



# The Price of Anarchy of Strategic Queuing Systems

JASON GAITONDE and ÉVA TARDOS, Cornell University

Bounding the price of anarchy, which quantifies the damage to social welfare due to selfish behavior of the participants, has been an important area of research in algorithmic game theory. Classical work on such bounds in repeated games makes the strong assumption that the subsequent rounds of the repeated games are independent beyond any influence on play from past history. This work studies such bounds in environments that themselves change due to the actions of the agents. Concretely, we consider this problem in discrete-time queuing systems, where competitive queues try to get their packets served. In this model, a queue gets to send a packet at each step to one of the servers, which will attempt to serve the oldest arriving packet, and unprocessed packets are returned to each queue. We model this as a repeated game where queues compete for the capacity of the servers, but where the state of the game evolves as the length of each queue varies.

We analyze this queuing system from multiple perspectives. As a baseline measure, we first establish precise conditions on the queuing arrival rates and service capacities that ensure all packets clear efficiently under centralized coordination. We then show that if queues strategically choose servers according to independent and stationary distributions, the system remains stable provided it would be stable under coordination with arrival rates scaled up by a factor of just  $\frac{e}{e-1}$ . Finally, we extend these results to no-regret learning dynamics: if queues use learning algorithms satisfying the no-regret property to choose servers, then the requisite factor increases to 2, and both of these bounds are tight. Both of these results require new probabilistic techniques compared to the classical price of anarchy literature and show that in such settings, no-regret learning can exhibit efficiency loss due to myopia.

CCS Concepts: • **Theory of computation** → **Convergence and learning in games**; **Quality of equilibria**;

Additional Key Words and Phrases: Price of anarchy, no-regret learning in games

## ACM Reference format:

Jason Gaitonde and Éva Tardos. 2023. The Price of Anarchy of Strategic Queuing Systems. *J. ACM* 70, 3, Article 20 (May 2023), 63 pages.  
<https://doi.org/10.1145/3587250>

## 1 INTRODUCTION

A fundamental aim at the intersection of economics and computer science is to understand the efficiency of systems when the dynamics are governed by the actions of strategic and competitive agents. In general, the outcomes reached by selfish agents at an equilibrium, or under some other

This work was supported by the National Science Foundation under grants CCF-1408673 and CCF-1563714, as well as the Air Force Office of Scientific Research under grants FA9550-19-1-0183 and FA9550-23-1-0068.

Authors' address: J. Gaitonde and É. Tardos, Cornell University, Gates Hall, Ithaca, New York, NY 14853; emails: [jsg355@cornell.edu](mailto:jsg355@cornell.edu), [eva.tardos@cs.cornell.edu](mailto:eva.tardos@cs.cornell.edu).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

0004-5411/2023/05-ART20 \$15.00

<https://doi.org/10.1145/3587250>

dynamics, need not necessarily align well with the design of the system. Real-world systems whose performance is provably near-optimal even in the presence of selfish agents can be viewed as *robust*: such systems can be safely decentralized without sacrificing much efficiency. However, systems whose behavior degrades dramatically in strategic settings require extra safeguards to prevent inferior outcomes.

To make this more concrete, consider the setting of routing in networks as considered in the work of Roughgarden and Tardos [34]. Formulating this as a game, each agent chooses a path in the network from their source vertex to their sink vertex with the objective of minimizing their experienced delay along this path. The delay of an edge is some monotonic function of the number of agents traversing this edge in their chosen path. At a *Nash equilibrium*, selfish agents each select the path with minimum delay *given the equilibrium behavior of the others*; in other words, each agent selfishly optimizes her own delay function given the actions of the others. These equilibrium outcomes can be quite different from the globally optimal choice of routes that minimizes the *total delay* of all agents. However, Roughgarden and Tardos show that in the case of atomless selfish routing with affine delay functions, the ratio between the total delay at any Nash equilibrium and that of the global optimum is bounded by  $4/3$ . More generally, if the delay functions are polynomials of degree at most  $d$ , then the ratio can behave as  $\Theta(d/\ln d)$  [31]. Such a result thus strongly characterizes the gap between selfish and optimal behavior, and identifies the *concrete obstruction* (here, *nonlinearity*) to efficiency.

In more general settings and games, given some quantitative notion of social welfare, the worst-case ratio between the social welfare at the global optimum with respect to this metric and that of any Nash equilibrium (defined in an analogous way as in the routing setting) is known as the *price of anarchy* [24]. More generally, we will refer to any sort of quantitative comparison between socially optimal outcomes and equilibrium outcomes as a price of anarchy analysis. A large body of work has established price of anarchy bounds for various well-studied games like routing, scheduling, and auctions, among others. Establishing price of anarchy results is important for several reasons. First, price of anarchy analyses help us understand the performance of real-world systems that already exist “in the wild” [33]. Just as importantly, the quantitative understanding given by price of anarchy-style analyses yield crucial insights toward the design of more robust systems with respect to selfish behaviors. For instance, Roughgarden and Tardos [34] further showed that in routing games, the cost of any Nash equilibrium is no more than that of the centralized optimum for twice as much flow. Such an analysis gives clearly actionable prescriptions: to attain good performance, one can simply augment the amount of resources in the system relative to the competition.

In many settings, agents repeatedly play against each other in the same game, not just once. In the most commonly studied models of repeated games, a critical assumption is that in each round, agents play an “independent” copy of the *same* game. In other words, the past sequence of play does not fundamentally change the nature of the game that is repeated in each round. In some settings, this approximation might be valid. For instance, in the preceding routing setting, consider routing on the scale of the morning rush-hour traffic. In this case, it usually holds that any traffic on Monday morning will have cleared by Tuesday morning, at which point agents again choose paths in a fresh version of the game. However, consider modeling packet routing in computer networks. If a packet gets dropped, then this packet must be re-sent in future rounds and thereby increases the total congestion going forward, thus violating the assumption that the game itself has not changed.

Therefore, developing a deeper theory of the efficiency of strategic agents in repeated games that retain state is of large importance. This motivates the first main question that guides our work in this article.

QUESTION 1. Do natural repeated games with state admit strong price of anarchy-style guarantees?

In this article, we extensively consider this problem of price of anarchy bounds in systems with state that impacts future rounds. Concretely, we study Question 1 in a queuing system with queues sending packets to servers as a simplified model of a network of queues as previously considered by Krishnasamy et al. [25]. In their work, the authors study the performance of a *centralized* learner in the same queuing system that finds the best server with respect to a more refined notion of “queue-regret,” which measures the expected difference in queue sizes to that of a genie strategy that knows the optimal server.

In contrast, our work studies a *decentralized and strategic multi-queue* version of the same system, where each queue selfishly competes with each other for service. Our primary focus is on precisely characterizing what conditions ensure that this queuing system remains stable (in a quantitative sense that will be formalized later) even under strategic assumptions, a concern that does not arise in the learning problem with centralized scheduling. In particular, we consider conditions that guarantee the efficiency of stochastic queuing systems when queues choose servers with the aim of getting their packets served in minimal time, and where queues must repeatedly resend their packets until the packet gets served.

To do so, we study the amount of extra resources required for stability in queuing systems under a variety of different behavioral structures. We first consider this question in the completely centralized setting, where a central coordinator can specify which queues send to which servers at each time. We first show, using the theory of majorization, that the obvious necessary condition is in fact sufficient: namely, the sum of arrival rates of the top  $k$  queues must be at most the sum of the capacities of the top  $k$  servers, for each  $k$ . This result establishes a baseline measure on the feasibility of queuing systems to compare against the rest of our results in strategic setting.

Once we establish this baseline, we may then proceed to ask: how many extra resources (server capacity) are needed relative to the coordinated setting to ensure that the system remains stable under some notion of game-theoretic equilibrium? We first study this question under the natural assumption that the behavior is stationary over time for each queue; this situation models the outcomes of systems that have reached a long-term equilibrium. We do so by defining a one-shot version of the queuing dynamics, which we call the *patient queuing game*, where each queue must choose a fixed distribution over servers. The term *patience* refers to the fact that the incentives for the queues in this game are their *long-run growth rates*, measured in terms of the limiting number of uncleared packets normalized by time. Thus, the values obtained in this game correspond to the long-run growth rates obtained in systems when the distribution of actions (choice of server) has stabilized for each queue, and we then ask for the values attained at any Nash equilibrium.

To answer our main question, we establish several delicate probabilistic properties of such systems and show that the long-run behavior of such systems, as a function of the choices of the queues, admits a number of favorable analytic properties. These techniques and results allow us to directly connect these probabilistic properties of the queuing system with the *game-theoretic incentives* of the agents to show that a factor of just  $\frac{e}{e-1}$  extra server capacity is needed to ensure stability of every Nash equilibrium of this game. Along the way, we also consider, but do not resolve, the separate question of the optimal rates attained by *any* independent strategies. We show that this question is equivalent to the classical notion of price of stability<sup>1</sup> [1, 2] of these games.

Although our results in this setting are tight, the assumption that agents have reached a stationary equilibrium is quite strong. In fact, this problem applies equally to the “independent game” setting. To address this deficiency, recent work in the independent game setting has shown that

<sup>1</sup>Note that “stability” here refers to game-theoretic stability of behavior, not the stability notion in our queuing system.

price of anarchy bounds often seamlessly generalize from the single-round equilibrium setting to the repeated games setting where agents employ *no-regret learning algorithms* [7, 26, 32]. In this model, agents repeatedly play the same game against each other and use learning algorithms to adapt to each other's behavior. Such extensions are crucial in capturing the outcomes that might arise in practice, as there are known obstructions to the predictive power of Nash equilibria; to name just a few, Nash equilibria need not be unique, may be computationally infeasible to compute [13, 35], and may require stringent assumptions on the knowledge of the agents. Taken together, these deficiencies may collectively prevent real-world agents from reaching equilibrium outcomes. By contrast, no-regret algorithms are simple, computationally efficient, and encode natural and minimal behavioral assumptions of strategic agents, all while enjoying sufficient provable guarantees to enable a price of anarchy analysis. This guarantee can be ensured by running any of a large set of learning algorithms [38]. The study of learning in games has a long history, dating back to the early work of Brown [10] and Robinson [30] (also see the work of Fudenberg and Levine [19]). In traditional repeated games, if all players employ a no-regret learning strategy, then the play converges to a form of correlated equilibrium of the game [20] (players correlating their play by each of them using the history of play to decide their next action), and price of anarchy analyses often extend also to the correlated play.

Therefore, to complement our equilibrium results, we turn to analyzing the outcomes that arise under learning dynamics. A fundamental question in our setting with strong inter-round dependencies is whether price of anarchy bounds similarly extend to natural learning dynamics.

**QUESTION 2.** *To what extent do price of anarchy-style bounds extend in repeated games with state to learning dynamics?*

To formally address this question, we consider the performance of queuing systems where each queue uses no-regret learning to determine which servers to send packets to over time. Because we only assume that the queues satisfy the no-regret property in their choices, much of the analytic structure developed for stationary strategies in the patient queuing game is no longer applicable. In contrast to more classical settings with repeated games, where such price of anarchy bounds often naturally extend from equilibrium notions to dynamic learning outcomes, we show that no-regret learning exhibits *myopia*, as it generally cannot consider the long-run dependencies in the system. We develop new probabilistic machinery from the equilibrium setting to prove that the corresponding queuing systems require twice the amount of server capacity needed for centralized stability, and this bound is again tight. When taken together with our equilibrium bounds, a key conceptual contribution of our work is thus that standard no-regret algorithms can indeed attain nontrivial performance guarantees in these repeated games with state, but that these guarantees are not entirely lossless. Moreover, the analyses themselves appear to require a fundamentally different set of tools rather than being corollaries of the same generic framework [32].

In the next section, we elaborate on the model, our results, and techniques.

## 1.1 Overview of Results and Techniques

**1.1.1 Strategic Queuing Model.** Before discussing our results in more detail, we first briefly describe the queuing model that will be the focus of this work.<sup>2</sup> As mentioned earlier, we consider the discrete-time queuing system studied by Krishnasamy et al. [25], where  $n$  queues receive packets at heterogeneous rates  $\lambda = (\lambda_1, \dots, \lambda_n) \in (0, 1)^n$  so that queue  $i$  receives a new packet at each time with probability  $\lambda_i$ . In each round, any queue that has any remaining packets must select exactly one of  $m$  servers with heterogeneous success probabilities  $\mu = (\mu_1, \dots, \mu_m) \in [0, 1]^m$ , to

<sup>2</sup>See Section 2.2 for precise definitions and the full specification.

attempt to clear a single packet. Each server can only succeed in clearing at most one packet in each round and, most importantly, returns each unprocessed packet to the original queue, assuming for simplicity that servers have no buffer.<sup>3</sup> We will also assume that each queue receives only bandit feedback in each round, meaning that it only observes whether it succeeded in clearing the packet it sent in the current round at the server it attempted.

Queue lengths can grow arbitrarily, so the efficiency we consider is under what conditions on the service and arrival rates can it be ensured that the system remains *stable*? We provide formal definitions in Section 2 of the various stability notions we will consider, but informally, stability corresponds to the number of uncleared packets growing *sublinearly* over time. To establish a baseline measure on what is possible under a *centralized* coordination algorithm, we will prove the following theorem.

**THEOREM 1.1 (THEOREM 2.2, INFORMAL).** *Let  $\lambda \in (0, 1)^n$  and  $\mu \in [0, 1]^m$  be the arrival and service rates, respectively. Then the preceding queuing system is stable for some centralized (coordinated) scheduling policy if and only if for all  $1 \leq k \leq n$ ,*

$$\sum_{j=1}^{\min\{k, m\}} \mu_j \geq \sum_{i=1}^k \lambda_i. \quad (1)$$

Because the focus of our work is on outcomes with *strategic queues* under different behavioral assumptions, an important feature of our model is how conflicts are resolved when multiple queues send to the same server in a time period. In decentralized settings, queues sending a packet at each round can and will often collide at a server, necessitating a choice of which (if any) packet a server attempts to serve. There are at least two natural choices. As a first natural choice, a server may choose a packet to attempt to clear *uniformly at random* among those that it receives in a round. Although this is a plausible modeling choice, we will actually show (Theorem 2.4) that in this case, the number of uncleared packets in the system can increase linearly over time when queues are strategic unless the success rates of the servers are prohibitively larger than the arrival rates of the queues. Roughly speaking, uniform randomization by the servers is not sufficiently adapted to our system objective of stability; even if there exists a simple coordinated strategy that would ensure queues remain bounded, strategic behavior by the queues can prevent queues with higher arrival rates from exploiting the necessary servers. In particular, such a model precludes the possibility of a constant factor of resource augmentation ensuring stability of selfish queuing systems.

To address this immediate difficulty, we turn to a second natural choice: instead of choosing a packet to serve uniformly at random, we will assume that packets are labeled with timestamps and that servers attempt to serve the received packet with *oldest* timestamp (breaking ties arbitrarily). This choice immediately induces significant dependencies between rounds, for queues that for some reason have not successfully cleared many packets will have priority over queues that have been successful in previous rounds. Although this is a natural choice to facilitate stability in such systems, we will require delicate probabilistic reasoning to study such processes. A simple but key idea that will enable our analysis will be to study a *Geometric* version of this model that is more tailored to this priority scheme via the principle of deferred decisions, where “Geometric” refers to the type of random variables governing the evolution of the system. Namely, we reduce the analysis of this complicated system to studying a *single parameter* for each queue, the age of the oldest packet in the queue, that meshes well with the priority structure. After formulating and

<sup>3</sup>We remark that in other queuing systems, the servers may also have a bounded size queue and would only send back (or drop) packets when they no longer fit on the queue; our simpler model without server queues makes the tradeoffs we want to study cleaner. A packet sent to a server is either served or returned and offers instantaneous feedback to the learning algorithms of the queues, in contrast to the bit more informative but delayed feedback available in real systems.



proving this equivalence in Section 2.4, all of our subsequent results will be in this Geometric system.

**1.1.2 Patient Selfishness.** To understand these systems game theoretically, our first main contribution is to understand the performance of systems that are at long-run equilibrium even though agents nonetheless compete to optimize their own long-run growth rate. We do this by formulating a one-shot version of the queuing dynamics. We define the *patient queuing game* where queues choose *fixed* randomized strategies over servers to be played at each round. Each queue  $i$ 's choice over servers can be described by a fixed vector  $p_i \in \Delta^{m-1}$ , where  $\Delta^{m-1}$  is the probability simplex over the  $m$  servers. We study this as a traditional game and consider the resulting Nash equilibria when each queue aims to choose their fixed randomization to minimize their *long-run aging rate* (equivalently, their long-run growth rate, see Section 2.2) conditioned on the others. Our main interest is understanding under what conditions on the service rates and arrival rates the system will remain stable in *every Nash equilibrium*. To study this, we face significant probabilistic and game-theoretic challenges: probabilistic challenges to determine and prove a closed form of asymptotic growth rates for given strategies exists, and game-theoretic challenges in showing that Nash equilibria exist and bounding their quality. The techniques we use will prove useful in addressing these conceptually distinct difficulties, thereby unifying the game-theoretic and probabilistic properties of our systems.

**Asymptotic Growth Rates.** In the preceding discussion, we stated that each queue aims to select a fixed randomization over servers to minimize their long-run aging rate in this system given the randomizations of the others. Our first task, to do any game-theoretic analysis of this system, is to analyze the long-run properties of this random process of queue ages (which typically will not even be recurrent). A major technical component of our work is showing that for any fixed, independent randomizations  $\mathbf{p}$  by the queues over servers, not only do these long-run growth rates exist almost surely, they are *deterministic* and can be explicitly computed as a function of the strategies.

**THEOREM 1.2 (THEOREM 3.3, INFORMAL).** *There exists an explicit, continuous function  $r : (\Delta^{m-1})^n \rightarrow \mathbb{R}_{\geq 0}^n$  such that if queues independently randomize over servers according to  $\mathbf{p} \in (\Delta^{m-1})^n$ , then the (random) long-run growth rate of each queue  $i$  is  $r_i(\mathbf{p})$  almost surely.*

To prove this result, we provide an alternate, *algorithmic* description of the long-run rates in Section 3.1, which we use for all of our subsequent game-theoretic results. Working just with this alternative definition, we show that the queues partition into groups such that all queues in a group age asymptotically at the same rate. We will return to the task of establishing that the true, long-run asymptotic aging rates of the queues for any choice of strategies coincide with the output of the algorithm in Section 6.

The key technical difficulty in proving this alternate characterization is that the priority structure via timestamps changes rapidly round to round and depends crucially on past successes by the queues. To overcome this, we use a rather delicate inductive argument that accounts for these changes in a controlled fashion that enables us to keep track of the evolution of the queuing system at a less granular level while still being sharp enough to prove the precise quantitative rates. Once we have done so, only then can we repeatedly appeal to concentration bounds to argue that each subset in the partition grows at the desired rate. To conclude, we carefully apply the Borel-Cantelli lemma to establish the result. For ease of exposition, some of the highly nontrivial and technical details are deferred to Appendix F.

**Game-Theoretic Properties: Equilibria and Price of Anarchy.** Once we show that these limits almost surely are equal to an explicit, deterministic function of  $\mathbf{p}$ , it might still not be the case that a Nash equilibrium exists in the induced game. However, we show that the cost function exhibits

significant analytic properties, which lets us reason about the structure of the sets that arise in the partition for any fixed strategy profile. More precisely, we show that each level set of the cost function corresponds to the minimizing subset of the ratio of a submodular and modular set function; this significant structure allows us to show that the subsets that minimize this ratio are closed under union and nonempty intersection (and thus essentially form a Boolean lattice). These considerations will be enough to show continuity as a function of the strategies (Proposition 3.7), which along with other properties will enable us to show that an equilibrium exists using Kakutani's theorem (Theorem 3.8). Although we show that the cost function of our game has significant structure, the correspondence between actions (randomizations) and costs is quite nonlinear, imposing new technical challenges.

Recall that our goal is to ensure stability in any Nash equilibrium, assuming some relationship on the service rates to the arrival rates. Our main result, proven in Section 4, shows that the correct constant of system slack is  $\frac{e}{e-1} \approx 1.58$ .

**THEOREM 1.3 (MAIN, COROLLARY 4.2, INFORMAL).** *If the service capacity is large enough so that the system would remain feasible when centrally managed even if server rates are scaled down by  $\frac{e}{e-1}$ , then in every equilibrium of this game, all queues are stable.*

This result is tight—in a symmetric system where  $n = m$ , each queue has the same arrival rate, each server has the same success rate, and each queue chooses to uniformly randomize over servers, a simple balls-in-bins analysis yields this constant as  $m, n \rightarrow \infty$ .

To prove this theorem, we provide a novel argument that establishes the result by continuously deforming any Nash profile toward a carefully constructed strategy profile while only monotonically decreasing the rate at which the top group clears. We then analyze the resulting profile to give a lower bound on the value of the Nash profile. The key difficulty is that the relevant incentives for each queue correspond to possibly *many different* subsets of queues that have a maximal aging rate. These constraints are difficult to directly compare; different choices of deviations in the strategy by a queue at any Nash equilibrium may violate distinct constraints, making it unclear how to argue about the quality of these equilibria. In particular, there does not seem to be a direct analogue of the Nash indifference principle in finite-action games where utilities are affine in the randomizations of each agent (recall Example 1.4, where the queue moving to the lesser server will still appear to prefer the better server).

To overcome these difficulties, we show that one can significantly reduce the number of incentive constraints one must consider for each queue (Proposition 4.3). This part of the argument crucially uses the lattice structure of the subsets of queues that ages fastest. We can then carefully perform our deformation of the collective strategy vector of the queues according to the structure of these sparsified constraints, and show that our deformation only hurts the quality of the Nash solution to provide a valid lower bound.

In contrast, almost every known price of anarchy-style result can be viewed via the very general *smoothness* framework of Roughgarden [32], which connects an equilibrium with the social optimum via discrete changes in the strategy profile. Our argument instead relies on a careful equilibrium analysis that smoothly interpolates between the equilibrium and a “good” profile that is easy to explicitly bound; however, during these deformations, these intermediate strategy profiles will *not* be equilibria. To prove the monotonicity of this deformation, we connect the incentives at Nash to the structure of the subset of maximizers of the long-run rate function and show that the Nash constraints still hold along the directions we deform.

**1.1.3 No-Regret Learning in Queuing Systems.** As mentioned earlier, no-regret learning is a classical modeling assumption in the context of repeated, independent games that often attains equivalent price of anarchy-style guarantees to that of Nash equilibria. Moreover, the

convergence of no-regret learning to a correlated equilibrium gives an intrinsic game-theoretic justification for using it as a behavioral model of the agents. In our setting, because queues only receive bandit feedback on their actions, but otherwise may not know the service rates  $\mu$  nor any other information on the number or identities of the other queues in the systems or their choices of servers, it is natural to assume that they may use an adaptive learning algorithm to select servers in each round. Indeed, in large systems, reasoning about explicit game-theoretic actions on a round-by-round basis may prove difficult or impossible due to these information constraints, but no-regret learning is nonetheless achievable and gives useful performance guarantees.

However, we now provide a simple example showing that no-regret behavior can exhibit surprisingly *myopic behavior* in our queuing model. In this example, there is a unique no-regret policy for each agent, given the behavior of the other agent, but both agents would have been better off long-term had one even *slightly* deviated to an inferior server and the other stayed the same. In the classical setting with “independent” repeated games, this cannot occur.

*Example 1.4.* Suppose that there are two queues with arrival rates  $\lambda_1 = \lambda_2 = .51$  and two servers with  $\mu_1 = 1$  and  $\mu_2 = .49$ . In this case, each queue receives a new packet roughly once every two periods on average. One can show that if both servers send to the top rate server every period, the sequence of play will satisfy the no-regret property, as they roughly split the top server equally. Each server then roughly clears at a rate of  $1/2$ , which is strictly better than deviating to the lower server, but this system will not be stable; packets arrive at a total rate of  $1.02$  but are cleared at a rate of  $1$  in expectation, leading to linear growth. However, if one queue commits to slightly deviate toward the inferior server, one can show that the resulting system will be stable.

Note that this behavior cannot arise in the patient version of the queuing game considered earlier, because the incentives explicitly favor smaller *long-run* growth rates. This example shows that no-regret learning, which considers less sophisticated measures of the efficiency of the servers without considering the long-run effects, exhibits myopia. The queue sending to the second server in the preceding example does have regret in the classical sense, despite doing better in the long term.

We thus see that no-regret outcomes could be susceptible to performance losses compared to the patient setting. Our second main result precisely captures the performance of generic no-regret dynamics. We show that under no-regret dynamics, the system still remains stable when there is a factor 2 extra slack in server capacity: if the system has enough capacity to serve all packets when they are centrally coordinated even with *half the service rates*, then no-regret learning of the queues guarantees that queue lengths stay bounded in expectation across time. We will prove the following theorem.

**THEOREM 1.5 (THEOREM 5.2, INFORMAL).** *If each queue in the queuing system satisfies the no-regret property in choosing servers at each round, and if for all  $k \leq n$ ,*

$$\sum_{j=1}^{\min\{k, m\}} \mu_j > 2 \sum_{i=1}^k \lambda_i,$$

*then the queuing system under these dynamics is stable.*

We complement these results by providing a partial converse that this factor of 2 on the required service rate is tight in Theorem 5.4, in that with less than a factor of 2 higher service rate, no-regret outcomes do not necessarily guarantee stability. Taken in tandem with our results on the patient queuing game, we thus observe rather subtle behavior that can arise in games with state. No-regret learning can indeed attain nontrivial performance guarantees, but the myopia induced by



the local dynamics may lead to performance losses compared to the setting where agents compete with stationary, patient strategies.

Our key technique in proving Theorem 1.5 is to use a delicate potential function argument. The main idea of our proof is to argue that when some potential function, which we must construct, has a high enough value, then it must have negative drift. To conclude that the queue sizes remain bounded in expectation, we can then employ a powerful theorem of Pemantle and Rosenthal [29] showing that a sufficiently regular stochastic process with negative drift must have moments uniformly bounded over time.

To carry out this approach, the key difficulty now becomes: what potential function (related to queue lengths/ages) should we choose that provably has negative drift when it is large? The simplest possible choice is the maximum queue age, which is an  $\ell_\infty$  potential. Indeed, we will argue that the oldest queues, by virtue of the slack, no-regret condition, and the priority, will tend to decrease in age *in aggregate*. However, the dependencies from the priority scheme and the learning dynamics make arguing about how this decrease is spread among old queues rather tricky. Moreover, this potential does not alone sufficiently account for the full state of the system; to try to benefit from the performance of all queues, one could instead try an  $\ell_1$ -style potential function. However, this potential has a different problem; the gains by older queues could be washed out by the aging of young queues that do not have priority for this choice of potential.

To that end, it will become most convenient to instead study an  $\ell_2^2$  potential function of queue ages. This potential function construction can be motivated in multiple ways. First, squaring queue ages naturally biases the potential to the older queues as for an  $\ell_\infty$  potential, which we will argue are decreasing in aggregate. Moreover, we will want to benefit at *all scales* like in an  $\ell_1$  potential: to do so, one is naturally led to summing an  $\ell_1$ -potential over just those queues above each age threshold. Upon doing some algebra, one arrives at our  $\ell_2^2$  potential function of queue ages. For technical reasons, we eventually translate back to a suitable  $\ell_2$ -style norm of queue ages. It is our hope that the kinds of qualitative features we establish and the methods of proof for these results will be of interest in the future study of repeated games that similarly relax the independence assumptions of the games played at each round.

## 1.2 Organization

We first formalize the strategic queuing model in Section 2. In doing so, we will formally establish our notions of stability and feasibility that will give the benchmark for our results with strategic agents. We then turn to our first main result on patient queuing systems at equilibrium. In Section 3, we introduce and prove several properties of the relevant queuing game. We provide an alternate, algorithmic description of the cost functions for the queues, and use the resulting description to prove various game-theoretic and structural properties of these systems. Using these results, we prove our tight bound on the price of anarchy in this patient queuing game in Section 4 assuming that this alternative characterization of long-run rates is valid. In Section 5, we prove our first main result showing that with a factor 2 extra slack, no-regret queues will remain stable. In Section 6, we return to formally showing that the algorithmic description of the cost functions indeed coincides with their original definition as long-run growth rates of the queues.

## 1.3 Related Work

Our work falls in a long tradition of establishing price of anarchy bounds for various games [24], but is one of the first examples of studying the effect of learning in games with carryover effects between rounds. Compared to classical price of anarchy bounds in repeated games [4, 32, 40], we no longer make the assumption that games at different rounds are independent. Studying this model requires us to combine ideas from the price of anarchy analysis of games with the theory of

stochastic processes. Another important repeated game setting with such carryover effect is the repeated ad-auction game with limited budgets. Some works [8, 11, 12] consider such games and offers results on convergence to equilibrium as well as understanding equilibria in the first-price auction settings under a particular behavioral model of the agents. Analyzing such systems for the more commonly used second-price auction system is an important open problem.

Although our stability objective differs from usual objectives in this literature, our results qualitatively also resemble the bicriteria result of Roughgarden and Tardos [34], which shows that in nonatomic routing, the cost incurred at any Nash flow is at most the optimal cost when twice the flow is routed. Unlike most such bounds that follow the *smoothness* framework of Roughgarden [32], our second main result is an *equilibrium analysis* that is more similar to that of Johari and Tsitsiklis [23], who establish equilibrium conditions and modify their problem while maintaining the equilibrium condition to arrive at a version that is easy to analyze. In our argument, we also modify the equilibrium itself toward a more tractable solution, but the intermediate points in this deformation will *not* be Nash, requiring additional arguments.

Our queuing setting also bears resemblance to *stochastic games*, a generalization of Markov decision processes where multiple players competitively and jointly control the actions and transitions (e.g., see [16, 28]). However, our work differs from this line in multiple ways: in our second model, queues are unaware of the system state and parameters, and most importantly, we are interested in explicit bounds to derive price of anarchy-style results for stability.

Although the goal of our work is in establishing price of anarchy-style bounds in dependent systems, this necessitates a careful understanding of the analytic properties of our random queuing dynamics. Among the large body of literature studying highly dependent random processes, closest to our work is the adversarial queuing systems of Borodin et al. [9], who also use the Pemantle and Rosenthal [29] theorem to establish bounded queue sizes in expectation.

The classical focus of work on scheduling in queuing systems is to identify policies that achieve optimal throughput (e.g., see the textbook of Shortle et al. [39]). There has also been work both on evaluating efficiency loss due to selfishness in different classical queuing systems, as well as the role of learning in such systems. For work on price of anarchy in queuing systems, see the book of Hassin [21] and the survey of Hassin and Haviv [22], and for a very recent tutorial on the role of learning and information in queuing systems, see the work of Walton and Xu [41]. Closest to our model from this literature is the work of Krishnasamy et al. [25], which characterizes the queue-regret of learning algorithms that only seek to identify the best servers but does not consider competition between selfish learners. Their primary goal is to study this more refined notion in the queuing setting for this classical stochastic bandit problem, which exhibits more complicated behavior than standard no-regret bounds that grow at least logarithmically with time. They characterize queue-regret for the case of a single queue aiming to find the best server, and extend the result to the case of multiple queues scheduled by a single coordinated scheduling algorithm, assuming that there is a perfect matching between queues and optimal servers that can serve them. In contrast, we assume that each queue separately learns to selfishly make sure its own packets are served at the highest possible rate, offering a strategic model of scheduling packets in a queuing system. Furthermore, we do not make the matching assumption on queues and servers.

