Nudge: Stochastically Improving upon FCFS

ISAAC GROSOF, Carnegie Mellon University, USA
KUNHE YANG, Tsinghua University, China
ZIV SCULLY, Carnegie Mellon University, USA
MOR HARCHOL-BALTER, Carnegie Mellon University, USA

The First-Come First-Served (FCFS) scheduling policy is the most popular scheduling algorithm used in practice. Furthermore, its usage is theoretically validated: for light-tailed job size distributions, FCFS has weakly optimal asymptotic tail of response time. But what if we don't just care about the asymptotic tail? What if we also care about the 99th percentile of response time, or the fraction of jobs that complete in under one second? Is FCFS still best? Outside of the asymptotic regime, only loose bounds on the tail of FCFS are known, and optimality is completely open.

In this paper, we introduce a new policy, Nudge, which is the first policy to provably stochastically improve upon FCFS. We prove that Nudge simultaneously improves upon FCFS at *every* point along the tail, for light-tailed job size distributions. As a result, Nudge outperforms FCFS for every moment and every percentile of response time. Moreover, Nudge provides a multiplicative improvement over FCFS in the asymptotic tail. This resolves a long-standing open problem by showing that, counter to previous conjecture, FCFS is not strongly asymptotically optimal.

1 INTRODUCTION

1.1 The Case for FCFS

While advanced scheduling algorithms are a popular topic in theory papers, it is unequivocal that the most popular scheduling policy used in practice is still First-Come First-Served (FCFS). There are many reasons for the popularity of FCFS. From a practical perspective, FCFS is easy to implement. Additionally, FCFS has a feeling of being fair.

However, there are also theoretical arguments for why one should use FCFS. For one thing, FCFS minimizes the *maximum* response time across jobs for *any* finite arrival sequence of jobs. By *response time* we mean the time from when a job arrives until it completes service.

For another thing, in an M/G/1 with a light-tailed job size distribution, FCFS is known to have a *weakly optimal* asymptotic tail of response time [5, 36]. Specifically, using T to denote response time, the asymptotic tail under FCFS is of the form:

$$\mathbf{P}\{T^{\text{FCFS}} > t\} \sim C_{\text{FCFS}}e^{-\theta^*t}, \tag{1}$$

where "~" indicates that the ratio of the two quantities converges to 1 in the $t \to \infty$ limit.

The exponent θ^* in (1) is known to be optimal, while the optimality of C_{FCFS} is an open problem [5]. The asymptotic tail growth under FCFS has been compared with more sophisticated policies [5]. It has been shown that, for light-tailed job size distributions, the tail of response time under

---

Processor-Sharing (PS), Preemptive Last-Come-First-Served (PLCFS), and Shortest-Remaining-Processing-Time (SRPT) each take the asymptotic form of

$$\mathbf{P}\{T > t\} \sim C'e^{-\theta't},$$

where θ' is the *worst possible* exponential decay rate [25] over all work-conserving scheduling policies. Roughly, FCFS's tail exponent θ^* arises from the tail of the workload distribution, while the other policies' tail exponent θ' arises from the tail of the busy period distribution, which is much larger under light-tailed job size distributions.

1.2   The Case For Light-Tailed Job Size Distributions

In this paper, we choose to focus on the case of light-tailed job size distributions. Light-tailed job size distributions show up naturally in workloads where all the transactions are of the same type (say shopping); while there is some variability in the time it takes to purchase an item, even high-variability distributions that arise in such settings are often light-tailed. Also, many natural distributions, like the Normal distribution, Exponential distribution, and all Phase-type distributions, are light-tailed. Finally, while heavy-tailed job size distributions are certainly prevalent in empirical workloads (see for example [9, 14, 37]), in practice, these heavy-tailed workloads are often *truncated*, which immediately makes them light-tailed. Such truncation can happen because there is a limit imposed on how long jobs are allowed to run. Alternatively, truncation can occur when a heavy-tailed job size distribution is divided into a few size classes as in [15, 18] where the smaller size classes end up being truncated distributions.

1.3   The Case for Non-Asymptotic Tails

Within the world of light-tailed job size distributions, FCFS is viewed as the best policy. However, while FCFS has a weakly optimal *asymptotic* tail, it is not best at minimizing $\mathbf{P}\{T > t\}$ for *all* t. In practice, one cares less about the asymptotic case than about particular t [17]. For example, one might want to minimize the fraction of response times that exceed $t = 0.5$ seconds, because such response times are noticeable by users. One might also want to meet several additional Service Level Objectives (SLOs) where one is charged for exceeding particular response time values, such as $t = 1$ minute, or $t = 1$ hour. SLOs are very common in the computing literature [8, 17, 24], in service industries [10, 34, 38], and in healthcare [4, 19]. Unfortunately, different applications have different SLOs. This leads us to ask:

> *When considering* $\mathbf{P}\{T > t\}$, *is it possible to strictly improve upon FCFS for* <u>all</u> *values of* t*?*

We are motivated by the fact that, for lower values of t, Shortest-Remaining-Processing-Time (SRPT) is better than FCFS, although FCFS clearly beats SRPT for higher values of t, as FCFS is weakly asymptotically optimal while SRPT is asymptotically pessimal [25, 26]. SRPT also minimizes mean response time [29], which is closely related to lower values of t. This motivates us to consider whether prioritizing small jobs might have some benefit, even in the world of light-tailed job size distributions.

We ask more specifically:

> *Can partial prioritization of small jobs lead to a strict improvement over FCFS? Specifically, is there a scheduling policy which strictly improves upon FCFS with respect to* $\mathbf{P}\{T > t\}$, *for* every *possible* t *including large* t*?*

1.4   Our Answer: Nudge

This paper answers the above question in the affirmative. We will define a policy, which we call *Nudge*, whose response time tail is *provably* better than that of FCFS for *every* value of t, assuming a
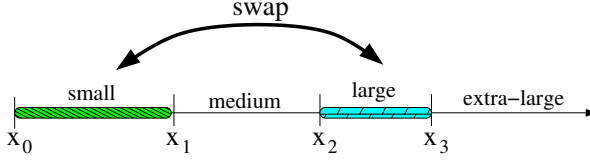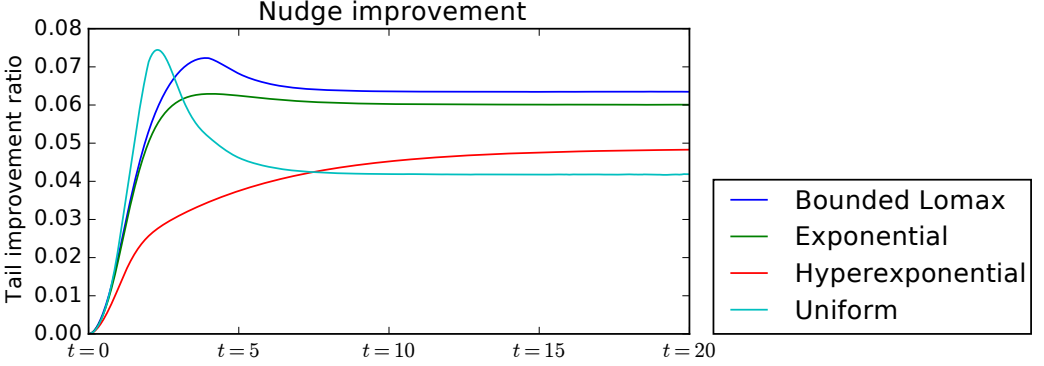
Fig. 1. The Nudge algorithm.

Fig. 2. Empirical tail improvement of Nudge over FCFS in an M/G/1. The tail improvement ratio (TIR) is defined as $1 - \mathbf{P}\{T^{\text{Nudge}} > t\} / \mathbf{P}\{T^{\text{FCFS}} > t\}$. Specific job size distributions, each with mean 1: Uniform(0,2); Exponential with mean 1; Hyperexponential with branches drawn from Exp(2) and Exp(1/3), where the first branch has probability 0.8 and where $C^2 = 3$; BoundedLomax($\lambda = 2$, max= 4, $\alpha = 2$). Distributions in legend ordered by asymptotic improvement. Simulations run with 10 billion arrivals. Load $\rho = 0.8$. Nudge parameters: $x_1 = 1, x_2 = 1, x_3 = \infty$. Note that $x_1 = x_2$, so there are no medium-sized jobs. Empirically, Nudge often achieves its best performance with $x_1 = x_2$, though our proofs involve setting $x_1 < x_2$. See Section 9.

light-tailed job size distribution[1] (see Theorem 4.2). We say that Nudge's response time *stochastically improves upon* that of FCFS, in the sense of stochastic dominance. Moreover, we prove that the asymptotic tail of response time of Nudge is of the form

$$P\{T^{\text{Nudge}} > t\} \sim C_{\text{Nudge}} e^{-\theta^* t},$$

with optimal decay rate θ^* and a superior leading constant $C_{\text{Nudge}} < C_{\text{FCFS}}$ (see Corollary 4.4). Thus, we demonstrate that FCFS is *not* strongly optimal, answering an open problem posed by Boxma and Zwart [5]. In particular, this is contrary to a conjecture of Wierman and Zwart [40] (see Section 2.1).

The intuition behind the Nudge algorithm is that we'd like to basically stick to FCFS, which we know is great for handling the extreme tail (high t), while at the same time incorporating a little bit of prioritization of small jobs, which we know can be helpful for the mean and lower t. We need to be careful, however, not to make too much use of size, because Nudge still needs to beat FCFS for high t; hence we want just a little "nudge" towards prioritizing small jobs.

We now describe the Nudge algorithm. Imagine that the job size distribution is divided into size regions, as shown in Fig. 1, consisting of small, medium, large, and extra large jobs. Most of the time, Nudge defaults to FCFS. However, when a "small" job arrives and finds a "large" job immediately ahead of it in the queue, we swap the positions of the small and large job in the queue.

---

[1]Technically, a Class I job size distribution. See Definition 3.3.

The one caveat is that a job which has already swapped is ineligible for further swaps. The size cutoffs defining small and large jobs will be defined later in the paper.

The degree of the tail improvement of Nudge over FCFS is non-trivial. In Fig. 2, we see that for many common light-tailed job size distributions, Nudge results in a multiplicative improvement of 4-7% throughout the tail. In Section 9.2, we show that with low load and a high-variability job size distribution, Nudge's improvement can be as much as 10-15% throughout the tail. The magnitude of these improvements highlights the importance of scheduling, even in the light-tailed setting.

We additionally present an exact analysis of the performance of Nudge. Nudge does not fit into any existing framework for M/G/1 transform analysis, including the recently developed SOAP framework [32] (see Section 2.3). Nonetheless, we derive a tagged-job analysis of Nudge in Theorem 4.5, deriving the Laplace-Stieltjes transform of response time of Nudge.

1.5 Contributions and Roadmap

This paper makes the following contributions.

- In Section 3.5, we introduce the Nudge policy.
- In Sections 4 to 6 we prove that with appropriately chosen parameters, Nudge stochastically improves upon FCFS for light-tailed[2] job size distributions; we also give a simple expression for such parameters. Moreover, in Section 8, we prove that Nudge achieves a multiplicative asymptotic improvement over FCFS.
- In Section 7, we derive the Laplace-Stieltjes transform of response time under Nudge.
- In Section 9, we empirically demonstrate the magnitude of Nudge's stochastic improvement over FCFS. We also discuss how to tune Nudge's parameters for best performance.
- In Section 10, we discuss practical considerations for using Nudge.
- In Section 11, we discuss several notable variants of Nudge.

2 PRIOR WORK

Most prior work on scheduling to optimize the tail of response time focuses on the *asymptotic* case, characterizing $\mathbf{P}\{T > t\}$ in the $t \to \infty$ limit. We review these results in Section 2.1.

Our main result, Theorem 4.2, is a *non-asymptotic* statement, characterizing the behavior of $\mathbf{P}\{T > t\}$ for *all* t, not just the $t \to \infty$ limit. There is much less prior work on the tail of response time outside of the asymptotic regime. We review the few results in this area in Section 2.2.

In addition to characterizing Nudge's tail of response time, we also give an exact transform analysis of Nudge's response time. Our analysis requires a novel approach that significantly differs from traditional analyses, as we discuss in Section 2.3.

Our paper's focus is the M/G/1 queue. All of the results cited in this section apply to the M/G/1, with some also applying to more general models, such as the GI/GI/1.

2.1 Asymptotic Tails: Extensive Theory, but Open Problems Remain

When optimizing the asymptotic tail, the goal is to find a policy π^* such that for all scheduling policies π,

$$\limsup_{t \to \infty} \frac{\mathbf{P}\{T^{\pi^*} > t\}}{\mathbf{P}\{T^{\pi} > t\}} \le c$$

for some constant $c \in [1, \infty)$. Such a policy π^* is called *weakly optimal*; if $c = 1$, then π^* is called *strongly optimal* [5]. While weak optimality has been well studied, proving or disproving strong optimality is much harder.

---

[2]Technically, any Class I job size distribution. See Definition 3.3.

One major theme of the prior work is that optimizing the asymptotic tail looks very different depending on the job size distribution.

- For light-tailed job sizes, FCFS is weakly optimal [5]. Specifically, the tail of response time has a form given in (1). Moreover, many popular preemptive policies such as PS, SRPT, and Foreground-Background (FB)[3] are "weakly pessimal": they have the *maximum* possible asymptotic tail, up to a constant factor, of any work-conserving scheduling policy [25, 26].
- For heavy-tailed job sizes, the reverse is true: PS, SRPT, FB, and similar policies are all weakly optimal [5, 26, 33], while FCFS is weakly pessimal.

This state of affairs prompts a question: is any policy weakly optimal in both the light-tailed and heavy-tailed cases? Nair et al. [25] show that a variant of PS achieves this, but their variant requires knowledge of the system's load. Wierman and Zwart [40] show that *any* policy that is weakly optimal in both the light- and heavy-tailed cases requires knowing some information about the system parameters, such as the load.

The above results mostly characterize weakly optimal scheduling policies, but the problem of strongly optimizing the tail remains open. Boxma and Zwart [5] pose the strong optimality of FCFS as an open problem. Wierman and Zwart [40] go further and conjecture that FCFS is in fact strongly optimal for light-tailed job size distributions. Despite a large of body of work analyzing the tail asymptotics of FCFS [1–3, 28], the problem has remained open. We solve the problem by showing that FCFS is *not* strongly optimal. Specifically, our Corollary 4.4 implies

$$\lim_{t \to \infty} \frac{\mathbf{P}\{T^{\mathsf{Nudge}} > t\}}{\mathbf{P}\{T^{\mathsf{FCFS}} > t\}} = \frac{C_{\mathsf{Nudge}}}{C_{\mathsf{FCFS}}} < 1.$$

2.2 Non-asymptotic Tails: Few Optimality Results

Characterizing $\mathbf{P}\{T > t\}$ outside the asymptotic regime is a much harder problem than characterizing the asymptotic tail. As such, the strongest results in this area are for relatively simple scheduling policies. For FCFS under light-tailed job size distributions, it is known that $\mathbf{P}\{T > t\} < e^{-\theta^* t}$ for the same constant θ^* as appears in FCFS's asymptotic tail formula [21, 22]. As a result, this bound is tight up to a constant factor [22], subject to subtleties discussed in Section 3.3. Beyond FCFS, one of the few known results gives an improved characterization of response time under preemptive-priority scheduling policies [3, Section 2].

