRPT}}(\mathcal{I})$. In particular,

$$\begin{aligned}\frac{1}{\mathbb{E}[N^{(n)}]}\mathbb{E}\left[\sum_{i=1}^{N^{(n)}} T_{\mathrm{eRMLF},i}^{(n)}\mathbb{1}(N^{(n)} \le N_0^{(n)})\right] &\le \frac{C_1}{\mathbb{E}[N^{(n)}]}\log(N_0^{(n)})\mathbb{E}\left[\sum_{i=1}^{N^{(n)}} T_{\mathrm{SRPT},i}^{(n)}\mathbb{1}(N^{(n)} \le N_0^{(n)})\right]\\ &\le \frac{C_1}{\mathbb{E}[N^{(n)}]}\log(N_0^{(n)})\mathbb{E}\left[\sum_{i=1}^{N^{(n)}} T_{\mathrm{SRPT},i}^{(n)}\right]\\ &= C_1\log(N_0^{(n)})\cdot\bar{T}_{\mathrm{SRPT}}^{(n)}.\end{aligned}$$

The proof is complete once we show that the second term in (9) is dominated by $\log(N_0^{(n)})\bar{T}_{\mathrm{SRPT}}^{(n)}$ as $n \to \infty$.

5.4. Large Busy Periods: Hölders Inequality

For any work-conserving scheduling policy, the sojourn time of an individual job is bounded by the duration $P^{(n)}$ of the busy period. Therefore the second term in (9) is bounded by

$$\frac{1}{\mathbb{E}[N^{(n)}]}\mathbb{E}\left[\sum_{i=1}^{N^{(n)}} T^{(n)}_{\mathrm{eRMLF},\,i}\mathbb{1}(N^{(n)} > N_0^{(n)})\right] \le \frac{1}{\mathbb{E}[N^{(n)}]}\mathbb{E}[N^{(n)}P^{(n)}\mathbb{1}(N^{(n)} > N_0^{(n)})]. \tag{10}$$

The functionals $N^{(n)}$ and $P^{(n)}$ are dependent, which makes an exact analysis of the expectation troublesome. This complication is avoided by applying Hölder's inequality, which allows us to approximate the dependent expectation by the product of two expectations. In particular, for $\tilde{s} = s/(s-1) \in (2, \alpha)$, we have $1/\tilde{s} + 1/s = 1$, and hence

$$\mathbb{E}\left[\sum_{i=1}^{N^{(n)}} T^{(n)}_{\mathrm{eRMLF},\,i}\mathbb{1}(N^{(n)} > N_0^{(n)})\right] \le \mathbb{E}[(P^{(n)})^{s/(s-1)}]^{(s-1)/s}\mathbb{E}[(N^{(n)})^s\mathbb{1}(N^{(n)} > N_0^{(n)})]^{1/s}. \tag{11}$$

Applying Hölder's inequality once more with parameters $2/s$ and $2/(2-s)$, we get

$$\mathbb{E}\left[\sum_{i=1}^{N^{(n)}} T^{(n)}_{\mathrm{eRMLF},\,i}\mathbb{1}(N^{(n)} > N_0^{(n)})\right] \le \mathbb{E}[(P^{(n)})^{s/(s-1)}]^{(s-1)/s}\mathbb{E}[(N^{(n)})^2]^{1/2}\mathbb{P}(N^{(n)} > N_0^{(n)})^{(2-s)/(2s)}. \tag{12}$$

Finally, the tail probability of $N^{(n)}$ is bounded by Markov's inequality. We therefore obtain the following upper bound for the second term in (9):

$$\frac{1}{\mathbb{E}[N^{(n)}]}\mathbb{E}\left[\sum_{i=1}^{N^{(n)}} T^{(n)}_{\mathrm{eRMLF},\,i}\mathbb{1}(N^{(n)} > N_0^{(n)})\right] \le \mathbb{E}[(P^{(n)})^{s/(s-1)}]^{(s-1)/s}\mathbb{E}[(N^{(n)})^2]^{1/2}\frac{\mathbb{E}[N^{(n)}]^{(2-s)/(2s)-1}}{(N_0^{(n)})^{(2-s)/(2s)}}. \tag{13}$$

The analysis of the average sojourn time for large busy periods is now reduced to the analysis of moments of $N^{(n)}$ and $P^{(n)}$. The following two propositions quantify the behaviour of these moments.

Proposition 1. *Assume* $\sup_{n\in\{1,2,\dots\}}\mathbb{E}[(B^{(n)})^\alpha] < \infty$ *for some* $\alpha \ge 2$. *Then,*

$$\mathbb{E}[(P^{(n)})^\kappa] = O((1-\rho^{(n)})^{1-2\kappa}) \tag{14}$$

for all $\kappa \in [1, \alpha]$. *Moreover,* $\mathbb{E}[P^{(n)}] = \Theta((1-\rho^{(n)})^{-1})$.

Proposition 2. *Assume* $\sup_{n\in\{1,2,\dots\}}\mathbb{E}[(B^{(n)})^\alpha] < \infty$ *for some* $\alpha \ge 2$. *Then,*

$$\mathbb{E}[(N^{(n)})^\kappa] = O((1-\rho^{(n)})^{1-2\kappa}) \tag{15}$$

for all $\kappa \in [1, \alpha]$. *Moreover,* $\mathbb{E}[N^{(n)}] = \Theta((1-\rho^{(n)})^{-1})$.

Both propositions are proven in Section 6.

Remark 3. When applied to the Example Model, Proposition 1 states that $\mathbb{E}[(P^{(r)})^\kappa]$ is uniformly bounded from above by $C_2(1-r)^{1-2\kappa}$ for some constant C_2. Alternatively, the integer moments of the busy period duration in an M/GI/1 queue can be calculated explicitly from its Laplace-Stieltjes transform, yielding $\mathbb{E}[P^{(r)}] = \mathbb{E}[B]/(1-r)$ and $\mathbb{E}[(P^{(r)})^2] = \mathbb{E}[B^2]/(1-r)^3$. One may therefore conclude that the asymptotic behaviour of the bound in Proposition 1 is, in fact, sharp for the first two moments of the busy period duration $P^{(r)}$ in the Example Model.

From Propositions 1 and 2, it follows that, for some constant C_3,

$$\begin{aligned}\frac{1}{\mathbb{E}[N^{(n)}]}\mathbb{E}[P^{(n)}N^{(n)}\mathbb{1}(N^{(n)} > N_0^{(n)})] &\le C_3(1-\rho^{(n)})^{(1-2(s/(s-1)))((s-1)/s)}(1-\rho^{(n)})^{-3/2}(1-\rho^{(n)})^{1-(2-s)/(2s)}(1-\rho^{(n)})^{((2-s)/(2s))\zeta}\\ &= C_3(1-\rho^{(n)})^{-1-1/s}(1-\rho^{(n)})^{-3/2}(1-\rho^{(n)})^{3/2-1/s}(1-\rho^{(n)})^{((2-s)/(2s))\zeta}\\ &= C_3(1-\rho^{(n)})^{((2-s)/(2s))\zeta-(2+s)/s}.\end{aligned} \tag{16}$$

Now, by choice of ζ, this expression tends to zero as $n \to \infty$. In particular, the contribution of large busy periods to the overall average sojourn time is negligible to the contribution of small busy periods. This completes the proof of Theorem 3.