*Subsequent Work.* After a subset of these results originally appeared in conference form, several works studied several game- and learning-theoretic questions arising from our model and technical results. On the strategic side, Fu et al. [18] show that the probabilistic techniques we introduce for analyzing these queuing systems extend to more general queuing networks once the feasibility conditions under coordinated scheduling are suitably adjusted, which naturally arise via duality

arguments. Baudin et al. [5] propose an alternative, episodic queuing system where agents have incentives to hold jobs in an episode before sending to a central server, but suffer penalties should their jobs not be completed before the end of the episode. Their main conclusion, similar to ours, is that both equilibrium and no-regret outcomes ensure stability as long as these costs are sufficiently large. On the learning-theoretic side, Sentenac et al. [37], as well as Freund et al. [17], consider the problem of decentralized learning dynamics in bipartite queuing systems that attain near-optimal performance, extending the original work of centralized learning by Krishnasamy et al. [25]. These algorithmic advances are incomparable with our results, which are inherently noncooperative and strategic in nature. The former work also shows that the more refined guarantee of *policy regret* [14] is not sufficient to bridge the quantitative gap between our positive results for no-regret learning and patient equilibria.

## 2 PRELIMINARIES

### 2.1 Notation

In general, random variables will be denoted by capital letters (i.e.,  $X, Y, Z, \dots$ ), whereas vectors will be bolded (i.e.,  $\boldsymbol{\mu}, \boldsymbol{\lambda}$ , etc). If a random variable  $X$  has some distribution  $\mathcal{D}$ , we write  $X \sim \mathcal{D}$ . We use the notation  $\text{Geom}(p)$  to denote a geometric distribution with parameter  $p$ ,  $\text{Bern}(p)$  for a Bernoulli distribution that is 1 with probability  $p$  and 0 otherwise, and  $\text{Bin}(n, p)$  for a binomial distribution with parameters  $n$  and  $p$ .

We say that an event occurs *almost surely* if it has probability 1. We use standard  $O(\cdot)$ ,  $o(\cdot)$ , and  $\Theta(\cdot)$  notation. We will sometimes write  $f(n) \asymp g(n)$  if  $f(n) = \Theta(g(n))$ . We will also consider the following norms: for a positive vector  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)$ , with  $\lambda_1 \geq \dots \geq \lambda_n > 0$ , we define the following two weighted  $\ell_p$  norms on  $\mathbb{R}^n$ :  $\|\mathbf{x}\|_{\lambda,1} \triangleq \sum_{i=1}^n \lambda_i |x_i|$  and  $\|\mathbf{x}\|_{\lambda,2} \triangleq \sqrt{\sum_{i=1}^n \lambda_i x_i^2}$ . It is easily seen that for any  $\mathbf{x}$ ,  $\|\mathbf{x}\|_{\lambda,1} \asymp \|\mathbf{x}\|_{\lambda,2}$  (where the constants depend on  $\boldsymbol{\lambda}$ ) via Cauchy-Schwarz (see Lemma A.1). We use the following *fractional sum* operation  $\oplus : \mathbb{R}_{\geq 0}^2 \times \mathbb{R}_{\geq 0}^2 \rightarrow \mathbb{R}_{\geq 0}$ :

$$\frac{a}{b} \oplus \frac{c}{d} \triangleq \frac{a+c}{b+d}.$$

We will later repeatedly use the following simple fact.

FACT 2.1. *For all  $a_1, \dots, a_n \geq 0$  and  $b_1, \dots, b_n > 0$ ,*

$$\min_{i \in [n]} \frac{a_i}{b_i} \leq \frac{a_1}{b_1} \oplus \dots \oplus \frac{a_n}{b_n} = \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \leq \max_{i \in [n]} \frac{a_i}{b_i}.$$

*Moreover, equality holds in either of the inequalities if and only if both inequalities are tight.*

Given a  $n \cdot m$ -dimensional vector  $\mathbf{p} = (p_1, \dots, p_n)$ , where  $p_i \in \mathbb{R}^m$ , we will write  $p_{ij}$  for the  $j$ th element of  $p_i$ . Given a vector  $\mathbf{x} \in \mathbb{R}^n$  and a subset  $I \subseteq [n]$ , we write  $\mathbf{x}_I$  to denote the vector restricted to the components in  $I$ . Given a set  $S$ , we will write  $\mathcal{P}(S)$  to denote the power set.

### 2.2 Bernoulli Queuing Model

We consider the following discrete-time queuing system illustrated by Figure 1, which is a decentralized, competitive version of the model considered by Krishnasamy et al. [25]: there is a system of  $n$  queues and  $m$  servers. During each discrete timestep  $t = 0, 1, \dots$ , the following occurs:

- (1) Each queue  $i$  receives a new packet with a fixed, time-independent probability  $\lambda_i$ . We model this via an independent random variable  $B_t^i \sim \text{Bern}(\lambda_i)$ . This packet has a timestamp that indicates that it was generated at time  $t$ . We label queues such that  $1 > \lambda_1 \geq \dots \geq \lambda_n > 0$ , writing  $\boldsymbol{\lambda}$  for the vector of arrival rates.

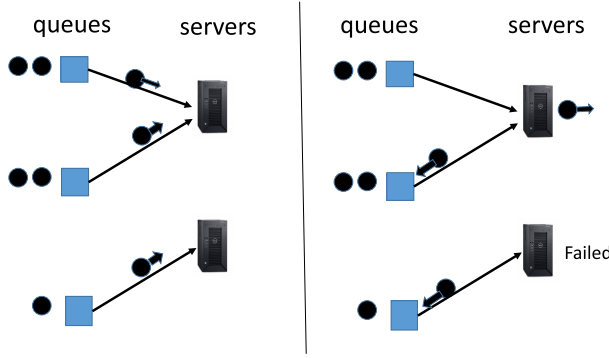


Fig. 1. Left: Three queues compete for two servers. Right: Unserved packets in each round return to their queue.

- (2) Each queue that currently has an uncompleted packet chooses one server to send their oldest unprocessed packet (in terms of timestamp) to.
- (3) Each server  $j$  that receives a packet does the following. First, it only considers the packet it receives with the oldest timestamp (breaking ties arbitrarily). It then processes this packet with a fixed, time-independent probability  $\mu_j$ . We again label servers so that  $\mu_1 \geq \dots \geq \mu_m \geq 0$ , writing  $\boldsymbol{\mu}$  for the vector of service rates.
- (4) All unprocessed packets, possibly including the packets that were selected if the corresponding server failed to process it, are then sent back to their respective queues still uncompleted. Queues receive bandit feedback on whether their packet cleared at their chosen server.

At each round of this process, the queues with packets aim to maximize the probability that their packet gets served, describing the incentives of the stage game.

We write  $Q_t^i$  for the number of unprocessed packets of queue  $i$  at the beginning of time  $t$  (before sampling new packets) and  $\mathbf{Q}_t = (Q_t^1, \dots, Q_t^n)$  for the vector of queue sizes at time  $t$ . Define  $Q_t = \sum_{i=1}^n Q_t^i$  as the total number of unprocessed packets in the system at time  $t$ . Formally, if  $X_t^i$  is the indicator event that queue  $i$  clears a packet at time  $t$  and  $B_t^i$  is again the indicator queue  $i$  received a new packet at time  $t$ , then we have the recurrence as random variables with  $Q_0^i = 0$  and

$$Q_{t+1}^i = Q_t^i + B_t^i - X_t^i, \quad (2)$$

where we note that  $X_t^i$  is necessarily 0 if  $Q_t^i + B_t^i = 0$  (i.e., queue  $i$  had no packets and did not receive a new one in the round, so it does not send nor clear a packet this time period). This ensures that each  $Q_t^i$  is integral and nonnegative. We call the preceding random process the *Bernoulli model*. We will be interested in the stability of this system in the following sense.

**Definition 2.1.** The preceding system is *strongly stable* under some given dynamics if, for any fixed  $r \geq 0$ , the random process  $Q_t$  satisfies  $\mathbb{E}[(Q_t)^r] \leq C_r$  for some absolute constant  $C_r$  depending only on  $r$  and the parameters  $\boldsymbol{\mu}$  and  $\boldsymbol{\lambda}$ , but not on  $t$ . If instead we only have  $\mathbb{E}[Q_t] = o(t)$  (i.e., sublinear expected growth), we say that the system is *weakly stable*.<sup>4</sup>

To get a baseline measure for the outcomes that could arise under strategic behavior, we must first understand when a queuing system is stable under centralized coordination: it turns out that an obvious necessary condition on  $\boldsymbol{\mu}$  and  $\boldsymbol{\lambda}$  is also sufficient. An instructive example to keep in

<sup>4</sup>While not immediately obvious, we will show in Appendix C.3 that strong stability implies almost sure subpolynomial growth.

mind is a single-queue, single-server system. Of course, there is no learning or competition in such a process. If  $0 < \lambda < \mu \leq 1$ , it is well known that  $Q_t^1$  follows a random walk on the nonnegative integers biased toward 0, and moreover is geometrically ergodic.<sup>5</sup> This in particular implies strong stability. However, if  $0 < \lambda = \mu < 1$ , then the corresponding unbiased random walk  $Q_t^1$  satisfies  $\mathbb{E}[Q_t^1] = \Theta(\sqrt{t})$ . Therefore, there is a sharp threshold for strong stability at  $\lambda = \mu$ . Our first result, which provides a natural baseline to compare with all of the subsequent results in this work, offers an extension of this for the feasibility of any coordinated policy for multiple queues and servers.

**THEOREM 2.2.** *Let  $\lambda \in (0, 1)^n$  and  $\mu \in [0, 1]^m$  be the arrival and service rates, respectively. Then the preceding queuing system is strongly stable for some centralized (coordinated) scheduling policy if and only if for all  $1 \leq k \leq n$ ,*

$$\sum_{j=1}^{\min\{k, m\}} \mu_j > \sum_{i=1}^k \lambda_i. \quad (3)$$

When Equation (3) holds, we say that the queuing system is (centrally) feasible.<sup>6</sup>

We give the proof of sufficiency in this section while deferring the (straightforward) proof of necessity to Appendix C, which requires a submartingale argument in the edge cases where some inequality is tight. To prove sufficiency, we require the well-known notions of majorization of vectors.

**Definition 2.3.** Let  $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$ , and assume that  $x_1 \geq x_2 \geq \dots \geq x_n \geq 0$  and analogously for  $\mathbf{y}$ . Then  $\mathbf{x}$  *majorizes*  $\mathbf{y}$  (written  $\mathbf{x} \succeq \mathbf{y}$ ) if for each  $1 \leq k \leq n$ ,

$$\sum_{i=1}^k x_i \geq \sum_{i=1}^k y_i.$$

If the preceding inequalities are strict for each  $1 \leq k \leq n$ , then  $\mathbf{x}$  *strictly majorizes*  $\mathbf{y}$  (written  $\mathbf{x} \succ \mathbf{y}$ ).<sup>7</sup>

With this definition in hand, we can prove sufficiency using standard results on the relation between majorization and doubly stochastic matrices (see Appendix C for these standard facts).

**PROOF OF SUFFICIENCY IN THEOREM 2.2.** First, suppose that  $\mu$  strictly majorizes  $\lambda$ , when appropriately appending zeros if needed to make the vectors of same length. By Corollary C.3, there exists some doubly stochastic  $P$  such that  $P\mu > \lambda$  elementwise. By the Birkhoff-von Neumann theorem, the set of doubly stochastic matrices is the convex hull of the set of permutation matrices  $\mathcal{P}$ . This implies that there exists a distribution  $\pi$  over  $\mathcal{P}$  such that  $\lambda < P\mu$ , interpreted componentwise, where  $P = \mathbb{E}_{\Pi \sim \pi}[\Pi]$ .

Consider the following oblivious scheduling algorithm: at each time  $t$ , independently sample a permutation matrix  $\Pi$  from  $\pi$ , and schedule queues via the associated matching on the bipartite graph of queues and servers (where we pad with zeros if necessary, and even if some queues have no available packets to send). For each queue  $i$ , the associated marginal distribution on servers it sends to in each round is given by the  $i$ th row of  $P$ . Given that queue  $i$  has a packet to send at time  $t$ , the probability of successfully clearing a packet is exactly  $(P\mu)_i > \lambda_i$ , as this scheduling scheme

<sup>5</sup>Namely, this random process mixes to a stationary distribution on  $\mathbb{N}$  with geometrically decreasing tail probabilities.

<sup>6</sup>We remark that the analogous statement for weak stability holds if the inequalities in Equation (3) only weakly hold by the same proof.

<sup>7</sup>Majorization is often used for probabilities, and hence defined so that the total sums are equal; we omit this condition in our work.



ensures that each queue is alone at each server it sends to. As a result, the packet clears as long as the server is successful.

Therefore, the random process  $Q_t^i$  of number of packets by queue  $i$  at time  $t$  follows a homogeneous random walk on the half-line biased toward 0, which is ergodic with a stationary distribution with geometric tails. It is not difficult to show that any distribution on the natural numbers with geometric tails has bounded  $r$ th moments for any  $r \geq 0$ .<sup>8</sup> This then extends to the  $r$ th moment of the sum by Minkowski's inequality. This proves strong stability when strict majorization holds.  $\square$

### 2.3 The Need for Packet Priorities

Recall that we will be interested in proving statements of the following form:

Given a queuing system that is centrally feasible (in the sense of Theorem 2.2) even when  $\mu$  is scaled down by some constant  $c \geq 1$  independent of the parameters of the system, then a random process where queues are decentralized and strategic under certain conditions remains stable.

To see the necessity of timestamps, consider instead a simpler model where there are no timestamps and priorities, and instead each server uniformly randomly picks which packet to process among those that are sent to it in each step. It is easy to see that if a queuing system is feasible even if  $\mu$  is scaled down by a factor of  $n$ , then it will remain a stable queuing system with reasonably strategic queues. Indeed, by this feasibility assumption,  $\mu_1 > n \cdot \lambda_1$  so that  $\mu_1 > \sum_{i=1}^n \lambda_i$ . Therefore, if every queue just always sends to the largest server whenever they have a packet, they will succeed in clearing a packet with probability at least  $\mu_1/n > \lambda_i$ , and it is not too difficult to prove that this results in a strongly stable process by comparing to a random walk biased toward the origin. It is natural to ask if a better factor is attainable in this alternate model, perhaps even a constant. It turns out that in general, a polynomial in  $n$  is required, as we formally show in Appendix C.

**THEOREM 2.4.** *In this alternate model, for large enough  $n$ , there exists a centrally feasible queuing system with  $n$  queues and servers with the following property: the system remains feasible even if  $\mu$  is scaled down by  $\Omega(n^{1/3})$  and it is possible for all queues to be in a Nash equilibrium of the stage game at each step<sup>9</sup> at each timestep (and in particular, satisfy no-regret properties as in Assumption 5.1), yet the system is not weakly stable.*

The basic reason this can occur is that low arrival rate queues can saturate the high success rate servers, making it impossible for high arrival rate queues to clear fast enough to offset their higher arrival. Our priority model, although more difficult to analyze, results in older queues gaining an advantage on young queues, causing the young queues to prefer lower-quality servers. In other words, our model implicitly helps fast-growing queues get better service, as long as queues are sufficiently adaptive to take advantage of it.

### 2.4 Geometric Queuing Model via Deferred Decisions

Because of the significant dependencies induced by the timestamp priority scheme in our model, a direct analysis of these systems proves quite unwieldy. In this section, we use the principle of deferred decision to give an alternate description of the Bernoulli system given in Section 2.2 that will prove significantly more amenable to analysis in the rest of this article.

To describe this Geometric model, suppose that each queue chooses which server to send to at time  $t$  only depending on past feedback and their history of oldest timestamps, but *not on the  $Q_t^i$* .

<sup>8</sup>This can also easily be seen directly using Theorem 5.3. Negative drift when exceeding  $Q_t = 0$  is obvious, and as queue sizes can change by at most  $n$  in total between steps, increments are clearly bounded in  $L^p$  for any  $p \geq 0$ .

<sup>9</sup>By this, we mean that conditioned on the (randomized) strategies of all other queues in a given timestep, each queue sends to a server with the highest probability of success.

In this case, we can equivalently characterize the evolution of this system keeping only the oldest timestamp of a packet at each queue. To do this, instead of randomly generating new packets at each timestep according to a Bernoulli process, each queue only maintains the timestamp of their current oldest unprocessed packet. Once this packet is successfully cleared, the new current oldest unprocessed packet has a timestamp generated by sampling a geometric random variable with parameter  $\lambda_i$  and adding this to the timestamp of the just-completed packet. If this number exceeds the current timestep  $t$ , this corresponds to having processed all packets that arrived before the current timestep, and receiving the next packet in the future. We will call this random process the *Geometric process*. Because the gap between successes in repeated independent  $\text{Bern}(\lambda_i)$  trials is given by a  $\text{Geom}(\lambda_i)$  random variable, the Bernoulli and Geometric processes can be completely coupled, as described in the following.

*Definition 2.5.* When the queues use strategies with the preceding property, the *Geometric process* can be described using the following notation:

- (1) Time progresses in discrete steps  $t = 0, 1, \dots$ . At each time  $t$ ,  $\tilde{T}_t^i$  is the *timestamp* of the oldest unprocessed packet of queue  $i$  at time  $t$ .  $T_t^i = \max\{0, t - \tilde{T}_t^i\}$  is the *age* of the current oldest packet of queue  $i$  at timestep  $t$ . In other words,  $T_t^i$  measures how old the current oldest unprocessed packet for queue  $i$  is.<sup>10</sup>
- (2) Queue  $i$  can send a packet to any server  $j$  in this timestep if  $t - \tilde{T}_t^i \geq 0$ . Each server  $j$  attempts to serve only the oldest packet it receives and succeeds with probability  $\mu_j$ . If queue  $i$ 's packet is successfully served, set  $\tilde{T}_{t+1}^i = \tilde{T}_t^i + G^i$ , where  $G^i \sim \text{Geom}(\lambda_i)$  is independent of all past events, and otherwise  $\tilde{T}_{t+1}^i = \tilde{T}_t^i$ .

We write  $\mathbf{T}_t = (T_t^1, \dots, T_t^n) \in \mathbb{N}^n$  for the vector of current ages of oldest packets. To see the equivalence, consider any Bernoulli queuing system with Bernoulli random variables  $\{B_t^i\}_{i \in [n], t \geq 0}$  for packet generation. To get a coupled Geometric system for the same system, use an independent sequence  $\{G_s^i\}_{i \in [n], s \geq 0}$  with the interpretation that  $G_s^i \sim \text{Geom}(\lambda_i)$  is the size of the  $s$ th gap between successes in the  $B_t^i$ . When queue  $i$  clears her  $s$ th packet, her new oldest timestamp increases by  $G^i = G_s^i$  as described earlier. As such gaps between timestamps in the Bernoulli model have  $\text{Geom}(\lambda_i)$  distributions, the Geometric system gives the ages of each queue in the Bernoulli system at all times and gives an explicit coupling.

The key feature is that, under the assumption that each queue  $i$  chooses servers at time  $t$  only based on at most the  $T_t^i$ , not on  $Q_t^i$ , all choices by queues are the same conditioned on just the current timestamp and past feedback as it is conditioned on all the past information in the Bernoulli model (which includes arrivals received after the current oldest packet). In other words, if  $\mathcal{G}_t$  denotes the information available to the queues in the Bernoulli model at time  $t$ , and  $\mathcal{F}_t$  for the Geometric model, then all choices by the queues at time  $t + 1$  are the same conditioned on either history. The point of doing so is that  $G_s^i$  will be independent of  $\mathcal{G}_t$  until the queue clears her  $s$ th packet (namely, the timestamp of queue  $i$ 's  $s + 1$ -th packet is not known until the time queue  $i$  clears her  $s$ th packet).

In the Geometric system, we define stability in the same way as before.

*Definition 2.6.* The Geometric system is *strongly stable* if, for any  $r \geq 0$ ,  $\mathbb{E}[(\sum_{i=1}^n T_t^i)^r] \leq C_r$ , where  $C_r$  is a constant depending only on  $r$  and the parameters  $\mu$  and  $\lambda$ , but not on  $t$ . The system is *weakly stable* if just  $\mathbb{E}[\sum_{i=1}^n T_t^i] = o(t)$ .

<sup>10</sup>Note that although  $T_t^i \geq 0$  by definition, it is possible that  $\tilde{T}_t^i > t$ . The interpretation of this is that the queue has cleared all of her packets at time  $t$  and will receive her next one at time  $t = \tilde{T}_t^i$ , or equivalently, in  $\tilde{T}_t^i - t$  steps in the future from the perspective at time  $t$ .

Because heuristically  $Q_t^i \approx \lambda_i T_t^i$ , it is intuitive that our notions of strong stability are equivalent whenever both systems correspond to the same random process. This is indeed the case, and furthermore, strong stability implies almost sure subpolynomial asymptotic growth. The basic idea is to use Markov's inequality and the Borel-Cantelli lemma along an appropriately chosen subsequence of times and then interpolate to the rest. We defer this equivalence to Appendix C.3.

**LEMMA 2.7.** *If the Bernoulli and Geometric models characterize the same queuing dynamics, then strong stability in the Bernoulli system is equivalent to strong stability in the Geometric system (and the same holds for weak stability). Moreover, if this holds, then strong stability in either system implies almost sure subpolynomial asymptotic growth.*

### 3 PATIENCE IN QUEUING SYSTEMS

In this section, we introduce and establish structural results about a patient version of our queuing game. In this game, each queue chooses a *fixed* distribution over servers that will be used in all rounds to optimize the long-run age of the packets in the queue. To begin, we formulate the game in a manner that is well defined *a priori*.

**Definition 3.1.** The *patient queuing game*  $\mathcal{G} = ([n], (c_i)_{i=1}^n, \mu, \lambda)$  is defined as follows: the strategy space for each queue is  $\Delta^{m-1}$ . Let  $\mathbf{p} = (p_1, \dots, p_n) \in (\Delta^{m-1})^n$  denote the vector of fixed distributions for all queues over servers. The cost function  $c_i$  for queue  $i$  is  $c_i(p_i, p_{-i}) \triangleq \limsup_{t \rightarrow \infty} \mathbb{E}[\frac{T_t^i}{t}]$ ,

where  $T_t^i$  is again the age of queue  $i$  at time  $t$  in the random queuing process induced by running the queuing system with  $\mu$  and  $\lambda$  the system parameters when each queue chooses a server by independently randomizing according to  $\mathbf{p}$  in each timestep she has a packet.

We say that  $\mathbf{p}$  is a *Nash equilibrium* of  $\mathcal{G}$  if for all  $i \in [n]$ ,  $p_i \in \arg \min_{p'_i \in \Delta^{m-1}} c_i(p'_i, p_{-i})$ —that is, each queue chooses  $p_i$  to minimize their cost function conditioned on the strategies  $p_{-i}$  of the other queues.

Note that the cost function defined previously is clearly well defined as the lim sup of expected values. However, we will actually show that the limit of the random quantity  $T_t^i/t$  (without expectations) is almost surely equal to a deterministic constant depending on  $\mathbf{p}, \lambda, \mu$  (see Theorem 3.3). By deriving an alternate, explicit characterization of these values, we show that Nash equilibria exist in Theorem 3.8. Because the cost functions are explicit functions of the randomizations and the parameters  $\mu$  and  $\lambda$ , we omit the notation  $(c_i)_{i=1}^n$  when instantiating a game  $\mathcal{G}$ .

Our main focus in Section 4 will be to give guarantees on the quality of all Nash equilibria in this game. In a slight abuse of the price of anarchy terminology, we make the following definition.

**Definition 3.2.** Let  $\lambda$  and  $\mu$  satisfy the conditions of Theorem 2.2 with weak inequalities so that  $\mu$  majorizes  $\lambda$ . For  $\alpha \geq 1$ , let  $\mathcal{G}(\alpha) = ([n], \mu, \alpha^{-1}\lambda)$ , and let  $\mathcal{N}(\alpha)$  denote the set of Nash equilibria of  $\mathcal{G}(\alpha)$ . The *price of anarchy* of  $\mathcal{G} = \mathcal{G}(1)$  is defined as the supremum of  $\alpha$  values such that there exists a Nash equilibrium  $\mathbf{p}^* \in \mathcal{N}(\alpha)$  and some  $i \in [n]$  such that  $c_i(\mathbf{p}^*) > 0$  in  $\mathcal{G}(\alpha)$ .<sup>11</sup> The price of anarchy of the patient queuing game is the supremum over all instances of  $\mathcal{G}$ .

In this section, we extensively study the properties of the cost function  $c(\mathbf{p})$ , which is currently written as the lim sup of the expected value of the random linear aging rate of each queue. By taking the lim sup and expected values, the cost function is well defined, albeit quite unwieldy at present. Our first task is thus to provide an alternative, algorithmic description of  $c(\mathbf{p})$ , which we

<sup>11</sup>In other words, the price of anarchy of a centrally feasible system is the supremum of values of  $\alpha$  such that when all queue arrival rates are scaled down by  $\alpha$ , there nonetheless exists a Nash equilibrium and some queue that suffers nonzero linear aging.

initially denote  $r(\mathbf{p})$  (for “rates”) in Section 3.1. We show that  $r$  has significant analytic structure that will help establish various game-theoretic properties of this system. In particular, we show that the level subsets (in  $[n]$ ) of  $r(\mathbf{p})$  enjoy convenient closure properties, which will be enough to establish continuity and other properties, which we use to prove the existence of equilibria. We will return to proving that this function is equal to  $c$  in Section 6.

*The Need for Packet Priorities with Patience.* We note that even with this restriction to stationary strategies in a patient queuing game, the priority scheme by servers to attempt to serve the oldest packet is necessary to obtain constant price of anarchy bounds. It is not too difficult to see that if servers choose packets uniformly at random among those it receives in each round, the price of anarchy can be polynomially large in  $n$  in the sense of Theorem 2.4 when the costs are defined to be the asymptotic aging rates as in Definition 3.1. More specifically, consider a queuing system with one queue with arrival rate  $C/n^{1/3}$  for some large constant  $C > 0$  to be determined, and  $n - 1$  queues with arrival rate  $1/n^{2/3}$ . Suppose that there is one server with success rate 1 and then  $n$  servers with success rate  $1/n^{1/3}$ . There exist bad Nash equilibria in stationary strategies of the following form: each small queue evenly mixes between a personal server with success rate  $1/n^{2/3}$  and the top server with success rate 1. It is clear that each small queue will be stable for large enough  $n$  (provided no other queue shares her personal server) by this choice of constants, so she has no incentive to deviate because her cost in this game is her long-run aging rate of zero.

By standard Chernoff bounds, in any given round, there are at least  $cn^{1/3}$  small queues that will send to the top server under these stationary strategies with overwhelming probability for some absolute constant  $c > 0$ ; therefore, with very high probability in each round, the large queue can succeed in clearing a packet by sending to the large server with probability at most  $1/cn^{1/3}$ . This is strictly smaller than her rate if  $C > 0$  is taken sufficiently large. At any other server, the queue can get rate at best  $1/n^{1/3}$ , which clearly does not ensure stability for large enough  $C > 0$ . Therefore, the system must remain unstable if she best responds to this behavior by the other queues (note further that this best response will never involve sending to a personal server of any of the other queues, as there are  $n$  such servers). No queue has any incentive to deviate according to the cost functions as defined in Definition 3.1 under this alternate server selection choice, so these strategies constitute a Nash equilibrium. This system would remain centrally feasible even if server rates were scaled down by  $\Omega(n^{1/3})$ , and hence the price of anarchy of such a patient queuing game can be polynomially large in  $n$ .

In this example, small queues mix between the top server and their own server equally because their long-run growth rate is zero in any case. It is possible to modify this example so that all queues mix among servers that offer equal long-run probability of success. This can be done by having each small queue send to the top server with probability  $p(n)$  and to a personal server with rate  $1/n^{1/3}$  with probability  $1 - p(n)$ , while the top queue sends deterministically to the last server with rate  $1/n^{1/3}$  (note that the effectiveness of the top server is at best  $1/n^{1/3}$  if the top queue were to deviate to there, by construction, so this is a best response). The parameter  $p(n)$  can be chosen so that the long-run average number of small queues sending to the top server is precisely  $n^{1/3}$  almost surely using the strong law of large numbers for Markov chains, so that each queue indeed mixes among servers that offer long-run success probability  $1/n^{1/3}$ .

### 3.1 Algorithmic Description of Costs

As stated, we now construct a function  $r : (\Delta^{m-1})^n \rightarrow [0, 1]$  that we will show is equivalent to  $c$ . We will show that for any fixed  $\mathbf{p}$ , the set  $[n]$  of queues partitions into subsets  $S_1, S_2, \dots$ , where each queue in  $S_i$  group has the same aging rate and  $S_1$  ages the fastest, then  $S_2$ , and so on, according to  $r$  (and so for  $c$  as well). To get a sense of the quantities that will arise before considering the general case, consider the simplest setting of a single queue and a single server (where there

are no nontrivial strategies nor competition), with rates  $\lambda > \mu$ . In any round where the queue has an uncleared packet, the age will first increase by 1 deterministically. With probability  $\mu$ , the queue will succeed in clearing this packet, and the age will go down in expectation by  $\mathbb{E}[G] = 1/\lambda$ , where  $G \sim \text{Geom}(\lambda)$  is independent of whether or not the server succeeds. Therefore, the expected change in this queue's age will be  $1 - \mu/\lambda > 0$ , and we expect that the queue will asymptotically age at this rate.

In general, with multiple queues and servers, the actual values of  $c_i$  are best described via a recursive algorithm that computes the rates, which we give in the following. The intuition is that  $S_1(\mathbf{p})$  will be the subset that minimizes the ratio of expected packets they clear collectively given  $\mathbf{p}$ , assuming that they have priority over all other queues, divided by their sum of arrival rates. This quantity arises by viewing this subset as a single large queue as in the preceding single queue example. Conditioned on this set  $S_1$  of queues growing fastest, they will typically have priority, and then we recurse to find the lower groups. The algorithm begins by initializing  $k = 1$  and  $I = [n]$ :

- (1) Compute the minimum value over all nonempty subsets  $S \subseteq I$  of

$$\frac{\sum_{j=1}^m \mu_j (1 - \prod_{i \in S} (1 - p_{i,j}))}{\sum_{i \in S} \lambda_i}.$$

This gives the expected number of packets cleared by  $S$  if all queues in  $S$  send in a timestep and they have priority over all other queues, divided by their sum of arrival rates.

- (2) If this value is at least 1, then no subset of queues will have linear aging, so set  $S_k = I$ ,  $r_i(\mathbf{p}) = 0$  for all  $i \in S_k$ , and terminate. Otherwise, set  $S_k$  to be the minimizer of the previous quantity over all nontrivial subsets of  $I$ , chosen to be of largest cardinality in the case of degeneracies.<sup>12</sup> In this case, for each  $i \in S_k$ ,  $r_i(\mathbf{p})$  gets set to

$$1 - \frac{\sum_{j=1}^m \mu_j (1 - \prod_{i \in S_k} (1 - p_{i,j}))}{\sum_{i \in S_k} \lambda_i}.$$

For  $k = 1$ , we refer to *any* subset with the minimum ratio as a *tight*, or *minimizing*, subset.

- (3) Update the server rates  $\mu_j$  as  $\mu_j \leftarrow \mu_j \prod_{i \in S_k} (1 - p_{i,j})$ . In other words,  $\mu_j$  gets discounted by the probability a queue from  $S_k$  sends to server  $j$  (assuming that all of these queues are sending). Update  $I \leftarrow I \setminus S_k$ ,  $k \leftarrow k + 1$ , and recurse on  $I$  with  $\boldsymbol{\mu}$  and  $\mathbf{p}_I$  if nonempty.