Very little is known about more complicated scheduling policies. While the Laplace-Steiltjes transform of T is known for a wide variety of scheduling policies [32, 35], these transforms do not readily yield useful bounds on $\mathbf{P}\{T > t\}$ for general job size distributions.

Given that characterizing $\mathbf{P}\{T > t\}$ is difficult, it comes as no surprise that optimizing $\mathbf{P}\{T > t\}$ is harder still. As such, rather than trying to crown a single optimal policy, we focus on a *relative* measure. Specifically, as we define in Definition 3.1, we say that policy π_1 *stochastically improves upon* another policy π_2 if $\mathbf{P}\{T^{\pi_1} > t\} \leq \mathbf{P}\{T^{\pi_2} > t\}$ for all t.

There are two stochastic improvement results in the literature, but both are much simpler than our Nudge result. Both results start with a well-known policy that does not use job sizes and show that a variation that does use job sizes stochastically improves response time.

- Nuyens et al. [26] show that SRPT and similar policies stochastically improve upon FB.
- Friedman and Henderson [11] and Friedman and Hurley [12] show that one can stochastically improve upon PS by using job sizes. Their policy, Fair Sojourn Protocol (FSP), guarantees in a sample-path sense that no job departs later than it would if the server were using PS.

---

[3]FB at all times serves the jobs that have received the least service so far.

The results above fit a common theme. Both FB and PS often share the server between multiple jobs. Sharing the server is fundamentally suboptimal. For example, when sharing the server between jobs 1 and 2, if we knew that we would finish job 1 first, then it would be better to devote the entire server to job 1 at first. Doing so improves the response time of job 1 without harming the response time of job 2. Roughly speaking, when FB and PS would share the server between jobs, SRPT and FSP serve the jobs one at a time, using job size information to choose the ordering.

FCFS is more difficult to stochastically improve upon than FB and PS. For one thing, FCFS never shares the server, removing this easy opportunity for stochastic improvement. Moreover, there is a sense in which FCFS is unimprovable: on any specific finite arrival sequence, FCFS minimizes the sorted vector of response times, where we order vectors lexicographically. For example, FCFS minimizes the maximum response time. As a result, the sample path arguments that work for improving FB and PS do not apply to improving FCFS.

In spite of these obstacles, we show in Theorem 4.2 that Nudge stochastically improves upon FCFS. Rather than reasoning in terms of sample paths, we take a fundamentally stochastic approach from the beginning. See our proofs in Section 5.

2.3 Transform of Response Time: Nudge Needs a Novel Approach

In Theorem 4.5, we give a closed-form expression for the Laplace-Stieltjes transform of Nudge's response time. There has been much prior work on analyzing the transform of response time of the M/G/1 under various scheduling policies. Some analysis techniques cover a wide variety of scenarios [13, 32]. However, as we explain below, none of these prior techniques can analyze Nudge.

SOAP Policies. Policies in the SOAP class, introduced by Scully et al. [32], schedule jobs based on an index calculated from each job's size and attained service[4], and their response time can be analyzed via the SOAP framework [32]. These include SRPT [31], FB [30], some multi-level processor sharing policies [23], and certain cases of the Gittins policy [27]. Unfortunately, Nudge is not a SOAP policy, so we cannot leverage this analysis method. This is because whether Nudge will swap a small job s with a large job ℓ depends in part on whether any other jobs arrive between ℓ and s. In contrast, a SOAP policy would make such a decision based on properties of ℓ and s alone.

Variations on FCFS. Nudge serves jobs in FCFS order by default and only ever swaps adjacent arrivals. One might therefore hope that Nudge could be analyzed as a variation on FCFS. There are many papers analyzing a variety of M/G/1 variants under FCFS scheduling. These include systems with generalized vacations [13] and exceptional first service [39]. Unfortunately, to the best of our knowledge, no prior analysis of a variation of FCFS applies to Nudge.

Other Analysis Techniques. There are a number of scheduling policies whose transform analyses do not fit into either of the previous categories, such as random order of service [20] and systems with accumulating priority [35]. However, these policies do not resemble Nudge, and the techniques used in their analyses do not readily apply to Nudge.

3 MODEL

3.1 Notation

We consider the M/G/1 queue in which job sizes are known. Let λ be the arrival rate, S be the job size distribution, and $s_{\min}$ be the minimum possible job size. Specifically, let $s_{\min}$ be the infimum of the support of S. We denote the load by $\rho = \lambda \mathbf{E}[S]$ and assume $0 < \rho < 1$.

---

[4]The index can also depend on certain other characteristics of the job, e.g. its class if there are multiple classes of jobs. Size and attained service are the attributes relevant to Nudge.

The queueing time, T_Q, is the time from when a job arrives until it first receives service. The response time, T, is the time from when a job arrives until it completes. We write T_Q^{Alg} and T^{Alg} for the queueing time and response time under scheduling algorithm Alg, respectively.

For any continuous random variable V, we will use $f_V(\cdot)$ to denote the probability density function (p.d.f.) of V. We write the $\widetilde{V}(s)$ for the Laplace-Stieltjes transform of V.

3.2 Stochastic Improvement

In this paper, our goal is to prove that the Nudge policy stochastically improves upon the FCFS policy. We now define stochastic improvement, along with the related notion of tail improvement ratio.

Definition 3.1 (Stochastic Improvement). For two scheduling algorithms Alg_1 and Alg_2, we say that Alg_1 (strictly) *stochastically improves* upon Alg_2 if, for any response time cutoff $t > s_{\min}$, the probability that response time of Alg_1 exceeds t is smaller than the probability that Alg_2's response time exceeds t, i.e.,

$$\forall t > s_{\min}, \quad \mathbf{P}\{T^{\text{Alg}_1} > t\} < \mathbf{P}\{T^{\text{Alg}_2} > t\}.$$

Definition 3.2 (Tail improvement ratio). For any response time cutoff t, the *tail improvement ratio* of Alg_1 versus Alg_2 at t, denoted $\text{TIR}(t)$, is defined as

$$\text{TIR}(t) \triangleq 1 - \frac{\mathbf{P}\{T^{\text{Alg}_1} > t\}}{\mathbf{P}\{T^{\text{Alg}_2} > t\}}.$$

The *asymptotic tail improvement ratio*, denoted AsymTIR, is defined as

$$\text{AsymTIR} \triangleq \liminf_{t \to \infty} \text{TIR}(t) = 1 - \limsup_{t \to \infty} \frac{\mathbf{P}\{T^{\text{Alg}_1} > t\}}{\mathbf{P}\{T^{\text{Alg}_2} > t\}}.$$

3.3 Class I "Light-Tailed" Distributions

In this paper, we focus on job size distributions for which the FCFS policy has an asymptotically exponential waiting time distribution. This property of FCFS will be crucial for our analysis. Prior work has exactly characterized the job size distributions for which FCFS has this property. These distributions are known as "class I" distributions [1, 3, 28].

Definition 3.3 (Class I Distribution). For a distribution S, let $-s^*$ be the rightmost singularity of $\widetilde{S}(s)$, with $-s^* = -\infty$ if $\widetilde{S}(s)$ is analytic everywhere. S is a *class I distribution* if and only if $s^* > 0$ and $\widetilde{S}(-s^*) = \infty$.

Class I distributions can roughly be thought of as "well-behaved" light-tailed distributions. In contrast, class II distributions, the other class of light-tailed distributions, are very unusual and "paradoxical", and rarely occur as job size distributions.

For our paper, the key property of class I job size distributions is that they cause FCFS to have an asymptotically exponential waiting time distribution for all loads [1, 2]. However, as shown by [1, 2], the waiting time also exhibits an exponential tail for light load if the job size is class II. For this reason, while we focus only on class I distributions, we believe that our results also hold for class II under light load. In Section 3.4, we characterize the exponential waiting time in more detail.

3.4 Characterizing the FCFS Waiting Time Distribution

In this paper, we care about the exponential tail of the FCFS response time distribution. It turns out to be simpler to focus on the FCFS waiting time distribution, which is closely related. We will make use of two key concepts regarding the waiting time distribution. The first concept is the asymptotic

exponential decay rate, as investigated in [2, 5]. We refer to this quantity as θ^* and formally define it to be the negative of the rightmost singularity of $\widetilde{T_Q^{\text{FCFS}}}$. Based on the Cramer-Lundberg theory, the waiting time distribution T_Q^{FCFS} takes an asymptotic exponential tail:

$$\mathbf{P}\left\{T_Q^{\text{FCFS}} > t\right\} \sim Ce^{-\theta^* t}. \tag{2}$$

The quantity θ^* is the least positive real solution to the equation

$$\widetilde{S}(-\theta^*) = \frac{\lambda + \theta^*}{\lambda}.$$

We also define the *normalized p.d.f.* to be

$$g(t) \triangleq f_{T_Q^{\text{FCFS}}}(t) \cdot e^{\theta^* t}. \tag{3}$$

Note that (2) relates to the c.d.f. of waiting time, while (3) relates to the p.d.f. of waiting time.

We characterize three important properties of the normalized p.d.f., namely its maximum, minimum, and asymptotic limit. Let $g_{\max}$, $g_{\min}$, g^* denote respectively the maximum, minimum and asymptotically limiting values of $g(\cdot)$ over $[0, \infty)$:

$$g_{\max} \triangleq \sup_{t \in [0,\infty)} g(t); \qquad g_{\min} \triangleq \inf_{t \in [0,\infty)} g(t); \qquad g^* \triangleq \lim_{t \to \infty} g(t).$$

The following lemma, proven in Appendix A, implies these quantities are well defined.

LEMMA 3.1. *Suppose S is a continuous class I job size distribution. For any load ρ, the normalized p.d.f. $g(t)$ is bounded above and below by positive constants, and $\lim_{t \to \infty} g(t)$ exists.*

The ratio $g_{\max}/g_{\min}$ will be particularly important in our analysis. Intuitively, we can think of the ratio as measuring the deviation of the queueing time T_Q^{FCFS} from a perfect exponential distribution. The queueing time distribution is exactly an exponential distribution in an $M/M/1$, and diverges from an exponential to greater or lesser degree under any class I job size distribution. The degree of divergence will show up in our later results.

3.5 Scheduling Algorithm: Nudge

We now formally define the Nudge algorithm. Nudge(x_1, x_2, x_3) first divides jobs into four regions based on their sizes:

- "small": $0 = x_0 \leq S < x_1$.
- "medium": $x_1 \leq S < x_2$.
- "large": $x_2 \leq S < x_3$.
- "very large": $x_3 \leq S < x_4 = \infty$.

Throughout the paper, we concentrate mostly on the "small" and the "large" jobs. For conciseness, we define S_{small}, S_{large}, p_{small}, p_{large} as follows.

Definition 3.4. We define S_{small} and S_{large} to be the distribution of small and large jobs, respectively. We also define p_{small} and p_{large} to be the fraction of small and large jobs, respectively.

$$S_{\text{small}} \sim [S|S < x_1], \qquad S_{\text{large}} \sim [S|x_2 \leq S < x_3]$$
$$p_{\text{small}} \triangleq \mathbf{P}\{S < x_1\}, \qquad p_{\text{large}} \triangleq \mathbf{P}\{x_2 \leq S < x_3\}.$$

To determine which job to serve, Nudge maintains an ordering over jobs which have not yet entered service. We call this ordering the "queue". For each job, we also track whether or not it each has already been "swapped".

Whenever a job completes, Nudge serves the job at the front of the queue (if any), and serves it to completion. By default, newly arriving jobs are placed at the back of the queue, resulting in FCFS scheduling by default. However, if *three conditions* are satisfied, then a "swap" is performed. If

(1) the arriving job is a small job, j_s,
(2) the job at the back of queue is a large job, j_ℓ, and
(3) the job j_ℓ at the back of queue has never been swapped,

then Nudge places the small job j_s just ahead of j_ℓ, in the second-to-last position in the queue. This is called a *swap*, and both j_ℓ and j_s are now marked as having been "swapped."

Because Nudge never swaps the same job twice, a job is only in the last position in the queue and eligible to be swapped immediately after it arrives. As a result, Nudge only ever swaps a job with the job that arrives immediately before or after it.

4 MAIN RESULTS

4.1 Nudge Improves upon FCFS Non-Asymptotically

Our main goal is to show that Nudge stochastically improves upon FCFS. Nudge's performance crucially depends on the choice of parameters x_1, x_2, and x_3, i.e. which jobs are small and which jobs are large. We begin by asking: given job size distribution S and load ρ, for what choices of parameters x_1, x_2, and x_3 does Nudge stochastically improve upon FCFS? We answer this in Theorem 4.1, which gives sufficient conditions on the parameters for Nudge to stochastically improve upon FCFS. We prove Theorem 4.1 in Section 5.

THEOREM 4.1 (STOCHASTIC IMPROVEMENT REGIME). *Suppose S is a continuous class I job size distribution. Then* Nudge(x_1, x_2, x_3) *stochastically improves upon* FCFS *for any $s_{\min} < x_1 \leq x_2 \leq x_3$ satisfying*[5]

$$\bullet \quad \frac{g_{\max}}{g_{\min}} \frac{\lambda + \theta^*}{\lambda} < \frac{1 - \widetilde{S_{\text{large}}}(-\theta^*)^{-1}}{1 - \widetilde{S_{\text{small}}}(-\theta^*)^{-1}}, \tag{4}$$

$$\bullet \quad x_1 + x_3 \leq 2x_2. \tag{5}$$

With Theorem 4.1 in hand, our goal reduces to the following question: given S and ρ, do there exist parameters satisfying the sufficient condition from Theorem 4.1? We answer this in Theorem 4.2, showing that as long as the minimum job size $s_{\min} = 0$, such parameters always exist. Our proof of Theorem 4.2 in Section 6 gives a simple construction of those parameters.

THEOREM 4.2 (EXISTENCE OF STOCHASTIC IMPROVEMENT). *For any continuous class I job size distribution S with $s_{\min} = 0$ and any load $0 < \rho < 1$, there exist x_1, x_2, x_3 satisfying (4) and (5), implying that* Nudge(x_1, x_2, x_3) *stochastically improves upon* FCFS.

4.2 Nudge Improves upon FCFS Asymptotically

Having shown that Nudge stochastically improves upon FCFS, we ask: is Nudge's improvement non-negligible in the asymptotic limit? We answer this in Theorem 4.3. Specifically, recall that in the $t \to \infty$ limit, $\mathbf{P}\{T^{\text{FCFS}} > t\} \sim C_{\text{FCFS}} e^{-\theta^* t}$. We show that $\mathbf{P}\{T^{\text{Nudge}} > t\} \sim C_{\text{Nudge}} e^{-\theta^* t}$ and that, with appropriately set parameters, $C_{\text{Nudge}} < C_{\text{FCFS}}$. This implies that FCFS is *not* strongly optimal for asymptotic tail behavior (see Section 2.1), resolving a long-standing open problem [5, 40]. We also exactly derive the difference $C_{\text{FCFS}} - C_{\text{Nudge}}$. We prove Theorem 4.3 in Section 8, making use of Theorem 4.5.

---

[5]Recall from Definition 3.4 that S_{small} and S_{large} depend on x_1, x_2, and x_3. This applies throughout the paper.