6. Moments of Busy Period Functionals

This section proves several results on the moments of functionals. First, we introduce some new notation in Section 6.1. Then, we state and prove two lemmas in Section 6.2 to prove Propositions 1 and 2. Sequentially, the propositions are proven in Sections 6.3 and 6.4. We emphasise that all of the functionals considered are independent of the scheduling policy, provided that it is work-conserving.

6.1. Counting and Netput Processes

For any nonnegative random variable Y, we define a random variable Y_e that is distributed as the excess of Y; i.e., $\mathbb{P}(Y_e \leq x) = \int_0^x \mathbb{P}(Y > x)\,dx/\mathbb{E}[Y]$. Next, we define two counting processes in the GI/GI/1 queue under consideration. The first process $N^{(n)}(t) := \inf\{m \in \{1,2,\ldots\}\colon A_1^{(n)} + \cdots + A_m^{(n)} \geq t\}, t \geq 0$ counts the number of arrivals in t time units, starting from a reference arrival that is also the first count. The second process $\tilde{N}^{(n)}(t)$, $t \geq 0$ is similar and only differs by initializing the count at an arbitrary point in time. Specifically,

$$\tilde{N}^{(n)}(t) := \begin{cases} 0 & \text{if } t < A_e^{(n)}, \\ \inf\{m \in \{1,2,\ldots\}\colon A_e^{(n)} + A_1^{(n)} + \cdots + A_m^{(n)} \geq t\} & \text{otherwise.} \end{cases}$$

These counting processes allow us to introduce two netput processes $X^{(n)}(t) := \sum_{i=1}^{N^{(n)}(t)} B_i^{(n)} - t$ and $\tilde{X}^{(n)}(t) := \sum_{i=1}^{\tilde{N}^{(n)}(t)} B_i^{(n)} - t$ that quantify the net amount of work that could have been processed by the server in the t time units after an arrival, or respectively, after an arbitrary point in time. Note that $X(t)$ becomes negative right after the first time that the queue is emptied. One may verify that $\mathbb{P}(\tilde{X}^{(n)}(t) > x) \leq \mathbb{P}(X^{(n)}(t) > x)$ for all $t \geq 0$, which will be denoted by $\tilde{X}^{(n)}(t) \leq_{st} X^{(n)}(t)$ in the remainder of this paper.

Similarly, we define two discrete processes that quantify the netput process *at an arrival instance*, denoted by $S_m^{(n)}$ and $\tilde{S}_m^{(n)}$, $m \geq 0$. The process $S_m^{(n)}$ is defined as $S_0^{(n)} := 0$, $S_m^{(n)} := \sum_{i=1}^m [B_i^{(n)} - A_i^{(n)}]$ and quantifies all the work that the server has received between the arrival of the reference job and the m-th next arrival, minus the work that it could have addressed during this time. The process $\tilde{S}_m^{(n)}$ starts observing at an arbitrary point in time instead of at the arrival of a reference job. It is defined as $\tilde{S}_0^{(n)} := -A_e^{(n)}$, $\tilde{S}_m^{(n)} := -A_e^{(n)} + \sum_{i=1}^m [B_i^{(n)} - A_i^{(n)}]$. Again, we obtain the relation $\tilde{S}_m^{(n)} \leq_{st} S_m^{(n)}$. Also, one may verify that $\sup_{t\geq 0} X^{(n)}(t) = \sup_{m\in\{1,2,\ldots\}} S_m^{(n)}$, and hence by Asmussen [2, corollary III.6.5], we have $\sup_{t\geq 0} X^{(n)}(t) = \sup_{m\in\{0,1,\ldots\}} S_m^{(n)} \overset{d}{=} W^{(n)}$. Here, $\cdot \overset{d}{=} \cdot$ denotes equality in distribution. All sums $\sum_{i=1}^0$ are understood to be zero.

6.2. Preliminary Lemmas

The following two lemmas facilitate the proof of Propositions 1 and 2. Lemma 1 concerns the first moment of $N^{(n)}$ and $I^{(n)}$, whereas Lemma 2 considers general moments of $W^{(n)}$.

Lemma 1. *The relations*

$$(1 - \rho^{(n)})\mathbb{E}[N^{(n)}] = \Theta(1), \qquad \text{and} \qquad \mathbb{E}[I^{(n)}] = \Theta(1) \tag{17}$$

hold as $n \longrightarrow \infty$.

Proof of Lemma 1. Since we have $\mu^{(n)} := \mathbb{E}[A^{(n)}](1 - \rho^{(n)}) = \Theta(1 - \rho^{(n)})$, it suffices to prove the relation $\mu^{(n)}\mathbb{E}[N^{(n)}] = \Theta(1)$. Proposition X.3.1 in Asmussen [2], stating

$$\mathbb{E}[I^{(n)}] = \mu^{(n)}\mathbb{E}[N^{(n)}], \tag{18}$$

then implies that this is equivalent to the relation $\mathbb{E}[I^{(n)}] = \Theta(1)$.

The upper and lower bound follow from Lotov [16], who considers the ladder height of a random walk. Specifically, Lotov obtains upper bounds for the moments of the ladder epochs and the moments of overshoot over an arbitrary nonnegative level if the expectation of jumps is positive and close to zero. As such, his results apply to the random walk $-S(n)$ with ladder epochs $N^{(n)}$.

The upper bound is implied by theorem 2 in Lotov [16], which claims that

$$\mu^{(n)}\mathbb{E}[N^{(n)}] \leq C_4 \tag{19}$$

for some constant C_4 and all n, provided that $\sup_{n\in\{1,2,\ldots\}} \mathbb{E}[(\max\{A^{(n)} - B^{(n)}, 0\})^2] < \infty$. Accordance with this condition follows directly from $\sup_{n\in\{1,2,\ldots\}} \mathbb{E}[(A^{(n)})^2] < \infty$.

The lower bound is implied by inequality (2) in Lotov [16]. In our model, we assumed that there exist constants $\delta > 0$, $\gamma > 0$ such that $\mathbb{P}(B^{(n)} - A^{(n)} \geq \delta) \geq \gamma$ for all n. Lotov [16] then states

$$\mu^{(n)}\mathbb{E}[N^{(n)}] \geq \int_0^\infty x \,\mathrm{d}\mathbb{P}(B^{(n)} - A^{(n)} \leq x) \geq \delta\gamma$$

for all n. This completes the proof. □

Lemma 2. *Let $p > 0$ and define $q = \max\{2, p+1\}$. Assume that $\sup_{n\in\{1,2,\ldots\}} \mathbb{E}[(B^{(n)})^q] < \infty$. Then,*

$$\limsup_{n\to\infty} (1-\rho^{(n)})^p \mathbb{E}[(W^{(n)})^p] < \infty. \tag{20}$$

Remark 4. Consider the Example Model, and assume that jobs are served according to the FIFO discipline. Then, $W^{(r)}$ is just the waiting time of a job, and hence for some constant C_5, the average sojourn time $\bar{T}^{(r)}_{\text{FIFO}} = \mathbb{E}[W^{(r)}_{\text{FIFO}}] + \mathbb{E}[B] \leq C_5/(1-r) + \mathbb{E}[B]$ scales no worse than $1/(1-r)$. Lemma 2 provides bounds on the work $W^{(r)}$ at an arrival for more general moments, provided that a sufficiently high moment of the job size distribution exists.

Proof of Lemma 2. Since $\sup_{n\in\{1,2,\ldots\}} \mathbb{E}[A^{(n)}] < \infty$, relation (20) is equivalent to

$$\limsup_{n\to\infty} (\mu^{(n)})^p \mathbb{E}[(W^{(n)})^p] < \infty, \tag{21}$$

which is proven below.