As many of these quantities will appear often, we make the following conventions: for any subsets  $S, S'$  such that  $S \subseteq [n] \setminus S'$ , define  $\lambda(S)$  as the sum of arrival rates of packets to a set of queues  $S$ , and  $\alpha(S|\mathbf{p}, \boldsymbol{\mu}, S')$  as the expected number of packets cleared from queues in  $S$  with service rates  $\boldsymbol{\mu}$ , if the queues in  $S'$  have priority,  $S$  has priority over all other queues, and all queues in  $S \cup S'$  send packets in the round:

$$\alpha(S|\mathbf{p}, \boldsymbol{\mu}, S') \triangleq \sum_{j=1}^m \mu_j \prod_{i \in S'} (1 - p_{i,j}) \left( 1 - \prod_{i \in S} (1 - p_{i,j}) \right), \quad \lambda(S) \triangleq \sum_{i \in S} \lambda_i,$$

and then let

$$f(S|\mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\lambda}, S') \triangleq \frac{\sum_{j=1}^m \mu_j \prod_{i \in S'} (1 - p_{i,j}) (1 - \prod_{i \in S} (1 - p_{i,j}))}{\sum_{i \in S} \lambda_i} = \frac{\alpha(S|\mathbf{p}, \boldsymbol{\mu}, S')}{\lambda(S)}$$

denote the ratio of expected number of packets cleared by  $S$  when having priority over all members but  $S'$ , normalized by the expected number of new packets received in each round by  $S$ .

<sup>12</sup>We show in Lemma 3.5 that this choice is unique and canonical.



Let  $S_k(\mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\lambda})$  be the  $k$ th set output by the preceding algorithm. When  $\mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\lambda}$  are clear from context, we will suppress them. We write  $U_k = \cup_{\ell=1}^k S_\ell$  as the set of queues in the top  $k$  groups outputted by the algorithm, with  $U_0 = \emptyset$ . We will write  $f_k = f(S_k | U_{k-1})$ , and we use  $g_k = \max\{0, 1 - f_k\}$  for the rate of the  $k$ th outputted set, which is equal to  $r_i(\mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\lambda})$  for any  $i \in S_k(\mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\lambda})$ . From the recursive construction,

$$S_{k+1}(\mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = S_1(\mathbf{p}_{[n] \setminus U_k}, \boldsymbol{\mu}', \boldsymbol{\lambda}_{[n] \setminus U_k}), \quad (4)$$

where  $\mu'_j = \mu_j \prod_{i \in U_k(\mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\lambda})} (1 - p_{ij})$  for all  $j \in [m]$ . In other words, having found  $U_k$ ,  $S_{k+1}$  is the largest minimal set among the remaining elements, but where the  $\boldsymbol{\mu}$  rates have been reweighed by the probability no element of  $U_k$  sends to each server. These quantities are compiled in a table in Appendix D for easy reference.

Our main probabilistic result about the function  $r$  is that this is indeed equivalent to the cost function  $c$  of the patient queuing games. In fact, we prove the following, stronger result.

**THEOREM 3.3 (ALMOST SURE ASYMPTOTIC CONVERGENCE).** *Let  $\mathcal{G} = ([n], \boldsymbol{\mu}, \boldsymbol{\lambda})$  be a one-shot patient queuing game. For each fixed  $\mathbf{p}$  and all  $i \in [n]$ , almost surely the long-run aging rate of queue  $i$  satisfies*

$$c_i(\mathbf{p}) = \lim_{t \rightarrow \infty} \frac{T_t^i}{t} = r_i(\mathbf{p}).$$

We will prove our game-theoretic results assuming this theorem; however, as the proof is quite nontrivial and rather involved, we defer the proof to Section 6.

### 3.2 Properties of Rate Function

We first establish basic properties of the output of the algorithm that will be useful in studying the analytic properties, as well as in proving that this algorithm gives the correct asymptotic rates. Throughout, we will view  $f$  as the quotient  $\alpha/\lambda$  when invoking Fact 3.1.

Clearly, for fixed  $S$ , the function  $f(S|T)$  is nonincreasing in  $T$  as a set function. We repeatedly use the following fact, which can be seen simply by expanding the definition of  $f$ .

**FACT 3.1.** *Suppose  $S, S', T$  are such that  $S, S' \subseteq [n] \setminus T$  and are disjoint. Writing  $f$  in the form of the quotient  $\alpha/\lambda$ , then*

$$f(S \cup S' | T) = f(S|T) \oplus f(S' | S \cup T).$$

Next, we characterize some structure in the minimizing subsets at each step of the algorithm, which will allow us to choose the  $S_k$  canonically as the largest cardinality minimizer. To do this, we first show that the function  $\alpha(\cdot)$  is *submodular*.

**LEMMA 3.4 (SUBMODULARITY).** *For fixed  $S', \mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\lambda}$  the function  $\alpha(S|\mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\lambda}, S')$  is submodular is  $S$ —that is, for any  $S, T \subseteq [n] \setminus S'$ ,  $\alpha(S \cap T | S') + \alpha(S \cup T | S') \leq \alpha(S | S') + \alpha(T | S')$ .*

**PROOF.** Fix  $S', \mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\lambda}$ , and we suppress the dependence in  $\alpha$ . To prove submodularity, recall that an equivalent definition is that for any  $S, T$  satisfying  $S \subseteq T$ , and  $i \notin T$ , then  $\alpha(S \cup \{i\}) - \alpha(S) \geq \alpha(T \cup \{i\}) - \alpha(T)$  [36]. Let  $S \subseteq T$  and  $i \notin T$ . A simple computation shows that for any  $V \subseteq [n] \setminus S'$ ,

$$\alpha(V \cup \{i\}) - \alpha(V) = \sum_{j=1}^m \mu'_j p_{ij} \prod_{k \in V} (1 - p_{kj}),$$

where  $\mu'_j$  is  $\mu_j \cdot \prod_{i \in S'} (1 - p_{ij})$  from priority. As each factor in the product is at most 1, this is clearly decreasing in  $V$  as a set function, establishing submodularity.  $\square$

Now, recall that the relevant functions in the construction of the preceding algorithm is the set function  $f = \alpha/\lambda$ . As a consequence of the fact that this function is the ratio of a submodular function with a modular function, we will be able to gain significant closure properties of the tight subsets (as defined earlier), which will end up being critical in establishing both game-theoretic and probabilistic properties of our systems.

**LEMMA 3.5 (CLOSURE).** *For each fixed  $p$  and  $k \geq 1$ , the set of minimizers of  $f(\cdot|U_{k-1})$  in  $\mathcal{P}([n] \setminus U_{k-1})$  is closed under union and nondisjoint intersection—that is, if  $S, S' \subseteq [n] \setminus U_{k-1}$  are minimizers, then so is  $S \cup S'$ , as well as  $S \cap S'$  if nonempty. Moreover, if  $S \cap S'$  is empty, then the queues in  $S$  and  $S'$  must send to disjoint subsets of servers.*

*In particular, the minimizing set with largest cardinality is unique, and is the union of all minimizing sets at step  $k$ . If  $S$  is considered at step  $k$  of the algorithm but  $S$  is not a subset of  $S_k$ , then  $f(S|U_{k-1}) > f(S_k|U_{k-1})$ .*

**PROOF.** The last statement is an immediate consequence of the first, so we focus on the first part. We show this just for  $k = 1$ ; the general case for  $k > 1$  follows from the recurrence in Equation (4).

By Lemma 3.4,  $\alpha$  is submodular, and it is immediate that  $\lambda(\cdot)$  is a *modular* function—that is,  $\lambda(S \cup S') + \lambda(S \cap S') = \lambda(S) + \lambda(S')$ . We claim that if a function  $f$  is a ratio of a nonnegative submodular function and a nonnegative supermodular function, then the set of minimizers of  $f$  is closed under union and nondisjoint intersection. To see this, suppose that  $S, S'$  are minimizers. Then we see the following inequalities, writing  $f$  always in the form of the quotient  $\alpha/\lambda$ , and using Fact 3.1 and Fact 2.1,

$$\max\{f(S), f(S')\} \geq f(S) \oplus f(S') \geq f(S \cap S') \oplus f(S \cup S') \geq \min\{f(S \cap S'), f(S \cup S')\}$$

(where we omit  $S \cap S'$  if it is empty), as the inequalities in the numerator and denominator go the correct way by Lemma 3.4, and then using Fact 2.1. But as  $S, S'$  were minimizers, these must be equalities, which occurs if and only if  $S \cap S'$  (if nonempty) and  $S \cup S'$  are both minimizers. As  $f = \alpha/\lambda$  here, this applies for our functions.

If  $S, S'$  are both minimizers and are disjoint, then from Fact 3.1, it follows that  $f(S \cup S') = f(S) \oplus f(S'|S)$ . As  $S, S', S \cup S'$  are evidently minimal, this equation can only hold if  $f(S'|S) = f(S)$ , which occurs if and only if  $S$  and  $S'$  disjointly mix among servers.  $\square$

From Lemma 3.5, it will nearly immediately follow that the outputted rates are strictly monotonic decreasing in the groups: as mentioned,  $[n] = S_1 \sqcup S_2 \sqcup \dots$  is meant to give a partition into groups that age together, where  $S_1$  is the fastest aging group,  $S_2$  the next fastest, and so on. As such, the disjoint subsets iteratively output by the algorithm satisfy the intuition that motivates the construction.

**LEMMA 3.6 (MONOTONICITY).** *Let  $S_1, S_2, \dots$  be the outputs of the algorithm in order. Then  $g_k > g_{k+1}$  for each  $k \geq 1$ .*

**PROOF.** Consider  $S_k \cup S_{k+1}$ : this set is considered at stage  $k$  of the algorithm and evidently was not selected. From the previous lemma, we must have

$$f(S_k|U_{k-1}) < f(S_k \cup S_{k+1}|U_{k-1}) < f(S_{k+1}|U_{k-1} \cup S_k) = f(S_{k+1}|U_k).$$

The first inequality follows from the selection criteria of the algorithm (and maximality of  $S_k$ ), whereas the second follows from writing (via Fact 3.1)

$$f(S_k \cup S_{k+1}|U_{k-1}) = f(S_k|U_{k-1}) \oplus f(S_{k+1}|U_{k-1} \cup S_k).$$

As we have already proven the first inequality, the second must follow from Fact 2.1. This yields the claim.  $\square$

With these basic properties, we can obtain an important structural result that will prove fruitful in establishing the existence of equilibria. We defer the somewhat technical proof to Appendix D.1.

**PROPOSITION 3.7 (CONTINUITY).** *The function  $r : (\Delta^{m-1})^n \rightarrow [0, 1]^n$  given by  $r(\mathbf{p}) = (r_1(\mathbf{p}), \dots, r_n(\mathbf{p}))$  is continuous.*

With these structural results, we can turn to showing our first game-theoretic property of this game, for now assuming that the costs are given by  $r$ , the output of the algorithm of Section 3.1: namely, that equilibria exist. Although the cost functions are not quite convex, by restricting each component to a line that varies only a single queue's strategy, one can deduce enough structure that allows for an application of Kakutani's theorem. We record this result here while deferring the proof to Appendix D.1.

**THEOREM 3.8 (EXISTENCE OF EQUILIBRIA).** *There exists an equilibrium of the game with costs given by  $r : (\Delta^{m-1})^n \rightarrow [0, 1]^n$ .*

### 3.3 Price of Stability and Independence

In Section 4, we will establish one of our main results, a tight bound of  $\frac{e}{e-1}$  on the price of anarchy in the patient queuing game. Recall that this bound asserts that in any patient queuing system that is centrally feasible, if the arrival rates are decreased by a factor of  $\frac{e}{e-1}$ , then *every* Nash equilibria of the resulting game will be stable. In this section, we consider the intermediate and complementary notion of the *price of stability*, which controls the efficiency loss of the *best* Nash equilibrium. Formally, we have the following definition.

**Definition 3.9.** Let  $\lambda$  and  $\mu$  satisfy the conditions of Theorem 2.2 with weak inequalities so that  $\mu$  majorizes  $\lambda$ . For  $\alpha \geq 1$ , let  $\mathcal{G}(\alpha) = ([n], \mu, \alpha^{-1}\lambda)$ , and let  $\mathcal{N}(\alpha)$  denote the set of Nash equilibria of  $\mathcal{G}(\alpha)$ . The *price of stability* of  $\mathcal{G} = \mathcal{G}(1)$  is defined as the supremum of  $\alpha$  values such that in *every* Nash equilibrium  $\mathbf{p}^* \in \mathcal{N}(\alpha)$ , there exists  $i \in [n]$  such that  $c_i(\mathbf{p}^*) > 0$  in  $\mathcal{G}(\alpha)$ . The price of stability of the patient queuing game is the supremum over all instances of  $\mathcal{G}$ .

By definition, the price of stability lies between 1 and the price of anarchy.

Another measure is the *price of independence*. Recall that the coordinated queuing strategy in the proof of Theorem 2.2 is highly centralized, in the sense that queues never collide due to the central scheduling.<sup>13</sup> How well can all agents do when they are not necessarily selfish but are restricted to stationary product strategies? In other words, how well can all agents do when they attempt to make the system as stable as possible for all agents but are restricted to product strategies? Such a measure decouples the contribution in the price of anarchy of selfishness and independence (i.e., the inability to coordinate due to product distributions). Formally, we have the following definition.

**Definition 3.10.** Let  $\lambda$  and  $\mu$  satisfy the conditions of Theorem 2.2 with weak inequalities so that  $\mu$  strictly majorizes  $\lambda$ . For  $\alpha \geq 1$ , let  $\mathcal{G}(\alpha) = ([n], \mu, \alpha^{-1}\lambda)$ . The *price of independence* of  $\mathcal{G} = \mathcal{G}(1)$  is defined as the supremum of  $\alpha$  values such that for all  $\mathbf{p}^* \in (\Delta^{m-1})^n$ , there exists  $i \in [n]$  such that  $c_i(\mathbf{p}^*) > 0$  in  $\mathcal{G}(\alpha)$ . The price of independence of the patient queuing game is the supremum over all instances of  $\mathcal{G}$ .

<sup>13</sup>We remark that because all queues are stable in the coordinated solution and thus no deviation is profitable, this strategy can be interpreted as a correlated equilibrium of the patient queuing game where at each time, each agent is told to play a (random, coordinated) Dirac strategy (and possibly abstaining from sending in that round). Therefore, if we allow for public randomness to induce coordination, the price of stability with respect to the larger class of correlated equilibria is simply 1.

In our setting, it is actually not very difficult to see that these two quantities are precisely the same. One direction is trivial, whereas if  $\alpha \geq 1$  is such that there exists a  $\mathbf{p}^*$  with  $c_i(\mathbf{p}^*) = 0$  in  $\mathcal{G}(\alpha)$  for all  $i \in [n]$ ,  $\mathbf{p}^*$  is clearly a Nash equilibrium in  $\mathcal{G}(\alpha)$  as well as it trivially holds that no agent can possibly improve their rate. In fact, we observe that there is a more general correspondence, as shown in the following proposition.

**PROPOSITION 3.11.** *For any patient queuing game  $\mathcal{G}$ , let  $\text{OPTIND}(\mathcal{G}) = \min_{\mathbf{p}} \max_{i \in [n]} c_i(\mathbf{p})$ . Then there exists a (possibly different) patient queuing game  $\mathcal{G}'$  and a Nash equilibrium  $\mathbf{p}'$  of  $\mathcal{G}'$  such that  $c_{i, \mathcal{G}'}(\mathbf{p}') = \text{OPTIND}(\mathcal{G})$  for all  $i$ . Moreover,  $\mathcal{G}'$  is a restriction of  $\mathcal{G}$  to a subset of queues and servers.*

**PROOF.** Fix  $\mathcal{G} = ([n], \boldsymbol{\mu}, \boldsymbol{\lambda})$ , and let  $\mathbf{p}$  denote an optimizer in the definition of  $\text{OPTIND}(\mathcal{G})$  with a *minimal* number of tight subsets achieving the *maximum* rate. If  $\text{OPTIND}(\mathcal{G}) = 0$ , the claim is trivial with  $\mathcal{G}' = \mathcal{G}$  and  $\mathbf{p}' = \mathbf{p}$  by the reasoning prior to the proposition statement, so assume that it is strictly positive. Let  $S_1$  denote the maximal tight subset, as seen in the previous section, with size  $n'$ . Let  $\boldsymbol{\lambda}'$  be the restriction of  $\boldsymbol{\lambda}$  to  $S_1$ ,  $\boldsymbol{\mu}' = \boldsymbol{\mu}$ , and  $\mathbf{p}'$  the restriction of  $\mathbf{p}$  to  $S_1$ . Note that  $\boldsymbol{\mu}'$  clearly still strictly majorizes  $\boldsymbol{\lambda}'$ .

We claim that  $\mathcal{G}' = (S_1, \boldsymbol{\mu}', \boldsymbol{\lambda}')$  is such that  $\mathbf{p}'$  is a Nash equilibrium of  $\mathcal{G}'$ . Note that this implies the conclusion of the proposition. Fix an  $i \in S_1$  and suppose by way of contradiction that there exists a deviation to  $p_i''$  such that the rate of queue  $i$  strictly improves. If  $i$  is part of every tight subset of the original game, this implies that the strategy vector  $\mathbf{p}^* = ((1-t)p_i + tp_i'', \mathbf{p}_{-i})$  for  $t > 0$  sufficiently close to zero in the original game attains a strictly smaller value than  $\mathbf{p}$ . This holds because for each subset  $S \subseteq [n]$  containing  $i$ , the function  $f(S|\cdot)$  is a linear function of  $i$ 's randomization, fixing  $\mathbf{p}_{-i}$ . Because all tight subsets contain  $i$  by assumption, and all of the other finitely many subsets attain strictly smaller rates, we may take  $t > 0$  sufficiently small to obtain a smaller value for  $\text{OPTIND}(\mathcal{G})$ , violating the definition of  $\mathbf{p}$ .

Therefore, suppose instead that there exists a tight subset  $S$  of  $S_1$  in the original game such that  $i \notin S$ . Again, suppose by way of contradiction that there exists a deviation to  $p_i''$  such that the rate of queue  $i$  strictly improves in  $\mathcal{G}'$ . Considering the same deviation as in the previous case, note that there is no contradiction to the optimality of the value attained by  $\mathbf{p}$  because the rate of the subset  $S$  is unchanged. But for sufficiently small  $t > 0$ , the number of tight subsets attaining the maximum rate strictly decreases as the rate of every subset containing  $i$  strictly decreases, contradicting the minimality of the number of tight subsets attaining the maximum rate in  $\mathbf{p}^*$ . As  $i \in S_1$  was arbitrary, we conclude that  $\mathbf{p}'$  is a Nash equilibrium of  $\mathcal{G}'$  attaining the same value of the optimizer in  $\mathcal{G}$ .  $\square$

We now give a simple argument showing that the price of independence, and therefore also the price of stability, is at most  $\frac{e}{e-1}$ . This will later also be a consequence of our bound on the price of anarchy, but the argument is substantially simpler.

**PROPOSITION 3.12.** *The price of independence/stability of the patient queuing game is at most  $\frac{e}{e-1}$ .*

**PROOF.** Let  $\mathcal{G} = ([n], \boldsymbol{\mu}, \boldsymbol{\lambda})$  such that  $\boldsymbol{\mu} \geq \boldsymbol{\lambda}$ . We first assume that  $m = n$ . Recall that this implies there exists a doubly stochastic matrix  $X \in \mathbb{R}^{n \times n}$  such that  $X\boldsymbol{\mu} \geq \boldsymbol{\lambda}$ , where the inequality is taken elementwise. We set  $x_i$  to be the  $i$ th column of  $X$  and consider the strategy  $\mathbf{x} = (x_1, \dots, x_n)$ . We write  $x_{ij}$  for the  $j$ th element of  $x_i$ , which is the  $(j, i)$  entry of the matrix  $X$ .

Consider any subset  $S \subseteq [n]$ . Our goal is to show that

$$f(S|\mathbf{x}) = \frac{\sum_{j=1}^n \mu_j (1 - \prod_{i \in S} (1 - x_{ij}))}{\sum_{i \in S} \lambda_i} \geq \frac{e-1}{e}.$$

If this holds, then scaling down  $\boldsymbol{\lambda}$  by  $\frac{e}{e-1}$  and using the same strategy tuple yields the desired result.

We have

$$\prod_{i \in S} (1 - x_{ij}) \leq \exp\left(-\sum_{i \in S} x_{ij}\right) \leq 1 - \frac{e}{e-1} \sum_{i \in S} x_{ij}.$$

Here, we use the inequalities  $1 - x \leq \exp(-x)$  and  $\exp(-x) \leq 1 - \frac{e-1}{e}x$  on  $[0, 1]$ , which can be easily obtained by convexity and observing that the left and right sides are equal at the endpoints. We also use the doubly stochasticity of  $X$  so that the sum is indeed at most 1. Plugging this in, we find that

$$f(S|\mathbf{x}) \geq \left(\frac{e-1}{e}\right) \frac{\sum_{j=1}^n \mu_j \sum_{i \in S} x_{ij}}{\sum_{i \in S} \lambda_i} = \left(\frac{e-1}{e}\right) \frac{\sum_{i \in S} \sum_{j=1}^m \mu_j x_{ij}}{\sum_{i \in S} \lambda_i} \geq \left(\frac{e-1}{e}\right),$$

where we use the fact that  $\sum_{j=1}^n \mu_j x_{ij} \geq \lambda_i$  (this is the  $i$ th inequality in  $X\boldsymbol{\mu} \geq \boldsymbol{\lambda}$ ).

To extend this argument to arbitrary  $m, n$ , simply observe that if  $m > n$ , one may simply truncate to the  $n$  largest servers. If instead  $m < n$ , one can instead argue as follows: consider  $\boldsymbol{\mu}'$  that extends  $\boldsymbol{\mu}$  by padding with zeros. The preceding argument shows that there exists the desired strategy vector  $\mathbf{x}^*$  attaining the bound where queues may send to dummy servers. In the original game, any strategy profile  $\mathbf{x}$  such that  $x_{ij} \geq x_{ij}^*$  for all  $j \leq m$  and  $x_{ij} = 0$  for  $j > m$  attains at least as good a value as the dummy servers confer no value to the rate functions. This concludes the proof.  $\square$

It is not difficult to see that the price of independence is strictly greater than 1; almost any nontrivial example will certify this. For instance, the following example shows that it is at least  $9/8$ .

*Example 3.13.* Consider  $\boldsymbol{\mu} = (1, 1/2)$  and  $\boldsymbol{\lambda} = (3/4, 3/4)$ . Let  $x, y \in [0, 1]$  denote the probabilities that the first and second queues send to the top server. Then it is easy to see that the optimal rate that can be attained is

$$\max_{x, y \in [0, 1]} 1 - \left( \min \left\{ \frac{1/2 + x/2}{3/4}, \frac{1/2 + y/2}{3/4}, \frac{1/2 + x + y - 3xy/2}{3/2} \right\} \right).$$

One can show that the minimum is attained at  $(x, y) = (1/3, 1)$  so that the rate of both queues is  $1 - (8/9) = 1/9$ . This implies that the system remains unstable unless the arrival rates are scaled down by  $9/8$ , so the price of independence is at least  $9/8$ .

We suspect that the price of independence is significantly less than  $\frac{e}{e-1}$ , but we leave the determination of its value as an interesting analytical question for future work.

**QUESTION 3.** Determine the solution of

$$\max_{\boldsymbol{\mu} \geq \boldsymbol{\lambda}} \min_{\mathbf{x} \in (\Delta^{m-1})^n} \max_{S \subseteq [n]} \left[ 1 - \frac{\sum_{j=1}^m \mu_j \left(1 - \prod_{i \in S} (1 - x_{ij})\right)}{\sum_{i \in S} \lambda_i} \right].$$

#### 4 PRICE OF ANARCHY OF PATIENT QUEUING

In this section, we turn to the game-theoretic problem of understanding what condition ensures the stability at any equilibrium profile assuming Theorem 3.3—we return to proving that this is valid in the next section. By considering the quality of deviations by a queue at a Nash equilibrium to a single other server, it is possible to show that the price of anarchy is always at most 2. With more careful, continuous deviations, we in fact show that this factor is loose with patience, and the correct bound is  $\frac{e}{e-1} \approx 1.58$ .

The following simple example shows that this is the best possible constant factor: fix  $\epsilon > 0$  small and suppose that there are  $n$  queues and  $n$  servers, with  $\boldsymbol{\lambda} = (1 - 1/e + \epsilon, \dots, 1 - 1/e + \epsilon)$  and  $\boldsymbol{\mu} = (1, \dots, 1)$ , and  $\mathbf{p}$  has every queue uniformly mixing among the servers. It is easy to see



by symmetry that this system is Nash with  $S_1 = [n]$ , for if a queue deviates from this uniform distribution, this does not change the worst ratio in the algorithm. Moreover, for any fixed  $\epsilon > 0$ , as  $n \rightarrow \infty$ , this system becomes unstable. One can check that

$$f(S_1|\mathbf{p}) = f([n]|\mathbf{p}) = \frac{\sum_{j=1}^n (1 - \prod_{i=1}^n (1 - 1/n))}{n(1 - 1/e + \epsilon)} \rightarrow \frac{1 - 1/e}{1 - 1/e + \epsilon} < 1$$

so that  $r(S_1) = \max_i c_i(\mathbf{p}) > 0$ . Our main result asserts that this is the worst case, where every queue is maximally colliding subject to being Nash. Concretely, we prove the following instance-dependent bound from which the claimed factor immediately follows.

**THEOREM 4.1 (MAIN).** *Let  $\mathbf{p}$  be any Nash equilibrium of  $\mathcal{G}$ , and let  $S_1$  be as defined before. Then<sup>14</sup>*

$$f(S_1|\mathbf{p}) \geq \min \left\{ 1, \min_{k \leq n} \max_{\mathbf{x} \in \mathbb{R}_{\geq 0}^m, \sum_{j=1}^m x_j = k} \frac{\sum_{j=1}^m \mu_j (1 - (1 - x_j/k)^k)}{\sum_{i=1}^k \lambda_i} \right\}.$$

**COROLLARY 4.2.** *Let  $\mathbf{p}$  be a Nash equilibrium of  $\mathcal{G}$ , and suppose that for each  $1 \leq k \leq n$ ,*

$$\sum_{j=1}^{\min\{k, m\}} \mu_j \geq \left( \frac{e}{e-1} \right) \sum_{i=1}^k \lambda_i.$$

*Then every queue is weakly stable at  $\mathbf{p}$ . In particular, the price of anarchy of the patient queuing game is exactly  $\frac{e}{e-1}$ .*

**PROOF.** In Theorem 4.1, for any  $k \leq n$ , we may set  $x_j = 1$  for  $1 \leq j \leq k$ . Note that  $(1 - x_j/k)^k < e^{-1}$ , so if  $\boldsymbol{\mu}$  majorizes  $\boldsymbol{\lambda}$  by a factor of at least  $\frac{e}{e-1}$ , Theorem 4.1 implies that  $f(S_1|\mathbf{p}) \geq 1$ . From Lemma 3.6 and Theorem 3.3, we conclude that all queues are weakly stable.  $\square$

We now prepare for the proof of Theorem 4.1. The idea will be to continuously deform the Nash profile toward a highly symmetrized strategy vector while only weakly decreasing  $f(S_1)$ . At the end of this process, we obtain a lower bound on this value at Nash. To do this deformation and ensure monotonicity of the growth rate, we must at some point use the Nash property. The difficulty lies in the form of the  $f$  functions; recall that as  $S_1$  is the set of all queues growing at the fastest rate as the union of all tight subsets, it can have many proper tight subsets, and each queue  $i \in S_1$  thus has to locally optimize all of the functions  $f(S|\mathbf{p})$  with  $S \ni i$  simultaneously at Nash (see Figure 2 for an interesting example). In particular, if queue  $i \in S_1$  at Nash, one possible deviation may weakly decrease  $f(S)$  for some tight subset  $S \ni i$ , whereas another deviation may weakly decrease  $f(S')$  for some different tight subset  $S' \ni i$ . In other words, each queue may be constrained by multiple different objective functions at Nash, making it difficult to generically argue about why any given deviation decreases performance. We overcome this barrier via Proposition 4.3 by connecting the incentives for each queue in  $S_1$  with the structure guaranteed by Lemma 3.5.

**PROPOSITION 4.3.** *Let  $\mathbf{p}$  be any arbitrary strategy vector by the queues, and without loss of generality, let  $[k]$  be the maximal tight subset after relabeling. Then, for some  $s > 0$ , there exists a level partition of  $[k]$  into  $s$  levels with the following property: if a queue  $i \in [k]$  belongs to a level- $\ell$  subset, then for any deviation by  $i$  that shifts probability mass from one server to another and does not increase  $f(S)$  for some tight subset  $S \ni i$ , there exists a tight subset  $S'$  containing all queues at all levels  $j \leq \ell$  such that  $f(S')$  must not increase.*

<sup>14</sup>For any fixed  $k$ , it is not difficult to determine the optimal value of  $\mathbf{x}$  to give the tightest lower bound. It suffices to maximize the numerator, which is concave. By standard Karush-Kuhn-Tucker conditions at optimality, for all  $j, j' \in [m]$  such that  $x_j > 0$ , we must have  $\mu_j(1 - x_j/k)^{k-1} = \mu_{j'}(1 - x_{j'}/k)^{k-1}$ , and  $x_\ell = 0$  for all lower indices.

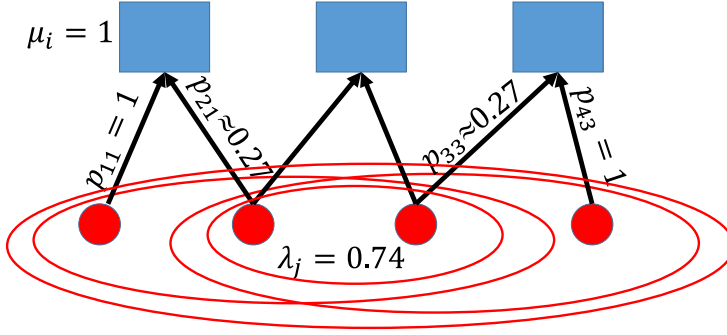


Fig. 2. Four queues compete for three servers, with maximally tight sets marked. The two outer queues exclusively send to the outer servers, and the inner queues send to two servers each, as indicated by the figure. One can verify that this profile is Nash and the subsets marked are tight. Consider whether one of the inner queues would want to deviate from the current profile. If she shifts probability to the inner server, the smallest tight set will rise in the aging rate. However, if she instead attempts to deviate to one of the outer servers, the rate of all four queues will rise.

PROOF. We construct the desired decomposition of  $[k]$  by instead recursively considering *minimal* tight subsets. In particular, let the level-1 subsets be the set of all minimal tight subsets (notice that by Lemma 3.5 and minimality, each subset of queues must be disjoint and disjointly mixing). Note that to clear packets at the rate of these subsets, they all must have priority over any other queues. Then given the level- $j$  subsets for all  $j \leq \ell$  for some  $\ell \geq 1$ , we recursively define the level- $(\ell + 1)$  subsets as the set of minimal tight subsets of the remaining queues conditioned on all subsets at lower levels having priority. The same argument as used by Lemma 3.5 implies that these subsets at each level must all be disjoint and disjointly mixing.<sup>15</sup> It is easy to verify by iteratively using Fact 3.1 that for any level  $\ell$ , the union of all subsets at levels  $j < \ell$  with any subset of the level- $\ell$  subsets must be tight. As  $[k]$  is tight and any tight subset must be contained in  $[k]$  by the maximality guaranteed in Lemma 3.5, it is immediate that this decomposition exhausts  $[k]$ .

Let  $i \in [k]$  belong to some level- $\ell$  subset, for some  $\ell \geq 1$ . Suppose that  $i$  has some deviation from server  $j$  to another server  $j'$  that causes  $f(S)$  to not increase for some tight subset  $S \subseteq [k]$ . As  $f(S)$  is a linear function in the randomizations by queue  $i$  holding each other queue fixed, this must hold for any sufficiently small deviation as well. By expanding the partial derivatives  $\frac{\partial f(S)}{\partial p_{ij}}$  and  $\frac{\partial f(S)}{\partial p_{ij'}}$  and clearing denominators, this holds if and only if

$$\mu_j \prod_{r \in S \setminus \{i\}} (1 - p_{rj}) \geq \mu_{j'} \prod_{r \in S \setminus \{i\}} (1 - p_{rj'}). \quad (5)$$

Define  $V_{\leq \ell}^i$  to be the set of all queues at level at most  $\ell$  that do not belong to the level- $\ell$  subset  $i$  belongs to. Then  $V_{\leq \ell}^i$  is tight, and by construction  $S \setminus V_{\leq \ell}^i$  is nonempty as  $i$  belongs to this set. As tight subsets are preserved under unions and intersections, we have that  $S \cup V_{\leq \ell}^i$  and  $S \cap V_{\leq \ell}^i$  are tight. On the one hand, we have by Fact 3.1 that

$$f(S) = f(S \cap V_{\leq \ell}^i) \oplus f(S \setminus V_{\leq \ell}^i | S \cap V_{\leq \ell}^i).$$

<sup>15</sup>In the example given before of a system with multiple tight subsets, the level-1 subset is the subset of two queues that split between the inner and outer servers. The level-2 subsets are the two singleton sets of the outer queues sending each to their own outer server. Notice that these two subsets indeed disjointly mix.