THEOREM 4.3 (ASYMPTOTIC IMPROVEMENT REGIME). *Suppose S is a continuous class I job size distribution. For any $s_{\min} < x_1 \leq x_2 \leq x_3$, the asymptotic tail improvement ratio of $\mathrm{Nudge}(x_1, x_2, x_3)$ compared to FCFS is*

$$\mathrm{AsymTIR} = p_{\mathrm{small}} p_{\mathrm{large}} \frac{\lambda}{\lambda + \theta^*} \left(\widetilde{S_{\mathrm{large}}}(-\theta^*) - \frac{\lambda}{\lambda + \theta^*} \widetilde{S_{\mathrm{small}}}(-\theta^*) - \frac{\theta^*}{\lambda + \theta^*} \widetilde{S_{\mathrm{large}}}(-\theta^*) \widetilde{S_{\mathrm{small}}}(-\theta^*) \right).$$

Furthermore, $\mathrm{AsymTIR}$ *is positive, meaning* $C_{\mathrm{Nudge}} < C_{\mathrm{FCFS}}$, *if*

$$\frac{\lambda + \theta^*}{\lambda} < \frac{1 - \widetilde{S_{\mathrm{large}}}(-\theta^*)^{-1}}{1 - \widetilde{S_{\mathrm{small}}}(-\theta^*)^{-1}}.$$

Note that the asymptotic improvement regime in Theorem 4.3 is a superset of the non-asymptotic improvement regime in Theorem 4.1, because $g_{\max}/g_{\min} \geq 1$. Thus, whenever Theorem 4.1 guarantees a stochastic improvement, we also have $C_{\mathrm{FCFS}} > C_{\mathrm{Nudge}}$. Thus, by Theorem 4.2, there exists an asymptotic improvement whenever $s_{\min} = 0$.

COROLLARY 4.4 (EXISTENCE OF ASYMPTOTIC IMPROVEMENT). *For any continuous class I job size distribution S with $s_{\min} = 0$ and any load $0 < \rho < 1$, there exist x_1, x_2, x_3 such that $C_{\mathrm{Nudge}} < C_{\mathrm{FCFS}}$.*

While Theorem 4.3 shows that there is a multiplicative improvement in the asymptotic tail, we find empirically that the same multiplicative improvement exists throughout nearly the entire tail. See Fig. 2 and Section 9.

4.3 Exact Analysis of Nudge

All of the above results compare Nudge's performance to that of FCFS. In particular, none of these results give an exact analysis of Nudge's response time. We give such an analysis in Theorem 4.5, in which we exactly derive $\widetilde{T^{\mathrm{Nudge}}}(s)$. This result is nontrivial, because Nudge does not fall into any class of policies with known analyses (see Section 2.3). We prove Theorem 4.5 in Section 7.

THEOREM 4.5 (TRANSFORM OF RESPONSE TIME). *The response time of Nudge has Laplace-Stieltjes transform*

$$\widetilde{T^{\mathrm{Nudge}}}(s) = \widetilde{T^{\mathrm{FCFS}}}(s) + p_{\mathrm{small}} p_{\mathrm{large}} \left(\widetilde{S_{\mathrm{large}}}(s)(1 - \widetilde{S_{\mathrm{small}}}(s)) \left(\widetilde{T_Q^{\mathrm{FCFS}}}(\lambda + s) - \widetilde{T_Q^{\mathrm{FCFS}}}(s) \right) \right.$$
$$\left. + \widetilde{S_{\mathrm{small}}}(s)(1 - \widetilde{S_{\mathrm{large}}}(s)) \left(\frac{\widetilde{T_Q^{\mathrm{FCFS}}}(s)}{\widetilde{S}(s)} - (1 - \rho) \frac{\lambda/\widetilde{S}(\lambda) - s/\widetilde{S}(s)}{\lambda - s} \right) \right).$$

5 PROOF OF THEOREM 4.1: STOCHASTIC IMPROVEMENT REGIME

Our goal in this section is to prove Theorem 4.1, which gives sufficient conditions on the parameters x_1, x_2, and x_3 for Nudge to stochastically improve upon FCFS. To do so, we employ a tagged job approach. In particular, we follow an arbitrary tagged job i making its way through a pair of coupled systems, one employing the FCFS policy and one employing the Nudge policy, both with the same arrival process and job sizes.

We focus on one particular response time threshold t, and in particular on the events $D_{i,t}$ and $I_{i,t}$, where the tagged job i has response time greater than t in one system and below in the other system. In (6), we write the difference in the response time tails of Nudge and FCFS in terms of the events $D_{i,t}$ and $I_{i,t}$. In Lemma 5.1, we derive formulas for the probabilities of these events.

Using these formulas, in Lemma 5.2, we derive a sufficient condition for Nudge to improve upon FCFS relative to a specific threshold t. This sufficient condition is dependent on the threshold t.

In order to remove this dependence, we prove Lemma 5.3, a technical lemma regarding arbitrary random variables.

Finally, in Section 5.2, we prove Theorem 4.1, by demonstrating that the conditions given in Theorem 4.1 ensure that the sufficient condition in Lemma 5.2 holds relative to every response time threshold t, making use of Lemma 5.3 to do so.

5.1 Intermediate Lemmas

Consider a tagged job i that arrives into the steady state of the pair of coupled systems. We write T_i^{Nudge} and T_i^{FCFS} for job i's response time in the Nudge and FCFS systems, respectively. For any $t \geq 0$, define the events

$$I_{i,t} \triangleq \left\{ T_i^{\text{FCFS}} \leq t < T_i^{\text{Nudge}} \right\}, \qquad D_{i,t} \triangleq \left\{ T_i^{\text{Nudge}} \leq t < T_i^{\text{FCFS}} \right\}.$$

Intuitively, $D_{i,t}$ is the event in which Nudge *decreases* job i's response time relative to FCFS, specifically from above t to below t. Similarly, $I_{i,t}$ is the event in which Nudge *increases* job i's response time relative to FCFS. We can write

$$\mathbf{P}\left\{ T_i^{\text{Nudge}} > t \right\} = \mathbf{P}\left\{ T_i^{\text{FCFS}} > t \right\} + \mathbf{P}\left\{ I_{i,t} \right\} - \mathbf{P}\left\{ D_{i,t} \right\}. \tag{6}$$

The events $D_{i,t}$ and $I_{i,t}$ are defined using the Nudge and FCFS systems. Our next step is to express them in terms of only the FCFS system, which we understand well.

We begin by defining the relevant quantities in the FCFS system. Let i^- be the arrival immediately before job i, and let i^+ be the arrival immediately after, and let

$\quad W_i \triangleq$ amount of work in the system (either Nudge or FCFS) when job i arrives,

$\quad A_i \triangleq$ interarrival time between jobs i and i^+,

$\quad S_i \triangleq$ size of job i.

We define analogous quantities for i^- and i^+. The work is the same in both systems because both Nudge and FCFS are work-conserving.

Note that Nudge will only ever swap job i with one of the adjacent arrivals, i^- or i^+ (see Section 3.5). Under what condition do we swap job i with job i^+? This happens if and only if the following events occur:

(a) Job i is large, which is when $x_2 \leq S_i < x_3$.
(b) Job i^+ is small, which is when $S_{i^+} < x_1$.
(c) Job i^+ arrives before job i enters service in the Nudge system.
(d) Job i has not swapped with any other job, namely job i^-.

Because job i cannot be both large and small, (a) implies (d). But (d) implies that (c) happens when $A_i \leq W_i$. This is because in the absence of swaps, job i would enter service after W_i time. Therefore, the event that job i swaps with job i^+ is

$$\text{swap}_{i,i^+} \triangleq \{(x_2 \leq S_i < x_3) \ \wedge \ (S_{i^+} \leq x_1) \ \wedge \ (A_i \leq W_i)\}. \tag{7}$$

Crucially, this definition of swap_{i,i^+} depends only on quantities in the FCFS system. We define $\text{swap}_{i^-,i}$ analogously.

LEMMA 5.1 (EVALUATING $\mathbf{P}\left\{ I_{i,t} \right\}$ AND $\mathbf{P}\left\{ D_{i,t} \right\}$). *We have*

$$\mathbf{P}\left\{ I_{i,t} \right\} = \mathbf{P}\left\{ \text{swap}_{i,i^+} \ \wedge \ (W_i + S_i \leq t < W_i + S_i + S_{i^+}) \right\}, \tag{8}$$

$$\mathbf{P}\left\{ D_{i,t} \right\} = \mathbf{P}\left\{ \text{swap}_{i^-,i} \ \wedge \ (W_{i^-} - A_{i^-} + S_i \leq t < W_{i^-} - A_{i^-} + S_i + S_{i^-}) \right\}. \tag{9}$$

PROOF. We begin by computing $\mathbf{P}\{I_{i,t}\}$. The event $I_{i,t}$ occurs only if $T_i^{\text{Nudge}} > T_i^{\text{FCFS}}$, which in turn occurs only if job i swaps with the next arrival, namely job i^+. If this swap occurs, then $T_i^{\text{Nudge}} = T_i^{\text{FCFS}} + S_{i^+}$. We know that $T_i^{\text{FCFS}} = W_i + S_i$, so (8) follows from

$$
\begin{aligned}
I_{i,t} &= \text{swap}_{i,i^+} \ \wedge \ (T_i^{\text{FCFS}} \le t < T_i^{\text{Nudge}}) \\
&= \text{swap}_{i,i^+} \ \wedge \ (T_i^{\text{FCFS}} \le t < T_i^{\text{FCFS}} + S_{i^+}) \\
&= \text{swap}_{i,i^+} \ \wedge \ (W_i + S_i \le t < W_i + S_i + S_{i^+}).
\end{aligned}
$$

We now compute $\mathbf{P}\{D_{i,t}\}$. By similar reasoning to the above, the event $D_{i,t}$ occurs only if job i swaps with job i^-. If this swap occurs, then $T_i^{\text{Nudge}} = T_i^{\text{FCFS}} - S_{i^-}$. We again have $T_i^{\text{FCFS}} = W_i + S_i$, so

$$
\begin{aligned}
D_{i,t} &= \text{swap}_{i^-,i} \ \wedge \ (T_i^{\text{Nudge}} \le t < T_i^{\text{FCFS}}) \\
&= \text{swap}_{i^-,i} \ \wedge \ (T_i^{\text{FCFS}} - S_{i^-} \le t < T_i^{\text{FCFS}}) \\
&= \text{swap}_{i^-,i} \ \wedge \ (W_i + S_i - S_{i^-} \le t < W_i + S_i).
\end{aligned}
$$

To obtain (9), observe that conditioned on $\text{swap}_{i^-,i}$, we have $W_i = W_{i^-} + S_{i^-} - A_{i^-}$.   □

Now, we give sufficient conditions for Nudge to improve upon FCFS relative to a particular threshold t.

LEMMA 5.2 (STRICT IMPROVEMENT AT A GIVEN THRESHOLD). *Given any $t > s_{\min}$, where $s_{\min}$ is the smallest value of S,*

$$
\mathbf{P}\{T^{\text{Nudge}} > t\} < \mathbf{P}\{T^{\text{FCFS}} > t\}
$$

if the following inequality in terms of t holds:

$$
\frac{g_{\max}}{g_{\min}} \frac{\lambda + \theta^*}{\lambda} < \frac{\mathbf{E}\left[e^{\theta^* \min(t, S_{\text{small}} + S_{\text{large}})} - e^{\theta^* \min(t, S_{\text{small}})} \right]}{\mathbf{E}\left[e^{\theta^* \min(t, S_{\text{small}} + S_{\text{large}})} - e^{\theta^* \min(t, S_{\text{large}})} \right]}. \tag{10}
$$

PROOF. Because the tagged job i is a random sample arriving to the steady state of the system, by (6), we have $\mathbf{P}\{T^{\text{Nudge}} > t\} < \mathbf{P}\{T^{\text{FCFS}} > t\}$ if and only if $\mathbf{P}\{I_{i,t}\} < \mathbf{P}\{D_{i,t}\}$. Our approach is to use Lemma 5.1 to bound each of $\mathbf{P}\{I_{i,t}\}$ and $\mathbf{P}\{D_{i,t}\}$, from which the desired sufficient condition follows.

We begin by computing $\mathbf{P}\{I_{i,t}\}$:

$$
\begin{aligned}
\mathbf{P}\{I_{i,t}\} &= \mathbf{P}\{\text{swap}_{i,i^+} \ \wedge \ (W_i + S_i \le t < W_i + S_i + S_{i^+})\} && \text{[by Lemma 5.1]} \\
&= \mathbf{P}\{(A_i \le W_i) \wedge (W_i + S_i \le t < W_i + S_i + S_{i^+}) \wedge (x_2 \le S_i < x_3) \wedge (S_{i^+} < x_1)\} && \text{[by (7)]} \\
&\le \mathbf{P}\{(W_i + S_i \le t < W_i + S_i + S_{i^+}) \wedge (x_2 \le S_i < x_3) \wedge (S_{i^+} < x_1)\} && \text{[discarding } A_i \le W_i] \\
&= \mathbf{P}\left\{\left((t - S_i - S_{i^+})^+ \le W_i \le (t - S_i)^+\right) \wedge (x_2 \le S_i < x_3) \wedge (S_{i^+} < x_1)\right\} \\
&= p_{\text{large}} p_{\text{small}} \cdot \mathbf{E}_{S_i \sim S_{\text{large}}, S_{i^+} \sim S_{\text{small}}}\left[\int_{w=(t-S_i-S_{i^+})^+}^{(t-S_i)^+} f_{W_i}(w) \, \mathrm{d}w \right] && \begin{bmatrix} \text{change of measure for } S_i \text{ and } S_i^+, \\ \text{independence of } S_i, S_{i^+}, \text{ and } W_i \end{bmatrix} \\
&\le p_{\text{large}} p_{\text{small}} \cdot \mathbf{E}_{S_i \sim S_{\text{large}}, S_{i^+} \sim S_{\text{small}}}\left[\int_{w=(t-S_i-S_{i^+})^+}^{(t-S_i)^+} g_{\max} e^{-\theta^* w} \, \mathrm{d}w \right] && \begin{bmatrix} \text{by Lemma 3.1 and the} \\ \text{fact that } W_i \sim T_Q^{\text{FCFS}} \end{bmatrix} \\
&= p_{\text{large}} p_{\text{small}} \cdot \frac{e^{-\theta^* t}}{\theta^*} \cdot g_{\max} \mathbf{E}\left[e^{\theta^* \min(t, S_{\text{small}} + S_{\text{large}})} - e^{\theta^* \min(t, S_{\text{large}})} \right]. && (11)
\end{aligned}
$$

