Random Walk Based Sampling for Load Balancing in Multi-Server Systems

DENGWANG TANG, University of Michigan, USA

VIJAY G. SUBRAMANIAN, University of Michigan, USA

In multi-server systems, a classical job assignment algorithm works as follows: at the arrival of each job, pick d servers independently and uniformly at random and send the job to the least loaded server among the d servers. This model is known as the *power-of- d choices* algorithm. In this paper, we analyze a variant of this algorithm, where d servers are sampled through d independent non-backtracking random walks on a k -regular graph. The random walkers are periodically reset to independent uniform random positions. Under some assumptions on the underlying graph, we show that the system dynamics under this new algorithm converges to the solution of a deterministic ordinary differential equation (ODE), which is the same ODE as the classical *power-of- d choices*. We also show that the new algorithm stabilizes the system, and the stationary distribution of the system converges to the stationary solution of the ODE. The new scheme can be considered as a derandomized version of *power-of- d choices* as it reduces the use of randomness while maintaining the performance of *power-of- d choices*.

CCS Concepts: • **Mathematics of computing** → **Probability and statistics**; *Discrete mathematics*; • **Theory of computation** → *Randomness, geometry and discrete structures*;

Additional Key Words and Phrases: Queueing Theory; Random Walks; Load Balancing; Martingale Concentration Inequality; Fluid Limit Approximation

ACM Reference Format:

Dengwang Tang and Vijay G. Subramanian. 2019. Random Walk Based Sampling for Load Balancing in Multi-Server Systems. *Proc. ACM Meas. Anal. Comput. Syst.* 3, 1, Article 14 (March 2019), 44 pages. <https://doi.org/10.1145/3311085>

1 INTRODUCTION

In the era of Big Data, the computing systems handling the computations typically consist of multiple servers. These computing systems are able to handle multiple jobs at the same time. As the scale of such systems grows larger and larger, load balancing algorithms, which determine how to route jobs to servers, require careful design. A good algorithm can balance the load of the servers and hence use the computational power of the system efficiently. It is well known that the Join-the-Shortest-Queue (JSQ) Algorithm is optimal in several settings [11]. However, JSQ is not practical since it has a huge messaging cost (sending messages at each arrival to all servers to request their queue lengths). This cost is huge especially for today's cloud computing systems, where there can be tens of thousands of servers. Another classical load balancing algorithm, called the *power-of- d -choices* algorithm [22, 23, 29], is as follows: At each job arrival, pick d servers independently and uniformly at random, and dispatch the job to the server with the shortest queue. There has been

Authors' addresses: Dengwang Tang, University of Michigan, Ann Arbor, Michigan, USA, dwtang@umich.edu; Vijay G. Subramanian, University of Michigan, Ann Arbor, Michigan, USA, vgsubram@umich.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 ACM 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.

© 2019 Association for Computing Machinery.

2476-1249/2019/3-ART14 \$15.00

<https://doi.org/10.1145/3311085>

significant work on this algorithm that we briefly outline below. It was shown that, when $d \geq 2$, the queuing delay can be reduced significantly compared to the case where $d = 1$ [20–23, 29]. The analysis for *power-of- d -choices* has been extended to general service time distributions by Aghajani and Ramanan [1]. Subsequent works proposed and analyzed several variants of *power-of- d -choices* model. Mitzenmacher et al. proposed a scheme where the dispatcher keeps m servers in memory at the same time of probing d random servers [24]. The fluid limit approximation result for this scheme is established in [19]. Mukherjee et al. [26] analyzed the power-of- d scheme where $d = d(n)$ scales with the number of systems. Yang, Srikanth, and Kang [32] analyzed a variant of *power-of- d -choices* for batch arrivals, where slightly more than 1 servers is sampled for each task in a job. There are also a series of works on the setting where the servers are interconnected through a graph. Gast [13] investigated asymptotic independence under the scheme where $d = 2$ servers are sampled through a random edge of the underlying graph for several graph topologies. In a recent work, Mukherjee, Borst and Leeuwarden [25] and Budhiraja, Mukherjee, and Wu [6] analyzed a scheme where d random servers are probed through a random vertex and its neighbors. Specifically, arrivals occur at a random vertex i and the job is assigned to the least loaded server of i and its neighbors or a randomly selected subset of its neighbors of size $d - 1$. Under certain assumptions on the graph, it was shown that this system has similar dynamics to that of *power-of- d choices* [6], where the result is insensitive to the specific graph topology. Whereas the schemes considered here use graphs for assignment, there isn't an exploration process such as a random walk, that determines and couples the load assignment strategy across arrivals. Ganesh et al. [12] also considered a model of load balancing on graphs, where the tasks are allowed to switch queues after assignment. In Yun and Proutiere [33], the result is extended to the cases of heterogeneous systems, where the service time distribution could vary based on specific job and server. In this problem setting too, while there's an underlying graph for load assignment/rebalancing, there isn't an exploration process that further couples past load assignments. Many variants of *power-of- d -choices* algorithms are also analyzed in the simplistic and static ball-in-bins model, where the maximum load, or the number of balls in the fullest bin, is analyzed. See [28] for a list of references.

In many of the schemes described above, the sampling is via independent uniformly random variables, which uses $\Theta(\log(n))$ randomness, and hence would be considered inefficient in the computer science literature [16]. In the same literature, random walks on graphs have been suggested as an efficient way of sampling [16]. Alon et al. analyzed the number of visits a non-backtracking random walk on an high-girth expander graph [2] [3]. The result shows that a sample path of a non-backtracking random walk on certain graphs has very similar behaviour to independent random choices in terms of statistics in a time window of certain length, while simple random walks do not share this property. With this as inspiration, in this paper we will analyze a variant of *power-of- d -choices* algorithm based on random walks on graphs. More specifically, based on our recent work in [28], in which three variants of *power-of- d -choices* are analyzed in balls-in-bins model, we seek to investigate the performance of these schemes in queuing system models. With the dynamic nature of the queueing model owing to service completions, in contrast to the static setting of the balls-in-bins model, it is not clear whether the dependent sampling of random walks will still yield good performance. One of the schemes proposed in [28] is as follows: assuming that n servers are connected through a high-girth k -regular graph G , d bins are sampled using d independent non-backtracking random walks on G , where the random walkers are reset to independent uniform random positions periodically. We will call this scheme the *Non-Backtracking Random Walk with Restart based Power-of- d -choices* (NBRWR-Pod) algorithm. In this paper, we analyze the NBRWR-Pod algorithm in the standard low traffic queueing system setting of load $\rho < 1$. The NBRWR-Pod algorithm can be considered as a derandomization of *power-of- d choices* algorithm, since it achieves the same performance of *power-of- d choices* while using less randomness.

The work in [6] and [12] are the closest to ours, where load balancing processes on graphs are analyzed. The key differences are the following: First, in our work, the underlying graph creates interdependence between the choices of different jobs owing to the random walk based sampling, while in [6] and [12], the graph encodes the interdependence between d choices of the same job. Secondly, in [12], the jobs are allowed to switch queues after assignment, while in our work (and also [6]), the jobs stay at the queues they are assigned to until service completion. Finally, our proof technique is different from [6] and [12].

The performance characterization of the NBRWR-Pod algorithm is presented via the following results:

- (1) We analyze the NBRWR-Pod algorithm with the method of *fluid-limit approximation*. We show in Section 3.1 that the system dynamics for a large-scale multi-server system up to any finite time $T > 0$ can be approximated by the solution to a deterministic ordinary differential equation (ODE), which is the same ODE for the regular *power-of- d choices*.
- (2) We show in Section 3.2 that, for any finite n , the NBRWR-Pod algorithm stabilizes the system.
- (3) We show in Section 3.3 that the stationary distribution of the queue lengths under NBRWR-Pod algorithm converges to the stationary solution of the limiting differential equation in L_1 .
- (4) We provide simulation-based results in Section 4 to illustrate that the differential equation can also provide a good approximation for the dynamics of relatively small systems. In there we also study the performance of our algorithm using many graph families, with an eye towards understanding both the strengths and the limitations of our assumptions and our analysis approach.

Our paper makes several methodological and technical contributions to the state of the art, which may be useful in other contexts. These are as follows:

- (1) Since the queue length statistics process for our algorithm is not a Markov process, the standard Kurtz's Theorem [10][9] based fluid limit approximation does not apply. The one-step drift method of Wormald [30] for more general processes does not apply either as we need to consider drift over a period that increases with the number of server nodes (albeit slowly). More interestingly, the use of random walks to sample queues introduces a dependence between queues of neighboring nodes. This we believe precludes the use of the *propagation-of-chaos* method [27], which was used by [6] for their model of *power-of- d choices* on graphs; we offer some evidence of this in Appendix B. Motivated by the proof of Kurtz's theorem, our proof for the fluid limit result is based on a martingale plus compensator decomposition using judiciously coupled processes obtained by sub-sampling, and then using martingale concentration inequalities and Gronwall's lemma. Methodological similarities in using the martingale plus compensator method can also be found with [8] and [19], both of which analyze Markov processes.
- (2) For the stability result, our proof utilizes a novel technique of applying the Foster-Lyapunov Theorem on a sub-sampled process, which is much simpler than the more generally applicable renovating events method [5].
- (3) We follow the technique of [32] for proving the result on the convergence of stationary distributions, with an additional innovation of approximating the stationary random vector with a "close-by" random vector.

We organize the paper as follows: In Section 2 we specify our model and notations. In Section 3 we state and prove our main results. While the proofs are outlined in the main text, the proofs for most lemmas are in Appendix A. In Section 4 we provide simulation results. We conclude in Section 6. Appendix B contains some additional simulation results.

2 MODEL

We consider following standard queuing system model: jobs arrives according to a Poisson process with rate λn where $\lambda < 1$ is a constant. There are n servers/queues in the system. At the time of an arrival, the dispatcher samples the queues of a subset of servers and dispatches the arrival to one of the sampled queues to wait for service. The service time for each job at each server is *i.i.d.* exponentially distributed with mean 1.

The NBRWR-Pod algorithm is defined as follows: At each time of the j -th job arrival, allocate the job to the least loaded server among $W_1\{j\}, \dots, W_d\{j\}$, which are the d sampled servers at this time. Ties are broken arbitrarily.

Assuming that the servers are interconnected by a k -regular graph $G = ([n], \mathcal{E})$, where $k \geq 3$ is a constant, the discrete time processes $(W_l\{j\})_{l=1}^d$ are defined as following. Starting from independent uniform random vertices, d random walkers perform independent non-backtracking random walks on G . After every $\lfloor c \log n \rfloor$ steps, the d random walkers are reset to independent uniform vertices on the graph, and thereafter, perform independent non-backtracking random walks as before. The parameter c is assumed to be a positive constant. The resets are introduced to ensure that the random walkers do not stay in a small subset of vertices for too long, hence ensuring that the servers are efficiently explored.

In order for this dispatching scheme to perform well, we need the following assumption on the graph G :

Assumption 1 (G is of high girth). The girth of the graph G is greater than or equal to $2\lceil \alpha \log_{k-1} n \rceil + 1$ for some constant $\alpha > 0$.

By using non-backtracking random walks on high-girth graphs instead of simple random walks or non-backtracking random walks on small-girth graphs, it is ensured that the random walkers are not likely to visit a recently visited vertex. In other words, the random walkers tend to visit vertices that haven't been visited in a long time, which tend to have relatively small queue length.

Remark 1. Graphs that satisfies Assumption 1 do exist. For example, Lubotzky, Phillips, and Sarnak's Ramanujan Graphs [18] form a sequence of regular graphs that satisfies Assumption 1 (with any $\alpha < \frac{2}{3}$) eventually as the number of vertices goes to infinity.

Remark 2. The graph G can be connected or disconnected, as long as every connected component has girth larger than or equal to $2\lceil \alpha \log_{k-1} n \rceil + 1$. For disconnected graphs, periodic resets allows the dispatcher to sample servers from the entire graph efficiently.

Remark 3. The results in the paper will also hold for 2-regular graphs (disjoint union of cycles) where each connected component contains strictly more than $\lfloor c \log n \rfloor$ vertices.

2.1 Notations

We list the notations that we will use in this paper in Table 1. To simplify notations, we will drop the superscript (n) in the proofs when considering a fixed n .

We will use four different brackets to indicate different time index systems: (t) stands for the original, continuous time; $[j]$ stands for the time of the j -th arrival or potential service (i.e. j -th "jump"); $\{j\}$ stands for the time of the j -th arrival; $\langle j \rangle$ stands for the time of the j -th restart (i.e. $j\lfloor c \log n \rfloor$ -th arrival). In the proof of stability of queuing systems, $\langle j \rangle$ stand for a system of time indices downsampled from the restart times, see Section 3.2 for detail.

We will use " $\stackrel{d}{\sim}$ " to indicate "equal in distribution." We use " $\Rightarrow$ " to denote weak convergence.

n	Number of servers, also the scaling parameter for the system
G	A k -regular graph of n vertices.
$Q_i^{(n)}(t)$	Queue length of server i at (continuous) time t .
$Q_i^{(n)}[j]$	Queue length of server i at the j -th event (i.e. arrival/potential service)
$Q_i^{(n)}\{j\}$	Queue length of server i as seen by the j -th arrival job.
$W_l\{j\}$	Position of random walker l at the arrival times of the j -th job.
$T_j^{(n)}$	Arrival time of j -th arrival job.
$\tau_j^{(n)}$	Inter-arrival time between $(j-1)$ -th and j -th job.
$X_i^{(n)}(t)$	Proportion of bins with load at least i at time t .
$\mathcal{F}_t$	A filtration that the random walk and queuing process are adapted to.

Table 1. Notations that will be used in this paper

2.2 Preliminaries: Large Deviation Inequalities

We will use a version of Azuma-Hoeffding Inequality for martingales, which we shall cite here for completeness.

LEMMA 1 (AZUMA-HOEFFDING). *Let $(X_j)_{j=0}^\infty$ be a martingale with respect to $(\mathcal{F}_j)_{j=1}^\infty$. Let C be an event such that*

$$\forall 1 \leq j \leq N, \quad |X_j - X_{j-1}| \leq c_j$$

is true on C . Then we have

$$\Pr \left(\max_{0 \leq j \leq N} |X_j - X_0| \geq \lambda, C \right) \leq 2 \exp \left(- \frac{\lambda^2}{2 \sum_{i=1}^N c_i^2} \right)$$

PROOF. See Appendix A.1. □

3 MAIN RESULTS

3.1 Fluid Limit Approximation

In this section, we will show that the limiting dynamics of the n -server system under NBRWR-Pod algorithm up to any finite time $T > 0$ can be approximated by the solution to a differential equation. It turns out that the limiting differential equation is the same as that of *power-of- d choices*.

THEOREM 1. *Consider the dynamic system $\mathbf{x}(t) \in [0, 1]^{\mathbb{Z}_+}$ described by the following differential equations:*

$$\begin{aligned} \frac{dx_i}{dt} &= \lambda(x_{i-1}^d - x_i^d) - (x_i - x_{i+1}) & i \geq 1 \\ x_0(t) &\equiv 1 \end{aligned} \tag{1}$$

Let $\mathbf{X}^{(n)}(t) = (X_i^{(n)}(t))_{i \in \mathbb{Z}_+}$ be an infinite dimensional vector, where $X_i^{(n)}(t)$ is the proportion of queues with length exceeding (or equal to) i at time t . Suppose that

- (a) $\mathbf{X}^{(n)}(0)$ is deterministic
- (b) $\lim_{n \rightarrow \infty} \|\mathbf{X}^{(n)}(0) - \mathbf{x}(0)\|_1 = 0$.

(c) $\|\mathbf{x}(0)\|_1 < +\infty$.

then for every finite $T > 0$

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} \|\mathbf{X}^{(n)}(t) - \mathbf{x}(t)\|_1 = 0 \quad a.s.$$

The proof is inspired by the results on *power-of- d choices* algorithm [22][23][29] along with literature of fluid limit approximation [10][8][19]. However, since the process $\mathbf{X}(t)$ is not a Markov process, Kurtz's theorem for density dependent Markov Chains [10][9] does not apply here. The dependence between random walker positions and queue length vector makes the proof non-trivial. The proof is based on large deviation estimates of martingales along with Gronwall's lemma. Methodology of our proof is closely related to that of [8] and [19] for general Markov processes, our proof is independent of the proof in [8] and [19], and in addition, is more intuitive to understand in a queuing system setting.

PROOF OF THEOREM 1. Let $\mathbf{a}, \mathbf{b} : [0, 1]^{\mathbb{Z}_+} \mapsto [0, 1]^{\mathbb{Z}_+}$ be defined by

$$a_i(\mathbf{x}) = \begin{cases} \lambda(x_{i-1}^d - x_i^d) & i \geq 1 \\ 0 & i = 0 \end{cases}$$

$$b_i(\mathbf{x}) = \begin{cases} x_i - x_{i+1} & i \geq 1 \\ 0 & i = 0 \end{cases}$$

Both $\mathbf{a}$ and $\mathbf{b}$ are Lipschitz continuous operators (under ℓ_1 norm), where $\mathbf{a}$ has Lipschitz constant $2\lambda d$ and $\mathbf{b}$ has Lipschitz constant 2.

For $i \geq 1$, define $A_i(t)$ to be the number of job arrivals that are allocated to a server with load $i - 1$ (just before this arrival) before (including) time t . Define $A_0(t) \equiv 0$. For all $i \geq 0$, also define $B_i(t)$ to be the number of departures from queues with load i (just before this departure) before (including) time t . Let $\mathbf{A}(t), \mathbf{B}(t)$ denote the corresponding infinite dimensional vectors. We have the relation

$$\mathbf{X}(t) = \mathbf{X}(0) + \frac{\mathbf{A}(t)}{n} - \frac{\mathbf{B}(t)}{n}$$

Now, define

$$\mathbf{M}(t) = \mathbf{X}(t) - \mathbf{X}(0) - \int_0^t [\mathbf{a}(\mathbf{X}(u)) - \mathbf{b}(\mathbf{X}(u))] du$$

The idea of the proof is to bound $\|\mathbf{M}(t)\|_1$ and then apply Gronwall's lemma to bound $\|\mathbf{X}(t) - \mathbf{x}(t)\|_1$. We write

$$\begin{aligned} \mathbf{M}(t) &= \frac{\mathbf{A}(t)}{n} - \frac{\mathbf{B}(t)}{n} - \int_0^t [\mathbf{a}(\mathbf{X}(u)) - \mathbf{b}(\mathbf{X}(u))] du \\ &= \left[\frac{\mathbf{A}(t)}{n} - \int_0^t \mathbf{a}(\mathbf{X}(u)) du \right] - \left[\frac{\mathbf{B}(t)}{n} - \int_0^t \mathbf{b}(\mathbf{X}(u)) du \right] \\ &=: \mathbf{M}^a(t) - \mathbf{M}^b(t) \end{aligned}$$

where $\mathbf{M}^a(t)$ is the "arrival part", and $\mathbf{M}^b(t)$ is the "service" part. We will bound $\|\mathbf{M}^a(t)\|_1$ and $\|\mathbf{M}^b(t)\|_1$ separately.

The idea to bound $\|\mathbf{M}^a(t)\|_1$ is as following: For every i , a downsampled version of $M_i^a(t)$ can be approximated by a martingale process. The martingale process can be shown to be small by martingale concentration inequalities. For large i , $M_i^a(t)$ is small since there are not too many queues with length greater than i . Combining the two estimates, we can obtain concentration results on $\|\mathbf{M}^a(t)\|_1$.

Now we define an auxiliary queuing process $\tilde{Q}(t)$, which is a process coupled with $Q(t)$.