On the other hand, we have

$$f(S \cup V_{\leq \ell}^i) = f(V_{\leq \ell}^i) \oplus f(S \setminus V_{\leq \ell}^i | V_{\leq \ell}^i).$$

Comparing these two expressions, the left-hand sides are minimal, as well as the first terms on the right, which by Fact 2.1 implies that the latter terms are equal. But we know from definition that

$$f(S \setminus V_{\leq \ell}^i | S \cap V_{\leq \ell}^i) \geq f(S \setminus V_{\leq \ell}^i | V_{\leq \ell}^i),$$

with equality if and only if  $V_{\leq \ell}^i \setminus S$  disjointly mixes from  $S \setminus V_{\leq \ell}^i$  in  $\mathbf{p}$ . As  $i$  belongs to this set, we can combine this with Equation (5) to see

$$\mu_j \prod_{r \in (V_{\leq \ell}^i \cup S) \setminus \{i\}} (1 - p_{rj}) = \mu_j \prod_{r \in S \setminus \{i\}} (1 - p_{rj}) \geq \mu_{j'} \prod_{r \in S \setminus \{i\}} (1 - p_{rj'}) \geq \mu_{j'} \prod_{r \in (V_{\leq \ell}^i \cup S) \setminus \{i\}} (1 - p_{rj'}).$$

This implies that  $f(S \cup V_{\leq \ell}^i)$  must also not increase from this deviation. The last thing to check is that the level- $\ell$  subset  $i$  belongs to is contained in  $S$  so that  $S \cup V_{\leq \ell}^i$  contains all queues up to level  $\ell$ . But this follows because  $S \setminus V_{\leq \ell}^i$  consists of queues at a level at least  $\ell$  and intersects this level- $\ell$  subset nontrivially at  $i$ , and therefore must contain it by minimality in our construction. Therefore, for this choice of deviation by queue  $i$ , the subset  $S \cup V_{\leq \ell}^i$  certifies the desired claim.  $\square$

With this result, we may finally return to the proof of Theorem 4.1.

**PROOF OF THEOREM 4.1.** Let  $\mathbf{p}$  be any Nash equilibrium, and suppose that  $S_1$  is the maximal tight subset. If  $f(S_1) \geq 1$ , then we are done, so suppose that  $f(S_1) < 1$ . The Nash assumption then implies that any deviation by a queue in  $S_1$  cannot decrease the value of each tight subset it is part of; note that we need the  $f(S_1) < 1$  assumption, as incentives are about the rates, and when  $f(S_1) \geq 1$ , the rate may remain 0 even if  $f$  decreases.

For convenience, reindex and relabel so that  $|S_1| = k$  and  $S_1 = [k]$ . Fix any  $\mathbf{x} \in \mathbb{R}_{\geq 0}^m$  such that  $\sum_{i=1}^m x_i = k$ . It suffices to show that

$$f([k]|\mathbf{p}) \geq \frac{\sum_{j=1}^m \mu_j (1 - (1 - x_j/k)^k)}{\sum_{i=1}^k \lambda_i}.$$

From now on, we omit the dependence on  $\mathbf{p}$  in  $f$  unless explicitly needed.

Consider the level partition of  $[k]$  guaranteed by Proposition 4.3, and suppose that there are  $s$  levels. We continuously deform the Nash solution while monotonically decreasing  $f([k])$  so that at the end of this process, we have a lower bound on the value of Nash. Given any strategy profile  $\mathbf{p}$ , we say that a server  $j$  is *oversaturated* if  $\sum_{i \in [k]} p_{ij} > x_j$  and *undersaturated* if  $\sum_{i \in [k]} p_{ij} < x_j$ . We will continuously move probability mass from the queues from oversaturated to undersaturated servers. If no server is oversaturated, we will be done; notice that if a server is oversaturated, an easy averaging argument implies that some server must be undersaturated.

Suppose that there exists some oversaturated server. Let  $i \in [k]$  be a queue at level  $s$ , the top level. If  $i$  nontrivially sends to an oversaturated server  $j$  so that  $p_{ij} > 0$ , we continuously decrease  $p_{ij}$  and increase  $p_{ij'}$  for some undersaturated server  $j'$ , until either  $j$  stops being oversaturated,  $j'$  stops being undersaturated, or  $p_{ij}$  hits zero. We claim that this deformation cannot increase  $f([k])$ . To see this, observe that because  $\mathbf{p}$  is Nash, we know that any deviation by  $i$  from one server it is nontrivially mixing at to another cannot increase  $f(S)$  for all tight subsets  $S$  containing  $i$ , hence there must be some  $S \ni i$  such that  $f(S)$  does not increase. But then Proposition 4.3 implies that some subset  $S'$  containing all queues up to level  $s$  must have  $f(S')$  not increase either. As  $[k]$  is the only such subset, this deformation could not have actually increased  $f([k])$ .

Moreover, we claim that we can do this for all level- $s$  queues one by one without increasing  $f([k])$ . Although the intermediate profiles are not Nash, because we only move probability mass from oversaturated to undersaturated servers, each oversaturated queue only has at most the same probability mass as it did at Nash while each undersaturated queue only has additional probability mass compared to what it had at Nash. As we have shown any such deviation by a level- $s$  server from an oversaturated queue to an undersaturated queue at Nash cannot increase  $f([k])$ , and now deviations are only worse at this intermediate stage while deforming the level- $k$  queue strategies, each such deformation still cannot increase  $f([k])$ . Therefore, we can continuously shift all probability mass from level- $s$  queues at oversaturated servers to undersaturated servers while never increasing  $f([k])$ .

Suppose that we have now done this for all levels at least  $\ell + 1$  for some  $\ell < s$  while not increasing  $f([k])$ , and we want to continue this process at level  $\ell$ . Let  $\mathbf{p}'$  be this intermediate strategy vector, where we note that for any queue  $i$  below level  $\ell + 1$ ,  $\mathbf{p}'_i = \mathbf{p}_i$ . Again, if no server is oversaturated, we are done. Otherwise, suppose that some queue  $i$  at level  $\ell$  still sends to an oversaturated server  $j$ , and we again try to decrease  $p'_{ij}$  and increase  $p'_{ij'}$  for some undersaturated server  $j'$  as before until the same stopping criterion. We must show that this too cannot increase  $f([k])$ .

Suppose otherwise that it did indeed increase  $f([k])$  with respect to  $\mathbf{p}'$ . For a contradiction, it suffices to show that this implies that this same deviation, with respect to the original Nash solution  $\mathbf{p}$ , must have increased  $f(S)$  for every subset  $S$  containing all queues up to level  $\ell$ . This is sufficient to obtain a contradiction, as then Proposition 4.3 implies that *every* subset containing  $i$  improves at Nash with respect to this deviation, which violates the Nash property.

To prove this claim, let  $S \subseteq [k]$  be any arbitrary tight subset at Nash containing all queues up to level  $\ell$ . Because we assume that this deviation improves  $f([k])$  with respect to  $\mathbf{p}'$ , it follows from taking partial derivatives that

$$\mu_j \prod_{r \in [k] \setminus \{i\}} (1 - p'_{rj}) < \mu_{j'} \prod_{r \in [k] \setminus \{i\}} (1 - p'_{rj'}).$$

However, note that at  $\mathbf{p}'$ , as  $j$  is still oversaturated,  $p'_{rj} = 0$  for all queues  $r$  that are at strictly higher levels. As all queues at level  $\ell$  and below have  $\mathbf{p}'_i = \mathbf{p}_i$  and  $S$  contains all such queues, this inequality implies that

$$\mu_j \prod_{r \in S \setminus \{i\}} (1 - p_{rj}) < \mu_{j'} \prod_{r \in [k] \setminus \{i\}} (1 - p'_{rj'}).$$

Moreover, as  $j'$  is undersaturated, we must have  $p'_{rj'} \geq p_{rj'}$  for all  $r \in [k]$  from the construction of this process, and removing terms in the product only increases the right side. Therefore, we deduce that

$$\mu_j \prod_{r \in S \setminus \{i\}} (1 - p_{rj}) < \mu_{j'} \prod_{r \in S \setminus \{i\}} (1 - p_{rj'}).$$

This implies that this deviation also increases  $f(S)$  at Nash. As  $S$  was an arbitrary tight subset containing all queues up to level  $\ell$ , the claim is proved. By the reduction described earlier, this is a contradiction, and therefore  $f([k])$  must further decrease with respect to  $\mathbf{p}'$ . The argument extends analogously at all intermediate points of this process at level  $\ell$  by the same reasoning as before.

Therefore, by induction, it follows that we may continuously deform probability mass from oversaturated servers to undersaturated servers while only decreasing  $f([k])$ . At the end of this process, there cannot be any oversaturated servers, as otherwise the process could have continued. In particular, if  $\mathbf{p}''$  is the final probability vector at the end of this process, we have shown that

$\sum_{i \in [k]} p''_{ij} = x_j$  for all servers  $j$  and that  $f([k]|\mathbf{p}'') \leq f([k]|\mathbf{p})$ . We have

$$\begin{aligned}
 f([k]|\mathbf{p}) &= \frac{\sum_{j=1}^m \mu_j (1 - \prod_{i=1}^k (1 - p_{ij}))}{\sum_{i=1}^k \lambda_i} && \text{(by definition)} \\
 &\geq \frac{\sum_{j=1}^m \mu_j (1 - \prod_{i=1}^k (1 - p''_{ij}))}{\sum_{i=1}^k \lambda_i} && \text{(by construction)} \\
 &\geq \frac{\sum_{j=1}^m \mu_j (1 - (1 - x_j/k)^k)}{\sum_{i=1}^k \lambda_i} && \text{(as symmetric profile maximizes product).}
 \end{aligned}$$

The second inequality holds because given  $\sum_{i \in [k]} p''_{ij}$ , the maximizer of  $\prod_{i \in [k]} (1 - p''_{ij})$  is attained when each term is equal. As  $\mathbf{x}$  was arbitrary, we may take the maximum of the right side over all  $\mathbf{x}$  satisfying the constraints and then the minimum over  $k$ . As  $\mathbf{p}$  was an arbitrary Nash profile, this concludes the proof.  $\square$

## 5 NO-REGRET LEARNING IN STRATEGIC QUEUING SYSTEMS

Our work in the previous sections provides tight bounds for strategic but *stationary* behavior in this queuing model. However, it is unclear how agents might reach such a state, let alone as the result of natural dynamics. In this section, we therefore turn to a *dynamical* setting where agents *adaptively* update their behavior using no-regret learning algorithms. In the no-regret queuing system we analyze, each queue aims to get their packets served as efficiently as possible. At each timestep, they aim to maximize the probability that their packet gets served, measuring value for a time period with the number of packets served. The effective rate of service at a server  $j$  depends on its rate  $\mu_j$ , as well as on the competition for service by the other queues.

In this section, we will model queues as learners, and will make the assumption that for a parameter  $w$ , each queue will satisfy a no-regret learning guarantee on the number of packets served during each window of  $w$  time length. To formalize this assumption, we need a few definitions.

*Definition 5.1.* Fix some time window of length  $w$ , and for convenience reindex time so that the beginning of this window is at  $t = 1$ . Let  $X_t^{i,j}$  be the indicator variable that queue  $i$  would have succeeded in clearing a packet at server  $j$  at time  $t$  (had she sent there), and let  $\sigma_i(t)$  be the identity of the server that queue  $i$  chooses at time  $t$ . Note that if queue  $i$  has no packets at time  $t$ , then  $X_t^{i,j} = 0$  for all  $j$ . Then the *regret* of queue  $i$  on this window, denoted  $\text{Reg}_i(w)$ , is defined as

$$\text{Reg}_i(w) \triangleq \max_{j \in [m]} \sum_{t=1}^w X_t^{i,j} - \sum_{t=1}^w X_t^{i, \sigma_i(t)}. \quad (6)$$

Note that all of these random variables are with respect to the same sample path; the  $X_t^{i,j}$  will depend on all previous randomizations and choices by the queues, as these implicitly yield the priorities of the queues. In other words,  $\text{Reg}_i(w)$  of queue  $i$  on some fixed window of length  $w$  is defined to be the (random) difference between the number of packets queue  $i$  cleared on these  $w$  periods compared to the backward-looking number of packets she would have cleared had she simply always sent to the best single server, where the comparison is in hindsight to the *best* single server on the realized sample path—that is, for each time  $t$  and server  $j$ , we have  $X_t^{i,j} = 1$  if at time  $t$  server  $j$  was successful (regardless of if a packet was sent there then), and the packet that queue  $i$  sent had priority over any packet sent there at that time.<sup>16</sup>

We now make the following assumption on the regret of queuing strategies.

<sup>16</sup>Note that this notion of regret does *not* take into account that had a queue  $i$  cleared a packet at at server  $j$  instead of another queue  $i'$ , at a later time  $i'$  would have had an older packet and therefore higher priority.



**ASSUMPTION 5.1 (QUEUES SATISFY HIGH-PROBABILITY NO-REGRET).** *All queues select servers using a strategy or algorithm satisfying the following no-regret guarantee: given fixed  $\delta \in (0, 1)$  and a fixed window with length  $w$ , the regret  $\text{Reg}_i(w)$  of queue  $i$  on this given window of  $w$  consecutive timesteps satisfies  $\text{Reg}_i(w) \leq \varphi_\delta(w)$  with probability at least  $1 - \delta$  only over their own randomizations during this window, where  $\varphi_\delta(w) = o(w)$  is some explicit function. Here,  $o(\cdot)$  hides constant factors depending on  $\delta$  and  $m$ , but not  $w$ .*

Moreover, we require that the choices of the queue depend only on their past bandit feedback and their past history of ages, but not on their history of queue sizes.

For instance, this assumption holds with EXP3.P.1 with the form of the regret scaling like  $\sqrt{wm \ln(mw/\delta)} = o(w)$  [3]. Note that this high-probability guarantee is possible in our setting even in the priority model where the random variables of success at each server from the perspective of each queue at each timestep depend on all previous actions (via timestamps and priorities), as well as the actions of the other queues in the current time period (e.g., see the discussion in Section 9 of Auer et al. [3]). This property is standard and necessary in applying learning algorithms to multi-player games. Using EXP3.P.1 ensures that such a guarantee holds simultaneously for each window of this length, and not only a fixed window, so the players would not have to be aware which window of size  $w$  is relevant for our analysis. This is true as EXP3.P.1 mixes in uniform exploration to guarantee that the probabilities remain high enough throughout the algorithm, allowing us to adapt the classical no-regret analysis starting at any timestep for the window of the next  $w$  timesteps.

## 5.1 Stability of No-Regret Queuing Systems

Our second main result shows that if the queuing system has enough slack and all queues satisfy an appropriate high-probability no-regret guarantee, then the queuing system is strongly stable. To this end, we make the following feasibility assumption asserting that a queuing system with servers scaled down by  $1/2$  would remain feasible.

**ASSUMPTION 5.2 (FEASIBILITY).** *There exists  $\eta > 0$  such that for all  $k \in [n]$ ,*

$$\frac{1}{2}(1 - \eta) \sum_{j=1}^{\min\{k, m\}} \mu_j \geq \sum_{i=1}^k \lambda_i.$$

We will use  $\eta$  to denote the maximum such value that this inequality holds. The parameter  $\eta$  controls the quality of learning required for our results. With these definitions in order, we may formally state our main result on stability of no-regret learners.

**THEOREM 5.2.** *Suppose that Assumption 5.2 holds with parameter  $\eta$ . Set the following parameters:  $\delta = \frac{\eta}{4}$ ,  $\epsilon = \frac{\delta \mu_1}{4n}$ , and  $\epsilon_i = \frac{\epsilon}{\lambda_i}$  for  $1 \leq i \leq n$ . Let  $w$  be large enough so that the following holds:*

$$n\varphi_{\frac{\eta}{256n}}(w) + n \leq \frac{w\delta\mu_1}{4}, \quad (7)$$

*and over  $w$  steps of our process the sum of the geometric variables of subsequent packet arrivals, and the sum of the Bernoulli server successes concentrate around their expectation with an error probability of at most  $\eta/256$  with the preceding values of  $\delta, \epsilon_i, \lambda$ , and  $\mu$  (see the required inequalities at Equations (10) and (11)).<sup>17</sup>*

*Then, if each queue satisfies Assumption 5.1 on each consecutive time interval of length  $w$  with probability at least  $1 - \frac{\eta}{256n}$ , then the random process  $\mathbf{T}_t$  under these dynamics is strongly stable.*

<sup>17</sup>Note that this is possible as  $\varphi_\gamma(w) = o(w)$  for any fixed  $\gamma$ , as well as the exponential decay in  $w$  of the concentration bounds in Equations (29) and (30) in the appendix.

The technical tool we use to prove Theorem 5.2 is the following result of Pemantle and Rosenthal [29]:

**THEOREM 5.3 (THEOREM 1 IN THE WORK OF PEMANTLE AND ROSENTHAL [29]).** *Let  $X_1, X_2, \dots$  be a sequence of nonnegative random variables with the following property:*

- (1) *There exist constants  $\alpha, \beta > 0$  such that if  $x_t > \beta$ , then*

$$\mathbb{E}[X_{t+1} - X_t | \mathcal{F}_t] < -\alpha,$$

*where the  $\sigma$ -algebra  $\mathcal{F}_t = \sigma(X_0, \dots, X_t)$  is the history until period  $t$  and  $X_t = x_t$ .*

- (2) *There exists  $p > 2$  and a constant  $\theta > 0$  a constant such that for any history,*

$$\mathbb{E}[|X_{t+1} - X_t|^p | \mathcal{F}_t] \leq \theta.$$

*Then, for any  $0 < r < p - 1$ , there exists an absolute constant  $M = M(\alpha, \beta, \theta, p, r)$  not depending on  $t$  such that  $\mathbb{E}[X_t^r] \leq M$  for all  $t$ .*

To apply this theorem, we must define an appropriate potential function of queue ages that satisfies the negative drift and bounded moments condition. We define for  $\tau \in \mathbb{N}$  the following potential functions that will feature prominently in the proof:

$$\Phi_\tau(\mathbf{T}_t) \triangleq \sum_{i \in [n]: T_t^i \geq \tau} \lambda_i (T_t^i - \tau), \quad (8)$$

$$\Phi(\mathbf{T}_t) \triangleq \sum_{\tau=1}^{\infty} \Phi_\tau(\mathbf{T}_t) = \sum_{\tau=1}^{\infty} \sum_{i \in [n]: T_t^i \geq \tau} \lambda_i (T_t^i - \tau) = \frac{1}{2} \sum_{i=1}^n \lambda_i T_t^i (T_t^i - 1). \quad (9)$$

In other words,  $\Phi_\tau(\cdot)$  denotes the expected number of total packets in the system aged above  $\tau$ , conditioned just on the ages  $\mathbf{T}_t$ .

**PROOF OF THEOREM 5.2.** To apply Theorem 5.3, we define the stochastic process  $Z_0, Z_1, \dots$  by  $Z_\ell = \sqrt{\Phi(\mathbf{T}_{w \cdot \ell})}$ . In other words,  $Z_\ell$  is the “snapshot” of the potential function  $\sqrt{\Phi}$  when evaluated on  $\mathbf{T}_{w \cdot \ell}$  that occurs every  $w$  steps. The filtration is given by  $\mathcal{H}_\ell = \mathcal{F}_{\ell \cdot w}$ , where  $\mathcal{F}_t$  is the corresponding information of the Geometric system at time  $t$  available to the queues.

*Summary of the Main Ideas.* Before we go through the detailed proof, we offer an outline of the main ideas. To establish the negative drift, we will focus on the  $w$ -long interval between two  $Z_\ell$  and  $Z_{\ell+1}$ . In this  $w$ -long window, we use the no-regret condition, as well as concentration bounds on the behavior of queues and servers. The main idea of the proof is to consider all queues that have retained an old packet throughout the period. A server either clears many such old packets or many times during this period no old packet is sent to it. In the second case, we can use the no-regret condition for any queue that still has very old packets, as they would have priority at the server, so these bounds in tandem will imply that many old packets must have cleared. To aid the analysis, we also lower bound the total decrease in ages from clearing packets on this window *before* accounting for the  $w$  extra steps of aging, only accounting for aging at the end; this allows us to consider the clearing process and aging from time passage separately. Finally, when concentration or the no-regret condition fails, we can trivially upper bound what this contributes to the expected drift, and this will be subsumed by the low probability that this occurs in the overall expectation.

*Organizing Randomness.* Let us first set up how we model the actual queuing process on each consecutive window of  $w$  steps between  $Z_\ell$  and  $Z_{\ell+1}$  for the probabilistic analysis. In the spirit of

“organizing randomness,” at step  $\ell$  of this process (step  $\ell \cdot w$  of the actual queuing process), sample up front an independent geometric ensemble  $\{G_k^i\}_{i \in [n], k \in [w]}$  with

$$G_k^i \sim \text{Geom}(\lambda_i), k = 1, \dots, w$$

as well as an independent Bernoulli ensemble  $\{I_k^j\}_{j \in [m], k \in [w]}$  with

$$I_k^j \sim \text{Bern}(\mu_j), k = 1, \dots, w.$$

The interpretation is that the  $I_k^j$  are random indicators if the  $j$ th server is able to clear a packet, *regardless of whether a packet is sent there*, at the  $k$ th step of this block of  $w$  steps. The  $G_k^i$  have the interpretation that when queue  $i$  clears her  $k$ th packet on this window, her age decreases by  $G_k^i$  (without accounting for the aging from passage of time). Crucially, as queue  $i$  clears packets on this window of  $w$  steps, her age decreases *by a sum of a prefix of  $G_1^i, \dots, G_k^i$*  (before accounting for aging as time passes). Observe that this independence arises precisely because of the independence of the geometric ensemble of timestamp differences from the filtration  $\mathcal{G}_{\ell \cdot w}$  of the Geometric system that only conditions on past feedback and the realized past sequence of oldest timestamps.

First, we define the “good” event asserting that the realized sequences of independent random variables on this window are close to their expected values. By our choices of parameters  $\delta$  and  $w$ , we have that with probability at least  $1 - \eta/128$ , all of the following “good” events hold on this window:

$$\left| \sum_{k=1}^q G_k^i - \frac{q}{\lambda_i} \right| < \epsilon_i w \quad \forall i \in [n], q \in [w], \quad (10)$$

$$\sum_{j=1}^q \sum_{k=1}^w I_k^j > w(1 - \delta) \sum_{j=1}^q \mu_j \quad \forall q \in [m], \quad (11)$$

$$\text{Reg}_i(w) \leq \varphi_{\eta/(128n)}(w) \quad \forall i \in [n] \implies n + \sum_{i=1}^n \text{Reg}_i(w) \leq n + n\varphi_{\eta/(128n)}(w) \leq \frac{w\delta\mu_1}{4}. \quad (12)$$

This simply follows from our assumption that  $w$  was chosen large enough so that the first two lines hold simultaneously with probability at least  $1 - \eta/256$  via Corollary B.6 and Lemma B.7 in the appendix, the fact that the no-regret bound held with probability at least  $1 - \eta/(256n)$  for each queue, and taking a union bound. Notice that Equation (10) asserts that every prefix of each of the geometric ensembles is additively not too far from the expectation, relative to  $w$ .

*Threshold Value for  $Z_\ell$ .* We will show that under the threshold assumption that

$$Z_\ell > \frac{w}{\sqrt{2\lambda_n}} \max\left(\frac{16}{\eta} \left(\sum_{i=1}^n \lambda_i\right), 16n^2\right), \quad (13)$$

then the drift condition in the Pemantle-Rosenthal result holds. The significance of this value is that it will imply that there exists at least one rather old queue, and that the expected total number of packets in the system is large relative to the window size. We will later use these facts as demonstrated by the following simple claim, whose proof we defer to Appendix E.

CLAIM 5.1. *Under Equation (13), both of the following statements hold:*

- (1) *There exists some  $i \in [n]$  such that  $\lambda_i T_{\ell \cdot w}^i > 16nw$ .*
- (2)  *$\sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i \geq \frac{16}{\eta} w \sum_{i=1}^n \lambda_i$ .*

*Using No-Regret to Bound the Number of Old Packets Cleared.* Continuing with the proof, we first analyze what happens on the “good” event of Equations (10), (11), and (12). Let  $\tau_i$  be the age of the oldest unprocessed packet of queue  $i$  at the end of this window of  $w$  steps, measured with respect

to the beginning of the window without accounting for the  $w$  steps of aging. If queue  $i$  cleared all her packets that were received before the beginning of this window, then we say that  $\tau_i = 0$ . Let  $J_\tau$  be the set of queues that at the end of the  $w$  steps still have packets that are at least  $\tau$ -old with respect to the beginning of the window. Let  $Y_{j,k}^\tau$  be an indicator variable that some packet that was at least  $\tau$ -old with respect to the beginning of the considered interval was sent to server  $j$  at the  $k$ th step in this window. For any queue in  $J_\tau$ , the priority scheme and the regret bound in Equation (12) implies that for any server  $j \in [m]$

$$\sum_{k=1}^w I_k^j (1 - Y_{j,k}^\tau) - \varphi_{\eta/(128n)}(w). \quad (14)$$

This is because a queue that is always at least  $\tau$ -old throughout the interval would succeed on any server  $j$  that is successful on a timestep  $k$  (indicated by  $I_k^j$ ) when no  $\tau$ -old packets were sent there.

Let  $N_\tau$  be the number of packets that were at least  $\tau$ -old with respect to the beginning of the interval that were cleared in the interval and  $N_\tau^i$  the number of such packets cleared by queue  $i$ . Then we clearly have

$$N_\tau = \sum_{j \in [m]} \sum_{k \in [w]} I_k^j Y_{j,k}^\tau \geq \sum_{j=1}^{\min\{m, |J_\tau|\}} \sum_{k \in [w]} I_k^j Y_{j,k}^\tau. \quad (15)$$

As every packet processed by queues in  $J_\tau$  contributes to  $N_\tau$ , by instantiating Equation (14) for each queue in  $J_\tau$  with each of the top  $\min\{m, |J_\tau|\}$  servers and summing, we also obtain

$$\min\{m, |J_\tau|\} \cdot N_\tau \geq \min\{m, |J_\tau|\} \sum_{i \in J_\tau} N_\tau^i \geq |J_\tau| \sum_{j=1}^{\min\{m, |J_\tau|\}} \left( \sum_{k=1}^w I_k^j (1 - Y_{j,k}^\tau) - \varphi_{\eta/(128n)}(w) \right). \quad (16)$$

Multiplying Equation (15) by  $|J_\tau|$  and summing with the previous equation, we obtain

$$N_\tau \geq \left( \frac{|J_\tau|}{|J_\tau| + \min\{m, |J_\tau|\}} \right) \sum_{j=1}^{\min\{m, |J_\tau|\}} \left( \sum_{k=1}^w I_k^j - \varphi_{\eta/(128n)}(w) \right) \quad (17)$$

$$\geq \frac{1}{2} \sum_{j=1}^{\min\{m, |J_\tau|\}} \sum_{k \in [w]} I_k^j - n \varphi_{\eta/(128n)}(w) \quad (18)$$

$$\geq \frac{1}{2} w(1 - \delta) \sum_{j=1}^{\min\{m, |J_\tau|\}} \mu_j - n \varphi_{\eta/(128n)}(w), \quad (19)$$

where the last inequality uses Equation (11).

*Bounding the Expected Drift in  $\Phi$  Assuming the “Good Event.”* Observe that from the construction of  $\Phi_\tau$ , when queue  $i$  manages to process a packet that is at least  $\tau$ -old,  $\Phi_\tau$  decreases either by  $\lambda_i G_k^i$  for some  $k$  if the new age remains above  $\tau$ , or the term vanishes, in which case  $\Phi_\tau$  may decrease by less. Crucially, this latter possibility can only happen at most once. Again, write  $N_\tau^i$  for the number of packets that queue  $i$  clears during this interval that are at least  $\tau$ -old. Then as  $\sum_{i=1}^n N_\tau^i = N_\tau$ , the decrease in  $\Phi$  from  $\Phi_\tau$ , denoted  $\Delta_\tau$ , satisfies

$$\Delta_\tau \geq \sum_{i=1}^n \lambda_i \left( \sum_{k=1}^{N_\tau^i-1} G_k^i \right) \geq \sum_{i=1}^n \lambda_i \left( \frac{N_\tau^i - 1}{\lambda_i} - \epsilon_i w \right) = \sum_{i=1}^n (N_\tau^i - 1 - \lambda_i \epsilon_i w) = N_\tau - n - n \epsilon w,$$

where we apply Equation (10) in the second inequality. By applying the lower bound derived in Equation (19), and then Equation (7) with the definition of  $\epsilon$ , we further obtain

$$N_\tau - n - n\epsilon w \geq \frac{1}{2}w(1-\delta) \sum_{j=1}^{\min\{m, |J_\tau|\}} \mu_j - n\varphi_{\eta/(128n)}(w) - n - n\epsilon w \geq \frac{1}{2}w(1-\delta) \sum_{j=1}^{\min\{m, |J_\tau|\}} \mu_j - \frac{\delta w \mu_1}{2}.$$

By absorbing the last term into the sum, we thus conclude that

$$\Delta_\tau \geq \frac{1}{2}w(1-2\delta) \sum_{j=1}^{\min\{m, |J_\tau|\}} \mu_j. \quad (20)$$

Summing this over all  $\tau$ , the decrease in  $\Phi$  before considering aging satisfies

$$\Delta\Phi \triangleq \sum_{\tau=1}^{\infty} \Delta_\tau \geq \frac{1}{2}w(1-2\delta) \sum_{\tau=1}^{\infty} \sum_{j=1}^{\min\{m, |J_\tau|\}} \mu_j = \frac{1}{2}w(1-2\delta) \sum_{j=1}^{\min\{m, n\}} \tau^{(j)} \mu_j, \quad (21)$$

where  $\tau^{(j)}$  is the  $j$ th largest of the  $\tau_j$ .

*Effect of Aging.* We now account for the increase due to aging by  $w$  over the course of this interval. The increase in  $\Phi$  from this is upper bounded by

$$\frac{1}{2} \sum_{i=1}^n \lambda_i (\tau_i + w)(\tau_i + w - 1) - \frac{1}{2} \sum_{i=1}^n \lambda_i \tau_i (\tau_i - 1) = w \sum_{i=1}^n \lambda_i \tau_i + \frac{1}{2} \sum_{i=1}^n \lambda_i w(w - 1) \leq w \sum_{i=1}^n \lambda_i \tau_i + \frac{1}{2} \sum_{i=1}^n \lambda_i w^2.$$

Note that this is only exact for those  $\tau_i$  that are nonzero but is an upper bound for those that are zero. Combining this with Equation (21), we see that the potential decrease is at least

$$\frac{1}{2}w(1-2\delta) \sum_{j=1}^{\min\{m, n\}} \tau^{(j)} \mu_j - \left( w \sum_{i=1}^n \lambda_i \tau_i + \frac{1}{2} \sum_{i=1}^n \lambda_i w^2 \right) \geq \frac{\eta}{4} w \sum_{i=1}^n \lambda_i \tau_i - \frac{1}{2} \sum_{i=1}^n \lambda_i w^2, \quad (22)$$

as  $2\delta = \eta/2$  and using Assumption 5.2 with the fact that the sum of a product of two nonnegative sequences is maximal when both are in the same sorted order (see Lemma C.1).