Similarly, we compute $\mathbf{P}\{D_{i,t}\}$:

$$
\begin{aligned}
\mathbf{P}\{D_{i,t}\} &= \mathbf{P}\{\text{swap}_{i^-,i} \ \wedge \ (W_{i^-} - A_{i^-} + S_i \leq t < W_{i^-} - A_{i^-} + S_i + S_{i^-})\} \\
&= \mathbf{P}\{(A_{i^-} \leq W_{i^-}) \wedge (W_{i^-} - A_{i^-} + S_i \leq t < W_{i^-} - A_{i^-} + S_i + S_{i^-}) \wedge (S_i < x_1) \wedge (x_2 \leq S_{i^-} < x_3)\} \\
&= \mathbf{P}\{(A_{i^-} + (t - S_i - S_{i^-})^+ \leq W_{i^-} \leq A_{i^-} + (t - S_i)^+) \ \wedge \ (S_i < x_1) \ \wedge \ (x_2 \leq S_{i^-} < x_3)\} \\
&= p_{\text{small}} p_{\text{large}} \cdot \mathbf{E}_{S_i \sim S_{\text{small}}, S_{i^-} \sim S_{\text{large}}, A_{i^-} \sim \text{Exp}(\lambda)} \left[\int_{w=A_{i^-}+(t-S_i-S_{i^-})^+}^{A_{i^-}+(t-S_i)^+} f_{W_{i^-}}(w)\, \mathrm{d}w \right] \begin{bmatrix} \text{mutual independence} \\ \text{of } S_i, S_{i^+}, A_{i^-}, \text{ and } W_i \end{bmatrix} \\
&\geq p_{\text{small}} p_{\text{large}} \cdot \mathbf{E}_{S_i \sim S_{\text{small}}, S_{i^-} \sim S_{\text{large}}, A_{i^-} \sim \text{Exp}(\lambda)} \left[\int_{w=A_{i^-}+(t-S_i-S_{i^-})^+}^{A_{i^-}+(t-S_i)^+} g_{\min} e^{-\theta^* w}\, \mathrm{d}w \right] \begin{bmatrix} \text{by Lemma 3.1 and the} \\ \text{fact that } W_i \sim T_Q^{\text{FCFS}} \end{bmatrix} \\
&\geq p_{\text{small}} p_{\text{large}} \cdot \mathbf{E}_{S_i \sim S_{\text{small}}, S_{i^-} \sim S_{\text{large}}} \left[\int_{a=0}^{\infty} \int_{w=a+(t-S_i-S_{i^-})^+}^{a+(t-S_i)^+} g_{\min} e^{-\theta^* w} \cdot \lambda e^{-\lambda a}\, \mathrm{d}w\, \mathrm{d}a \right] \quad [A_{i^-} \sim \text{Exp}(\lambda)] \\
&= p_{\text{small}} p_{\text{large}} \cdot \frac{e^{-\theta^* t}}{\theta^*} \cdot g_{\min} \frac{\lambda}{\lambda + \theta^*} \mathbf{E}\left[e^{\theta^* \min(t, S_{\text{small}} + S_{\text{large}})} - e^{\theta^* \min(t, S_{\text{small}})} \right]. \tag{12}
\end{aligned}
$$

Combining the bounds (11) and (12), we find that $\mathbf{P}\{I_{i,t}\} < \mathbf{P}\{D_{i,t}\}$ holds if

$$
g_{\max} \mathbf{E}\left[e^{\theta^* \min(t, S_{\text{small}} + S_{\text{large}})} - e^{\theta^* \min(t, S_{\text{large}})} \right] < g_{\min} \frac{\lambda}{\lambda + \theta^*} \mathbf{E}\left[e^{\theta^* \min(t, S_{\text{small}} + S_{\text{large}})} - e^{\theta^* \min(t, S_{\text{small}})} \right]. \tag{13}
$$

$\square$

Having proven Lemma 5.2, we have a sufficient condition, namely (10), for Nudge to improve upon FCFS at a specific value of t. But our goal is to improve upon FCFS for *all* values of t. We therefore seek a condition which implies that (10) holds for all t.

We start by simplifying (10). Let $A = e^{\theta^* S_{\text{small}}}$, $B = e^{\theta^* S_{\text{large}}}$, and $c = e^{\theta^* t}$. Then (10) becomes

$$
\frac{g_{\max}}{g_{\min}} \frac{\lambda + \theta^*}{\lambda} < \frac{\mathbf{E}[\min(AB, c) - \min(A, c)]}{\mathbf{E}[\min(AB, c) - \min(B, c)]}. \tag{14}
$$

Here the only appearance of the specific value of t is via c. Our strategy is to lower-bound the right-hand side of (14) by a quantity that does not include c. The following lemma, which we prove in Appendix A, helps accomplish this under an additional assumption.

LEMMA 5.3. *Let A, B be two independent real-valued random variables and c be a fixed constant. Suppose $1 \leq A \leq c$ and $A < B$. Under these assumptions, if $\mathbf{P}\{B > c\} > 0$ and $c\mathbf{E}[B] \geq \mathbf{E}[A]\,\mathbf{E}[B|B > c]$, then*

$$
\frac{\mathbf{E}[\min(AB, c) - \min(A, c)]}{\mathbf{E}[\min(AB, c) - \min(B, c)]} = \frac{\mathbf{E}[\min(AB, c) - A]}{\mathbf{E}[\min(AB, c) - \min(B, c)]} \geq \frac{\mathbf{E}[AB - A]}{\mathbf{E}[AB - B]}.
$$

5.2 Main Proof

THEOREM 4.1. *Suppose S is a continuous class I job size distribution. Then $\text{Nudge}(x_1, x_2, x_3)$ stochastically improves upon FCFS for any $s_{\min} < x_1 \leq x_2 \leq x_3$ satisfying*

$$
\bullet \quad \frac{g_{\max}}{g_{\min}} \frac{\lambda + \theta^*}{\lambda} < \frac{1 - \widetilde{S_{\text{large}}}(-\theta^*)^{-1}}{1 - \widetilde{S_{\text{small}}}(-\theta^*)^{-1}}, \tag{15}
$$

$$
\bullet \quad x_1 + x_3 \leq 2x_2. \tag{16}
$$

PROOF. We prove Theorem 4.1 by verifying the condition in Lemma 5.2. For every $t > s_{\min}$, we will show that Inequalities (15) and (16) together imply (10).

(i) When $s_{\min} < t \leq x_2$, the denominator of the right hand side of (10) becomes

$$\mathbf{E}\left[e^{\theta^* \min(t, S_{\text{small}} + S_{\text{large}})} - e^{\theta^* \min(t, S_{\text{large}})}\right] = \mathbf{E}\left[e^t - e^t\right] = 0.$$

Thus (10) is not well defined, but (13) holds trivially, which is sufficient.

(ii) When $x_2 < t < x_3$, we will apply Lemma 5.3 with $A = e^{\theta^* S_{\text{small}}}$, $B = e^{\theta^* S_{\text{large}}}$ and $c = e^{\theta^* t}$. Note that $1 \leq A \leq c$ and $A < B$. Now, let us verify that $c\mathbf{E}[B] \geq \mathbf{E}[A]\,\mathbf{E}[B|B > c]$. By (16), we know that

$$c\mathbf{E}[B] = e^{\theta^* t}\mathbf{E}\left[e^{\theta^* S_{\text{large}}}\right] \geq e^{\theta^* (2x_2)} \geq e^{\theta^* (x_1 + x_3)}.$$

Noting that $S_{\text{small}} \leq x_1$ and $S_{\text{large}} \leq x_3$, we have

$$c\mathbf{E}[B] \geq e^{\theta^* (x_1 + x_3)} \geq \mathbf{E}\left[e^{\theta^* S_{\text{small}}}\right]\mathbf{E}\left[e^{\theta^* S_{\text{large}}} | S_{\text{large}} > t\right] = \mathbf{E}[A]\,\mathbf{E}[B|B > c].$$

Thus, we can apply Lemma 5.3 to obtain

$$\frac{\mathbf{E}\left[e^{\theta^* \min(S_{\text{small}} + S_{\text{large}}, t)} - e^{\theta^* S_{\text{small}}}\right]}{\mathbf{E}\left[e^{\theta^* \min(S_{\text{small}} + S_{\text{large}}, t)} - e^{\theta^* \min(S_{\text{large}}, t)}\right]} \geq \frac{\mathbf{E}\left[e^{\theta^* (S_{\text{small}} + S_{\text{large}})} - e^{\theta^* S_{\text{small}}}\right]}{\mathbf{E}\left[e^{\theta^* (S_{\text{small}} + S_{\text{large}})} - e^{\theta^* S_{\text{large}}}\right]}. \tag{17}$$

Moreover, condition (15) implies that

$$\frac{g_{\max}}{g_{\min}} \frac{\lambda + \theta^*}{\lambda} < \frac{1 - \widetilde{S_{\text{large}}}(-\theta^*)^{-1}}{1 - \widetilde{S_{\text{small}}}(-\theta^*)^{-1}} = \frac{\mathbf{E}\left[e^{\theta^* (S_{\text{small}} + S_{\text{large}})} - e^{\theta^* S_{\text{small}}}\right]}{\mathbf{E}\left[e^{\theta^* (S_{\text{small}} + S_{\text{large}})} - e^{\theta^* S_{\text{large}}}\right]} \tag{18}$$

Combining (17) with (18) establishes (10).

(iii) When $t \geq x_3$, we have $\min(t, S_{\text{small}}) = S_{\text{small}}$ and $\min(t, S_{\text{large}}) = S_{\text{large}}$. Therefore, condition (15) is equivalent to (10).

Therefore, for every $t > s_{\min}$, we have proven that $\mathbf{P}\{T^{\text{Nudge}} > t\} < \mathbf{P}\{T^{\text{FCFS}} > t\}$.   □

6 PROOF OF THEOREM 4.2: EXISTENCE OF STOCHASTIC IMPROVEMENT

THEOREM 4.2. *For any continuous class I job size distribution S with $s_{\min} = 0$ and any load $0 < \rho < 1$, there exist x_1, x_2, x_3 satisfying (15) and (16), implying that $\text{Nudge}(x_1, x_2, x_3)$ stochastically improves upon FCFS.*

PROOF. We start by constructing x_1, x_2, x_3 that satisfy both Inequality (15) and (16). For notational convenience, let $M = \frac{g_{\max}}{g_{\min}} \frac{\lambda + \theta^*}{\lambda}$. First fix an arbitrary $x_3 > 0$ and let $x_2 = \frac{3}{4}x_3$, then compute $\widetilde{S_{\text{large}}}(-\theta^*)$ and choose a small enough x_1 such that

$$x_1 < \min\left\{-\frac{1}{\theta^*}\ln\left(1 - \frac{1 - \widetilde{S_{\text{large}}}(-\theta^*)^{-1}}{M}\right), \frac{1}{2}x_3\right\}. \tag{19}$$

Clearly, such $x_1 > s_{\min} = 0$ in (19) exists because $M > 1$, so we have

• $x_1 + x_3 < \frac{3}{2}x_3 = 2x_2,$ \hfill (20)

• $\dfrac{1 - \widetilde{S_{\text{large}}}(-\theta^*)^{-1}}{1 - \widetilde{S_{\text{small}}}(-\theta^*)^{-1}} = \dfrac{1 - \widetilde{S_{\text{large}}}(-\theta^*)^{-1}}{1 - \mathbf{E}\left[e^{\theta^* S_{\text{small}}}\right]^{-1}} \geq \dfrac{1 - \widetilde{S_{\text{large}}}(-\theta^*)^{-1}}{1 - e^{-\theta^* x_1}} > M = \dfrac{g_{\max}}{g_{\min}} \dfrac{\lambda + \theta^*}{\lambda}.$ \hfill (21)

By Theorem 4.1, (20) and (21) together imply that $\mathbf{P}\{T^{\text{Nudge}} \geq t\} < \mathbf{P}\{T^{\text{FCFS}} \geq t\}$ for every $t > s_{\min} = 0$. Therefore,

$$\forall t > 0, \quad \mathbf{P}\{T^{\text{Nudge}} > t\} < \mathbf{P}\{T^{\text{FCFS}} > t\}. \tag{□}$$

7 PROOF OF THEOREM 4.5: TRANSFORM OF RESPONSE TIME

In this section we compute an exact formula for $\widetilde{T^{\text{Nudge}}}(s)$. The formula holds for arbitrary job size distributions, not just those of class I.

At a high level, our analysis works by considering two systems experiencing identical arrivals: one using Nudge, and one using FCFS. We consider a tagged job arriving to this pair of systems in equilibrium and determine how its Nudge queueing time relates to its FCFS queueing time.

- *Small jobs:* Nudge queueing time is FCFS queueing time, possibly minus a large job's size.
- *Large jobs:* Nudge queueing time is FCFS queueing time, possibly plus a small job's size.
- *Other jobs:* Nudge queueing time is identical to FCFS queueing time.

We will determine $\widetilde{T_{Q,\text{small}}^{\text{Nudge}}}(s)$ and $\widetilde{T_{Q,\text{large}}^{\text{Nudge}}}(s)$, from which $\widetilde{T^{\text{Nudge}}}(s)$ easily follows.

7.1 Probabilistic Interpretation of the Laplace-Stieltjes Transform

Before jumping into the Nudge queueing time analysis, we recall a probabilistic interpretation of the Laplace-Stieltjes transform.

Let V be a nonnegative random variable. Consider a time interval of length V and a Poisson process of rate s that is independent of V. We call the increments of the Poisson process "interruptions". Let NoPoisson(V, s) be the event that there are no interruptions during the time interval. Then [16, Exercise 25.7]

$$\widetilde{V}(s) = \mathbf{P}\{\text{NoPoisson}(V, s)\}. \tag{22}$$

The interpretation in (22) necessarily requires $s \geq 0$. Fortunately, formulas we derive using (22) are still valid for $s < 0$ because Laplace transforms are uniquely defined by their value on any bounded interval on the real line [7].

7.2 Transform for Large Jobs

LEMMA 7.1. *The queueing time of large jobs under Nudge has Laplace-Stieltjes transform*

$$\widetilde{T_{Q,\text{large}}^{\text{Nudge}}}(s) = \left(1 - p_{\text{small}}(1 - \widetilde{S_{\text{small}}}(s))\right)\widetilde{T_Q^{\text{FCFS}}}(s) + p_{\text{small}}(1 - \widetilde{S_{\text{small}}}(s))\widetilde{T_Q^{\text{FCFS}}}(\lambda + s).$$

PROOF. Consider a large tagged job arriving to the pair of systems, one using Nudge and the other using FCFS, in equilibrium. We can think of the job's Nudge queueing time as the time it takes to do the following two steps:

(a) We first wait for its FCFS queueing time, namely T_Q^{FCFS}.

(b) If during that T_Q^{FCFS} time there has been at least one arrival, and if the first such arrival is a small job, we then wait for that small job's service, which takes S_{small} time. Note that the small job's size is independent of the FCFS queueing time.

We will use (22) to compute $\widetilde{T_{Q,\text{large}}^{\text{Nudge}}}(s)$. To that end, we associate each of the Nudge and FCFS systems with a Poisson "interruption" process of rate s. The interruption processes are independent of the arrival times and job sizes of each system, but they are coupled to each other in the following way. At any moment in time when the systems are busy, some job j has been in service for some amount of time t. We couple the interruption processes such that interruptions occur at the same (j, t) pairs in both systems.

By (22), $\widetilde{T_{Q,\text{large}}^{\text{Nudge}}}(s)$ is the probability that no interruptions occur during steps (a) and (b). We compute this probability by conditioning on the following event:

$$E = \left\{ \begin{array}{l} \text{the next arrival after the tagged job is small,} \\ \text{and an interruption occurs during its service} \end{array} \right\}$$

Note that E does not consider whether the next arrival occurs before the tagged job exits the queue. Therefore, it is independent of the length T_Q^{FCFS} of step (a).

If E does not occur, then there are no interruptions during step (b). Therefore, there are no interruptions if and only if there are no interruptions during step (a). By (22), this happens with probability $\widetilde{T_Q^{\mathrm{FCFS}}}(s)$.