Recall that an alternative discription of $Q(t)$ is as follows: $Q(t)$ is obtained from a discrete-time process $Q[j]$ with holding times, where the holding times are *i.i.d.* exponentially distributed with mean $\frac{1}{(\lambda+1)n} \cdot Q[j]$ that can be described as follows

$$Q[j+1] = (Q[j] + R[j](1 - \Lambda[j]) - S[j]\Lambda[j])_+$$

where $\Lambda[j]$, $j = 0, 1, 2, \dots$ are *i.i.d.* Bernoulli random variables with mean $\frac{1}{\lambda+1}$. $S[j]$ are *i.i.d.* uniformly distributed on $\{e_i\}_{i=1}^n$ which indicates potential services. $R[j] \in \{e_i\}_{i=1}^n$ are potential arrivals to the system. $\{\Lambda[j]\}_{j=0}^\infty$ and $\{S[j]\}_{j=0}^\infty$ are mutually independent, and are also independent of $\{Q[j']\}_{j'=0}^j$.

In the discrete process, let $\tilde{T}[s]$ be the arrival timestamp of the s -th job (i.e. the s -th timestamp such that $\Lambda[j] = 0$). Set $\tilde{T}[0] = 0$. Define

$$\tilde{Q}[j+1] = \begin{cases} Q[j] & j+1 = \tilde{T}[s \lfloor c \log n \rfloor] \text{ for some } s \in \mathbb{Z}_+ \\ (\tilde{Q}[j] - S[j]\Lambda[j])_+ & \text{otherwise} \end{cases}$$

Now, let $\tilde{Q}(t)$ be obtained from holding the discrete-time process $\tilde{Q}[j]$ using the same holding times for $Q(t)$.

In words, $\tilde{Q}(t)$ is a coupled system such that it copies $Q(t)$ at every time that the random walkers reset. Between random walker resets, $\tilde{Q}(t)$ takes the same service schedule as $Q(t)$ but does not accept any arrivals. Notice that

$$\tilde{Q}_i(t) \leq Q_i(t) \quad \forall i \in [n] \quad \forall t \geq 0 \quad a.s.$$

and

$$\sum_{i=1}^n (Q_i(t) - \tilde{Q}_i(t)) \leq \lfloor c \log n \rfloor \quad a.s.$$

since $\sum_{i=1}^n (Q_i(t) - \tilde{Q}_i(t))$ is bounded by the number of arrivals by time t since last reset, which is at most $\lfloor c \log n \rfloor$.

Further, we define $\tilde{X}(t)$ to be the vector of proportions of bins corresponding to $\tilde{Q}(t)$, i.e.

$$\tilde{X}_i(t) = \frac{1}{n} \sum_{l=1}^n \mathbb{1}_{\{\tilde{Q}_l(t) \geq i\}} \quad i \in \mathbb{Z}_+$$

We then have

$$\begin{aligned} & \|X(t) - \tilde{X}(t)\|_1 \\ & \leq \frac{1}{n} \sum_{l=1}^n |Q_l(t) - \tilde{Q}_l(t)| \\ & = \frac{1}{n} \sum_{l=1}^n (Q_l(t) - \tilde{Q}_l(t)) \leq \frac{\lfloor c \log n \rfloor}{n} \quad \forall t \geq 0 \end{aligned} \tag{2}$$

We further define $\bar{X}(t)$ as follows:

$$\bar{X}(t) = \tilde{X}(T_j) \quad T_j \leq t < T_{j+1}$$

where T_j is the time of the j -th arrival. In other words, $\bar{X}(t)$ is obtained from sampling and holding $\tilde{X}(t)$ at arrival events. Effectively, $\bar{X}(t)$ is a process which accumulate services at arrival times.

Now, we write

$$\begin{aligned} \mathbf{M}^a(t) &= \left[\frac{\mathbf{A}(t)}{n} - \int_0^t \mathbf{a}(\bar{\mathbf{X}}(u)) du \right] + \int_0^t [\mathbf{a}(\bar{\mathbf{X}}(u)) - \mathbf{a}(\tilde{\mathbf{X}}(u))] du \\ &\quad + \int_0^t [\mathbf{a}(\tilde{\mathbf{X}}(u)) - \mathbf{a}(\mathbf{X}(u))] du \\ &=: \mathbf{M}^{a,1}(t) + \mathbf{M}^{a,2}(t) + \mathbf{M}^{a,3}(t) \end{aligned}$$

Using (2) and the fact that $\mathbf{a}$ is $2\lambda d$ -Lipschitz continuous, we can immediately bound

$$\|\mathbf{M}^{a,3}(t)\|_1 \leq \int_0^t 2\lambda d \frac{\lfloor c \log n \rfloor}{n} du = \frac{2\lambda d \lfloor c \log n \rfloor}{n} t$$

Hence

$$\sup_{t \in [0, T]} \|\mathbf{M}^{a,3}(t)\|_1 \leq \frac{2\lambda d T \lfloor c \log n \rfloor}{n} \text{ a.s.}$$

Now we try to bound $\|\mathbf{M}^{a,2}(t)\|_1$ and $\|\mathbf{M}^{a,1}(t)\|_1$. We now introduce two high probability events

$$\begin{aligned} \mathcal{A} &:= \{\text{there are strictly less than } e\lambda T n \text{ job arrivals before } T\} \\ \mathcal{B} &:= \mathcal{A} \cap \{\text{no queue has accepted more than } \kappa L \log n \\ &\quad \text{arrival jobs within the first } \lfloor e\lambda T n \rfloor \text{ arrivals}\} \end{aligned}$$

where $L := d \left(\frac{c \log(k-1)}{2\alpha} + 1 \right)$ and $\kappa = 3 \left[\log(1 + \frac{1}{e\lambda d T}) \right]^{-1}$.

LEMMA 2. *We have*

$$\Pr(\mathcal{A}^c) \leq e^{-\lambda n T}, \quad \Pr(\mathcal{B}^c) \leq e^{-\lambda n T} + n^{-2}$$

PROOF. See Appendix A.2 □

LEMMA 3. *For any $\varepsilon_2 > 0$,*

$$\Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^{a,2}(t)\|_1 \geq 2\lambda d T \varepsilon_2, \mathcal{A} \right) \leq e(1 + \lambda) \lambda T n \left(\frac{1}{1 + \lambda} \right)^{n\varepsilon_2}$$

PROOF. See Appendix A.3. □

The rest of the proof is to bound $\|\mathbf{M}^{a,1}(t)\|_1$. To achieve this goal, we first sample the continuous-time process $\mathbf{M}^{a,1}(t)$ at times $T_{j \lfloor c \log n \rfloor}$, the time of the $j \lfloor c \log n \rfloor$ -th job arrival. Denote

$$\mathbf{M}^{a,1}\langle j \rangle := \mathbf{M}^{a,1}(T_{j \lfloor c \log n \rfloor} -)$$

Let $\mathcal{G}_j := \mathcal{F}_{T_{j \lfloor c \log n \rfloor}} \cdot \{\mathcal{G}_j\}_{j=0}^\infty$ forms a filtration. The key lemma for the proof is stated as follows:

LEMMA 4. *For all $i \geq 1$, we have*

$$\begin{aligned} &|E[M_i^{a,1}\langle j+1 \rangle \mid \mathcal{G}_j] - M_i^{a,1}\langle j \rangle| \\ &\leq \frac{d \lfloor c \log n \rfloor^2}{2n} \left(\frac{1}{n^\alpha} + \frac{d-1}{n} \right) \quad \text{a.s.} \end{aligned}$$

PROOF. See Appendix A.4. □

Lemma 4 basically stated that $\mathbf{M}^{a,1}\langle j \rangle$ is “approximately” a martingale, which allows us to apply martingale concentration inequality to bound each coordinate of $\mathbf{M}^{a,1}$. Note that [30] uses the Lipschitz property of one-step drifts but here we consider a $\lfloor c \log n \rfloor$ -step drift.

COROLLARY 1. *Let*

$$N = \frac{e\lambda Tn}{\lfloor c \log n \rfloor}, \quad \delta = \frac{e\lambda dT \lfloor c \log n \rfloor}{2} \left(\frac{1}{n^\alpha} + \frac{1}{n} \right).$$

For all $i \geq 1$, for any $\varepsilon_1 > 0$ and $r > 1$,

$$\Pr \left(\sup_{t \in [0, T]} |M_i^{a,1}(t)| \geq \varepsilon_1 + K_r + \delta, \mathcal{A} \right) \leq \exp \left(-\frac{\varepsilon_1^2}{NK_r^2} \right) + eNn^{-r}$$

where $K_r := \frac{\lfloor c \log n \rfloor}{n} + \frac{r \log n (\lfloor c \log n \rfloor + 1)}{\lambda n}$.

PROOF. See Appendix A.5. □

However, since the vector $\mathbf{M}^{a,1}(t)$ is infinitely dimensional, a bound on each coordinate of $\mathbf{M}^{a,1}(t)$ does not naturally provide a bound on $\|\mathbf{M}^{a,1}(t)\|_1$. Hence we need the following lemma to bound the tail of $\|\mathbf{M}^{a,1}(t)\|_1$.

LEMMA 5. *Let $\rho > 0$ be any constant. Set $b := \lceil (\rho + \kappa L) \log n + L \rceil$ and $\varphi = 2 \lceil N \rceil \lfloor c \log n \rfloor \left(\frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n} \right)^d$.*

For sufficiently large n ,

$$\begin{aligned} & \Pr \left(\sup_{t \in [0, T]} \sum_{i=b+1}^{\infty} |M_i^{a,1}(t)| \geq \frac{\varphi}{n} + \lambda T \left(\frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right)^d, \mathcal{B} \right) \\ & \leq \exp \left(-\frac{3\varphi}{16 \lfloor c \log n \rfloor} \right) \end{aligned}$$

PROOF. See Appendix A.6. □

COROLLARY 2. *For any $\varepsilon_1 > 0$, let $\tilde{\varepsilon}_0 = b(\varepsilon_1 + K_r + \delta) + \frac{\varphi}{n} + \lambda T \left(\frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right)^d$, we have*

$$\begin{aligned} & \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^{a,1}(t)\|_1 \geq \tilde{\varepsilon}_0, \mathcal{B} \right) \\ & \leq b \left[2 \exp \left(-\frac{\varepsilon_1^2}{2NK_r^2} \right) + eNn^{-r} \right] + \exp \left(-\frac{3\varphi}{16 \lfloor c \log n \rfloor} \right) \end{aligned}$$

PROOF. By Union Bound, we have

$$\begin{aligned} & \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^{a,1}(t)\|_1 \geq \tilde{\varepsilon}_0, \mathcal{B} \right) \\ & \leq \Pr \left(\sup_{t \in [0, T]} \sum_{i=1}^b |M_i^{a,1}(t)| \geq b(\varepsilon_1 + K_r + \delta), \mathcal{B} \right) \\ & \quad + \Pr \left(\sup_{t \in [0, T]} \sum_{i=b+1}^{\infty} |M_i^{a,1}(t)| \geq \frac{\varphi}{n} + \lambda T \left(\frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right)^d, \mathcal{B} \right) \\ & \leq \sum_{i=1}^b \Pr \left(\sup_{t \in [0, T]} |M_i^{a,1}(t)| \geq \varepsilon_1 + K_r + \delta, \mathcal{A} \right) \\ & \quad + \exp \left(-\frac{3\varphi}{16 \lfloor c \log n \rfloor} \right) \end{aligned}$$

$$\leq b \left[2 \exp \left(-\frac{\varepsilon_1^2}{2NK_r^2} \right) + eNn^{-r} \right] + \exp \left(-\frac{3\varphi}{16\lfloor c \log n \rfloor} \right)$$

□

Now we have provided bounds for $\|\mathbf{M}^{a,1}(t)\|_1$, $\|\mathbf{M}^{a,2}(t)\|_1$, $\|\mathbf{M}^{a,3}(t)\|_1$. Combine all the above, we obtain that

$$\begin{aligned} & \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^a(t)\|_1 \geq \tilde{\varepsilon}_0 + 2\lambda dT\varepsilon_2 + \frac{2\lambda dT \lfloor c \log n \rfloor}{n}, \mathcal{B} \right) \\ & \leq \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^{a,1}(t)\|_1 \geq \tilde{\varepsilon}_0, \mathcal{B} \right) \\ & \quad + \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^{a,2}(t)\|_1 \geq 2\lambda dT\varepsilon_2, \mathcal{A} \right) \\ & \leq b \left[2 \exp \left(-\frac{\varepsilon_1^2}{2NK_r^2} \right) + eNn^{-r} \right] + \exp \left(-\frac{3\varphi}{16\lfloor c \log n \rfloor} \right) \\ & \quad + e(1+\lambda)\lambda Tn \left(\frac{1}{1+\lambda} \right)^{n\varepsilon_2} \end{aligned} \tag{3}$$

Now we need to bound $\|\mathbf{M}^b(t)\|_1$. Different from $\mathbf{M}^a(t)$, $\mathbf{M}^b(t)$ can be bounded using the methodology used to prove Kurtz's Theorem [9][10].

LEMMA 6.

$$\begin{aligned} & \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^b(t)\|_1 \geq b\varepsilon_3 + \frac{(e+1)T\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L}, \mathcal{B} \right) \\ & \leq 2be^{-nTh(\varepsilon_3/T)} + \exp \left(-nT \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right) \end{aligned} \tag{4}$$

PROOF. See Appendix A.7.

□

Combining (3) and (4), denote

$$\begin{aligned} \varepsilon_0 &:= b(\varepsilon_1 + K_r + \delta) + \frac{\varphi}{n} + \lambda T \left(\frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right)^d + 2\lambda dT\varepsilon_2 \\ & \quad + \frac{2\lambda dT \lfloor c \log n \rfloor}{n} + b\varepsilon_3 + \frac{(e+1)T\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \end{aligned}$$

we have

$$\begin{aligned} & \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}(t)\|_1 \geq \varepsilon_0 \right) \\ & \leq \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^a(t)\|_1 \geq \tilde{\varepsilon}_0 + 2\lambda dT\varepsilon_2 + \frac{2\lambda dT \lfloor c \log n \rfloor}{n}, \mathcal{B} \right) \\ & \quad + \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^b(t)\|_1 \geq b\varepsilon_3 + \frac{(e+1)T\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L}, \mathcal{B} \right) + \Pr(\mathcal{B}^c) \\ & \leq b \left[2 \exp \left(-\frac{\varepsilon_1^2}{2NK_r^2} \right) + eNn^{-r} \right] + \exp \left(-\frac{3\varphi}{16\lfloor c \log n \rfloor} \right) \end{aligned}$$

$$\begin{aligned}
& + e(1+\lambda)\lambda T n \left(\frac{1}{1+\lambda} \right)^{n\varepsilon_2} + 2be^{-nTh(\varepsilon_3/T)} \\
& + \exp \left(-nT \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right) + e^{-\lambda nT} + n^{-2} \\
& =: p_0
\end{aligned}$$

Recall that $N = \Theta(\frac{n}{\log n})$, $K_r = \Theta(\frac{r(\log n)^2}{n})$, $\delta = \Theta(\frac{\log n}{n^{\alpha \wedge 1}})$, $b = \Theta(\log n)$, $\varphi = \Theta(\frac{n}{(\log n)^d})$.

Select $r = 3$, $\varepsilon_1 = 2\sqrt{N}K_r \log n = \Theta\left(\frac{(\log n)^{5/2}}{n^{1/2}}\right)$,

$\varepsilon_2 = \frac{3}{\log(1+\lambda)} \frac{\log n}{n^{\alpha \wedge 1}}$, $\varepsilon_3 = 2\sqrt{\frac{T \log n}{n}}$, using the fact that

$$h(t) = (1+t)\log(1+t) - t = \frac{t^2}{2} + o(t^2)$$

we arrive at

$$\varepsilon_0 = o(1), p_0 = O\left(\frac{\log n}{n^2}\right)$$

The rest of the proof is the application of Gronwall's lemma:

LEMMA 7 (GRONWALL'S LEMMA). [9][10] Let g be a bounded real-valued function satisfying

$$g(t) \leq C + D \int_0^t g(u) du \quad \forall t \geq 0$$

where C, D are positive constants. Then

$$g(t) \leq Ce^{Dt} \quad \forall t \geq 0$$

PROOF. See [9] or [10]. □

Now, observe that

$$\begin{aligned}
& \mathbf{X}^{(n)}(t) - \mathbf{x}(t) \\
& = \mathbf{X}^{(n)}(0) + \mathbf{M}(t) + \int_0^t [\mathbf{a}(\mathbf{X}^{(n)}(u)) - \mathbf{b}(\mathbf{X}^{(n)}(u))] du \\
& \quad - \mathbf{x}(0) - \int_0^t [\mathbf{a}(\mathbf{x}(u)) - \mathbf{b}(\mathbf{x}(u))] du \\
& = [\mathbf{X}^{(n)}(0) - \mathbf{x}(0)] + \mathbf{M}(t) + \int_0^t [\mathbf{a}(\mathbf{X}^{(n)}(u)) - \mathbf{a}(\mathbf{x}(u))] du \\
& \quad + \int_0^t [\mathbf{b}(\mathbf{X}^{(n)}(u)) - \mathbf{b}(\mathbf{x}(u))] du
\end{aligned}$$

Using the fact that $\mathbf{a}$ is $2\lambda d$ -Lipschitz continuous and $\mathbf{b}$ is 2-Lipschitz continuous, we have

$$\begin{aligned}
& \|\mathbf{X}^{(n)}(t) - \mathbf{x}(t)\|_1 \\
& \leq \|\mathbf{X}^{(n)}(0) - \mathbf{x}(0)\|_1 + \sup_{t \in [0, T]} \|\mathbf{M}(t)\|_1 \\
& \quad + (2\lambda d + 2) \int_0^t \|\mathbf{X}^{(n)}(u) - \mathbf{x}(u)\|_1 du
\end{aligned}$$

For any fixed $\varepsilon > 0$, there exist $\bar{n} \in \mathbb{N}$ such that

$$\begin{cases} \|\mathbf{X}^{(n)}(0) - \mathbf{x}(0)\|_1 \leq \frac{\varepsilon e^{-(2\lambda d+2)T}}{2} \\ \varepsilon_0 \leq \frac{\varepsilon e^{-(2\lambda d+2)T}}{2} \end{cases} \quad \forall n \geq \bar{n}$$

For all $n \geq \bar{n}$, by Gronwall's lemma,

$$\begin{aligned} \sup_{t \in [0, T]} \|\mathbf{M}(t)\|_1 &\leq \varepsilon_0 \\ \Rightarrow \|\bar{\mathbf{X}}^{(n)}(t) - \mathbf{x}(t)\|_1 &\leq \varepsilon e^{-(2\lambda d+2)T} \\ &\quad + (2\lambda d + 2) \int_0^t \|\bar{\mathbf{X}}^{(n)}(u) - \mathbf{x}(u)\|_1 du \\ &\stackrel{(G)}{\Rightarrow} \|\bar{\mathbf{X}}^{(n)}(t) - \mathbf{x}(t)\|_1 \leq \varepsilon \quad \forall t \in [0, T] \end{aligned}$$

Hence for sufficiently large n ,

$$\begin{aligned} \Pr \left(\sup_{t \in [0, T]} \|\bar{\mathbf{X}}^{(n)}(t) - \mathbf{x}(t)\|_1 > \varepsilon \right) &\leq \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}(t)\|_1 \geq \varepsilon_0 \right) \\ &\leq p_0 = O \left(\frac{\log n}{n^2} \right) \end{aligned}$$

Almost sure convergence then follows from the Borel-Cantelli Lemma. $\square$

If the system is initialized to be empty, using the proof of the main theorem we can also obtain an alternative version of the fluid limit result which has a better deviation estimate.