*Relating  $\tau_i$  and  $T_{\ell \cdot w}^i$ .* We now need the following claim that relates the  $\tau_i$  and  $T_{\ell \cdot w}^i$ .

CLAIM 5.2. *If there exists an  $i \in [n]$  such that  $\lambda_i T_{\ell \cdot w}^i > 16nw$  under the good event assumptions, then*

$$\sum_{i=1}^n \lambda_i \tau_i \geq \frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i.$$

By Claim 5.1, the precondition of Claim 5.2 holds for our threshold value, so Claim 5.1, Claim 5.2, and Equation (22) imply that the decrease in  $\Phi$  on this good event is at least

$$\frac{\eta}{8} w \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i - \frac{1}{2} \sum_{i=1}^n \lambda_i w^2 \geq \frac{\eta}{16} w \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i.$$

Translating this into the decrease in  $\sqrt{\Phi}$ , Fact A.2 implies that the contribution toward the expected decrease on this event, which occurs with probability at least  $1 - \eta/128 \geq 1/2$ , is at least

$$\left(\frac{1}{2}\right) \frac{\frac{\eta}{16} w \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i}{2 \cdot \sqrt{\frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1)}} = \frac{\eta w \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i}{64 \cdot \sqrt{\frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1)}}. \quad (23)$$

*Considering “Bad” Events.* We now analyze the bad event where any of these assumptions fails: the worst case is that all queues clear no packets, and so each  $T_{\ell \cdot w}^i$  increases by  $w$  on the next  $w$  steps. The increase in  $\Phi$  is thus at most

$$\frac{1}{2} \sum_{i=1}^n \lambda_i (T_{\ell \cdot w}^i + w)(T_{\ell \cdot w}^i + w - 1) - \frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1) \leq w \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i + \frac{1}{2} \sum_{i=1}^n \lambda_i w^2 \leq 2w \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i,$$

where the last inequality is again Claim 5.1. Translating to square roots again, on this bad event that occurs with probability at most  $\eta/128$ , the contribution of increase to the expected change in  $\sqrt{\Phi}$  is at most

$$\left( \frac{\eta}{128} \right) \frac{2w \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i}{2 \cdot \sqrt{\frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1)}} = \frac{\eta w \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i}{128 \cdot \sqrt{\frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1)}}, \quad (24)$$

by Fact A.1. Summing Equations (23) and (24), it follows that  $\sqrt{\Phi}$  decreases in expectation by at least

$$\frac{\eta w \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i}{128 \cdot \sqrt{\frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1)}} \geq \frac{\eta w \sqrt{\lambda_n}}{128},$$

where the last inequality is Lemma A.1. This proves that the drift condition holds for this stochastic process with the threshold given previously.

*Bounded  $p$ th Moments.* The last thing to check to apply Theorem 5.3 is that the increments  $Z_{\ell+1} - Z_\ell$  have conditionally bounded  $p$ th moments for each even integer  $p \geq 2$  to obtain boundedness of our sequence in  $L^r$  for all  $r \geq 0$ . This is relatively straightforward but tedious: therefore, we prove the following claim in Appendix E.

**CLAIM 5.3.** *For each even integer  $p \geq 2$ , there exists a constant  $C_{p,n,w} > 0$  that depends only on  $n, w, p$  and the parameters of the system such that for all  $\ell \geq 0$ ,*

$$\mathbb{E}[|Z_{\ell+1} - Z_\ell|^p | \mathcal{H}_\ell] \leq C_{p,n,w}.$$

Therefore, Theorem 5.3 applies to the random process  $Z_\ell$ , and we conclude that for each  $r \geq 0$ , there exists some absolute constant  $C_r$  such that for all  $\ell = 0, 1, \dots$ ,

$$\mathbb{E}[Z_\ell^r] \leq C_r.$$

In particular, this means that for each  $\ell \geq 0$ ,

$$\mathbb{E} \left[ \left( \sqrt{\sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1)} \right)^r \right] \leq C_r.$$

To extend this to all  $t \geq 0$  not necessarily a multiple of  $w$ , it is clear that deterministically,  $\sqrt{\Phi(\mathbf{T}_t)} \asymp \|\mathbf{T}_t\|_{\lambda,2} \asymp \|\mathbf{T}_t\|_{\lambda,1}$  up to additive and multiplicative constants as in Lemma A.1, from which we can conclude that

$$\mathbb{E} \left[ \left( \sum_{i=1}^n \lambda_i T_t^i \right)^r \right] \leq C'_r$$

for some other constant  $C'_r$  for each  $t = \ell w$ . Note that for  $\ell w \leq t < (\ell+1)w$ , each term can increase by at most  $w$  compared to the value at  $\ell w$ , and therefore we can conclude that for *all*  $t \geq 0$ ,

$$\mathbb{E} \left[ \left( \sum_{i=1}^n T_t^i \right)^r \right] \leq C''_r$$

for some constant  $C''_r$  independent of  $t$ . This concludes the proof of strong stability.  $\square$



*Remark 5.1.* This analysis crucially relies on using the Geometric system as opposed to the Bernoulli system. The reason is that the preconditions in Theorem 5.3 must hold conditioned on any history, even low probability events. In the Bernoulli system, this would require us to condition on too much. For instance, it is possible for there to be a queue with a very old packet and yet have received no other packets until the current timestep. Although unlikely to actually ever happen, this is a perfectly valid potential history. In this case, clearing this packet would lead to an arbitrarily large  $p$ th moment change, as her age would drastically decrease, and therefore the moment condition of Theorem 5.3 would be violated. Although intuitively this should only help the stability of the random process, the conditions in Theorem 5.3 are subtle (see the discussion in the work of Pemantle and Rosenthal [29]).

Even if that obstruction can be managed suitably, the extra conditional information in the Bernoulli system highly complicates the analysis, as then one must reason about the priorities of the packets that have already been received before the present timestep. These could in principle be quite arbitrary. We avoid these complications in the Geometric system, as it allows us to only condition on current ages.

We now provide a simple construction showing that a partial converse holds:  $\frac{1}{2}$  is the best constant that can appear in Assumption 5.2 for a similar no-regret condition to be sufficient for stability as in Theorem 5.2. To set it up, let  $w_k = k^2$  for each  $k \geq 1$ . Then we have the following theorem.

**THEOREM 5.4.** *Partition time  $t = 0, 1, \dots$  into consecutive windows, where the  $k$ th window has length  $w_k = k^2$ . Then there exists a family of queuing systems with  $n$  queues and servers for each  $n \geq 1$  satisfying Assumption 5.2 with  $\frac{1}{2} + o_n(1)$  in place of  $\frac{1}{2}$  with the following properties: almost surely, each queue has zero regret on all but at most finitely many of the windows, but the system is not weakly stable.*

The formal details are slightly technical, and therefore the proof is deferred to Appendix E, but the high-level idea is quite natural: for each  $n \geq 1$ , consider the following system on  $n$  queues and  $n$  servers where we set  $\lambda = (\frac{n+1}{n^2}, \dots, \frac{n+1}{n^2})$  and  $\mu = (1, \frac{n-1}{n^2}, \dots, \frac{n-1}{n^2})$ . Consider the strategy where every queue always sends to the rate 1 server. It is easy to see purely from expectations that the queue lengths are unbounded in expectation, as the sum of arrival rates strictly exceeds 1. However, it is intuitive that this strategy will “usually” be zero regret; if all the queues are similarly aged at the start of some window, then they should expect to clear roughly  $1/n$  fraction of the time on this window using this strategy, which strictly exceeds what they would get at any other server. We use standard concentration arguments and the Borel-Cantelli lemma to argue that this situation will happen all but finitely many times almost surely, thereby obtaining the claim.

## 6 ASYMPTOTIC CONVERGENCE

In this section, we finally return to proving Theorem 3.3, which asserted the equivalence between the long-run rates of queue ages that form the cost functions for the patient queuing game and the output of the algorithm given in Section 3.1. The high-level idea is to show that this identity holds for all  $i \in S_1$ , then  $S_2$ , and so on. The first step is showing that the maximum queue age grows by at most the desired rate on each long-enough window with high probability.

**PROPOSITION 6.1.** *Fix  $\epsilon > 0$ . For any integer  $a \in \mathbb{N}$ , let  $w = a \cdot \lceil \frac{a}{\epsilon} \rceil^{n-1}$ . Suppose that it holds at time  $t$  that  $\max_{i \in [n]} T_t^i \geq w \cdot f_1$ . Then*

$$\max_{i \in [n]} T_{t+w}^i - \max_{i \in [n]} T_t^i \leq (1 - (1 - \epsilon) \cdot f_1) \cdot w$$

with probability at least  $1 - C_1 \exp(-C_2 a)$ , where  $C_1, C_2 > 0$  are absolute constants depending only on  $n, \epsilon, \lambda, \mu, p$ , but not on  $a$ . More generally, for each  $s \geq 1$ , if  $\max_{i \notin U_{s-1}} T_t^i \geq w \cdot f_s$ , then

$$\max_{i \notin U_{s-1}} T_{t+w}^i - \max_{i \notin U_{s-1}} T_t^i \leq (1 - (1 - \epsilon) \cdot f_s) \cdot w$$

with probability at least  $1 - C_1 \exp(-C_2 a)$ , where  $C_1, C_2 > 0$  are absolute constants depending only on  $n, \epsilon, \lambda, \mu, p$ , but not on  $a$ .

We prove Proposition 6.1 in Appendix F using a delicate argument relying on a variety of concentration bounds. The key insight is that if a subset of much older queues  $S$  is likely to have priority on a long window of length  $w$ , the quantity  $w \cdot f(S_1) \cdot \lambda(S)$  is a *lower bound* on the expected number of packets cleared *collectively* by  $S$  on this window by definition of  $S_1$ . The analysis gets complicated when there are multiple old queues, as although we know these queues *collectively* have priority over all young queues, we must argue about priorities *within* this subset to bound the growth of the maximum queue age. We deal with this by induction by carefully chaining together large windows to obtain a win-win analysis.

For Proposition 6.1 to yield anything useful, we will need a corresponding lower bound asserting roughly that if groups have separated according to what the algorithm asserts, then the aging rate of the *average* queue in a group grows at the conjectured rate. To that end, we prove the following result in Appendix F, which shows that *if* we have the conjectured aging separation between groups  $U_{k-1}$  and  $S_k$ , then some weighted combination of the queue ages in  $S_k$  (whose significance will prove apparent momentarily) must rise quickly.

**PROPOSITION 6.2.** *For any  $s \geq 1$  and any fixed  $\epsilon > 0$ , the following holds: suppose that at time  $t$ , it holds that*

$$\min_{i \in U_s} T_t^i - \max_{i \in S_{s+1}} T_t^i \geq 2 \cdot \frac{w}{\lambda_n}.$$

*Then with probability  $1 - A \exp(-Bw)$  where  $A, B > 0$  are absolute constants not depending on  $w$ , we have*

$$\sum_{i \in S_{s+1}} \lambda_i T_{t+w}^i - \sum_{i \in S_{s+1}} \lambda_i T_t^i \geq (1 - (1 + \epsilon) f_{s+1}) \cdot w \cdot \left( \sum_{i \in S_{s+1}} \lambda_i \right).$$

*Moreover, for any fixed  $\epsilon > 0$ , with probability at least  $1 - A \exp(-Bw)$  it holds that*

$$\sum_{i \in S_1} \lambda_i T_w^i \geq (1 - (1 + \epsilon) f_1) \cdot w \cdot \left( \sum_{i \in S_1} \lambda_i \right).$$

Combined with Proposition 6.1, this will allow us to conclude that because the average queue and oldest queue in  $S_1$  ages at the desired rate almost surely, *all* queues in  $S_1$  must age at this rate almost surely. To extend this analysis to lower groups  $S_2$  and so on, we will use a similar analysis to show that the maximum age over every queue not in  $S_1$  grows at most like  $r(S_2)$ . Then, because we know that every queue in  $S_1$  grows by a  $r(S_1) > r(S_2)$  rate, almost surely, eventually every queue in  $S_1$  will be much older than every queue not in  $S_1$ , giving priority. We leverage this fact to show again that the average queue in  $S_2$  must grow by at least  $r(S_2)$ , and therefore every queue in  $S_2$  grows at this rate almost surely. The proof for the lower groups  $S_3, \dots$  is completely analogous. We now proceed to make this argument formal to prove Theorem 3.3.

**PROOF OF THEOREM 3.3.** By the dominated convergence theorem, it suffices to show the second equality. We will show that the desired statement holds for each  $i \in S_1$ , then  $S_2$ , and so on. We first treat the case that the last outputted group  $S_k$  satisfies  $g_k = 0$ , or equivalently that  $f_k \geq 1$ . Fix

$\epsilon > 0$  and partition time into consecutive windows of size  $w_\ell = \ell \cdot \lceil \frac{6}{\epsilon} \rceil^{n-1}$ . Let  $W_\ell = \sum_{q=1}^{\ell-1} w_q$  be the time period at the beginning of the  $\ell$ th window, and note that  $w_\ell = \Theta(W_\ell^{1/2})$ .

Consider the following events for  $\ell = 1, 2, \dots$

$$A_\ell = \left\{ \max_{i \in [n] \setminus U_{k-1}} T_{W_{\ell+1}}^i - \max_{i \in [n] \setminus U_{k-1}} T_{W_\ell}^i > (1 - (1 - \epsilon) \cdot f_k) \cdot w_\ell \right\} \quad B_\ell = \left\{ \max_{i \in [n] \setminus U_{k-1}} T_{W_\ell}^i \geq w_\ell \cdot f_k \right\}$$

$$C_\ell = A_\ell \cap B_\ell.$$

Clearly,  $\Pr(C_\ell) \leq \Pr(A_\ell | B_\ell)$ . But by Proposition 6.1, we know that for some constants  $C_1, C_2 > 0$  independent of  $\ell$ ,

$$\Pr(A_\ell | B_\ell) \leq C_1 \exp(-C_2 \cdot \ell).$$

The sum over  $\ell$  is thus finite, and so the first Borel-Cantelli lemma (Lemma B.1) implies that almost surely at most finitely many of the  $C_\ell$  occur. Equivalently, almost surely, for all but finitely many of the  $\ell$ , either

$$\max_{i \in [n] \setminus U_{k-1}} T_{W_{\ell+1}}^i - \max_{i \in [n] \setminus U_{k-1}} T_{W_\ell}^i \leq (1 - (1 - \epsilon) \cdot f_k) \cdot w_\ell \quad \text{or} \quad \max_{i \in [n] \setminus U_{k-1}} T_{W_\ell}^i < w_\ell \cdot f_k.$$

Observe that for each of the intervals where the latter holds, the value during the interval is at most  $w_\ell \cdot f_k + w_{\ell+1} = O(W_\ell^{1/2})$ . In particular, it is not difficult to see that almost surely  $\max_{i \in [n] \setminus U_{k-1}} T_{W_\ell}^i$  is either  $o(W_\ell)$ , in which case we are done, or grows by at most a rate of  $(1 - (1 - \epsilon) \cdot f_k) \cdot w_\ell$ . Either way, as  $\epsilon > 0$  was arbitrary, we may take  $\epsilon \rightarrow 0$  to deduce the desired result that almost surely<sup>18</sup>

$$\limsup_{t \rightarrow \infty} \frac{\max_{i \in [n] \setminus U_{k-1}} T_t^i}{t} = 0 = g_k,$$

using  $f_k \geq 1$ . As ages of queues are nonnegative, the lower bound of 0 is trivial.

Now we turn to the rest of the groups, and we now assume that  $g_k > 0$ . We do this inductively. For  $S_1$ , fix  $\epsilon > 0$  and define

$$A_t = \left\{ \sum_{i \in S_1} \lambda_i T_t^i < (1 - (1 + \epsilon)f_1) \cdot t \cdot \left( \sum_{i \in S_1} \lambda_i \right) \right\}.$$

By Proposition 6.2, we know that  $\Pr(A_t) \leq A \exp(-Bt)$  for some constants  $A, B > 0$  independent of  $t$ . Therefore,  $\sum_{t=1}^{\infty} \Pr(A_t) < \infty$ , from which the Borel-Cantelli lemma implies that almost surely, for all but finitely many  $t$ ,

$$\sum_{i \in S_1} \lambda_i T_t^i \geq (1 - (1 + \epsilon)f_1) \cdot t \cdot \left( \sum_{i \in S_1} \lambda_i \right).$$

Taking  $\epsilon \rightarrow 0$ , we obtain almost surely that

$$\liminf_{t \rightarrow \infty} \frac{\sum_{i \in S_1} \lambda_i T_t^i}{t} \geq g_1 \cdot \left( \sum_{i \in S_1} \lambda_i \right). \quad (25)$$

<sup>18</sup>For any  $\epsilon > 0$ , we have directly shown that the statement holds for  $t$  of the form  $t = W_\ell$  for  $\ell \geq 1$ . For any  $t$  such that  $W_\ell \leq t < W_{\ell+1}$ ,  $T_t^i$  cannot be more than  $w_\ell$  from the value at  $W_\ell$ , as ages can increase by at most 1 in each period. This implies that on any such intermediate time, the difference in the numerator from the value at  $t = W_\ell$  is  $O(w_\ell) = O(t^{1/2}) = o(t)$  and thus vanishes in the limit when divide by  $t$ , so the lim sup may be taken over all  $t$ , not just the sparsified sequence.

Next, note that deterministically, we have from Fact 2.1 that

$$\min_{i \in S_1} T_t^i \leq \frac{\sum_{i \in S_1} \lambda_i T_t^i}{\sum_{i \in S_1} \lambda_i} \leq \max_{i \in S_1} T_t^i. \quad (26)$$

In particular, we deduce that almost surely,

$$\liminf_{t \rightarrow \infty} \frac{\max_{i \in S_1} T_t^i}{t} \geq g_1 > 0. \quad (27)$$

For the upper bound, let the  $w_\ell$  and  $W_\ell$  be as before, and now define

$$A_\ell = \left\{ \max_{i \in [n]} T_{W_{\ell+1}}^i - \max_{i \in [n]} T_{W_\ell}^i > (1 - (1 - \epsilon) \cdot f_1) \cdot w_\ell \right\} \quad B_\ell = \left\{ \max_{i \in [n]} T_{W_\ell}^i \geq w_\ell \cdot f_1 \right\} \\ C_\ell = A_\ell \cap B_\ell.$$

Again,  $\Pr(C_\ell) \leq \Pr(A_\ell | B_\ell)$ . By Proposition 6.1, a now routine application of the Borel-Cantelli lemma implies that almost surely, for all but finitely many  $\ell$ , either

$$\max_{i \in [n]} T_{W_{\ell+1}}^i - \max_{i \in [n]} T_{W_\ell}^i < (1 - (1 - \epsilon) \cdot f_1) \cdot w_\ell \quad \text{or} \quad \max_{i \in [n]} T_{W_\ell}^i < w_\ell \cdot f_1.$$

But the latter event cannot happen infinitely often with positive probability, as this would imply  $\max_{i \in [n]} T_{W_\ell}^i = o(W_\ell)$  infinitely often with nonzero probability, which violates (27). Therefore, it must be the case that almost surely, for all but finitely many  $\ell$ , the former event holds. This implies that almost surely,

$$\limsup_{t \rightarrow \infty} \frac{\max_{i \in [n]} T_t^i}{t} \leq (1 - (1 - \epsilon) \cdot f_1);$$

taking  $\epsilon \rightarrow 0$  implies that

$$\limsup_{t \rightarrow \infty} \frac{\max_{i \in [n]} T_t^i}{t} \leq g_1.$$

As clearly the left side is an upper bound for the lim sup of only those queues in  $S_1$ , almost surely

$$\limsup_{t \rightarrow \infty} \frac{\max_{i \in S_1} T_t^i}{t} \leq g_1.$$

Combining this with (27), we finally deduce that almost surely,

$$\lim_{t \rightarrow \infty} \frac{\max_{i \in S_1} T_t^i}{t} = g_1.$$

By Equations (25) and (26), we can also conclude that almost surely the same holds for the minimum  $i \in S_1$ . Thus,  $r_i(\mathbf{p}) = g_1$  for all  $i \in S_1(\mathbf{p})$  by definition of  $g_1$ , proving the theorem for all queues in  $S_1$ .

We now show how to extend this inductively to higher values of  $k$  with  $g_k > 0$ . Suppose that we have shown for all  $i \in U_{k-1}$  that the desired almost sure limit holds, and now consider  $S_k$ . A completely analogous argument using the windows  $w_\ell$  as earlier with Proposition 6.1 via the Borel-Cantelli lemma implies that almost surely,

$$\limsup_{t \rightarrow \infty} \frac{\max_{i \in [n] \setminus U_{k-1}} T_t^i}{t} \leq g_k. \quad (28)$$

Now, with these same windows, fix  $\epsilon > 0$  and let

$$\begin{aligned} A_\ell &= \left\{ \sum_{i \in S_k} \lambda_i T_{W_{\ell+1}}^i - \sum_{i \in S_k} \lambda_i T_{W_\ell}^i < (1 - (1 + \epsilon)f_k) \cdot w_\ell \cdot \left( \sum_{i \in S_k} \lambda_i \right) \right\} \\ B_\ell &= \left\{ \min_{i \in U_{k-1}} T_{W_\ell}^i - \max_{i \in S_k} T_t^i \geq 2 \cdot \frac{w_\ell}{\lambda_n} \right\} \\ C_\ell &= A_\ell \cap B_\ell. \end{aligned}$$

Another completely analogous application of Proposition 6.2 and the Borel-Cantelli lemma implies that almost surely, at most finitely many of the  $C_\ell$  occur. In other words, almost surely, for all but finitely many  $\ell$ , either

$$\sum_{i \in S_k} \lambda_i T_{W_{\ell+1}}^i - \sum_{i \in S_k} \lambda_i T_{W_\ell}^i \geq (1 - (1 + \epsilon)f_k) \cdot w_\ell \cdot \left( \sum_{i \in S_k} \lambda_i \right) \quad \text{or} \quad \min_{i \in U_{k-1}} T_{W_\ell}^i - \max_{i \in S_k} T_t^i < 2 \cdot \frac{w_\ell}{\lambda_n}.$$

But the latter event cannot happen infinitely often with any nonzero probability by virtue of the inductive hypothesis and (28), as  $g_{k-1} > g_k$  by Lemma 3.6, which implies that these timestamps cannot be so close infinitely often. Therefore, it must be the case that for all but finitely many of the  $\ell$ , the former event holds. As usual, this immediately implies that

$$\liminf_{t \rightarrow \infty} \frac{\sum_{i \in S_k} \lambda_i T_t^i}{t} \geq (1 - (1 + \epsilon)f_k) \left( \sum_{i \in S_k} \lambda_i \right).$$

Again taking  $\epsilon \rightarrow 0$  thus implies that almost surely,

$$\liminf_{t \rightarrow \infty} \frac{\sum_{i \in S_k} \lambda_i T_t^i}{t} \geq g_k \left( \sum_{i \in S_k} \lambda_i \right),$$

which again coupled with (28) and Fact 2.1 yields that almost surely,

$$\lim_{t \rightarrow \infty} \frac{\min_{i \in S_k} T_t^i}{t} = \lim_{t \rightarrow \infty} \frac{\max_{i \in S_k} T_t^i}{t} = g_k.$$

The extension to all  $i \in S_k$  follows in the same manner as before by comparing with the average, completing the proof.  $\square$

Observe that Theorem 3.3 rather strongly and explicitly characterizes the *linear* almost sure asymptotic growth rates of each queue for any choices of randomizations. Our main result in Theorem 4.1 showed that with a small slack in the system capacity, each queue will be guaranteed *sublinear* asymptotic growth almost surely in any equilibrium. Although our objective function emphasizes the physical interpretation for each queue as an asymptotic linear growth rate, these incentives impose that queues are indifferent between sublinear growth rates. One could instead define the game just using the  $f_k$  quantities directly rather than taking the max with 0 as is needed to argue about the asymptotic growth rates via  $r$ . If queues started out equally backed up, the  $f_k$  quantities measure the linear speed at which their ages descend to zero. In this setting, we provide the following stronger conclusion, whose proof is deferred to Appendix F.

**COROLLARY 6.3.** *Fix  $\mathbf{p}$  and suppose that for some group  $S_k$  output by the algorithm,  $f_k > 1$ , so that  $1 - f_k < 0$ . Then, for each  $i \in S_k$ ,  $T_t^i$  is strongly stable.*

## 7 CONCLUSION AND OPEN QUESTIONS

In this article, we have studied the outcomes of strategic queuing under multiple behavioral assumptions. When considering a patient queuing game restricted to stationary equilibria, we were able to use careful probabilistic arguments to establish the incentive structure of the game. We show that the correct bicriteria factor in this setting is  $\frac{e}{e-1}$  via a novel deformation argument. Turning then to no-regret dynamics, we show that the factor remains constant but degrades to a factor of 2. In total, our work shows that price of anarchy-style bounds are attainable in such repeated games with state in both equilibrium outcomes and learning outcomes, but that these results require substantially different techniques from the existing literature.

Our work leaves open several enticing questions. The most immediate technical question left open is Question 3 to determine the price of stability of the patient queuing game. Doing so would give a more fine-grained analysis of the cost of selfishness rather than independence here. Moreover, although our restriction to time-independent policies in the patient queuing game exhibits quite rich behavior while enabling us to completely characterize the game-theoretic properties, perhaps there is a larger space of strategies where similar results hold.

But a more pressing direction, in our view, is that the gap we show in equilibrium and no-regret outcomes suggests that no-regret behavior can (surprisingly) yield reasonable outcomes here, but may not necessarily be the correct notion of agent behavior in repeated games that carry strong interdependencies between rounds. It is an interesting question as to whether a natural form of non-cooperative learning can arrive at a Nash equilibrium of the patient version or at least result in stable outcomes without reaching an equilibrium. To that end, it may be necessary to explore the theoretical properties of more powerful learning algorithms in such settings that get the best of both worlds: namely, balancing current rewards while maintaining long-run perspective. Whether such results are possible is an exciting open direction in resolving some of the deficiencies of traditional price of anarchy results.

Equally as important, we wonder if it is possible to obtain a more general theory of price of anarchy in stochastic games. If so, it would be quite interesting to see whether such an analysis extends naturally to a reasonable learning dynamic as in the classical setting. The techniques in this work appear quite specialized to the queuing setting and differ significantly in the equilibrium and learning settings. We leave a systematic exploration of this question to future work.

## APPENDICES

### A BASIC INEQUALITIES

FACT A.1. *Suppose that  $a, b, c \geq 0$  and that  $a - b \leq c$ . Then*

$$\sqrt{a} - \sqrt{b} \leq \min \left\{ \frac{c}{2\sqrt{b}}, \sqrt{c} \right\}.$$

PROOF. The first inequality arises from rearranging and concavity of the square root function:

$$\sqrt{a} \leq \sqrt{b}\sqrt{1 + c/b} \leq \sqrt{b}(1 + c/2b).$$

The second follows from assuming without loss of generality that  $a \geq b$  and observing that the claim is implied by  $\sqrt{a} - \sqrt{b} \leq \sqrt{a - b}$ , which holds by squaring and simple algebra.  $\square$

FACT A.2. *Suppose that  $a, b, c \geq 0$ . Then  $a - b \geq c$  implies that*

$$\sqrt{a} - \sqrt{b} \geq \frac{c}{2\sqrt{a}}.$$



PROOF.

$$a - b = (\sqrt{a} - \sqrt{b})(\sqrt{a} + \sqrt{b}) \geq c \implies \sqrt{a} - \sqrt{b} \geq \frac{c}{\sqrt{a} + \sqrt{b}} \geq \frac{c}{2\sqrt{a}}. \quad \square$$

Recall that we defined the following two weighted  $\ell_p$  norms on  $\mathbb{R}^n$ :  $\|x\|_{\lambda,1} \triangleq \sum_{i=1}^n \lambda_i |x_i|$  and  $\|x\|_{\lambda,2} \triangleq \sqrt{\sum_{i=1}^n \lambda_i x_i^2}$ .

LEMMA A.1. For all  $x \in \mathbb{R}^n$ ,

$$\sqrt{\lambda_n} \|x\|_{\lambda,2} \leq \|x\|_{\lambda,1} \leq \sqrt{\sum_{i=1}^n \lambda_i} \|x\|_{\lambda,2}.$$

PROOF. For the first inequality,

$$\|x\|_{\lambda,1}^2 = \sum_{i,j=1}^n \lambda_i \lambda_j |x_i| |x_j| \geq \sum_{i=1}^n \lambda_i^2 x_i^2 \geq \lambda_n \sum_{i=1}^n \lambda_i x_i^2 = \lambda_n \|x\|_{\lambda,2}^2.$$

The second is a routine application of Cauchy-Schwarz:

$$\sum_{i=1}^n \lambda_i |x_i| = \sum_{i=1}^n \sqrt{\lambda_i} (\sqrt{\lambda_i} |x_i|) \leq \sqrt{\sum_{i=1}^n \lambda_i} \sqrt{\sum_{i=1}^n \lambda_i x_i^2} = \sqrt{\sum_{i=1}^n \lambda_i} \|x\|_{\lambda,2}. \quad \square$$

## B PROBABILITY TOOLS

We will use the following concentration results throughout the article.

LEMMA B.1 (FIRST BOREL-CANTELLI LEMMA, THEOREM 2.3.1 OF DURRETT [15]). Let  $A_1, A_2, \dots$  be a sequence of events with  $\sum_{i=1}^{\infty} \Pr(A_i) < \infty$ . Then with probability 1 at most finitely many of the  $A_i$  occur.

LEMMA B.2 (AZUMA-HOEFFDING). Let  $\{\mathcal{F}_k\}_{k \leq n}$  be any filtration and let  $A_k, B_k, \Delta_k$  satisfy the following conditions:

- (1)  $\Delta_k$  is  $\mathcal{F}_k$ -measurable and  $\mathbb{E}[\Delta_k | \mathcal{F}_{k-1}] = 0$ . In other words, the  $\Delta_k$  form a martingale difference sequence.
- (2)  $A_k, B_k$  are  $\mathcal{F}_{k-1}$ -measurable and satisfy  $A_k \leq \Delta_k \leq B_k$  almost surely.

Then

$$\Pr\left(\sum_{k=1}^n \Delta_k \geq t\right) \leq \exp\left(\frac{-2t^2}{\sum_{k=1}^n \|B_k - A_k\|_{\infty}}\right).$$

LEMMA B.3 (ETEMADI, THEOREM 22.5 IN THE WORK OF BILLINGSLEY [6]). Suppose that  $X_1, \dots, X_n$  are independent random variables. Then for any  $x \geq 0$ ,

$$\Pr\left(\max_{1 \leq k \leq n} |Z_k| \geq 3x\right) \leq 3 \max_{1 \leq i \leq n} \Pr(|Z_i| \geq x),$$

where  $Z_k$  is the  $k$ th partial sum of the  $X_i$  (i.e.,  $Z_k = \sum_{i=1}^k X_i$ ).