If E does occur, then an interruption will occur during step (b) if and only if a new job arrives during step (a). That is, by conditioning on E, we have predetermined that the next arrival will be small and, if it swaps with the tagged job, will cause an interruption. Therefore, to avoid interruptions, we need to avoid interruptions *and arrivals* during step (a). Merging the arrival and interruption processes yields a Poisson process of rate $\lambda + s$, so avoiding interruptions corresponds to the event $\mathrm{NoPoisson}(T_Q^{\mathrm{FCFS}}, \lambda + s)$. By (22), this happens with probability $\widetilde{T_Q^{\mathrm{FCFS}}}(\lambda + s)$.

Conditioning on whether E occurs and using (22) to compute $\mathbf{P}\{E\} = p_{\mathrm{small}}(1 - \widetilde{S_{\mathrm{small}}}(s))$, we obtain the desired expression.   □

7.3   Transform for Small Jobs

LEMMA 7.2.   *The queueing time of small jobs under Nudge has Laplace-Stieltjes transform*

$$\widetilde{T_{Q,\mathrm{small}}^{\mathrm{Nudge}}}(s) = \widetilde{T_Q^{\mathrm{FCFS}}}(s)\left(1 + \frac{p_{\mathrm{large}}(1 - \widetilde{S_{\mathrm{large}}}(s))}{\widetilde{S}(s)}\right) - p_{\mathrm{large}}(1 - \widetilde{S_{\mathrm{large}}}(s))(1 - \rho) \cdot \frac{\lambda/\widetilde{S}(\lambda) - s/\widetilde{S}(s)}{\lambda - s}.$$

The analysis of small jobs is more involved than the analysis of large jobs. We therefore state and prove several more intermediate results before proving Lemma 7.2.

Consider a small tagged job arriving to the pair of systems, one using Nudge and the other using FCFS, in equilibrium. The main question we need to answer is whether the tagged job will swap with a large job in the Nudge system. Our main insight is that we can tell whether the swap will occur by examining just the FCFS system. Because we understand FCFS well, this makes it relatively simple to tell whether a swap will occur.

LEMMA 7.3.   *The small tagged job swaps with a large job in the Nudge system if and only if, when it arrives, the FCFS system has a nonempty queue whose last job is large.*

The proof of Lemma 7.3 follows very similar reasoning to our analysis at the start of Section 5.1. For completeness, we provide a proof in Appendix B.

Thanks to Lemma 7.3, we can determine the queueing time of the small tagged job by looking at the state of the FCFS system when it arrives. We describe the equilibrium state of the FCFS system with the following quantities:

- W: the amount of work in the system.
- N_Q: the number of jobs in the queue.
- W_{most}: the amount of work in the system, excluding the last job in the queue if $N_Q \geq 1$.
- S_{last}: the size of the last job in the queue, or 0 if $N_Q = 0$.

Note that these quantities are not independent. In particular, $W = W_{\mathrm{most}} + S_{\mathrm{last}}$. However, W_{most} and S_{last} are *conditionally* independent given $N_Q \geq 1$.

Armed with Lemma 7.3 and the system state notation, we are ready to compute $T_{Q,\mathrm{small}}^{\mathrm{Nudge}}(s)$, thus proving Lemma 7.2. Our computation will make use of an additional lemma which we state after the proof and prove in Appendix B.

Proof of Lemma 7.2. Consider the small tagged job arriving to the pair of systems in equilibrium. By Lemma 7.3, its Nudge queueing time is

$$T_{Q,\text{small}}^{\text{Nudge}} = \begin{cases} W_{\text{most}} & \text{if } N_Q = 0 \\ W_{\text{most}} + S_{\text{last}}\mathbf{1}(\neg(x_2 \leq S_{\text{last}} < x_3)) & \text{if } N_Q \geq 1. \end{cases}$$

Applying (22) and the conditional independence of W_{most} and S_{last} yields

$$\begin{aligned} \widetilde{T_{Q,\text{small}}^{\text{Nudge}}}(s) &= \mathbf{P}\big\{\text{NoPoisson}(W_{\text{most}}, s) \,\wedge\, N_Q = 0\big\} \\ &\quad + \mathbf{P}\big\{\text{NoPoisson}(W_{\text{most}}, s) \,\wedge\, N_Q \geq 1 \,\wedge\, (\text{NoPoisson}(S_{\text{last}}, s) \,\vee\, x_2 \leq S_{\text{last}} < x_3)\big\} \\ &= \mathbf{P}\big\{\text{NoPoisson}(W_{\text{most}}, s) \,\wedge\, N_Q = 0\big\} \\ &\quad + \mathbf{P}\big\{\text{NoPoisson}(W_{\text{most}}, s) \,\wedge\, N_Q \geq 1\big\} \cdot \big(\widetilde{S}(s) + p_{\text{large}}(1 - \widetilde{S_{\text{large}}}(s))\big). \end{aligned} \tag{23}$$

It remains only to compute the two probabilities in (23). Let

$$q \triangleq \mathbf{P}\big\{\text{NoPoisson}(W, s) \,\wedge\, N_Q = 0\big\} = \mathbf{P}\big\{\text{NoPoisson}(W_{\text{most}}, s) \,\wedge\, N_Q = 0\big\}, \tag{24}$$

making q the first probability in (23). We now compute the second probability in (23) in terms of q. First, note that T_Q^{FCFS} and W are identically distributed. Recalling the conditional independence of W_{most} and S_{last}, we have, using (22) throughout,

$$\begin{aligned} \widetilde{T_Q^{\text{FCFS}}}(s) - q &= \widetilde{W}(s) - q = \mathbf{P}\big\{\text{NoPoisson}(W, s) \,\wedge\, N_Q \geq 1\big\} \\ &= \mathbf{P}\big\{\text{NoPoisson}(W_{\text{most}}, s) \,\wedge\, N_Q \geq 1 \,\wedge\, \text{NoPoisson}(S_{\text{last}}, s)\big\} \\ &= \mathbf{P}\big\{\text{NoPoisson}(W_{\text{most}}, s) \,\wedge\, N_Q \geq 1\big\} \cdot \widetilde{S}(s). \end{aligned} \tag{25}$$

Plugging (24) and (25) into (23) yields

$$\widetilde{T_{Q,\text{small}}^{\text{Nudge}}}(s) = q + \big(\widetilde{T_Q^{\text{FCFS}}}(s) - q\big)\left(1 + \frac{p_{\text{large}}(1 - \widetilde{S_{\text{large}}}(s))}{\widetilde{S}(s)}\right).$$

Lemma 7.4 below, proven in Appendix B, computes the value of q, yielding the desired result. □

Lemma 7.4. Let $q \triangleq \mathbf{P}\big\{\text{NoPoisson}(W, s) \,\wedge\, N_Q = 0\big\}$. We have

$$q = \widetilde{T_Q^{\text{FCFS}}}(\lambda) \cdot \frac{\lambda\widetilde{S}(s) - s\widetilde{S}(\lambda)}{\lambda - s} = \frac{1 - \rho}{\widetilde{S}(\lambda)} \cdot \frac{\lambda\widetilde{S}(s) - s\widetilde{S}(\lambda)}{\lambda - s}.$$

7.4 Overall Response Time Transform

Theorem 4.5. *The response time of Nudge has Laplace-Stieltjes transform*

$$\begin{aligned} \widetilde{T^{\text{Nudge}}}(s) = \widetilde{T^{\text{FCFS}}}(s) &+ p_{\text{small}}p_{\text{large}}\left(\widetilde{S_{\text{large}}}(s)(1 - \widetilde{S_{\text{small}}}(s))\left(\widetilde{T_Q^{\text{FCFS}}}(\lambda + s) - \widetilde{T_Q^{\text{FCFS}}}(s)\right) \right. \\ &\left. + \widetilde{S_{\text{small}}}(s)(1 - \widetilde{S_{\text{large}}}(s))\left(\frac{\widetilde{T_Q^{\text{FCFS}}}(s)}{\widetilde{S}(s)} - (1 - \rho)\frac{\lambda/\widetilde{S}(\lambda) - s/\widetilde{S}(s)}{\lambda - s}\right)\right). \end{aligned}$$

Proof. The expression follows by plugging the results of Lemmas 7.1 and 7.2 into

$$\widetilde{T^{\mathsf{Nudge}}}(s) = p_{\mathsf{small}} \cdot \widetilde{T_{Q,\mathsf{small}}^{\mathsf{Nudge}}}(s) \cdot \widetilde{S_{\mathsf{small}}}(s) + p_{\mathsf{large}} \cdot \widetilde{T_{Q,\mathsf{large}}^{\mathsf{Nudge}}}(s) \cdot \widetilde{S_{\mathsf{large}}}(s)$$
$$+ (1 - p_{\mathsf{small}} - p_{\mathsf{large}}) \cdot \widetilde{T_Q^{\mathsf{FCFS}}}(s) \cdot \frac{\widetilde{S}(s) - p_{\mathsf{small}}\widetilde{S_{\mathsf{small}}}(s) - p_{\mathsf{large}}\widetilde{S_{\mathsf{large}}}(s)}{1 - p_{\mathsf{small}} - p_{\mathsf{large}}}$$

and simplifying the resulting expression. One key step is recognizing $\widetilde{T^{\mathsf{FCFS}}}(s) = \widetilde{T_Q^{\mathsf{FCFS}}}(s) \cdot \widetilde{S}(s)$. □

8   PROOF OF THEOREM 4.3: ASYMPTOTIC IMPROVEMENT

Theorem 4.3. *Suppose S is a continuous class I job size distribution. For any $s_{\min} < x_1 \leq x_2 \leq x_3$, the asymptotic tail improvement ratio of* Nudge(x_1, x_2, x_3) *compared to* FCFS *is*

$$\mathsf{AsymTIR} = p_{\mathsf{small}} p_{\mathsf{large}} \frac{\lambda}{\lambda + \theta^*} \left(\widetilde{S_{\mathsf{large}}}(-\theta^*) - \frac{\lambda}{\lambda + \theta^*} \widetilde{S_{\mathsf{small}}}(-\theta^*) - \frac{\theta^*}{\lambda + \theta^*} \widetilde{S_{\mathsf{large}}}(-\theta^*)\widetilde{S_{\mathsf{small}}}(-\theta^*) \right).$$

Furthermore, AsymTIR *is positive, meaning* $C_{\mathsf{Nudge}} < C_{\mathsf{FCFS}}$, *if*

$$\frac{\lambda + \theta^*}{\lambda} < \frac{1 - \widetilde{S_{\mathsf{large}}}(-\theta^*)^{-1}}{1 - \widetilde{S_{\mathsf{small}}}(-\theta^*)^{-1}}.$$

Below we give a high-level overview of the proof. We provide full computations in Appendix B.

Proof sketch. Using the Final Value Theorem, one can show that for Alg $\in \{\mathsf{Nudge}, \mathsf{FCFS}\}$,

$$C_{\mathsf{Alg}} = \frac{1}{\theta^*} \lim_{s \to 0} s\widetilde{T^{\mathsf{Alg}}}(s - \theta^*).$$

Combining this with Theorem 4.5, which relates $\widetilde{T^{\mathsf{Nudge}}}(s)$ to $\widetilde{T^{\mathsf{FCFS}}}(s)$, will relate C_{Nudge} to C_{FCFS}.

To obtain AsymTIR $= 1 - C_{\mathsf{Nudge}}/C_{\mathsf{FCFS}}$, we compute $\lim_{s \to 0}(s\widetilde{T^{\mathsf{FCFS}}}(s - \theta^*) - s\widetilde{T^{\mathsf{Nudge}}}(s - \theta^*))$ via Theorem 4.5. Each non-vanishing term has an $sT_Q^{\mathsf{FCFS}}(s - \theta^*)$ factor. Because $\widetilde{T_Q^{\mathsf{FCFS}}}(s) = \widetilde{T^{\mathsf{FCFS}}}(s)/\widetilde{S}(s)$, we can express $C_{\mathsf{FCFS}} - C_{\mathsf{Nudge}}$ as a constant times C_{FCFS}; this constant is AsymTIR. □

9   EMPIRICAL LESSONS

This paper proves that Nudge stochastically improves upon FCFS under the correct choice of parameters, and achieves multiplicative improvement in the asymptotic tail. However, there are a few practical questions remaining. These questions center around finding Nudge parameters in practice. In this section, we demonstrate several practical lessons on choosing Nudge parameters.

(1) (Section 9.1) We find that Nudge typically achieves its greatest improvement over FCFS when the Nudge parameters specify that all jobs are either large or small (i.e. $x_1 = x_2$, $x_3 = \infty$).
(2) (Section 9.2) We find that when load is low, Nudge can dramatically improve upon FCFS (10-20%) in the common case where job size variability is relatively high (i.e. $C^2 > 1$). When job size variability is lower and load is low, improvement is smaller.
(3) (Section 9.3) We find that the space of parameters that lead Nudge to *asymptotically* improve upon FCFS typically also cause Nudge to *stochastically* improve upon FCFS. This is serendipitous, because Theorem 4.3 provides a simple, exact method to check whether given Nudge parameters will achieve asymptotic improvement over FCFS.
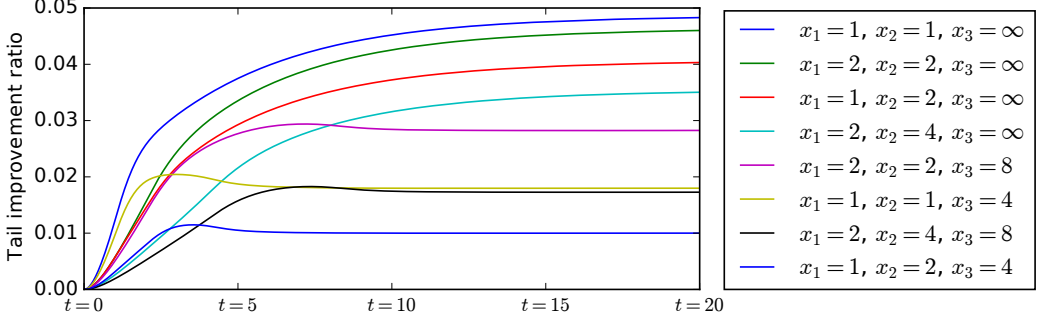
Fig. 3. Empirical tail improvement of Nudge over FCFS under a variety of Nudge parameter choices. Highest improvement occurs when $x_1 = x_2$, $x_3 = \infty$. Job size distribution is hyperexponential with branches drawn from $\text{Exp}(2)$ and $\text{Exp}(1/3)$, where the first branch has probability 0.8. $\text{E}[S] = 1$, $C^2 = 3$. Simulations run with 10 billion arrivals. Load $\rho = 0.8$. Parameter choices are listed in order of asymptotic improvement.

9.1 All Jobs Should Be Either Large or Small

When evaluating Nudge on common job size distributions, we have found that the greatest improvement over FCFS is achieved by setting the Nudge parameters such that all jobs are either large or small (i.e. $x_1 = x_2$, $x_3 = \infty$), with no medium or very large jobs. This is a pattern we have seen with great consistency across a variety of job size distributions.

In Fig. 3, we show one instance of this pattern, for the case of a particular hyperexponential distribution. We see that the two choices of Nudge parameters that display the *least* improvement over FCFS are those where both medium and very large jobs exist, i.e. $x_1 \neq x_2$ and $x_3 \neq \infty$.

To explain this phenomenon, note that when we remove medium and very large jobs, we end up with more swaps. Empirically, we have found that the *quantity* of swaps is more important than the *quality* of those swaps, and thus maximizing the number of swaps leads to the largest improvement. While empirically removing medium and very large jobs improves performance, our analytical result in Theorem 4.2 requires medium and very large jobs.