PROPOSITION 1. *Let $\mathbf{X}^{(n)}(t)$ and $\mathbf{x}(t)$ be as defined in Theorem 1. For an initially empty system, there exist constants $\kappa_2, \kappa_3 > 0$ such that for sufficiently large n ,*

$$\Pr \left(\sup_{t \in [0, T]} \|\mathbf{X}^{(n)}(t) - \mathbf{x}(t)\|_1 \geq \kappa_2 \frac{(\log n)^{7/2}}{n^{\alpha \wedge (1/2)}} \right) \leq \kappa_3 \frac{\log n}{n^2}$$

PROOF. Similar to the proof of Theorem 1. Omitted. $\square$

3.2 Stability of the Queuing System

We first need to show that the proposed dispatching scheme actually stabilizes the system for every value of n . More specifically, as the time goes to infinity, the queue length vector $\mathbf{Q}(t)$ will converge in distribution to some random vector $\hat{\mathbf{Q}}$.

THEOREM 2. *For any initial queue length $\mathbf{Q}(0)$, $\mathbf{Q}(t) \xrightarrow{t \rightarrow \infty} \hat{\mathbf{Q}}$.*

Unlike *power-of- d* choices and many other related schemes, the proof for this part is not trivial since the process $\mathbf{Q}(t)$ is not a Markov process on its own, so we cannot directly apply the Foster-Lyapunov Theorem. Our approach for this section is to pass to a subprocess and prove that the subprocess is stable, and to then use the stability of the subprocess to argue for the stability of $\mathbf{Q}(t)$. The stability of the subprocess is obtained by comparing the NBRWR-Pod algorithm ($d \geq 2$) with the NBRWR-Po1 algorithm.

Recall that, from Section 3.1, the process $\mathbf{Q}[j]$ can be described as follows

$$\mathbf{Q}[j+1] = (\mathbf{Q}[j] + \mathbf{R}[j](1 - \Lambda[j]) - \mathbf{S}[j]\Lambda[j])_+$$

where $\mathbf{R}[j], \Lambda[j], \mathbf{S}[j]$ are defined in Section 3.1.

Consider non-backtracking random walks as a simple random walk on oriented edges of graph G . Let $\bar{W}_l[j]$ be the l -th random walker's position (oriented edge). Let $\Xi[j] \in \{0, 1, \dots, \lfloor c \log n \rfloor - 1\}$ be a circular counter variable which count the number of arrivals already added to the system since last reset. We can now have a description for $\bar{W}_l[j]$ and $R[j]$ as follows:

- If $\Xi[j] \neq 0$, $R[j] = e_i$ where i is the the least loaded queue among $W_l[j]$ (which is the destination vertex of $\bar{W}_l[j]$), $l = 1, \dots, d$. At the event of an arrival, $\bar{W}[j+1]$ is obtained from $\bar{W}[j]$ by taking one step of random walk independently and $\Xi[j+1] = (\Xi[j] + 1) \bmod \lfloor c \log n \rfloor$.
- If $\Xi[j] = 0$, $R[j] = e_{i^*}$ where i^* is the the least loaded queue among d independent uniform random queues at the sampled vertices. At the event of an arrival, $\bar{W}[j+1]$ will be reset to random oriented edges starting from these uniform random queues/vertices and $\Xi[j+1] = (\Xi[j] + 1) \bmod \lfloor c \log n \rfloor$.

Now, we have $(\Omega[j])_{j=0}^\infty = (Q[j], \bar{W}[j], \Xi[j])_{j=0}^\infty$ to be a countable state space Markov Chain. The state space of the chain is $\mathbb{Z}_+^n \times \tilde{\mathcal{E}}^d \times \{0, 1, \dots, \lfloor c \log n \rfloor - 1\}$. It can be verified that this chain is irreducible.

Now, sampling $(Q[j], \bar{W}[j], \Xi[j])$ at arrival times (i.e. times where $\Lambda[j] = 0$), we can further obtain a discrete-time process $(Q\{j\}, \bar{W}\{j\}, \Xi\{j\})$ where $Q\{j\}$ that can be described as follows:

$$Q\{j+1\} = (Q\{j\} + R\{j\} - S\{j\})_+$$

where $R\{j\} \in \{e_i\}_{i=1}^n$ are arrivals to the system, and $S_i\{j\}, i \in [n], j \geq 0$ are i.i.d. Geometric random variables with mean $\frac{1}{\lambda n}$ which indicates potential services. Let $W_l\{j\}, l = 1, 2, \dots, d$ be the position of random walkers at the time of the j -th arrival.

The key lemma we take advantage of is the following: Under certain condition, positive recurrence in a down-sampled process can imply positive recurrence in original process.

LEMMA 8. *Let $(X_j)_{j=0}^\infty$ be a discrete-time Markov Chain adapted to filtration $(\mathcal{F}_j)_{j=0}^\infty$. Let $T_1 < T_2 < T_3 < \dots$ be stopping times w.r.t. $(\mathcal{F}_j)_{j=0}^\infty$. We further assume that*

- $T_j - T_{j-1}$ is independent of $\mathcal{F}_{T_{j-1}}$.
- $E[T_j - T_{j-1}] < C < +\infty$ for some constant $C > 0$ for all j .

Then, if x is a positive recurrent state for the Markov Chain $(X_{T_j})_{j=0}^\infty$, so is x for $(X_j)_{j=0}^\infty$.

PROOF. Use Wald's Identity to establish that the expected return time of state x for Markov chain $(X_j)_{j=0}^\infty$ is finite. $\square$

The crucial observation is as follows:

OBSERVATION 1. $(Q\{j\lfloor c \log n \rfloor\})_{j=0}^\infty$ is a countable state Markov Chain.

LEMMA 9. $(Q\{j\lfloor c \log n \rfloor\})_{j=0}^\infty$ is irreducible and positive recurrent.

PROOF. It is clear that this Markov Chain is irreducible:

- For any $\mathbf{q} \in \mathbb{Z}_+^n$ and any $i \in [n]$, there is positive probability to transit to $\mathbf{q} + \mathbf{e}^{(i)}$ in one step: There is positive probability that all d random walkers restart at vertex i , and all $\lfloor c \log n \rfloor$ allocated jobs except the first job are immediately served.
- For any $\mathbf{q} \in \mathbb{Z}_+^n$ and any $i \in [n]$ such that $q_i \geq 1$, there is positive probability to transit to $\mathbf{q} - \mathbf{e}^{(i)}$ in one step: There is positive probability that all allocated jobs are immediately served, and one extra service is applied on server i .

It remains to prove that the chain is positive recurrent. Define a quadratic Lyapunov function $V : \mathbb{Z}_+^n \mapsto \mathbb{R}_+$

$$V(\mathbf{q}) := \sum_{i=1}^n q_i^2$$

Now we compute the drift. For any $j' \geq 0$,

$$\begin{aligned} & V(\mathbf{Q}\{j' + 1\}) - V(\mathbf{Q}\{j'\}) \\ &= \sum_{i=1}^n [(Q_i\{j'\} + R_i\{j'\} - S_i\{j'\})_+^2 - (Q_i\{j'\})^2] \\ &\leq \sum_{i=1}^n [(Q_i\{j'\} + R_i\{j'\} - S_i\{j'\})^2 - (Q_i\{j'\})^2] \\ &= 2 \sum_{i=1}^n Q_i\{j'\}(R_i\{j'\} - S_i\{j'\}) + \sum_{i=1}^n (R_i\{j'\} - S_i\{j'\})^2 \\ &= -2 \sum_{i=1}^n Q_i\{j'\}S_i\{j'\} + \sum_{i=1}^n (S_i\{j'\})^2 + 2Q_{i^*}\{j'\} + 1 - 2S_{i^*}\{j'\} \\ &\leq -2 \sum_{i=1}^n Q_i\{j'\}S_i\{j'\} + \sum_{i=1}^n (S_i\{j'\})^2 + \frac{2}{d} \sum_{l=1}^d Q_{W_l\{j'\}}\{j'\} \\ &\quad + 1 - 2S_{i^*}\{j'\} \end{aligned}$$

where i^* is the queue that the $(j' + 1)$ -th job is dispatched to, i.e. $\mathbf{R}\{j'\} = e_{i^*}$. By the construction of the scheme we know that

$$Q_{i^*}\{j'\} = \min_{1 \leq l \leq d} Q_{W_l\{j'\}}\{j'\}$$

For any $0 \leq j' < \lfloor c \log n \rfloor$,

$$\begin{aligned} & \mathbb{E}[V(\mathbf{Q}\{j' + 1\}) - V(\mathbf{Q}\{j'\}) \mid \mathbf{Q}\{0\}] \\ &\leq -2\mathbb{E}\left[\sum_{i=1}^n Q_i\{j'\}S_i\{j'\} \mid \mathbf{Q}\{0\}\right] + \mathbb{E}\left[\sum_{i=1}^n (S_i\{j'\})^2 \mid \mathbf{Q}\{0\}\right] \\ &\quad + \frac{2}{d} \sum_{l=1}^d \mathbb{E}[Q_{W_l\{j'\}}\{j'\} \mid \mathbf{Q}\{0\}] + 1 - 2\mathbb{E}[S_{i^*}\{j'\} \mid \mathbf{Q}\{0\}] \\ &= -\frac{2}{\lambda n} \mathbb{E}\left[\sum_{i=1}^n Q_i\{j'\} \mid \mathbf{Q}\{0\}\right] + n \left[\frac{1}{\lambda n} + \frac{2}{(\lambda n)^2} \right] \\ &\quad + \frac{2}{d} \sum_{l=1}^d \mathbb{E}[Q_{W_l\{j'\}}\{j'\} \mid \mathbf{Q}\{0\}] + 1 - \frac{2}{\lambda n} \end{aligned} \tag{5}$$

Now, recall that any vertex cannot be visited by the random walkers by more than $L := d \left(\frac{c \log(k-1)}{2\alpha} + 1 \right)$ times between resets. Also observe that $\mathbf{Q}\{j\lfloor c \log n \rfloor\}$ is independent of $W_l\{j'\}$, and $W_l\{j'\}$ is uniformly distributed, we further have

$$\begin{aligned} & \mathbb{E}[Q_{W_l\{j'\}}\{j'\} \mid \mathbf{Q}\{0\}] \\ &\leq \mathbb{E}[Q_{W_l\{j'\}}\{0\} \mid \mathbf{Q}\{0\}] + L = \frac{1}{n} \sum_{i=1}^n Q_i\{0\} + L \end{aligned} \tag{6}$$

Using the dynamics of $Q_i\{j'\}$ and the fact that $(a)_+ \geq a$, we have

$$\sum_{i=1}^n Q_i\{j'\} \geq \sum_{i=1}^n Q_i\{0\} + j' - \sum_{s=0}^{j'-1} \sum_{i=1}^n S_i\{s\}$$

Hence

$$\begin{aligned} \mathbb{E} \left[\sum_{i=1}^n Q_i\{j'\} \mid \mathbf{Q}\{0\} \right] &\geq \sum_{i=1}^n Q_i\{0\} + j' - \mathbb{E} \left[\sum_{s=0}^{j'-1} \sum_{i=1}^n S_i\{s\} \mid \mathbf{Q}\{0\} \right] \\ &= \sum_{i=1}^n Q_i\{0\} + j' - j'n \cdot \frac{1}{\lambda n} = \sum_{i=1}^n Q_i\{0\} - \left(\frac{1}{\lambda} - 1 \right) j' \end{aligned} \quad (7)$$

Combining (5), (6), and (7) we have

$$\begin{aligned} &\mathbb{E}[V(\mathbf{Q}\{j' + 1\}) - V(\mathbf{Q}\{j'\}) \mid \mathbf{Q}\{0\}] \\ &\leq -\frac{2}{\lambda n} \left[\sum_{i=1}^n Q_i\{0\} - \left(\frac{1}{\lambda} - 1 \right) j' \right] + n \left[\frac{1}{\lambda n} + \frac{2}{(\lambda n)^2} \right] \\ &\quad + \frac{2}{n} \sum_{i=1}^n Q_i\{0\} + 2L + 1 - \frac{2}{\lambda n} \\ &= -2 \left(\frac{1}{\lambda} - 1 \right) \frac{1}{n} \sum_{i=1}^n Q_i\{0\} + \frac{2}{\lambda n} \left(\frac{1}{\lambda} - 1 \right) j' \\ &\quad + \frac{1}{\lambda} + \frac{2}{\lambda^2 n} + 2L + 1 - \frac{2}{\lambda n} \end{aligned} \quad (8)$$

Summing (8) for $0 \leq j' < \lfloor c \log n \rfloor$ we obtain

$$\begin{aligned} &\Delta V(\mathbf{q}) \\ &:= \mathbb{E}[V(\mathbf{Q}\{\lfloor c \log n \rfloor\}) - V(\mathbf{Q}\{0\}) \mid \mathbf{Q}\{0\} = \mathbf{q}] \\ &\leq -2 \left(\frac{1}{\lambda} - 1 \right) \frac{\lfloor c \log n \rfloor}{n} \sum_{i=1}^n q_i + \frac{1}{\lambda n} \left(\frac{1}{\lambda} - 1 \right) \lfloor c \log n \rfloor (\lfloor c \log n \rfloor - 1) \\ &\quad + \lfloor c \log n \rfloor \left(\frac{1}{\lambda} + \frac{2}{\lambda^2 n} + 2L + 1 - \frac{2}{\lambda n} \right) \end{aligned} \quad (9)$$

Hence, for all $\mathbf{q} \in \mathbb{Z}_+^n$ such that $\sum_{i=1}^n q_i$ is sufficiently large, we have $\Delta V(\mathbf{q}) \leq -1$. Therefore, by the Foster-Lyapunov Theorem, the chain $(\mathbf{Q}\{\lfloor c \log n \rfloor\})_{j=0}^\infty$ is positive recurrent. $\square$

A byproduct of the above proof is a Kingman-type bound for the stationary queue lengths. The next Corollary is established in preparation for the proof in Section 3.3.

COROLLARY 3. *Let $\check{\mathbf{Q}}^{(n)}$ be the stationary queue vector with respect to the Markov Chain $(\mathbf{Q}^{(n)}\{\lfloor c \log n \rfloor\})_{j=0}^\infty$. For any $\varepsilon > 0$, there exist $\bar{n}(\varepsilon) \in \mathbb{N}$ such that for all $n \geq \bar{n}(\varepsilon)$,*

$$\mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \check{Q}_i^{(n)} \right] \leq \frac{1 + \lambda + 2\lambda L}{2(1 - \lambda)} + \varepsilon$$

PROOF. Using (9) and $E[\Delta(\check{Q}^{(n)})] = 0$ we obtain

$$\begin{aligned} & -2 \left(\frac{1}{\lambda} - 1 \right) \lfloor c \log n \rfloor E \left[\frac{1}{n} \sum_{i=1}^n \check{Q}_i^{(n)} \right] \\ & + \frac{1}{\lambda n} \left(\frac{1}{\lambda} - 1 \right) \lfloor c \log n \rfloor (\lfloor c \log n \rfloor - 1) \\ & + \lfloor c \log n \rfloor \left(\frac{1}{\lambda} + \frac{2}{\lambda^2 n} + 2L + 1 - \frac{2}{\lambda n} \right) \geq 0 \end{aligned}$$

Hence

$$\begin{aligned} E \left[\frac{1}{n} \sum_{i=1}^n \check{Q}_i^{(n)} \right] & \leq \frac{\lfloor c \log n \rfloor - 1}{2\lambda n} + \frac{\frac{1}{\lambda} + \frac{2}{\lambda^2 n} + 2L + 1 - \frac{2}{\lambda n}}{2(\frac{1}{\lambda} - 1)} \\ & \xrightarrow{n \rightarrow \infty} \frac{1 + \lambda + 2\lambda L}{2(1 - \lambda)} \end{aligned}$$

proving the result. $\square$

COROLLARY 4. $(Q\{j\lfloor c \log n \rfloor\}, \bar{W}\{j\lfloor c \log n \rfloor\}, \Xi\{j\lfloor c \log n \rfloor\})_{j=0}^{\infty}$ is positive recurrent on $\mathbb{Z}_+^n \times \tilde{\mathcal{E}}^d \times \{0\}$.

PROOF. Consider any $\mathbf{q} \in \mathbb{Z}_+^n$. Sample the process

$(Q\{j\lfloor c \log n \rfloor\}, \bar{W}\{j\lfloor c \log n \rfloor\}, \Xi\{j\lfloor c \log n \rfloor\})_{j=0}^{\infty}$ at j 's such that

$Q\{j\lfloor c \log n \rfloor\} = \mathbf{q}$ to obtain a Markov Chain $(Q\langle j \rangle, \bar{W}\langle j \rangle, \Xi\langle j \rangle) = (\mathbf{q}, \bar{W}\langle j \rangle, 0)$.

For any $\bar{\mathbf{w}}, \bar{\mathbf{w}}' \in \tilde{\mathcal{E}}^d$, the transition probability from $\bar{\mathbf{w}}$ to $\bar{\mathbf{w}}'$ is strictly positive: At $(\mathbf{q}, \bar{\mathbf{w}}, 0)$, there is positive probability that the random walkers restarts and walks in the manner that $\bar{\mathbf{w}}'$ is reached after $\lfloor c \log n \rfloor - 1$ steps of random walk, and all newly added jobs are immediately served with no other jobs served, which implies that we end up at $(\mathbf{q}, \bar{\mathbf{w}}', 0)$ after $\lfloor c \log n \rfloor$ arrivals. Hence the chain is irreducible. (Actually, it can be shown that $\bar{W}\langle j \rangle$ are i.i.d.) As $\bar{W}\langle j \rangle$ is a finite state space Markov Chain, we conclude that $(Q\langle j \rangle, \bar{W}\langle j \rangle, \Xi\langle j \rangle)$ is positive recurrent.

Denote $\tau^q\langle s \rangle$ as the s -th smallest j such that $Q\{j\lfloor c \log n \rfloor\} = \mathbf{q}$. $\tau^q\langle s \rangle$ are increasing stopping times with $\tau^q\langle s \rangle - \tau^q\langle s-1 \rangle$ being i.i.d. Thanks to Lemma 9, which states that the chain $Q\{j\lfloor c \log n \rfloor\}$ is positive recurrent, we know that $\tau^q\langle s \rangle - \tau^q\langle s-1 \rangle$ has bounded mean. By Strong Markov Property we also know that $\tau^q\langle s \rangle - \tau^q\langle s-1 \rangle$ is independent of the process before time $\tau^q\langle s-1 \rangle$. Hence the conditions in Lemma 8 are satisfied, we conclude that the process $(Q\{j\lfloor c \log n \rfloor\}, \bar{W}\{j\lfloor c \log n \rfloor\}, \Xi\{j\lfloor c \log n \rfloor\})_{j=0}^{\infty}$ is positive recurrent at states $(\mathbf{q}, \bar{\mathbf{w}}, 0)$ for all $\bar{\mathbf{w}} \in \tilde{\mathcal{E}}^d$. Since $\mathbf{q}$ is arbitrary, we conclude that $(Q\{j\lfloor c \log n \rfloor\}, \bar{W}\{j\lfloor c \log n \rfloor\}, \Xi\{j\lfloor c \log n \rfloor\})_{j=0}^{\infty}$ is positive recurrent. $\square$

PROOF OF THEOREM 2.