LEMMA B.4 (THEOREM 1 IN THE WORK OF WITT [42]). Let  $X_1, \dots, X_n$  be i.i.d.  $\text{Geom}(\lambda)$  random variables so that  $\mathbb{E}[X_i] = \frac{1}{\lambda}$ . Let  $s = \frac{n}{\lambda^2}$  and  $Z_n = \sum_{i=1}^n X_i$ . Then for all  $\delta > 0$ ,

$$\Pr\left(Z_n - \frac{n}{\lambda} < -\delta\right) \leq \exp\left(\frac{-\delta^2}{2s}\right) \quad \text{and} \quad \Pr\left(Z_n - \frac{n}{\lambda} > \delta\right) \leq \exp\left(\frac{-\delta}{4} \min\{\delta/s, \lambda\}\right).$$

COROLLARY B.5. *Under the assumptions and notation of Lemma B.4, for any  $\epsilon \in [0, 1]$ ,*

$$\Pr\left(\max_{1 \leq j \leq n} \left|Z_j - \frac{j}{\lambda}\right| > \frac{\epsilon n}{\lambda}\right) \leq 6 \exp\left(\frac{-\epsilon^2 n}{36}\right).$$

PROOF. First apply Lemma B.4 for each partial sum  $Z_j$  and  $\delta = \epsilon n/\lambda$ . By considering the cases  $j \leq \epsilon n$  and  $j > \epsilon n$ , respectively, it follows for all  $j \leq n$  that  $\min\{\delta/s, \lambda\} \geq \epsilon\lambda$ . Lemma B.4 implies that for all  $j \leq n$ ,

$$\Pr\left(\left|Z_j - \frac{j}{\lambda}\right| > \frac{\epsilon n}{\lambda}\right) \leq 2 \exp\left(\frac{-\epsilon^2 n}{4}\right).$$

Now apply Lemma B.3 using the centered random variables  $Y_i = X_i - 1/\lambda$ . This yields

$$\Pr\left(\max_{1 \leq j \leq n} \left|Z_j - \frac{j}{\lambda}\right| > \frac{\epsilon n}{\lambda}\right) \leq 3 \max_{1 \leq j \leq n} \Pr\left(\left|Z_j - \frac{j}{\lambda}\right| > \frac{\epsilon n}{3\lambda}\right) \leq 6 \exp\left(\frac{-\epsilon^2 n}{36}\right). \quad \square$$

COROLLARY B.6. *Let  $\{G_k^i\}_{i \in [n], k \in [w]}$  be a family of independent geometric random variables such that for all  $i, k$ ,  $G_k^i \sim \text{Geom}(\lambda_i)$ . Let  $Z_q^i = \sum_{k=1}^q G_k^i$ . Then for any  $\epsilon \in [0, 1]$ ,*

$$\Pr\left(\exists i \in [n], q \in [w] : \left|Z_q^i - \frac{q}{\lambda_i}\right| \geq \frac{\epsilon w}{\lambda_i}\right) \leq 6n \exp\left(\frac{-\epsilon^2 w}{36}\right). \quad (29)$$

PROOF. This follows immediately from Corollary B.5 and a union bound.  $\square$

LEMMA B.7. *Let  $\{I_k^j\}_{j \in [m], k \in [w]}$  be an independent Bernoulli ensemble such that for all  $j, k$   $I_k^j \sim \text{Bern}(\mu_j)$  with  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_m$ . Then for all  $\delta \in [0, 1]$ ,*

$$\Pr\left(\exists q \in [m] : \sum_{j=1}^q \sum_{k=1}^w I_k^j \leq (1 - \delta)w \left(\sum_{j=1}^q \mu_j\right)\right) \leq m \exp\left(\frac{-\delta^2 w \mu_1}{2}\right). \quad (30)$$

PROOF. The multiplicative form of the Chernoff bound immediately implies that for each  $q \in [m]$ ,

$$\Pr\left(\sum_{j=1}^q \sum_{k=1}^w I_k^j \leq (1 - \delta)w \left(\sum_{j=1}^q \mu_j\right)\right) \leq \exp\left(\frac{-\delta^2 w \sum_{j=1}^q \mu_j}{2}\right) \leq \exp\left(\frac{-\delta^2 w \mu_1}{2}\right).$$

The result then follows from a union bound over all  $q \in [m]$ .  $\square$

The following characterizes the moments of geometric distributions.

LEMMA B.8. *Let  $X \sim \text{Geom}(\lambda)$ . Then for all  $k \geq 1$ ,  $\mathbb{E}[X^k] \leq \frac{c_k}{\lambda^k}$ , where  $c_k$  is a constant depending on  $k$  but not on  $\lambda$ .*

LEMMA B.9. *Let  $X \sim \text{Bin}(n, p)$ , where  $p \in (0, 1]$  is considered fixed. Then, for any fixed integer  $k \geq 0$ ,  $\mathbb{E}[X^k] \asymp n^k$ , where the implicit constants depend on  $p$  and  $k$ , but not  $n$ .*

PROOF. By definition,  $X = \sum_{i=1}^n X_i$ , where  $X_i \sim \text{Bern}(p)$  are i.i.d. We clearly have

$$X^k = \sum_{1 \leq i_1, \dots, i_k \leq n} \prod_{j=1}^k X_{i_j}.$$

Note that products of these indicator variables remain indicator random variables, and it is easy to see that for any indices  $1 \leq i_1, \dots, i_k \leq n$ ,  $p^k \leq \mathbb{E}\left[\prod_{j=1}^k X_{i_j}\right] \leq p$ . Taking expectations and summing, we obtain  $p^k n^k \leq \mathbb{E}[X^k] \leq p n^k$ .  $\square$

## C PROOFS DEFERRED FROM SECTION 2

### C.1 Central Feasibility

LEMMA C.1. *Suppose that  $\mathbf{x}$  majorizes  $\mathbf{y}$ . Then for any nonnegative, monotone decreasing sequence  $z_1 \geq \dots \geq z_n \geq 0$ ,*

$$\sum_{i=1}^n z_i x_i \geq \sum_{i=1}^n z_i y_i.$$

PROOF. Multiply the equations by appropriate scalars in the definition and sum to obtain the inequality.  $\square$

LEMMA C.2 (THEOREM B.2. IN THE WORK OF MARSHALL ET AL. [27]). *Suppose that  $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$  are in sorted order, have equal total sums, and  $\mathbf{x}$  majorizes  $\mathbf{y}$ . Then  $\mathbf{y} = P\mathbf{x}$  for some doubly stochastic matrix  $P$ .*

COROLLARY C.3. *Suppose that  $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$  are in sorted order and  $\mathbf{x}$  strictly majorizes  $\mathbf{y}$ . Then there exists a doubly stochastic matrix  $P$  such that  $P\mathbf{x}$  is strictly greater than  $\mathbf{y}$  componentwise.*

PROOF. By continuity and strict majorization, it is possible to scale all entries of  $\mathbf{x}$  by nonnegative factors strictly less than 1 to obtain a vector  $\mathbf{x}'$  that majorizes  $\mathbf{y}$  and so that the sums are equal. Applying the previous result, we have  $P\mathbf{x}' = \mathbf{y}$  for some doubly stochastic  $P$ . But  $P\mathbf{x}$  strictly exceeds  $P\mathbf{x}'$  componentwise, giving the result.  $\square$

THEOREM C.4 (THEOREM 2.2, RESTATED). *Let  $\lambda \in (0, 1)^n$  and  $\mu \in [0, 1]^m$  be the arrival and service rates, respectively. Then the preceding queuing system is strongly stable for some centralized (coordinated) scheduling policy if and only if for all  $1 \leq k \leq n$ ,*

$$\sum_{j=1}^{\min\{k, m\}} \mu_j > \sum_{i=1}^k \lambda_i. \quad (31)$$

PROOF. The proof of sufficiency was given in the main text, so we just show necessity. It suffices to show that if one of the preceding inequalities fails, the first moment of  $Q_t$  is unbounded over time. To that end, first suppose that strict majorization is strictly violated, namely there is some  $k \leq n$  such that  $\sum_{i=1}^k \lambda_i > \sum_{i=1}^k \mu_i$ . Let  $Q_t^{\leq k} = \sum_{i=1}^k Q_t^i$  be the total number of packets at the  $k$  queues with highest arrival rate. Under any scheduling policy, the difference between  $Q_{t+1}^{\leq k}$  and  $Q_t^{\leq k}$  is bounded below in expectation by  $\sum_{i=1}^k \lambda_i - \sum_{i=1}^k \mu_i > 0$ , as  $\sum_{i=1}^k \lambda_i$  new packets arrive for these queues at each step in expectation, and at most  $\sum_{i=1}^k \mu_i$  packets can be cleared in expectation. In particular, as  $Q_t := \sum_{i=1}^n Q_t^i \geq Q_t^{\leq k}$  surely by nonnegativity of queue sizes, telescoping gives

$$\mathbb{E}[Q_t] \geq \mathbb{E}[Q_t^{\leq k}] = \sum_{s=0}^{t-1} \mathbb{E}[Q_{s+1}^{\leq k} - Q_s^{\leq k}] \geq t \cdot \left( \sum_{i=1}^k \lambda_i - \sum_{i=1}^k \mu_i \right) \rightarrow \infty.$$

Now suppose that strict majorization is only weakly violated, namely there is some  $k \leq n$  such that  $\sum_{i=1}^k \lambda_i = \sum_{i=1}^k \mu_i$ . Again, it is sufficient to show that  $\mathbb{E}[Q_t^{\leq k}] \rightarrow \infty$ . The previous argument shows that  $Q_t^{\leq k}$  is a nonnegative submartingale for any measurable scheduling policy. If  $\lim_{t \rightarrow \infty} \mathbb{E}[Q_t^{\leq k}] = \sup_t \mathbb{E}[Q_t^{\leq k}] < \infty$ , then the Martingale convergence theorem (Theorem 4.2.11 of Durrett [15]) implies that there exists an almost surely finite random variable  $Q_\infty^{\leq k}$  such that  $\lim_{t \rightarrow \infty} Q_t^{\leq k} \rightarrow Q_\infty^{\leq k}$  almost surely. But  $Q_{t+1}^{\leq k} - Q_t^{\leq k}$  is integer valued and not equal to zero with nonzero probability unless  $\mu$  and  $\lambda$  are degenerate, which violates our assumption. This implies that the pointwise limit cannot exist unless it is infinite. This violates the almost sure finiteness of  $Q_\infty^{\leq k}$ .  $\square$

## C.2 Impossibility for No-Priority Model

Next, we give the promised example that the simpler queuing model is too weak to give any sub-polynomial bicriterion result.

**THEOREM C.5 (THEOREM 2.4, RESTATED).** *In the alternate model, for large enough  $n$ , there exists a centrally feasible queuing system with  $n$  queues and servers with the following property: the system remains feasible even if  $\mu$  is scaled down by  $\Omega(n^{1/3})$  and it is possible for all queues to be in a Nash equilibrium of the stage game at each timestep (and in particular, satisfy no-regret properties as in Assumption 5.1), yet the system is not strongly stable.*

**PROOF.** Let  $\lambda_1 = 2/n^{1/3}$ , whereas  $\lambda_2 = \dots = \lambda_n = 1/n^{2/3}$ ; let  $\mu_1 = 1/2$  and  $\mu_2 = \dots = \mu_n = c/n^{1/3}$ , where  $c = c(n) = \Theta(1)$  is such that

$$\frac{1}{n^{1/3} + 2} < \frac{c}{n^{1/3}} < \frac{1}{n^{1/3}}.$$

We proceed by considering an adversarial, centralized scheduler that suggests actions for each queue in each round while ensuring that each agent achieves no regret. The schedule is as follows: in each round, the scheduler chooses  $n^{1/3}/2 - 1$  of the low rate agents arbitrarily to send to the unique high rate server, if that many low rate agents have packets, as well as the high rate queue. All other low rate agents send to distinct low rate servers. If fewer than  $n^{1/3}/2 - 1$  low rate servers are active, then the scheduler schedules all active queues to the high rate server.

By standard Chernoff bounds, the number of low rate queues that receive a packet in a given round is at least  $n^{1/3}/2 - 1$  with probability at least  $1 - \exp(-\Omega(n^{1/3}))$ , so with at least this probability there are enough low rate queues for the first case to hold. The preceding inequalities show that in such a round where there are at least  $n^{1/3}/2 - 1$  active low agents, the suggested schedule is a Nash equilibrium, and the probability of success for each queue sending to the high server is exactly  $1/n^{1/3}$  in such rounds. When this does not occur, the suggested schedule is still Nash, and the probability of success for any queue sending to the high rate server is at most  $1/2$ . Therefore, in any timestep where the high rate queue has a packet, by the Law of Total Probability, her probability of clearing is upper bounded by

$$\frac{1}{n^{1/3}} + \exp(-\Omega(n^{1/3})) \cdot (1/2) < \frac{1.5}{n^{1/3}},$$

where the inequality is for sufficiently large  $n$ . As a result, in expectation  $Q_{t+1}^1 - Q_t^1$  is lower bounded by a nonzero constant (depending on  $n$  but not on  $t$ ), and therefore  $Q_t^1$  diverges with  $t$  in expectation by telescoping. This shows that this system is not strongly stable, even though every queue plays a Nash strategy at each time. Note that this system would still be centrally feasible if all queues were scaled up by a factor of  $\Theta(n^{1/3})$ , giving the result.

To see that this is no regret with high probability on each fixed window, define  $X_t^{i,j}$  to be the indicator variable that queue  $i$  would succeed in clearing a packet at server  $j$  at time  $t$ , and let  $\sigma_i(t)$  be the identity of the server that queue  $i$  chooses at time  $t$ . Note that if queue  $i$  is empty at time  $t$ , then  $X_t^{i,j} = 0$  for all  $j$  and  $\sigma_i(t)$  can be arbitrary. Then define  $\Delta_t^{i,j} = X_t^{i,\sigma_i(t)} - X_t^{i,j}$ . By the preceding Nash discussion,  $\mathbb{E}[\Delta_t^{i,j} | \mathcal{F}_{t-1}] \geq 0$  for all  $t$  in both cases as described earlier, where  $\mathcal{F}_t$  denotes the past history of this process up to time  $t$ . This holds regardless of if queue  $i$  is really sending in that round (in which case the quantity is just 0).

Therefore, as  $|\Delta_t^{i,j}| \leq 2$  surely, we may apply the Azuma-Hoeffding inequality (Lemma B.2) to see that on any fixed window of length  $w$  (and reindexing time so that time progresses  $t = 1, \dots, w$

on this window for notational ease),

$$\Pr\left(\sum_{t=1}^w \Delta_t^{i,j} \leq -\alpha\right) \leq \exp\left(\frac{-\alpha^2}{w}\right). \quad (32)$$

By a union bound, for each queue  $i$ , this holds for all servers  $j \in [m]$  with probability at most  $m \cdot \exp(-\alpha^2/w)$ . Note that if  $\alpha = \sqrt{w \ln(m/\delta)}$ , this quantity is at most  $\delta$ . As such, by definition of regret, on any fixed period of length  $w$ , with probability at least  $1 - \delta$ , this strategy satisfies  $\text{Reg}_i(w) \leq \sqrt{w \ln(m/\delta)} = o(w)$ , as needed.  $\square$

### C.3 Proofs Deferred from Section 2.4

We now show the desired relations between our notions of stability and almost sure subpolynomial growth. We need the following technical lemma.

**LEMMA C.6.** *Suppose that a nonnegative sequence of random variables  $X_1, X_2, \dots$  satisfies  $X_t \leq X_{t-1} + L$  surely for some fixed  $L \geq 0$  and any  $t$ , as well as the moment condition  $\sup_t \mathbb{E}[X_t^p] \leq C_p$  for some constant  $C_p \geq 0$  for each  $p \geq 1$ . Then, for any  $c > 0$ , almost surely,  $X_t = o(t^c)$ .*

**PROOF.** Fix  $\epsilon > 0$ . It suffices to prove the lemma for  $0 < c < 1$ , so take  $0 < d < c$  and set  $p = d^{-1}$ . We do this by proving the desired asymptotics on a conveniently chosen subsequence, then interpolate to intermediate values. Indeed, by Markov's inequality, for each  $k \geq 1$ ,

$$\Pr(X_{k^{1+\epsilon}} > k^{(1+\epsilon)d}) = \Pr(X_{k^{1+\epsilon}}^p > k^{1+\epsilon}) \leq \frac{C_p}{k^{1+\epsilon}}.$$

Summing over  $k$  and observing that the right side is summable, we deduce from the first Borel-Cantelli lemma that almost surely, for all sufficiently large  $k$ ,  $X_{k^{1+\epsilon}} \leq k^{(1+\epsilon)d}$ . To extend this to all large enough  $t$ , suppose that  $t$  is such that  $k^{1+\epsilon} \leq t < (k+1)^{1+\epsilon}$ . By the one-sided boundedness, we know that almost surely, for such  $t$  and all large enough  $k$ ,

$$X_t \leq L \cdot (t - k^{1+\epsilon}) + X_{k^{1+\epsilon}} \leq L \cdot (1 + \epsilon)(k + 1)^\epsilon + k^{(1+\epsilon)d} \leq L \cdot (1 + \epsilon)(t^{1/(1+\epsilon)} + 1)^\epsilon + t^d,$$

where the bound on  $t - k^{1+\epsilon}$  arises from the mean value theorem. Clearly, this last expression is  $O(t^{\epsilon/(1+\epsilon)} + t^d)$ . As this holds for arbitrary  $\epsilon > 0$ , we may take  $\epsilon$  small enough so that this expression is  $o(t^c)$ , as claimed.  $\square$

**LEMMA C.7 (LEMMA 2.7, RESTATED).** *If the Bernoulli and Geometric models characterize the same queuing dynamics, then strong (weak) stability in the Bernoulli system is equivalent to strong (weak) stability in the Geometric system. Moreover, if this holds, then strong stability in either system implies almost sure subpolynomial growth.*

**PROOF.** Suppose that the dynamics are as stated so that the Bernoulli and Geometric dynamics yield completely equivalent processes. Then the distribution of  $Q_t^i$  conditioned on the value of  $T_t^i$  at time  $t$  is  $\text{Bin}(T_t^i, \lambda_i)$ . Note that by the Law of Iterated Expectations,  $\mathbb{E}[(Q_t^i)^p] = \mathbb{E}[\mathbb{E}[(Q_t^i)^p | T_t^i]]$ . But by Lemma B.9,  $\mathbb{E}[(Q_t^i)^p | T_t^i] \asymp (T_t^i)^p$  up to absolute constants depending only on  $p$  and  $\lambda_i$ . Therefore, by taking expectations, the Bernoulli system and the Geometric system have equivalent stability properties. The second claim now follows from either form of strong stability from Lemma C.6, noting that either  $Q_t$  or  $T_t$  can increase by at most  $L = n$  in each timestep.  $\square$

## D PROOFS AND NOTATION FOR SECTION 3

For convenience, we repeat the definitions of various expressions considered in Section 3.

Symbol	Formula	Definition
$\lambda$		Vector of length $n$ of queue arrival rates in descending order
$\Delta^{m-1}$		Probability simplex over the $m$ element set
$\mu$		Vector of length $m$ of server success rates in descending order
$\mathbf{p}$		Vector of queue randomizations over servers in $(\Delta^{m-1})^n$
$T_i^i$		Timestamp of the oldest packet at queue $i$ at timestep $t$
$T_i^i$	$\max\{0, t - T_i^i\}$	Age of queue $i$ at time $t$
$\alpha(S \mathbf{p}, \mu, S')$	$\sum_{j=1}^m \mu_j \prod_{i \in S'} (1 - p_{i,j}) (1 - \prod_{i \in S} (1 - p_{i,j}))$	Expected number of packets cleared by queues in $S$ if all have packets in a round and have priority over all queues except for those in $S'$ and each such queue also has packets in the round
$\lambda(S)$	$\sum_{i \in S} \lambda_i$	Sum of arrival rates of queues in $S$
$f(S \mathbf{p}, \mu, \lambda, S')$	$\alpha(S \mathbf{p}, \mu, S')/\lambda(S)$	Ratio of expected packets cleared by $S$ with priority over all queues except $S'$ to the total arrival rate of $S$
$S_k(\mathbf{p}, \mu, \lambda)$		The $k$ th subset output in the algorithm of Section 3.2
$U_k(\mathbf{p}, \mu, \lambda)$	$\cup_{\ell=1}^k S_\ell(\mathbf{p}, \mu, \lambda)$	Union of the first $k$ outputted subsets in the algorithm of Section 3.2
$r_i(\mathbf{p}, \mu, \lambda)$		Outputted aging rate of queue $i$ in the algorithm of Section 3.2
$f_k(\mathbf{p}, \mu, \lambda)$	$f(S_k(\mathbf{p}, \mu, \lambda) \mathbf{p}, \mu, \lambda, U_{k-1})$	Value of $f$ for $S_k$ when $U_{k-1}$ has priority
$g_k(\mathbf{p}, \mu, \lambda)$	$\max\{0, 1 - f_k(\mathbf{p}, \mu, \lambda)\}$	Outputted rate for $S_k$ ; equivalently, value of $r_i$ for any $i \in S_k$

When the values of, or dependencies on,  $\mathbf{p}, \mu, \lambda$  are clear from context, we omit them for notational ease.

### D.1 Proofs Deferred from Section 3

In this section, we fill in the deferred proofs showing analytic properties of  $r$ , the output of the algorithm described in Section 3.1.

**PROPOSITION D.1 (PROPOSITION 3.7, RESTATED).** *The function  $r : (\Delta^{m-1})^n \rightarrow [0, 1]^n$  given by  $r(\mathbf{p}) = (r_1(\mathbf{p}), \dots, r_n(\mathbf{p}))$  is continuous.*

**PROOF.** Fix  $\mathbf{p}^*$ , a point we wish to show continuity at, and let  $\mathbf{p}^k \rightarrow \mathbf{p}^*$  be a convergent sequence in  $(\Delta^{m-1})^n$ . It is easy to see that because the function  $\mathbf{p} \mapsto \min_{S \subseteq [n]} f(S|\mathbf{p})$  is clearly continuous as the minimum of finitely many continuous functions, the function  $\mathbf{p} \mapsto \max\{1 - f(S_1(\mathbf{p})|\mathbf{p}), 0\} = g(S_1(\mathbf{p}))$  is continuous. Therefore, if  $\max_{i \in [n]} r_i(\mathbf{p}^*) = g_1(\mathbf{p}^*) = 0$ , then by Lemma 3.6, as  $\max_i r_i(\mathbf{p}^k) \rightarrow 0$ , monotonicity yields  $r_i(\mathbf{p}^k) \rightarrow 0$  along the sequence for every  $i \in [n]$ , proving continuity. We will now assume the harder case  $g_1(\mathbf{p}^*) > 0$ .



Before proceeding, define  $\delta > 0$  to be the minimal nonzero gap between  $f(S|S', \mathbf{p}^*)$  and  $f(T|S', \mathbf{p}^*)$  over all choices of  $S, T, S'$  such that  $S, T \subseteq [n] \setminus S'$ —that is,

$$\delta \triangleq \min_{S, T, S': f(S|S', \mathbf{p}^*) \neq f(T|S', \mathbf{p}^*)} |f(S|S', \mathbf{p}^*) - f(T|S', \mathbf{p}^*)|. \quad (33)$$

Note that  $\delta$  is strictly positive, as there are only finitely many choices of  $S, T, S'$ .

Fix  $0 < \epsilon < \delta$ . Clearly, for any fixed  $S, S'$  such that  $S \subseteq [n] \setminus S'$ , the function  $f(S|\mathbf{p}, S')$  is continuous as a function of  $\mathbf{p}$ . For this choice of  $\epsilon$ , we may restrict to a tail of the sequence  $\{\mathbf{p}^k\}$  and reindex so that for all  $k \geq 1$ , and all  $S, S'$ ,

$$|f(S|\mathbf{p}^k, S') - f(S|\mathbf{p}^*, S')| < \epsilon/3. \quad (34)$$

We now claim that for every  $k \geq 1$ , the following holds in the algorithm's outputted rates on  $\mathbf{p}^k$ : although there exists any element  $i \in S_1(\mathbf{p}^*)$  that has not been outputted, the union of outputted subsets to that point must itself be a minimizing subset with respect to  $f(\cdot)$  evaluated with profile  $\mathbf{p}^*$ , and that each element outputted so far has  $f$  value at  $\mathbf{p}^k$  at most  $f(S_1|\mathbf{p}^*) + \epsilon/3$ .

To see this, we proceed inductively: at the beginning of the algorithm, for every tight subset  $S \subseteq S_1(\mathbf{p}^*)$  (by Lemma 3.5), we have by Equation (34) that

$$f(S|\mathbf{p}^k) < f(S|\mathbf{p}^*) + \epsilon/3 = f(S_1(\mathbf{p}^*)|\mathbf{p}^*) + \epsilon/3,$$

whereas for every subset  $T$  that is not minimal, we have by Equations (33) and (34) that

$$f(T|\mathbf{p}^k) > f(T|\mathbf{p}^*) - \epsilon/3 \geq f(S_1(\mathbf{p}^*)|\mathbf{p}^*) + \delta - \epsilon/3 > f(S_1(\mathbf{p}^*)|\mathbf{p}^*) + 2\epsilon/3.$$

In particular, the first outputted subset must be a tight subset for  $\mathbf{p}^*$ , and the rate of each element in that subset is at least the desired amount.

Suppose that this holds inductively, and now let  $S \subseteq S_1(\mathbf{p}^*)$  be the union of the initial outputted sets, which we know is tight. If  $S = S_1(\mathbf{p}^*)$ , we are done, so suppose that there exists  $i \in S_1(\mathbf{p}^*) \setminus S$ . Suppose that  $T$  is disjoint and such that  $S \cup T \subseteq S_1(\mathbf{p}^*)$  is tight. Such sets exist—for instance,  $T = S_1(\mathbf{p}^*) \setminus S$ . From Fact 3.1,

$$f(S \cup T|\mathbf{p}^*) = f(S|\mathbf{p}^*) \oplus f(T|S, \mathbf{p}^*),$$

and by Fact 2.1, the fact that both the left side and the first term on the right are minimal implies that  $f(T|S, \mathbf{p}^*) = f(S_1(\mathbf{p}^*)|\mathbf{p}^*)$ . Then at the next step of the algorithm, we again have by Equation (34) that

$$f(T|S, \mathbf{p}^k) < f(T|S, \mathbf{p}^*) + \epsilon/3 = f(S_1(\mathbf{p}^*)|\mathbf{p}^*) + \epsilon/3.$$

For any disjoint subset  $T'$  such that  $S \cup T'$  is not tight, note that by Fact 3.1,

$$f(S \cup T'|\mathbf{p}^*) = f(S|\mathbf{p}^*) \oplus f(T'|S, \mathbf{p}^*).$$

By Fact 2.1, minimality of the first term on the right, and the fact that the left term is *not* minimal, it follows that  $f(T'|S, \mathbf{p}^*) \geq f(S_1(\mathbf{p}^*)|\mathbf{p}^*) + \delta$  by definition. We then have by Equation (34) that

$$f(T'|S, \mathbf{p}^k) > f(T'|S, \mathbf{p}^*) - \epsilon/3 \geq f(S_1(\mathbf{p}^*)|\mathbf{p}^*) + \delta - \epsilon/3 > f(S_1(\mathbf{p}^*)|\mathbf{p}^*) + 2\epsilon/3.$$

In particular, the next outputted set is such that  $S \cup T \subseteq S_1(\mathbf{p}^*)$  is tight, and the rate condition holds as well. We can iteratively apply this while  $S_1(\mathbf{p}^*)$  is not exhausted, proving that every element of  $S_1(\mathbf{p}^*)$  is outputted before any other element when running the algorithm on  $\mathbf{p}^k$ , with the rate within  $\epsilon/3$  of  $r_i(\mathbf{p}^*)$ . As  $\epsilon$  was arbitrary, this shows continuity on all components of  $S_1(\mathbf{p}^*)$ .

Because we have shown that every queue of  $S_1(\mathbf{p}^*)$  is outputted before every queue not in  $S_1$ , we can apply the recurrence as discussed in Equation (4) to show continuity for each queue in  $S_2(\mathbf{p}^*)$ , just discounting  $\mu$  as usual. The same argument restricted to  $[n] \setminus S_1(\mathbf{p}^*)$  nearly shows continuity; the only difference is that the discounting of  $\mu$  by the queues in  $S_1(\mathbf{p}^*)$  depends on  $\mathbf{p}^k$ , not  $\mathbf{p}^*$ , but as each  $f(S|S_1(\mathbf{p}^*), \mathbf{p}, \mu')$  is jointly continuous in  $\mathbf{p}, \mu'$ , and the composition of continuous functions

is continuous, the same argument holds with minimal modification. This proves continuity for the components of each subsequent group recursively, and thus of each component in  $[n]$ .  $\square$

With these results, we complete the proof of Theorem 3.8. We proceed as follows: fix any queue  $i$ , as well as any fixed probability choices  $p_{-i} \in (\Delta^{m-1})^{n-1}$  by the other players, and any two  $p, p' \in \Delta^{m-1}$ . Define for  $t \in [0, 1]$ ,

$$h(t) = r_i(tp + (1-t)p', p_{-i}).$$

LEMMA D.2. *For any fixed  $i$ ,  $p_{-i} \in (\Delta^{m-1})^{n-1}$ , and  $p_i, p'_i \in \Delta^{m-1}$ , the function  $h(t)$  is piecewise linear and has no local maxima on the interior.*

PROOF. Let  $\mathbf{p}(t) = (tp_i + (1-t)p'_i, p_{-i})$ . By Proposition 3.7,  $h$  is continuous as the restriction of a continuous function, and it is easy to see that it must be piecewise linear in  $t$  by inspection. Indeed, as the algorithmic description of  $c$  takes minimums and maximums of finitely many linear functions, this yields a piecewise linear function with no jump discontinuities.

We now prove the last claim. It is sufficient to show that if  $h$  is increasing at  $t'$ , then it is increasing for all  $t'' > t'$ . Suppose that this is violated for some  $t' < t''$ ; by piecewise linearity, there must exist some  $t^*$  such that as  $t' < t^* < t''$  where two lines intersect in the graph, and so that as  $t \rightarrow t^{*-}$ , the slope is increasing while it is nonincreasing for  $t \rightarrow t^{*+}$ .

Suppose that for all  $t$  that are sufficiently close to  $t^*$  from the left,  $i$  is outputted at step  $k$  of the algorithm. The only reason the slope can go from positive to nonpositive at  $t^*$  is that there is a change in which sets are outputted in the algorithm at some step  $\ell \leq k$ , which can happen only if some new set  $S$  including  $i$  gets selected for  $t \geq t^*$ . However, as the rates of all sets not including  $i$  fixing any other disjoint set having priority are all constants with respect to  $t$ , this can only occur because at  $t^*$ , some linear function  $f(S|\mathbf{p}(t), S')$  went below the  $f(S''|\mathbf{p}(t), S')$  that was previously selected at step  $\ell$ , where  $S'$  is the union of all sets outputted prior at  $t$  and  $S''$  is the set that was outputted next for all  $t$  close enough to the left of  $t^*$ . If  $S''$  included  $i$ , this could only occur if  $r(S|\mathbf{p}(t), S') := \max\{0, 1 - f(S|\mathbf{p}(t), S')\} = 1 - f(S|\mathbf{p}(t), S')$  has larger positive slope than  $r(S''|\mathbf{p}(t), S')$ , so the slope of  $h$  would be strictly larger (and in particular, increasing) for all  $t$  sufficiently close to  $t^*$  on the right, contradicting our assumption that it is nonincreasing. If  $S''$  does not include  $i$ , then  $r(S''|\mathbf{p}(t), S')$  is a constant with respect to  $t$ , so for  $r(S|\mathbf{p}(t), S')$  to exceed it for  $t$  larger than  $t^*$  but be lower for  $t$  less than  $t^*$ , the slope of  $r(S|\mathbf{p}(t), S')$  must also be positive, another contradiction. Both cases lead to a contradiction, proving the claim.  $\square$

COROLLARY D.3. *Fix  $p_{-i} \in (\Delta^{m-1})^{n-1}$ . Then the set of global minimizers of  $r_i(\cdot, p_{-i})$  form a nonempty, closed convex set.*

PROOF. Note that global minima exist by continuity from Proposition 3.7 and the extreme value theorem. Let  $p_i, p'_i$  be global minimizers; if we form the line between them in  $\Delta^{n-1}$  (which is of course convex) and consider the function  $h$  defined on this line, then as there are no local maxima in the interior of  $h$  by the previous lemma, the maximum must lie at an endpoint. This immediately implies that every point on this line is also a global minimizer. Closedness of the set of global minimizers follows immediately from the continuity guaranteed in Proposition 3.7.  $\square$

THEOREM D.4 (THEOREM 3.8, RESTATED). *There exists a pure equilibrium of the game with costs given by  $r : (\Delta^{m-1})^n \rightarrow [0, 1]^n$ .*

PROOF. We will prove this by appealing to Kakutani's theorem. Let  $B : (\Delta^{m-1})^n \rightrightarrows (\Delta^{m-1})^n$  be the correspondence that maps  $\mathbf{p} \in (\Delta^{m-1})^n$  to the set  $B(\mathbf{p}) \subseteq (\Delta^{m-1})^n$ , where  $B(\mathbf{p}) = \{\mathbf{p}' \in (\Delta^{m-1})^n : p'_i \in \arg \min_{x \in \Delta^{m-1}} r_i(x, p_{-i})\}$  is the best-response correspondence.