Assume $p \geq 1$ and let $E_i^{(n)}$, $i \in \{1, 2, \ldots\}$ be independent exponentially distributed random variables with mean

$$\mathbb{E}[E^{(n)}] = \frac{\mathbb{E}[A^{(n)}] + \mathbb{E}[B^{(n)}]}{2} < \mathbb{E}[A^{(n)}].$$

We define $E^{(n)} := E_1^{(n)}$ and note that $\sup_{n\in\{1,2,\ldots\}} \mathbb{E}[E^{(n)}] \leq \sup_{n\in\{1,2,\ldots\}} \mathbb{E}[A^{(n)}] < \infty$.

By Asmussen [2, corollary III.6.5] and subadditivity of suprema, $W^{(n)}$ is upper bounded as

$$W^{(n)} \overset{d}{=} \sup_{m\in\{0,1,2,\ldots\}} \sum_{i=1}^m [B_i^{(n)} - A_i^{(n)}] \leq \sup_{m\in\{0,1,2,\ldots\}} \sum_{i=1}^m [B_i^{(n)} - E_i^{(n)}] + \sup_{m\in\{0,1,2,\ldots\}} \sum_{i=1}^m [E_i^{(n)} - A_i^{(n)}] =: W_1^{(n)} + W_2^{(n)},$$

where $W_1^{(n)}$ can be interpreted as the total work in an M/GI/1 queue as observed by an arrival, and $W_2^{(n)}$ as the total work in an GI/M/1 queue as observed by an arrival. As a consequence, $\mathbb{P}(W^{(n)} > x) \leq \mathbb{P}(W_1^{(n)} + W_2^{(n)} > x) \leq \mathbb{P}(W_1^{(n)} > x/2) + \mathbb{P}(W_2^{(n)} > x/2)$ and thus

$$\begin{aligned}\mathbb{E}[(W^{(n)})^p] &= p\int_0^\infty x^{p-1}\mathbb{P}(W^{(n)} > x)\,\mathrm{d}x \\ &\leq p\int_0^\infty x^{p-1}\mathbb{P}(W_1^{(n)} > x/2)\,\mathrm{d}x + p\int_0^\infty x^{p-1}\mathbb{P}(W_2^{(n)} > x/2)\,\mathrm{d}x \\ &= 2^p(\mathbb{E}[(W_1^{(n)})^p] + \mathbb{E}[(W_2^{(n)})^p]).\end{aligned}$$

First, we consider $W_1^{(n)}$. Define the geometrically distributed random variable $K_1^{(n)}$ with support $\{0, 1, \ldots\}$ and fail parameter

$$\xi_1^{(n)} := \frac{\mathbb{E}[B^{(n)}]}{\mathbb{E}[E^{(n)}]} = \frac{2\mathbb{E}[B^{(n)}]}{\mathbb{E}[A^{(n)}] + \mathbb{E}[B^{(n)}]}.$$

For notational convenience, we drop the superscript (n) of $\xi_1^{(n)}$ for the remainder of this section. Theorem VIII.5.7 in Asmussen [2] presents a random sum representation of the functional $W_1^{(n)}$ in terms of $K_1^{(n)}$ and $B_{e,i}^{(n)}$:

$$W_1^{(n)} \overset{d}{=} \sum_{i=1}^{K_1^{(n)}} B_{e,i}^{(n)}.$$

Since $f(x) = x^p$ is a convex function for all $p \geq 1$, lemma 5 in Remerova et al. [24] implies

$$\mathbb{E}[(W_1^{(n)})^p] \leq \mathbb{E}[(K_1^{(n)})^p]\mathbb{E}[(B_e^{(n)})^p]. \tag{22}$$

The conditions of Lemma 2 ensure that the p-th moment of $B_e^{(n)}$ is finite as $n \to \infty$:

$$\begin{aligned}\mathbb{E}[(B_e^{(n)})^p] &= \int_0^\infty x^p \,\mathrm{d}\mathbb{P}(B_e^{(n)} \le x) = \frac{1}{\mathbb{E}[B^{(n)}]} \int_0^\infty x^p \mathbb{P}(B^{(n)} > x)\,\mathrm{d}x \\ &= \frac{1}{(p+1)\mathbb{E}[B^{(n)}]} \int_0^\infty x^{p+1} \,\mathrm{d}\mathbb{P}(B^{(n)} \le x) = \frac{1}{(p+1)\mathbb{E}[B^{(n)}]} \mathbb{E}[(B^{(n)})^{p+1}]. \end{aligned} \tag{23}$$

Therefore we need to show that $(\mu^{(n)})^p \mathbb{E}[(K_1^{(n)})^p]$ is uniformly bounded as $n \to \infty$. Let $k = \lfloor p \rfloor$. Then,

$$\begin{aligned}\mathbb{E}[(K_1^{(n)})^p] &= \frac{1-\xi_1}{(1-\xi_1)^p} \sum_{m=0}^{\infty} ((1-\xi_1)m)^p \xi_1^m \\ &\le \frac{1-\xi_1}{(1-\xi_1)^p} \sum_{m=0}^{\lfloor 1/(1-\xi_1) \rfloor} ((1-\xi_1)m)^k \xi_1^m + \frac{1-\xi_1}{(1-\xi_1)^p} \sum_{m=\lfloor 1/(1-\xi_1) \rfloor + 1}^{\infty} ((1-\xi_1)m)^{k+1} \xi_1^m \\ &\le \frac{(1-\xi_1)^{k+1}}{(1-\xi_1)^p} \sum_{m=0}^{\infty} m^k \xi_1^m + \frac{(1-\xi_1)^{k+2}}{(1-\xi_1)^p} \sum_{m=0}^{\infty} m^{k+1} \xi_1^m. \end{aligned} \tag{24}$$

On the one hand, for any $l \in \{1, 2, \ldots\}$, we have

$$(1-\xi_1)^{l+1} \xi_1^l \frac{\mathrm{d}^l}{\mathrm{d}\xi_1^l} \sum_{m=0}^{\infty} \xi_1^m = (1-\xi_1)^{l+1} \xi_1^l \frac{\mathrm{d}^l}{\mathrm{d}\xi_1^l} (1-\xi_1)^{-1} = l!\xi_1^l. \tag{25}$$

On the other hand we have

$$\begin{aligned}(1-\xi_1)^{l+1} \xi_1^l \frac{\mathrm{d}^l}{\mathrm{d}\xi_1^l} \sum_{m=0}^{\infty} \xi_1^m &= (1-\xi_1)^{l+1} \sum_{m=l}^{\infty} m(m-1)\cdots(m-l+1)\xi_1^m \\ &= (1-\xi_1)^{l+1} \sum_{m=0}^{\infty} m^l \xi_1^m - (1-\xi_1)^{l+1} \sum_{m=0}^{l-1} m^l \xi_1^m + (1-\xi_1)^{l+1} \sum_{m=l}^{\infty} o(m^l) \xi_1^m. \end{aligned} \tag{26}$$

Combining equalities (25) and (26), we find that

$$(1-\xi_1)^{l+1} \sum_{m=0}^{\infty} m^l \xi_1^m = l!\xi_1^l + (1-\xi_1)^{l+1} \sum_{m=0}^{l-1} m^l \xi_1^m + (1-\xi_1)^{l+1} \sum_{m=l}^{\infty} o(m^l) \xi_1^m.$$