Setting the Nudge parameters so that all jobs are either large or small dramatically simplifies the problem of choosing Nudge parameters, in addition to achieving consistently strong performance. Now, only *one free parameter* remains: the cutoff between small and large jobs.

9.2 Low Load: Dramatic Improvement when Variability is High

At low load, Nudge has the potential for dramatic improvement over FCFS (>10% throughout the tail), in the common case where the job size distribution is more variable than an exponential distribution, i.e. $C^2 > 1$. On the other hand, under low-variability job size distributions ($C^2 < 1$), we find that Nudge's improvement shrinks at lower loads; here it helps to set the x_1 cutoff close to 0.

In Fig. 4 we show these patterns for a wide variety of distributions at relatively low load $\rho = 0.4$. In Fig. 4(a), we have four high-variance job size distributions, each with $C^2 \in [2, 3.33]$. In every case, Nudge dramatically improves upon FCFS, with TIR in the range of 10-15%. In Fig. 4(b), we have four low-variance job size distributions, each with $C^2 \in [1/5, 1/2]$. In these cases, we reduce the cutoff x_1 to 0.2 for best performance, and Nudge's improvement over FCFS is under 3%.

Intuitively, when load is low, each job waits behind fewer other jobs on average, so Nudge's one swap per job has a greater relative impact. When those swaps are broadly beneficial for the tail, as occurs when job size variance is high, Nudge achieves the most dramatic improvement over FCFS. When job size variance is low, swaps involving small jobs that are near the mean job size cause the
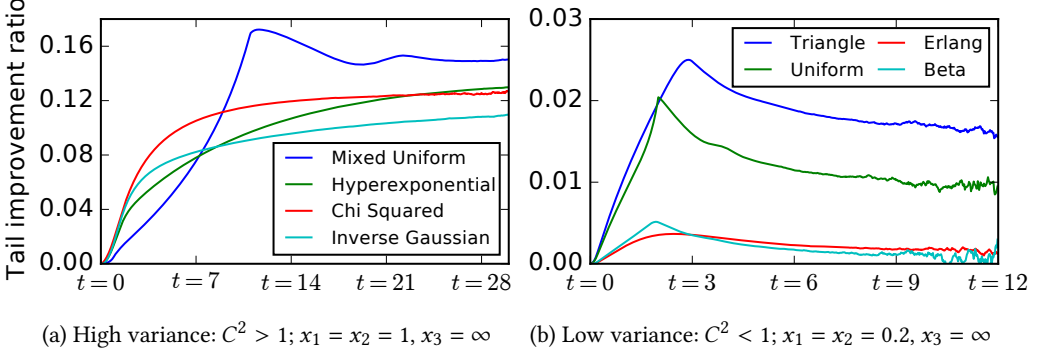
(a) High variance: $C^2 > 1$; $x_1 = x_2 = 1$, $x_3 = \infty$ (b) Low variance: $C^2 < 1$; $x_1 = x_2 = 0.2$, $x_3 = \infty$

Fig. 4. Empirical tail improvement of Nudge over FCFS at low load ($\rho = 0.4$) under a variety of job size distributions with $\mathbf{E}[S] = 1$. (a) Higher variance distributions show dramatic improvement. (b) Lower variance distributions show modest improvement. Specific distributions: In (a), Mixed Uniform: Uniform$(0, 1)$ w/prob. 0.9, else Uniform$(0, 11)$, $C^2 = 3.33$; Hyperexponential: Exp(2) w/ prob. 0.8, else Exp$(1/3)$, $C^2 = 3$; ChiSquared(1), $C^2 = 2$; InverseGaussian$(\mu = 1, \lambda = 1/2)$, $C^2 = 2$. In (b), Triangle w/ min $= 0$, mode $= 0$, max $= 3$, $C^2 = 1/2$; Uniform$(0, 2)$, $C^2 = 1/3$; Erlang$(k = 3, \lambda = 1/3$, $C^2 = 1/3$; Beta$(\alpha = 2, \beta = 2)$ scaled by a factor of 2, $C^2 = 1/5$. Distributions listed in order of asymptotic improvement. Simulations run with 10 billion arrivals.

Job size dist.	x_1	Asym.	Stoc.	x_1	Asym.	Stoc.	x_1	Asym.	Stoc.	x_1	Asym.	Stoc.	x_1	Asym.	Stoc.
Exponential	0.5	✓	✓	1	✓	✓	2	✗	✗	4	✗	✗	8	✗	✗
Hyperexponential	0.5	✓	✓	1	✓	✓	2	✓	✓	4	✓	✓	8	✓	✓
Bounded Lomax	0.5	✓	✓	1	✓	✓	1.5	✗	✗	2	✗	✗	3	✗	✗
Uniform	0.1	✓	✓	0.2	✓	✓	0.5	✗	✗	0.75	✗	✗	1	✗	✗
Beta	0.1	✓	✓	0.2	✓	✓	0.3	✗	✗	0.4	✗	✗	0.5	✗	✗

Table 1. Presence or absence of asymptotic and stochastic improvement of Nudge over FCFS under a variety of job size distributions and Nudge parameter choices. Stochastic improvement occurs whenever asymptotic improvement occurs. Each row gives a distinct job size distribution, and each cell gives a distinct Nudge parameter setting. In every case, $x_2 = x_1$ and $x_3 = \infty$, so only x_1 is specified. Load $\rho = 0.4$. Specific job size distributions, each with mean 1: Exponential with mean 1, Uniform$(0, 2)$, Hyperexponential drawn from Exp(2) w/ prob. 0.8 and Exp$(1/3)$ with prob. 0.2, BoundedLomax$(\lambda = 2, \text{max} = 4, \alpha = 2)$, Beta$(\alpha = 2, \beta = 2)$ scaled by a factor of 2.

response time of the large jobs to suffer too much. To alleviate this, we reduce the small job cutoff x_1 to maintain stochastic improvement over FCFS.

9.3 Asymptotic Improvement Means Stochastic Improvement

After extensively simulating Nudge under different loads and job size distributions, we have found that the space of parameters under which Nudge asymptotically improves upon FCFS typically matches the space under which Nudge stochastically improves upon FCFS.

In Table 1, we show the consistency of this relationship across a wide variety of job size distributions and choices of Nudge parameters. The distributions range from a low-variance Beta distribution with $C^2 = 1/5$ to a hyperexponential distribution with $C^2 = 3$. Across the spectrum, Nudge stochastically improves over FCFS whenever it asymptotically improves over FCFS.

This connection between asymptotic and stochastic improvement is surprising given that the conditions that we need to prove stochastic improvement (Theorem 4.1) are much more stringent

than what we need to prove asymptotic improvement (Theorem 4.3). Nonetheless, the connection is highly useful because we have provided a simple analytical formula for determining when Nudge asymptotically improves upon FCFS (Theorem 4.3).

10 NUDGE IN PRACTICE

Nudge can be used in practice even if some of the assumptions made in this paper are not perfectly satisfied.

In this paper, we assume that exact job size information is known to the scheduler. However, such information is only used to determine which size class (small, large, etc.) a job should be placed in. In practice, only estimates of job size may be known. In such a setting, the scheduler could assign jobs that are clearly above or below a size threshold to the large and small classes, while placing ambiguous jobs in the medium class. If the estimates are reasonably accurate, we would expect such a Nudge policy to stochastically improve upon FCFS.

We also assume that the exact job size distribution is known to the scheduler. This assumption is needed to choose the Nudge parameters for our proofs in Section 4. However, our empirical results in Section 9 show that much less information is needed in practice to choose good Nudge parameters. For instance, as we saw in Section 9.2, the following choice of parameters works well empirically: By default, set $x_1 = x_2 = E[S]$, $x_3 = \infty$. However, if load is low and job size variability (C^2) is low, set $x_1 = x_2 = E[S]/5$, $x_3 = \infty$.

11 VARIANTS ON NUDGE

As Nudge is such a simple policy, there are many interesting variants of Nudge that one could investigate. We now discuss the advisability of several such variants.

Recall that Nudge only ever swaps a job at most once. One might consider allowing a job to swap a second or third time with new arrivals, or even an unlimited number of times. Unfortunately, this change could ruin Nudge's stochastic improvement over FCFS, if implemented poorly. In particular, under a Nudge variant where large jobs can be swapped with an unlimited number of small arrivals, such highly-swapped large jobs will typically dominate the response time tail, dramatically worsening the variant's tail performance. A wiser variant might be to allow large jobs to be swapped with a bounded number of small jobs, or to allow only the small jobs to be swapped an unlimited number of times.

Another interesting variant of Nudge would only swap in a probabilistic fashion, such as with an i.i.d. coin flip. We believe such a policy could achieve stochastic improvement over FCFS. However, proving such a result would be no simpler than for Nudge, because probabilistic swapping does not change the shape of the distribution of swaps. Moreover, the variant's tail improvement ratios would likely be smaller than those of Nudge, because a smaller fraction of jobs are involved in swaps.

Finally, one could design a more complicated variant of Nudge which would consider a job's exact size when deciding whether to swap, rather than simply comparing the job's size to a threshold. For instance, one might decide to swap all pairs of jobs whose sizes differ by a factor of 2, as long as neither job has yet been swapped. These more complicated Nudge variants might achieve even larger stochastic improvements over FCFS than Nudge. Beyond FCFS, such Nudge variants might be able to stochastically improve upon some or even all Nudge policies. We leave this possibility as an open question.

12 CONCLUSION

We introduce Nudge, the first scheduling policy whose response time distribution stochastically improves upon that of FCFS. Specifically, we prove that with appropriately chosen parameters,

Nudge stochastically improves upon FCFS for light-tailed job size distributions[6]. From an asymptotic viewpoint, we prove that Nudge achieves a multiplicative improvement over FCFS, disproving the strong asymptotic optimality conjecture for FCFS. Finally, we derive the Laplace-Stieltjes transform of response time under Nudge, using a novel technique. Nudge is simple to implement and is a practical drop-in replacement for FCFS when job sizes are known.

One of the major insights of this paper is that improving the tail does not follow the same intuitions that we use in improving the mean. While improving mean response time is often a matter of helping small jobs jump ahead of large ones, when it comes to the tail, this has to be done in a very measured way. Too much help to the small jobs causes the tail to get a lot worse. Nudge finds the exactly appropriate way to do this.

One direction for future work is further exploring the stochastic improvement frontier. Can we stochastically improve upon other commonly used scheduling policies? Can we improve upon Nudge itself, such as with a more complicated variant of Nudge (see Section 11)? One policy which cannot be stochastically improved upon is SRPT, due to its optimal mean response time. Can we prove that other policies are unimprovable?

Another direction is simplifying the definition of Nudge. Our empirical results in Section 9 indicate that in practice, Nudge can always stochastically improve upon FCFS with only two classes of jobs: small and large. It would be of practical importance to figure out how to extend the theorems in this paper to hold for this simplified definition of Nudge.

ACKNOWLEDGEMENTS

REFERENCES

[1] Joseph Abate, Gagan L Choudhury, and Ward Whitt. 1994. Waiting-time tail probabilities in queues with long-tail service-time distributions. *Queueing systems* 16, 3-4 (1994), 311–338.
[2] Joseph Abate, Gagan L. Choudhury, and Ward Whitt. 1995. Exponential Approximations for Tail Probabilities in Queues, I: Waiting Times. 43, 5 (1995), 885–901.
[3] Joseph Abate and Ward Whitt. 1997. Asymptotics for M/G/1 low-priority waiting-time tail probabilities. *Queueing Systems* 25, 1-4 (1997), 173–233.
[4] John T. Blake and Michael W. Carter. 1996. An analysis of emergency room wait time issues via computer simulation. *INFOR* 34, 4 (November 1996), 263–273.
[5] Onno Boxma and Bert Zwart. 2007. Tails in Scheduling. *ACM SIGMETRICS Performance Evaluation Review* 34, 4 (March 2007), 13–20.
[6] Percy H Brill. 2000. A brief outline of the level crossing method in stochastic models. *CORS Bulletin* 34, 4 (2000), 9–21.
[7] Patrick Chareka. 2007. A Finite-Interval Uniqueness Theorem for Bilateral Laplace Transforms. *International Journal of Mathematics and Mathematical Sciences* 2007 (2007), 6 pages.
[8] Y. Chen, S. Iyer, X. Liu, D. Milojicic, and A. Sahai. 2007. SLA Decomposition: Translating Service Level Objectives to System Level Thresholds. In *Fourth International Conference on Autonomic Computing (ICAC'07)*. 10 pages.
[9] Mark E. Crovella, Murad S. Taqqu, and Azer Bestavros. 1998. Heavy-Tailed Probability Distributions in the World Wide Web. In *A Practical Guide To Heavy Tails*. Chapman & Hall, New York, Chapter 1, 1–23.
[10] Mark M. Davis. 1991. How Long Should a Customer Wait for Service? *Decision Sciences* 22, 2 (1991), 421–434.
[11] Eric J. Friedman and Shane G. Henderson. 2003. Fairness and Efficiency in Web Server Protocols. In *Proceedings of the 2003 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems* (San Diego, CA, USA) *(SIGMETRICS '03)*. Association for Computing Machinery, New York, NY, USA, 229–237.
[12] Eric J Friedman and Gavin Hurley. 2003. *Protective scheduling*. Technical Report. Cornell University Operations Research and Industrial Engineering.
[13] S. W. Fuhrmann and Robert B. Cooper. 1985. Stochastic Decompositions in the M/G/1 Queue with Generalized Vacations. *Operations Research* 33, 5 (Oct. 1985), 1117–1129.

---

[6]More specifically, continuous class I job size distributions with positive density at 0.

[14] Mor Harchol-Balter. 1999. The Effect of Heavy-Tailed Job Size Distributions on Computer System Design. In *Proceedings of the ASA-IMS Conference on Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics.* Washington, DC.

[15] Mor Harchol-Balter. 2002. Task Assignment with Unknown Duration. *J. ACM* 49, 2 (March 2002), 260–288.

[16] Mor Harchol-Balter. 2013. *Performance Modeling and Design of Computer Systems: Queueing Theory in Action.* Cambridge University Press, Cambridge.

[17] Mor Harchol-Balter. 2021. Open poblems in queueing theory inspired by datacenter computing. *Queueing Systems: Theory and Applications* 97, 1 (2021), 3–37.

[18] Mor Harchol-Balter, Mark Crovella, and Cristina Murta. 1999. On Choosing a Task Assignment Policy for a Distributed Server System. *IEEE Journal of Parallel and Distributed Computing* 59 (1999), 204–228.

[19] Leora I. Horwitz, Jeremy Green, and Elizabeth H. Bradley. 2010. US Emergency Department Performance on Wait Time and Length of Visit. *Annals of Emergency Medicine* 55, 2 (2010), 133 – 141.

[20] John F. C. Kingman. 1962. On Queues in Which Customers Are Served in Random Order. *Mathematical Proceedings of the Cambridge Philosophical Society* 58, 1 (Jan. 1962), 79–91.

[21] J. F. C. Kingman. 1964. A martingale inequality in the theory of queues. *Mathematical Proceedings of the Cambridge Philosophical Society* 60, 2 (1964), 359–361.

[22] J. F. C. Kingman. 1970. Inequalities in the Theory of Queues. *Journal of the Royal Statistical Society: Series B (Methodological)* 32, 1 (1970), 102–110.