We must verify the preconditions of Kakutani's theorem.  $(\Delta^{m-1})^n$  is clearly compact and convex, and we have shown that  $B(\mathbf{p})$  is nonempty and is a convex set by Corollary D.3. The final condition to show is that  $r$  has closed graph, which can be done by a completely standard argument; we must show that if  $(\mathbf{p}^k, \mathbf{s}^k) \rightarrow (\mathbf{p}, \mathbf{s})$ , where  $\mathbf{s}^k \in B(\mathbf{p}^k)$ , then  $\mathbf{s} \in B(\mathbf{p})$ . Suppose for a contradiction that this does not hold for some such convergent sequence. This implies that for some  $i \in [n]$ , there exists some  $s'_i$  and  $\epsilon > 0$  such that

$$r_i(s'_i, p_{-i}) + 3\epsilon < r_i(s_i, p_{-i}).$$

As  $p_{-i}^k \rightarrow p_{-i}$ , the continuity of  $r$  from Proposition 3.7 gives for large enough  $k$  that  $r_i(s'_i, p_{-i}^k) \leq r_i(s'_i, p_{-i}) + \epsilon$ . Thus,

$$r_i(s'_i, p_{-i}^k) < r_i(s_i, p_{-i}) - 2\epsilon < r_i(s_i^k, p_{-i}^k) - \epsilon,$$

where the last inequality holds for all large enough  $k$  by continuity of  $r$ . This contradicts the optimality of  $\mathbf{s}^k \in B(\mathbf{p}^k)$ , proving that  $r$  has a closed graph. Kakutani's theorem then immediately implies the existence of a pure equilibrium—that is,  $\mathbf{p} \in (\Delta^{m-1})^n$  such that  $\mathbf{p} \in B(\mathbf{p})$ .  $\square$

## E PROOFS DEFERRED FROM SECTION 5

### E.1 Deferred Claims

CLAIM E.1 (CLAIM 5.1, RESTATED). *Under Equation (13), both of the following statements hold:*

- (1) *There exists some  $i \in [n]$  such that  $\lambda_i T_{\ell \cdot w}^i > 16nw$ .*
- (2)  $\sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i \geq \frac{16}{\eta} w \sum_{i=1}^n \lambda_i$ .

PROOF. Equation (13) immediately implies by definition of  $Z_\ell$  and  $\Phi$  that

$$\sqrt{\sum_{i=1}^n \lambda_i (T_{\ell \cdot w}^i)^2} \geq \frac{w}{\sqrt{\lambda_n}} \max \left( \frac{16}{\eta} \left( \sum_{i=1}^n \lambda_i \right), 16n^2 \right).$$

From Lemma A.1, this implies that

$$\sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i \geq w \max \left( \frac{16}{\eta} \left( \sum_{i=1}^n \lambda_i \right), 16n^2 \right),$$

from which both parts follow, the first from averaging.  $\square$

CLAIM E.2 (CLAIM 5.2, RESTATED). *If there exists an  $i \in [n]$  such that  $\lambda_i T_{\ell \cdot w}^i > 16nw$  under the good event assumptions, then*

$$\sum_{i=1}^n \lambda_i \tau_i \geq \frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i.$$

PROOF. By Equation (10), we must have

$$\lambda_i \tau_i \geq \lambda_i T_{\ell \cdot w}^i - w - \lambda_i \epsilon_i w,$$

even if queue  $i$  clears a packet every step in the window. Observe that if  $\lambda_i T_{\ell \cdot w}^i \geq 8w$ , then

$$w + \lambda_i \epsilon_i w \leq 2w \leq \frac{1}{4} \lambda_i T_{\ell \cdot w}^i,$$

and so  $\lambda_i \tau_i \geq \frac{3}{4} \lambda_i T_{\ell \cdot w}^i$ . We also have

$$\frac{1}{2} \sum_{i: \lambda_i T_{\ell \cdot w}^i < 8w} \lambda_i T_{\ell \cdot w}^i < 4nw.$$

In particular, if there exists some  $i$  such that  $\lambda_i T_{\ell \cdot w}^i > 16nw$ , then

$$\lambda_i \tau_i \geq \frac{3}{4} \lambda_i T_{\ell \cdot w}^i > \frac{1}{2} \lambda_i T_{\ell \cdot w}^i + 4nw \geq \frac{1}{2} \lambda_i T_{\ell \cdot w}^i + \frac{1}{2} \sum_{i: \lambda_i T_{\ell \cdot w}^i < 8w} \lambda_i T_{\ell \cdot w}^i.$$

It follows that if this holds, then indeed,

$$\sum_{i=1}^n \lambda_i \tau_i = \sum_{i: \lambda_i T_{\ell \cdot w}^i \geq 8w} \lambda_i \tau_i + \sum_{i: \lambda_i T_{\ell \cdot w}^i < 8w} \lambda_i \tau_i \geq \frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i.$$

□

CLAIM E.3 (CLAIM 5.3, RESTATED). *For each even integer  $p \geq 2$ , there exists a constant  $C_{p,n,w} > 0$  that depends only on  $n, w, p$  and the parameters of the system such that for all  $\ell \geq 0$ ,*

$$\mathbb{E}[|Z_{\ell+1} - Z_\ell|^p | \mathcal{H}_\ell] \leq C_{p,n,w}.$$

PROOF. By the triangle inequality, it is easy to see that as random variables, the change in  $T_{\ell \cdot w}^i$  is at most  $G^i \triangleq \sum_{k=1}^w G_k^i$ . Then the change in  $\Phi$  is again at most

$$\frac{1}{2} \sum_{i=1}^n \lambda_i (T_{\ell \cdot w}^i + G^i)(T_{\ell \cdot w}^i + G^i - 1) - \frac{1}{2} \sum_{i=1}^n \lambda_i (T_{\ell \cdot w}^i)(T_{\ell \cdot w}^i - 1) \leq \sum_{i=1}^n \lambda_i G^i T_{\ell \cdot w}^i + \frac{1}{2} \sum_{i=1}^n \lambda_i (G^i)^2,$$

as random variables. We treat two different cases separately:

- (1) Suppose that there does not exist  $i \in [n]$  such that  $\lambda_i T_{\ell \cdot w}^i > 1$ . Then the change in  $\Phi$  is at most

$$\sum_{i=1}^n G^i + \frac{1}{2} \sum_{i=1}^n \lambda_i (G^i)^2.$$

From Fact A.1, this means that the change in  $\sqrt{\Phi}$  is upper bounded as random variables by

$$\sqrt{\sum_{i=1}^n G^i + \frac{1}{2} \sum_{i=1}^n \lambda_i (G^i)^2}.$$

Raising this to the  $p$ th power, expanding, and taking expectations, this term is at most  $C_{p,n,w}/\lambda_n^{2p}$  for some constant  $C_{p,n,w}$  depending only on  $n, w$ , and  $p$  by Lemma B.8.

- (2) Suppose that there does exist  $i \in n$  such that  $\lambda_i T_{\ell \cdot w}^i > 1$ . We claim that this implies that for all  $j \in [n]$ ,

$$\frac{\lambda_j T_{\ell \cdot w}^j}{2\sqrt{\frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1)}} \leq \sqrt{\lambda_j}.$$

First, note that for any  $i \in [n]$ ,  $T_{\ell \cdot w}^i \geq 2$  implies that

$$\frac{1}{2} \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1) \geq \frac{1}{4} \lambda_i (T_{\ell \cdot w}^i)^2, \quad (35)$$

as can be confirmed from basic algebra. As  $\lambda_i \leq 1/2$  by feasibility (as  $\mu_1 \leq 1$ ), our assumption implies that  $T_{\ell \cdot w}^i > 2$ , and so

$$2\sqrt{\frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1)} > 1.$$

To prove the claim, we split into more cases: if  $T_{\ell \cdot w}^j \leq 1/\sqrt{\lambda_j}$ , the claim holds using the last inequality in the denominator. Otherwise, we must have  $T_{\ell \cdot w}^j \geq 2$  by integrality, in which case by Equation (35),

$$\frac{\lambda_j T_{\ell \cdot w}^j}{2 \cdot \sqrt{\frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1)}} \leq \frac{\lambda_j T_{\ell \cdot w}^j}{\sqrt{\lambda_j (T_{\ell \cdot w}^j)^2}} = \sqrt{\lambda_j}.$$

Thus, in this case, we have

$$\frac{\sum_{i=1}^n \lambda_i G^i T_{\ell \cdot w}^i + \frac{1}{2} \sum_{i=1}^n \lambda_i (G^i)^2}{2 \cdot \sqrt{\frac{1}{2} \sum_{i=1}^n \lambda_i T_{\ell \cdot w}^i (T_{\ell \cdot w}^i - 1)}} \leq \sum_{i=1}^n \sqrt{\lambda_i} G^i + \frac{1}{2} \sum_{i=1}^n \lambda_i (G^i)^2.$$

By Fact A.1, this is an upper bound as random variables of the change in  $\sqrt{\Phi}$ , so taking  $p$ th powers, expanding, and taking expectations, we get an upper bound of  $C_{p,n,w}/\lambda_n^{2p}$  by Lemma B.8 for some constant  $C_{p,n,w}$  depending only on  $n, w, p, \lambda$ .

□

## E.2 Tightness of Factor 2

**THEOREM E.1 (THEOREM 5.4, RESTATED).** *Partition time  $t = 0, 1, \dots$  into consecutive windows, where the  $k$ th window has length  $w_k = k^2$ . Then there exists a family of queuing systems with  $n$  queues and servers for each  $n \geq 1$  satisfying Assumption 5.2 with  $\frac{1}{2} + o_n(1)$  in place of  $\frac{1}{2}$  with the following properties: almost surely, each queue has zero regret on all but at most finitely many of the windows, but the system is not weakly stable.*

**PROOF.** Define  $W_k = \sum_{i=1}^{k-1} w_i$ . Note that  $W_k = \Theta(k^3) = \Theta(w_k^{3/2})$ .  $W_k$  is the actual timestep at the end of  $k-1$  of the consecutive windows of length  $w_i$  for  $i = 1, \dots, k-1$ . Note also that  $W_{k+1} - W_k = w_k$ .

For each  $n \geq 1$ , consider the following system on  $n$  queues and  $n$  servers: set  $\lambda = (\frac{n+1}{n^2}, \dots, \frac{n+1}{n^2})$  and  $\mu = (1, \frac{n-1}{n^2}, \dots, \frac{n-1}{n^2})$ . This system satisfies Assumption 5.2 with factor  $\frac{1}{2} - o_n(1)$ . We consider the simple strategy where every queue always sends to the rate 1 server. Under these oblivious dynamics, in expectation the total number of packets grows by  $\frac{1}{n}$  with every step, and therefore this system is not even weakly stable. What we must show is that almost surely, this fixed strategy is zero regret for every queue for all but finitely many of the windows.

We first show almost sure concentration of the arrivals of new packets. Let  $\{B_t^i\}_{i \in [n], t \geq 1}$  be the independent random variables for arrivals as usual. Now, for each queue  $i \in [n]$  and  $\ell \geq 0$ , we have

$$\Pr\left(\left|\sum_{t=1}^{\ell} B_t^i - \lambda_i \ell\right| \geq \sqrt{\ell \ln(\ell)}\right) \leq \frac{2}{\ell^2}, \quad (36)$$

where we use the additive form of the Chernoff bound. As the same holds for all queues, the probability this event happens for any of the  $n$  queues is at most  $2n/\ell^2$ . As this is summable in  $\ell$ , we may sum over all  $\ell \geq 1$  to deduce from the Borel-Cantelli lemma that almost surely, for all sufficiently large  $\ell$ , all  $i \in [n]$  satisfy

$$\sum_{t=1}^{\ell} B_t^i = \lambda_i \ell \pm O(\sqrt{\ell \ln(\ell)}). \quad (37)$$

Note that this also implies that almost surely, for all large  $\ell$ ,  $\sum_{t=1}^{\ell} \sum_{i=1}^n B_t^i \geq (1 + \frac{1}{2n}) \cdot \ell$  by the choice of  $\lambda_i$ . Moreover, under this fixed strategy where everyone always sends to the rate 1 server, at most  $\ell$  packets can be cleared by time  $\ell$ .

Next, we show that almost surely, there is a large backup proportional to the current time period. Let  $t_k$  be the last timestamp the rate 1 server clears up to time  $W_k$ . As all queues send there under this fixed strategy, at this point, all queues only have packets that were received after  $t_k$  by priority. On the one hand, it is not difficult to see that deterministically,  $t_k \geq W_k/n$  (equality happens in the worst case where every queue received a packet in every step up to  $W_k$ ). On the other hand, in light of our preceding results, almost surely, for all but finitely many of the  $k$ ,

$$t_k < \frac{1}{1 + \frac{1}{2n}} W_k = (1 - \Omega(1)) W_k. \quad (38)$$

This is because at least  $W_k$  packets have been received up to time  $\frac{1}{1 + \frac{1}{2n}} W_k$ , and because the server can only have cleared at most  $W_k$  packets up to time  $W_k$ , the oldest timestamp the server could have cleared by time  $W_k$  can be at most this quantity.

Next, we show almost sure concentration of the nontrivial server success rates. Let  $I_t^j$  be the indicator that server  $j$  *would* succeed at clearing a packet at time  $t$  (regardless of if one is sent there; under this strategy, no queue ever sends to  $j \neq 1$ ). A similar application of the Chernoff bound and union bound with the Borel-Cantelli lemma implies that almost surely, for all but finitely many of the  $k$ , and for each server  $j \in [n]$ , we have

$$\sum_{t=W_k+1}^{W_{k+1}} I_t^j = w_k \mu_i \pm O(\sqrt{w_k \ln w_k}). \quad (39)$$

Note that the increasing nature of the  $w_k$  is needed here for this to be valid (and in fact, this statement will be false with probability 1 if interval sizes are kept fixed by independence and the second Borel-Cantelli lemma). Thus, almost surely, for all large enough  $\ell$  and  $k$ , all of these events happen simultaneously.

As  $t_k \geq W_k/n$ , almost surely for large enough  $k$ ,  $t_k$  eventually exceeds the random time  $\ell$  at which Equation (37) holds. Consider any subsequent window of length  $w_k$ . Our goal is to use these facts to show that on these windows, all queues have zero regret. First, we show that each queue clears  $(\frac{1}{n} - o(1))w_k$  packets on each such window. Let  $c = \frac{1}{n\lambda_i} < 1$  (note that this is independent of  $i$ ). We know from Equation (38) that  $t_k + w_k < (1 - \Omega(1))W_k + w_k < W_k$ ; moreover, using Equation (37) and the fact that  $t_k \geq \ell$ ,

$$\begin{aligned} \sum_{t=t_k+1}^{t_k+c \cdot w_k} B_t^i &= \sum_{t=1}^{t_k+c \cdot w_k} B_t^i - \sum_{t=1}^{t_k} B_t^i \\ &= \frac{1}{n} \cdot w_k \pm O(\sqrt{(t_k + c \cdot w_k) \ln(t_k + c \cdot w_k)}) \\ &= \frac{1}{n} \cdot w_k \pm O(\sqrt{W_k \ln W_k}) \\ &= \frac{1}{n} \cdot w_k \pm \tilde{O}(w_k^{3/4}), \end{aligned}$$

where the last line uses the relationship between  $W_k$  and  $w_k$ . As  $t_k + w_k < W_k$ , all of these packets were evidently received before the start of the given window, and therefore every queue is backed up throughout the period, and by virtue of the previous equation, each queue has  $\frac{1}{n} - o(1)$  fraction of the next  $w_k$  packets that will be cleared by this top server on this window. Thus, each queue clears at least  $(\frac{1}{n} - o(1)) \cdot w_k$  packets on such windows under this fixed strategy.



Finally, had any queue deviated on such a window to a single fixed low rate server, in light of Equation (39), she would have cleared  $(\frac{n-1}{n^2} + o(1)) \cdot w_k$  packets, which is linearly smaller than the amount she actually cleared. Therefore, almost surely, on all but finitely many of the windows, every queue actually has zero regret.  $\square$

## F PROOFS DEFERRED FROM SECTION 6

This section is devoted to proving the intermediate claims in the proof of Theorem 3.3. Our main first technical result of this section asserts that with high probability, the maximum queue age increases at a rate of at most  $(1 - (1 - \epsilon) \cdot f_1)$  on the next  $w$  steps for a large enough  $w$ . In fact, more generally, the following holds.

**PROPOSITION F.1 (PROPOSITION 6.1, RESTATED).** *Fix  $\epsilon > 0$ . For any integer  $a \in \mathbb{N}$ , let  $w = a \cdot \lceil \frac{6}{\epsilon} \rceil^{n-1}$ . Suppose that it holds at time  $t$  that  $\max_{i \in [n]} T_t^i \geq w \cdot f_1$ . Then*

$$\max_{i \in [n]} T_{t+w}^i - \max_{i \in [n]} T_t^i \leq (1 - (1 - \epsilon) \cdot f_1) \cdot w$$

*with probability at least  $1 - C_1 \exp(-C_2 a)$ , where  $C_1, C_2 > 0$  are absolute constants depending only on  $n, \epsilon, \lambda, \mu, \mathbf{p}$ , but not on  $a$ . More generally, for each  $s \geq 1$ , if  $\max_{i \notin U_{s-1}} T_t^i \geq w \cdot f_s$ , then*

$$\max_{i \notin U_{s-1}} T_{t+w}^i - \max_{i \notin U_{s-1}} T_t^i \leq (1 - (1 - \epsilon) \cdot f_s) \cdot w$$

*with probability at least  $1 - C_1 \exp(-C_2 a)$ , where  $C_1, C_2 > 0$  are absolute constants depending only on  $n, \epsilon, \lambda, \mu, \mathbf{p}$ , but not on  $a$ .*

Note that the first part is simply the  $j = 1$  case of the more general statement. This proposition will follow from the following lemma that will prove inductively.

**LEMMA F.2.** *Fix  $s \geq 1$  and  $\epsilon > 0$ . Then, for each  $1 \leq \tau \leq n$  and all  $a \in \mathbb{N}$ , the following holds: let  $w = a \cdot \lceil \frac{6}{\epsilon} \rceil^{\tau-1}$ , and suppose that at time  $t$ ,  $M^* := \max_{i \notin U_{s-1}} T_t^i \geq w \cdot f_s$ , and that the set*

$$J = \{i \notin U_{s-1} : T_t^i \geq M^* - w \cdot f_s\}$$

*has  $|J| \leq \tau$ . Then, with probability at least  $1 - C_1 \exp(-C_2 a)$ , we have*

$$\max_{i \notin U_{s-1}} T_{t+w}^i - \max_{i \notin U_{s-1}} T_t^i \leq (1 - (1 - \epsilon) \cdot f_s) \cdot w,$$

*where  $C_1, C_2 > 0$  are absolute constants depending only on  $n, \epsilon, \lambda, \mu, \mathbf{p}$ , but not  $a$ .*

Proposition 6.1 follows as for any  $s \geq 1$  and  $a \in \mathbb{N}$ ,  $|J| \leq n$ . We now turn to proving Lemma F.2 inductively. The case for  $\tau = 1$  will turn out to be relatively easy; this case just says that there is a single very old queue among those not in  $U_{s-1}$ , so we will be able to lower bound the number of packets she clears by simply assuming that every queue in  $U_{s-1}$  is older than her. To extend this to higher  $\tau$  will be more difficult. To do this, we will chunk together many windows that we know have this property for smaller values of  $\tau$  and then leverage two facts to get a win-win situation. We will be able to easily show that at least one queue in  $J$  is always decreasing at the correct rate. If all queues in  $J$  are “close,” they also are at the correct rate. If not, then on the next chunk, they will clear at the correct rate on the next chunk inductively with high probability.

We now carry out this high-level plan. For reference, we will use the following similar notation to that used in the main text, but extended to more general windows:

- (1)  $w := B \cdot L$  will denote a given *window* length composed of  $B$  consecutive *blocks* of  $L$  steps. As we will be considering the behavior of the process on some fixed window, we may as well reindex  $t = 1$  for convenience so that each window we consider will go from  $t = 1$  to  $w$ . We will reserve the superscript  $t = 0$  to denote the value of the ages at the very beginning of the window we consider.

- (2) Recall the shorthand  $f_s := f(S_s | U_{s-1})$  and  $g_s := \max\{0, 1 - f_s\}$ .  
 (3) With this convention, we will often define (and will make clear from context) at the beginning of some considered window of fixed length  $w$ , fixed  $s \geq 1$ , and fixed  $\epsilon > 0$ :

$$M^* := \max_{i \in [n] \setminus U_{s-1}} T_0^i \quad T^* := M^* - (1 - \epsilon) \cdot w \cdot f_s.$$

We will often refer to  $T^*$  as the *target* value for this window, which does not change over the course of the window (notice that it is measured at the beginning of the window). Then, define

$$J = \{i \notin U_{s-1} : T_0^i \geq M^* - w \cdot f_s\}.$$

In other words,  $J$  is the set of queues whose age is within  $w \cdot f_s$  of the oldest age, measured at the beginning of the window. Our goal will be to eventually show that if  $w$  is sufficiently large, then with high probability, every queue in  $J$  has age below  $T^*$  at the end of the next  $w$  steps before accounting for  $w$  steps of aging, and of course all queues not in  $J$  are already strictly below  $T^*$  by definition. This will imply that the maximum age grows by at most  $(1 - (1 - \epsilon) \cdot f_s) \cdot w$  once we account for the  $w$  steps of aging over this window.

- (4) Given a window of length  $w = B \cdot L$ ,  $\mathcal{F}_\ell^{(b)}$  is the filtration of  $\sigma$ -algebras generated up to step  $\ell$  in the  $b$ th block, for  $b = 1, \dots, B$ . In particular,  $\mathcal{F}_0^{(1)} \subseteq \mathcal{F}_1^{(1)} \subseteq \dots \subseteq \mathcal{F}_L^{(1)} \subseteq \mathcal{F}_0^{(2)} \subseteq \dots \subseteq \mathcal{F}_\ell^{(b)} \subseteq \dots \subseteq \mathcal{F}_L^{(B)}$ .  
 (5)  $X_{b,\ell}^i$  will be the indicator that queue  $i$  cleared a packet in timestep  $\ell$  in the  $b$ th block.  $X_{b,\ell}^i$  is  $\mathcal{F}_\ell^{(b)}$ -measurable.  
 (6)  $Y_{b,\ell}^i$  will be a sequence of random variables, with same interpretation of the indices, defined as follows: for  $b, \ell$  such that every queue in  $J$  is still above  $M^* - w \cdot f_s$  at the start of the  $\ell$ th step of the  $b$ th block, set  $Y_{b,\ell}^i = X_{b,\ell}^i$ . If this does not hold for some  $b, \ell$ , then let the  $Y_{b,\ell}^i$  be arbitrary indicator random variables satisfying

$$\mathbb{E} \left[ \sum_{i \in I} Y_{b,\ell}^i \middle| \mathcal{F}_{\ell-1}^{(b)} \right] \geq f_s \cdot \left( \sum_{i \in I} \lambda_i \right).$$

Note that  $Y_{b,\ell}^i$  is  $\mathcal{F}_\ell^{(b)}$ -measurable. These random variables are purely for technical convenience because they have an *a priori* lower bound on the conditional expectation, which is not always true of  $X$  (i.e., if queues have already cleared a lot and thus some queues in  $J$  have lost priority over those not in  $J$ ).

- (7)  $G_{b,\ell}^i$  will be i.i.d.  $\text{Geom}(\lambda_i)$  random variables in the same way for  $\ell \in [L], b \in [B]$ . We define the partial sums on each window  $Z_{b,k}^i := \sum_{\ell=1}^k G_{b,\ell}^i$ . For each  $b = 1, \dots, B$ , we make the convention that the  $G_{b,\ell}^i$  are sampled for all  $i \in [n]$  and  $\ell \in [L]$  at the beginning of the  $b$ th block so that they are all  $\mathcal{F}_0^{(b)}$ -measurable; there is no corresponding queuing step. When queue  $i$  clears her  $k$ th packet in the  $b$ th block, her timestamp decreases by  $G_{b,k}^i$ ; equivalently, when this happens, her timestamp will have decreased on the  $b$ th block by  $Z_{b,k}^i$  so far.

**PROOF OF LEMMA F.2.** Fix any group  $s \geq 1$  and  $\epsilon > 0$ . Note that if  $f_s = 0$  (i.e., queues never clear packets), then the lemma holds trivially with probability 1 as every queue in this subset increases deterministically by  $w$  in age on any window of length  $w$ , so we suppose otherwise. The proof is by induction on  $\tau = |J|$ .

*Base Case:  $\tau = 1$  (One Old Queue).* We first consider  $\tau = 1$ . Let  $a \in \mathbb{N}$  be arbitrary, and then set  $w = a$ . If  $i^* \in J$  is the unique queue in the subset, then  $i^* = \arg \max_{i \in [n] \setminus U_{s-1}} T_0^i$ . We must show that at the end of  $w$  steps, this queue has decreased age (before accounting for aging) by at least

$(1 - \epsilon) \cdot w \cdot f_s$  with high probability; the desired conclusion then follows from adding back in the  $w$  steps of aging. For simplicity, write  $\lambda = \lambda_{i^*}$  as the arrival rate of this unique queue. We prove the claim in the following two steps.

*First, we show that with high probability, the number of packets this queue must clear on this window to get below the target is not too large.* Recall that to model this random process on the next  $w$  steps, let  $G_1, \dots, G_w$  be i.i.d.  $\text{Geom}(\lambda)$  random variables (write  $w = 1 \cdot w$  to indicate that our window is just one block of  $w$  steps, so we omit the superscripts for blocks). As usual, write  $Z_k = \sum_{\ell=1}^k G_\ell$ . Recall that when this queue clears her  $k$ th packet on this window, her timestamp decreases by  $G_k$ ; equivalently, clearing  $k$  packets decreases her timestamp by  $Z_k$  collectively. Sample all of these random variables beforehand, and let  $\mathcal{F}_0$  be the  $\sigma$ -algebra generated by the previous history, as well as these random variables.  $\mathcal{F}_\ell$  will be the filtration generated by all prior events up to the  $\ell$ th step in this window.

Next, define the random variable  $K^* = \min\{k \in [w] : Z_k \geq (1 - \epsilon) \cdot w \cdot f_s\}$ , with the convention that if this set is empty, then  $K^* = w + 1$ . Observe that  $K^*$  is exactly the number of packets that this queue must clear to clear the target. We claim that with high probability,  $K^* \leq K := \lambda \cdot (1 - \epsilon/2) \cdot w \cdot f_s$ . To see this, apply Lemma B.4 with  $K = (1 - \epsilon/2) \cdot w \cdot f_s$  and  $\delta = \frac{wf_s\epsilon}{2}$  to see that for this choice of  $K$ ,

$$\Pr(K^* > K) = \Pr(Z_K < (1 - \epsilon) \cdot w \cdot f_s) \quad (40)$$

$$= \Pr\left(Z_K < \frac{\lambda \cdot (1 - \epsilon/2) \cdot w \cdot f_s}{\lambda} - \frac{\epsilon \cdot w \cdot f_s}{2}\right) \quad (41)$$

$$\leq \exp\left(\frac{-w^2 f_s^2 \lambda^2 \epsilon^2}{8\lambda w f_s (1 - \epsilon/2)}\right) \quad (42)$$

$$\leq \exp\left(\frac{-\lambda w f_s \epsilon^2}{8}\right). \quad (43)$$

As  $f_s > 0$ , with high probability, the queue will only need to clear at most  $K$  packets to get below the desired target.

*Next, we show that with high probability, this queue will clear at least as many packets as required by the previous claim.* Let  $X_1, \dots, X_w$  be the indicator variables for if the queue cleared a packet on each step of the window. Note that these random variables are path dependent, but  $X_\ell$  is  $\mathcal{F}_\ell$ -measurable. Then the queue's timestamp decreases by at least  $(1 - \epsilon) \cdot w \cdot f_s$  if and only if  $\sum_{\ell=1}^w X_\ell \geq K^*$  by definition. Therefore, the probability the queue's timestamp decreases by at least  $(1 - \epsilon) \cdot w \cdot f_s$  on the next  $w$  steps is

$$\Pr\left(\sum_{\ell=1}^w X_\ell \geq K^*\right) = \sum_{k=1}^w \Pr\left(\sum_{\ell=1}^w X_\ell \geq K^* \mid K^* = k\right) \Pr(K^* = k).$$

Consider the family  $Y_1, \dots, Y_w$  of indicator random variables that we couple with  $X_1, \dots, X_w$  as follows: while  $\sum_{q=1}^{\ell-1} X_q < K^*$ , set  $Y_\ell = X_\ell$ . While  $\sum_{q=1}^{\ell-1} X_q \geq K^*$  on a sample path, let  $Y_\ell$  be an arbitrary indicator random variable satisfying  $\mathbb{E}[Y_\ell | \mathcal{F}_{\ell-1}] \geq \lambda \cdot f_s$ . Notice that by construction, we always have  $\mathbb{E}[Y_\ell | \mathcal{F}_{\ell-1}] \geq \lambda \cdot f_s$ ; this is because if  $\sum_{q=1}^{\ell-1} X_q < K^*$ , then the queue is still above the target, and therefore by assumption has priority. As this queue is the oldest not in  $U_{s-1}$ , she has priority over all other queues. If  $V \subseteq U_{s-1}$  is some subset of queues with priority over her before she reaches her target, we know that in this case,

$$\mathbb{E}[Y_\ell | \mathcal{F}_{\ell-1}] \mathbb{E}[X_\ell | \mathcal{F}_{\ell-1}] = \lambda \cdot f(\{i^*\} | V) \geq \lambda \cdot f(\{i^*\} | U_{s-1}) \geq \lambda \cdot f_s,$$

where we use set monotonicity and the fact that  $f_s$  is the minimal value of  $f(\cdot|U_{s-1})$  over all subsets contained in the complement.