$(Q\{j\lfloor c \log n \rfloor\}, \bar{W}\{j\lfloor c \log n \rfloor\}, \Xi\{j\lfloor c \log n \rfloor\})_{j=0}^{\infty}$ is a downsampled version of $(Q[j], \bar{W}[j], \Xi[j])_{j=0}^{\infty}$, where the sampling times are

$$\tau^r\{j\} = (\text{discrete}) \text{ time of the } j\lfloor c \log n \rfloor\text{-th arrival}$$

We know that $\tau^{\{r\}}\{j\} - \tau^{\{r\}}\{j-1\}$ are i.i.d. Pascal random variables (sum of $\lfloor c \log n \rfloor$ independent Geometric random variables with mean $\frac{1+\lambda}{\lambda}$), hence $E[\tau^r\{j\} - \tau^r\{j-1\}] = \frac{1+\lambda}{\lambda} \lfloor c \log n \rfloor < +\infty$. By Strong Markov Property, $\tau^r\{j\} - \tau^r\{j-1\}$ is independent of the process before $\tau^r\{j-1\}$. Hence Lemma 9 applies. Irreducibility along with Corollary 4 imply that $(Q[j], \bar{W}[j], \Xi[j])_{j=0}^{\infty}$ is positive recurrent.

Hence, through standard arguments relating a continuous time Markov Chain (CTMC) with its uniformized chain, the CTMC $(Q(t), \bar{W}(t), \Xi(t))_{t \geq 0}$ is positive recurrent. Hence $(Q(t), \bar{W}(t), \Xi(t))$ converges in distribution to its stationary distribution. In particular, we have $Q(t) \xrightarrow{t \rightarrow \infty} \hat{Q}$ for some random vector $\hat{Q}$.

□

Now, define the random vectors

$$\check{X}_i^{(n)} = \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{\check{Q}_j^{(n)} \geq i\}}, \quad \hat{X}_i^{(n)} = \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{\hat{Q}_j^{(n)} \geq i\}}$$

We will provide an auxiliary result, to show that $\check{X}$ and $\hat{X}$ are “close” to each other. The proposition is established for the proof in Section 3.3.

PROPOSITION 2. *Let $d_{W_1, \|\cdot\|_1}$ denote the first Wasserstein distance of two probability measures on metric space $([0, 1]^{\mathbb{Z}^+}, \|\cdot\|_1)$, we have*

$$\lim_{n \rightarrow \infty} d_{W_1, \|\cdot\|_1}(\mathcal{L}(\hat{X}^{(n)}), \mathcal{L}(\check{X}^{(n)})) = 0$$

PROOF. See Appendix A.8

□

3.3 Convergence of Stationary Distributions

In this section we will prove the following theorem which shows that the stationary solution of the differential equation is indeed a good approximation for the stationary distribution of queue lengths in a large-scale system.

THEOREM 3. $\lim_{n \rightarrow \infty} E[\|\hat{X}^{(n)} - \hat{x}\|_1] = 0$, where $\hat{x}$ is the unique fixed point of (1).

Our proof uses the method outlined by [32] along with the idea of approximating $\hat{X}^{(n)}$ with a “close by” random vector $\check{X}^{(n)}$ (thanks to Proposition 2 above). We start with establishing the uniform convergence of the tails of $\check{X}^{(n)}$. Then using this result, we show the tightness of the sequence $\hat{X}^{(n)}$. Finally, using an exchange of limits argument, we show the L_1 convergence of the sequence $\hat{X}^{(n)}$.

LEMMA 10. *For some $\bar{n} \in \mathbb{N}$,*

$$\lim_{b \rightarrow \infty} \sup_{n \geq \bar{n}} E \left[\sum_{i=b+1}^{\infty} \check{X}_i^{(n)} \right] = 0$$

PROOF. See Appendix A.9

□

COROLLARY 5. *Every subsequence of $(\hat{X}^{(n)})$ has a further subsequence that converges in $d_{W_1, \|\cdot\|_1}$.*

PROOF. Endow $[0, 1]^{\mathbb{Z}^+}$ with the metric

$$\rho(x, y) = \sup_{i \geq 0} \frac{|x_i - y_i|}{i + 1}$$

Under metric ρ , $[0, 1]^{\mathbb{Z}^+}$ is a compact separable metric space. Any collection of probability measures is trivially tight in $([0, 1]^{\mathbb{Z}^+}, \rho)$. By Prokhorov Theorem, every subsequence of $(\hat{X}^{(n)})_n$ has a further subsequence that converges weakly in $([0, 1]^{\mathbb{Z}^+}, \rho)$.

Now, let $(\hat{X}^{(n_k)})$ be a subsequence which converges weakly to some Y in $([0, 1]^{\mathbb{Z}^+}, \rho)$.

Observe that $\rho(x, y) \leq \|x - y\|_1$ for all $x, y \in [0, 1]^{\mathbb{Z}^+}$. Hence Proposition 2 implies that

$$\lim_{k \rightarrow \infty} d_{W_1, \rho}(\mathcal{L}(\hat{X}^{(n_k)}), \mathcal{L}(\check{X}^{(n_k)})) = 0$$

By Theorem 3.1 of [4], we conclude that $\check{X}^{(n_k)}$ converges weakly to Y in $([0, 1]^{\mathbb{Z}_+}, \rho)$. By Skorokhod's representation theorem, there exist a sequence of random vectors $(\check{X}^{(n_k)})$ such that $\check{X}^{(n_k)} \stackrel{d}{\sim} \check{X}^{(n_k)}$ with $(\check{X}^{(n_k)})$ converges in $([0, 1]^{\mathbb{Z}_+}, \rho)$ almost surely. Without loss of generality, assume that $(\check{X}^{(n_k)})$ converges almost surely to Y in $([0, 1]^{\mathbb{Z}_+}, \rho)$.

CLAIM 1. $\sum_{i=1}^{\infty} E[Y_i] < +\infty$

PROOF OF CLAIM. By Fatou's Lemma, we have

$$\sum_{i=1}^{\infty} Y_i = \sum_{i=1}^{\infty} \liminf_{k \rightarrow \infty} \check{X}_i^{(n_k)} \leq \liminf_{k \rightarrow \infty} \sum_{i=1}^{\infty} \check{X}_i^{(n_k)} \quad a.s. \quad (10)$$

Again, by Fatou's Lemma, we have

$$E \left[\liminf_{k \rightarrow \infty} \sum_{i=1}^{\infty} \check{X}_i^{(n_k)} \right] \leq \liminf_{k \rightarrow \infty} E \left[\sum_{i=1}^{\infty} \check{X}_i^{(n_k)} \right] \quad (11)$$

By Corollary 3 we know that

$$\begin{aligned} \liminf_{k \rightarrow \infty} E \left[\sum_{i=1}^{\infty} \check{X}_i^{(n_k)} \right] &= \liminf_{k \rightarrow \infty} E \left[\frac{1}{n_k} \sum_{i=1}^{n_k} \check{Q}_i^{(n_k)} \right] \\ &\leq \frac{1 + \lambda + 2\lambda L}{2(1 - \lambda)} \end{aligned} \quad (12)$$

Combining (10)(11)(12) we have

$$E \left[\sum_{i=1}^{\infty} Y_i \right] \leq \frac{1 + \lambda + 2\lambda L}{2(1 - \lambda)}$$

□

For any $\varepsilon > 0$, by Lemma 10 and Claim 1, there exist $b \in \mathbb{Z}_+$ such that

$$\sup_{k: n_k \geq \bar{n}} E \left[\sum_{i=b+1}^{\infty} \check{X}_i^{(n_k)} \right] \leq \frac{\varepsilon}{3}, \quad E \left[\sum_{i=b+1}^{\infty} Y_i \right] \leq \frac{\varepsilon}{3}$$

which implies

$$E \left[\sum_{i=b+1}^{\infty} |\check{X}_i^{(n_k)} - Y_i| \right] \leq \frac{2\varepsilon}{3}$$

for all k such that $n_k \geq \bar{n}$.

Since $\check{X}_i^{(n_k)} - Y_i$ converges to 0 a.s., and $|\check{X}_i^{(n_k)} - Y_i| \leq 2$ a.s., by Bounded Convergence Theorem we have

$$\lim_{k \rightarrow \infty} E \left[\sum_{i=0}^b |\check{X}_i^{(n_k)} - Y_i| \right] = 0$$

Thus, for sufficiently large k , $E \left[\sum_{i=0}^b |\check{X}_i^{(n_k)} - Y_i| \right] \leq \frac{\varepsilon}{3}$ and

$$\begin{aligned} &E[\|\check{X}^{(n_k)} - Y\|_1] \\ &= E \left[\sum_{i=0}^b |\check{X}_i^{(n_k)} - Y_i| \right] + E \left[\sum_{i=b+1}^{\infty} |\check{X}_i^{(n_k)} - Y_i| \right] \leq \varepsilon \end{aligned}$$

Hence, we established that

$$\lim_{k \rightarrow \infty} \mathbb{E} \left[\|\check{\mathbf{X}}^{(n_k)} - \mathbf{Y}\|_1 \right] = 0$$

By Triangle Inequality of Wasserstein distances and Proposition 2, we conclude

$$\begin{aligned} & d_{W_1, \|\cdot\|_1}(\mathcal{L}(\hat{\mathbf{X}}^{(n_k)}), \mathcal{L}(\mathbf{Y})) \\ & \leq d_{W_1, \|\cdot\|_1}(\mathcal{L}(\hat{\mathbf{X}}^{(n_k)}), \mathcal{L}(\check{\mathbf{X}}^{(n_k)})) + d_{W_1, \|\cdot\|_1}(\mathcal{L}(\check{\mathbf{X}}^{(n_k)}), \mathcal{L}(\mathbf{Y})) \\ & \leq d_{W_1, \|\cdot\|_1}(\mathcal{L}(\hat{\mathbf{X}}^{(n_k)}), \mathcal{L}(\check{\mathbf{X}}^{(n_k)})) + \mathbb{E} \left[\|\check{\mathbf{X}}^{(n_k)} - \mathbf{Y}\|_1 \right] \\ & \xrightarrow{k \rightarrow \infty} 0 \end{aligned}$$

Therefore, we have proved that any subsequence of $(\hat{\mathbf{X}}^{(n)})$ that weakly converges in $([0, 1]^{\mathbb{Z}_+}, \rho)$ also converges in $d_{W_1, \|\cdot\|_1}$. We conclude that every subsequence of $(\hat{\mathbf{X}}^{(n)})$ has a further subsequence that converges in $d_{W_1, \|\cdot\|_1}$. $\square$

Recall the definition of the extended system $(\mathbf{Q}(t), \bar{\mathbf{W}}(t), \Xi(t))_{t \geq 0}$. Now, we allow $\Xi(0)$ to be an arbitrary random variable with support on $\{0, 1, \dots, \lfloor c \log n \rfloor - 1\}$. i.e. the first reset could happen before the arrival of the $\lfloor c \log n \rfloor$ -th job, and the random walkers do not necessarily start from uniform random positions.

PROPOSITION 3. *Let $(\mathbf{Q}^{(n)}(0), \bar{\mathbf{W}}^{(n)}(0), \Xi^{(n)}(0))$ be random vectors such that*

- (a) $\mathbf{X}^{(n)}(0)$ converges to $\mathbf{Y}(0)$ weakly in $([0, 1]^{\mathbb{Z}_+}, \|\cdot\|_1)$
- (b) $\|\mathbf{Y}(0)\|_1 < +\infty$ a.s.

then for every finite $t > 0$, $\mathbf{X}^{(n)}(t)$ converges weakly to $\mathbf{Y}(t)$ in $([0, 1]^{\mathbb{Z}_+}, \|\cdot\|_1)$, where $\mathbf{Y}(t)$ is the state of the limit dynamical system in 1 when the initial state $\mathbf{x}(0) = \mathbf{Y}(0)$.

PROOF. See Appendix A.10 $\square$

PROOF OF THEOREM 3. Let $(\hat{\mathbf{Q}}^{(n)}, \hat{\mathbf{W}}^{(n)}, \hat{\Xi}^{(n)})$ be a random vector whose distribution is the stationary distribution with respect to the n -server extended system. Set $(\mathbf{Q}^{(n)}(0), \bar{\mathbf{W}}^{(n)}(0), \Xi^{(n)}(0)) \stackrel{d}{\sim} (\hat{\mathbf{Q}}^{(n)}, \hat{\mathbf{W}}^{(n)}, \hat{\Xi}^{(n)})$ for every n . We have $(\mathbf{Q}^{(n)}(t), \bar{\mathbf{W}}^{(n)}(t), \Xi^{(n)}(t)) \stackrel{d}{\sim} (\hat{\mathbf{Q}}^{(n)}, \hat{\mathbf{W}}^{(n)}, \hat{\Xi}^{(n)})$ for every n and all $t \geq 0$. In particular,

$$\mathbf{X}^{(n)}(t) \stackrel{d}{\sim} \hat{\mathbf{X}}^{(n)}$$

for all n and $t \geq 0$.

Let $(\check{\mathbf{X}}^{(n_k)})_{k=1}^\infty$ be a subsequence of $(\hat{\mathbf{X}}^{(n)})_n$ that converges in $d_{W_1, \|\cdot\|_1}$. Denote its limit by $\mathbf{Y}$. In the proof of Corollary 5 we have shown that $\mathbb{E}[\|\mathbf{Y}\|_1] < +\infty$. Hence $\|\mathbf{Y}\|_1 < +\infty$ a.s.

By Proposition 3,

$$\mathbf{X}^{(n_k)}(t) \xrightarrow{k \rightarrow \infty} \mathbf{Y}(t) \quad \forall t \geq 0$$

where $\mathbf{Y}(t)$ is the state of the limiting dynamical system (specified by the differential equations in (1)) with initial state given by the random variable $\mathbf{x}(0) \stackrel{d}{\sim} \mathbf{Y}$.

We hence know that

$$\mathbf{Y}(t) \stackrel{d}{\sim} \mathbf{Y} \quad \forall t \geq 0 \tag{13}$$

It is proved in [23] that the limiting dynamics in (1) is global asymptotically stable. For any $\mathbf{x}(0) \in \ell_1([0, 1])$, we have $\mathbf{x}(t) \xrightarrow{t \rightarrow \infty} \hat{\mathbf{x}}$. Hence, we have

$$\lim_{t \rightarrow \infty} \mathbf{Y}(t) = \hat{\mathbf{x}} \quad a.s. \tag{14}$$

Combining (13)(14), we have $Y = \hat{x}$ a.s. We have proved that every subsequence of $(\hat{X}^{(n)})_n$ which converges in $d_{W_1, \|\cdot\|_1}$ converges *weakly* to $\hat{x}$.

Since the weak limit of a sequence of random vectors is unique, we conclude that every subsequence of $(\hat{X}^{(n)})_n$ which converges in $d_{W_1, \|\cdot\|_1}$ converges in $d_{W_1, \|\cdot\|_1}$ to $\hat{x}$. Since $\hat{x}$ is deterministic, we further have this subsequence to converge to $\hat{x}$ in L_1 .

Combining Corollary 5, we conclude that every subsequence of $(\hat{X}^{(n)})_n$ has a further subsequence that converges to $\hat{x}$ in L_1 .

Therefore we conclude that $\hat{X}^{(n)}$ converges to $\hat{x}$ in L_1 . $\square$

Given the theorem, many results on *power-of-d choices* algorithm could also apply on NBRWR-Pod.

COROLLARY 6. *The average waiting time for a job in a stationary system of n servers converges to $\sum_{i=1}^{\infty} \lambda^{\frac{d^i-d}{d-1}}$ as n goes to infinity.*

PROOF. Mitzenmacher proved in [23] that the fixed point $\hat{x}$ is given by

$$\hat{x}_i = \lambda^{\frac{d^i-1}{d-1}} \quad i \geq 1$$

By Little's Law, the average waiting time for a job in the n -server system is given by

$$\begin{aligned} \mathcal{T}^{(n)} &= \frac{\mathbb{E} \left[\sum_{i=1}^n \hat{Q}_i^{(n)} \right]}{\lambda n} = \frac{\mathbb{E} \left[n \sum_{i=1}^{\infty} \hat{X}_i^{(n)} \right]}{\lambda n} \\ &= \frac{1}{\lambda} \mathbb{E} \left[\sum_{i=1}^{\infty} \hat{X}_i^{(n)} \right] \xrightarrow{n \rightarrow \infty} \frac{1}{\lambda} \sum_{i=1}^{\infty} \hat{x}_i = \sum_{i=1}^{\infty} \lambda^{\frac{d^i-d}{d-1}} \end{aligned}$$

$\square$

4 SIMULATION RESULTS

We conduct extensive simulations to investigate the dynamics of a system of finitely many servers. Specifically, we would like to investigate how different graph families affects the dynamics.

We use Lubotzky, Phillips, and Sarnak's Ramanujan Graphs [18] as the underlying graphs in many simulations. As pointed out in Remark 1, these graphs are known to satisfy the high girth assumption asymptotically [18]. Specifically, we use a 6-regular LPS graph of $n = 12180$ vertices, which has girth 9. The simulation results in Figure 1 suggest that the differential equation provides a good approximation for the scheme even when n is relatively small (on the order of tens of thousands, which is the typical scale of a cloud-computing center).

We also conduct simulations for the system when the underlying graph has small girth. The simulation results are provided for a "slash" torus graph¹ (in Figure 2) and a torus graph (in Figures 3), where the former has girth 3 and the latter has girth 4. The simulation results show that the dynamics of the system under these two torus graphs deviates from the solution of the differential equation. This result suggests that when the girth assumption (Assumption 1) is not satisfied, there are examples where the differential equation could fail to approximate the dynamics of the NBRWR-Pod algorithm. This result also suggest that when the underlying graph has small girth, the limiting dynamics is sensitive to the actual girth of the graph, which contrasts with the case of large girth graphs, where the limiting dynamics hold irrespective of the specific graph family.

We also conduct simulations for the NBRWR-Pod algorithm on a family of random graphs. The family of random graphs, called the 2-cycle graph, is obtained from superposition of two random

¹A "slash" torus graph is a 6-regular graph, where the vertex set is $\mathbb{Z}_m \times \mathbb{Z}_n$, and (i, j) is connected to $(i-1, j)$, $(i+1, j)$, $(i, j-1)$, $(i, j+1)$, $(i-1, j-1)$, $(i+1, j+1)$, i.e. a torus graph with an extra "slash" in each grid.

Fig. 1. Queue length statistics evolution for NBRWR-Pod algorithm with $d = 2$ under $\lambda = 0.95$. The underlying graph G is 6-regular LPS Graph with $n = 12180$ vertices. The initial queue length is set to 0 (left) and 5 (right) for all queues respectively.

Fig. 2. Queue length statistics evolution for NBRWR-Pod algorithm with $d = 2$ under $\lambda = 0.95$. The underlying graph G is the “slash” torus graph $\mathbb{Z}_{105} \times \mathbb{Z}_{116}$. The initial queue length is set to 0 (left) and 5 (right) for all queues respectively.

Fig. 3. Queue length statistics evolution for NBRWR-Pod algorithm with $d = 2$ under $\lambda = 0.95$. The underlying graph G is the torus graph $\mathbb{Z}_{15} \times \mathbb{Z}_{28} \times \mathbb{Z}_{29}$. The initial queue length is set to 0 (left) and 5 (right) for all queues respectively.

Hamiltonian cycles. This family of random graphs has been used in peer-to-peer streaming systems [17]. In our settings, we need to make sure that the graph is a regular simple graph, hence we modify the procedure as follows: the second Hamiltonian cycle is chosen uniformly at random over the cycles which have no edge overlap with the first cycle. Given this construction, the resulting

graph is always a 4-regular graph. We generated a random graph for $n = 12180$ vertices which has girth 3. Figure 4 shows the simulation results on the 2-cycle graph with $n = 12180$ vertices. The results suggest that the differential equation approximation can still be accurate for small girth graphs, as long as there are not too many small cycles in the graph. Such performance is also observed in k -regular random graphs, where the comment above on the girth also holds. These results indicate that the girth assumption is not necessary for the fluid limit approximation result. The general necessary and sufficient conditions are still to be uncovered.