[23] Leonard Kleinrock. 1976. *Queueing Systems, Volume 2: Computer Applications.* Wiley, New York, NY.

[24] Jeffrey C. Mogul and John Wilkes. 2019. Nines are not enough: Meaningful metrics for clouds. In *Proceedings of the Workshop on Hot Topics in Operating Systems (HotOS19).* USA, 136 – 141.

[25] Jayakrishnan Nair, Adam Wierman, and Bert Zwart. 2010. Tail-robust scheduling via limited processor sharing. *Performance Evaluation* 67, 11 (2010), 978 – 995.

[26] Misja Nuyens, Adam Wierman, and Bert Zwart. 2008. Preventing Large Sojourn Times Using SMART Scheduling. *Operations Research* 56, 1 (2008), 88–101.

[27] Natalia Osipova, Urtzi Ayesta, and Konstantin Avrachenkov. 2009. Optimal Policy for Multi-Class Scheduling in a Single Server Queue. In *2009 21st International Teletraffic Congress.* IEEE, Paris, France, 1–8.

[28] T. Sakurai. 2004. Approximating M/G/1 Waiting Time Tail Probabilities. *Stochastic Models* 20, 2 (2004), 173–191.

[29] Linus Schrage. 1968. A proof of the optimality of the shortest remaining processing time discipline. *Operations Research* 16 (1968), 687–690.

[30] Linus E. Schrage. 1967. The Queue M/G/1 with Feedback to Lower Priority Queues. *Management Science* 13, 7 (March 1967), 466–474.

[31] Linus E. Schrage and Louis W. Miller. 1966. The Queue M/G/1 with the Shortest Remaining Processing Time Discipline. *Operations Research* 14, 4 (Aug. 1966), 670–684.

[32] Ziv Scully, Mor Harchol-Balter, and Alan Scheller-Wolf. 2018. SOAP: One Clean Analysis of All Age-Based Scheduling Policies. *Proc. ACM Meas. Anal. Comput. Syst.* 2, 1, Article 16 (April 2018), 30 pages.

[33] Ziv Scully, Lucas van Kreveld, Onno J. Boxma, Jan-Pieter Dorsman, and Adam Wierman. 2020. Characterizing Policies with Optimal Response Time Tails under Heavy-Tailed Job Sizes. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 4, 2, Article 30 (June 2020), 33 pages.

[34] Kut C. So and Jing-Sheng Song. 1998. Price, delivery time guarantees and capacity selection. *European Journal of Operational Research* 111, 1 (1998), 28 – 49.

[35] David A Stanford, Peter Taylor, and Ilze Ziedins. 2014. Waiting time distributions in the accumulating priority queue. *Queueing Systems* 77, 3 (2014), 297–330.

[36] Alexander L. Stolyar and Kavita Ramanan. 2001. Largest Weighted Delay First Scheduling: Large Deviations and Optimality. *Annals of Applied Probability* 11, 1 (2001), 1–48.

[37] Muhammad Tirmazi, Adam Barker, Nan Deng, MD E. Haque, Zhijing Gene Qin, Steven Hand, Mor Harchol-Balter, and John Wilkes. 2020. Borg: The next generation. In *Proceedings of the Fifteenth European Conference on Computer Systems (EuroSys '20).* Greece, 1–14.

[38] Timothy L. Urban. 2009. Establishing delivery guarantee policies. *European Journal of Operational Research* 196, 3 (2009), 959 – 967.

[39] Peter D. Welch. 1964. On a Generalized M/G/1 Queuing Process in Which the First Customer of Each Busy Period Receives Exceptional Service. *Operations Research* 12, 5 (Oct. 1964), 736–752.

[40] Adam Wierman and Bert Zwart. 2012. Is Tail-Optimal Scheduling Possible? *Operations Research* 60, 5 (Oct. 2012), 1249–1257.

[41] Ronald W. Wolff. 1982. Poisson Arrivals See Time Averages. *Operations Research* 30, 2 (1982), 223–231.

A   PROOFS FOR STOCHASTIC IMPROVEMENT

LEMMA 3.1. *Suppose S is a continuous class I job size distribution. For any load ρ, the normalized p.d.f. $g(t)$ is bounded above and below by positive constants, and $\lim_{t \to \infty} g(t)$ exists.*

PROOF. First we show (following prior work [1, 3, 28]) that $\widetilde{T_Q^{\text{FCFS}}}$ has a simple pole $-\theta^*$ as its rightmost singularity.

We let $-\theta^*$ be the root of the denominator of $\widetilde{T_Q^{\text{FCFS}}}(s)$, which is

$$\lambda \widetilde{S}(s) - \lambda + s = 0 \iff \widetilde{S}(s) = \frac{\lambda - s}{\lambda}.$$

Since the left hand $\widetilde{S}(s)$ is convex in s[7], and the right hand $\frac{\lambda-s}{\lambda}$ is only linear in s, their intersection $s = -\theta^*$ must be a simple root. Moreover, such an intersection exists for $s < 0$ because

- $\widetilde{S}(0) = \frac{\lambda-0}{\lambda} = 1$;
- $\widetilde{S}'(0) = -\frac{1}{\mu} > -\frac{1}{\lambda}$;
- $\widetilde{S}(s) \to \infty$ when s approaches the rightmost singularity of $\widetilde{S}$ (since S is a class I distribution).

Now we use final value theorem to establish the limit of the ratio between the p.d.f. $f_{T_Q^{\text{FCFS}}}$ and the exponential function $e^{-\theta^* t}$. Recall the function $g(t) = f_{T_Q^{\text{FCFS}}}(e)e^{\theta^* t}$ and consider its Laplace transform $\widetilde{G}(s) = \widetilde{T_Q^{\text{FCFS}}}(s - \theta^*)$. Since the poles of $\widetilde{G}(s)$ map one-to-one to the poles of $\widetilde{T_Q^{\text{FCFS}}}(s - \theta^*)$, the above arguments show that every pole of $\widetilde{G}(s)$ is either in the open left half plane or at the origin, and the origin is a simple pole. Therefore, the Final Value Theorem for $g(t)$ tells us

$$\lim_{t \to \infty} f_{T_Q}^{\text{FCFS}}(t)e^{\theta^* t} = \lim_{t \to \infty} g(t) = \lim_{s \to 0} s\widetilde{G}(s) = \frac{(1-\rho)\theta^*}{-\lambda\widetilde{S}'(-\theta^*) - 1} \triangleq g^* > 0. \tag{26}$$

Since the limit g^* exists, for $\epsilon = \frac{g^*}{2}$, there exists $N_\epsilon < \infty$ such that $\forall t \geq N_\epsilon$,

$$|g(t) - g^*| \leq \frac{g^*}{2} \implies \frac{g^*}{2}e^{-\theta^* t} \leq f_{T_Q}^{\text{FCFS}}(t) \leq \frac{3g^*}{2}e^{-\theta^* t}.$$

Next, we want to show that

$$\min_{0 \leq t < N_\epsilon} g(t) > 0 \quad \text{and} \quad \max_{0 \leq t < N_\epsilon} g(t) < \infty. \tag{27}$$

First, note that $f_{T_Q}^{\text{FCFS}}$ satisfies the following level-crossing differential equations [6] (we abbreviate $f_{T_Q}^{\text{FCFS}}$ to f):

$$\begin{cases} f(0) = (1-\rho)\lambda; \\ f'(t) = \lambda f(t) - (1-\rho)\lambda f_S(t) - \lambda \int_{j=0}^{t} f(t-j)f_S(j)\,\mathrm{d}j. \end{cases}$$

To begin with, $f(t)$ is continuous because $f_S(t)$ is continuous. If $f(t) = 0$ for some $t < N_\epsilon$, we let $t_0 = \inf\{t < N_\epsilon : f(t) = 0\}$. Clearly $t_0 > 0$ because $f(0) = (1-\rho)\lambda > 0$. Note also that $f(t_0) = 0$, because f is continuous. Since $f(t_0) < f(0)$, $\exists 0 < s_0 < t_0$ s.t. $f_S(s_0) > 0$. Then $\exists a, b$ where $0 \leq a < s_0 < b \leq t_0$, s.t. $f(t) > 0$ for all $t \in [a, b]$. Now we have

$$f'(t_0) = -(1-\rho)\lambda f_S(t_0) - \lambda \int_{j=0}^{t_0} f(t_0 - j)f_S(j)\,\mathrm{d}j < 0$$

---

[7]We have $\widetilde{S}''(s) = \int_{t=0}^{\infty} t^2 e^{-st} f_S(t)\,\mathrm{d}t > 0$ for every s in the convergence region of $\widetilde{S}(\cdot)$.

because the first term $-(1 - \rho)\lambda f_S(t_0) \leq 0$ and the second term

$$-\lambda \int_{j=0}^{t_0} f(t_0 - j) f_S(j) \, \mathrm{d}j \leq -\lambda \int_{j=a}^{b} f(t_0 - j) f_S(j) \, \mathrm{d}j < 0.$$

But $f'(t_0) < 0$ is impossible, because we assumed that $f(t_0) = 0$. The implication that $f'(t_0) < 0$ contradicts the fact that f is a non-negative probability density function. Therefore,

$$\min_{0 \leq t < N_\epsilon} g(t) \geq \min_{0 \leq t < N_\epsilon} g(t) > 0.$$

On the other hand, since $f'(t) \leq \lambda f(t)$ everywhere, we have $f(t) \leq (1 - \rho)\lambda e^{\lambda t}$. Note that this bound holds even if S has infinite density at one or more points. As a result,

$$\max_{0 \leq t < N_\epsilon} g(t) \leq \max_{0 \leq t < N_\epsilon} \left[f_{T_Q}^{\mathrm{FCFS}}(t) e^{\theta^* N_\epsilon} \right] \leq (1 - \rho)\lambda e^{\lambda N_\epsilon} e^{\theta^* N_\epsilon} < \infty.$$

Finally, note that

$$\inf_{t \in [0, \infty)} g(t) \triangleq g_{\min} \geq \min \left\{ \inf_{0 \leq t < N_\epsilon} \left[f_{T_Q}^{\mathrm{FCFS}}(t) e^{\theta^* t} \right], \frac{g^*}{2} \right\} > 0$$

$$\sup_{t \in [0, \infty)} g(t) \triangleq g_{\max} \leq \max \left\{ \sup_{0 \leq t < N_\epsilon} \left[f_{T_Q}^{\mathrm{FCFS}}(t) e^{\theta^* t} \right], \frac{3g^*}{2} \right\} < \infty$$

which indicates both $g_{\min}$ and $g_{\max}$ are well-defined and nonzero. This completes the proof. □

LEMMA 5.3. *Let A, B be two independent real-valued random variables and c be a fixed constant. Suppose $1 \leq A \leq c$ and $A < B$. Under these assumptions, if $\mathbf{P}\{B > c\} > 0$ and*

$$c\mathbf{E}[B] \geq \mathbf{E}[A] \, \mathbf{E}[B | B > c], \tag{28}$$

then

$$\frac{\mathbf{E}[\min(AB, c) - \min(A, c)]}{\mathbf{E}[\min(AB, c) - \min(B, c)]} = \frac{\mathbf{E}[\min(AB, c) - A]}{\mathbf{E}[\min(AB, c) - \min(B, c)]} \geq \frac{\mathbf{E}[AB - A]}{\mathbf{E}[AB - B]}. \tag{29}$$

PROOF. First we observe $\mathbf{E}[\min(AB, c) - A] > \mathbf{E}[\min(AB, c) - \min(B, c)]$ because $A < \min(B, c)$. Based on this, we can shrink the left hand side of inequality (29) by adding the same positive term to both the denominator and numerator. We compute

$$\frac{\mathbf{E}[\min(AB, c) - A]}{\mathbf{E}[\min(AB, c) - \min(B, c)]} = \frac{\mathbf{E}[AB - A] - \mathbf{E}[(AB - c)\mathbf{1}_{AB>c}]}{\mathbf{E}[AB - B] - \mathbf{E}[(AB - c)\mathbf{1}_{AB>c}] + \mathbf{E}[(B - c)\mathbf{1}_{B>c}]}. \tag{30}$$

Since $A \geq 1$ and $B > 0$, we have $AB \geq B$. Therefore,

$$0 \leq \mathbf{1}_{AB>c} - \mathbf{1}_{B>c} \leq \mathbf{1}_{AB>c} = \mathbf{1}_{AB-c>0}.$$

We proceed by adding a positive term, $\mathbf{E}[(AB - c)\,(\mathbf{1}_{AB>c} - \mathbf{1}_{B>c})]$, to both the denominator and numerator of the right hand side of (30) and obtain

$$\frac{\mathbf{E}[\min(AB, c) - A]}{\mathbf{E}[\min(AB, c) - \min(B, c)]} \geq \frac{\mathbf{E}[AB - A] - \mathbf{E}[(AB - c)\mathbf{1}_{B>c}]}{\mathbf{E}[AB - B] - \mathbf{E}[(AB - B)\mathbf{1}_{B>c}]}.$$

Note that the mediant of this fraction and $\frac{\mathbf{E}[(AB-c)\mathbf{1}_{B>c}]}{\mathbf{E}[(AB-B)\mathbf{1}_{B>c}]}$ is $\frac{\mathbf{E}[AB-A]}{\mathbf{E}[AB-B]}$. The mediant inequality states that the mediant of two fractions lies between those fractions: If $\frac{w}{y} > \frac{x}{z}$ and $y, z > 0$, then $\frac{w}{y} > \frac{w+x}{y+z} > \frac{x}{z}$. Because $1 \leq A < B$ and $\mathbf{P}\{B > c\} > 0$, we have $\mathbf{E}[AB - B] - \mathbf{E}[(AB - B)\mathbf{1}_{B>c}] > 0$

and $\mathbf{E}[(AB - B)\mathbf{1}_{B>c}] > 0$, so we may apply the mediant inequality. In particular, assume that $\frac{\mathbf{E}[(AB-c)\mathbf{1}_{B>c}]}{\mathbf{E}[(AB-B)\mathbf{1}_{B>c}]} \leq \frac{\mathbf{E}[AB-A]}{\mathbf{E}[AB-B]}$. Then, by the mediant inequality,

$$\frac{\mathbf{E}[(AB-c)\mathbf{1}_{B>c}]}{\mathbf{E}[(AB-B)\mathbf{1}_{B>c}]} \leq \frac{\mathbf{E}[AB-A]}{\mathbf{E}[AB-B]}$$

$$\implies \frac{\mathbf{E}[AB-A]}{\mathbf{E}[AB-B]} \leq \frac{\mathbf{E}[AB-A] - \mathbf{E}[(AB-c)\mathbf{1}_{B>c}]}{\mathbf{E}[AB-B] - \mathbf{E}[(AB-B)\mathbf{1}_{B>c}]} \leq \frac{\mathbf{E}[\min(AB,c) - A]}{\mathbf{E}[\min(AB,c) - \min(B,c)]},$$

which is the inequality we are trying to prove, (29).