Now, for any $\nu > 0$, there exists a $M_\nu \in \{1, 2, \ldots\}$ independent of the system index n such that for all $m \ge M_\nu$, the $o(m^l)$ term is dominated by νm^l. Fix such $\nu \in (0,1)$ and M_ν. Then,

$$(1-\xi_1)^{l+1} \sum_{m=0}^{\infty} m^l \xi_1^m \le l! + l^{l+1} + (1-\xi_1)^{l+1} \nu \sum_{m=M_\nu}^{\infty} m^l \xi_1^m + (1-\xi_1)^{l+1} \sum_{m=0}^{M_\nu} o(m^l) \xi_1^m \le C_6 + (1-\xi_1)^{l+1} \nu \sum_{m=0}^{\infty} m^l \xi_1^m$$

for some constant $C_6 > 0$, and hence

$$(1-\xi_1)^{l+1} \sum_{m=0}^{\infty} m^l \xi_1^m \le \frac{C_6}{1-\nu}. \tag{27}$$

Since $(\mu^{(n)}/(1-\xi_1))^p = (\mathbb{E}[A^{(n)}] + \mathbb{E}[B^{(n)}])^p$, we may conclude from relations (24) and (27) that $(\mu^{(n)})^p \mathbb{E}[(K_1^{(n)})^p]$ is uniformly bounded from above as $n \to \infty$, and so is $(\mu^{(n)})^p \mathbb{E}[(W_1^{(n)})^p]$ by (22).

Second, we consider the functional $W_2^{(n)}$. Recall that $W_2^{(n)}$ denotes the steady-state work load in a GI/M/1 queue upon arrival. Theorem VIII.5.8 and p. 296 in Asmussen [2] together state that

$$W_2^{(n)} \overset{d}{=} \sum_{i=1}^{K_2^{(n)}} E_i^{(n)}, \tag{28}$$

where $K_2^{(n)}$ is a geometrically distributed random variable with support $\{0, 1, \ldots\}$ and unknown fail parameter $\xi_2^{(n)}$. Remerova et al. [24] again ensure that

$$\mathbb{E}[(W_2^{(n)})^p] \le \mathbb{E}[(K_2^{(n)})^p] \mathbb{E}[(E^{(n)})^p], \tag{29}$$

where the latter expectation is finite uniformly in n as a property of exponential distributions. The p-th moment of $K_2^{(n)}$ is bounded by (24) and (27), so that

$$(\mu^{(n)})^p \mathbb{E}[(K_2^{(n)})^p] \le C_7 \left(\frac{\mu^{(n)}}{1-\xi_2^{(n)}} \right)^p \tag{30}$$

for some constant $C_7 > 0$. The proof is complete once we show $\mu^{(n)}/(1-\xi_2^{(n)}) = O(1)$.

From (28), it can be verified that $\mathbb{P}(W_2^{(n)} = 0) = \mathbb{P}(K_2^{(n)} = 0) = 1 - \xi_2^{(n)}$. Additionally, by theorem VIII.2.3 in Asmussen [2], we have $\mathbb{P}(W_2^{(n)} = 0) = 1/\mathbb{E}[N_2^{(n)}]$, and hence $\mathbb{E}[N_2^{(n)}] = 1/(1-\xi_2^{(n)})$. Here, $N_2^{(n)}$ is the steady-state number of jobs in a busy period of the GI/M/1 queue. Applying Lemma 1 with interarrival times $A_i^{(n)}$, job sizes $E_i^{(n)}$, and mean service between arrivals $\frac{1}{2}\mu^{(n)}$ yields $\frac{1}{2}\mu^{(n)}\mathbb{E}[N_2^{(n)}] = \Theta(1)$, and therefore $\mu^{(n)}/(1-\xi_2^{(n)}) = O(1)$.

Finally, for $0 < p < 1$, the lemma follows directly from the case $p = 1$ after observing that $(\mu^{(n)})^p \mathbb{E}[(W^{(n)})^p] \le (\mu^{(n)}\mathbb{E}[W^{(n)}])^p$ by Jensen's inequality. □

Lemmas 1 and 2 provide the asymptotic behaviour of functionals that are closely related to $P^{(n)}$ and $N^{(n)}$. The remainder of this section uses these results to prove Propositions 1 and 2.

6.3. Busy Period Duration $P^{(n)}$

This section is devoted to the proof of Proposition 1. We wish to show that

$$\mathbb{E}[(P^{(n)})^\kappa] = O((1-\rho^{(n)})^{1-2\kappa}) \tag{14, revisited}$$

for all $\kappa \in [1, \alpha]$, provided that $\sup_{n \in \{1,2,\ldots\}} \mathbb{E}[(B^{(n)})^\alpha] < \infty$ for some $\alpha \ge 2$. Moreover, we claim that $\mathbb{E}[P^{(n)}] = \Theta((1-\rho^{(n)})^{-1})$.

First, consider $\kappa = 1$. Due to Little's law for a busy server, we have

$$1 - \rho^{(n)} = \frac{\mathbb{E}[I^{(n)}]}{\mathbb{E}[I^{(n)}] + \mathbb{E}[P^{(n)}]}, \tag{31}$$

so that

$$\mathbb{E}[P^{(n)}] = \frac{\rho^{(n)}\mathbb{E}[I^{(n)}]}{1-\rho^{(n)}}. \tag{32}$$

The result now follows from Lemma 1.

Second, consider $\kappa > 1$. Similar to (23), one obtains $\mathbb{E}[(P_e^{(n)})^{\kappa-1}] = \mathbb{E}[(P^{(n)})^\kappa]/(\kappa\mathbb{E}[P^{(n)}])$, and hence it suffices to show that $\mathbb{E}[(P_e^{(n)})^{\kappa-1}] = O((1-\rho^{(n)})^{2(1-\kappa)})$. We have the following convenient representation for $P_e^{(n)}$ (Asmussen [2, theorem X.3.4]):

$$P_e^{(n)} \overset{d}{=} \inf\{\tau \ge 0\colon \tilde{X}^{(n)}(\tau) \le -V^{(n)} \mid V^{(n)} > 0\} \overset{d}{=} \inf\{\tau \ge 0\colon B_e^{(n)} + W^{(n)} + \tilde{X}^{(n)}(\tau) \le 0\}.$$

The above relation allows us to bound $\mathbb{P}(P_e^{(n)} > t)$ as

$$\begin{aligned}
\mathbb{P}(P_e^{(n)} > t) &= \mathbb{P}(\inf\{\tau \ge 0\colon B_e^{(n)} + W^{(n)} + \tilde{X}^{(n)}(\tau) \le 0\} > t) \\
&= \mathbb{P}(B_e^{(n)} + W^{(n)} + \tilde{X}^{(n)}(\tau) > 0,\ \forall\, \tau \le t) \\
&\le \mathbb{P}(B_e^{(n)} + W^{(n)} + \tilde{X}^{(n)}(t) + (1-\rho^{(n)})t/2 > (1-\rho^{(n)})t/2) \\
&\le \mathbb{P}(B_e^{(n)} > (1-\rho^{(n)})t/6) + \mathbb{P}(W^{(n)} > (1-\rho^{(n)})t/6) \\
&\quad + \mathbb{P}\Big(\sup_{\tau \ge 0}[\tilde{X}^{(n)}(\tau) + (1-\rho^{(n)})\tau/2] > (1-\rho^{(n)})t/6\Big).
\end{aligned}$$