Fig. 4. Queue length statistics evolution for NBRWR-Pod algorithm with $d = 2$ under $\lambda = 0.95$. The underlying graph G are 4-regular 2-Cycle Random Graphs with $n = 12180$ vertices. The system is initialized to be empty.

For all the simulations discussed above, we set the reset interval to be 50, that is, the random walkers are reset to independent uniform positions for every 50 job arrivals. We also conduct simulations to investigate the influence of restart period on the system. For this purpose, we test the NBRWR-Pod algorithm on a cycle graph. For a fixed cycle graph, we vary the restart interval (Figure 5). From the simulation results, we see that as the restart period goes up, the process becomes more and more noisy and the estimation error for the differential equation goes up. This result suggests that resets are necessary for the fluid-limit approximation to work.

In all figures, the first four coordinates of the queue length proportion vector $\mathbf{X}(t)$ and differential equation solution $\mathbf{x}(t)$ are compared with each other. Additional simulation results are included in Appendix B.

5 DISCUSSIONS

For a practical implementation of the NBRWR-Pod algorithm, one may think that large memory will be required to store the graph G . This is true if we use unstructured graphs. For graphs with special structures, the graph need not be stored in physical memory and the neighbors of any vertex can be computed easily, instead of performing a memory table lookup. For example, the LPS graph [18] is a Cayley graph of a PSL group. One can find the neighbors of any vertex just by multiplying 2×2 matrices. Another example is the cycle graph, where a non-backtracking random walk on a cycle graph is equivalent to round-robin, once the direction is set.

Our result is a significant improvement on previous low-randomness load balancing algorithms in terms of performance. For example, the Round-Robin scheme, as a deterministic scheme, has larger expected delay than power-of- d . Also, to the best of our knowledge, our result is the first low-randomness variant of power-of- d , as many other variants of power-of- d introduce only

Fig. 5. Queue length statistics evolution for NBRWR-Pod algorithm for cycle graph. The restart period is set to 50 (left), 200 (right), 1000 (bottom) respectively.

dependence between the d choices of the same job, while each job still independently samples a subset of d servers from n servers.

6 CONCLUSIONS AND FUTURE WORK

In this paper we analyzed the performance of a new load balancing algorithm for the balls-in-bins problem proposed in [28], in the context of multi-server queuing systems. The new algorithm is a variant of the classical *power-of- d choices* algorithm, which replaces uniform independent sampling with random walks on k -regular graph based sampling. We show that this new algorithm yields a similar behavior as *power-of- d choices* for large scale systems. Hence, the new algorithm can be considered as a derandomization of *power-of- d choices*.

The study of the coupling between queue assignments engendered by the random walk based sampling required the development of many novel proof techniques, which we argue are also a noteworthy contribution to the state of the art.

This paper suggests the following future work. First, the performance analysis for the other two proposed schemes in [28] in queuing systems are of interest, especially the scheme without resets. Second, it is also interesting to see if the NBRWR-Pod algorithm could behave similarly as the regular *power-of- d choices* scheme in heavy traffic. Third, developing necessary and other sufficient conditions for good performance of the random walks based sampling scheme would also be of interest. Fourth, understanding whether the propagation of chaos property holds or not is also worth studying. Finally, the results of this paper suggest that random walk based sampling can be used in many other dynamic contexts, with good performance when some properties hold for the underlying graph, and it would be worth exploring this possibility more broadly.

ACKNOWLEDGMENTS

The authors would like acknowledge support by NSF via grants AST-1343381, AST-1516075, IIS-1538827 and ECCS-1608361, and also Harsha Honappa and Yang Xiao for early discussions.

REFERENCES

- [1] Reza Aghajani and Kavita Ramanan. 2017. The hydrodynamic limit of a randomized load balancing network. *arXiv preprint arXiv:1707.02005* (2017).
- [2] Noga Alon, Itai Benjamini, Eyal Lubetzky, and Sasha Sodin. 2007. Non-backtracking random walks mix faster. *Communications in Contemporary Mathematics* 9, 04 (2007), 585–603.
- [3] Noga Alon and Eyal Lubetzky. 2009. Poisson approximation for non-backtracking random walks. *Israel Journal of Mathematics* 174, 1 (2009), 227–252.
- [4] Patrick Billingsley. 2013. *Convergence of probability measures*. John Wiley & Sons.
- [5] Aleksandr Alekseevich Borovkov. 1998. Ergodicity and stability of stochastic processes. (1998).
- [6] Amarjit Budhiraja, Debankur Mukherjee, and Ruoyu Wu. 2017. Supermarket model on graphs. *arXiv preprint arXiv:1712.07607* (2017).
- [7] Fan Chung and Linyuan Lu. 2006. Concentration inequalities and martingale inequalities: a survey. *Internet Mathematics* 3, 1 (2006), 79–127.
- [8] R.W.R. Darling and James R. Norris. 2008. Differential equation approximations for Markov chains. *Probability surveys* 5 (2008), 37–79.
- [9] Moez Draief and Laurent Massouli. 2010. *Epidemics and rumours in complex networks*. Cambridge University Press.
- [10] Stewart N. Ethier and Thomas G. Kurtz. 2009. *Markov processes: characterization and convergence*. Vol. 282. John Wiley & Sons.
- [11] G. Foschini and Jack Salz. 1978. A basic dynamic routing problem and diffusion. *IEEE Transactions on Communications* 26, 3 (1978), 320–327.
- [12] Ayalvadi Ganesh, Sarah Lilienthal, D Manjunath, Alexandre Proutiere, and Florian Simatos. 2010. Load balancing via random local search in closed and open systems. In *ACM SIGMETRICS Performance Evaluation Review*, Vol. 38. ACM, 287–298.
- [13] Nicolas Gast. 2015. The Power of Two Choices on Graphs: the Pair-Approximation is Accurate? *ACM SIGMETRICS Performance Evaluation Review* 43, 2 (2015), 69–71.
- [14] Nicolas Gast. 2017. Expected values estimated via mean-field approximation are $1/N$ -accurate. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 1, 1 (2017), 17.
- [15] Michel Habib, Colin McDiarmid, Jorge Ramirez-Alfonsin, and Bruce Reed. 2013. *Probabilistic methods for algorithmic discrete mathematics*. Vol. 16. Springer Science & Business Media.
- [16] Shlomo Hoory, Nathan Linial, and Avi Wigderson. 2006. Expander graphs and their applications. *Bull. Amer. Math. Soc.* 43, 4 (2006), 439–561.
- [17] Joohean Kim and R Srikanth. 2013. Real-time peer-to-peer streaming over multiple random hamiltonian cycles. *IEEE Transactions on Information Theory* 59, 9 (2013), 5763–5778.
- [18] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. 1988. Ramanujan graphs. *Combinatorica* 8, 3 (1988), 261–277.
- [19] MJ Luczak and JR Norris. 2013. Averaging over fast variables in the fluid limit for Markov chains: application to the supermarket model with memory. *The Annals of Applied Probability* 23, 3 (2013), 957–986.
- [20] Malwina J Luczak and Colin McDiarmid. 2006. On the maximum queue length in the supermarket model. *The Annals of Probability* 34, 2 (2006), 493–527.
- [21] Malwina J Luczak and James Norris. 2005. Strong approximation for the supermarket model. *The Annals of Applied Probability* 15, 3 (2005), 2038–2061.
- [22] Michael Mitzenmacher. 1996. Load balancing and density dependent jump Markov processes. In *IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE, 213.
- [23] Michael Mitzenmacher. 2001. The power of two choices in randomized load balancing. *IEEE Transactions on Parallel and Distributed Systems* 12, 10 (2001), 1094–1104.
- [24] Michael Mitzenmacher, Balaji Prabhakar, and Devavrat Shah. 2002. Load balancing with memory. In *IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE, 799.
- [25] Debankur Mukherjee, Sem C Borst, and Johan SH van Leeuwen. 2018. Asymptotically optimal load balancing topologies. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 2, 1 (2018), 14.
- [26] Debankur Mukherjee, Sem C Borst, Johan SH van Leeuwen, and Philip A Whiting. 2016. Universality of Power-of-d Load Balancing in Many-Server Systems. *arXiv preprint arXiv:1612.00723* (2016).
- [27] Alain-Sol Sznitman. 1991. Topics in propagation of chaos. In *Ecole d’été de probabilités de Saint-Flour XIX—1989*. Springer, 165–251.

- [28] Dengwang Tang and Vijay G. Subramanian. 2018. Balanced Allocation with Random Walk Based Sampling. *arXiv preprint arXiv:1810.02722* (2018).
- [29] Nikita Dmitrievna Vvedenskaya, Roland L'vovich Dobrushin, and Fridrikh Izrailevich Karpelevich. 1996. Queueing system with selection of the shortest of two queues: An asymptotic approach. *Problemy Peredachi Informatsii* 32, 1 (1996), 20–34.
- [30] Nicholas C Wormald. 1995. Differential equations for random processes and random graphs. *The annals of applied probability* 5, 4 (1995), 1217–1235.
- [31] Lei Ying. 2016. On the approximation error of mean-field models. In *ACM SIGMETRICS Performance Evaluation Review*, Vol. 44. ACM, 285–297.
- [32] Lei Ying, R. Srikant, and Xiaohan Kang. 2015. The power of slightly more than one sample in randomized load balancing. In *Computer Communications (INFOCOM), 2015 IEEE Conference on*. IEEE, 1131–1139.
- [33] Se-Young Yun and Alexandre Proutiere. 2015. Distributed proportional fair load balancing in heterogenous systems. In *ACM SIGMETRICS Performance Evaluation Review*, Vol. 43. ACM, 17–30.

A OMITTED PROOFS

A.1 Proof of Lemma 1

By Lemma 8.2 in [7], there exist random variables $(Y_j)_{j=0}^N$ such that

- $(Y_j)_{j=0}^N$ is a martingale w.r.t. $(\mathcal{F}_j)_{j=1}^N$
- $|Y_j - Y_{j-1}| \leq c_j$
- $Y_j = X_j$ under event C .

By Azuma-Hoeffding Inequality for Maxima (Eq. 3.30 of [15]),

$$\Pr \left(\max_{0 \leq j \leq N} |Y_j - Y_0| \geq \lambda \right) \leq 2 \exp \left(-\frac{\lambda^2}{2 \sum_{i=1}^N c_i^2} \right)$$

Hence

$$\begin{aligned} \Pr \left(\max_{0 \leq j \leq N} |X_j - X_0| \geq \lambda, C \right) &= \Pr \left(\max_{0 \leq j \leq N} |Y_j - Y_0| \geq \lambda, C \right) \\ &\leq 2 \exp \left(-\frac{\lambda^2}{2 \sum_{i=1}^N c_i^2} \right) \end{aligned}$$

□

A.2 Proof of Lemma 2

By Chernoff bound we know that

$$\begin{aligned} \Pr(\mathcal{A}^c) &= \Pr(\text{Poisson}(\lambda n T) \geq e \lambda n T) \\ &\leq e^{-se \lambda n T} \exp(\lambda n T(e^s - 1)) \quad \forall s > 0 \end{aligned}$$

pick $s = 1$ results in

$$\Pr(\mathcal{A}^c) \leq e^{-\lambda n T}$$

Now we estimate the probability of $\mathcal{B}^c$: If $\mathcal{B}$ is not true, then either $\mathcal{A}$ is not true, or some queue is allocated with more than $\kappa L \log n$ arrival jobs while $\mathcal{A}$ is true. Let $V_i[j]$ denote the sum of number of visits to queue i by the random walkers between (discrete) time $j \lfloor c \log n \rfloor$ and $(j+1) \lfloor c \log n \rfloor$. Let $N = \frac{e \lambda T n}{\lfloor c \log n \rfloor}$

We have

$$\begin{aligned}\Pr(\mathcal{B}^c) &\leq \Pr(\mathcal{A}^c) + \Pr\left(\exists i \in [n], \sum_{j=0}^{\lfloor N \rfloor} V_i[j] \geq \kappa L \log n\right) \\ &\leq \Pr(\mathcal{A}^c) + n \Pr\left(\sum_{j=0}^{\lfloor N \rfloor} V_1[j] \geq \kappa L \log n\right)\end{aligned}$$

By the girth assumption, a random walker cannot visit a certain vertex more than

$$\left\lceil \frac{\lfloor c \log n \rfloor}{2\lceil \alpha \log_{k-1} n \rceil + 1} \right\rceil \leq \frac{c \log(k-1)}{2\alpha} + 1$$

times in $\lfloor c \log n \rfloor$ timestamps. Hence we have $V_1[j] \leq L$. Hence

$$\begin{aligned}\Pr(\mathcal{B}^c) &\leq \Pr(\mathcal{A}^c) + n \Pr\left(\sum_{j=0}^{\lfloor N \rfloor} L \mathbb{1}_{V_1[j]>0} \geq \kappa L \log n\right) \\ &= \Pr(\mathcal{A}^c) + n \Pr\left(\sum_{j=0}^{\lfloor N \rfloor} \mathbb{1}_{V_1[j]>0} \geq \kappa \log n\right)\end{aligned}$$

We know that $\mathbb{1}_{V_1[j]>0}, j = 0, 1, \dots$ are i.i.d. Bernoulli random variables with parameter p where

$$p \leq \mathbb{E}[V_1[j]] \leq d \lfloor c \log n \rfloor \cdot \frac{1}{n}$$

Hence

$$\begin{aligned}\Pr(\mathcal{B}^c) &\leq \Pr(\mathcal{A}^c) + n \Pr\left(\text{Binomial}\left(\lfloor N \rfloor, \frac{d \lfloor c \log n \rfloor}{n}\right) \geq \kappa \log n\right) \\ &\leq \Pr(\mathcal{A}^c) + n \exp(-s\kappa \log n) \left(1 - \frac{d \lfloor c \log n \rfloor}{n} + \frac{d \lfloor c \log n \rfloor}{n} e^s\right)^{\lfloor N \rfloor} \\ &\leq \Pr(\mathcal{A}^c) + n \exp(-s\kappa \log n) \exp\left(\frac{d \lfloor c \log n \rfloor}{n} \lfloor N \rfloor (e^s - 1)\right) \\ &\leq \Pr(\mathcal{A}^c) + n \exp(-s\kappa \log n) \exp(ed\lambda T(e^s - 1))\end{aligned}$$

Recall that $\kappa = \frac{3}{\log(1 + \frac{1}{ed\lambda T})}$ and $s = \log(1 + \frac{1}{ed\lambda T})$, we have

$$\Pr(\mathcal{B}^c) \leq \Pr(\mathcal{A}^c) + n^{-2}$$

□

A.3 Proof of Lemma 3

Recall the definition of $\tilde{T}[s]$ in the discrete time system. Let $H_j = \tilde{T}[j+1] - \tilde{T}[j] - 1$. H_j equals to the number of potential services between the j -th and $(j+1)$ -th job arrival. A standard result tells us that $(H_j)_{j=1}^\infty$ are i.i.d. Geometric random variables with mean $\frac{1}{\lambda}$.

Now, observe that by construction, for $T_j \leq t < T_{j+1}$

$$\|\bar{\mathbf{X}}(t) - \tilde{\mathbf{X}}(t)\|_1 = \|\tilde{\mathbf{X}}(T_j) - \tilde{\mathbf{X}}(t)\|_1 \leq \frac{1}{n} \sum_{i=1}^n |\tilde{Q}_i(T_j) - \tilde{Q}_i(t)| \leq \frac{H_j}{n}$$

Hence, conditioned on $\mathcal{A}$ we have

$$\sup_{t \in [0, T]} \|\bar{\mathbf{X}}(t) - \tilde{\mathbf{X}}(t)\|_1 \leq \frac{1}{n} \max_{0 \leq j \leq \lfloor e\lambda Tn \rfloor} H_j$$

Thus for any $\varepsilon_2 > 0$

$$\begin{aligned} & \Pr \left(\sup_{t \in [0, T]} \|\bar{\mathbf{X}}(t) - \tilde{\mathbf{X}}(t)\|_1 \geq \varepsilon_2, \mathcal{A} \right) \\ & \leq \Pr \left(\frac{1}{n} \max_{0 \leq j \leq \lfloor e\lambda Tn \rfloor} H_j \geq \varepsilon_2 \right) \\ & \leq \sum_{j=0}^{\lfloor e\lambda Tn \rfloor} \Pr(H_j \geq n\varepsilon_2) \\ & \leq \lfloor e\lambda Tn \rfloor \cdot \left(\frac{1}{1+\lambda} \right)^{\lfloor n\varepsilon_2 \rfloor} \\ & \leq e(1+\lambda)\lambda Tn \left(\frac{1}{1+\lambda} \right)^{n\varepsilon_2} \end{aligned}$$

By Lipschitz continuity of $\mathbf{a}$, $\sup_{t \in [0, T]} \|\bar{\mathbf{X}}(t) - \tilde{\mathbf{X}}(t)\|_1 \leq \varepsilon_2$ implies that $\sup_{t \in [0, T]} \|\mathbf{a}(\bar{\mathbf{X}}(t)) - \mathbf{a}(\tilde{\mathbf{X}}(t))\|_1 \leq 2\lambda d\varepsilon_2$, which further implies that $\sup_{t \in [0, T]} \|\mathbf{M}^{a,2}(t)\|_1 \leq 2\lambda dT\varepsilon_2$. Hence

$$\Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^{a,2}(t)\|_1 \geq 2\lambda dT\varepsilon_2, \mathcal{A} \right) \leq e(1+\lambda)\lambda Tn \left(\frac{1}{1+\lambda} \right)^{n\varepsilon_2}$$

□

A.4 Proof of Lemma 4

For $0 \leq s < \lfloor c \log n \rfloor$, define $I_i[j, s]$ to be the indicator of the event that the $(j\lfloor c \log n \rfloor + s)$ -th arrival is allocated to a queue of load $i - 1$.

Recall that τ_j is the inter-arrival time between job $j - 1$ and job j (τ_1 is defined to be the arrival time of the first job). We have

$$\begin{aligned} & M_i^{a,1}\langle j+1 \rangle - M_i^{a,1}\langle j \rangle \\ &= \frac{1}{n} \sum_{s=0}^{\lfloor c \log n \rfloor - 1} I_i[j, s] - \int_{T_j\lfloor c \log n \rfloor}^{T_{j+1}\lfloor c \log n \rfloor} a_i(\bar{\mathbf{X}}(u)) du \\ &= \frac{1}{n} \sum_{s=0}^{\lfloor c \log n \rfloor - 1} I_i[j, s] - \sum_{s=0}^{\lfloor c \log n \rfloor - 1} a_i(\tilde{\mathbf{X}}(T_j\lfloor c \log n \rfloor + s)) \tau_j\lfloor c \log n \rfloor + s + 1 \\ &= \sum_{s=0}^{\lfloor c \log n \rfloor - 1} \left(\frac{1}{n} I_i[j, s] - a_i(\tilde{\mathbf{X}}(T_j\lfloor c \log n \rfloor + s)) \tau_j\lfloor c \log n \rfloor + s + 1 \right) \end{aligned} \tag{15}$$

Define $\mathcal{D}_{j,s}$ to be the event that either one of $W_1[j\lfloor c \log n \rfloor + s]$ or $W_2[j\lfloor c \log n \rfloor + s]$ has been visited by at least one of the random walkers at some timestamps $j\lfloor c \log n \rfloor \leq t < j\lfloor c \log n \rfloor + s$.

For ease of notation, define $W_l\{j, s\} := W_l\{j\lfloor c \log n \rfloor + s\}$, $T_{j,s} := T_j\lfloor c \log n \rfloor + s$, $\tau_{j,s} := \tau_j\lfloor c \log n \rfloor + s$.