As a result, to prove (29), it suffices to show that $\frac{\mathbf{E}[(AB-c)\mathbf{1}_{B>c}]}{\mathbf{E}[(AB-B)\mathbf{1}_{B>c}]} \leq \frac{\mathbf{E}[AB-A]}{\mathbf{E}[AB-B]}$. We simplify this inequality as follows:

$$\frac{\mathbf{E}[(AB-c)\mathbf{1}_{B>c}]}{\mathbf{E}[(AB-B)\mathbf{1}_{B>c}]} \leq \frac{\mathbf{E}[AB-A]}{\mathbf{E}[AB-B]}$$

$$\iff \mathbf{E}[(AB-c)\mathbf{1}_{B>c}]\,\mathbf{E}[AB-B] \leq \mathbf{E}[(AB-B)\mathbf{1}_{B>c}]\,\mathbf{E}[AB-A]$$

$$\iff (\mathbf{E}[A]\,\mathbf{E}[B|B>c] - c)\,\mathbf{P}\{B>c\}\,\mathbf{E}[AB-B] \leq (\mathbf{E}[A]-1)\,\mathbf{E}[B|B>c]\,\mathbf{P}\{B>c\}\,\mathbf{E}[AB-A]$$

$$\iff (\mathbf{E}[A]\,\mathbf{E}[B|B>c] - c)\,\mathbf{E}[B] \leq \mathbf{E}[B|B>c]\,\mathbf{E}[AB-A]$$

$$\iff \mathbf{E}[A]\,\mathbf{E}[B|B>c]\,\mathbf{E}[B] - c\mathbf{E}[B] \leq \mathbf{E}[B|B>c]\,\mathbf{E}[A]\,\mathbf{E}[B] - \mathbf{E}[B|B>c]\,\mathbf{E}[A]$$

$$\iff -c\mathbf{E}[B] \leq -\mathbf{E}[A]\,\mathbf{E}[B|B>c]$$

$$\iff c\mathbf{E}[B] \geq \mathbf{E}[A]\,\mathbf{E}[B|B>c]$$

which is precisely the condition provided in (28).   □

B PROOFS FOR TRANSFORM ANALYSIS

LEMMA 7.3. *The small tagged job swaps with a large job in the Nudge system if and only if, when it arrives, the FCFS system has a nonempty queue whose last job is large.*

PROOF. By definition of Nudge, the tagged job swaps if and only if, when it arrives, the Nudge system has a nonempty queue whose last job is a large job that has not been swapped. It therefore suffices to show that at any moment in time, the FCFS system has a nonempty queue whose last job is large if and only if the Nudge system has a nonempty queue whose last job is a large job that has not been swapped.

We first note that the total amount of work in both systems is the same at every moment in time, because both FCFS and Nudge are work conserving.

Suppose the FCFS system has a nonempty queue whose last job j is large. Because it is the last job in the FCFS queue, there have been no new arrivals after j. In the Nudge system, this means j has not been swapped, so either j is the last job in the Nudge queue or has entered service. By work conservation, both systems had the same amount of work when j arrived, so j must still be in the Nudge queue, as desired.

Suppose the Nudge system has a nonempty queue whose last job j is a large job that has not been swapped. We argue similarly to the previous direction: there have been no arrivals since j because it is at the end of the Nudge queue without being swapped, and j cannot have entered service in the FCFS system by work conservation, so j must be the last job in the FCFS queue, as desired.   □

LEMMA 7.4. *Let $q \triangleq \mathbf{P}\big\{\mathrm{NoPoisson}(W, s) \,\wedge\, N_Q = 0\big\}$. We have*

$$q = \widetilde{T_Q^{\mathrm{FCFS}}}(\lambda) \cdot \frac{\lambda\widetilde{S}(s) - s\widetilde{S}(\lambda)}{\lambda - s} = \frac{1-\rho}{\widetilde{S}(\lambda)} \cdot \frac{\lambda\widetilde{S}(s) - s\widetilde{S}(\lambda)}{\lambda - s}.$$

To prove Lemma 7.4, we require an additional lemma.

LEMMA B.1. *Let V be a nonnegative random variable, and let $\text{Exp}(r)$ and $\text{Exp}(s)$ be exponentially distributed random variables of rates r and s, respectively. Suppose V, $\text{Exp}(r)$, and $\text{Exp}(s)$ are mutually independent. Then*

$$\mathbf{P}\{V < \text{Exp}(r) + \text{Exp}(s)\} = \frac{r\widetilde{V}(s) - s\widetilde{V}(r)}{r - s}.$$

PROOF. We compute

$$
\begin{aligned}
\mathbf{P}\{V < \text{Exp}(r) + \text{Exp}(s)\} &= \int_{v=0}^{\infty} \mathbf{P}\{v < \text{Exp}(r) + \text{Exp}(s)\} f_V(v) \, \mathrm{d}v \\
&= \int_{v=0}^{\infty} \left(\int_{u=0}^{\infty} \int_{t=0}^{\infty} \mathbf{1}(v < t + u) \cdot re^{-rt} \cdot se^{-su} \, \mathrm{d}t \, \mathrm{d}u \right) f_V(v) \, \mathrm{d}v \\
&= \int_{v=0}^{\infty} \frac{re^{-rv} - se^{-sv}}{r - s} f_V(v) \, \mathrm{d}v \\
&= \frac{r\widetilde{V}(s) - s\widetilde{V}(r)}{r - s}. \qquad \qquad \square
\end{aligned}
$$

PROOF OF LEMMA 7.4. Consider a FCFS system in equilibrium along with an independent Poisson "interruption" process of rate s. Call a job *lucky* if it enters the system while $N_Q = 0$ and experiences no interruptions during its queueing time. Because Poisson arrivals see time averages [41], q is probability an arriving job is lucky.

We compute q in an unusual way. Let a job's *departure period* be the time interval starting when the job enters service and ending when the next job enters service. Jobs enter service at average rate λ, so q is the average number of lucky jobs that arrive a departure period. More formally, by renewal-reward theorem,

$$
\begin{aligned}
q &= \mathbf{P}\{\text{arrival is lucky}\} \\
&= \frac{\text{average rate of lucky arrivals}}{\lambda} \\
&= \frac{\text{average rate of lucky arrivals}}{\text{average rate of departure periods}} \\
&= \mathbf{E}[\text{number of lucky arrivals during a departure period}].
\end{aligned}
$$

Moreover, because a job is lucky only if $N_Q = 0$, only the first arrival in a departure period can possibly be lucky, so

$$q = \mathbf{P}\{\text{first arrival in a departure period is lucky}\}.$$

Consider a job j. The first arrival in j's departure period is lucky if both of the following events occur:

E_1 = there are no arrivals during j's queueing time

E_2 = j completes before the first interruption after the first arrival of j's departure period.

By (22), $\mathbf{P}\{E_1\} = \widetilde{T_Q^{\text{FCFS}}}(\lambda)$. We compute $\mathbf{P}\{E_2 \mid E_1\}$ below.

Conditioned on E_1, the queue is empty when j enters service, so the first arrival during j's departure period is simply the first arrival after j enters service. Let $\text{Exp}(\lambda)$ be the amount of time between when j enters service and the next arrival, and let $\text{Exp}(s)$ be the amount of time between that next arrival and the first interruption after it. Both $\text{Exp}(\lambda)$ and $\text{Exp}(s)$ are exponentially

distributed with rates λ and s, respectively, and they and j's size are mutually independent. Because j's size is distributed as S, we have

$$\mathbf{P}\{E_2 \mid E_1\} = \mathbf{P}\{S < \text{Exp}(\lambda) + \text{Exp}(s)\} = \frac{\lambda \widetilde{S}(s) - s\widetilde{S}(\lambda)}{\lambda - s},$$

where the latter equality follows from Lemma B.1 below.

It remains only to show $\widetilde{T_Q^{\text{FCFS}}}(\lambda) = (1-\rho)/\widetilde{S}(\lambda)$. Because response time T^{FCFS} is a sum of independent random variables with distributions T_Q^{FCFS} and S, we have $\widetilde{T_Q^{\text{FCFS}}}(\lambda) = \widetilde{T^{\text{FCFS}}}(\lambda)/\widetilde{S}(\lambda)$. By (22), $\widetilde{T^{\text{FCFS}}}(\lambda)$ is the probability that no arrivals occur during a job's response time. But this is simply the probability that a job leaves an empty system when it departs, which is $1-\rho$.   □

THEOREM 4.3. *Suppose S is a continuous class I job size distribution. For any $s_{\min} < x_1 \leq x_2 \leq x_3$, the asymptotic tail improvement ratio of $\text{Nudge}(x_1, x_2, x_3)$ compared to FCFS is*

$$\text{AsymTIR} = p_{\text{small}} p_{\text{large}} \frac{\lambda}{\lambda + \theta^*} \left(\widetilde{S_{\text{large}}}(-\theta^*) - \frac{\lambda}{\lambda + \theta^*} \widetilde{S_{\text{small}}}(-\theta^*) - \frac{\theta^*}{\lambda + \theta^*} \widetilde{S_{\text{large}}}(-\theta^*)\widetilde{S_{\text{small}}}(-\theta^*) \right).$$

Furthermore, AsymTIR is positive, meaning $C_{\text{Nudge}} < C_{\text{FCFS}}$, if

$$\frac{\lambda + \theta^*}{\lambda} < \frac{1 - \widetilde{S_{\text{large}}}(-\theta^*)^{-1}}{1 - \widetilde{S_{\text{small}}}(-\theta^*)^{-1}}.$$

PROOF. We prove this theorem using the Laplace-Stieltjes transform derived in Theorem 4.5. The transform of the tail of T^{Nudge} can be calculated as

$$\begin{aligned}
\int_{t=0}^{\infty} e^{-st} \mathbf{P}\{T^{\text{Nudge}} > t\} \, dt &= -\int_{t=0}^{\infty} \mathbf{P}\{T^{\text{Nudge}} > t\} \, d\left(\frac{e^{-st}}{s}\right) \\
&= -\mathbf{P}\{T^{\text{Nudge}} > t\}\left(\frac{e^{-st}}{s}\right)\Big|_0^{\infty} + \int_{t=0}^{\infty} \frac{e^{-st}}{s} \, d\mathbf{P}\{T^{\text{Nudge}} > t\} \\
&= \frac{1}{s}\left(1 - \int_{t=0}^{\infty} e^{-st} f_{T^{\text{Nudge}}}(t) \, dt\right) \\
&= \frac{1 - \widetilde{T^{\text{Nudge}}}(s)}{s}.
\end{aligned} \tag{31}$$

Then the transform of $e^{\theta^* t} \mathbf{P}\{T^{\text{Nudge}} > t\}$ is obtained by translating (31) horizontally through θ^* units:

$$\int_{t=0}^{\infty} e^{-st} \left(e^{\theta^* t} \mathbf{P}\{T^{\text{Nudge}} > t\}\right) dt = \frac{1 - \widetilde{T^{\text{Nudge}}}(s - \theta^*)}{s - \theta^*}.$$

Now, we are ready to calculate C_{Nudge} using this transform. From Final Value Theorem,

$$\begin{aligned}
C_{\text{Nudge}} &= \lim_{t \to \infty} e^{\theta^* t} \mathbf{P}\{T^{\text{Nudge}} > t\} \\
&= \lim_{s \to 0} s \frac{1 - \widetilde{T^{\text{Nudge}}}(s - \theta^*)}{s - \theta^*} \\
&= \frac{1}{\theta^*} \lim_{s \to 0} s \widetilde{T^{\text{Nudge}}}(s - \theta^*)
\end{aligned}$$

Now, we substitute the expression from Theorem 4.5, and drop terms that are negligible in $s \to 0$ limit.

$$\frac{1}{\theta^*} \lim_{s \to 0} s \widetilde{T^{\text{Nudge}}}(s - \theta^*)$$

$$= C_{\text{FCFS}} + \frac{1}{\theta^*} p_{\text{small}} p_{\text{large}} \left(\widetilde{S_{\text{large}}}(-\theta^*)(\widetilde{S_{\text{small}}}(-\theta^*) - 1) C_{Q,\text{FCFS}} + \widetilde{S_{\text{small}}}(-\theta^*)(1 - \widetilde{S_{\text{large}}}(-\theta^*)) \frac{C_{Q,\text{FCFS}}}{\widetilde{S}(-\theta^*)} \right),$$
(32)

where

$$C_{Q,\text{FCFS}} = \lim_{s \to 0} s \widetilde{T_Q^{\text{FCFS}}}(s - \theta^*) = \lim_{s \to 0} s \frac{\widetilde{T^{\text{FCFS}}}(s - \theta^*)}{\widetilde{S}(s - \theta^*)} = \frac{\theta^* C_{\text{FCFS}}}{\widetilde{S}(-\theta^*)}.$$
(33)

We recall that $-\theta^*$ is the rightmost singularity of $\widetilde{T_Q^{\text{FCFS}}}(s) = \frac{(1-\rho)s}{\lambda \widetilde{S}(s) - \lambda + s}$, which indicates that θ^* is the smallest positive value that satisfies

$$\lambda \widetilde{S}(-\theta^*) - \lambda - \theta^* = 0 \quad \text{and} \quad \widetilde{S}(-\theta^*) = \frac{\lambda + \theta^*}{\lambda}.$$
(34)

Using (33) and (34) to simplify (32), we obtain

$$C_{\text{Nudge}}$$

$$= C_{\text{FCFS}} \left(1 - p_{\text{small}} p_{\text{large}} \frac{\lambda}{\lambda + \theta^*} \left(\widetilde{S_{\text{large}}}(-\theta^*) - \frac{\lambda}{\lambda + \theta^*} \widetilde{S_{\text{small}}}(-\theta^*) - \frac{\theta^*}{\lambda + \theta^*} \widetilde{S_{\text{large}}}(-\theta^*) \widetilde{S_{\text{small}}}(-\theta^*) \right) \right).$$

This gives us

$$\text{AsymTIR} = 1 - \frac{C_{\text{Nudge}}}{C_{\text{FCFS}}}$$

$$= p_{\text{small}} p_{\text{large}} \frac{\lambda}{\lambda + \theta^*} \left(\widetilde{S_{\text{large}}}(-\theta^*) - \frac{\lambda}{\lambda + \theta^*} \widetilde{S_{\text{small}}}(-\theta^*) - \frac{\theta^*}{\lambda + \theta^*} \widetilde{S_{\text{large}}}(-\theta^*) \widetilde{S_{\text{small}}}(-\theta^*) \right).$$

By assumption,

$$\frac{\lambda + \theta^*}{\lambda} < \frac{1 - \widetilde{S_{\text{large}}}(-\theta^*)^{-1}}{1 - \widetilde{S_{\text{small}}}(-\theta^*)^{-1}},$$

so we have

$$\text{AsymTIR} = p_{\text{small}} p_{\text{large}} \frac{\lambda}{\lambda + \theta^*} \widetilde{S_{\text{large}}}(-\theta^*) \widetilde{S_{\text{small}}}(-\theta^*) \left(\widetilde{S_{\text{small}}}(-\theta^*)^{-1} - \frac{\lambda}{\lambda + \theta^*} \widetilde{S_{\text{large}}}(-\theta^*)^{-1} - \frac{\theta^*}{\lambda + \theta^*} \right)$$

$$= p_{\text{small}} p_{\text{large}} \frac{\lambda}{\lambda + \theta^*} \widetilde{S_{\text{large}}}(-\theta^*) \widetilde{S_{\text{small}}}(-\theta^*) \left(\frac{\lambda}{\lambda + \theta^*} \left(1 - \widetilde{S_{\text{large}}}(-\theta^*)^{-1} \right) - \left(1 - \widetilde{S_{\text{small}}}(-\theta^*)^{-1} \right) \right) > 0.$$

Hence $C_{\text{Nudge}} < C_{\text{FCFS}}$. □