In Section 6.1, we derived the relations $\tilde{X}^{(n)}(\tau) \le_{st} X^{(n)}(\tau)$, and $W^{(n)} \overset{d}{=} \sup_{\tau \ge 0} X^{(n)}(\tau)$. These relations imply

$$\mathbb{P}(P_e^{(n)} > t) \le \mathbb{P}(B_e^{(n)} > (1-\rho^{(n)})t/6) + 2\mathbb{P}\Big(\sup_{\tau \ge 0}[X^{(n)}(\tau) + (1-\rho^{(n)})\tau/2] > (1-\rho^{(n)})t/6\Big). \tag{33}$$

The last inequality suggests that

$$\mathbb{E}[(P_e^{(n)})^{\kappa-1}] = (\kappa-1)\int_0^\infty t^{\kappa-2}\mathbb{P}(P_e^{(n)} > t)\,\mathrm{d}t$$

$$\begin{aligned}\le (\kappa-1)\int_0^\infty & t^{\kappa-2}\mathbb{P}(B_e^{(n)} > (1-\rho^{(n)})t/6)\,\mathrm{d}t\\ &+2(\kappa-1)\int_0^\infty t^{\kappa-2}\mathbb{P}\Big(\sup_{\tau\ge 0}[X^{(n)}(\tau)+(1-\rho^{(n)})\tau/2] > (1-\rho^{(n)})t/6\Big)\,\mathrm{d}t.\end{aligned}$$

To deal with the first term, note that

$$\begin{aligned}(\kappa-1)\int_0^\infty t^{\kappa-2}\mathbb{P}(B_e^{(n)} > (1-\rho^{(n)})t/6)\,\mathrm{d}t &= (\kappa-1)\int_0^\infty t^{\kappa-2}\mathbb{P}(6B_e^{(n)}/(1-\rho^{(n)}) > t)\,\mathrm{d}t\\ &= \mathbb{E}\bigg[\bigg(\frac{6B_e^{(n)}}{1-\rho^{(n)}}\bigg)^{\kappa-1}\bigg] = O((1-\rho^{(n)})^{1-\kappa})\mathbb{E}[(B_e^{(n)})^{\kappa-1}] = o((1-\rho^{(n)})^{2(1-\kappa)})\end{aligned}$$

since $\mathbb{E}[(B^{(n)})^\kappa] < \infty$ implies $\mathbb{E}[(B_e^{(n)})^{\kappa-1}] < \infty$ (cf. 23).

For the second term, observe that

$$\begin{aligned}\sup_{\tau\ge 0}[X^{(n)}(\tau)+(1-\rho^{(n)})\tau/2] &= \sup_{\tau\ge 0}\Bigg[\sum_{i=1}^{N^{(n)}(\tau)} B_i^{(n)} - \tau + (1-\rho^{(n)})\tau/2\Bigg]\\ &= \sup_{\tau\ge 0}\Bigg[\sum_{i=1}^{N^{(n)}(\tau)} B_i^{(n)} - \frac{1+\rho^{(n)}}{2}\tau\Bigg] = \sup_{\eta\in\{0,1,2,\ldots\}}\Bigg[\sum_{i=1}^{\eta}\bigg\{B_i^{(n)} - \frac{1+\rho^{(n)}}{2}A_i^{(n)}\bigg\}\Bigg] =: \tilde{W}^{(n)},\end{aligned}$$

where $\tilde{W}^{(n)}$ is equal in distribution to the steady-state cumulative amount of work at arrival in a GI/GI/1 queue with job sizes $B_i^{(n)}$ and interarrival times $((1+\rho^{(n)})/2)A_i^{(n)}$, $i\in\{1,2,\ldots\}$. The mean amount of work that the server completes between two consecutive arrivals in this system is then given by $\tilde{\mu}^{(n)} := ((1+\rho^{(n)})/2)\mathbb{E}[A^{(n)}] - \mathbb{E}[B^{(n)}] = (1-\rho^{(n)})/2)\mathbb{E}[A^{(n)}]$. We therefore obtain

$$\begin{aligned}(\kappa-1)\int_0^\infty t^{\kappa-2}\mathbb{P}\Big(\sup_{\tau\ge 0}[X^{(n)}(\tau)+(1-\rho^{(n)})\tau/2] > (1-\rho^{(n)})t/6\Big)\,\mathrm{d}t &= (\kappa-1)\int_0^\infty t^{\kappa-2}\mathbb{P}\bigg(\frac{6}{1-\rho^{(n)}}\tilde{W}^{(n)} > t\bigg)\,\mathrm{d}t\\ = \frac{6^{\kappa-1}}{(1-\rho^{(n)})^{\kappa-1}(\tilde{\mu}^{(n)})^{\kappa-1}}\mathbb{E}[(\tilde{\mu}^{(n)}\tilde{W}^{(n)})^{\kappa-1}] \le \frac{C_8}{(1-\rho^{(n)})^{2(\kappa-1)}}\mathbb{E}[(\tilde{\mu}^{(n)}\tilde{W}^{(n)})^{\kappa-1}]&\end{aligned}$$

for some constant $C_8 > 0$. Finally, $\mathbb{E}[(\tilde{\mu}^{(n)}\tilde{W}^{(n)})^{\kappa-1}]$ is bounded due to Lemma 2, which completes the proof of the proposition.

6.4. Arrivals in a Busy Period $N^{(n)}$

This section contains the proof of Proposition 2. The proposition states that

$$\mathbb{E}[(N^{(n)})^\kappa] = O((1-\rho^{(n)})^{1-2\kappa}) \qquad (15, \text{revisited})$$

for all $\kappa\in[1,\alpha]$, provided that $\sup_{n\in\{1,2,\ldots\}}\mathbb{E}[(B^{(n)})^\alpha] < \infty$ for some $\alpha \ge 2$. Moreover, we claim that $\mathbb{E}[N^{(n)}] = \Theta((1-\rho^{(n)})^{-1})$.

The structure of the proof is identical to the proof of Proposition 1. For $\kappa = 1$, the result follows directly from Lemma 1. Therefore we consider $\mathbb{E}[(N^{(n)})^\kappa]$ for $\kappa > 1$. Similar to the proof of Proposition 1, we use the relation

$$\mathbb{E}[(N_e^{(n)})^{\kappa-1}] = \frac{\mathbb{E}[(N^{(n)})^\kappa]}{\kappa\mathbb{E}[N^{(n)}]} \qquad (34)$$

and note that

$$N_e^{(n)} \overset{d}{=} \inf\{\eta\in\{0,1,2,\ldots\}\colon \tilde{S}_\eta^{(n)} \le -V^{(n)} \mid V^{(n)} > 0\} \overset{d}{=} \inf\{\eta\in\{0,1,2,\ldots\}\colon B_e^{(n)} + W^{(n)} + \tilde{S}_\eta^{(n)} \le 0\}.$$

As before, the relations $\tilde{S}_\eta^{(n)} \le_{st} S_\eta^{(n)}$, and $W^{(n)} \overset{d}{=} \sup_{\eta\in\{0,1,\ldots\}} S_\eta^{(n)}$ are exploited to obtain an equivalent of (33):

$$\mathbb{P}(N_e^{(n)} > m) \le \mathbb{P}(B_e^{(n)} > (1-\rho^{(n)})m/6) + 2\mathbb{P}\Big(\sup_{\eta\in\{0,1,2,\ldots\}}[S_\eta^{(n)} + (1-\rho^{(n)})\eta/2] > (1-\rho^{(n)})m/6\Big).$$

For any $\eta\in\{0,1,2,\ldots\}$, consider $\tau_\eta^{(n)} := A_1^{(n)} + \cdots + A_\eta^{(n)}$. Then, $S_\eta^{(n)} = X^{(n)}(\tau_\eta^{(n)})$, therefore, in particular,

$$\mathbb{P}(N_e^{(n)} > m) \le \mathbb{P}(B_e^{(n)} > (1-\rho^{(n)})m/6) + 2\mathbb{P}\Big(\sup_{\tau\ge 0}[X^{(n)}(\tau) + (1-\rho^{(n)})\tau/2] > (1-\rho^{(n)})m/6\Big).$$

The remainder of the proof is identical to that of Proposition 1.