Of course, if  $\sum_{q=1}^{\ell-1} X_q \geq K^*$ ,  $\mathbb{E}[Y_\ell|\mathcal{F}_{\ell-1}] \geq \lambda f_s$  simply by construction. However, because  $X_\ell$  and  $Y_\ell$  are equal while the queue is above the target, or equivalently before having cleared  $K^*$  packets, it follows that the events that  $\sum_{\ell=1}^w X_\ell \geq K^*$  and  $\sum_{\ell=1}^w Y_\ell \geq K^*$  have the same probability. Recall again that  $K := \lambda \cdot (1 - \epsilon/2) \cdot w \cdot f_s$ . We obtain the probability the queue's timestamp gets below the target is at least

$$\begin{aligned} \sum_{k=1}^K \Pr\left(\sum_{\ell=1}^w X_\ell \geq k \middle| K^* = k\right) \Pr(K^* = k) &= \sum_{k=1}^K \Pr\left(\sum_{\ell=1}^w Y_\ell \geq k \middle| K^* = k\right) \Pr(K^* = k) \\ &\geq \sum_{k=1}^K \Pr\left(\sum_{\ell=1}^w Y_\ell \geq \sum_{\ell=1}^w \mathbb{E}[Y_\ell|\mathcal{F}_{\ell-1}] - \frac{\lambda w f_s \epsilon}{2} \middle| K^* = k\right) \Pr(K^* = k) \\ &\geq \left(1 - \exp\left(-\lambda^2 w f_s^2 \epsilon^2 / 4\right)\right) \Pr(K^* \leq K) \\ &\geq \left(1 - \exp\left(-\lambda^2 w f_s^2 \epsilon^2 / 4\right)\right) \left(1 - \exp\left(\frac{-\lambda w f_s \epsilon^2}{8}\right)\right) \\ &\geq 1 - 2 \exp\left(\frac{-\lambda^2 w f_s^2 \epsilon^2}{8}\right). \end{aligned}$$

We use Azuma-Hoeffding in the fourth line applied to  $\Delta_\ell := Y_\ell - \mathbb{E}[Y_\ell|\mathcal{F}_{\ell-1}]$ , which is surely between  $-1$  and  $1$ . As  $w = a = a \cdot \lceil \frac{6}{\epsilon} \rceil^{1-1}$ , it is clear that the probability this occurs is of the claimed form. To be safe, one should take  $\lambda = \min_i \lambda_i$  to make the bound hold uniformly, independent of the identity of this queue.

*Inductive Step for  $\tau > 1$ .* Suppose that the proposition holds up to  $\tau$ , and now we must show that it holds for  $\tau + 1$ . Let  $a \in \mathbb{N}$ , and then  $w = a \cdot \lceil \frac{6}{\epsilon} \rceil^\tau = \lceil \frac{6}{\epsilon} \rceil \cdot (a \cdot \lceil \frac{6}{\epsilon} \rceil^{\tau-1})$ . In our notation, we have  $w = B \cdot L$ , where  $B = \lceil \frac{6}{\epsilon} \rceil$  and  $L = a \cdot \lceil \frac{6}{\epsilon} \rceil^{\tau-1}$ .

Define as usual

$$M^* = \max_{i \in [n] \setminus U_{s-1}} T_0^i \quad T^* := M^* - (1 - \epsilon) \cdot w \cdot f_s \quad J = \{i \in [n] \setminus U_{s-1} : T_t^i \geq M^* - w \cdot f_s\},$$

and suppose now that  $|J| \leq \tau + 1$  and  $M^* \geq w \cdot f_s$ . For all  $i \in J$ ,  $\ell = 1, \dots, L$ , and  $b = 1, \dots, B$ , define the random variables  $X_{b,\ell}^i$  and  $Y_{b,\ell}^i$  as described earlier, as well as the  $\sigma$ -algebras  $\mathcal{F}_\ell^{(b)}$ .

First, note that for any  $b, \ell$ , if no queue in  $J$  has timestamp below  $M^* - w \cdot f_s$  at the  $\ell$ th step of the  $b$ th block (without accounting for aging), then every queue in  $J$  has priority over queues not in  $J$ . Then for the same reason as in the base case, we have

$$\mathbb{E}\left[\sum_{i \in J} X_{b,\ell}^i \middle| \mathcal{F}_{\ell-1}^{(b)}\right] \geq f_s \cdot \left(\sum_{i \in J} \lambda_i\right).$$

Then by construction, we always have

$$\mathbb{E}\left[\sum_{i \in J} Y_{b,\ell}^i \middle| \mathcal{F}_{\ell-1}^{(b)}\right] \geq f_s \cdot \left(\sum_{i \in J} \lambda_i\right),$$

and this holds regardless of the conditioning. In particular, it follows that for every  $b = 1, \dots, B$ , we have

$$\sum_{\ell=1}^L \mathbb{E}\left[\sum_{i \in J} Y_{b,\ell}^i \middle| \mathcal{F}_{\ell-1}^{(b)}\right] \geq L \cdot f_s \cdot \left(\sum_{i \in J} \lambda_i\right).$$

Now, we define the following events for  $b = 1, \dots, B$ :

$$\begin{aligned} D_b &= \left\{ T_{b,L}^i \leq M^* - (b-3) \cdot L \cdot f_s(1 - \epsilon/2), \forall i \in J \right\} \\ E_b &= \left\{ Z_{b,\ell}^i \geq \frac{\ell}{\lambda} - \frac{\epsilon L f_s}{4}, \forall i \in J, \ell = 1, \dots, L \right\} \\ F_b &= \left\{ \sum_{\ell=1}^L \sum_{i \in J} Y_{b,\ell}^i \geq L \cdot f_s(1 - \epsilon/4) \left( \sum_{i \in J} \lambda_i \right) \right\}. \end{aligned}$$

We make the convention that  $D_0, E_0, F_0$  are just the trivial event  $\Omega$  that happens surely. Then define inductively:

$$A_0 = \Omega \quad A_{b+1} = A_b \cap D_{b+1} \cap E_{b+1} \cap F_{b+1}, \quad b < B.$$

We note that  $A_B$  implies  $D_B$  by construction so that if  $A_B$  holds, then at the end of this window of  $w$  steps (and again without accounting for  $w$  steps of aging), for all  $i \in J$ ,

$$T_w^i = T_{B,L}^i \leq M^* - (B-3) \cdot L \cdot f_s(1 - \epsilon/2).$$

Recall that our goal is that  $T_w^i \leq M^* - B \cdot L \cdot f_s(1 - \epsilon)$  (before accounting for  $w$  steps of aging); note that our choice of  $B = \lceil \frac{6}{\epsilon} \rceil$  satisfies

$$(B-3) \cdot L \cdot f_s \cdot (1 - \epsilon/2) \geq B \cdot L \cdot f_s \cdot (1 - \epsilon).$$

Now, we lower bound the probability that  $A_b$  holds inductively. As  $A_{b+1} \subseteq A_b$ , we have

$$\Pr(A_{b+1}) = \Pr(D_{b+1} \cap E_{b+1} \cap F_{b+1} | A_b) \Pr(A_b) \geq (\Pr(D_{b+1} | A_b) - \Pr(E_{b+1}^c \cup F_{b+1}^c | A_b)) \Pr(A_b),$$

where we just use a union bound. By a union bound and Corollary B.6,

$$\Pr(E_{b+1}^c | A_b) \leq 6(\tau + 1) \exp\left(\frac{-L \cdot \lambda^2 f_s^2 \epsilon^2}{144}\right),$$

where  $\lambda = \min_{i \in [n]} \lambda_i$ . A familiar argument by Azuma-Hoeffding gives

$$\Pr(F_{b+1}^c | A_b) \leq \exp\left(-\left(\sum_{i \in J} \lambda_i\right)^2 L f_s^2 \epsilon^2 / 16n\right) \leq \exp(-\lambda^2 L f_s^2 \epsilon^2 / 16n),$$

where we note that  $\sum_{i \in J} Y_{b,\ell}^i - \mathbb{E}[\sum_{i \in J} Y_{b,\ell}^i]$  is surely between  $-n$  and  $n$  for that extra factor. Therefore, a union bound implies that

$$\Pr(A_{b+1}) \geq \left( \Pr(D_{b+1} | A_b) - 6(\tau + 2) \exp\left(\frac{-\lambda^2 L f_s^2 \epsilon^2}{144n}\right) \right) \Pr(A_b).$$

We now show that  $\Pr(D_{b+1} | A_b)$  is large using a case analysis.

*Case 1: No Gap.* First suppose that after the  $b$ th block, there is no large gap between the maximum and minimum timestamp in  $J$ —that is (without accounting for aging, which affects queues equally),

$$\max_{i \in J} T_{b,L}^i - \min_{i \in J} T_{b,L}^i \leq L \cdot f_s(1 - \epsilon/2).$$

We show that this along with the other assumptions in  $A_b$  already imply the event  $D_{b+1}$ , so there is no need to analyze what happens on the  $b+1$ th block. Note that because there is no large gap,  $D_{b+1}$  will be implied by

$$\min_{i \in J} T_{b,L}^i \leq M^* - b \cdot L \cdot f_s(1 - \epsilon/2), \quad (44)$$

because then (again, as we only account for aging at the end).

$$\max_{i \in J} T_{(b+1) \cdot L}^i \leq \max_{i \in J} T_{b \cdot L}^i \leq M^* - (b-1) \cdot L \cdot f_s(1 - \epsilon/2).$$

We now show that Equation (44) is indeed implied by  $A_b$ , which we recall implies  $E_b$  and  $F_b$  for all  $q \leq b$ . This means that for each  $q \leq b, i \in J, \ell = 1, \dots, L$ ,

$$Z_{q,\ell}^i \geq \frac{\ell}{\lambda} - \frac{\epsilon L f_s}{4} \quad (45)$$

$$\sum_{q=1}^b \sum_{\ell=1}^L \sum_{i \in J} Y_{q,\ell}^i \geq b \cdot L \cdot f_s(1 - \epsilon/4) \left( \sum_{i \in J} \lambda_i \right). \quad (46)$$

If  $\min_{i \in J} T_{b \cdot L}^i \leq M^* - B \cdot L \cdot f_s$ , we are done by the preceding discussion, so suppose not; this means that no queue in  $J$  has a timestamp below  $M^* - w f_s$  (before accounting for aging). By our coupling, we have  $X_{q,\ell}^i = Y_{q,\ell}^i$  up until now, as no queue in  $J$  has lost priority to those outside of  $J$ , so this last inequality is equivalent to

$$\sum_{q=1}^b \sum_{\ell=1}^L \sum_{i \in J} X_{q,\ell}^i \geq b \cdot L \cdot f_s(1 - \epsilon/4) \left( \sum_{i \in J} \lambda_i \right).$$

By averaging, this implies that there exists some  $i \in J$  such that

$$\sum_{q=1}^b \sum_{\ell=1}^L X_{q,\ell}^i \geq b \cdot \lambda_i \cdot L \cdot f_s(1 - \epsilon/4).$$

By Equation (45), this implies that this queue  $i$  has decreased in age by at least

$$b \cdot L \cdot f_s(1 - \epsilon/4) - \frac{b \cdot \epsilon \cdot L f_s}{4} = b \cdot L \cdot f_s(1 - \epsilon/2).$$

Because all queues in  $J$  have a timestamp at most  $M^*$  at the beginning of this process, this implies that

$$\min_{i \in J} T_{b \cdot L}^i \leq M^* - b \cdot L \cdot f_s(1 - \epsilon/2)$$

(before accounting for aging), which combined with the no-gap assumption implies  $D_{b+1}$ . Thus, the no-gap condition implies that the conditional probability of  $D_{b+1}$  is 1.

*Case 2: Large Gap.* Suppose after the  $b$ th block that there is a large gap between the maximum and minimum timestamp in  $J$ ; specifically, suppose that (again, without aging)

$$\max_{i \in J} T_{b \cdot L}^i - \min_{i \in J} T_{b \cdot L}^i > L \cdot f_s(1 - \epsilon/2).$$

The previous case showed that  $A_b$  itself implies that

$$\min_{i \in J} T_{b \cdot L}^i \leq M^* - b \cdot L \cdot f_s(1 - \epsilon/2).$$

As we have conditioned on  $A_b$ ,  $\max_{i \in J} T_{b \cdot L}^i \leq M^* - (b-3) \cdot L f_s(1 - \epsilon/2)$ ; if this held instead with  $b-2$ , we would already be done. If not, as we have set  $L = a \cdot \lceil \frac{6}{\epsilon} \rceil^{\tau-1}$  and evidently the set of queues with timestamp within  $L \cdot f_s$  of the maximum in  $J$  has size at most  $\tau$ , we may apply the inductive hypothesis to see that with probability at least  $1 - C_1 \exp(-C_2 a)$ , for some absolute constants  $C_1$  and  $C_2$  depending only on  $n, \epsilon, \lambda, \mu, \mathbf{p}$  that every such queue decreases by at least



$L \cdot f_s \cdot (1 - \epsilon/2)$  on this block by our choice of  $L$ ; therefore, conditioned on  $A_b$ , with probability at least  $1 - C_1 \exp(-C_2 a)$ , these queues decrease by enough to satisfy  $D_{b+1}$ .

Combining these cases, we conclude that

$$\Pr(A_{b+1}) \geq \left(1 - C_1 \exp(-C_2 a) - 6(\tau + 2) \exp\left(\frac{-\lambda^2 L f_s^2 \epsilon^2}{144n}\right)\right) \Pr(A_b).$$

Unravelling this recurrence and using  $\Pr(A_0) = 1$ , we conclude that

$$\Pr(A_B) \geq 1 - B(C_1 \exp(-C_2 a)) - 6B(\tau + 2) \exp\left(\frac{-\lambda^2 L f_s^2 \epsilon^2}{144n}\right).$$

As  $L = a \cdot \lceil \frac{6}{\epsilon} \rceil^{\tau-1}$  and  $B = \lceil \frac{6}{\epsilon} \rceil$ , this evidently has the claimed form of  $1 - C'_1 \exp(-C'_2 a)$  for absolute constants  $C'_1, C'_2 > 0$  depending only on system parameters and not  $a$ , completing the inductive step. With this, the lemma is proved.  $\square$

**PROPOSITION F.3 (PROPOSITION 6.2, RESTATED).** *For any  $s \geq 1$  and any fixed  $\epsilon > 0$ , the following holds: suppose that at time  $t$ , it holds that*

$$\min_{i \in U_s} T_t^i - \max_{i \in S_{s+1}} T_t^i \geq 2 \cdot \frac{w}{\lambda_n}.$$

*Then with probability  $1 - A \exp(-Bw)$ , where  $A, B > 0$  are absolute constants not depending on  $w$ , we have*

$$\sum_{i \in S_{s+1}} \lambda_i T_{t+w}^i - \sum_{i \in S_{s+1}} \lambda_i T_t^i \geq (1 - (1 + \epsilon)f_{s+1}) \cdot w \cdot \left(\sum_{i \in S_{s+1}} \lambda_i\right).$$

*Moreover, for any fixed  $\epsilon > 0$ , with probability at least  $1 - A \exp(-Bw)$  it holds that*

$$\sum_{i \in S_1} \lambda_i T_w^i \geq (1 - (1 + \epsilon)f_1) \cdot w \cdot \left(\sum_{i \in S_1} \lambda_i\right).$$

**PROOF.** We prove the second statement first, as it is slightly simpler and the main idea will reappear. As usual, let  $X_t^i$  denote the indicator variable that queue  $i$  cleared a packet at time  $t$ . Similar to the previous proof, we have for every  $t \geq 1$ ,

$$\mathbb{E} \left[ \sum_{i \in S_1} X_t^i \middle| \mathcal{F}_{t-1} \right] \leq \left( \sum_{i \in S_1} \lambda_i \right) \cdot f_1.$$

This is an upper bound, as other queues may have priority at time  $t$  and some queues in  $S_1$  may be empty at time  $t$ .

Recall that  $G_\ell^i$  are i.i.d. geometric random variables with parameter  $\lambda_i$  for  $\ell = 1, \dots, w$ . Again, the interpretation is that when queue  $i$  clears her  $\ell$ th packet, her age decreases by  $G_\ell^i$ ; in particular, the cumulative decrease from clearing  $k$  packets is then  $Z_k^i := \sum_{\ell=1}^k G_\ell^i$ . Another familiar application of Corollary B.6 implies that with probability at least  $1 - A \exp(-Bw)$  (where  $A, B$  are absolute constants depending only on  $n, \epsilon, \lambda$ , not  $w$ ),

$$\left| Z_k^i - \frac{k}{\lambda_i} \right| \leq \frac{(f_1 \cdot (\sum_{i \in S_1} \lambda_i) \cdot \epsilon/2) \cdot w}{\lambda_i}.$$

Because the expected number of packets cleared by the queues in  $S_1$  is at most  $\lambda(S_1) \cdot f_1$  by definition, another familiar application of the Azuma-Hoeffding inequality also implies with probability

at least  $1 - A' \exp(-B'w)$  (where  $A', B'$  do not depend on  $w$ ) that

$$\sum_{t=1}^w \sum_{i \in S_1} X_t^i \leq (1 + \epsilon/2) \cdot \left( \sum_{i \in S_1} \lambda_i \right) \cdot f_1 \cdot w.$$

Combining these two estimates, we recall that

$$T_w^i = \max \left\{ 0, w - \sum_{\ell=1}^{\Gamma_w^i} G_{\ell}^i \right\} \geq w - \sum_{\ell=1}^{\Gamma_w^i} G_{\ell}^i,$$

where  $\Gamma_w^i := \sum_{t=1}^w X_t^i$ . Thus, it follows under these good events that

$$\begin{aligned} \sum_{i \in S_1} \lambda_i T_w^i &\geq w \cdot \left( \sum_{i \in S_1} \lambda_i \right) - \sum_{i \in S_1} \lambda_i \left( \sum_{\ell=1}^{\Gamma_w^i} G_{\ell}^i \right) \\ &\geq w \cdot \left( \sum_{i \in S_1} \lambda_i \right) - \sum_{i \in S_1} \lambda_i \cdot \left( \frac{\Gamma_w^i}{\lambda_i} + \frac{(f_1 \cdot (\sum_{i \in S_1} \lambda_i) \cdot \epsilon/2) \cdot w}{\lambda_i} \right) \\ &= w \cdot \left( \sum_{i \in S_1} \lambda_i \right) - \sum_{i \in S_1} \left( \Gamma_w^i + \left( f_1 \cdot \left( \sum_{i \in S_1} \lambda_i \right) \cdot \epsilon/2 \right) \cdot w \right) \\ &\geq w \cdot \left( \sum_{i \in S_1} \lambda_i \right) - w \cdot f_1 \cdot \left( \sum_{i \in S_1} \lambda_i \right) \cdot \epsilon/2 - \sum_{t=1}^w \sum_{i \in S_1} X_t^i \\ &\geq w \cdot \left( \sum_{i \in S_1} \lambda_i \right) - w \cdot f_1 \cdot \left( \sum_{i \in S_1} \lambda_i \right) \epsilon/2 - w \cdot (1 + \epsilon/2) \cdot \left( \sum_{i \in S_1} \lambda_i \right) \cdot f_1 \\ &= w \cdot \left( \sum_{i \in S_1} \lambda_i \right) \cdot (1 - (1 + \epsilon) \cdot f_1). \end{aligned}$$

By a union bound, we thus find that this occurs with probability  $1 - A \exp(-Bw)$  for some possibly different constants  $A, B > 0$  that do not depend on  $w$ , thus concluding the proof of the second statement.

The proof for  $s \geq 1$  is quite similar, just with one extra condition. Suppose that at time  $t$  we have

$$\min_{i \in U_s} T_t^i - \max_{i \in S_{s+1}} T_t^i \geq 2 \cdot \frac{w}{\lambda_n}.$$

Consider now the next  $w$  steps, and let  $G_{\ell}^i$  be as earlier, with  $\ell = 1$  to  $w$  being the same geometric random variables, with  $Z_k^i$  being the  $k$ th partial sum. Another application of Corollary B.6 implies that with probability at least  $1 - A_1 \exp(-A_2 w)$ ,

$$Z_w^i < 1.5 \cdot \frac{w}{\lambda_n}, \quad \forall i \in U_s.$$

Moreover, this event occurring implies that no queue in  $U_s$  can possibly become younger than the any queue in  $S_{s+1}$  on the next  $w$  steps, even if they clear a packet in every step by our assumption. Therefore, on this event, we can apply the same analysis with the variables  $X_t^i$  as earlier, just noting

that conditioned on this occurring,

$$\mathbb{E} \left[ \sum_{i \in S_{s+1}} X_{t+\ell}^i \middle| \mathcal{F}_{t+\ell-1} \right] \leq \left( \sum_{i \in S_{s+1}} \lambda_i \right) \cdot f_{s+1},$$

as every queue in  $U_s$  will have priority over every queue in  $S_{s+1}$  on this window. An extremely similar argument via Azuma-Hoeffding and Corollary B.6 and taking a union bound so that the concentration of queues in  $U_s$  earlier implies that the desired result holds with probability at least  $1 - A \exp(-Bw)$  for some constants  $A, B > 0$  not depending on  $w$  (but again, on  $n, \epsilon, \lambda, \mu, p$ ).  $\square$

**COROLLARY F.4 (COROLLARY 6.3, RESTATED).** *Fix  $p$  and suppose that for some group  $S_k$  output by the algorithm,  $f_k > 1$ , so that  $1 - f_k < 0$ . Then, for each  $i \in S_k$ ,  $T_t^i$  is strongly stable.*

**PROOF SKETCH.** It suffices to show this for the random variable  $\max_{i \notin U_{k-1}} T_t^i$ . Let  $\epsilon > 0$  be small enough such that  $(1 - (1 - \epsilon) \cdot f_k) < \eta < 0$ . Then let  $w = a \cdot \lceil \frac{6}{\epsilon} \rceil^{n-1}$  be large enough so, on the event that  $\max_{i \notin U_{k-1}} T_{\ell \cdot w}^i \geq f_k \cdot w$ , then  $\mathbb{E}[\max_{i \notin U_{k-1}} T_{(\ell+1) \cdot w}^i - \max_{i \notin U_{k-1}} T_{\ell \cdot w}^i \middle| \mathcal{F}_{\ell \cdot w}] < \beta < 0$  for some  $\beta < 0$ , where  $\mathcal{F}_{\ell \cdot w}$  is the filtration of the Geometric system; this can be done by Proposition 6.1, noting that on the event where the proposition fails, the queue age can increase at most by  $w$ , and this can be drowned out in the expectation by the exponential decay of the probability bound by taking  $w$  large enough. This yields negative drift for the random process  $Y_\ell := \max_{i \notin U_{k-1}} T_{\ell \cdot w}^i$  with threshold value  $f_k \cdot w$ .

Then, for any even  $p \geq 0$ ,  $\mathbb{E}[|\max_{i \notin U_{k-1}} T_{(\ell+1) \cdot w}^i - \max_{i \notin U_{k-1}} T_{\ell \cdot w}^i|^p \middle| \mathcal{F}_{\ell \cdot w}]$  is bounded by some constant  $C_p > 0$  for each  $p$  depending only on  $n, w, \lambda$ . This is because the difference between these is crudely upper bounded as random variables by a sum of at most  $n \cdot w$  geometric random variables in the case that queues somehow clear a packet every round, which are easily seen to have bounded moments. By Theorem 5.3, this implies stochastic stability as the  $p$ th moment condition holds for arbitrarily large  $p$ .  $\square$

## ACKNOWLEDGMENTS

We thank Christos Papadimitriou for valuable discussions and insights in the early stages of this work.

## REFERENCES

- [1] Elliot Anshelevich, Anirban Dasgupta, Jon M. Kleinberg, Éva Tardos, Tom Wexler, and Tim Roughgarden. 2008. The price of stability for network design with fair cost allocation. *SIAM J. Comput.* 38, 4 (2008), 1602–1623. <https://doi.org/10.1137/070680096>
- [2] Elliot Anshelevich, Anirban Dasgupta, Éva Tardos, and Tom Wexler. 2008. Near-optimal network design with selfish agents. *Theory Comput.* 4, 1 (2008), 77–109. <https://doi.org/10.4086/toc.2008.v004a004>
- [3] Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. 2002. The nonstochastic multiarmed bandit problem. *SIAM J. Comput.* 32, 1 (2002), 48–77. <https://doi.org/10.1137/S0097539701398375>
- [4] Maria-Florina Balcan, Avrim Blum, and Yishay Mansour. 2013. Circumventing the price of anarchy: Leading dynamics to good behavior. *SIAM J. Comput.* 42, 1 (2013), 230–264. <https://doi.org/10.1137/110821317>
- [5] Lucas Baudin, Marco Scarsini, and Xavier Venel. 2023. Strategic behavior and no-regret learning in queueing systems. *CoRR abs/2302.03614* (2023). <https://doi.org/10.48550/arXiv.2302.03614>
- [6] P. Billingsley. 2008. *Probability and Measure* (3rd ed.). Wiley.
- [7] Avrim Blum, MohammadTaghi Hajiaghayi, Katrina Ligett, and Aaron Roth. 2008. Regret minimization and the price of total anarchy. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*. ACM, New York, NY, 373–382. <https://doi.org/10.1145/1374376.1374430>

- [8] Christian Borgs, Jennifer T. Chayes, Nicole Immorlica, Kamal Jain, Omid Etesami, and Mohammad Mahdian. 2007. Dynamics of bid optimization in online advertisement auctions. In *Proceedings of the 16th International Conference on World Wide Web (WWW'07)*. ACM, New York, NY, 531–540. <https://doi.org/10.1145/1242572.1242644>
- [9] Allan Borodin, Jon M. Kleinberg, Prabhakar Raghavan, Madhu Sudan, and David P. Williamson. 2001. Adversarial queuing theory. *J. ACM* 48, 1 (2001), 13–38. <https://doi.org/10.1145/363647.363659>
- [10] G. W. Brown. 1951. Iterative solutions of games by fictitious play. In *Activity Analysis of Production and Allocation*, Tjalling C. Koopmans (Ed.). Wiley, New York, NY, 374–376.
- [11] Vincent Conitzer, Christian Kroer, Debmalaya Panigrahi, Okke Schrijvers, Eric Sodomka, Nicolás E. Stier Moses, and Chris Wilkens. 2019. Pacing equilibrium in first-price auction markets. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC'19)*. ACM, New York, NY, 587. <https://doi.org/10.1145/3328526.3329600>
- [12] Vincent Conitzer, Christian Kroer, Eric Sodomka, and Nicolás E. Stier Moses. 2018. Multiplicative pacing equilibria in auction markets. In *Web and Internet Economics. Lecture Notes in Computer Science*, Vol. 11316. Springer, 443. <https://link.springer.com/content/pdf/bbm%3A978-3-030-04612-5%2F1.pdf>.
- [13] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. 2009. The complexity of computing a nash equilibrium. *SIAM J. Comput.* 39, 1 (2009), 195–259. <https://doi.org/10.1137/070699652>
- [14] Ofer Dekel, Ambuj Tewari, and Raman Arora. 2012. Online bandit learning against an adaptive adversary: From regret to policy regret. In *Proceedings of the 29th International Conference on Machine Learning (ICML'12)*. 1–8. <http://icml.cc/2012/papers/749.pdf>.
- [15] Rick Durrett. 2019. *Probability: Theory and Examples*. Vol. 49. Cambridge University Press.
- [16] Jerzy Filar and Koos Vrieze. 2012. *Competitive Markov Decision Processes*. Springer Science & Business Media.
- [17] Daniel Freund, Thodoris Lykouris, and Wentao Weng. 2022. Efficient decentralized multi-agent learning in asymmetric queuing systems. In *Proceedings of the 35th Conference on Learning Theory*, Vol. 178. 4080–4084. <https://proceedings.mlr.press/v178/freund22a.html>.
- [18] Hu Fu, Qun Hu, and Jia'nan Lin. 2022. Stability of decentralized queueing networks beyond complete bipartite cases. In *Web and Internet Economics. Lecture Notes in Computer Science*, Vol. 13778. Springer, 96–114. [https://doi.org/10.1007/978-3-031-22832-2\\_6](https://doi.org/10.1007/978-3-031-22832-2_6)
- [19] Drew Fudenberg and David Levine. 1998. *The Theory of Learning in Games*. MIT Press, Cambridge, MA.
- [20] S. Hart and A. M. Colell. 2000. A simple adaptive procedure leading to correlated equilibrium. *Econometrica* 68, 5 (2000), 1127–1150.
- [21] Refael Hassin. 2020. *Rational Queueing*. CRC Press, Boca Raton, FL. <https://books.google.com/books?id=M3G1zQEACAAJ>.
- [22] Refael Hassin and Moshe Haviv. 2003. *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*. Springer US. <https://books.google.com/books?id=K5KhJSZkhHQC>.
- [23] Ramesh Johari and John N. Tsitsiklis. 2004. Efficiency loss in a network resource allocation game. *Math. Oper. Res.* 29, 3 (2004), 407–435. <https://doi.org/10.1287/moor.1040.0091>
- [24] Elias Koutsoupias and Christos H. Papadimitriou. 1999. Worst-case equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS'99)*. 404–413. [https://doi.org/10.1007/3-540-49116-3\\_38](https://doi.org/10.1007/3-540-49116-3_38)
- [25] Subhashini Krishnasamy, Rajat Sen, Ramesh Johari, and Sanjay Shakkottai. 2016. Regret of queueing bandits. In *Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS'16)*. 1669–1677. <http://papers.nips.cc/paper/6370-regret-of-queueing-bandits>.
- [26] Thodoris Lykouris, Vasilis Syrgkanis, and Éva Tardos. 2016. Learning and efficiency in games with dynamic population. In *Proceedings of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'16)*. 120–129. <https://doi.org/10.1137/1.9781611974331.ch9>
- [27] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. 1979. *Inequalities: Theory of Majorization and Its Applications*. Vol. 143. Springer.
- [28] Abraham Neyman and Sylvain Sorin. 2003. *Stochastic Games and Applications*. Vol. 570. Springer Science & Business Media.
- [29] Robin Pemantle and Jeffrey S. Rosenthal. 1999. Moment conditions for a sequence with negative drift to be uniformly bounded in  $L^r$ . *Stoch. Proc. Appl.* 82, 1 (1999), 143–155.
- [30] Julia Robinson. 1851. An iterative method of solving a game. *Ann. Math. Stat.* 54 (1851), 296–301.
- [31] Tim Roughgarden. 2003. The price of anarchy is independent of the network topology. *J. Comput. Syst. Sci.* 67, 2 (2003), 341–364. [https://doi.org/10.1016/S0022-0000\(03\)00044-8](https://doi.org/10.1016/S0022-0000(03)00044-8)
- [32] Tim Roughgarden. 2015. Intrinsic robustness of the price of anarchy. *J. ACM* 62, 5 (2015), Article 32, 42 pages. <https://doi.org/10.1145/2806883>
- [33] Tim Roughgarden. 2016. *Twenty Lectures on Algorithmic Game Theory*. Cambridge University Press, Cambridge, NY.
- [34] Tim Roughgarden and Éva Tardos. 2002. How bad is selfish routing? *J. ACM* 49, 2 (2002), 236–259. <https://doi.org/10.1145/506147.506153>

- [35] Aviad Rubinstein. 2018. Inapproximability of Nash equilibrium. *SIAM J. Comput.* 47, 3 (2018), 917–959. <https://doi.org/10.1137/15M1039274>
- [36] Alexander Schrijver. 2003. *Combinatorial Optimization: Polyhedra and Efficiency*. Vol. 24. Springer Science & Business Media.
- [37] Flore Sentenac, Etienne Boursier, and Vianney Perchet. 2021. Decentralized learning in online queuing systems. In *Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS'21)*. 18501–18512. <https://proceedings.neurips.cc/paper/2021/hash/99ef04eb612baf0e86671a5109e22154-Abstract.html>.
- [38] Shai Shalev-Shwartz and Shai Ben-David. 2014. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.
- [39] John F. Shortle, James M. Thompson, Donald Gross, and Carl M. Harris. 2018. *Fundamentals of Queueing Theory*. Wiley.
- [40] Vasilis Syrgkanis and Éva Tardos. 2013. Composable and efficient mechanisms. In *Proceedings of the 45th Annual Symposium on Theory of Computing (STOC'13)*. ACM, New York, NY, 211–220. <https://doi.org/10.1145/2488608.2488635>
- [41] Neil Walton and Kuang Xu. Learning and information in stochastic networks and queues. In *Tutorials in Operations Research: Emerging Optimization Methods and Modeling Techniques with Applications*, John Gunnar Carlsson (Ed.). INFORMS, 161–198. <https://doi.org/10.1287/educ.2021.0235> <https://pubsonline.informs.org/doi/pdf/10.1287/educ.2021.0235>
- [42] Carsten Witt. 2014. Fitness levels with tail bounds for the analysis of randomized search heuristics. *Inf. Process. Lett.* 114, 1-2 (2014), 38–41. <https://doi.org/10.1016/j.ipl.2013.09.013>

Received 15 December 2021; revised 11 October 2022; accepted 12 January 2023