Conditioned on $\mathcal{D}_{j,s}^c$, the sampled queues have not accepted new jobs since T_j , we have $Q_{W_l\{j,s\}}(T_{j,s}-) = \tilde{Q}_{W_l\{j,s\}}(T_{j,s})$, as $\tilde{Q}(T_{j,s})$ is coupled to $Q(T_{j,s})$ such that no arrivals are accepted between resets.

Therefore

$$\begin{aligned}
& \mathbb{E}[I_i[j, s] \mid \mathcal{G}_j] \\
&= \Pr \left(\min_{1 \leq l \leq d} Q_{W_l\{j, s\}}(T_{j, s-}) = i - 1 \mid \mathcal{G}_j \right) \\
&\leq \Pr \left(\min_{1 \leq l \leq d} Q_{W_l\{j, s\}}(T_{j, s-}) = i - 1, \mathcal{D}_{j, s}^c \mid \mathcal{G}_j \right) + \Pr(\mathcal{D}_{j, s} \mid \mathcal{G}_j) \\
&= \Pr \left(\min_{1 \leq l \leq d} \tilde{Q}_{W_l\{j, s\}}(T_{j, s}) = i - 1, \mathcal{D}_{j, s}^c \mid \mathcal{G}_j \right) + \Pr(\mathcal{D}_{j, s} \mid \mathcal{G}_j) \\
&\leq \Pr \left(\min_{1 \leq l \leq d} \tilde{Q}_{W_l\{j, s\}}(T_{j, s}) = i - 1 \mid \mathcal{G}_j \right) + \Pr(\mathcal{D}_{j, s})
\end{aligned}$$

Similarly we have a lower bound

$$\begin{aligned}
& \mathbb{E}[I_i[j, s] \mid \mathcal{G}_j] \\
&= \Pr \left(\min_{1 \leq l \leq d} Q_{W_l\{j, s\}}(T_{j, s-}) = i - 1 \mid \mathcal{G}_j \right) \\
&\geq \Pr \left(\min_{1 \leq l \leq d} Q_{W_l\{j, s\}}(T_{j, s-}) = i - 1, \mathcal{D}_{j, s}^c \mid \mathcal{G}_j \right) \\
&= \Pr \left(\min_{1 \leq l \leq d} \tilde{Q}_{W_l\{j, s\}}(T_{j, s}) = i - 1, \mathcal{D}_{j, s}^c \mid \mathcal{G}_j \right) \\
&\geq \Pr \left(\min_{1 \leq l \leq d} \tilde{Q}_{W_l\{j, s\}}(T_{j, s}) = i - 1 \mid \mathcal{G}_j \right) - \Pr(\mathcal{D}_{j, s})
\end{aligned}$$

Now, we take advantage of the following important observation:

OBSERVATION 2. *Given $\mathcal{G}_j$, $(\tilde{\mathbf{Q}}(T_{j, s}))_{0 \leq s < \lfloor c \log n \rfloor}$ is conditionally independent of $(\mathbf{W}\{j, s\})_{s \geq 0}$*

The observation is true since $(\tilde{\mathbf{Q}}(T_{j, s}))_{0 \leq s < \lfloor c \log n \rfloor}$ is obtained from applying the potential service schedules to $\mathbf{Q}(T_{j, 0})$, where since the random walkers reset, both $\mathbf{Q}(T_{j, 0})$ and the service schedule are conditionally independent of the random walker positions given $\mathcal{G}_j$.

Using the observation, we conclude that $\tilde{\mathbf{Q}}(T_{j, s})$ is independent of $\mathbf{W}\{j, s\}$ conditioning on $\mathcal{G}_j$. We then conclude that

$$\begin{aligned}
& \mathbb{E}[I_i[j, s] \mid \mathcal{G}_j] \\
&\leq \Pr \left(\min_{1 \leq l \leq d} \tilde{Q}_{W_l\{j, s\}}(T_{j, s}) = i - 1 \mid \mathcal{G}_j \right) + \Pr(\mathcal{D}_{j, s} \mid \mathcal{G}_j) \\
&= \mathbb{E} \left[\Pr \left(\min_{1 \leq l \leq d} \tilde{Q}_{W_l\{j, s\}}(T_{j, s}) = i - 1 \mid \tilde{\mathbf{Q}}(T_{j, s}), \mathcal{G}_j \right) \mid \mathcal{G}_j \right] \\
&\quad + \Pr(\mathcal{D}_{j, s}) \\
&= \mathbb{E} \left[(\tilde{X}_{i-1}(T_{j, s}))^d - (\tilde{X}_i(T_{j, s}))^d \mid \mathcal{G}_j \right] + \Pr(\mathcal{D}_{j, s}) \\
&= \frac{1}{\lambda} \mathbb{E}[a_i(\tilde{\mathbf{X}}(T_{j, s})) \mid \mathcal{G}_j] + \Pr(\mathcal{D}_{j, s})
\end{aligned} \tag{16}$$

and

$$\mathbb{E}[I_i[j, s] \mid \mathcal{G}_j] \geq \frac{1}{\lambda} \mathbb{E}[a_i(\tilde{\mathbf{X}}(T_{j, s})) \mid \mathcal{G}_j] - \Pr(\mathcal{D}_{j, s}) \tag{17}$$

By construction, $(\tilde{X}(T_{j,s}))$ is independent of $\tau_{j,s+1}$ conditioned on $\mathcal{G}_j$ (Actually, $\tau_{j,s+1}$ is independent of $\mathcal{F}_{T_{j,s}}$). Recall that $\tau_{j,s+1}$ is an exponential random variable with mean $\frac{1}{\lambda n}$. We have

$$\begin{aligned} & \mathbb{E}[a_i(\tilde{X}(T_{j,s})\tau_{j,s+1} \mid \mathcal{G}_j)] \\ &= \mathbb{E}[\tau_{j,s+1} \mid \mathcal{G}_j] \mathbb{E}[a_i(\tilde{X}(T_{j,s}) \mid \mathcal{G}_j)] \\ &= \frac{1}{\lambda n} \mathbb{E}[a_i(\tilde{X}(T_{j,s}) \mid \mathcal{G}_j)] \end{aligned} \tag{18}$$

Combining (15)(16)(17)(18) we have

$$\begin{aligned} & |\mathbb{E}[M_i^{a,1}\langle j+1 \rangle - M_i^{a,1}\langle j \rangle \mid \mathcal{G}_j]| \\ &\leq \sum_{s=0}^{\lfloor c \log n \rfloor - 1} \left| \frac{1}{n} \mathbb{E}[I_i[j, s] \mid \mathcal{G}_j] - \mathbb{E}[a_i(\tilde{X}(T_{j,s})\tau_{j,s+1} \mid \mathcal{G}_j)] \right| \\ &\leq \sum_{s=0}^{\lfloor c \log n \rfloor - 1} \frac{1}{n} \Pr(\mathcal{D}_{j,s}) \end{aligned}$$

CLAIM 2.

$$\Pr(\mathcal{D}_{j,s}) \leq \frac{ds}{n^\alpha} + \frac{d(d-1)s}{n}$$

PROOF OF CLAIM. By Union Bound we have

$$\begin{aligned} \Pr(\mathcal{D}_{j,s}) &\leq \sum_{l=1}^d \sum_{r=0}^{s-1} \Pr(W_l\{j, r\} = W_l\{j, s\}) \\ &\quad + \sum_{l_1 \neq l_2} \sum_{r=0}^{s-1} \Pr(W_{l_1}\{j, r\} = W_{l_2}\{j, s\}) \\ &= d \sum_{r=0}^{s-1} \Pr(W_1\{j, r\} = W_1\{j, s\}) \\ &\quad + d(d-1) \sum_{r=0}^{s-1} \Pr(W_1\{j, r\} = W_2\{j, s\}) \end{aligned} \tag{19}$$

Here we will use the result from [28] that characterizes the probability of a random walker visiting the same node at two points of time in between resets. For completeness we state the result below:

LEMMA 11. *Let $V(t)$ be a non-backtracking random walk on G . Then for all $s > t$*

$$\Pr(V(s) = V(t) \mid V(t), V(t-1)) \leq \frac{1}{n^\alpha} \quad a.s.$$

Using the lemma we have

$$\Pr(W_1\{j, r\} = W_1\{j, s\}) \leq \frac{1}{n^\alpha} \tag{20}$$

$W_1\{j, r\}, W_2\{j, s\}$ are independent uniform random vertices, hence

$$\Pr(W_1\{j, r\} = W_2\{j, s\}) = \frac{1}{n} \tag{21}$$

Combining (19)(20)(21) we prove the claim. $\square$

Hence we conclude

$$\begin{aligned}
 & |E[M_i^{a,1}\langle j+1 \rangle - M_i^{a,1}\langle j \rangle \mid \mathcal{G}_j]| \\
 &= |E[M_i^{a,1}\langle j+1 \rangle \mid \mathcal{G}_j] - M_i^{a,1}\langle j \rangle| \\
 &\leq \frac{d\lfloor c \log n \rfloor^2}{2n} \left(\frac{1}{n^\alpha} + \frac{d-1}{n} \right)
 \end{aligned}$$

□

A.5 Proof of Corollary 1

We first prove a large deviation result for sum of exponential random variables.

LEMMA 12. *Let Z be the sum of $\lfloor c \log n \rfloor$ independent exponential random variables with mean $\frac{1}{\lambda n}$, then*

$$\Pr \left(Z \geq \frac{r \log n (\lfloor c \log n \rfloor + 1)}{\lambda n} \right) \leq e n^{-r}$$

PROOF. The moment generating function of Z is

$$M_Z(s) = \left(\frac{\lambda n}{\lambda n - s} \right)^{\lfloor c \log n \rfloor} = \left(1 + \frac{1}{\frac{\lambda n}{s} - 1} \right)^{\lfloor c \log n \rfloor}$$

Set $s = \frac{\lambda n}{\lfloor c \log n \rfloor + 1}$ we obtain

$$M_Z(s) = \left(1 + \frac{1}{\lfloor c \log n \rfloor} \right)^{\lfloor c \log n \rfloor} \leq e$$

By Markov Inequality, for any $\xi > 0$ we have

$$\Pr(Z \geq \xi) \leq e^{-s\xi} M_Z(s) \leq e^{-(s\xi-1)}$$

Pick $\xi = \frac{r \log n (\lfloor c \log n \rfloor + 1)}{\lambda n}$ we have

$$\Pr \left(Z \geq \frac{r \log n (\lfloor c \log n \rfloor + 1)}{\lambda n} \right) \leq e^{-(r \log n - 1)} = e n^{-r}$$

□

Decompose $M_i^{a,1}\langle j \rangle = \tilde{M}_i^{a,1}\langle j \rangle + \Delta_i^{a,1}\langle j \rangle$ where

$$\tilde{M}_i^{a,1}\langle j \rangle := M_i^{a,1}\langle 0 \rangle + \sum_{l=1}^j (M_i^{a,1}\langle l \rangle - E[M_i^{a,1}\langle l \rangle \mid \mathcal{G}_{l-1}])$$

is the martingale part, and

$$\Delta_i^{a,1}\langle j \rangle = \sum_{l=1}^j (E[M_i^{a,1}\langle l \rangle \mid \mathcal{G}_{l-1}] - M_i^{a,1}\langle l-1 \rangle)$$

is the compensator, where we also have

$$|\Delta_i^{a,1}\langle j \rangle| \leq j \cdot \frac{d\lfloor c \log n \rfloor^2}{2n} \left(\frac{1}{n^\alpha} + \frac{d-1}{n} \right) \quad \forall j \quad a.s.$$

Thus,

$$\begin{aligned}
& \sup_{0 \leq j < N} |M_i^{a,1}\langle j \rangle| \\
& \leq \sup_{0 \leq j < N} |\tilde{M}_i^{a,1}\langle j \rangle| + \frac{e\lambda Tn}{\lfloor c \log n \rfloor} \frac{d \lfloor c \log n \rfloor^2}{2n} \left(\frac{1}{n^\alpha} + \frac{d-1}{n} \right) \\
& = \sup_{0 \leq j < N} |\tilde{M}_i^{a,1}\langle j \rangle| + \delta \quad \forall j \quad a.s.
\end{aligned} \tag{22}$$

Now we need to provide a bound for the difference of the martingale sequence $\tilde{M}_i^{a,1}[j]$. Notice that

$$\tilde{M}_i^{a,1}\langle j+1 \rangle - \tilde{M}_i^{a,1}\langle j \rangle = Z_i^{a,1}\langle j \rangle - \mathbb{E}[Z_i^{a,1}\langle j \rangle \mid \mathcal{G}_j]$$

where

$$\begin{aligned}
Z_i^{a,1}\langle j \rangle &:= M_i^{a,1}\langle j+1 \rangle - M_i^{a,1}\langle j \rangle \\
&= \frac{1}{n} \sum_{s=0}^{\lfloor c \log n \rfloor - 1} I_i[j, s] - \sum_{s=0}^{\lfloor c \log n \rfloor - 1} a_i(\tilde{X}(T_{j,s}-)) \tau_{j,s+1}
\end{aligned}$$

We know that

$$- \sum_{s=0}^{\lfloor c \log n \rfloor - 1} \tau_{j \lfloor c \log n \rfloor + s + 1} \leq Z_i^{a,1}\langle j \rangle \leq \frac{\lfloor c \log n \rfloor}{n} \quad a.s.$$

where the lower bound is obtained using $I_i[j, s] \geq 0$ and $a_i(\tilde{X}(T_{j,s}-)) \leq 1$, and the upper bound is obtained using $I_i[j, s] \leq 1$ and $a_i(\tilde{X}(T_{j,s}-)) \geq 0$.

Hence

$$\mathbb{E}[Z_i^{a,1}\langle j \rangle \mid \mathcal{G}_{j-1}] \geq \mathbb{E} \left[- \sum_{s=0}^{\lfloor c \log n \rfloor - 1} \tau_{j,s+1} \mid \mathcal{G}_{j-1} \right] = - \frac{\lfloor c \log n \rfloor}{\lambda n} \quad a.s.$$

and

$$\mathbb{E}[Z_i^{a,1}\langle j \rangle \mid \mathcal{G}_{j-1}] \leq \frac{\lfloor c \log n \rfloor}{n} \quad a.s.$$

Therefore

$$\begin{aligned}
& |\tilde{M}_i^{a,1}\langle j+1 \rangle - \tilde{M}_i^{a,1}\langle j \rangle| \\
& = |Z_i^{a,1}\langle j \rangle - \mathbb{E}[Z_i^{a,1}\langle j \rangle \mid \mathcal{G}_{j-1}]| \\
& \leq \frac{\lfloor c \log n \rfloor}{n} + \max \left\{ \sum_{s=0}^{\lfloor c \log n \rfloor - 1} \tau_{j,s+1}, \frac{\lfloor c \log n \rfloor}{\lambda n} \right\} \quad a.s.
\end{aligned}$$

Define C to be the event that

$$\forall 0 \leq j < N, \quad \sum_{s=0}^{\lfloor c \log n \rfloor - 1} \tau_{j \lfloor c \log n \rfloor + s + 1} \leq \frac{r \log n (\lfloor c \log n \rfloor + 1)}{\lambda n}$$

We immediately see that C is a “good” event (as mentioned in Theorem 1) in the sense that:

$$\forall 0 \leq j < N, \quad |\tilde{M}_i^{a,1}[j+1] - \tilde{M}_i^{a,1}[j]| \leq K_r$$

is true on C .

By Azuma-Hoeffding Inequality (Theorem 1) we have

$$\Pr \left(\sup_{0 \leq j < N} |\tilde{M}_i^{a,1}(j)| \geq \varepsilon_1, C \right) \leq 2 \exp \left(-\frac{\varepsilon_1^2}{2NK_r^2} \right)$$

Combining (22) we obtain

$$\Pr \left(\sup_{0 \leq j < N} |M_i^{a,1}(j)| \geq \varepsilon_1 + \delta, C \right) \leq 2 \exp \left(-\frac{\varepsilon_1^2}{2NK_r^2} \right)$$

When C is true, we have that, for any t such that $T_j \leq t < T_{j+1}$ for some $0 \leq j < N$,

$$|M_i^{a,1}(t) - M_i^{a,1}(j)| \leq \frac{\lfloor c \log n \rfloor}{n} + \sum_{s=0}^{\lfloor c \log n \rfloor - 1} \tau_{j, s+1} \leq K_r$$

almost surely. Hence

$$\Pr \left(\sup_{t \in [0, T_{\lfloor N \rfloor \lfloor c \log n \rfloor}]} |M_i^{a,1}(t)| \geq \varepsilon_1 + K_r + \delta, C \right) \leq 2 \exp \left(-\frac{\varepsilon_1^2}{2NK_r^2} \right)$$

For the bad event C^c we can bound

$$\begin{aligned} & \Pr(C^c) \\ & \leq \Pr \left(\exists 0 \leq j < N, \sum_{s=0}^{\lfloor c \log n \rfloor - 1} \tau_{j, \lfloor c \log n \rfloor + s+1} \geq \frac{r \log n (\lfloor c \log n \rfloor + 1)}{\lambda n} \right) \\ & \leq \sum_{j=0}^{\lfloor N \rfloor} \Pr \left(\sum_{s=0}^{\lfloor c \log n \rfloor - 1} \tau_{j, \lfloor c \log n \rfloor + s+1} \geq \frac{r \log n (\lfloor c \log n \rfloor + 1)}{\lambda n} \right) \\ & \stackrel{(*)}{\leq} \lfloor N \rfloor \cdot e n^{-r} \leq e N n^{-r} \end{aligned}$$

where in $(*)$ we used Lemma 12.

Finally, under event $\mathcal{A}$, $T_{\lfloor N \rfloor \lfloor c \log n \rfloor} \geq T$. Thus

$$\begin{aligned} & \Pr \left(\sup_{t \in [0, T]} |M_i^{a,1}(t)| \geq \varepsilon_1 + K_r + \delta, \mathcal{A} \right) \\ & \leq \Pr \left(\sup_{t \in [0, T_{\lfloor N \rfloor \lfloor c \log n \rfloor}]} |M_i^{a,1}(t)| \geq \varepsilon_1 + K_r + \delta, \mathcal{A}, C \right) + \Pr(C^c) \\ & \leq 2 \exp \left(-\frac{\varepsilon_1^2}{2NK_r^2} \right) + e N n^{-r} \end{aligned}$$

□

A.6 Proof of Lemma 5

We will need the following large deviation result in this proof, which we state here for completeness.

LEMMA 13 (COROLLARY 1 OF [28]). *Let $\{Z_j\}_{j=1}^\infty$ be an adapted process w.r.t. the filtration $\{\mathcal{F}_j\}_{j=1}^\infty$. If*

$$0 \leq Z_j \leq B \quad a.s.$$

and

$$\mathbb{E}[Z_j \mid \mathcal{F}_{j-1}] \leq m \quad a.s.$$

then for any $\lambda \geq 2Nm$, we have

$$\Pr \left(\sum_{j=1}^N Z_j \geq \lambda \right) \leq \exp \left(-\frac{3\lambda}{16B} \right)$$

By assumption, $\lim_{n \rightarrow \infty} \|\mathbf{X}^{(n)}(0) - \mathbf{x}(0)\|_1 = 0$, hence, for sufficiently large n , we have

$$\frac{1}{n} \sum_{i=1}^n Q_i^{(n)}(0) = \|\mathbf{X}^{(n)}(0)\|_1 \leq \|\mathbf{x}(0)\|_1 + 1$$

Under event $\mathcal{B}$, the queues with length at least $(\rho + \kappa L) \log n$ at some time before the $\lfloor e\lambda Tn \rfloor$ -th arrival must have an initial length of at least $\rho \log n$. Hence

$$\begin{aligned} \sup_{0 \leq j < e\lambda Tn} X_{\lceil (\rho + \kappa L) \log n \rceil \{j\}} &\leq X_{\lceil \rho \log n \rceil \{0\}} \\ &\leq \frac{\frac{1}{n} \sum_{i=0}^n Q_i^{(n)}(0)}{\rho \log n} \leq \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n} \end{aligned}$$

where the first inequality is just Markov Inequality.