7. Conclusion

In this paper, we proved a result about the average case performance of (an extension of) the Randomized Multilevel Feedback (RMLF) algorithm in a GI/GI/1 queue. Specifically, the gap in average sojourn time between the RMLF algorithm and the Shortest Remaining Processing Time algorithm behaves like $O(\log(1/(1-\rho^{(n)})))$ and this bound is tight for the M/M/1 queue. An appealing property of the RMLF algorithm is that its implementation does not depend on the input distributions $F_A^{(n)}$ and $F_B^{(n)}$; however, if $F_A^{(n)}$ and $F_B^{(n)}$ are known then there can be blind algorithms with a better performance than RMLF (e.g., Foreground-Background if $F_B^{(n)}$ has decreasing failure rate). The result was established by using techniques from competitive analysis and applied probability. As the structure of the proof is quite general, it would be interesting to explore other possibilities in the intersection of these areas.

Acknowledgment

This work is part of the free competition research programme with project number 613-001-219, which is financed by the Netherlands Organisation for Scientific Research (NWO). The authors thank Adam Wierman for his encouragements to work on this research topic.

Appendix

The eRMLF algorithm is presented after the introduction of some notation. Define the queues $\tilde{Q}_z$, $z \in \mathbb{Z}$, and a "new job" queue $\tilde{Q}^*$. Let the targets $\tilde{U}_{z,j}$ be given by $\tilde{U}_{z,j} = 2^z \max\{1, 2-\tilde{\beta}_j\}$, where the $\tilde{\beta}_j$'s are independent random variables with exponential cumulative distribution function $\mathbb{P}(\tilde{\beta}_j \le x) = 1 - \exp[-\theta x \ln j]$. Identical to the RMLF algorithm, θ is a symbolic constant fixed at $\theta = 4/3$. All symbols $\tilde{J}_j$, $\tilde{r}_j$, $\tilde{B}_j$, and $\tilde{w}_j(t)$ are defined analogue to the symbols without accent in the RMLF algorithm. All release times $\tilde{r}_j$ must be distinct (e.g., all interarrival times are strictly positive), and jobs may have any size $\tilde{B}_j \ge 0$. Note that the original RMLF algorithm requires the job sizes to be uniformly bounded from below, but does not restrict the interarrival times to be nonzero.

Every job $\tilde{J}_h$ is assigned an initial target $\tilde{U}_{*,h}$ upon arrival, after which it is immediately served in $\tilde{Q}^*$ by a dedicated server. It departs from $\tilde{Q}^*$ on three occasions:

- The amount of service received equals the size $\tilde{B}_h$ of the job. In this case, $\tilde{J}_h$ is completed and leaves the system.
- A new job enters the system. In this case, $\tilde{J}_h$ is moved to a queue $\tilde{Q}_z$, $z \in \mathbb{Z}$, that it naturally belongs to based on the amount of service $\tilde{w}_h(t)$ it has obtained thus far; that is, it is moved to the unique queue $\tilde{Q}_{z_h^*}$ that satisfies $\tilde{U}_{z_h^*-1,h} \le \tilde{w}_h(t) < \tilde{U}_{z_h^*,h}$.
- The amount of service received equals the initial target $\tilde{U}_{*,h}$. In this case, similar to the previous case, $\tilde{J}_h$ is moved to a queue $\tilde{Q}_z$ that it naturally belongs to.

The choice of the initial target $\tilde{U}_{*,h}$ depends on the system state:

- If the system is empty upon arrival, then the server is dedicated to $\tilde{J}_h$ regardless of the queue that $\tilde{J}_h$ is in. In this case, the target can be chosen arbitrarily; we set it to $\tilde{U}_{*,h} = \tilde{U}_{0,h}$.
- If the system is not empty upon arrival, then there must be a lowest index nonempty queue $\tilde{Q}_{z_h^*}$ (possibly after moving the job originally in $\tilde{Q}^*$ to another queue). $\tilde{J}_h$ may now experience a dedicated server until the moment when it would enter queue $\tilde{Q}_{z_h^*}$ based on its obtained service and the (z_h^*-1)-th target $\tilde{U}_{z_h^*-1,h}$. Therefore $\tilde{J}_h$ should be moved no later then after $\tilde{U}_{*,h} = \tilde{U}_{z_h^*-1,h}$ units of obtained service.

If $\tilde{Q}^*$ is empty, then eRMLF always works on the earliest released job in the nonempty queue $\tilde{Q}_z$ with the lowest index $z \in \mathbb{Z}$.

The eRMLF is formally presented in Figure A.1. Observe that RMLF and eRMLF preserve the ordering of the jobs; that is, if job $\tilde{J}_j$ is released prior to job $\tilde{J}_k$, then as long as both jobs are incomplete:

- job $\tilde{J}_j$ will never be in a lower queue than job $\tilde{J}_k$, and
- if both jobs are in the same queue, then job $\tilde{J}_j$ is closer to the front of the queue than job $\tilde{J}_k$.

We are now ready to prove Theorem 2, stating that

$$\mathbb{E}[\bar{T}_{\mathrm{eRMLF}}(\mathcal{J})] \le C_1 \log(m) \cdot \bar{T}_{\mathrm{SRPT}}(\mathcal{J}), \tag{5, revisited}$$

for all instances, $\mathcal{J}$ of size at most m for a universal constant C_1. This constant is identical to the constant C_1 in Theorem 1.

Proof of Theorem 2. Consider any instance $\tilde{\mathcal{J}}$ for eRMLF of size at most m. Since all jobs of size zero are immediately served in queue $\tilde{Q}^*$ upon arrival, we assume without loss of generality that the instance does not contain any jobs of size zero. As a consequence, the minimum job size $\tilde{B}_{\min} = \min_{j=1,\dots,|\tilde{\mathcal{J}}|} \tilde{B}_j$ is strictly positive. We now transform the instance $\tilde{\mathcal{J}}$ for eRMLF to a corresponding instance $\mathcal{J}$ for RMLF.

Define the scaling parameter $g := \lfloor \log_2(\tilde{B}_{\min}) \rfloor - 1 \in \mathbb{Z}$, satisfying $2^{-g}\tilde{B}_{\min} \ge 2$. The instance $\mathcal{J}$ consists out of $|\tilde{\mathcal{J}}|$ jobs that are scaled versions of the original $|\tilde{\mathcal{J}}|$ jobs; specifically, job J_j has size $B_j := 2^{-g}\tilde{B}_j$ and release date $r_j := 2^{-g}\tilde{r}_j$. Then, the smallest job is of size at least 2 and the RMLF algorithm may be applied to the instance $\mathcal{J}$.

Since the jobs are released in the same order as in the original instance, we note that the random variables β_j assigned by RMLF have the same distribution as the $\tilde{\beta}_j$ assigned by eRMLF. We therefore couple these random variables in a trivial

Figure A.1. Formal Statement of eRMLF Algorithm

Algorithm eRMLF: At all times, the collection of released but uncompleted jobs is partitioned into queues, $\tilde{Q}^*$, $\tilde{Q}_z$, $z \in \mathbb{Z}$. We say that $\tilde{Q}_i$ is lower than $\tilde{Q}_j$ for $i < j$. $\tilde{Q}^*$ is the lowest queue. For each job $\tilde{J}_j \in \tilde{Q}_i$, $\tilde{U}_{i,j} \in [2^i, 2^{i+1}]$ when it entered $\tilde{Q}_i$. eRMLF maintains the invariant that it is always running the earliest released job in the lowest nonempty queue.

When a job $\tilde{J}_h$ is released at time $\tilde{r}_h$, eRMLF takes the following actions:

- If, just prior to $\tilde{r}_h$, all queues were empty, then
  —Job $\tilde{J}_h$ is enqueued on $\tilde{Q}^*$.
  —The initial target $\tilde{U}_{*,h}$ is set to $\tilde{U}_{0,h} = \max\{1, 2 - \tilde{\beta}_h\}$.
- If, just prior to $\tilde{r}_h$, there are unfinished jobs in the system but $\tilde{Q}^*$ is empty, then
  —Job $\tilde{J}_h$ is enqueued on $\tilde{Q}^*$.
  —The initial target $\tilde{U}_{*,h}$ is set to $\tilde{U}_{z_h^*-1,h} = 2^{z_h^*-1}\max\{1, 2 - \tilde{\beta}_h\}$, where the queue index $z_h^* = \min\{z \in \mathbb{Z}\colon \tilde{Q}_z \text{ nonempty at time } t\}$ corresponds to the lowest nonempty queue.
- If, just prior to $\tilde{r}_h$, $\tilde{Q}^*$ is nonempty, then $\tilde{Q}^* = \{\tilde{J}_{h-1}\}$ at that time. Now,
  —The target $\tilde{U}_{z_h^*,h-1} = 2^{z_h^*}\max\{1, 2 - \tilde{\beta}_{h-1}\}$ with $z_h^* := \min\{z \in \mathbb{Z}\colon \tilde{w}_{h-1}(\tilde{r}_h) \le \tilde{U}_{z,h-1}\}$ is the lowest target not yet reached by job $\tilde{J}_{h-1}$.
  —Job $\tilde{J}_{h-1}$ is dequeued from $\tilde{Q}^*$.
  —Job $\tilde{J}_{h-1}$ is enqueued on $\tilde{Q}_{z_h^*}$.
  —Job $\tilde{J}_h$ is enqueued on $\tilde{Q}^*$.
  —The initial target $\tilde{U}_{*,h}$ is set to $\tilde{U}_{z_h^*-1,h} = 2^{z_h^*-1}\max\{1, 2 - \tilde{\beta}_h\}$.
- If, just prior to $\tilde{r}_h$, it was the case that eRMLF was running a job $\tilde{J}_j$, then $\tilde{J}_j$ is preempted.
- eRMLF begins running $\tilde{J}_h$.

If at some time t, a job $\tilde{J}_j \in \tilde{Q}_{z-1}$ is being run when $\tilde{w}_j(t)$ becomes equal to $\tilde{U}_{z-1,j}$, then eRMLF takes the following actions:

- Job $\tilde{J}_j$ is dequeued from $\tilde{Q}_{z-1}$.
- Job $\tilde{J}_j$ is enqueued on $\tilde{Q}_z$.
- The target $\tilde{U}_{z,j}$ is set to $2\tilde{U}_{z-1,j} = 2^z\max\{1, 2 - \tilde{\beta}_j\}$.

If at some time t, a job $\tilde{J}_j \in \tilde{Q}^*$ is being run when $\tilde{w}_j(t)$ becomes equal to $\tilde{U}_{*,j}$, then eRMLF takes the following actions:

- Job $\tilde{J}_j$ is dequeued from $\tilde{Q}^*$.
- Job $\tilde{J}_j$ is enqueued on $\tilde{Q}_{z_h^*}$, where $z_h^* = \log_2(\tilde{w}_j(t)/\max\{1, 2 - \tilde{\beta}_j\}) + 1$.
- The target $\tilde{U}_{z_h^*,j}$ is set to $2\tilde{U}_{*,j} = 2^{z_h^*}\max\{1, 2 - \tilde{\beta}_j\}$.

Whenever a job is completed, it is removed from its queue.

way: $\beta_j \equiv \tilde{\beta}_j$ for all $j = 1, \ldots, |\tilde{\mathcal{J}}|$. It immediately follows that the targets $\tilde{U}_{z,j}$ as assigned to $\tilde{\mathcal{J}}$ by eRMLF and the targets $U_{i,j}$ as assigned to $\mathcal{J}$ by RMLF satisfy $\tilde{U}_{i,j} = U_{i,j}$ for all $i \in \{0, 1, 2, \ldots\}$. Additionally, the initial RMLF target $U_{0,j}$ satisfies $U_{0,j} = \tilde{U}_{0,j} = 2^{g-g}\max\{1, 2 - \tilde{\beta}_j\} = 2^{-g}\tilde{U}_{g,j}$.

We will show that the above construction implies an equivalence between RMLF and eRMLF. For all $z \in \mathbb{Z}$ and $t \ge 0$, define the sets

$$\tilde{Q}_z(t) := \{\tilde{J}_j\colon \tilde{U}_{z-1,j} \le \tilde{w}_j(t) < \tilde{U}_{z,j}\} \tag{A.1}$$

that contain all jobs in the system at time t that are in queue $\tilde{Q}_z$, *or the most recently released job if it has received a similar amount of service*. The equivalence is first observed between the initial RMLF queue $Q_0(t)$ and the augmented eRMLF queue $\hat{Q}(t)$, defined as

$$Q_0(t) := \{J_j\colon w_j(t) < U_{0,j}\}, \tag{A.2}$$

and

$$\hat{Q}(t) := \bigcup_{z=-\infty}^{g} \tilde{Q}_z(t) = \{\tilde{J}_j\colon \tilde{w}_j(t) < \tilde{U}_{g,j}\}, \tag{A.3}$$

respectively. One may observe that since Q_0 is the highest priority queue, it experiences a dedicated, work-conserving server that works at unit speed on jobs J_j with sizes $U_{0,j}$. Therefore the event that the RMLF server works on Q_0 is equivalent to the event $\{Q_0(t) > 0\}$.

We assume without loss of generality that job J_1 arrives at time $r_1 = 0$. The arrival of this job initiates a first busy period for $Q_0(t)$ of N_1 jobs, where N_1 is such that the cumulative targets $U_{0,j}$ of the first N_1 jobs can be served before the (N_1+1)-th job is released. It is defined as $N_1 = \inf\{k \ge 1\colon \sum_{j=1}^{k} U_{0,j} - r_{j+1} \le 0\}$, where $r_{|\tilde{\mathcal{J}}|+1}$ is understood as plus infinity. The duration of the busy period is given by $P_1 = \sum_{j=1}^{N_1} U_{0,j} = \sum_{j=1}^{N_1} 2^{-g}\tilde{U}_{g,j}$. The server may then work on jobs in higher queues (perceived as idle time by Q_0) until time r_{N_1+1}, when a new busy period is initiated. For $t \in [0, r_{N_1+1})$, we have now obtained

$$Q_0(t) > 0 \quad\Leftrightarrow\quad t \le P_1 \quad\Leftrightarrow\quad 2^g t \le \sum_{j=1}^{N_1} \tilde{U}_{g,j}. \tag{A.4}$$

By a similar analysis of the augmented queue $\hat{Q}$, we find that for all $t \in [0, 2^g r_{N_1+1}) = [0, \tilde{r}_{N_1+1})$, the relation $\hat{Q}(t) > 0 \Leftrightarrow Q_0(2^{-g}t) > 0$ holds, and for all $t \ge 0$ by a straightforward generalisation of the above procedure. Observing that both algorithms preserve the ordering of jobs, we may similarly show that the eRMLF server processes job $\tilde{J}_j$ in queue $\tilde{Q}_{g+i}$ at time t if and only if the RMLF server processes job J_j in queue Q_i at time $2^g t$ for all $i \in \{1, 2, \ldots\}$, and $t \ge 0$.