Let $A_{\geq b}(t) := \sum_{i=b+1}^{\infty} A_i(t)$ to be the number of arrivals dispatched to queues with length at least b before time t . Let $\tilde{A}_{\geq b}(t)$ be the number of arrivals satisfying the following conditions before time t :

- The assigned queue has length at least b at the arrival time.
- Let τ be the last reset time before/including this arrival, we require

$$X_{\lceil (\rho + \kappa L) \log n \rceil}(\tau) \leq \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n}$$

Denote $\tilde{A}_{\geq b}\langle j \rangle := \tilde{A}_{\geq b}(T_j \lfloor c \log n \rfloor -)$. Set $b := \lceil (\rho + \kappa L) \log n + L \rceil$. Using the same proof for Lemma 4, Eq. 10 in [28] we have

$$\mathbb{E}[\tilde{A}_{\geq b}\langle j+1 \rangle - \tilde{A}_{\geq b}\langle j \rangle \mid \mathcal{G}_j] \leq \lfloor c \log n \rfloor \left(\frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n} \right)^d$$

We further have

$$\tilde{A}_{\geq b}\langle j+1 \rangle - \tilde{A}_{\geq b}\langle j \rangle \leq \lfloor c \log n \rfloor \quad a.s.$$

Set $\varphi = 2\lceil N \rceil \lfloor c \log n \rfloor \left(\frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n} \right)^d$. Using Lemma 13 we obtain

$$\Pr \left(\tilde{A}_{\geq b}\langle \lceil N \rceil \rangle \geq \varphi \right) \leq \exp \left(-\frac{3\varphi}{16\lfloor c \log n \rfloor} \right)$$

Under event $\mathcal{B}$, we have $A_{\geq b}(t) = \tilde{A}_{\geq b}(t)$ for $t \in [0, T]$, and $T_{\lceil N \rceil \lfloor c \log n \rfloor} \geq T$. Hence

$$\Pr \left(\sup_{t \in [0, T]} \sum_{i=b+1}^{\infty} A_i(t) \geq \varphi, \mathcal{B} \right) \leq \exp \left(-\frac{3\varphi}{16\lfloor c \log n \rfloor} \right) \quad (23)$$

Under event $\mathcal{B}$, we also have

$$\bar{X}_{b+1}(t) \leq X_{\lceil \rho \log n + L \rceil} \{0\} \leq \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \quad \forall t \in [0, T]$$

which implies that

$$\begin{aligned}
 & \sup_{t \in [0, T]} \sum_{i=b+1}^{\infty} \int_0^t a_i(\bar{X}(u)) du \\
 &= \sup_{t \in [0, T]} \int_0^t \lambda(\bar{X}_{b+1}(u))^d du \\
 &\leq \lambda T \left(\frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right)^d
 \end{aligned} \tag{24}$$

We also know that

$$\begin{aligned}
 & \sup_{t \in [0, T]} \sum_{i=b+1}^{\infty} |M_i^{a,1}(t)| \\
 &\leq \sup_{t \in [0, T]} \sum_{i=b+1}^{\infty} \frac{A_i^{a,1}(t)}{n} + \sup_{t \in [0, T]} \sum_{i=b+1}^{\infty} \int_0^t a_i(\bar{X}(u)) du
 \end{aligned} \tag{25}$$

Combining (23)(24)(25) we obtain

$$\begin{aligned}
 & \Pr \left(\sup_{t \in [0, T]} \sum_{i=b+1}^{\infty} |M_i^{a,1}(t)| \geq \frac{\varphi}{n} + \lambda T \left(\frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right)^d \right) \\
 &\leq \exp \left(-\frac{3\varphi}{16 \lfloor c \log n \rfloor} \right)
 \end{aligned}$$

□

A.7 Proof of Lemma 6

The service times are exponentially distributed with parameter 1. Then for any state $\mathbf{x}$ the transition rate from $\mathbf{x}$ to $\mathbf{x} - \frac{1}{n} \mathbf{e}^{(i)}$ (which indicates a service completion at a queue of length i) is equal to the number of queues of length i , which is $nb_i(\mathbf{x})$. Hence we have the following random time change representation

$$\mathbf{B}(t) = \sum_{i=1}^{\infty} \mathbf{e}^{(i)} N_i \left(n \int_0^t b_i(\mathbf{X}(u)) du \right)$$

where $N_i(\cdot)$, $i = 1, 2, \dots$ are independent unit rate Poisson processes.

Using this representation we have

$$\mathbf{M}^b(t) = \frac{1}{n} \sum_{i=1}^{\infty} \mathbf{e}^{(i)} \bar{N}_i \left(n \int_0^t b_i(\mathbf{X}(u)) du \right)$$

where $\bar{N}_i$ is the centered Poisson process defined as

$$\bar{N}_i(s) := N_i(s) - s$$

Now, we will use the following large deviation result for centered Poisson processes:

LEMMA 14 (PROPOSITION 5.2 OF [9]). *Let $\bar{N}(t)$ be a centered Poisson process of unit rate, then*

$$\Pr \left(\sup_{t \in [0, M]} |\bar{N}(t)| \geq \varepsilon \right) \leq 2e^{-Mh(\varepsilon/M)}$$

where $h(t) = (1+t)\log(1+t) - t$.

Using the lemma we obtain that for any $i \geq 1$ and any $\varepsilon_3 > 0$,

$$\begin{aligned}
 & \Pr \left(\sup_{t \in [0, T]} |M_i^b(t)| \geq \varepsilon_3 \right) \\
 &= \Pr \left(\sup_{t \in [0, T]} \left| \bar{N}_i \left(n \int_0^t b_i(\mathbf{X}(u)) du \right) \right| \geq n\varepsilon_3 \right) \\
 &\leq \Pr \left(\sup_{t' \in [0, nT]} \left| \bar{N}_i(t') \right| \geq n\varepsilon_3 \right) \\
 &\leq 2e^{-nTh(\varepsilon_3/T)}
 \end{aligned} \tag{26}$$

Now, assuming that event $\mathcal{B}$ is true, we have

$$\sup_{t \in [0, T]} X_{b+1}(t) \leq X_{\lceil \rho \log n + L \rceil}(0) \leq \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \tag{27}$$

(See Appendix A.6 for derivation of (27).)

Consequently, for $t \in [0, T]$,

$$\begin{aligned}
 & \sum_{i=b+1}^{\infty} \int_0^t b_i(\mathbf{X}(u)) du = \int_0^t \sum_{i=b+1}^{\infty} b_i(\mathbf{X}(u)) du \\
 &= \int_0^t X_{b+1}(u) du \leq T \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L}
 \end{aligned}$$

Therefore

$$\begin{aligned}
 & \Pr \left(\sup_{t \in [0, T]} \sum_{i=b+1}^{\infty} |M_i^b(t)| \geq (e+1)T \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L}, \mathcal{B} \right) \\
 &\leq \Pr \left(\sup_{t \in [0, T]} \frac{1}{n} \sum_{i=b+1}^{\infty} N_i \left(n \int_0^t b_i(\mathbf{X}(u)) du \right) \geq eT \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L}, \mathcal{B} \right) \\
 &= \Pr \left(\sup_{t \in [0, T]} N_1 \left(\sum_{i=b+1}^{\infty} n \int_0^t b_i(\mathbf{X}(u)) du \right) \geq enT \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L}, \mathcal{B} \right) \\
 &\leq \Pr \left(\sup_{t \in [0, T]} N_1 \left(nT \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right) \geq enT \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right) \\
 &\leq \exp \left(-nT \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right)
 \end{aligned} \tag{28}$$

Combining (26)(28) we obtain

$$\begin{aligned}
& \Pr \left(\sup_{t \in [0, T]} \|\mathbf{M}^b(t)\|_1 \geq b\varepsilon_3 + \frac{(e+1)T\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L}, \mathcal{B} \right) \\
& \leq \sum_{i=1}^b \Pr \left(\sup_{t \in [0, T]} |M_i^b(t)| \geq \varepsilon_3 \right) \\
& \quad + \Pr \left(\sup_{t \in [0, T]} \sum_{i=b+1}^{\infty} |M_i^b(t)| \geq (e+1)T \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L}, \mathcal{B} \right) \\
& \leq 2be^{-nTh(\varepsilon_3/T)} + \exp \left(-nT \frac{\|\mathbf{x}(0)\|_1 + 1}{\rho \log n + L} \right)
\end{aligned}$$

A.8 Proof of Proposition 2

Assume that $\mathbf{Q}(0) \stackrel{d}{\sim} \check{\mathbf{Q}}$. By PASTA, we know that for all $\mathbf{q} \in \mathbb{Z}_+^n$,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbb{1}_{\{\mathbf{Q}(t)=\mathbf{q}\}} dt = \lim_{J \rightarrow \infty} \frac{1}{J} \sum_{j=0}^{J-1} \mathbb{1}_{\{\mathbf{Q}\{j\}=\mathbf{q}\}}$$

Since $(\mathbf{Q}(t), \overline{\mathbf{W}}(t), \Xi(t))_{t \geq 0}$ is an irreducible and ergodic countinuous time Markov process, we have

$$\Pr(\hat{\mathbf{Q}} = \mathbf{q}) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbb{1}_{\{\mathbf{Q}(t)=\mathbf{q}\}} dt$$

We also know that $\frac{1}{J} \sum_{j=0}^{J-1} \mathbb{1}_{\{\mathbf{Q}\{j\}=\mathbf{q}\}}$ converges a.s. to a number since $(\mathbf{Q}\{j\}, \overline{\mathbf{W}}\{j\}, \Xi\{j\})_{j=0}^{\infty}$ is an irreducible and ergodic Markov Chain.

Consider $\mathbb{E} \left[\frac{1}{J} \sum_{j=0}^{J-1} \mathbb{1}_{\{\mathbf{Q}\{j\}=\mathbf{q}\}} \right]$. Observe that since we start from $\check{\mathbf{Q}}$, the stationary distribution of $\mathbf{Q}\{s \lfloor c \log n \rfloor\}$, we have $\mathbf{Q}\{s \lfloor c \log n \rfloor\} \stackrel{d}{\sim} \check{\mathbf{Q}}$ for all $s \geq 0$ and hence $\mathbf{Q}\{j\} \stackrel{d}{\sim} \mathbf{Q}\{j \bmod \lfloor c \log n \rfloor\}$ for all j .

We have

$$\begin{aligned}
& \mathbb{E} \left[\frac{1}{J} \sum_{j=0}^{J-1} \mathbb{1}_{\{\mathbf{Q}\{j\}=\mathbf{q}\}} \right] \\
& = \frac{1}{J} \sum_{j=0}^{J-1} \Pr(\mathbf{Q}\{j\} = \mathbf{q}) \\
& = \frac{1}{J} \sum_{j=0}^{J-1} \Pr(\mathbf{Q}\{j \bmod \lfloor c \log n \rfloor\} = \mathbf{q}) \\
& \xrightarrow{J \rightarrow \infty} \frac{1}{\lfloor c \log n \rfloor} \sum_{j=0}^{\lfloor c \log n \rfloor - 1} \Pr(\mathbf{Q}\{j\} = \mathbf{q}) \\
& = \Pr(\mathbf{Q}\{U\} = \mathbf{q})
\end{aligned}$$

where U is a uniform random variable on $\{0, 1, \dots, \lfloor c \log n \rfloor - 1\}$.

By the Bounded Convergence Theorem, we have

$$\begin{aligned} & \mathbb{E} \left[\lim_{J \rightarrow \infty} \frac{1}{J} \sum_{j=0}^{J-1} \mathbb{1}_{\{Q\{j\}=\mathbf{q}\}} \right] \\ &= \lim_{J \rightarrow \infty} \mathbb{E} \left[\frac{1}{J} \sum_{j=0}^{J-1} \mathbb{1}_{\{Q\{j\}=\mathbf{q}\}} \right] = \Pr(Q\{U\} = \mathbf{q}) \end{aligned}$$

Hence we conclude that

$$\Pr(\hat{\mathbf{Q}} = \mathbf{q}) = \Pr(Q\{U\} = \mathbf{q})$$

which means that $Q\{U\} \stackrel{d}{\sim} \hat{\mathbf{Q}}$.

Hence

$$\begin{aligned} & d_{W_1, \|\cdot\|_1}(\mathcal{L}(\hat{\mathbf{X}}), \mathcal{L}(\check{\mathbf{X}})) \\ &= \inf \{ \mathbb{E}[\|\hat{\mathbf{Y}} - \check{\mathbf{Y}}\|_1] \mid \hat{\mathbf{Y}} \stackrel{d}{\sim} \hat{\mathbf{X}}, \check{\mathbf{Y}} \stackrel{d}{\sim} \check{\mathbf{X}} \} \\ &\leq \mathbb{E}[\|\mathbf{X}\{U\} - \mathbf{X}\{0\}\|_1] \leq \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n |Q_i\{U\} - Q_i\{0\}| \right] \\ &\leq \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \sum_{j=0}^{U-1} (R_i\{j\} + S_i\{j\}) \right] \\ &= \frac{1}{n} \mathbb{E} \left[U + \sum_{i=1}^n \sum_{j=0}^{U-1} S_i\{j\} \right] \\ &= \frac{1}{n} \left(\frac{\lfloor c \log n \rfloor - 1}{2} + \frac{1}{\lambda} \frac{\lfloor c \log n \rfloor - 1}{2} \right) \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

□

A.9 Proof of Lemma 10

Recall that $\check{\mathbf{X}}$ is the queue length proportion vector for queue length vector $\check{\mathbf{Q}}$, where $\check{\mathbf{Q}}$ is distributed according to the stationary distribution of the Markov Chain $(Q\{j \lfloor c \log n \rfloor\})_{j=0}^{\infty}$.

Define a quadratic Lyapunov function $V^b : \mathbb{Z}_+^n \mapsto \mathbb{R}_+$

$$V^b(\mathbf{q}) := \sum_{i=1}^n v^b(q_i) := \sum_{i=1}^n (q_i - b)_+^2$$

LEMMA 15. For any $x, y \in \mathbb{R}$, $(x + y)_+^2 \leq ((x)_+ + y)^2$.

PROOF OF LEMMA 15. • If $x + y \leq 0$, then $(x + y)_+^2 = 0$ and the claim is trivially true.

- If $x \geq 0$, then $(x + y)_+^2 \leq (x + y)^2 = ((x)_+ + y)^2$.
- Otherwise $x < 0$ and $x + y > 0$, which implies that $0 < x + y < y$, hence

$$(x + y)_+^2 = (x + y)^2 \leq y^2 = ((x)_+ + y)^2$$

□

Now, for any $j \geq 0$ and $i \in [n]$ we have

$$\begin{aligned}
 & v^b(Q_i\{j+1\}) - v^b(Q_i\{j\}) \\
 &= (Q_i\{j\} - b + R_i\{j\} - S_i\{j\})_+^2 - (Q_i\{j\} - b)_+^2 \\
 &\leq ((Q_i\{j\} - b)_+ + R_i\{j\} - S_i\{j\})_+^2 - (Q_i\{j\} - b)_+^2 \\
 &\leq 2(Q_i\{j\} - b)_+(R_i\{j\} - S_i\{j\}) + (R_i\{j\} - S_i\{j\})^2
 \end{aligned}$$

Notice that if $Q_i\{j\} \leq b-1$, then $v^b(Q_i\{j+1\}) = v^b(Q_i\{j\}) = 0$. Therefore

$$\begin{aligned}
 & v^b(Q_i\{j+1\}) - v^b(Q_i\{j\}) \\
 &\leq 2(Q_i\{j\} - b)_+(R_i\{j\} - S_i\{j\}) + (R_i\{j\} - S_i\{j\})^2 \cdot \mathbb{1}_{\{Q_i\{j\} \geq b\}}
 \end{aligned} \tag{29}$$

Summing (29) up we obtain

$$\begin{aligned}
 & V^b(Q\{j+1\}) - V^b(Q\{j\}) \\
 &\leq 2 \sum_{i=1}^n (Q_i\{j\} - b)_+(R_i\{j\} - S_i\{j\}) \\
 &\quad + \sum_{i=1}^n (R_i\{j\} - S_i\{j\})^2 \cdot \mathbb{1}_{\{Q_i\{j\} \geq b\}} \\
 &= -2 \sum_{i=1}^n (Q_i\{j\} - b)_+ S_i\{j\} + \sum_{i=1}^n (S_i\{j\})^2 \cdot \mathbb{1}_{\{Q_i\{j\} \geq b\}} \\
 &\quad + 2(Q_{i^*}\{j\} - b)_+ + (1 - 2S_{i^*}\{j\}) \mathbb{1}_{\{Q_{i^*}\{j\} \geq b\}} \\
 &\leq -2 \sum_{i=1}^n (Q_i\{j\} - b)_+ S_i\{j\} + \sum_{i=1}^n (S_i\{j\})^2 \cdot \mathbb{1}_{\{Q_i\{j\} \geq b\}} \\
 &\quad + 2(Q_{W_1\{j\}}\{j\} - b)_+ + (1 - 2S_{i^*}\{j\}) \mathbb{1}_{\{Q_{i^*}\{j\} \geq b\}}
 \end{aligned}$$

where i^* is the queue that the $(j+1)$ -th job is dispatched to.