From the above results, one may deduce that the average sojourn time $\mathbb{E}[\bar{T}_{\mathrm{eRMLF}}(\tilde{\mathcal{J}})]$ of instance $\tilde{\mathcal{J}}$ under algorithm eRMLF equals 2^g times the average sojourn time $\mathbb{E}[\bar{T}_{\mathrm{RMLF}}(\mathcal{J})]$ of instance $\mathcal{J}$ under RMLF. The competitive ratio of RMLF as stated in Theorem 1 hence guarantees that, for all instances $\tilde{\mathcal{J}}$ of size at most m,

$$\mathbb{E}[\bar{T}_{\mathrm{eRMLF}}(\tilde{\mathcal{J}})] = 2^g \mathbb{E}[\bar{T}_{\mathrm{RMLF}}(\mathcal{J})] \leq C_1 \log(m) 2^g \cdot \bar{T}_{\mathrm{SRPT}}(\mathcal{J}). \tag{A.5}$$

The competitive ratio of eRMLF is concluded by verifying

$$2^g \cdot \bar{T}_{\mathrm{SRPT}}(\mathcal{J}) = \bar{T}_{\mathrm{SRPT}}(\tilde{\mathcal{J}}), \tag{A.6}$$

which is a direct consequence of our scaling. In particular, the constant C_1 in the upper bound is the same for RMLF and eRMLF. □

Endnote

[1] Note that SRPT is deterministic, but not blind.

References

[1] Aalto S, Ayesta U (2006) Mean delay analysis of multi level processor sharing disciplines. *Proc. 25th IEEE Internat. Conf. Comput. Comm., INFOCOM* (IEEE, Piscataway, NJ), , 1–11.
[2] Asmussen S (2003) *Applied Probability and Queues* (Springer, New York).
[3] Bansal N (2005) On the average sojourn time under M/M/1/SRPT. *Oper. Res. Lett.* 33(2):195–200.
[4] Bansal N, Gamarnik D (2006) Handling load with less stress. *Queueing Systems* 54(1):45–54.
[5] Bansal N, Harchol-Balter M (2001) Analysis of SRPT scheduling: Investigating unfairness. *Proc.* 2001 *ACM SIGMETRICS Internat. Conf. Measurement and Modeling Comput. Systems*, SIGMETRICS '01 (ACM, New York), 279–290.
[6] Becchetti L, Leonardi S (2004) Nonclairvoyant scheduling to minimize the total flow time on single and parallel machines. *J. ACM* 51(4):517–539.
[7] Borodin A, El-Yaniv R (1998) *Online Computation and Competitive Analysis* (Cambridge University Press, Cambridge, UK).
[8] Boxma O, Zwart B (2007) Tails in scheduling. *SIGMETRICS Perform. Eval. Rev.* 34(4):13–20.
[9] Conway RW, Maxwell WL, Miller LW (1967) *Theory of Scheduling* (Addison-Wesley Publishing Company, Reading, MA).
[10] Corbató FJ, Merwin-Daggett M, Daley RC (1962) An experimental time-sharing system. *Proc. May* 1–3, 1962, *Spring Joint Comput. Conf.*, AIEE-IRE '62 (Spring) (ACM, New York), 335–344.
[11] Fiat A, Woeginger GJ (1998) Online algorithms: The state of the art (Springer, New York), 1442.
[12] Gromoll HC (2004) Diffusion approximation for a processor sharing queue in heavy traffic. *Ann. Appl. Probab.* 14(2):555–611.
[13] Kalyanasundaram B, Pruhs KR (2003) Minimizing flow time nonclairvoyantly. *J. ACM* 50(4):551–567.
[14] Kleinrock L (1976) *Queueing Systems* II*: Computer Theory* (John Wiley & Sons, New York).
[15] Lin M, Wierman A, Zwart B (2011) Heavy-traffic analysis of mean response time under shortest remaining processing time. *Performance Eval.* 68(10):955–966.
[16] Lotov VI (2002) Inequalities for the moments and distribution of the ladder height of a random walk. *Siberian Math. J.* 43(4):655–660.
[17] Lu D, Sheng H, Dinda P (2004) Size-based scheduling policies with inaccurate scheduling information. DeGroot D, Harrison P, Wijshoff H, Segall Z, eds. *Proc. 12th Internat. Workshop Modeling, Anal., and Simulation Comput. Telecomm. Systems, MASCOTS '04* (IEEE Computer Society, Washington, DC), 31–38.
[18] Motwani R, Phillips S, Torng E (1994) Nonclairvoyant scheduling. *Theoretical Comput. Sci.* 130(1):17–47.
[19] Nair J, Wierman A, Zwart B (2010) Tail-robust scheduling via limited processor sharing. *Performance Eval.* 67(11):978–995.
[20] Nuyens M, Wierman A (2008) The foreground–background queue: A survey. *Performance Eval.* 65(3):286–307.
[21] Nuyens M, Wierman A, Zwart B (2008) Preventing large sojourn times using SMART scheduling. *Oper. Res.* 56(1):88–101.
[22] Pruhs K, Sgall J, Torng E (2004) Online scheduling. Leung J Y-T, ed. *Handbook of Scheduling: Algorithms, Models, and Performance Analysis* (CRC Press, Boca Raton, FL), 15-1.
[23] Puha AL, et al. (2015) Diffusion limits for shortest remaining processing time queues under nonstandard spatial scaling. *Ann. Appl. Probab.* 25(6):3381–3404.
[24] Remerova M, Foss S, Zwart B (2014) Random fluid limit of an overloaded polling model. *Adv. Appl. Probab.* 46(1):76–101.
[25] Righter R, Shanthikumar JG, Yamazaki G (1990) On extremal service disciplines in single-stage queueing systems. *J. Appl. Probab.* 27(2):409–416.
[26] Schrage L (1968) Letter to the editor—A proof of the optimality of the shortest remaining processing time discipline. *Oper. Res.* 16(3):687–690.
[27] Stolyar AL, Ramanan K (2001) Largest weighted delay first scheduling: Large deviations and optimality. *Ann. Appl. Probab.* 11(1):1–48.
[28] Wierman A, Nuyens M (2008) Scheduling despite inexact job-size information. *SIGMETRICS Perform. Eval. Rev.* 36(1):25–36.
[29] Wierman A, Zwart AP (2012) Is tail-optimal scheduling possible? *Oper. Res.* 60(5):1249–1257.
[30] Wierman A, Harchol-Balter M, Osogami T (2005) Nearly insensitive bounds on SMART scheduling. Eager DL, Williamson CL, Borst SC, Lui JCS, eds. *Proc. Internat. Conf. Measurements Modeling Comput. Systems, SIGMETRICS '05*, (ACM, New York), 205–216.