Now, assume that $0 \leq j < \lfloor c \log n \rfloor$. Recall that a server can be visited by the random walkers for at most L times. Hence $\mathbb{1}_{\{Q_i\{0\} \geq b-L\}} \geq \mathbb{1}_{\{Q_i\{j\} \geq b\}}$

Therefore

$$\begin{aligned}
 & \mathbb{E}[V^b(Q\{j+1\}) - V^b(Q\{j\})] \\
 &\leq -2\mathbb{E}\left[\sum_{i=1}^n (Q_i\{j\} - b)_+ S_i\{j\}\right] + \sum_{i=1}^n \mathbb{E}\left[(S_i\{j\})^2 \cdot \mathbb{1}_{\{Q_i\{0\} \geq b-L\}}\right] \\
 &\quad + \mathbb{E}\left[(Q_{W_1\{j\}}\{j\} - b)_+\right] + \mathbb{E}\left[(1 - 2S_{i^*}\{j\}) \mathbb{1}_{\{Q_{i^*}\{j\} \geq b\}}\right] \\
 &= -\frac{2}{\lambda n} \mathbb{E}\left[\sum_{i=1}^n (Q_i\{j\} - b)_+\right] + \left(\frac{1}{\lambda n} + \frac{2}{\lambda^2 n^2}\right) \sum_{i=1}^n \mathbb{E}\left[\mathbb{1}_{\{Q_i\{0\} \geq b-L\}}\right] \\
 &\quad + \mathbb{E}\left[(Q_{W_1\{j\}}\{j\} - b)_+\right] + \left(1 - \frac{2}{\lambda n}\right) \mathbb{E}\left[\mathbb{1}_{\{Q_{i^*}\{j\} \geq b\}}\right]
 \end{aligned} \tag{30}$$

Using the fact that $(x+y)_+ \leq (x)_+ + (y)_+ \mathbb{1}_{\{x+y \geq 0\}}$ we have

$$\begin{aligned}
 (Q_{W_1\{j\}}\{j\} - b)_+ &\leq (Q_{W_1\{j\}}\{0\} + L - b)_+ \\
 &\leq (Q_{W_1\{j\}}\{0\} - b)_+ + L \mathbb{1}_{\{Q_{W_1\{j\}}\{0\} \geq b-L\}}
 \end{aligned}$$

Using the fact that $W_1\{j\}$ is independent of $Q\{0\}$, we have

$$\begin{aligned}
 & \mathbb{E}[(Q_{W_1\{j\}}\{j\} - b)_+] \\
 & \leq \mathbb{E}[(Q_{W_1\{j\}}\{0\} - b)_+] + L\mathbb{E}[\mathbb{1}_{\{Q_{W_1\{j\}}\{0\} \geq b-L\}}] \\
 & = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[(Q_i\{0\} - b)_+] + \frac{L}{n} \sum_{i=1}^n \mathbb{E}[\mathbb{1}_{\{Q_i\{0\} \geq b-L\}}]
 \end{aligned} \tag{31}$$

As before we have

$$\begin{aligned}
 & \sum_{i=1}^n (Q_i\{j\} - b)_+ \\
 & \geq \sum_{i=1}^n \left(Q_i\{0\} - \sum_{s=0}^{j-1} S_i\{s\} - b \right)_+ \\
 & \geq \sum_{i=1}^n (Q_i\{0\} - b)_+ - \sum_{i=1}^n \sum_{s=0}^{j-1} S_i\{s\} \mathbb{1}_{\{Q_i\{0\} \geq b-L\}}
 \end{aligned}$$

thus

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i=1}^n (Q_i\{j\} - b)_+ \right] \\
 & \geq \mathbb{E} \left[\sum_{i=1}^n (Q_i\{0\} - b)_+ \right] - \mathbb{E} \left[\sum_{i=1}^n \sum_{s=0}^{j-1} S_i\{s\} \mathbb{1}_{\{Q_i\{0\} \geq b-L\}} \right] \\
 & = \mathbb{E} \left[\sum_{i=1}^n (Q_i\{0\} - b)_+ \right] - \frac{j}{\lambda n} \mathbb{E} \left[\sum_{i=1}^n \mathbb{1}_{\{Q_i\{0\} \geq b-L\}} \right]
 \end{aligned} \tag{32}$$

We also have

$$\begin{aligned}
 & \mathbb{E} \left[\mathbb{1}_{\{Q_{i^*}\{j\} \geq b\}} \right] \leq \mathbb{E} \left[\mathbb{1}_{\{Q_{W_1\{j\}}\{j\} \geq b\}} \right] \\
 & \leq \mathbb{E} \left[\mathbb{1}_{\{Q_{W_1\{j\}}\{0\} \geq b-L\}} \right] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\mathbb{1}_{\{Q_i\{0\} \geq b-L\}}]
 \end{aligned} \tag{33}$$

Combining (30)(31)(32)(33), for $n \geq \frac{2}{\lambda}$ we obtain

$$\begin{aligned}
 & \mathbb{E}[V^b(Q\{j+1\}) - V^b(Q\{j\})] \\
 & \leq -\frac{2}{\lambda n} \mathbb{E} \left[\sum_{i=1}^n (Q_i\{0\} - b)_+ \right] + \frac{2j}{\lambda^2 n^2} \mathbb{E} \left[\sum_{i=1}^n \mathbb{1}_{\{Q_i\{0\} \geq b-L\}} \right] \\
 & \quad + \left(\frac{1}{\lambda n} + \frac{2}{\lambda^2 n^2} \right) \sum_{i=1}^n \mathbb{E} \left[\mathbb{1}_{\{Q_i\{0\} \geq b-L\}} \right] + \frac{2}{n} \sum_{i=1}^n \mathbb{E}[(Q_i\{0\} - b)_+] \\
 & \quad + \frac{2L}{n} \sum_{i=1}^n \mathbb{E}[\mathbb{1}_{\{Q_i\{0\} \geq b-L\}}] + \left(1 - \frac{2}{\lambda n} \right) \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\mathbb{1}_{\{Q_i\{0\} \geq b-L\}}] \\
 & = -2 \left(\frac{1}{\lambda} - 1 \right) \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (Q_i\{0\} - b)_+ \right]
 \end{aligned}$$

$$+ \left(\frac{2j}{\lambda^2 n} + \frac{1}{\lambda} + \frac{2}{\lambda^2 n} + 2L + 1 - \frac{2}{\lambda n} \right) \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Q_i\{0\} \geq b-L\}} \right]$$

Summing over $j = 0, 2, \dots, \lfloor c \log n \rfloor - 1$ results in

$$\begin{aligned} & \mathbb{E}[V^b(Q\{\lfloor c \log n \rfloor\}) - V^b(Q\{0\})] \\ & \leq -2\lfloor c \log n \rfloor \left(\frac{1}{\lambda} - 1 \right) \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (Q_i\{0\} - b)_+ \right] \\ & \quad + \lfloor c \log n \rfloor \left(\frac{\lfloor c \log n \rfloor - 1}{\lambda^2 n} + \frac{1}{\lambda} + \frac{2}{\lambda^2 n} + 2L + 1 - \frac{2}{\lambda n} \right) \\ & \quad \times \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Q_i\{0\} \geq b-L\}} \right] \end{aligned}$$

Set $Q\{0\} \stackrel{d}{\sim} \check{Q}$, we will have $Q\{\lfloor c \log n \rfloor\} \stackrel{d}{\sim} \check{Q}$. Hence

$$\begin{aligned} 0 &= \mathbb{E}[V^b(Q\{\lfloor c \log n \rfloor\}) - V^b(Q\{0\})] \\ & \leq -2\lfloor c \log n \rfloor \left(\frac{1}{\lambda} - 1 \right) \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (\check{Q}_i - b)_+ \right] \\ & \quad + \lfloor c \log n \rfloor \left(\frac{\lfloor c \log n \rfloor - 1}{\lambda^2 n} + \frac{1}{\lambda} + \frac{2}{\lambda^2 n} + 2L + 1 - \frac{2}{\lambda n} \right) \\ & \quad \times \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\check{Q}_i \geq b-L\}} \right] \end{aligned}$$

which implies

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (\check{Q}_i - b)_+ \right] \\ & \leq \frac{\frac{\lfloor c \log n \rfloor - 1}{\lambda^2 n} + \frac{1}{\lambda} + \frac{2}{\lambda^2 n} + 2L + 1 - \frac{2}{\lambda n}}{2(\frac{1}{\lambda} - 1)} \cdot \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\check{Q}_i \geq b-L\}} \right] \\ & \leq \kappa_5 \cdot \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\check{Q}_i \geq b-L\}} \right] \end{aligned} \tag{34}$$

for some constant $\kappa_5 > 0$ for all $n \geq \frac{2}{\lambda}$.

Now, by Markov Inequality and Corollary 3, for $b > L$ we have

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\check{Q}_i \geq b-L\}} \right] \leq \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \frac{\check{Q}_i}{b-L} \right] \\ & = \frac{1}{b-L} \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \check{Q}_i \right] \leq \frac{1}{b-L} \left(\frac{1 + \lambda + 2\lambda L}{2(1-\lambda)} + \varepsilon \right) \end{aligned} \tag{35}$$

for $n \geq \bar{n}(\varepsilon)$.

Hence set $\bar{n} = \bar{n}(\varepsilon) \vee \frac{2}{\lambda}$, combining (34)(35) we have

$$\sup_{n \geq \bar{n}} \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (\check{Q}_i - b)_+ \right] \leq \frac{1}{b-L} \cdot \kappa_5 \left(\frac{1 + \lambda + 2\lambda L}{2(1-\lambda)} + \varepsilon \right) \\ \xrightarrow{b \rightarrow \infty} 0$$

Observe that

$$\sum_{i=b+1}^{\infty} \check{X}_i = \frac{1}{n} \sum_{i=1}^n (\check{Q}_i - b)_+$$

We conclude

$$\lim_{b \rightarrow \infty} \sup_{n \geq \bar{n}} \mathbb{E} \left[\sum_{i=b+1}^{\infty} \check{X}_i \right] = 0$$

□

A.10 Proof of Proposition 3

By the assumption, $Y(0) \in \ell_1([0, 1])$ a.s. Since $\ell_1([0, 1]) \subset [0, 1]^{\mathbb{Z}_+}$ is separable with respect to the $\|\cdot\|_1$ norm, we conclude that Y has a separable support. By Skorokhod representation theorem, without loss of generality, we can assume that

$$\lim_{k \rightarrow \infty} \|X^{(n)}(0) - Y(0)\|_1 \xrightarrow{n \rightarrow \infty} 0 \quad a.s.$$

(The random vectors $(Q^{(n)}(0), \bar{W}^{(n)}(0), \Xi^{(n)}(0))_n$ are constructed accordingly given $(X^{(n)})_n$.)

Let $\bar{\tau}^{(n)} := \min\{t \geq 0 : \Xi^{(n)}(t) = 0\}$.

There are at most $\lfloor c \log n \rfloor - 1$ arrivals before $\bar{\tau}^{(n)}$, hence we can bound

$$\sup_{0 \leq t \leq \bar{\tau}^{(n)}} \|X^{(n)}(t) - X^{(n)}(0)\|_1 \\ \leq \frac{\lfloor c \log n \rfloor - 1}{n} + \frac{1}{n} \sum_{l=1}^{\lfloor c \log n \rfloor - 1} S_l^{(n)} \xrightarrow{n \rightarrow \infty} 0 \quad a.s.$$

where $S_l^{(n)}, l = 1, 2, \dots$ are i.i.d. Geometric random variables with mean $\frac{1}{\lambda}$. Almost sure convergence can be shown from large deviation estimates and the Borel-Cantelli Lemma.

Along with the assumption that $\lim_{n \rightarrow \infty} \|X^{(n)}(0) - Y(0)\|_1 = 0$ almost surely, we have

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq \bar{\tau}^{(n)}} \|X^{(n)}(t) - Y(0)\|_1 = 0 \quad a.s. \quad (36)$$

In particular, we have

$$\lim_{n \rightarrow \infty} \|X^{(n)}(\bar{\tau}^{(n)}) - Y(0)\|_1 = 0 \quad a.s. \quad (37)$$

Viewing the system as starting from time $\bar{\tau}^{(n)}$, WLOG assume that the systems evolves independently for different n , applying Theorem 1 we obtain

$$\Pr \left(\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} \|X^{(n)}(t + \bar{\tau}^{(n)}) - Y(t)\|_1 = 0 \mid (X^{(n)}(\bar{\tau}^{(n)}))_n \right) = 1$$

for ω 's such that $[X^{(n)}(\bar{\tau})](\omega)$ converges in ℓ_1 to $Y(0)(\omega)$ and $\|Y(0)(\omega)\|_1 < +\infty$.

By (37), $\mathbf{X}^{(n)}(\bar{\tau})(\omega)$ converges in ℓ_1 to $\mathbf{Y}(0)(\omega)$ and $\|\mathbf{Y}(0)(\omega)\|_1 < +\infty$ for almost all ω , hence we have

$$\Pr \left(\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} \|\mathbf{X}^{(n)}(t + \bar{\tau}^{(n)}) - \mathbf{Y}(t)\|_1 = 0 \right) = 1$$

which means that

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} \|\mathbf{X}^{(n)}(t + \bar{\tau}^{(n)}) - \mathbf{Y}(t)\|_1 = 0 \quad a.s. \quad (38)$$

Now we need a lemma stating the continuity of the limiting dynamics.

LEMMA 16. For any $s_1, s_2 > 0$, $\|\mathbf{x}(s_1) - \mathbf{x}(s_2)\| \leq (\lambda + 1)|s_1 - s_2|$.

PROOF. Assume that $s_1 < s_2$. We know that

$$\mathbf{x}(s_2) - \mathbf{x}(s_1) = \int_{s_1}^{s_2} [\mathbf{a}(\mathbf{x}(u)) - \mathbf{b}(\mathbf{x}(u))] du$$

Hence

$$\begin{aligned} \|\mathbf{x}(s_1) - \mathbf{x}(s_2)\|_1 &\leq \int_{s_1}^{s_2} [\|\mathbf{a}(\mathbf{x}(u))\|_1 + \|\mathbf{b}(\mathbf{x}(u))\|_1] du \\ &= \int_{s_1}^{s_2} (\lambda x_0^d + x_1) du \\ &\leq \int_{s_1}^{s_2} (\lambda + 1) du = (\lambda + 1)(s_2 - s_1) \end{aligned}$$

□

Given the lemma, we have

$$\sup_{0 \leq t \leq T} \|\mathbf{x}(t) - \mathbf{x}((t - \bar{\tau}^{(n)})_+)\|_1 \leq (\lambda + 1)\bar{\tau}^{(n)} \quad (39)$$

where we have

$$\bar{\tau}^{(n)} \leq \sum_{j=1}^{\lfloor c \log n \rfloor} \tau_j^{(n)} \xrightarrow[n \rightarrow \infty]{a.s.} 0$$

where $\tau_j^{(n)}, j = 1, 2, \dots$ are i.i.d. exponential random variables with mean $\frac{1}{\lambda n}$ and almost sure convergence follows from the large deviation estimate in Lemma 12 and Borel-Cantelli Lemma.

Combining (36)(38)(39) we obtain

$$\begin{aligned} &\sup_{0 \leq t \leq T} \|\mathbf{X}^{(n)}(t) - \mathbf{Y}(t)\|_1 \\ &\leq \sup_{0 \leq t \leq T} \|\mathbf{X}^{(n)}(t) - \mathbf{Y}((t - \bar{\tau}^{(n)})_+)\|_1 + \sup_{0 \leq t \leq T} \|\mathbf{Y}(t) - \mathbf{Y}((t - \bar{\tau}^{(n)})_+)\|_1 \\ &\leq \max \left\{ \sup_{0 \leq t \leq \bar{\tau}^{(n)}} \|\mathbf{X}^{(n)}(t) - \mathbf{Y}(0)\|_1, \sup_{\bar{\tau}^{(n)} \leq t \leq T + \bar{\tau}^{(n)}} \|\mathbf{X}^{(n)}(t) - \mathbf{Y}(t - \bar{\tau}^{(n)})\|_1 \right\} \\ &\quad + \sup_{0 \leq t \leq T} \|\mathbf{Y}(t) - \mathbf{Y}((t - \bar{\tau}^{(n)})_+)\|_1 \\ &= \max \left\{ \sup_{0 \leq t \leq \bar{\tau}^{(n)}} \|\mathbf{X}^{(n)}(t) - \mathbf{Y}(0)\|_1, \sup_{0 \leq t \leq T} \|\mathbf{X}^{(n)}(t + \bar{\tau}^{(n)}) - \mathbf{Y}(t)\|_1 \right\} \\ &\quad + \sup_{0 \leq t \leq T} \|\mathbf{Y}(t) - \mathbf{Y}((t - \bar{\tau}^{(n)})_+)\|_1 \end{aligned}$$

$$\xrightarrow[n \rightarrow \infty]{a.s.} 0,$$

and in particular,

$$\mathbf{X}^{(n)}(t) \xrightarrow[n \rightarrow \infty]{} \mathbf{Y}(t) \quad \forall t \geq 0$$

B ADDITIONAL SIMULATION RESULTS

Besides the statistics of queue lengths, i.e. the number of queues with certain length, we are also interested in the evolution of individual queues. More specifically, we are interested in whether the lengths of two arbitrary queues are approximately independent under the stationary distribution of the system, where asymptotic independence is true for *power-of- d* choices [23, 29].

We conduct simulations on cycle graphs with $n = 12180$ vertices and $n = 60900$ vertices respectively. The restart period is set to 50, and $d = 2$, $\lambda = 0.95$. For each graph, we start the system from empty and run the system for 1000 units of time. We then look at the queue lengths of two adjacent vertices within time $[500, 1000]$. We then obtain an empirical joint distribution of the two queue lengths using the observed sample path within time $[500, 1000]$. The results shows that there exist significant positive correlation between the lengths of two adjacent queues on the graph. However, this correlation becomes weaker for large n . One of the future work is to identify if propagation of chaos happens in our system.

	0	1	2	3	4	5	6	avg
0	0.2447	0.2325	0.2460	0.1249	0.1265	0.0252	0	1.7318
1	0.1231	0.2338	0.2089	0.1849	0.1449	0.1034	0.0012	2.3096
2	0.0996	0.1417	0.2503	0.2084	0.1741	0.1209	0.0050	2.5982
3	0.0408	0.1083	0.1993	0.2370	0.2423	0.1532	0.0190	3.0675
4	0.0079	0.0589	0.1152	0.2435	0.3407	0.1941	0.0397	3.5913
5	0.0054	0.0165	0.0692	0.2235	0.3403	0.2678	0.0771	3.9888
6	0.0124	0.0377	0.0351	0.0648	0.3330	0.3615	0.1555	4.3747

Table 2. Empirical conditional distribution of the queue length of queue 2 condition on the length of queue 1. The underlying graph is the cycle graph with $n = 12180$ vertices. The last column gives the average queue length of queue 2 given queue 1's length.

	0	1	2	3	4	5	6	avg
0	0.0388	0.2191	0.1753	0.3336	0.1782	0.0459	0.0090	2.5672
1	0.1385	0.1551	0.1477	0.2928	0.2132	0.0392	0.0136	2.4587
2	0.1124	0.1471	0.1420	0.2354	0.2361	0.0759	0.0512	2.7680
3	0.0420	0.1055	0.1484	0.2227	0.2890	0.1477	0.0448	3.2332
4	0.0370	0.0997	0.1388	0.2110	0.2996	0.1656	0.0483	3.3265
5	0.0218	0.0402	0.1428	0.1778	0.2978	0.2712	0.0354	3.7092
6	0.0000	0.0148	0.2001	0.1965	0.3133	0.1657	0.1096	3.7437

Table 3. Empirical conditional distribution of the queue length of queue 2 condition on the length of queue 1. The underlying graph is the cycle graph with $n = 60900$ vertices. The last column gives the average queue length of queue 2 given queue 1's length.

The fluid limit estimation error has been analyzed in [31] and [14] for general continuous time Markov Chains. Since our process is not Markov, the results in [31] and [14] cannot be directly applied. Therefore, we conduct simulations to investigate the relationship between the system size n and the fluid-limit estimation error, i.e. $\sup_{[0, T]} \|\mathbf{X}(t) - \mathbf{x}(t)\|_1$, for a fixed graph family. We generate 200 sample paths for cycle graphs with different n 's. For each sample path, we find the estimation error. We set $\lambda = 0.7$, and each system is simulated for $T = 20$ units of time. The

systems are initialized to be empty. The reset periods are set to be $\lfloor c \log n \rfloor$ where $c = 5.4$. The average approximation error of 200 sample paths is computed and plotted in Figure 6. The results suggests that the mean estimation error is proportional to $n^{-1/2}$, which suggests that a Central Limit Theorem approximation should hold. This result suggests a similar behavior of NBRWR-Pod with regular *power-of-d-choices* in terms of the scaling of approximation error [14, 31].

We also conduct simulations to investigate if the estimation error of the ODE depends on the graph topology. We generate 500 sample paths for each of the following four graphs: LPS and cycle graphs with $n = 12180$ and $n = 34440$ vertices. For each sample path, we find the L_1 error between the system dynamics and the ODE, i.e. $\sup_{[0,T]} \|X(t) - \mathbf{x}(t)\|_1$. We set $\lambda = 0.7$, and each system is simulated for $T = 20$ units of time. The systems are initialized to be empty. The histograms of the approximation errors are shown in Figure 7. We can see that while the mean approximation error seems to be the same for different graph families, the variances of the errors show a significant difference. Therefore, while it is clear that a Central Limit Theorem approximation should hold, additional graph properties might influence the variance. Determining this is for future work.

Fig. 6. Log-log plot of Average estimation error of the ODE with respect to the system dynamics vs. system size, with error bars showing the standard deviation of the 200 data points

Fig. 7. Histogram of absolute approximation errors of the ODE with respect to the system dynamics. The reset period is set to 200 for all plots.

Received November 2018; revised December 2018; accepted January 2019